Simplified Algorithms for Optimizing Multiuser Multi-Hop MIMO Relay Systems

Yue Rong, *Senior Member, IEEE*

***Abstract*—In this paper, we address the issue of multiaccess communication through multi-hop linear non-regenerative relays, where all users, all relay nodes, and the destination node may have multiple antennas. Using a linear minimal mean-squared error (MMSE) receiver at the destination node, we demonstrate that the optimal amplifying matrix at each relay node can be viewed as a linear MMSE filter concatenated with another linear filter. As a consequence, the MSE matrix of the signal waveform estimation at the destination node is decomposed into the sum of the MSE matrices at all relay nodes. We show that at a high signal-to-noise ratio (SNR) environment, this MSE matrix decomposition significantly simplifies the solution to the problem of optimizing the source precoding matrices and relay amplifying matrices. Simulation results show that even at the low to medium SNR range, the simplified optimization algorithms have only a marginal performance degradation but a greatly reduced computational complexity and signalling overhead compared with the existing optimal iterative algorithm, and thus are of great interest for practical relay systems.**

***Index Terms*—MIMO relay, multi-hop relay, MMSE, multiuser.**

I. INTRODUCTION

NON-REGENERATIVE multiple-input multiple-output (MIMO) relay communication systems recently have attracted much research interest [1]-[7]. For a single-user multi-hop MIMO relay system with any number of hops, the optimality of channel diagonalization has been proved in [3]. Based on this diagonalization property, an iterative algorithm was developed in [3] to optimize the power allocation at all data streams and all nodes. For a multiuser relay system, the achievable sum rate has been derived in [4], assuming that each user is equipped with a single antenna. In [5], the optimal relay amplifying matrix and source precoding matrices were developed to maximize the sum source-destination mutual information of a two-hop multiuser relay system, where the users and the relay node are equipped with multiple antennas. Recently, a minimal mean-squared error (MMSE)-based optimal multiuser MIMO relay system has been proposed [6]. The quality-of-service constraints in a multi-antenna relay broadcast channel were investigated in [7].

In this paper, we focus on multiaccess communication through multi-hop linear non-regenerative relays. In contrast to [4], we consider a relay system where all users, all relay nodes, and the destination node may have multiple antennas. Using a linear MMSE receiver at the destination node, we show that the optimal amplifying matrix at each relay node can be viewed as a linear MMSE filter concatenated with another linear filter. As a consequence, the MSE matrix of the signal waveform estimation at the destination node can be decomposed into the sum of the MSE matrices at all relay nodes. A very useful application of such decomposition is that it greatly simplifies the source precoding matrices and relay amplifying matrices optimization problem at a (moderately) high signal-to-noise ratio (SNR) environment. In particular, it enables the power allocation optimization to be performed at each relay node in a distributed manner, which requires only local channel state information (CSI) knowledge. Thus the simplified relay algorithms proposed in this paper have a significant reduction in both the computational complexity and the signalling overhead compared with the iterative algorithm developed in [6]. Simulation results show that even at the low to medium SNR range, the simplified optimization algorithms only slightly increase the MSE of the signal waveform estimation and the system bit-error-rate (BER), but greatly reduce the computational complexity (less than the complexity of carrying out one iteration of the algorithm in [6]). Thus, the simplified optimization algorithms are of great interest for practical relay systems. The multiuser multi-hop relay algorithms developed in this paper can be applied in multi-hop wireless backhaul networks where high data rate wireless links need to be established over a long distance with the aid of fixed relay nodes. Note that multi-hop wireless backhaul networks are being considered by several industry standards such as IEEE802.16j [8], [9].

We would like to mention that the decomposition of the MSE matrix was first discovered in [10] for a single-user two-hop MIMO relay system. Our paper generalizes [10] from single-user two-hop MIMO relay system to multiuser multi-hop MIMO relay systems with any number of hops and any number of users. Note that due to the introduction of multiusers and multiple relay nodes, a rigorous proof of the MSE matrix decomposition for multi-hop MIMO relay system is much more challenging than that for the two-hop MIMO channel, and is one contribution of this paper. The generalization from a single-user two-hop MIMO system to multiuser multi-hop MIMO relay systems is significant. Note that although in this paper we focus on uplink multiaccess systems, the downlink broadcast system can be designed by exploiting the uplink-downlink duality for multi-hop linear non-regenerative MIMO relay systems established in [11].

Paper approved by O. Oyman, the Editor for Cooperative and Heterogeneous Networks of the IEEE Communications Society. Manuscript received December 7, 2010; revised March 15, 2011 and May 9, 2011.

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

This research was supported under the Australian Research Council's *Discovery Projects* funding scheme (project numbers DP110100736, DP110102076).

Digital Object Identifier 10.1109/TCOMM.2011.081111.100756

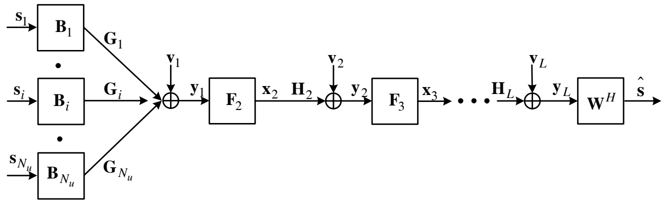


Fig. 1. Block diagram of an N_u-user L-hop linear non-regenerative MIMO relay communication system.

In this paper, for notational convenience, we consider a narrow band single-carrier system. However, our results can be straightforwardly generalized to broadband multi-carrier multi-hop MIMO relay systems as in the case of two-hop MIMO relay system shown in [2].

The rest of this paper is organized as follows. In Section II, we introduce the model of a multi-hop linear non-regenerative multiaccess MIMO relay communication system. The proposed source and relay design algorithms are presented in Section III. In Section IV, we show some numerical examples. Conclusions are drawn in Section V.

II. System Model

We consider a multiaccess (uplink) system where N_u users simultaneously transmit information to a common destination node equipped with a linear receiver as shown in Fig. 1. Due to the long source-destination distance, $L-1$ relay nodes are applied in serial to relay signals from all users to the destination node, where the lth relay node is equipped with N_l antennas, $l=1,\cdots,L-1$, and the destination node has N_L antennas. The ith user transmits M_i independent data streams using M_i antennas, $i=1,\cdots,N_u$. We denote $N_0=\sum_{i=1}^{N_u} M_i$ as the total number of independent data streams from all users. For a linear non-regenerative MIMO relay system, there should be $N_0 \leq \min(N_1,\cdots,N_L)$, since otherwise the system can not support N_0 active symbols in each transmission. Such condition is imposed by the inherent physical property of the MIMO channel (which is true also for classical single-hop MIMO communication systems [12]), but not from our algorithms developed later. In systems with more than $L-1$ potential relays, relay nodes having enough number of antennas should be chosen to establish an L-hop communication system.

In this paper, we focus on the non-regenerative relay strategy as in [1]-[7] and [10]-[11] due to the following two reasons. First, in the non-regenerative strategy, the relay node only amplifies and retransmits its received signal. Thus, the complexity of the non-regenerative strategy is much lower than that of the regenerative strategy. This advantage is particularly important when all nodes are equipped with multiple antennas, since decoding multiple data streams involves much more computational efforts than decoding a single data stream. Second, for multi-hop relay networks, the delay introduced by the regenerative strategy due to decoding is much larger than that of the non-regenerative strategy.

At the ith user, the $M_i \times 1$ modulated signal vector $\mathbf{s}_i$ is linearly precoded by the $M_i \times M_i$ source precoding matrix $\mathbf{B}_i$, and the precoded signal vector $\mathbf{u}_i=\mathbf{B}_i\mathbf{s}_i$ is transmitted to the first relay node. The received signal vector at the first relay node is given by

$$\mathbf{y}_1=\sum_{i=1}^{N_u}\mathbf{G}_i\mathbf{u}_i+\mathbf{v}_1 \triangleq \mathbf{H}_1\mathbf{x}_1+\mathbf{v}_1 \tag{1}$$

where $\mathbf{G}_i$, $i=1,\cdots,N_u$, is the $N_1\times M_i$ MIMO channel matrix between the first relay node and the ith user, $\mathbf{v}_1$ is the $N_1\times 1$ independent and identically distributed (i.i.d.) additive white Gaussian noise (AWGN) vector at the first relay node, $\mathbf{x}_1=\mathbf{F}_1\mathbf{s}$, $\mathbf{s}\triangleq\left[\mathbf{s}_1^T,\cdots,\mathbf{s}_{N_u}^T\right]^T$, and

$$\mathbf{H}_1\triangleq[\mathbf{G}_1,\cdots,\mathbf{G}_{N_u}], \qquad \mathbf{F}_1\triangleq \mathrm{bd}(\mathbf{B}_1,\cdots,\mathbf{B}_{N_u}). \tag{2}$$

Here $\mathbf{H}_1$ is the equivalent $N_1\times N_0$ first-hop MIMO channel, $\mathbf{F}_1$ is the equivalent $N_0\times N_0$ block diagonal source precoding matrix, $\mathbf{s}$ is an $N_0\times 1$ vector containing source symbols from all users, $\mathrm{bd}(\cdot)$ stands for a block diagonal matrix, and $(\cdot)^T$ denotes matrix (vector) transpose. We assume that $\mathrm{E}[\mathbf{s}\mathbf{s}^H]=\mathbf{I}_{N_0}$, where $\mathrm{E}[\cdot]$ stands for the statistical expectation, $(\cdot)^H$ denotes the Hermitian transpose, and $\mathbf{I}_n$ is an $n\times n$ identity matrix.

Due to its simplicity, a linear nonregenerative relay matrix is used at each relay as in [1]-[7]. The input-output relationship at the lth relay nodes is

$$\mathbf{x}_{l+1}=\mathbf{F}_{l+1}\mathbf{y}_l, \qquad l=1,\cdots,L-1 \tag{3}$$

where $\mathbf{F}_{l+1}$, $l=1,\cdots,L-1$, is the $N_l\times N_l$ amplifying matrix at the lth relay node, and $\mathbf{y}_l$, $l=1,\cdots,L-1$, is the $N_l\times 1$ signal vector received at the lth relay node written as

$$\mathbf{y}_l=\mathbf{H}_l\mathbf{x}_l+\mathbf{v}_l, \qquad l=1,\cdots,L-1 \tag{4}$$

where $\mathbf{H}_l$, $l=1,\cdots,L-1$, is the $N_l\times N_{l-1}$ MIMO channel matrix of the lth hop, and $\mathbf{v}_l$ is the i.i.d. AWGN vector at the lth relay node. Finally, at the last hop, the signal vector received at the destination node is given by (4) with $l=L$. We assume that all noises are complex circularly symmetric with zero mean and unit variance. From (1)-(4), we have [3]

$$\mathbf{y}_l=\mathbf{A}_l\mathbf{s}+\bar{\mathbf{v}}_l, \qquad l=1,\cdots,L$$

where $\mathbf{A}_l$ is the equivalent MIMO channel matrix from the source to the lth hop, and $\bar{\mathbf{v}}_l$ is the equivalent noise vector given by [3]

$$\mathbf{A}_l=\bigotimes_{i=l}^{1}(\mathbf{H}_i\mathbf{F}_i), \qquad l=1,\cdots,L \tag{5}$$

$$\bar{\mathbf{v}}_1=\mathbf{v}_1,\ \ \bar{\mathbf{v}}_l=\sum_{j=2}^{l}\Big(\bigotimes_{i=l}^{j}(\mathbf{H}_i\mathbf{F}_i)\mathbf{v}_{j-1}\Big)+\mathbf{v}_l,\ \ l=2,\cdots,L. \tag{6}$$

Here for matrices $\mathbf{X}_i$, $\bigotimes_{i=l}^{k}(\mathbf{X}_i)\triangleq\mathbf{X}_l\cdots\mathbf{X}_k$.

From (6), the covariance matrix of $\bar{\mathbf{v}}_l$, $\mathbf{C}_l=\mathrm{E}[\bar{\mathbf{v}}_l\bar{\mathbf{v}}_l^H]$, $l=1,\cdots,L$, is given by

$$\mathbf{C}_1=\mathbf{I}_{N_1}$$

$$\mathbf{C}_l=\sum_{j=2}^{l}\Big(\bigotimes_{i=l}^{j}(\mathbf{H}_i\mathbf{F}_i)\bigotimes_{i=j}^{l}(\mathbf{F}_i^H\mathbf{H}_i^H)\Big)+\mathbf{I}_{N_l},\ l=2,\cdots,L.$$

We would like to mention that the system model in this paper extends the model of multiuser MIMO relay systems in [4],

[5], [7] from two hops to multiple hops. Such extension is important in the case of long source-destination distance where a two-hop relay is not sufficient and a multi-hop relay is necessary to establish a reliable source-destination link.

III. Proposed Source and Relay Design Algorithms

With a linear receiver at the destination node, the estimated signal vector is given by $\hat{\mathbf{s}} = \mathbf{W}_L^H \mathbf{y}_L$, where $\mathbf{W}_L$ is the $N_L \times N_0$ weight matrix. The weight matrix of the linear MMSE receiver [13] is $\mathbf{W}_L = (\mathbf{A}_L\mathbf{A}_L^H + \mathbf{C}_L)^{-1}\mathbf{A}_L$ [3], where $(\cdot)^{-1}$ stands for the matrix inversion. Using this MMSE receiver, the MSE matrix $\mathbf{E}_L$ at the destination node is given by [3]

$$\begin{aligned}\mathbf{E}_L &= \left(\mathbf{I}_{N_0} + \mathbf{A}_L^H\mathbf{C}_L^{-1}\mathbf{A}_L\right)^{-1} \\ &= \Bigg[\mathbf{I}_{N_0} + \bigotimes_{i=1}^{L}(\mathbf{F}_i^H\mathbf{H}_i^H)\Bigg(\sum_{l=2}^{L}\Big(\bigotimes_{i=L}^{l}(\mathbf{H}_i\mathbf{F}_i) \\ &\qquad \bigotimes_{i=l}^{L}(\mathbf{F}_i^H\mathbf{H}_i^H)\Big) + \mathbf{I}_{N_L}\Bigg)^{-1}\bigotimes_{i=L}^{1}(\mathbf{H}_i\mathbf{F}_i)\Bigg]^{-1}. \end{aligned} \quad (7)$$

Let us introduce matrices

$$\mathbf{D}_l \triangleq \mathbf{A}_l\mathbf{A}_l^H + \mathbf{C}_l = \sum_{j=1}^{l}\Big(\bigotimes_{i=l}^{j}(\mathbf{H}_i\mathbf{F}_i)\bigotimes_{i=j}^{l}(\mathbf{F}_i^H\mathbf{H}_i^H)\Big) + \mathbf{I}_{N_l}, \quad l = 1, \cdots, L-1. \quad (8)$$

It can be shown from [3] that the transmission power consumed by the lth relay node is

$$\mathrm{tr}\left(\mathrm{E}\left[\mathbf{x}_{l+1}\mathbf{x}_{l+1}^H\right]\right) = \mathrm{tr}\left(\mathbf{F}_{l+1}\mathbf{D}_l\mathbf{F}_{l+1}^H\right), \; l = 1, \cdots, L-1 \quad (9)$$

where $\mathrm{tr}(\cdot)$ denotes matrix trace.

Using (7) and (9), the problem of minimizing the MSE of the signal waveform estimation at the destination node can be written as

$$\min_{\{\mathbf{F}_l\},\{\mathbf{B}_i\}} \quad \mathrm{tr}\Big(\left(\mathbf{I}_{N_0} + \mathbf{A}_L^H\mathbf{C}_L^{-1}\mathbf{A}_L\right)^{-1}\Big) \quad (10)$$

$$\text{s.t.} \quad \mathrm{tr}\left(\mathbf{F}_l\mathbf{D}_{l-1}\mathbf{F}_l^H\right) \le p_l, \qquad l = 2, \cdots, L \quad (11)$$

$$\mathrm{tr}(\mathbf{B}_i\mathbf{B}_i^H) \le q_i, \qquad i = 1, \cdots, N_u \quad (12)$$

where (11) and (12) are the transmission power constraint at each relay node and each user, respectively, p_l and q_i are the corresponding power budget, $\{\mathbf{F}_l\} \triangleq [\mathbf{F}_2, \cdots, \mathbf{F}_L]$, and $\{\mathbf{B}_i\} \triangleq [\mathbf{B}_1, \cdots, \mathbf{B}_{N_u}]$. The problem (10)-(12) is nonconvex with matrix variables, and a globally optimal solution is very difficult to obtain with a reasonable computational complexity (non-exhaustive searching). In [6], an iterative procedure was developed to obtain (at least) a locally optimal solution of the problem (10)-(12), where in each iteration, the relay amplifying matrices are optimized with fixed source precoding matrices, and then the source precoding matrices are updated with the given relay amplifying matrices. However, the computational complexity and the signalling overhead of the iterative algorithm is quite high for practical relay systems. In the following, we propose simplified algorithms to solve an approximation of the problem (10)-(12). The proposed algorithms have much smaller computational complexity and signalling overhead than the iterative algorithm in [6] as analyzed and shown later.

A. Optimal Structure of Relay Amplifying Matrices

By introducing $N_{l-1} \times N_0$ matrices $\mathbf{T}_l$, $l = 2, \cdots, L$, the following theorem establishes the structure of the optimal relay amplifying matrices, and demonstrates that the MSE matrix at the destination node can be decomposed into the sum of the MSE matrices at all relay nodes.

THEOREM 1: The optimal relay amplifying matrices have the following structure

$$\mathbf{F}_l = \mathbf{T}_l\mathbf{A}_{l-1}^H\mathbf{D}_{l-1}^{-1}, \qquad l = 2, \cdots, L. \quad (13)$$

Using (13), the MSE matrix at the destination node can be equivalently decomposed to

$$\mathbf{E}_L = \left(\mathbf{I}_{N_0} + \mathbf{F}_1^H\mathbf{H}_1^H\mathbf{H}_1\mathbf{F}_1\right)^{-1} + \sum_{l=2}^{L}\left(\mathbf{R}_l^{-1} + \mathbf{T}_l^H\mathbf{H}_l^H\mathbf{H}_l\mathbf{T}_l\right)^{-1} \quad (14)$$

where

$$\mathbf{R}_l \triangleq \mathbf{A}_{l-1}^H\mathbf{D}_{l-1}^{-1}\mathbf{A}_{l-1}, \qquad l = 2, \cdots, L. \quad (15)$$

PROOF: See Appendix A. □

Interestingly, it can be seen from (13) that the optimal relay amplifying matrices $\mathbf{F}_l$, $l = 2, \cdots, L$, can be decomposed into $\mathbf{F}_l = \mathbf{T}_l\mathbf{W}_l^H$, where $\mathbf{W}_l = (\mathbf{A}_{l-1}\mathbf{A}_{l-1}^H + \mathbf{C}_{l-1})^{-1}\mathbf{A}_{l-1}$, $l = 2, \cdots, L$, is the weight matrix of the linear MMSE filter for the received signal vector at the $(l-1)$-th relay node given by $\mathbf{y}_{l-1} = \mathbf{A}_{l-1}\mathbf{s} + \bar{\mathbf{v}}_{l-1}$, and the linear filter $\mathbf{T}_l$ will be designed later. The term $\left(\mathbf{R}_l^{-1} + \mathbf{T}_l^H\mathbf{H}_l^H\mathbf{H}_l\mathbf{T}_l\right)^{-1}$, $l = 2, \cdots, L$, in (14) is the increment of the MSE matrix introduced by the $(l-1)$-th relay node as detailed in Appendix A. It is worth noting that $\mathbf{R}_l$ is in fact the covariance matrix of $\mathbf{W}_l^H\mathbf{y}_{l-1}$ as $\mathbf{R}_l = \mathbf{W}_l^H\mathrm{E}[\mathbf{y}_{l-1}\mathbf{y}_{l-1}^H]\mathbf{W}_l$. It can be seen from (14) that the effect of noise in the first hop is reflected by $\mathbf{I}_{N_0}$ in the first term. As SNR increases, $\mathbf{F}_1^H\mathbf{H}_1^H\mathbf{H}_1\mathbf{F}_1$ increases to infinity, and $\left(\mathbf{I}_{N_0} + \mathbf{F}_1^H\mathbf{H}_1^H\mathbf{H}_1\mathbf{F}_1\right)^{-1}$ approaches $\mathbf{0}_{N_0\times N_0}$. Similarly, the effect of noise in the lth hop, $l = 2, \cdots, L$, is reflected by $\mathbf{R}_l$. Interestingly, as SNR increases, $\mathbf{R}_l$ approaches $\mathbf{I}_{N_0}$, $\mathbf{T}_l^H\mathbf{H}_l^H\mathbf{H}_l\mathbf{T}_l$ increases to infinity, and consequently, $\left(\mathbf{R}_l^{-1} + \mathbf{T}_l^H\mathbf{H}_l^H\mathbf{H}_l\mathbf{T}_l\right)^{-1}$ approaches $\mathbf{0}_{N_0\times N_0}$.

By exploiting (13), the transmission power consumed by each relay node can be written as

$$\begin{aligned}\mathrm{tr}\left(\mathbf{F}_l\mathbf{D}_{l-1}\mathbf{F}_l^H\right) &= \mathrm{tr}\left(\mathbf{T}_l\mathbf{A}_{l-1}^H\mathbf{D}_{l-1}^{-1}\mathbf{D}_{l-1}\mathbf{D}_{l-1}^{-1}\mathbf{A}_{l-1}\mathbf{T}_l^H\right) \\ &= \mathrm{tr}\left(\mathbf{T}_l\mathbf{R}_l\mathbf{T}_l^H\right), \qquad l = 2, \cdots, L. \end{aligned} \quad (16)$$

From (2) we have

$$\begin{aligned}&\mathrm{tr}\Big(\left(\mathbf{I}_{N_0} + \mathbf{F}_1^H\mathbf{H}_1^H\mathbf{H}_1\mathbf{F}_1\right)^{-1}\Big) \\ &\quad = \mathrm{tr}\left(\left(\mathbf{I}_{N_1} + \sum_{i=1}^{N_u}\mathbf{G}_i\mathbf{Q}_i\mathbf{G}_i^H\right)^{-1}\right) + N_0 - N_1 \end{aligned} \quad (17)$$

where $\mathbf{Q}_i \triangleq \mathrm{E}[\mathbf{u}_i\mathbf{u}_i^H] = \mathbf{B}_i\mathbf{B}_i^H$, $i = 1, \cdots, N_u$, is the covariance matrix of the signal transmitted by the ith user. Now by using (14)-(17), the problem (10)-(12) can be equivalently rewritten as

$$\min_{\{\mathbf{Q}_i\},\{\mathbf{T}_l\}} \quad \mathrm{tr}\Bigg(\bigg(\mathbf{I}_{N_1}+\sum_{i=1}^{N_u}\mathbf{G}_i\mathbf{Q}_i\mathbf{G}_i^H\bigg)^{-1} + \sum_{l=2}^{L}\big(\mathbf{R}_l^{-1}+\mathbf{T}_l^H\mathbf{H}_l^H\mathbf{H}_l\mathbf{T}_l\big)^{-1}\Bigg) \tag{18}$$

$$\text{s.t.} \quad \mathrm{tr}\big(\mathbf{T}_l\mathbf{R}_l\mathbf{T}_l^H\big)\le p_l, \qquad l=2,\cdots,L \tag{19}$$

$$\mathrm{tr}(\mathbf{Q}_i)\le q_i, \quad \mathbf{Q}_i\succeq 0, \qquad i=1,\cdots,N_u \tag{20}$$

where $\{\mathbf{T}_l\}\triangleq[\mathbf{T}_2,\cdots,\mathbf{T}_L]$, $\{\mathbf{Q}_i\}\triangleq[\mathbf{Q}_1,\cdots,\mathbf{Q}_{N_u}]$, and $\succeq$ stands for the matrix positive semi-definiteness.

B. Proposed Algorithm 1

Using the matrix inversion lemma $(\mathbf{A}+\mathbf{BCD})^{-1}=\mathbf{A}^{-1}-\mathbf{A}^{-1}\mathbf{B}(\mathbf{DA}^{-1}\mathbf{B}+\mathbf{C}^{-1})^{-1}\mathbf{DA}^{-1}$, it can be seen from (15) that

$$\begin{aligned}\mathbf{R}_l &= \mathbf{A}_{l-1}^H\Big(\mathbf{C}_{l-1}^{-1}-\mathbf{C}_{l-1}^{-1}\mathbf{A}_{l-1}\big(\mathbf{A}_{l-1}^H\mathbf{C}_{l-1}^{-1}\mathbf{A}_{l-1}+\mathbf{I}_{N_0}\big)^{-1} \\ &\quad\times\mathbf{A}_{l-1}^H\mathbf{C}_{l-1}^{-1}\Big)\mathbf{A}_{l-1} \\ &= \mathbf{A}_{l-1}^H\mathbf{C}_{l-1}^{-1}\mathbf{A}_{l-1}\big(\mathbf{A}_{l-1}^H\mathbf{C}_{l-1}^{-1}\mathbf{A}_{l-1}+\mathbf{I}_{N_0}\big)^{-1}, \ l=2,\cdots,L.\end{aligned}$$

In the case of (moderately) high SNR where $\mathbf{A}_{l-1}^H\mathbf{C}_{l-1}^{-1}\mathbf{A}_{l-1}\gg\mathbf{I}_{N_0}$, we can approximate $\mathbf{R}_l$ as $\mathbf{I}_{N_0}$, $l=2,\cdots,L$. In other words, in such case, the value of $\mathbf{Q}_i$, $i=1,\cdots,N_u$, does not affect $\mathbf{R}_l$, $l=2,\cdots,L$, and $\mathbf{T}_l$ does not affect $\mathbf{R}_j$, $j=l+1,\cdots,L$. This fact implies that the objective function (18) and the constraints in (19) are decoupled with respect to the variables $\{\mathbf{Q}_i\}$ and $\{\mathbf{T}_l\}$. Therefore, the problem (18)-(20) can be approximated and decomposed into the following relay amplifying matrix optimization problem for each $l=2,\cdots,L$

$$\min_{\mathbf{T}_l} \quad \mathrm{tr}\Big(\big(\mathbf{R}_l^{-1}+\mathbf{T}_l^H\mathbf{H}_l^H\mathbf{H}_l\mathbf{T}_l\big)^{-1}\Big) \tag{21}$$

$$\text{s.t.} \quad \mathrm{tr}\big(\mathbf{T}_l\mathbf{R}_l\mathbf{T}_l^H\big)\le p_l \tag{22}$$

and the source covariance matrices optimization problem

$$\min_{\{\mathbf{Q}_i\}} \quad \mathrm{tr}\Bigg(\bigg(\mathbf{I}_{N_1}+\sum_{i=1}^{N_u}\mathbf{G}_i\mathbf{Q}_i\mathbf{G}_i^H\bigg)^{-1}\Bigg) \tag{23}$$

$$\text{s.t.} \quad \mathrm{tr}(\mathbf{Q}_i)\le q_i, \qquad \mathbf{Q}_i\succeq 0, \qquad i=1,\cdots,N_u. \tag{24}$$

In the following, we show that the problem (21)-(22) has a water-filling solution. Let us introduce the eigenvalue decomposition (EVD) of $\mathbf{H}_l^H\mathbf{H}_l=\mathbf{V}_l\mathbf{\Lambda}_l\mathbf{V}_l^H$ and $\mathbf{R}_l=\mathbf{U}_l\mathbf{\Sigma}_l\mathbf{U}_l^H$, $l=2,\cdots,L$, where the dimensions of $\mathbf{V}_l$ and $\mathbf{\Lambda}_l$ are $N_{l-1}\times N_{l-1}$, the dimensions of $\mathbf{U}_l$ and $\mathbf{\Sigma}_l$ are $N_0\times N_0$, and the diagonal elements in $\mathbf{\Lambda}_l$ and $\mathbf{\Sigma}_l$ are sorted in increasing orders. By introducing $\tilde{\mathbf{T}}_l\triangleq\mathbf{T}_l\mathbf{R}^{\frac{1}{2}}$, the problem (21)-(22) can be rewritten as

$$\min_{\tilde{\mathbf{T}}_l} \quad \mathrm{tr}\Big(\mathbf{R}_l^{\frac{1}{2}}\big(\mathbf{I}_{N_0}+\tilde{\mathbf{T}}_l^H\mathbf{H}_l^H\mathbf{H}_l\tilde{\mathbf{T}}_l\big)^{-1}\mathbf{R}_l^{\frac{1}{2}}\Big) \tag{25}$$

$$\text{s.t.} \quad \mathrm{tr}\big(\tilde{\mathbf{T}}_l\tilde{\mathbf{T}}_l^H\big)\le p_l. \tag{26}$$

Using Lemma 2 in Appendix A, the solution to the problem (25)-(26) in terms of the singular value decomposition (SVD) of $\tilde{\mathbf{T}}_l$ is given by $\tilde{\mathbf{T}}_l=\mathbf{V}_{l,1}\mathbf{\Omega}_l\mathbf{U}_l^H$, where $\mathbf{V}_{l,1}$ contains the rightmost N_0 columns of $\mathbf{V}_l$. Thus the structure of the optimal linear filter $\mathbf{T}_l$ is given by

$$\mathbf{T}_l=\mathbf{V}_{l,1}\mathbf{\Delta}_l\mathbf{U}_l^H, \quad \mathbf{\Delta}_l=\mathbf{\Omega}_l\mathbf{\Sigma}_l^{-\frac{1}{2}}, \quad l=2,\cdots,L \tag{27}$$

where $\mathbf{\Delta}_l$ is an $N_0\times N_0$ diagonal matrix that remains to be optimized.

Interestingly, it can be seen from (27) that at the $(l-1)$-th relay node, the linear filter $\mathbf{T}_l$ first performs beamforming to the direction of the eigenvectors of $\mathbf{R}_l$, then it allocates power to N_0 streams through $\mathbf{\Delta}_l$, and finally beamforms to the direction of the eigenvectors of $\mathbf{H}_l^H\mathbf{H}_l$. Substituting (27) back into (21)-(22), we find that the matrix-variable optimization problem (21)-(22) is converted to the following optimal power loading problem with scalar variables

$$\min_{\delta_{l,1},\cdots,\delta_{l,N_0}} \quad \sum_{i=1}^{N_0}\frac{1}{\sigma_{l,i}^{-1}+\delta_{l,i}^2\lambda_{l,i}} \tag{28}$$

$$\text{s.t.} \quad \sum_{i=1}^{N_0}\delta_{l,i}^2\sigma_{l,i}\le p_l \tag{29}$$

where $\delta_{l,i},\sigma_{l,i},\lambda_{l,i}$, $i=1,\cdots,N_0$, denote the ith diagonal element of $\mathbf{\Delta}_l$, $\mathbf{\Sigma}_l$, $\mathbf{\Lambda}_l$, respectively. Using the Lagrange multiplier method [14], it can be shown that the problem (28)-(29) has a water-filling solution given by

$$\delta_{l,i}^2=\frac{1}{\lambda_{l,i}}\left(\sqrt{\frac{\lambda_{l,i}}{\mu_l\,\sigma_{l,i}}}-\frac{1}{\sigma_{l,i}}\right)^+, \qquad i=1,\cdots,N_0$$

where $(x)^+\triangleq\max(x,0)$, and $\mu_l>0$ is the Lagrangian multiplier and the solution to the nonlinear equation of $\sum_{i=1}^{N_0}\frac{\sigma_{l,i}}{\lambda_{l,i}}\left(\sqrt{\frac{\lambda_{l,i}}{\mu_l\,\sigma_{l,i}}}-\frac{1}{\sigma_{l,i}}\right)^+=p_l$.

Substituting (27) back into (13), the structure of the optimal relay amplifying matrices is given by

$$\mathbf{F}_l=\mathbf{V}_{l,1}\mathbf{\Delta}_l\mathbf{U}_l^H\mathbf{A}_{l-1}^H\mathbf{D}_{l-1}^{-1}, \qquad l=2,\cdots,L. \tag{30}$$

Interestingly, although (30) is derived under a high SNR assumption, this structure is in fact optimal for the whole SNR region as stated by the following theorem.

THEOREM 2: The structure of the relay amplifying matrix given in (30) can be equivalently written as $\mathbf{F}_l=\mathbf{V}_{l,1}\mathbf{\Upsilon}_l\mathbf{U}_{H_{l-1},1}^H$, $l=2,\cdots,L$, which is optimal for multi-hop MIMO relay systems as proved in [3]. Here $\mathbf{\Upsilon}_l$, $l=2,\cdots,L$, are $N_0\times N_0$ diagonal matrices, $\mathbf{H}_1\mathbf{F}_1=\mathbf{U}_{H_1}\mathbf{\Gamma}_1\mathbf{V}_1^H$, $\mathbf{H}_l=\mathbf{U}_{H_l}\mathbf{\Gamma}_l\mathbf{V}_l^H$, $l=2,\cdots,L$, are SVDs of $\mathbf{H}_1\mathbf{F}_1$ and $\mathbf{H}_l$ with the diagonal elements of $\mathbf{\Gamma}_l$ sorted in increasing orders, and $\mathbf{U}_{H_l,1}$ contains the rightmost N_0 columns of $\mathbf{U}_{H_l}$.

PROOF: See Appendix B. □

Finally, the source covariance matrices optimization problem (23)-(24) can be solved as follows. By introducing a positive semi-definite (PSD) matrix $\mathbf{X}$ with $\mathbf{X}\succeq\big(\mathbf{I}_{N_1}+\sum_{i=1}^{N_u}\mathbf{G}_i\mathbf{Q}_i\mathbf{G}_i^H\big)^{-1}$ and using the Schur complement [14], the problem (23)-(24) can be converted to the problem of

$$\min_{\{\mathbf{Q}_i\},\mathbf{X}} \quad \mathrm{tr}(\mathbf{X}) \tag{31}$$

$$\text{s.t.} \quad \begin{pmatrix}\mathbf{X} & \mathbf{I}_{N_1}\\ \mathbf{I}_{N_1} & \mathbf{I}_{N_1}+\sum_{i=1}^{N_u}\mathbf{G}_i\mathbf{Q}_i\mathbf{G}_i^H\end{pmatrix}\succeq 0 \tag{32}$$

$$\mathrm{tr}(\mathbf{Q}_i)\le q_i, \qquad \mathbf{Q}_i\succeq 0, \quad i=1,\cdots,N_u. \tag{33}$$

TABLE I
PROCEDURE OF OPTIMIZING THE SOURCE AND RELAY MATRICES.

1) Solve the SDP problem (31)-(33) to obtain $\{\mathbf{Q}_i\}$.
2) For $l = 2 : L$
   Compute $\mathbf{R}_l$; Solve the problem (21)-(22) to have $\mathbf{T}_l$; Obtain $\mathbf{F}_l$ as in (13).
   End.

The problem (31)-(33) is a convex semi-definite programming (SDP) problem which can be efficiently solved by the interior-point method [14]. Then the optimal $\mathbf{B}_i$ can be obtained as $\mathbf{B}_i = \mathbf{U}_{Q,i}\mathbf{\Lambda}_{Q,i}^{\frac{1}{2}}\mathbf{U}_b$, where $\mathbf{U}_{Q,i}\mathbf{\Lambda}_{Q,i}\mathbf{U}_{Q,i}^H$ is the EVD of $\mathbf{Q}_i$, and $\mathbf{U}_b$ is an arbitrary $M_i \times M_i$ unitary matrix. We would like to mention that when the number of independent data streams transmitted by the ith user (defined as $N_{b,i}$) is smaller than the number of transmit antennas M_i, solving the problem (31)-(33) might yield $\mathbf{Q}_i$ whose rank is larger than $N_{b,i}$. In such case, the randomization technique [15] can be applied to obtain a possibly suboptimal solution of $\mathbf{B}_i$ with rank $N_{b,i}$. Nevertheless, in this paper, we focus on $N_{b,i} = M_i$, $i = 1, \cdots, N_u$. Thus, $\mathbf{B}_i$ obtained through solving the problem (31)-(33) is optimal.

The procedure of optimizing all source precoding matrices and relay amplifying matrices is described in Table I. We would like to mention that in each iteration of the algorithm in [6], the complexity of updating the source precoding matrices is similar to that of solving the problem (31)-(33). While at each iteration of [6], an alternating power loading algorithm is applied to update the relay amplifying matrices, which has a higher computational complexity than that of solving the problem (21)-(22). Therefore, the computational complexity of carrying out the procedure in Table I is less than that of each iteration in [6]. Interestingly, the procedure in Table I can be carried out in a distributed manner where each relay node performs the necessary optimization procedure locally. In particular, the first relay node optimizes all source precoding matrices and sends back $\mathbf{B}_i$ to user i. The first relay node also computes the optimal $\mathbf{F}_2$. Then at the lth relay node, $l = 2, \cdots, L-1$, the optimal $\mathbf{F}_{l+1}$ is computed based on $\mathbf{H}_{l+1}$, $\mathbf{C}_l$, and $\mathbf{A}_l$. The CSI of $\mathbf{H}_{l+1}$ can be first estimated at the $(l+1)$-th relay node through channel training [16], and then fed back to the lth relay node. The knowledge of $\mathbf{C}_l$ and $\mathbf{A}_l$ is forwarded from the $(l-1)$-th relay node. Note that due to its iterative nature, the algorithm in [6] requires centralized processing. Obviously, compared with the centralized method, the distributed approach requires much less information exchange and signalling overhead among different nodes, and thus, is preferred in practical relay systems.

C. Proposed Algorithm 2

In this algorithm, the source precoding matrices are optimized by solving the problem (31)-(33). However, since $\mathbf{R}_l$ approaches $\mathbf{I}_{N_0}$ as SNR increases, at a high SNR environment, the relay amplifying matrices optimization can be further simplified by substituting $\mathbf{R}_l$ in (21) and (22) with $\mathbf{I}_{N_0}$. Then we have the following optimization problem for each $\mathbf{T}_l$, $l = 2, \cdots, L$

$$\min_{\mathbf{T}_l} \quad \operatorname{tr}\Big(\big(\mathbf{I}_{N_0} + \mathbf{T}_l^H\mathbf{H}_l^H\mathbf{H}_l\mathbf{T}_l\big)^{-1}\Big) \tag{34}$$

$$\text{s.t.} \quad \operatorname{tr}\big(\mathbf{T}_l\mathbf{T}_l^H\big) \leq p_l. \tag{35}$$

Interestingly, (34) is in fact an upper-bound of (21). Applying Lemma 2 in Appendix A, the solution to the problem (34)-(35) is given by

$$\mathbf{T}_l = \mathbf{V}_{l,1}\mathbf{\Theta}_l\mathbf{\Pi}, \qquad l = 2, \cdots, L \tag{36}$$

where $\mathbf{\Pi}$ can be any $N_0 \times N_0$ unitary matrix, and $\mathbf{\Theta}_l$ is an $N_0 \times N_0$ diagonal matrix. Substituting (36) back into (34)-(35), we find that the ith diagonal element of $\mathbf{\Theta}_l$ is given by $\theta_{l,i} = \left[\frac{1}{\lambda_{l,i}}\left(\sqrt{\frac{\lambda_{l,i}}{\nu_l}}-1\right)^+\right]^{\frac{1}{2}}$, $i = 1, \cdots, N_0$. Here $\nu_l > 0$ is the solution to the nonlinear equation of $\sum_{i=1}^{N_0} \frac{1}{\lambda_{l,i}}\left(\sqrt{\frac{\lambda_{l,i}}{\nu_l}}-1\right)^+ = p_l$.

Compared with the problem (21)-(22), the relay amplifying matrices designed by the problem (34)-(35) has a smaller computational complexity, since the latter algorithm does not need to compute $\mathbf{R}_l$ and its SVD. Comparing (36) with (27), we can choose $\mathbf{\Pi} = \mathbf{U}_l^H = \mathbf{V}_1^H$ which provides an optimal structure of $\mathbf{F}_l$. $l = 2, \cdots, L$, as proved in Appendix B. However, in this case, all relay nodes need to know $\mathbf{V}_1$, which increases the signalling overhead. For the reason of simplicity, we choose $\mathbf{\Pi} = \mathbf{I}_{N_0}$. Through numerical simulations in Section IV, we will see that there is only a negligible increase in MSE and BER by using $\mathbf{\Pi} = \mathbf{I}_{N_0}$ instead of $\mathbf{\Pi} = \mathbf{V}_1^H$.

IV. NUMERICAL EXAMPLES

In this section, we study the performance of the proposed multiuser multi-hop MIMO relay design algorithms through numerical simulations. We simulate a flat Rayleigh fading environment where all channel matrices have entries with zero mean. In particular, the variance of entries in $\mathbf{G}_i$ is $1/M_i$, $i = 1, \cdots, N_u$, and the variance of entries in $\mathbf{H}_l$ is $1/N_{l-1}$, $l = 2, \cdots, L$. All noises are complex circularly symmetric with zero mean and unit variance. We also assume that $p_l = P$, $l = 2, \cdots, L$, $q_i = Q$, $i = 1, \cdots, N_u$. We would like to mention that in order to minimize (10), each user should use its maximal power Q. In fact, (10) can be equivalently rewritten as

$$\text{MSE} = \operatorname{tr}\left(\left(\mathbf{I}_{N_1} + \mathbf{\Psi}^H \sum_{i=1}^{N_u} \mathbf{G}_i\mathbf{B}_i\mathbf{B}_i^H\mathbf{G}_i^H\mathbf{\Psi}\right)^{-1}\right) - N_1 + N_0 \tag{37}$$

where

$$\mathbf{\Psi}\mathbf{\Psi}^H = \bigotimes_{i=2}^{L}(\mathbf{F}_i^H\mathbf{H}_i^H)\left(\sum_{l=2}^{L}\Big(\bigotimes_{i=L}^{l}(\mathbf{H}_i\mathbf{F}_i) \bigotimes_{i=l}^{L}(\mathbf{F}_i^H\mathbf{H}_i^H)\Big) + \mathbf{I}_{N_L}\right)^{-1} \bigotimes_{i=L}^{2}(\mathbf{H}_i\mathbf{F}_i).$$

Assuming that the ith user applies $\tilde{\mathbf{B}}_i$ with $\operatorname{tr}(\tilde{\mathbf{B}}_i\tilde{\mathbf{B}}_i^H) = \tilde{q}_i < Q$, obviously, the value of (37) can be reduced by using $\mathbf{B}_i = \alpha_i\tilde{\mathbf{B}}_i$, where $\alpha_i = \sqrt{Q/\tilde{q}_i} > 1$.

All simulation results are averaged over 5000 independent channel realizations. The CVX convex optimization software

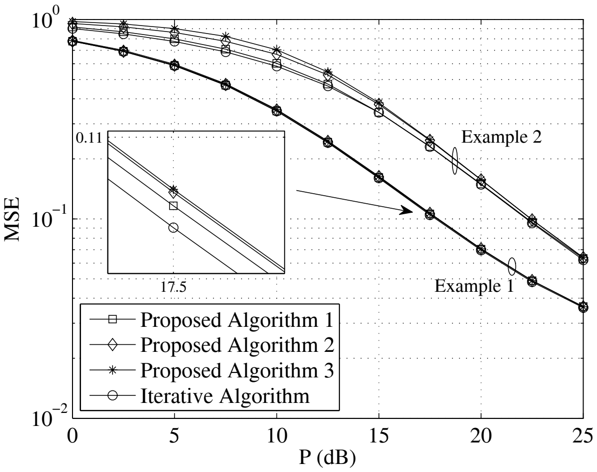


Fig. 2. MSE versus P. Example 1: $L = 2$, $N_u = 2$, $M_1 = 3$, $M_2 = 2$, $N_1 = 8$, and $N_2 = 7$; Example 2: $L = 3$, $N_u = 3$, $M = 2$, and $N = 8$.

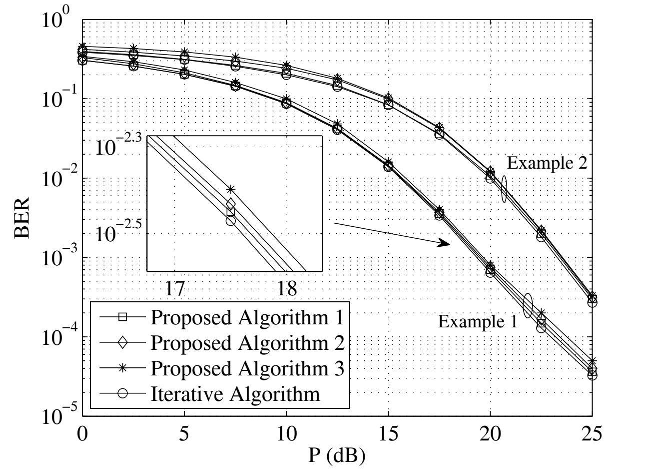


Fig. 3. BER versus P. Example 1: $L = 2$, $N_u = 2$, $M_1 = 3$, $M_2 = 2$, $N_1 = 8$, and $N_2 = 7$; Example 2: $L = 3$, $N_u = 3$, $M = 2$, and $N = 8$.

package [17] is applied to solve the SDP problem (31)-(33). For all examples, we set $Q = 20$dB and compare the performance of the algorithm described in Table I (denoted as Proposed Algorithm 1), the algorithm where the relay amplifying matrices are designed by solving the problem (34)-(35) using $\mathbf{\Pi} = \mathbf{V}_1^H$ (denoted as Proposed Algorithm 2), the algorithm of solving the problem (34)-(35) with $\mathbf{\Pi} = \mathbf{I}_{N_0}$ (denoted as Proposed Algorithm 3), and the optimal iterative algorithm developed in [6] (denoted as Iterative Algorithm).

In our first example, we simulate a two-hop relay system with $N_u = 2$, $M_1 = 3$, $M_2 = 2$, $N_1 = 8$, and $N_2 = 7$. Fig. 2 shows the MSE performance of all algorithms versus P, and the system BER yielded by all algorithms with QPSK constellations are illustrated in Fig. 3 versus P. Our results clearly demonstrate that the Proposed Algorithms 1-3 only have slightly higher MSE and BER than the Iterative Algorithm. Note that three proposed algorithms have a much smaller computational complexity and signalling overhead than the Iterative Algorithm as analyzed in Subsection III-B. We also observe from Figs. 2 and 3 that there is a significant performance improvement for all algorithms when P increases from 20dB to 25dB. The reason is that the MIMO relay system simulated is under-loaded in terms of data streams since $\sum_{i=1}^{N_u} M_i = N_0 < N_l$, $l = 1, \cdots, L$. In such scenario, the system has a high spatial diversity order, which overcomes the saturation effect caused by the fixed Q. In MIMO relay systems with a low spatial diversity order, one can observe the saturation effect as P increases when Q is fixed.

We simulate multi-hop ($L \geq 3$) multiuser relay systems in the following two examples. Since there are many parameters on the system setup for multi-hop relays, for simplicity, we consider relay systems where all users have the same number of antennas (i.e., $M_i = M$, $i = 1, \cdots, N_u$) and all relay nodes and the destination node have the same number of antennas (i.e., $N_l = N$, $l = 1, \cdots, N_L$). The extension to systems where different nodes have different number of antennas is straight-forward. A three-hop ($L = 3$) MIMO relay system is simulated in the second example with $N_u = 3$, $M = 2$, and $N = 8$. Fig. 2 and Fig. 3 show the MSE and BER comparisons among four algorithms, respectively. It can be seen that due to the approximation from the problem (21)-(22) to the problem (34)-(35), the MSE and BER gaps between the Proposed Algorithm 2, the Proposed Algorithm 3, and the other two algorithm increase at low to medium P. But the performance of the Proposed Algorithm 1 is very close to that of the Iterative Algorithm. It can also be observed from Figs. 2 and 3 that both the MSE and BER values of Example 1 are lower than those of Example 2, indicating that the algorithm yielding a lower MSE indeed guarantees a better BER performance.

In the third example, a five-hop ($L = 5$) relay system is simulated with $N_u = 3$, $M = 2$, and $N = 9$. The MSE and BER comparisons of four algorithms are shown in Fig. 4 and Fig. 5, respectively. It can be clearly seen that the Proposed Algorithm 1 yields almost the same BER as the Iterative Algorithm. It can also be observed from Figs. 2-5 that there is only a small gap in both the MSE and BER performance between the Proposed Algorithm 2 and the Proposed Algorithm 3. The reason is that the performance of both algorithms depend not only on $\mathbf{\Pi}$, but also on the power loading matrix $\mathbf{\Theta}_l$ in (36). From the fact that the gap between the Proposed Algorithms 2 and 3 is smaller than that between the Proposed Algorithms 1 and 2, it shows that the choice of the power loading matrix plays a more important role than that of $\mathbf{\Pi}$. Since both the Proposed Algorithms 2 and 3 use the same $\mathbf{\Theta}_l$, the performance gap caused by using different $\mathbf{\Pi}$ is small. Based on the simulation results and taking into account the complexity-performance tradeoff, the Proposed Algorithm 1 is most suitable for practical multiuser multi-hop MIMO relay systems.

V. Conclusions

We addressed the issue of multiaccess communication through multi-hop non-regenerative MIMO relays. It has been shown that the MSE matrix of the signal waveform estimation at the destination node can be decomposed into the sum of the MSE matrices at all relay nodes. Simplified source and relay optimization algorithms have been proposed which greatly

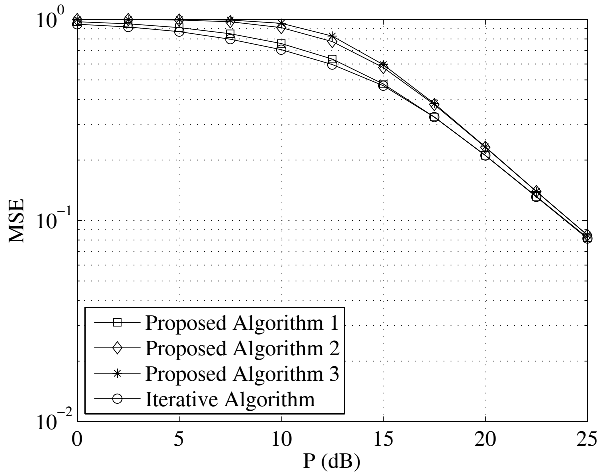


Fig. 4. MSE versus P. Example 3: $L = 5$, $N_u = 3$, $M = 2$, $N = 9$.

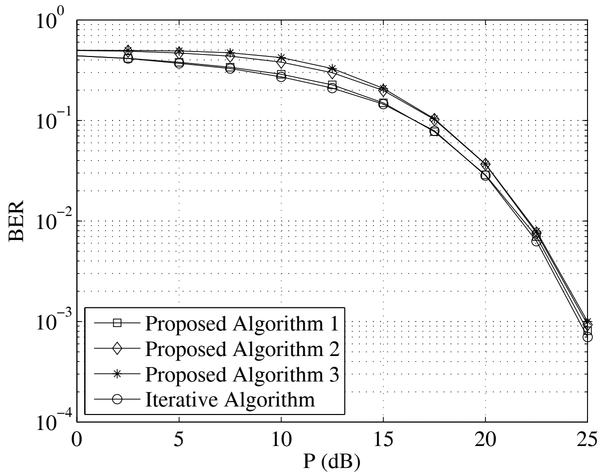


Fig. 5. BER versus P. Example 3: $L = 5$, $N_u = 3$, $M = 2$, $N = 9$.

reduce the computational complexity and signalling overhead with only a negligible MSE and BER degradation.

APPENDIX A
PROOF OF THEOREM 1

To prove Theorem 1, we need the following two lemmas.

LEMMA 1 [18, 9.H.1.h]: For two $N \times N$ positive semidefinite matrices $\mathbf{A}$ and $\mathbf{B}$ with eigenvalues $\lambda_{a,i}$ and $\lambda_{b,i}$, $i = 1, \cdots, N$, arranged in the same order, respectively, it follows that $\mathrm{tr}(\mathbf{AB}) \geq \sum_{i=1}^{N} \lambda_{a,i}\lambda_{b,N+1-i}$.

LEMMA 2: For matrices $\mathbf{A}$, $\mathbf{T}$, $\mathbf{H}$ with dimensions of $m \times n$, $l \times m$, and $k \times l$, respectively, where $k, l, m \geq n$, $r \triangleq \mathrm{rank}(\mathbf{H}) \geq n$, and $\mathrm{rank}(\mathbf{T}) = n$, the solution to the problem of

$$\min_{\mathbf{T}} \ \mathrm{tr}\big(\mathbf{A}^H\big(\mathbf{I}_m + \mathbf{T}^H\mathbf{H}^H\mathbf{HT}\big)^{-1}\mathbf{A}\big) \quad \text{s.t. } \mathrm{tr}\big(\mathbf{TT}^H\big) \leq p \tag{38}$$

is given by $\mathbf{T} = \mathbf{V}_{h,1}\mathbf{\Lambda}\mathbf{U}_a^H$ in terms of the SVD of $\mathbf{T}$. Here

$$\mathbf{H} = \mathbf{U}_h\mathbf{\Sigma}_h\mathbf{V}_h^H, \qquad \mathbf{A} = \mathbf{U}_a\mathbf{\Sigma}_a\mathbf{V}_a^H \tag{39}$$

are the SVDs of $\mathbf{H}$ and $\mathbf{A}$, with dimensions of $\mathbf{U}_h$, $\mathbf{\Sigma}_h$, $\mathbf{V}_h$, $\mathbf{U}_a$, $\mathbf{\Sigma}_a$, and $\mathbf{V}_a$ being $k \times k$, $k \times l$, $l \times l$, $m \times n$, $n \times n$, and $n \times n$, respectively. The diagonal elements of $\mathbf{\Sigma}_h$ and $\mathbf{\Sigma}_a$ are sorted in increasing orders, and $\mathbf{V}_{h,1}$ contains the rightmost n columns of $\mathbf{V}_h$.

PROOF: Let us introduce the EVD of

$$\mathbf{T}^H\mathbf{H}^H\mathbf{HT} = [\mathbf{U}_{\bar{x}} \ \ \mathbf{U}_x]\,\mathrm{bd}\big(\mathbf{0}_{(m-n)\times(m-n)}, \mathbf{\Lambda}_x\big)[\mathbf{U}_{\bar{x}} \ \ \mathbf{U}_x]^H \tag{40}$$

where the dimensions of $\mathbf{U}_{\bar{x}}$, $\mathbf{U}_x$, and $\mathbf{\Lambda}_x$ are $m \times (m-n)$, $m \times n$, and $n \times n$, respectively, $\mathbf{0}_{p\times m}$ denotes a $p \times m$ matrix with all zero entries. Substituting (39) and (40) into (38), the objective function in (38) can be written as

$$\mathrm{tr}\big(\mathbf{\Sigma}_a\mathbf{U}_a^H\big(\mathbf{U}_x(\mathbf{I}_n + \mathbf{\Lambda}_x)^{-1}\mathbf{U}_x^H + \mathbf{U}_{\bar{x}}\mathbf{U}_{\bar{x}}^H\big)\mathbf{U}_a\mathbf{\Sigma}_a\big). \tag{41}$$

From Lemma 1 we know that (41) is minimized if the diagonal elements of $\mathbf{\Lambda}_x$ are sorted in increasing order, and $\mathbf{U}_x^H\mathbf{U}_a = \mathbf{\Phi}_n$, where $\mathbf{\Phi}_n$ stands for an arbitrary $n \times n$ diagonal matrix with unit-norm main diagonal elements, i.e., $|[\mathbf{\Phi}_n]_{i,i}| = 1, [\mathbf{\Phi}_n]_{i,j} = 0, i, j = 1, \cdots, n, i \neq j$. Without affecting the objective function in (38), we choose $\mathbf{U}_x = \mathbf{U}_a$. It will be seen later on that the power constraint in (38) is invariant to $\mathbf{U}_x$.

From (40) we have

$$\mathbf{HT} = \mathbf{V}_x\mathbf{\Lambda}_x^{\frac{1}{2}}\mathbf{U}_a^H \tag{42}$$

where $\mathbf{V}_x$ is a $k \times n$ (semi)-unitary matrix with $\mathbf{V}_x^H\mathbf{V}_x = \mathbf{I}_n$. Substituting the SVD of $\mathbf{H}$ in (39) into (42) and left multiplying by $\mathbf{U}_h^H$ on both sides, we have

$$\begin{bmatrix} \mathbf{0}_{(k-r)\times(l-r)} & \mathbf{0}_{(k-r)\times r} \\ \mathbf{0}_{r\times(l-r)} & \mathbf{\Sigma}_{h,r} \end{bmatrix} \hat{\mathbf{T}} = \begin{bmatrix} \mathbf{U}_{h,\bar{r}}^H \\ \mathbf{U}_{h,r}^H \end{bmatrix} \mathbf{V}_x\mathbf{\Lambda}_x^{\frac{1}{2}}\mathbf{U}_a^H \tag{43}$$

where $\hat{\mathbf{T}} = \mathbf{V}_h^H\mathbf{T}$, $\mathbf{\Sigma}_{h,r}$ is a $r \times r$ diagonal matrix containing the nonzero singular values of $\mathbf{H}$, and $\mathbf{U}_{h,\bar{r}}$ and $\mathbf{U}_{h,r}$ contain columns of $\mathbf{U}_h$ associated with the zero and nonzero singular values of $\mathbf{H}$, respectively. If $k = l = r$, (43) holds if and only if

$$\hat{\mathbf{T}} = \mathbf{\Sigma}_{h,r}^{-1}\mathbf{U}_{h,r}^H\mathbf{V}_x\mathbf{\Lambda}_x^{\frac{1}{2}}\mathbf{U}_a^H\,. \tag{44}$$

If $l > k = r$, then (43) holds if and only if

$$\big[\ \mathbf{0}_{r\times(l-r)} \quad \mathbf{\Sigma}_{h,r}\ \big]\,\hat{\mathbf{T}} = \mathbf{U}_{h,r}^H\mathbf{V}_x\mathbf{\Lambda}_x^{\frac{1}{2}}\mathbf{U}_a^H\,. \tag{45}$$

Finally, if $k > r$ and $l > r$, (43) is true if and only if $\mathbf{U}_{h,\bar{r}}^H\mathbf{V}_x = \mathbf{0}_{(k-r)\times n}$ and (45) holds.

From (45), we see that in the latter two cases, there are many solutions for $\hat{\mathbf{T}}$. We should choose $\hat{\mathbf{T}}$ that $\mathrm{tr}(\hat{\mathbf{T}}\hat{\mathbf{T}}^H) = \mathrm{tr}(\mathbf{TT}^H)$ is minimized. Such $\hat{\mathbf{T}}$ is the minimum norm solution given by

$$\hat{\mathbf{T}} = \big[\ \mathbf{0}_{r\times(l-r)} \quad \mathbf{\Sigma}_{h,r}^{-1}\ \big]^T\,\mathbf{U}_{h,r}^H\mathbf{V}_x\mathbf{\Lambda}_x^{\frac{1}{2}}\mathbf{U}_a^H\,. \tag{46}$$

Interestingly, both (44) and (46) lead to the same transmission power, given by

$$\mathrm{tr}\big(\mathbf{TT}^H\big) = \mathrm{tr}\Big(\mathbf{\Lambda}_x^{\frac{1}{2}}\mathbf{V}_x^H\mathbf{U}_{h,r}\mathbf{\Sigma}_{h,r}^{-2}\mathbf{U}_{h,r}^H\mathbf{V}_x\mathbf{\Lambda}_x^{\frac{1}{2}}\Big)\,. \tag{47}$$

Obviously, (47) is invariant to $\mathbf{U}_x$. Using Lemma 1, we know that (47) is minimized by $\mathbf{U}_{h,r}^H\mathbf{V}_x = \mathbf{\Phi}_r$. Without loss of generality, we choose $\mathbf{V}_x = \mathbf{U}_{h,r}$. From (44) and (46), we find that $\mathbf{T} = \mathbf{V}_{h,1}\mathbf{\Sigma}_{h,r}^{-1}\mathbf{\Lambda}_x^{\frac{1}{2}}\mathbf{U}_a^H$. Thus, we prove the optimal structure of $\mathbf{T}$ with $\mathbf{\Lambda} = \mathbf{\Sigma}_{h,r}^{-1}\mathbf{\Lambda}_x^{\frac{1}{2}}$. □

Now we start to prove Theorem 1. The MSE matrix (7) can be rewritten as

$$\begin{aligned}\mathbf{E}_L &= \mathbf{I}_{N_0} - \mathbf{A}_{L-1}^H\mathbf{F}_L^H\mathbf{H}_L^H\big(\mathbf{H}_L\mathbf{F}_L\mathbf{D}_{L-1}\mathbf{F}_L^H\mathbf{H}_L^H + \mathbf{I}_{N_L}\big)^{-1} \\ &\quad \times\mathbf{H}_L\mathbf{F}_L\mathbf{A}_{L-1} \qquad (48)\\ &= \mathbf{I}_{N_0} - \mathbf{A}_{L-1}^H\Big[\mathbf{D}_{L-1}^{-1} - \big(\mathbf{D}_{L-1}\mathbf{F}_L^H\mathbf{H}_L^H\mathbf{H}_L\mathbf{F}_L\mathbf{D}_{L-1} \\ &\quad +\mathbf{D}_{L-1}\big)^{-1}\Big]\mathbf{A}_{L-1} \qquad (49)\\ &= \big(\mathbf{I}_{N_0} + \mathbf{A}_{L-1}^H\mathbf{C}_{L-1}^{-1}\mathbf{A}_{L-1}\big)^{-1} + \mathbf{A}_{L-1}^H \\ &\quad \times\big(\mathbf{D}_{L-1}\mathbf{F}_L^H\mathbf{H}_L^H\mathbf{H}_L\mathbf{F}_L\mathbf{D}_{L-1} + \mathbf{D}_{L-1}\big)^{-1}\mathbf{A}_{L-1} \qquad (50)\end{aligned}$$

where the matrix inversion lemma is used to obtain (48) and (50), and the identity $\mathbf{B}^H(\mathbf{B}\mathbf{C}\mathbf{B}^H + \mathbf{I})^{-1}\mathbf{B} = \mathbf{C}^{-1} - (\mathbf{C}\mathbf{B}^H\mathbf{B}\mathbf{C}+\mathbf{C})^{-1}$ is applied to get (49). Interestingly, the first term in (50) is the MSE matrix associated with a MIMO relay constituted by the first $L-1$ hops of the original relay, and the second term in (50) is the increment of the MSE matrix introduced by the last-hop of the original relay. Moreover, since the first term in (50) is irrelevant to $\mathbf{F}_L$, the optimization problem in (10)-(12) for $\mathbf{F}_L$ can be equivalently written as

$$\min_{\mathbf{F}_L} \ \mathrm{tr}\Big(\mathbf{A}_{L-1}^H\big(\mathbf{D}_{L-1}\mathbf{F}_L^H\mathbf{H}_L^H\mathbf{H}_L\mathbf{F}_L\mathbf{D}_{L-1}+\mathbf{D}_{L-1}\big)^{-1}\mathbf{A}_{L-1}\Big) \qquad (51)$$

$$\text{s.t.} \ \ \mathrm{tr}\big(\mathbf{F}_L\mathbf{D}_{L-1}\mathbf{F}_L^H\big) \leq p_L. \qquad (52)$$

By introducing $\tilde{\mathbf{F}}_L = \mathbf{F}_L\mathbf{D}_{L-1}^{\frac{1}{2}}$, the problem (51)-(52) can be rewritten as

$$\min_{\tilde{\mathbf{F}}_L} \ \mathrm{tr}\Big(\mathbf{A}_{L-1}^H\mathbf{D}_{L-1}^{-\frac{1}{2}}\big(\tilde{\mathbf{F}}_L^H\mathbf{H}_L^H\mathbf{H}_L\tilde{\mathbf{F}}_L+\mathbf{I}_{L-1}\big)^{-1}\mathbf{D}_{L-1}^{-\frac{1}{2}}\mathbf{A}_{L-1}\Big) \qquad (53)$$

$$\text{s.t.} \ \ \mathrm{tr}\big(\tilde{\mathbf{F}}_L\tilde{\mathbf{F}}_L^H\big) \leq p_L. \qquad (54)$$

Let us introduce the following EVD and SVD: $\mathbf{H}_L^H\mathbf{H}_L = \mathbf{V}_L\mathbf{\Lambda}_L\mathbf{V}_L^H$, $\mathbf{D}_{L-1}^{-\frac{1}{2}}\mathbf{A}_{L-1} = \mathbf{U}_d\mathbf{\Lambda}_d\mathbf{V}_d^H$, where $\mathbf{\Lambda}_L$ and $\mathbf{V}_L$ are $N_{L-1}\times N_{L-1}$ matrices, the dimension of $\mathbf{U}_d$, $\mathbf{\Lambda}_d$, $\mathbf{V}_d$ are $N_{L-1}\times N_0$, $N_0\times N_0$, $N_0\times N_0$, respectively, and the diagonal elements of $\mathbf{\Lambda}_L$ and $\mathbf{\Lambda}_d$ are sorted in increasing order. It can be seen from Lemma 2 that the SVD of the optimal $\tilde{\mathbf{F}}_L$ in (53)-(54) is given by $\tilde{\mathbf{F}}_L = \mathbf{V}_{L,1}\mathbf{\Lambda}_f\mathbf{U}_d^H$, where $\mathbf{\Lambda}_f$ is the $N_0 \times N_0$ diagonal singular value matrix, and $\mathbf{V}_{L,1}$ contains the rightmost N_0 columns of $\mathbf{V}_L$. With simple manipulations, we obtain

$$\tilde{\mathbf{F}}_L = \mathbf{V}_{L,1}\mathbf{\Lambda}_f\mathbf{\Lambda}_d^{-1}\mathbf{V}_d^H\mathbf{V}_d\mathbf{\Lambda}_d\mathbf{U}_d^H = \mathbf{T}_L\mathbf{A}_{L-1}^H\mathbf{D}_{L-1}^{-\frac{1}{2}}$$

where $\mathbf{T}_L \triangleq \mathbf{V}_{L,1}\mathbf{\Lambda}_f\mathbf{\Lambda}_d^{-1}\mathbf{V}_d^H$. Therefore, we have $\mathbf{F}_L = \mathbf{T}_L\mathbf{A}_{L-1}^H\mathbf{D}_{L-1}^{-1}$.

Now by using the matrix inversion lemma, the second term in (50) can be written as

$$\begin{aligned}&\mathbf{A}_{L-1}^H\big(\mathbf{A}_{L-1}\mathbf{T}_L^H\mathbf{H}_L^H\mathbf{H}_L\mathbf{T}_L\mathbf{A}_{L-1}^H + \mathbf{D}_{L-1}\big)^{-1}\mathbf{A}_{L-1}\\ &= \mathbf{A}_{L-1}^H\Big[\mathbf{D}_{L-1}^{-1} - \mathbf{D}_{L-1}^{-1}\mathbf{A}_{L-1}\big(\mathbf{A}_{L-1}^H\mathbf{D}_{L-1}^{-1}\mathbf{A}_{L-1}\\ &\quad +(\mathbf{T}_L^H\mathbf{H}_L^H\mathbf{H}_L\mathbf{T}_L)^{-1}\big)^{-1}\mathbf{A}_{L-1}^H\mathbf{D}_{L-1}^{-1}\Big]\mathbf{A}_{L-1}\\ &= \Big[\mathbf{T}_L^H\mathbf{H}_L^H\mathbf{H}_L\mathbf{T}_L + \big(\mathbf{A}_{L-1}^H\mathbf{D}_{L-1}^{-1}\mathbf{A}_{L-1}\big)^{-1}\Big]^{-1}. \qquad (55)\end{aligned}$$

Substituting (55) back into (50) and using (15), we obtain

$$\mathbf{E}_L = \mathbf{E}_{L-1} + \big(\mathbf{T}_L^H\mathbf{H}_L^H\mathbf{H}_L\mathbf{T}_L + \mathbf{R}_L^{-1}\big)^{-1} \qquad (56)$$

where $\mathbf{E}_{L-1} = \big(\mathbf{I}_{N_0} + \mathbf{A}_{L-1}^H\mathbf{C}_{L-1}^{-1}\mathbf{A}_{L-1}\big)^{-1}$ is the MSE matrix at the $(L-1)$-th hop. Similar to (50)-(56), we can show that the optimal $\mathbf{F}_l$ is given by $\mathbf{F}_l = \mathbf{T}_l\mathbf{A}_{l-1}^H\mathbf{D}_{l-1}^{-1}$, $l = 2, \cdots, L-1$, and

$$\mathbf{E}_l = \mathbf{E}_{l-1} + \big(\mathbf{T}_l^H\mathbf{H}_l^H\mathbf{H}_l\mathbf{T}_l + \mathbf{R}_l^{-1}\big)^{-1}, \ \ l = 2, \cdots, L-1 \qquad (57)$$

$$\mathbf{E}_1 = \big(\mathbf{I}_{N_0} + \mathbf{A}_1^H\mathbf{C}_1^{-1}\mathbf{A}_1\big)^{-1} = \big(\mathbf{I}_{N_0} + \mathbf{F}_1^H\mathbf{H}_1^H\mathbf{H}_1\mathbf{F}_1\big)^{-1}. \qquad (58)$$

By combining (56)-(58), we have $\mathbf{E}_L = \big(\mathbf{I}_{N_0} + \mathbf{F}_1^H\mathbf{H}_1^H\mathbf{H}_1\mathbf{F}_1\big)^{-1} + \sum_{l=2}^{L}\big(\mathbf{R}_l^{-1} + \mathbf{T}_l^H\mathbf{H}_l^H\mathbf{H}_l\mathbf{T}_l\big)^{-1}$.

APPENDIX B
PROOF OF THEOREM 2

The proof is conducted through mathematical induction. First, for $l = 1$, we have $\mathbf{A}_1 = \mathbf{H}_1\mathbf{F}_1$, $\mathbf{D}_1 = \mathbf{H}_1\mathbf{F}_1\mathbf{F}_1^H\mathbf{H}_1^H + \mathbf{I}_{N_1}$, and $\mathbf{R}_2 = \mathbf{F}_1^H\mathbf{H}_1^H(\mathbf{H}_1\mathbf{F}_1\mathbf{F}_1^H\mathbf{H}_1^H + \mathbf{I}_{N_1})^{-1}\mathbf{H}_1\mathbf{F}_1$. By using $\mathbf{H}_1\mathbf{F}_1 = \mathbf{U}_{H_1}\mathbf{\Gamma}_1\mathbf{V}_1^H$, we obtain that $\mathbf{U}_2 = \mathbf{V}_1$, $\mathbf{A}_1^H\mathbf{D}_1^{-1} = \mathbf{V}_1\mathbf{\Gamma}_{1,1}(\mathbf{\Gamma}_{1,1}^2 + \mathbf{I}_{N_0})^{-1}\mathbf{U}_{H_1,1}^H$, where $\mathbf{\Gamma}_{1,1}$ is an $N_0 \times N_0$ diagonal matrix containing the nonzero singular values of $\mathbf{H}_1\mathbf{F}_1$. Thus from (30) we have $\mathbf{F}_2 = \mathbf{V}_{2,1}\mathbf{\Upsilon}_2\mathbf{U}_{H_1,1}^H$, where $\mathbf{\Upsilon}_2 \triangleq \mathbf{\Delta}_2\mathbf{\Gamma}_{1,1}(\mathbf{\Gamma}_{1,1}^2 + \mathbf{I}_{N_0})^{-1}$ is an $N_0 \times N_0$ diagonal matrix. The optimality of the structure of such $\mathbf{F}_2$ has been proven in [3] for multi-hop MIMO relay systems.

Second, assuming that for $l \geq 2$, the optimal $\mathbf{F}_j$ given in (30) can be written as $\mathbf{F}_j = \mathbf{V}_{j,1}\mathbf{\Upsilon}_j\mathbf{U}_{H_{j-1},1}^H$, $j = 2, \cdots, l$, we now show that $\mathbf{F}_{l+1}$ in (30) can be written as $\mathbf{F}_{l+1} = \mathbf{V}_{l+1,1}\mathbf{\Upsilon}_{l+1}\mathbf{U}_{H_l,1}^H$. In fact, from (5) we can write $\mathbf{A}_l = \mathbf{U}_{H_l,1}\bigotimes_{i=l}^{2}(\mathbf{\Gamma}_{i,1}\mathbf{\Upsilon}_i)\mathbf{\Gamma}_{1,1}\mathbf{V}_1^H$, where $\mathbf{\Gamma}_{i,1}$ is an $N_0 \times N_0$ diagonal matrix containing the largest N_0 singular values of $\mathbf{H}_i$. Then from (8) we have $\mathbf{D}_l = \mathbf{U}_{H_l,1}\sum_{j=1}^{l}\big(\bigotimes_{i=l}^{j}(\mathbf{\Gamma}_{i,1}^2\mathbf{\Upsilon}_i^2)\mathbf{\Gamma}_{1,1}^2\big)\mathbf{U}_{H_l,1}^H + \mathbf{I}_{N_l}$, from (15) we get $\mathbf{R}_{l+1} = \mathbf{V}_1\bigotimes_{i=l}^{2}(\mathbf{\Gamma}_{i,1}^2\mathbf{\Upsilon}_i^2)\mathbf{\Gamma}_{1,1}^2\big(\sum_{j=1}^{l}\big(\bigotimes_{i=l}^{j}(\mathbf{\Gamma}_{i,1}^2\mathbf{\Upsilon}_i^2)\mathbf{\Gamma}_{1,1}^2\big) + \mathbf{I}_{N_0}\big)^{-1}\mathbf{V}_1^H$, and thus $\mathbf{U}_{l+1} = \mathbf{V}_1$. Finally, from (30) we have $\mathbf{F}_{l+1} = \mathbf{V}_{l+1,1}\mathbf{\Upsilon}_{l+1}\mathbf{U}_{H_l,1}^H$, where

$$\mathbf{\Upsilon}_{l+1} = \mathbf{\Delta}_{l+1}\bigotimes_{i=l}^{2}(\mathbf{\Gamma}_{i,1}\mathbf{\Upsilon}_i)\mathbf{\Gamma}_{1,1} \times\left(\sum_{j=1}^{l}\Big(\bigotimes_{i=l}^{j}(\mathbf{\Gamma}_{i,1}^2\mathbf{\Upsilon}_i^2)\mathbf{\Gamma}_{1,1}^2\Big) + \mathbf{I}_{N_0}\right)^{-1}.$$

Therefore, based on the principle of mathematical induction, we have $\mathbf{F}_l = \mathbf{V}_{l,1}\mathbf{\Upsilon}_l\mathbf{U}_{H_{l-1},1}^H$, $l = 2, \cdots, L$.

REFERENCES

[1] X. Tang and Y. Hua, "Optimal design of non-regenerative MIMO wireless relays," *IEEE Trans. Wireless Commun.*, vol. 6, pp. 1398-1407, Apr. 2007.

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

[3] Y. Rong and Y. Hua, "Optimality of diagonalization of multi-hop MIMO relays," *IEEE Trans. Wireless Commun.*, vol. 8, pp. 6068-6077, Dec. 2009.

[4] C.-B. Chae, T. Tang, R. W. Heath Jr., and S. Cho, "MIMO relaying with linear processing for multiuser transmission in fixed relay networks," *IEEE Trans. Signal Process.*, vol. 56, pp. 727-738, Feb. 2008.

[5] Y. Yu and Y. Hua, "Power allocation for a MIMO relay system with multiple-antenna users," *IEEE Trans. Signal Process.*, vol. 58, pp. 2823-2835, May 2010.

[6] M. R. A. Khandaker and Y. Rong, “Joint source and relay optimization for multiuser MIMO relay communication systems," in *Proc. 4th Intern. Conf. Signal Process. Commun. Sys.*, Dec. 2010.

[7] R. Zhang, C. C. Chai, and Y. C. Liang, “Joint beamforming and power control for multi-antenna relay broadcast channel with QoS constraints," *IEEE Trans. Signal Process.*, vol. 57, pp. 726-737, Feb. 2009.

[8] M. Cao, X. Wang, S.-J. Kim, and M. Madihian, “Multi-hop wireless backhaul networks: a cross-layer design paradigm," *IEEE J. Sel. Areas Commun.*, vol. 25, pp. 738-748, May 2007.

[9] V. Genc, S. Murphy, Y. Yu, and J. Murphy, “IEEE 802.16j relay-based wireless access networks: an overview," *IEEE Wireless Commun.*, vol. 15, pp. 56-63, Oct. 2008.

[10] C. Song, K.-J. Lee, and I. Lee, “MMSE based transceiver designs in closed-loop non-regenerative MIMO relaying systems," *IEEE Trans. Wireless Commun.*, vol. 9, pp. 2310-2319, July 2010.

[11] Y. Rong and M. R. A. Khandaker, “On uplink-downlink duality of multi-hop MIMO relay channel," *IEEE Trans. Wireless Commun.*, vol. 10, pp. 1923-1931, June 2011.

[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. Signal Process.*, vol. 51, pp. 2381-2401, Sep. 2003.

[13] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory.* Prentice Hall, 1993.

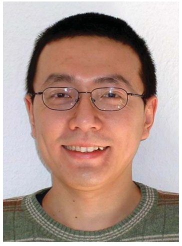
[14] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[15] P. Tseng, “Further results on approximating nonconvex quadratic optimization by semidefinite programming relaxation," *SIAM J. Optim.*, vol. 14, no. 1, pp. 268-283, July 2003.

[16] P. Lioliou and M. Viberg, “Least-squares based channel estimation for MIMO relays," in *Proc. IEEE WSA 2008*, Feb. 2008, pp. 90-95.

[17] M. Grant and S. Boyd, “Cvx: Matlab software for disciplined convex programming (webpage and software)." Available: http://cvxr.com/cvx, Apr. 2010.

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



Yue Rong (S'03-M'06-SM'11) received the B.E. degree from Shanghai Jiao Tong University, China, the M.Sc. degree from the University of Duisburg-Essen, Duisburg, Germany, and the Ph.D. degree (summa cum laude) from Darmstadt University of Technology, Darmstadt, Germany, all in electrical engineering, in 1999, 2002, and 2005, respectively.

From April 2001 to October 2001, he was a research assistant at the Fraunhofer Institute of Microelectronic Circuits and Systems, Duisburg, Germany. From October 2001 to March 2002, he was with Nokia Ltd., Bochum, Germany. From November 2002 to March 2005, he was a Research Associate at the Department of Communication Systems in the University of Duisburg-Essen. From April 2005 to January 2006, he was with the Institute of Telecommunications at Darmstadt University of Technology, as a Research Associate. From February 2006 to November 2007, he was a Postdoctoral Researcher with the Department of Electrical Engineering, University of California, Riverside. Since December 2007, he has been with the Department of Electrical and Computer Engineering, Curtin University of Technology, Perth, Australia, where he is now a Senior Lecturer. His research interests include signal processing for communications, wireless communications, wireless networks, applications of linear algebra and optimization methods, and statistical and array signal processing.

Dr. Rong received the Best Paper Award at the 16th Asia-Pacific Conference on Communications, Auckland, New Zealand, 2010, the 2010 Young Researcher of the Year Award of the Faculty of Science and Engineering at Curtin University, the 2004 Chinese Government Award for Outstanding Self-Financed Students Abroad (China), and the 2001-2002 DAAD/ABB Graduate Sponsoring Asia Fellowship (Germany). He has co-authored more than 50 referred IEEE journal and conference papers. He is a Guest Editor of the IEEE JSAC special issue on Theories and Methods for Advanced Wireless Relays to be published in 2012, a Guest Editor of the EURASIP JASP Special Issue on Signal Processing Methods for Diversity and Its Applications to be published in 2012, and has served as a TPC member for IEEE ICC, IWCMC, and ChinaCom.