

# Multiuser Multi-Hop MIMO Relay Systems with Correlated Fading Channels

Yue Rong, *Senior Member, IEEE*, and Yong Xiang

**Abstract**—In this letter, we address multiuser multi-hop multiple-input multiple-output (MIMO) relay communication systems with correlated MIMO fading channels. In particular, we consider the practical scenario where the channel fading is fast and thus the instantaneous channel state information (CSI) is only available at the destination node, but unknown at all users and all relay nodes. We derive the structure of the optimal user precoding matrices and relay amplifying matrices that maximizes the users-destination ergodic sum mutual information. Compared with existing works, our results are more general, since we address multiuser scenarios, consider MIMO relays with a finite dimension, and take into account the noise vector at each relay node.

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

## I. INTRODUCTION

NON-regenerative multiple-input multiple-output (MIMO) relay communication systems have been recently investigated under various assumptions on the availability of channel state information (CSI) [1]–[8]. Given the instantaneous CSI of all hops, the optimal source precoding matrix and relay amplifying matrices of a single-user multi-hop relay system were developed in [2] for a broad class of commonly used objective functions in MIMO system design. A minimal mean-squared error (MMSE)-based optimal two-hop multiuser MIMO relay system has been recently proposed in [3] where the instantaneous CSI is known. In [4], the source and relay matrices were optimized for a single-user two-hop relay system where the instantaneous source-relay CSI is available, while only the covariance matrix of the relay-destination channel is known at the relay node. The ergodic mutual information (MI) of a two-hop MIMO relay channel with a large number of antennas has been analyzed in [5] and [6] using the information on the covariance matrices of MIMO channels. The exact analytic form of the ergodic capacity of a two-hop amplify-and-forward MIMO relay channel with finite dimension has been recently derived in [7]. In [8], by neglecting the noise at all the relay nodes, Fawaz *et al.* derived the optimal source and relay matrices of a single-user multi-hop MIMO relay system with correlated fading channels.

Manuscript received January 24, 2011; revised April 4, 2011; accepted June 19, 2011. The associate editor coordinating the review of this letter and approving it for publication was A. Nosratinia.

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

Y. Xiang is with the School of Engineering, Deakin University, Geelong, VIC 3217, Australia (e-mail: yong.xiang@deakin.edu.au).

This work was supported under the Australian Research Council's Discovery Projects funding scheme (project numbers DP110100736, DP110102076). Digital Object Identifier 10.1109/TWC.2011.070711.110160

In this letter, we focus on multiaccess communication through multi-hop linear non-regenerative relays. We consider the practical scenario where the channel fading is fast and thus the instantaneous CSI is only available at the destination node, but unknown at other nodes. We assume that each MIMO channel is correlated at both the transmit and the receive side, each user knows the correlation matrix of the first hop channel, and each relay node has the knowledge on the correlation matrices of its direct backward and forward channel. Based on the knowledge of the channel correlation matrices, we show that to optimize the users-destination ergodic sum MI, the optimal precoding matrix at each user is the product of the eigenvector matrix of the transmit side correlation matrix of the first-hop channel and a diagonal power loading matrix. And the optimal amplifying matrix at each relay node is the product of the eigenvector matrix of the transmit side correlation matrix of the forward channel, a diagonal power loading matrix, and Hermitian transpose of the eigenvector matrix of the receive side correlation matrix of the backward channel. Compared with existing works in this area (for example [5], [6], [8]), our results in this letter are more general, since we address multiuser scenarios, consider MIMO relays with a finite dimension, and take into account the noise vector at all relay nodes. Another advantage of our algorithm is that power allocation is computed locally at each node based on the local knowledge on the channel correlation.

The rest of this letter 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 user and relay matrices design algorithm is 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 system with  $N_u$  users simultaneously transmitting information to a common destination node through  $L - 1$  relay nodes as illustrated in Fig. 1. We assume that the  $i$ th user has  $M_i$  antennas,  $i = 1, \dots, N_u$ , the  $l$ th relay node is equipped with  $N_l$  antennas,  $l = 1, \dots, L - 1$ , and the destination node has  $N_L$  antennas. 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 with a linear receiver at the destination node, there should be  $N_0 \leq \min(N_1, \dots, N_L)$ , since otherwise the system can not support  $N_0$  active symbols in each transmission. This condition establishes the upper-bound for the total number of data streams that can be concurrently transmitted from all users.

The  $M_i \times 1$  modulated signal vector  $\mathbf{s}_i$  at the  $i$ th user is linearly precoded by the  $M_i \times M_i$  user precoding matrix  $\mathbf{B}_i$ ,

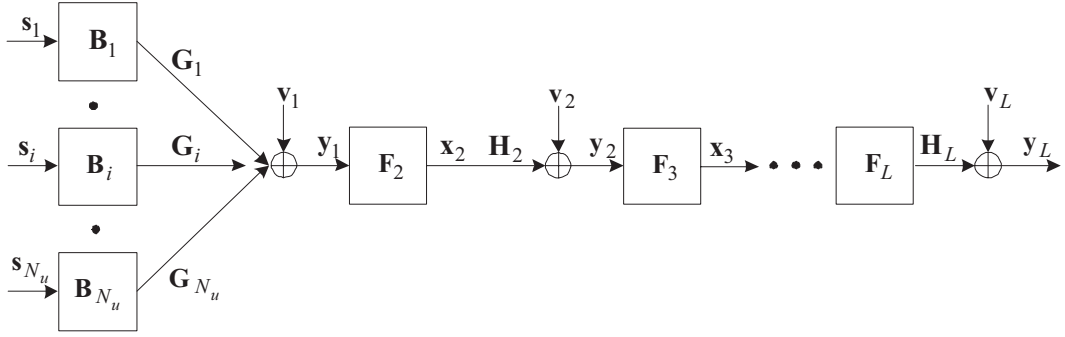


Fig. 1. Block diagram of an  $N_u$ -user  $L$ -hop linear non-regenerative MIMO relay communication system.

and the precoded signal vector  $\mathbf{u}_i = \mathbf{B}_i \mathbf{s}_i$  is transmitted to the first relay node. The  $N_1 \times 1$  signal vector  $\mathbf{y}_1$  received at the first relay node is given by

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

where  $\mathbf{G}_i$ ,  $i = 1, \dots, N_u$ , is the  $N_1 \times M_i$  MIMO channel matrix between the first relay node and the  $i$ th 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,  $\mathbf{x}_1 = \mathbf{F}_1 \mathbf{s}$ ,  $\mathbf{s} \triangleq [\mathbf{s}_1^T, \dots, \mathbf{s}_{N_u}^T]^T$ , and

$$\mathbf{H}_1 \triangleq [\mathbf{G}_1, \dots, \mathbf{G}_{N_u}], \quad \mathbf{F}_1 \triangleq \text{bd}(\mathbf{B}_1, \dots, \mathbf{B}_{N_u}). \quad (2)$$

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 user precoding matrix,  $\mathbf{s}$  is an  $N_0 \times 1$  vector containing source symbols from all users,  $\text{bd}(\cdot)$  stands for a block diagonal matrix, and  $(\cdot)^T$  denotes matrix (vector) transpose. We assume that  $\mathbb{E}[\mathbf{s}\mathbf{s}^H] = \mathbf{I}_{N_0}$ , where  $\mathbb{E}[\cdot]$  stands for the statistical expectation,  $(\cdot)^H$  denotes the Hermitian transpose, and  $\mathbf{I}_n$  is an  $n \times n$  identity matrix.

Using a linear nonregenerative relay matrix at each relay as in [1]-[8], the input-output relationship at the  $l$ th relay nodes is

$$\mathbf{x}_{l+1} = \mathbf{F}_{l+1} \mathbf{y}_l, \quad l = 1, \dots, L-1 \quad (3)$$

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

$$\mathbf{y}_l = \mathbf{H}_l \mathbf{x}_l + \mathbf{v}_l, \quad l = 1, \dots, L-1 \quad (4)$$

where  $\mathbf{H}_l$ ,  $l = 1, \dots, L-1$ , is the  $N_l \times N_{l-1}$  MIMO channel matrix of the  $l$ th hop, and  $\mathbf{v}_l$  is the i.i.d. AWGN vector at the  $l$ th relay node. We assume that all noises are complex circularly symmetric with zero mean and unit variance.

From (1)-(4), the received signal vector at the destination node is given by

$$\mathbf{y}_L = \mathbf{A} \mathbf{s} + \bar{\mathbf{v}} \quad (5)$$

where  $\mathbf{A}$  is the equivalent MIMO channel matrix from all users to the destination, and  $\bar{\mathbf{v}}$  is the equivalent noise vector at the destination node given by [2]

$$\mathbf{A} = \bigotimes_{i=L}^1 (\mathbf{H}_i \mathbf{F}_i), \quad \bar{\mathbf{v}} = \sum_{l=2}^L \left( \bigotimes_{i=L}^l (\mathbf{H}_i \mathbf{F}_i) \mathbf{v}_{l-1} \right) + \mathbf{v}_L. \quad (6)$$

Here for matrices  $\mathbf{X}_i$ ,  $\bigotimes_{i=l}^k (\mathbf{X}_i) \triangleq \mathbf{X}_l \cdots \mathbf{X}_k$ . The covariance matrix of  $\bar{\mathbf{v}}$  in (6) is given by

$$\mathbf{C}_{\bar{\mathbf{v}}} = \sum_{l=2}^L \left( \bigotimes_{i=L}^l (\mathbf{H}_i \mathbf{F}_i) \bigotimes_{i=l}^L (\mathbf{F}_i^H \mathbf{H}_i^H) \right) + \mathbf{I}_{N_L}.$$

We would like to mention that the system model (3)-(5) is applicable for relays working in full-duplex mode as in [8] or half-duplex mode as in [1]-[3].

We assume that the channel fading is fast and thus the instantaneous CSI of  $\mathbf{A}$  is only available at the destination node, but unknown at all users and all relay nodes. We also assume that each MIMO channel is correlated at both the transmit and receive side. Let us introduce  $\Theta_{t,l}$  and  $\Theta_{r,l}$  as the correlation matrix at the transmit and receive side of  $\mathbf{H}_l$ ,  $l = 2, \dots, L$ , respectively, and  $\Phi_{t,i}$  as the correlation matrix at the transmit side of  $\mathbf{G}_i$ ,  $i = 1, \dots, N_u$ . We assume that the correlation matrix  $\Phi_r$  at the receive side of  $\mathbf{G}_i$ ,  $i = 1, \dots, N_u$  are identical. Thus the instantaneous channel matrices can be represented as

$$\mathbf{G}_i = \Phi_r^{\frac{1}{2}} \mathbf{G}_i^w \Phi_{t,i}^{\frac{1}{2}}, \quad i = 1, \dots, N_u \quad (7)$$

$$\mathbf{H}_l = \Theta_{r,l}^{\frac{1}{2}} \mathbf{H}_l^w \Theta_{t,l}^{\frac{1}{2}}, \quad l = 2, \dots, L \quad (8)$$

where  $\mathbf{G}_i^w$  and  $\mathbf{H}_l^w$  are Gaussian random matrices with i.i.d. zero mean and unit variance entries and unknown to all users and all relay nodes. We assume that each node only has the local channel correlation knowledge. In particular, the  $i$ th user knows  $\Phi_i$ ,  $i = 1, \dots, N_u$ , the first relay node knows  $\Phi_r$  and  $\Theta_{t,2}$ , and the  $l$ th relay node has the knowledge of  $\Theta_{r,l}$  and  $\Theta_{t,l+1}$ ,  $l = 2, \dots, L-1$ . Since channel correlation matrices vary much slower than the instantaneous channel, they can be estimated at the corresponding node with a reasonably high precision. We also assume that channel correlation matrices  $\Theta_{r,l}$ ,  $\Theta_{t,l}$ ,  $l = 2, \dots, L$ ,  $\Phi_r$ , and  $\Phi_{t,i}$ ,  $i = 1, \dots, N_u$ , are all full rank.

The sum MI of the users-destination channel (5) for one channel realization is given by

$$\begin{aligned} \text{MI} &= \log_2 |\mathbf{I}_{N_0} + \mathbf{A}^H \mathbf{C}_{\bar{\mathbf{v}}}^{-1} \mathbf{A}| \\ &= \log_2 \left| \mathbf{I}_{N_0} + \bigotimes_{i=1}^L (\mathbf{F}_i^H \mathbf{H}_i^H) \left( \sum_{l=2}^L \left( \bigotimes_{i=L}^l (\mathbf{H}_i \mathbf{F}_i) \right. \right. \right. \\ &\quad \left. \left. \left. \bigotimes_{i=l}^L (\mathbf{F}_i^H \mathbf{H}_i^H) \right) + \mathbf{I}_{N_L} \right)^{-1} \bigotimes_{i=L}^1 (\mathbf{H}_i \mathbf{F}_i) \right| \quad (9) \end{aligned}$$

where  $|\cdot|$ ,  $(\cdot)^{-1}$  denote matrix determinant and inversion, respectively.

### III. OPTIMAL STRUCTURE OF USER AND RELAY MATRICES

We consider a practical MIMO relay system where the precoding/amplifying matrix is designed locally at each user/relay node. Obviously, since the channel fading is fast and the instantaneous CSI is unknown at users and relay nodes, it is impossible to design  $\{\mathbf{B}_i\}$  and  $\{\mathbf{F}_l\}$  to maximize MI in (9), where  $\{\mathbf{F}_l\} \triangleq \{\mathbf{F}_l, l = 2, \dots, L\}$ , and  $\{\mathbf{B}_i\} \triangleq \{\mathbf{B}_i, i = 1, \dots, N_u\}$ . In this letter, we aim at maximizing  $E_H[\text{MI}]$  by exploiting the channel statistics, where the expectation  $E_H$  is over the realization of all channels. The maximal ergodic sum MI-based optimal multiuser multi-hop MIMO relay design problem is given as

$$\max_{\{\mathbf{F}_l\}, \{\mathbf{B}_i\}} E_H[\text{MI}] \quad (10)$$

$$\text{s.t. } E_H \left[ \text{tr} \left( \mathbf{F}_{l+1} \left( \sum_{j=1}^l \left( \bigotimes_{i=l}^j (\mathbf{H}_i \mathbf{F}_i) \bigotimes_{i=j}^l (\mathbf{F}_i^H \mathbf{H}_i^H) \right) + \mathbf{I}_{N_l} \right) \mathbf{F}_{l+1}^H \right) \right] \leq p_{l+1}, \quad l = 1, \dots, L-1 \quad (11)$$

$$\text{tr}(\mathbf{B}_i \mathbf{B}_i^H) \leq q_i, \quad i = 1, \dots, N_u \quad (12)$$

where  $\text{tr}(\cdot)$  denotes matrix trace, (11) contains the averaged transmission power constraint at each relay node, (12) includes the transmission power constraint at each user,  $p_l$  and  $q_i$  are the corresponding power budget.

Let us introduce the eigenvalue decompositions (EVDs) of

$$\Theta_{r,l} = \mathbf{V}_{\theta,l} \mathbf{\Lambda}_{\theta,l} \mathbf{V}_{\theta,l}^H, \quad \Theta_{t,l} = \mathbf{U}_{\theta,l} \mathbf{\Sigma}_{\theta,l} \mathbf{U}_{\theta,l}^H, \quad l = 2, \dots, L \quad (13)$$

$$\Phi_r = \mathbf{V}_{\theta,1} \mathbf{\Lambda}_{\theta,1} \mathbf{V}_{\theta,1}^H, \quad \Phi_{t,i} = \mathbf{U}_{\phi,i} \mathbf{\Sigma}_{\phi,i} \mathbf{U}_{\phi,i}^H, \quad i = 1, \dots, N_u \quad (14)$$

where the main diagonal elements of the eigenvalue matrices are sorted in a decreasing order. The following Theorem establishes the structure of the optimal user precoding matrices and relay amplifying matrices.

**THEOREM 1:** The optimal structure of  $\mathbf{B}_i$  and  $\mathbf{F}_l$  in the form of their singular value decompositions (SVDs) is given by

$$\mathbf{B}_i = \mathbf{U}_{\phi,i} \mathbf{\Delta}_{b,i}, \quad i = 1, \dots, N_u, \quad (15)$$

$$\mathbf{F}_l = \mathbf{U}_{\theta,l} \mathbf{\Delta}_{f,l} \mathbf{V}_{\theta,l-1}^H, \quad l = 2, \dots, L. \quad (16)$$

**PROOF:** See Appendix A.  $\square$

Interestingly, from (15) and (16) we find that the optimal precoding matrix at each user is the product of the eigenvector matrix of the transmit side correlation matrix of the first-hop channel and a diagonal power loading matrix. And the optimal amplifying matrix at each relay node is the product of the eigenvector matrix of the transmit side correlation matrix of the forward channel, a diagonal power loading matrix, and Hermitian transpose of the eigenvector matrix of the receive side correlation matrix of the backward channel. Compared with [8], our result in (15)-(16) is more general, since it holds for multiuser scenarios by considering MIMO relays with a finite dimension, and taking into account the noise vector at all

relay nodes. Using (15)-(16), the optimal relay design problem (10)-(12) is converted to the following power loading problem

$$\max_{\{\Delta_{b,i}\}, \{\Delta_{f,l}\}} E_H \left[ \log_2 \left| \mathbf{I}_{N_0} + \bigotimes_{i=1}^L (\mathbf{\Sigma}_i \bar{\mathbf{H}}_i^H) \left( \sum_{l=2}^L \bigotimes_{i=L}^l (\bar{\mathbf{H}}_i \mathbf{\Sigma}_i) \right) \times \mathbf{\Lambda}_{\theta,l-1}^{-1} \bigotimes_{i=l}^L (\mathbf{\Sigma}_i \bar{\mathbf{H}}_i^H) + \mathbf{\Lambda}_{\theta,L}^{-1} \bigotimes_{i=L}^L (\bar{\mathbf{H}}_i \mathbf{\Sigma}_i) \right| \right] \quad (17)$$

$$\text{s.t. } \text{tr}(\mathbf{\Delta}_{f,l+1}^2 \mathbf{\Lambda}_{\theta,l}) \sum_{j=1}^l \alpha_j \text{tr}(\mathbf{\Sigma}_j^2 \mathbf{\Pi}_j) + \text{tr}(\mathbf{\Delta}_{f,l+1}^2) \leq p_{l+1}, \quad l = 1, \dots, L-1 \quad (18)$$

$$\text{tr}(\mathbf{\Delta}_{b,i}^2) \leq q_i, \quad i = 1, \dots, N_u \quad (19)$$

where the definitions of  $\bar{\mathbf{H}}_l$ ,  $\alpha_l$ ,  $\mathbf{\Pi}_l$ , and  $\mathbf{\Sigma}_l$ ,  $l = 1, \dots, L$ , are given in Appendix A.

For a multi-hop MIMO relay system with a finite dimension, the analytic expression of (17) is still an open problem. Thus, the problem (17)-(19) is very difficult to solve. In this letter, we take a simple (but possibly) suboptimal solution such that  $\mathbf{\Delta}_{f,l}^2 = \gamma_l \mathbf{I}_{N_{l-1}}$ ,  $l = 2, \dots, L$ , and  $\mathbf{\Delta}_{b,i}^2 = \beta_i \mathbf{I}_{M_i}$ ,  $i = 1, \dots, N_u$ . From (33) in Appendix A, we obtain that  $\alpha_j = \prod_{i=l}^{j+1} \gamma_i \text{tr}(\mathbf{\Sigma}_{\theta,i} \mathbf{\Lambda}_{\theta,i-1})$ ,  $j = 1, \dots, l-1$ , and thus we have that

$$\sum_{j=1}^l \alpha_j \text{tr}(\mathbf{\Sigma}_j^2 \mathbf{\Pi}_j) = \sum_{j=1}^{l-1} \alpha_j \text{tr}(\mathbf{\Sigma}_j^2 \mathbf{\Pi}_j) + \text{tr}(\mathbf{\Sigma}_l^2 \mathbf{\Lambda}_{\theta,l-1}^{-1}) = \sum_{j=1}^{l-1} \prod_{i=l}^{j+1} \gamma_i \text{tr}(\mathbf{\Sigma}_{\theta,i} \mathbf{\Lambda}_{\theta,i-1}) \text{tr}(\mathbf{\Delta}_{f,j}^2 \mathbf{\Sigma}_{\theta,j}) + \text{tr}(\mathbf{\Delta}_{f,l}^2 \mathbf{\Sigma}_{\theta,l}) = \sum_{j=1}^{l-1} \prod_{i=l}^{j+1} \gamma_i \text{tr}(\mathbf{\Sigma}_{\theta,i} \mathbf{\Lambda}_{\theta,i-1}) \gamma_j \text{tr}(\mathbf{\Sigma}_{\theta,j}) + \gamma_l \text{tr}(\mathbf{\Sigma}_{\theta,l}), \quad l \geq 2 \quad (20)$$

$$\sum_{j=1}^l \alpha_j \text{tr}(\mathbf{\Sigma}_j^2 \mathbf{\Pi}_j) = \sum_{i=1}^{N_u} \beta_i \text{tr}(\mathbf{\Sigma}_{\phi,i}), \quad l = 1 \quad (21)$$

where for simplicity, we define  $\gamma_1 \text{tr}(\mathbf{\Sigma}_{\theta,1}) \triangleq \sum_{i=1}^{N_u} \beta_i \text{tr}(\mathbf{\Sigma}_{\phi,i})$ . From (18)-(21), we now obtain that

$$\gamma_{l+1} = \frac{p_{l+1}}{\text{tr}(\mathbf{\Lambda}_{\theta,l}) \sum_{j=1}^l \alpha_j \text{tr}(\mathbf{\Sigma}_j^2 \mathbf{\Pi}_j) + N_l}, \quad l = 1, \dots, L-1$$

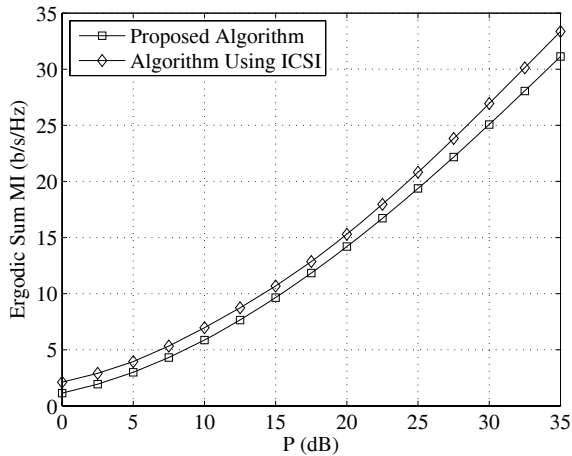
$$\beta_i = \frac{q_i}{M_i}, \quad i = 1, \dots, N_u.$$

Finally, from (15)-(16), the user precoding matrices and relay amplifying matrices are given by

$$\mathbf{B}_i = \sqrt{\beta_i} \mathbf{U}_{\phi,i}, \quad i = 1, \dots, N_u, \quad (22)$$

$$\mathbf{F}_l = \sqrt{\gamma_l} \mathbf{U}_{\theta,l} \mathbf{V}_{\theta,l-1}^H, \quad l = 2, \dots, L. \quad (23)$$

A distinct advantage of the user and relay matrices design in (22)-(23) is that it is robust and very simple to implement. It does not require any instantaneous CSI. In fact, each node only needs the local knowledge on the channel correlation matrix, performing the EVD and computing the power allocation scaler locally. Such simple design is of great importance for

Fig. 2. Ergodic sum MI versus  $P$ .  $L = 2$ ,  $N = 4$ , and  $N_u = 2$ .

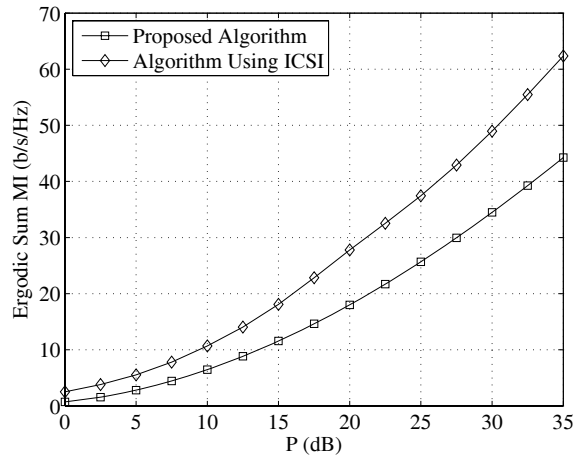
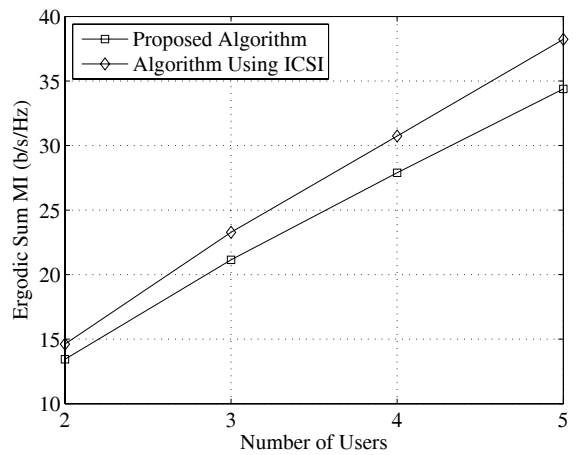
practical multiuser multi-hop MIMO relay systems. Interestingly, in Section IV we will see that for a two-hop multiuser MIMO relay system, the user and relay matrices design in (22)-(23) only has a small performance degradation compared with the optimal user and relay design using the instantaneous CSI.

#### IV. NUMERICAL EXAMPLES

In this section, we study the performance of the proposed multiuser multi-hop MIMO relay design algorithm through numerical simulations. For simplicity, we consider systems where all users have the same number of antennas (i.e.,  $M_i = M$ ,  $i = 1, \dots, N_u$ ), all relay nodes and the destination node are equipped with the same number of antennas (i.e.,  $N_l = N$ ,  $l = 1, \dots, N_L$ ), and  $MN_u = N$ . We assume that all users have an identical transmission power  $q_i = P/N_u$ ,  $i = 1, \dots, N_u$ , and all relay nodes have the same transmission power  $p_l = P$ ,  $l = 2, \dots, L$ . We also assume that the channel correlation matrices have the commonly used exponential Toeplitz structure [5], [8] such that  $[\Theta_{t,l}]_{m,n} = \theta_{t,l}^{|m-n|}$ ,  $[\Theta_{r,l}]_{m,n} = \theta_{r,l}^{|m-n|}$ ,  $[\Phi_{t,i}]_{m,n} = \phi_{t,i}^{|m-n|}$ , and  $[\Phi_r]_{m,n} = \phi_r^{|m-n|}$ . The ergodic sum MI is obtained by averaging over  $10^4$  channel realizations. In the simulations, we compare the proposed algorithm with the optimal user and relay matrices design for multiuser multi-hop MIMO relay systems using the instantaneous CSI (ICSI). The ICSI-based scheme uses the iterative algorithm developed in [3] with the source-destination MI as the objective function.

In our first example, we set  $L = 2$ ,  $N = 4$ , and  $N_u = 2$ . The channel correlation parameters are chosen as  $\phi_{t,1} = 0.5$ ,  $\phi_{t,2} = 0.6$ ,  $\phi_r = 0.2$ ,  $\theta_{t,2} = 0.3$ ,  $\theta_{r,2} = 0.5$ . Fig. 2 shows the ergodic sum MI comparison between two algorithms. It can be seen that the performance of the proposed algorithm is very close to that of the optimal algorithm using the ICSI.

In the second example, we set  $L = 4$ ,  $N = 9$ ,  $N_u = 3$ , and choose  $\phi_{t,1} = 0.5$ ,  $\phi_{t,2} = 0.6$ ,  $\phi_{t,3} = 0.4$ ,  $\phi_r = 0.2$ ,  $\theta_{t,2} = 0.3$ ,  $\theta_{t,3} = \theta_{t,4} = 0.4$ ,  $\theta_{r,2} = 0.5$ ,  $\theta_{r,3} = 0.3$ ,  $\theta_{r,4} = 0.4$  as the channel correlation parameters. It can be seen that with the increasing number of hops, the gap between the

Fig. 3. Ergodic sum MI versus  $P$ .  $L = 4$ ,  $N = 9$ , and  $N_u = 3$ .Fig. 4. Ergodic sum MI versus  $N_u$ .  $L = 2$ ,  $M = 2$ , and  $P = 20\text{dB}$ .

proposed algorithm and the optimal algorithm using the ICSI increases. However, the former algorithm has a much lower computational complexity and signalling overhead compared with the latter algorithm.

We investigate the ergodic sum MI of both algorithms versus the number of users in our third example. We set  $L = 2$ ,  $M = 2$ ,  $P = 20\text{dB}$ ,  $\phi_{t,i} = 0.5$ ,  $i = 1, \dots, N_u$ ,  $\phi_r = 0.2$ ,  $\theta_{t,2} = 0.3$ , and  $\theta_{r,2} = 0.5$ . It can be seen from Fig. 4 that the MI gap between the proposed algorithm and the ICSI-based algorithm increases slightly as the number of users increases.

#### V. CONCLUSIONS

We developed the optimal structure of user precoding matrices and relay amplifying matrices to optimize the ergodic sum mutual information of multiuser multi-hop MIMO relay communication systems. A simple power loading algorithm was proposed. The proposed algorithm only has a small performance degradation compared with the optimal user and relay design using the instantaneous CSI, but greatly reduced computational complexity and signalling overhead.

APPENDIX A  
PROOF OF THEOREM 1

The proof is conducted in three steps: First, we show that  $\mathbf{B}_i$  in (15) minimizes the left-hand side of (12). Second, we prove that  $\mathbf{F}_l$  in (16) minimizes the left-hand side of (11). At last, we show that (15)-(16) is also optimal for the objective function (10).

Using (13) and (14), we can write the channel matrices in (7) and (8) as

$$\begin{aligned} \mathbf{G}_i &= \mathbf{V}_{\theta,1} \Lambda_{\theta,1}^{\frac{1}{2}} \mathbf{V}_{\theta,1}^H \mathbf{G}_i^w \mathbf{U}_{\phi,i} \Sigma_{\phi,i}^{\frac{1}{2}} \mathbf{U}_{\phi,i}^H \\ &= \mathbf{V}_{\theta,1} \Lambda_{\theta,1}^{\frac{1}{2}} \bar{\mathbf{G}}_i \Sigma_{\phi,i}^{\frac{1}{2}} \mathbf{U}_{\phi,i}^H, \quad i = 1, \dots, N_u \end{aligned} \quad (24)$$

$$\begin{aligned} \mathbf{H}_l &= \mathbf{V}_{\theta,l} \Lambda_{\theta,l}^{\frac{1}{2}} \mathbf{V}_{\theta,l}^H \mathbf{H}_l^w \mathbf{U}_{\theta,l} \Sigma_{\theta,l}^{\frac{1}{2}} \mathbf{U}_{\theta,l}^H \\ &= \mathbf{V}_{\theta,l} \Lambda_{\theta,l}^{\frac{1}{2}} \bar{\mathbf{H}}_l \Sigma_{\theta,l}^{\frac{1}{2}} \mathbf{U}_{\theta,l}^H, \quad l = 2, \dots, L \end{aligned} \quad (25)$$

where  $\bar{\mathbf{G}}_i \triangleq \mathbf{V}_{\theta,1}^H \mathbf{G}_i^w \mathbf{U}_{\phi,i}$  and  $\bar{\mathbf{H}}_l \triangleq \mathbf{V}_{\theta,l}^H \mathbf{H}_l^w \mathbf{U}_{\theta,l}$  have the same statistics as  $\mathbf{G}_i^w$  and  $\mathbf{H}_l^w$ , respectively. From (24) and (25) we have

$$\begin{aligned} \mathbf{H}_1 \mathbf{F}_1 &= \mathbf{V}_{\theta,1} \Lambda_{\theta,1}^{\frac{1}{2}} [\bar{\mathbf{G}}_1, \dots, \bar{\mathbf{G}}_{N_u}] \\ &\quad \times \text{bd}(\Sigma_{\phi,1}^{\frac{1}{2}} \mathbf{U}_{\phi,1}^H \mathbf{B}_1, \dots, \Sigma_{\phi,N_u}^{\frac{1}{2}} \mathbf{U}_{\phi,N_u}^H \mathbf{B}_{N_u}) \\ \mathbf{H}_l \mathbf{F}_l &= \mathbf{V}_{\theta,l} \Lambda_{\theta,l}^{\frac{1}{2}} \bar{\mathbf{H}}_l \Sigma_{\theta,l}^{\frac{1}{2}} \mathbf{U}_{\theta,l}^H \mathbf{F}_l, \quad l = 2, \dots, L. \end{aligned}$$

Let us also introduce the following SVDs

$$\Sigma_{\phi,i}^{\frac{1}{2}} \mathbf{U}_{\phi,i}^H \mathbf{B}_i = \mathbf{Q}_i \Lambda_i \mathbf{P}_i^H, \quad i = 1, \dots, N_u \quad (26)$$

$$\Sigma_{\theta,l}^{\frac{1}{2}} \mathbf{U}_{\theta,l}^H \mathbf{F}_l \mathbf{V}_{\theta,l-1} \Lambda_{\theta,l-1}^{\frac{1}{2}} = \mathbf{U}_l \Sigma_l \mathbf{V}_l^H, \quad l = 2, \dots, L \quad (27)$$

where  $\mathbf{Q}_i$  and  $\mathbf{P}_i$  are  $M_i \times M_i$  unitary matrices,  $\mathbf{U}_l$  and  $\mathbf{V}_l$  are  $N_{l-1} \times N_{l-1}$  unitary matrices, and the main diagonal elements of  $\Lambda_i$  and  $\Sigma_l$  are sorted in a decreasing order. By using (26) and (27) we have

$$\begin{aligned} \bigotimes_{i=L}^1 (\mathbf{H}_i \mathbf{F}_i) &= \mathbf{V}_{\theta,L} \Lambda_{\theta,L}^{\frac{1}{2}} \bigotimes_{i=L}^2 \left( \bar{\mathbf{H}}_i \Sigma_{\theta,i}^{\frac{1}{2}} \mathbf{U}_{\theta,i}^H \mathbf{F}_i \mathbf{V}_{\theta,i-1} \Lambda_{\theta,i-1}^{\frac{1}{2}} \right) \\ &\quad \times \bar{\mathbf{H}}_1 \text{bd}(\Sigma_{\phi,1}^{\frac{1}{2}} \mathbf{U}_{\phi,1}^H \mathbf{B}_1, \dots, \Sigma_{\phi,N_u}^{\frac{1}{2}} \mathbf{U}_{\phi,N_u}^H \mathbf{B}_{N_u}) \\ &= \mathbf{V}_{\theta,L} \Lambda_{\theta,L}^{\frac{1}{2}} \bigotimes_{i=L}^1 \left( \bar{\mathbf{H}}_i \mathbf{U}_i \Sigma_i \mathbf{V}_i^H \right) \end{aligned} \quad (28)$$

where for the notational simplicity, we define  $\bar{\mathbf{H}}_1 \triangleq [\bar{\mathbf{G}}_1, \dots, \bar{\mathbf{G}}_{N_u}]$ ,  $\mathbf{U}_1 \triangleq \text{bd}(\mathbf{Q}_1, \dots, \mathbf{Q}_{N_u})$ ,  $\Sigma_1 \triangleq \text{bd}(\Lambda_1, \dots, \Lambda_{N_u})$ , and  $\mathbf{V}_1 \triangleq \text{bd}(\mathbf{P}_1, \dots, \mathbf{P}_{N_u})$ . From (26) and (27), we can also obtain that for  $l = 2, \dots, L$ ,

$$\begin{aligned} \bigotimes_{i=L}^l (\mathbf{H}_i \mathbf{F}_i) \bigotimes_{i=l}^L (\mathbf{F}_i^H \mathbf{H}_i^H) &= \mathbf{V}_{\theta,L} \Lambda_{\theta,L}^{\frac{1}{2}} \bigotimes_{i=L}^l \left( \bar{\mathbf{H}}_i \mathbf{U}_i \Sigma_i \mathbf{V}_i^H \right) \\ &\quad \times \Lambda_{\theta,l-1}^{-1} \bigotimes_{i=l}^L \left( \mathbf{V}_i \Sigma_i \mathbf{U}_i^H \bar{\mathbf{H}}_i^H \right) \Lambda_{\theta,L}^{\frac{1}{2}} \mathbf{V}_{\theta,L}^H. \end{aligned} \quad (29)$$

Now we consider the constraints in (11) and (12). From (26), the power consumed at each user is given by

$$\text{tr}(\mathbf{B}_i \mathbf{B}_i^H) = \text{tr} \left( \Sigma_{\phi,i}^{-\frac{1}{2}} \mathbf{Q}_i \Lambda_i^2 \mathbf{Q}_i^H \Sigma_{\phi,i}^{-\frac{1}{2}} \right), \quad i = 1, \dots, N_u. \quad (30)$$

It can be seen from [9, 9.H.1.h] that  $\text{tr}(\mathbf{B}_i \mathbf{B}_i^H) \geq \text{tr}(\Sigma_{\phi,i}^{-1} \Lambda_i^2)$ , where the equality holds if and only if  $\mathbf{Q}_i$  is an  $M_i \times M_i$

diagonal matrix with unit-norm main diagonal elements, i.e.,  $[[\mathbf{Q}_i]_{m,m}] = 1, [\mathbf{Q}_i]_{m,n} = 0, m, n = 1, \dots, M_i, m \neq n$ . For the sake of simplicity, we choose  $\mathbf{Q}_i = \mathbf{I}_{M_i}$ , and from (26), we have  $\mathbf{B}_i = \mathbf{U}_{\phi,i} \Sigma_{\phi,i}^{-\frac{1}{2}} \Lambda_i \mathbf{P}_i^H$ . Since the objective function (10) is invariant to  $\mathbf{P}_i$ , for simplicity, we also choose  $\mathbf{P}_i = \mathbf{I}_{M_i}$ . We will show later that such structure of  $\mathbf{B}_i$  is also optimal for the objective function (10).

Using (26) and (27), the averaged power consumption at the  $l$ th relay node in (11),  $l = 1, \dots, L-1$ , can be written as

$$\begin{aligned} &E_H \left[ \text{tr} \left( \mathbf{F}_{l+1} \left( \sum_{j=1}^l \left( \bigotimes_{i=l}^j (\mathbf{H}_i \mathbf{F}_i) \bigotimes_{i=j}^l (\mathbf{F}_i^H \mathbf{H}_i^H) \right) + \mathbf{I}_{N_l} \right) \mathbf{F}_{l+1}^H \right) \right] \\ &= E_H \left[ \text{tr} \left( \mathbf{F}_{l+1} \left( \sum_{j=1}^l \left( \mathbf{V}_{\theta,l} \Lambda_{\theta,l}^{\frac{1}{2}} \bigotimes_{i=l}^j (\bar{\mathbf{H}}_i \mathbf{U}_i \Sigma_i \mathbf{V}_i^H) \Pi_j \right. \right. \right. \right. \\ &\quad \left. \left. \left. \bigotimes_{i=j}^l \left( \mathbf{V}_i \Sigma_i \mathbf{U}_i^H \bar{\mathbf{H}}_i^H \right) \Lambda_{\theta,l}^{\frac{1}{2}} \mathbf{V}_{\theta,l}^H \right) + \mathbf{I}_{N_l} \right) \mathbf{F}_{l+1}^H \right) \right] \\ &= E_H \left[ \text{tr} \left( \Sigma_{\theta,l+1}^{-\frac{1}{2}} \mathbf{U}_{l+1} \Sigma_{l+1} \left( \sum_{j=1}^l \bigotimes_{i=l}^j (\bar{\mathbf{H}}_i \Sigma_i) \mathbf{V}_j^H \Pi_j \mathbf{V}_j \right. \right. \right. \\ &\quad \left. \left. \left. \bigotimes_{i=j}^l \left( \Sigma_i \bar{\mathbf{H}}_i^H \right) + \mathbf{V}_{l+1}^H \Lambda_{\theta,l}^{-1} \mathbf{V}_{l+1} \right) \Sigma_{l+1} \mathbf{U}_{l+1}^H \Sigma_{\theta,l+1}^{-\frac{1}{2}} \right) \right] \end{aligned} \quad (31)$$

where  $\tilde{\mathbf{H}}_l \triangleq \mathbf{V}_{l+1}^H \bar{\mathbf{H}}_l \mathbf{U}_l$ ,  $l = 1, \dots, L-1$  and

$$\Pi_l \triangleq \begin{cases} \mathbf{I}_{N_0}, & l = 1; \\ \Lambda_{\theta,l-1}^{-1}, & l = 2, \dots, L. \end{cases}$$

From [10], we know that for an  $N \times N$  Gaussian random matrix  $\mathbf{H}$  with i.i.d. zero mean and unit variance entries, there is  $E_H[\mathbf{H}\mathbf{A}\mathbf{H}^H] = \text{tr}(\mathbf{A})\mathbf{I}_N$ . Thus we have

$$\begin{aligned} &E_H \left[ \bigotimes_{i=l}^j (\tilde{\mathbf{H}}_i \Sigma_i) \mathbf{V}_j^H \Pi_j \mathbf{V}_j \bigotimes_{i=j}^l (\Sigma_i \tilde{\mathbf{H}}_i^H) \right] \\ &= \alpha_j \text{tr}(\Sigma_j \mathbf{V}_j^H \Pi_j \mathbf{V}_j \Sigma_j) \mathbf{I}_{N_l}, \quad j = 1, \dots, l \end{aligned} \quad (32)$$

where

$$\alpha_j \triangleq \begin{cases} 1, & j = l; \\ \prod_{i=l}^{j+1} \text{tr}(\Sigma_i^2), & j = 1, \dots, l-1, (l \geq 2). \end{cases} \quad (33)$$

Using (32), (31) can be rewritten as

$$\begin{aligned} &\text{tr} \left( \Sigma_{\theta,l+1}^{-\frac{1}{2}} \mathbf{U}_{l+1} \Sigma_{l+1}^2 \mathbf{U}_{l+1}^H \Sigma_{\theta,l+1}^{-\frac{1}{2}} \right) \sum_{j=1}^l \alpha_j \text{tr}(\Sigma_j \mathbf{V}_j^H \Pi_j \mathbf{V}_j \Sigma_j) \\ &+ \text{tr} \left( \Sigma_{\theta,l+1}^{-\frac{1}{2}} \mathbf{U}_{l+1} \Sigma_{l+1} \mathbf{V}_{l+1}^H \Lambda_{\theta,l}^{-1} \mathbf{V}_{l+1} \Sigma_{l+1} \mathbf{U}_{l+1}^H \Sigma_{\theta,l+1}^{-\frac{1}{2}} \right). \end{aligned} \quad (34)$$

Similar to (30), (34) is minimized if and only if  $\mathbf{U}_{l+1}$  is an  $N_l \times N_l$  diagonal matrix with unit-norm main diagonal elements, and  $\mathbf{V}_j$  is an  $N_{j-1} \times N_{j-1}$ ,  $j = 1, \dots, l+1$ , diagonal matrix with unit-norm main diagonal elements. For simplicity, we choose  $\mathbf{U}_{l+1} = \mathbf{I}_{N_l}$  and  $\mathbf{V}_j = \mathbf{I}_{N_{j-1}}$ ,  $j = 1, \dots, l+1$ .

From (27), we obtain  $\mathbf{F}_l = \mathbf{U}_{\theta,l} \Sigma_{\theta,l}^{-\frac{1}{2}} \Sigma_l \Lambda_{\theta,l-1}^{-\frac{1}{2}} \mathbf{V}_{\theta,l-1}^H$ .

Now we start to consider the objective function. Since  $\bar{\mathbf{H}}_L \mathbf{U}_L$  has the same distribution as  $\bar{\mathbf{H}}_L$ , and  $\mathbf{V}_{l+1}^H \bar{\mathbf{H}}_l \mathbf{U}_l$  has an identical distribution as  $\bar{\mathbf{H}}_l$ ,  $l = 1, \dots, L-1$ , by

substituting (28) and (29) back into (10), the objective function (10) can be equivalently rewritten as

$$E_H[\text{MI}] = E_H \left[ \log_2 \left| \mathbf{I}_{N_0} + \bigotimes_{i=1}^L (\boldsymbol{\Sigma}_i \bar{\mathbf{H}}_i^H) \left( \sum_{l=2}^L \bigotimes_{i=L}^l (\bar{\mathbf{H}}_i \boldsymbol{\Sigma}_i) \right. \right. \right. \\ \left. \left. \left. \times \mathbf{V}_l^H \boldsymbol{\Lambda}_{\theta, l-1}^{-1} \mathbf{V}_l \bigotimes_{i=l}^L (\boldsymbol{\Sigma}_i \bar{\mathbf{H}}_i^H) + \boldsymbol{\Lambda}_{\theta, L}^{-1} \right) \bigotimes_{i=L}^1 (\bar{\mathbf{H}}_i \boldsymbol{\Sigma}_i) \right| \right]. \quad (35)$$

It can be seen from (35) that the objective function is irrelevant to  $\mathbf{P}_i$ ,  $\mathbf{Q}_i$ ,  $i = 1, \dots, N_u$ , and  $\mathbf{U}_l$ ,  $l = 2, \dots, L$ . Thus, the optimal  $\mathbf{B}_i$  is indeed  $\mathbf{B}_i = \mathbf{U}_{\phi, i} \boldsymbol{\Delta}_{b, i}$  with  $\boldsymbol{\Delta}_{b, i} = \boldsymbol{\Sigma}_{\phi, i}^{-\frac{1}{2}} \boldsymbol{\Lambda}_i$ ,  $i = 1, \dots, N_u$ .

Let us introduce the EVD of

$$\boldsymbol{\Sigma}_l \mathbf{V}_l^H \boldsymbol{\Lambda}_{\theta, l-1}^{-1} \mathbf{V}_l \boldsymbol{\Sigma}_l = \mathbf{S}_l \boldsymbol{\Gamma}_l \mathbf{S}_l^H, \quad l = 2, \dots, L \quad (36)$$

where  $\mathbf{S}_l$  is an  $N_{l-1} \times N_{l-1}$  unitary matrix, and  $\boldsymbol{\Gamma}_l$  is the  $N_{l-1} \times N_{l-1}$  diagonal eigenvalue matrix. We can rewrite (35) as

$$E_H[\text{MI}] = E_H \left[ \log_2 \left| \mathbf{I}_{N_0} + \mathbf{M}^H \right. \right. \\ \left. \left. \times \left( \boldsymbol{\Lambda}_{\theta, L}^{-1} + \sum_{l=2}^L \mathbf{T}_l \check{\mathbf{H}}_l \boldsymbol{\Gamma}_l \check{\mathbf{H}}_l^H \mathbf{T}_l^H \right)^{-1} \mathbf{M} \right| \right] \quad (37)$$

where for notational simplicity, we define  $\mathbf{M} \triangleq \bigotimes_{i=L}^1 (\bar{\mathbf{H}}_i \boldsymbol{\Sigma}_i)$ ,  $\check{\mathbf{H}}_l \triangleq \bar{\mathbf{H}}_l \mathbf{S}_l$ ,  $l = 2, \dots, L$ ,  $\mathbf{T}_L \triangleq \mathbf{I}_{N_L}$ , and  $\mathbf{T}_l \triangleq \bigotimes_{i=L}^{l+1} (\bar{\mathbf{H}}_i \boldsymbol{\Sigma}_i)$ ,  $l = 2, \dots, L-1$ . It can be seen that  $\mathbf{M}$  does not depend on  $\mathbf{V}_l$ ,  $l = 2, \dots, L$ , thus  $\mathbf{M}$  does not affect the structure of the user and relay matrices.

Due to the contribution of the noise at all relay nodes to the destination node (which is ignored in [8]), the closed-form expression of the optimal  $\boldsymbol{\Gamma}_l$  that maximizes (37) by satisfying (36) is very difficult to obtain. On the other hand, it can be shown from [11] that  $\log_2 |\mathbf{I} + \mathbf{X}^{-1}|$  is a convex function of  $\mathbf{X}$ , and from Jensen's inequality we have  $E[\log_2 |\mathbf{I} + \mathbf{X}^{-1}|] \geq$

$\log_2 |\mathbf{I} + (E[\mathbf{X}])^{-1}|$ . Thus (37) can be lower-bounded by

$$E_H[\text{MI}] \geq \\ E_H \left[ \log_2 \left| \mathbf{I}_{N_0} + \mathbf{M}^H \left( \boldsymbol{\Lambda}_{\theta, L}^{-1} + \sum_{l=2}^L \mathbf{T}_l E_{\check{\mathbf{H}}_l} [\check{\mathbf{H}}_l \boldsymbol{\Gamma}_l \check{\mathbf{H}}_l^H] \mathbf{T}_l^H \right)^{-1} \mathbf{M} \right| \right] \\ = E_H \left[ \log_2 \left| \mathbf{I}_{N_0} + \mathbf{M}^H \left( \boldsymbol{\Lambda}_{\theta, L}^{-1} + \sum_{l=2}^L \text{tr}(\boldsymbol{\Gamma}_l) \mathbf{T}_l \mathbf{T}_l^H \right)^{-1} \mathbf{M} \right| \right] \quad (38)$$

where we applied Jensen's inequality to matrix variables  $\check{\mathbf{H}}_l \boldsymbol{\Gamma}_l \check{\mathbf{H}}_l^H$ ,  $l = 2, \dots, L$ . Obviously, to maximize (38),  $\text{tr}(\boldsymbol{\Gamma}_l)$  should be minimized. From [9, 9.H.1.h] and (36), we know that  $\text{tr}(\boldsymbol{\Gamma}_l) \geq \text{tr}(\boldsymbol{\Sigma}_l^2 \boldsymbol{\Lambda}_{\theta, l-1}^{-1})$ , where the equality holds if and only if  $\mathbf{V}_l$  is an  $N_{l-1} \times N_{l-1}$  diagonal matrix with unit-norm main diagonal elements. Therefore, we prove the optimal structure of  $\mathbf{B}_i$  and  $\mathbf{F}_l$  in (15) and (16) with  $\boldsymbol{\Delta}_{b, i} \triangleq \boldsymbol{\Sigma}_{\phi, i}^{-\frac{1}{2}} \boldsymbol{\Lambda}_i$ ,  $i = 1, \dots, N_u$ , and  $\boldsymbol{\Delta}_{f, l} \triangleq \boldsymbol{\Sigma}_{\theta, l}^{-\frac{1}{2}} \boldsymbol{\Sigma}_l \boldsymbol{\Lambda}_{\theta, l-1}^{-\frac{1}{2}}$ ,  $l = 2, \dots, L$ .

## 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 and Y. Hua, "Optimality of diagonalization of multi-hop MIMO relays," *IEEE Trans. Wireless Commun.*, vol. 8, pp. 6068-6077, Dec. 2009.
- [3] M. R. A. Khandaker and Y. Rong, "Joint source and relay optimization for multiuser MIMO relay communication systems," in *Proc. 4th Int. Conf. Signal Process. Commun. Sys.*, Dec. 2010.
- [4] C. Jeong and H.-M. Kim, "Precoder design of non-regenerative relays with covariance feedback," *IEEE Commun. Lett.*, vol. 13, pp. 920-922, Dec. 2009.
- [5] J. Wagner, B. Rankov, and A. Wittneben, "Large  $n$  analysis of amplify-and-forward MIMO relay channels with correlated Rayleigh fading," *IEEE Trans. Inf. Theory*, vol. 54, pp. 5735-5746, Dec. 2008.
- [6] C.-K. Wen, K.-K. Wong, and C. T. K. Ng, "On the asymptotic properties of amplify-and-forward MIMO relay channels," *IEEE Trans. Commun.*, vol. 59, pp. 590-602, Feb. 2011.
- [7] S. Jin, M. R. McKay, C. Zhong, and K.-K. Wong, "Ergodic capacity analysis of amplify-and-forward MIMO dual-hop systems," *IEEE Trans. Inf. Theory*, vol. 56, pp. 2204-2224, May 2010.
- [8] N. Fawaz, K. Zarifi, M. Debbah, and D. Gesbert, "Asymptotic capacity and optimal precoding in MIMO multi-hop relay networks," *IEEE Trans. Inf. Theory*, vol. 57, pp. 2050-2069, Apr. 2011.
- [9] A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and its Applications*. Academic Press, 1979.
- [10] A. Gupta and D. Nagar, *Matrix Variate Distributions*. Chapman & Hall/CRC, 2000.
- [11] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.