

Bayesian Robust Linear Transceiver Design for Dual-Hop Amplify-and-Forward MIMO Relay Systems

Chengwen Xing, Shaodan Ma and Yik-Chung Wu
 Department of Electrical and Electronic Engineering
 The University of Hong Kong, Hong Kong
 Email: {cwxing, sdma, ycwu}@eee.hku.hk

Abstract—In this paper, we address the robust linear transceiver design for dual-hop amplify-and-forward (AF) MIMO relay systems, where both transmitters and receivers have imperfect channel state information (CSI). With the statistics of channel estimation errors in the two hops being Gaussian, we formulate the robust linear-minimum-mean-square-error (LMMSE) transceiver design problem using the Bayesian framework, and derive a closed-form solution. Simulation results show that the proposed algorithm reduces the sensitivity of the relay system to channel estimation errors, and performs better than the algorithm using estimated channel only.

I. INTRODUCTION

Recently, cooperative communication has gained significant interest, due to its great potentials to improve reliability, coverage and capacity of wireless links [1] [2]. Generally speaking, there are three kinds of relay protocols, amplify-and-forward (AF), compress-and-forward (CF) and decode-and-forward (DF). Among the three schemes, AF is conceptually the simplest one, in which the relay just scales the signal transmitted from the source, and then transmits to the destination. Due to its simplicity and low implementation complexity, AF strategy has received many researchers' attention.

On the other hand, it is well-known that in fully scattered environments, multiantenna systems provide spatial diversity and multiplexing gains. This kind of benefits can be directly introduced into cooperative communications via deployment of multiple antennas at transmitters and receivers. The combination of AF and MIMO systems brings great potentials in performance improvement.

Linear transceiver design for AF MIMO relay systems has been addressed in [2], [3], [4] and [5]. The capacity scaling law of MIMO relay networks has been discussed in [2]. The linear transceiver design for the fixed relay in cellular networks has been addressed in [3]. Joint linear-minimum-mean-square-error (LMMSE) transceiver design for AF MIMO relay systems is considered in [4] and [5]. However, all of the above mentioned works require the channel state information (CSI) perfectly known at the transmitters and receivers.

Unfortunately, in practical systems, channel estimation errors are inevitable, which should be taken into account in transceiver design. In this paper, we propose a robust linear transceiver design method for AF MIMO relay systems. The

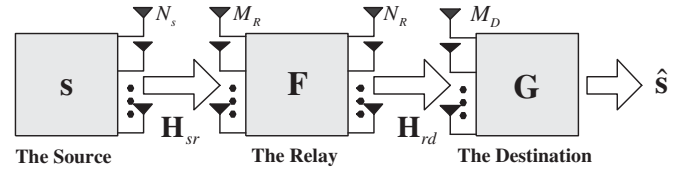


Fig. 1. Amplify-and-forward MIMO relay diagram

channel estimation errors are modeled as Gaussian random variables. The statistics of the channel estimation errors are incorporated into the design using the Bayesian framework, and a closed-form solution is obtained. Simulation results show that the proposed algorithm performs better than the algorithm using estimated channel only.

The following notations are used throughout this paper. Boldface lowercase letters denote vectors, while boldface uppercase letters denote matrices. The notation \mathbf{Z}^H denotes the Hermitian of the matrix \mathbf{Z} , and $\text{Tr}(\mathbf{Z})$ 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.

II. PROBLEM FORMULATION

In this paper, a dual-hop amplify-and-forward (AF) 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

$$\mathbf{x} = \mathbf{H}_{sr}\mathbf{s} + \mathbf{n}_1 \quad (1)$$

where \mathbf{s} is the data vector transmitted by the source with the covariance matrix $\mathbf{R}_s = \mathbb{E}\{\mathbf{s}\mathbf{s}^H\}$. The matrix \mathbf{H}_{sr} is the MIMO channel matrix between the source and the relay. Symbol \mathbf{n}_1 is the additive Gaussian noise with covariance matrix \mathbf{R}_{n_1} . At the relay, the received signal \mathbf{x} is multiplied by a precoder matrix \mathbf{F} , under a power constraint $\text{Tr}(\mathbf{F}\mathbf{R}_x\mathbf{F}^H) \leq$

P where $\mathbf{R}_x = \mathbb{E}\{\mathbf{x}\mathbf{x}^H\}$ and P is the maximum transmit power. Then the resulting signal is transmitted to the destination. The received signal at the destination, \mathbf{y} , can be written as

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

where \mathbf{H}_{rd} is the MIMO channel matrix between the relay and the destination, and \mathbf{n}_2 is the additive Gaussian noise vector at the second hop with covariance matrix \mathbf{R}_{n_2} . In order to guarantee 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 [4].

It is assumed that both the relay and destination have the estimated channel state information (CSI). When channel estimation errors are considered, we have

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

where the symbols $\bar{\mathbf{H}}_{sr}$ and $\bar{\mathbf{H}}_{rd}$ are the estimated CSI, while $\Delta\mathbf{H}_{sr}$ and $\Delta\mathbf{H}_{rd}$ are the corresponding channel estimation errors 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 $\Delta\mathbf{H}_{sr} = \Sigma_{sr}^{\frac{1}{2}} \mathbf{H}_W \Psi_{sr}^{\frac{1}{2}}$ 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 mean and unit variance. The $M_R \times M_R$ matrix Σ_{sr} and $N_S \times N_S$ matrix Ψ_{sr}^T are the row and column covariance matrices of $\Delta\mathbf{H}_{sr}$, respectively [6]. It is easy to see that $\text{vec}(\Delta\mathbf{H}_{sr}^T) \sim \mathcal{CN}(\mathbf{0}_{M_R \times N_S}, \Sigma_{sr} \otimes \Psi_{sr}^T)$ based on which $\Delta\mathbf{H}_{sr}$ is said to have a matrix-variate complex Gaussian distribution, which can be written as [7]

$$\Delta\mathbf{H}_{sr} \sim \mathcal{CN}_{M_R, N_S}(\mathbf{0}_{M_R, N_S}, \Sigma_{sr} \otimes \Psi_{sr}^T), \quad (4)$$

with the probability density function (p.d.f.) given by [8] [9]

$$f(\Delta\mathbf{H}_{sr}) = \frac{\exp(-\text{Tr}((\Delta\mathbf{H}_{sr} - \mathbf{0})^H \Sigma_{sr}^{-1} (\Delta\mathbf{H}_{sr} - \mathbf{0}) \Psi_{sr}^{-1}))}{(\pi)^{N_S M_R} \det(\Sigma_{sr})^{N_S} \det(\Psi_{sr})^{M_R}}. \quad (5)$$

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

$$\Delta\mathbf{H}_{rd} \sim \mathcal{CN}_{M_D, N_R}(\mathbf{0}_{M_D, N_R}, \Sigma_{rd} \otimes \Psi_{rd}^T) \quad (6)$$

where the $M_D \times M_D$ matrix Σ_{rd} and $N_R \times N_R$ matrix Ψ_{rd}^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. 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 mean square errors (MSE) of the received data at the destination:

$$\text{MSE}(\mathbf{F}, \mathbf{G}) = \mathbb{E}\{\text{Tr}((\mathbf{G}\mathbf{y} - \mathbf{s})(\mathbf{G}\mathbf{y} - \mathbf{s})^H)\}, \quad (7)$$

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 .

III. ROBUST TRANSCEIVER DESIGN FOR MIMO RELAY

A. MSE Averaged over Channel Uncertainties

Since \mathbf{s} , \mathbf{n}_1 and \mathbf{n}_2 are independent, the MSE expression (7) can be written as

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

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

$$\begin{aligned} \mathbb{E}_{\Delta\mathbf{H}_{sr}, \Delta\mathbf{H}_{rd}}\{\text{Tr}((\mathbf{G}\mathbf{H}_{rd}\mathbf{F}\mathbf{H}_{sr})\mathbf{R}_s(\mathbf{G}\mathbf{H}_{rd}\mathbf{F}\mathbf{H}_{sr})^H)\} \\ = \text{Tr}(\mathbf{G}\mathbb{E}_{\Delta\mathbf{H}_{rd}}\{\mathbf{H}_{rd}\mathbf{F}\mathbb{E}_{\Delta\mathbf{H}_{sr}}\{\mathbf{H}_{sr}\mathbf{R}_s\mathbf{H}_{sr}^H\}\mathbf{F}^H\mathbf{H}_{rd}^H\}\mathbf{G}^H). \end{aligned} \quad (9)$$

For the inner expectation, due to the fact that the distribution of $\Delta\mathbf{H}_{sr}$ is matrix-variate complex Gaussian with zero mean, the following equation holds [7]

$$\begin{aligned} \mathbb{E}_{\Delta\mathbf{H}_{sr}}\{\mathbf{H}_{sr}\mathbf{R}_s\mathbf{H}_{sr}^H\} &= \mathbb{E}_{\Delta\mathbf{H}_{sr}}\{(\bar{\mathbf{H}}_{sr} + \Delta\mathbf{H}_{sr})\mathbf{R}_s(\bar{\mathbf{H}}_{sr} + \Delta\mathbf{H}_{sr})^H\} \\ &= \text{Tr}(\mathbf{R}_s \Psi_{sr}) \Sigma_{sr} + \bar{\mathbf{H}}_{sr} \mathbf{R}_s \bar{\mathbf{H}}_{sr}^H \\ &\triangleq \Pi_0. \end{aligned} \quad (10)$$

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

$$\begin{aligned} \text{Tr}(\mathbf{G}\mathbb{E}_{\Delta\mathbf{H}_{rd}}\{\mathbf{H}_{rd}\mathbf{F}\mathbb{E}_{\Delta\mathbf{H}_{sr}}\{\mathbf{H}_{sr}\mathbf{R}_s\mathbf{H}_{sr}^H\}\mathbf{F}^H\mathbf{H}_{rd}^H\}\mathbf{G}^H) \\ = \text{Tr}(\mathbf{G}(\text{Tr}(\mathbf{F}\Pi_0\mathbf{F}^H\Psi_{rd})\Sigma_{rd} + \bar{\mathbf{H}}_{rd}\mathbf{F}\Pi_0\mathbf{F}^H\bar{\mathbf{H}}_{rd}^H)\mathbf{G}^H). \end{aligned} \quad (11)$$

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

$$\begin{aligned} \text{Tr}(\mathbf{G}\mathbb{E}_{\Delta\mathbf{H}_{rd}}\{\mathbf{H}_{rd}\mathbf{F}\mathbf{R}_{n_1}\mathbf{F}^H\mathbf{H}_{rd}^H\}\mathbf{G}^H) \\ = \text{Tr}(\mathbf{G}(\text{Tr}(\mathbf{F}\mathbf{R}_{n_1}\mathbf{F}^H\Psi_{rd})\Sigma_{rd} + \bar{\mathbf{H}}_{rd}\mathbf{F}\mathbf{R}_{n_1}\mathbf{F}^H\bar{\mathbf{H}}_{rd}^H)\mathbf{G}^H). \end{aligned} \quad (12)$$

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

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

where

$$\mathbf{R}_x = \Pi_0 + \mathbf{R}_{n_1} \quad (14)$$

$$\mathbf{K} = \text{Tr}(\mathbf{F}(\Pi_0 + \mathbf{R}_{n_1})\mathbf{F}^H\Psi_{rd})\Sigma_{rd} + \mathbf{R}_{n_2}. \quad (15)$$

Notice that the matrix \mathbf{R}_x is the autocorrelation matrix of the receive signal \mathbf{x} at the relay.

B. Joint Robust Design of Equalizer and Precoder

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

$$\begin{aligned} \min_{\mathbf{F}, \mathbf{G}} \quad & \text{MSE}(\mathbf{F}, \mathbf{G}) \\ \text{s.t.} \quad & \text{Tr}(\mathbf{F}\mathbf{R}_x\mathbf{F}^H) \leq P. \end{aligned} \quad (16)$$

Since the constraint does not involve the equalizer \mathbf{G} , when the precoder matrix \mathbf{F} is fixed, the optimal linear equalizer, \mathbf{G}_{opt} , satisfies the following condition

$$\frac{\partial \text{MSE}(\mathbf{F}, \mathbf{G})}{\partial \mathbf{G}^*} = 0, \quad (17)$$

based on which we have

$$\mathbf{G}_{\text{opt}} = \mathbf{R}_s(\bar{\mathbf{H}}_{rd}\mathbf{F}\bar{\mathbf{H}}_{sr})^H(\bar{\mathbf{H}}_{rd}\mathbf{F}\mathbf{R}_x\mathbf{F}^H\bar{\mathbf{H}}_{rd}^H + \mathbf{K})^{-1}. \quad (18)$$

Substituting (18) into (13), the MSE at the destination equals to

$$\overline{\text{MSE}}(\mathbf{F}) = \text{Tr}(\mathbf{R}_s) - \text{Tr}(\mathbf{R}_s\bar{\mathbf{H}}_{sr}^H[\mathbf{F}^H\bar{\mathbf{H}}_{rd}^H(\bar{\mathbf{H}}_{rd}\mathbf{F}\mathbf{R}_x\mathbf{F}^H\bar{\mathbf{H}}_{rd}^H + \mathbf{K})^{-1}\bar{\mathbf{H}}_{rd}\mathbf{F}]\bar{\mathbf{H}}_{sr}\mathbf{R}_s). \quad (19)$$

Since $\mathbf{K}^{\frac{1}{2}}$ and $\mathbf{R}_x^{\frac{1}{2}}$ are both Hermitian matrices, exploiting the matrix inversion lemma, we have

$$\begin{aligned} & \mathbf{F}^H\bar{\mathbf{H}}_{rd}^H(\bar{\mathbf{H}}_{rd}\mathbf{F}\mathbf{R}_x\mathbf{F}^H\bar{\mathbf{H}}_{rd}^H + \mathbf{K})^{-1}\bar{\mathbf{H}}_{rd}\mathbf{F} \\ &= \mathbf{R}_x^{-\frac{H}{2}}\mathbf{R}_x^{\frac{H}{2}}\mathbf{F}^H\bar{\mathbf{H}}_{rd}^H\mathbf{K}^{-\frac{H}{2}}(\mathbf{K}^{-\frac{1}{2}}\bar{\mathbf{H}}_{rd}\mathbf{F}\mathbf{R}_x^{\frac{1}{2}}\mathbf{R}_x^{\frac{H}{2}} \\ & \mathbf{F}^H\bar{\mathbf{H}}_{rd}^H\mathbf{K}^{-\frac{H}{2}} + \mathbf{I}_{M_D})^{-1}\mathbf{K}^{-\frac{1}{2}}\bar{\mathbf{H}}_{rd}\mathbf{F}\mathbf{R}_x^{\frac{1}{2}}\mathbf{R}_x^{-\frac{1}{2}} \\ &= \mathbf{R}_x^{-1} - \mathbf{R}_x^{-\frac{H}{2}}(\mathbf{R}_x^{\frac{H}{2}}\mathbf{F}^H\bar{\mathbf{H}}_{rd}^H\mathbf{K}^{-1}\bar{\mathbf{H}}_{rd}\mathbf{F}\mathbf{R}_x^{\frac{1}{2}} \\ & + \mathbf{I}_{M_R})^{-1}\mathbf{R}_x^{-\frac{1}{2}}. \end{aligned} \quad (20)$$

Putting (20) into (19), and defining the constant part $c \triangleq \text{Tr}(\mathbf{R}_s) - \text{Tr}(\mathbf{R}_s\bar{\mathbf{H}}_{sr}^H\mathbf{R}_x^{-1}\bar{\mathbf{H}}_{sr}\mathbf{R}_s)$, equation (19) can be rewritten as

$$\overline{\text{MSE}}(\mathbf{F}) = \text{Tr}(\mathbf{R}_s\bar{\mathbf{H}}_{sr}^H\mathbf{R}_x^{-\frac{H}{2}}(\mathbf{R}_x^{\frac{H}{2}}\mathbf{F}^H\bar{\mathbf{H}}_{rd}^H\mathbf{K}^{-1}\bar{\mathbf{H}}_{rd}\mathbf{F}\mathbf{R}_x^{\frac{1}{2}} + \mathbf{I}_{M_R})^{-1}\mathbf{R}_x^{-\frac{1}{2}}\bar{\mathbf{H}}_{sr}\mathbf{R}_s) + c. \quad (21)$$

From (15), $\mathbf{K} = \text{Tr}(\mathbf{F}\mathbf{R}_x\mathbf{F}^H\Psi_{rd})\Sigma_{rd} + \mathbf{R}_{n_2}$, so $\overline{\text{MSE}}(\mathbf{F})$ is a high order function of \mathbf{F} and the problem of minimizing (21) is very difficult to solve. In order to proceed, notice that [10]

$$\text{Tr}(\mathbf{F}\mathbf{R}_x\mathbf{F}^H)\lambda_{\max}(\Psi_{rd})\Sigma_{rd} + \mathbf{R}_{n_2} \succeq \mathbf{K}, \quad (22)$$

where $\lambda_{\max}(\mathbf{Z})$ denotes the largest eigenvalue of \mathbf{Z} . Since for the minimum MSE, the corresponding transmit power must be on the boundary (i.e., $\text{Tr}(\mathbf{F}\mathbf{R}_x\mathbf{F}^H) = P$), we have $P\lambda_{\max}(\Psi_{rd})\Sigma_{rd} + \mathbf{R}_{n_2} \succeq \mathbf{K}$. As shown in Appendix I, when \mathbf{K} in (21) is replaced by its upper-bound $\Phi \triangleq P\lambda_{\max}(\Psi_{rd})\Sigma_{rd} + \mathbf{R}_{n_2}$, the resultant MSE expression, denoted as $\overline{\text{MSE}}_U(\mathbf{F})$ and defined in (43), is an upper-bound of $\overline{\text{MSE}}(\mathbf{F})$ (i.e., $\overline{\text{MSE}}_U(\mathbf{F}) \geq \overline{\text{MSE}}(\mathbf{F})$). Therefore, we

propose to design the precoder \mathbf{F} by minimizing $\overline{\text{MSE}}_U(\mathbf{F})$, which corresponds to

$$\begin{aligned} \min_{\mathbf{F}} \quad & \text{Tr}(\underbrace{\mathbf{R}_x^{-\frac{1}{2}}\bar{\mathbf{H}}_{sr}\mathbf{R}_s\mathbf{R}_s\bar{\mathbf{H}}_{sr}^H\mathbf{R}_x^{-\frac{H}{2}}}_{\triangleq \mathbf{T}}(\mathbf{R}_x^{\frac{H}{2}}\mathbf{F}^H \\ & \underbrace{\bar{\mathbf{H}}_{rd}^H\Phi^{-1}\bar{\mathbf{H}}_{rd}}_{\triangleq \Theta}\mathbf{F}\mathbf{R}_x^{\frac{1}{2}} + \mathbf{I}_{M_R})^{-1}) \\ \text{s.t.} \quad & \text{Tr}(\mathbf{F}\mathbf{R}_x\mathbf{F}^H) \leq P \end{aligned} \quad (23)$$

where the constant c is neglected, which does not affect the optimization problem. Notice that when $\Psi_{rd} \propto \mathbf{I}$, the replacement involves no approximation.

Based on eigendecomposition, we have

$$\mathbf{T} = \mathbf{U}_T\Lambda_T\mathbf{U}_T^H, \quad (24)$$

$$\Theta = \mathbf{U}_\Theta\Lambda_\Theta\mathbf{U}_\Theta^H, \quad (25)$$

where the matrices \mathbf{U}_T and \mathbf{U}_Θ consist of the eigenvectors of \mathbf{T} and Θ , respectively, while the diagonal matrices Λ_T and Λ_Θ contains the eigenvalues of \mathbf{T} and Θ , respectively. Without loss of generality, it is assumed that the diagonal elements of Λ_T and Λ_Θ are in decreasing order. Substituting (24) and (25) into (23) and defining

$$\tilde{\mathbf{F}} \triangleq \mathbf{U}_\Theta^H\mathbf{F}\mathbf{R}_x^{\frac{1}{2}}\mathbf{U}_T, \quad (26)$$

the optimization problem can be written in a compact form as

$$\begin{aligned} \min_{\tilde{\mathbf{F}}} \quad & \text{Tr}(\Lambda_T(\tilde{\mathbf{F}}^H\Lambda_\Theta\tilde{\mathbf{F}} + \mathbf{I}_{M_R})^{-1}) \\ \text{s.t.} \quad & \text{Tr}(\tilde{\mathbf{F}}\tilde{\mathbf{F}}^H) \leq P. \end{aligned} \quad (27)$$

For the objective function of (27), notice that

$$\text{Tr}(\Lambda_T\mathbf{B}) \geq \sum_{i=1}^{M_R} \lambda_{T,i}\lambda_{B,M_R-i+1} \quad (28)$$

where \mathbf{B} is defined as $\mathbf{B} \triangleq (\tilde{\mathbf{F}}^H\Lambda_\Theta\tilde{\mathbf{F}} + \mathbf{I}_{M_R})^{-1}$, $\lambda_{B,i}$ is the i^{th} largest eigenvalue of \mathbf{B} , and the symbol $\lambda_{T,i}$ denotes the i^{th} diagonal element of Λ_T . In (28), the equality holds when the matrix \mathbf{B} is diagonal with diagonal elements in increasing order [11, 9.H.1.h]. Therefore, for the optimal solution, $(\tilde{\mathbf{F}}^H\Lambda_\Theta\tilde{\mathbf{F}} + \mathbf{I}_{M_R})^{-1}$ must be diagonal with diagonal elements in increasing order.

Based on the diagonal structure, introducing a permutation matrix with dimension $M \times M$ as

$$\mathbf{Q}_M = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{1} \\ \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{1} & \mathbf{0} & \mathbf{0} \end{bmatrix}_M, \quad (29)$$

the objective function of the optimization problem (27) can be rewritten as

$$\begin{aligned} & \text{Tr}(\Lambda_T(\tilde{\mathbf{F}}^H\Lambda_\Theta\tilde{\mathbf{F}} + \mathbf{I}_{M_R})^{-1}) \\ &= \text{Tr}(\Lambda'_T(\mathcal{F}^H\Lambda'_\Theta\mathcal{F} + \mathbf{I}_{M_R})^{-1}) \end{aligned} \quad (30)$$

where $\Lambda'_T = \mathbf{Q}_{M_R}\Lambda_T\mathbf{Q}_{M_R}$ and $\Lambda'_\Theta = \mathbf{Q}_{N_R}\Lambda_\Theta\mathbf{Q}_{N_R}$ are Λ_T and Λ_Θ with diagonal elements in reverse order, and $\mathcal{F} =$

$\mathbf{Q}_{N_R} \tilde{\mathbf{F}} \mathbf{Q}_{M_R}$. With the fact that $\mathbf{\Lambda}_T$ is a diagonal matrix, the optimization problem (27) can be reformulated as

$$\begin{aligned} \min \quad & f_0(\mathbf{b}) = \mathbf{d}^T \{ \underbrace{\mathbf{\Lambda}'_T \{ (\mathcal{F}^H \mathbf{\Lambda}'_{\Theta} \mathcal{F} + \mathbf{I}_{M_R})^{-1} \}}_{\triangleq \mathbf{b}} \} \\ \text{s.t.} \quad & \text{Tr}(\mathcal{F}^H \mathcal{F}) \leq P, \end{aligned} \quad (31)$$

where the symbol $\mathbf{d}\{\mathbf{Z}\}$ denotes the vector formed from the main diagonal of \mathbf{Z} . Notice that because of the permutation matrices, the order of \mathbf{b} is the reverse to that of the main diagonal of \mathbf{B} (i.e., the elements of \mathbf{b} are in decreasing order). Together with the fact that the diagonal elements of $\mathbf{\Lambda}'_T$ are in increasing order, the function $f_0(\mathbf{b})$ is Schur-concave [11, 3.H.3]. Based on [12, Theorem 1], the optimal $\mathcal{F} = \mathbf{Q}_{N_R} \tilde{\mathbf{F}} \mathbf{Q}_{M_R}$ for the problem (31) has zero elements except along the rightmost main diagonal. Defining $N = \min(\text{Rank}(\mathbf{\Lambda}_{\Theta}), M_R)$, the optimal $\tilde{\mathbf{F}}$ has the following structure

$$\tilde{\mathbf{F}} = \begin{bmatrix} \text{diag}(f_1, \dots, f_N) & \mathbf{0}_{N \times (M_R - N)} \\ \mathbf{0}_{(N_R - N) \times N} & \mathbf{0}_{(N_R - N) \times (M_R - N)} \end{bmatrix}. \quad (32)$$

With (32), the optimization problem (27) can be rewritten as

$$\begin{aligned} \min_{f_i^2} \quad & \sum_{i=1}^N \frac{\lambda_{T,i}}{\lambda_{\Theta,i} f_i^2 + 1} + \sum_{i=N+1}^{M_R} \lambda_{T,i} \\ \text{s.t.} \quad & \sum_{i=1}^N f_i^2 \leq P \end{aligned} \quad (33)$$

where $\lambda_{\Theta,i}$ denotes the i^{th} diagonal element of $\mathbf{\Lambda}_{\Theta}$. Obviously, the solution of the problem (33) is the modified water-filling [13], and based on the Karush-Kuhn-Tucker (KKT) conditions of (33), we have [14]

$$f_{i,\text{opt}}^2 = \left(\sqrt{\frac{\lambda_{T,i}}{\mu \lambda_{\Theta,i}}} - \frac{1}{\lambda_{\Theta,i}} \right)^+ \quad i = 1, \dots, N \quad (34)$$

where $\mu > 0$ is the Lagrangian multiplier such that $\sum_{i=1}^N f_{i,\text{opt}}^2 = P$ holds. From the definition of $\tilde{\mathbf{F}}$ in (26), (32) and (34), we can write the optimal \mathbf{F} compactly as

$$\mathbf{F}_{\text{opt}} = \mathbf{U}_{\Theta,N} \left(\left(\frac{1}{\sqrt{\mu}} \tilde{\mathbf{\Lambda}}_{\Theta}^{-\frac{1}{2}} \tilde{\mathbf{\Lambda}}_T^{\frac{1}{2}} - \tilde{\mathbf{\Lambda}}_{\Theta}^{-1} \right)^+ \right)^{\frac{1}{2}} \mathbf{U}_{T,N}^H \mathbf{R}_x^{-\frac{1}{2}} \quad (35)$$

where $[(\mathbf{Z})^+]_{i,j} = \max(0, (\mathbf{Z})_{i,j})$. The matrices $\tilde{\mathbf{\Lambda}}_{\Theta}$ and $\tilde{\mathbf{\Lambda}}_T$ are the principle submatrices of $\mathbf{\Lambda}_{\Theta}$ and $\mathbf{\Lambda}_T$ with dimensions $N \times N$. The matrices $\mathbf{U}_{\Theta,N}$ and $\mathbf{U}_{T,N}$ are the first N cloumns of \mathbf{U}_{Θ} and \mathbf{U}_T , respectively. Notice that when the source-relay link is noiseless and the channel realization is perfectly known, equation (35) reduces to the point-to-point MIMO robust LMMSE transceiver [15]. If both two channels are exactly known, (35) is exactly the solution in [5].

IV. NUMERICAL EXPERIMENTS

In this section, simulation results will be shown to verify the effectiveness of the proposed algorithm. In this paper, the source, relay and destination are all equipped with 3 antennas.

At the source, it is assumed that the transmit power $\text{Tr}(\mathbf{R}_s) = 20\text{dB}$ and the modulation scheme is QPSK. The estimated channel matrices, $\tilde{\mathbf{H}}_{sr}$ and $\tilde{\mathbf{H}}_{rd}$, are

$$\begin{aligned} \tilde{\mathbf{H}}_{sr} &= \begin{bmatrix} 0.2714 - 0.3487i & -0.6170 - 0.4784i & -0.2315 + 0.5103i \\ 0.2354 + 0.2462i & -0.3534 + 0.1253i & 0.1964 - 0.7238i \\ 1.1809 - 0.3305i & 0.3179 + 2.3439i & 0.1989 - 1.1954i \end{bmatrix} \\ \tilde{\mathbf{H}}_{rd} &= \begin{bmatrix} 0.9002 - 0.4583i & -0.9646 - 0.6782i & 0.9360 + 1.1348i \\ 0.9969 + 0.1589i & 0.2910 + 0.3071i & 0.6035 - 0.4315i \\ 0.6798 - 1.1627i & 0.7557 + 0.3929i & -0.3742 - 0.0623i \end{bmatrix}. \end{aligned} \quad (36)$$

The estimation error correlation matrices are assumed to be

$$\begin{aligned} \Sigma_{sr} &= \begin{bmatrix} 1 & \beta & \beta^2 \\ \beta & 1 & \beta \\ \beta^2 & \beta & 1 \end{bmatrix} & \Sigma_{rd} &= \begin{bmatrix} 1 & \beta & \beta^2 \\ \beta & 1 & \beta \\ \beta^2 & \beta & 1 \end{bmatrix} \\ \Psi_{sr} &= 0.03 \begin{bmatrix} 1 & \alpha & \alpha^2 \\ \alpha & 1 & \alpha \\ \alpha^2 & \alpha & 1 \end{bmatrix} & \Psi_{rd} &= 0.04 \begin{bmatrix} 1 & \alpha & \alpha^2 \\ \alpha & 1 & \alpha \\ \alpha^2 & \alpha & 1 \end{bmatrix}. \end{aligned} \quad (37)$$

In each simulation run, channel estimation errors, $\Delta \mathbf{H}_{sr}$ and $\Delta \mathbf{H}_{rd}$, are generated independently, according to (4) and (6), respectively, and 1000 trials are averaged to give each point in the figures.

Fig. 2 shows the MSE of the received signal at the destination versus the transmit power at the relay P , for the algorithm using estimated channel matrices only and the proposed Bayesian algorithm, with different values of β , when $\alpha = 0.4$. It can be seen that in general, the whole system performance degrades when the correlation factor β increases. This is due to the fact that channel correlations reduce the number of effective eigenchannels [6]. However, the performance of the proposed algorithm is significantly better than the algorithm using estimated channel matrices only, regardless of the value of β . Fig. 3 shows the corresponding results for different values of α , when $\beta = 0.4$. A similar conclusion to that of Fig. 2 can be drawn.

V. CONCLUSIONS

In this paper, we presented the joint design of linear transceivers for AF MIMO relay systems under the knowledge of estimated channel and error covariance matrices. The statistics of channel estimation errors were incorporated into the transceiver design using the Bayesian framework. A closed-form solution has been derived and two existing algorithms were shown to be special cases of our framework. From the simulations, it was found that the proposed algorithm reduces the sensitivity of the relay system to channel estimation errors, and improves the system performance greatly, compared to the algorithm using estimated channel only.

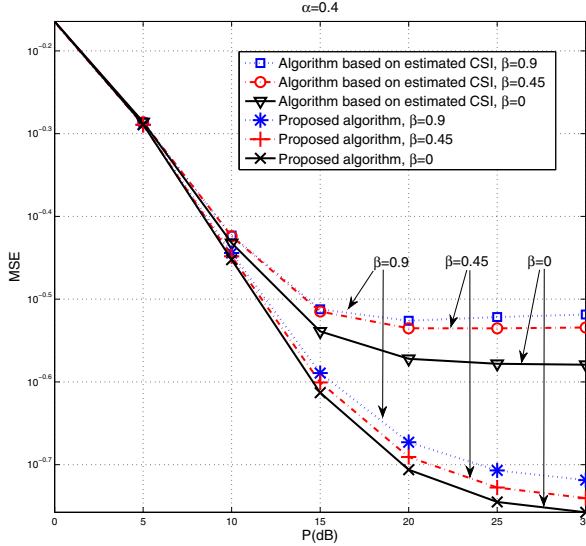


Fig. 2. MSE versus transmit power at the relay for the algorithm based on estimated channel and the proposed Bayesian algorithm, with different values of β , when $\alpha = 0.4$

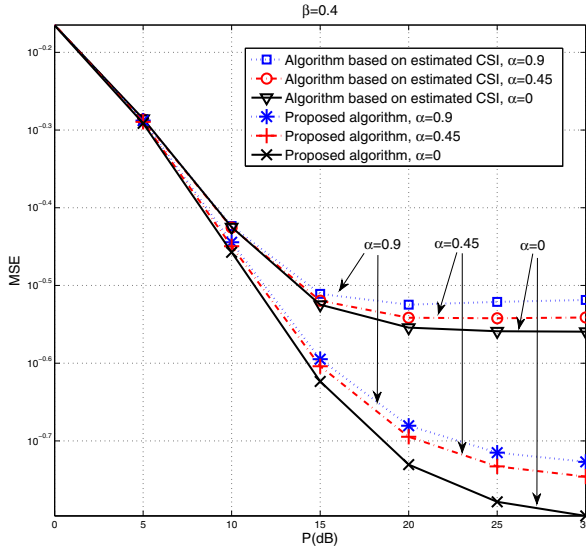


Fig. 3. MSE versus transmit power at the relay for the algorithm based on estimated channel and the proposed Bayesian algorithm with different values of α , when $\beta = 0.4$

APPENDIX I

For positive Hermitian matrices, \mathbf{M} and \mathbf{N} , if $\mathbf{M} \succeq \mathbf{N}$, the following inequality holds [10, 7.7.4]

$$\mathbf{N}^{-1} \succeq \mathbf{M}^{-1}. \quad (38)$$

Furthermore, for any matrix \mathbf{A} , the inequality

$$\mathbf{A}^H \mathbf{N}^{-1} \mathbf{A} \succeq \mathbf{A}^H \mathbf{M}^{-1} \mathbf{A} \quad (39)$$

always holds [10, 7.7.3.a]. Adding an identity matrix on both sides of (39), the inequality sign does not change. Together with (38), we have

$$(\mathbf{A}^H \mathbf{M}^{-1} \mathbf{A} + \mathbf{I})^{-1} \succeq (\mathbf{A}^H \mathbf{N}^{-1} \mathbf{A} + \mathbf{I})^{-1}. \quad (40)$$

With the result in (39), for an arbitrary matrix \mathbf{B} , we have

$$\mathbf{B}^H (\mathbf{A}^H \mathbf{M}^{-1} \mathbf{A} + \mathbf{I})^{-1} \mathbf{B} \succeq \mathbf{B}^H (\mathbf{A}^H \mathbf{N}^{-1} \mathbf{A} + \mathbf{I})^{-1} \mathbf{B}. \quad (41)$$

Putting $\mathbf{A} = \bar{\mathbf{H}}_{rd} \mathbf{F} \mathbf{R}_x^{\frac{1}{2}}$, $\mathbf{B} = \mathbf{R}_x^{-\frac{1}{2}} \bar{\mathbf{H}}_{sr}^H \mathbf{R}_s$, $\mathbf{N} = \mathbf{K}$ and $\mathbf{M} = \text{Tr}(\mathbf{F} \mathbf{R}_x \mathbf{F}^H) \lambda_{\max}(\Psi_{rd}) \Sigma_{rd} + \mathbf{R}_{n2}$, and taking the trace on both sides of (41), we have

$$\overline{\text{MSE}}_U(\mathbf{F}) \geq \overline{\text{MSE}}(\mathbf{F}) \quad (42)$$

where $\overline{\text{MSE}}_U(\mathbf{F})$ is defined as

$$\overline{\text{MSE}}_U(\mathbf{F}) = \text{Tr}(\mathbf{R}_s \bar{\mathbf{H}}_{sr}^H \mathbf{R}_x^{-\frac{H}{2}} (\mathbf{R}_x^{\frac{H}{2}} \mathbf{F}^H \bar{\mathbf{H}}_{rd}^H \mathbf{M}^{-1} \bar{\mathbf{H}}_{rd} \mathbf{F} \mathbf{R}_x^{\frac{1}{2}} + \mathbf{I}_{M_R})^{-1} \mathbf{R}_x^{-\frac{1}{2}} \bar{\mathbf{H}}_{sr} \mathbf{R}_s) + c, \quad (43)$$

and $\overline{\text{MSE}}(\mathbf{F})$ is defined in (21).

ACKNOWLEDGEMENT

This study was partially supported by a grant from the Research Grants Council of the Hong Kong SAR.

REFERENCES

- [1] A. Scaglione, D. L. Goeckel, and J. N. Laneman, "Cooperative communications in mobile Ad Hoc networks," *IEEE Signal Processing Magazine*, pp. 18–29, Sept. 2006.
- [2] H. Bolcskei, R. U. Nabar, O. Oyman, and A. J. Paulraj, "Capacity scaling laws in MIMO relay networks," *IEEE Trans. on Wireless Comm.*, vol. 5, no. 6, pp. 1433–1443, June 2006.
- [3] C.-B. Chae, T. Tang, R. W. Heath, and S. Cho, "MIMO relaying with linear processing for multiuser transmission in fixed relay networks," *IEEE Trans. on Signal Processing*, vol. 56, no. 2, pp. 727–738, Feb. 2008.
- [4] A. S. Behbahani, R. Merched, and A. M. Eltawil, "Optimizations of a MIMO relay network," *IEEE Trans. on Signal Processing*, vol. 56, no. 10, part 2, pp. 5062–5073, Oct. 2008.
- [5] W. Guan and H. Luo, "Joint MMSE transceiver design in non-regenerative MIMO relay systems," *IEEE Communications Letters*, vol. 12, issue 7, pp. 517–519, July 2008.
- [6] E. G. Larsson and P. Stoica, *Space-Time Block Coding for Wireless Communications*, Cambridge University Press, 2003.
- [7] A. Gupta and D. Nagar, *Matrix Variates Distributions*, London, U.K., Chapman&Hall/CRC, 2000.
- [8] A. P. Dawid, "Some matrix-variate distribution theory: Notational considerations and a Bayesian application," *Oxford Journals, Mathematics & Physical Sciences*, vol. 68, no. 1, pp. 265–274, 1981.
- [9] A. T. James, "Distributions of matrix variates and latent roots derived from normal samples," *Ann. Math. Statistics*, vol. 35, pp. 475–501, 1964.
- [10] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, 1985.
- [11] A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Application*, New York, Academic Press, 1979.
- [12] D. P. Palomar, J. M. Cioffi, and M. A. Lagunas, "Joint Tx-Rx beamforming design for multicarrier MIMO channels: A unified framework for convex optimization," *IEEE Trans. on Signal Processing*, vol. 51, no. 9, pp. 2381–2401, Sep. 2003.
- [13] F. Frachinei and J. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems*, Spring Series in Operation Research, Volume I, 2003.
- [14] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.
- [15] X. Zhang, D. P. Palomar, and B. Ottersten, "Statistically robust design of linear MIMO transceiver," *IEEE Trans. on Signal Processing*, vol. 56, no. 8, pp. 3678–3689, Aug. 2008.