Linear MMSE Transceiver Design in Amplify-and-Forward MIMO Relay Systems

Fan-Shuo Tseng, *Student Member, IEEE*, and Wen-Rong Wu, *Member, IEEE*

***Abstract*—We consider the precoding problem in an amplify-and-forward (AF) multiple-input–multiple-output (MIMO) relay system in which multiple antennas are equipped at the source, the relay, and the destination. Most existing methods for this problem only consider the design of the relay precoder, and some even ignore the direct link. In this paper, we consider a joint source/relay precoder design problem, taking both the direct and the relay links into account. Using a minimum-mean-square-error (MMSE) criterion, we first formulate the problem as a constrained optimization problem. However, it is found that the mean square error (MSE) is a highly nonlinear function of the precoders, and a direct optimization is difficult to conduct. We then design the precoders to diagonalize the MSE matrix in the cost function. To do that, we pose certain structural constraints on the precoders and derive an MSE upper bound. It turns out that minimization with respect to this bound becomes simple and straightforward. Using the standard Lagrange technique, we can finally obtain the solution with an iterative water-filling method. Simulation results show that the proposed method, with an additional precoder, outperforms the existing methods, in terms of either the MSE or the bit error rate (BER).**

***Index Terms*—Amplify and forward (AF), cooperative communication, minimum mean square error (MMSE), multiple input–multiple output (MIMO), precoder.**

I. Introduction

DIVERSITY is a common technique to overcome the multipath channel fading in wireless communications. Popular diversity schemes include time diversity, frequency diversity, and spatial diversity. Among these techniques, the spatial diversity is particularly attractive. This is because this technique can combine with the other two diversity techniques with no expansion of time and bandwidth. The conventional way to obtain the spatial diversity is the use of multiple transmit or multiple receive antennas. When both multiple transmit and receive antennas are used, the system is referred to as a multiple-input–multiple-output (MIMO) system [1]–[13]. MIMO systems have widely been studied in the literature since they can enhance the diversity or spectral efficiency in an efficient way [2]–[13].

User cooperation is an alternative way to obtain spatial diversity [14]–[21]. With the aid of additional relay nodes, spatial diversity can be achieved in a distributed manner. In typical cooperative communications, the source and relay nodes share the same spectrum. As a result, half-duplexing is often used for signal transmission between the source and the relays. In a typical three-node system, signal transmission is generally divided into two time phases [15]–[28]. In the first phase, the signal at the source is transmitted to the relay and the destination. In the second time phase, the relay forwards its received signal to the destination. Finally, the destination combines the received signals to achieve the spatial diversity. There are several protocols defining retransmission strategies at the relays. Two relay strategies are well known, i.e., amplify-and-forward (AF) and decode-and-forward (DF). In AF, the relay receives the signal from the source and retransmits it to the destination with signal amplification only. The system with the protocol is also called a nonregenerate cooperative system [23]–[25], [27]. In DF, the relay decodes the received signals, reencodes the information bits, and retransmits the resultant signal to the destination. The system is also called a regenerative cooperative system. It is simple to see that the DF protocol requires higher computational complexity and a larger processing delay at the relay nodes. In this paper, we only consider the AF-based cooperative system.

Recently, the MIMO technique has been introduced to cooperative systems as a means for further performance enhancement. With the multiple antennas equipped at the source and each node, a MIMO relay system is constructed [22]–[28]. Capacity bounds for a single-relay MIMO channel was first addressed in [22]. Similar to conventional MIMO systems, the precoding operation can be conducted in a MIMO relay system. The relay precoder in an AF-based MIMO relay system was first designed in [23] and [24] to enhance the overall channel capacity. In most of those approaches, only the relay link (the source-to-relay and relay-to-destination links) is considered. It was shown that the capacity can further be increased if the direct link (the link between the source and the destination) is taken into account [24]. Apart from the capacity, the link quality is another criterion that has been considered [25], [26]. In these works, a relay precoder is designed using a minimum-mean-square-error (MMSE) criterion. Most of the MIMO relay systems considered above use one relay. Precoding in multiple-relay MIMO systems was investigated in [26]. Note that the aforementioned works all address the spatial-multiplexing scenario. Recently, the design for the transmission of a single data stream, which is referred to as beamforming, has also been considered. For example, [28] derives the optimal source and relay beamformers using a maximal-SNR criterion. In this paper, the optimal solution is derived for the relay-link-only

Manuscript received April 13, 2009; revised July 14, 2009 and September 23, 2009. First published October 30, 2009; current version published February 19, 2010. This work was supported by the National Science Council under Grant NSC-98-2221-E-009-046-MY3. The review of this paper was coordinated by Dr. A. Ghrayeb.

The authors are with the Department of Electrical Engineering, National Chiao Tung University, Hsinchu 300, Taiwan (e-mail: stron.cm93g@nctu.edu.tw; wrwu@faculty.nctu.edu.tw).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TVT.2009.2034971

system. In addition to beamforming, antenna selection in MIMO relay systems was also studied. With the MMSE criterion, an optimum selection scheme was developed in [27]. In this approach, only one antenna is selected at the source and the relay, respectively, for signal transmission.

As mentioned, in the precoder design for spatial-multiplexing AF-based MIMO relay systems, the existing works only consider the precoder at the relay. Furthermore, the direct link is frequently ignored [23]–[26]. In this paper, we propose a new design method to solve the problems. Similar to previous works, we assume a linear receiver at the destination. Our approach is to seek precoders such that the MMSE of the linear receiver is minimized. However, we found that the MMSE is a complicated function of precoding matrices, and a direct minimization is almost impossible to conduct. To overcome the difficulty, we propose a structural constraint on the precoders to diagonalize the mean-square-error (MSE) matrix in the cost function. With the precoders, we can then derive a tractable MSE upper bound. Minimization with this upper bound, instead of the original MMSE, then becomes feasible. The proposed precoders can finally be computed via an iterative water-filling technique [9], [31], [32]. Note that the MSE criterion to minimize is the total MSE of the multiplexed signal streams. With the specially designed structure, the proposed precoders can make the individual MSEs of all signal streams equal, indicating that the bit error rate (BER) of the proposed precoded system will be the minimal among all precoded systems with the same minimum total MSE [5]. The rest of this paper is organized as follows. Section II gives the system model and problem formulation, Section III derives the proposed joint design in detail, Section IV reports simulation results for various applications, and finally, Section V concludes this paper.

II. Proposed System Model and Problem Formulation

A. Precoders for the AF MIMO Relay System

We consider a typical three-node half-duplex cooperative AF MIMO relay system where multiple antennas are placed at each node. Under this scenario, signals can be transmitted from the source to the destination and from the source to the relay and then to the destination. To avoid the interference between direct and relay links, we consider the time-division-duplexing scheme [23]–[26] used in a typical two-phase transmission aforementioned (see Fig. 1). Let N, R, and M denote the number of antennas at the source, the relay, and the destination, and assume that all channels are flat fading. For the first phase, the received signals at the destination and the relay can be expressed as

$$\mathbf{y}_{D,1} = \mathbf{H}_{SD}\mathbf{F}_S\mathbf{s} + \mathbf{n}_{D,1} \tag{1}$$

$$\mathbf{y}_{R} = \mathbf{H}_{SR}\mathbf{F}_S\mathbf{s} + \mathbf{n}_{R} \tag{2}$$

respectively, where $\mathbf{s} \in \mathbb{C}^{L\times 1}$ is the transmitted signal vector, with L being the number of the substreams, $\mathbf{F}_S \in \mathbb{C}^{N\times L}$ is the precoding matrix at the source, $\mathbf{H}_{SR} \in \mathbb{C}^{R\times N}$ and $\mathbf{H}_{SD} \in \mathbb{C}^{M\times N}$ are the channel matrices corresponding to the source-to-relay and source-to-destination channels, respectively, $\mathbf{n}_{D,1} \in \mathbb{C}^{M\times 1}$ is the first-phase received noise vector at the destination, and $\mathbf{n}_R \in \mathbb{C}^{R\times 1}$ is the received noise vector at the relay. Here, we assume that $L \le \min\{N, M\}$ provides sufficient degrees of freedom for signal detection.

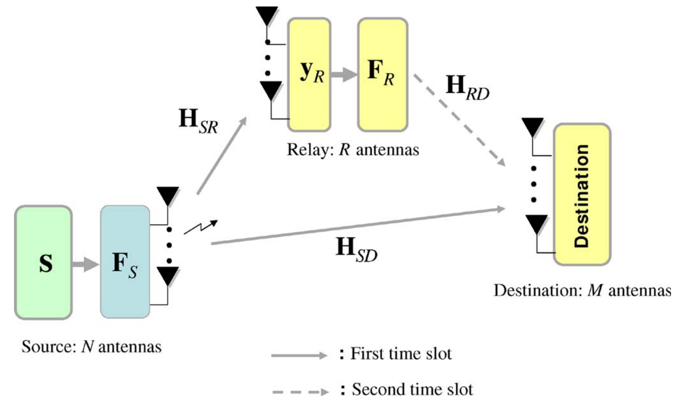


Fig. 1. Three-node AF-based MIMO relay system with multiple antennas allocated at each node.

In the second phase of the transmission, the relay retransmits the received signal with another precoding matrix. Thus, the received signals at the destination can be expressed as

$$\begin{aligned}\mathbf{y}_{D,2} &= \mathbf{H}_{RD}\mathbf{F}_R\mathbf{y}_R + \mathbf{n}_{D,2}\\ &= \mathbf{H}_{RD}\mathbf{F}_R\mathbf{H}_{SR}\mathbf{F}_S\mathbf{s} + (\mathbf{H}_{RD}\mathbf{F}_R\mathbf{n}_R + \mathbf{n}_{D,2})\end{aligned} \tag{3}$$

where $\mathbf{F}_R \in \mathbb{C}^{R\times R}$ is the precoding matrix at the relay, $\mathbf{H}_{RD} \in \mathbb{C}^{M\times R}$ is the channel matrix corresponding to the relay-to-destination channel, and $\mathbf{n}_{D,2} \in \mathbb{C}^{M\times 1}$ is the second-phase received noise vector at the destination. Here, we assume that each element in $\mathbf{n}_{D,1}$ has a zero-mean circularly symmetric Gaussian distribution, and all the elements are independent identically distributed (i.i.d.). The same assumption is applied for $\mathbf{n}_{D,2}$ and $\mathbf{n}_R$. As a result, the received signal vectors $\mathbf{y}_{D,1}$ and $\mathbf{y}_{D,2}$ for the two phases can be combined into a single vector, which is denoted as $\mathbf{y}_D \in \mathbb{C}^{2M\times 1}$. Consequently, we have

$$\mathbf{y}_D := \begin{bmatrix}\mathbf{y}_{D,1}\\ \mathbf{y}_{D,2}\end{bmatrix} = \mathbf{H}\mathbf{F}_S\mathbf{s} + \mathbf{n} \tag{4}$$

where

$$\mathbf{H} = \begin{bmatrix}\mathbf{H}_{SD}\\ \mathbf{H}_{RD}\mathbf{F}_R\mathbf{H}_{SR}\end{bmatrix} \tag{5}$$

$$\mathbf{n} = \begin{bmatrix}\mathbf{n}_{D,1}\\ \mathbf{H}_{RD}\mathbf{F}_R\mathbf{n}_R + \mathbf{n}_{D,2}\end{bmatrix}. \tag{6}$$

Here, $\mathbf{H}$ is the equivalent channel matrix with rank $(\mathbf{H}) = N$, and $\mathbf{n}$ is the equivalent noise vector at the destination. It is noteworthy that the noise received at the relay is amplified by the relay precoder and the relay-to-destination channel. Furthermore, the equivalent channel matrix in (5) is a function of the relay precoder $\mathbf{F}_R$. This is quite different from the scenario considered in conventional MIMO systems. The precoder design problem is actually a joint-transceiver-design problem. In other words, the optimum precoders depend on not only the channels but on the receiver as well. Similar to previous works, we will consider the linear MMSE receiver in our design [25], [26].

B. MMSE Receiver and Related MSE Matrix

Let $\mathbf{R}_{\mathbf{n}_{D,1}} = E[\mathbf{n}_{D,1}\mathbf{n}_{D,1}^H] = \sigma_{n,d}^2\mathbf{I}_M$, $\mathbf{R}_{\mathbf{n}_{D,2}} = E[\mathbf{n}_{D,2}\mathbf{n}_{D,2}^H] = \sigma_{n,d}^2\mathbf{I}_M$, and $\mathbf{R}_R = E[\mathbf{n}_R\mathbf{n}_R^H] = \sigma_{n,r}^2\mathbf{I}_R$, where $\sigma_{n,d}^2$ and $\sigma_{n,r}^2$ are the noise variances at the destination and the relay, respectively. Furthermore, the elements of the transmitted symbols are i.i.d. with zero mean and covariance matrix $\mathbf{R}_s = \sigma_s^2\mathbf{I}_L$, where σ_s^2 is the transmitted symbol power.

Using the setting, we can have the covariance matrix of the equivalent noise vector as

$$\begin{aligned}\mathbf{R}_n &= E\left[\mathbf{n}\mathbf{n}^H\right]\\ &= \begin{bmatrix}\sigma_{n,d}^2\mathbf{I}_M & \mathbf{0}\\ \mathbf{0} & \sigma_{n,r}^2\mathbf{H}_{RD}\mathbf{F}_R\mathbf{F}_R^H\mathbf{H}_{RD}^H + \sigma_{n,d}^2\mathbf{I}_M\end{bmatrix}. \end{aligned}\tag{7}$$

Let $\mathbf{G}$ be the equalization matrix in the receiver. Then, the MSE for recovering $\mathbf{s}$, which is denoted as J, is given by

$$J = E\left\{\|\mathbf{G}\mathbf{y}_D - \mathbf{s}\|^2\right\}. \tag{8}$$

The minimization of (8) leads to the optimal equalization matrix [4] as

$$\mathbf{G}_{\text{opt}} = \sigma_s^2\mathbf{F}_S^H\mathbf{H}^H\left(\sigma_s^2\mathbf{H}\mathbf{F}_S\mathbf{F}_S^H\mathbf{H}^H + \mathbf{R}_n\right)^{-1}. \tag{9}$$

Substituting (9) into (8) and invoking the matrix inversion lemma [29], we can then have the MMSE, which is denoted by $J_{\min}$, as

$$J_{\min} = \text{tr}\{\mathbf{E}\} \tag{10}$$

where

$$\mathbf{E} = \left(\sigma_s^{-2}\mathbf{I}_L + \mathbf{E}_S + \mathbf{E}_R\right)^{-1}. \tag{11}$$

In (11)

$$\mathbf{E}_S = \sigma_{n,d}^{-2}\mathbf{F}_S^H\mathbf{H}_{SD}^H\mathbf{H}_{SD}\mathbf{F}_S \tag{12}$$

$$\begin{aligned}\mathbf{E}_R = \mathbf{F}_S^H\mathbf{H}_{SR}^H\mathbf{F}_R^H\mathbf{H}_{RD}^H\left(\sigma_{n,r}^2\mathbf{H}_{RD}\mathbf{F}_R\mathbf{F}_R^H\mathbf{H}_{RD}^H + \sigma_{n,d}^2\mathbf{I}_M\right)^{-1}\\ \times\mathbf{H}_{RD}\mathbf{F}_R\mathbf{H}_{SR}\mathbf{F}_S.\end{aligned}\tag{13}$$

As we can see from (11), the MMSE is a function of $\mathbf{F}_S$ and $\mathbf{F}_R$. It is also simple to see that $\mathbf{E}_S$ accounts for the MMSE contributed in the direct link, and $\mathbf{E}_R$ accounts for that in the relay link. If we ignore the direct link and only consider the relay precoder, the problem will be degenerated to the case considered in [25].

C. Problem Formulation

As shown in (10), the MMSE is a function of the two precoding matrices, i.e., $\mathbf{F}_S$ and $\mathbf{F}_R$. Our task here is to design these two matrices such that the MSE in (10) can be minimized. The optimization problem can then be formulated as (14), shown at the bottom of the page. The inequalities in (14) indicate that the precoders have to satisfy the transmit power constraints at both the source and the relay.

From (14), we can readily find that (14) is not a convex optimization. Furthermore, the cost function involves a series of matrix multiplications and inversions, it is a complicated and nonlinear function of $\mathbf{F}_S$ and $\mathbf{F}_R$. The cost function may have many local minimums, and the optimal solution, even with numerical methods [30], is difficult to derive. We will propose a method, which is described below, to solve these problems.

III. Joint Source/Relay Precoder Design With MMSE Receiver

As aforementioned, the optimum solution for (14) is difficult to derive. In this section, we then propose a method to seek for a suboptimum solution. One difficulty in (14) is that the number of unknown parameters in $\mathbf{F}_R$ and $\mathbf{F}_S$ can be large. The first idea of our approach is to use a constrained precoder structure such that the number of unknowns can effectively be reduced. The other difficulty in (14) is that the formulas are too complicated to work with. Our second idea is to derive an MMSE upper bound having a simple expression and conduct minimization with this upper bound. Even though the cost function can dramatically be simplified with the proposed method, a closed-form solution is still difficult to obtain. We then use an iterative water-filling method to solve the problem.

$$\begin{aligned}
&\min_{\mathbf{F}_S,\mathbf{F}_R}\quad \text{tr}\{\mathbf{E}\} = \sum_{i=1}^{L}\mathbf{E}_{(i,i)}\\
&\text{s.t.}\\
&\mathbf{E} = \left(\sigma_s^{-2}\mathbf{I}_L + \underbrace{\sigma_{n,d}^{-2}\mathbf{F}_S^H\mathbf{H}_{SD}^H\mathbf{H}_{SD}\mathbf{F}_S}_{:=\mathbf{E}_S} + \underbrace{\mathbf{F}_S^H\mathbf{H}_{SR}^H\mathbf{F}_R^H\mathbf{H}_{RD}^H\left(\sigma_{n,r}^2\mathbf{H}_{RD}\mathbf{F}_R\mathbf{F}_R^H\mathbf{H}_{RD}^H + \sigma_{n,d}^2\mathbf{I}_M\right)^{-1}\mathbf{H}_{RD}\mathbf{F}_R\mathbf{H}_{SR}\mathbf{F}_S}_{:=\mathbf{E}_R}\right)^{-1}\\
&\text{tr}\left\{E\left[\mathbf{F}_R\mathbf{y}_R\mathbf{y}_R^H\mathbf{F}_R^H\right]\right\} = \text{tr}\left\{\mathbf{F}_R\left(\sigma_{n,r}^2\mathbf{I}_R + \sigma_s^2\mathbf{H}_{SR}\mathbf{F}_S\mathbf{F}_S^H\mathbf{H}_{SR}^H\right)\mathbf{F}_R^H\right\} \le P_{R,T}\\
&\text{tr}\left\{\mathbf{F}_S E[\mathbf{s}\mathbf{s}^H]\mathbf{F}_S^H\right\} = \sigma_s^2\text{tr}\left\{\mathbf{F}_S\mathbf{F}_S^H\right\} \le P_{S,T}
\end{aligned}\tag{14}$$

A. Proposed Approach

When the direct link is ignored and only a relay precoder is considered, the optimal MMSE precoder can analytically be obtained through an MSE matrix diagonalization procedure [25]. Motivated by this fact, we propose to conduct a similar matrix diagonalization in our design. Indeed, if the error matrix $\mathbf{E}$ in (14) can be diagonalized, the trace operation can easily be conducted, and the whole problem can be greatly simplified. To do that, we first consider the following singular value decomposition (SVD) for the channel matrices in all links:

$$\mathbf{H}_{SD} = \mathbf{U}_{sd}\mathbf{\Sigma}_{sd}\mathbf{V}_{sd}^H \quad (15)$$
$$\mathbf{H}_{SR} = \mathbf{U}_{sr}\mathbf{\Sigma}_{sr}\mathbf{V}_{sr}^H \quad (16)$$
$$\mathbf{H}_{RD} = \mathbf{U}_{rd}\mathbf{\Sigma}_{rd}\mathbf{V}_{rd}^H \quad (17)$$

where $\mathbf{U}_{sd} \in \mathbb{C}^{M\times M}$, $\mathbf{U}_{sr} \in \mathbb{C}^{R\times R}$, and $\mathbf{U}_{rd} \in \mathbb{C}^{M\times M}$ are the left singular matrices of $\mathbf{H}_{SD}$, $\mathbf{H}_{SR}$, and $\mathbf{H}_{RD}$, respectively; $\mathbf{\Sigma}_{sd} \in \mathbb{R}^{M\times N}$, $\mathbf{\Sigma}_{sr} \in \mathbb{R}^{R\times N}$, and $\mathbf{\Sigma}_{rd} \in \mathbb{R}^{M\times R}$ are diagonal singular-value matrices of $\mathbf{H}_{SD}$, $\mathbf{H}_{SR}$, and $\mathbf{H}_{RD}$, respectively; and $\mathbf{V}_{sd}^H \in \mathbb{C}^{N\times N}$, $\mathbf{V}_{sr}^H \in \mathbb{C}^{N\times N}$, and $\mathbf{V}_{rd}^H \in \mathbb{C}^{R\times R}$ are the right singular matrices of $\mathbf{H}_{SD}$, $\mathbf{H}_{SR}$, and $\mathbf{H}_{RD}$, respectively.

Observing (14), we will readily find that a complete diagonalization of $\mathbf{E}$ will be difficult. We then first consider the diagonalization of $(\sigma_{n,r}^2\mathbf{H}_{RD}\mathbf{F}_R\mathbf{F}_R^H\mathbf{H}_{RD}^H + \sigma_{n,d}^2\mathbf{I}_M)^{-1}$ using $\mathbf{F}_R$ so that the inverse operation can easily be tackled. Such an approach, although suboptimal, will considerably simplify our derivation. It also allows us to derive an MSE upper bound and then obtain a scalar-valued optimization problem. With the SVD in (17), an immediate choice for $\mathbf{F}_R$ to diagonalize $(\sigma_{n,r}^2\mathbf{H}_{RD}\mathbf{F}_R\mathbf{F}_R^H\mathbf{H}_{RD}^H + \sigma_{n,d}^2\mathbf{I}_M)$ is

$$\mathbf{F}_R = \mathbf{V}_{rd}\mathbf{\Sigma}_r\mathbf{U}_r \quad (18)$$

where $\mathbf{\Sigma}_r \in \mathbb{R}^{R\times R}$ is a diagonal matrix, and $\mathbf{U}_r \in \mathbb{C}^{R\times R}$ is a unitary matrix to be determined. With (18), we have

$$\left(\sigma_{n,r}^2\mathbf{H}_{RD}\mathbf{F}_R\mathbf{F}_R^H\mathbf{H}_{RD}^H + \sigma_{r,d}^2\mathbf{I}_M\right)^{-1} = \mathbf{U}_{rd}\left(\sigma_{n,r}^2\mathbf{\Sigma}_{rd}\mathbf{\Sigma}_r^2\mathbf{\Sigma}_{rd}^H + \sigma_{r,d}^2\mathbf{I}_M\right)^{-1}\mathbf{U}_{rd}^H. \quad (19)$$

To further diagonalize $\mathbf{F}_S^H\mathbf{H}_{SR}^H\mathbf{F}_R^H\mathbf{H}_{RD}^H(\sigma_{n,r}^2\mathbf{H}_{RD}\mathbf{F}_R \times \mathbf{F}_R^H\mathbf{H}_{RD}^H + \sigma_{n,d}^2\mathbf{I}_M)^{-1}\mathbf{H}_{RD}\mathbf{F}_R\mathbf{H}_{SR}\mathbf{F}_S$, we can select

$$\mathbf{U}_r = \mathbf{U}_{sr}^H \quad (20)$$
$$\mathbf{F}_S = \mathbf{V}_{sr}\mathbf{\Sigma}_s\mathbf{U}_s. \quad (21)$$

$\mathbf{\Sigma}_s \in \mathbb{R}^{N\times L}$ is a diagonal matrix, and $\mathbf{U}_s \in \mathbb{C}^{L\times L}$ is a unitary matrix yet to be specified. From (18) and (20), we have

$$\mathbf{F}_R = \mathbf{V}_{rd}\mathbf{\Sigma}_r\mathbf{U}_{sr}^H. \quad (22)$$

After some manipulations, we can obtain the MSE in (14) as (23), shown at the bottom of the page, where

$$\mathbf{V} = \mathbf{V}_{sd}^H\mathbf{V}_{sr} \quad (24)$$

is a constant matrix related to the channels. Note that the inclusion of the unitary matrix $\mathbf{U}_s$ in (23) will not change the cost function at all. However, by an appropriate design of $\mathbf{U}_s$, we can make the diagonal components of $\mathbf{E}$ equal. It has been shown that, under a fixed MSE, i.e., $\mathrm{tr}\{\mathbf{E}\}$, the receiver that makes the MSEs of the MIMO components equal has the lowest BER performance [5]. From (23), we now have some observations in order. First, we see that (23) is obtained with the constrained structure of the precoding matrices specified in (21) and (22). The MMSE obtained with the precoders can serve as an upper bound of the true MMSE. Second, the unknown matrices become $\mathbf{\Sigma}_r$ and $\mathbf{\Sigma}_s$, which are diagonal, and the whole problem is easier to handle. Finally, the matrix $\mathbf{E}_S$ cannot be diagonalized. However, starting from (23) and exploiting the diagonal nature of $\mathbf{E}_R$, we can further derive an MSE upper bound and use it to diagonalize $\mathbf{E}_S$.

To proceed, let us use the matrix inverse lemma to rewrite (23) as

$$\begin{aligned}\mathrm{tr}(\mathbf{E}) &= \mathrm{tr}\left(\left[\underbrace{\left(\sigma_s^{-2}\mathbf{I}_L+\mathbf{E}_R\right)}_{:=\mathbf{A}}+\mathbf{\Sigma}_s^H\underbrace{\left(\sigma_{n,d}^{-2}\mathbf{V}^H\mathbf{\Sigma}_{sd}^H\mathbf{\Sigma}_{sd}\mathbf{V}\right)}_{:=\mathbf{B}}\mathbf{\Sigma}_s\right]^{-1}\right)\\ &= \mathrm{tr}(\mathbf{A}^{-1}) - \mathrm{tr}\left(\mathbf{A}^{-1}\mathbf{\Sigma}_s^H\left(\mathbf{B}^{-1}+\mathbf{\Sigma}_s\mathbf{A}^{-1}\mathbf{\Sigma}_s^H\right)^{-1}\mathbf{\Sigma}_s\mathbf{A}^{-1}\right).\end{aligned} \quad (25)$$

It is noted here that to make sure that the inverse of $\mathbf{B}$ exists, $\mathbf{B}$ should be positive definite. To achieve that, we assume that $N \le M$. Based on (25), the desired MSE upper bound can be obtained by the aid of the next lemma.

$$\begin{aligned}\mathrm{tr}\{\mathbf{E}\} &= \mathrm{tr}\Bigg\{\left(\sigma_s^{-2}\mathbf{I}_L + \mathbf{U}_s^H\mathbf{\Sigma}_s^H\mathbf{\Sigma}_{sr}^H\mathbf{\Sigma}_r^H\mathbf{\Sigma}_{rd}^H\left(\sigma_{n,r}^2\mathbf{\Sigma}_{rd}\mathbf{\Sigma}_r^2\mathbf{\Sigma}_{rd}^H + \sigma_{n,d}^2\mathbf{I}_M\right)^{-1}\mathbf{\Sigma}_{rd}\mathbf{\Sigma}_r\mathbf{\Sigma}_{sr}\mathbf{\Sigma}_s\mathbf{U}_s\right.\\ &\qquad\left. + \sigma_{n,d}^{-2}\mathbf{U}_s^H\mathbf{\Sigma}_s^H\mathbf{V}^H\mathbf{\Sigma}_{sd}^H\mathbf{\Sigma}_{sd}\mathbf{V}\mathbf{\Sigma}_s\mathbf{U}_s\right)^{-1}\Bigg\}\\ &= \mathrm{tr}\left\{\left(\sigma_s^{-2}\mathbf{I}_L + \underbrace{\mathbf{\Sigma}_s^H\mathbf{\Sigma}_{sr}^H\mathbf{\Sigma}_r^H\mathbf{\Sigma}_{rd}^H\left(\sigma_{n,r}^2\mathbf{\Sigma}_{rd}\mathbf{\Sigma}_r^2\mathbf{\Sigma}_{rd}^H + \sigma_{n,d}^2\mathbf{I}_M\right)^{-1}\mathbf{\Sigma}_{rd}\mathbf{\Sigma}_r\mathbf{\Sigma}_{sr}\mathbf{\Sigma}_s}_{:=\mathbf{E}_R} + \underbrace{\sigma_{n,d}^{-2}\mathbf{\Sigma}_s^H\mathbf{V}^H\mathbf{\Sigma}_{sd}^H\mathbf{\Sigma}_{sd}\mathbf{V}\mathbf{\Sigma}_s}_{:=\mathbf{E}_S}\right)^{-1}\right\}\end{aligned} \quad (23)$$

Lemma: Let $\mathbf{D}_1$ and $\mathbf{D}_2$ be diagonal matrices, with the diagonal entries of $\mathbf{D}_2$ being positive. Then, for any positive definite matrix $\mathbf{X}$, we have

$$\mathrm{tr}\left(\mathbf{D}_1^H(\mathbf{X}+\mathbf{D}_2)^{-1}\mathbf{D}_1\right) \geq \mathrm{tr}\left(\mathbf{D}_1^H\left(\mathrm{diag}(\mathbf{X})+\mathbf{D}_2\right)^{-1}\mathbf{D}_1\right) \tag{26}$$

where $\mathrm{diag}(\mathbf{X})$ is obtained from $\mathbf{X}$ by setting its off-diagonal entries to zero. The equality in (26) holds if $\mathbf{X}$ is diagonal.

Proof: See the Appendix.

By the lemma, it follows that

$$\begin{aligned}&\mathrm{tr}\left(\mathbf{A}^{-1}\boldsymbol{\Sigma}_s^H\left(\mathbf{B}^{-1}+\boldsymbol{\Sigma}_s\mathbf{A}^{-1}\boldsymbol{\Sigma}_s^H\right)^{-1}\boldsymbol{\Sigma}_s\mathbf{A}^{-1}\right)\\ &\qquad\geq \mathrm{tr}\left(\mathbf{A}^{-1}\boldsymbol{\Sigma}_s^H\left(\mathrm{diag}(\mathbf{B}^{-1})+\boldsymbol{\Sigma}_s\mathbf{A}^{-1}\boldsymbol{\Sigma}_s^H\right)^{-1}\boldsymbol{\Sigma}_s\mathbf{A}^{-1}\right).\end{aligned} \tag{27}$$

Using (25) and (27), we can have the following key result:

$$\begin{aligned}\mathrm{tr}(\mathbf{E}) \leq\, &\mathrm{tr}(\mathbf{A}^{-1})\\ &-\mathrm{tr}\left(\mathbf{A}^{-1}\boldsymbol{\Sigma}_s^H\left(\mathrm{diag}(\mathbf{B}^{-1})+\boldsymbol{\Sigma}_s\mathbf{A}^{-1}\boldsymbol{\Sigma}_s^H\right)^{-1}\boldsymbol{\Sigma}_s\mathbf{A}^{-1}\right)\\ =&\sum_{i=1}^{L}\frac{1}{\sigma_s^{-2}+\frac{\sigma_{s,i}^2\sigma_{r,i}^2\sigma_{sr,i}^2\sigma_{rd,i}^2}{\sigma_{n,r}^2\sigma_{r,i}^2\sigma_{rd,i}^2+\sigma_{n,d}^2}+\sigma_{s,i}^2\left(\mathbf{B}^{-1}(i,i)\right)^{-1}}.\end{aligned} \tag{28}$$

Compared with the original MSE function (14), the upper bound in (28) admits a much simpler form and is analytically tractable. Hence, we propose to design the precoder by minimizing the upper bound in (28). For convenience, let $p_{s,i}=\sigma_{s,i}^2$ and $p_{r,i}=\sigma_{r,i}^2$ in (28). The optimization can finally be formulated as

$$\begin{aligned}\min_{\substack{p_{s,i},p_{r,i},\\ i=1,\cdots,L}}\quad &\sum_{i=1}^{L}\frac{1}{\sigma_s^{-2}+\frac{p_{s,i}p_{r,i}\sigma_{sr,i}^2\sigma_{rd,i}^2}{\sigma_{n,r}^2p_{r,i}\sigma_{rd,i}^2+\sigma_{n,d}^2}+p_{s,i}\left(\mathbf{B}^{-1}(i,i)\right)^{-1}}\\ \text{s.t.}\quad &\mathrm{tr}\left\{\boldsymbol{\Sigma}_r\left(\sigma_{n,r}^2\mathbf{I}_R+\sigma_s^2\boldsymbol{\Sigma}_{sr}\boldsymbol{\Sigma}_s\boldsymbol{\Sigma}_s^H\boldsymbol{\Sigma}_{sr}^H\right)\boldsymbol{\Sigma}_r^H\right\}\\ &\quad=\sum_{i=1}^{L}p_{r,i}\left(\sigma_{n,r}^2+\sigma_s^2p_{s,i}\sigma_{sr,i}^2\right)\leq P_{R,T}\\ &\sigma_s^2\mathrm{tr}\left\{\boldsymbol{\Sigma}_s\boldsymbol{\Sigma}_s^H\right\}=\sigma_s^2\sum_{i=1}^{L}p_{s,i}\leq P_{S,T},\\ &p_{s,i}\geq 0, p_{r,i}\geq 0,\qquad \forall i.\end{aligned} \tag{29}$$

It is simple to see that the problem in (29) is not a convex optimization problem either, and the optimum solution is still difficult to find. However, note that if one of $p_{r,i}$ and $p_{s,i}$ is given, (29) will become a convex optimization problem. This suggests a method, which is referred to as the iterative water-filling method [9], [31], [32], to find a suboptimum solution. For a given $p_{s,i}$, the optimum $p_{r,i}$ can be expressed as (30) (see Appendix B), shown at the bottom of the page, where $[y]^+=\max[0,y]$, and μ_r is the water level chosen to satisfy the power constraint at the relay, i.e., $\sum_{i=1}^{L}p_{r,i}(\sigma_{n,r}^2+\sigma_s^2p_{s,i}\sigma_{sr,i}^2)=P_{R,T}$. With $p_{r,i}=\sigma_{r,i}^2$ in (30), the relay precoder can be obtained by (22). For a given $p_{r,i}$, the optimum $p_{s,i}$ can be expressed as (31), shown at the bottom of the page, where μ_s is the water level chosen to meet the power constraint at the source, i.e., $\sum_{i=1}^{L}p_{s,i}=P_{S,T}$, and

$$\begin{aligned}\beta_i=&\left(\sigma_{n,d}^2+p_{r,i}\sigma_{n,r}^2\sigma_{rd,i}^2\right)\\ &\times\left(\left(\mathbf{B}^{-1}(i,i)\right)^{-1}\left(\sigma_{n,d}^2+p_{r,i}\sigma_{n,r}^2\sigma_{rd,i}^2\right)+p_{r,i}\sigma_{sr,i}^2\sigma_{rd,i}^2\right).\end{aligned} \tag{32}$$

Thus, we can use (30) and (31) to iteratively solve $p_{r,i}$ and $p_{s,i}$. To determine the $\mathbf{U}_s$, we first substitute (30) and (31) into (22) and (21), respectively, and express the error matrix in (11) as

$$\mathbf{E}=\left(\sigma_s^{-2}\mathbf{I}_L+\mathbf{U}_s^H\tilde{\mathbf{E}}\mathbf{U}_s\right)^{-1} \tag{33}$$

where

$$\begin{aligned}\tilde{\mathbf{E}}=\,&\boldsymbol{\Sigma}_s^H\boldsymbol{\Sigma}_{sr}^H\boldsymbol{\Sigma}_r^H\boldsymbol{\Sigma}_{rd}^H\left(\sigma_{n,r}^2\boldsymbol{\Sigma}_{rd}\boldsymbol{\Sigma}_r^2\boldsymbol{\Sigma}_{rd}^H+\sigma_{n,d}^2\mathbf{I}_M\right)^{-1}\\ &\times\boldsymbol{\Sigma}_{rd}\boldsymbol{\Sigma}_r\boldsymbol{\Sigma}_{sr}\boldsymbol{\Sigma}_s+\sigma_{n,d}^{-2}\boldsymbol{\Sigma}_s^H\mathbf{V}^H\boldsymbol{\Sigma}_{sd}^H\boldsymbol{\Sigma}_{sd}\mathbf{V}\boldsymbol{\Sigma}_s.\end{aligned} \tag{34}$$

Our task now is to design $\mathbf{U}_s$ such that (33) has equal diagonal MSE values. To do that, we consider the following eigendecomposition:

$$\tilde{\mathbf{E}}=\mathbf{V}_{\tilde{\mathbf{E}}}\mathbf{D}_{\tilde{\mathbf{E}}}\mathbf{V}_{\tilde{\mathbf{E}}}^H \tag{35}$$

where $\mathbf{V}_{\tilde{\mathbf{E}}}\in\mathbb{C}^{L\times L}$ is a matrix with the eigenvectors of $\tilde{\mathbf{E}}$ as its columns, and $\mathbf{D}_{\tilde{\mathbf{E}}}\in\mathbb{R}^{L\times L}$ is a diagonal matrix with the eigenvalues of $\tilde{\mathbf{E}}$ as its diagonal components. Therefore, if we let

$$\mathbf{U}_s=\mathbf{V}_{\tilde{\mathbf{E}}}\mathbf{F}_L \tag{36}$$

$$p_{r,i}=\left[\frac{\mu_r\sigma_{n,d}\sqrt{p_{s,i}}\sigma_{sr,i}\sigma_{rd,i}\left(\sigma_s^2p_{s,i}\sigma_{sr,i}^2+\sigma_{n,r}^2\right)^{-1/2}-\sigma_{n,d}^2\left(\sigma_s^{-2}+p_{s,i}\left(\mathbf{B}^{-1}(i,i)\right)^{-1}\right)}{\sigma_{rd,i}^2\left(\sigma_{n,r}^2\left(\sigma_s^{-2}+p_{s,i}\left(\mathbf{B}^{-1}(i,i)\right)^{-1}\right)+p_{s,i}\sigma_{sr,i}^2\right)}\right]^+ \tag{30}$$

$$p_{s,i}=\left[\frac{\mu_s\sqrt{\beta_i}-\sigma_s^{-2}\left(\sigma_{n,d}^2+p_{r,i}\sigma_{n,r}^2\sigma_{rd,i}^2\right)}{\left(\left(\mathbf{B}^{-1}(i,i)\right)^{-1}\left(\sigma_{n,d}^2+p_{r,i}\sigma_{n,r}^2\sigma_{rd,i}^2\right)+p_{r,i}\sigma_{sr,i}^2\sigma_{rd,i}^2\right)}\right]^+ \tag{31}$$

TABLE I
COMPLEXITY OF THE PROPOSED JOINT PRECODERS

Operation	FLOPs
SVD, (15)-(17)	$(14MN^2+8N^3)+(14RN^2+8N^3)+(14MR^2+8R^3)$
$\mathbf{B}^{-1}$, (25)	$2MN^2 + 2MN + 2N^3 + 13/4N^2 + N^2$
$p_{s,i}$ and $p_{r,i}$, (30)-(31)	$(21LI_r+14LI_s)I_i$
$\tilde{\mathbf{E}}$, (34)	$14L+10M+4NL+2L^2N$
SVD of $\tilde{\mathbf{E}}$, (35)	$22L^3$
$\mathbf{U}_S$, (36)	$2L^3$
$\mathbf{F}_S$ and $\mathbf{F}_R$, (21)-(22)	$(2NL+2NL^2)+(2R^2+2R^3)$
N: number of transmit antennas R: number of relay antennas M: number of receive antennas L: number of transmitted symbol streams I_r: number of iteration for computing $p_{r,i}$ I_s: number of iteration for computing $p_{s,i}$ I_i: number of iteration of the water-filling process	

where $\mathbf{F}_L$ is the L-point discrete-Fourier-transform matrix, (33) can be reexpressed as

$$\mathbf{E} = \mathbf{F}_L^H \left(\sigma_s^2 \mathbf{I}_L + \mathbf{D}_{\tilde{\mathbf{E}}}\right)^{-1} \mathbf{F}_L \tag{37}$$

which reveals that $\mathbf{E}$ is a circulant matrix with equal diagonal elements. It is simple to check the unitary property that $\mathbf{U}_s\mathbf{U}_s^H = \mathbf{U}_s^H\mathbf{U}_s = \mathbf{I}_L$.

The proposed scheme mainly involves the operations of the SVD in (15)–(17) and (35) and the inversion of the matrix $\mathbf{B}$ in (28). The computational complexity of the proposed scheme, which is measured in terms of floating-point operations (FLOPs), is summarized in Table I.

B. Special Case: Cooperative Beamforming

In this section, we consider the cooperative beamforming in a two-hop cooperative system. This is a special case of our precoding problem in which $L=1$ and the direct link is not considered (i.e., $\mathbf{H}_{SD} = \mathbf{0}$).

For a given source beamforming vector $\mathbf{f}_S$, the optimal relay precoder can be derived by [25]

$$\mathbf{F}_R = \mathbf{V}_{rd}\mathbf{\Sigma}_r\mathbf{U}_{sr}^H \tag{38}$$

where $\mathbf{\Sigma}_r = \mathrm{diag}\{\sigma_{r,1},\ldots,\sigma_{r,R}\}$, with $\sigma_{r,1} \geq \cdots \geq \sigma_{r,R}$.

Let $\mathbf{f}_S = \sqrt{\alpha_s}\mathbf{v}_S \in \mathbb{C}^{N\times 1}$, where $\mathbf{v}_S$ is a unit vector, and $\mathbf{f}_S$ satisfies the transmit power constraint, i.e., $\sigma_s^2\alpha_s\|\mathbf{v}_S\|^2 \leq P_{S,T}$. Substituting the beamformer and (38) into (10) with $\mathbf{H}_{SD} = \mathbf{0}$, we have (39), shown at the bottom of the page, where $\mathbf{w}_{sr} = \mathbf{V}_{sr}^H\mathbf{v}_S = [w_{sr,1},\ldots,w_{sr,N}]^T$ and $\|\mathbf{w}_{sr}\|^2 = 1$, $\mathbf{\Sigma}_{sr}$, $\mathbf{\Sigma}_{rd}$, $\mathbf{\Sigma}_r$ are diagonal matrices, with their diagonal elements arranged in decreasing order. The beamforming problem can then be formulated as (40), shown at the bottom of the page.

Theorem 1: The optimal beamforming vector denoted by $\mathbf{f}_S^*$ and the optimal relay precoder denoted

$$\begin{aligned}
J_{\min} &= \mathrm{tr}\left\{\left(\sigma_s^{-2} + \mathbf{f}_S^H\mathbf{V}_{sr}\mathbf{\Sigma}_{sr}^H\mathbf{\Sigma}_r^H\mathbf{\Sigma}_{rd}^H\left(\sigma_{n,d}^2\mathbf{I}_M + \sigma_{n,r}^2\mathbf{\Sigma}_{rd}\mathbf{\Sigma}_r\mathbf{\Sigma}_r^H\mathbf{\Sigma}_{rd}^H\right)^{-1}\mathbf{\Sigma}_{rd}\mathbf{\Sigma}_r\mathbf{\Sigma}_{sr}\mathbf{V}_{sr}^H\mathbf{f}_S\right)^{-1}\right\} \\
&= \mathrm{tr}\left\{\left(\sigma_s^{-2} + \alpha_s \underbrace{\mathbf{v}_S^H\mathbf{V}_{sr}}_{:=\mathbf{w}_{sr}^H}\mathbf{\Sigma}_{sr}^H\mathbf{\Sigma}_r^H\mathbf{\Sigma}_{rd}^H\left(\sigma_{n,d}^2\mathbf{I}_M + \sigma_{n,r}^2\mathbf{\Sigma}_{rd}\mathbf{\Sigma}_r\mathbf{\Sigma}_r^H\mathbf{\Sigma}_{rd}^H\right)^{-1}\mathbf{\Sigma}_{rd}\mathbf{\Sigma}_r\mathbf{\Sigma}_{sr}\underbrace{\mathbf{V}_{sr}^H\mathbf{v}_S}_{:=\mathbf{w}_{sr}}\right)^{-1}\right\} \\
&= \frac{1}{\sigma_s^{-2} + \alpha_s\sum_{i=1}^{\min\{N,M,R\}}|w_{sr,i}|^2\frac{\sigma_{r,i}^2\sigma_{sr,i}^2\sigma_{rd,i}^2}{\sigma_{n,d}^2+\sigma_{n,r}^2\sigma_{r,i}^2\sigma_{rd,i}^2}}
\end{aligned} \tag{39}$$

$$\begin{aligned}
\min_{\alpha_s, w_{sr,i}, \sigma_{r,i}, \forall i} \quad & \frac{1}{\sigma_s^{-2} + \alpha_s\sum_{i=1}^{\min\{N,M,R\}}|w_{sr,i}|^2\frac{\sigma_{r,i}^2\sigma_{sr,i}^2\sigma_{rd,i}^2}{\sigma_{n,d}^2+\sigma_{n,r}^2\sigma_{r,i}^2\sigma_{rd,i}^2}} \\
\text{s.t.} \quad & \sigma_s^2\|\mathbf{f}_S\|^2 = \sigma_s^2\alpha_s\|\mathbf{v}_S\|^2 \leq P_{S,T}, \qquad \sum_{i=1}^{N}|w_{sr,i}|^2 = 1, \\
& \mathrm{tr}\left\{\mathbf{\Sigma}_r\left(\sigma_{n,r}^2\mathbf{I}_R + \sigma_s^2\mathbf{\Sigma}_{sr}\mathbf{V}_{sr}^H\mathbf{f}_S\mathbf{f}_S^H\mathbf{V}_{sr}\mathbf{\Sigma}_{sr}^H\right)\mathbf{\Sigma}_r^H\right\} \leq P_{R,T}
\end{aligned} \tag{40}$$

by $\mathbf{F}_R^*$ for (40) are $\sqrt{(P_{S,T}/\sigma_s^2)}\mathbf{V}_{sr}(:,1)$ and $\sqrt{(P_{R,T}/(\sigma_{n,r}^2+P_{S,T}\sigma_{sr,1}^2))}\mathbf{V}_{rd}(:,1)[\mathbf{U}_{sr}(:,1)]^H$, where $\mathbf{V}_{rd}(:,i)$ and $\mathbf{U}_{sr}(:,i)$ denote the ith columns of $\mathbf{V}_{rd}$ and $\mathbf{U}_{sr}$, respectively.

Proof: We first derive the optimal $\mathbf{w}_{sr}$ for given α_s and $\sigma_{r,i},\ i=1,\ldots,R$. From (40), it is simple to see that the optimal $\mathbf{w}_{sr}$ can be derived by the following equivalent problem:

$$\begin{aligned}\max_{\mathbf{w}_{sr}} \quad & \sum_{i=1}^{\min\{N,M,R\}} |w_{sr,i}|^2 \frac{\sigma_{r,i}^2\sigma_{sr,i}^2\sigma_{rd,i}^2}{\sigma_{n,d}^2+\sigma_{n,r}^2\sigma_{r,i}^2\sigma_{rd,i}^2}\\ \text{s.t.} \quad & \sum_{i=1}^{N} |w_{sr,i}| = 1. \end{aligned} \tag{41}$$

From (41), it is obvious that the optimum $\mathbf{w}_{sr}$ is $[1,0,\ldots,0]^T$. This can easily be checked by

$$\frac{\sigma_{r,i}^2\sigma_{sr,i}^2\sigma_{rd,i}^2}{\sigma_{n,d}^2+\sigma_{n,r}^2\sigma_{r,i}^2\sigma_{rd,i}^2} \geq \frac{\sigma_{r,j}^2\sigma_{sr,j}^2\sigma_{rd,j}^2}{\sigma_{n,d}^2+\sigma_{n,r}^2\sigma_{r,j}^2\sigma_{rd,j}^2}, \qquad i \geq j. \tag{42}$$

The solution implies that the optimal $\mathbf{v}_S$, which is denoted by $\mathbf{v}_S^*$, is

$$\mathbf{v}_S^* = \mathbf{V}_{sr}(:,1). \tag{43}$$

As a result, the design problem can therefore be expressed as

$$\begin{aligned}\max_{\alpha_s,\sigma_{r,i},\forall i} \quad & \alpha_s \frac{\sigma_{r,1}^2\sigma_{sr,1}^2\sigma_{rd,1}^2}{\sigma_{n,d}^2+\sigma_{n,r}^2\sigma_{r,1}^2\sigma_{rd,1}^2}\\ \text{s.t.} \quad & 0 \leq \alpha_s \leq \frac{P_{S,T}}{\sigma_s^2}\\ & \operatorname{tr}\left\{\mathbf{\Sigma}_r\left(\sigma_{n,r}^2\mathbf{I}_R+\alpha_s\sigma_s^2\mathbf{\Sigma}_{sr}\mathbf{V}_{sr}^H\mathbf{v}_S\mathbf{v}_S^H\mathbf{V}_{sr}\mathbf{\Sigma}_{sr}^H\right)\mathbf{\Sigma}_r^H\right\}\\ & = \left(\sigma_{n,r}^2\sum_{i=1}^{R}\sigma_{r,i}^2\right) + \alpha_s\sigma_s^2\sigma_{r,1}^2\sigma_{sr,1}^2 \leq P_{R,T}. \end{aligned} \tag{44}$$

Taking a close look at (44), we first find that the cost function is only related to α_s and $\sigma_{r,1}^2$. Then, we can have the following observation:

$$\alpha_s \frac{\sigma_{r,1}^2\sigma_{sr,1}^2\sigma_{rd,1}^2}{\sigma_{n,d}^2+\sigma_{n,r}^2\sigma_{r,1}^2\sigma_{rd,1}^2} \quad \text{is monotonous in} \quad \sigma_{r,1}^2 \text{ and } \alpha_s. \tag{45}$$

From (45) and (44), we can rewrite the power constraint for the relay as

$$\sigma_{r,1}^2\left(\sigma_{n,r}^2+\alpha_s^2\sigma_s^2\sigma_{sr,1}^2\right) \leq P_{R,T} \tag{46}$$

and consequently have the following relationship:

$$\sigma_{r,1}^2 \leq \frac{P_{R,T}}{\left(\sigma_{n,r}^2+\alpha_s^2\sigma_s^2\sigma_{sr,1}^2\right)}. \tag{47}$$

It is noteworthy that (46) also implies that the optimal $\mathbf{\Sigma}_r$, which is denoted by $\mathbf{\Sigma}_r^*$, is

$$\begin{aligned}\mathbf{\Sigma}_r^* &= \begin{bmatrix} \sigma_{r,1} & 0 & \cdots & 0\\ 0 & \sigma_{r,2} & \ddots & \vdots\\ \vdots & \ddots & \ddots & 0\\ 0 & \cdots & 0 & \sigma_{r,R}\end{bmatrix}\\ &= \begin{bmatrix} \sigma_{r,1} & 0 & \cdots & 0\\ 0 & 0 & \ddots & \vdots\\ \vdots & \ddots & \ddots & 0\\ 0 & \cdots & 0 & 0\end{bmatrix}. \end{aligned} \tag{48}$$

Substituting (47) into the cost function in (44), we have

$$\alpha_s \frac{\sigma_{r,1}^2\sigma_{sr,1}^2\sigma_{rd,1}^2}{\sigma_{n,d}^2+\sigma_{n,r}^2\sigma_{r,1}^2\sigma_{rd,1}^2} \leq \frac{\alpha_s\sigma_{sr,1}^2\sigma_{rd,1}^2}{\sigma_{n,d}^2\frac{\left(\sigma_{n,r}^2+\alpha_s\sigma_s^2\sigma_{sr,1}^2\right)}{P_{R,T}}+\sigma_{n,r}^2\sigma_{rd,1}^2} \tag{49}$$

where the upper bound of the cost function can be achieved if

$$\sigma_{r,1}^2 = \frac{P_{R,T}}{\left(\sigma_{n,r}^2+\alpha_s\sigma_s^2\sigma_{sr,1}^2\right)}. \tag{50}$$

Therefore, via (50), the problem can finally be expressed as the minimization of a function of α_s given by

$$\begin{aligned}\max_{\alpha_s} \quad & \frac{\sigma_{sr,1}^2\sigma_{rd,1}^2\alpha_s}{\frac{\sigma_{n,d}^2\sigma_s^2\sigma_{sr,1}^2}{P_{R,T}}\alpha_s+\sigma_{n,r}^2\sigma_{rd,1}^2+\frac{\sigma_{n,d}^2\sigma_{n,r}^2}{P_{R,T}}}\\ \text{s.t.} \quad & 0 \leq \alpha_s \leq \frac{P_{S,T}}{\sigma_s^2}. \end{aligned} \tag{51}$$

Since the cost function in (51) is monotonically increasing in α_s, it is clear that the optimal α_s, which is denoted by α_s^*, is

$$\alpha_s^* = \frac{P_{S,T}}{\sigma_s^2}. \tag{52}$$

Combining (52) and (43), we finally obtain the optimal beamforming vector as

$$\mathbf{f}_s^* = \sqrt{\frac{P_{S,T}}{\sigma_s^2}}\mathbf{V}_{sr}(:,1). \tag{53}$$

Substituting (52) into (50) and combining (48) and (38), we have the optimal relay precoder as

$$\sqrt{\frac{P_{R,T}}{\left(\sigma_{n,r}^2+P_{S,T}\sigma_{sr,1}^2\right)}}\mathbf{V}_{rd}(:,1)\left[\mathbf{U}_{sr}(:,1)\right]^H. \tag{54}$$

It is noteworthy that the result of Theorem 1 is the same as that in [28], in which the criterion for the beamformer design is the maximization of the received SNR. Here, we use the MMSE criterion and obtain the same solution.

IV. APPLICATIONS

The proposed precoding scheme can be used in many scenarios. In this section, we conduct simulations to evaluate the

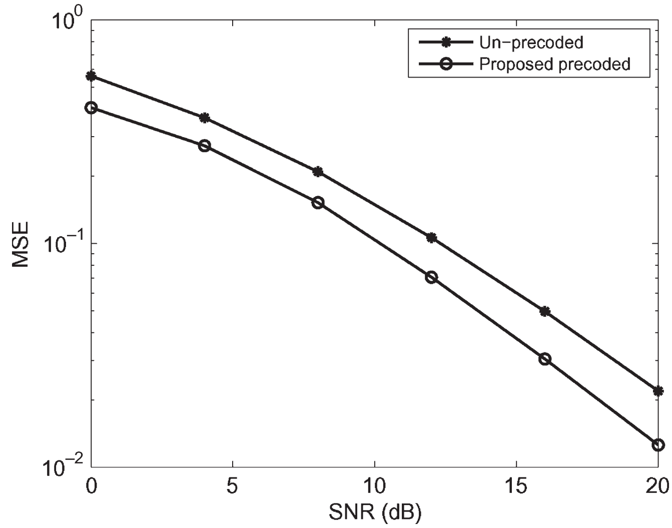


Fig. 2. MSE performance comparison for the unprecoded and proposed precoded schemes in an AF-based SISO OFDM cooperative system.

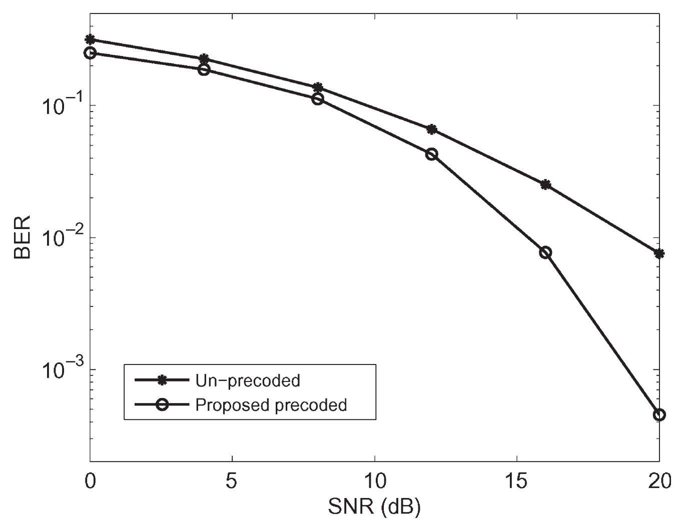


Fig. 3. BER performance comparison for the unprecoded and proposed precoded schemes in an AF-based SISO OFDM cooperative system.

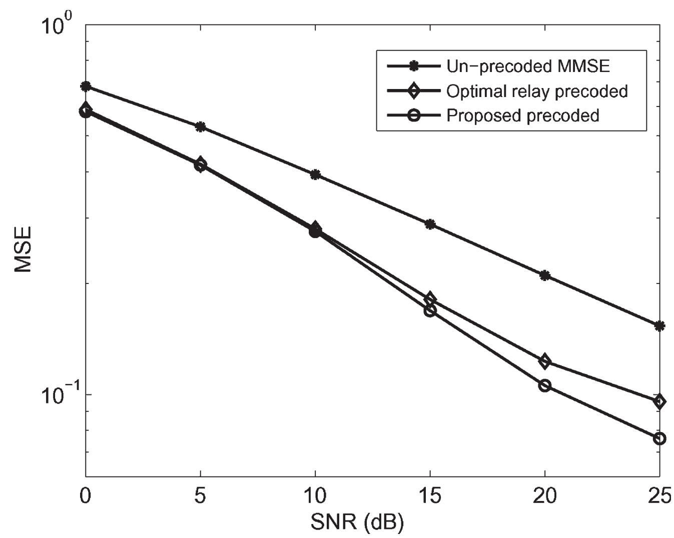


Fig. 4. MSE performance comparison for existing the unprecoded/precoded and proposed precoded schemes in an AF-based two-hop MIMO relay system.

performance of the proposed scheme in three different applications, namely, a single-input–single-output (SISO) orthogonal-frequency-division-multiplexing (OFDM) system, a two-hop MIMO relay system (where only the relay link is considered), and a general MIMO relay system. Assume that all channel state information (CSI) of all the links are available at all nodes and that perfect synchronization can be achieved. For the first case, the channel is assumed to be frequency-selective fading, and for the rest of the cases, the channel is assumed to be flat fading. Furthermore, the modulation scheme is quaternary phase-shift keying.

A. SISO OFDM Relay System

Assume that the cyclic prefix length is longer than the overall channel delay spread such that intersymbol interference will not occur. Furthermore, the channel is assumed to be quasi-static, which means that its response remains constant during each OFDM symbol. Note that each node only has one antenna. As a result, the equivalent frequency-domain channel matrices of all links are diagonal. The proposed precoders in (21) and (22) therefore become $\mathbf{F}_S = \mathbf{F}_L^H \boldsymbol{\Sigma}_s \mathbf{F}_L$ and $\mathbf{F}_R = \boldsymbol{\Sigma}_r$, where the relay precoder becomes a subcarrier power-allocation problem. Let $h_{sr}(l)$, $h_{rd}(l)$, and $h_{sd}(l)$ be the channel impulse responses for the source-to-relay, relay-to-destination, and source-to-destination channels, respectively. The channel taps $h_{sr}(l)$, $h_{rd}(l)$, and $h_{sd}(l)$, $0 \le l \le 5$, are generated from i.i.d. complex Gaussian random variables with zero mean and a variance of 1/6, such that $E\{\sum_{l=0}^{5} |h_{sr}(l)|^2\} = E\{\sum_{l=0}^{5} |h_{sd}(l)|^2\} E\{\sum_{l=0}^{5} |h_{rd}(l)|^2\} = 1$. Furthermore, let $N = 64$, the total available powers at the source and the relay be equal, and SNR_{sr}, SNR_{rd}, and SNR_{sd} be defined as the received SNR at the source-to-relay, relay-to-destination, and source-to-destination links. Here, we let $\mathrm{SNR}_{sr} = \mathrm{SNR}_{rd} = \mathrm{SNR}_{sd} = \mathrm{SNR}$. Figs. 2 and 3 show the MSE and BER comparisons for the unprecoded and proposed precoded systems, respectively. As shown in the figures, the proposed precoded system significantly outperforms the unprecoded system because the proposed system considers all the link resources and properly allocates the power.

B. Two-Hop MIMO Relay System

In this scenario, the channel condition in the direct link is poor such that the destination only receives the signal from the relay link. Here, we first consider the case that $N = R = M = L = 4$. Let the elements of each channel matrix be i.i.d. complex Gaussian random variables with zero mean and unity variance. Let SNR_{sr} and SNR_{rd} denote, respectively, the SNR per receive antenna of the source-to-relay and relay-to-destination links. Here, we set $\mathrm{SNR}_{sr} = 20$ dB and vary SNR_{rd}. Figs. 4 and 5 show the MSE and BER comparisons, respectively, for 1) an unprecoded system with a zero-forcing (ZF) receiver; 2) an unprecoded system with

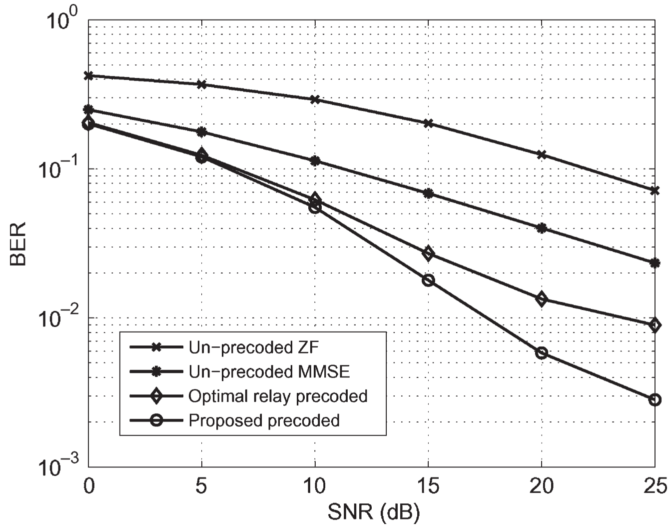


Fig. 5. BER performance comparison for the existing unprecoded/precoded and proposed precoded schemes in an AF-based two-hop MIMO relay system.

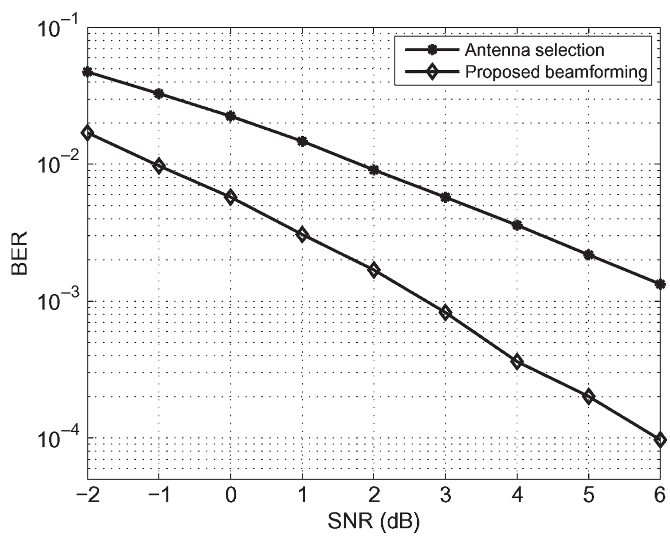


Fig. 6. BER performance comparison for the antenna selection [27] and proposed beamforming schemes in an AF-based two-hop MIMO relay system ($L = 1$ and $N = R = M = 4$).

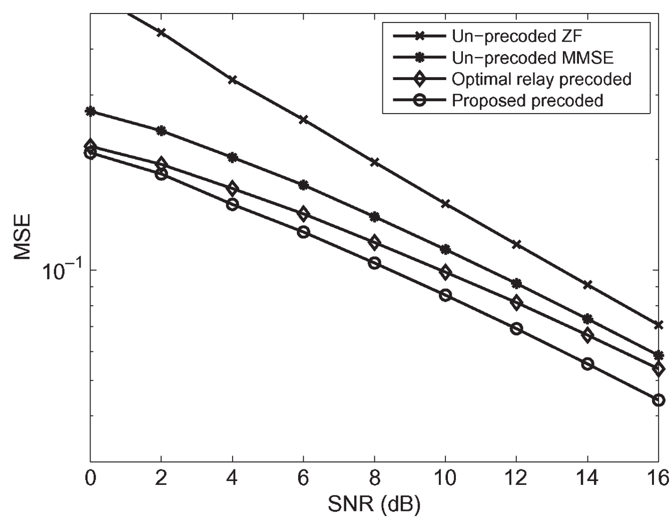


Fig. 7. MSE performance comparison for the existing unprecoded/precoded and proposed precoded schemes in an AF-based MIMO relay system.

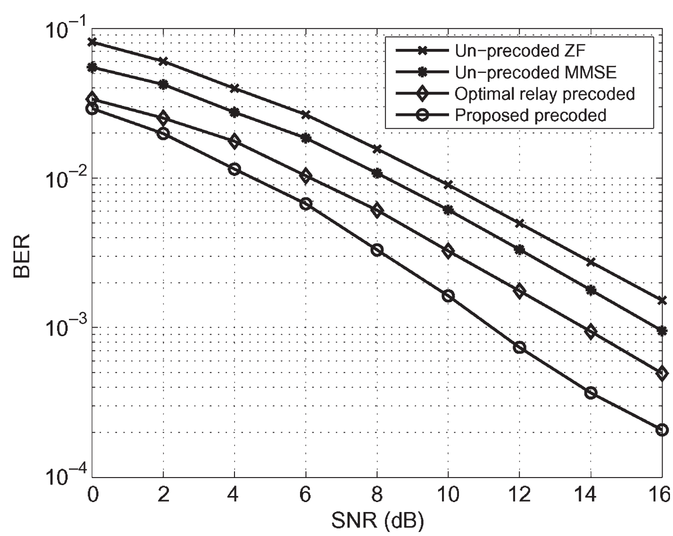


Fig. 8. BER performance comparison for the existing unprecoded/precoded and proposed precoded schemes in an AF-based MIMO relay system.

an MMSE receiver; 3) a precoded system with the optimal relay precoder [25]; and 4) the proposed precoded system. From those figures, we can see that the proposed precoded system outperforms not only the unprecoded system but the precoded system in [25] as well. This is because the proposed method incorporates an additional source precoder such that the performance can be enhanced even if the direct link is not considered.

We also report the simulation result for cooperative beamforming, i.e., $L = 1$. As discussed in Theorem 1, our design for this case is optimal. We let $N = R = M = 4$ and $\mathrm{SNR}_{sr} = 5$ dB. Fig. 6 shows the BER comparison for the antenna-selection method in [27] and the proposed method. From the figure, we can see that the proposed method is superior to the antenna selection. This is expected since the proposed beamforming scheme is optimal.

C. General MIMO Relay Channel

In this scenario, we consider a symmetric MIMO relay system, i.e., $N = M = R = L = 4$. As in the previous case, each element of the channel matrices is assumed to be i.i.d. complex Gaussian random variables with zero mean and the same variance. We let SNR_{sr} and SNR_{rd} be the same as those defined in Section IV-B and SNR_{sd} be the SNR per receive antenna for the source-to-destination link. Here, we set $\mathrm{SNR}_{sr} = 15$ dB and $\mathrm{SNR}_{rd} = 10$ dB and vary SNR_{sd}. Figs. 7 and 8 show the MSE and BER comparisons, respectively, for the proposed system and other systems described in Section IV-B. Note that the optimal relay precoder in [25] only considers the two-hop relay system. For fair comparison, we include the direct link at the destination when implementing the MMSE receiver. As expected, the proposed method outperforms all the other systems.

V. Conclusion

In this paper, we have proposed a joint design method for precoders in a half-duplex AF-based MIMO relay system. In the system, the MMSE receiver is used at the destination, and the precoders at the source and at the relay are determined to minimize the MMSE. Since the MMSE is a complicated function of the precoding matrices, a direct minimization is not feasible. To solve the problem, we have first used a constrained precoder structure and derived an MMSE upper bound. Since the upper bound has a simple expression, minimization with the upper bound becomes feasible. Resorting to the Karush–Kuhn–Tucker (KKT) optimality conditions, we have used an iterative water-filling algorithm to obtain a suboptimal solution. The proposed scheme can be applied in various kinds of communication systems as long as the channel effect can be described with an equivalent channel matrix. We have then considered the applications of the proposed scheme in the SISO OFDM, two-hop MIMO relay channel, and general MIMO relay systems. Simulations show that the proposed scheme outperforms the existing unprecoded/precoded systems in terms of either the MSE or the BER. In practical systems, perfect CSI may not be available. How to design the robust precoders for imperfect CSI can be a topic for further research.

Appendix A
Proof of Lemma (26)

Let us first rewrite

$$\mathrm{tr}\left(\mathbf{D}_1^H(\mathbf{X}+\mathbf{D}_2)^{-1}\mathbf{D}_1\right) = \sum_{i=1}^{L} |\mathbf{D}_1(i,i)|^2 (\mathbf{X}+\mathbf{D}_2)^{-1}(i,i) \tag{55}$$

where $\mathbf{D}_1 \in \mathbb{R}^{N\times L}$ and $\mathbf{D}_2 \in \mathbb{R}^{N\times N}$, $N \geq L$, are diagonal matrices with positive elements; $\mathbf{X} \in \mathbb{C}^{N\times N}$ is a Hermitian matrix. Therefore, $(\mathbf{X}+\mathbf{D}_2)$ is a positive definite matrix.

We claim that

$$(\mathbf{X}+\mathbf{D}_2)^{-1}(i,i) \geq \frac{1}{(\mathbf{X}+\mathbf{D}_2)(i,i)} \tag{56}$$

which will be proved in the next paragraph. From (55) and (56), we immediately have

$$\begin{aligned}\mathrm{tr}\left(\mathbf{D}_1^H(\mathbf{X}+\mathbf{D}_2)^{-1}\mathbf{D}_1\right) &\geq \sum_{i=1}^{L} \frac{|\mathbf{D}_1(i,i)|^2}{(\mathbf{X}+\mathbf{D}_2)(i,i)} \\ &= \sum_{i=1}^{L} \frac{|\mathbf{D}_1(i,i)|^2}{(\mathrm{diag}(\mathbf{X})+\mathbf{D}_2)(i,i)} \\ &= \mathrm{tr}\left(\mathbf{D}_1^H \left(\mathrm{diag}(\mathbf{X})+\mathbf{D}_2\right)^{-1}\mathbf{D}_1\right)\end{aligned} \tag{57}$$

which proves the lemma.

Proof of (56): Let $\mathbf{Z} := (\mathbf{X}+\mathbf{D}_2) = \mathbf{U\Sigma U}^H$ be the eigendecomposition of the positive definite matrix $\mathbf{Z}$. Since $1 = \mathbf{e}_i^T\mathbf{I}\mathbf{e}_i = \mathbf{e}_i^T\mathbf{Z}^{1/2}\mathbf{Z}^{-1/2}\mathbf{e}_i$, where $\mathbf{e}_i$ is the ith unit standard vector, we then have

$$1 = \left\|\mathbf{e}_i^T\mathbf{Z}^{1/2}\mathbf{Z}^{-1/2}\mathbf{e}_i\right\|_2^2 \leq \left\|\mathbf{e}_i^T\mathbf{Z}^{1/2}\right\|_2^2 \|\mathbf{Z}^{-1/2}\mathbf{e}_i\|_2^2 \tag{58}$$

where the inequality in (58) follows from the submultiplicative property of the matrix norm [29]. Since $\|\mathbf{e}_i^T\mathbf{Z}^{1/2}\|_2^2 = (\mathbf{e}_i^T\mathbf{Z}^{1/2}\mathbf{Z}^{1/2}\mathbf{e}_i) = \mathbf{Z}(i,i)$ and $\|\mathbf{e}_i^T\mathbf{Z}^{-1/2}\|_2^2 = (\mathbf{e}_i^T\mathbf{Z}^{-1/2}\mathbf{Z}^{-1/2}\mathbf{e}_i) = \mathbf{Z}^{-1}(i,i)$, the inequality in (58) thus leads to $1 \leq \mathbf{Z}(i,i)\mathbf{Z}^{-1}(i,i)$ or, equivalenty, $\mathbf{Z}^{-1}(i,i) \geq (1/\mathbf{Z}(i,i))$.

Appendix B
Derivation of (30) and (31)

The Lagrangian function with respect to (29) can be written as

$$\begin{aligned}L = &\sum_{i=1}^{L} \frac{1}{\sigma_s^{-2} + \frac{p_{s,i}p_{r,i}\sigma_{sr,i}^2\sigma_{rd,i}^2}{\sigma_{n,r}^2 p_{r,i}\sigma_{rd,i}^2+\sigma_{n,d}^2} + p_{s,i}\left(\mathbf{B}^{-1}(i,i)\right)^{-1}} \\ &+ \lambda_s\left(\sigma_s^2\sum_{i=1}^{L} p_{s,i} - P_{S,T}\right) \\ &+ \lambda_r\left(\sum_{i=1}^{L} p_{r,i}\left(\sigma_{n,r}^2 + \sigma_s^2 p_{s,i}\sigma_{sr,i}^2\right) - P_{R,T}\right) \\ &- \sum_{i=1}^{L}\mu_{s,i}p_{s,i} - \sum_{i=1}^{L}\mu_{r,i}p_{r,i}.\end{aligned} \tag{59}$$

As mentioned, if $p_{s,i}$ is given, (29) is a convex optimization problem (for $p_{r,i}$). Thus, we can obtain the optimum $p_{r,i}$ using the KKT conditions [30]. The KKT optimality conditions for solving $p_{r,i}$, $1 \leq i \leq L$ are given as follows:

$$-\frac{p_{s,i}\sigma_{n,d}^2\sigma_{sr,i}^2\sigma_{rd,i}^2}{c(p_{s,i},p_{r,i})} + \lambda_r\left(\sigma_{n,r}^2 + \sigma_s^2 p_{s,i}\sigma_{sr,i}^2\right) - \mu_{r,i} = 0 \tag{60}$$

where

$$\begin{aligned}c(p_{s,i},p_{r,i}) = \Big[&\left(\sigma_s^{-2} + p_{s,i}\left(\mathbf{B}^{-1}(i,i)\right)^{-1}\right) \\ &\times\left(p_{r,i}\sigma_{n,r}^2\sigma_{rd,i}^2 + \sigma_{n,d}^2\right) + p_{s,i}p_{r,i}\sigma_{sr,i}^2\sigma_{rd,i}^2\Big]^2\end{aligned} \tag{61}$$

$$\mu_{r,i} \geq 0 \tag{62}$$

$$\lambda_r \geq 0 \tag{63}$$

$$\mu_{r,i}p_{r,i} = 0 \tag{64}$$

$$\lambda_r\left(\sum_{i=1}^{L} p_{r,i}\left(\sigma_{n,r}^2 + \sigma_s^2 p_{s,i}\sigma_{sr,i}^2\right) - P_{R,T}\right) = 0. \tag{65}$$

$$p_{r,i} = \frac{\frac{\sqrt{p_{s,i}}\sigma_{n,d}\sigma_{sr,i}\sigma_{rd,i}}{\lambda_r^{1/2}\left(\sigma_{n,r}^2+\sigma_s^2 p_{s,i}\sigma_{sr,i}^2\right)^{1/2}} - \sigma_{n,d}^2\left(\sigma_s^{-2}+p_{s,i}\left(\mathbf{B}^{-1}(i,i)\right)^{-1}\right)}{\sigma_{rd,i}^2\left(\sigma_{n,r}^2\left(\sigma_s^{-2}+p_{s,i}\left(\mathbf{B}^{-1}(i,i)\right)^{-1}\right)+p_{s,i}\sigma_{sr,i}^2\right)} \tag{68}$$

Combining (60) and (62), we have

$$\lambda_r\left(\sigma_{n,r}^2+\sigma_s^2 p_{s,i}\sigma_{sr,i}^2\right) \geq \frac{p_{s,i}\sigma_{n,d}^2\sigma_{sr,i}^2\sigma_{rd,i}^2}{c(p_{s,i},p_{r,i})}. \tag{66}$$

Substituting (60) into (64) leads to

$$p_{r,i}\left(\lambda_r\left(\sigma_{n,r}^2+\sigma_s^2 p_{s,i}\sigma_{sr,i}^2\right) - \frac{p_{s,i}\sigma_{n,d}^2\sigma_{sr,i}^2\sigma_{rd,i}^2}{c(p_{s,i},p_{r,i})}\right) = 0. \tag{67}$$

To satisfy (67), we then have the following.

1) If

$$\lambda_r\left(\sigma_{n,r}^2+\sigma_s^2 p_{s,i}\sigma_{sr,i}^2\right) > \frac{p_{s,i}\sigma_{n,d}^2\sigma_{sr,i}^2\sigma_{rd,i}^2}{c(p_{s,i},p_{r,i})}$$

then $p_{r,i}=0$.

2) If

$$\lambda_r\left(\sigma_{n,r}^2+\sigma_s^2 p_{s,i}\sigma_{sr,i}^2\right) = \frac{p_{s,i}\sigma_{n,d}^2\sigma_{sr,i}^2\sigma_{rd,i}^2}{c(p_{s,i},p_{r,i})}$$

then (68), shown at the top of the page, holds.

Considering items 1 and 2 and $p_{r,i}\geq 0$, we then find the solution of $p_{r,i}$ as (30), where $\mu_r=\lambda_r^{-1/2}$ is the water level that should be chosen to satisfy the power constraint at the relay. Similarly, we can obtain the optimum $p_{s,i}$ for a given $p_{r,i}$, as shown in (31). The details, however, are omitted.

References

[1] I. E. Telatar, "Capacity of multiantenna Gaussian channel," *Eur. Trans. Telecommun.*, vol. 6, pp. 311–335, 1998.

[2] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 5, pp. 684–702, Jun. 2003.

[3] M. Joham, W. Utschick, and J. A. Nossek, "Linear transmit processing in MIMO communications systems," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 2700–2712, Aug. 2005.

[4] 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.

[5] D. P. Palomar, J. M. Cioffi, and M. A. Laguna, "Joint Tx-Rx beamforming design for multicarrier MIMO channels: A unified framework for convex optimization," *IEEE Trans. Signal Process.*, vol. 51, no. 9, pp. 2381–2401, Sep. 2003.

[6] D. P. Palomar and S. Barbarossa, "Designing MIMO communication systems: Constellation choice and linear transceiver design," *IEEE Trans. Signal Process.*, vol. 53, no. 10, pp. 3804–3818, Oct. 2005.

[7] Y. Jiang, J. Li, and W. W. Hager, "Joint transceiver design for MIMO communications using geometric mean decomposition," *IEEE Trans. Signal Process.*, vol. 53, no. 10, pp. 3791–3803, Oct. 2005.

[8] J. K. Zhang, A. Kavcic, and K. M. Wong, "Equal-diagonal QR decomposition and its application to precoder design for successive cancellation detection," *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 154–172, Jan. 2005.

[9] N. Jindal, "Sum power iterative water-filling for multi-antenna Gaussian broadcast channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 4, pp. 1570–1580, Apr. 2005.

[10] C. B. Chae and R. W. Heath, Jr., "On the optimality of linear multiuser MIMO beamforming for a two-user two-input multiple-output broadcast system," *IEEE Signal Process. Lett.*, vol. 16, no. 2, pp. 117–120, Feb. 2009.

[11] A. Tolli, M. Codreanu, and M. Juntti, "Linear multiuser MIMO transceiver design with quality of service and per-antenna constraints," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 3049–3055, Jul. 2008.

[12] X. Zhang, Z. Lv, and W. Wang, "Performance analysis of multiuser diversity in MIMO systems with antenna selection," *IEEE Trans. Wireless Commun.*, vol. 7, no. 1, pp. 15–21, Jan. 2008.

[13] R. Chen, Z. Shen, J. G. Andrews, and R. W. Heath, "Multimode transmission for multiuser MIMO systems with block diagonalization," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 3294–3302, Jul. 2008.

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

[15] Y. Jing and H. Jafarkhani, "Network beamforming with channel mean and covariance at relays," in *Proc. IEEE ICC*, Beijing, China, May 19–23, 2008, pp. 3743–3747.

[16] Y. Jing and H. Jafarkhani, "Beamforming in wireless relay networks," in *Proc. ITA Workshop*, La Jolla, CA, Jan. 27–Feb. 1, 2008, pp. 142–150.

[17] Y. Jing and H. Jafarkhani, "Network beamforming using relays with perfect channel information," in *Proc. IEEE ICASSP*, Honolulu, HI, Apr. 15–20, 2007, pp. III-473–III-476.

[18] Z. Yi and I. M. Kim, "Joint optimization of relay-precoders and decoders with partial channel side information in cooperative networks," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 2, pp. 447–458, Feb. 2007.

[19] E. Koyuncu, Y. Jing, and H. Jafarkhani, "Distributed beamforming in wireless relay networks with quantized feedback," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 8, pp. 1429–1439, Oct. 2008.

[20] Y. Jing and B. Hassibi, "Distributed space-time coding in wireless relay networks," *IEEE Trans. Wireless Commun.*, vol. 5, no. 12, pp. 1524–1536, Dec. 2006.

[21] Y. Ding, J. K. Zhang, and K. M. Wong, "The amplify-and-forward half-duplex cooperative system: Pairwise error probability and precoder design," *IEEE Trans. Signal Process.*, vol. 55, no. 2, pp. 605–617, Feb. 2007.

[22] B. Wang, J. Zhang, and A. Høst-Madsen, "On the capacity of MIMO relay channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 29–43, Jan. 2005.

[23] 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.

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

[25] W. Guan and H. Luo, "Joint MMSE transceiver design in non-regenerative MIMO relay systems," *IEEE Commun. Lett.*, vol. 12, no. 7, pp. 517–519, Jul. 2008.

[26] A. S. Behbahani, R. Merched, and A. M. Eltawil, "Optimization of a MIMO relay network," *IEEE Trans. Signal Process.*, vol. 56, no. 10, pp. 5062–5073, Oct. 2008.

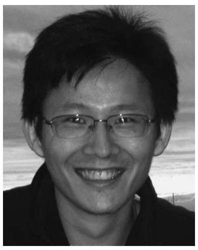
[27] S. W. Peters and R. W. Heath, "Nonregenerative MIMO relaying with optimal transmit antenna selection," *IEEE Signal Process. Lett.*, vol. 15, pp. 421–424, 2008.

[28] B. Khoshnevis and R. Adve, "Grassmannian beamforming for MIMO amplify-and-forward relaying," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 8, pp. 1397–1407, Oct. 2008.

[29] D. S. Bernstein, *Matrix Mathematics*. Princeton, NJ: Princeton Univ. Press, 2005.

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

[31] W. Yu, “Multiuser waterfilling in the presence of crosstalk,” in *Proc. Inf. Theory Appl. Workshop*, San Diego, CA, Jan./Feb. 2007, pp. 414–420.
[32] G. Scutari, D. P. Palomar, and S. Barbarossa, “The MIMO iterative waterfilling algorithm,” *IEEE Trans. Signal Process.*, vol. 57, no. 5, pp. 1917–1935, May 2009.



Fan-Shuo Tseng (S’08) was born in Kaohsiung, Taiwan, in November 1980. He received the B.S. degree from Fu Jen Catholic University, Taipei, Taiwan, in 2002 and the M.S. degree from National Central University, Taoyuan, Taiwan, in 2004. He is currently working toward the Ph.D. degree with the Department of Electrical Engineering, National Chiao Tung University, Hsinchu, Taiwan.

His research interests include transceiver design in cooperative communication and statistical signal processing in digital communication.

Wen-Rong Wu (M’89) received the B.S. degree in mechanical engineering from Tatung Institute of Technology, Taipei, Taiwan, in 1980 and the M.S. degree in mechanical engineering, the M.S. degree in electrical engineering, and the Ph.D. degree in electrical engineering from the State University of New York, Buffalo, in 1985, 1986, and 1989, respectively.

Since August 1989, he has been a faculty member with the Department of Electrical Engineering, National Chiao Tung University, Hsinchu, Taiwan. His research interests include statistical signal processing and digital communication.