JOINT SOURCE AND RELAY OPTIMIZATION FOR A NON-REGENERATIVE MIMO RELAY

Zheng Fang, Yingbo Hua

Department of Electrical Engineering
University of California
Riverside, CA 92521
{zhfang,yhua}@ee.ucr.edu

John C. Koshy

Advanced Wireless Signal Processing Research
Telcordia Technologies, Inc. NVC 2X-439
331 Newman Springs Road, Red Bank, NJ
koshy@research.telcordia.com

ABSTRACT

We consider a non-regenerative MIMO relay system where the source, relay and destination are all equipped with multiple antennas. The relay does not decode the packets but performs a multi-dimensional amplify-and-forward function (a relay matrix) on the baseband signals. Under the condition that the source is white, the relay matrix that maximizes the capacity between the source and the destination has been previously found. In this paper, we show a new result on how the source covariance matrix and the relay matrix can be jointly optimized to maximize the source-destination capacity. It is shown that the optimal coordinate system governed by the previously discovered relay matrix is still valid under the joint optimization, and the joint optimization yields a further capacity gain when the SNR at the relay is low.

1. INTRODUCTION

Wireless relays are important for wireless ad hoc communication networks. Deploying a relay between a source and a destination can reduce the (required) transmitted power from the source, and hence reduce the interference to other neighboring nodes. A relay may also be necessary when there is strong shadowing between the source and the destination. Relays can be regenerative or non-regenerative. The former performs decoding and then re-encoding while the latter only performs an amplify-and-forward function on the baseband symbols. Because of the above difference, a non-regenerative relay generally causes a much smaller delay than a regenerative relay. A non-regenerative relay is also more flexible as it does not need to know the code used at the source and the destination. Also note that "regenerative relay and non-regenerative relay" can be treated as two functionalities that can be embedded in a single physical node, and which to use can be adaptively governed by a higher layer networking protocol.

Design of MIMO relays is important for a network of nodes equipped with multiple antennas. A regenerative MIMO relay system is studied in [1] where the relay is assumed to be able to simultaneously transmit and receive at a single frequency. This type of full duplex relay is difficult to implement in practice. A non-regenerative MIMO relay system is previously studied in [2] where the relay transmits and receives in two orthogonal channels (i.e., half duplex). But in [2], the source covariance matrix is assumed to be proportional to the identity matrix, and only the relay matrix is optimized. In this paper, we show that the source covariance matrix and the relay matrix can be jointly optimized to maximize the source-destination capacity. The direct link between the source and the destination is assumed to be weak and will not be considered in this paper. But the effect of the direct link and a comparison to the result shown in [1] will be given in a forthcoming paper. With the direct link, finding the optimal relay matrix is still an open problem.

This work was supported in part by the National Science Foundation under Grant No. ECS-0219377, the U. S. Army Research Laboratory under the CTA Program Cooperative Agreement DAAD19-01-2-0011, and the U. S. Army Research Office under the MURI Grant No. W911NF-04-1-0224. The U. S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation thereon.

In Section 2, the non-regenerative MIMO relay system is formulated. In Section 3, we show that the coordinate system governed by the relay matrix shown in [2] remains optimal under the joint optimization of source and relay. In Section 4, we show how the power at the source and the relay can be optimally distributed along the optimal coordinates. Numerical results are given in Section 5. Section 6 concludes the paper.

2. PROBLEM FORMULATION

A non-regenerative MIMO relay system is depicted in Fig. 1 where $\boldsymbol{s}$ denotes the source vector, $\boldsymbol{H}_1$ the channel matrix between the source and the relay, $\boldsymbol{H}_2$ the channel matrix between the relay and the destination, $\boldsymbol{F}$ the relay matrix, $\boldsymbol{y}$ the signal vector received at the destination, $\boldsymbol{n}_1$ and $\boldsymbol{n}_2$ are the noise at the relay and the destination, respectively. The input-channel and output-channel of the relay are assumed to be orthogonal to each other in time and/or frequency (although frequency division appears the most desirable in this context).

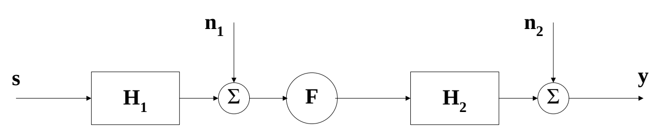

Fig. 1. A non-regenerative MIMO relay system

It follows that the signal vector received at the destination from the relay can be written as

$$\boldsymbol{y} = \boldsymbol{H}_2\boldsymbol{F}\boldsymbol{H}_1\boldsymbol{s} + \boldsymbol{H}_2\boldsymbol{F}\boldsymbol{n}_1 + \boldsymbol{n}_2 \tag{1}$$

where the (weak) signal arriving at the destination from the source is not included. The noise vectors $\boldsymbol{n}_1$ and $\boldsymbol{n}_2$ are assumed to be zero-mean white Gaussian with the covariance σ_1^2 and σ_2^2, respectively. The numbers of antennas equipped at the source, relay and destination are denoted by M, L and N, respectively, and hence $\boldsymbol{H}_1 \in C^{L\times M}$, $\boldsymbol{H}_2 \in C^{N\times L}$, and $\boldsymbol{F} \in C^{L\times L}$.

If $\boldsymbol{n}_1$ and $\boldsymbol{n}_2$ are colored with known covariance matrices, i.e., $E(\boldsymbol{n}_1\boldsymbol{n}_1^H) = \boldsymbol{R}_1$, and $E(\boldsymbol{n}_2\boldsymbol{n}_2^H) = \boldsymbol{R}_2$, then the above system (1) is equivalent to

$$\bar{\boldsymbol{y}} = \bar{\boldsymbol{H}}_2\bar{\boldsymbol{F}}\bar{\boldsymbol{H}}_1\bar{\boldsymbol{s}} + \bar{\boldsymbol{H}}_2\bar{\boldsymbol{F}}\bar{\boldsymbol{n}}_1 + \bar{\boldsymbol{n}}_2 \tag{2}$$

where $\bar{\boldsymbol{y}} = \boldsymbol{R}_2^{-1/2}\boldsymbol{y}$, $\bar{\boldsymbol{H}}_2 = \boldsymbol{R}_2^{-1/2}\boldsymbol{H}_2$, $\bar{\boldsymbol{F}} = \boldsymbol{F}\boldsymbol{R}_1^{1/2}$, $\bar{\boldsymbol{H}}_1 = \boldsymbol{R}_1^{-1/2}\boldsymbol{H}_1$, $\bar{\boldsymbol{s}} = \boldsymbol{s}$, $\bar{\boldsymbol{n}}_1 = \boldsymbol{R}_1^{-1/2}\boldsymbol{n}_1$, and $\bar{\boldsymbol{n}}_2 = \boldsymbol{R}_2^{-1/2}\boldsymbol{n}_2$. Here, the noise vectors are whitened as $E(\bar{\boldsymbol{n}}_1\bar{\boldsymbol{n}}_1^H) = \boldsymbol{I}$ and $E(\bar{\boldsymbol{n}}_2\bar{\boldsymbol{n}}_2^H) = \boldsymbol{I}$. Therefore, without loss of generality, we will assume that the noise vectors $\boldsymbol{n}_1$ and $\boldsymbol{n}_2$ are white.

The source-destination capacity of the MIMO relay system (1) is the maximal mutual information between $\boldsymbol{s}$ and $\boldsymbol{y}$, which is known to be as follows [3] (a factor 1/2 penalty due to the orthogonal channels is ignored here as it does not affect the optimization problem):

$$C(\boldsymbol{F},\boldsymbol{Q}) = \log_2 \det[\boldsymbol{I} + \boldsymbol{H}_2\boldsymbol{F}\boldsymbol{H}_1\boldsymbol{Q}\boldsymbol{H}_1^H\boldsymbol{F}^H\boldsymbol{H}_2^H \times (\sigma_1^2\boldsymbol{H}_2\boldsymbol{F}\boldsymbol{F}^H\boldsymbol{H}_2^H + \sigma_2^2\boldsymbol{I})^{-1}] \tag{3}$$

where $\boldsymbol{Q} = E(\boldsymbol{s}\boldsymbol{s}^H)$ is the source covariance matrix. We assume that $\boldsymbol{H}_1$ and $\boldsymbol{H}_2$ are known to all nodes. In [2], it is assumed that $\boldsymbol{Q} = \frac{P_1}{M}\boldsymbol{I}$. In this paper, we consider a more general problem as follows:

$$\max_{F,Q} \quad C(\boldsymbol{F},\boldsymbol{Q}) \tag{4}$$

$$\text{s.t.} \quad \text{tr}(\boldsymbol{Q}) \le P_1$$

$$\text{tr}(\boldsymbol{F}\boldsymbol{H}_1\boldsymbol{Q}\boldsymbol{H}_1^H\boldsymbol{F}^H + \sigma_1^2\boldsymbol{F}\boldsymbol{F}^H) \le P_2 \tag{5}$$

3. OPTIMAL COORDINATES OF JOINT SOURCE AND RELAY DESIGN

Let the singular value decomposition (SVD) of $\boldsymbol{H}_1$ and $\boldsymbol{H}_2$ be

$$\boldsymbol{H}_1 = \boldsymbol{U}_1\boldsymbol{\Sigma}_1\boldsymbol{V}_1^H$$
$$\boldsymbol{H}_2 = \boldsymbol{U}_2\boldsymbol{\Sigma}_2\boldsymbol{V}_2^H$$

Here, $\boldsymbol{U}_1$, $\boldsymbol{U}_2$, $\boldsymbol{V}_1$, and $\boldsymbol{V}_2$ are unitary matrices of singular vectors, and $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$ are diagonal matrices of singular values of $\boldsymbol{H}_1$ and $\boldsymbol{H}_2$, respectively, in descending order. We also define $\boldsymbol{\Lambda}_1 = \boldsymbol{\Sigma}_1^2$ and $\boldsymbol{\Lambda}_2 = \boldsymbol{\Sigma}_2^2$. Our main theorem is:

Theorem 1 *The capacity $C(\boldsymbol{F},\boldsymbol{Q})$ can achieve its maximum when the source covariance matrix $\boldsymbol{Q}$ and the relay matrix $\boldsymbol{F}$ are constructed as follows:*

$$\boldsymbol{Q} = \boldsymbol{V}_1\boldsymbol{\Lambda}_Q\boldsymbol{V}_1^H$$
$$\boldsymbol{F} = \boldsymbol{V}_2\boldsymbol{\Sigma}_F\boldsymbol{U}_1^H$$

where $\boldsymbol{\Sigma}_F$ and $\boldsymbol{\Lambda}_Q$ are diagonal matrices.

Remark: The structure of the optimal relay matrix is the same as that shown in [2] under the constraint $\boldsymbol{Q} = \frac{P_1}{M}\boldsymbol{I}$. The above result turns the original MIMO channel into a set of parallel SISO channels. With the coordinates of $\boldsymbol{Q}$ and $\boldsymbol{F}$ given as in the theorem, the optimization of $\boldsymbol{Q}$ and $\boldsymbol{F}$ now comes down to the optimization of $\boldsymbol{\Lambda}_Q$ and $\boldsymbol{\Sigma}_F$. The power distribution at the source and the relay (along the optimal coordinates) is governed by the diagonal entries of $\boldsymbol{\Lambda}_Q$ and $\boldsymbol{\Sigma}_F$, respectively.

Proof: Recall that given two $N \times N$ positive semi-definite Hermitian matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ with eigenvalues $\lambda_k(\boldsymbol{A})$ and $\lambda_k(\boldsymbol{B})$ arranged in the descending order respectively, we have

$$\sum_{k=1}^{N}\lambda_k(\boldsymbol{A})\lambda_{N+1-k}(\boldsymbol{B}) \le \text{tr}(\boldsymbol{A}\boldsymbol{B}) \le \sum_{k=1}^{N}\lambda_k(\boldsymbol{A})\lambda_k(\boldsymbol{B}) \tag{6}$$

Also recall from [2] that when the source signal is white, the optimal weighting matrix at the relay can be chosen as $\boldsymbol{F} = \boldsymbol{V}_2\boldsymbol{\Sigma}_F\boldsymbol{U}_1^H$. If we define an equivalent channel $\tilde{\boldsymbol{H}}_1 = \boldsymbol{H}_1\boldsymbol{Q}^{1/2}$, the capacity formulation (3) is equivalent to the relay-only design formulation in [2]. Hence, by following [2], for any given ("source-relay") pair $\tilde{\boldsymbol{Q}}$ and $\tilde{\boldsymbol{F}}$, there always exists another pair $\tilde{\boldsymbol{Q}}$ and $\tilde{\boldsymbol{F}}_o$ that achieves better or equal capacity with the same power constraints, and $\tilde{\boldsymbol{F}}_o$ can be represented as

$$\tilde{\boldsymbol{F}}_o = \boldsymbol{V}_2\boldsymbol{\Sigma}_F\tilde{\boldsymbol{U}}_1^H \tag{7}$$

where $\tilde{\boldsymbol{U}}_1$ is dependent on $\tilde{\boldsymbol{Q}}$, i.e., $\boldsymbol{H}_1\tilde{\boldsymbol{Q}}\boldsymbol{H}_1^H = \tilde{\boldsymbol{U}}_1\tilde{\boldsymbol{\Lambda}}_1\tilde{\boldsymbol{U}}_1^H$. Using (7) in (3), we have a partially optimized source-destination capacity:

$$C(\tilde{\boldsymbol{F}}^{(1)},\tilde{\boldsymbol{Q}}) = \log_2 \det[\boldsymbol{I} + \boldsymbol{\Lambda}_2\boldsymbol{\Lambda}_F\tilde{\boldsymbol{\Lambda}}_1(\sigma_1^2\boldsymbol{\Lambda}_2\boldsymbol{\Lambda}_F + \sigma_2^2\boldsymbol{I})^{-1}]$$

The power constraint on the relay becomes $\text{tr}(\boldsymbol{\Lambda}_F\tilde{\boldsymbol{\Lambda}}_1 + \sigma_1^2\boldsymbol{\Lambda}_F) \le P_2$. The power constraint on the source is still $\text{tr}(\boldsymbol{Q}) \le P_1$.

Note that the capacity and the power constraint are only dependent on $\tilde{\boldsymbol{\Lambda}}_1$ but not on $\tilde{\boldsymbol{U}}_1$. It follows that for any matrix $\boldsymbol{Q}$ satisfying

$$\boldsymbol{H}_1\boldsymbol{Q}\boldsymbol{H}_1^H = \hat{\boldsymbol{U}}_1\tilde{\boldsymbol{\Lambda}}_1\hat{\boldsymbol{U}}_1^H \tag{8}$$

where $\hat{\boldsymbol{U}}_1$ is any orthogonal matrix, the optimal capacity is the same as that for $\tilde{\boldsymbol{Q}}$. Therefore, (7) can be replaced by

$$\boldsymbol{F}_o = \boldsymbol{V}_2\boldsymbol{\Sigma}_F\hat{\boldsymbol{U}}_1^H \tag{9}$$

We need now to determine the optimal structure for $\boldsymbol{Q}$. Denote $r = \text{rank}(\boldsymbol{H}_1) \le \min(M, L)$. Then the SVD of $\boldsymbol{H}_1$ with rank r can be written as

$$\boldsymbol{H}_1 = \boldsymbol{U}_1\boldsymbol{\Sigma}_1\boldsymbol{V}_1^H = \begin{bmatrix} \boldsymbol{U}_{1,1} & \boldsymbol{U}_{1,2} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Sigma}_{1,1} & \\ & \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{V}_{1,1} & \boldsymbol{V}_{1,2} \end{bmatrix}^H$$

where $\boldsymbol{\Sigma}_{1,1}$ has the dimension $r \times r$. It follows that

$$\boldsymbol{H}_1\boldsymbol{Q}\boldsymbol{H}_1^H = \hat{\boldsymbol{U}}_1\tilde{\boldsymbol{\Lambda}}_1\hat{\boldsymbol{U}}_1^H = \begin{bmatrix} \hat{\boldsymbol{U}}_{1,1} & \hat{\boldsymbol{U}}_{1,2} \end{bmatrix} \begin{bmatrix} \tilde{\boldsymbol{\Lambda}}_{1,1} & \\ & \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{V}}_{1,1} & \hat{\boldsymbol{V}}_{1,2} \end{bmatrix}^H$$

where $\tilde{\mathbf{\Lambda}}_{1,1}$ has the dimension $r \times r$. It is easy to verify that $\boldsymbol{U}_{1,1} \perp \hat{\boldsymbol{U}}_{1,2}$ and $\boldsymbol{U}_{1,2} \perp \hat{\boldsymbol{U}}_{1,1}$. Hence, from (8), we have

$$
\begin{aligned}
& \boldsymbol{H}_1^{+} \boldsymbol{H}_1 \boldsymbol{Q} \boldsymbol{H}_1^H \boldsymbol{H}_1^{H+} \\
= & \boldsymbol{H}_1^{+} \hat{\boldsymbol{U}}_1 \tilde{\mathbf{\Lambda}}_1 \hat{\boldsymbol{U}}_1^H \boldsymbol{H}_1^{H+} \\
= & \begin{bmatrix} \boldsymbol{V}_{1,1} & \boldsymbol{V}_{1,2} \end{bmatrix} \begin{bmatrix} \mathbf{\Sigma}_{1,1}^{-1} & \\ & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{U}_{1,1} & \boldsymbol{U}_{1,2} \end{bmatrix}^H \\
& \times \begin{bmatrix} \hat{\boldsymbol{U}}_{1,1} & \hat{\boldsymbol{U}}_{1,2} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{\Lambda}}_{1,1} & \\ & \mathbf{0} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{U}}_{1,1} & \hat{\boldsymbol{U}}_{1,2} \end{bmatrix}^H \\
& \times \begin{bmatrix} \boldsymbol{U}_{1,1} & \boldsymbol{U}_{1,2} \end{bmatrix} \begin{bmatrix} \mathbf{\Sigma}_{1,1}^{-1} & \\ & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{V}_{1,1} & \boldsymbol{V}_{1,2} \end{bmatrix}^H \\
= & \begin{bmatrix} \boldsymbol{V}_{1,1} & \boldsymbol{V}_{1,2} \end{bmatrix} \begin{bmatrix} \mathbf{\Sigma}_{1,1}^{-1} \boldsymbol{U}_{1,1}^H \hat{\boldsymbol{U}}_{1,1} \tilde{\mathbf{\Lambda}}_{1,1} \hat{\boldsymbol{U}}_{1,1}^H \boldsymbol{U}_{1,1} \mathbf{\Sigma}_{1,1}^{-1} & \\ & \mathbf{0} \end{bmatrix} \\
& \times \begin{bmatrix} \boldsymbol{V}_{1,1} & \boldsymbol{V}_{1,2} \end{bmatrix}^H
\end{aligned}
$$

where $\boldsymbol{H}_1^{+}$ denotes the pseudo-inverse of $\boldsymbol{H}_1$. It can be verified that $\boldsymbol{U}_{1,1}^H \hat{\boldsymbol{U}}_{1,1}$ is unitary since $\boldsymbol{U}_1^H \hat{\boldsymbol{U}}_1$ is unitary. Hence, by denoting $\mathbf{\Lambda}_{1,1} = \mathbf{\Sigma}_{1,1}^2$, we have

$$
\begin{aligned}
\text{tr}(\boldsymbol{Q}) & \geq \text{tr}(\boldsymbol{H}_1^{+} \boldsymbol{H}_1 \boldsymbol{Q} \boldsymbol{H}_1^H \boldsymbol{H}_1^{H+}) & (10) \\
& = \text{tr}(\mathbf{\Sigma}_{1,1}^{-1} \boldsymbol{U}_{1,1}^H \hat{\boldsymbol{U}}_{1,1} \tilde{\mathbf{\Lambda}}_{1,1} \hat{\boldsymbol{U}}_{1,1}^H \boldsymbol{U}_{1,1} \mathbf{\Sigma}_{1,1}^{-1}) & \\
& \geq \text{tr}(\tilde{\mathbf{\Lambda}}_{1,1} \mathbf{\Lambda}_{1,1}^{-1}) & (11)
\end{aligned}
$$

The first inequality (10) in the above utilizes the second inequality in (6) and the fact that $\boldsymbol{H}_1^H \boldsymbol{H}_1^{H+} \boldsymbol{H}_1^{+} \boldsymbol{H}_1$ is a project matrix with eigenvalues being only 1 and 0. The second inequality (11) comes from the first inequality in (6).

Now we note that the following $\boldsymbol{Q}_o$:

$$
\boldsymbol{Q}_o = \begin{bmatrix} \boldsymbol{V}_{1,1} & \boldsymbol{V}_{1,2} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{\Lambda}}_{1,1} \mathbf{\Lambda}_{1,1}^{-1} & \\ & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{V}_{1,1} & \boldsymbol{V}_{1,2} \end{bmatrix}^H
$$

satisfies

$$
\boldsymbol{H}_1 \boldsymbol{Q}_o \boldsymbol{H}_1^H = \boldsymbol{U}_1 \begin{bmatrix} \tilde{\mathbf{\Lambda}}_{1,1} & \\ & \mathbf{0} \end{bmatrix} \boldsymbol{U}_1^H.
$$

and hence satisfies (8). Furthermore, $\boldsymbol{Q}_o$ has the minimum trace $\text{tr}(\tilde{\mathbf{\Lambda}}_{1,1} \mathbf{\Lambda}_{1,1}^{-1})$. Therefore, the theorem is proved.

4. OPTIMAL POWER DISTRIBUTION

Under the theorem, we can now rewrite (1) as

$$
\tilde{\boldsymbol{y}} = \mathbf{\Sigma}_2 \mathbf{\Sigma}_F \mathbf{\Sigma}_1 \tilde{\boldsymbol{s}} + \mathbf{\Sigma}_2 \mathbf{\Sigma}_F \tilde{\boldsymbol{n}}_1 + \tilde{\boldsymbol{n}}_2 \qquad (12)
$$

where $\tilde{\boldsymbol{y}} = \boldsymbol{U}_2^H \boldsymbol{y}$, $\tilde{\boldsymbol{s}} = \boldsymbol{V}_1^H \boldsymbol{s}$, $\tilde{\boldsymbol{n}}_1 = \boldsymbol{U}_1^H \boldsymbol{n}_1$, and $\tilde{\boldsymbol{n}}_2 = \boldsymbol{U}_2^H \boldsymbol{n}_2$. Note that $\tilde{\boldsymbol{s}}$, $\tilde{\boldsymbol{n}}_1$, and $\tilde{\boldsymbol{n}}_2$ are all white. The original MIMO relay channel has thus been decomposed into a set of parallel SISO subchannels. The corresponding source-destination capacity is

$$
C = \log_2 \det[\boldsymbol{I} + \mathbf{\Lambda}_1 \mathbf{\Lambda}_2 \mathbf{\Lambda}_F \mathbf{\Lambda}_Q (\sigma_1^2 \mathbf{\Lambda}_2 \mathbf{\Lambda}_F + \sigma_2^2 \boldsymbol{I})^{-1}] \qquad (13)
$$

where $\mathbf{\Lambda}_F = \mathbf{\Sigma}_F^2$, $\mathbf{\Lambda}_1 = \mathbf{\Sigma}_1^2$, and $\mathbf{\Lambda}_2 = \mathbf{\Sigma}_2^2$. We will use

$$
\begin{aligned}
\mathbf{\Lambda}_1 &= \text{diag}(\alpha_1, \alpha_2, \cdots, \alpha_L) \\
\mathbf{\Lambda}_2 &= \text{diag}(\beta_1, \beta_2, \cdots, \beta_L) \\
\mathbf{\Lambda}_Q &= \text{diag}(q_1, q_2, \cdots, q_L) \\
\mathbf{\Lambda}_F &= \text{diag}(f_1, f_2, \cdots, f_L)
\end{aligned}
$$

The capacity can be further expressed as

$$
C = \sum_{k=1}^{L} \log_2(1 + \frac{\alpha_k q_k \beta_k f_k}{\sigma_1^2 \beta_k f_k + \sigma_2^2})
$$

The power constraints (5) become $\sum_{k=1}^{L} q_k \leq P_1$ and $\sum_{k=1}^{L} f_k(\alpha_k q_k + \sigma_1^2) \leq P_2$.

Let $d_k = f_k(\alpha_k q_k + \sigma_1^2)$. By some simple derivations, the original optimization problem is equivalent to

$$
\begin{aligned}
\max_{\boldsymbol{q}, \boldsymbol{d}} \quad & \sum_{k=1}^{L} \log_2 \frac{(1 + \frac{\alpha_k}{\sigma_1^2} q_k)(1 + \frac{\beta_k}{\sigma_2^2} d_k)}{1 + \frac{\alpha_k}{\sigma_1^2} q_k + \frac{\beta_k}{\sigma_2^2} d_k} & (14) \\
\text{s.t.} \quad & \sum_{k=1}^{L} q_k \leq P_1 \ \& \ \sum_{k=1}^{L} d_k \leq P_2 & (15)
\end{aligned}
$$

Once d_k and q_k are obtained, f_k can be calculated via $f_k = d_k / (\alpha_k q_k + \sigma_1^2)$. For convenience, we define $\boldsymbol{q} = [q_1, q_2, \cdots, q_L]^T$, and $\boldsymbol{f}$ and $\boldsymbol{d}$ are similarly defined.

The optimal solution to the above problem is still unknown as it is nonconvex. But we propose two algorithms that we believe can yield a "near" optimal solution if not the optimal solution.

4.1. Iterative Algorithm

By observing the above optimization problem (14), we see that the roles of $\boldsymbol{q}$ and $\boldsymbol{d}$ are virtually symmetric. If we fix either $\boldsymbol{q}$ or $\boldsymbol{d}$, the problem is equivalent to the relay-only design problem in [2]. Thus an iterative procedure can be designed by estimating $\boldsymbol{q}$ and $\boldsymbol{d}$ alternately. Once converged, $\boldsymbol{q}$ and $\boldsymbol{d}$ can be used to compute $\boldsymbol{f}$. According to [2], we have the following iterative algorithm:

1. Determine an initial value of $\boldsymbol{q}$, satisfying the power constraint (15).
2. Calculate $\boldsymbol{d}$ with the given $\boldsymbol{q}$ as follows:

$$
d_k = \frac{\sigma_2^2}{2\beta_k} \left[\sqrt{(\frac{\alpha_k}{\sigma_1^2} q_k)^2 + 4 \frac{\alpha_k}{\sigma_1^2} q_k \frac{\beta_k}{\sigma_2^2} \mu} - \frac{\alpha_k}{\sigma_1^2} q_k - 2 \right]^+
$$

   where $[x]^+$ denotes $\max\{0, x\}$, and μ is decided by

$$
\frac{1}{2} \sum_{k=1}^{L} \frac{\sigma_2^2}{\beta_k} \left[\sqrt{(\frac{\alpha_k}{\sigma_1^2} q_k)^2 + 4 \frac{\alpha_k}{\sigma_1^2} q_k \frac{\beta_k}{\sigma_2^2} \mu} - \frac{\alpha_k}{\sigma_1^2} q_k - 2 \right]^+ = P_2
$$

3. Calculate $\boldsymbol{q}$ with the new $\boldsymbol{d}$ fixed via a similar set of formula.
4. Go back to Step 2 until convergence.

4.2. Dual Decomposition Algorithm

This algorithm is via dual decomposition (see [4] or [5]). Instead of focusing on the primal problem (14), the dual objection function $g(\boldsymbol{\lambda})$ is considered, which can be decoupled into L independent problems with respect to q_k and d_k. That is,

$$
g(\boldsymbol{\lambda}) = \max_{\boldsymbol{q}, \boldsymbol{d}} J(\boldsymbol{q}, \boldsymbol{d}, \boldsymbol{\lambda})
$$

where

$$J(\boldsymbol{q},\boldsymbol{d},\boldsymbol{\lambda}) = \sum_{k=1}^{L}\log_2\frac{(1+\frac{\alpha_k}{\sigma_1^2}q_k)(1+\frac{\beta_k}{\sigma_2^2}d_k)}{1+\frac{\alpha_k}{\sigma_1^2}q_k+\frac{\beta_k}{\sigma_2^2}d_k} + \lambda_1(P_1-\sum_{k=1}^{L}q_k)+\lambda_2(P_2-\sum_{k=1}^{L}d_k) = \sum_{k=1}^{L}J_k(q_k,d_k,\boldsymbol{\lambda})+\lambda_1P_1+\lambda_2P_2$$

with

$$J_k(q_k,d_k,\boldsymbol{\lambda}) = \log_2\frac{(1+\frac{\alpha_k}{\sigma_1^2}q_k)(1+\frac{\beta_k}{\sigma_2^2}d_k)}{1+\frac{\alpha_k}{\sigma_1^2}q_k+\frac{\beta_k}{\sigma_2^2}d_k} - \lambda_1 q_k - \lambda_2 d_k$$

The dual optimization problem is

$$\min_{\lambda\geq 0} \quad g(\boldsymbol{\lambda}) \tag{16}$$

The solution to the dual problem provides an upper bound to the primal problem (14) (see [6]).

One can observe that the dual problem is decomposed with respect to q_k and d_k by absorbing the (originally coupled) power constraints into the Lagrangian $g(\boldsymbol{\lambda})$. Thus the task now is to solve the unconstrained problem (16).

The dual decomposition method [4] provides a computationally tractable way to solve this problem. The idea is to perform a global search to find the optimal values of λ_1 and λ_2. For every fixed λ_1 and λ_2, the optimal values of q_k and d_k are calculated by maximizing $J_k(q_k,d_k,\boldsymbol{\lambda})$. However, the optimal solution of maximizing $J_k(q_k,d_k,\boldsymbol{\lambda})$ is not easy to find, since it is not a concave function. Here we again use iterative searching by alternately calculating q_k and d_k. Since each iteration increases the objective function $J_k(q_k,d_k,\boldsymbol{\lambda})$, it is guaranteed to converge to a local maximum.

5. NUMERICAL EXAMPLES

In this section, we compare the capacities of the relay system under three different schemes:

1. C_{Naive}: Capacity of the relay system without power distribution control neither at the source nor at the relay. This scheme is called a "naive" scheme.
2. $C_{RelayOnly}$: Capacity of the relay system with power distribution control only at the relay.
3. C_{Joint}: Capacity of the relay system with joint power distribution control at both the source and the relay.

For the relay-only scheme and the naive scheme, the source covariance matrix is fixed to be a scaled identity matrix $\frac{P_1}{M}\boldsymbol{I}$.

For the naive scheme, the relay simply normalizes the received signal to meet the power constraint and then forward the signal to the destination. In this case, the weighting matrix at the relay is

$$\boldsymbol{F} = \eta\boldsymbol{I}$$

Given the power constraint P_2 at the relay, we can have

$$\eta = \sqrt{\frac{P_2}{\text{tr}(\frac{P_1}{M}\boldsymbol{H}_1\boldsymbol{H}_1^H+\sigma_1^2\boldsymbol{I})}}$$

In all simulations, the channels are assumed to be independent Raleigh fading channels (i.e, all entries in the channel matrices are independent and complex Gaussian with zero mean and unit variance). The number of antennas is chosen to be $L = M = N = 4$. We will use SNR1 $= \frac{P_1}{M\sigma_1^2}$ and SNR2 $= \frac{P_2}{L\sigma_2^2}$.

For all cases considered, the two optimization algorithms yielded the same results, which seems to suggest that the optimal solution had most probably been found in each case. 20000 Monte Carlo runs were done for each pair of SNR1 and SNR2.

Fig. 2 shows the probability density functions (PDFs) of C_{Naive}. Fig. 3 shows the PDFs of the capacity gain $C_{RelayOnly}/C_{Naive}$. Fig. 4 shows the PDFs of the capacity gain C_{Joint}/C_{Naive}. Fig. 5 shows the PDFs of the capacity gain $C_{Joint}/C_{RelayOnly}$.

From these figures, we see that when SNR2 (SNR at the destination) is low, the relay-only scheme yields a large capacity gain over the naive scheme. And when SNR1 (SNR at the relay) is low, the joint scheme yields an additional capacity gain. This observation is supported by the fact that capacity is a logarithmic function of power and hence is more sensitive to change of power in low power region than in high power region.

6. CONCLUSION

We have studied a joint optimization of the source covariance matrix and the relay matrix of a non-regenerative MIMO relay system. The results shown here extend the previous work in [2] where the source covariance matrix was assumed to be proportional to the identity matrix. It is shown that the structure (or coordinates) of the optimal relay matrix given in [2] is still valid in the current setting. Optimization algorithms are given to compute the optimal power distribution along the optimal coordinates of both the source covariance matrix and the relay matrix (although the exact optimal power distribution remains an open problem). Our results show that the joint source and relay optimization yields a further capacity gain beyond the relay-only optimization when SNR at the relay is low.

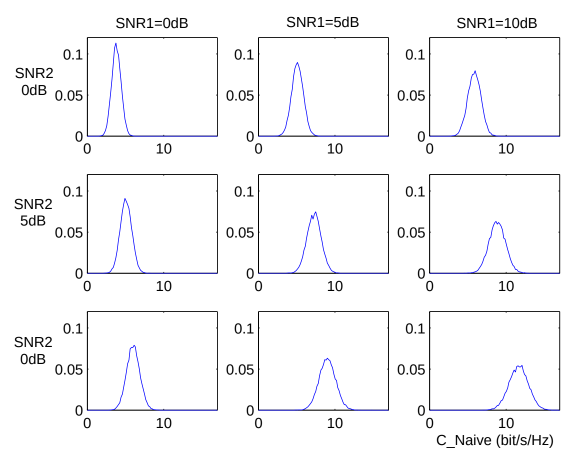

Fig. 2. PDF of the capacity of the naive scheme

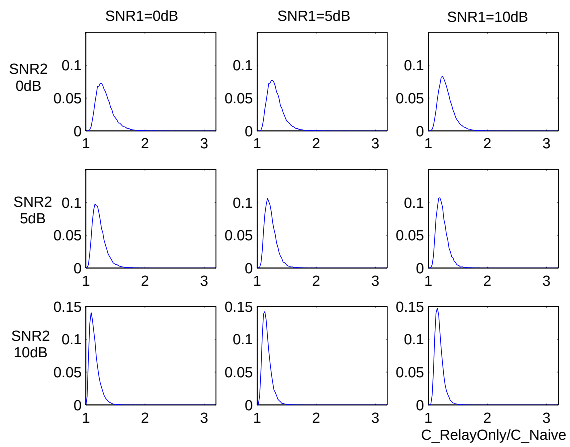

Fig. 3. PDF of the capacity gain of the relay-only scheme over the naive scheme

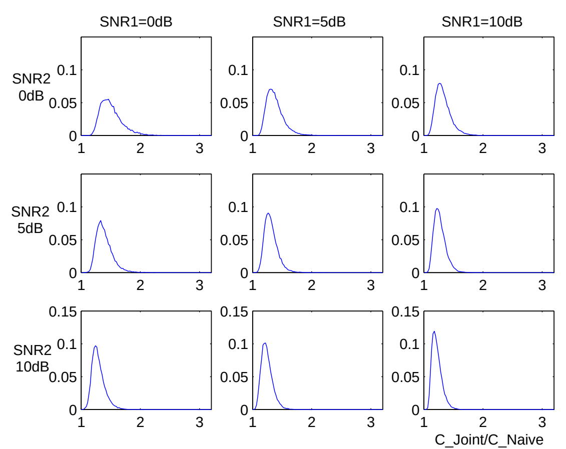

Fig. 4. PDF of the capacity gain of the joint scheme over the naive scheme

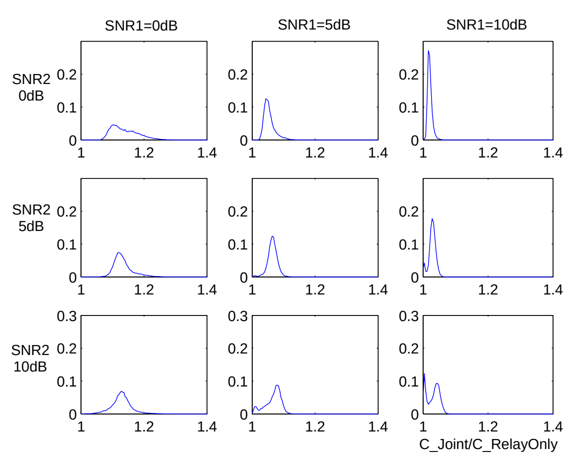

Fig. 5. PDF of the capacity gain of the joint scheme over the relay-only scheme.

7. REFERENCES

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

[2] X. Tang and Y. Hua, "Optimal design of non-regenerative MIMO wireless relays," *IEEE Transactions on Wireless Communication*, to appear.

[3] I. E. Telatar, "Capacity of multi-antenna gaussian channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, 1999.

[4] R. Cendrillon, M. Moonen, J. Verliden, T. Bostoen, and W. Yu, "Optimal multiuser spectrum management for digital subscriber lines," *IEEE International Conference on Communications*, vol. 1, pp. 1–5, 2004.

[5] W. Yu and R. Lui, "Dual methods for non-convex spectrum optimization of multi-carrier systems," *IEEE Transaction on Communication*, to appear.

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

The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory or the U. S. Government.