

Wireless MIMO Switching

Fanggang Wang^{*†} and Soung Chang Liew[†]

^{*}State Key Lab of Rail Traffic Control and Safety, EIE, Beijing Jiaotong Univ., Beijing, China

[†]Institute of Network Coding, The Chinese Univ. of Hong Kong, Hong Kong, China

Abstract— In a generic switching problem, a switching pattern consists of a one-to-one mapping from a set of inputs to a set of outputs (i.e., a permutation). We propose and investigate a wireless switching framework in which a multi-antenna relay is responsible for switching traffic among a set of N stations. We refer to such a relay as a MIMO switch. With beamforming and linear detection, the MIMO switch controls which stations are connected to which stations. Each beamforming matrix realizes a permutation pattern among the stations. We refer to the corresponding permutation matrix as a switch matrix. By scheduling a set of different switch matrices, full connectivity among the stations can be established. In this paper, we focus on “fair switching” in which equal amounts of traffic are to be delivered for all the $N(N-1)$ ordered pairs of stations. In particular, we investigate how the system throughput can be maximized. In general, for large N the number of possible switch matrices $N!$ is huge, making the scheduling problem combinatorially challenging. We show that for $N = 4$ and 5 , only a subset of $N-1$ of the $N!$ switch matrices need to be considered in the scheduling problem to achieve good throughput. We conjecture that this will be the case for large N as well. This conjecture, if valid, implies that for practical purposes, fair-switching scheduling is not an intractable problem.

Index Terms—Fairness, MIMO switching, physical-layer network coding, relay, scheduling.

I. INTRODUCTION

Relaying in wireless networks plays a key role in various communication applications [1]. The use of relays can extend coverage as well as improve energy efficiency [2]. In this paper, we study a set-up in which N stations communicate with each other via a multi-antenna relay. With beamforming, the relay controls which stations are connected to which other stations. Each beamforming matrix realizes a permutation among the stations represented by a switch matrix. By scheduling a set of different switch matrices, full connectivity among the stations can be established.

Prior work that investigated N stations exchanging data via a relay includes [2]–[4]. Ref. [2] studied “pairwise data exchange”. For pairwise data exchange, if station i transmits to station j , then station j transmits to station i as well, and the two stations communicate with each other only. In [2], MIMO relays with different forwarding strategies were considered. Ref. [3] also studied pairwise data exchange, but the relay

adopts the decode-and-forward strategy only. The diversity-multiplexing tradeoffs under reciprocal and non-reciprocal channels were analyzed. Pairwise data exchange as studied in [2], [3] assumes a station communicates with one other station only. In a general setting, a station could have data for more than one station. In this paper, we focus on a uniform traffic setting in which the amounts of traffic from station i to station j are the same for all $i, j \in \{1, \dots, N\}$, $i \neq j$. “Fair switching” is used to meet the uniform traffic requirement. Specifically, fair switching is realized by scheduling a set of switch matrices. To the best of our knowledge, the framework of fair switching has not been considered in the existing literature.

Refs. [4] investigated the case of full data exchange, in which all stations want to broadcast their data to all other stations.¹ Data transmissions in [4] can be summarized as follows: in the first slot, all stations transmit to the relay simultaneously; the first slot is followed by multiple slots for downlink transmissions; in each downlink slot, the relay multiplies the signal received in the first time slot by a different beamforming matrix, such that at the end of all downlink slots, all stations receive the broadcast data from all other stations. By contrast, the framework investigated in this paper is more general in that it can accommodate the pure unicast case, the mixed unicast-multicast case, as well as the pure broadcast case as in [4]. In particular, a station i can have M_i data streams, and each station $j \neq i$ is a target receiver of one of the M_i streams.

In our framework, the MIMO relay serves as a general switch that switches traffic among the stations. We use beamforming at the relay and linear detection to realize different connectivity patterns among the stations. Each beamforming matrix realizes a permutation among the stations represented by a switch matrix. By scheduling a set of switch matrices, the MIMO switching system can realize any general traffic-flow pattern (unicast, multicast, broadcast, or a mixture of them) among the stations.

Before delving into technical details, we provide a simple example to illustrate the scenario of interest to us here. Consider a network with three stations, 1, 2, and 3. The traffic flows among them are shown in Fig. 1: station 1 wants to deliver “ a ” to both stations 2 and 3; station 2 wants to deliver “ b ” and “ c ” to stations 1 and 3, respectively; and station 3

¹Note that full data exchange is also discussed in [2]. But only a single-antenna relay was considered.

This work was partially supported by grants from the Univ. Grants Committee of the Hong Kong, China (AoE/E-02/08; 414812); the State Key Lab of Rail Traffic Control and Safety (RCS2011ZT011); the Fundamental Research Funds for the Central Universities (2010RC007, 2011JBM203); Program for Changjiang Scholars and Innovative Research Team in Univ. (IRT0949); the Joint Funds of State Key Program of NSFC (60830001); the China 973 Program (2012CB315904).

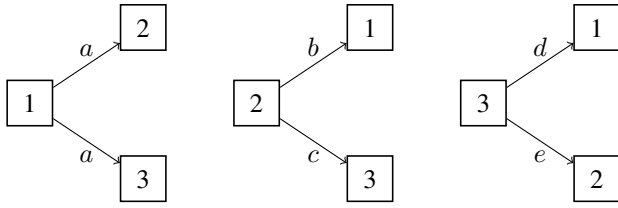


Fig. 1. Traffic demand of a three stations example.

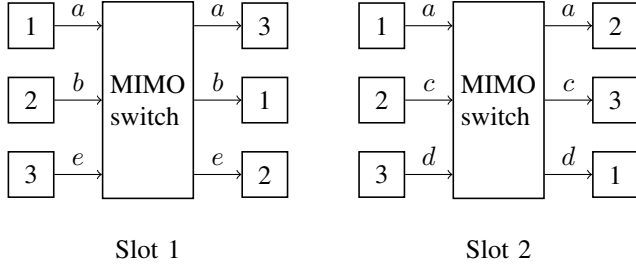


Fig. 2. A transmission established by two slots of unicast connectivity realizes the traffic demand in Fig. 1.

wants to deliver “ d ” and “ e ” to stations 1 and 2, respectively. Pairwise data exchange as in [2] and [3] is not effective in this case because when the number of stations is odd, one station will always be left out when forming pairs. That is, when the number of stations is odd, the connectivity pattern realized by a switch/permutation matrix does not correspond to pairwise communication. Full data exchange is not appropriate either, since in our example, station 2 (as well as station 3) wants to deliver different data to the other two stations. Under our framework, the traffic flows among stations can be met as shown in Fig. 2. In the first slot, station 1 transmits “ a ” to station 3; station 2 transmits “ b ” to station 1; and station 3 transmits “ e ” to station 2. In the second slot, station 1 transmits “ a ” to station 2; station 2 transmits “ c ” to station 3; and station 3 transmits “ d ” to station 1. In Section III.C, we will present the details on how to realize the switch matrices. To limit the scope, this paper focuses on the use of amplify-and-forward relaying and zero forcing (ZF) to establish the permutations among stations.

The rest of the paper is organized as follows: Section II describes the framework of wireless MIMO switching and introduces the ZF relaying method for establishing permutations among stations. The fair switching framework is presented in Section III. Section IV discusses our simulation results. Section V concludes this paper.

II. SYSTEM DESCRIPTION

A. System Model

Consider N stations, S_1, \dots, S_N , each with one antenna, as shown in Fig. 3. The stations communicate via a relay R with K antennas and there is no direct link between any two stations. In this paper, we assume $K \geq N$ so that the relay has enough degree of freedom to switch N flows at a time. Without loss of generality, we further assume $K = N$ since

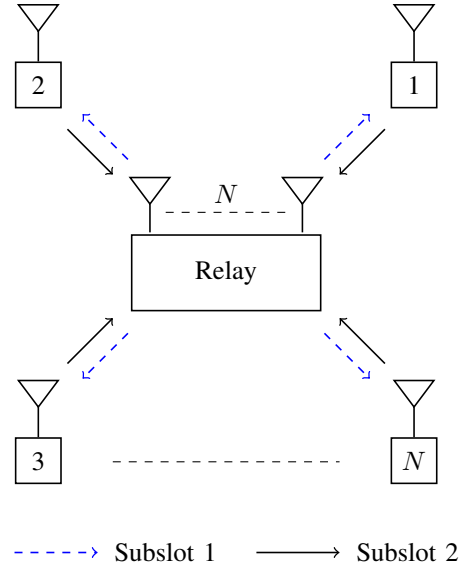


Fig. 3. Wireless MIMO switching.

the case of $K > N$ can be dealt with in a similar way, with the difference that instead of matrix inverses, sometimes pseudo inverses are used in the signal processing of the $K > N$ case, as will be explained later. Each time slot is divided into two subslots. The first subslot is for uplink transmissions from the stations to the relay; the second subslot is for downlink transmissions from the relay to the stations. We assume the two subslots are of equal duration. Each time slot realizes a switching permutation, as described below.

Consider one time slot. Let $\mathbf{x} = \{x_1, \dots, x_N\}^T$ be the vector representing the signals transmitted by the stations. We assume all stations use the same transmit power, normalized to one. Thus, $\mathbb{E}\{x_i^2\} = 1, \forall i$. We also assume that $\mathbb{E}\{x_i\} = 0, \forall i$, and that there is no cooperative coding among the stations so that $\mathbb{E}\{x_i x_j\} = 0, \forall i \neq j$. Let $\mathbf{y} = \{y_1, \dots, y_N\}^T$ be the received signals at the relay, and $\mathbf{v} = \{v_1, \dots, v_N\}^T$ be the noise vector with i.i.d. noise samples following the complex Gaussian distribution, i.e., $v_n \sim \mathcal{N}_c(0, \sigma_r^2)$. Then

$$\mathbf{y} = \mathbf{H}_u \mathbf{x} + \mathbf{v}, \quad (1)$$

where \mathbf{H}_u is the uplink channel gain matrix. The relay multiplies \mathbf{y} by a beamforming matrix \mathbf{G} before relaying the signals. We impose a power constraint on the signals transmitted by the relay so that

$$\mathbb{E}\{\|\mathbf{G}\mathbf{y}\|^2\} \leq p. \quad (2)$$

Combining (1) and (2), we have

$$\text{Tr}[\mathbf{H}_u^H \mathbf{G}^H \mathbf{G} \mathbf{H}_u] + \text{Tr}[\mathbf{G}^H \mathbf{G}] \sigma_r^2 \leq p. \quad (3)$$

Let \mathbf{H}_d be the downlink channel matrix. Then, the received signals at the stations in vector form are

$$\mathbf{r} = \mathbf{H}_d \mathbf{G} \mathbf{y} + \mathbf{w} = \mathbf{H}_d \mathbf{G} \mathbf{H}_u \mathbf{x} + \mathbf{H}_d \mathbf{G} \mathbf{v} + \mathbf{w}, \quad (4)$$

where \mathbf{w} is the noise vector at the receiver, with the i.i.d. noise samples following the complex Gaussian distribution, i.e., $w_n \sim \mathcal{N}_c(0, \sigma^2)$.

B. MIMO Switching

Suppose that the purposes of \mathbf{G} are to realize a particular permutation represented by the permutation matrix \mathbf{P} , and to amplify the signals coming from the stations. That is,

$$\mathbf{H}_d \mathbf{G} \mathbf{H}_u = \mathbf{A} \mathbf{P}, \quad (5)$$

where $\mathbf{A} = \text{diag}\{a_1, \dots, a_N\}$ is an ‘‘amplification’’ diagonal matrix. Given \mathbf{A} and \mathbf{P} , when \mathbf{H}_d and \mathbf{H}_u are invertible, the beamforming matrix \mathbf{G} is calculated as²

$$\mathbf{G} = \mathbf{H}_d^{-1} \mathbf{A} \mathbf{P} \mathbf{H}_u^{-1}. \quad (6)$$

Define $\hat{\mathbf{r}} = \mathbf{A}^{-1} \mathbf{r}$, i.e., station S_j divides its received signal by a_j . We can rewrite (4) as

$$\hat{\mathbf{r}} = [x_{i_1}, \dots, x_{i_j}, \dots, x_{i_N}]^T + \mathbf{P} \mathbf{H}_u^{-1} \mathbf{v} + \mathbf{A}^{-1} \mathbf{w}, \quad (7)$$

where S_{i_j} is the station transmitting to S_j under the permutation \mathbf{P} (i.e., in row j of \mathbf{P} , element i_j is one, and all other elements are zero). Suppose that we require the received signal-to-noise ratio (SNR) of each station to be the same. Let $h_{u,(i,j)}^{(-1)}$ be element (i, j) in \mathbf{H}_u^{-1} . Then

$$\sigma_r^2 \sum_k |h_{u,(i,j,k)}^{(-1)}|^2 + \frac{\sigma^2}{|a_j|^2} = \sigma_e^2, \quad \forall j. \quad (8)$$

σ_e^2 is the effective noise power for each station under unit signal power. Note that we actually use a zero-forcing relaying scheme in (5). For other schemes such as MMSE, it is difficult to require the interference plus noise power of each station to be the same.

Substituting (5) into (3), we have

$$q \triangleq \sum_{i,j} |h_{d,(i,j)}^{(-1)}|^2 |a_j|^2 + \sigma_r^2 \sum_{i,k} \left| \sum_j h_{d,(i,j)}^{(-1)} a_j h_{u,(i,j,k)}^{(-1)} \right|^2 \leq p, \quad (9)$$

where $h_{d,(i,j)}^{(-1)}$ is element (i, j) in \mathbf{H}_d^{-1} . Let $a_j = |a_j| e^{i\theta_j}$, then combining (8) and (9) gives

$$q = \sum_{i,j} \frac{|h_{d,(i,j)}^{(-1)}|^2 \sigma^2}{\sigma_e^2 - \sigma_r^2 \sum_k |h_{u,(i,j,k)}^{(-1)}|^2} + \sigma_r^2 \sum_{i,k} \left| \sum_j \frac{h_{d,(i,j)}^{(-1)} h_{u,(i,j,k)}^{(-1)} \sigma e^{i\theta_j}}{\sqrt{\sigma_e^2 - \sigma_r^2 \sum_k |h_{u,(i,j,k)}^{(-1)}|^2}} \right|^2 \leq p. \quad (10)$$

Problem Definition 1: Given $\mathbf{H}_u, \mathbf{H}_d, p, \sigma^2, \sigma_r^2$, and a desired permutation \mathbf{P} , solve for minimum σ_e^2 and the corresponding \mathbf{G} .

This problem is non-convex. We can use a well-known global optimization algorithm, particle swarm optimization, to

solve this problem. However, the method of particle swarm optimization is generally computationally complex. We propose a random-phase algorithm, as follows, to reduce complexity.

Random-phase Algorithm: For a given set of θ_j , $j = 1, \dots, N$, according to (10), $\lim_{\sigma_e^2 \rightarrow +\infty} q = 0$, and $\lim_{\sigma_e^2 \rightarrow 0} q = +\infty$. Furthermore, q is a continuous function of σ_e . Thus, there exists a σ_e such that $q = p$. Denote such a σ_e by $\sigma_e(\theta_1, \dots, \theta_N)$. The problem consists of finding

$$\sigma_e^* = \arg \min_{\theta_1, \dots, \theta_N} \sigma_e(\theta_1, \dots, \theta_N). \quad (11)$$

We note that σ_e is a complicated nonlinear function of θ_j . A time-consuming exhaustive search can be used to find the solution to (11). We use a random-phase algorithm to reduce the complexity. We quantize the interval of $[0, 2\pi)$ equally into M bins with the values of $0, \frac{2\pi}{M}, \dots, \frac{2(M-1)\pi}{M}$ respectively and randomly pick among them to set the value of θ_j for each and every $j = 1, \dots, N$. After that, we compute the corresponding $\sigma_e(\theta_1, \dots, \theta_N)$ by solving (10) with the inequality set to equality. We perform L trials of these random phase assignments to obtain L different values of σ_e . We choose the smallest among them as our estimate for σ_e^* . Substituting the estimated σ_e^* into (8) yields $|a_j|$ for all j ; hence \mathbf{G} . Note that the solution found is a feasible solution and is in general larger than the actual optimal σ_e^* . In Section IV, we will show that large gains can be achieved with only small M and L . Moreover, increasing M and L further yields very little improvement, suggesting that the estimated σ_e^* with small M and L is near optimal.

III. FAIR SWITCHING

In this section, we study the fair switching scenario in which each station has an equal amount of traffic for every other station. The data from station i to station j could be different for different j , so this is not restricted to the multicast or broadcast setting. As has been described in the previous section, in each time slot, the stations transmit to one another according to a switch matrix. To achieve fair switching, multiple transmissions using a succession of different switch matrices over different time slots will be needed. We next discuss the set of switch matrices.

A. Derangement

We assume a station does not deliver traffic flow to itself. A derangement is a permutation in which i is not mapped to itself [5]. While the number of distinct permutations with N stations is $N!$, the number of derangements is given in [5] as

$$d_N = N! \sum_{i=0}^N \frac{(-1)^i}{i!}. \quad (12)$$

For example, $d_4 = 9$ although the number of permutations is $4! = 24$. It can be shown that $\lim_{N \rightarrow \infty} \frac{d_N}{N!} = e^{-1}$ and the limit is approached quite quickly. Thus, the number of derangements is in general still very large for large N .

²When $K > N$, we use pseudo inverse of \mathbf{H}_d and \mathbf{H}_u in (6).

TABLE I
DERANGEMENTS OF $N = 4$.

$$\begin{aligned}
 P_1 &= \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}, & P_2 &= \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, & P_3 &= \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \\
 P_4 &= \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}, & P_5 &= \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, & P_6 &= \begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \\
 P_7 &= \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}, & P_8 &= \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, & P_9 &= \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}.
 \end{aligned}$$

Performing optimization over this large combinatorial set of derangements is a formidable task. For example, in our fair switching problem, we want to maximize the system throughput by scheduling over a subset of derangements. It would be nice if for our problem, the optimal solution is not very sensitive to the particular selection of derangements. In Part B, we will formalize the concept of ‘‘condensed derangement sets’’.

B. Condensed Derangement Set

Definition 1: A set of $N - 1$ derangements, D_1, D_2, \dots, D_{N-1} , is said to be a *condensed derangement set* if

$$\sum_{n=1}^{N-1} D_n = J - I, \quad (13)$$

where J is a matrix with all ‘‘1’’ elements, and I is the identity matrix.

The four condensed derangement sets for $N = 4$ are $\mathcal{Q}_1 = \{P_1, P_5, P_9\}$, $\mathcal{Q}_2 = \{P_1, P_6, P_8\}$, $\mathcal{Q}_3 = \{P_2, P_4, P_9\}$, and $\mathcal{Q}_4 = \{P_3, P_5, P_7\}$, where P_n are listed in TABLE I. There are $d_5 = 44$ derangements for $N = 5$ and the number of condensed derangement sets is 56.

In fair switching, we want to switch an equal amount of traffic from any station i to any station j , $i \neq j$. This can be achieved by scheduling the derangements in the condensed derangement set in a weighted round-robin manner (as detailed in ‘‘Approach to Problem 2’’ below). Given a condensed set, the scheduling to achieve fair switching is rather simple, as elaborated in ‘‘Approach to Problem 2’’ below. However, different condensed sets could potentially yield solutions of different performance. And the number of condensed derangement sets is huge for large N . We define a problem as follows.

Problem Definition 2: Suppose that we want to deliver equal amounts of traffic from S_i to $S_j \forall i \neq j$. Which condensed derangement sets should be used to schedule transmissions to maximize throughput? Does it matter?

Approach to Problem 2: The derangements in a condensed derangement set are the building blocks for scheduling. For

a condensed derangement set, we may schedule derangement D_n for k_n time slots. The length of a complete round of transmissions for that condensed set will then be $\sum_{n=1}^{N-1} k_n$.

Consider the case of $N = 4$. There are four condensed derangement sets. The question is which condensed derangement set will result in the highest throughput. We could approach the problem as follows.

Let $\mathcal{Q}_m = \{D_1^m, D_2^m, \dots, D_{N-1}^m\}$ be a particular condensed derangement set. For each D_n^m , we use random-phase algorithm above to compute the corresponding σ_e^2 , denoted by $\sigma_{e,n,m}^2$. The Shannon rate is then

$$r_{n,m} = \log\left(1 + \frac{1}{\sigma_{e,n,m}^2}\right). \quad (14)$$

Because of the uniform traffic assumption, we require $k_{n,m} r_{n,m} = c, \forall n \in [1, \dots, N - 1]$, for some c . That is, c is the amount of traffic delivered from one station to another station in one round of transmissions. The effective throughput per station (i.e., the amount of traffic from a station to all other stations) is

$$T_m = \frac{(N - 1)c}{\sum_{n=1}^{N-1} k_{n,m}} = \frac{N - 1}{\sum_{n=1}^{N-1} 1/r_{n,m}}. \quad (15)$$

Numerically, we could first solve for $r_{n,m} \forall n$. Then, we apply (15) to find the throughput.

The question we want to answer is whether T_m for different \mathcal{Q}_m are significantly different. For the case of $N = 4$ and 5, we will show simulation results indicating that the throughputs of different \mathcal{Q}_m are rather close, and therefore it does not matter which \mathcal{Q}_m we use.

To summarize our discussion so far, recall that a central proposal of this paper is a framework for MIMO switching. In particular, to simplify the scheduling problem, we aim to achieve equal SNR for all connections in each derangement. Different derangements, however, may still have different SNR. Thus, in principle, the throughput T_m as in (15) may still be different for different condensed set m . Fortunately, our simulation results indicate that T_m for different m are roughly the same for $N = 4$ and $N = 5$.

C. Generalization

As mentioned in the introduction, most prior works for multi-way relay networks focus on two traffic-flow patterns. The first is pairwise unicast, in which stations form pairs, and the two stations of a pair communicate with each other only [2], [3]. The second is the full data exchange, in which each station needs to broadcast to all the other stations [4]. In practice, however, the actual traffic-flow patterns could be different from these two patterns. For example, for video conferencing, a subset of stations within the network forms a multicast group, and the traffic-flow pattern is somewhere between the two extremes above.

More generally, broadcast sessions, multicast sessions, pairwise unicast sessions, and unidirectional unicast sessions could co-exist in the same network. The MIMO switching framework here is flexible and encompasses this generality. For easy

explanation, our previous discussion in Part B has an implicit assumption that each station i wants to deliver different data to different stations $j \neq i$. If we examine the scheme carefully, this assumption is not necessary. In the scheme, a station will have chances to deliver the same data to all other stations. In particular, a station i will have chances to deliver the same data to two different stations j and k in two different derangements. If so desired, station i could transmit the same data to stations j and k in the two derangements. This observation implies that the general traffic-flow pattern can be realized.

For illustration, let us examine how the traffic pattern of Fig. 1 can be realized. This example is a pattern consisting of the co-existence of unicast and broadcast. As has been described, the data transmission can be realized by scheduling a condensed derangement set, which is $\mathbf{D}_1 = [e_3, e_1, e_2]$ and $\mathbf{D}_2 = [e_2, e_3, e_1]$, and e_n contains 1 in the n th position and zeros elsewhere. The transmitted data of station 1, 2 and 3 are respectively $[a, b, e]^T$ for \mathbf{D}_1 and $[a, c, d]^T$ for \mathbf{D}_2 .

We remark that there is an alternative way to realize a mix of unicast and broadcast sessions. The derangement/switch matrix \mathbf{P} discussed in Part A of this section can be generalized to accommodate multicast sessions. In particular, to support multicast sessions, instead of a permutation matrix, each column of the switch matrix can contain more than one “1” element. Each row, however, still has at most one “1” element. To limit the scope of this paper, we defer the study of this alternative method for the future.

Furthermore, the MIMO switch so far makes use of zero-forcing relaying. In general, there is no need to force a diagonal element to zero because the self-interference is known by the transmitter, and can be removed out by the transmitter. This is the basic idea behind physical-layer network coding [6], which underlies many other works, e.g., [7]–[9]. By removing the constraint on the diagonals, there is extra potential in our design to improve the system performance. In this case we rewrite (5) as

$$\mathbf{H}^T \mathbf{G} \mathbf{H} = \mathbf{A}(\mathbf{P} + \mathbf{B}), \quad (16)$$

where \mathbf{B} is a diagonal matrix $\mathbf{B} = \text{diag}\{b_1, \dots, b_N\}$. More details are available in [10].

IV. SIMULATION

In this section, we evaluate the throughputs achieved by different condensed sets. We assume that the uplink channel \mathbf{H}_u and downlink channel \mathbf{H}_d are reciprocal, i.e., $\mathbf{H}_d = \mathbf{H}_u^T$, and that they both follow the complex Gaussian distribution $\mathcal{N}_c(\mathbf{0}, \mathbf{I})$. We assume the relay has the same transmit power as all the stations, i.e., $p = 1$.

We will answer the question raised in Problem Definition 2. We analyze the scenarios where $N = 4$ and $N = 5$. The four different condensed derangement sets of $N = 4$, \mathbb{Q}_1 , \mathbb{Q}_2 , \mathbb{Q}_3 and \mathbb{Q}_4 , are considered for fair switching. For each channel realization, we evaluate the throughput per station T_m as defined by (15). We simulated a total of 10000 channel realizations and computed $\mathbb{E}\{T_m\}$ averaged over the channel realizations. Recall that for the random-phase algorithm, there

are two associated parameters: number of trials L and number of bins M (see Section II). We find that for a fixed L and a fixed M , the four condensed derangement sets yield essentially the same throughput (within 1% in the medium and high SNR regimes and within 2% in the low SNR regime). Fig. 4 plots the throughput for one of the condensed derangement set for different L and M . For $N = 5$ there are 56 different condensed derangement sets. As with the $N = 4$ case, all the sets have roughly the same throughput (within 1%). Fig. 5 also plots the results of one set. We also note that increasing L and M beyond 10 and 8 respectively yields little throughput gain. This implies that our heuristic yields near optimal result when $L = 10$ and $M = 8$.

For $N = 6$ there are 265 derangements and with 9408 corresponding condensed derangement sets. The simulation to evaluate different condensed sets is very time consuming. Furthermore, as mentioned in Section III.A, the number of derangements increases in the order of $N!$. It is a formidable task to optimize the system throughput over the huge number of different condensed sets.

We conjecture that different condensed derangement sets achieve roughly the same throughput for N larger than 5 as well. A concrete proof remains an open problem. The ramification of this result, if valid, is as follows. For large N , the number of condensed derangement set is huge, and choosing the optimal set is a complex combinatorial problem. However, if their relative performances do not differ much, choosing any one of them in our engineering design will do, significantly simplifying the problem.

A scheme proposed in [11] investigates a similar problem as ours. It simply uses a positive scalar weight to control the relay power consumption instead of our diagonal \mathbf{A} . As a comparison, we also plot the throughputs of the scalar scheme in [11]. Our scheme with diagonal \mathbf{A} outperforms the scalar scheme by 1 dB and 0.5 dB in Fig. 4 and Fig. 5. Besides the advantage in throughput, our scheme has another advantage over the scalar scheme in that our scheme guarantees fairness. That is, in our scheme, each station has exactly the same throughput, while the stations in the scheme in [11] could have varying throughputs. The scalar scheme in [11] focuses on optimizing the sum rate of all stations; the individual rates of the stations may vary widely with only one degree of freedom given by the scalar. It is worth emphasizing that how the throughputs vary among the stations in [11] is not controllable, whereas our fair switching scheme can be easily generalized to the case where the relative throughputs among the stations are specified. We also plot the throughputs of the computationally intensive particle swarm optimization scheme in Fig. 4 and Fig. 5 for benchmarking purposes. With 20 particles and 20 iterations, the particle swarm optimization scheme has roughly the same throughput as our random-phase scheme, suggesting that our random-phase scheme is near optimal while being less complex.

The following general observation and conjecture sum up our investigation:

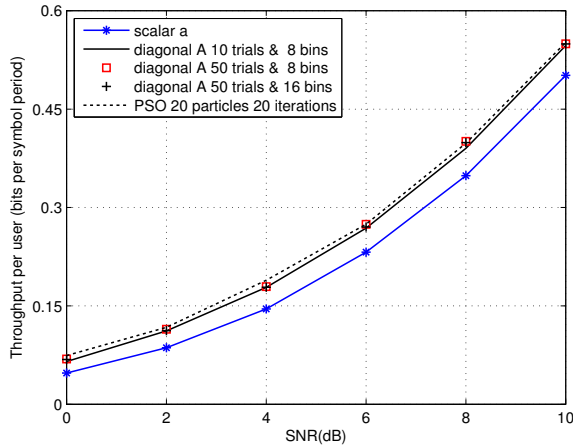


Fig. 4. Throughput per station under MIMO fair switching when $N = 4$. In each case only the result of one condensed derangement set is presented because the results of other derangement sets are within 2% of the results shown here.

General Observation: In our framework of MIMO fair switching with 4 or 5 stations, any condensed derangement set can be used because different condensed derangement sets achieve roughly the same throughput. We conjecture that this will be the case when the number of stations is large as well. If this conjecture holds, then the issue of condensed set selection will go away, and the complexity of the optimization problem will be greatly reduced. This conjecture remains to be proven.

V. CONCLUSIONS

We have proposed a framework for wireless MIMO switching to facilitate communications among multiple wireless stations. A salient feature of our MIMO switching framework is that it can cater to general traffic patterns consisting of a mixture of unicast traffic, multicast traffic, and broadcast traffic flows among the stations.

There are many nuances and implementation variations arising out of our MIMO switching framework. In this paper, we have only studied the “fair switching” setting in which each station wants to deliver equal amounts of traffic to all other stations. In this setting, we aim to deliver the same amount of data from each station i to each station $j \neq i$ by scheduling a set of switch matrices. In general, many sets of switch matrices could be used for such scheduling. The problem of finding the set that yields optimal throughput is a very challenging problem combinatorially. Fortunately, for number of stations $N = 4$ or 5 , our simulation results indicate that different sets of switch matrices achieve roughly the same throughput, essentially rendering the selection of the optimal set a non-issue. We conjecture this will be the case for larger N as well. If this conjecture holds, then the complexity of the optimization problem can be decreased significantly as far as engineering design is concerned.

There are many future directions going forward. For example, the beamforming matrices used in our simulation studies could be further optimized. Physical-layer network coding

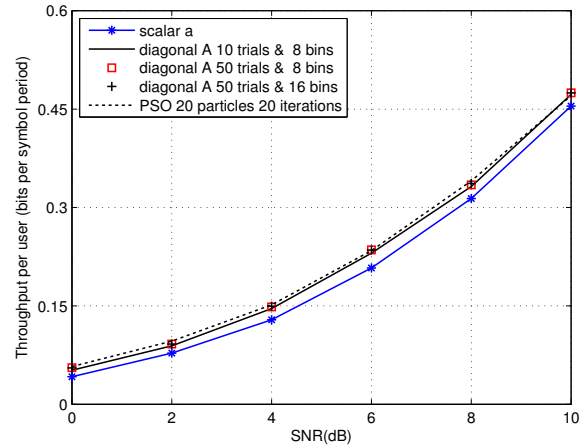


Fig. 5. Throughput per station under MIMO fair switching when $N = 5$. In each case only the result of one condensed derangement set is presented because the results of other derangement sets are within 1% of the results shown here.

could be considered to improve throughput performance [6]. In addition, the setting in which there are unequal amounts of traffic between stations will be interesting to explore. Also, this paper has only considered switch matrices that realize full permutations in which all stations participate in transmission and reception in each slot; it will be interesting to explore switch matrices that realize connectivities among stations that are not full permutations. Finally, future work could also explore the case where the number of antenna at the relay is not exactly N .

REFERENCES

- [1] T. Cover and A. E. Gamal, “Capacity theorems for the relay channel,” *IEEE Trans. Inform. Theory*, vol. 25, no. 5, pp. 572–584, 1979.
- [2] D. Gunduz, A. Yener, A. Goldsmith, and H. V. Poor, “The multi-way relay channel,” in *Proc. IEEE Int’l Symp. Inform. Theory*, 2009, pp. 339–343.
- [3] Y. Mhassab, H. Ghazlan, G. Kramer, and H. El Gamal, “The MIMO wireless switch: Relaying can increase the multiplexing gain,” in *Proc. IEEE Int’l Symp. Inform. Theory*, 2009, pp. 1448–1552.
- [4] T. Cui, T. Ho, and J. Kliewer, “Space-time communication protocols for N-way relay networks,” in *IEEE Global Telecommunications Conference*, 2008, pp. 1–5.
- [5] M. Hassani, “Derangements and applications,” *Journal of Integer Sequences*, vol. 6, no. 1, 2003.
- [6] S. Zhang, S. C. Liew, and P. P. Lam, “Hot topic: physical-layer network coding,” in *Proc. of ACM Mobicom*, 2006.
- [7] K. H. Hui, Y. E. Sagduyu, D. Guo, and R. A. Berry, “The maximum stable broadcast throughput for wireless line networks with network coding and topology control,” in *Proc. Conf. Inform. Sci. Sys.*, Princeton, NJ, USA, Mar. 2010, pp. 1–6.
- [8] T. Yang, X. Yuan, L. Ping, I. B. Collings, and J. Yuan, “A new eigen-direction alignment algorithm for physical layer network coding in MIMO two-way relay channels,” in *Proc. IEEE Int’l Symp. Inform. Theory*, Jul. 2011.
- [9] C. K. Ho, K. T. Gowda, and S. Sun, “Relaying for pair-wise information exchange,” in *Proc. IEEE Inform. Theory Workshop*, Taormina, Sicily, Italy, Oct. 2009, pp. 421–425.
- [10] F. Wang, S. C. Liew, and D. Guo, “Wireless MIMO Switching with Zero-forcing Relaying and Network-coded Relaying,” to appear in *IEEE J. Sel. Area Comm.*, Aug. [Online]. Available: <http://arxiv.org/abs/1108.1522>
- [11] A. U. T. Amah and A. Klein, “Non-regenerative multi-way relaying with linear beamforming,” in *Proc. 20th IEEE Int’l Symp. on Personal, Indoor and Mobile Radio Commun.*, 2009.