ADAPTIVE TRANSMIT ANTENNA SELECTION IN MIMO AMPLIFY-AND-FORWARD RELAY CHANNELS

Kien T. Truong and Robert W. Heath, Jr.

The University of Texas at Austin
Department of Electrical & Computer Engineering
1 University Station C0806
Austin, TX 78712-0240
kientruong@mail.utexas.edu, rheath@ece.utexas.edu

ABSTRACT

Obtaining the most from multiple-antenna relay systems requires adapting the number of substreams based on channel conditions. In this paper, we develop algorithms for selecting dynamically the number of data streams and the transmit antenna subsets at both the source and relay to minimize the vector symbol error rate at a fixed overall rate. An antenna selection mode of operation is defined by the number of transmit antennas and the substream-to-antenna mapping at the source and relay. We propose three suboptimal multimode algorithms that aim to select the best mode based on the knowledge of channel conditions. The algorithms are shown by Monte Carlo simulations to achieve the full diversity order and to provide considerable gains over the existing designs.

Index Terms— MIMO relay, antenna selection, amplify-and-forward, multi-antenna relay

1. INTRODUCTION

Multiple-input multiple-output (MIMO) relay systems combine the benefits of both MIMO communication and relay transmission to provide high data rates and large coverage [1, 2]. Taking full advantage of MIMO relay systems requires adaptive algorithms that configure transmission parameters to the changing channel conditions. This paper presents algorithms for selecting dynamically the number of data streams as well as transmit antenna subsets in half-duplex amplify-and-forward (AF) MIMO relay systems with linear receivers. AF relays are attractive in practice thanks to their low complexity and faster processing since they apply linear filtering to the received signal while decode-and-forward (DF) relays have to decode the signal. Moreover, AF relays can provide better diversity performance and in some cases higher throughput than DF relays in single-antenna half-duplex relay systems [3, 4].

Most prior work on MIMO AF relay systems assumes a fixed number of data streams [4–7], thus working well only in certain channel conditions. This motivates the development of relaying protocols that can adapt to the changes in channel conditions. To the best of our knowledge, [8] is the only work so far on adaptive MIMO AF relaying. Assuming fixed-rate transmission and maximum-likelihood receivers, the authors of [8] propose an adaptive dualmode antenna selection that is based on a sufficient condition for full spatial multiplexing (i.e., all transmit antennas are used) to have a lower vector symbol error rate (VSER) than does full selection diversity (i.e., single antenna transmission with selection diversity where the best antenna is used at both the source and relay).

In this paper, the number of data streams and the antenna subsets at both the source and relay are controlled dynamically according to the current channel conditions based on feedback from the sink. We define an antenna selection mode of operation, or a mode for short, by the number of transmit antennas and the substream-to-antenna mapping at the source and relay. In particular, having the perfect instantaneous information of all constituent channels by assumption [4, 5], the sink selects the mode that minimizes the VSER for a given overall data rate. The selected mode index is sent back to the source and relay via zero-delay, error-free, low-rate feedback channels. The three proposed multimode algorithms are shown by simulation to achieve full diversity performance and to provide considerable array gains over full selection diversity [8] and limited feedback beamforming [6]. Our work is a natural generalization of the results in [8, 9]. Compared to the algorithm in [8], our algorithms allow for more freedom in antenna subset selection at the source and relay assuming linear receivers at the sink.

Note that the objective of adaptation is to minimize error rate at a fixed rate. This is motivated by the need for reliable transmission at a guaranteed data rate. This approach is relevant for providing the constant bit-rate services, like voice and audio, where the error rate performance needs to be improved. This kind of adaptation is also used to provide services to cell-edge users in cellular networks, to whom maintaining a reliable connection at a low fixed-rate rate may be preferred to having an unsteady connection at higher date rates. The problem of error-rate minimization at a fixed rate constraint has also been investigated for both point-to-point MIMO transmission [10, 11] and MIMO relay transmission [8].

Notation: h is a scalar, h is a column vector and H is a matrix. $H^{-1}$, $H^*$, $||H||_F$, $\text{tr}(H)$, and $\text{rank}(H)$ denote the inverse, complex conjugate, Frobenius norm and rank of H. $[H]_{ij}$ is the entry at the $i$th row and $j$th column of H. $I_n$ is the $n \times n$ identity matrix and $0_n$ is the $n \times 1$ zero column vector. $E(\cdot)$ denotes the expectation operator. $CN(a, R)$ is the complex Gaussian distribution with mean $a$ and covariance matrix $R$.

2. SYSTEM MODEL

In this paper we consider a half-duplex nonregenerative multiple-antenna relay system with linear receivers, where the source $S$ sends a message to the sink $D$ with the aid of the relay $R$. Due to the half-duplex nature of the relay, the transmission procedure occurs in two phases. In the first phase, the source broadcasts the message to
both the relay and the sink. In the second phase, the relay forward the message to the sink while the source is assumed to remain silent. Note that in the relaying protocol, the source and relay transmit on orthogonal channels in time as in [4–8].

We assume that the source, sink, and relay are equipped with \( N_S, N_R, \) and \( N_A \) antennas, respectively. In this paper we consider slowly-varying frequency-flat block fading channels. Let \( H_{SD} \in \mathbb{C}^{N_{SD} \times N_{S}} \) denote the \( S \rightarrow D \) channel, \( H_{SR} \in \mathbb{C}^{N_{SR} \times N_{S}} \) the \( S \rightarrow R \) channel, and \( H_{RD} \in \mathbb{C}^{N_{RD} \times N_{R}} \) the \( R \rightarrow D \) channel. Let \( n_{D1}, n_{D2} \) and \( n_R \) be the additive white Gaussian noise at the sink in two phases and at the relay, where \( n_{D1}, n_{D2} \sim \mathcal{CN}(0, N_0) \) and \( n_R \sim \mathcal{CN}(0, \sigma^2 I_{N_R}) \).

We focus on transmit antenna subset selection [12] at both the source and relay. Specifically, we develop adaptive antenna selection algorithms with the aim to reduce the VSER at a given data rate. This kind of adaptation is useful in the cases where the reliability of data transmission is of importance, for example, to provide services to cell-edge users in cellular networks. We assume that the source and relay are equipped with \( M_S \) and \( M_R \) transmit radio frequency (RF) chains, where \( M_S \leq N_S \) and \( M_R \leq N_R \). Based on the antenna selection decision, \( m_S \) out of the \( M_S \) RF chains at the source and \( m_R \) out of the \( M_R \) RF chains at the relay are actually active.

The source consists of a spatial multiplexer that transforms \( R \) bits into \( m_S \) different bit substreams, i.e., the overall data rate is fixed regardless of the number of substreams (\( R \) is assumed to be divisible by \( m_S \)). Let \( M_S \) be the set of supported numbers of substreams. The bit substreams are modulated independently using the same constellation of size \( 2^{R/m_S} \) to form an \( m_S \)-dimensional complex symbol vector \( x_S \in \mathbb{C}^{m_S} \times 1 \), such that \( E_S[x_S^n x_S^m] = (E_S/m_S)I_{m_S} \), where \( E_S \) is the average transmit energy at the source. After going through the transmit RF chains, the \( m_S \) substreams are mapped to a subset of \( m_S \) transmit antennas under a substream-to-antenna mapping matrix. The \( (m_S) \) possible mapping matrices, formed by \( m_S \) columns of \( I_{N_S} \), are indexed and denoted by \( W_{S,(m_S,p)} \), where \( p = 1, 2, \ldots, (m_S) \). In addition, \( H_{SD,(m_S,p)} \triangleq H_{SD}W_{S,(m_S,p)} \).

The substreams are mapped to a subset of \( m_R \) transmit antenna by a spatial mapper. The mapping matrices at the relay, formed by \( m_R \) columns of \( I_{N_R} \), are indexed and denoted by \( W_{R,(m_R,q)} \), where \( q = 1, 2, \ldots, (m_R) \). We also define \( H_{RD,(m_R,q)} \triangleq H_{RD}W_{R,(m_R,q)} \).

In this phase, the sink receives the signal

\[
y_{D2} = \alpha H_{RD} F H_{SR} x_S + \alpha H_{RD} F n_R + n_{D2}. \tag{5}
\]

Using the MIMO notation, the stacked signals available at the sink after two phases are

\[
y = \begin{pmatrix} H_{SD} \alpha \end{pmatrix} x_S + \begin{pmatrix} \alpha H_{RD} F n_R + n_{D2} \end{pmatrix} \tag{6}
\]

A linear equalizer matrix \( G \) is applied to \( y \). The equalized substreams \( G y \) are then detected independently. To support linear detection, \( H \) must have at least as many rows as columns, that is, \( M_S \leq N_D \). For tractability and without any loss in diversity gain, in this paper we focus only on zero-forcing (ZF) receivers. The equalizer is

\[
G^{(ZF)} = \left( R^{-1/2} H \right)^\dagger, \tag{7}
\]

where \( R \) is the covariance matrix of \( n \).

Note that an antenna selection mode is identified by the quadrature \( \omega = (m_S, p, m_R, q) \). We denote the set of all supported modes of operation as \( \Omega \). Let \( \gamma_S = E_S/\sigma^2 \) and \( \gamma_R = E_R/\sigma^2 \). We use the SNR-scaled singular values by denoting \( \lambda_k(H) = \sqrt{\gamma_S}\sigma_k(H_{SD}), \lambda_k(H_{SR}) = \sqrt{\gamma_R}\sigma_k(H_{SR}), \) and \( \lambda_k(H_{RD}) = \sqrt{\gamma_R}\sigma_k(H_{RD}) \), where \( \sigma_k(H) \) is the \( k \)th singular value of \( H \) assuming that \( \sigma_{\text{max}}(H) = \sigma_1(H) \geq \cdots \geq \sigma_k(H) \geq \cdots \geq \sigma_{\text{min}}(H) \).

3. MULTIMODE ANTENNA SELECTION

The objective of multimode antenna selection is to select the mode that minimizes the VSER for a fixed overall data rate. Note that full selection diversity can be supported in all MIMO relay systems. The following theorem guarantees full diversity performance of the optimal multimode algorithm, if any, in Rayleigh fading channels.

**Theorem 1** Selection of the optimal mode of operation that minimizes the vector symbol error rate provides full diversity advantage.

**Proof:** The single best antenna selection strategy, also full selection diversity, is shown to achieve full diversity advantage on the order \( N_S N_R + N_R \min(N_S, N_D) \) [7]. Since full selection diversity is included among the set of possible modes, the optimal selection can only perform better than single-antenna selection diversity for all channel realizations.

Since the closed-form expressions of VSER are not always available, it is difficult to develop the optimal multimode algorithm. In this paper, we develop three suboptimal adaptive antenna selection algorithms by using the nearest neighbor upper bound (NNUB) to derive the closed-form bounds on the VSER.

The post-processing SNRs of the \( k \)th substream for \( k, k = 1, 2, \ldots, m_S \) is

\[
\text{SNR}_k(\omega) = \frac{E_S/m_S}{\|H_{SR}\|_F^2 + m_R}, \tag{7}
\]

where \( H \) is defined in (6). The minimum post-processing SNR over \( m_S \) substreams is defined as \( \text{SNR}_{\min}(\omega) \triangleq \min_{1 \leq k \leq m_S} \text{SNR}_k(\omega) \). Let \( d_{\min}(\omega) = d_{\min}(m_S, R) \) be the minimum distance and \( N_c(\omega) = N_c(m_S, R) \) the average number of nearest neighbors of the transmit
The NNUB bound on the conditional VSER is \[ P_e(\omega) = 1 - \left( 1 - N_e(\omega)Q \left( \frac{\text{SNR}_{\min}(\omega) d_{\min}^2(\omega)}{2} \right) \right)^{m_S}. \]

The first multimode antenna selection is based directly on the closed-form NNUB bounds on the VSER. Specifically, the selection criterion is simply choose the mode with the lowest NNUB bound.

**VSER-Based Selection Criterion:** Choose \( \omega^* \) that solves
\[
\arg \min_{\omega \in \Omega} 1 - \left( 1 - N_e(\omega)Q \left( \frac{\text{SNR}_{\min}(\omega) d_{\min}^2(\omega)}{2} \right) \right)^{m_S}.
\]

The main computational issue with the VSER-based algorithms is the implementation of the \( Q \)-function. The following observations may help reduce the computation load. First, for small \( x \) and positive \( n \), we have \( 1 - (1 - x)^n \approx nx \), therefore, \( P_e(\omega) \) can be approximated by the products of a nearest neighbor term and a \( Q \)-function. Next, if the overall rate is fixed, the number of points in the vector constellation is the same, thus we can neglect the nearest neighbor terms in the relative comparison of VSERs. Finally, since the \( Q \)-function decreases with increasing arguments, we have the following algorithm.

**SNR-Based Selection Criterion:** Choose \( \omega^* \) that solves
\[
\arg \max_{\omega \in \Omega} d_{\min}^2(\omega) \left( \min_{1 \leq k \leq m_S} SNR_k(\omega) \right).
\]

Note that the algorithm still requires the computation of the post-processing SNR for each stream in any possible mode of operation. Moreover, along with the VSER-based algorithm, the SNR-based algorithm does not reveal much about the spatial characteristics of the constituent channel affect the mode selection. Following the idea of the condition-number-based selection criterion for the point-to-point MIMO channel in [9], we will develop a selection criterion that is based on a lower bound on the minimum post-processing stream SNR, which is a function of the singular values of the constituent channels. Such a lower bound is given in the following theorem.

**Theorem 2** When the direct link is considered, the minimum substream post-processing SNR satisfies
\[
\text{SNR}_{\min}(\omega) \geq \frac{1}{m_S} \lambda_{m_S}(\mathcal{H}(\omega)),
\]
where \( \lambda_{m_S}(\mathcal{H}(\omega)) \) is the SNR-scaled “minimum singular value” of the effective channel \( \mathcal{H} \) and is defined as
\[
\lambda_{m_S}(\mathcal{H}(\omega)) \triangleq \lambda_{m_S}(\mathcal{H}_{SD}) + \lambda_{m_S}(\mathcal{H}_{SR}) + \frac{1}{m_S} \sum_{i=1}^{m_S} \lambda_i(\mathcal{H}_{SR}) + \lambda_{m_S}(\mathcal{H}_{RD}) + m_R, \]
where \( \lambda_{m_S}(\mathcal{H}(\omega)) = 0 \) if \( \text{rank}(\mathcal{H}(\omega)) < m_S \).

**Proof:** See the proof of Theorem 2 in [14].

Note that the first term in the expression of \( \lambda(\mathcal{H}) \) is in fact the SNR-scaled minimum singular value of the direct channel \( \mathcal{H}_{SD} \). The second term can be treated as the SNR-scaled minimum singular value of the two-hop channel, \( S \rightarrow R \rightarrow D \). In addition, if either \( \text{rank}(\mathcal{H}_{SR}) < m_S \) or \( m_R < m_S \), then the second term is zero, meaning that the two-hop channel cannot support the transmission of \( m_S \) substreams and the source cannot use the help of the relay. This means that we should consider only the mode corresponding to \( m_R \geq m_S \). We notice that, however, if \( m_R \) increases, then \( \lambda_{m_S}(\mathcal{H}) \) decreases. Therefore, we should choose \( m_R = m_S \). This result is intuitive since with uniform power allocation at the relay, choosing \( m_R < m_S \) makes at least one substream lost in the null space, while choosing \( m_R > m_S \) decreases the transmit power per substream.

**Theorem 2** leads to a straightforward simplification of the SNR-based Selection Criterion, where \( \min_{1 \leq k \leq m_S} SNR_k(\omega) \) is replaced by its lower bound. Nevertheless, this algorithms still has high computational load since it requires computing all the singular values of \( \mathcal{H} \). To further simplify the multimode selection, we relate the smallest scaled singular value of \( \mathcal{H} \) to the eigenmodes of \( \mathcal{H}_{SD}, \mathcal{H}_{SR}, \) and \( \mathcal{H}_{RD} \). From Theorem 3 in [9], we have \( \lambda_2(\mathcal{H}_{SD}) \geq \lambda_2(\mathcal{H}_{SR}), \lambda_2(\mathcal{H}_{SR}) \geq \lambda_2(\mathcal{H}_{RD}) \) and \( \lambda_2(\mathcal{H}_{RD}) \geq \lambda_2(\mathcal{H}_{RD}) \), for all \( 1 \leq i \leq m_S \). Using the relationship to approximate \( \lambda_{m_S}(\mathcal{H}) \), leading to an eigenmode-based selection criterion where the computation of the optimal number of data substreams can thus be isolated from the computation of other parameters. The eigenmode-based multimode antenna selection is stated as follows.

**Eigenmode-Based Selection Criterion:** The selection procedure consists of three steps:

1. **Step 1:** Solve for
   \[
   m_S^* = \arg \max_{m_S \in \mathbb{M}} \left( \frac{\lambda_{m_S}(\mathcal{H}_{SD})}{m_S} + \frac{\lambda_{m_S}(\mathcal{H}_{SR})}{m_S} + \frac{\lambda_{m_S}(\mathcal{H}_{RD})}{m_S} + m_R \right).
   \]
   Note that \( m_R = m_S \).

2. **Step 2:** Find \( q^* \) that solve
   \[
   \arg \max_{1 \leq q \leq (m_S^*)} \lambda_{m_S^*}(\mathcal{H}(m_S^*, q^*)).
   \]

3. **Step 3:** Choose \( p^* \) that solves
   \[
   \arg \max_{1 \leq p \leq (m_S^*)} \lambda_{m_S^*}(\mathcal{H}(m_S^*, p, m_R, q^*)).
   \]

The eigenmode-based algorithm has much lower complexity than other multimode algorithms at the price of accuracy as we use several approximations in the derivation. Notably, this algorithm reveals how the quality of original constituent channels (i.e., \( \mathcal{H}_{SD}, \mathcal{H}_{SR}, \) and \( \mathcal{H}_{RD} \)) affects mode selection. Specifically, the selection of the number of streams \( m_S \) based on the singular values of the original channels provides intuition about multimode selection. If we define \( r = \max(\text{rank}(\mathcal{H}_{SD}), \text{min}(\text{rank}(\mathcal{H}_{SR}), \text{rank}(\mathcal{H}_{RD}))) \), then there is no reason to consider \( m_S > r \) since at least one data stream is lost in the null space with linear detection. Note that for a given rate \( R \), there is a tradeoff between the minimum distance of the constellation (i.e., \( d_{\min}(\omega) \)) and the number of streams \( m_S \). A larger value of \( m_S \) corresponds to a smaller constellation size, and hence a larger value of \( d_{\min}(\omega) \). Depending on \( R \), the tradeoff between \( d_{\min}(\omega) \) and \( m_S \), and the distribution of the singular values of the original constituent channels, any \( m_S < r \) may be optimal.

4. **SIMULATIONS**

This section provides Monte Carlo simulations to evaluate the VSER performance of the proposed antenna selection criteria when using the i.i.d. Rayleigh fading model. We assume block fading with the block size of 100 vector symbols. The data rate is fixed at \( R = 8 \) bits and symbols are drawn from QAM constellations with unit power. We assume \( N_S = N_R = N_D = 4 \) and the set of possible numbers of
source substreams is $M_s \in \{1, 2, 4\}$ with 256-QAM, 16-QAM, and 4-QAM. Let $\text{SNR}_{SR}$, $\text{SNR}_{RD}$ and $\text{SNR}_{SD}$ denote the mean SNR values of the corresponding channels. We fix $\text{SNR}_{SR} = 20 \text{ dB}$ and $\text{SNR}_{RD} = 10 \text{ dB}$ and produce the performance curves of VSER against $\text{SNR}_{SD}$ as shown in Fig. 1.

First, notice that the proposed multimode algorithms have the same diversity order as full selection diversity, which is shown to achieve full diversity order performance [7]. Therefore, although suboptimal in term of VSER minimization, our multimode algorithms can still obtain the full diversity advantage, thus providing significant diversity gain over full spatial multiplexing at high SNR. Notably, the multimode algorithms provide large array gains, e.g., around 6 dB at $10^{-2}$ for the eigenmode-based multimode algorithm, over single-stream transmission strategies, namely full antenna selection diversity and limited feedback Grassmannian beamforming. This can be explained by the fact that by supporting a varying number of data streams, multimode transmission provides better coupling of signs and channels than single-stream transmission. Finally, the SNR-based algorithm and the VSER-based algorithm give nearly the same VSER performance, which is better than the performance of the eigenmode-based algorithm. This means that when the complexity is not an issue, we should use the SNR-based algorithm since it has lower complexity than the VSER-based algorithm. On the other hand, when the complexity is critical, the eigenmode-based algorithm should be used since it provides a good tradeoff between complexity and VSER performance.

5. CONCLUSIONS

We proposed various multimode antenna selection criteria for half-duplex MIMO AF relay systems. Although suboptimal in term of VSER minimization, the proposed algorithms can achieve the full diversity gain of the corresponding channel, thus significantly improve the diversity performance of the plain spatial multiplexing. They also provide considerable array gains over the single antenna transmission, such as the full selection diversity (i.e., single antenna transmission where the best antenna is used at both the source and relay) and limited feedback Grassmannian beamforming. Our multimode algorithms also allow for more freedom in operation mode selection than the existing dualmode algorithm.

6. REFERENCES