User Selection Methods for Multiuser Two-Way Relay Communications Using Space Division Multiple Access

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

***Abstract*—In this paper, we design a multiuser two-way relay system using space division multiple access (SDMA) communications and devise an optimal scheduling method that maximizes the sum rate while ensuring fairness among users. To reduce the computational load at the relays, we propose rate- and angle-based suboptimal scheduling methods. The numerical results illustrate tradeoff between complexity and the performance. Specifically, when the relay has two antennas, we verify that the rate-based method can provide significant computational savings at the cost of a rate reduction of less than 4% when compared with the optimal scheduling method.**

***Index Terms*—Space division multiple access (SDMA), multiuser communications, two-way relay systems, scheduling.**

I. Introduction

TWO-WAY relay communications allows the exchange of data between two users (denoted by U_1 and U_2) with the assistance of a relay node (denoted by R). When a relay is employed, four phases of communications generally arise to support two data streams: $\mathrm{U}_1 \rightarrow \mathrm{R}$, $\mathrm{R} \rightarrow \mathrm{U}_2$, $\mathrm{U}_2 \rightarrow \mathrm{R}$, $\mathrm{R} \rightarrow \mathrm{U}_1$. Various protocols have been proposed to improve the use of channel resources such as: physical layer network coding (PNC) requiring three phases ($\mathrm{U}_1 \rightarrow \mathrm{R}$, $\mathrm{U}_2 \rightarrow \mathrm{R}$, $\mathrm{R} \rightarrow \mathrm{U}_1\&\mathrm{U}_2$) [1], [2] and analog network coding (ANC) requiring two phases ($\mathrm{U}_1\&\mathrm{U}_2 \rightarrow \mathrm{R}$, $\mathrm{R} \rightarrow \mathrm{U}_1\&\mathrm{U}_2$) [3], [4]. Also, a hybrid PNC and ANC method sharing time resources was proposed in [5] and an opportunistic source selection (OSS) protocol considering a *direct* path between U_1 and U_2 was studied in [6]. In the OSS protocol, multiuser diversity can be exploited by selecting a communication mode between ($\mathrm{U}_1\&\mathrm{R} \rightarrow \mathrm{U}_2$) and ($\mathrm{U}_2\&\mathrm{R} \rightarrow \mathrm{U}_1$), according to the signal-to-noise ratio (SNR) at the user.

By using code division multiple access (CDMA) or space division multiple access (SDMA) schemes, *multiuser* two-way relay communications have been proposed for decode-and-forward [7], [8] and amplify-and-forward [9], [10] relay systems for $2K$ users (K pairs). Every user transmits signals to the relay simultaneously in a multiple-access (MAC) phase, and the relay retransmits the received signals to every user in a broadcast (BC) phase similar to the ANC protocol. The SDMA method makes it possible to reuse the conventional channels constructed by time, frequency, or code, at the cost of knowing the channel state information (CSI) at the transmitter. In multiuser two-way communications, CSIs are required at the relay for the SDMA processing and they can be estimated through the MAC phase by using orthogonal training sequences transmitted from the users to the relays [6], [8]–[11]. Zero-forcing (ZF)- and minimum mean-square-error (MMSE)-based SDMA relaying methods have been studied under the assumption that the number of users ($2K$) is less than or equal to the number of relay antennas (N) [9], [10]. The condition that $2K \leq N$ is necessary and sufficient to cancel the interferences perfectly for ZF-based SDMA relaying when each user transmits one data stream. Therefore, when $2K > N$, selecting (scheduling) affordable users among $2K$ users is required to efficiently reduce the interference and fairly support all users.

In this paper, we derive both ZF- and MMSE-based SDMA relaying matrices for a general number of users and introduce user selection schemes for multiuser two-way relay communications. To serve all users fairly, multiple SDMA user groups are selected and served through different time slots, i.e., a time-division multiple access (TDMA) method is used. An optimal method selecting M_t users for the tth SDMA group is presented to maximize the sum rate of the system. The optimal method requires a search whose complexity increases combinatorially with K since it considers every possible combination of all SDMA groups. Moreover, for a given M_t, $\mathcal{O}(M_t^2 N)$ computations are needed for calculating the sum rate of each search. To avoid combinatorial search, we propose a rate-based suboptimal method, which sequentially selects SDMA groups to achieve the largest rate for *part* of the time slots. To further reduce the computational load, we introduce an angle-based suboptimal method selecting one user occupying the most orthogonal channels to a given user channels. Computing the orthogonality between two channel vectors requires only $\mathcal{O}(N)$ computations. Simulations are conducted to evaluate performance in terms of the average sum rate. As a result, an average rate loss of less than 4% compared to the optimal method is observed with considerable computational reduction for the rate-based suboptimal method when $N = 2$, though the loss increases as N increases. For the angle-based method, the performance loss is not negligible; however, the computational complexity is reduced dramatically.

Notation: The superscripts 'T' and '$*$' denote transposition and complex conjugate transposition for any vector or

Manuscript received July 16, 2009; revised January 7, 2010 and April 11, 2010; accepted April 20, 2010. The associate editor coordinating the review of this letter and approving if for publication was I.-M. Kim.

J. Joung is with the Institute for Infocomm Research (I^{2}R), A*STAR, Singapore 138632 (e-mail: jgjoung@i2r.a-star.edu.sg). This work was performed while J. Joung was a post-doctoral researcher at the UCLA Adaptive Systems Laboratory.

A. H. Sayed is with the Department of Electrical Engineering, University of California (UCLA), Los Angeles, CA 90095, USA (e-mail: sayed@ee.ucla.edu).

This work was supported in part by NSF grants ECS-0601266 and ECS-0725441 and by the Korea Research Foundation Grant funded by the Korean Government [KRF-2008-357-D00179].

Digital Object Identifier 10.1109/TWC.2010.07.091054

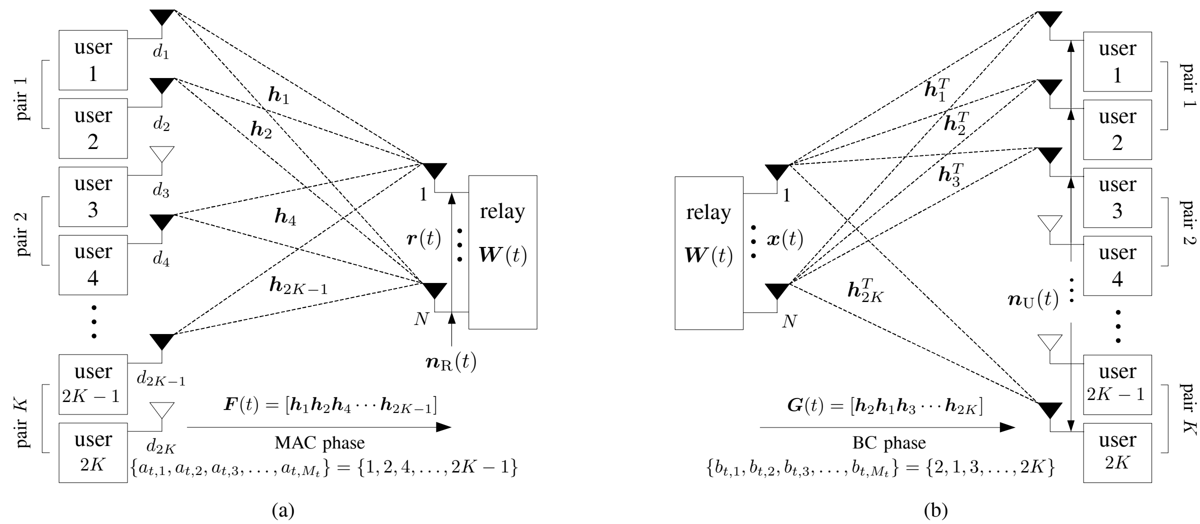


Fig. 1. Multiuser relay system model at the tth slot. (a) The MAC phase: transmission from the selected $\{a_{t,1},\dots,a_{t,M_t}\}$th users to the relay. (b) The BC phase: transmission from the relay to the selected $\{b_{t,1},\dots,b_{t,M_t}\}$th users.

matrix, respectively; $\boldsymbol{A}^{-1}$ and $\boldsymbol{A}^{+}$ denote matrix inversion and pseudoinversion of $\boldsymbol{A}$, respectively; $\boldsymbol{I}_a$ represents an a-by-a identity matrix; $\mathrm{tr}(\boldsymbol{A})$ represents the trace of matrix $\boldsymbol{A}$; 'E' stands for expectation of a random variable; for any scalar a, vector $\boldsymbol{a}$, and matrix $\boldsymbol{A}$, the notation $|a|$, $\|\boldsymbol{a}\|$, and $\|\boldsymbol{A}\|_F$ denote the absolute value of a, 2-norm of $\boldsymbol{a}$, and Frobenius-norm of $\boldsymbol{A}$, respectively; $\mathrm{diag}(\boldsymbol{A})$ and $\mathrm{offd}(\boldsymbol{A})$ are the diagonal and off-diagonal matrices of a square matrix $\boldsymbol{A}$, respectively; $\mathrm{mod}(a,b)$ is a modulo operation finding the remainder of division of a by b; $\binom{a}{b}$ represents the number of b-combinations from a set with a elements, i.e., $\frac{a!}{b!(a-b)!}$, where $a!$ means the factorial of a; $\lceil a \rceil$ is the smallest integer larger than a; $\mathcal{A} \subseteq \mathcal{B}$ means $\mathcal{A}$ is a subset of $\mathcal{B}$; and $\bigcup_i \mathcal{A}_i$ denotes a union of sets $\{\mathcal{A}_i\}$.

II. Multiuser Two-Way Relay System Description

There are $2K$ user nodes having one antenna each and one relay node having N antennas as shown in Fig. 1. The $2K$ users result in K pairs of two users exchanging data with each other through the relay. Without loss of generality, it is assumed that the $(2k-1)$th and the $(2k)$th users communicate with each other ($k \in \{1,\cdots,K\}$). The vector channel between the jth user ($j \in \{1,\dots,2K\}$) and the relay node is represented by $\boldsymbol{h}_j \in \mathbb{C}^{N\times 1}$, where the ith element is the channel gain between the ith antenna of the relay and the jth user. We assume that the elements of $\boldsymbol{h}_j$ are independent and identically distributed (i.i.d.) and zero-mean complex Gaussian random variables with unit variance[1]. We also assume that every channel remains static during one scheduling period, i.e., a quasi-static channel. One scheduling period is divided into T slots ($t \in \{1,\dots,T\}$) and each slot t is composed of MAC and BC phases. In the MAC phase at the tth slot, the selected M_t users, $a_{t,1} < a_{t,2} < \cdots < a_{t,M_t}$ and $a_{t,m} \in \{1,\dots,2K\}$, construct one SDMA group and transmit their data simultaneously to the relay as shown in Fig. 1(a). In the BC phase at the same slot, the relay retransmits (broadcasts) the received M_t data streams to the $\{b_{t,m}\}$th users ($b_{t,m} \in \{1,\dots,2K\}$ and $m = \{1,\dots,M_t\}$) as shown in Fig. 1(b). For the data exchange between two-way communication users, $a_{t,m}$ and $b_{t,m}$, the user indices $\{b_{t,m}\}$ in BC phase are determined according to $\{a_{t,m}\}$ as follows:

$$b_{t,m} = a_{t,m} + 1 - 2\,\mathrm{mod}(a_{t,m}+1, 2). \tag{1}$$

To avoid ambiguity and to effectively mitigate co-channel interferences (CCIs) among the M_t data streams, as we mentioned previously, the number of supported data streams M_t at one slot t should be less than or equal to the number of relay antennas [9], [10]; also, to enable the two-way communications protocol, M_t should be larger than two, i.e.,

$$2 \le M_t \le N. \tag{2}$$

Though there is no restriction on the maximum number of users for MMSE-based SDMA systems, the interferences can be effectively mitigated when (2) is satisfied [9], [10]. The different SDMA user groups are time-duplexed and supported through T different slots as TDMA. Here, note that T depends on M_t. For example, when $2K = 8$ and $N = 4$, four scenarios are possible for SDMA groups: $\{M_t = 2\}$, $\{M_{t_1} = 2, M_{t_2} = M_{t_3} = 3\}$, $\{M_{t_1} = M_{t_2} = 2, M_{t_3} = 4\}$, and $\{M_t = 4\}$ yield $T = 4$, 3, 3, and 2, respectively. Throughout this paper, we assume that the coherent time of the channel is long enough to support all users within any possible T scheduling time[2].

Let d_j denote the data symbol for the jth user. The received signal at the relay, at the MAC phase of the tth slot, can be written as follows:

$$\boldsymbol{r}(t) = \boldsymbol{F}(t)\boldsymbol{d}(t) + \boldsymbol{n}_{\mathrm{R}}(t) \in \mathbb{C}^{N\times 1} \tag{3}$$

[1]Using a transmit power control mechanism for the users (relay) [10], the average received power at the relay (each user) can be assumed to be identical. Therefore, we can set the variances of the channel elements to one.

[2]Otherwise, the previously unsupported users might be scheduled in the next scheduling period with higher priority than the supported users for fairness. Also, additional resources such as code or frequency can be used for the unsupported users in the same scheduling period. Namely, CDMA and frequency-division multiple-access (FDMA) can be directly combined with the SDMA-based TDMA method.

where the multiuser transmit signal vector $\boldsymbol{d}(t) = [d_{a_{t,1}} \cdots d_{a_{t,M_t}}]^T \in \mathbb{C}^{M_t \times 1}$ satisfies $\mathrm{E}\,\boldsymbol{d}(t)\boldsymbol{d}(t)^* = \boldsymbol{I}_{M_t}$; the multiuser channel matrix $\boldsymbol{F}(t) = [\boldsymbol{h}_{a_{t,1}} \cdots \boldsymbol{h}_{a_{t,M_t}}] \in \mathbb{C}^{N \times M_t}$; and $\boldsymbol{n}_{\mathrm{R}}(t) \in \mathbb{C}^{N \times 1}$ is a zero-mean additive white Gaussian noise (AWGN) at the relay and $\mathrm{E}\,\boldsymbol{n}_{\mathrm{R}}(t)\boldsymbol{n}_{\mathrm{R}}^*(t) = \sigma_{n_{\mathrm{R}}}^2 \boldsymbol{I}_N$. The relay multiplies $\boldsymbol{r}(t)$ by a relay processing matrix $\boldsymbol{W}(t) \in \mathbb{C}^{N \times N}$, and forwards

$$\boldsymbol{x}(t) = \boldsymbol{W}(t)\boldsymbol{r}(t) \in \mathbb{C}^{N \times 1} \tag{4}$$

during the BC phase. Here, the transmit power of the relay is bounded by P_{R} as

$$\mathrm{E}\,\|\boldsymbol{x}(t)\|^2 \leq P_{\mathrm{R}}. \tag{5}$$

Denoting the received signal at the selected $b_{t,m}$th user by $y_{b_{t,m}}$, the received signal vector of the selected users is written as

$$\boldsymbol{y}(t) = \boldsymbol{G}^T(t)\boldsymbol{W}(t)\boldsymbol{F}(t)\boldsymbol{d}(t) + \boldsymbol{G}^T(t)\boldsymbol{W}(t)\boldsymbol{n}_{\mathrm{R}}(t) + \boldsymbol{n}_{\mathrm{U}}(t), \tag{6}$$

where $\boldsymbol{y}(t) = [y_{b_{t,1}}, \ldots, y_{b_{t,M_t}}]^T \in \mathbb{C}^{M_t \times 1}$; the multiuser channel matrix $\boldsymbol{G}(t)$ can be represented as $\boldsymbol{G}(t) = [\boldsymbol{h}_{b_{t,1}} \cdots \boldsymbol{h}_{b_{t,M_t}}] \in \mathbb{C}^{N \times M_t}$ from the reciprocity between MAC and BC channels in the same scheduling period as the up- and down-link channels in time division duplex (TDD) systems; and $\boldsymbol{n}_{\mathrm{U}}(t) \in \mathbb{C}^{M_t \times 1}$ is a multiuser AWGN satisfying $\mathrm{E}\,\boldsymbol{n}_{\mathrm{U}}(t)\boldsymbol{n}_{\mathrm{U}}^*(t) = \sigma_{n_{\mathrm{U}}}^2 \boldsymbol{I}_{M_t}$.

III. SDMA-Based Two-Way Relay Processing Matrix Design

In this section, we design the relay transceiver processing matrix $\boldsymbol{W}(t)$ based on both ZF and MMSE criteria. Contrary to the design of $\boldsymbol{W}(t)$ in [9], [10], we derive $\boldsymbol{W}(t)$ here for the cases of a general number of users. Although the SDMA relay system is designed for single-antenna users in this paper, it is straightforward to extend the method to the case of multiple-antenna users with beamforming.

A. ZF Design

In order to perfectly cancel CCIs, the effective channel matrix in (6) should be reduced to a diagonal matrix[3] as

$$\boldsymbol{G}^T(t)\boldsymbol{W}(t)\boldsymbol{F}(t) = q(t)\boldsymbol{I}_{M_t} \tag{7}$$

where $q(t)$ is an effective channel gain. Under the condition (2), the minimum norm solution for the ZF relay processing matrix is obtained from (7) as

$$\boldsymbol{W}_{ZF}(t) = q(t)(\boldsymbol{G}^T(t))^+\boldsymbol{F}^+(t). \tag{8}$$

Using ZF-based SDMA relay processing in (8), the received signal in (6) becomes

$$\boldsymbol{y}(t) = q(t)\boldsymbol{d}(t) + q(t)\boldsymbol{F}^+(t)\boldsymbol{n}_{\mathrm{R}}(t) + \boldsymbol{n}_{\mathrm{U}}(t). \tag{9}$$

[3]If there is no power constraint on the relay, i.e., $P_{\mathrm{R}} = \infty$ in (5), we can find a feasible solution that $\boldsymbol{G}^T(t)\boldsymbol{W}(t)\boldsymbol{F}(t) = \boldsymbol{I}_{M_t}$ instead of (7). However, due to (5), we need to relax the ZF condition as in (7). This relaxation means that the users require the information $q(t)$ to equalize the received signal as shown later. Thus, $q(t)$ should be broadcast from the relay to the users since it will be derived as a function of the multiuser channels shown later.

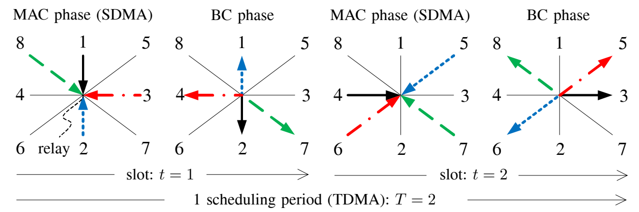


Fig. 2. Examples of multiuser two-way communications when $2K = 8$, $N = 4$ and $M_t = 4$.

From (9), we can see that $q(t)$ is the effective channel gain for each data stream. After equalization with $q^{-1}(t)$ at the users' side, the estimates of the transmitted data can be written as

$$\begin{aligned}\hat{\boldsymbol{d}}(t) &= q^{-1}(t)\boldsymbol{y}(t) && (10a)\\ &= \boldsymbol{d}(t) + \boldsymbol{F}^+(t)\boldsymbol{n}_{\mathrm{R}}(t) + q^{-1}(t)\boldsymbol{n}_{\mathrm{U}}(t) && (10b)\end{aligned}$$

where $\hat{\boldsymbol{d}}(t) \triangleq [\hat{d}_{b'_{t,1}} \cdots \hat{d}_{b'_{t,M_t}}]^T$ and $\hat{d}_{b'_{t,m}}$ is the estimate at the selected $b_{t,m}$th user. Here, the subscript $b'_{t,m}$ represents the index of the pair of the $b_{t,m}$th user; thus, we have $b'_{t,m} = a_{t,m}$ since the estimate of the $b_{t,m}$th user is the transmitted data from the $a_{t,m}$th user. Refer to the following example.

Example: Figure 2 illustrates an example of one scheduling period when $2K = 8$ and $N = 4$. For simple description, we fix $M_t = 4$. Thus, the required scheduling time $T = 2$ in this example. In the MAC phase of the first slot ($t = 1$), users 1, 2, 3, and 8 transmit data to the relay, simultaneously, i.e., $(a_{1,1}, \ldots, a_{1,4})$=$(1, 2, 3, 8)$. In the BC phase of the first slot, the relay retransmits four data streams of users 1, 2, 3, and 8 to users 2, 1, 4, and 7, respectively, i.e., $(b_{1,1}, \ldots, b_{1,4})$=$(2, 1, 4, 7)$. Similarly, in the second slot ($t = 2$), $(a_{2,1}, \ldots, a_{2,4})$=$(4, 5, 6, 7)$ and $(b_{2,1}, \ldots, b_{2,4})$=$(3, 6, 5, 8)$. All users' data exchanges are completed through two slots ($T = 2$). From (10), the estimates at each slot can be written as

$$\begin{aligned}q^{-1}(1)\boldsymbol{y}(1) &= q^{-1}(1)\,[\,y_2\ y_1\ y_4\ y_7\,]^T = \hat{\boldsymbol{d}}(1) = [\,\hat{d}_1\ \hat{d}_2\ \hat{d}_3\ \hat{d}_8\,]^T\\ &= [\,d_1\ d_2\ d_3\ d_8\,]^T + \boldsymbol{F}^+(1)\boldsymbol{n}_{\mathrm{R}}(1) + q^{-1}(1)\boldsymbol{n}_{\mathrm{U}}(1)\\ q^{-1}(2)\boldsymbol{y}(2) &= q^{-1}(2)\,[\,y_3\ y_6\ y_5\ y_8\,]^T = \hat{\boldsymbol{d}}(2) = [\,\hat{d}_4\ \hat{d}_5\ \hat{d}_6\ \hat{d}_7\,]^T\\ &= [\,d_4\ d_5\ d_6\ d_7\,]^T + \boldsymbol{F}^+(2)\boldsymbol{n}_{\mathrm{R}}(2) + q^{-1}(2)\boldsymbol{n}_{\mathrm{U}}(2).\end{aligned} \tag{11}$$

From (11), we can see that $(b'_{1,1}, \ldots, b'_{1,4})$=$(1, 2, 3, 8)$=$(a_{1,1}, \ldots, a_{1,4})$; $(b'_{2,1}, \ldots, b'_{2,4})$=$(4, 5, 6, 7)$=$(a_{2,1}, \ldots, a_{2,4})$; and according to $a_{t,m}$, the multiuser channel matrices $\boldsymbol{F}(1) = [\boldsymbol{h}_1\ \boldsymbol{h}_2\ \boldsymbol{h}_3\ \boldsymbol{h}_8]$ and $\boldsymbol{F}(2) = [\boldsymbol{h}_4\ \boldsymbol{h}_5\ \boldsymbol{h}_6\ \boldsymbol{h}_7]$. ■

In (10), we should note that the effective channel gain $q(t)$ is bounded due to the relay transmit power constraint (5). Substituting (3) and (8) into (4), the power constraint (5) gives

$$q(t) \leq \sqrt{\frac{P_{\mathrm{R}}}{\|(\boldsymbol{G}^T(t))^+\|_F^2 + \sigma_{n_{\mathrm{R}}}^2\|(\boldsymbol{G}^T(t))^+\boldsymbol{F}^+(t)\|_F^2}}. \tag{12}$$

Therefore, the relay processing matrix $\boldsymbol{W}(t)$, which maximizes the effective channel gain, can be obtained from (8) and (12) as

$$\boldsymbol{W}_{ZF}(t) = \frac{\sqrt{P_{\mathrm{R}}}(\boldsymbol{G}^T(t))^+\boldsymbol{F}^+(t)}{\sqrt{\|(\boldsymbol{G}^T(t))^+\|_F^2 + \sigma_{n_{\mathrm{R}}}^2\|(\boldsymbol{G}^T(t))^+\boldsymbol{F}^+(t)\|_F^2}}. \tag{13}$$

B. MMSE Design

We start from (6) and (10a), and omit henceforth the time index t for notational convenience whenever convenient. We define the MMSE formulation as

$$\arg\min_{\boldsymbol{W}} \mathrm{E}\left\|\boldsymbol{d}-\hat{\boldsymbol{d}}\right\|^2 \text{ s.t. } \mathrm{E}\left\|\boldsymbol{x}\right\|^2 \leq P_{\mathrm{R}}. \tag{14}$$

The minimization problem (14) with constraint can be transformed into

$$\arg\min_{\{\bar{\boldsymbol{W}},\lambda,q\}} \Big[\mathrm{E}\left\|\boldsymbol{d}-\boldsymbol{G}^T\bar{\boldsymbol{W}}\boldsymbol{F}\boldsymbol{d}-\boldsymbol{G}^T\bar{\boldsymbol{W}}\boldsymbol{n}_{\mathrm{R}}-q^{-1}\boldsymbol{n}_{\mathrm{U}}\right\|^2 + \lambda\left(\mathrm{E}\left\|q\bar{\boldsymbol{W}}\left(\boldsymbol{F}\boldsymbol{d}+\boldsymbol{n}_{\mathrm{R}}\right)\right\|^2 - P_{\mathrm{R}}\right)\Big] \tag{15}$$

with a non-negative Lagrange multiplier λ and a substitution of $\boldsymbol{W}$ by $q\bar{\boldsymbol{W}}$. Setting the derivatives of the Lagrange cost J in the square bracket of (15) with respect to $\{\bar{\boldsymbol{W}},\lambda,q\}$ to zero, we get the Karush-Kuhn-Tucker (KKT) conditions as

$$\frac{\partial J}{\partial \bar{\boldsymbol{W}}}=0\rightarrow\left((\boldsymbol{G}^*)^T\boldsymbol{G}^T+\lambda q^2\boldsymbol{I}_M\right)\bar{\boldsymbol{W}}\left(\boldsymbol{F}\boldsymbol{F}^*+\sigma_{n_{\mathrm{R}}}^2\boldsymbol{I}_M\right) = (\boldsymbol{G}^*)^T\boldsymbol{F}^* \tag{16a}$$

$$\frac{\partial J}{\partial \lambda}=0\rightarrow q^2\left(\mathrm{tr}\left(\boldsymbol{F}^*\bar{\boldsymbol{W}}^*\bar{\boldsymbol{W}}\boldsymbol{F}\right)+\sigma_{n_{\mathrm{R}}}^2\,\mathrm{tr}\left(\bar{\boldsymbol{W}}^*\bar{\boldsymbol{W}}\right)\right)=P_{\mathrm{R}} \tag{16b}$$

$$\frac{\partial J}{\partial q}=0\rightarrow q^4=\frac{\sigma_{n_{\mathrm{U}}}^2 M}{\lambda\left(\mathrm{tr}(\boldsymbol{F}^*\bar{\boldsymbol{W}}^*\bar{\boldsymbol{W}}\boldsymbol{F})+\sigma_{n_{\mathrm{R}}}^2\,\mathrm{tr}(\bar{\boldsymbol{W}}^*\bar{\boldsymbol{W}})\right)} \tag{16c}$$

To directly evaluate $\bar{\boldsymbol{W}}$ from (16), a numerical and iterative search over λ is required. To avoid the iterative procedure, we follow the optimization approach in [9], [10]. When $\lambda \neq 0$ and $\sigma_{n_{\mathrm{R}}}^2 \neq 0$, $\bar{\boldsymbol{W}}$ in (16a) can be represented as

$$\bar{\boldsymbol{W}}(\xi)=\left((\boldsymbol{G}^*)^T\boldsymbol{G}^T+\xi\boldsymbol{I}_M\right)^{-1}(\boldsymbol{G}^*)^T\boldsymbol{F}^*\left(\boldsymbol{F}\boldsymbol{F}^*+\sigma_{n_{\mathrm{R}}}^2\boldsymbol{I}_M\right)^{-1}, \tag{17}$$

which is a function of $\xi \triangleq \lambda q^2$. Substituting (17) into (16b), and using the cyclic property of the trace function, q is also represented as

$$q(\xi)=\sqrt{\frac{P_{\mathrm{R}}}{\mathrm{tr}\left(\bar{\boldsymbol{W}}(\xi)(\boldsymbol{F}\boldsymbol{F}^*+\sigma_{n_{\mathrm{R}}}^2\boldsymbol{I}_M)\bar{\boldsymbol{W}}^*(\xi)\right)}}. \tag{18}$$

Continuing from (17) and (18), which satisfy the conditions in (16a) and (16b), the problem in (15) can be rewritten as

$$\arg\min_{\xi}\left[\mathrm{E}\left\|\boldsymbol{d}-\boldsymbol{G}^T\bar{\boldsymbol{W}}(\xi)\boldsymbol{F}\boldsymbol{d}-\boldsymbol{G}^T\bar{\boldsymbol{W}}(\xi)\boldsymbol{n}_{\mathrm{R}}-q^{-1}(\xi)\boldsymbol{n}_{\mathrm{U}}\right\|^2\right]. \tag{19}$$

Here, we note that the second term multiplied by λ in (15) disappears due to (18) satisfying the power constraint (16b). Since the cost $J(\xi)$ in the square bracket of (19) is convex or strictly quasi-convex with respect to ξ [10], equating the derivative $\frac{\partial J(\xi)}{\partial \xi}$ to zero yields the optimal ξ_o as

$$\xi_o=\sigma_{n_{\mathrm{U}}}^2 P_{\mathrm{R}}^{-1} M. \tag{20}$$

The closed formed MMSE solution of $\boldsymbol{W}$ can then be obtained from (17), (18) and (20) as

$$\boldsymbol{W}_{MMSE}=q(\xi_o)\bar{\boldsymbol{W}}(\xi_o). \tag{21}$$

Note that the solution in (21) satisfies (16). From this fact, we can see that (21) is the solution of the original optimization problem in (14). Also, it can be easily shown that $\boldsymbol{W}_{MMSE}$ becomes identical to $\boldsymbol{W}_{ZF}$ in (13) when $\sigma_{n_{\mathrm{R}}}^2=\sigma_{n_{\mathrm{U}}}^2=0$.

IV. User Selection Algorithms

In this section, we propose optimal and suboptimal criteria for multiuser selection. Here, single user communications[4] are not considered due to low spectral efficiency. We assume that each user treats the interference as noise and the sum achievable rate at slot t is defined as

$$\mathcal{R}(t) \triangleq \frac{M_t}{2}\log_2\left(1+\mathrm{SNR}(t)\right) \tag{22}$$

where the pre-log term M_t appears from the fact that independent M_t data streams are transmitted through the tth slot; the pre-log term $\frac{1}{2}$ comes from the fact that each slot is composed of two phases; and the received SNR at slot t is expressed from (6) as

$$\mathrm{SNR}(t) = \frac{\mathrm{E}\left\|\mathrm{diag}(\boldsymbol{G}^T(t)\boldsymbol{W}(t)\boldsymbol{F}(t))\boldsymbol{d}(t)\right\|^2}{\mathrm{E}\|\mathrm{offd}(\boldsymbol{G}^T(t)\boldsymbol{W}(t)\boldsymbol{F}(t))\boldsymbol{d}(t)\|^2+\mathrm{E}\|\boldsymbol{G}^T(t)\boldsymbol{W}(t)\boldsymbol{n}_{\mathrm{R}}(t)\|^2+\mathrm{E}\|\boldsymbol{n}_{\mathrm{U}}\|^2} \tag{23}$$

Using (22), after supporting all users during one scheduling time T, the average sum rate per slot, i.e., the average sum rate per time, is given by

$$\bar{\mathcal{R}}=\frac{1}{T}\sum_{t=1}^{t=T}\mathcal{R}(t). \tag{24}$$

Noting that the SNR in (23) is a function of $\boldsymbol{F}(t)$ and $\boldsymbol{G}(t)$, we can see that the SNR depends only on $\{a_{t,m}\}$ since the $\{b_{t,m}\}$ are determined by $\{a_{t,m}\}$ as mentioned in (1). Accordingly, the index set $\boldsymbol{\Upsilon}^o$ for the optimal SDMA group selection in terms of $\bar{\mathcal{R}}$ can be obtained via the following optimization:

$$\boldsymbol{\Upsilon}^o=\arg\max_{\boldsymbol{\Upsilon}(M_1,\ldots,M_T)\subseteq\boldsymbol{\Omega}^o}\bar{\mathcal{R}}. \tag{25}$$

In (25), for the given $\{M_t\}$, the number of candidates for a subset $\boldsymbol{\Upsilon}(M_1,\ldots,M_T)=\{(a_{1,1},\ldots,a_{1,M_1}),\ldots,(a_{T,1},\ldots,a_{T,M_T})\}$ of $\boldsymbol{\Omega}^o=\{1,\ldots,2K\}$ is

$$Q^o=\prod_{t=1}^{t=T}\left\{\frac{1}{c(M_t)}\binom{2K-(t-1)M_t}{M_t}\right\}, \tag{26}$$

where $c(M_t)$ gives the number of such M_t-permutations that give the same M_t-combination when the order of M_t is ignored and it can then be expressed as

$$c(M_t)=\begin{cases}1, & \text{if } t=1 \text{ or } M_t=M_{t-1}\\ c(M_t)+1, & \text{if } t\geq 2 \text{ and } M_t\neq M_{t-1}\end{cases}$$

The computational complexity for the cost in (25) with (23) is $\mathcal{O}(M_t^2 N)$ and it might be moderate; however, the combinatorial number Q^o in (26) would be a burden on the relay since it increases exponentially as K increases. Regarding the training for CSI estimation and the computation of $\boldsymbol{W}(t)$ at the relay, the complexity can be assumed independent of the user selection methods. Therefore, to efficiently reduce the computational complexity, we propose suboptimal algorithms avoiding the combinatorial search with reasonable performance degradation.

[4]In single user communications, *every user* transmits by using different time resources or other orthogonal resources such as frequency and code.

A simple suboptimal choice is a rate-based sequential method, in which selects SDMA groups with $\frac{2K}{L}$ users instead of $2K$ users, where L is a positive devisor of $2K$ and $1 \leq L \leq K$. Hence, the optimization is sequentially performed throughout L steps. For the lth step, the rate-based suboptimal method is represented as

$$\boldsymbol{\Upsilon}_l^r = \arg \max_{\boldsymbol{\Upsilon}_l(M_1^l,\ldots,M_{T'}^l)\subseteq \boldsymbol{\Omega}_{l,T'}^r} \frac{1}{T'} \sum_{t=(l-1)T'+1}^{t=lT'} \mathcal{R}(t) \quad (27)$$

where M_t^l represents a number of selected users among $\frac{2K}{L}$ users at the tth slot of the lth step and T' is a slot number depending on M_t^l. In (27), $\boldsymbol{\Omega}_{l,T'}^r$ is an unselected user index set represented by

$$\boldsymbol{\Omega}_{l,T'}^r = \{1,2,\ldots,2K\} - \bigcup_{l'=1}^{l'=l-1} \boldsymbol{\Upsilon}_{l'}^r$$

since the selected users in the MAC phase of the previous $\{l'\}$th steps are discarded in the present lth step for fairness among users. Therefore, for a given $M_t^l = M_t$, the number of possible candidates for $\{\boldsymbol{\Upsilon}_1^r,\ldots,\boldsymbol{\Upsilon}_L^r\}$ can be written as

$$Q^r = \sum_{l=1}^{\max(1,L-1)} \prod_{t=1}^{T'} \left\{ \frac{1}{c(M_t)} \binom{2K-((l-1)T'+t-1)M_t}{M_t} \right\}. \quad (28)$$

Note that the rate-based suboptimal method is identical to the optimal method if we set $L = 1$, and it becomes more simple as L increases.

Another simple selecting choice is an angle-based method. Substituting $\boldsymbol{W}(t)$ in (23) with $\boldsymbol{W}_{ZF}(t)$ in (13), the received SNR in (23) is rewritten as (29) and we can get the lower bound of its denominator as (30), at the bottom of this page. In (30), $\lambda_m(\boldsymbol{A})$ is the mth largest singular value of $\boldsymbol{A}$. Here, the bound, which maximizes the SNR in (29), can be achieved when $\lambda_m(\boldsymbol{F}(t)) = \lambda_{\boldsymbol{F}}$ and $\lambda_m(\boldsymbol{G}(t)) = \lambda_{\boldsymbol{G}}$ for all m, i.e., $\boldsymbol{F}^*(t)\boldsymbol{F}(t) = \lambda_{\boldsymbol{F}}^2 \boldsymbol{I}_{M_t}$ and $\boldsymbol{G}^*(t)\boldsymbol{G}(t) = \lambda_{\boldsymbol{G}}^2 \boldsymbol{I}_{M_t}$. Equivalently, the upper bound of (29) can be achieved when the column vectors of $\boldsymbol{F}(t)$ and $\boldsymbol{G}(t)$ form an orthogonal basis. In accordance with this fact, the angle-based method, which selects $\{a_{t,m}, b_{t,m}\}$th users having the most orthogonal channel vectors relative to the previously selected channel vectors of the users, can be formulated as

$$a_{t,m} = \arg \max_{a_{t,m}\in \boldsymbol{\Omega}_{t,m}^a} \sum_{m'=1}^{m'=m-1} \left(\theta_{\boldsymbol{h}_{a_{t,m'}},\boldsymbol{h}_{a_{t,m}}} + \theta_{\boldsymbol{h}_{b_{t,m'}},\boldsymbol{h}_{b_{t,m}}} \right) \quad (31)$$

with $a_{1,1} = 1$ as an initial setup. In (31), the index set $\boldsymbol{\Omega}_{t,m}^a$ of unselected users is represented as

$$\boldsymbol{\Omega}_{t,m}^a = \{1,2,\ldots,2K\} - \{a_{1,1},\ldots,a_{1,M_1}\} - \cdots - \{a_{t-1,1},\ldots,a_{t-1,M_{t-1}}\} - \{a_{t,1},\ldots,a_{t,m-1}\},$$

and the orthogonality $\theta_{\boldsymbol{a},\boldsymbol{b}}$ between two complex vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ is defined by a Hermitian angle as [12]:

$$\theta_{\boldsymbol{a},\boldsymbol{b}} \triangleq \cos^{-1}\left(\frac{|\boldsymbol{a}^*\boldsymbol{b}|}{\|\boldsymbol{a}\|\|\boldsymbol{b}\|} \right), \; 0 \leq \theta_{\boldsymbol{a},\boldsymbol{b}} \leq \frac{\pi}{2}. \quad (32)$$

Using (32) in (31), the angle-based method can be reformulated as

$$a_{t,m} = \arg \min_{a_{t,m}\in \boldsymbol{\Omega}_{t,m}^a} \sum_{m'=1}^{m'=m-1} \left(\frac{|\boldsymbol{h}_{a_{t,m'}}^* \boldsymbol{h}_{a_{t,m}}|}{\|\boldsymbol{h}_{a_{t,m'}}\|\|\boldsymbol{h}_{a_{t,m}}\|} + \frac{|\boldsymbol{h}_{b_{t,m'}}^* \boldsymbol{h}_{b_{t,m}}|}{\|\boldsymbol{h}_{b_{t,m'}}\|\|\boldsymbol{h}_{b_{t,m}}\|} \right). \quad (33)$$

Contrary to the user selection algorithms in (25) and (27), that in (33) does not include the number of SDMA users (data streams) at the tth slot, i.e., M_t or M_t^l, as a variable. Therefore, M_t should be predetermined. For the low complexity with moderate performance degradation, we set M_t as its minimum or maximum value, respectively, 2 or N. Then, after comparing two $\bar{\mathcal{R}}$'s obtained when $M_t = 2$ and N, the relay decides M_t yielding the larger sum rate. Although the angle-based suboptimal method is designed for the ZF-based relay system, it also works for the MMSE-based relay system as shown later, and it needs to compare only

$$Q^a = \begin{cases} 1, & K = 1 \\ \sum_{t=1}^{t=\lceil \frac{2K}{N}-1 \rceil} \left(\sum_{k=1}^{N-1} (2K-(t-1)N-k) \right), & K > 1 \end{cases} \quad (34)$$

candidates for the SDMA groups. Moreover, the computational complexity for the cost in (33) is $\mathcal{O}(N)$.

In Fig. 3, we depict the numbers of candidates of available user groups, i.e., the number of comparisons, when $N = 2$. Note that $M_t = M_t^l = 2$ for all t since $N = 2$. Q^o in (26) increases exponentially, while Q^r in (28) and Q^a in (34) increase moderately as the number of users increases.

$$\mathrm{SNR}(t) = \frac{M_t}{\sigma_{n_\mathrm{R}}^2 \|\boldsymbol{F}^+(t)\|_F^2 + \frac{M_t\sigma_{n_\mathrm{U}}^2}{P_\mathrm{R}} \|\boldsymbol{G}^+(t)\|_F^2 + \frac{M_t\sigma_{n_\mathrm{R}}^2\sigma_{n_\mathrm{U}}^2}{P_\mathrm{R}} \left\| (\boldsymbol{G}^T(t))^+ \boldsymbol{F}^+(t) \right\|_F^2} \quad (29)$$

$$\begin{aligned} \text{denominator of (29)} &= \sigma_{n_\mathrm{R}}^2 \sum_{m=1}^{m=M_t} \lambda_m^2\left(\boldsymbol{F}^+(t)\right) + \frac{M_t\sigma_{n_\mathrm{U}}^2}{P_\mathrm{R}} \left(\sum_{m=1}^{m=M_t} \lambda_m^2\left(\boldsymbol{G}^+(t)\right) + \sigma_{n_\mathrm{R}}^2 \sum_{m=1}^{m=M_t} \lambda_m^2\left((\boldsymbol{G}^T(t))^+\boldsymbol{F}^+(t)\right) \right) \\ &= \sum_{m=1}^{m=M_t} \frac{\sigma_{n_\mathrm{R}}^2}{\lambda_m^2(\boldsymbol{F}(t))} + \frac{M_t\sigma_{n_\mathrm{U}}^2}{P_\mathrm{R}} \left(\sum_{m=1}^{m=M_t} \frac{1}{\lambda_m^2(\boldsymbol{G}(t))} + \sum_{m=1}^{m=M_t} \frac{\sigma_{n_\mathrm{R}}^2}{\lambda_m^2((\boldsymbol{F}(t)\boldsymbol{G}^T(t)))} \right) \\ &\geq \frac{M_t\sigma_{n_\mathrm{R}}^2}{\lambda_1^2(\boldsymbol{F}(t))} + \frac{M_t\sigma_{n_\mathrm{U}}^2}{P_\mathrm{R}} \left(\frac{M_t}{\lambda_1^2(\boldsymbol{G}(t))} + \frac{M_t\sigma_{n_\mathrm{R}}^2}{\lambda_1^2((\boldsymbol{F}(t)\boldsymbol{G}^T(t)))} \right) \end{aligned} \quad (30)$$

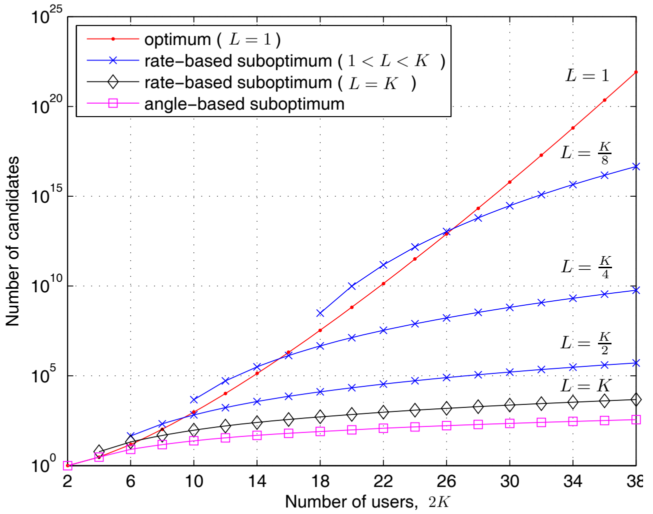


Fig. 3. Comparison of the number of candidate user groups for optimal (26), rate-based suboptimal (28), and angle-based suboptimal (34) when $N = 2$ and $M_t = M_t^l = 2$.

Obviously, $Q^r = Q^o$ when $L \leq 1$, which is not depicted. For a certain small number of users, it is observed that Q^r is larger than Q^o. As an example, when $2K = 4$, the optimal scheme compares three candidates $\{(1,2),(3,4)\}$, $\{(1,3),(2,4)\}$, and $\{(1,4),(2,3)\}$ for two SDMA groups, while the rate-based suboptimal scheme compares six candidates $\{(1,2)\}$, $\{(1,3)\}$, $\{(1,4)\}$, $\{(2,3)\}$, $\{(2,4)\}$, and $\{(3,4)\}$ for the first SDMA group. It is nevertheless obvious that the proposed suboptimal methods can substantially reduce the computational complexity at the relay as K increases. However, at the same time, it should be verified that the performance degradation of the suboptimal methods is not significant compared to the optimal method. To confirm it, we will evaluate and compare the performance of the optimal and suboptimal methods with respect to the achievable rate.

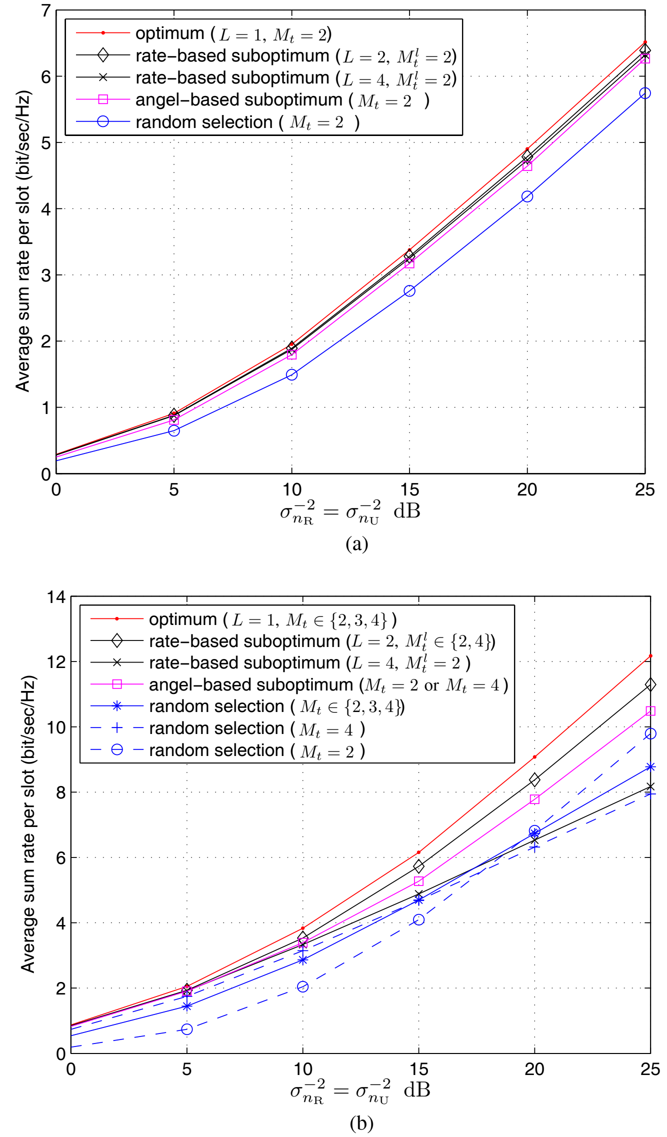


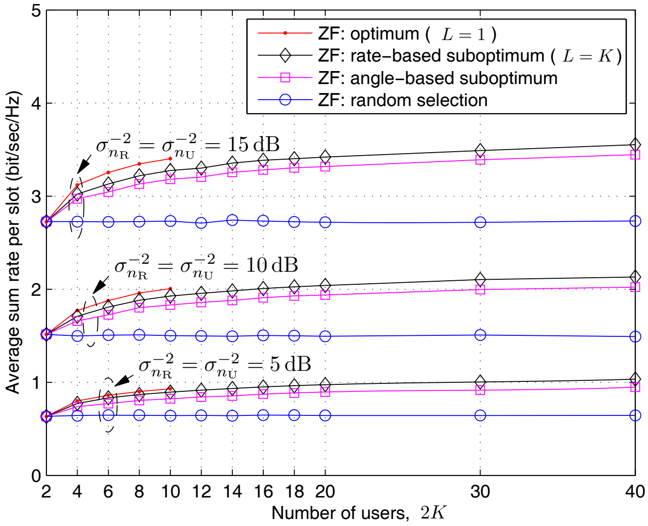
Fig. 4. Average sum rates per slot in (24) of ZF-based SDMA systems when $2K = 8$ and $P_R = 1$. (a) $N = 2$. (b) $N = 4$.

V. Simulation Results

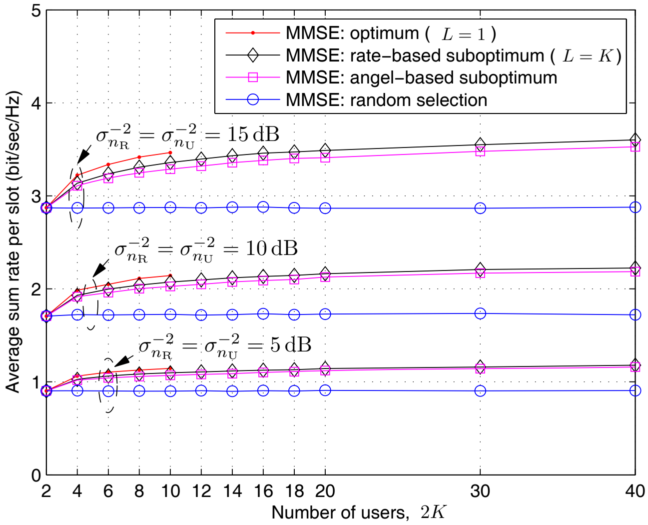
We compare the average sum rates per slot in (24) for four scheduling methods: optimal, rate-based suboptimal, angle-based suboptimal and random selection methods. The random selection method selects M_t SDMA users randomly but exclusively at each time slot. Letting $P_R = 1$, the received SNRs at the relay and the users are defined as $\sigma_{n_R}^{-2}$ and $\sigma_{n_U}^{-2}$, respectively.

In Fig. 4, the average sum rates of ZF-based systems are evaluated against the received SNRs when $2K = 8$. As expected, we can see a tradeoff between complexity and performance. When $N = 2$ as depicted in Fig. 4(a), the available number of SDMA users in each slot is two for all algorithms, i.e., $M_t = M_t^l = 2$. Hence, the suboptimum schemes achieve almost similar performance to the optimal scheme. The average loss rates of the suboptimal methods are $2.5(3.5)\%$ and 8.0% for the rate-based suboptimal scheme with $L = 2(4)$ and the angle-based suboptimal method, respectively. Note that the increase in rates compared to the random selection method are, respectively, $25.8(24.5)\%$ and 18.3%. However, when $N = 4$, the performance gap between the optimal and suboptimal schemes increases as shown in Fig. 4(b). The average loss rates are, respectively, $6.3(17.4)\%$ and 11.1%, while the increased rates are, respectively, $31.5(16.4)\%$ and 24.9%. In contrast to the optimal scheme, in which M_t can be any choice satisfying (2), the suboptimal schemes have a restriction on M_t as presented in Fig. 4(b), resulting in higher performance loss. From the random selection method with the values of M_t at 2 and 4 in our simulations, we can see the effect of M_t on the system performance.

Figures 5(a) and (b) show the average rates per slot versus the number of users in the ZF- and MMSE-based systems, respectively, when $N = 2$. As expected, the average rate of the proposed suboptimal scheduling methods place themselves between those of the optimal and the random selection methods. Due to the computational complexity, we show the average sum rate of the optimal scheduling method from 2 up to 10 users in simulation. When there is only one user pair ($2K = 2$), obviously the average sum rates of all schemes are identical. The average rates of the rate-based (angle-based)



(a)



(b)

Fig. 5. Average sum rates per slot in (24) for optimal ($L = 1$), rate-based ($L = K$), angle-based, and random selection schemes versus the number of users when $N = 2$ and $P_R = 1$. (a) ZF-based SDMA. (b) MMSE-based SDMA.

suboptimal method for the ZF-based system are decreased by $3.7(11.2)\%$, $3.9(8.4)\%$, and $3.7(6.4)\%$ compared to the optimal method, when both of $\sigma_{n_R}^{-2}$ and $\sigma_{n_U}^{-2}$ are $5\,\mathrm{dB}$, $10\,\mathrm{dB}$, and $15\,\mathrm{dB}$, respectively; however, these are increased by $35.7(26.2)\%$, $25.5(20.1)\%$, and $18.2(15.2)\%$ compared to the random selection method. For the MMSE-based system, the rate losses of the rate-based (angle-based) suboptimal method are $3.9(5.9)\%$, $3.1(8.4)\%$, and $3.7(6.4)\%$, respectively, while the gains are $19.6(17.3)\%$, $18.4(18.0)\%$, and $15.4(13.4)\%$ compared to the random selection method.

From these results, it can be surmised that the rate-based scheme with $L = K$ when $N = 2$ can achieve close performance in less than 4% loss to the optimal scheme with the extremely reduced complexity (see Fig. 3). It can be also seen that the average sum rates per slot, except that of the random selection method, increase as the number of total users increases, i.e., all schemes except the random selection method can obtain multiuser diversity gain.

VI. Conclusion

For multiuser two-way relay systems, SDMA-based relay processing matrices are designed. Also, an optimal scheduling method maximizing the average sum rate and its suboptimal methods reducing complexity are proposed. A trade-off between complexity and performance can be verified. Especially, when the relay has two antennas, it is shown that the proposed suboptimal scheduling methods can achieve significant complexity reduction with some tolerable sacrifice in performance.

References

[1] S. Zhang, S.-C. Liew, and P. Lam, “Physical layer network coding,” in *Proc. ACM Mobile Comput. Netw. (MobiCom)*, Los Angeles, USA, Sep. 2006, pp. 358–365.

[2] S. Katti, H. Rahul, W. Hu, D. Katabi, M. Medard, and J. Crowcroft, “XORs in the air: practical wireless network coding,” *IEEE/ACM Trans. Networking*, vol. 16, pp. 497–510, June 2008.

[3] S. Katti, S. Gollakota, and D. Katabi, “Embracing wireless interference: analog network coding,” in *Proc. ACM Special Interest Group Data Commun. (SIGCOMM)*, Kyoto, Japan, Aug. 2007, pp. 397–408.

[4] 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, June 2009.

[5] S. J. Kim, P. Mitran, and V. Tarokh, “Performance bounds for bidirectional coded cooperation protocols,” *IEEE Trans. Inf. Theory*, vol. 54, pp. 5235–5241, Nov. 2008.

[6] Z. Yi and I.-M. Kim, “An opportunistic-based protocol for bidirectional cooperative networks,” *IEEE Trans. Wireless Commun.*, vol. 8, pp. 4836–4847, Sep. 2009.

[7] M. Chen and A. Yener, “Multiuser two-way relaying: detection and interference management strategies,” *IEEE Trans. Wireless Commun.*, vol. 8, pp. 4296–4303, Aug. 2009.

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

[9] J. Joung and A. H. Sayed, “Multiuser two-way relaying method for beamforming systems,” in *Proc. IEEE International Workshop Signal Process. Advanced Wireless Commun. (SPAWC)*, Perugia, Italy, June 2009.

[10] J. Joung and A. H. Sayed, “Multiuser two-way amplify-and-forward relay processing and power control methods for beamforming systems,” *IEEE Trans. Signal Process.*, vol. 58, pp. 1833–1846, Mar. 2010.

[11] Z. Jian, M. Kuhn, A. Wittneben, and G. Bauch, “Self-interference aided channel estimation in two-way relaying systems,” in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, New Orleans, US, Nov. 2008, pp. 1–6.

[12] W. M. Goldman, *Complex Hyperbolic Geometry*. New York: Oxford University Press, 1999.