Robust Design for Linear Non-Regenerative MIMO Relays

Yue Rong
Department of Electrical and Computer Engineering, Curtin University of Technology, Bentley, WA 6102, Australia

***Abstract*—Optimal source precoding matrix and relay amplifying matrices have been developed in recent works for linear non-regenerative multiple-input multiple-output (MIMO) relay communication systems. All these works require the channel state information (CSI) knowledge in order to perform the optimization. However, in practical relay systems, the CSI is unknown and has to be estimated. There is always mismatch between the true and the estimated CSI. In this paper, the true CSI is modelled as a Gaussian random matrix with the estimated CSI as the mean value. First, by using the averaged mean-squared error (MSE) matrix of the signal waveform estimation, we develop the source and relay matrices which optimize most commonly used objective functions subjecting to the averaged transmission power constraints. Second, we consider the quality-of-service (QoS) issue by minimizing the averaged total transmission power subjecting to the averaged MSE constraints at each data stream. Simulation results demonstrate the improved robustness of the proposed algorithms against CSI errors.**

I. Introduction

It is well-known that wireless relays are useful in increasing the coverage of wireless communications under power and spectral constraints. Wireless relays can be regenerative or non-regenerative. For the non-regenerative relay strategy, the relay node only amplifies and retransmits its received signals. Thus, the complexity of the nonregenerative strategy is much lower than that of the regenerative strategy.

Recently, there have been many research efforts on non-regenerative multiple-input multiple-output (MIMO) relay systems where the relay nodes have multiple antennas [1]-[6]. For a three-node two-hop MIMO relay system where the direct-link is omitted, a unified framework is established for linear non-regenerative MIMO relay systems with a broad class of objective functions [3]. The framework in [3] has been further extended to multi-hop non-regenerative MIMO relay systems with arbitrary number of hops [4]. Both [3] and [4] consider a linear minimal mean-squared error (MMSE) receiver at the destination node. Recently, it has been shown in [5] that by using a nonlinear decision feedback equalizer (DFE) based on the MMSE criterion at the destination node, the system bit-error-rate (BER) performance can be significantly improved. The optimal source and relay matrices of a multi-hop MIMO relay system which guarantee that the predetermined quality-of-service (QoS) criteria be attained with the minimal total transmission power have been derived in [6].

In order to optimize the source and relay matrices, the channel state information (CSI) knowledge of all hops is required at the scheduler in [1]-[6]. However, in practical relay communication systems, the exact CSI is unknown and has to be estimated. There is always mismatch between the true and the estimated CSI. Obviously, the performance of the algorithms in [1]-[6] will degrade due to such CSI mismatch. In this paper, we develop the optimal source and relay matrices which are robust against CSI errors. In particular, 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 correlation of the CSI mismatch [7]-[9]. First, by using the averaged MSE matrix of the signal waveform estimation, we develop the source and relay matrices which optimize most commonly used objective functions when the objective is Schur-concave or Schur-convex [10]. It is shown that the structure of the optimal source and relay matrices includes the CSI mismatch information. Moreover, the available power at the source and the relay nodes is optimally distributed among all data stream in a robust fashion against the CSI mismatch. Second, we consider the QoS issue by minimizing the averaged transmission power subjecting to the averaged MSE constraints at each data stream. Simulation results demonstrate the improved robustness of the proposed approaches against the CSI mismatch.

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. The communication process between the source and destination nodes is completed in two time slots. In the first time slot, the $N_b \times 1$ signal vector $\mathbf{s}$ is linearly precoded as

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

where $\mathbf{F}_1$ is the $N_1 \times N_b$ precoding matrix. 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. The precoded signal vector $\mathbf{x}_1$ is transmitted to the relay node, and the received signal vector is given by

$$\mathbf{y}_2 = \mathbf{H}_1\mathbf{x}_1 + \mathbf{v}_2 \tag{2}$$

where $\mathbf{H}_1$ is the $N_2 \times N_1$ MIMO fading channel matrix between the source and relay nodes, and $\mathbf{v}_2$ is an $N_2 \times 1$ noise vector at the relay node.

In the second time slot, a linear non-regenerative relay matrix $\mathbf{F}_2$ is used at the relay node to amplify the received signal vector as in [1]-[6]. The input-output relationship at the relay node can be written as

$$\mathbf{x}_2 = \mathbf{F}_2\mathbf{y}_2. \tag{3}$$

The amplified signal vector $\mathbf{x}_2$ is transmitted to the destination, and the received signal vector is given by

$$\mathbf{y}_3 = \mathbf{H}_2\mathbf{x}_2 + \mathbf{v}_3 \tag{4}$$

where $\mathbf{H}_2$ is the $N_3 \times N_2$ MIMO fading channel matrix between the source and relay nodes, $\mathbf{v}_3$ is an $N_3 \times 1$ noise vector at the destination. We assume that $\mathbf{v}_i$, $i = 2, 3$ is independent and identically distributed (i.i.d.) additive white Gaussian noise (AWGN) with zero mean and unit variance.

Combining (1)-(4), the signal vector at the destination node can be written as

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

where $\mathbf{G} \triangleq \mathbf{H}_2\mathbf{F}_2\mathbf{H}_1\mathbf{F}_1$, $\bar{\mathbf{v}} \triangleq \mathbf{H}_2\mathbf{F}_2\mathbf{v}_2 + \mathbf{v}_3$. Using a linear receiver at the destination node, the estimated signal vector $\hat{\mathbf{s}}$ can be written as

$$\hat{\mathbf{s}} = \mathbf{W}^H\mathbf{y}_3$$

where $\mathbf{W}$ is the $N_3 \times N_b$ weight matrix. The MSE matrix of the signal waveform estimation $\mathbf{E} \triangleq \mathrm{E}[(\hat{\mathbf{s}} - \mathbf{s})(\hat{\mathbf{s}} - \mathbf{s})^H]$ can be written as

$$\begin{aligned}\mathbf{E} &= (\mathbf{W}^H\mathbf{G} - \mathbf{I}_{N_b})(\mathbf{W}^H\mathbf{G} - \mathbf{I}_{N_b})^H + \mathbf{W}^H\mathbf{C}_{\bar{v}}\mathbf{W}\\ &= \mathbf{W}^H(\mathbf{H}_2\mathbf{F}_2\mathbf{H}_1\mathbf{F}_1\mathbf{F}_1^H\mathbf{H}_1^H\mathbf{F}_2^H\mathbf{H}_2^H + \mathbf{H}_2\mathbf{F}_2\mathbf{F}_2^H\mathbf{H}_2^H\\ &\quad + \mathbf{I}_{N_3})\mathbf{W} - \mathbf{W}^H\mathbf{H}_2\mathbf{F}_2\mathbf{H}_1\mathbf{F}_1 - \mathbf{F}_1^H\mathbf{H}_1^H\mathbf{F}_2^H\mathbf{H}_2^H\mathbf{W} + \mathbf{I}_{N_b}\end{aligned} \tag{5}$$

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.

III. Robust MIMO Relay Design

When there is mismatch between the true and the estimated CSI, the true channel $\mathbf{H}_i$ can be represented by the well-known Kronecker model [7]-[9], where $\mathbf{H}_i$ is a complex-valued Gaussian random matrix with

$$\mathbf{H}_i \sim \mathcal{CN}(\bar{\mathbf{H}}_i, \mathbf{\Theta}_i \otimes \mathbf{\Phi}_i), \qquad i = 1, 2 \tag{6}$$

where the mean value is the estimated channel matrix $\bar{\mathbf{H}}_i$, $\otimes$ stands for the Kronecker product [11], $\mathbf{\Theta}_i$ denotes the covariance matrix at the receiver side, while $\mathbf{\Phi}_i$ is the covariance matrix seen from the transmitter side.

A. Robust MIMO Relay for Most Commonly Used Objectives

It has been shown in [3] that many commonly used MIMO relay system design objectives can be written as a function of the main diagonal elements of the MSE matrix $\mathbf{E}$. It can be seen from (5) that $\mathbf{E}$ is a function of channel matrices $\mathbf{H}_1$ and $\mathbf{H}_2$, whose true value is unknown in practice. Therefore, it is impossible to optimize $q(\mathbf{d}[\mathbf{E}])$, where as in [3], q denotes a general objective function, and $\mathbf{d}[\mathbf{E}]$ stands for the main diagonal elements of $\mathbf{E}$. If we design $\mathbf{F}_1$ 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 $\bar{\mathbf{H}}_i$, $i = 1, 2$.

In this subsection, instead of optimizing $q(\mathbf{d}[\mathbf{E}])$, we optimize the objective function of $q(\mathbf{d}[\mathrm{E}[\mathbf{E}]])$, where the statistical expectation $\mathrm{E}[\cdot]$ is carried out with respect to $\mathbf{H}_i$, $i = 1, 2$ with the distribution in (6).

THEOREM 1: The statistical expectation of $\mathbf{E}$ is given by

$$\mathrm{E}[\mathbf{E}] = \mathbf{W}^H\mathbf{A}\mathbf{W} - \mathbf{W}^H\bar{\mathbf{H}}_2\mathbf{F}_2\bar{\mathbf{H}}_1\mathbf{F}_1 - \mathbf{F}_1^H\bar{\mathbf{H}}_1^H\mathbf{F}_2^H\bar{\mathbf{H}}_2^H\mathbf{W} + \mathbf{I}_{N_b} \tag{7}$$

where

$$\begin{aligned}\mathbf{A} &\triangleq \bar{\mathbf{H}}_2\mathbf{F}_2\big(\bar{\mathbf{H}}_1\mathbf{F}_1\mathbf{F}_1^H\bar{\mathbf{H}}_1^H + \alpha_1\mathbf{\Phi}_1 + \mathbf{I}_{N_2}\big)\mathbf{F}_2^H\bar{\mathbf{H}}_2^H\\ &\quad + \alpha_2\mathbf{\Phi}_2 + \mathbf{I}_{N_3} \qquad (8)\\ \alpha_1 &\triangleq \mathrm{tr}(\mathbf{F}_1\mathbf{F}_1^H\mathbf{\Theta}_1^T) \qquad (9)\\ \alpha_2 &\triangleq \mathrm{tr}(\mathbf{F}_2(\bar{\mathbf{H}}_1\mathbf{F}_1\mathbf{F}_1^H\bar{\mathbf{H}}_1^H + \alpha_1\mathbf{\Phi}_1 + \mathbf{I}_{N_2})\mathbf{F}_2^H\mathbf{\Theta}_2^T). \qquad (10)\end{aligned}$$

Here $\mathrm{tr}(\cdot)$ denotes the matrix trace, and $(\cdot)^T$ stands for the matrix transpose.

PROOF: See Appendix A. □

The weight matrix $\mathbf{W}$ which minimizes (7) is the famous Wiener filter [12] given by

$$\mathbf{W} = \mathbf{A}^{-1}\bar{\mathbf{H}}_2\mathbf{F}_2\bar{\mathbf{H}}_1\mathbf{F}_1 \tag{11}$$

where $(\cdot)^{-1}$ denotes the matrix inversion. Substituting (11) back into (7), we have

$$\mathrm{E}[\mathbf{E}] = \mathbf{I}_{N_b} - \mathbf{F}_1^H\bar{\mathbf{H}}_1^H\mathbf{F}_2^H\bar{\mathbf{H}}_2^H\mathbf{A}^{-1}\bar{\mathbf{H}}_2\mathbf{F}_2\bar{\mathbf{H}}_1\mathbf{F}_1. \tag{12}$$

The transmission power consumed by the relay node can be written as [3]

$$\mathrm{tr}\big(\mathbf{F}_2\big(\mathbf{H}_1\mathbf{F}_1\mathbf{F}_1^H\mathbf{H}_1^H + \mathbf{I}_{N_2}\big)\mathbf{F}_2^H\big). \tag{13}$$

However, since the true $\mathbf{H}_1$ is unknown, (13) is also unknown. Thus, the power constraint can not be imposed for any fixed $\mathbf{H}_1$. In this paper, we consider the averaged transmission power at the relay node, which is given by

$$\begin{aligned}&\mathrm{E}[\mathrm{tr}\big(\mathbf{F}_2\big(\mathbf{H}_1\mathbf{F}_1\mathbf{F}_1^H\mathbf{H}_1^H + \mathbf{I}_{N_2}\big)\mathbf{F}_2^H\big)]\\ &= \mathrm{tr}\big(\mathbf{F}_2\big(\mathrm{E}[\mathbf{H}_1\mathbf{F}_1\mathbf{F}_1^H\mathbf{H}_1^H] + \mathbf{I}_{N_2}\big)\mathbf{F}_2^H\big)\\ &= \mathrm{tr}\big(\mathbf{F}_2\big(\bar{\mathbf{H}}_1\mathbf{F}_1\mathbf{F}_1^H\bar{\mathbf{H}}_1^H + \alpha_1\mathbf{\Phi}_1 + \mathbf{I}_{N_2}\big)\mathbf{F}_2^H\big).\end{aligned} \tag{14}$$

Combining (12) and (14), the robust relay optimization problem can be written as

$$\begin{aligned}\min_{\mathbf{F}_1,\mathbf{F}_2}\ & q\big(\mathbf{d}\big[\mathbf{I}_{N_b} - \mathbf{F}_1^H\bar{\mathbf{H}}_1^H\mathbf{F}_2^H\bar{\mathbf{H}}_2^H\mathbf{A}^{-1}\bar{\mathbf{H}}_2\mathbf{F}_2\bar{\mathbf{H}}_1\mathbf{F}_1\big]\big) \qquad (15)\\ \text{s.t.}\ & \mathrm{tr}\big(\mathbf{F}_2\big(\bar{\mathbf{H}}_1\mathbf{F}_1\mathbf{F}_1^H\bar{\mathbf{H}}_1^H + \alpha_1\mathbf{\Phi}_1 + \mathbf{I}_{N_2}\big)\mathbf{F}_2^H\big) \le P_2 \qquad (16)\\ & \mathrm{tr}\big(\mathbf{F}_1\mathbf{F}_1^H\big) \le P_1 \qquad (17)\end{aligned}$$

where $P_i > 0$, $i = 1, 2$, is the transmission power available at the ith node. Here (16) and (17) are the transmission power constraint at the relay and source node, respectively. The problem (15)-(17) provides a statistically robust MIMO relay system design.

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

$$\begin{aligned}\mathbf{\Phi}_i &= \mathbf{U}_{\Phi_i}\mathbf{\Lambda}_{\Phi_i}\mathbf{U}_{\Phi_i}^H \qquad (18)\\ \tilde{\mathbf{\Lambda}}_{\Phi_i} &\triangleq \alpha_i\mathbf{\Lambda}_{\Phi_i} + \mathbf{I}_{N_{i+1}} \qquad (19)\\ \tilde{\mathbf{H}}_i &\triangleq \tilde{\mathbf{\Lambda}}_{\Phi_i}^{-\frac{1}{2}}\mathbf{U}_{\Phi_i}^H\bar{\mathbf{H}}_i = \tilde{\mathbf{U}}_i\tilde{\mathbf{\Sigma}}_i\tilde{\mathbf{V}}_i^H \qquad (20)\end{aligned}$$

where $\mathbf{U}_{\Phi_i}$ and $\tilde{\mathbf{U}}_i$ are $N_{i+1} \times N_{i+1}$ unitary matrices, $\mathbf{\Lambda}_{\Phi_i}$ 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 $\tilde{\mathbf{\Sigma}}_i$ is an $N_{i+1} \times N_i$ singularvalue matrix. The following theorem establishes the structure of the optimal $\mathbf{F}_1$ and $\mathbf{F}_2$ which are robust against the CSI mismatch.

THEOREM 2: For the robust relay design problem (15)-(17), if q is Schur-concave, the optimal $\mathbf{F}_1$ and $\mathbf{F}_2$ are given by

$$\mathbf{F}_1 = \tilde{\mathbf{V}}_{1,1}\mathbf{\Lambda}_1, \quad \mathbf{F}_2 = \tilde{\mathbf{V}}_{2,1}\mathbf{\Lambda}_2\tilde{\mathbf{U}}_{1,1}^H\tilde{\mathbf{\Lambda}}_{\Phi_1}^{-\frac{1}{2}}\mathbf{U}_{\Phi_1}^H \tag{21}$$

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 singularvalues, respectively, and $\mathbf{\Lambda}_i$, $i = 1, 2$ are $N_b \times N_b$ diagonal matrices. If q is Schur-convex, the optimal $\mathbf{F}_2$ is given in (21), while the optimal $\mathbf{F}_1$ is $\mathbf{F}_1 = \tilde{\mathbf{V}}_{1,1}\mathbf{\Lambda}_1\mathbf{V}_0$, where $\mathbf{V}_0$ is an $N_b \times N_b$ unitary matrix such that $\mathrm{E}[\mathbf{E}]$ has identical main-diagonal elements.

PROOF: See Appendix B. □

From (9) and (10) we find that α_1 is a function of $\mathbf{F}_1$, and α_2 is a function of $\mathbf{F}_1$ and $\mathbf{F}_2$. Consequently, it can be seen from (19) and (20) that $\tilde{\mathbf{V}}_1$ and $\tilde{\mathbf{U}}_1$ depend on $\mathbf{F}_1$, and $\tilde{\mathbf{V}}_2$ depends on $\mathbf{F}_1$ and $\mathbf{F}_2$. Thus, from (21) we find that the explicit structure of the optimal $\mathbf{F}_1$ and $\mathbf{F}_2$ is very difficult to find for general $\mathbf{\Theta}_i$ and $\mathbf{\Phi}_i$. In the following, we show that if $\mathbf{\Theta}_i = \mathbf{I}_i$ and/or $\mathbf{\Phi}_i = \mathbf{I}_{i+1}$, $i = 1, 2$. The structure of the optimal $\mathbf{F}_1$ and $\mathbf{F}_2$ can be obtained explicitly.

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

$$\min_{\mathbf{F}_1, \mathbf{F}_2} \; q\big(\mathbf{d}\big[\mathbf{I}_{N_b} - \mathbf{F}_1^H\bar{\mathbf{H}}_1^H\mathbf{F}_2^H\bar{\mathbf{H}}_2^H\mathbf{B}^{-1}\bar{\mathbf{H}}_2\mathbf{F}_2\bar{\mathbf{H}}_1\mathbf{F}_1\big]\big) \tag{22}$$

$$\text{s.t.} \;\; \mathrm{tr}\Big(\mathbf{F}_2\big(\bar{\mathbf{H}}_1\mathbf{F}_1\mathbf{F}_1^H\bar{\mathbf{H}}_1^H + \beta_1\mathbf{I}_{N_2}\big)\mathbf{F}_2^H\Big) \leq P_2 \tag{23}$$

$$\mathrm{tr}\Big(\mathbf{F}_1\mathbf{F}_1^H\Big) \leq P_1. \tag{24}$$

where

$$\mathbf{B} \triangleq \bar{\mathbf{H}}_2\mathbf{F}_2\big(\bar{\mathbf{H}}_1\mathbf{F}_1\mathbf{F}_1^H\bar{\mathbf{H}}_1^H + \beta_1\mathbf{I}_{N_2}\big)\mathbf{F}_2^H\bar{\mathbf{H}}_2^H + \beta_2\mathbf{I}_{N_3} \tag{25}$$

$$\beta_1 \triangleq \mathrm{tr}(\mathbf{F}_1\mathbf{F}_1^H\mathbf{\Theta}_1^T) + 1 \tag{26}$$

$$\beta_2 \triangleq \mathrm{tr}(\mathbf{F}_2(\bar{\mathbf{H}}_1\mathbf{F}_1\mathbf{F}_1^H\bar{\mathbf{H}}_1^H + \beta_1\mathbf{I}_{N_2})\mathbf{F}_2^H\mathbf{\Theta}_2^T) + 1. \tag{27}$$

Let us introduce the SVDs of $\bar{\mathbf{H}}_i = \mathbf{U}_i\mathbf{\Sigma}_i\mathbf{V}_i^H$. It can be easily seen from (18)-(20) that for $\mathbf{\Phi}_i = \mathbf{I}_{N_{i+1}}$, we have $\tilde{\mathbf{V}}_i = \mathbf{V}_i$ and $\tilde{\mathbf{U}}_i = \mathbf{U}_i$. Consequently, for Schur-concave q, we have

$$\mathbf{F}_1 = \mathbf{V}_{1,1}\mathbf{\Lambda}_1, \quad \mathbf{F}_2 = \mathbf{V}_{2,1}\mathbf{\Lambda}_2\mathbf{U}_{1,1}^H \tag{28}$$

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 singularvalues, respectively. If q is Schur-convex, $\mathbf{F}_2$ is given in (28), 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 (28) back into (22)-(24), we obtain the following problem

$$\min_{\mathbf{\Lambda}_1, \mathbf{\Lambda}_2} \; q\big([\mathbf{I}_{N_b} + \mathbf{\Sigma}_{1,1}^2\mathbf{\Lambda}_1^2\mathbf{\Sigma}_{2,1}^2\mathbf{\Lambda}_2^2[\beta_1\mathbf{\Sigma}_{2,1}^2\mathbf{\Lambda}_2^2 + \beta_2\mathbf{I}_{N_b}]^{-1}]^{-1}\big) \tag{29}$$

$$\text{s.t.} \;\; \mathrm{tr}(\mathbf{\Lambda}_2^2(\mathbf{\Lambda}_1^2\mathbf{\Sigma}_{1,1}^2 + \beta_1\mathbf{I}_{N_b})) \leq P_2 \tag{30}$$

$$\mathrm{tr}(\mathbf{\Lambda}_1^2) \leq P_1 \tag{31}$$

where $\mathbf{\Sigma}_{i,1}$ is a diagonal matrix containing the largest N_b singularvalues in $\mathbf{\Sigma}_i$, $i = 1, 2$. Problem (29)-(31) can be equivalently written as

$$\min_{\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2} \; q\left(\left\{\left[1 + \frac{\sigma_{1,k}^2\lambda_{1,k}^2\sigma_{2,k}^2\lambda_{2,k}^2}{\beta_1\sigma_{2,k}^2\lambda_{2,k}^2 + \beta_2}\right]^{-1}\right\}\right) \tag{32}$$

$$\text{s.t.} \;\; \sum_{k=1}^{N_b}\lambda_{2,k}^2(\lambda_{1,k}^2\sigma_{1,k}^2 + \beta_1) \leq P_2 \tag{33}$$

$$\sum_{i=1}^{N_b}\lambda_{1,k}^2 \leq P_1 \tag{34}$$

$$\lambda_{1,k} \geq 0, \quad \lambda_{2,k} \geq 0, \qquad k = 1, \cdots, N_b \tag{35}$$

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

$$a_k \triangleq \sigma_{1,k}^2, \qquad x_k \triangleq \lambda_{1,k}^2 \tag{36}$$

$$b_k \triangleq \sigma_{2,k}^2, \qquad y_k \triangleq \lambda_{2,k}^2(\lambda_{1,k}^2\sigma_{1,k}^2 + \beta_1). \tag{37}$$

The problem (32)-(35) can be simplified to

$$\min_{\mathbf{x}, \mathbf{y}} \; q\left(\left\{1 - \frac{a_kx_kb_ky_k}{(a_kx_k + \beta_1)(b_ky_k + \beta_2)}\right\}\right) \tag{38}$$

$$\text{s.t.} \;\; \sum_{k=1}^{N_b}x_k \leq P_1, \quad x_k \geq 0, \qquad k = 1, \cdots, N_b \tag{39}$$

$$\sum_{k=1}^{N_b}y_k \leq P_2, \quad y_k \geq 0, \qquad k = 1, \cdots, N_b \tag{40}$$

where $\beta_1 = \sum_{k=1}^{N_b}x_k[\tilde{\mathbf{\Theta}}_1]_{k,k} + 1$, $\beta_2 = \sum_{k=1}^{N_b}y_k[\tilde{\mathbf{\Theta}}_2]_{k,k} + 1$, $\mathbf{x} \triangleq [x_1, \cdots, x_{N_b}]^T$, and $\mathbf{y} \triangleq [y_1, \cdots, y_{N_b}]^T$.

The problem (38)-(40) can be solved by an iterative method. To update $\mathbf{x}$, we solve the following problem with $\mu_k \triangleq \frac{b_ky_k}{b_ky_k + \beta_2}$, $k = 1, \cdots, N_b$

$$\min_{\mathbf{x}} \; q\left(\left\{1 - \frac{\mu_ka_kx_k}{a_kx_k + \beta_1}\right\}\right) \tag{41}$$

$$\text{s.t.} \;\; \sum_{k=1}^{N_b}x_k \leq P_1, \qquad x_k \geq 0, \qquad k = 1, \cdots, N_b. \tag{42}$$

The solution to the problem (41)-(42) follows the water-filling principle. Similarly, to update $\mathbf{y}$, we solve the following problem with $\nu_k \triangleq \frac{a_kx_k}{a_kx_k + \beta_1}$, $k = 1, \cdots, N_b$

$$\min_{\mathbf{y}} \; q\left(\left\{1 - \frac{\nu_kb_ky_k}{b_ky_k + \beta_2}\right\}\right)$$

$$\text{s.t.} \;\; \sum_{k=1}^{N_b}y_k \leq P_2, \qquad y_k \geq 0, \qquad k = 1, \cdots, N_b.$$

For Schur-convex q, it can be shown similar to [3] that the optimal power loading is obtained by solving (38)-(40) 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 $\mathbf{\Theta}_1 = \mathbf{I}_{N_1}$ and $\mathbf{\Theta}_2 = \mathbf{I}_{N_2}$, we have $\alpha_1 = \mathrm{tr}(\mathbf{F}_1\mathbf{F}_1^H)$, and $\alpha_2 = \mathrm{tr}(\mathbf{F}_2(\bar{\mathbf{H}}_1\mathbf{F}_1\mathbf{F}_1^H\bar{\mathbf{H}}_1^H + \alpha_1\mathbf{\Phi}_1 + \mathbf{I}_{N_2})\mathbf{F}_2^H)$.

Now we show that (12) is decreasing w.r.t. $\mathrm{tr}(\mathbf{F}_1\mathbf{F}_1^H)$. By introducing $\tilde{\mathbf{F}}_1 = \mathbf{F}_1/\sqrt{\mathrm{tr}(\mathbf{F}_1\mathbf{F}_1^H)}$, (12) can be written as

$$\mathrm{E}[\mathbf{E}] = \mathbf{I}_{N_b} - \tilde{\mathbf{F}}_1^H\bar{\mathbf{H}}_1^H\mathbf{F}_2^H\bar{\mathbf{H}}_2^H\bar{\mathbf{A}}^{-1}\bar{\mathbf{H}}_2\mathbf{F}_2\bar{\mathbf{H}}_1\tilde{\mathbf{F}}_1 \tag{43}$$

where

$$\bar{\mathbf{A}} = \bar{\mathbf{H}}_2\mathbf{F}_2\left(\bar{\mathbf{H}}_1\tilde{\mathbf{F}}_1\tilde{\mathbf{F}}_1^H\bar{\mathbf{H}}_1^H + \mathbf{\Phi}_1 + \frac{1}{\mathrm{tr}(\mathbf{F}_1\mathbf{F}_1^H)}\mathbf{I}_{N_2}\right)\mathbf{F}_2^H\bar{\mathbf{H}}_2^H + \tilde{\alpha}_2\mathbf{\Phi}_2 + \frac{1}{\mathrm{tr}(\mathbf{F}_1\mathbf{F}_1^H)}\mathbf{I}_{N_3} \tag{44}$$

$$\tilde{\alpha}_2 \triangleq \mathrm{tr}\left(\mathbf{F}_2\left(\bar{\mathbf{H}}_1\tilde{\mathbf{F}}_1\tilde{\mathbf{F}}_1^H\bar{\mathbf{H}}_1^H + \mathbf{\Phi}_1 + \frac{1}{\mathrm{tr}(\mathbf{F}_1\mathbf{F}_1^H)}\mathbf{I}_{N_2}\right)\mathbf{F}_2^H\right). \tag{45}$$

It can be clearly seen from (43)-(45) that for a given $\tilde{\mathbf{F}}_1$, $\mathrm{E}[\mathbf{E}]$ is a decreasing function of $\mathrm{tr}(\mathbf{F}_1\mathbf{F}_1^H)$. It can be shown in a similar way that $\mathrm{E}[\mathbf{E}]$ also decreases with respect to $\mathrm{tr}\big(\mathbf{F}_2(\bar{\mathbf{H}}_1\mathbf{F}_1\mathbf{F}_1^H\bar{\mathbf{H}}_1^H + \alpha_1\mathbf{\Phi}_1 + \mathbf{I}_{N_2})\mathbf{F}_2^H\big)$. Thus, the optimal solution of $\mathbf{F}_1$ and $\mathbf{F}_2$ occurs at $\alpha_1 = P_1$ and $\alpha_2 = P_2$. From (19) and (20), we find that in this case $\tilde{\mathbf{U}}_i$. 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 (21) back into (15)-(17), we have

$$\min_{\mathbf{\Lambda}_1,\mathbf{\Lambda}_2} \quad q\big([\mathbf{I}_{N_b} + \tilde{\mathbf{\Sigma}}_{1,1}^2\mathbf{\Lambda}_1^2\tilde{\mathbf{\Sigma}}_{2,1}^2\mathbf{\Lambda}_2^2[\tilde{\mathbf{\Sigma}}_{2,1}^2\mathbf{\Lambda}_2^2 + \mathbf{I}_{N_b}]^{-1}]^{-1}\big) \tag{46}$$

$$\text{s.t.} \quad \mathrm{tr}(\mathbf{\Lambda}_2^2(\mathbf{\Lambda}_1^2\tilde{\mathbf{\Sigma}}_{1,1}^2 + \mathbf{I}_{N_b})) \le P_2 \tag{47}$$

$$\mathrm{tr}(\mathbf{\Lambda}_1^2) \le P_1 \tag{48}$$

where $\tilde{\mathbf{\Sigma}}_{i,1}$ is a diagonal matrix containing the largest N_b singularvalues in $\tilde{\mathbf{\Sigma}}_i$, $i = 1, 2$. The problem (46)-(48) can be solved by the iterative method we just developed.

B. Robust MIMO Relay With QoS Constraints

The optimal robust relay scheme developed in the previous subsection does not consider any QoS constraints for each data stream. In practical communication systems, QoS criteria are very important. When the CSI is exactly known, the optimal source and relay matrices which minimize the total transmission power subjecting to MSE constraints at each data stream have been developed in [6]. However, due to the mismatch between the true and the estimated CSI, the algorithms in [6] can not guarantee the satisfaction of the QoS criteria. In this subsection, we develop robust source and relay matrices which guarantee that the predetermined QoS criteria be attained by the averaged MSE at each data stream with the minimal averaged total transmission power.

Using (15)-(17), the QoS-constrained robust relay design problem can be written as

$$\min_{\mathbf{F}_1,\mathbf{F}_2} \quad \mathrm{tr}\big(\mathbf{F}_2(\bar{\mathbf{H}}_1\mathbf{F}_1\mathbf{F}_1^H\bar{\mathbf{H}}_1^H + \alpha_1\mathbf{\Phi}_1 + \mathbf{I}_{N_2})\mathbf{F}_2^H + \mathbf{F}_1\mathbf{F}_1^H\big) \tag{49}$$

$$\text{s.t.} \quad \mathbf{d}\big[\mathbf{I}_{N_b} - \mathbf{F}_1^H\bar{\mathbf{H}}_1^H\mathbf{F}_2^H\bar{\mathbf{H}}_2^H\mathbf{A}^{-1}\bar{\mathbf{H}}_2\mathbf{F}_2\bar{\mathbf{H}}_1\mathbf{F}_1\big] \le \mathbf{g} \tag{50}$$

where $\mathbf{g} = [g_1, g_2, \cdots, g_{N_b}]^T$ is the QoS vector measured in terms of the averaged MSE of each data stream that must be satisfied, and $\mathbf{A}$ and α_1 are given in (8) and (9), respectively.

THEOREM 3: The optimal $\mathbf{F}_1$ and $\mathbf{F}_2$ as the solution to the problem (49)-(50) are given by

$$\mathbf{F}_1 = \tilde{\mathbf{V}}_{1,1}\mathbf{\Lambda}_1\mathbf{U}_0, \quad \mathbf{F}_2 = \tilde{\mathbf{V}}_{2,1}\mathbf{\Lambda}_2\tilde{\mathbf{U}}_{1,1}^H\tilde{\mathbf{\Lambda}}_{\Phi_1}^{-\frac{1}{2}}\mathbf{U}_{\Phi_1}^H \tag{51}$$

where $\mathbf{U}_0$ is an $N_b \times N_b$ unitary matrix such that $[\mathrm{E}[\mathbf{E}]]_{k,k} = g_k$, $k = 1, \cdots, N_b$.

PROOF: See Appendix C. $\square$

For the case of $\mathbf{\Phi}_i = \mathbf{I}_{N_{i+1}}$, $i = 1, 2$, using (22)-(24), the QoS-constrained relay design problem can be written as

$$\min_{\mathbf{F}_1,\mathbf{F}_2} \quad \mathrm{tr}\big(\mathbf{F}_2(\bar{\mathbf{H}}_1\mathbf{F}_1\mathbf{F}_1^H\bar{\mathbf{H}}_1^H + \beta_1\mathbf{I}_{N_2})\mathbf{F}_2^H\big) + \mathrm{tr}\big(\mathbf{F}_1\mathbf{F}_1^H\big) \tag{52}$$

$$\text{s.t.} \quad \mathbf{d}\big[\mathbf{I}_{N_b} - \mathbf{F}_1^H\bar{\mathbf{H}}_1^H\mathbf{F}_2^H\bar{\mathbf{H}}_2^H\mathbf{B}^{-1}\bar{\mathbf{H}}_2\mathbf{F}_2\bar{\mathbf{H}}_1\mathbf{F}_1\big] \le \mathbf{g} \tag{53}$$

where $\mathbf{B}$ and β_1 are given in (25) and (26), respectively. In this case, we have

$$\mathbf{F}_1 = \mathbf{V}_{1,1}\mathbf{\Lambda}_1\mathbf{U}_0, \quad \mathbf{F}_2 = \mathbf{V}_{2,1}\mathbf{\Lambda}_2\mathbf{U}_{1,1}^H.$$

Using the proof of Theorem 1 in [6], it can be shown that the optimal power loading matrices $\mathbf{\Lambda}_1$ and $\mathbf{\Lambda}_2$ are the solution to the following problem

$$\min_{\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2} \quad \sum_{k=1}^{N_b}(\lambda_{2,k}^2(\lambda_{1,k}^2\sigma_{1,k}^2 + \beta_1) + \lambda_{1,k}^2) \tag{54}$$

$$\text{s.t.} \quad \mathbf{g} \prec^{\mathrm{w}} \left\{\left(1 + \frac{\sigma_{1,k}^2\lambda_{1,k}^2\sigma_{2,k}^2\lambda_{2,k}^2}{\beta_1\sigma_{2,k}^2\lambda_{2,k}^2 + \beta_2}\right)^{-1}\right\} \tag{55}$$

$$\lambda_{1,k} > 0, \quad \lambda_{2,k} > 0, \qquad k = 1, \cdots, N_b \tag{56}$$

where $\prec^{\mathrm{w}}$ denotes weakly super-majorization [10]. Utilizing the variable substitution in (36)-(37) and the definition of $\prec^{\mathrm{w}}$, the problem (54)-(56) can be equivalently written as

$$\min_{\mathbf{x},\mathbf{y}} \quad \sum_{k=1}^{N_b}(x_k + y_k) \tag{57}$$

$$\text{s.t.} \quad \sum_{k=1}^{j}\left(1 - \frac{a_k x_k b_k y_k}{(a_k x_k + \beta_1)(b_k y_k + \beta_2)}\right) \le \sum_{k=1}^{j} g_k, \quad j = 1, \cdots, N_b \tag{58}$$

$$x_k > 0, \quad y_k > 0, \qquad k = 1, \cdots, N_b \tag{59}$$

Let us introduce $z_k \le \frac{a_k x_k b_k y_k}{(a_k x_k + \beta_1)(b_k y_k + \beta_2)}$, $k = 1, \cdots, N_b$. The problem (57)-(59) can be written as

$$\min_{\mathbf{x},\mathbf{y},\mathbf{z}} \quad \sum_{k=1}^{N_b}(x_k + y_k) \tag{60}$$

$$\text{s.t.} \quad \sum_{k=1}^{j} z_k \ge \sum_{k=1}^{j}(1 - g_k), \qquad j = 1, \cdots, N_b \tag{61}$$

$$z_k\frac{(a_k x_k + \beta_1)(b_k y_k + \beta_2)}{a_k x_k b_k y_k} \le 1, \quad k = 1, \cdots, N_b \tag{62}$$

$$x_k > 0, \quad y_k > 0, \qquad k = 1, \cdots, N_b \tag{63}$$

where $\mathbf{z} \triangleq [z_1, \cdots, z_{N_b}]^T$. If the constraints in (61) can be converted to posynomial upper-bound constraints, then the problem (60)-(63) becomes a geometric programming (GP) problem. Towards this end, we apply the geometric inequality to the left-hand side of (61) such that

$$\sum_{k=1}^{j} z_k \ge \prod_{k=1}^{j}\left(\frac{z_k}{\delta_{j,k}}\right)^{\delta_{j,k}} \tag{64}$$

where $\sum_{k=1}^{j}\delta_{j,k} = 1, j = 1,\cdots,N_b$, and $\delta_{j,k} > 0, k = 1,\cdots,j, j = 1,\cdots,N_b$. Substituting (61) with the inequalities $\prod_{k=1}^{j}\left(\frac{z_k}{\delta_{j,k}}\right)^{\delta_{j,k}} \geq \sum_{k=1}^{j}(1-g_k)$, we have

$$\min_{\mathbf{x},\mathbf{y},\mathbf{z}} \sum_{k=1}^{N_b}(x_k+y_k) \tag{65}$$

$$\text{s.t.}\quad \gamma_j\prod_{k=1}^{j} z_k^{-\delta_{j,k}} \leq 1, \qquad j=1,\cdots,N_b \tag{66}$$

$$z_k(1+\beta_1 a_k^{-1}x_k^{-1})(1+\beta_2 b_k^{-1}y_k^{-1}) \leq 1,$$
$$k=1,\cdots,N_b \tag{67}$$

$$x_k>0,\quad y_k>0,\qquad k=1,\cdots,N_b \tag{68}$$

where

$$\gamma_j \triangleq \sum_{k=1}^{j}(1-g_k)\prod_{k=1}^{j}\delta_{j,k}^{\delta_{j,k}}. \tag{69}$$

The problem (65)-(68) is a GP problem in standard form, which can be converted to a convex optimization problem and efficiently solved by interior-point method-based convex optimization toolbox such as MOSEK [13].

IV. Simulations

In this section, we study the performance of the proposed robust source and relay matrices. 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 $\bar{\mathbf{H}}_i$, $i=1,2$. The true channel matrices are modelled as (6) with $\mathbf{\Theta}_i = \sqrt{\varepsilon_i}\mathbf{I}_{N_i}$ and $\mathbf{\Phi}_i = \sqrt{\varepsilon_i}\mathbf{I}_{N_{i+1}}$, $i=1,2$. We choose $\varepsilon_1=\varepsilon_2=0.2$, which correspond to a severe CSI mismatch. All simulation results are averaged over 1000 independent realizations of the truce channel matrices $\mathbf{H}_1$ and $\mathbf{H}_2$. We set $N_1=N_2=N_3=4$, and define $\text{SNR}_i = \sigma_i^2P_iN_{i+1}/N_i$ as the signal-to-noise ratio (SNR) for the ith ($i=1,2$) hop.

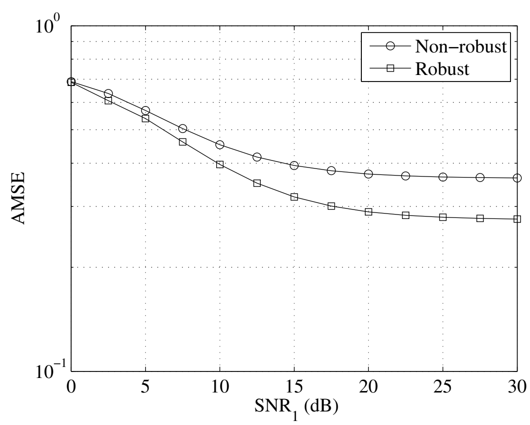

Fig. 1. Example 1: AMSE versus SNR_1; $N_b=3$, $\text{SNR}_2=20$dB.

In the first example, we study the performance of the robust relay algorithm developed in Section III-A. In particular, we choose q as the averaged MSE (AMSE) of all data streams, which is given as $\text{tr}(\text{E}[\mathbf{E}])/N_b$. It has been shown in [3] that $\text{tr}(\text{E}[\mathbf{E}])$ is a Schur-concave function of $\mathbf{d}[\text{E}[\mathbf{E}]]$. Fig. 1 shows the AMSE performance of both the robust and non-robust relay algorithms versus SNR_1 for $\text{SNR}_2=20$dB and $N_b=3$. It can be seen that the proposed algorithm has an improved robustness against the CSI mismatch.

Fig. 2 shows the BER performance of both algorithms versus SNR_1 for $N_b=3$ and $\text{SNR}_2=20$dB. The QPSK constellations are used in the simulation. We can clearly see that the robust relay algorithm has a better BER performance than the non-robust algorithm. In Figs. 1 and 2, the non-robust algorithm refers to the relay algorithm developed in [3].

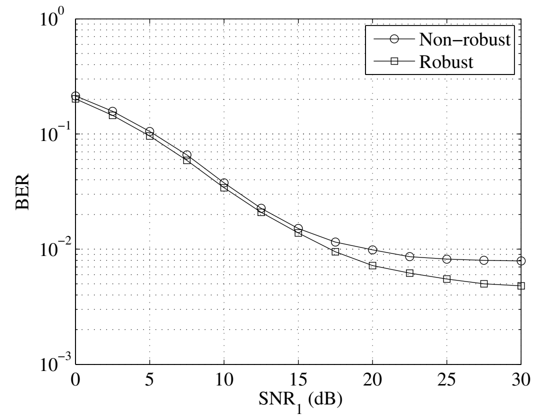

Fig. 2. Example 1: BER versus SNR_1; $N_b=2$, $\text{SNR}_2=20$dB.

In the second example, we simulate the robust relay algorithm with QoS constraints developed in Section III-B. We set $N_b=3$ and choose the same QoS criteria for all 3 data streams, i.e., $q_1=q_2=q_3=q$. Fig. 3 shows the total transmission power required versus MSE (q). It can be seen that the robust relay algorithm requires much less power than the non-robust algorithm developed in [6].

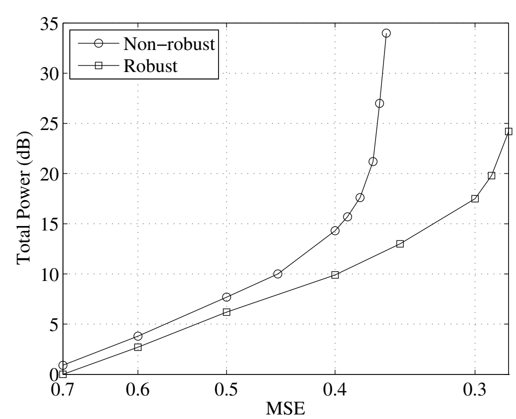

Fig. 3. Example 2: Total power versus MSE (q); $N_b=3$.

V. Conclusions

We have derived the optimal source and relay matrices for linear non-regenerative MIMO relay systems. The proposed source and relay matrices are robust against the CSI mismatch. We have considered most commonly used MIMO system design criteria and addressed the QoS issues. Simulation results show an improved robustness of the proposed algorithms against CSI errors.

Appendix A
Proof of Theorem 1

Lemma 1 [14]: The statistical expectation for the product of four Gaussian matrices $\mathbf{A}\mathbf{b}\mathbf{c}^T\mathbf{D}$ is given by $\mathrm{E}[\mathbf{A}\mathbf{b}\mathbf{c}^T\mathbf{D}] = \mathrm{E}[\mathbf{A}\mathbf{b}]\mathrm{E}[\mathbf{c}^T\mathbf{D}] + \mathrm{E}[\mathbf{c}^T \otimes \mathbf{A}]\mathrm{E}[\mathbf{D}\otimes\mathbf{b}] + \mathrm{E}[\mathbf{A}\mathrm{E}[\mathbf{b}\mathbf{c}^T]\mathbf{D}] - 2\mathrm{E}[\mathbf{A}]\mathrm{E}[\mathbf{b}]\mathrm{E}[\mathbf{c}^T]\mathrm{E}[\mathbf{D}]$.

Lemma 2 [15]: For $\mathbf{H} \sim \mathcal{CN}(\bar{\mathbf{H}}, \mathbf{\Theta}\otimes\mathbf{\Phi})$, there is $\mathrm{E}[\mathbf{H}\mathbf{A}\mathbf{H}^H] = \bar{\mathbf{H}}\mathbf{A}\bar{\mathbf{H}}^H + \mathrm{tr}(\mathbf{A}\mathbf{\Theta}^T)\mathbf{\Phi}$.

Using Lemma 1, we have

$$\begin{aligned}
&\mathrm{E}[\mathbf{H}_2\mathbf{F}_2\mathbf{H}_1\mathbf{F}_1\mathbf{F}_1^H\mathbf{H}_1^H\mathbf{F}_2^H\mathbf{H}_2^H] \\
&= \sum_{i=1}^{N_1}\big(\mathrm{E}[\mathbf{H}_2\mathbf{F}_2\mathbf{H}_1\mathbf{f}_{1,i}]\mathrm{E}[\mathbf{f}_{1,i}^H\mathbf{H}_1^H\mathbf{F}_2^H\mathbf{H}_2^H] \\
&\quad +\mathrm{E}[(\mathbf{f}_{1,i}^H\mathbf{H}_1^H)\otimes(\mathbf{H}_2\mathbf{F}_2)]\mathrm{E}[(\mathbf{F}_2^H\mathbf{H}_2^H)\otimes(\mathbf{H}_1\mathbf{f}_{1,i})] \\
&\quad +\mathrm{E}[\mathbf{H}_2\mathbf{F}_2\mathrm{E}[\mathbf{H}_1\mathbf{f}_{1,i}\mathbf{f}_{1,i}^H\mathbf{H}_1^H]\mathbf{F}_2^H\mathbf{H}_2^H]\big) \\
&\quad -2\bar{\mathbf{H}}_2\mathbf{F}_2\bar{\mathbf{H}}_1\mathbf{F}_1\mathbf{F}_1^H\bar{\mathbf{H}}_1^H\mathbf{F}_2^H\bar{\mathbf{H}}_2^H
\end{aligned} \tag{70}$$

Applying Lemma 2, the third term in (70) can be written as

$$\begin{aligned}
&\mathrm{E}[\mathbf{H}_2\mathbf{F}_2\mathrm{E}[\mathbf{H}_1\mathbf{f}_{1,i}\mathbf{f}_{1,i}^H\mathbf{H}_1^H]\mathbf{F}_2^H\mathbf{H}_2^H] \\
&= \mathrm{E}[\mathbf{H}_2\mathbf{F}_2(\bar{\mathbf{H}}_1\mathbf{f}_{1,i}\mathbf{f}_{1,i}^H\bar{\mathbf{H}}_1^H + \mathrm{tr}(\mathbf{f}_{1,i}\mathbf{f}_{1,i}^H\mathbf{\Theta}_1^T)\mathbf{\Phi}_1)\mathbf{F}_2^H\mathbf{H}_2^H] \\
&= \bar{\mathbf{H}}_2\mathbf{F}_2\bar{\mathbf{H}}_1\mathbf{f}_{1,i}\mathbf{f}_{1,i}^H\bar{\mathbf{H}}_1^H\mathbf{F}_2^H\bar{\mathbf{H}}_2^H \\
&\quad +\mathrm{tr}(\mathbf{F}_2\bar{\mathbf{H}}_1\mathbf{f}_{1,i}\mathbf{f}_{1,i}^H\bar{\mathbf{H}}_1^H\mathbf{F}_2^H\mathbf{\Theta}_2^T)\mathbf{\Phi}_2 + \mathrm{tr}(\mathbf{f}_{1,i}\mathbf{f}_{1,i}^H\mathbf{\Theta}_1^T) \\
&\quad \times[\bar{\mathbf{H}}_2\mathbf{F}_2\mathbf{\Phi}_1\mathbf{F}_2^H\bar{\mathbf{H}}_2^H + \mathrm{tr}(\mathbf{F}_2\mathbf{\Phi}_1\mathbf{F}_2^H\mathbf{\Theta}_2^T)\mathbf{\Phi}_2].
\end{aligned} \tag{71}$$

By substituting (71) back into (70) we obtain

$$\begin{aligned}
&\mathrm{E}[\mathbf{H}_2\mathbf{F}_2\mathbf{H}_1\mathbf{F}_1\mathbf{F}_1^H\mathbf{H}_1^H\mathbf{F}_2^H\mathbf{H}_2^H] \\
&= \bar{\mathbf{H}}_2\mathbf{F}_2\bar{\mathbf{H}}_1\mathbf{F}_1\mathbf{F}_1^H\bar{\mathbf{H}}_1^H\mathbf{F}_2^H\bar{\mathbf{H}}_2^H \\
&\quad +\mathrm{tr}(\mathbf{F}_2\bar{\mathbf{H}}_1\mathbf{F}_1\mathbf{F}_1^H\bar{\mathbf{H}}_1^H\mathbf{F}_2^H\mathbf{\Theta}_2^T)\mathbf{\Phi}_2 + \mathrm{tr}(\mathbf{F}_1\mathbf{F}_1^H\mathbf{\Theta}_1^T) \\
&\quad \times[\bar{\mathbf{H}}_2\mathbf{F}_2\mathbf{\Phi}_1\mathbf{F}_2^H\bar{\mathbf{H}}_2^H + \mathrm{tr}(\mathbf{F}_2\mathbf{\Phi}_1\mathbf{F}_2^H\mathbf{\Theta}_2^T)\mathbf{\Phi}_2].
\end{aligned} \tag{72}$$

From Lemma 2, we know that

$$\mathrm{E}[\mathbf{H}_2\mathbf{F}_2\mathbf{F}_2^H\mathbf{H}_2^H] = \bar{\mathbf{H}}_2\mathbf{F}_2\mathbf{F}_2^H\bar{\mathbf{H}}_2^H + \mathrm{tr}(\mathbf{F}_2\mathbf{F}_2^H\mathbf{\Theta}_2^T)\mathbf{\Phi}_2. \tag{73}$$

Now using (72), (73), and $\mathrm{E}[\mathbf{H}_2\mathbf{F}_2\mathbf{H}_1\mathbf{F}_1] = \bar{\mathbf{H}}_2\mathbf{F}_2\bar{\mathbf{H}}_1\mathbf{F}_1$, we can write the expectation of $\mathbf{E}$ in (5) as (7).

Appendix B
Proof of Theorem 2

From (18), we can write

$$\begin{aligned}
\alpha_i\mathbf{\Phi}_i + \mathbf{I}_{N_{i+1}} &= \mathbf{U}_{\Phi_i}(\alpha_i\mathbf{\Lambda}_{\Phi_i} + \mathbf{I}_{N_{i+1}})\mathbf{U}_{\Phi_i}^H \\
&\triangleq \mathbf{U}_{\Phi_i}\tilde{\mathbf{\Lambda}}_{\Phi_i}\mathbf{U}_{\Phi_i}^H, \qquad i=1,2.
\end{aligned}$$

By using (20) and introducing

$$\tilde{\mathbf{F}}_2 \triangleq \mathbf{F}_2\mathbf{U}_{\Phi_1}\tilde{\mathbf{\Lambda}}_{\Phi_1}^{\frac{1}{2}}, \tag{74}$$

the problem (15)-(17) can be equivalently written as

$$\min_{\mathbf{F}_1,\tilde{\mathbf{F}}_2} \; q\Big(\mathbf{d}\big[\mathbf{I}_{N_b} - \mathbf{F}_1^H\tilde{\mathbf{H}}_1^H\tilde{\mathbf{F}}_2^H\tilde{\mathbf{H}}_2^H\tilde{\mathbf{A}}^{-1}\tilde{\mathbf{H}}_2\tilde{\mathbf{F}}_2\tilde{\mathbf{H}}_1\mathbf{F}_1\big]\Big) \tag{75}$$

$$\text{s.t.} \;\; \mathrm{tr}\Big(\tilde{\mathbf{F}}_2\big(\tilde{\mathbf{H}}_1\mathbf{F}_1\mathbf{F}_1^H\tilde{\mathbf{H}}_1^H + \mathbf{I}_{N_2}\big)\tilde{\mathbf{F}}_2^H\Big) \le P_2 \tag{76}$$

$$\mathrm{tr}\Big(\mathbf{F}_1\mathbf{F}_1^H\Big) \le P_1 \tag{77}$$

where

$$\tilde{\mathbf{A}} = \tilde{\mathbf{H}}_2\tilde{\mathbf{F}}_2\big(\tilde{\mathbf{H}}_1\mathbf{F}_1\mathbf{F}_1^H\tilde{\mathbf{H}}_1^H + \mathbf{I}_{N_2}\big)\tilde{\mathbf{F}}_2^H\tilde{\mathbf{H}}_2^H + \mathbf{I}_{N_3}. \tag{78}$$

It can be shown using Theorem 1 in [3] that the solution to the problem (75)-(77) for Schur-concave q is

$$\mathbf{F}_1 = \tilde{\mathbf{V}}_{1,1}\mathbf{\Lambda}_1, \quad \tilde{\mathbf{F}}_2 = \tilde{\mathbf{V}}_{2,1}\mathbf{\Lambda}_2\tilde{\mathbf{U}}_{1,1}^H. \tag{79}$$

While for Schur-convex q, the optimal $\tilde{\mathbf{F}}_2$ is given in (79), while the optimal $\mathbf{F}_1$ is $\mathbf{F}_1 = \tilde{\mathbf{V}}_{1,1}\mathbf{\Lambda}_1\mathbf{V}_0$. Substituting (79) back into (74), we obtain (21).

Appendix C
Proof of Theorem 3

Using (18)-(20) and (74), the problem (49)-(50) can be equivalently written as

$$\min_{\mathbf{F}_1,\tilde{\mathbf{F}}_2} \; \mathrm{tr}\Big(\tilde{\mathbf{F}}_2\big(\tilde{\mathbf{H}}_1\mathbf{F}_1\mathbf{F}_1^H\tilde{\mathbf{H}}_1^H + \mathbf{I}_{N_2}\big)\tilde{\mathbf{F}}_2^H + \mathbf{F}_1\mathbf{F}_1^H\Big) \tag{80}$$

$$\text{s.t.} \;\; \mathbf{d}\big[\mathbf{I}_{N_b} - \mathbf{F}_1^H\tilde{\mathbf{H}}_1^H\tilde{\mathbf{F}}_2^H\tilde{\mathbf{H}}_2^H\tilde{\mathbf{A}}^{-1}\tilde{\mathbf{H}}_2\tilde{\mathbf{F}}_2\tilde{\mathbf{H}}_1\mathbf{F}_1\big] \le \mathbf{g} \tag{81}$$

where $\tilde{\mathbf{A}}$ is given in (78). It can be shown using Theorem 1 in [6] that the solution to the problem (80)-(81) is given by

$$\mathbf{F}_1 = \tilde{\mathbf{V}}_{1,1}\mathbf{\Lambda}_1\mathbf{U}_0, \quad \tilde{\mathbf{F}}_2 = \tilde{\mathbf{V}}_{2,1}\mathbf{\Lambda}_2\tilde{\mathbf{U}}_{1,1}^H. \tag{82}$$

By substituting (82) back into (74) we obtain (51).

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] Y. Fan and J. Thompson, "MIMO configurations for relay channels: Theory and practice," *IEEE Trans. Wireless Commun.*, vol. 6, pp. 1774-1786, May 2007.

[3] Y. Rong, X. Tang, and Y. Hua, "A unified framework for optimizing linear non-regenerative multicarrier MIMO relay communication systems," *IEEE Trans. Signal Processing*, vol. 57, pp. 4837-4851, Dec. 2009.

[4] Y. Rong and Y. Hua, "Optimality of diagonalization of multi-hop MIMO relays," *IEEE Trans. Wireless Commun.*, in press.

[5] Y. Rong, "Optimal linear non-regenerative multi-hop MIMO relays with MMSE-DFE receiver at the destination," *IEEE Trans. Wireless Commun.*, revised, 2009.

[6] Y. Rong, "Multi-hop non-regenerative MIMO relays – QoS considerations," *IEEE Trans. Commun.*, submitted, 2009.

[7] D. S. Shiu, G. Foschini, M. Gans, and J. Kahn, "Fading correlation and its effect on the capacity of multielement antenna systems," *IEEE Trans. Commun.*, vol. 48, pp. 503-513, Mar. 2000.

[8] F. Rey, M. Lamarca, and G. Vazquez, "Robust power allocation algorithms for MIMO OFDM systems with imperfect CSI," *IEEE Trans. Signal Proessing*, vol. 53, pp. 1070-1085, Mar. 2005.

[9] X. Zhang, D. P. Palomar, and B. Ottersten, "Statistically robust design of linear MIMO transceivers," *IEEE Trans. Signal Processing*, vol. 56, pp. 3678-3689, Aug. 2008.

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

[11] J. W. Brewer, "Kronecker products and matrix calculus in system theory," *IEEE Trans. Circuits Syst.*, vol. 25, pp. 772-781, Sept. 1978.

[12] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory.* Englewood Cilffs, NJ: Prentice Hall, 1993.

[13] The MOSEK optimization toolbox for MATLAB manual. Version 5.0. [Online]. Available: http://www.mosek.com.

[14] P. H. M. Janssen and P. Stoica, "On the expectation of the product of four matrix-valued Gaussian random variables," *IEEE Trans. Automatic Control*, vol. 33, pp. 867-870, Sep. 1988.

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