This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the WCNC 2010 proceedings.

# Linear Transceiver Design for Amplify-and-Forward MIMO Relay Systems under Channel Uncertainties

Chengwen Xing†, Shaodan Ma†, Yik-Chung Wu†, Tung-Sang Ng† and H. Vincent Poor§
†Department of Electrical and Electronic Engineering, The University of Hong Kong, Hong Kong
§ Department of Electrical Engineering, Princeton University, Princeton, NJ 08544, USA
Email: †{cwxing, sdma, ycwu, tsng}@eee.hku.hk §poor@princeton.edu

***Abstract*—In this paper, robust joint design of linear relay precoders and destination equalizers for amplify-and-forward (AF) MIMO relay systems under Gaussian channel uncertainties is investigated. After incorporating the channel uncertainties into the robust design based on the Bayesian framework, a closed-form solution is derived to minimize the mean-square-error (MSE) of the received signal at the destination. The effectiveness of the proposed robust transceiver is verified by simulations.**

## I. Introduction

Due to its potential to improve reliability and coverage of wireless communication systems, cooperative communications has received considerable attention. In cooperative systems, relay nodes are deployed to offer cooperative diversity and facilitate communications between the source and destination [1], [2]. Generally, there are three kinds of relay strategies, including decode-and-forward (DF), compress-and-forward (CF) and amplify-and-forward (AF). In terms of implementation complexity, AF strategy is preferable, since for this strategy relay nodes simply amplify the received signal and then forward it to the destination.

On the other hand, it is well-known that multiple antennas can bring spacial diversity and multiplexing gains to communication systems. This kind of benefit can be directly introduced into cooperative communication systems by deploying multiple antennas at the transceivers. The resulting AF multiple-input multiple-output (MIMO) relay systems have attracted considerable research interest in recent years [3]–[10].

Transceiver design for AF MIMO relay systems to minimize the mean-square-error (MSE) of the received signal at the destination has been discussed in [9], [10]. Perfect channel state information (CSI) is usually assumed for the transceiver design. Unfortunately, in practical systems, due to the time varying nature of wireless channels and limited length of training sequences, channel estimation errors are inevitable [11]. Therefore, robust designs that can improve the performance of wireless systems by taking channel uncertainties into account are of interest.

In this paper, we propose a robust linear transceiver design for AF MIMO relay systems under channel uncertainties. The channel estimation errors are modeled as Gaussian random variables and incorporated into the design using a Bayesian framework. A closed-form solution is derived to minimize the MSE of the received signal at the destination. Simulation results verify the robustness of the proposed robust design against channel uncertainties, and show that it performs better than the corresponding algorithm without taking channel estimation errors into account.

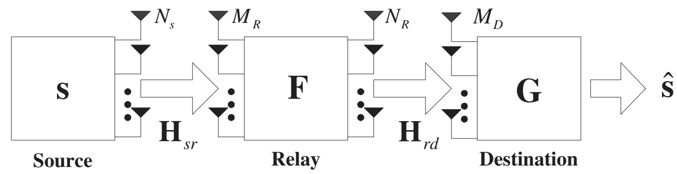


Fig. 1. Amplify-and-forward MIMO relay system.

The following notation is used throughout this paper. Boldface lowercase letters denote vectors, while boldface uppercase letters denote matrices. The notations $\mathbf{Z}^{\mathrm{T}}$ and $\mathbf{Z}^{\mathrm{H}}$ denote the transpose and Hermitian of the matrix $\mathbf{Z}$, respectively, and $\mathrm{Tr}(\mathbf{Z})$ is the trace of the matrix $\mathbf{Z}$. The symbol $\mathbf{I}_M$ denotes an $M\times M$ identity matrix, while $\mathbf{0}_{M,N}$ denotes an $M\times N$ all zero matrix. The notation $\mathbf{Z}^{\frac{1}{2}}$ is the Hermitian square root of the positive semidefinite matrix $\mathbf{Z}$, such that $\mathbf{Z}^{\frac{1}{2}}\mathbf{Z}^{\frac{1}{2}}=\mathbf{Z}$ and $\mathbf{Z}^{\frac{1}{2}}$ is also a Hermitian matrix. The operation $\mathrm{vec}(\mathbf{Z})$ stacks the columns of the matrix $\mathbf{Z}$ into a single column vector.

## II. System Model

In this paper, a dual-hop amplify-and-forward cooperative communication system is considered. In the considered system, there is one source with $N_S$ antennas, one relay with $M_R$ receive antennas and $N_R$ transmit antennas, and one destination with $M_D$ antennas, as shown in Fig. 1. At the first hop, the source transmits data to the relay. The received signal, $\mathbf{x}$, at the relay is

$$\mathbf{x}=\mathbf{H}_{sr}\mathbf{s}+\mathbf{n}_1, \tag{1}$$

where $\mathbf{s}$ is the data vector transmitted by the source with covariance matrix $\mathbf{R}_{\mathbf{s}}=\mathbb{E}\{\mathbf{s}\mathbf{s}^{\mathrm{H}}\}$. The matrix $\mathbf{H}_{sr}$ is the MIMO channel matrix between the source and the relay. Symbol $\mathbf{n}_1$ denotes zero-mean additive Gaussian noise with covariance matrix $\mathbf{R}_{n_1}=\sigma_{n_1}^2\mathbf{I}_{M_R}$. At the relay, the received signal $\mathbf{x}$ is multiplied by a precoder matrix $\mathbf{F}$, under a power constraint $\mathrm{Tr}(\mathbf{F}\mathbf{R}_{\mathbf{x}}\mathbf{F}^{\mathrm{H}})\leq P_r$ where $\mathbf{R}_{\mathbf{x}}=\mathbb{E}\{\mathbf{x}\mathbf{x}^{\mathrm{H}}\}$ and $P_r$ is the maximum transmit power. Then the resulting signal is transmitted to the destination. The received signal $\mathbf{y}$ at the destination can be written as

$$\mathbf{y}=\mathbf{H}_{rd}\mathbf{F}\mathbf{H}_{sr}\mathbf{s}+\mathbf{H}_{rd}\mathbf{F}\mathbf{n}_1+\mathbf{n}_2, \tag{2}$$

where $\mathbf{H}_{rd}$ is the MIMO channel matrix between the relay and the destination, and $\mathbf{n}_2$ denotes a zero-mean additive Gaussian noise vector at the second hop with covariance matrix $\mathbf{R}_{n_2} = \sigma_{n_2}^2 \mathbf{I}_{M_D}$. In order to guarantee that the transmitted data $\mathbf{s}$ can be recovered at the destination, it is assumed that $M_R$, $N_R$, and $M_D$ are greater than or equal to $N_S$ [9].

It is assumed that both the relay and destination have estimates of the channel state information. Thus, we can write

$$\begin{aligned} \mathbf{H}_{sr} &= \bar{\mathbf{H}}_{sr} + \Delta\mathbf{H}_{sr} \\ \text{and} \quad \mathbf{H}_{rd} &= \bar{\mathbf{H}}_{rd} + \Delta\mathbf{H}_{rd}, \end{aligned} \tag{3}$$

where the symbols $\bar{\mathbf{H}}_{sr}$ and $\bar{\mathbf{H}}_{rd}$ denote the estimated CSI matrices, while $\Delta\mathbf{H}_{sr}$ and $\Delta\mathbf{H}_{rd}$ denote the corresponding channel estimation error matrices whose elements are zero mean Gaussian random variables.

In general, the $M_R \times N_S$ matrix $\Delta\mathbf{H}_{sr}$ can be written as $\Delta\mathbf{H}_{sr} = \mathbf{\Sigma}_{sr}^{\frac{1}{2}} \mathbf{H}_W \mathbf{\Psi}_{sr}^{\frac{1}{2}}$ where the elements of the $M_R \times N_S$ matrix $\mathbf{H}_W$ are independent and identically distributed (i.i.d.) Gaussian random variables with zero means and unit variances [12]–[14]. The $M_R \times M_R$ matrix $\mathbf{\Sigma}_{sr}$ and $N_S \times N_S$ matrix $\mathbf{\Psi}_{sr}^{\mathrm{T}}$ are the row and column covariance matrices of $\Delta\mathbf{H}_{sr}$, respectively [15]. It is easy to see that $\mathrm{vec}(\Delta\mathbf{H}_{sr}^{\mathrm{T}}) \sim \mathcal{CN}(\mathbf{0}_{M_R N_S,1}, \mathbf{\Sigma}_{sr} \otimes \mathbf{\Psi}_{sr}^{\mathrm{T}})$ based on which $\Delta\mathbf{H}_{sr}$ is said to have a matrix-variate complex Gaussian distribution, which can be written as [16]

$$\Delta\mathbf{H}_{sr} \sim \mathcal{CN}_{M_R,N_S}(\mathbf{0}_{M_R,N_S}, \mathbf{\Sigma}_{sr} \otimes \mathbf{\Psi}_{sr}^{\mathrm{T}}), \tag{4}$$

with the probability density function (p.d.f.) given by [17]

$$f(\Delta\mathbf{H}_{sr}) = \frac{\exp(-\mathrm{Tr}(\Delta\mathbf{H}_{sr}^{\mathrm{H}} \mathbf{\Sigma}_{sr}^{-1} \Delta\mathbf{H}_{sr} \mathbf{\Psi}_{sr}^{-1}))}{(\pi)^{N_S M_R} \det(\mathbf{\Sigma}_{sr})^{N_S} \det(\mathbf{\Psi}_{sr})^{M_R}}. \tag{5}$$

Similarly, for the estimation error in the second hop, we have

$$\Delta\mathbf{H}_{rd} \sim \mathcal{CN}_{M_D,N_R}(\mathbf{0}_{M_D,N_R}, \mathbf{\Sigma}_{rd} \otimes \mathbf{\Psi}_{rd}^{\mathrm{T}}) \tag{6}$$

where the $M_D \times M_D$ matrix $\mathbf{\Sigma}_{rd}$ and $N_R \times N_R$ matrix $\mathbf{\Psi}_{rd}^{\mathrm{T}}$ are the row and column covariance matrices of $\Delta\mathbf{H}_{rd}$, respectively. It is assumed that the channel estimation errors, $\Delta\mathbf{H}_{sr}$ and $\Delta\mathbf{H}_{rd}$, are independent.

***Remark 1:*** In general, the expressions for $\mathbf{\Psi}_{sr}$, $\mathbf{\Sigma}_{sr}$, $\mathbf{\Psi}_{rd}$ and $\mathbf{\Sigma}_{rd}$ depend on specific channel estimation algorithms. If the channel estimation algorithm proposed in [12] is used, we have $\mathbf{\Psi}_{sr} = \mathbf{R}_{T,sr}$, $\mathbf{\Sigma}_{sr} = \sigma_{e,sr}^2 \mathbf{R}_{R,sr}$, $\mathbf{\Psi}_{rd} = \mathbf{R}_{T,rd}$ and $\mathbf{\Sigma}_{rd} = \sigma_{e,rd}^2 \mathbf{R}_{R,rd}$. The matrices $\mathbf{R}_{T,sr}$ and $\mathbf{R}_{R,sr}$ are the transmit and receive antenna correlation matrices at the source and the relay, respectively, and $\sigma_{e,sr}^2$ is the source-relay channel estimation error variance. Similarly, $\mathbf{R}_{T,rd}$, $\mathbf{R}_{R,rd}$ and $\sigma_{e,rd}^2$ are defined for the channel between the relay and the destination. On the other hand, when the channels are estimated based on the algorithm proposed in [13], we have $\mathbf{\Psi}_{sr} = \mathbf{R}_{T,sr}$, $\mathbf{\Sigma}_{sr} = \sigma_{e,sr}^2 (\mathbf{I}_{M_R} + \sigma_{e,sr}^2 \mathbf{R}_{R,sr}^{-1})^{-1}$, $\mathbf{\Psi}_{rd} = \mathbf{R}_{T,rd}$ and $\mathbf{\Sigma}_{rd} = \sigma_{e,rd}^2 (\mathbf{I}_{M_D} + \sigma_{e,rd}^2 \mathbf{R}_{R,rd}^{-1})^{-1}$. In the following, the proposed algorithm is developed without assuming any specific form of $\mathbf{\Psi}_{sr}$, $\mathbf{\Sigma}_{sr}$, $\mathbf{\Psi}_{rd}$ and $\mathbf{\Sigma}_{rd}$.

## III. Problem Formulation

At the destination, a linear equalizer $\mathbf{G}$ is adopted to detect the transmitted data $\mathbf{s}$. The problem is how to design the linear precoder matrix $\mathbf{F}$ at the relay and the linear equalizer $\mathbf{G}$ at the destination to minimize the MSE of the received data at the destination:

$$\mathrm{MSE}(\mathbf{F}, \mathbf{G}) = \mathbb{E}\{\mathrm{Tr}\left((\mathbf{G}\mathbf{y} - \mathbf{s})(\mathbf{G}\mathbf{y} - \mathbf{s})^{\mathrm{H}}\right)\}, \tag{7}$$

where the expectation is taken with respect to $\mathbf{s}$, $\Delta\mathbf{H}_{sr}$, $\Delta\mathbf{H}_{rd}$, $\mathbf{n}_1$ and $\mathbf{n}_2$. Since $\mathbf{s}$, $\mathbf{n}_1$ and $\mathbf{n}_2$ are independent, the MSE expression (7) can be written as

$$\begin{aligned} &\mathrm{MSE}(\mathbf{F}, \mathbf{G}) \\ &= \mathbb{E}\{\|(\mathbf{G}\mathbf{H}_{rd}\mathbf{F}\mathbf{H}_{sr} - \mathbf{I}_{N_S})\mathbf{s} + \mathbf{G}\mathbf{H}_{rd}\mathbf{F}\mathbf{n}_1 + \mathbf{G}\mathbf{n}_2\|^2\} \\ &= \mathbb{E}_{\Delta\mathbf{H}_{sr},\Delta\mathbf{H}_{rd}}\{\mathrm{Tr}((\mathbf{G}\mathbf{H}_{rd}\mathbf{F}\mathbf{H}_{sr} - \mathbf{I}_{N_S})\mathbf{R}_s \\ &\qquad\qquad \times (\mathbf{G}\mathbf{H}_{rd}\mathbf{F}\mathbf{H}_{sr} - \mathbf{I}_{N_S})^{\mathrm{H}})\} \\ &\quad + \mathbb{E}_{\Delta\mathbf{H}_{rd}}\{\mathrm{Tr}\left((\mathbf{G}\mathbf{H}_{rd}\mathbf{F})\mathbf{R}_{n_1}(\mathbf{G}\mathbf{H}_{rd}\mathbf{F})^{\mathrm{H}}\right)\} + \mathrm{Tr}(\mathbf{G}\mathbf{R}_{n_2}\mathbf{G}^{\mathrm{H}}) \\ &= \mathbb{E}_{\Delta\mathbf{H}_{sr},\Delta\mathbf{H}_{rd}}\{\mathrm{Tr}\left((\mathbf{G}\mathbf{H}_{rd}\mathbf{F}\mathbf{H}_{sr})\mathbf{R}_s(\mathbf{G}\mathbf{H}_{rd}\mathbf{F}\mathbf{H}_{sr})^{\mathrm{H}}\right)\} \\ &\quad + \mathrm{Tr}\left(\mathbf{G}\mathbb{E}_{\Delta\mathbf{H}_{rd}}\{\mathbf{H}_{rd}\mathbf{F}\mathbf{R}_{n_1}\mathbf{F}^{\mathrm{H}}\mathbf{H}_{rd}^{\mathrm{H}}\}\mathbf{G}^{\mathrm{H}}\right) \\ &\quad - \mathrm{Tr}\left(\mathbf{R}_s(\mathbf{G}\bar{\mathbf{H}}_{rd}\mathbf{F}\bar{\mathbf{H}}_{sr})^{\mathrm{H}}\right) - \mathrm{Tr}\left(\mathbf{G}\bar{\mathbf{H}}_{rd}\mathbf{F}\bar{\mathbf{H}}_{sr}\mathbf{R}_s\right) \\ &\quad + \mathrm{Tr}(\mathbf{R}_s) + \mathrm{Tr}(\mathbf{G}\mathbf{R}_{n_2}\mathbf{G}^{\mathrm{H}}). \end{aligned} \tag{8}$$

Because $\Delta\mathbf{H}_{sr}$ and $\Delta\mathbf{H}_{rd}$ are independent, the first term of MSE is

$$\begin{aligned} &\mathbb{E}_{\Delta\mathbf{H}_{sr},\Delta\mathbf{H}_{rd}}\{\mathrm{Tr}\left((\mathbf{G}\mathbf{H}_{rd}\mathbf{F}\mathbf{H}_{sr})\mathbf{R}_s(\mathbf{G}\mathbf{H}_{rd}\mathbf{F}\mathbf{H}_{sr})^{\mathrm{H}}\right)\} \\ &= \mathrm{Tr}\left(\mathbf{G}\mathbb{E}_{\Delta\mathbf{H}_{rd}}\left\{\mathbf{H}_{rd}\mathbf{F}\mathbb{E}_{\Delta\mathbf{H}_{sr}}\{\mathbf{H}_{sr}\mathbf{R}_s\mathbf{H}_{sr}^{\mathrm{H}}\}\mathbf{F}^{\mathrm{H}}\mathbf{H}_{rd}^{\mathrm{H}}\right\}\mathbf{G}^{\mathrm{H}}\right). \end{aligned} \tag{9}$$

For the inner expectation, due to the fact that the distribution of $\Delta\mathbf{H}_{sr}$ is a matrix-valued complex Gaussian with zero mean, the following equation holds [16]:

$$\begin{aligned} &\mathbb{E}_{\Delta\mathbf{H}_{sr}}\{\mathbf{H}_{sr}\mathbf{R}_s\mathbf{H}_{sr}^{\mathrm{H}}\} \\ &= \mathbb{E}_{\Delta\mathbf{H}_{sr}}\{(\bar{\mathbf{H}}_{sr} + \Delta\mathbf{H}_{sr})\mathbf{R}_s(\bar{\mathbf{H}}_{sr} + \Delta\mathbf{H}_{sr})^{\mathrm{H}}\} \\ &= \mathrm{Tr}(\mathbf{R}_s\mathbf{\Psi}_{sr})\mathbf{\Sigma}_{sr} + \bar{\mathbf{H}}_{sr}\mathbf{R}_s\bar{\mathbf{H}}_{sr}^{\mathrm{H}} \\ &\triangleq \mathbf{\Pi}_0. \end{aligned} \tag{10}$$

Applying (10) and the corresponding result for $\Delta\mathbf{H}_{rd}$ to (9), the first term of MSE becomes

$$\begin{aligned} &\mathrm{Tr}\left(\mathbf{G}\mathbb{E}_{\Delta\mathbf{H}_{rd}}\left\{\mathbf{H}_{rd}\mathbf{F}\mathbb{E}_{\Delta\mathbf{H}_{sr}}\{\mathbf{H}_{sr}\mathbf{R}_s\mathbf{H}_{sr}^{\mathrm{H}}\}\mathbf{F}^{\mathrm{H}}\mathbf{H}_{rd}^{\mathrm{H}}\right\}\mathbf{G}^{\mathrm{H}}\right) \\ &= \mathrm{Tr}(\mathbf{G}(\mathrm{Tr}(\mathbf{F}\mathbf{\Pi}_0\mathbf{F}^{\mathrm{H}}\mathbf{\Psi}_{rd})\mathbf{\Sigma}_{rd} + \bar{\mathbf{H}}_{rd}\mathbf{F}\mathbf{\Pi}_0\mathbf{F}^{\mathrm{H}}\bar{\mathbf{H}}_{rd}^{\mathrm{H}})\mathbf{G}^{\mathrm{H}}). \end{aligned} \tag{11}$$

Similarly, the second term of MSE in (8) can be simplified as

$$\begin{aligned} &\mathrm{Tr}\left(\mathbf{G}\mathbb{E}_{\Delta\mathbf{H}_{rd}}\{\mathbf{H}_{rd}\mathbf{F}\mathbf{R}_{n_1}\mathbf{F}^{\mathrm{H}}\mathbf{H}_{rd}^{\mathrm{H}}\}\mathbf{G}^{\mathrm{H}}\right) \\ &= \mathrm{Tr}(\mathbf{G}\left(\mathrm{Tr}(\mathbf{F}\mathbf{R}_{n_1}\mathbf{F}^{\mathrm{H}}\mathbf{\Psi}_{rd})\mathbf{\Sigma}_{rd} + \bar{\mathbf{H}}_{rd}\mathbf{F}\mathbf{R}_{n_1}\mathbf{F}^{\mathrm{H}}\bar{\mathbf{H}}_{rd}^{\mathrm{H}}\right)\mathbf{G}^{\mathrm{H}}). \end{aligned} \tag{12}$$

Based on (11) and (12), the MSE (8) equals

$$\begin{aligned} \mathrm{MSE}(\mathbf{F}, \mathbf{G}) = &\mathrm{Tr}\left(\mathbf{G}(\bar{\mathbf{H}}_{rd}\mathbf{F}\mathbf{R}_{\mathbf{x}}\mathbf{F}^{\mathrm{H}}\bar{\mathbf{H}}_{rd}^{\mathrm{H}} + \mathbf{K})\mathbf{G}^{\mathrm{H}}\right) \\ &- \mathrm{Tr}\left(\mathbf{R}_s\bar{\mathbf{H}}_{sr}^{\mathrm{H}}\mathbf{F}^{\mathrm{H}}\bar{\mathbf{H}}_{rd}^{\mathrm{H}}\mathbf{G}^{\mathrm{H}}\right) \\ &- \mathrm{Tr}\left(\mathbf{G}\bar{\mathbf{H}}_{rd}\mathbf{F}\bar{\mathbf{H}}_{sr}\mathbf{R}_s\right) + \mathrm{Tr}(\mathbf{R}_s) \end{aligned} \tag{13}$$

where

$$\mathbf{R_x} = \mathbf{\Pi}_0 + \mathbf{R}_{n_1}$$
$$\text{and} \quad \mathbf{K} = \text{Tr}(\mathbf{F}\mathbf{R_x}\mathbf{F}^{\text{H}}\mathbf{\Psi}_{rd})\mathbf{\Sigma}_{rd} + \mathbf{R}_{n_2}. \tag{14}$$

Notice that the matrix $\mathbf{R_x}$ is the autocorrelation matrix of the received signal at the relay.

Subject to the transmit power constraint at the relay, the joint design of the equalizer at the destination and the precoder at the relay can be formulated as the following optimization problem:

$$\min_{\mathbf{F},\mathbf{G}} \quad \text{MSE}(\mathbf{F},\mathbf{G})$$
$$\text{s.t.} \quad \text{Tr}(\mathbf{F}\mathbf{R_x}\mathbf{F}^{\text{H}}) \le P_r. \tag{15}$$

## IV. The Proposed Closed-Form Solution

It is difficult to find the optimal solution for the optimization problem (15), because $\text{MSE}(\mathbf{F},\mathbf{G})$ is a very complicated function of $\mathbf{F}$ and $\mathbf{G}$. However, from the definition of $\mathbf{K} = \text{Tr}(\mathbf{F}\mathbf{R_x}\mathbf{F}^{\text{H}}\mathbf{\Psi}_{rd})\mathbf{\Sigma}_{rd} + \mathbf{R}_{n_2}$, we have

$$\mathbf{K} \preceq \mathbf{K}_U \triangleq \underbrace{(\sigma_{n_2}^2 + \text{Tr}(\mathbf{F}\mathbf{R_x}\mathbf{F}^{\text{H}}\mathbf{\Psi}_{rd})\lambda_{\max}(\mathbf{\Sigma}_{rd}))}_{\triangleq \eta}\mathbf{I}_{M_D}, \tag{16}$$

where $\lambda_{\max}(\mathbf{Z})$ denotes the largest eigenvalue of $\mathbf{Z}$. It follows that

$$\mathbf{G}\mathbf{K}\mathbf{G}^{\text{H}} \preceq \mathbf{G}\mathbf{K}_U\mathbf{G}^{\text{H}}, \tag{17}$$

which implies

$$\text{Tr}(\mathbf{G}\mathbf{K}\mathbf{G}^{\text{H}}) \le \text{Tr}(\mathbf{G}\mathbf{K}_U\mathbf{G}^{\text{H}}). \tag{18}$$

Replacing $\mathbf{K}$ by $\mathbf{K}_U$, the corresponding MSE is

$$\begin{aligned} &\text{MSE}_U(\mathbf{F},\mathbf{G}) \\ &= \text{Tr}\left(\mathbf{G}(\bar{\mathbf{H}}_{rd}\mathbf{F}\mathbf{R_x}\mathbf{F}^{\text{H}}\bar{\mathbf{H}}_{rd}^{\text{H}} + \mathbf{K}_U)\mathbf{G}^{\text{H}}\right) + \text{Tr}(\mathbf{R_s}) \\ &- \text{Tr}\left(\mathbf{R_s}\bar{\mathbf{H}}_{sr}^{\text{H}}\mathbf{F}^{\text{H}}\bar{\mathbf{H}}_{rd}^{\text{H}}\mathbf{G}^{\text{H}}\right) - \text{Tr}\left(\mathbf{G}\bar{\mathbf{H}}_{rd}\mathbf{F}\bar{\mathbf{H}}_{sr}\mathbf{R_s}\right). \end{aligned} \tag{19}$$

It is obvious that $\text{MSE}_U$ is an upper-bound on MSE, i.e., $\text{MSE}(\mathbf{F},\mathbf{G}) \le \text{MSE}_U(\mathbf{F},\mathbf{G})$. Thus the optimization problem (15) can be relaxed to

$$\min_{\mathbf{F},\mathbf{G}} \quad \text{MSE}_U(\mathbf{F},\mathbf{G})$$
$$\text{s.t.} \quad \text{Tr}(\mathbf{F}\mathbf{R_x}\mathbf{F}^{\text{H}}) \le P_r. \tag{20}$$

Notice that when $\mathbf{\Sigma}_{rd} \propto \mathbf{I}_{M_D}$, $\text{MSE}(\mathbf{F},\mathbf{G}) = \text{MSE}_U(\mathbf{F},\mathbf{G})$, no relaxation is needed and the problem (20) is exactly equivalent to (15).

The corresponding Karush-Kuhn-Tucker (KKT) conditions for the optimization problem (20) are given as follows [18]:

$$\mathbf{G}(\bar{\mathbf{H}}_{rd}\mathbf{F}\mathbf{R_x}\mathbf{F}^{\text{H}}\bar{\mathbf{H}}_{rd}^{\text{H}} + \mathbf{K}_U) = \mathbf{R}_s(\bar{\mathbf{H}}_{rd}\mathbf{F}\bar{\mathbf{H}}_{sr})^{\text{H}}, \tag{21a}$$
$$\bar{\mathbf{H}}_{rd}^{\text{H}}\mathbf{G}^{\text{H}}\mathbf{G}\bar{\mathbf{H}}_{rd}\mathbf{F}\mathbf{R_x} + (\lambda_{\max}(\mathbf{\Sigma}_{rd})\text{Tr}(\mathbf{G}\mathbf{G}^{\text{H}})\mathbf{\Psi}_{rd} + \gamma)\mathbf{F}\mathbf{R_x} - \left(\bar{\mathbf{H}}_{sr}\mathbf{R_s}\mathbf{G}\bar{\mathbf{H}}_{rd}\right)^{\text{H}} = \mathbf{0}_{N_R,M_R}, \tag{21b}$$
$$\gamma(\text{Tr}(\mathbf{F}\mathbf{R_x}\mathbf{F}^{\text{H}}) - P_r) = 0 \quad \text{and} \quad \gamma \ge 0, \tag{21c}$$

where $\gamma$ is the Lagrange multiplier.

***Lemma 1:*** Based on the KKT conditions (21a)-(21c), the Lagrange multiplier satisfies

$$\gamma = \sigma_{n_2}^2 \frac{\text{Tr}(\mathbf{G}\mathbf{G}^{\text{H}})}{P_r}. \tag{22}$$

A proof is given in Appendix A.

Based on *Lemma 1*, the second KKT condition (21b) can be simplified as

$$\bar{\mathbf{H}}_{rd}^{\text{H}}\mathbf{G}^{\text{H}}\mathbf{G}\bar{\mathbf{H}}_{rd}\mathbf{F}\mathbf{R_x} + \underbrace{(P_r\mathbf{\Psi}_{rd}\lambda_{\max}(\mathbf{\Sigma}_{rd}) + \sigma_{n_2}^2\mathbf{I})}_{\triangleq \mathbf{M}} \times \mathbf{F}\mathbf{R_x}\frac{\gamma}{\sigma_{n_2}^2} - \left(\bar{\mathbf{H}}_{sr}\mathbf{R_s}\mathbf{G}\bar{\mathbf{H}}_{rd}\right)^{\text{H}} = \mathbf{0}. \tag{23}$$

Furthermore, we have the following lemma about the optimal precoder $\mathbf{F}$ and equalizer $\mathbf{G}$.

***Lemma 2:*** Based on the KKT conditions, the optimal precoder $\mathbf{F}$ and equalizer $\mathbf{G}$ for the optimization problem (20) are

$$\mathbf{F}_{\text{opt}} = \mathbf{M}^{-\frac{1}{2}}\mathbf{U}_{\mathbf{\Theta},N}\mathbf{\Lambda}_{\mathbf{F},\text{opt}}\mathbf{U}_{\mathbf{T},N}^{\text{H}}\mathbf{R_x}^{-\frac{1}{2}}, \tag{24}$$
$$\mathbf{G}_{\text{opt}} = \mathbf{V}_{\mathbf{T},N}\mathbf{\Lambda}_{\mathbf{G},\text{opt}}\mathbf{U}_{\mathbf{\Theta},N}^{\text{H}}\mathbf{M}^{-\frac{\text{H}}{2}}\bar{\mathbf{H}}_{rd}^{\text{H}}, \tag{25}$$
$$\mathbf{\Lambda}_{\mathbf{F},\text{opt}} = \left[\left(\sqrt{\frac{\sigma_{n_2}^2\eta}{\gamma}}\tilde{\mathbf{\Lambda}}_{\mathbf{\Theta}}^{-\frac{1}{2}}\tilde{\mathbf{\Lambda}}_{\mathbf{T}} - \eta\tilde{\mathbf{\Lambda}}_{\mathbf{\Theta}}^{-1}\right)^{+}\right]^{\frac{1}{2}} \quad \text{and} \tag{26}$$
$$\mathbf{\Lambda}_{\mathbf{G},\text{opt}} = \left[\left(\sqrt{\frac{\gamma}{\eta\sigma_{n_2}^2}}\tilde{\mathbf{\Lambda}}_{\mathbf{\Theta}}^{-\frac{1}{2}}\tilde{\mathbf{\Lambda}}_{\mathbf{T}} - \frac{\gamma}{\sigma_{n_2}^2}\tilde{\mathbf{\Lambda}}_{\mathbf{\Theta}}^{-1}\right)^{+}\right]^{\frac{1}{2}}\tilde{\mathbf{\Lambda}}_{\mathbf{\Theta}}^{-\frac{1}{2}}, \tag{27}$$

where $\mathbf{U}_{\mathbf{T},N}$, $\mathbf{V}_{\mathbf{T},N}$ and $\mathbf{U}_{\mathbf{\Theta},N}$ are the first $N$ columns of $\mathbf{U_T}$, $\mathbf{V_T}$ and $\mathbf{U_\Theta}$, respectively. The matrices $\tilde{\mathbf{\Lambda}}_{\mathbf{T}}$ and $\tilde{\mathbf{\Lambda}}_{\mathbf{\Theta}}$ are the principal sub-matrices of $\mathbf{\Lambda_T}$ and $\mathbf{\Lambda_\Theta}$ with dimension $N$, respectively. The matrices $\mathbf{U_T}$, $\mathbf{V_T}$, $\mathbf{\Lambda_T}$, $\mathbf{U_\Theta}$ and $\mathbf{\Lambda_\Theta}$, and the number $N$ are defined based on singular value decomposition as follows:

$$\mathbf{M}^{-\frac{\text{H}}{2}}\bar{\mathbf{H}}_{rd}^{\text{H}}\bar{\mathbf{H}}_{rd}\mathbf{M}^{-\frac{1}{2}} = \mathbf{U_\Theta}\mathbf{\Lambda_\Theta}\mathbf{U_\Theta}^{\text{H}}, \tag{28}$$
$$\mathbf{R_x}^{-\frac{1}{2}}\bar{\mathbf{H}}_{sr}\mathbf{R_s} = \mathbf{U_T}\mathbf{\Lambda_T}\mathbf{V_T}^{\text{H}} \tag{29}$$
$$\text{and} \quad N = \min\left(\text{rank}(\mathbf{\Lambda_T}), \text{rank}(\mathbf{\Lambda_\Theta})\right). \tag{30}$$

Without loss of generality, diagonal elements of the diagonal matrices $\mathbf{\Lambda_T}$ and $\mathbf{\Lambda_\Theta}$ are arranged in decreasing order.

A proof is given in Appendix B.

Based on *Lemma 2*, the remaining problem for finding the optimal $\mathbf{F}$ and $\mathbf{G}$ is to solve for the Lagrange multiplier $\gamma$ and the parameter $\eta$. From (21c) and (22), and together with the fact that $\mathbf{G}_{\text{opt}} \neq \mathbf{0}$, the optimal precoder and equalizer must satisfy the following two equations:

$$\text{Tr}(\mathbf{F}_{\text{opt}}\mathbf{R_x}\mathbf{F}_{\text{opt}}^{\text{H}}) = P_r \tag{31}$$
$$\text{and} \quad \text{Tr}(\mathbf{G}_{\text{opt}}\mathbf{G}_{\text{opt}}^{\text{H}}) = \gamma\frac{P_r}{\sigma_{n_2}^2}. \tag{32}$$

Substituting (24), (25), (26) and (27) into (31) and (32), $\eta$ and $\gamma$ can be found as

$$\eta = \frac{b_3P_r}{P_rb_1 + b_1b_4 - b_2b_3} \tag{33}$$
$$\text{and} \quad \gamma = \frac{b_3\sigma_{n_2}^2(P_rb_1 + b_1b_4 - b_2b_3)}{(P_r + b_4)^2P_r}, \tag{34}$$

where $b_1$, $b_2$, $b_3$ and $b_4$ are defined as

$$b_1 \triangleq \mathrm{Tr}(\mathbf{U}_{\mathbf{\Theta},L}^{\mathrm{H}}\mathbf{M}^{-1}\mathbf{U}_{\mathbf{\Theta},L}\breve{\mathbf{\Lambda}}_{\mathbf{T}}\breve{\mathbf{\Lambda}}_{\mathbf{\Theta}}^{-\frac{1}{2}}), \quad (35a)$$

$$b_2 \triangleq \mathrm{Tr}(\mathbf{U}_{\mathbf{\Theta},L}^{\mathrm{H}}\mathbf{M}^{-1}\mathbf{U}_{\mathbf{\Theta},L}\breve{\mathbf{\Lambda}}_{\mathbf{\Theta}}^{-1}), \quad (35b)$$

$$b_3 \triangleq \mathrm{Tr}(\breve{\mathbf{\Lambda}}_{\mathbf{T}}\breve{\mathbf{\Lambda}}_{\mathbf{\Theta}}^{-\frac{1}{2}}) \quad (35c)$$

$$\text{and} \quad b_4 \triangleq \mathrm{Tr}(\breve{\mathbf{\Lambda}}_{\mathbf{\Theta}}^{-1}). \quad (35d)$$

In (35), $\breve{\mathbf{\Lambda}}_{\mathbf{\Theta}}$ and $\breve{\mathbf{\Lambda}}_{\mathbf{T}}$ are the principal sub-matrices of $\tilde{\mathbf{\Lambda}}_{\mathbf{\Theta}}$ and $\tilde{\mathbf{\Lambda}}_{\mathbf{T}}$ with dimension $L$, respectively, and $L$ is the number of nonzero entries of $\mathbf{\Lambda}_{\mathbf{F},\mathrm{opt}}$, which can be computed easily by the algorithm proposed in [13]. Notice that when CSI is perfectly known, $\eta = \sigma_{n_2}^2$ and the proposed closed-form solution given by (24) and (25) is exactly the solution in [10].

## V. Simulation Results and Discussion

In this section, we investigate the performance of the proposed algorithm and for the purpose of comparison, the algorithm based on the estimated channel only (without taking the channel errors into account) [10] is also simulated. In the following, we consider an AF MIMO relay system in which the source, relay and destination are equipped with same number of antennas, i.e., $N_S = M_R = N_R = M_D = 4$. The widely used exponential model is chosen for both transmit and receive antenna correlation matrices [12]–[14]. More specifically, the channel correlation matrices are chosen as

$$\mathbf{R}_{T,sr} = \mathbf{R}_{T,rd} = \begin{bmatrix} 1 & \alpha & \alpha^2 & \alpha^3 \\ \alpha & 1 & \alpha & \alpha^2 \\ \alpha^2 & \alpha & 1 & \alpha \\ \alpha^3 & \alpha^2 & \alpha & 1 \end{bmatrix}$$

$$\text{and} \quad \mathbf{R}_{R,sr} = \mathbf{R}_{R,rd} = \begin{bmatrix} 1 & \beta & \beta^2 & \beta^3 \\ \beta & 1 & \beta & \beta^2 \\ \beta^2 & \beta & 1 & \beta \\ \beta^3 & \beta^2 & \beta & 1 \end{bmatrix}, \quad (36)$$

where $\alpha$ and $\beta$ are the correlation coefficients.

Here the channel estimation algorithm in [13] is adopted, and the correlation matrices of channel estimation errors are in the form:

$$\begin{aligned} \mathbf{\Psi}_{sr} &= \mathbf{R}_{T,sr}, \\ \mathbf{\Sigma}_{sr} &= \sigma_e^2(\mathbf{I}_{M_R} + \sigma_e^2\mathbf{R}_{R,sr}^{-1})^{-1}, \\ \mathbf{\Psi}_{rd} &= \mathbf{R}_{T,rd} \\ \text{and} \quad \mathbf{\Sigma}_{rd} &= \sigma_e^2(\mathbf{I}_{M_D} + \sigma_e^2\mathbf{R}_{R,rd}^{-1})^{-1}, \end{aligned} \quad (37)$$

where $\sigma_e^2$ is the variance of the channel estimation errors.

The signal-to-noise ratio for the source-relay link ($\mathrm{SNR}_{sr}$) is defined as $\mathrm{E}_s/\mathrm{N}_1 = \mathrm{Tr}(\mathbf{R}_s)/\mathrm{Tr}(\mathbf{R}_{n_1})$, and is fixed at 30dB. At the source, four independent data streams are transmitted by four antennas at the same power. For each data stream, $N_{Data} = 10000$ independent quadrature phase-shift keying(QPSK) symbols are transmitted and $\mathrm{Tr}(\mathbf{R}_s)$ is normalized to 1. Similarly, the SNR for the relay-destination link ($\mathrm{SNR}_{rd}$) is defined as $\mathrm{E}_r/\mathrm{N}_2 = P_r/\mathrm{Tr}(\mathbf{R}_{n_2})$. Each point in the following figure is an average of 10,000 independent trials.

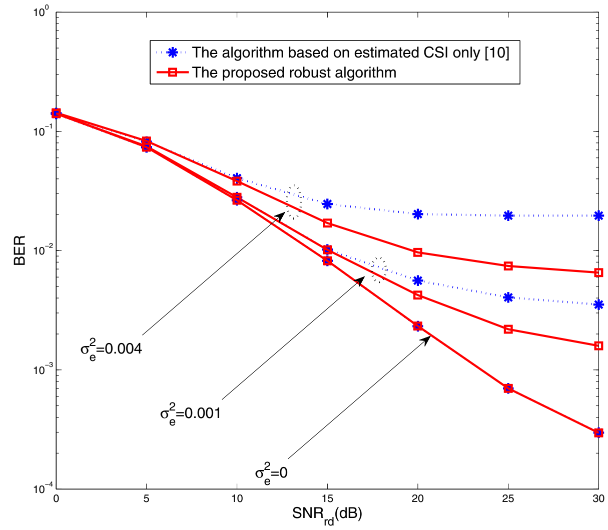


Fig. 2. BERs of the proposed robust algorithm and the algorithm based on estimated channel state information only, where $\alpha = 0.6$ and $\beta = 0.4$.

Fig. 2 shows the bit-error-rate (BER) performance of the proposed algorithm and the algorithm based on estimated channels only with different values of $\sigma_e^2$, when $\alpha = 0.6$ and $\beta = 0.4$. It can be seen that as the channel errors decrease, the performance of both algorithms improves. Moreover, since the proposed algorithm has taken the channel estimation errors into account, its performance is always better than that of the algorithm based on estimated channels only, when $\sigma_e^2 \neq 0$.

## VI. Conclusions

In this paper, a robust linear transceiver has been designed for dual-hop AF MIMO relay systems under channel uncertainties based on a minimum mean-square error criterion. The channel estimation errors are modeled to be Gaussian random variables and incorporated into the robust transceiver design based on a Bayesian framework. A closed-form solution has been derived and when $\mathbf{\Sigma}_{rd} \propto \mathbf{I}_{M_D}$, the proposed closed-form solution is exactly the optimal solution. It has been demonstrated by computer simulations that our proposed algorithm performs better than an algorithm based on estimated channels only.

## Appendix A
## Proof of Lemma 1

Right-multiplying both sides of (21a) by $\mathbf{G}^{\mathrm{H}}$, the following equality holds:

$$\mathbf{G}(\bar{\mathbf{H}}_{rd}\mathbf{F}\mathbf{R}_{\mathbf{x}}\mathbf{F}^{\mathrm{H}}\bar{\mathbf{H}}_{rd}^{\mathrm{H}} + \mathbf{K}_U)\mathbf{G}^{\mathrm{H}} = \mathbf{R}_s(\bar{\mathbf{H}}_{rd}\mathbf{F}\bar{\mathbf{H}}_{sr})^{\mathrm{H}}\mathbf{G}^{\mathrm{H}}. \quad (38)$$

Left-multiplying (21b) by $\mathbf{F}^{\mathrm{H}}$, we have

$$\begin{aligned} &\mathbf{F}^{\mathrm{H}}\bar{\mathbf{H}}_{rd}^{\mathrm{H}}\mathbf{G}^{\mathrm{H}}\mathbf{G}\bar{\mathbf{H}}_{rd}\mathbf{F}\mathbf{R}_{\mathbf{x}} + \mathbf{F}^{\mathrm{H}}\lambda_{\max}(\mathbf{\Sigma}_{rd})\mathrm{Tr}(\mathbf{G}\mathbf{G}^{\mathrm{H}}) \\ &\mathbf{\Psi}_{rd}\mathbf{F}\mathbf{R}_{\mathbf{x}} + \gamma\mathbf{F}^{\mathrm{H}}\mathbf{F}\mathbf{R}_{\mathbf{x}} = \mathbf{F}^{\mathrm{H}}\left(\bar{\mathbf{H}}_{sr}\mathbf{R}_s\mathbf{G}\bar{\mathbf{H}}_{rd}\right)^{\mathrm{H}}. \end{aligned} \quad (39)$$

After taking the traces of both sides of (38) and (39) and using the fact that the traces of their righthand sides are equivalent, i.e.,

$$\mathrm{Tr}(\mathbf{R}_s(\bar{\mathbf{H}}_{rd}\mathbf{F}\bar{\mathbf{H}}_{sr})^{\mathrm{H}}\mathbf{G}^{\mathrm{H}}) = \mathrm{Tr}(\mathbf{F}^{\mathrm{H}}(\bar{\mathbf{H}}_{sr}\mathbf{R}_s\mathbf{G}\bar{\mathbf{H}}_{rd})^{\mathrm{H}}), \quad (40)$$

we have

$$\begin{aligned}&\mathrm{Tr}(\mathbf{G}(\bar{\mathbf{H}}_{rd}\mathbf{F}\mathbf{R}_{\mathbf{x}}\mathbf{F}^{\mathrm{H}}\bar{\mathbf{H}}_{rd}^{\mathrm{H}} + \mathbf{K}_U)\mathbf{G}^{\mathrm{H}})\\ &= \mathrm{Tr}(\mathbf{F}^{\mathrm{H}}\bar{\mathbf{H}}_{rd}^{\mathrm{H}}\mathbf{G}^{\mathrm{H}}\mathbf{G}\bar{\mathbf{H}}_{rd}\mathbf{F}\mathbf{R}_{\mathbf{x}}) + \gamma\mathrm{Tr}(\mathbf{F}^{\mathrm{H}}\mathbf{F}\mathbf{R}_{\mathbf{x}})\\ &+ \lambda_{\max}(\mathbf{\Sigma}_{rd})\mathrm{Tr}(\mathbf{G}\mathbf{G}^{\mathrm{H}})\mathrm{Tr}(\mathbf{F}^{\mathrm{H}}\mathbf{\Psi}_{rd}\mathbf{F}\mathbf{R}_{\mathbf{x}}). \quad (41)\end{aligned}$$

Since the following equation always holds:

$$\mathrm{Tr}(\mathbf{G}(\bar{\mathbf{H}}_{rd}\mathbf{F}\mathbf{R}_{\mathbf{x}}\mathbf{F}^{\mathrm{H}}\bar{\mathbf{H}}_{rd}^{\mathrm{H}})\mathbf{G}^{\mathrm{H}}) = \mathrm{Tr}(\mathbf{F}^{\mathrm{H}}\bar{\mathbf{H}}_{rd}^{\mathrm{H}}\mathbf{G}^{\mathrm{H}}\mathbf{G}\bar{\mathbf{H}}_{rd}\mathbf{F}\mathbf{R}_{\mathbf{x}}),$$

(41) reduces to

$$\begin{aligned}&\mathrm{Tr}(\mathbf{G}\mathbf{K}_U\mathbf{G}^{\mathrm{H}})\\ &= \lambda_{\max}(\mathbf{\Sigma}_{rd})\mathrm{Tr}(\mathbf{G}\mathbf{G}^{\mathrm{H}})\mathrm{Tr}(\mathbf{F}^{\mathrm{H}}\mathbf{\Psi}_{rd}\mathbf{F}\mathbf{R}_{\mathbf{x}}) + \gamma\mathrm{Tr}(\mathbf{F}^{\mathrm{H}}\mathbf{F}\mathbf{R}_{\mathbf{x}}).\end{aligned} \quad (42)$$

On the other hand, based on the definition of $\mathbf{K}_U$ in (16) we have

$$\begin{aligned}&\mathrm{Tr}(\mathbf{G}\mathbf{K}_U\mathbf{G}^{\mathrm{H}})\\ &= \lambda_{\max}(\mathbf{\Sigma}_{rd})\mathrm{Tr}(\mathbf{G}\mathbf{G}^{\mathrm{H}})\mathrm{Tr}(\mathbf{F}^{\mathrm{H}}\mathbf{\Psi}_{rd}\mathbf{F}\mathbf{R}_{\mathbf{x}}) + \sigma_{n_2}^2\mathrm{Tr}(\mathbf{G}\mathbf{G}^{\mathrm{H}}).\end{aligned} \quad (43)$$

Comparing (42) with (43), it can be concluded that

$$\sigma_{n_2}^2\mathrm{Tr}(\mathbf{G}\mathbf{G}^{\mathrm{H}}) = \gamma\mathrm{Tr}(\mathbf{F}\mathbf{R}_{\mathbf{x}}\mathbf{F}^{\mathrm{H}}). \quad (44)$$

Furthermore, based on (21c) we have

$$\sigma_{n_2}^2\mathrm{Tr}(\mathbf{G}\mathbf{G}^{\mathrm{H}}) - \gamma P_r = \gamma(\mathrm{Tr}(\mathbf{F}\mathbf{R}_{\mathbf{x}}\mathbf{F}^{\mathrm{H}}) - P_r) = 0, \quad (45)$$

based on which the following equality holds:

$$\gamma = \sigma_{n_2}^2\frac{\mathrm{Tr}(\mathbf{G}\mathbf{G}^{\mathrm{H}})}{P_r}. \quad (46)$$

## Appendix B
## Proof of Lemma 2

Given an arbitrary $N_R \times M_R$ matrix $\mathbf{F}$, based on the second KKT condition, it can be reformulated as

$$\mathbf{F} = \mathbf{M}^{-\frac{1}{2}}\mathbf{U}_{\mathbf{\Theta},N}\mathbf{\Lambda}_{\mathbf{F}}\mathbf{U}_{\mathbf{T},N}^{\mathrm{H}}\mathbf{R}_{\mathbf{x}}^{-\frac{1}{2}}. \quad (47)$$

Substituting (47) into (21a), the equalizer $\mathbf{G}$ can be derived as

$$\begin{aligned}\mathbf{G} &= \mathbf{R}_s(\bar{\mathbf{H}}_{rd}\mathbf{F}\bar{\mathbf{H}}_{sr})^{\mathrm{H}}(\bar{\mathbf{H}}_{rd}\mathbf{F}\mathbf{R}_{\mathbf{x}}\mathbf{F}^{\mathrm{H}}\bar{\mathbf{H}}_{rd}^{\mathrm{H}} + \eta\mathbf{I}_{M_D})^{-1}\\ &= (\mathbf{R}_x^{-\frac{1}{2}}\bar{\mathbf{H}}_{sr}\mathbf{R}_s)^{\mathrm{H}}(\mathbf{R}_{\mathbf{x}}^{\frac{1}{2}}\mathbf{F}^{\mathrm{H}}\bar{\mathbf{H}}_{rd}^{\mathrm{H}}\bar{\mathbf{H}}_{rd}\mathbf{F}\mathbf{R}_{\mathbf{x}}^{\frac{1}{2}} + \eta\mathbf{I}_{M_R})^{-1}\\ &\qquad \times \mathbf{R}_{\mathbf{x}}^{\frac{1}{2}}\mathbf{F}^{\mathrm{H}}\bar{\mathbf{H}}_{rd}^{\mathrm{H}}\\ &= \mathbf{V}_{\mathbf{T},N}\underbrace{\tilde{\mathbf{\Lambda}}_{\mathbf{T}}^{\mathrm{H}}(\mathbf{\Lambda}_{\mathbf{F}}^{\mathrm{H}}\tilde{\mathbf{\Lambda}}_{\mathbf{\Theta}}\mathbf{\Lambda}_{\mathbf{F}} + \eta\mathbf{I}_N)^{-1}\mathbf{\Lambda}_{\mathbf{F}}^{\mathrm{H}}}_{\triangleq\mathbf{\Lambda}_{\mathbf{G}}}\mathbf{U}_{\mathbf{\Theta},N}^{\mathrm{H}}\mathbf{M}^{-\frac{\mathrm{H}}{2}}\bar{\mathbf{H}}_{rd}^{\mathrm{H}},\end{aligned} \quad (48)$$

where the second equality comes from the matrix inversion lemma.

Substituting (47) and (48) into (38) and (39), after a tedious derivation, we have

$$\mathbf{\Lambda}_{\mathbf{G}}\tilde{\mathbf{\Lambda}}_{\mathbf{\Theta}}\mathbf{\Lambda}_{\mathbf{F}}\mathbf{\Lambda}_{\mathbf{F}}^{\mathrm{H}}\tilde{\mathbf{\Lambda}}_{\mathbf{\Theta}}\mathbf{\Lambda}_{\mathbf{G}}^{\mathrm{H}} + \eta\mathbf{\Lambda}_{\mathbf{G}}\tilde{\mathbf{\Lambda}}_{\mathbf{\Theta}}\mathbf{\Lambda}_{\mathbf{G}}^{\mathrm{H}} = (\mathbf{\Lambda}_{\mathbf{G}}\tilde{\mathbf{\Lambda}}_{\mathbf{\Theta}}\mathbf{\Lambda}_{\mathbf{F}}\tilde{\mathbf{\Lambda}}_{\mathbf{T}})^{\mathrm{H}} \quad (49)$$

and

$$\mathbf{\Lambda}_{\mathbf{F}}^{\mathrm{H}}\tilde{\mathbf{\Lambda}}_{\mathbf{\Theta}}\mathbf{\Lambda}_{\mathbf{G}}^{\mathrm{H}}\mathbf{\Lambda}_{\mathbf{G}}\tilde{\mathbf{\Lambda}}_{\mathbf{\Theta}}\mathbf{\Lambda}_{\mathbf{F}} + \frac{\gamma}{\sigma_{n_2}^2}\mathbf{\Lambda}_{\mathbf{F}}^{\mathrm{H}}\mathbf{\Lambda}_{\mathbf{F}} = (\tilde{\mathbf{\Lambda}}_{\mathbf{T}}\mathbf{\Lambda}_{\mathbf{G}}\tilde{\mathbf{\Lambda}}_{\mathbf{\Theta}}\mathbf{\Lambda}_{\mathbf{F}})^{\mathrm{H}}. \quad (50)$$

Because the left hand side of (49) is a Hermitian matrix, the matrix $\mathbf{\Lambda}_{\mathbf{G}}\tilde{\mathbf{\Lambda}}_{\mathbf{\Theta}}\mathbf{\Lambda}_{\mathbf{F}}\tilde{\mathbf{\Lambda}}_{\mathbf{T}}$ on the right hand side must also be a Hermitian matrix. Similarly, $\tilde{\mathbf{\Lambda}}_{\mathbf{T}}\mathbf{\Lambda}_{\mathbf{G}}\tilde{\mathbf{\Lambda}}_{\mathbf{\Theta}}\mathbf{\Lambda}_{\mathbf{F}}$ is also a Hermitian matrix. Together with the fact that the matrix $\tilde{\mathbf{\Lambda}}_{\mathbf{T}}$ is real diagonal, $\mathbf{\Lambda}_{\mathbf{G}}\tilde{\mathbf{\Lambda}}_{\mathbf{\Theta}}\mathbf{\Lambda}_{\mathbf{F}}$ must also be diagonal [19]. Applying these results into (50), it follows that $\mathbf{\Lambda}_{\mathbf{F}}^{\mathrm{H}}\mathbf{\Lambda}_{\mathbf{F}}$ is also diagonal.

On the other hand, from the definition of $\mathbf{\Lambda}_{\mathbf{G}}$ in (48), we have

$$\mathbf{\Lambda}_{\mathbf{G}}\tilde{\mathbf{\Lambda}}_{\mathbf{\Theta}}\mathbf{\Lambda}_{\mathbf{F}} = \tilde{\mathbf{\Lambda}}_{\mathbf{T}}^{\mathrm{H}}(\mathbf{\Lambda}_{\mathbf{F}}^{\mathrm{H}}\tilde{\mathbf{\Lambda}}_{\mathbf{\Theta}}\mathbf{\Lambda}_{\mathbf{F}} + \eta\mathbf{I})^{-1}\mathbf{\Lambda}_{\mathbf{F}}^{\mathrm{H}}\tilde{\mathbf{\Lambda}}_{\mathbf{\Theta}}\mathbf{\Lambda}_{\mathbf{F}}. \quad (51)$$

As $\tilde{\mathbf{\Lambda}}_{\mathbf{T}}$ and $\mathbf{\Lambda}_{\mathbf{G}}\tilde{\mathbf{\Lambda}}_{\mathbf{\Theta}}\mathbf{\Lambda}_{\mathbf{F}}$ are diagonal, we can conclude that $\mathbf{\Lambda}_{\mathbf{F}}^{\mathrm{H}}\tilde{\mathbf{\Lambda}}_{\mathbf{\Theta}}\mathbf{\Lambda}_{\mathbf{F}}$ must be diagonal. Then, similarly to [19], from the diagonality of $\mathbf{\Lambda}_{\mathbf{F}}^{\mathrm{H}}\mathbf{\Lambda}_{\mathbf{F}}$ and $\mathbf{\Lambda}_{\mathbf{F}}^{\mathrm{H}}\tilde{\mathbf{\Lambda}}_{\mathbf{\Theta}}\mathbf{\Lambda}_{\mathbf{F}}$, it can be concluded that

$$\mathbf{\Lambda}_{\mathbf{F}} = \mathbf{P}\tilde{\mathbf{\Lambda}}_{\mathbf{F}}, \quad (52)$$

where $\tilde{\mathbf{\Lambda}}_{\mathbf{F}}$ is a diagonal matrix and $\mathbf{P}$ is a permutation matrix [20]. Substituting (52) into the definition of $\mathbf{\Lambda}_{\mathbf{G}}$ in (48), it can be derived that

$$\mathbf{\Lambda}_{\mathbf{G}} = \tilde{\mathbf{\Lambda}}_{\mathbf{G}}\mathbf{P}^{\mathrm{H}}, \quad (53)$$

where $\tilde{\mathbf{\Lambda}}_{\mathbf{G}}$ is also a diagonal matrix.

Putting (52) and (53) into (49) and (50), we can solve $\tilde{\mathbf{\Lambda}}_{\mathbf{F}}$ and $\tilde{\mathbf{\Lambda}}_{\mathbf{G}}$ as

$$\tilde{\mathbf{\Lambda}}_{\mathbf{F}} = \left[\left(\sqrt{\frac{\eta\sigma_{n_2}^2}{\gamma}}\tilde{\mathbf{\Lambda}}_{\mathbf{T}}(\mathbf{P}^{\mathrm{H}}\tilde{\mathbf{\Lambda}}_{\mathbf{\Theta}}\mathbf{P})^{-\frac{1}{2}} - \eta(\mathbf{P}^{\mathrm{H}}\tilde{\mathbf{\Lambda}}_{\mathbf{\Theta}}\mathbf{P})^{-1}\right)^{+}\right]^{\frac{1}{2}} \quad (54)$$

and

$$\begin{aligned}\tilde{\mathbf{\Lambda}}_{\mathbf{G}} = &\left[\left(\sqrt{\frac{\gamma}{\eta\sigma_{n_2}^2}}\tilde{\mathbf{\Lambda}}_{\mathbf{T}}(\mathbf{P}^{\mathrm{H}}\tilde{\mathbf{\Lambda}}_{\mathbf{\Theta}}\mathbf{P})^{-\frac{1}{2}} - \frac{\gamma}{\sigma_{n_2}^2}(\mathbf{P}^{\mathrm{H}}\tilde{\mathbf{\Lambda}}_{\mathbf{\Theta}}\mathbf{P})^{-1}\right)^{+}\right]^{\frac{1}{2}}\\ &\times(\mathbf{P}^{\mathrm{H}}\tilde{\mathbf{\Lambda}}_{\mathbf{\Theta}}\mathbf{P})^{-\frac{1}{2}}.\end{aligned} \quad (55)$$

Notice that as $\mathbf{P}$ is a permutation matrix , $\mathbf{P}^{\mathrm{H}}\tilde{\mathbf{\Lambda}}_{\mathbf{\Theta}}\mathbf{P}$ is also a diagonal matrix with a different ordering of the diagonal elements of $\tilde{\mathbf{\Lambda}}_{\mathbf{\Theta}}$ [20], and $\mathrm{Tr}((\mathbf{P}^{\mathrm{H}}\tilde{\mathbf{\Lambda}}_{\mathbf{\Theta}}\mathbf{P})^{-1}) = \mathrm{Tr}(\tilde{\mathbf{\Lambda}}_{\mathbf{\Theta}}^{-1})$.

In order to identify $\mathbf{P}$ for the optimal precoder and equalizer, it is assumed that all the eigenchannels in (54) and (55) are allocated nonzero power. Substituting (47), (48), and (52)-(55) into (19), together with (46) after a tedious derivation, we have

$$\mathrm{MSE}_U = \frac{\left[\mathrm{Tr}\left(\tilde{\mathbf{\Lambda}}_{\mathbf{T}}(\mathbf{P}^{\mathrm{H}}\tilde{\mathbf{\Lambda}}_{\mathbf{\Theta}}\mathbf{P})^{-\frac{1}{2}}\right)\right]^2}{\mathrm{Tr}\left(\tilde{\mathbf{\Lambda}}_{\mathbf{\Theta}}^{-1}\right) + P_r} + c, \quad (56)$$

where $c$ is a constant (i.e., independent of $\mathbf{P}$). Since $\tilde{\mathbf{\Lambda}}_{\mathbf{\Theta}}$ and $\tilde{\mathbf{\Lambda}}_{\mathbf{T}}$ are diagonal matrices with diagonal elements in decreasing order, and $\mathbf{P}^{\mathrm{H}}\tilde{\mathbf{\Lambda}}_{\mathbf{\Theta}}\mathbf{P}$ is also a diagonal matrix, the following inequality holds [6]:

$$\mathrm{MSE}_U \geq \frac{\left[\mathrm{Tr}\left(\tilde{\mathbf{\Lambda}}_{\mathbf{T}}\tilde{\mathbf{\Lambda}}_{\mathbf{\Theta}}^{-\frac{1}{2}}\right)\right]^2}{\mathrm{Tr}\left(\tilde{\mathbf{\Lambda}}_{\mathbf{\Theta}}^{-1}\right)+P_r} + c, \tag{57}$$

where the equality holds when $\mathbf{P}=\mathbf{I}_N$. It follows from (52)-(55) that for the minimal $\mathrm{MSE}_U$, $\mathbf{P}=\mathbf{I}_N$ and the optimal $\mathbf{\Lambda}_{\mathbf{F}}$ and $\mathbf{\Lambda}_{\mathbf{G}}$ are

$$\mathbf{\Lambda}_{\mathbf{F},\mathrm{opt}} = \left[\left(\sqrt{\frac{\sigma_{n_2}^2\eta}{\gamma}}\tilde{\mathbf{\Lambda}}_{\mathbf{\Theta}}^{-\frac{1}{2}}\tilde{\mathbf{\Lambda}}_{\mathbf{T}} - \eta\tilde{\mathbf{\Lambda}}_{\mathbf{\Theta}}^{-1}\right)^+\right]^{\frac{1}{2}} \tag{58}$$

and

$$\mathbf{\Lambda}_{\mathbf{G},\mathrm{opt}} = \left[\left(\sqrt{\frac{\gamma}{\eta\sigma_{n_2}^2}}\tilde{\mathbf{\Lambda}}_{\mathbf{\Theta}}^{-\frac{1}{2}}\tilde{\mathbf{\Lambda}}_{\mathbf{T}} - \frac{\gamma}{\sigma_{n_2}^2}\tilde{\mathbf{\Lambda}}_{\mathbf{\Theta}}^{-1}\right)^+\right]^{\frac{1}{2}}\tilde{\mathbf{\Lambda}}_{\mathbf{\Theta}}^{-\frac{1}{2}}. \tag{59}$$

Furthermore, the optimal precoder $\mathbf{F}_{\mathrm{opt}}$ and equalizer $\mathbf{G}_{\mathrm{opt}}$ are

$$\mathbf{F}_{\mathrm{opt}} = \mathbf{M}^{-\frac{1}{2}}\mathbf{U}_{\mathbf{\Theta},N}\mathbf{\Lambda}_{\mathbf{F},\mathrm{opt}}\mathbf{U}_{\mathbf{T},N}^{\mathrm{H}}\mathbf{R}_{\mathbf{x}}^{-\frac{1}{2}} \tag{60}$$

$$\text{and} \quad \mathbf{G}_{\mathrm{opt}} = \mathbf{V}_{\mathbf{T},N}\mathbf{\Lambda}_{\mathbf{G},\mathrm{opt}}\mathbf{U}_{\mathbf{\Theta},N}^{\mathrm{H}}\mathbf{M}^{-\frac{\mathrm{H}}{2}}\bar{\mathbf{H}}_{rd}^{\mathrm{H}}. \tag{61}$$

## ACKNOWLEDGEMENT

This work was supported by the Hong Kong Research Grants Council (Grant No. 7154/08E).

## REFERENCES

[1] A. Scaglione, D. L. Goeckel, and J. N. Laneman, "Cooperative communications in mobile ad hoc networks," *IEEE Signal Processing Magazine*, vol. 23, no. 5, pp. 18–29, Sep. 2006.

[2] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Information Theory*, vol. 50, no. 12 pp. 3062–3080, Dec. 2004.

[3] H. Bölcskei, R. U. Nabar, O. Oyman, and A. J. Paulraj, "Capacity scaling laws in MIMO relay networks," *IEEE Trans. Wireless Commun.*, vol. 5, no. 6, pp. 1433–1443, June 2006.

[4] C.-B. Chae, T. W. Tang, R. W. Health, and S.-Y. Cho, "MIMO relaying with linear processing for multiuser transmission in fixed relay networks," *IEEE Trans. Signal Processing*, vol. 56, no. 2, pp. 727–738, Feb. 2008.

[5] B. Wang, J. Zhang, and A. H$\phi$st-Madsen, "On the capacity of MIMO relay channels," *IEEE Trans. Information Theory*, vol. 51, no. 1, pp. 29–43, Jan. 2005.

[6] 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.

[7] O. Munoz-Medina, J. Vidal, and A. Agustin, "Linear transceiver design in nonregenerative relays with channel state information," *IEEE Trans. Signal Processing*, vol. 55, no. 6, pp. 2198–2206, June 2007.

[8] H. W. Je, B. Lee, S. Kim, and K. B. Lee, "Design of non-regenerative MIMO-relay systems with partial channel state information," *in Proc. IEEE Int'l Conf. Commun. 2008*, Beijing, China, May 2008.

[9] A. S. Behbahani, R. Merched, and A. J. Eltawil, "Optimizations of a MIMO relay network, " *IEEE Trans. Signal Processing*, vol. 56, no. 10, part 2, pp. 5062–5073, Oct. 2008.

[10] W. Guan and H. W. Luo, "Joint MMSE transceiver design in non-regenerative MIMO relay systems," *IEEE Commun. Letters*, vol. 12, issue 7, pp. 517–519, July 2008.

[11] X. Y. Wang and J. Wang, "Effect of imperfect channel estimation on transmit diversity in CDMA systems," *IEEE Trans. Veh. Technol.*, vol. 53, no. 5, pp. 1400-1412, Sept. 2004.

[12] L. Musavian, M. R. Nakhi, M. Dohler, and A. H. Aghvami, "Effect of channel uncertainty on the mutual information of MIMO fading channels, " *IEEE Trans. Veh. Technol.*, vol. 56, no. 5, pp. 2798–2806, Sep. 2007.

[13] M. Ding and S. D. Blostein, "MIMO minimum total MSE transceiver design with imperfect CSI at both ends," *IEEE Trans. Signal Processing*, vol. 57, no. 3, pp. 1141–1150, March 2009.

[14] X. Zhang, D. P. Palomar, and B. Ottersten, "Statistically robust design of linear MIMO transceiver," *IEEE Trans. Signal Processing*, vol. 56, no. 8, pp. 3678–3689, Aug. 2008.

[15] E. G. Larsson and P. Stoica, *Space-Time Block Coding for Wireless Communications*. Cambridge, UK: Cambridge University Press, 2003.

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

[17] A. T. James, "Dstributions of matrix variates and latent roots derived from normal samples," *Ann. Math. Stat.*, vol. 35, pp. 475–501, 1964.

[18] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, UK: Cambridge University Press, 2004.

[19] H. Sampath, P. Stoica, and A. Paulraj, "Generalized linear precoder and decoder design for MIMO channels using the weighted MMSE criterion," *IEEE Trans. Commun.*, vol. 49, no. 12, pp. 2198–2206, Dec. 2001.

[20] T. K. Moon and W. C. Stirling, *Mathematical Methods and Algorithms for Signal Processing*. Upper Saddle River, N.J.: Prentice-Hall, 2000.