Linear Non-Regenerative Multicarrier MIMO Relay Communications Based on MMSE Criterion

Yue Rong, *Member, IEEE*

Abstract—In this letter we propose linear non-regenerative multicarrier multiple-input multiple-output (MIMO) relay technique that aims to minimize the mean-squared error (MSE) of the signal waveform estimation at the destination. We generalize the existing result on the structure of the optimal relay amplifying matrix by considering the direct source-destination link. To minimize the MSE, a power loading algorithm is developed which has a significantly reduced computational complexity compared with existing techniques.

Index Terms—MIMO, non-regenerative relay, multicarrier, MMSE, power loading.

I. INTRODUCTION

RECENTLY, non-regenerative approaches for cooperative multiple-input multiple-output (MIMO) communications have received many research interests [1]-[12]. Optimal algorithms have been developed to maximize the mutual information (MI) between source and destination of singlecarrier MIMO relay systems [1]-[4] and multicarrier MIMO relay systems [5]. However, the maximal MI (MMI)-based algorithms in [1]-[5] are optimal only when the codewords are infinitely long. However, in practical communication systems, due to the delay constraint, codewords always have a finite length. Thus, the performance of the MMI-based algorithms will degrade in practical systems. An important criterion in communication systems is to minimize the raw bit-error-rate (BER) which is closely related to the mean-squared error (MSE) of the signal waveform estimation. Motivated by this fact, the optimal source beamforming vector and the optimal rank-1 relay amplifying matrix of half- and full-duplex relay system based on the minimal MSE (MMSE) criterion is derived in [6]. However, in such system only one symbol is transmitted at a time, which is strongly suboptimal in terms of the system data rate (multiplexing gain). Without the rank-1 constraint, MMSE-based approaches are developed in [7] and [8] to optimize the relay amplifying matrix of one-way MIMO relay system and two-way MIMO relay channel [9]. However, [7]-[9] assume that the source covariance matrix is an identity matrix, which is strongly suboptimal. Recently, a unified framework is established for jointly optimizing the source precoding matrix and the relay amplifying matrix of two-hop linear non-regenerative MIMO relay systems with a broad class of objective functions [10].

In this letter, we present our recent discoveries on the MSE-based non-regenerative multicarrier MIMO relay design that are not included in [10]. First, we rigorously derive the structure of the optimal relay amplifying matrix when the direct source-destination link is included. Note that in [1], [2], [5], [7]-[10], the optimal relay amplifying matrix is derived when the direct link is omitted. In [3], the structure of the relay amplifying matrix that maximizes the source-destination MI in the presence of the direct link was derived. However, such relay amplifying matrix is suboptimal, since it does not consider the structure of the transmission power constraint at the relay node. The relay amplifying matrix structure proposed in [4] is also suboptimal since it does not include the channel information of the direct link.

Second, when the direct link is sufficiently weak to be omitted as in [1]-[10], we develop a new power allocation algorithm to minimize the MSE. The proposed algorithm has a substantially reduced computational complexity compared with the iterative power allocation algorithm in [10], with almost no performance loss in BER and only a negligible MSE performance degradation. These two results are the main contributions of this letter. Numerical simulations show that our algorithm greatly outperforms [1]-[8] in terms of MSE and BER.

II. SYSTEM MODEL

We consider a three-node multicarrier MIMO communication system where the source node transmits information to the destination node with the aid of one relay node. The source, relay, and destination nodes are equipped with N_s, N_r, and N_d antennas, respectively. The communication process between the source and destination nodes is completed in two time slots. In the first time slot, the signal sequence is modulated by N_c subcarriers. We denote $N_b^{(n)}$, $n = 1, \cdots, N_c$, as the number of symbols in the nth subcarrier. Hereafter, the superscript (n) denotes the corresponding variables for the nth subcarrier. Then the $N_b^{(n)} \times 1$ signal vector $\mathbf{s}^{(n)}(t)$ is linearly precoded as

$$\mathbf{x}^{(n)}(t) = \mathbf{B}^{(n)}\mathbf{s}^{(n)}(t), \qquad n = 1, \cdots, N_c \qquad (1)$$

where $\mathbf{B}^{(n)}$ is an $N_s \times N_b^{(n)}$ ($N_s \geq N_b^{(n)}$) precoding matrix for the source signals at the nth sub-carrier. The precoded signal vector $\mathbf{x}^{(n)}(t)$ is transmitted to the relay and the destination nodes via the nth subcarrier. The received signal at the relay and the destination nodes can be respectively written as

$$\mathbf{y}_r^{(n)}(t) = \mathbf{H}_s^{(n)}\mathbf{x}^{(n)}(t) + \mathbf{v}_r^{(n)}(t), \qquad n = 1, \cdots, N_c \qquad (2)$$

$$\mathbf{y}_d^{(n)}(t) = \mathbf{H}_d^{(n)}\mathbf{x}^{(n)}(t) + \mathbf{v}_d^{(n)}(t), \qquad n = 1, \cdots, N_c \qquad (3)$$

where $\mathbf{H}_s^{(n)}$ is an $N_r \times N_s$ MIMO channel matrix between the source and relay nodes, $\mathbf{H}_d^{(n)}$ is an $N_d \times N_s$ MIMO channel matrix between the source and destination nodes, $\mathbf{y}_r^{(n)}(t)$ and $\mathbf{v}_r^{(n)}(t)$ are the received signal and the noise vectors at the relay node, respectively, $\mathbf{y}_d^{(n)}(t)$ and $\mathbf{v}_d^{(n)}(t)$ are the received signal and the noise vectors at the destination node at time t, respectively.

In the second slot, the relay node amplifies the received signal vector at the nth subcarrier with an $N_r \times N_r$ matrix $\mathbf{F}^{(n)}$ and transmits the amplified signal vector

$$\mathbf{x}_r^{(n)}(t+1) = \mathbf{F}^{(n)} \mathbf{y}_r^{(n)}(t), \qquad n = 1, \cdots, N_c \qquad (4)$$

to the destination node. Using (1), (2), (4), the received signal vector at the nth subcarrier of the destination node via the source-relay-destination link can be written as

$$\begin{aligned}
\mathbf{y}_d^{(n)}(t+1) &= \mathbf{H}_r^{(n)} \mathbf{x}_r^{(n)}(t+1) + \mathbf{v}_d^{(n)}(t+1) \\
&= \mathbf{H}_r^{(n)} \mathbf{F}^{(n)} \mathbf{H}_s^{(n)} \mathbf{x}^{(n)}(t) + \mathbf{H}_r^{(n)} \mathbf{F}^{(n)} \mathbf{v}_r^{(n)}(t) \\
&\quad + \mathbf{v}_d^{(n)}(t+1)
\end{aligned} \qquad (5)$$

where $\mathbf{H}_r^{(n)}$ is an $N_d \times N_r$ MIMO channel matrix between the relay and destination nodes, $\mathbf{y}_d^{(n)}(t+1)$ and $\mathbf{v}_d^{(n)}(t+1)$ are the received signal and the noise vectors at the destination node at time $t+1$, respectively.

Combining (3) and (5), the received signal vector at the destination node over two time slots is given by

$$\begin{aligned}
\mathbf{y}^{(n)}(t) &\triangleq \begin{bmatrix} \mathbf{y}_d^{(n)}(t+1) \\ \mathbf{y}_d^{(n)}(t) \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{H}_r^{(n)} \mathbf{F}^{(n)} \mathbf{H}_s^{(n)} \\ \mathbf{H}_d^{(n)} \end{bmatrix} \mathbf{B}^{(n)} \mathbf{s}^{(n)}(t) \\
&\quad + \begin{bmatrix} \mathbf{H}_r^{(n)} \mathbf{F}^{(n)} \mathbf{v}_r^{(n)}(t) + \mathbf{v}_d^{(n)}(t+1) \\ \mathbf{v}_d^{(n)}(t) \end{bmatrix}.
\end{aligned} \qquad (6)$$

We assume that $\mathbf{H}_s^{(n)}$, $\mathbf{H}_r^{(n)}$ and $\mathbf{H}_d^{(n)}$, $n = 1, \cdots, N_c$, are all quasi-static. We also assume that all additive noises are independent and identically distributed complex circularly symmetric Gaussian noise with zero mean and unit variance.

III. PROPOSED APPROACH

Due to its simplicity, a linear receiver is used at the destination node to retrieve the transmitted signals at all subcarriers. The estimated signal waveform is given by

$$\hat{\mathbf{s}}^{(n)}(t) = (\mathbf{W}^{(n)})^H \mathbf{y}^{(n)}(t), \qquad n = 1, \cdots, N_c \qquad (7)$$

where $\mathbf{W}^{(n)}$ is a $2N_d \times N_b^{(n)}$ weight matrix at the nth subcarrier, and $(\cdot)^H$ denotes the matrix Hermitian transpose. The MSE of the signal waveform estimation is

$$\begin{aligned}
&\sum_{n=1}^{N_c} \mathrm{E}\Big[\mathrm{tr}\big((\hat{\mathbf{s}}^{(n)}(t) - \mathbf{s}^{(n)}(t))(\hat{\mathbf{s}}^{(n)}(t) - \mathbf{s}^{(n)}(t))^H\big)\Big] \\
&= \sum_{n=1}^{N_c} \mathrm{tr}\Big(\big((\mathbf{W}^{(n)})^H \bar{\mathbf{H}}^{(n)} - \mathbf{I}_{N_b^{(n)}}\big)\big((\mathbf{W}^{(n)})^H \bar{\mathbf{H}}^{(n)} - \mathbf{I}_{N_b^{(n)}}\big)^H \\
&\quad + (\mathbf{W}^{(n)})^H \bar{\mathbf{C}}^{(n)} \mathbf{W}^{(n)}\Big)
\end{aligned} \qquad (8)$$

where we assume that $\mathrm{E}[\mathbf{s}^{(n)}(t)(\mathbf{s}^{(n)}(t))^H] = \mathbf{I}_{N_b^{(n)}}$, $n = 1, \cdots, N_c$, $\mathrm{tr}(\cdot)$ stands for the matrix trace, $\mathrm{E}[\cdot]$ denotes the statistical expectation, and

$$\bar{\mathbf{H}}^{(n)} \triangleq \begin{bmatrix} \mathbf{H}_r^{(n)} \mathbf{F}^{(n)} \mathbf{H}_s^{(n)} \\ \mathbf{H}_d^{(n)} \end{bmatrix} \mathbf{B}^{(n)}$$

$$\bar{\mathbf{C}}^{(n)} \triangleq \begin{bmatrix} \mathbf{H}_r^{(n)} \mathbf{F}^{(n)} (\mathbf{H}_r^{(n)} \mathbf{F}^{(n)})^H + \mathbf{I}_{N_d} & \mathbf{0}_{N_d \times N_d} \\ \mathbf{0}_{N_d \times N_d} & \mathbf{I}_{N_d} \end{bmatrix}.$$

Here $\mathbf{0}_{m \times n}$ denotes an $m \times n$ matrix with all zeros entries. We assume that $\bar{\mathbf{H}}^{(n)}$, $n = 1, \cdots, N_c$, are known by the destination node, $\mathbf{H}_s^{(n)}$, $n = 1, \cdots, N_c$, are available at the source node, and the relay node knows $\mathbf{H}_s^{(n)}$ and $\mathbf{H}_r^{(n)}$, $n = 1, \cdots, N_c$. Note that the same amount of channel state information (CSI) knowledge is required for the algorithms in [1]-[10]. In practice, $\mathbf{H}_s^{(n)}$ can be obtained at the relay node with standard training methods, and $\mathbf{H}_r^{(n)}$ can be acquired by a feedback from the destination node to the relay node. Similarly, $\bar{\mathbf{H}}^{(n)}$ is obtained at the destination node through channel training, and $\mathbf{H}_s^{(n)}$ is obtained at the source node by a feedback from the relay node. For wireless relays, the fading is often relatively slow whenever the mobility of the relays is relatively low, and for static relays, the CSI can be almost constant. Thus, in this way, the necessary CSI can be obtained at each node with a reasonably high precision.

The optimal linear receiver that minimizes (8) is the well-known MMSE receiver, whose weight matrix is given by

$$\mathbf{W}^{(n)} = \left(\bar{\mathbf{H}}^{(n)} (\bar{\mathbf{H}}^{(n)})^H + \bar{\mathbf{C}}^{(n)} \right)^{-1} \bar{\mathbf{H}}^{(n)}, \quad n = 1, \cdots, N_c \qquad (9)$$

where $(\cdot)^{-1}$ denotes the matrix inversion. Substituting (9) back into (8), we obtain the MSE given by (10) at the bottom of the next page, where the matrix inversion lemma $(\mathbf{A} + \mathbf{BCD})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1} \mathbf{B} (\mathbf{D} \mathbf{A}^{-1} \mathbf{B} + \mathbf{C}^{-1})^{-1} \mathbf{D} \mathbf{A}^{-1}$ is applied to obtain the second equation.

The optimal $\{\mathbf{B}^{(n)}\}$ and $\{\mathbf{F}^{(n)}\}$ that minimize (10) can be found by the following optimization problem

$$\min_{\{\mathbf{B}^{(n)}\}, \{\mathbf{F}^{(n)}\}} \mathrm{MSE} \qquad (11)$$

$$\text{s.t.} \sum_{n=1}^{N_c} \mathrm{tr}\big(\mathbf{B}^{(n)}(\mathbf{B}^{(n)})^H\big) \leq p_s \qquad (12)$$

$$\sum_{n=1}^{N_c} \mathrm{tr}\Big(\mathbf{F}^{(n)}\big[\mathbf{H}_s^{(n)} \mathbf{B}^{(n)}(\mathbf{H}_s^{(n)} \mathbf{B}^{(n)})^H$$
$$+ \mathbf{I}_{N_r}\big](\mathbf{F}^{(n)})^H\Big) \leq p_r \qquad (13)$$

where for a matrix $\mathbf{X}$, $\{\mathbf{X}^{(n)}\} \triangleq \mathbf{X}^{(1)}, \mathbf{X}^{(2)}, \cdots, \mathbf{X}^{(N_c)}$. Here (12) and (13) are constraints for the transmission power at the source and relay nodes, respectively, and $p_s > 0$, $p_r > 0$ are the corresponding available transmission power. It can be clearly seen from (12) and (13) that the power constraints are imposed on the sum power throughout all subcarriers. Thus, the power allocation is optimized throughout all subcarriers. In other words, the subcarriers are "coupled" through the power constraints at the source and relay nodes. Let us define the

following singular value decompositions (SVD)

$$\mathbf{H}_s^{(n)}\mathbf{B}^{(n)}=\mathbf{U}_1^{(n)}\mathbf{\Lambda}_1^{(n)}(\mathbf{V}_1^{(n)})^H, \quad \mathbf{H}_r^{(n)}=\mathbf{U}_r^{(n)}\mathbf{\Lambda}_r^{(n)}(\mathbf{V}_r^{(n)})^H$$
$$n=1,\cdots,N_c \quad (14)$$

where $\mathbf{\Lambda}_1^{(n)}$ and $\mathbf{\Lambda}_r^{(n)}$ are $R_1^{(n)}\times R_1^{(n)}$ and $R_r^{(n)}\times R_r^{(n)}$ square diagonal matrices (i.e., zero singularvalues are excluded). Here $R_1^{(n)}\triangleq\mathrm{rank}(\mathbf{H}_s^{(n)}\mathbf{B}^{(n)})$, $R_r^{(n)}\triangleq\mathrm{rank}(\mathbf{H}_r^{(n)})$, $\mathrm{rank}(\cdot)$ denotes the rank of a matrix.

THEOREM 1: The optimal $\mathbf{F}^{(n)}$ for (11)-(13) is given by

$$\mathbf{F}^{(n)}=\mathbf{V}_r^{(n)}\mathbf{A}^{(n)}(\mathbf{U}_1^{(n)})^H, \quad n=1,\cdots,N_c \quad (15)$$

where $\mathbf{A}^{(n)}$ is an $R_r^{(n)}\times R_1^{(n)}$ matrix.

PROOF: Without loss of generality, we write $\mathbf{F}^{(n)}$ as

$$\mathbf{F}^{(n)}=\begin{bmatrix}\mathbf{V}_r^{(n)} & \mathbf{V}_r^{(n)\perp}\end{bmatrix}\begin{bmatrix}\mathbf{A}^{(n)} & \mathbf{G}^{(n)}\\ \mathbf{C}^{(n)} & \mathbf{D}^{(n)}\end{bmatrix}\begin{bmatrix}(\mathbf{U}_1^{(n)})^H\\ (\mathbf{U}_1^{(n)\perp})^H\end{bmatrix} \quad (16)$$

where $\mathbf{V}_r^{(n)\perp}(\mathbf{V}_r^{(n)\perp})^H = \mathbf{I}_{N_r} - \mathbf{V}_r^{(n)}(\mathbf{V}_r^{(n)})^H$, $\mathbf{U}_1^{(n)\perp}(\mathbf{U}_1^{(n)\perp})^H = \mathbf{I}_{N_r} - \mathbf{U}_1^{(n)}(\mathbf{U}_1^{(n)})^H$, such that $\bar{\mathbf{V}}_r^{(n)} \triangleq \begin{bmatrix}\mathbf{V}_r^{(n)} & \mathbf{V}_r^{(n)\perp}\end{bmatrix}$ and $\bar{\mathbf{U}}_1^{(n)} \triangleq \begin{bmatrix}\mathbf{U}_1^{(n)} & \mathbf{U}_1^{(n)\perp}\end{bmatrix}$ are unitary matrices, $\mathbf{A}^{(n)}$, $\mathbf{G}^{(n)}$, $\mathbf{C}^{(n)}$, $\mathbf{D}^{(n)}$ are arbitrary matrices with dimensions of $R_r^{(n)}\times R_1^{(n)}$, $R_r^{(n)}\times(N_r-R_1^{(n)})$, $(N_r-R_r^{(n)})\times R_1^{(n)}$, $(N_r-R_r^{(n)})\times(N_r-R_1^{(n)})$, respectively. Using (14) and (16) we have

$$(\mathbf{H}_r^{(n)}\mathbf{F}^{(n)})^H\mathbf{H}_r^{(n)}\mathbf{F}^{(n)}$$
$$= \bar{\mathbf{U}}_1^{(n)}\begin{bmatrix}(\mathbf{A}^{(n)})^H & (\mathbf{C}^{(n)})^H\\ (\mathbf{G}^{(n)})^H & (\mathbf{D}^{(n)})^H\end{bmatrix}\begin{bmatrix}(\mathbf{V}_r^{(n)})^H\\ (\mathbf{V}_r^{(n)\perp})^H\end{bmatrix}\mathbf{V}_r^{(n)}(\mathbf{\Lambda}_r^{(n)})^2$$
$$\times(\mathbf{V}_r^{(n)})^H\begin{bmatrix}\mathbf{V}_r^{(n)} & \mathbf{V}_r^{(n)\perp}\end{bmatrix}\begin{bmatrix}\mathbf{A}^{(n)} & \mathbf{G}^{(n)}\\ \mathbf{C}^{(n)} & \mathbf{D}^{(n)}\end{bmatrix}(\bar{\mathbf{U}}_1^{(n)})^H$$
$$= \bar{\mathbf{U}}_1^{(n)}\begin{bmatrix}(\mathbf{A}^{(n)})^H(\mathbf{\Lambda}_r^{(n)})^2\mathbf{A}^{(n)} & (\mathbf{A}^{(n)})^H(\mathbf{\Lambda}_r^{(n)})^2\mathbf{G}^{(n)}\\ (\mathbf{G}^{(n)})^H(\mathbf{\Lambda}_r^{(n)})^2\mathbf{A}^{(n)} & (\mathbf{G}^{(n)})^H(\mathbf{\Lambda}_r^{(n)})^2\mathbf{G}^{(n)}\end{bmatrix}$$
$$\times(\bar{\mathbf{U}}_1^{(n)})^H. \quad (17)$$

From (10) and (17) we see that the objective function (11) does not depend on $\mathbf{C}^{(n)}$ and $\mathbf{D}^{(n)}$. Substituting (17) back into (10) we have (18) at the bottom of the next page, where we applied the matrix inversion lemma for partitioned matrix to obtain the second equation, and

$$\mathbf{M}^{(n)} \triangleq (\mathbf{A}^{(n)})^H(\mathbf{\Lambda}_r^{(n)})^2\mathbf{G}^{(n)}$$
$$\times\left((\mathbf{G}^{(n)})^H(\mathbf{\Lambda}_r^{(n)})^2\mathbf{G}^{(n)}+\mathbf{I}_{N_r-R_1^{(n)}}\right)^{-1}(\mathbf{G}^{(n)})^H(\mathbf{\Lambda}_r^{(n)})^2\mathbf{A}^{(n)}.$$

Obviously, from (18), MSE is minimized if $\mathbf{M}^{(n)} = \mathbf{0}_{R_1^{(n)}\times R_1^{(n)}}$, $n = 1,\cdots,N_c$, which holds if $\mathbf{G} = \mathbf{0}_{R_r^{(n)}\times(N_r-R_1^{(n)})}$.

Now we look at the constraint (13). The power consumed by the relay node can be rewritten as (19) at the bottom of the next page. Obviously, $\mathbf{G} = \mathbf{0}_{R_r^{(n)}\times(N_r-R_1^{(n)})}$, $\mathbf{C} = \mathbf{0}_{(N_r-R_r^{(n)})\times R_1^{(n)}}$, and $\mathbf{D} = \mathbf{0}_{(N_r-R_r^{(n)})\times(N_r-R_1^{(n)})}$ minimize the power consumption. Thus we have $\mathbf{F}^{(n)} = \mathbf{V}_r^{(n)}\mathbf{A}^{(n)}(\mathbf{U}_1^{(n)})^H$. $\square$

It can be seen from (18) that given (15), MSE depends on $\mathbf{B}^{(n)}$ and $\mathbf{A}^{(n)}$. For general $\mathbf{H}_d^{(n)}$, one can not obtain a closed-form solution for the structure of the optimal $\mathbf{A}^{(n)}$ and $\mathbf{B}^{(n)}$. In [11], an alternating algorithm is developed to optimize $\mathbf{F}^{(n)}$ and $\mathbf{B}^{(n)}$ without exploiting the optimal structure of $\mathbf{F}^{(n)}$ in (15). The algorithm in [11] is suboptimal and has a very high computational complexity. A locally optimal solution of $\mathbf{A}^{(n)}$ and $\mathbf{B}^{(n)}$ can be obtained using general gradient-based numerical methods. However, gradient-based methods usually have difficulties in tuning various parameters such as the initialization point and step size, especially for matrix variables. How to jointly optimize $\mathbf{B}^{(n)}$ and $\mathbf{A}^{(n)}$ with a reasonable computational complexity is an important future research topic.

When the direct link is sufficiently weak (eg. shadowed by obstacles such as mountain and building) that can be ignored ($\mathbf{H}_d^{(n)} = \mathbf{0}_{N_d\times N_s}$) as in [2], [5]-[10], the MSE (10) is equivalent to

$$\text{MSE} =\sum_{n=1}^{N_c} \text{tr}\Big(\Big[\mathbf{I}_{N_b^{(n)}} + (\mathbf{H}_s^{(n)}\mathbf{B}^{(n)})^H\mathbf{H}_s^{(n)}\mathbf{B}^{(n)}-\mathbf{V}_1^{(n)}\mathbf{\Lambda}_1^{(n)}$$
$$\times\Big((\mathbf{A}^{(n)})^H(\mathbf{\Lambda}_r^{(n)})^2\mathbf{A}^{(n)}+\mathbf{I}_{R_1^{(n)}}\Big)^{-1}\mathbf{\Lambda}_1^{(n)}(\mathbf{V}_1^{(n)})^H\Big]^{-1}\Big).$$

Obviously, in this case, for an efficient utilization of the available transmission power at the source and relay nodes, $N_b^{(n)}$ should satisfy $N_b^{(n)} \leq \min(R_s^{(n)}, R_r^{(n)})$, $n=1,\cdots,N_c$, with $R_s^{(n)} \triangleq \mathrm{rank}(\mathbf{H}_s^{(n)})$. Since the channel matrices are random, we have $\min(R_s^{(n)}, R_r^{(n)}) = \min(N_s, N_r, N_d)$. In fact, it can be shown that $R_1^{(n)} = N_b^{(n)}$ and the optimal $\mathbf{A}^{(n)}$ is given by

$$\mathbf{A}^{(n)} = \Big[\mathbf{A}_1^{(n)}, \mathbf{0}_{N_b^{(n)}\times(R_r^{(n)}-N_b^{(n)})}\Big]^T$$

where $\mathbf{A}_1^{(n)}$ is an $N_b^{(n)}\times N_b^{(n)}$ diagonal matrix. In particular, denoting $\mathbf{H}_s^{(n)} = \mathbf{U}_s^{(n)}\mathbf{\Lambda}_s^{(n)}(\mathbf{V}_s^{(n)})^H$, $n=1,\cdots,N_c$, as the SVD of $\mathbf{H}_s^{(n)}$, where $\mathbf{\Lambda}_s^{(n)}$ is a $R_s^{(n)}\times R_s^{(n)}$ diagonal matrix, it has been shown in [10] using the majorization theory [13] that the solution to the problem (11)-(13) is

$$\mathbf{B}_0^{(n)} = \mathbf{V}_{s,1}^{(n)}\mathbf{\Lambda}_b^{(n)}\mathbf{U}_0^{(n)}, \quad n=1,\cdots,N_c \quad (20)$$
$$\mathbf{F}_0^{(n)} = \mathbf{V}_{r,1}^{(n)}\mathbf{\Lambda}_f^{(n)}(\mathbf{U}_{s,1}^{(n)})^H, \quad n=1,\cdots,N_c. \quad (21)$$

$$\text{MSE} =\sum_{n=1}^{N_c}\text{tr}\Big(\Big[\mathbf{I}_{N_b^{(n)}} + (\mathbf{H}_d^{(n)}\mathbf{B}^{(n)})^H\mathbf{H}_d^{(n)}\mathbf{B}^{(n)}+(\mathbf{H}_r^{(n)}\mathbf{F}^{(n)}\mathbf{H}_s^{(n)}\mathbf{B}^{(n)})^H$$
$$\times(\mathbf{H}_r^{(n)}\mathbf{F}^{(n)}(\mathbf{H}_r^{(n)}\mathbf{F}^{(n)})^H+\mathbf{I}_{N_d})^{-1}\mathbf{H}_r^{(n)}\mathbf{F}^{(n)}\mathbf{H}_s^{(n)}\mathbf{B}^{(n)}\Big]^{-1}\Big)$$
$$=\sum_{n=1}^{N_c}\text{tr}\Big(\Big[\mathbf{I}_{N_b^{(n)}} + (\mathbf{H}_d^{(n)}\mathbf{B}^{(n)})^H\mathbf{H}_d^{(n)}\mathbf{B}^{(n)}+(\mathbf{H}_s^{(n)}\mathbf{B}^{(n)})^H\Big[\mathbf{I}_{N_r}-(\mathbf{I}_{N_r}+(\mathbf{H}_r^{(n)}\mathbf{F}^{(n)})^H\mathbf{H}_r^{(n)}\mathbf{F}^{(n)})^{-1}\Big]\mathbf{H}_s^{(n)}\mathbf{B}^{(n)}\Big]^{-1}\Big) (10)$$

Here $\mathbf{V}_{s,1}^{(n)}, \mathbf{V}_{r,1}^{(n)}, \mathbf{U}_{s,1}^{(n)}$ stand for the singular vector matrices corresponding to $N_b^{(n)}$ strongest singular values, $\mathbf{\Lambda}_b^{(n)}, \mathbf{\Lambda}_f^{(n)}$ are $N_b^{(n)} \times N_b^{(n)}$ diagonal matrices, and $\mathbf{U}_0^{(n)}$ is an $N_b^{(n)} \times N_b^{(n)}$ unitary matrix.

It is worth noting that (15) includes (21) as a special case. Thus, Theorem 1 generalizes the structure of the optimal relay amplifying matrix from the relay system without the direct link [10] to relay systems with the direct link. This is one contribution of this letter. In the following, we focus on relay system where the direct link is sufficiently weak that can be ignored and develop a power loading algorithm which has a greatly reduced computational complexity than the iterative power loading algorithm in [10].

Let us define for $i = 1, \cdots, N_b^{(n)}$, $n = 1, \cdots, N_c$

$$a_i^{(n)} \triangleq (\lambda_{s,i}^{(n)})^2, \quad b_i^{(n)} \triangleq (\lambda_{r,i}^{(n)})^2,$$
$$x_i^{(n)} \triangleq (\lambda_{b,i}^{(n)})^2, \quad y_i^{(n)} \triangleq (\lambda_{f,i}^{(n)})^2 [(\lambda_{s,i}^{(n)} \lambda_{b,i}^{(n)})^2 + 1]$$

where $\lambda_{r,i}^{(n)}, \lambda_{f,i}^{(n)}, \lambda_{s,i}^{(n)}, \lambda_{b,i}^{(n)}$ are the main diagonal elements of $\mathbf{\Lambda}_r^{(n)}, \mathbf{\Lambda}_f^{(n)}, \mathbf{\Lambda}_s^{(n)}, \mathbf{\Lambda}_b^{(n)}$, respectively. It is shown in [10] that $\mathbf{\Lambda}_b^{(n)}$ and $\mathbf{\Lambda}_f^{(n)}$ in (20) and (21) can be optimized through optimizing $\{x_i^{(n)}\}$ and $\{y_i^{(n)}\}$, which is given by the problem (22)-(25) at the bottom of the next page, where for a scalar a, $\{a_i^{(n)}\} \triangleq a_1^1, \cdots, a_{N_b^{(N_c)}}^{N_c}$.

Closed-form solution to problem (22)-(25) is intractable. In fact, since (22) is nonconvex with respect to $\{x_i^{(n)}\}$ and $\{y_i^{(n)}\}$, a global-optimal solution is hard to obtain with a practical computational complexity. In [12], a grid-searching-based algorithm is applied to find the global-optimal solution where the objective is to maximize the mutual information between source and destination. With some modifications, the algorithm in [12] can also be applied to solve the problem (22)-(25). However, the computational complexity of grid-searching is extremely high, since in order to obtain a reasonably good solution, searching over a high-dense grid must be employed. In [10], an iterative algorithm is developed to

find a locally optimal solution of problem (22)-(25), which has a much lower computational complexity than that of [12].

However, the iterative algorithm in [10] may still be computationally intensive for practical systems. Note that if the objective function (22) can be decoupled for $\{x_i^{(n)}\}$ and $\{y_i^{(n)}\}$, then the optimization of $\{x_i^{(n)}\}$ and $\{y_i^{(n)}\}$ can be independently conducted, since the constraints (23)-(25) are already decoupled for $\{x_i^{(n)}\}$ and $\{y_i^{(n)}\}$. Towards this end, we first derive an upper-bound of (22) based on its special structure. In particular, we have

$$\sum_{n=1}^{N_c} \sum_{i=1}^{N_b^{(n)}} \frac{a_i^{(n)} x_i^{(n)} + b_i^{(n)} y_i^{(n)} + 1 + \alpha}{a_i^{(n)} b_i^{(n)} x_i^{(n)} y_i^{(n)} + a_i^{(n)} x_i^{(n)} + b_i^{(n)} y_i^{(n)} + 1} =$$
$$\sum_{n=1}^{N_c} \sum_{i=1}^{N_b^{(n)}} \left(\frac{c_1}{a_i^{(n)} x_i^{(n)} + 1} + \frac{c_2}{b_i^{(n)} y_i^{(n)} + 1} \right) \quad (26)$$

where $\alpha > 0$. From (26) we see that $\alpha = 1$, and $c_1 = c_2 = 1$. Using (26) as the objective function and together with the constraints (23)-(25), we find that $\{x_i^{(n)}\}$ and $\{y_i^{(n)}\}$ can be optimized independently. In particular, the problem of optimizing $\{x_i^{(n)}\}$ is written as

$$\min_{\{x_i^{(n)}\}} \sum_{n=1}^{N_c} \sum_{i=1}^{N_b^{(n)}} \frac{1}{a_i^{(n)} x_i^{(n)} + 1} \quad (27)$$
$$\text{s.t.} \quad \sum_{n=1}^{N_c} \sum_{i=1}^{N_b^{(n)}} x_i^{(n)} \leq p_s \quad (28)$$
$$x_i^{(n)} \geq 0, \quad i = 1, \cdots, N_b^{(n)}, \quad n = 1, \cdots, N_c. \quad (29)$$

While the problem of optimizing $\{y_i^{(n)}\}$ is given by

$$\min_{\{y_i^{(n)}\}} \sum_{n=1}^{N_c} \sum_{i=1}^{N_b^{(n)}} \frac{1}{b_i^{(n)} y_i^{(n)} + 1} \quad (30)$$
$$\text{s.t.} \quad \sum_{n=1}^{N_c} \sum_{i=1}^{N_b^{(n)}} y_i^{(n)} \leq p_r \quad (31)$$

$$\text{MSE} = \sum_{n=1}^{N_c} \text{tr}\left(\left[\mathbf{I}_{N_b^{(n)}} + (\mathbf{H}_d^{(n)} \mathbf{B}^{(n)})^H \mathbf{H}_d^{(n)} \mathbf{B}^{(n)} + \mathbf{V}_1^{(n)} \mathbf{\Lambda}_1^{(n)} \left(\mathbf{I}_{R_1^{(n)}} - (\mathbf{U}_1^{(n)})^H \bar{\mathbf{U}}_1^{(n)} \right. \right. \right.$$
$$\left. \times \begin{bmatrix} (\mathbf{A}^{(n)})^H (\mathbf{\Lambda}_r^{(n)})^2 \mathbf{A}^{(n)} + \mathbf{I}_{R_1^{(n)}} & (\mathbf{A}^{(n)})^H (\mathbf{\Lambda}_r^{(n)})^2 \mathbf{G}^{(n)} \\ (\mathbf{G}^{(n)})^H (\mathbf{\Lambda}_r^{(n)})^2 \mathbf{A}^{(n)} & (\mathbf{G}^{(n)})^H (\mathbf{\Lambda}_r^{(n)})^2 \mathbf{G}^{(n)} + \mathbf{I}_{N_r - R_1^{(n)}} \end{bmatrix}^{-1} (\bar{\mathbf{U}}_1^{(n)})^H \mathbf{U}_1^{(n)} \right) \mathbf{\Lambda}_1^{(n)} (\mathbf{V}_1^{(n)})^H \Bigg]^{-1} \right)$$
$$= \left[\mathbf{I}_{N_b^{(n)}} + (\mathbf{B}^{(n)})^H \left((\mathbf{H}_d^{(n)})^H \mathbf{H}_d^{(n)} + (\mathbf{H}_s^{(n)})^H \mathbf{H}_s^{(n)} \right) \mathbf{B}^{(n)} - \mathbf{V}_1^{(n)} \mathbf{\Lambda}_1^{(n)} \right.$$
$$\left. \times \left((\mathbf{A}^{(n)})^H (\mathbf{\Lambda}_r^{(n)})^2 \mathbf{A}^{(n)} + \mathbf{I}_{R_1^{(n)}} - \mathbf{M}^{(n)} \right)^{-1} \mathbf{\Lambda}_1^{(n)} (\mathbf{V}_1^{(n)})^H \right]^{-1} \quad (18)$$

$$\text{tr}\left(\mathbf{F}^{(n)} \left(\mathbf{H}_s^{(n)} \mathbf{B}^{(n)} (\mathbf{H}_s^{(n)} \mathbf{B}^{(n)})^H + \mathbf{I}_{N_r} \right) (\mathbf{F}^{(n)})^H \right)$$
$$= \text{tr}\left(\begin{bmatrix} \mathbf{A}^{(n)} & \mathbf{G}^{(n)} \\ \mathbf{C}^{(n)} & \mathbf{D}^{(n)} \end{bmatrix} \begin{bmatrix} (\mathbf{\Lambda}_1^{(n)})^2 + \mathbf{I}_{R_1^{(n)}} & \mathbf{0}_{R_1^{(n)} \times (N_r - R_1^{(n)})} \\ \mathbf{0}_{(N_r - R_1^{(n)}) \times R_1^{(n)}} & \mathbf{I}_{N_r - R_1^{(n)}} \end{bmatrix} \begin{bmatrix} (\mathbf{A}^{(n)})^H & (\mathbf{C}^{(n)})^H \\ (\mathbf{G}^{(n)})^H & (\mathbf{D}^{(n)})^H \end{bmatrix} \right)$$
$$= \text{tr}\left(\mathbf{A}^{(n)} \left((\mathbf{\Lambda}_1^{(n)})^2 + \mathbf{I}_{R_1^{(n)}} \right) (\mathbf{A}^{(n)})^H + \mathbf{G}^{(n)} (\mathbf{G}^{(n)})^H + \mathbf{C}^{(n)} \left((\mathbf{\Lambda}_1^{(n)})^2 + \mathbf{I}_{R_1^{(n)}} \right) (\mathbf{C}^{(n)})^H + \mathbf{D}^{(n)} (\mathbf{D}^{(n)})^H \right). \quad (19)$$

$$y_i^{(n)} \geq 0, \quad i = 1, \cdots, N_b^{(n)}, \quad n = 1, \cdots, N_c . \quad (32)$$

Interestingly, both problems optimize the MSE of the signal waveform estimation of single-hop parallel scalar Gaussian channels. In particular, problem (27)-(29) minimizes the MSE of the source-relay channels, while problem (30)-(32) optimizes that of the relay-destination channels. The solutions to problem (27)-(29) and problem (30)-(32) follow the water-filling principle and are given respectively by

$$x_i^{(n)} = \frac{1}{a_i^{(n)}} \left[\sqrt{\frac{a_i^{(n)}}{\nu_1}} - 1 \right]^{\dagger}, \quad y_i^{(n)} = \frac{1}{b_i^{(n)}} \left[\sqrt{\frac{b_i^{(n)}}{\nu_2}} - 1 \right]^{\dagger}$$

$$i = 1, \cdots, N_b^{(n)}, \quad n = 1, \cdots, N_c$$

where $\nu_1 > 0$ and $\nu_2 > 0$ are the solution to the following nonlinear equations, respectively

$$\sum_{n=1}^{N_c} \sum_{i=1}^{N_b^{(n)}} \frac{1}{a_i^{(n)}} \left[\sqrt{\frac{a_i^{(n)}}{\nu_1}} - 1 \right]^{\dagger} = p_s$$

$$\sum_{n=1}^{N_c} \sum_{i=1}^{N_b^{(n)}} \frac{1}{b_i^{(n)}} \left[\sqrt{\frac{b_i^{(n)}}{\nu_2}} - 1 \right]^{\dagger} = p_r.$$

Here $[x]^{\dagger} \triangleq \max(0, x)$.

Finally, the optimal relay amplifying matrix is given by $\mathbf{F}_0^{(n)} = \mathbf{V}_{r,1}^{(n)} \mathbf{\Lambda}_f^{(n)} (\mathbf{U}_{s,1}^{(n)})^H$, with $\lambda_{f,i}^{(n)} = \sqrt{y_i^{(n)} / [(\lambda_{s,i}^{(n)})^2 x_i^{(n)} + 1]}$. The optimal source precoding matrix is

$$\mathbf{B}_0^{(n)} = \mathbf{V}_{s,1}^{(n)} \mathbf{\Lambda}_b^{(n)} \mathbf{U}_0^{(n)}, \quad \text{with} \quad \lambda_{b,i}^{(n)} = \sqrt{x_i^{(n)}}. \quad (33)$$

From (33) we see that the source node only requires the information of $\mathbf{H}_s^{(n)}$, but does not need to know $\mathbf{H}_r^{(n)}$. The relay node computes the diagonal elements of $\mathbf{\Lambda}_b^{(n)}$ and sends them to the source node. Note that this can be easily implemented since only up to $N_b^{(n)}$ real numbers are needed to be sent from the relay node to the source node at each subcarrier.

Compared with the iterative power loading algorithm in [10], the proposed upper-bound-based non-iterative algorithm is suboptimal. However, the proposed algorithm has only a computational complexity of *each iteration* of the iterative algorithm in [10]. Therefore, the proposed algorithm has a substantially reduced computational complexity. Note

that the upper-bound (26) is tight when the transmission power p_s and p_r are sufficiently high, since in such case, $a_i^{(n)} x_i^{(n)} + b_i^{(n)} y_i^{(n)} \gg 1$, and we have $a_i^{(n)} x_i^{(n)} + b_i^{(n)} y_i^{(n)} + 1 \approx a_i^{(n)} x_i^{(n)} + b_i^{(n)} y_i^{(n)} + 2$, $i = 1, \cdots, N_b^{(n)}, n = 1, \cdots, N_c$. In Section IV we will see that the proposed algorithm yields only a slight MSE increment and has almost the same BER performance compared with the iterative algorithm in [10]. Therefore, the proposed power loading algorithm is very useful for practical relay communication systems with a weak direct link, and is another contribution of this letter.

IV. SIMULATIONS

Similar to [2], [5]-[10], we simulate a two-hop MIMO relay system without the direct link. In the simulations, the channel between each transmit-receive antenna pair is modelled as the ETSI "Vehicular A" channel environment. An OFDM communication system with $N_c = 64$ subcarriers is assumed. The channel matrices have zero-mean entries with variances σ_s^2 and σ_r^2 for $\mathbf{H}_s^{(n)}$ and $\mathbf{H}_r^{(n)}$, respectively. The signal-to-noise ratio (SNR) is defined as $\mathrm{SNR}_s \triangleq \sigma_s^2 p_s / N_c$ and $\mathrm{SNR}_r \triangleq \sigma_r^2 p_r / N_c$ for the source-relay and relay-destination links, respectively. All simulation results are averaged over 2000 independent channel realizations. We compare the proposed upper-bound-based suboptimal (UBS) algorithm with the iterative optimal (ITO) algorithm in [10], the relay-only suboptimal (ROS) algorithm developed in [7], and the MMI algorithm developed in [2] which maximizes the mutual information between source and destination. Note that the ITO, UBS, and MMI algorithms require the knowledge of $\mathbf{H}_s^{(n)}$ at the source node, while the ROS algorithm does not. Moreover, the MMI algorithm [2] and the ITO algorithm [10] have a similar computational complexity, which is much higher than the proposed UBS algorithm. Both the ROS and the UBS algorithms are non-iterative and have the same complexity order.

In the first example, we choose $N_s = N_r = N_d = 3$. Fig. 1 displays the averaged MSE (AMSE) of different algorithms versus SNR_s for $\mathrm{SNR}_r = 20$dB and $N_b^{(n)} = 3$, $n = 1, \cdots, N_c$. It can be seen that the ITO algorithm consistently yields the lowest AMSE over the whole SNR_s range. The performance of the proposed UBS algorithm is very closed to the ITO algorithm. Since for practical communication systems the BER is an important criterion, the performance of all algorithms in terms of BER versus SNR_s is shown in Fig. 2 for $\mathrm{SNR}_r = 20$dB and $N_b^{(n)} = 2$, $n = 1, \cdots, N_c$. The QPSK constellations are used. In the simulation, after $\hat{\mathbf{s}}^{(n)}(t)$

$$\min_{\{x_i^{(n)}\}, \{y_i^{(n)}\}} \sum_{n=1}^{N_c} \sum_{i=1}^{N_b^{(n)}} \frac{a_i^{(n)} x_i^{(n)} + b_i^{(n)} y_i^{(n)} + 1}{a_i^{(n)} b_i^{(n)} x_i^{(n)} y_i^{(n)} + a_i^{(n)} x_i^{(n)} + b_i^{(n)} y_i^{(n)} + 1} \quad (22)$$

$$\text{s.t.} \quad \sum_{n=1}^{N_c} \sum_{i=1}^{N_b^{(n)}} x_i^{(n)} \leq p_s \quad (23)$$

$$\sum_{n=1}^{N_c} \sum_{i=1}^{N_b^{(n)}} y_i^{(n)} \leq p_r \quad (24)$$

$$x_i^{(n)} \geq 0, \quad y_i^{(n)} \geq 0, \quad i = 1, \cdots, N_b^{(n)}, \quad n = 1, \cdots, N_c. \quad (25)$$
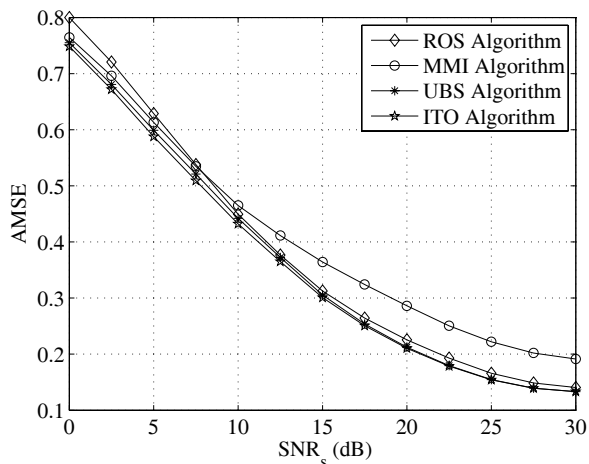
Fig. 1. AMSE versus SNR_s. $N_s = N_r = N_d = 3$, $N_b^{(n)} = 3$, $SNR_r = 20dB$.

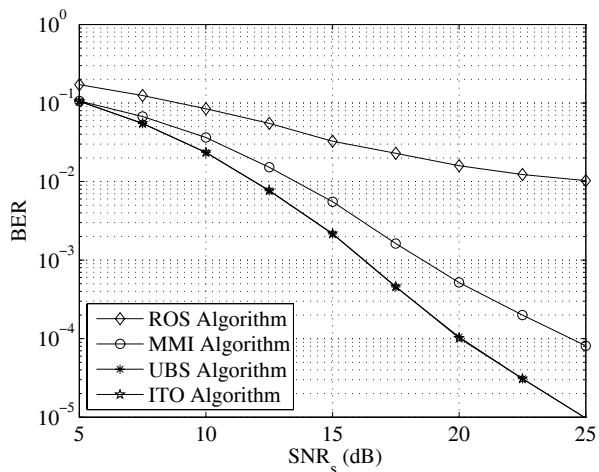Fig. 3. BER versus SNR_s. $N_s = 5, N_r = 6, N_d = 4$, $N_b^{(n)} = 3$, $SNR_r = 20dB$.

Fig. 2. BER versus SNR_s. $N_s = N_r = N_d = 3$, $N_b^{(n)} = 2$, $SNR_r = 20dB$.

is obtain by the linear MMSE receiver as in (7), a symbol-by-symbol demodulation is used to retrieve the source bits. We find from Fig. 2 that the UBS algorithm and the ITO algorithm almost have the same BER performance.

In the second example, we simulate a MIMO relay system with different number of antennas. We set $N_s = 5, N_r = 6, N_d = 4$, and $N_b^{(n)} = 3$, $n = 1, \cdots, N_c$. Fig. 3 shows the BER performance of all four algorithms. From Fig. 3, we find that the performance of the UBS algorithm is almost identical to that of the ITO algorithm. Thus, the UBS algorithm is very useful for practical systems. Note that in Figs. 2 and 3, the ROS algorithm yields a higher BER than that of the MMI algorithm. The reason is that the ROS algorithm does not optimize the source precoding matrix and thus has a poor performance when $N_b^{(n)}$ is less than $\min(N_s, N_r, N_d)$.

V. CONCLUSIONS

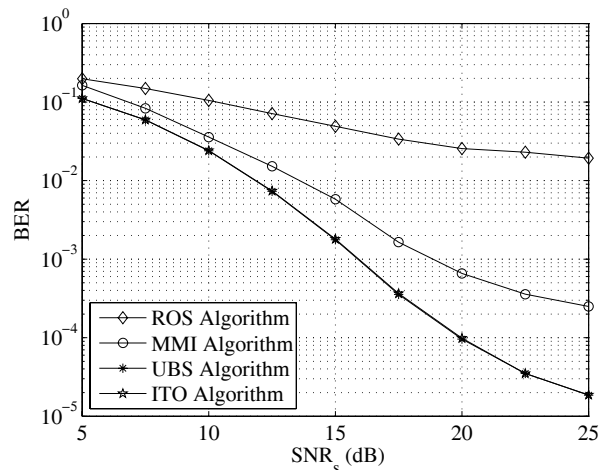We derived the optimal structure of the relay amplifying matrix for multicarrier MIMO relay systems with the direct source-destination link. This relay amplifying matrix minimizes the MSE of the signal waveform estimation. A power loading algorithm was developed which has a significantly reduced computational complexity than the existing approaches with almost no BER performance loss.

REFERENCES

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

[2] 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 Process.*, Waltham, WA, July 2006, pp. 239-243.

[3] 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.

[4] Q. Miao, A. Osseiran, and J. Gan, "MIMO amplify-and-forward relaying: spatial gain and filter matrix design," in *Proc. IEEE Int. Conf. Commun.*, Beijing, China, May 2008, pp. 26-30.

[5] 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.

[6] J. Joung and A. H. Sayed, "Design of half- and full-duplex relay systems based on the MMSE formulation," in *Proc. IEEE Workshop Statistical Signal Process.*, Cardiff, UK, Aug. 2009, pp. 281-284.

[7] W. Guan and H. Luo, "Joint MMSE transceiver design in non-regenerative MIMO relay systems," *IEEE Commun. Lett.*, vol. 12, pp. 517-519, July 2008.

[8] 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.

[9] N. Lee, H. Park, and J. Chun, "Linear precoder and decoder design for two-way AF MIMO relaying system," in *Proc. IEEE Veh. Technol. Conf.*, Singapore, May 2008, pp. 1221-1225.

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

[11] Y. Rong, "MMSE-based non-regenerative multicarrier MIMO wireless relay communications with direct source-destination link," in *Proc. IEEE Int. Conf. Commun.*, Dresden, Germany, June 2009.

[12] W. Zhang, U. Mitra, and M. Chiang, "Optimization of amplify-and-forward multicarrier two-hop transmission," submitted to *IEEE Trans. Commun.*, 2008.

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