Multiuser Two-Way Amplify-and-Forward Relay Processing and Power Control Methods for Beamforming Systems

Jingon Joung, *Member, IEEE*, and Ali H. Sayed, *Fellow, IEEE*

Abstract—In this paper, multiple-input multiple-output (MIMO) relay transceiver processing is proposed for multiuser two-way relay communications. The relay processing is optimized based on both zero-forcing (ZF) and minimum mean-square-error (MMSE) criteria under relay power constraints. Various transmit and receive beamforming methods are compared including eigen beamforming, antenna selection, random beamforming, and modified equal gain beamforming. Local and global power control methods are designed to achieve fairness among all users and to maximize the system signal-to-noise ratio (SNR). Numerical results show that the proposed multiuser two-way relay processing can efficiently eliminate both co-channel interference (CCI) and self-interference (SI).

Index Terms—Beamforming, minimum mean-square error (MMSE), multiple-input multiple-output (MIMO), multiuser, power control, two-way relay, zero-forcing (ZF).

I. INTRODUCTION

VARIOUS relay communication techniques have been studied to improve the capacity and/or expand the coverage of wireless networks [1]–[9]. In relay communications, multiple relays create a multiple-input multiple-output (MIMO) environment at the relay layer and exploit various MIMO techniques such as array processing [1], space-frequency coding [2], and beamforming [3]–[5]. By using multiple antennas at the transmit or receive nodes, the source as well as the other nodes can also take advantage of MIMO techniques, such as spatial multiplexing [6], [7] and beamforming based on either maximizing the signal-to-noise ratio (SNR) [8] or minimizing the mean-square-error (MMSE) [9]. For general relay communications between two user nodes, four phases are typically used: from user 1 to the relay node, from the relay node to
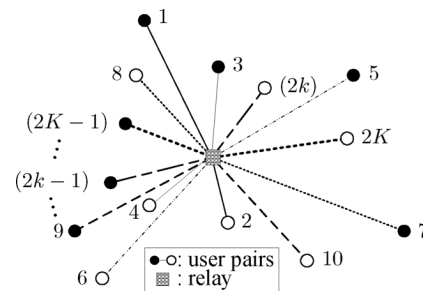
Fig. 1. Multiuser two-way relay system with $2K$-users and one-relay node.

user 2, and two other phases for reverse link communications. In contrast, for direct communications without relay, two phases are used: from user 1 to user 2 and vice versa. Since an orthogonal channel is required to implement each phase, relay communications spend twice as much channel resources compared to direct communications.

To reduce the use of extra channel resources and improve spectral efficiency, *two-way* relay methods (bidirectional communications) between two users have been studied in [10]–[17]. Two-way communications complete the data exchange between two users through two phases: receive phase (first phase) and transmit phase (second phase). In the receive phase, the relay receives data from the two users simultaneously, and in the transmit phase, it retransmits the received signals to the two users. A signal transmitted by one user may return to the same user resulting in self-interference (SI) [10]. Fortunately, each user can cancel its SI since it knows its transmitted signal. To cancel SI efficiently and to improve system performance, bit-level coding techniques, such as XOR and superposition coding, have been proposed for the decode-and-forward (DF) relay systems [10]–[12], as well as linear preprocessing at the amplify-and-forward (AF) relay [13]. MIMO techniques have also been vigorously studied in two-way relay communications. Distributed space-time coding for multiple single-antenna relays was realized in partial decode-and-forward protocol in [14], and optimal beamforming for two-way relay communications was presented in [15]. Furthermore, optimal distributed beamforming methods were proposed in [16] and [17] by using multiple single-antenna relays.

In [18]–[21], two-way DF-relay communications involving two users has been extended to multiple pairs of users, as illustrated in Fig. 1 for $2K$ users (K pairs). The multiple pairs cause co-channel interference (CCI) among the pairs. The CCI can be mitigated by using orthogonal channels realized through the code [18], space [19], [20], or frequency domain [21], i.e.,
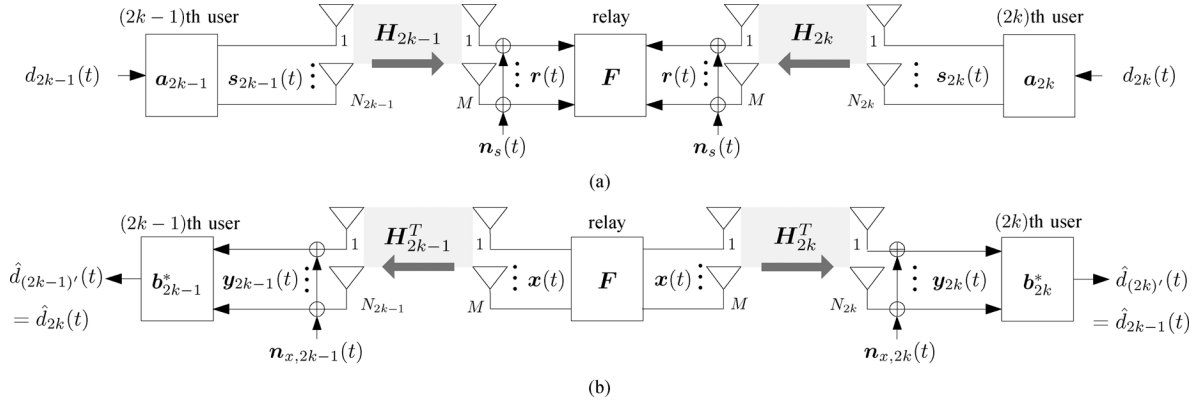
Fig. 2. Multiuser two-way relay system model, $k \in \{1, \ldots, K\}$. (a) First phase communication from the users to the relay. (b) Second phase communication from the relay to the users.

by using code division multiple access (CDMA), space-division multiple access (SDMA), or orthogonal-frequency-division multiple access (OFDMA), respectively. These orthogonal multiple access methods can accommodate multiple pairs without the CCIs; therefore, the conventional SI cancellation techniques can be applied for each user-pair on two-way communications.

In this paper, we focus on AF MIMO relay systems as shown in Fig. 2. First, we propose a multiuser two-way relay transceiver processing that can mitigate both SI and CCI by using steered beams through multiple antennas at the relay. This processing is designed optimally by using zero-forcing (ZF) and MMSE formulations under relay power constraints and predetermined transmit- and receive-beamforming of users [20]. The systems proposed in this paper can be considered as a general extension of [13] to SDMA systems with multiple beamforming users. The SDMA framework using the spatial resource (antennas) is attractive because it becomes possible to reuse the conventional channels constructed by time, frequency, or code. Various beamforming methods, such as eigen beamforming [22], antenna selection [23], random beamforming [24], and modification of equal gain beamforming [25] are introduced as the predetermined beamforming. Next, we propose local and global power control methods for the ZF-based system. In the local power control case, the multiuser power is allocated to each user for fairness among all users. In the global power control case, the total network power is divided into the multiuser power and the relay power for maximizing the system SNR under the high SNR assumption. Numerical results for the system SNR, sum rate, and bit-error-rate (BER) are shown. The performance of the power control methods are also compared with respect to the system sum rate and fairness among all users. Simulation results illustrate that i) the proposed multiuser two-way relay processing can efficiently eliminate both CCI and SI and ii) the proposed power control methods can effectively serve their purpose.

The rest of the paper is organized as follows. In Section II, the beamforming relay network model is described. Section III presents multiuser signal model; designs ZF- and MMSE-based two-way relay processing matrices and beamforming vectors; and also contains some useful remarks. In Section IV, local and global power control methods are designed for the ZF-based

system. Simulation results are shown in Section V. Section VI provides the conclusion. The specific derivations and proofs are relegated to Appendixes.

Notation

Throughout this paper, for any vector or matrix, the superscripts "T" and "$*$" denote transposition and complex conjugate transposition, respectively; $\mathrm{tr}(\boldsymbol{W})$ represents the trace of matrix $\boldsymbol{W}$; "E" stands for expectation of a random variable; for any scalar w, vector $\boldsymbol{w}$, and matrix $\boldsymbol{W}$, the notations $|w|$, $\|\boldsymbol{w}\|$, and $\|\boldsymbol{W}\|_F$ denote the absolute value of w, 2-norm of $\boldsymbol{w}$, and Frobenius-norm of $\boldsymbol{W}$, respectively; $\boldsymbol{0}_w$ is w-dimensional zero column vector; $\boldsymbol{I}_w$ is a w-by-w identity matrix; and $[\boldsymbol{W}]_w$ denotes the wth column vector of $\boldsymbol{W}$.

II. RELAY SYSTEM AND SIGNAL MODEL

There are $2K$ user nodes and one relay node in a multiuser two-way communications system as shown in Fig. 1. The $2K$ users result in K pairs of two users each performing two-way communication. Without loss of generality, it is assumed that the $(2k-1)$th and the $(2k)$th users communicate with each other ($k \in \{1, \ldots, K\}$ for user pairs) through two phases. In the first phase, the $2K$ users transmit their data simultaneously to the relay node as shown in Fig. 2(a). In the second phase, the relay retransmits (broadcasts) the received data to the users as shown in Fig. 2(b). Each kth user has N_k antennas and the relay has M antennas ($k \in \{1, \ldots, 2K\}$ for users). The MIMO channel matrix between the kth user and the relay node is represented by $\boldsymbol{H}_k \in \mathbb{C}^{M \times N_k}$ where the (m, n)th element is the channel gain between the nth antenna of the user and the mth antenna of the relay. Here, $\boldsymbol{H}_k$ is represented by

$$\boldsymbol{H}_k = \sigma_{H,k} \bar{\boldsymbol{H}}_k \tag{1}$$

where the elements of $\bar{\boldsymbol{H}}_k$ are i.i.d and zero-mean complex Gaussian random variables with unit variance and $\sigma_{H,k}$ is a path loss from the effects of shadowing and large scale fading. Let $d_k(t)$ denote the data symbol at time t for the kth user. The kth user performs transmit beamforming with vector $\boldsymbol{a}_k \in \mathbb{C}^{N_k \times 1}$ and transmits the vector

$$\boldsymbol{s}_k(t) = \boldsymbol{a}_k d_k(t) \in \mathbb{C}^{N_k \times 1} \tag{2}$$

to the relay node, through the first phase. In (2), the beam-forming vector will be represented as $\boldsymbol{a}_k = \gamma_k \bar{\boldsymbol{a}}_k$, where γ_k is a positive scalar to control the transmit power of $d_k(t)$ and $\bar{\boldsymbol{a}}_k$ is a normalized beamforming vector satisfying $\|\bar{\boldsymbol{a}}_k\|^2 = 1$. The received signal at the relay is then

$$\boldsymbol{r}(t) = \sum_{k=1}^{2K} \boldsymbol{H}_k \boldsymbol{s}_k(t) + \boldsymbol{n}_s(t) \in \mathbb{C}^{M \times 1} \qquad (3)$$

where $\boldsymbol{n}_s(t) \in \mathbb{C}^{M \times 1}$ is a zero-mean additive white Gaussian noise (AWGN) at the relay and $\mathrm{E}\,\boldsymbol{n}_s(t)\boldsymbol{n}_s^*(t) = \sigma_{n_s}^2 \boldsymbol{I}_M$. The relay multiplies $\boldsymbol{r}(t)$ by the relay processing matrix $\boldsymbol{F} \in \mathbb{C}^{M \times M}$, and forwards $\boldsymbol{x}(t) \in \mathbb{C}^{M \times 1}$ during the second phase, where

$$\boldsymbol{x}(t) = \boldsymbol{F}\boldsymbol{r}(t). \qquad (4)$$

Due to the reciprocity[1] between the first- and second phase channels, the received signal vector $\boldsymbol{y}_k(t) \in \mathbb{C}^{N_k \times 1}$, at the kth user, can be written as

$$\boldsymbol{y}_k(t) = \boldsymbol{H}_k^T \boldsymbol{x}(t) + \boldsymbol{n}_{x,k}(t) \qquad (5)$$

where $\boldsymbol{n}_{x,k}(t) \in \mathbb{C}^{N_k \times 1}$ is an AWGN at the kth user and $\mathrm{E}\,\boldsymbol{n}_{x,k}(t)\boldsymbol{n}_{x,k}^*(t) = \sigma_{n_x}^2 \boldsymbol{I}_{N_k}$. The kth user combines its own received signal (5) by using a receive beamforming vector $\boldsymbol{b}_k^* \in \mathbb{C}^{1 \times N_k}$ to get the estimate

$$\hat{d}_{k'}(t) = \boldsymbol{b}_k^* \boldsymbol{y}_k(t) \qquad (6)$$

where the subscript k' represents the index of the pair of the kth user. Thus, we have $(2k)' = 2k - 1$ and $(2k - 1)' = 2k$ since the estimate of the kth user is the transmitted data from the k'th user. Similarly to $\boldsymbol{a}_k$, the receive beamforming vector will be represented as $\boldsymbol{b}_k = q^{-1}\bar{\boldsymbol{b}}_k$, where $\bar{\boldsymbol{b}}_k$ is a normalized beam-forming vector satisfying $\|\bar{\boldsymbol{b}}_k\|^2 = 1$ and q^{-1} is a scaling factor that will be the same for all users. For notational convenience, the time index t is henceforth omitted. Now, we proceed to design the relay processing matrix $\boldsymbol{F}$ and the user beamforming vectors $\{\boldsymbol{a}_k, \boldsymbol{b}_k\}$.

III. MULTIUSER TWO-WAY RELAY SYSTEM DESIGN

In this section, we design the relay transceiver processing matrix $\boldsymbol{F}$ based on both ZF and MMSE criteria, under the assumption that the transmit and receive-beamforming vectors are pre-determined.

A. Multiuser Signal Model

From (2) and (3), the received signal at the relay can be rewritten in matrix form as follows:

$$\boldsymbol{r} = \boldsymbol{H}\boldsymbol{s} + \boldsymbol{n}_s = \boldsymbol{H}\boldsymbol{A}\boldsymbol{d} + \boldsymbol{n}_s \qquad (7)$$

where the multiuser transmit signal vector $\boldsymbol{s} = [\boldsymbol{s}_1^T \cdots \boldsymbol{s}_{2K}^T]^T$. Here, the multiuser data symbol vector $\boldsymbol{d} = [d_1 \cdots d_{2K}]^T \in$

$\mathbb{C}^{2K \times 1}$ satisfying $\mathrm{E}\,\boldsymbol{d}\boldsymbol{d}^* = \sigma_d^2 \boldsymbol{I}_{2K}$; the multiuser channel matrix is $\boldsymbol{H} = [\boldsymbol{H}_1 \quad \cdots \quad \boldsymbol{H}_{2K}] \in \mathbb{C}^{M \times \sum N_k}$; and the multiuser transmit-beamforming matrix is

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{0}_{N_1} & \cdots & \boldsymbol{0}_{N_1} \\ \boldsymbol{0}_{N_2} & \boldsymbol{a}_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \boldsymbol{0}_{N_{2K-1}} \\ \boldsymbol{0}_{N_{2K}} & \cdots & \boldsymbol{0}_{N_{2K}} & \boldsymbol{a}_{2K} \end{bmatrix}$$

$$= \begin{bmatrix} \bar{\boldsymbol{a}}_1 \gamma_1 & \boldsymbol{0}_{N_1} & \cdots & \boldsymbol{0}_{N_1} \\ \boldsymbol{0}_{N_2} & \bar{\boldsymbol{a}}_2 \gamma_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \boldsymbol{0}_{N_{2K-1}} \\ \boldsymbol{0}_{N_{2K}} & \cdots & \boldsymbol{0}_{N_{2K}} & \bar{\boldsymbol{a}}_{2K} \gamma_{2K} \end{bmatrix}$$

$$= \bar{\boldsymbol{A}}\boldsymbol{\Gamma} \in \mathbb{C}^{\sum N_k \times 2K}. \qquad (8)$$

In (8), $\boldsymbol{\Gamma}$ is a $2K$-dimensional diagonal matrix whose kth diagonal element is γ_k. Without loss of generality, the power ratio factors γ_k will be assumed to satisfy

$$\sum_{k=1}^{2K} \gamma_k^2 = 2K. \qquad (9)$$

Later, we shall show how to select $\boldsymbol{\Gamma}$ and σ_d^2 in order to ensure fairness among the users and system SNR, respectively. Constructing the vector $\hat{\boldsymbol{d}} = [\hat{d}_{1'}\ \hat{d}_{2'} \cdots \hat{d}_{(2K)'}]^T \in \mathbb{C}^{2K \times 1}$ with $\hat{d}_{k'}$ from (6), the multiuser received signal is represented as follows:

$$\hat{\boldsymbol{d}} = \boldsymbol{B}^* \boldsymbol{H}^T \boldsymbol{x} + \boldsymbol{B}^* \boldsymbol{n}_x. \qquad (10)$$

In (10), the multiuser noise vector is $\boldsymbol{n}_x = [\boldsymbol{n}_{x,1}^T \cdots \boldsymbol{n}_{x,2K}^T]^T \in \mathbb{C}^{\sum N_k \times 1}$; the retransmitted signal is

$$\boldsymbol{x} = \boldsymbol{F}\boldsymbol{r} = \boldsymbol{F}\boldsymbol{H}\boldsymbol{A}\boldsymbol{d} + \boldsymbol{F}\boldsymbol{n}_s \qquad (11)$$

from (4) and (7); and the multiuser receive-beamforming matrix is

$$\boldsymbol{B}^* = \begin{bmatrix} \boldsymbol{b}_1^* & \boldsymbol{0}_{N_2}^T & \cdots & \boldsymbol{0}_{N_{2K}}^T \\ \boldsymbol{0}_{N_1}^T & \boldsymbol{b}_2^* & \ddots & \vdots \\ \vdots & \ddots & \ddots & \boldsymbol{0}_{N_{2K}}^T \\ \boldsymbol{0}_{N_1}^T & \cdots & \boldsymbol{0}_{N_{2K-1}}^T & \boldsymbol{b}_{2K}^* \end{bmatrix}$$

$$= \begin{bmatrix} q^{-1}\bar{\boldsymbol{b}}_1^* & \boldsymbol{0}_{N_2}^T & \cdots & \boldsymbol{0}_{N_{2K}}^T \\ \boldsymbol{0}_{N_1}^T & q^{-1}\bar{\boldsymbol{b}}_2^* & \ddots & \vdots \\ \vdots & \ddots & \ddots & \boldsymbol{0}_{N_{2K}}^T \\ \boldsymbol{0}_{N_1}^T & \cdots & \boldsymbol{0}_{N_{2K-1}}^T & q^{-1}\bar{\boldsymbol{b}}_{2K}^* \end{bmatrix}$$

$$= q^{-1}\bar{\boldsymbol{B}}^* \in \mathbb{C}^{2K \times \sum N_k}. \qquad (12)$$

Using (11) in (10) yields the multiuser received signal vector:

$$\hat{\boldsymbol{d}} = \boldsymbol{B}^* \boldsymbol{H}^T \boldsymbol{F}\boldsymbol{H}\boldsymbol{A}\boldsymbol{d} + \boldsymbol{B}^* \boldsymbol{H}^T \boldsymbol{F}\boldsymbol{n}_s + \boldsymbol{B}^* \boldsymbol{n}_x. \qquad (13)$$

Due to the specific structure of the transmit and receive-beam-forming matrices in (8) and (12), it is difficult to jointly design $\{\boldsymbol{A}, \boldsymbol{B}\}$ with $\boldsymbol{F}$. This difficulty encourages us to optimize
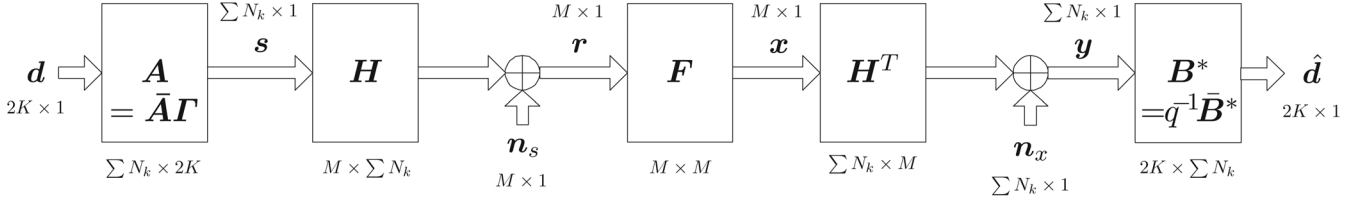
Fig. 3.   Beamforming, relaying and equalization model in multiuser MIMO two-way channels.

sequentially for $\{\boldsymbol{F}, \boldsymbol{a}_k, \boldsymbol{b}_k\}$. The sequential optimization for $\{\boldsymbol{F}, \boldsymbol{a}_k, \boldsymbol{b}_k\}$ yields coupled solutions and every node will need to know the *full* channel state information (CSI) $\{\boldsymbol{H}_k$ for all $k\}$, resulting in a burden on the network. To avoid this problem, we optimize $\boldsymbol{F}$ for predetermined transmit- and receive-beamforming matrices $\{\boldsymbol{A}, \boldsymbol{B}\}$ under the following relaxed assumptions: i) the kth node can estimate its own MIMO channel matrix $\boldsymbol{H}_k$ (termed as *perfect* CSI to distinguish it from *partial* or *no* CSI later) by training through the second phase whenever required, and ii) the relay node can estimate the full CSI $\{\boldsymbol{H}_k\}$ by training through the first phase. Although these assumptions could yield performance degradation compared to systems with full CSIs at every node, they are nevertheless more practical. Later, we employ various beamforming methods to evaluate the performance of the designed relay processing. With the predetermined beamforming matrices $\{\boldsymbol{A}, \boldsymbol{B}\}$ and full CSI $\{\boldsymbol{H}_k\}$ at the relay, the relay transceiver processing matrix $\boldsymbol{F}$ will be designed by using both ZF and MMSE criteria. Refer to Fig. 3, which shows the overall procedure for transmit- and receive-beamforming and relaying for multiuser two-way systems.

B. ZF-Based Design

Keeping in mind that the ZF criterion eliminates interferences, such as CCIs and SIs, while ignoring the noise, and that the $(2k-1)$th and $(2k)$th users' data should be exchanged after two-way communications, the ZF optimization is formulated as follows:

$$\boxed{\arg \min_{\boldsymbol{F}} \mathrm{E} \left\| \boldsymbol{d} - \boldsymbol{P}\hat{\boldsymbol{d}}_f \right\|^2 \text{ s.t. } \mathrm{E} \left\| \boldsymbol{x} \right\|^2 \leq P_R}. \tag{14}$$

In (14), the transmit power of the relay signal is limited by P_R; $\hat{\boldsymbol{d}}_f = [\hat{d}_{f,1'} \cdots \hat{d}_{f,(2K)'}]$ is the noise-free signal from (13), namely,

$$\hat{\boldsymbol{d}}_f = \boldsymbol{B}^* \boldsymbol{H}^T \boldsymbol{F} \boldsymbol{H} \boldsymbol{A} \boldsymbol{d}; \tag{15}$$

and $\boldsymbol{P}$ is a $2K$-dimensional block diagonal matrix with kth block diagonal matrix $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ exchanging data of the $(2k-1)$th and the $(2k)$th users, i.e.,

$$\boldsymbol{P} = \mathrm{blockdiag} \left(\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \dots, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \right).$$

The constraint in (14) can be merged into a minimization cost by the Lagrangian method, and the optimal $\boldsymbol{F}$ can be obtained from the Karush–Kuhn–Tucker (KKT) conditions [26]. By using $\hat{\boldsymbol{d}}_f$ from (15) and the retransmitted signal $\boldsymbol{x}$ in (11), the minimization problem (14) with constraint can be transformed into (16), shown at the bottom of this page, with a non-negative Lagrange multiplier λ. The cost J in (16) can be expanded as

$$\begin{aligned} J = {} & 2\sigma_d^2 K - \sigma_d^2 \mathrm{tr} \left(\boldsymbol{P}\boldsymbol{B}^* \boldsymbol{H}^T \boldsymbol{F} \boldsymbol{H} \boldsymbol{A} \right) \\ & - \sigma_d^2 \mathrm{tr} \left(\boldsymbol{A}^* \boldsymbol{H}^* \boldsymbol{F}^* (\boldsymbol{H}^*)^T \boldsymbol{B} \boldsymbol{P} \right) \\ & + \sigma_d^2 \mathrm{tr} \left(\boldsymbol{A}^* \boldsymbol{H}^* \boldsymbol{F}^* (\boldsymbol{H}^*)^T \boldsymbol{B} \boldsymbol{B}^* \boldsymbol{H}^T \boldsymbol{F} \boldsymbol{H} \boldsymbol{A} \right) \\ & + \lambda \left(\sigma_d^2 \mathrm{tr}(\boldsymbol{A}^* \boldsymbol{H}^* \boldsymbol{F}^* \boldsymbol{F} \boldsymbol{H} \boldsymbol{A}) + \sigma_{n_s}^2 \mathrm{tr}(\boldsymbol{F}^* \boldsymbol{F}) - P_R \right) \end{aligned} \tag{17}$$

under the assumption that the data symbols, channel elements, and noises are independent of one another. Using the techniques of complex matrix derivatives in [27] and setting the derivatives of (17) with respect to $\{\boldsymbol{F}, \lambda\}$ to zero, we get

$$\begin{aligned} \frac{\partial J}{\partial \boldsymbol{F}} = 0 \rightarrow {} & (\boldsymbol{H}^*)^T \boldsymbol{B} \boldsymbol{B}^* \boldsymbol{H}^T \boldsymbol{F} \boldsymbol{H} \boldsymbol{A} \boldsymbol{A}^* \boldsymbol{H}^* \\ & + \lambda \boldsymbol{F} \left(\boldsymbol{H} \boldsymbol{A} \boldsymbol{A}^* \boldsymbol{H}^* + \sigma_d^{-2} \sigma_{n_s}^2 \boldsymbol{I}_M \right) \\ & = (\boldsymbol{H}^*)^T \boldsymbol{B} \boldsymbol{P} \boldsymbol{A}^* \boldsymbol{H}^* \end{aligned} \tag{18a}$$

$$\begin{aligned} \frac{\partial J}{\partial \lambda} = 0 \rightarrow {} & \sigma_d^2 \mathrm{tr} \left(\boldsymbol{A}^* \boldsymbol{H}^* \boldsymbol{F}^* \boldsymbol{F} \boldsymbol{H} \boldsymbol{A} \right) + \sigma_{n_s}^2 \mathrm{tr} \left(\boldsymbol{F}^* \boldsymbol{F} \right) \\ & = P_R. \end{aligned} \tag{18b}$$

Due to the non-zero term $\sigma_{n_s}^2$ in (18a), it is difficult to find a feasible solution for $\boldsymbol{F}$ directly from (18a). To circumvent this problem, we define

$$\boldsymbol{F} \triangleq q\bar{\boldsymbol{F}} \tag{19}$$

and rewrite the original problem in (16) as (20), shown at the bottom of the next page. Due to the scaling factor q^{-1} in the received beamforming matrix $\boldsymbol{B}$, the first part of the cost function in (20) becomes independent of q; therefore, we can design

$$\arg \min_{\{\bar{\boldsymbol{F}}, \lambda\}} \underbrace{\mathrm{E} \left\| \boldsymbol{d} - \boldsymbol{P} \boldsymbol{B}^* \boldsymbol{H}^T \boldsymbol{F} \boldsymbol{H} \boldsymbol{A} \boldsymbol{d} \right\|^2 + \lambda \left(\mathrm{E} \left\| \boldsymbol{F} (\boldsymbol{H} \boldsymbol{A} \boldsymbol{d} + \boldsymbol{n}_s) \right\|^2 - P_R \right)}_{J}. \tag{16}$$

q from the second part of the cost, which relates to the power constraint. The modified cost J in (20) can be expanded as

$$
\begin{aligned}
J = {} & 2\sigma_d^2 K - \sigma_d^2 \operatorname{tr}\left(P\bar{B}^*H^T\bar{F}HA\right) \\
& - \sigma_d^2 \operatorname{tr}\left(A^*H^*\bar{F}^*(H^*)^T\bar{B}P\right) \\
& + \sigma_d^2 \operatorname{tr}\left(A^*H^*\bar{F}^*(H^*)^T\bar{B}\bar{B}^*H^T\bar{F}HA\right) \\
& + \lambda\left(q^2\sigma_d^2 \operatorname{tr}\left(A^*H^*\bar{F}^*\bar{F}HA\right)\right. \\
& \left. + q^2\sigma_{n_s}^2 \operatorname{tr}\left(\bar{F}^*\bar{F}\right) - P_R\right).
\end{aligned}
\tag{21}
$$

Now, equating the derivatives of the modified cost function J in (21) with respect to $\{\bar{F}, \lambda, q\}$ to zero, we get

$$
\begin{aligned}
\frac{\partial J}{\partial \bar{F}} = 0 \rightarrow {} & (H^*)^T\bar{B}\bar{B}^*H^T\bar{F}HAA^*H^* \\
& + \lambda q^2 \bar{F}\left(HAA^*H^* + \sigma_d^{-2}\sigma_{n_s}^2 I_M\right) \\
& = (H^*)^T\bar{B}PA^*H^*
\end{aligned}
\tag{22a}
$$

$$
\begin{aligned}
\frac{\partial J}{\partial \lambda} = 0 \rightarrow {} & q^2\left(\sigma_d^2 \operatorname{tr}\left(A^*H^*\bar{F}^*\bar{F}HA\right) + \sigma_{n_s}^2 \operatorname{tr}\left(\bar{F}^*\bar{F}\right)\right) \\
& = P_R
\end{aligned}
\tag{22b}
$$

$$
\begin{aligned}
\frac{\partial J}{\partial q} = 0 \rightarrow {} & 2\lambda q\left(\sigma_d^2 \operatorname{tr}\left(A^*H^*\bar{F}^*\bar{F}HA\right) + \sigma_{n_s}^2 \operatorname{tr}\left(\bar{F}^*\bar{F}\right)\right) \\
& = 0.
\end{aligned}
\tag{22c}
$$

Setting λ to zero, which satisfies (22c), we get from (22a) as

$$
\left(\bar{B}^*H^T\right)^*\left(\bar{B}^*H^T\right)\bar{F}(HA)(HA)^* = \left(\bar{B}^*H^T\right)^*P(HA)^*.
\tag{23}
$$

When $M < 2K$, $\bar{B}^*H^T \in \mathbb{C}^{2K \times M}$ and $HA \in \mathbb{C}^{M \times 2K}$ become tall and fat full-rank matrices, respectively, in which case we can get $\bar{F}$ from (23) as follows:

$$
\begin{aligned}
\bar{F} = {} & \left(\left(\bar{B}^*H^T\right)^*\left(\bar{B}^*H^T\right)\right)^{-1} \\
& \times \left(\bar{B}^*H^T\right)^* P (HA)^* \left((HA)(HA)^*\right)^{-1}.
\end{aligned}
\tag{24}
$$

On the other hand, when $M > 2K$, $\bar{B}^*H^T$ and HA become fat and tall full-rank matrices, respectively. In this case, (23) becomes under-determined over $\bar{F}$ and it can be solved to yield the following minimum norm solution for $\bar{F}$:

$$
\begin{aligned}
\bar{F} = {} & \left(\bar{B}^*H^T\right)^*\left(\left(\bar{B}^*H^T\right)\left(\bar{B}^*H^T\right)^*\right)^{-1}P \\
& \times \left((HA)^*(HA)\right)^{-1}(HA)^*.
\end{aligned}
\tag{25}
$$

Including the case $M = 2K$, (24) and (25) can be generally expressed as[2]

$$
\bar{F} = \left(\bar{B}^*H^T\right)^+ P(HA)^+
\tag{26}
$$

[2]After (52), we explain why the condition $M \geq 2K$ is a necessary condition to avoid CCI and SI; it is also a sufficient condition [28].

where $(W)^+ = (W^*W)^{-1}W^*$ if $W \in \mathbb{C}^{a \times b}$ is a tall or square matrix, i.e., $a \geq b$, and $(W)^+ = W^*(WW^*)^{-1}$ if W is a fat matrix, i.e., $a < b$. Using the linearity and the cyclic properties of the trace function, i.e., $\omega_1 \operatorname{tr}(W_1) + \omega_2 \operatorname{tr}(W_2) = \operatorname{tr}(\omega_1 W_1 + \omega_2 W_2)$ and $\operatorname{tr}(W_1 W_2) = \operatorname{tr}(W_2 W_1)$, respectively, we can get q satisfying (22b) as

$$
q = \sqrt{\frac{P_R}{\operatorname{tr}\left(\bar{F}\Omega\bar{F}^*\right)}}
\tag{27}
$$

where

$$
\Omega = \sigma_d^2 HAA^*H^* + \sigma_{n_s}^2 I_M.
\tag{28}
$$

Consequently, combining (26) and (27), the ZF-based relaying matrix satisfying (22a) can be obtained as

$$
\boxed{F_{ZF} = \sqrt{\frac{P_R}{\operatorname{tr}(\bar{F}\Omega\bar{F}^*)}} \cdot \bar{F}.}
\tag{29}
$$

C. MMSE-Based Design

Similarly, the optimization problem for minimizing the total MSE under the relay power constraint can be written as

$$
\boxed{\arg\min_{F} \operatorname{E}\left\|d - P\hat{d}\right\|^2 \quad \text{s.t.} \quad \operatorname{E}\|x\|^2 \leq P_R}
\tag{30}
$$

where we are now using $\hat{d}$ from (13) with noise instead of the $\hat{d}_f$ from (15). Similar to the ZF optimization procedure, the minimization problem (30) with constraint can be transformed into (31), shown at the bottom of this page, with a non-negative Lagrange multiplier λ. In (31), we modified the optimization problem with $\bar{F}$ as in the ZF case. The Lagrange cost J in (31) can be expanded as

$$
\begin{aligned}
J = {} & 2\sigma_d^2 K - \sigma_d^2 \operatorname{tr}\left(P\bar{B}^*H^T\bar{F}HA\right) - \sigma_d^2 \operatorname{tr}\left(A^*H^*\bar{F}^*\right. \\
& \times (H^*)^T\bar{B}P) + \sigma_d^2 \operatorname{tr}\left(A^*H^*\bar{F}^*(H^*)^T\bar{B}\bar{B}^*H^T\bar{F}\right. \\
& \times HA) + \sigma_{n_s}^2 \operatorname{tr}\left(\bar{F}^*(H^*)^T\bar{B}\bar{B}^*H^T\bar{F}\right) \\
& + q^{-2}\sigma_{n_x}^2 \operatorname{tr}\left(\bar{B}\bar{B}^*\right) + \lambda\left(\sigma_d^2 q^2 \operatorname{tr}\left(A^*H^*\bar{F}^*\bar{F}HA\right)\right. \\
& \left. + \sigma_{n_s}^2 q^2 \operatorname{tr}\left(\bar{F}^*\bar{F}\right) - P_R\right)
\end{aligned}
\tag{32}
$$

under the same assumption as in the ZF case. Setting the derivatives of J in (32) with respect to $\{\bar{F}, \lambda, q\}$ to zero, we get

$$
\begin{aligned}
\frac{\partial J}{\partial \bar{F}} = 0 \rightarrow {} & \left((H^*)^T\bar{B}\bar{B}^*H^T + \lambda q^2 I_M\right)\bar{F} \\
& \times \left(\sigma_d^2 HAA^*H^* + \sigma_{n_s}^2 I_M\right) \\
& = \sigma_d^2 (H^*)^T\bar{B}PA^*H^*
\end{aligned}
\tag{33a}
$$

$$
\arg\min_{\{\bar{F}, \lambda, q\}} \underbrace{\operatorname{E}\left\|d - q^{-1}P\bar{B}^*H^T q\bar{F}HAd\right\|^2 + \lambda\left(\operatorname{E}\left\|q\bar{F}(HAd + n_s)\right\|^2 - P_R\right)}_{J}.
\tag{20}
$$

$$
\arg\min_{\{\bar{F}, \lambda, q\}} \underbrace{\operatorname{E}\left\|d - P\left(\bar{B}^*H^T\bar{F}HAd + \bar{B}^*H^T\bar{F}n_s + q^{-1}\bar{B}^*n_x\right)\right\|^2 + \lambda\left(\operatorname{E}\left\|q\bar{F}(HAd + n_s)\right\|^2 - P_R\right)}_{J}.
\tag{31}
$$

$$\frac{\partial J}{\partial \lambda} = 0 \rightarrow q^2 \left(\sigma_d^2 \operatorname{tr} \left(\boldsymbol{A}^* \boldsymbol{H}^* \bar{\boldsymbol{F}}^* \bar{\boldsymbol{F}} \boldsymbol{H} \boldsymbol{A} \right) + \sigma_{n_s}^2 \operatorname{tr} \left(\bar{\boldsymbol{F}}^* \bar{\boldsymbol{F}} \right) \right) = P_R \tag{33b}$$

$$\frac{\partial J}{\partial q} = 0 \rightarrow q^4 = \sigma_{n_x}^2 \operatorname{tr}(\bar{\boldsymbol{B}}^* \bar{\boldsymbol{B}}) \lambda^{-1} \Big(\sigma_d^2 \operatorname{tr}(\boldsymbol{A}^* \boldsymbol{H}^* \bar{\boldsymbol{F}}^* \bar{\boldsymbol{F}} \boldsymbol{H} \boldsymbol{A})$$
$$+ \sigma_{n_s}^2 \operatorname{tr}(\bar{\boldsymbol{F}}^* \bar{\boldsymbol{F}}) \Big)^{-1}. \tag{33c}$$

Directly evaluating $\bar{\boldsymbol{F}}$ from (33a) is formidable due to the Lagrange multiplier λ. To avoid a numerical search over λ, we follow the optimization procedure in [29]. When $\lambda \neq 0$ and $\sigma_{n_s}^2 \neq 0$, $\bar{\boldsymbol{F}}$ in (33a) can be obtained as

$$\bar{\boldsymbol{F}} = \sigma_d^2 \left((\boldsymbol{H}^*)^T \bar{\boldsymbol{B}} \bar{\boldsymbol{B}}^* \boldsymbol{H}^T + \lambda q^2 \boldsymbol{I}_M \right)^{-1} (\boldsymbol{H}^*)^T \bar{\boldsymbol{B}} \boldsymbol{P} \boldsymbol{A}^* \boldsymbol{H}^*$$
$$\times \left(\sigma_d^2 \boldsymbol{H} \boldsymbol{A} \boldsymbol{A}^* \boldsymbol{H}^* + \sigma_{n_s}^2 \boldsymbol{I}_M \right)^{-1} \tag{34}$$

which is a function of $\lambda q^2 \triangleq \xi$. We note this fact explicitly by writing

$$\bar{\boldsymbol{F}} \triangleq \bar{\boldsymbol{F}}(\xi) \tag{35}$$

where

$$\bar{\boldsymbol{F}}(\xi) = \sigma_d^2 \boldsymbol{\Phi}^{-1}(\xi) \boldsymbol{\Psi} \boldsymbol{\Omega}^{-1} \tag{36}$$

and

$$\boldsymbol{\Phi}(\xi) = (\boldsymbol{H}^*)^T \bar{\boldsymbol{B}} \bar{\boldsymbol{B}}^* \boldsymbol{H}^T + \xi \boldsymbol{I}_M \tag{37a}$$
$$\boldsymbol{\Psi} = (\boldsymbol{H}^*)^T \bar{\boldsymbol{B}} \boldsymbol{P} \boldsymbol{A}^* \boldsymbol{H}^*. \tag{37b}$$

Substituting (35) into (33b), and using the cyclic property of the trace function, q is represented as

$$q = \sqrt{\frac{P_R}{\operatorname{tr} \left(\bar{\boldsymbol{F}}(\xi) \boldsymbol{\Omega} \bar{\boldsymbol{F}}^*(\xi) \right)}}. \tag{38}$$

Continuing from (35) and (38), when the conditions in (33a) and (33b) are satisfied, the problem in (31) can be rewritten as (39), shown at the bottom of the page. Here, we note that the second term multiplied by λ in (31) disappears due to (38) satisfying the power constraint (33b). Substituting (36) into (39), and using the cyclic property of the trace function and the property that $(\boldsymbol{W}^* \boldsymbol{W} + \omega \boldsymbol{I})^{-1} \boldsymbol{W}^* = \boldsymbol{W}^* (\boldsymbol{W} \boldsymbol{W}^* + \omega \boldsymbol{I})^{-1}$ in [30], we can write

$$J(\xi) = 2\sigma_d^2 K - 2\sigma_d^4 \operatorname{tr} \left(\boldsymbol{\Phi}^{-1}(\xi) \boldsymbol{\Psi} \boldsymbol{\Omega}^{-1} \boldsymbol{\Psi}^* \right)$$
$$+ \sigma_d^4 \sigma_{n_s}^2 \operatorname{tr} \left(\boldsymbol{\Phi}^{-2}(\xi) \boldsymbol{\Psi} \boldsymbol{\Omega}^{-2} \boldsymbol{\Psi}^* (\boldsymbol{H}^*)^T \bar{\boldsymbol{B}} \bar{\boldsymbol{B}}^* \boldsymbol{H}^T \right)$$
$$+ \sigma_d^6 \operatorname{tr} \left(\boldsymbol{\Phi}^{-2}(\xi) \boldsymbol{\Psi} \boldsymbol{\Omega}^{-1} \boldsymbol{H} \boldsymbol{A} \boldsymbol{A}^* \boldsymbol{H}^* \boldsymbol{\Omega}^{-1} \boldsymbol{\Psi}^* (\boldsymbol{H}^*)^T \right.$$
$$\left. \times \bar{\boldsymbol{B}} \bar{\boldsymbol{B}}^* \boldsymbol{H}^T \right) + \frac{\sigma_d^4 \sigma_{n_x}^2 \operatorname{tr}(\bar{\boldsymbol{B}} \bar{\boldsymbol{B}}^*)}{P_R}$$
$$\times \operatorname{tr} \left(\boldsymbol{\Phi}^{-2}(\xi) \boldsymbol{\Psi} \boldsymbol{\Omega}^{-1} \boldsymbol{\Psi}^* \right). \tag{40}$$

The derivative of (40) with respect to ξ is

$$\frac{\partial J(\xi)}{\partial \xi} = 2\sigma_d^4 \operatorname{tr} \left(\boldsymbol{\Phi}^{-2}(\xi) \boldsymbol{\Psi} \boldsymbol{\Omega}^{-1} \boldsymbol{\Psi}^* \right)$$
$$- 2\sigma_d^4 \sigma_{n_s}^2 \operatorname{tr} \left(\boldsymbol{\Phi}^{-3}(\xi) \boldsymbol{\Psi} \boldsymbol{\Omega}^{-2} \boldsymbol{\Psi}^* (\boldsymbol{H}^*)^T \bar{\boldsymbol{B}} \bar{\boldsymbol{B}}^* \boldsymbol{H}^T \right)$$
$$- 2\sigma_d^6 \operatorname{tr} \left(\boldsymbol{\Phi}^{-3}(\xi) \boldsymbol{\Psi} \boldsymbol{\Omega}^{-1} \boldsymbol{H} \boldsymbol{A} \boldsymbol{A}^* \boldsymbol{H}^* \boldsymbol{\Omega}^{-1} \boldsymbol{\Psi}^* \right.$$
$$\left. \times (\boldsymbol{H}^*)^T \bar{\boldsymbol{B}} \bar{\boldsymbol{B}}^* \boldsymbol{H}^T \right) - 2\sigma_d^4 \sigma_{n_x}^2 P_R^{-1} \operatorname{tr} \left(\bar{\boldsymbol{B}} \bar{\boldsymbol{B}}^* \right)$$
$$\times \operatorname{tr} \left(\boldsymbol{\Phi}^{-3}(\xi) \boldsymbol{\Psi} \boldsymbol{\Omega}^{-1} \boldsymbol{\Psi}^* \right). \tag{41}$$

Following the procedure outlined in Appendix A, (41) can be expressed as

$$\frac{\partial J(\xi)}{\partial \xi} = 2\sigma_d^4 \left(\xi - \sigma_{n_x}^2 P_R^{-1} \operatorname{tr} \left(\bar{\boldsymbol{B}} \bar{\boldsymbol{B}}^* \right) \right) \operatorname{tr} \left(\boldsymbol{\Phi}^{-3}(\xi) \boldsymbol{\Psi} \boldsymbol{\Omega}^{-1} \boldsymbol{\Psi}^* \right). \tag{42}$$

Since the cost in (39) is convex or strictly quasi-convex [31] with respect to ξ (see Appendix B), equating the derivative (42) to zero yields the optimal ξ_o as

$$\xi_o = 2\sigma_{n_x}^2 P_R^{-1} K. \tag{43}$$

The closed formed MMSE solution of $\boldsymbol{F}$ can be obtained from (35)–(38) and (43) as follows:

$$\boxed{\boldsymbol{F}_{\text{MMSE}} = \sqrt{\frac{P_R}{\operatorname{tr} \left(\bar{\boldsymbol{F}}(\xi_o) \boldsymbol{\Omega} \bar{\boldsymbol{F}}^*(\xi_o) \right)}} \cdot \bar{\boldsymbol{F}}(\xi_o).} \tag{44}$$

Note that the solution in (44) satisfies (33a), (33b), and (33c) [from (33b) and (33c), we can easily show (43)]. From this fact, we can see that (44) is the solution of the original optimization problem in (30).

D. Transmit and Receive-Beamforming Vectors

The relay needs to know the beamforming vectors $\{\boldsymbol{a}_k, \boldsymbol{b}_k\}$ as well as full CSI $\{\boldsymbol{H}_k\}$ to obtain the relay processing matrix. Fortunately, the relay is able to obtain beamforming vectors from the CSI. To begin with, since the two-way communication channel is assumed reciprocal, we assume that the normalized transmit and receive-beamforming vectors are also reciprocal, i.e., $\bar{\boldsymbol{a}}_k$ for $\boldsymbol{H}_k$ and $\bar{\boldsymbol{a}}_k^T$ for $\boldsymbol{H}_k^T$, so that

$$\boxed{\bar{\boldsymbol{b}}_k^* = \bar{\boldsymbol{a}}_k^T.} \tag{45}$$

For the normalized transmit and receive-beamforming vectors $\{\bar{\boldsymbol{a}}_k, \bar{\boldsymbol{b}}_k\}$, we employ four methods according to the required information at each user.

1) Eigen Beamforming (Perfect CSI): The eigen beamforming vector of the kth user is defined as a right singular

---

$$\arg \min_{\xi} \underbrace{\operatorname{E} \left\| \boldsymbol{d} - \boldsymbol{P} \bar{\boldsymbol{B}}^* \left(\boldsymbol{H}^T \bar{\boldsymbol{F}}(\xi) \boldsymbol{H} \boldsymbol{A} \boldsymbol{d} + \boldsymbol{H}^T \bar{\boldsymbol{F}}(\xi) \boldsymbol{n}_s + \sigma_d^2 P_R^{-1/2} \sqrt{\operatorname{tr} \left(\boldsymbol{\Psi}^* \boldsymbol{\Phi}^{-2}(\xi) \boldsymbol{\Psi} \boldsymbol{\Omega}^{-1} \right)} \boldsymbol{n}_x \right) \right\|^2}_{J(\xi)}. \tag{39}$$

vector corresponding to the largest singular value of $\boldsymbol{H}_k$[22]. Therefore, the kth user needs to know its channel $\boldsymbol{H}_k$ to generate the eigen beamforming vector.

2) Antenna Selection (Partial CSI): If the nth antenna of the kth user is selected for communication, the transmit beamforming vector is an N_k-dimensional vector whose nth element is 1 and the rest are 0s. Due to the reciprocity of the transmit and receive-beamforming vectors, the same antenna of each user is selected for transmission and reception in the first and second phases, respectively. Norm-based selection is employed as a selection criterion [23], i.e., n is $\arg\max_n \|[\boldsymbol{H}_k]_n\|$. Consequently, the kth user requires only selected antenna index n as partial CSI, which can be fed back from the relay.

3) Random Beamforming (No CSI): A simple beamforming choice without CSI at the users is a random beamforming [24]. It can achieve multiuser diversity efficiently when there is no CCIs as in ZF-based multiuser systems. The random beamforming vector is a eigen beamformer generated from an arbitrary random matrix being the same size and the same distribution as $\boldsymbol{H}_k$. Therefore, the users do not need to know the CSI. However, to generate the same random beamformer at the relay, a common signal between the users and the relay, such as a synch information, is required.

4) Equal Gain Beamforming (No CSI): Equal gain beamforming is obtained as

$$\bar{\boldsymbol{a}}_k = (1/\sqrt{N_k}) \begin{bmatrix} e^{j\theta_{k,1}} & \cdots & e^{j\theta_{k,N_k}} \end{bmatrix}^T$$

where $\theta_{k,n}$ can be designed from the CSI to obtain diversity gain [25]. For comparison with the spatial multiplexing system later, we assume that $\theta_{k,n} = 0$. Though this assumption does not provide spatial diversity gain, multiuser diversity gain can be expected as random beamforming and CSI is not required for the users. However, in contrast to the random beamforming system, extra information is not required for the relay to generate the beamforming vectors.

Remarks

- *Remark 1*: The inequality power constraint in (30) can be reduced to an equality power constraint. Since ξ in (43) is positive, $\lambda = \xi/q^2$ is also positive, whereby $\mathrm{E}\left\|q\bar{\boldsymbol{F}}(\boldsymbol{HAd} + \boldsymbol{n}_s)\right\|^2 = P_R$ in (31). This result is reasonable due to the reciprocity of the two-way communications. This means that there would be no power saving effect [9] under an inequality power constraint; as a result, the inequality power constraint can be substituted by the equality power constraints in (30).
- *Remark 2*: When $\sigma_{n_x}^2 = 0$, ξ in (43) is zero and also $\lambda = 0$ since $q > 0$ in $\xi = \lambda q^2$. In addition, when $\sigma_{n_s}^2 = 0$, the equality (33a) becomes identical to (23). From these and the fact that (33b) is identical to (22b), we can see that the MMSE solution in (44) is reduced to the ZF solution in (29) under the high SNR assumption.
- *Remark 3*: If each user has just one antenna, i.e., no beamforming, in the proposed $2K$-user two-way system, this system can be modeled equivalently as a two-user two-way system, where each user has K antennas and transmits independent K data streams by spatial multiplexing. There-

fore, the two-user two-way relay system in [13] exchanging the spatially multiplexed K data can be interpreted as a special case of the proposed method under $2K$ users having one antenna without power control, i.e., $\boldsymbol{\Gamma} = \boldsymbol{I}_{2K}$. In other words, (29) and (44) degenerate to the ZF and the MMSE solution in [13], respectively, when $N_k = 1$, $k \in \{1,\ldots,K\}$ and $\boldsymbol{\Gamma} = \boldsymbol{I}_{2K}$ (see Appendix C).

- *Remark 4*: For the multiuser two-way communications, additional resources, such as frequency bandwidth, orthogonal codes, and time slots (phases), are required to achieve the CSI and the equalization factors (or power control factors shown later). First, in order to obtain the multiuser CSI at the relay, the training sequences have to be transmitted from every user to the relay through the first phase. Meanwhile, the multiuser training sequences should be decomposable without the CCI, and an orthogonality is required among the multiuser training sequences. This orthogonality can be obtained by using additional resources mentioned before. Second, for the users, the relay broadcasts a common training signal through the second phase, whereby each user can estimate their own channel $\boldsymbol{H}_k^T$ if it is required. Lastly, since q in (27) and (38) is generated from the multiuser channels with a relay noise variance, the users cannot compute it, thereby causing it to be fed back from the relay.

IV. POWER CONTROL FOR ZF SYSTEMS

We now propose two power control methods for the ZF-based system. For the transmit power of the relay and users, local and global power control methods are proposed. With the network total power P_T defined as

$$P_T = P_U + P_R. \tag{46}$$

In (46), the multiuser power P_U is as follows (see Fig. 4):

$$P_U = \mathrm{E}\|\boldsymbol{s}\|_2^2 = \sigma_d^2 \,\mathrm{tr}\,(\boldsymbol{A}^*\boldsymbol{A}) = \sigma_d^2 \,\mathrm{tr}\,\left(\boldsymbol{\Gamma}\bar{\boldsymbol{A}}^*\bar{\boldsymbol{A}}\boldsymbol{\Gamma}\right)$$
$$= \sigma_d^2 \sum_{k=1}^{2K} \gamma_k^2 \|\bar{\boldsymbol{a}}_k\|^2 = 2\sigma_d^2 K. \tag{47}$$

The local power control divides the multiuser power P_U among users according to their channel gains, to yield the same average SNR for each user resulting in fairness among all users. In other words, the local power control will decide the user power ratio γ_k in (47) in order to guarantee fairness among users with some sacrifice in performance. On the other hand, the global power control will divide the network total power P_T between the relay and the multiuser into P_R and P_U, respectively, to maximize the system SNR under the high SNR assumption. Namely, the global power control decides σ_d^2 in (47).

A. Local Power Control

Since the average channel gains $\{\sigma_{H,k}\}$ in (1) are different for different users, in practice, the effective received SNR at each user is also different. For fairness among the users, we propose a local power control method that leads to the same average SNR at the users by adjusting the power ratio of the transmit beamforming vectors $\{\boldsymbol{a}_k\}$, i.e., γ_k. Due to the reciprocity of
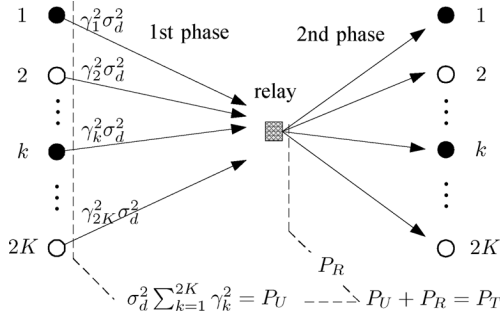
Fig. 4. Transmit power of $2K$-users and one-relay node in multiuser two-way relay system.

two-way communications, the same average SNR at the relay from each user yields the same average SNR at the users as follows. From (3), the effective SNR at the relay for the kth user and for the given channel is given by

$$\mathrm{SNR}_{R,k} \triangleq \frac{\mathrm{E}\,\|\boldsymbol{H}_k \boldsymbol{a}_k d_k\|^2}{\mathrm{E}\,\|\boldsymbol{n}_s\|^2} = \frac{\gamma_k^2 \sigma_d^2 \sigma_{H,k}^2 \bar{\boldsymbol{a}}_k^* \bar{\boldsymbol{H}}_k^* \bar{\boldsymbol{H}}_k \bar{\boldsymbol{a}}_k}{M \sigma_{n_s}^2}. \quad (48)$$

The average SNR for the kth user is therefore

$$\overline{\mathrm{SNR}}_{R,k} \triangleq \frac{\mathrm{E}\left(\gamma_k^2 \sigma_d^2 \sigma_{H,k}^2 \bar{\boldsymbol{a}}_k^* \bar{\boldsymbol{H}}_k^* \bar{\boldsymbol{H}}_k \bar{\boldsymbol{a}}_k\right)}{M \sigma_{n_s}^2} = \frac{\gamma_k^2 \sigma_d^2 \sigma_{H,k}^2 \rho_k}{M \sigma_{n_s}^2} \quad (49)$$

where

$$\rho_k \triangleq \mathrm{E}\left(\bar{\boldsymbol{a}}_k^* \bar{\boldsymbol{H}}_k^* \bar{\boldsymbol{H}}_k \bar{\boldsymbol{a}}_k\right). \quad (50)$$

From (49), we can select γ_k to ensure a desired average SNR as $\overline{\mathrm{SNR}}_{R,k} = \epsilon$ as

$$\gamma_k = \sqrt{\frac{\epsilon M \sigma_{n_s}^2}{\rho_k \sigma_d^2 \sigma_{H,k}^2}}.$$

Using the normalization condition (9) we arrive at

$$\boxed{\gamma_k = \sqrt{\frac{2K}{\rho_k \sigma_{H,k}^2 \sum\limits_{n=1}^{2K} \rho_n^{-1} \sigma_{H,n}^{-2}}}.} \quad (51)$$

These values of γ_k ensure that the average SNR at the relay is the same for all users. Now, we verify that we can also achieve the same average SNR at the users through (51). Indeed, from (12) and (19), the multiuser received signal model (13) is rewritten as

$$\hat{\boldsymbol{d}} = \bar{\boldsymbol{B}}^* \boldsymbol{H}^T \bar{\boldsymbol{F}} \boldsymbol{H} \boldsymbol{A} \boldsymbol{d} + \bar{\boldsymbol{B}}^* \boldsymbol{H}^T \bar{\boldsymbol{F}} \boldsymbol{n}_s + q^{-1} \bar{\boldsymbol{B}}^* \boldsymbol{n}_x. \quad (52)$$

From (52), it can be easily verified that the ZF-based solution $\bar{\boldsymbol{F}}$ from (25) ensures an interference-free path from $\boldsymbol{d}$ to $\hat{\boldsymbol{d}}$ as follows:

$$\hat{\boldsymbol{d}} = \boldsymbol{P} \boldsymbol{d} + \boldsymbol{P} \left(\boldsymbol{H} \bar{\boldsymbol{A}} \boldsymbol{\Gamma}\right)^+ \boldsymbol{n}_s + q^{-1} \bar{\boldsymbol{B}}^* \boldsymbol{n}_x. \quad (53)$$

From this reason, we shall now assume that the number of relay antennas is greater than or equal to the number of users, i.e.,

$$M \geq 2K. \quad (54)$$

For the $(2k)$th user, the received signal model is then written from (53) as follows:

$$\hat{d}_{(2k)'} = \hat{d}_{2k-1} = d_{2k-1} + \left[\left(\boldsymbol{H} \bar{\boldsymbol{A}} \boldsymbol{\Gamma}\right)^+\right]_{2k-1}^T \boldsymbol{n}_s + q^{-1} \bar{\boldsymbol{b}}_{2k}^* \boldsymbol{n}_{x,2k}. \quad (55)$$

From (55), the $(2k)$th user's SNR for a given channel is

$$\mathrm{SNR}_{U,2k} \triangleq \frac{\mathrm{E}\,|d_{2k-1}|^2}{\mathrm{E}\left|\left[\left(\boldsymbol{H} \bar{\boldsymbol{A}} \boldsymbol{\Gamma}\right)^+\right]_{2k-1}^T \boldsymbol{n}_s + q^{-1} \bar{\boldsymbol{b}}_{2k}^* \boldsymbol{n}_{x,2k}\right|^2}. \quad (56)$$

Similarly, $\hat{d}_{(2k-1)'}$ and $\mathrm{SNR}_{U,2k-1}$ can be obtained. For the $(2k)$th user SNR in (56) we have

$$\mathrm{SNR}_{U,2k}$$
$$= \frac{\sigma_d^2}{\sigma_{n_s}^2 \,\mathrm{tr}\left(\left(\left[\left(\boldsymbol{H} \bar{\boldsymbol{A}} \boldsymbol{\Gamma}\right)^+\right]_{2k-1}^T\right)^* \left[\left(\boldsymbol{H} \bar{\boldsymbol{A}} \boldsymbol{\Gamma}\right)^+\right]_{2k-1}^T\right) + q^{-2} \sigma_{n_x}^2} \quad (57)$$

so that the average SNR in (57) over all channels is

$$\overline{\mathrm{SNR}}_{U,2k} \triangleq \frac{\sigma_d^2}{\mathrm{E}\left(\sigma_{n_s}^2 \varrho_{2k-1} + q^{-2} \sigma_{n_x}^2\right)}$$
$$= \frac{\sigma_d^2}{\sigma_{n_s}^2 \,\mathrm{E}(\varrho_{2k-1}) + q^{-2} \sigma_{n_x}^2} \quad (58)$$

where $\varrho_{2k-1} = \left[\left(\boldsymbol{H} \bar{\boldsymbol{A}} \boldsymbol{\Gamma}\right)^+\right]_{2k-1}^* \left[\left(\boldsymbol{H} \bar{\boldsymbol{A}} \boldsymbol{\Gamma}\right)^+\right]_{2k-1}$. Using $\boldsymbol{\Gamma}$ in (51), we can rewrite ϱ_{2k-1} as

$$\varrho_{2k-1} = \left[\left(\bar{\boldsymbol{H}} \bar{\boldsymbol{A}} \bar{\boldsymbol{\Gamma}}\right)^+\right]_{2k-1}^* \left[\left(\bar{\boldsymbol{H}} \bar{\boldsymbol{A}} \bar{\boldsymbol{\Gamma}}\right)^+\right]_{2k-1} \quad (59)$$

where $\bar{\boldsymbol{\Gamma}}$ is a diagonal matrix whose kth diagonal element is

$$\bar{\gamma}_k = \sqrt{\frac{2K}{\rho_k \sum\limits_{n=1}^{2K} \rho_n^{-1} \sigma_{H,n}^{-2}}}. \quad (60)$$

In (60), if $\rho_k = \rho$ for all users k, then every user achieves the same average SNR in (58). Actually, when every user employs the same beamforming method with the same number of antennas ($N_k = N$), then we can set $\rho_k = \rho$. To see this, when eigen beamforming is used, ρ_k in (50) is an expectation of the square of the largest singular value of $\bar{\boldsymbol{H}}_k$; when antenna selection is performed, ρ_k is an expectation of the largest norm among column vectors of $\bar{\boldsymbol{H}}_k$; and when equal gain beamforming is used $\rho_k = (1/N)\,\mathrm{E}\,\|\bar{\boldsymbol{H}}_k[1 \cdots 1]^T\|^2 = N$. Since the dimension and distribution of $\bar{\boldsymbol{H}}_k \in \mathbb{C}^{M \times N}$ are independent of k, we get $\rho_k = \rho$. Now, when $\rho_k = \rho$, we can also set $\bar{\gamma}_k = \bar{\gamma}$ in (60) and express ϱ_{2k-1} in (59) as

$$\varrho_{2k-1} = \bar{\gamma}^{-2} \left[\left(\bar{\boldsymbol{H}} \bar{\boldsymbol{A}}\right)^+\right]_{2k-1}^* \left[\left(\bar{\boldsymbol{H}} \bar{\boldsymbol{A}}\right)^+\right]_{2k-1}. \quad (61)$$

Since $\mathrm{E}\left(\left[(\bar{H}\bar{A})^+\right]_{2k-1}^* \left[(\bar{H}\bar{A})^+\right]_{2k-1}\right)$ from (61) is independent of k, the denominator of the last equality in (58) is also independent of k. Consequently, we achieve the same average SNR at the users, resulting in fairness among users.

B. Global Power Control

The global power control is performed to achieve the maximum system SNR by allocating P_U and P_R for the users and the relay, respectively. From (53), the system SNR is defined as

$$\mathrm{SNR}_{\mathrm{sys}} \triangleq \frac{\mathrm{E}\,\|\boldsymbol{P}\boldsymbol{d}\|^2}{\mathrm{E}\left\|\boldsymbol{P}\left(\boldsymbol{H}\bar{\boldsymbol{A}}\boldsymbol{\Gamma}\right)^+ \boldsymbol{n}_s + q^{-1}\bar{\boldsymbol{B}}^*\boldsymbol{n}_x\right\|^2}. \quad (62)$$

The system SNR refers to the average SNR over the users. Using the symmetry of the beamforming vectors in (45), the system SNR (62) can be approximated under a high SNR assumption as (see Appendix D):

$$\mathrm{SNR}_{\mathrm{sys}} \simeq \frac{P_U P_R}{\mu_1 \sigma_{n_s}^2 P_R + \mu_2 \sigma_{n_x}^2 P_U} \quad (63)$$

where

$$\begin{aligned} \mu_1 &= \mathrm{tr}\left(\left(\boldsymbol{\Gamma}\bar{\boldsymbol{A}}^*\boldsymbol{H}^*\boldsymbol{H}\bar{\boldsymbol{A}}\boldsymbol{\Gamma}\right)^{-1}\right) \\ \mu_2 &= \mathrm{tr}\left(\left(\bar{\boldsymbol{A}}^*\boldsymbol{H}^*\boldsymbol{H}\bar{\boldsymbol{A}}\right)^{-1}\right). \end{aligned} \quad (64)$$

Denoting the received SNRs at the relay and at the users by

$$\begin{aligned} \mathrm{SNR}_{RU} &\triangleq \frac{P_U}{\sigma_{n_s}^2} \\ \mathrm{SNR}_{UR} &\triangleq \frac{P_R}{\sigma_{n_x}^2} \end{aligned} \quad (65)$$

respectively, the system SNR in (63) can be expressed as

$$\mathrm{SNR}_{\mathrm{sys}} = \frac{\mathrm{SNR}_{UR}\mathrm{SNR}_{RU}}{\mu_1 \mathrm{SNR}_{UR} + \mu_2 \mathrm{SNR}_{RU}}$$

and its bound can be derived as follows:

$$\mathrm{SNR}_{\mathrm{sys}} \leq \sqrt{\frac{\mathrm{SNR}_{RU}\mathrm{SNR}_{UR}}{4\mu_1\mu_2}}. \quad (66)$$

The inequality in (66) comes from the inequality between arithmetic and harmonic means, i.e., $2/(w_1 + w_2) \leq (w_1 + w_2)/2w_1w_2$. Since the equality holds when $w_1 = w_2$, the system SNR bound is achieved when

$$\mu_1 \mathrm{SNR}_{UR} = \mu_2 \mathrm{SNR}_{RU}. \quad (67)$$

Consequently, from (46), (65), and (67), the global power control is performed as

$$\boxed{P_U = \frac{\mu_1 \sigma_{n_s}^2 P_T}{\mu_1 \sigma_{n_s}^2 + \mu_2 \sigma_{n_x}^2} \quad \text{and} \quad P_R = \frac{\mu_2 \sigma_{n_x}^2 P_T}{\mu_1 \sigma_{n_s}^2 + \mu_2 \sigma_{n_x}^2}.} \quad (68)$$

Using the global power control (68), the SNRs between two hops are balanced, resulting in the system SNR being maximized. This balancing condition is reported in [9], where the balancing minimizes the system MSE.

Remarks

- *Remark 5*: The power control is initiated by the relay due to the convenience of gathering the required CSIs for achieving power control factors. We can implement four different power control methods as follows:
  — none: $\gamma_k = 1$ for all k. $P_R = P_U = P_T/2$, resulting in $\sigma_d^2 = P_T/4K$.
  — local: $\{\gamma_k\}$ are obtained from (51). $P_R = P_U = P_T/2$, resulting in $\sigma_d^2 = P_T/4K$.
  — global: First, $\gamma_k = 1$ for all k. Next, $\{P_R, P_U\}$ are obtained from (64) and (68).
  — local + global: First, $\{\gamma_k\}$ are obtained from (51). Next, $\{P_R, P_U\}$ are obtained from (64) and (68).
  After computing the power distribution of the two-way relay network, the relay broadcasts the power control information $\{\gamma_k\}$ and P_U (or σ_d^2) to the users.

- *Remark 6*: The proposed power control methods match well for the MMSE-based relay system under high SNR since MMSE and ZF solutions become identical as SNR increases as mentioned in *Remark 2*. This is verified by simulation later.

- *Remark 7*: Eigen beamforming at all users maximizes the upper bound of system SNR in (66) for the multiuser two-way relay communications (see Appendix E). Although eigen beamforming may not be optimal in the sense of system performance due to the tightness of the upper bound of the system SNR, simulation results indicate that eigen beamforming can achieve the best performance compared the other methods in Section III-E.

- *Remark 8*: The performance of the ZF relay is worse than that of the MMSE relay since it does not consider noise effect. However, the ZF solution has some advantages such as requiring less information and the convenient design of resource allocation. The ZF relay does not require noise variance information to achieve the relay processing matrix. Furthermore, it is convenient for the ZF systems to design power control methods as derived in this section and in [21] and also to develop scheduling algorithms, since the multiuser channels are completely orthogonalized resulting in uncoupled problems of resource allocation.

V. SIMULATION AND DISCUSSION

Computer simulations are conducted to examine the performance of the designed systems. In the first set of simulation, the performance of the proposed system is evaluated and discussed when $\{\sigma_{H,k}^2\}$ are the same. Then, in the second set, the behavior of the proposed power control methods is examined when the user channel gains $\{\sigma_{H,k}^2\}$ are different.

A. Performance Evaluation

In this subsection, system SNR, sum rate, and system BER are evaluated when every channel gain of users is identical. The sum rate is $\sum_{k=1}^{2K} \mathcal{R}_k$, where the rate of the kth user is defined as $\mathcal{R}_k \triangleq \mathrm{E}\log_2\left(1 + \mathrm{SNR}_{U,k}\right)$ by using a point-to-point communication rate [33]. Here, $\mathrm{SNR}_{U,k}$ is the single user (stream) SNR obtained from (6). The system BER is defined as the BER averaged over the users. The modulated symbols are grouped into frames consisting of 100 symbols. For each frame of the kth user, a flat fading MIMO channel matrix $\boldsymbol{H}_k$ is generated from independent Gaussian random variables with zero mean and unit variance (i.e., $\sigma_{H,k}^2 = 1$, for all k). Channels $\{\boldsymbol{H}_k\}$ are fixed
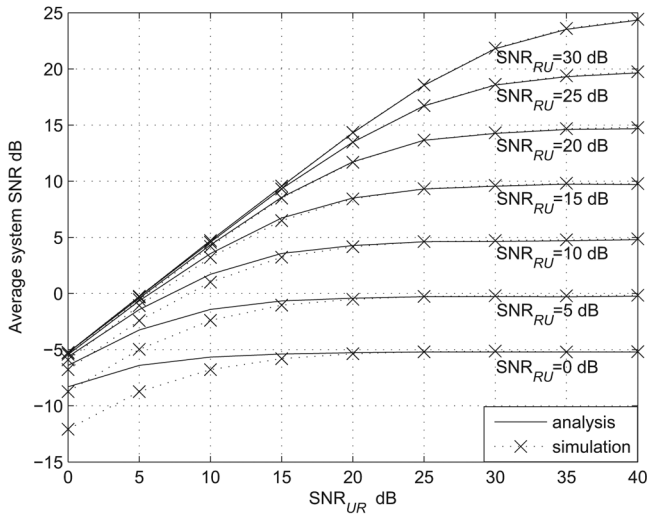
Fig. 5. The average system SNR simulation from (62) and its analysis in (63) when $N_k = 2$, $2K = 4$, and $M = 4$.

Fig. 6. The average system SNR in (63) and its bound in (66) when $N_k = 2$, $2K = 4$, $M = 4$, and $\text{SNR}_{RU} = 25$ dB.

during two consecutive phases, but they vary independently over two phases. The results shown below are the averages over 10^5 independent trials. The number of relay antennas is M and it is greater than or equal to the number of users, $2K$, while there is no restriction on the number of user antennas $\{N_k\}$.

Fig. 5 shows the average system SNRs obtained from the analysis in (63) and simulation when $N_k = 2$, $2K = 4$, $M = 4$, $\sigma_d^2 = 1$, and $P_R = 1$. Since the system SNR in (63) is derived under the assumption of high SNR, a remarkably good agreement is observed between the analysis and the simulation especially when the SNRs are high.

In Fig. 6, the average system SNR in (63) and its bound in (66) are compared when $\text{SNR}_{RU} = 25$ dB, $N_k = 2$, $2K = 4$, and $M = 4$. Various beamformings, such as eigen beamforming, antenna selection, random beamforming, and equal gain beamforming are employed at the users. Here, the effect of the global power control can be verified. Obviously, when $\text{SNR}_{UR} < \text{SNR}_{RU}$ the system SNR increases as SNR_{UR} increases. As a result, the system SNR bound is almost perfectly achieved when $\text{SNR}_{RU} = \text{SNR}_{UR}$, i.e., 25 dB in our simulation. This is true because $\mu_1 = \mu_2$ in (67) in this simulation environment [note that $\mathbf{\Gamma} = \mathbf{I}_{2K}$ since $\sigma_{H,k}^2 = 1$ and ρ_k is an invariable for all k in (51)]. The average system SNR of the eigen beamforming system is the largest; however, perfect CSI is required at the users as well as the equalization factor. The performance of the antenna selection method using partial CSI is between that of eigen beamforming using the perfect CSI and of the other methods using no CSI. For the equal gain and random beamforming methods, only the equalization factor is needed for the users, and simple implementation is possible. It can be also seen that the average system SNR of the equal gain and the random beamforming systems are almost identical. Therefore, the equal gain beamforming system is more attractive than the random beamforming system since its relay can easily achieve the beamforming vectors as mentioned before.

Fig. 7 shows the sum rate. Interestingly, the sum rate with a fixed number of relay antennas increases as the number of users
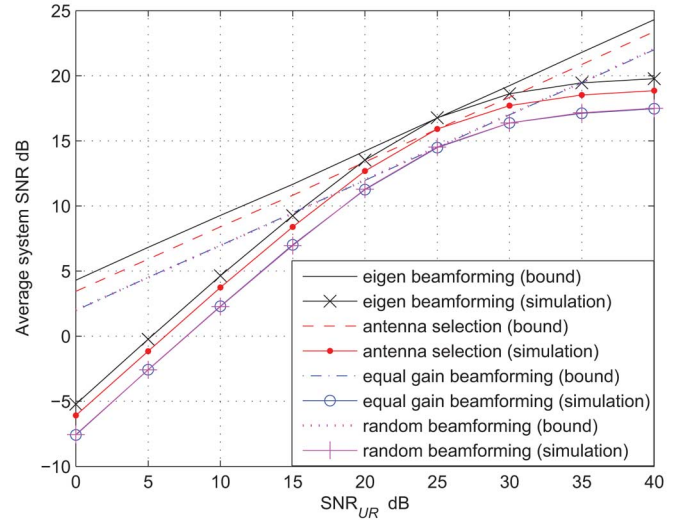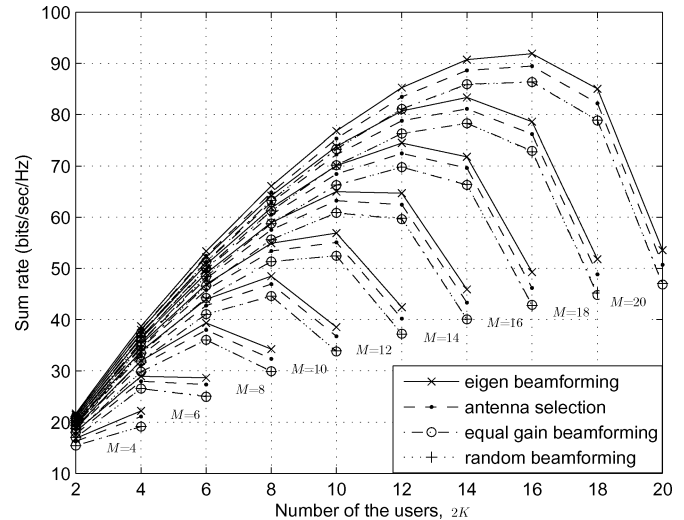

Fig. 7. Sum rate when $N_k = 2$ and $\text{SNR}_{RU} = \text{SNR}_{UR} = 25$ dB.

increases up to a certain number. This result is reasonable since the system SNR decreases due to the increased CCIs. Consequently, an optimal supportable number of users, with respect to the sum rate, can be determined with given parameters, such as the number of antennas and the average received SNRs. As an example, when $M = 14$, the optimal supportable number of users is 10, although the maximum number of the supportable users is 14 from (54).

In Fig. 8, the system BER performance is evaluated over the number of users when $\text{SNR}_{UR} = \text{SNR}_{RU} = 15$ dB, $N_k = 2$, and quadrature PSK (QPSK) modulation is used. As expected, the eigen beamforming method is superior to the other methods; the MMSE relay performs better than the ZF relay; and the BER performance becomes better as the number of the relay antennas increases. It is further observed that the performance gap between the ZF and MMSE relays increases as the number of users increases. In addition, the system BERs becomes worse as the number of users $2K$ increases due to the increase of the CCIs; however, saturation of BERs is observed at a certain point. From
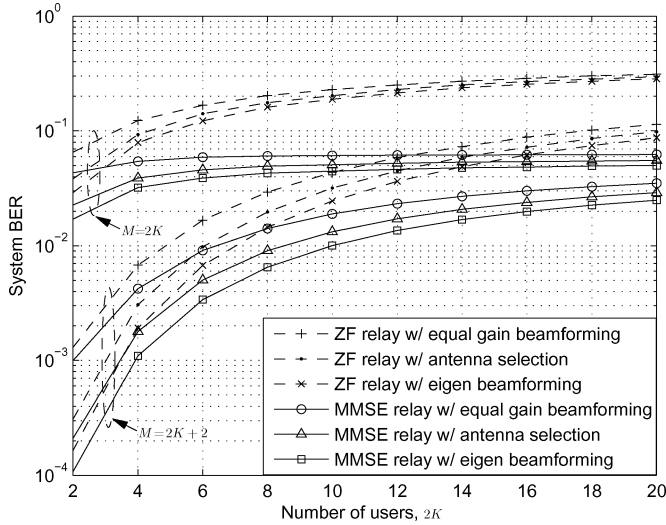
Fig. 8. System BER over the number of users when $N_k = 2$ and $\mathrm{SNR}_{RU} = \mathrm{SNR}_{UR} = 15$ dB.

Fig. 9. System BER comparison of beamforming and spatial multiplexing when $\mathrm{SNR}_{RU} = 25$ dB and $2K = 2$.

this observation, it can be surmised that CCIs and SIs are canceled efficiently though the number of users increases.

In Fig. 9, the BER performance of the proposed beamforming systems and the spatial multiplexing system is compared when $\mathrm{SNR}_{RU} = 25$ dB. For the beamformers, we consider equal gain beamforming as well as eigen beamforming. The equal gain beamforming requires the same quantity of information as the spatial multiplexing system, i.e., no CSI at the users and full CSI at the relay. For comparing with the conventional two-user two-way multiplexing systems [13], we fix the total number of users at two and $M = \sum_{k=1}^{k=2K} N_k$. Generally, the performance gap between the ZF and MMSE relays does not decrease as SNR_{UR} increases since SNR_{RU} is fixed by 25 dB. When $N_k = 1$, the proposed systems become identical to the spatial multiplexing systems, as mentioned in *Remark 3*. Here, every user transmits binary PSK (BPSK) symbols without spatial multiplexing and beamforming. On the other hand, when $N_k = 2$, considerable difference between the spatial multiplexing and the beamforming systems is observed. Here, since spatial multiplexing with two transmit antennas can transmit two different data symbols simultaneously, the BPSK and QPSK modulations are used for the spatial multiplexing and beamforming systems, respectively, for a fair comparison. The poor performance of the spatial multiplexing systems comes from the fact that the spatially multiplexed signals interfere with each other through the AF relay processing [9], particularly in the interference limited two-way systems. From this result, we can surmise that the beamforming method at the user is proper than the spatial multiplexing method for the two-way relay system.

B. Power Control Comparison

For comparison of the power control methods, we set the simulation environment as follows: number of users is 20; channel gain of kth user is defined by $\sigma_{H,k}^2 \triangleq k$; the noise variance of the relay is four times less than that of the users, i.e., $\sigma_{n_s}^2 = \sigma_{n_x}^2/4$; every user has two antennas; the relay has 20 antennas; total system power is 100; and the eigen beamforming is used for all users. The results are the averages over 10^7 independent
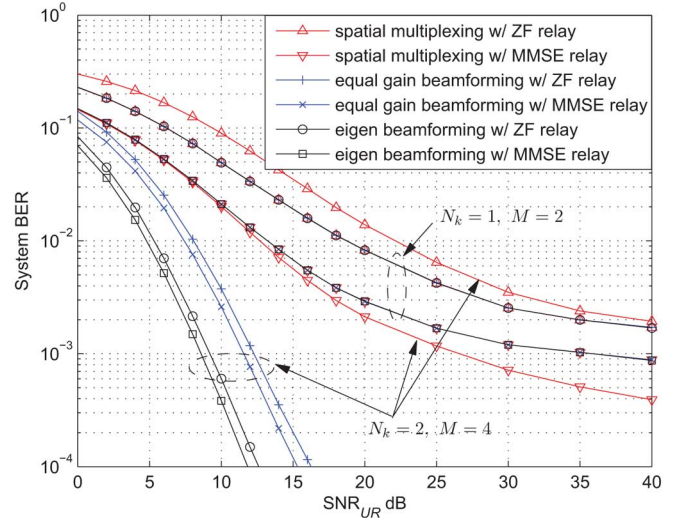
trials. In the simulation, the fairness index of [34] was used to compare fairness among users. The index is bounded between 0 and 1. The higher index represents higher fairness among users. When there are K users, the fairness index is obtained by $\left(\sum_{k=1}^{K} r_k\right)^2 \left(K \sum_{k=1}^{K} r_k^2\right)^{-1}$, where the normalized throughput $r_k = T_k/O_k$, T_k is the measured throughput, and O_k is the fair throughput, for the kth user. Here, we assumed that all O_k are identical for all k, i.e., traffic characteristics for each user are identical and required rates for each user are identical also.

In Fig. 10, the sum rate and the fairness index are examined for the ZF- and MMSE-based two-way relay systems with four power control methods in *Remark 5*. Fig. 10(a)–(d) and (e)–(h) illustrates the results for ZF and MMSE systems, respectively. From these results, we can observe that the local power control method can guarantee fairness among users with a certain sacrifice of the sum rate, regardless of the relaying method or SNR region. In the low SNR region, i.e., Fig. 10(e) and (f), it is observed that the global power control does not work for the MMSE system. However, the global power control yields the largest sum rate for the ZF system as we designed. Also, it can be verified that the ZF and MMSE methods yield similar performance in high SNR region from Fig. 10(c), (d), (g), and (h). Consequently, the power control methods can be determined according to the system environment and requirement such as fairness and system performance.

VI. CONCLUSION

In this paper, we proposed relay transceiver processing for multiuser two-way relay systems. The optimal transceiver processing matrices were designed based on both the ZF and MMSE criteria under the assumption that every user employs spatial beamforming. For ZF systems, local and global power control methods were proposed to achieve fairness among all users and maximum system SNR. The average system SNR, sum rate, and BER were evaluated with various beamforming methods. The performance of the power control methods were also simulated with respect to the system sum rate and fairness
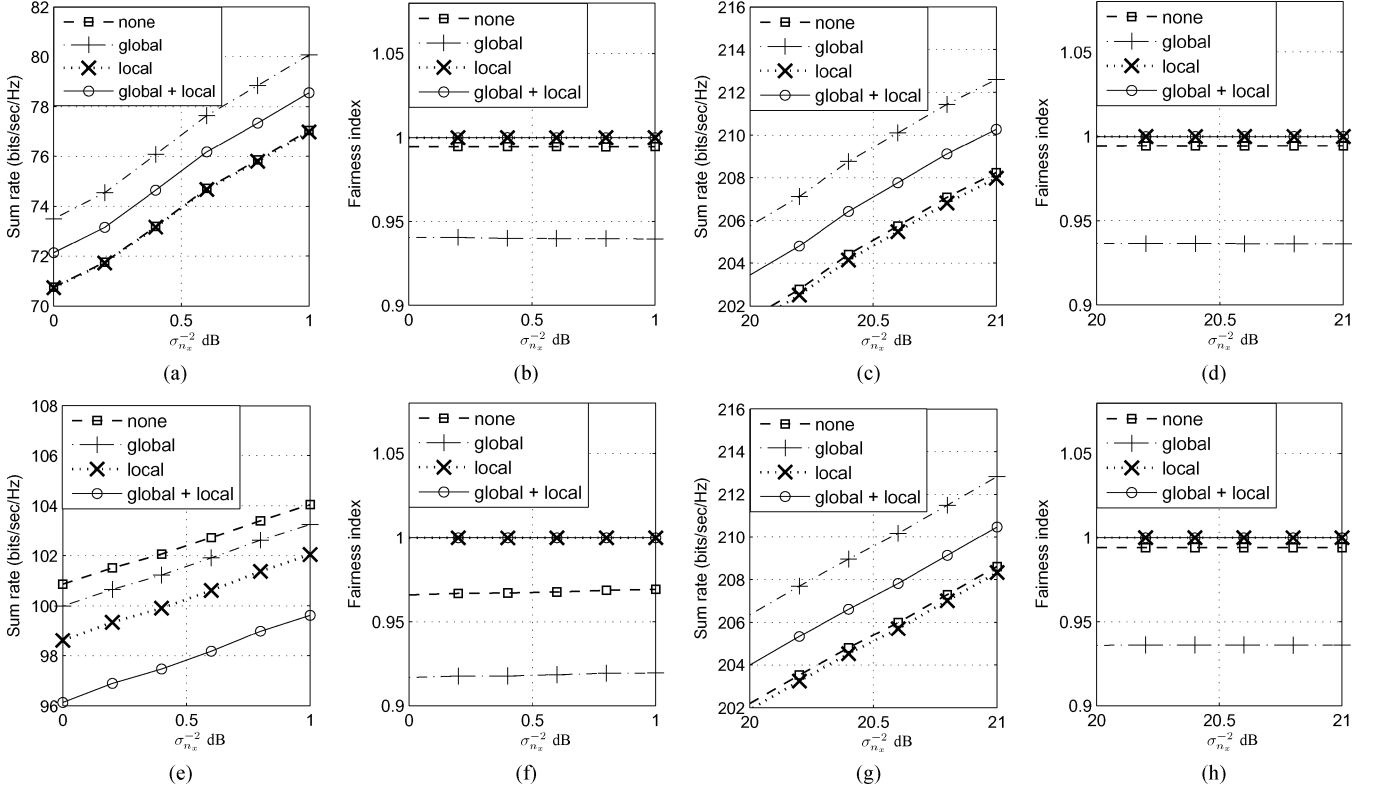
Fig. 10.   Comparison among power control schemes when $2K = 20$, $N_k = 2$, $M = 20$, $P_T = 100$, and eigen beamforming is employed. (a)–(d) Rate and fairness comparison of ZF-based methods. (e)–(h) Rate and fairness comparison of MMSE-based methods.

among users. Consequently, it was verified that the proposed beamforming systems with multiuser two-way AF relay can efficiently cancel out the CCIs and SIs and the proposed power control methods can effectively serve their purpose.

APPENDIX A
DERIVATION OF (42)

Using (28), (37a), and the linearity of the trace function, the right-hand side (RHS) of (41) can be derived as follows:

RHS of (41)

$$= 2\sigma_d^4 \operatorname{tr} \Big(\boldsymbol{\Phi}^{-2}(\xi) \boldsymbol{\Psi} \boldsymbol{\Omega}^{-1} \boldsymbol{\Psi}^* - P_R^{-1} \sigma_{n_x}^2 \operatorname{tr}(\bar{\boldsymbol{B}}^* \bar{\boldsymbol{B}}) \boldsymbol{\Phi}^{-3}(\xi) \boldsymbol{\Psi}$$
$$\times \boldsymbol{\Omega}^{-1} \boldsymbol{\Psi}^* \Big) - 2\sigma_d^4 \operatorname{tr} \Big(\sigma_{n_s}^2 \boldsymbol{\Phi}^{-3}(\xi) \boldsymbol{\Psi} \boldsymbol{\Omega}^{-2} \boldsymbol{\Psi}^* (\boldsymbol{H}^*)^T$$
$$\times \bar{\boldsymbol{B}} \bar{\boldsymbol{B}}^* \boldsymbol{H}^T + \sigma_d^2 \boldsymbol{\Phi}^{-3}(\xi) \boldsymbol{\Psi} \boldsymbol{\Omega}^{-1} \boldsymbol{H} \boldsymbol{A} \boldsymbol{A}^* \boldsymbol{H}^* \boldsymbol{\Omega}^{-1}$$
$$\times \boldsymbol{\Psi}^* (\boldsymbol{H}^*)^T \bar{\boldsymbol{B}} \bar{\boldsymbol{B}}^* \boldsymbol{H}^T \Big)$$
$$= 2\sigma_d^4 \operatorname{tr} \Big((\boldsymbol{\Phi} - P_R^{-1} \sigma_{n_x}^2 \operatorname{tr}(\bar{\boldsymbol{B}}^* \bar{\boldsymbol{B}}) \boldsymbol{I}_M) \boldsymbol{\Phi}^{-3}(\xi) \boldsymbol{\Psi} \boldsymbol{\Omega}^{-1} \boldsymbol{\Psi}^* \Big)$$
$$- 2\sigma_d^4 \operatorname{tr} \Big(\boldsymbol{\Phi}^{-3}(\xi) \boldsymbol{\Psi} \boldsymbol{\Omega}^{-1} (\sigma_{n_s}^2 \boldsymbol{I}_M + \sigma_d^2 \boldsymbol{H} \boldsymbol{A} \boldsymbol{A}^* \boldsymbol{H}^*) \boldsymbol{\Omega}^{-1}$$
$$\times \boldsymbol{\Psi}^* (\boldsymbol{H}^*)^T \bar{\boldsymbol{B}} \bar{\boldsymbol{B}}^* \boldsymbol{H}^T \Big)$$
$$= 2\sigma_d^4 \operatorname{tr} \Big(((\boldsymbol{H}^*)^T \bar{\boldsymbol{B}} \bar{\boldsymbol{B}}^* \boldsymbol{H}^T + \xi \boldsymbol{I}_M$$
$$- P_R^{-1} \sigma_{n_x}^2 \operatorname{tr}(\bar{\boldsymbol{B}}^* \bar{\boldsymbol{B}}) \boldsymbol{I}_M) \boldsymbol{\Phi}^{-3}(\xi) \boldsymbol{\Psi} \boldsymbol{\Omega}^{-1} \boldsymbol{\Psi}^* \Big)$$
$$- 2\sigma_d^4 \operatorname{tr} \Big(\boldsymbol{\Phi}^{-3}(\xi) \boldsymbol{\Psi} \boldsymbol{\Omega}^{-1} \boldsymbol{\Omega} \boldsymbol{\Omega}^{-1} \boldsymbol{\Psi}^* (\boldsymbol{H}^*)^T \bar{\boldsymbol{B}} \bar{\boldsymbol{B}}^* \boldsymbol{H}^T \Big)$$
$$= 2\sigma_d^4 \operatorname{tr} \Big(((\boldsymbol{H}^*)^T \bar{\boldsymbol{B}} \bar{\boldsymbol{B}}^* \boldsymbol{H}^T + \xi \boldsymbol{I}_M - \sigma_{n_x}^2 P_R^{-1} \operatorname{tr}(\bar{\boldsymbol{B}} \bar{\boldsymbol{B}}^*) \boldsymbol{I}_M$$

$$- (\boldsymbol{H}^*)^T \bar{\boldsymbol{B}} \bar{\boldsymbol{B}}^* \boldsymbol{H}^T)(\boldsymbol{\Phi}^{-3}(\xi) \boldsymbol{\Psi} \boldsymbol{\Omega}^{-1} \boldsymbol{\Psi}^*) \Big)$$
$$= 2\sigma_d^4 \operatorname{tr} \Big((\xi \boldsymbol{I}_M - \sigma_{n_x}^2 P_R^{-1} \operatorname{tr}(\bar{\boldsymbol{B}} \bar{\boldsymbol{B}}^*) \boldsymbol{I}_M)$$
$$\times (\boldsymbol{\Phi}^{-3}(\xi) \boldsymbol{\Psi} \boldsymbol{\Omega}^{-1} \boldsymbol{\Psi}^*) \Big)$$
$$= 2\sigma_d^4 \big(\xi - \sigma_{n_x}^2 P_R^{-1} \operatorname{tr}(\bar{\boldsymbol{B}} \bar{\boldsymbol{B}}^*)\big) \operatorname{tr} \big(\boldsymbol{I}_M (\boldsymbol{\Phi}^{-3}(\xi) \boldsymbol{\Psi} \boldsymbol{\Omega}^{-1} \boldsymbol{\Psi}^*) \big).$$
$$\text{(A1)}$$

The RHS of the last equality in (A1) is identical to (42).

APPENDIX B
PROOF OF STRICT QUASI--CONVEXITY OF (39)

The second derivative of $J(\xi)$ is derived from (42) as

$$\frac{\partial^2 J(\xi)}{\partial^2 \xi} = 2\sigma_d^4 \operatorname{tr} \big(\boldsymbol{\Phi}^{-3}(\xi) \boldsymbol{\Psi} \boldsymbol{\Omega}^{-1} \boldsymbol{\Psi}^* \big)$$
$$- 6\sigma_d^4 (\xi - \xi_o) \operatorname{tr}(\boldsymbol{\Phi}^{-4}(\xi) \boldsymbol{\Psi} \boldsymbol{\Omega}^{-1} \boldsymbol{\Psi}^*)$$
$$= 2\sigma_d^2 \operatorname{tr} \big((\boldsymbol{\Phi}(\xi) + 3(\xi_o - \xi) \boldsymbol{I}_M) \boldsymbol{\Phi}^{-4}(\xi) \boldsymbol{\Psi} \boldsymbol{\Omega}^{-1} \boldsymbol{\Psi}^* \big)$$
$$= 2\sigma_d^2 \operatorname{tr} \Big(\big((\boldsymbol{H}^*)^T \boldsymbol{B} \boldsymbol{B}^* \boldsymbol{H}^T$$
$$+ 3\big(\xi_o - \frac{2}{3}\xi\big) \boldsymbol{I}_M \big) \boldsymbol{\Phi}^{-4}(\xi) \boldsymbol{\Psi} \boldsymbol{\Omega}^{-1} \boldsymbol{\Psi}^* \Big). \text{ (B1)}$$

Due to the quadratic matrices in the trace function in (42) and (B1), it can be shown that

$$\begin{cases} J'(\xi) < 0 \text{ and } J''(\xi) > 0, & \text{for } 0 < \xi < \xi_o \\ J'(\xi) = 0 \text{ and } J''(\xi) > 0, & \text{for } \xi = \xi_o \\ J'(\xi) > 0 \text{ and } J''(\xi) > 0, & \text{for } \xi_o < \xi < \frac{3}{2}\xi_o \\ J'(\xi) > 0 \text{ and } J''(\xi) \in \mathbb{R}, & \text{for } \xi > \frac{3}{2}\xi_o. \end{cases}$$

Consequently, $J(\xi)$ is convex or strictly quasi-convex with respect to ξ.

APPENDIX C
PROOF OF REMARK 3

When $N_k = 1$ and $\boldsymbol{\Gamma} = \boldsymbol{I}_{2K}$, both beamforming matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ in (8) and (12), respectively, degenerate to $\boldsymbol{I}_{2K}$. Accordingly, the ZF solution in (26) becomes

$$\bar{\boldsymbol{F}} = \left(\boldsymbol{H}^T\right)^+ \boldsymbol{P} \left(\boldsymbol{H}^*\right)^+. \qquad (C1)$$

Noting the normalization and exchange matrix form in this paper, $\bar{\boldsymbol{F}}$ in (C1) can be easily shown to be identical to the ZF solution in [13]. In the MMSE case, the solution $\bar{\boldsymbol{F}}(\xi)$ in (36) becomes

$$\bar{\boldsymbol{F}}(\xi) = \sigma_d^2 \left((\boldsymbol{H}\boldsymbol{H}^*)^T + \frac{2\sigma_{n_x}^2 K}{P_R}\boldsymbol{I}_M\right)^{-1} (\boldsymbol{H}^*)^T \boldsymbol{P}\boldsymbol{H}^*$$
$$\times \left(\sigma_d^2 \boldsymbol{H}\boldsymbol{H}^* + \sigma_{n_s}^2 \boldsymbol{I}_M\right)^{-1}. \quad (C2)$$

Considering the normalization factor, (C2) can be easily shown to be identical to the MMSE solution in [13].

APPENDIX D
DERIVATION OF (63)

The numerator of (62) yields

$$\mathrm{E}\,\|\boldsymbol{P}\boldsymbol{d}\|^2 = \sigma_d^2 \,\mathrm{tr}\left(\boldsymbol{P}^T\boldsymbol{P}\right) = 2\sigma_d^2 K = P_U. \qquad (D1)$$

The denominator of (62) is derived as

$$\mathrm{E}\,\left\|\boldsymbol{P}(\boldsymbol{H}\bar{\boldsymbol{A}}\boldsymbol{\Gamma})^+\boldsymbol{n}_s + q^{-1}\bar{\boldsymbol{B}}^*\boldsymbol{n}_x\right\|^2$$
$$= \mathrm{E}\,\left\|\boldsymbol{P}(\boldsymbol{H}\bar{\boldsymbol{A}}\boldsymbol{\Gamma})^+\boldsymbol{n}_s + q^{-1}\bar{\boldsymbol{A}}^T\boldsymbol{n}_x\right\|^2$$
$$= \sigma_{n_s}^2 \,\mathrm{tr}\left((\boldsymbol{\Gamma}\bar{\boldsymbol{A}}^*\boldsymbol{H}^*)^+ \boldsymbol{P}^T\boldsymbol{P}(\boldsymbol{H}\bar{\boldsymbol{A}}\boldsymbol{\Gamma})^+\right)$$
$$+ P_R^{-1}\sigma_d^2 \,\mathrm{tr}\left((\bar{\boldsymbol{A}}^T\boldsymbol{H}^T)^+((\bar{\boldsymbol{A}}^T\boldsymbol{H}^T)^+)^*\right)\sigma_{n_x}^2 \,\mathrm{tr}\left(\bar{\boldsymbol{A}}\bar{\boldsymbol{A}}^*\right)$$
$$+ P_R^{-1}\sigma_{n_x}^2\sigma_{n_s}^2 \,\mathrm{tr}\left(\left((\bar{\boldsymbol{A}}^*\boldsymbol{H}^*\boldsymbol{H}\bar{\boldsymbol{A}})^{-1}\right)^T \boldsymbol{P}\right.$$
$$\left.\times (\boldsymbol{\Gamma}\bar{\boldsymbol{A}}^*\boldsymbol{H}^*\boldsymbol{H}\bar{\boldsymbol{A}}\boldsymbol{\Gamma})^{-1}\boldsymbol{P}\right)\,\mathrm{tr}\left(\bar{\boldsymbol{A}}\bar{\boldsymbol{A}}^*\right)$$
$$\simeq \sigma_{n_s}^2 \,\mathrm{tr}\left((\boldsymbol{\Gamma}\bar{\boldsymbol{A}}^*\boldsymbol{H}^*\boldsymbol{H}\bar{\boldsymbol{A}}\boldsymbol{\Gamma})^{-1}\right)$$
$$+ P_R^{-1}\sigma_d^2\sigma_{n_x}^2 \,\mathrm{tr}\left(\left((\bar{\boldsymbol{A}}^*\boldsymbol{H}^*\boldsymbol{H}\bar{\boldsymbol{A}})^{-1}\right)^T\right)\,\mathrm{tr}\left(\bar{\boldsymbol{A}}\bar{\boldsymbol{A}}^*\right)$$
$$= \sigma_{n_s}^2 \,\mathrm{tr}\left((\boldsymbol{\Gamma}\bar{\boldsymbol{A}}^*\boldsymbol{H}^*\boldsymbol{H}\bar{\boldsymbol{A}}\boldsymbol{\Gamma})^{-1}\right)$$
$$+ P_R^{-1}P_U\sigma_{n_x}^2 \,\mathrm{tr}\left((\bar{\boldsymbol{A}}^*\boldsymbol{H}^*\boldsymbol{H}\bar{\boldsymbol{A}})^{-1}\right). \qquad (D2)$$

In (D2), the approximation comes from the high SNR assumption, i.e., $P_R^{-1}\sigma_{n_x}^2\sigma_{n_s}^2 \simeq 0$, and the third equality comes from the definitions $\|\bar{\boldsymbol{a}}\|^2 = 1$ and $2\sigma_d^2 K = P_U$ in (47). Consequently, we can achieve (63) from (D1) and (D2).

APPENDIX E
PROOF OF REMARK 7

Letting $\{\kappa_k\}$ and $\{\lambda_k\}$ be the $2K$ singular values of $\boldsymbol{H}\bar{\boldsymbol{A}}\boldsymbol{\Gamma}$ and $\boldsymbol{H}\bar{\boldsymbol{A}}$, respectively, and denoting $\sqrt{\mathrm{SNR}_{RU}\mathrm{SNR}_{UR}}/2$ by c, the system SNR upper bound in (66) is derived as follows:

$$\frac{c}{\sqrt{\mu_1\mu_2}} = c\sqrt{\left(\sum_{k=1}^{2K}\kappa_k^{-2}\right)^{-1}\left(\sum_{k=1}^{2K}\lambda_k^{-2}\right)^{-1}}$$
$$= c\sqrt{\left(2K\,\mathrm{E}(\kappa_k^{-2})\right)^{-1}\left(2K\,\mathrm{E}(\lambda_k^{-2})\right)^{-1}}$$
$$\leq c\sqrt{\frac{\mathrm{E}(\kappa_k^2)}{2K}\frac{\mathrm{E}(\lambda_k^2)}{2K}} = c\sqrt{\frac{\sum_{k=1}^{2K}\kappa_k^2}{(2K)^2}\frac{\sum_{k=1}^{2K}\lambda_k^2}{(2K)^2}}$$
$$= \frac{c}{4K^2}\sqrt{\mathrm{tr}(\boldsymbol{\Gamma}\bar{\boldsymbol{A}}^*\boldsymbol{H}^*\boldsymbol{H}\bar{\boldsymbol{A}}\boldsymbol{\Gamma})\,\mathrm{tr}(\bar{\boldsymbol{A}}^*\boldsymbol{H}^*\boldsymbol{H}\bar{\boldsymbol{A}})}$$
$$= \frac{c}{4K^2}\sqrt{\sum_{k=1}^{2K}\gamma_k^2\bar{\boldsymbol{a}}_k^*\boldsymbol{H}_k^*\boldsymbol{H}_k\bar{\boldsymbol{a}}_k\sum_{k=1}^{2K}\bar{\boldsymbol{a}}_k^*\boldsymbol{H}_k^*\boldsymbol{H}_k\bar{\boldsymbol{a}}_k} \quad (E1)$$

where the inequality comes from Jensen's inequality [32] by using the convexity of the function $1/w$, i.e., $\mathrm{E}^{-1}(w) \leq \mathrm{E}(w^{-1})$. Therefore, maximizing the upper bound in (E1) is equivalent as

$$\max_{\bar{\boldsymbol{a}}_k}\bar{\boldsymbol{a}}_k^*\boldsymbol{H}_k^*\boldsymbol{H}_k\bar{\boldsymbol{a}}_k, \;\forall k.$$

Consequently, the maximum upper bound of $\sqrt{\mathrm{SNR}_{RU}\mathrm{SNR}_{UR}/4\mu_1\mu_2}$ in (66) is achieved when $\bar{\boldsymbol{a}}_k$ is the eigen vector corresponding to the largest singular value of $\boldsymbol{H}_k$.

REFERENCES

[1] N. Khajehnouri and A. H. Sayed, "Distributed MMSE relay strategies for wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 55, no. 7, pp. 3336–3348, Jul. 2007.

[2] K. G. Seddik and K. J. R. Liu, "Distributed space-frequency coding over broadband relay channels," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 4748–4759, Nov. 2008.

[3] V. Havary-Nassab, S. Shahbazpanahi, A. Grami, and Z.-Q. Luo, "Distributed beamforming for relay networks based on second-order statistics of the channel state information," *IEEE Trans. Signal Process.*, vol. 56, no. 9, pp. 4306–4316, Sep. 2008.

[4] M. M. Abdallah and H. C. Papadopoulos, "Beamforming algorithms for information relaying in wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 56, no. 10, pp. 4772–4784, Oct. 2008.

[5] E. Koyuncu, Y. Jing, and H. Jafarkhani, "Distributed beamforming in wireless relay networks with quantized feedback," *IEEE J. Sel. Areas Commun.*, vol. 26, pp. 1429–1439, Oct. 2008.

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

[7] A. S. Behbahani, R. Merched, and A. M. Eltawil, "Optimizations of a MIMO relay network," *IEEE Trans. Signal Process.*, vol. 56, no. 10, pp. 5062–5072, Oct. 2008.

[8] B. Khoshnevis, W. Yu, and R. Adve, "Grassmannian beamforming for MIMO amplify-and-forward relaying," *IEEE J. Sel. Areas Commun.*, vol. 26, pp. 1397–1407, Oct. 2008.

[9] J. Joung and A. H. Sayed, "Power allocation for beamforming relay networks under channel uncertainties," in *Proc. IEEE Global Commun. Conf. (IEEE GLOBECOM)*, Honolulu, HI, Nov.–Dec. 2009.

[10] B. Rankov and A. Wittneben, "Spectral efficient protocol for half-duplex fading relay channels," *IEEE J. Sel. Areas Commun.*, vol. 25, pp. 379–389, Feb. 2007.

[11] I. Hammerström, M. Kuhn, C. Eşli, J. Zhao, A. Wittneben, and G. Bauch, "MIMO two-way relaying with transmit CSI at the relay," in *Proc. IEEE Signal Processing Adv. Wireless Commun., (SPAWC)*, Helsinki, Finland, Jun. 2007, pp. 1–5.

[12] H. J. Yang and J. Chun, "Generalized Schur decomposition-based two-way relaying for wireless MIMO systems," in *Proc. IEEE Global Commun. Conf., (IEEE GLOBECOM)*, New Orleans, LA, Nov. 2008, pp. 1–6.

[13] T. Unger and A. Klein, "Duplex schemes in multiple antenna two-hop relaying," *EURASIP J. Adv. Signal Process.*, vol. 2008, no. ID 128592.

[14] T. Cui, F. Gao, T. Ho, and A. Nallanathan, "Distributed space-time coding for two-way wireless relay networks," *IEEE Trans. Signal Process.*, vol. 57, no. 2, pp. 658–671, Feb. 2009.

[15] R. Zhang, Y.-C. Liang, C. C. Chai, and S. Cui, "Optimal beamforming for two-way multi-antenna relay channel with analogue network coding," *IEEE J. Sel. Areas Commun.*, vol. 27, pp. 699–712, Jun. 2009.

[16] V. Havary-Nassab, S. Shahbazpanahi, and A. Grami, "Optimal network beamforming for bi-directional relay networks," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Taipei, Taiwan, Apr. 2009, pp. 2277–2280.

[17] V. Havary-Nassab, S. Shahbazpanahi, and A. Grami, "Optimal distributed beamforming for two-way relay networks," *IEEE Trans. Signal Process.*, to be published.

[18] M. Chen and A. Yener, "Multiuser tow-way relaying: Detection and interference management strategies," *IEEE Trans. Wireless Commun.*, vol. 8, pp. 4296–4303, Aug. 2009.

[19] C. Eşli and A. Wittneben, "One- and two-way decode-and-forward relaying for wireless multiuser MIMO network," in *Proc. IEEE Global Commun. Conf., (IEEE GLOBECOM)*, New Orleans, LA, Nov. 2008, pp. 1–6.

[20] J. Joung and A. H. Sayed, "Multiuser two-way relaying method for beamforming systems," in *Proc. IEEE Int. Workshop Signal Processing Adv. Wireless Commun. (SPAWC)*, Perugia, Italy, Jun. 2009, pp. 280–284.

[21] K. Jitvanichphaibool, R. Zhang, and Y.-C. Liang, "Optimal resource allocation for two-way relay-assisted OFDMA," *IEEE Trans. Veh. Technol.*, vol. 58, pp. 3311–3321, Sep. 2009.

[22] A. Abdel-Samad, T. N. Davidson, and A. B. Gershman, "Robust transmit eigen beamforming based on imperfect channel state information," *IEEE Trans. Signal Process.*, vol. 54, no. 5, pp. 1596–1609, May 2006.

[23] S. Sanayei and A. Nosratinia, "Antenna selection in MIMO systems," *IEEE Commun. Mag.*, vol. 42, pp. 68–73, Oct. 2004.

[24] J. Chung, C.-S. Hwang, K. Kim, and Y. K. Kim, "A random beamforming technique in MIMO systems exploiting multiuser diversity," *IEEE J. Sel. Areas Commun.*, vol. 21, pp. 848–855, Jun. 2003.

[25] D. J. Love and R. W. Heath, Jr., "Equal gain transmission in multiple-input multiple-output wireless systems," *IEEE Trans. Commun.*, vol. 51, pp. 1102–1110, Jul. 2003.

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

[27] A. Hjørungnes and D. Gesbert, "Complex-valued matrix differentiation: Techniques and key results," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2740–2746, Jun. 2007.

[28] Z. Pan, K. K. Wong, and T. S. Ng, "Generalized multiuser orthogonal space-division multiplexing," *IEEE Trans. Wireless Commun.*, vol. 3, pp. 1969–1973, Nov. 2004.

[29] J. Joung and Y. H. Lee, "Regularized channel diagonalization for multiuser MIMO downlink using a modified MMSE criterion," *IEEE Trans. Signal Process.*, vol. 55, no. 4, pp. 1573–1579, Apr. 2007.

[30] A. H. Sayed, *Adaptive Filters*. New York: Wiley, 2008.

[31] L. S. Lasdon, *Optimization Theory for Large Systems*. New York: Dover, 2002.

[32] G. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 1978.

[33] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. New York: Wiley, 2006.

[34] R. K. Jain, D.-M. W. Chiu, and W. R. Hawe, "A quantitative measure of fairness and discrimination for resource allocation in shared computer systems," Digital Equipment Corp. (DEC), Tech. Rep. TR301, 1984.

Jingon Joung (S'04–M'07) received the B.S. degree in electrical engineering from Yonsei University, Seoul, Korea, in 2001 and the M.S. and Ph.D. degrees in electrical engineering and computer science from the Korea Advance Science and Technology (KAIST), Daejeon, Korea, in 2003 and 2007, respectively.

From March 2007 to August 2008, he was a Postdoctoral Research Scientist with the Department of Electrical Engineering, KAIST. From April 2007 to August 2008, he worked as a Commissioned Researcher at Lumicomm, Inc., Daejeon, Korea. From October 2008 to September 2009, he was a Postdoctoral Fellow with the Department of Electrical Engineering at the University of California, Los Angeles (UCLA). Since October 2009, he has been working as a Research Fellow at the Institute for Infocomm Research (I^{2}R), Singapore. His research activities are in the area of multiuser systems, multiple-input multiple-output (MIMO) communications, and cooperative systems.

Dr. Joung was the recipient of a Gold Prize at the Intel-ITRC Student Paper Contest in 2006.

Ali H. Sayed (S'90–M'92–SM'99–F'01) received the Ph.D. degree from Stanford University, Stanford, CA, in 1992.

He is Professor and Chairman of Electrical Engineering at the University of California, Los Angeles (UCLA) and Principal Investigator of the Adaptive Systems Laboratory. He has published widely, with over 300 articles and five books, in the areas of statistical signal processing, estimation theory, adaptive filtering, signal processing for communications and wireless networking, and fast algorithms for large structured problems. He is coauthor of the textbook *Linear Estimation* (Prentice-Hall, 2000) and of the research monograph *Indefinite Quadratic Estimation and Control* (SIAM, 1999), and he is co-editor of *Fast Algorithms for Matrices with Structure* (SIAM, 1999). He is also the author of the textbooks *Fundamentals of Adaptive Filtering* (Wiley, 2003) and *Adaptive Filters* (Wiley, 2008). He has contributed several encyclopedia and handbook articles.

Dr. Sayed has served on the editorial boards of the *IEEE Signal Processing Magazine*, the *European Signal Processing Journal*, the *International Journal on Adaptive Control and Signal Processing*, and the *SIAM Journal on Matrix Analysis and Applications*. He also served as the Editor-in-Chief of the IEEE TRANSACTIONS ON SIGNAL PROCESSING from 2003 to 2005 and the *EURASIP Journal on Advances in Signal Processing* from 2006 to 2007. He has served on the Publications (2003–2005), Awards (2005), and Conference (2007–present) Boards of the IEEE Signal Processing Society. He served on the Board of Governors (2007–2008) of the same Society and is now serving as Vice-President of Publications (2009–present). His work has received several recognitions, including the 1996 IEEE Donald G. Fink Award, the 2002 Best Paper Award from the IEEE Signal Processing Society, the 2003 Kuwait Prize in Basic Sciences, the 2005 Terman Award, the 2005 Young Author Best Paper Award from the IEEE Signal Processing Society, and two Best Student Paper Awards at international meetings (1999 and 2001). He has served as a 2005 Distinguished Lecturer of the IEEE Signal Processing Society and as General Chairman of ICASSP 2008.