©2008 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

Optimality of Diagonalization of Multi-Hop MIMO Relays

Yue Rong, *Member, IEEE,* and Yingbo Hua, *Fellow, IEEE*

***Abstract*—For a two-hop linear non-regenerative multiple-input multiple-output (MIMO) relay system where the direct link between source and destination is negligible, the optimal design of the source and relay matrices has been recently established for a broad class of objective functions. The optimal source and relay matrices jointly diagonalize the MIMO relay system into a set of parallel scalar channels. In this paper, we show that this diagonalization is also optimal for a multi-hop MIMO relay system with any number of hops, which is a further generalization of several previously established results. Specifically, for Schur-concave objective functions, the optimal source precoding matrix, the optimal relay amplifying matrices and the optimal receiving matrix jointly diagonalize the multi-hop MIMO relay channel. And for Schur-convex objectives, such joint diagonalization along with a rotation of the source precoding matrix is also shown to be optimal. We also analyze the system performance when each node has the same transmission power budget and the same asymptotically large number of antennas. The asymptotic analysis shows a good agreement with numerical results under a finite number of antennas.**

***Index Terms*—MIMO relay network, multi-hop relay, linear non-regenerative relay, majorization.**

I. Introduction

It is well-known that multiple-input multiple-output (MIMO) wireless communication techniques enhance system reliability and increase system capacity. To efficiently exploit the multi-antenna hardware, an important issue in MIMO system design is to optimize the source precoding matrix [1], [2]. A general framework of optimizing the source precoding matrix has been developed in [2] by using the majorization theory [3]. It has been shown that the optimal source precoding matrix and the optimal receiving matrix diagonalize the MIMO source-destination channel for Schur-concave objective functions. And for Schur-convex objectives, the MIMO channel is also diagonalized by the optimal source matrix and the optimal receiving matrix except for a special rotation matrix at the source node.

In the case of a long source-destination distance, single or multiple MIMO relay nodes may be necessary to relay signals from the source node to the destination node [4]-[15]. In this scenario, the source signals travel through two or multiple hops before they are received by the destination node. We call such system a MIMO relay system. When the non-regenerative strategy is used, each relay node amplifies its received signal vector with a matrix (known as the relay amplifying matrix) and retransmits the amplified signal vector. Obviously, for non-regenerative MIMO relay systems, in addition to the source precoding matrix, it is crucial to optimize the relay amplifying matrices, in order to achieve an optimal system performance. Recently, it has been shown in [6] that for a three-node two-hop linear non-regenerative MIMO relay system where the direct link between source and destination is negligible, the optimal source, relay and receiving matrices jointly diagonalize the source-relay-destination channel for Schur-concave objective functions. And for Schur-convex objectives, such joint diagonalization along with a rotation of the source matrix is also shown to be optimal. The above result is a generalization of that in [2] from a one-hop MIMO link to a two-hop MIMO relay system.

In this paper, we show that the above stated results are also true for a multi-hop non-regenerative MIMO relay system with any number of hops using linear relaying and the linear minimal mean-squared error (MMSE) processing at the destination. Note that although the structures of the optimal source and relay matrices are similar for both two-hop and multi-hop systems, the proof of the main theorem is much more involved for the multi-hop system than for the two-hop system [6]. In fact, for a multi-hop system, the objective function depends on the amplifying matrices at *all* nodes. Moreover, the transmission power constraint at each node is a function of the amplifying matrices of *all backward* nodes. It will be seen that the introduction of multi-hops greatly complicates the proof of the theorem. A rigorous proof of the main theorem in this paper is technically challenging. The generalization from two-hop system to multi-hop systems is significant and it is one major contribution of this paper. In this paper, for notational convenience, we consider a narrow band single-carrier system. However, our results can be straightforwardly generalized to wide band multi-carrier multi-hop MIMO relay systems, as in the case of two-hop MIMO relay system shown in [6].

Another contribution of this paper is a performance analysis of the multi-hop MIMO relays when all nodes have the same transmission power constraint and the same asymptotically large number of antennas. Our numerical results indicate that the asymptotic analysis serves as a good approximation even for situations where each node has only a small number of antennas.

We would like to mention a few other recent works on

Manuscript received March 15, 2009; revised June 27, 2009; accepted September 7, 2009. The associate editor coordinating the review of this paper and approving it for publication was W. Choi.

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

Y. Hua is with the Department of Electrical Engineering, University of California, Riverside, CA 92521 (e-mail: yhua@ee.ucr.edu).

Y. Hua's work was in part supported by the U.S. Army Research Laboratory under the Collaborative Technology Alliance Program, and the U.S. Army Research Office under the MURI Grant W911NF-04-1-0224.

Digital Object Identifier 10.1109/TWC.2009.12.090379

1536-1276/09$25.00 © 2009 IEEE

Authorized licensed use limited to: CURTIN UNIVERSITY OF TECHNOLOGY. Downloaded on December 13, 2009 at 21:17 from IEEE Xplore. Restrictions apply.

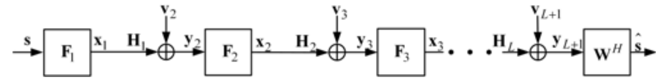


Fig. 1. Block diagram of an L-hop linear non-regenerative MIMO relay communication system.

multi-hop non-regenerative MIMO relay systems [7]-[10]. Under the assumption that the relay matrices are scaled identity matrices, the asymptotic capacity of multi-hop MIMO relay system is derived in [7]. The capacity scaling of multi-hop amplify-and-forward MIMO relay system with an asymptotically large number of hops is derived in [8]. In [9], the authors investigated the diversity gain of multi-hop MIMO relay channel when the relays use diagonal amplifying matrices. In [10], by neglecting the noise at the relay nodes, the authors derived the optimal relay matrices. Compared with those results in [7]-[10], our results in this paper are more general.

Compared with regenerative strategies (for example, decode-and-forward), the complexity of the linear non-regenerative MIMO relay system is much lower, since decoding multiple data streams involves much more computational efforts and processing latency than simply amplifying them. However, when the source-destination distance is very large, a combination of regenerative and non-regenerative relays should be used to provide a good tradeoff between the end-to-end delay and the end-to-end error rate. The more regenerative relays, the less end-to-end error rate. The more non-regenerative relays, the less end-to-end delay.

The rest of this paper is organized as follows. In Section II, we introduce the model of a multi-hop linear non-regenerative MIMO relay communication system. The structures of the optimal source and relay matrices are shown in Section III. An asymptotic performance analysis is developed in Section IV. In Section V, we show some numerical examples. Conclusions are drawn in Section VI.

II. System Model

We consider a wireless communication system with one source node, one destination node, and $L-1$ relay nodes. In this paper, we consider the scenario where $L \geq 2$. The case of $L = 1$ has been investigated in [2]. We assume that due to the propagation path-loss, the signal transmitted by the ith node can only be received by its direct forward node, i.e., the $(i+1)$-th node. Thus, signals transmitted by the source node pass through L hops until they reach the destination node. We also assume that the number of antennas at each node is N_i, $i = 1, \cdots, L+1$, and the number of source symbols in each transmission is N_b. Like [5], [6], [11]-[14], a linear non-regenerative relay matrix is used at each relay. The system block diagram is shown in Fig. 1.

The $N_1 \times 1$ signal vector transmitted by the source node is

$$\mathbf{x}_1 = \mathbf{F}_1 \mathbf{s} \tag{1}$$

where $\mathbf{s}$ is the $N_b \times 1$ source symbol vector, and $\mathbf{F}_1$ is the $N_1 \times N_b$ source precoding matrix. We assume that $\mathrm{E}[\mathbf{s}\mathbf{s}^H] = \mathbf{I}_{N_b}$, 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. The $N_i \times 1$ signal vector received at the ith node is written as

$$\mathbf{y}_i = \mathbf{H}_{i-1}\mathbf{x}_{i-1} + \mathbf{v}_i, \qquad i = 2, \cdots, L+1 \tag{2}$$

where $\mathbf{H}_{i-1}$ is the $N_i \times N_{i-1}$ MIMO channel matrix between the ith and the $(i-1)$-th nodes, i.e., the $(i-1)$-th hop, $\mathbf{v}_i$ is the $N_i \times 1$ independent and identically distributed (i.i.d.) additive white Gaussian noise (AWGN) vector at the ith node, and $\mathbf{x}_{i-1}$ is the $N_{i-1} \times 1$ signal vector transmitted by the $(i-1)$-th node. We assume that the noises are complex circularly symmetric with zero mean and unit variance.

The input-output relationship at node i is given by

$$\mathbf{x}_i = \mathbf{F}_i \mathbf{y}_i, \qquad i = 2, \cdots, L \tag{3}$$

where $\mathbf{F}_i$ is the $N_i \times N_i$ amplifying matrix at node i. Combining (1)-(3), we obtain the received signal vector at the destination node (the $(L+1)$-th node) as

$$\mathbf{y}_{L+1} = \bar{\mathbf{H}}\mathbf{s} + \bar{\mathbf{v}} \tag{4}$$

where $\bar{\mathbf{H}}$ and $\bar{\mathbf{v}}$ are the equivalent MIMO channel matrix and the noise vector, and given respectively by

$$\bar{\mathbf{H}} = \mathbf{H}_L\mathbf{F}_L \cdots \mathbf{H}_1\mathbf{F}_1 = \bigotimes_{i=L}^{1}(\mathbf{H}_i\mathbf{F}_i) \tag{5}$$

$$\begin{aligned}\bar{\mathbf{v}} &= \mathbf{H}_L\mathbf{F}_L \cdots \mathbf{H}_2\mathbf{F}_2\mathbf{v}_2 + \cdots + \mathbf{H}_L\mathbf{F}_L\mathbf{v}_L + \mathbf{v}_{L+1} \\ &= \sum_{l=2}^{L}\Big(\bigotimes_{i=L}^{l}(\mathbf{H}_i\mathbf{F}_i)\mathbf{v}_l\Big) + \mathbf{v}_{L+1}.\end{aligned} \tag{6}$$

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

We assume that without wasting transmission power at any node, the number of source symbols at each transmission satisfies $N_b \leq \min(r_1, r_2, \cdots, r_L)$, where $r_i \triangleq \mathrm{rank}(\mathbf{H}_i)$, and $\mathrm{rank}(\cdot)$ denotes the rank of a matrix. We also assume that $\mathrm{rank}(\mathbf{H}_i\mathbf{F}_i) = \mathrm{rank}(\mathbf{F}_i) \triangleq \xi_i$, $i = 1, \cdots, L$, equal N_b. The reason is that if $\xi_i < N_b$, then the system can not support N_b active symbol in each transmission. On the other hand, if $\xi_i > N_b$, some transmission power must be wasted.

From (1), we know that the power of the signal transmitted by the source node is $\mathrm{tr}(\mathbf{F}_1\mathbf{F}_1^H)$, where $\mathrm{tr}(\cdot)$ denotes the trace of a matrix. Based on (2) and (3), the power of the signal transmitted by the relay node i, $i = 2, \cdots, L$, is given by

$$\begin{aligned}&\mathrm{tr}\left(\mathrm{E}\left[\mathbf{x}_i\mathbf{x}_i^H\right]\right) \\ &= \mathrm{tr}\left(\mathbf{F}_i\mathrm{E}\left[\mathbf{y}_i\mathbf{y}_i^H\right]\mathbf{F}_i^H\right) \\ &= \mathrm{tr}\left(\mathbf{F}_i\left(\sum_{l=1}^{i-1}\Big(\bigotimes_{k=i-1}^{l}(\mathbf{H}_k\mathbf{F}_k)\bigotimes_{k=l}^{i-1}\left(\mathbf{F}_k^H\mathbf{H}_k^H\right)\Big) + \mathbf{I}_{N_i}\right)\mathbf{F}_i^H\right).\end{aligned}$$

We assume that the source node has the channel state information (CSI) knowledge of $\mathbf{H}_1$, the destination node knows $\bar{\mathbf{H}}$, and the ith node, $i = 2, \cdots, L$, knows the CSI of its backward channel $\mathbf{H}_{i-1}$ and its forward channel $\mathbf{H}_i$. In practice, the backward CSI can be obtained through standard training methods. The forward CSI required at the ith node ($\mathbf{H}_i$) is exactly the backward CSI at the $(i+1)$-th node, and thus can be obtained by a feedback from the $(i+1)$-th node. For wireless relays, the fading is often relatively slow whenever the mobility of the relays is relatively low, and for static relays, the channel state information can be almost constant. Thus, in this way, the necessary CSI can be obtained at each node with a reasonably high precision.

Authorized licensed use limited to: CURTIN UNIVERSITY OF TECHNOLOGY. Downloaded on December 13, 2009 at 21:17 from IEEE Xplore. Restrictions apply.

III. Optimal Source and Relay Matrices

It has been shown in [2], [6] that many practical objectives for MIMO systems such as the maximal mutual information (MI) between $\mathbf{s}$ and $\mathbf{y}_{L+1}$ can be represented as functions of the main diagonal elements of the MMSE matrix. The MMSE matrix is the error matrix of the linear MMSE estimates of the elements of $\mathbf{s}$ using $\mathbf{y}_{L+1}$. With a linear receiver at the destination node, the estimated signal vector is

$$\hat{\mathbf{s}} = \mathbf{W}^H \mathbf{y}_{L+1} \tag{7}$$

where $\mathbf{W}$ is the $N_{L+1} \times N_b$ weight matrix of the linear receiver. The weight matrix of the linear MMSE receiver is [2], [6]

$$\mathbf{W} = (\bar{\mathbf{H}}\bar{\mathbf{H}}^H + \mathbf{C}_{\bar{v}})^{-1}\bar{\mathbf{H}} \tag{8}$$

where $\mathbf{C}_{\bar{v}}$ is the noise covariance matrix, and $(\cdot)^{-1}$ denotes the matrix inversion. The MMSE matrix denoted as $\mathbf{E}(\{\mathbf{F}_i\})$, is given by [2], [6]

$$\mathbf{E}(\{\mathbf{F}_i\}) = \left(\mathbf{I}_{N_b} + \bar{\mathbf{H}}^H \mathbf{C}_{\bar{v}}^{-1} \bar{\mathbf{H}}\right)^{-1} \tag{9}$$

where $\{\mathbf{F}_i\} \triangleq \{\mathbf{F}_i, i = 1, \cdots, L\}$. From (6) we have

$$\mathbf{C}_{\bar{v}} = \sum_{l=2}^{L} \Big(\bigotimes_{i=L}^{l} (\mathbf{H}_i\mathbf{F}_i) \bigotimes_{i=l}^{L} (\mathbf{F}_i^H \mathbf{H}_i^H) \Big) + \mathbf{I}_{N_{L+1}}. \tag{10}$$

Substituting (5) and (10) into (9), we obtain

$$\mathbf{E}(\{\mathbf{F}_i\}) = \Bigg[\mathbf{I}_{N_b} + \bigotimes_{i=1}^{L} (\mathbf{F}_i^H \mathbf{H}_i^H) \Bigg(\sum_{l=2}^{L} \Big(\bigotimes_{i=L}^{l} (\mathbf{H}_i\mathbf{F}_i) \bigotimes_{i=l}^{L} (\mathbf{F}_i^H \mathbf{H}_i^H) \Big) + \mathbf{I}_{N_{L+1}} \Bigg)^{-1} \bigotimes_{i=L}^{1} (\mathbf{H}_i\mathbf{F}_i) \Bigg]^{-1}. \tag{11}$$

The multi-hop linear non-regenerative MIMO relay design problem can be summarized as

$$\min_{\{\mathbf{F}_i\}} \ q(\mathbf{d}[\mathbf{E}(\{\mathbf{F}_i\})]) \tag{12}$$

$$\text{s.t.} \ \operatorname{tr}\big(\mathbf{F}_1\mathbf{F}_1^H\big) \le p_1 \tag{13}$$

$$\operatorname{tr}\Bigg(\mathbf{F}_i \Bigg(\sum_{l=1}^{i-1} \Big(\bigotimes_{k=i-1}^{l} (\mathbf{H}_k\mathbf{F}_k) \bigotimes_{k=l}^{i-1} (\mathbf{F}_k^H\mathbf{H}_k^H) \Big) + \mathbf{I}_{N_i} \Bigg) \mathbf{F}_i^H \Bigg) \le p_i, \ i = 2, \cdots, L \tag{14}$$

where $q(\cdot)$ stands for a unified objective function, for a matrix $\mathbf{A}$, $\mathbf{d}[\mathbf{A}]$ is a column vector containing all main diagonal elements of $\mathbf{A}$, and $p_i > 0$, $i = 1, \cdots, L$, is the transmission power available at the ith node. Here (13) is the power constraint at the source node, and (14) are the power constraints at all relay nodes.

Before stating the key theorem on the solution of problem (12)-(14), we introduce two important definitions from [3].

Definition 1 [3, 1.A.1]: Consider any two real-valued $N \times 1$ vectors $\mathbf{x}, \mathbf{y}$, let $x_{[1]} \ge x_{[2]} \ge \cdots \ge x_{[N]}$, $y_{[1]} \ge y_{[2]} \ge \cdots \ge y_{[N]}$ denote the elements of $\mathbf{x}$ and $\mathbf{y}$ sorted in decreasing order, respectively. Then we say that vector $\mathbf{x}$ is majorized by vector $\mathbf{y}$, denoted as $\mathbf{x} \prec \mathbf{y}$, if $\sum_{i=1}^{n} x_{[i]} \le \sum_{i=1}^{n} y_{[i]}$, for $n = 1, \cdots, N-1$, and $\sum_{i=1}^{N} x_{[i]} = \sum_{i=1}^{N} y_{[i]}$.

Definition 2 [3, 3.A.1]: A real-valued function f is called Schur-convex if $f(\mathbf{x}) \le f(\mathbf{y})$ for $\mathbf{x} \prec \mathbf{y}$, or called Schur-concave if $f(\mathbf{x}) \ge f(\mathbf{y})$ for $\mathbf{x} \prec \mathbf{y}$.

Let us write the singular value decomposition (SVD) of $\mathbf{H}_i$ as

$$\mathbf{H}_i = \mathbf{U}_i \mathbf{\Sigma}_i \mathbf{V}_i^H, \qquad i = 1, \cdots, L \tag{15}$$

where the dimensions of $\mathbf{U}_i$, $\mathbf{\Sigma}_i$, $\mathbf{V}_i$ are $N_{i+1} \times N_{i+1}$, $N_{i+1} \times N_i$, $N_i \times N_i$, respectively. We assume that the main diagonal elements of $\mathbf{\Sigma}_i$, $i = 1, \cdots, L$, are arranged in the *increasing* order. The following theorem is a main result of this paper.

Theorem 1: Assume that the following three conditions hold: (1) $N_b \le \min(r_1, r_2, \cdots, r_L)$; (2) $N_b = \operatorname{rank}(\mathbf{F}_i), i = 1, \cdots, L$; (3) $q(\mathbf{d}[\mathbf{E}])$ is an increasing function with respect to each element of $\mathbf{d}[\mathbf{E}]$. Then for the linear non-regenerative multi-hop MIMO relay design problem (12)-(14), if the objective function (12) with respect to $\mathbf{d}[\mathbf{E}]$ is Schur-concave, the optimal source and relay matrices $\mathbf{F}_i$, $i = 1, \cdots, L$, are given by

$$\mathbf{F}_1 = \mathbf{V}_{1,1}\mathbf{\Lambda}_1, \quad \mathbf{F}_i = \mathbf{V}_{i,1}\mathbf{\Lambda}_i\mathbf{U}_{i-1,1}^H, \quad i = 2, \cdots, L \tag{16}$$

where $\mathbf{\Lambda}_i$, $i = 1, \cdots, L$, are $N_b \times N_b$ diagonal matrices, and $\mathbf{U}_{i,1}$ and $\mathbf{V}_{i,1}$ contain the rightmost N_b vectors of $\mathbf{U}_i$ and $\mathbf{V}_i$, respectively. And if the objective function (12) with respect to $\mathbf{d}[\mathbf{E}]$ is Schur-convex, the optimal $\mathbf{F}_i$ are

$$\mathbf{F}_1 = \mathbf{V}_{1,1}\mathbf{\Lambda}_1\mathbf{U}_0, \quad \mathbf{F}_i = \mathbf{V}_{i,1}\mathbf{\Lambda}_i\mathbf{U}_{i-1,1}^H, \quad i = 2, \cdots, L \tag{17}$$

where $\mathbf{U}_0$ is an $N_b \times N_b$ unitary rotation matrix, such that $\mathbf{d}[\mathbf{E}(\{\mathbf{F}_i\})]$ has identical elements.

Proof: See Appendix A. □

The condition 1 is motivated by the fact that under the criterion of the maximal MI between source and destination, the maximal number of independent data streams that can be sent from source to destination for any given $\{\mathbf{F}_i\}$ is no more than $\min(r_1, r_2, \cdots, r_L)$. The condition 2 is motivated by the fact that under the criterion of the maximal MI between source and destination, conditions 1 and 2 are sufficient to allow N_b independent data streams to be sent from source to destination. The condition 3 is a natural choice for any practical purpose.

Theorem 1 generalizes the results obtained in [2] and [6]. Similar to the examples shown in [2], the Schur-concave objective functions include for example the arithmetic mean of the mean-squared errors (AMSE) of the MMSE estimates of the elements of $\mathbf{s}$ using $\mathbf{y}_{L+1}$, the negative of the MI between $\mathbf{s}$ and $\mathbf{y}_{L+1}$, and the negative of the geometric mean of the signal to interference and noise ratio of $\mathbf{y}_{L+1}$. And the Schur-convex functions include for example the maximum of the mean-squared errors of the MMSE estimates of the elements of $\mathbf{s}$ using $\mathbf{y}_{L+1}$. In the next two subsections, we discuss the remaining design issues under the optimal structures of the source matrix and the relay matrices given in Theorem 1.

A. MIMO Relay Design with Schur-Concave Objective Functions

For Schur-concave objective functions, substituting (16) into (5) and (10), we have

$$\bar{\mathbf{H}} = \mathbf{U}_{L,1}\mathbf{D}_h \tag{18}$$

$$\mathbf{C}_{\bar{v}} = \mathbf{U}_{L,1}\mathbf{D}_c\mathbf{U}_{L,1}^H + \mathbf{I}_{N_{L+1}} \tag{19}$$

Authorized licensed use limited to: CURTIN UNIVERSITY OF TECHNOLOGY. Downloaded on December 13, 2009 at 21:17 from IEEE Xplore. Restrictions apply.

where $\mathbf{D}_h$ and $\mathbf{D}_c$ are $N_b \times N_b$ diagonal matrices with the kth diagonal elements $k=1,\cdots,N_b$, given by

$$[\mathbf{D}_h]_{k,k} = \prod_{l=1}^{L} \lambda_{l,k}\sigma_{l,k} \qquad [\mathbf{D}_c]_{k,k} = \sum_{l=2}^{L}\prod_{i=l}^{L} \lambda_{i,k}^2\sigma_{i,k}^2 \,.$$

Here $\lambda_{i,k}$ and $\sigma_{i,k}$, $i=1,\cdots,L, k=1,\cdots,N_b$, are the kth main diagonal elements of $\mathbf{\Lambda}_i$ and $\mathbf{\Sigma}_i$, respectively. Note that in order to achieve the optimal performance, strong subchannels of $\sigma_{i,k}$ in all hops should be paired together, while the weak subchannels of $\sigma_{i,k}$ should be coupled together [6]. Substituting (18) and (19) back into (8), we obtain

$$\mathbf{W} = \left[\mathbf{U}_{L,1}(\mathbf{D}_h^2+\mathbf{D}_c)\mathbf{U}_{L,1}^H + \mathbf{I}_{N_{L+1}}\right]^{-1}\mathbf{U}_{L,1}\mathbf{D}_h \,. \quad (20)$$

From (4), (7), (20), we can write

$$\begin{aligned}\hat{\mathbf{s}} &= \mathbf{D}_h\mathbf{U}_{L,1}^H\left[\mathbf{U}_{L,1}(\mathbf{D}_h^2+\mathbf{D}_c)\mathbf{U}_{L,1}^H+\mathbf{I}_{N_{L+1}}\right]^{-1} \\ &\quad \times \mathbf{U}_{L,1}\mathbf{D}_h\mathbf{s} + \mathbf{W}^H\bar{\mathbf{v}} \\ &\triangleq \mathbf{D}_s\mathbf{s} + \tilde{\mathbf{v}} \end{aligned} \quad (21)$$

where $\mathbf{D}_s$ is a diagonal matrix with

$$[\mathbf{D}_s]_{k,k} = \frac{\prod_{l=1}^{L}\lambda_{l,k}^2\sigma_{l,k}^2}{\sum_{l=1}^{L}\prod_{i=l}^{L}\lambda_{i,k}^2\sigma_{i,k}^2+1}, \qquad k=1,\cdots,N_b\,.$$

In (21), $\tilde{\mathbf{v}} \triangleq \mathbf{W}^H\bar{\mathbf{v}}$ is the noise vector after the receiver processing, and its covariance matrix is given by

$$\begin{aligned}\mathbf{C}_{\tilde{v}} &= \mathbf{W}^H\mathbf{C}_{\bar{v}}\mathbf{W} \\ &= \mathbf{D}_h\mathbf{U}_{L,1}^H\left[\mathbf{U}_{L,1}(\mathbf{D}_h^2+\mathbf{D}_c)\mathbf{U}_{L,1}^H+\mathbf{I}_{N_{L+1}}\right]^{-1} \\ &\quad \times\left(\mathbf{U}_{L,1}\mathbf{D}_c\mathbf{U}_{L,1}^H+\mathbf{I}_{N_{L+1}}\right) \\ &\quad \times\left[\mathbf{U}_{L,1}(\mathbf{D}_h^2+\mathbf{D}_c)\mathbf{U}_{L,1}^H+\mathbf{I}_{N_{L+1}}\right]^{-1}\mathbf{U}_{L,1}\mathbf{D}_h \\ &= \mathbf{D}_h(\mathbf{D}_h^2+\mathbf{D}_c+\mathbf{I}_{N_b})^{-1}(\mathbf{D}_c+\mathbf{I}_{N_b}) \\ &\quad \times(\mathbf{D}_h^2+\mathbf{D}_c+\mathbf{I}_{N_b})^{-1}\mathbf{D}_h \\ &\triangleq \mathbf{D}_v \end{aligned} \quad (22)$$

where the matrix inversion lemma $(\mathbf{A}+\mathbf{BCD})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{DA}^{-1}\mathbf{B}+\mathbf{C}^{-1})^{-1}\mathbf{DA}^{-1}$ is applied to obtain the third equation from the second equation, and $\mathbf{D}_v$ is a diagonal matrix with

$$[\mathbf{D}_v]_{k,k} = \frac{\prod_{l=1}^{L}\lambda_{l,k}^2\sigma_{l,k}^2\left(\sum_{l=2}^{L}\prod_{i=l}^{L}\lambda_{i,k}^2\sigma_{i,k}^2+1\right)}{\left(\sum_{l=1}^{L}\prod_{i=l}^{L}\lambda_{i,k}^2\sigma_{i,k}^2+1\right)^2}, \qquad k=1,\cdots,N_b\,.$$

From (21) and (22) we see that the optimal source, relay, and destination matrices jointly diagonalize the L-hop MIMO relay channel between $\mathbf{s}$ and $\hat{\mathbf{s}}$, and the effective noise $\tilde{\mathbf{v}}$ is white. Substituting (16) back into (11), we find that $\mathbf{E}$ is diagonal with

$$[\mathbf{E}]_{k,k} = \left(1+\frac{\prod_{l=1}^{L}\lambda_{l,k}^2\sigma_{l,k}^2}{1+\sum_{l=2}^{L}\prod_{i=l}^{L}\lambda_{i,k}^2\sigma_{i,k}^2}\right)^{-1}, \quad k=1,\cdots,N_b\,. \quad (23)$$

Using the optimal source and relay matrices (16), the transmission power constraints (13)-(14) are equivalent to

$$\sum_{k=1}^{N_b}\lambda_{1,k}^2 \le p_1 \quad (24)$$

$$\sum_{k=1}^{N_b}\lambda_{i,k}^2\left(\sum_{j=1}^{i-1}\prod_{l=j}^{i-1}\lambda_{l,k}^2\sigma_{l,k}^2+1\right) \le p_i, \quad i=2,\cdots,L\,. \quad (25)$$

To simplify notations, let us introduce the following variable substitutions for $k=1,\cdots,N_b$

$$a_{i,k} \triangleq \sigma_{i,k}^2, \quad i=1,\cdots,L \quad (26)$$

$$x_{1,k} \triangleq \lambda_{1,k}^2 \quad (27)$$

$$x_{i,k} \triangleq \lambda_{i,k}^2(a_{i-1,k}x_{i-1,k}+1), \quad i=2,\cdots,L\,. \quad (28)$$

Then we obtain

$$\prod_{l=1}^{L}\lambda_{l,k}^2\sigma_{l,k}^2 = a_{1,k}x_{1,k}\prod_{i=2}^{L}\frac{a_{i,k}x_{i,k}}{1+a_{i-1,k}x_{i-1,k}} \quad (29)$$

$$\begin{aligned}&\sum_{l=2}^{L}\prod_{i=l}^{L}\lambda_{i,k}^2\sigma_{i,k}^2 \\ &= \sum_{l=2}^{L}\prod_{i=l}^{L}\frac{a_{i,k}x_{i,k}}{1+a_{i-1,k}x_{i-1,k}} \\ &= \sum_{l=2}^{L}\prod_{i=l}^{L}\frac{a_{i,k}x_{i,k}}{1+a_{i-1,k}x_{i-1,k}} + a_{1,k}x_{1,k} \\ &\quad \times\prod_{i=2}^{L}\frac{a_{i,k}x_{i,k}}{1+a_{i-1,k}x_{i-1,k}} - a_{1,k}x_{1,k}\prod_{i=2}^{L}\frac{a_{i,k}x_{i,k}}{1+a_{i-1,k}x_{i-1,k}} \\ &= a_{L,k}x_{L,k} - a_{1,k}x_{1,k}\prod_{i=2}^{L}\frac{a_{i,k}x_{i,k}}{1+a_{i-1,k}x_{i-1,k}}. \end{aligned} \quad (30)$$

Substituting (29) and (30) back into (23) we have

$$\begin{aligned}[\mathbf{E}]_{k,k} &= \left(1+\frac{a_{1,k}x_{1,k}\prod_{i=2}^{L}\frac{a_{i,k}x_{i,k}}{1+a_{i-1,k}x_{i-1,k}}}{a_{L,k}x_{L,k}-a_{1,k}x_{1,k}\prod_{i=2}^{L}\frac{a_{i,k}x_{i,k}}{1+a_{i-1,k}x_{i-1,k}}}\right)^{-1} \\ &= 1-\prod_{i=1}^{L}\frac{a_{i,k}x_{i,k}}{1+a_{i,k}x_{i,k}}, \quad k=1,\cdots,N_b\,. \end{aligned} \quad (31)$$

Using (26)-(28), the power constraints (24), (25) can be summarized as

$$\sum_{k=1}^{N_b}x_{i,k} \le p_i, \quad x_{i,k}\ge 0, \quad i=1,\cdots,L, \quad k=1,\cdots,N_b\,. \quad (32)$$

Using (31) and (32), problem (12)-(14) is equivalently written as

$$\min_{\{x_{i,k}\}} \quad q\left(\left\{1-\prod_{i=1}^{L}\frac{a_{i,k}x_{i,k}}{1+a_{i,k}x_{i,k}}\right\}\right) \quad (33)$$

$$\text{s.t.} \quad \sum_{k=1}^{N_b}x_{i,k}\le p_i, \quad x_{i,k}\ge 0, \quad i=1,\cdots,L, \quad k=1,\cdots,N_b \quad (34)$$

where we define

Authorized licensed use limited to: CURTIN UNIVERSITY OF TECHNOLOGY. Downloaded on December 13, 2009 at 21:17 from IEEE Xplore. Restrictions apply.

$$\{x_{i,k}\} \triangleq \{x_{i,k}, 1 \leq i \leq L, 1 \leq k \leq N_b\}$$
$$\left\{1 - \prod_{i=1}^{L} \frac{a_{i,k}x_{i,k}}{1+a_{i,k}x_{i,k}}\right\} \triangleq \left\{1 - \prod_{i=1}^{L} \frac{a_{i,k}x_{i,k}}{1+a_{i,k}x_{i,k}}, 1 \leq k \leq N_b\right\}.$$

When $L = 1$, as shown in [2], the problem of (33) and (34) with respect to $\{x_{i,k}\}$ is convex for most common objective functions and adopts a water-filling type solution. However, when $L = 2$, as illustrated in [6], the problem of (33) and (34) with respect to $\{x_{i,k}\}$ is nonconvex. Obviously, the non-convexity of the problem of (33) and (34) also holds for $L > 2$. Thus for $L \geq 2$, a globally optimal solution is difficult to obtain especially when L is large. However, the problem of (33) and (34) has a conditional convexity, i.e., it is convex with respect to $\{x_{i,k}\}$ for a fixed i. Hence, a locally optimal solution of this problem can be obtained by using the alternating algorithm as shown in [6], [13], [14]. This algorithm starts at a random feasible $\{x_{i,k}\}$ and updates $\{x_{i,k}\}$ in an alternating fashion. Each time we update $x_{i,k}$, $k = 1, \cdots, N_b$, by fixing $x_{j,k}$, $j = 1, \cdots, L, j \neq i, k = 1, \cdots, N_b$. In particular, to update $x_{i,k}$, $k = 1, \cdots, N_b$, we solve the following problem

$$\min_{x_{i,1},\cdots,x_{i,N_b}} \quad q\left(\left\{1 - \frac{\beta_{i,k}a_{i,k}x_{i,k}}{1+a_{i,k}x_{i,k}}\right\}\right) \tag{35}$$
$$\text{s.t.} \quad \sum_{k=1}^{N_b} x_{i,k} \leq p_i, \quad x_{i,k} \geq 0, \quad k = 1, \cdots, N_b \tag{36}$$

where

$$\beta_{i,k} \triangleq \prod_{j=1, j\neq i}^{L} \frac{a_{j,k}x_{j,k}}{1+a_{j,k}x_{j,k}}.$$

For most common q, the problem of (35) and (36) is convex and has a water-filling type solution. Since the conditional update of $x_{i,k}$, $k = 1, \cdots, N_b$, may either decrease or maintain but cannot increase the objective function (33), monotonic convergence of $\{x_{i,k}\}$ follows directly from this observation. After the convergence of the alternating algorithm, $\lambda_{i,k}$ is obtained from (26)-(28) as

$$\lambda_{1,k} = \sqrt{x_{1,k}}, \quad \lambda_{i,k} = \sqrt{x_{i,k}/(\sigma_{i-1,k}^2 x_{i-1,k} + 1)},$$
$$k = 1, \cdots, N_b, \quad i = 2, \cdots, L.$$

Note that once a local optimum is reached, the updating process will terminate. Therefore, it is not guaranteed that the alternating algorithm will achieve the globally optimal solution.

B. MIMO Relay Design with Schur-Convex Objective Functions

For all Schur-convex objective functions, from (5) and (17) we obtain

$$\bar{\mathbf{H}} = \mathbf{U}_{L,1}\mathbf{D}_h\mathbf{U}_0 \tag{37}$$

and $\mathbf{C}_{\bar{v}}$ is given by (19). Substituting (37) and (19) into (8), we have

$$\mathbf{W} = \left[\mathbf{U}_{L,1}(\mathbf{D}_h^2 + \mathbf{D}_c)\mathbf{U}_{L,1}^H + \mathbf{I}_{N_{L+1}}\right]^{-1}\mathbf{U}_{L,1}\mathbf{D}_h\mathbf{U}_0.$$

Therefore, $\hat{\mathbf{s}}$ is given as

$$\hat{\mathbf{s}} = \mathbf{U}_0^H\mathbf{D}_s\mathbf{U}_0\mathbf{s} + \mathbf{U}_0^H\tilde{\mathbf{v}}. \tag{38}$$

From (38) we find that for Schur-convex objective functions, the equivalent channel between $\mathbf{s}$ and $\hat{\mathbf{s}}$ is diagonalized by the source, relay, and receiving matrices after a rotation $\mathbf{U}_0$ of the source matrix. Moreover, the effective noise $\mathbf{U}_0^H\tilde{\mathbf{v}}$ is no longer white, and its covariance matrix is given by $\mathbf{U}_0^H\mathbf{C}_{\tilde{v}}\mathbf{U}_0$. By substituting (17) back into (11), we obtain

$$[\mathbf{E}]_{k,k} = \frac{1}{N_b}\sum_{j=1}^{N_b}\left(1 + \frac{\prod_{l=1}^{L}\lambda_{l,j}^2\sigma_{l,j}^2}{1+\sum_{l=2}^{L}\prod_{i=l}^{L}\lambda_{i,j}^2\sigma_{i,j}^2}\right)^{-1},$$
$$k = 1, \cdots, N_b. \tag{39}$$

Substituting (29) and (30) into (39), we obtain

$$[\mathbf{E}]_{k,k} = \frac{1}{N_b}\sum_{j=1}^{N_b}\left(1 - \prod_{i=1}^{L}\frac{a_{i,j}x_{i,j}}{1+a_{i,j}x_{i,j}}\right), \quad k = 1, \cdots, N_b. \tag{40}$$

Interestingly, for all Schur-convex objectives, since the MMSE matrix $\mathbf{E}$ has identical diagonal entries, we only need to minimize $\mathrm{tr}(\mathbf{E})$, despite the specific form of the objective function. From (40) we see that the relay optimization problem is equivalent to

$$\min_{\{x_{i,k}\}} \quad \sum_{k=1}^{N_b}\left(1 - \prod_{i=1}^{L}\frac{a_{i,k}x_{i,k}}{1+a_{i,k}x_{i,k}}\right) \tag{41}$$
$$\text{s.t.} \quad \sum_{k=1}^{N_b} x_{i,k} \leq p_i, \quad x_{i,k} \geq 0,$$
$$i = 1, \cdots, L, \quad k = 1, \cdots, N_b. \tag{42}$$

Similar to Section III-A, the problem of (41) and (42) with respect to $\{x_{i,k}\}$ is conditional convex and hence can be solved by alternatingly updating $\{x_{i,k}\}$.

IV. Performance Analysis

In this section, we conduct performance analysis of multi-hop linear non-regenerative MIMO systems under some special circumstances. From (31) and (40) we find that for both Schur-concave and Schur-convex objective functions, $[\mathbf{E}]_{k,k}$ increases with L. This indicates that the system performance degrades with increasing number of hops. This is due to the linear non-regenerative strategy used at each relay node, where noises at all relay nodes are amplified and superimposed at the destination node. In particular, when $L \to \infty$, $[\mathbf{E}]_{k,k} \to 1$. In such extreme case, the source signal can not be correctly recovered at the destination node. We should note that when the source-destination distance is very large, digital repeaters should be deployed. In fact, a combination of digital repeaters and the non-regenerative relays can provide a good tradeoff between the end-to-end delay and the end-to-end error rate. The more digital repeaters, the less end-to-end error rate. The more non-regenerative relays, the less end-to-end delay.

The capacity of a two-hop amplify-and-forward MIMO relay system with a large number of antennas is analyzed in [15]. For a multi-hop amplify-and-forward MIMO relay system, the system capacity scaling is derived in [8] for $L \to \infty$. In the following, we assume a finite L, and study the system performance when each node has the same power budget and the same asymptotically large number of antennas, i.e., $N_i = N$, $i = 1, \cdots, L+1$. When $N \to \infty$, the distribution

Authorized licensed use limited to: CURTIN UNIVERSITY OF TECHNOLOGY. Downloaded on December 13, 2009 at 21:17 from IEEE Xplore. Restrictions apply.

of the square of the singular values of $\mathbf{H}_i, i = 1, \cdots, L$, (i.e., the eigenvalues of $\mathbf{H}_i\mathbf{H}_i^H$) denoted as λ^2, does not depend on i. In fact, λ^2 follows the quarter-circle law [16] with the following probability density function

$$f_{\lambda^2}(a) = \frac{1}{2\pi}\sqrt{\frac{4-a}{a}}, \qquad 0 < a \le 4. \tag{43}$$

Since a is independent of i, the power allocated to each data stream, denoted as x, is also independent of i. Therefore, for Schur-concave objective functions, by using (31), the diagonal elements of the MMSE matrix is given by

$$[\mathbf{E}]_a = 1 - \left(1 + a^{-1}x^{-1}\right)^{-L}. \tag{44}$$

While for Schur-convex objectives, we have

$$[\mathbf{E}]_a = \int_0^4 \left[1 - \left(1 + a^{-1}x^{-1}\right)^{-L}\right] f_{\lambda^2}(a)da. \tag{45}$$

For each Schur-concave objective function, using (43) and (44), we can write the specific optimization problem. For example, choosing the AMSE of the signal waveform estimation as the criterion, we have

$$\text{AMSE} = N\int_0^4 \left[1 - \left(1 + a^{-1}x^{-1}\right)^{-L}\right] f_{\lambda^2}(a)da.$$

Thus, the AMSE optimization problem can be formulated as

$$\min_x \quad \int_0^4 \left[1 - \left(1 + a^{-1}x^{-1}\right)^{-L}\right] f_{\lambda^2}(a)da \tag{46}$$

$$\text{s.t.} \quad N\int_0^4 x f_{\lambda^2}(a)da \le P, \qquad x \ge 0 \tag{47}$$

where (47) is the (same) power constraint at the source node and each relay node, and P denotes the transmission power budget. Another commonly applied criterion in MIMO relay design is the MI between the source and received signals, which is

$$\text{MI} = -N\int_0^4 \log_2\!\left[1 - \left(1 + a^{-1}x^{-1}\right)^{-L}\right] f_{\lambda^2}(a)da.$$

The corresponding optimization problem is given by

$$\min_x \quad \int_0^4 \log_2\!\left[1 - \left(1 + a^{-1}x^{-1}\right)^{-L}\right] f_{\lambda^2}(a)da \tag{48}$$

$$\text{s.t.} \quad N\int_0^4 x f_{\lambda^2}(a)da \le P, \qquad x \ge 0. \tag{49}$$

Both the problem of (46) and (47) and the problem of (48) and (49) are convex and have water-filling type solutions. The solutions can be obtained by the Lagrangian multiplier method [17].

For all Schur-convex objective functions, based on (41) and (45), the optimization problem is identical to the optimal AMSE problem given by (46) and (47).

V. Numerical Examples

To verify the validity of our asymptotic performance analysis, we carry out numerical simulations. In the simulations, all nodes are equipped with N antennas, and the number of source symbols is $N_b = N$. The MIMO channel matrices $\mathbf{H}_i$, $i = 1, \cdots, L$, have i.i.d. Gaussian entries with zero mean and normalized variance $1/N$. We assume that all hops have equal distance, and the propagation path-loss is included in the channel matrices. The source node and all relay nodes have the same transmission power P. All simulation results are averaged over 100 channel realizations.

In the first example, we consider a two-hop MIMO relay system. Fig. 2 shows the normalized per-antenna MSE (AMSE divided by N) versus P for different N. Here, AMSE is selected as the objective function. Thus, the asymptotic results are computed by solving the problem of (46) and (47). While the simulation results are obtained by alternatingly solving the problem of (35) and (36) for $i = 1, \cdots, L$, where q is the AMSE function. From Fig. 2 we see that the asymptotic results agree with the simulation results. We also observe from Fig. 2 that the normalized per-antenna MSE increases with increasing number of antennas.

Fig. 3 displays the normalized per-antenna MI (MI divided by N) versus P, when the MI between the source and destination signal is chosen as the design objective. We compute the asymptotic results by solving the problem of (48) and (49). The simulation results are obtained by solving the problem of (35) and (36) with q being the negative MI function. Fig. 3 verifies the validity of our analysis. We find that the normalized per-antenna MI decreases with the number of antennas at each node. From Figs. 2 and 3 we also observe that although the analysis is conducted under the assumption of very large N, the results are valid even for very small number of antennas such as $N = 5$.

In the second example, we fix the number of antennas at each node to be $N = 10$ and study the relationship of system performance with respect to the number of hops. Fig. 4 shows the normalized per-antenna MSE versus P for different L. While Fig. 5 displays the normalized per-antenna MI versus P. Similar to Figs. 2 and 3, we see that the asymptotic results agree with the simulation results. From Figs. 4 and 5 we find that the system performance in terms of normalized per-antenna MSE and normalized per-antenna MI degrades with increasing number of hops.

From Figs. 2 and 4, we see that the normalized per-antenna MSE obtained by the simulation is slightly higher than that by the analysis for $L > 1$. We also observe from Figs. 3 and 5 that for $L > 1$ the normalized per-antenna MI obtained by the simulation is slighter lower than that of the analysis. The reason is that the alternating algorithm used in the simulation only finds a locally optimal solution.

VI. Conclusions

In this paper, the previous results on the optimal source and relay matrices for two-hop linear non-regenerative MIMO relay systems have been generalized to multi-hop MIMO relay systems. The structures of the optimal source and relay matrices have been shown to diagonalize a multi-hop MIMO relay system into a set of parallel scalar multi-hop relay channels. Performance analysis has been conducted when each node has the same power budget and the same asymptotically large number of antennas. The validity of the asymptotic results are verified by numerical simulations. Our results can be straightforwardly generalized to multi-carrier multi-hop linear non-regenerative MIMO relay systems.

Authorized licensed use limited to: CURTIN UNIVERSITY OF TECHNOLOGY. Downloaded on December 13, 2009 at 21:17 from IEEE Xplore. Restrictions apply.

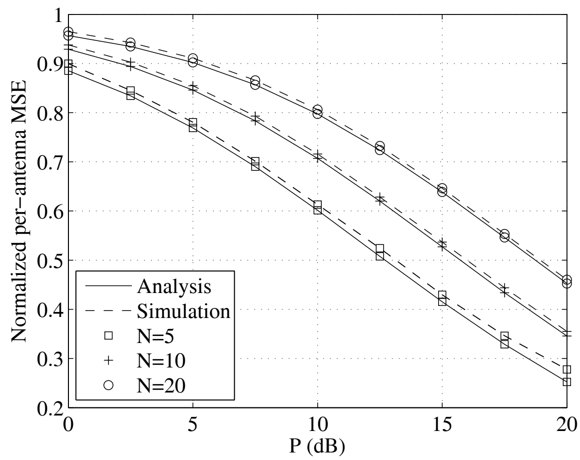


Fig. 2. Example 1: Normalized per-antenna MSE versus P; $L = 2$.

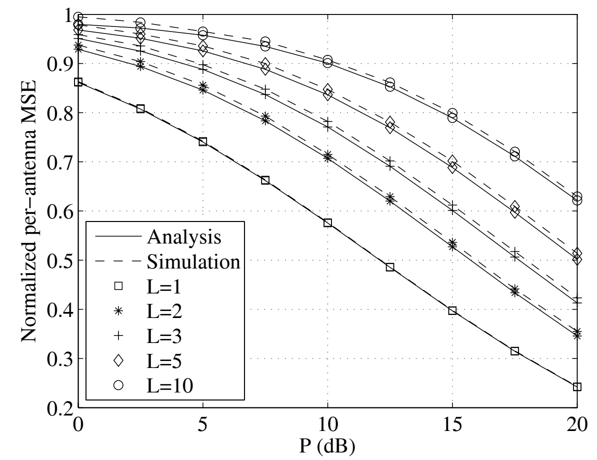


Fig. 4. Example 2: Normalized per-antenna MSE versus P; $N = 10$.

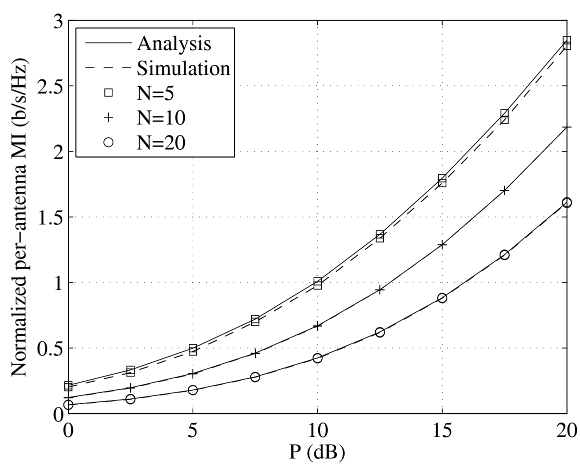


Fig. 3. Example 1: Normalized per-antenna MI versus P; $L = 2$.

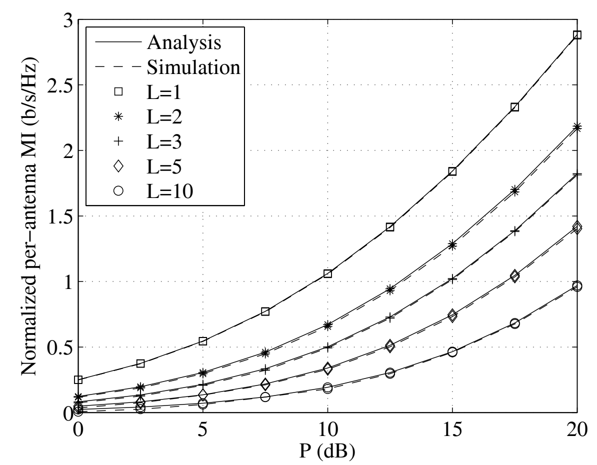


Fig. 5. Example 2: Normalized per-antenna MI versus P; $N = 10$.

APPENDIX A
PROOF OF THEOREM 1

To prove Theorem 1, we need the following definitions and lemmas from [3].

DEFINITION 3 [3, 1.A.2]: Consider any two real-valued $N \times 1$ vectors $\mathbf{x}, \mathbf{y}$, let $x_{[1]} \geq x_{[2]} \geq \cdots \geq x_{[N]}$, $y_{[1]} \geq y_{[2]} \geq \cdots \geq y_{[N]}$ denote the elements of $\mathbf{x}$ and $\mathbf{y}$ sorted in decreasing order, respectively. Then we say that $\mathbf{x}$ is weakly submajorized by $\mathbf{y}$, denoted as $\mathbf{x} \prec_{\mathrm{w}} \mathbf{y}$, if $\sum_{i=1}^{n} x_{[i]} \leq \sum_{i=1}^{n} y_{[i]}$, for $n = 1, \cdots, N$.

DEFINITION 4 [3, 2.A.1]: An $N \times N$ matrix $\mathbf{S}$ is doubly stochastic if $[\mathbf{S}]_{i,j} \geq 0, i, j = 1, \cdots, N$, $\sum_{i=1}^{N} [\mathbf{S}]_{i,j} = 1, j = 1, \cdots, N$, and $\sum_{j=1}^{N} [\mathbf{S}]_{i,j} = 1, i = 1, \cdots, N$.

LEMMA 1 [3, 2.B.2]: A necessary and sufficient condition that $\mathbf{x} \prec \mathbf{y}$ is that there exists a doubly stochastic matrix $\mathbf{S}$ such that $\mathbf{x} = \mathbf{S}\mathbf{y}$.

LEMMA 2 [3, 9.B.1]: For a Hermitian matrix $\mathbf{A}$ with the vector of its main diagonal elements $\mathbf{d}[\mathbf{A}]$ and the vector of its eigenvalues $\boldsymbol{\lambda}[\mathbf{A}]$, it follows that $\mathbf{d}[\mathbf{A}] \prec \boldsymbol{\lambda}[\mathbf{A}]$.

LEMMA 3 [3, 9.H.2]: For m $N \times N$ complex matrices $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_m$, let $\mathbf{B} = \bigotimes_{i=1}^{m} \mathbf{A}_i$, then $\boldsymbol{\sigma}_b \prec_{\mathrm{w}} (\boldsymbol{\sigma}_{a_1} \odot \boldsymbol{\sigma}_{a_2} \odot \cdots \odot \boldsymbol{\sigma}_{a_m})$, where $\boldsymbol{\sigma}_b$, and $\boldsymbol{\sigma}_{a_i}$, $i = 1, \cdots, m$, denote $N \times 1$ vectors containing the singular values of $\mathbf{B}$ and $\mathbf{A}_i$ arranged in the same order, respectively, and $\odot$ denotes the Schur (element-wise) product of two vectors.

LEMMA 4 [3, 3.A.8]: A real-valued function f satisfies $\mathbf{x} \prec_{\mathbf{w}} \mathbf{y} \Rightarrow f(\mathbf{x}) \leq f(\mathbf{y})$ if and only if f is increasing with respect to each variable and Schur-convex.

LEMMA 5 [3, 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 6 [3, p.7]: For an $N \times 1$ real-valued vector $\mathbf{x}$, let us define an $N \times 1$ vector $\underline{\mathbf{x}}$ with identical elements of $\sum_{i=1}^{N} x_i / N$, there is $\underline{\mathbf{x}} \prec \mathbf{x}$.

LEMMA 7 [3, 9.B.2]: For any $N \times 1$ real-valued vector $\mathbf{x}$, there exists a real symmetric (and thus Hermitian) matrix $\mathbf{A}$ with equal main diagonal elements and eigenvalues given by $\mathbf{x}$. Equivalently, there is a unitary matrix $\mathbf{U}_0$ such that $\mathbf{A} = \mathbf{U}_0^H \mathcal{D}(\mathbf{x}) \mathbf{U}_0$. Here $\mathcal{D}(\mathbf{x})$ denotes a diagonal matrix taking $\mathbf{x}$ as the main diagonal.

The following two lemmas are also required to prove Theorem 1.

LEMMA 8: If $q(\mathbf{x})$ is Schur-concave with respect to $\mathbf{x}$, and

Authorized licensed use limited to: CURTIN UNIVERSITY OF TECHNOLOGY. Downloaded on December 13, 2009 at 21:17 from IEEE Xplore. Restrictions apply.

$\mathbf{y} = \mathbf{1} - \mathbf{x}$, where $\mathbf{1}$ is a vector of all ones, then $q(\mathbf{1} - \mathbf{y})$ is also Schur-concave with respect to $\mathbf{y}$.

PROOF: Assume that $\mathbf{x} \prec \mathbf{x}_1$. Based on Lemma 1, we have

$$\mathbf{x} \prec \mathbf{x}_1 \Leftrightarrow \mathbf{x} = \mathbf{S}\mathbf{x}_1. \tag{50}$$

Since $\mathbf{S}$ is doubly stochastic, it follows from Definition 4 that $\mathbf{S1} = \mathbf{1}$, and consequently,

$$\mathbf{y} = \mathbf{1} - \mathbf{x} = \mathbf{S1} - \mathbf{S}\mathbf{x}_1 = \mathbf{S}(\mathbf{1} - \mathbf{x}_1) = \mathbf{S}\mathbf{y}_1 \Leftrightarrow \mathbf{y} \prec \mathbf{y}_1. \tag{51}$$

Since $q(\mathbf{x})$ is Schur-concave with respect to $\mathbf{x}$, it follows from Definition 2 that $\mathbf{x} \prec \mathbf{x}_1 \Rightarrow q(\mathbf{x}) \geq q(\mathbf{x}_1)$. Moreover, from (50) and (51) we have $\mathbf{x} \prec \mathbf{x}_1 \Leftrightarrow \mathbf{y} \prec \mathbf{y}_1$, hence $\mathbf{y} \prec \mathbf{y}_1 \Rightarrow q(\mathbf{1} - \mathbf{y}) \geq q(\mathbf{1} - \mathbf{y}_1)$. □

LEMMA 9: For any $n \times n$ positive definite $\mathbf{A}$, the solution to the problem

$$\min_{\mathbf{F}} \quad \mathrm{tr}(\mathbf{F}\mathbf{A}\mathbf{F}^H) \qquad \text{s.t.} \quad \mathbf{B}\mathbf{F} = \mathbf{C} \tag{52}$$

is $\mathbf{F} = \mathbf{B}^H(\mathbf{B}\mathbf{B}^H)^{-1}\mathbf{C}$, where the dimensions of $\mathbf{F}$, $\mathbf{B}$, and $\mathbf{C}$ are $m \times n$, $p \times m$, and $p \times n$, respectively, and $\mathrm{rank}(\mathbf{B}) = p < m$.

PROOF: The complete solution to the linear constraint in (52) is

$$\mathbf{F} = \mathbf{B}^H(\mathbf{B}\mathbf{B}^H)^{-1}\mathbf{C} + \left(\mathbf{I}_m - \mathbf{B}^H(\mathbf{B}\mathbf{B}^H)^{-1}\mathbf{B}\right)\mathbf{X} \tag{53}$$

where $\mathbf{X}$ is an arbitrary $m \times n$ matrix. Substituting (53) back into the objective function of (52), and using the fact that $\mathbf{B}(\mathbf{I}_m - \mathbf{B}^H(\mathbf{B}\mathbf{B}^H)^{-1}\mathbf{B}) = \mathbf{0}_{p\times m}$, where $\mathbf{0}_{p\times m}$ denotes a $p \times m$ matrix with all zero entries, we have

$$\begin{aligned}&\mathrm{tr}(\mathbf{F}\mathbf{A}\mathbf{F}^H)\\ &= \mathrm{tr}\left(\mathbf{B}^H(\mathbf{B}\mathbf{B}^H)^{-1}\mathbf{C}\mathbf{A}\mathbf{C}^H(\mathbf{B}\mathbf{B}^H)^{-1}\mathbf{B}\right) + \mathrm{tr}\Big(\left(\mathbf{I}_m - \mathbf{B}^H\right.\\ &\quad \left.\times(\mathbf{B}\mathbf{B}^H)^{-1}\mathbf{B}\right)\mathbf{X}\mathbf{A}\mathbf{X}^H\left(\mathbf{I}_m - \mathbf{B}^H(\mathbf{B}\mathbf{B}^H)^{-1}\mathbf{B}\right)\Big). \end{aligned} \tag{54}$$

Since $\mathbf{A}$ is positive definite, (54) is minimized if and only if $\mathbf{X} = \mathbf{0}_{m\times n}$. Thus, we obtain $\mathbf{F} = \mathbf{B}^H(\mathbf{B}\mathbf{B}^H)^{-1}\mathbf{C}$. □

We start to prove Theorem 1 for Schur-concave objective functions. Let us define

$$\mathbf{A}_1 = \mathbf{H}_1\mathbf{F}_1\mathbf{F}_1^H\mathbf{H}_1^H \tag{55}$$
$$\mathbf{A}_i = \mathbf{H}_i\mathbf{F}_i(\mathbf{A}_{i-1} + \mathbf{I}_{N_i})\mathbf{F}_i^H\mathbf{H}_i^H, \quad i = 2, \cdots, L \tag{56}$$

and write $\mathbf{A}_i = \mathbf{U}_{A_i}\mathbf{\Lambda}_{A_i}\mathbf{U}_{A_i}^H, i = 1, \cdots, L$, as the eigendecomposition of $\mathbf{A}_i$, where $\mathbf{\Lambda}_{A_i}$ is an $N_b \times N_b$ diagonal matrix containing all nonzero eigenvalues of $\mathbf{A}_i$ sorted in the *increasing* order for all i, and $\mathbf{U}_{A_i}$ is the associated $N_{i+1} \times N_b$ matrix of eigenvectors. From (55) and (56), we have

$$\mathbf{H}_1\mathbf{F}_1 = \mathbf{U}_{A_1}\mathbf{\Lambda}_{A_1}^{\frac{1}{2}}\mathbf{Q}_1 \tag{57}$$
$$\mathbf{H}_i\mathbf{F}_i(\mathbf{A}_{i-1} + \mathbf{I}_{N_i})^{\frac{1}{2}} = \mathbf{U}_{A_i}\mathbf{\Lambda}_{A_i}^{\frac{1}{2}}\mathbf{Q}_i, \quad i = 2, \cdots, L \tag{58}$$

where $\mathbf{Q}_1$ is an $N_b \times N_b$ unitary matrix, $\mathbf{Q}_i$ $i = 2, \cdots, L$, are $N_b \times N_i$ semi-unitary matrices with $\mathbf{Q}_i\mathbf{Q}_i^H = \mathbf{I}_{N_b}$. It will be seen that the power constraints (13) and (14) are invariant to $\mathbf{Q}_i$, $i = 1, \cdots, L$. We obtain from (58) that

$$\mathbf{H}_i\mathbf{F}_i = \mathbf{U}_{A_i}\mathbf{\Lambda}_{A_i}^{\frac{1}{2}}\mathbf{Q}_i(\mathbf{A}_{i-1} + \mathbf{I}_{N_i})^{-\frac{1}{2}}, \quad i = 2, \cdots, L. \tag{59}$$

Applying the matrix inversion lemma to (11), the MMSE matrix $\mathbf{E}$ can be written as

$$\begin{aligned}\mathbf{E} = \mathbf{I}_{N_b} - \bigotimes_{i=1}^{L}(\mathbf{F}_i^H\mathbf{H}_i^H)\Bigg(\sum_{l=1}^{L}\Big(\bigotimes_{i=L}^{l}(\mathbf{H}_i\mathbf{F}_i)\\ \bigotimes_{i=l}^{L}(\mathbf{F}_i^H\mathbf{H}_i^H)\Big) + \mathbf{I}_{N_{L+1}}\Bigg)^{-1}\bigotimes_{i=L}^{1}(\mathbf{H}_i\mathbf{F}_i). \end{aligned}\tag{60}$$

Substituting (56), (57) and (59) to (60), we have

$$\begin{aligned}\mathbf{E} &= \mathbf{I}_{N_b} - \mathbf{Q}_1^H\mathbf{\Lambda}_{A_1}^{\frac{1}{2}}\mathbf{U}_{A_1}^H\bigotimes_{i=2}^{L}\Big((\mathbf{A}_{i-1} + \mathbf{I}_{N_i})^{-\frac{1}{2}}\\ &\quad\times\mathbf{Q}_i^H\mathbf{\Lambda}_{A_i}^{\frac{1}{2}}\mathbf{U}_{A_i}^H\Big)(\mathbf{A}_L + \mathbf{I}_{N_{L+1}})^{-1}\\ &\quad\bigotimes_{i=L}^{2}\Big(\mathbf{U}_{A_i}\mathbf{\Lambda}_{A_i}^{\frac{1}{2}}\mathbf{Q}_i(\mathbf{A}_{i-1} + \mathbf{I}_{N_i})^{-\frac{1}{2}}\Big)\mathbf{U}_{A_1}\mathbf{\Lambda}_{A_1}^{\frac{1}{2}}\mathbf{Q}_1\\ &= \mathbf{I}_{N_b} - \mathbf{Q}_1^H\mathbf{\Lambda}_{A_1}^{\frac{1}{2}}\mathbf{U}_{A_1}^H\bigotimes_{i=2}^{L}\Big(\mathbf{U}_{A_{i-1}}(\mathbf{\Lambda}_{A_{i-1}} + \mathbf{I}_{N_b})^{-\frac{1}{2}}\mathbf{U}_{A_{i-1}}^H\\ &\quad\mathbf{Q}_i^H\mathbf{\Lambda}_{A_i}^{\frac{1}{2}}\mathbf{U}_{A_i}^H\Big)\mathbf{U}_{A_L}(\mathbf{\Lambda}_{A_L} + \mathbf{I}_{N_b})^{-1}\mathbf{U}_{A_L}^H\bigotimes_{i=L}^{2}\Big(\mathbf{U}_{A_i}\mathbf{\Lambda}_{A_i}^{\frac{1}{2}}\\ &\quad\times\mathbf{Q}_i\mathbf{U}_{A_{i-1}}(\mathbf{\Lambda}_{A_{i-1}} + \mathbf{I}_{N_b})^{-\frac{1}{2}}\mathbf{U}_{A_{i-1}}^H\Big)\mathbf{U}_{A_1}\mathbf{\Lambda}_{A_1}^{\frac{1}{2}}\mathbf{Q}_1 \qquad (61)\\ &\triangleq \mathbf{I}_{N_b} - \mathbf{G}. \end{aligned}$$

Applying Lemmas 2 and 3 to $\mathbf{G}$, we obtain

$$\mathbf{d}[\mathbf{G}] \prec \boldsymbol{\lambda}[\mathbf{G}] \prec_{\mathrm{w}} \mathbf{d}[\tilde{\mathbf{G}}] \tag{62}$$

where $\tilde{\mathbf{G}}$ is a diagonal matrix given by

$$\begin{aligned}\tilde{\mathbf{G}} &\triangleq \mathbf{\Lambda}_{A_1}^{\frac{1}{2}}\bigotimes_{i=2}^{L}\Big((\mathbf{\Lambda}_{A_{i-1}} + \mathbf{I}_{N_b})^{-\frac{1}{2}}\mathbf{\Lambda}_{A_i}^{\frac{1}{2}}\Big)(\mathbf{\Lambda}_{A_L} + \mathbf{I}_{N_b})^{-1}\\ &\quad\bigotimes_{i=L}^{2}\Big(\mathbf{\Lambda}_{A_i}^{\frac{1}{2}}(\mathbf{\Lambda}_{A_{i-1}} + \mathbf{I}_{N_b})^{-\frac{1}{2}}\Big)\mathbf{\Lambda}_{A_1}^{\frac{1}{2}}\\ &= \bigotimes_{i=1}^{L}(\mathbf{\Lambda}_{A_i}(\mathbf{\Lambda}_{A_i} + \mathbf{I}_{N_b})^{-1}). \end{aligned}$$

Since $q(\mathbf{d}[\mathbf{E}])$ is Schur-concave and increasing with respect to $\mathbf{d}[\mathbf{E}]$, from Lemma 8 we find that $q(\mathbf{d}[\mathbf{I}_{N_b} - \mathbf{G}])$ is Schur-concave and decreasing with respect to $\mathbf{d}[\mathbf{G}]$. Obviously, $-q(\mathbf{d}[\mathbf{I}_{N_b} - \mathbf{G}])$ is Schur-convex and increasing with respect to $\mathbf{d}[\mathbf{G}]$. Based on Lemma 4 and (62), we obtain $-q(\mathbf{d}[\mathbf{I}_{N_b} - \mathbf{G}]) \leq -q(\mathbf{d}[\mathbf{I}_{N_b} - \tilde{\mathbf{G}}])$, hence $q(\mathbf{d}[\mathbf{I}_{N_b} - \mathbf{G}]) \geq q(\mathbf{d}[\mathbf{I}_{N_b} - \tilde{\mathbf{G}}])$. From (61), we see that the minimum is obtained at

$$\mathbf{Q}_1 = \mathbf{\Phi}, \qquad \mathbf{Q}_i = \mathbf{\Phi}\mathbf{U}_{A_{i-1}}^H, \qquad i = 2, \cdots, L$$

where $\mathbf{\Phi}$ stands for an arbitrary $N_b \times N_b$ diagonal matrix with unit-norm main diagonal elements, i.e., $|[\mathbf{\Phi}]_{i,i}| = 1, [\mathbf{\Phi}]_{i,j} = 0, i, j = 1, \cdots, N_b, i \neq j$. Without affecting $\min q(\mathbf{d}[\mathbf{E}])$, we choose $\mathbf{Q}_1 = \mathbf{I}_{N_b}$, and $\mathbf{Q}_i = \mathbf{U}_{A_{i-1}}^H$, $i = 2, \cdots, L$.

Now we set out to consider the power constraints. First, we introduce some notations: for $i = 1, \cdots, L$, $\hat{\mathbf{F}}_i \triangleq \mathbf{V}_i^H\mathbf{F}_i$, $\mathbf{U}_i \triangleq [\mathbf{U}_{i,\bar{r}_i}, \mathbf{U}_{i,r_i}]$, where $\mathbf{U}_{i,\bar{r}_i}$ and $\mathbf{U}_{i,r_i}$ contain the left singular vectors of $\mathbf{H}_i$ associated with the zero and nonzero singular values of $\mathbf{H}_i$, respectively, $\mathbf{\Sigma}_{i,r_i}$ is a diagonal matrix

Authorized licensed use limited to: CURTIN UNIVERSITY OF TECHNOLOGY. Downloaded on December 13, 2009 at 21:17 from IEEE Xplore. Restrictions apply.

containing the nonzero singular values of $\mathbf{H}_i$, $\mathbf{\Sigma}_{i,1}$ contains the largest N_b singular values of $\mathbf{H}_i$ sorted in the same order as the diagonal elements of $\mathbf{\Lambda}_{A_i}$. Substituting the SVD of $\mathbf{H}_1$ in (15) into (57) and left multiplying by $\mathbf{U}_1^H$ on both sides, we have

$$\begin{bmatrix} \mathbf{0}_{(N_2-r_1)\times(N_1-r_1)} & \mathbf{0}_{(N_2-r_1)\times r_1} \\ \mathbf{0}_{r_1\times(N_1-r_1)} & \mathbf{\Sigma}_{1,r_1} \end{bmatrix} \hat{\mathbf{F}}_1 = \mathbf{U}_1^H \mathbf{U}_{A_1} \mathbf{\Lambda}_{A_1}^{\frac{1}{2}} \mathbf{Q}_1 \,. \tag{63}$$

If $N_1 = N_2 = r_1$, (63) holds if and only if

$$\hat{\mathbf{F}}_1 = \mathbf{\Sigma}_{1,r_1}^{-1} \mathbf{U}_{1,r_1}^H \mathbf{U}_{A_1} \mathbf{\Lambda}_{A_1}^{\frac{1}{2}} \mathbf{Q}_1 \,. \tag{64}$$

If $N_1 > N_2 = r_1$, then (63) holds if and only if

$$\left[\ \mathbf{0}_{r_1\times(N_1-r_1)} \quad \mathbf{\Sigma}_{1,r_1}\ \right] \hat{\mathbf{F}}_1 = \mathbf{U}_{1,r_1}^H \mathbf{U}_{A_1} \mathbf{\Lambda}_{A_1}^{\frac{1}{2}} \mathbf{Q}_1 \,. \tag{65}$$

Finally, if $N_1 > r_1, N_2 > r_1$, (63) is true if and only if $\mathbf{U}_{1,\bar{r}_1}^H \mathbf{U}_{A_1} = \mathbf{0}_{(N_2-r_1)\times N_b}$ and (65) holds. From (65), we see that in the latter two cases, there are many solutions for $\hat{\mathbf{F}}_1$. We should choose $\hat{\mathbf{F}}_1$ such that the transmission power at the source node is minimized. Since $\mathrm{tr}(\mathbf{F}_1\mathbf{F}_1^H) = \mathrm{tr}(\hat{\mathbf{F}}_1\hat{\mathbf{F}}_1^H)$, the transmission power minimization problem is written as

$$\min_{\hat{\mathbf{F}}_1} \ \mathrm{tr}\left(\hat{\mathbf{F}}_1\hat{\mathbf{F}}_1^H\right) \tag{66}$$

$$\text{s.t.} \ \left[\ \mathbf{0}_{r_1\times(N_1-r_1)} \quad \mathbf{\Sigma}_{1,r_1}\ \right] \hat{\mathbf{F}}_1 = \mathbf{U}_{1,r_1}^H \mathbf{U}_{A_1} \mathbf{\Lambda}_{A_1}^{\frac{1}{2}} \mathbf{Q}_1 . \tag{67}$$

From Lemma 9, the solution to the problem of (66) and (67) is given by

$$\hat{\mathbf{F}}_1 = \left[\ \mathbf{0}_{r_1\times(N_1-r_1)} \quad \mathbf{\Sigma}_{1,r_1}^{-1}\ \right]^T \mathbf{U}_{1,r_1}^H \mathbf{U}_{A_1} \mathbf{\Lambda}_{A_1}^{\frac{1}{2}} \mathbf{Q}_1 \,. \tag{68}$$

To determine $\mathbf{U}_{A_1}$ in (64) and (68), we substitute (64) and (68) into the objective function of (66). Interestingly, both (64) and (68) lead to the same transmission power, given by

$$\mathrm{tr}\left(\mathbf{F}_1\mathbf{F}_1^H\right) = \mathrm{tr}\left(\mathbf{\Lambda}_{A_1}^{\frac{1}{2}} \mathbf{U}_{A_1}^H \mathbf{U}_{1,r_1} \mathbf{\Sigma}_{1,r_1}^{-2} \mathbf{U}_{1,r_1}^H \mathbf{U}_{A_1} \mathbf{\Lambda}_{A_1}^{\frac{1}{2}}\right) . \tag{69}$$

We note that the transmission power (69) is invariant to $\mathbf{Q}_i$, $i = 1, \cdots, L$. Using Lemma 5, we know that under $\mathrm{rank}(\mathbf{F}_1) = N_b$, (69) is minimized if and only if $\mathbf{U}_{A_1}^H \mathbf{U}_{1,r_1} = [\mathbf{0}_{N_b\times(r_1-N_b)}, \mathbf{\Phi}]$. The minimum of (69) is $\mathrm{tr}(\mathbf{\Lambda}_{A_1}\mathbf{\Sigma}_{1,1}^{-2})$. Without loss of generality, we choose $\mathbf{\Phi} = \mathbf{I}_{N_b}$. Therefore, we have $\mathbf{U}_{A_1} = \mathbf{U}_{1,1}$, and together with $\mathbf{Q}_1 = \mathbf{I}_{N_b}$, we obtain

$$\mathbf{F}_1 = \mathbf{V}_1 \left[\mathbf{0}_{N_b\times(N_1-N_b)}, \mathbf{\Sigma}_{1,1}^{-1}\mathbf{\Lambda}_{A_1}^{\frac{1}{2}}\right]^T = \mathbf{V}_{1,1}\mathbf{\Sigma}_{1,1}^{-1}\mathbf{\Lambda}_{A_1}^{\frac{1}{2}} \,.$$

We have now proved that the optimal structure of $\mathbf{F}_1$ is as in (16) with $\mathbf{\Lambda}_1 = \mathbf{\Sigma}_{1,1}^{-1}\mathbf{\Lambda}_{A_1}^{\frac{1}{2}}$.

Now we consider the power constraints (14). Similar to steps (63)-(68), for $i = 2, \cdots, L$, we have $\mathbf{U}_{i,\bar{r}_i}^H \mathbf{U}_{A_i} = \mathbf{0}_{(N_{i+1}-r_i)\times N_b}$ when $N_{i+1} > r_i$ and

$$\hat{\mathbf{F}}_i = \mathbf{\Sigma}_{i,r_i}^{-1} \mathbf{U}_{i,r_i}^H \mathbf{U}_{A_i} \mathbf{\Lambda}_{A_i}^{\frac{1}{2}} \mathbf{Q}_i (\mathbf{A}_{i-1} + \mathbf{I}_{N_i})^{-\frac{1}{2}}, \qquad r_i = N_i$$

$$\hat{\mathbf{F}}_i = \left[\ \mathbf{0}_{r_i\times(N_i-r_i)} \quad \mathbf{\Sigma}_{i,r_i}^{-1}\ \right]^T \mathbf{U}_{i,r_i}^H \mathbf{U}_{A_i} \mathbf{\Lambda}_{A_i}^{\frac{1}{2}} \times \mathbf{Q}_i (\mathbf{A}_{i-1} + \mathbf{I}_{N_i})^{-\frac{1}{2}}, \qquad r_i < N_i$$

where we solved the following problem using Lemma 9

$$\min_{\hat{\mathbf{F}}_i} \ \mathrm{tr}\left(\hat{\mathbf{F}}_i \left(\sum_{l=1}^{i-1} \left(\bigotimes_{n=i-1}^{l} (\mathbf{H}_n\mathbf{F}_n) \bigotimes_{n=l}^{i-1} \left(\mathbf{F}_n^H\mathbf{H}_n^H\right)\right) + \mathbf{I}_{N_i}\right) \hat{\mathbf{F}}_i^H\right)$$

$$\text{s.t.} \ \left[\ \mathbf{0}_{r_i\times(N_i-r_i)} \quad \mathbf{\Sigma}_{i,r_i}\ \right] \hat{\mathbf{F}}_i = \mathbf{U}_{i,r_i}^H \mathbf{U}_{A_i} \mathbf{\Lambda}_{A_i}^{\frac{1}{2}} \mathbf{Q}_i (\mathbf{A}_{i-1} + \mathbf{I}_{N_i})^{-\frac{1}{2}} \,.$$

The transmission power at the ith node is

$$\mathrm{tr}\left(\mathbf{F}_i(\mathbf{A}_{i-1} + \mathbf{I}_{N_i})\mathbf{F}_i^H\right) = \mathrm{tr}\left(\mathbf{\Sigma}_{i,r_i}^{-1} \mathbf{U}_{i,r_i}^H \mathbf{U}_{A_i} \mathbf{\Lambda}_{A_i} \mathbf{U}_{A_i}^H \mathbf{U}_{i,r_i} \mathbf{\Sigma}_{i,r_i}^{-1}\right) . \tag{70}$$

Obviously, (70) is also invariant to $\mathbf{Q}_i$, $i = 1, \cdots, L$. Similar to (69), (70) is minimized by $\mathbf{U}_{A_i} = \mathbf{U}_{i,1}$, and together with $\mathbf{Q}_i = \mathbf{U}_{A_{i-1}}^H$, we obtain

$$\mathbf{F}_i = \mathbf{V}_{i,1}\mathbf{\Sigma}_{i,1}^{-1}\mathbf{\Lambda}_{A_i}^{\frac{1}{2}}(\mathbf{\Lambda}_{A_{i-1}} + \mathbf{I}_{N_b})^{-\frac{1}{2}}\mathbf{U}_{i-1,1}^H \,.$$

Thus, the optimal structure of $\mathbf{F}_i$ is given by (16) with $\mathbf{\Lambda}_i = \mathbf{\Sigma}_{i,1}^{-1}\mathbf{\Lambda}_{A_i}^{\frac{1}{2}}(\mathbf{\Lambda}_{A_{i-1}} + \mathbf{I}_{N_b})^{-\frac{1}{2}}$. Therefore, we have now proved the optimal structures of $\mathbf{F}_i$, $i = 1, \cdots, L$, for Schur-concave objective functions.

The proof for the case of Schur-convex objective functions is given as follows. Based on Definition 2 and Lemma 6, the objective function (12) is minimized when $\mathbf{E}(\{\mathbf{F}_i\})$ has identical diagonal elements. Let us introduce the following eigendecomposition

$$\bigotimes_{i=1}^{L}(\mathbf{F}_i^H\mathbf{H}_i^H)\left(\sum_{l=2}^{L}\left(\bigotimes_{i=L}^{l}(\mathbf{H}_i\mathbf{F}_i)\bigotimes_{i=l}^{L}(\mathbf{F}_i^H\mathbf{H}_i^H)\right) + \mathbf{I}_{N_{L+1}}\right)^{-1} \bigotimes_{i=L}^{1}(\mathbf{H}_i\mathbf{F}_i) = \mathbf{U}_E\tilde{\mathbf{\Lambda}}_E\mathbf{U}_E^H$$

where the dimensions of $\mathbf{U}_E$ and $\tilde{\mathbf{\Lambda}}_E$ are both $N_b \times N_b$. From (11), we have $\mathbf{E}(\{\mathbf{F}_i\}) = \mathbf{U}_E\mathbf{\Lambda}_E\mathbf{U}_E^H$, where $\mathbf{\Lambda}_E = (\mathbf{I}_{N_b} + \tilde{\mathbf{\Lambda}}_E)^{-1}$. Based on Lemma 7, we know that there is a unitary $\mathbf{U}_0$ such that $\mathbf{U}_0^H\mathbf{\Lambda}_E\mathbf{U}_0$ has identical main diagonal elements. Therefore, for any given $\{\mathbf{F}_i\}$, we can use $\tilde{\mathbf{F}}_1 = \mathbf{F}_1\mathbf{U}_F$, where $\mathbf{U}_F = \mathbf{U}_E\mathbf{U}_0$ to have an $\mathbf{E}$ with identical main diagonal elements, and hence improve the performance. Since $\mathbf{U}_F$ is unitary, rotating $\mathbf{F}_1$ by $\mathbf{U}_F$ does not affect $\mathrm{tr}(\mathbf{E}(\{\mathbf{F}_i\}))$ and the power constraints. Using such $\tilde{\mathbf{F}}_1$, we have

$$\left[\mathbf{E}\left(\tilde{\mathbf{F}}_1, \mathbf{F}_2, \cdots, \mathbf{F}_L\right)\right]_{k,k} = \frac{1}{N_b}\mathrm{tr}(\mathbf{E}(\mathbf{F}_1, \mathbf{F}_2, \cdots, \mathbf{F}_L)) \,, \quad k = 1, \cdots, N_b \,. \tag{71}$$

Matrix $\mathbf{U}_0$ can be any rotation matrix that satisfies $|[\mathbf{U}_0]_{i,k}| = |[\mathbf{U}_0]_{i,l}|, \forall i, k, l$. When the dimensions are appropriate such as a power of two, the discrete Fourier transform matrix can be chosen for $\mathbf{U}_0$. While for general case, $\mathbf{U}_0$ can be computed using the method developed in [18].

From (71), we see that $\{\mathbf{F}_i\}$ should be chosen to minimize $\mathrm{tr}(\mathbf{E}(\{\mathbf{F}_i\}))$. Since $\mathrm{tr}(\mathbf{E}(\{\mathbf{F}_i\}))$ is a Schur-concave[1]

[1] In fact, $\mathrm{tr}(\mathbf{d}[\mathbf{A}])$ is both Schur-convex and Schur-concave with respect to $\mathbf{d}[\mathbf{A}]$.

Authorized licensed use limited to: CURTIN UNIVERSITY OF TECHNOLOGY. Downloaded on December 13, 2009 at 21:17 from IEEE Xplore. Restrictions apply.

and increasing function of $\mathbf{d}[\mathbf{E}(\{\mathbf{F}_i\})]$, the previous results for Schur-concave objective functions can be applied here. Therefore, there are two steps in the optimal relay design with Schur-convex objective functions. First, we compute the optimal $\{\mathbf{F}_i\}$ according to (16) using $\text{tr}(\mathbf{E})$ as the objective function. After the first step, we obtain a diagonal $\mathbf{E}$ (i.e., $\mathbf{U}_E = \mathbf{I}_{N_b}$) with minimal $\text{tr}(\mathbf{E})$. In the second step, $\mathbf{F}_1$ is rotated by $\mathbf{U}_F = \mathbf{U}_0$ such that the new $\mathbf{E}$ has identical main diagonal elements. Therefore, for Schur-convex objective functions, the optimal $\{\mathbf{F}_i\}$ are given by (17). □

REFERENCES

[1] A. Scaglione, P. Stoica, S. Barbarossa, G. B. Giannakis, and H. Sampath, "Optimal designs for space-time linear precoders and decoders," *IEEE Trans. Signal Process.*, vol. 50, pp. 1051-1064, May 2002.

[2] 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.

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

[4] B. Wang, J. Zhang, and A. Høst-Madsen, "On the capacity of MIMO relay channels," *IEEE Trans. Inf. Theory*, vol. 51, pp. 29-43, Jan. 2005.

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

[6] Y. Rong, X. Tang, and Y. Hua, "A unified framework for optimizing linear non-regenerative multicarrier MIMO relay communication systems," *IEEE Trans. Signal Process.*, to appear.

[7] S. Yeh and O. Lévêque, "Asymptotic capacity of multi-level amplify-and-forward relay networks," in *Proc. IEEE ISIT*, Nice, France, June 2007, pp. 1436-1440.

[8] J. Wagner and A. Wittneben, "On capacity scaling of (long) MIMO amplify-and-forward multi-hop networks," in *Proc. IEEE Asilomar Conf. Signals, Systems Computers*, Pacific Grove, CA, USA, Oct. 2008, pp. 346-350.

[9] S. Yang and J.-C. Belfiore, "Diversity of MIMO multihop relay channels," submitted to *IEEE Trans. Inf. Theory*. [Online]. Avaliable: http://arxiv.org/abs/0708.0386v1.

[10] N. Fawaz, K. Zarifi, M. Debbah, and D. Gesbert, "Asymptotic capacity and optimal precoding strategy of multi-level precode & forward in correlated channels," in *Proc. IEEE ITW*, Porto, Portugal, May 2008, pp. 209-213.

[11] O. Muñoz-Medina, J. Vidal, and A. Agustín, "Linear transceiver design in nonregenerative relays with channel state information," *IEEE Trans. Signal Process.*, vol. 55, pp. 2593-2604, June 2007.

[12] A. S. Behbahani, R. Merched, and A. M. Eltawil, "Optimizations of a MIMO relay network," *IEEE Trans. Signal Process.*, vol. 56, pp. 5062-5073, Oct. 2008.

[13] Z. Fang, Y. Hua, and J. C. Koshy, "Joint source and relay optimization for a non-regenerative MIMO relay," in *Proc. IEEE Workshop Sensor Array Multi-Channel Signal Processing*, Waltham, WA, July 2006, pp. 239-243.

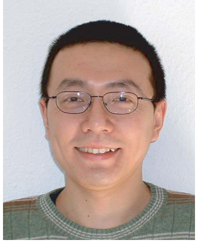
[14] I. Hammerström and A. Wittneben, "Power allocation schemes for amplify-and-forward MIMO-OFDM relay links," *IEEE Trans. Wireless Commun.*, vol. 6, pp. 2798-2802, Aug. 2007.

[15] 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.

[16] A. M. Tulino and S. Verdú, *Random Matrix Theory and Wireless Communications*. Now Publishers, 2004.

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

[18] P. Viswanath and V. Anantharam, "Optimal sequences and sum capacity of synchronous CDMA systems," *IEEE Trans. Inf. Theory*, vol. 45, pp. 1984-1993, Sept. 1999.



Yue Rong (S'03-M'06) 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, as a 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 2001-2002 Graduate Sponsoring Asia scholarship of DAAD/ABB (Germany) and the 2004 Chinese Government Award for Outstanding Self-Financed Students Abroad (China).

Yingbo Hua (S'86-M'88-SM'92-F'02) received B.S. degree from Nanjing Institute of Technology, Nanjing, China, in Feb. 1982, and M.S. and Ph.D. degrees from Syracuse University, Syracuse, NY, in 1983 and 1988, respectively. From 1988 to 1989, he was a research fellow at Syracuse, consulting for Syracuse Research Co., NY, and Aeritalia Co., Italy. He was Lecturer from Feb. 1990 to 1992, Senior Lecturer from 1993 to 1995, and Reader and Associate Professor from 1996 to 2001, with the University of Melbourne, Australia. He served as a visiting professor with Hong Kong University of Science and Technology from 1999 to 2000, and consulted for Microsoft Research Co., WA, summer 2000. Since Feb. 2001, he has been Professor of Electrical Engineering with the University of California, Riverside, CA. He is an author/coauthor of numerous articles in journals, conference proceedings and books, which span the fields of sensor array signal processing, channel and system identification, wireless communications and networking, and distributed computations in sensor networks. He is a co-editor of *Signal Processing Advances in Wireless and Mobile Communications* (Prentice-Hall, 2001), and *High-Resolution and Robust Signal Processing* (Marcel Dekker, 2003). He has served on the Editorial Boards for IEEE TRANSACTIONS ON SIGNAL PROCESSING, IEEE SIGNAL PROCESSING LETTERS, IEEE SIGNAL PROCESSING MAGAZINE, and SIGNAL PROCESSING (EURASIP). He also served on IEEE SPS Technical Committees for Underwater Acoustic Signal Processing, Sensor Array and Multi-channel Signal Processing, and Signal Processing for Communications and Networking, and on numerous international conference organization committees. He is a Fellow of IEEE.

Authorized licensed use limited to: CURTIN UNIVERSITY OF TECHNOLOGY. Downloaded on December 13, 2009 at 21:17 from IEEE Xplore. Restrictions apply.