On the Achievable Diversity-Multiplexing Tradeoff in Half-Duplex Cooperative Channels

Kambiz Azarian, Hesham El Gamal, and Philip Schniter
Dept. of Electrical and Computer Engineering
The Ohio State University
{azariany,helgamal,schniter}@ece.osu.edu

July 16, 2004

Abstract

In this paper, we propose novel cooperative transmission protocols for delay limited coherent fading channels consisting of N (half-duplex and single-antenna) partners and one cell site. In our work, we differentiate between the relay, cooperative broadcast (down-link), and cooperative multiple-access (up-link) channels. The proposed protocols are evaluated using Zheng-Tse diversity-multiplexing tradeoff. For the relay channel, we investigate two classes of cooperation schemes; namely, Amplify and Forward (AF) protocols and Decode and Forward (DF) protocols. For the first class, we establish an upper bound on the achievable diversity-multiplexing tradeoff with a single relay. We then construct a new AF protocol that achieves this upper bound. The proposed algorithm is then extended to the general case with $(N-1)$ relays where it is shown to outperform the space-time coded protocol of Laneman and Worenell without requiring decoding/encoding at the relays. For the class of DF protocols, we develop a dynamic decode and forward (DDF) protocol that achieves the optimal tradeoff for multiplexing gains $0 \leq r \leq 1/N$. Furthermore, with a single relay, the DDF protocol is shown to dominate the class of AF protocols for all multiplexing gains. The superiority of the DDF protocol is shown to be more significant in the cooperative broadcast channel. The situation is reversed in the cooperative multiple-access channel where we propose a new AF protocol that achieves the optimal tradeoff for all multiplexing gains. A distinguishing feature of the proposed protocols in the three scenarios is that they do not rely on orthogonal subspaces, allowing for a more efficient use of resources. In fact, using our results one can argue that the sub-optimality of previously proposed protocols stems from their use of orthogonal subspaces rather than the half-duplex constraint.

1 Introduction

Recently, there has been a growing interest in the design and analysis of cooperative transmission protocols for wireless fading channels (e.g., [1, 2, 3, 4, 5, 6]). In delay limited (e.g., quasi-static) coherent channels, the basic idea is to leverage the antennas available at the other nodes in the network as a source of *virtual* spatial diversity. Here, we use the same setup as considered by Laneman, Tse, and Wornell in [3]. There, the authors imposed the half-duplex constraint (either transmit or receive, but not both) on the cooperating nodes and proposed several cooperative transmission protocols. The proposed protocols in [3] were classified as either Amplify and Forward (AF), where the helping node retransmits a scaled version of its soft observation, or Decode and Forward (DF), where the helping node attempts first to decode the information stream and then re-encodes it using (a possibly different) code-book. All the proposed schemes in [3] used a Time Division Multiple Access (TDMA) strategy, where the two partners relied on the use of orthogonal subspaces to repeat each other's signals. Later, Laneman and Wornell extended their DF strategy to the N partners scenario [4]. Other follow-up works have focused on developing practical coding schemes that attempt to exploit the promised information theoretic gains (e.g., [6, 7]).

As observed in [3, 4], previously proposed cooperation protocols suffer from a significant loss of performance in high spectral efficiency scenarios. In fact, the authors of [3] posed the following open problem: *"a key area of further research is exploring cooperative diversity protocols in the high spectral efficiency regime."* This remark motivates our work here, where we present more efficient (and in some cases optimal) AF and DF protocols for the relay, cooperative broadcast (CB), and cooperative multiple-access (CMA) channels. To establish the gain offered by the proposed protocols, we adopt the diversity-multiplexing tradeoff as our measure of performance. This powerful tool was introduced by Zheng and Tse for point-to-point multi-input-multi-output (MIMO) channels in [8] and later used by Tse, Viswanath, and Zheng to study the (non-cooperative) multiple-access channel in [9].

In the following, we summarize the main results of this paper, some of which were initially reported in [10, 11, 12].

1. For the single relay channel, we establish an upper bound on the achievable diversity-multiplexing tradeoff by the class of AF protocols. We then identify a variant within this class, referred to as the Nonorthogonal Amplify and Forward (NAF) protocol, that achieves this upper bound. We then propose a dynamic decode and forward (DDF) protocol and show that it achieves the **optimal** tradeoff for multiplexing gains $0 \leq r \leq 0.5$[1]. Furthermore, the DDF protocol is shown to outperform all AF protocols for arbitrary multiplexing gains. Finally, the two protocols (i.e., NAF and DDF) are extended to the scenario with $N-1$ relays where we characterize their tradeoff curves. Notably, the NAF protocol is shown to outperform the space-time coded protocol of Laneman and Wornell [4] without requiring decoding/encoding at the relays.

2. For the cooperative broadcast channel, we present a modified version of the DDF protocol to allow for reliable transmission of the common information. We then characterize the tradeoff curve of this protocol and use this characterization to establish its superiority compared to AF protocols. In fact, we argue that the gain offered by the DDF is more significant in this scenario (as compared to the relay channel).

3. For the symmetric multiple-access scenario, we propose a novel AF cooperative

[1]The multiplexing gain "r" will be defined rigorously in the sequel

protocol where an *artificial* inter-symbol-interference (ISI) channel is created. We prove the optimality (in the sense of the diversity-multiplexing tradeoff) of this protocol by showing that, for all multiplexing gains (i.e., $0 \le r \le 1$), it achieves the diversity-multiplexing tradeoff of the corresponding $N \times 1$ point-to-point channel. One can then use this result to argue that the sub-optimality of previously proposed schemes was dictated by the use of orthogonal subspaces rather than the half-duplex constraint. We also utilize this result to shed more light on the fundamental difference between half-duplex relay and cooperative multiple-access channels.

Before proceeding further, a brief remark about notation is in order. We use $(x)^+$ to mean $\max\{x, 0\}$, $(x)^-$ to mean $\min\{x, 0\}$ and $\lceil x \rceil$ to mean nearest integer to x towards plus infinity. $\mathbb{R}^N$ and $\mathbb{C}^N$ denote the set of real and complex N-tuples, respectively, while $\mathbb{R}^{N+}$ denotes the set of non-negative N-tuples. I_N denotes the $N \times N$ identity matrix, $\Sigma_{\mathbf{x}}$ denotes the autocovariance matrix of vector $\mathbf{x}$, and $\log(.)$ denotes the base-2 logarithm.

The rest of the paper is organized as follows. In Section 2, we detail our modeling assumptions and review, briefly, some results that will be extensively used in the sequel. The half-duplex relay channel is investigated in Section 3 where we describe the NAF and DDF protocols and derive their tradeoff curves. In Section 4, we extend the DDF protocol to the cooperative broadcast channel. Section 5 is devoted to the cooperative multiple-access channel where we propose a new AF protocol and establish its optimality, in the symmetric scenario, with respect to the diversity-multiplexing tradeoff. In Section 6, we present numerical results that show the SNR gains offered by the proposed schemes in certain representative scenarios. Finally, we offer some concluding remarks in Section 7. To enhance the flow of the paper, we collect all the proofs in the Appendix.

2 Background

First we state the general assumptions that apply to the three scenarios considered in this paper (i.e., relay, broadcast, and multiple-access). Assumptions pertaining to a specific scenario will be given in the related section.

1. All channels are assumed to be flat Rayleigh-fading and quasi-static: the channel gains remain constant during a coherence-interval and change independently from one coherence-interval to another. Furthermore, the channel gains are mutually independent with unit variance. The additive noises at different nodes are zero-mean, mutually-independent, circularly-symmetric and white complex-Gaussian. Furthermore, the variances of these noises are proportional to one another such that there will always be *fixed* offsets between the different channels signal to noise ratios (SNRs).

2. All nodes have the same power constraint, have a single antenna, and operate synchronously. Only the receiving node of any link knows the channel gain; no feedback to the transmitting node is permitted (the incremental relaying protocol proposed in [3] can not, therefore, be considered in our framework). Following in the footsteps of [3], all cooperating partners operate in the half-duplex mode, i.e., at any point in time, a node can either transmit or receive, but not both. This constraint is motivated by, e.g., the typically large difference between the incoming and outgoing signal power levels. Though this half-duplex constraint is

quite restrictive to protocol development, it is nevertheless assumed throughout the paper.

3. Throughout the paper, we assume the use of random Gaussian code-books where a codeword spans the entire coherence-interval of the channel. Furthermore, we assume an asymptotically large codeword size[2]. Results related to the design of practical coding/decoding schemes that approach the fundamental limits established here will be reported elsewhere.

Next we summarize several important definitions and results that will be used throughout the paper.

1. The SNR of a link, ρ, is defined as

$$\rho \triangleq \frac{E}{\sigma_v^2}, \tag{1}$$

where E denotes the average energy available for transmission of a symbol across the link and σ_v^2 denotes the variance of the noise observed at the receiving end of the link. We say that $f(\rho)$ is *exponentially equal to* ρ^b, denoted by $f(\rho) \doteq \rho^b$, when

$$\lim_{\rho\to\infty} \frac{\log(f(\rho))}{\log(\rho)} = b. \tag{2}$$

In (2), b is called the *exponential order* of $f(\rho)$. $\dot{\leq}$ and $\dot{\geq}$ are defined similarly.

2. Consider a family of codes $\{C_\rho\}$ indexed by operating SNR ρ, such that the code C_ρ has a rate of $R(\rho)$ bits per channel use (BPCU) and a maximum likelihood (ML) error probability $P_e(\rho)$. For this family, the *multiplexing gain* "r" and the *diversity gain* "d" are defined as

$$r \triangleq \lim_{\rho\to\infty} \frac{R(\rho)}{\log \rho}, \qquad d \triangleq -\lim_{\rho\to\infty} \frac{\log(P_e(\rho))}{\log \rho}. \tag{3}$$

3. The problem of characterizing the optimal tradeoff between the reliability and throughput of a point-to-point communication system over a coherent quasi-static flat Rayleigh-fading channel was posed and solved by Zheng and Tse in [8]. For a MIMO communication system with M transmit and N receive antennas, they showed that, for any $r \leq \min\{M, N\}$, the optimal diversity gain $d^*(r)$ is given by the piecewise linear function joining the (r, d) pairs $(k, (M-k)(N-k))$ for $k = 0, ..., \min\{M, N\}$, provided that the code-length l satisfies $l \geq M + N - 1$.

4. We say that protocol A *uniformly dominates* protocol B if, for any multiplexing gain r, $d_A(r) \geq d_B(r)$.

5. Assume that g is a Gaussian random variable with zero mean and unit variance. If v denotes the exponential order of $1/|g|^2$, i.e.,

$$v = -\lim_{\rho\to\infty} \frac{\log(|g|^2)}{\log(\rho)}, \tag{4}$$

[2] In practice, codeword sizes in the order of one hundred bits will only entail a very small loss in throughput compared to the idealistic case.

then the probability density function (PDF) of v can be shown to be:

$$p_v = \lim_{\rho\to\infty} \ln(\rho)\rho^{-v}\exp(-\rho^{-v}).$$

Careful examination of the previous expression reveals that

$$p_v \dot{=} \begin{cases} \rho^{-\infty} = 0, & \text{for } v < 0, \\ \rho^{-v}, & \text{for } v \geq 0 \end{cases}. \tag{5}$$

Thus, for independent random variables $\{v_j\}_{j=1}^N$ distributed identically to v, the probability P_O that $(v_1, \ldots, v_N)$ belongs to set O can be characterized by

$$P_O \dot{=} \rho^{-d} \quad \text{for} \quad d = \inf_{(v_1,\ldots,v_N)\in O^+} \sum_{j=1}^{N} v_j, \tag{6}$$

where $O^+ \triangleq O \cap \mathbb{R}^{N+}$. It should be noted that P_O only depends on O^+; by increasing SNR, the probability of $(v_1, \ldots, v_N)$ with a negative element can be made arbitrarily small.

6. Consider a coherent linear Gaussian channel where a random Gaussian code-book is used. The pairwise error probability (PEP) of the ML decoder, denoted as P_{pe}, averaged over the ensemble of random Gaussian codes, is upper bounded by

$$P_{pe} \leq \det(I_N + \frac{1}{2}\Sigma_{\mathbf{s}}\Sigma_{\mathbf{n}}^{-1})^{-1}, \tag{7}$$

where $\mathbf{s} \in \mathbb{C}^N$ and $\mathbf{n} \in \mathbb{C}^N$ denote the signal and noise components of the observed vector, respectively (i.e., $\mathbf{y} = \mathbf{s} + \mathbf{n}$).

3 The Half-Duplex Relay Channel

In this section, we consider the relay scenario in which $N-1$ relays help a single source to better transmit its message to the destination. As the vague descriptions "help" and "better transmit" suggest, the general relay problem is rather broad and only certain sub-problems have been studied (for example see [13]). In this work, we focus on two important classes of relay protocols. The first is the class of Amplify and Forward (AF) protocols, where a relaying node can only process the observed signal linearly before re-transmitting it. The second is the class of Decode and Forward (DF) protocols, where the relays are allowed to decode and re-encode the message using (a possibly different) code-book. Here we emphasize that, a priori, it is not clear whether the DF protocols offer better performance than AF protocols or not (e.g., [3]).

3.1 Amplify and Forward Protocols

We first consider the single relay scenario (i.e., $N = 2$). For this scenario, we derive the optimal diversity-multiplexing tradeoff and identify a specific protocol within this class, i.e., the NAF protocol, that achieves this optimal tradeoff. We then extend the NAF protocol to the general case with an arbitrary number of relays.

Under the half-duplex constraint, it is easy to see that any single-relay AF protocol can be mathematically described by some choice of the matrices A_1, A_2, and B in the following model

$$\mathbf{y} = \begin{bmatrix} g_1 A_1 & 0 \\ g_2 h B & g_1 A_2 \end{bmatrix} \mathbf{x} + \begin{bmatrix} 0 \\ g_2 B \end{bmatrix} \mathbf{w} + \mathbf{v}. \tag{8}$$

In (8), $\mathbf{y} \in \mathbb{C}^l$ represents the vector of observations at the destination, $\mathbf{x} \in \mathbb{C}^l$ the vector of symbols transmitted by the source, $\mathbf{w} \in \mathbb{C}^{l'}$ the vector of noise samples (of variance σ_w^2) observed by the relay, and $\mathbf{v} \in \mathbb{C}^l$ the vector of noise samples (of variance σ_v^2) observed by the destination. The variables h, g_1 and g_2 denote the source-relay channel gain, source-destination channel gain, and relay-destination channel gain, respectively. $A_1 \in \mathbb{C}^{l' \times l'}$ and $A_2 \in \mathbb{C}^{(l-l') \times (l-l')}$ are diagonal matrices. In this protocol, the source can potentially transmit a new symbol in every symbol-interval of the codeword, while the relay listens during the first l' symbols and then, for the remaining $l - l'$ symbols, transmits linear combinations of the l' noisy observations using the coefficients in $B \in \mathbb{C}^{(l-l') \times l'}$. In fact, by letting $l' = l/2$, $A_1 = I_{l'}$, $A_2 = 0$ and $B = I_{l'}$ we obtain Laneman-Tse-Wornell Amplify and Forward (LTW-AF) protocol [3]. Finally, we note that when the source symbols are independent, the average energy constraint translates to

$$\sum_{i=1}^{l'} |b_{ji}|^2 \leq \frac{E}{|h|^2 E + \sigma_w^2}, \quad j = 1, \ldots, l - l', \tag{9}$$

where $B = [b_{ji}]$.

Theorem 1 *The optimal diversity gain for the cooperative relay scenario with a single AF relay is upper-bounded by*

$$d^*(r) \leq (1 - r) + (1 - 2r)^+. \tag{10}$$

Proof: Please refer to the Appendix.

The upper-bound on $d^*(r)$, as given by (10), is shown in Fig. 1. Having Theorem 1 at hand, it now suffices to identify an AF protocol that achieves this upper-bound in order to establish its optimality. Towards this end, we observe that, in the proof of Theorem 1, the only requirements on B such that the protocol described by (8) could *potentially* achieve the optimal diversity-multiplexing tradeoff are for B to be square (of dimension $l/2 \times l/2$) and full-rank. Furthermore, B should not violate the relay average energy constraint as given by (9). Thus, the simple choices

$$A_1 = I_{l/2} \qquad A_2 = I_{l/2} \qquad B = bI_{l/2} \text{ for } b \leq \sqrt{\frac{E}{|h|^2 E + \sigma_w^2}} \tag{11}$$

inspire our NAF protocol. In particular, the source transmits on every symbol-interval in a cooperation frame, where a cooperation frame is defined as two consecutive symbol-intervals. The relay, on the other hand, transmits only once per cooperation frame; it simply repeats the (noisy) signal it observed during the previous symbol-interval. It is important to realize that this design is dictated by the half-duplex constraint, which implies that the relay can repeat at most once per cooperation frame. We denote the repetition gain by b and, for frame k, we denote the information symbols by $\{x_{j,k}\}_{j=1}^{2}$. The signals received by the destination during frame k are thus:

$$y_{1,k} = g_1 x_{1,k} + v_{1,k}$$
$$y_{2,k} = g_1 x_{2,k} + g_2 b(h x_{1,k} + w_{1,k}) + v_{2,k}$$

where the repetition gain b must satisfy (11). Now, we are ready to establish the optimality of the NAF protocol with respect to the diversity-multiplexing tradeoff.

Theorem 2 *The NAF protocol achieves the optimal diversity-multiplexing tradeoff for the AF single-relay scenario, which is:*

$$d^*(r) = (1-r) + (1-2r)^+. \tag{12}$$

Proof: Please refer to the Appendix.

Three remarks are now in order:

1. As shown in Fig. 1, the NAF protocol enjoys uniform dominance over the direct transmission scheme (i.e., no cooperation) and LTW-AF protocol. This dominance can be attributed to relaxing the orthogonality constraint whereby one can reap two distinct benefits: rate enhancement via continuous transmission and diversity enhancement via cooperation. It is interesting to note that this dominance is achieved while only half of the symbols are repeated by the relay.

2. From Fig. 1, one can see that for multiplexing gains greater that 0.5, the diversity gain achieved by the proposed NAF relay protocol is identical to that of the non-cooperative protocol. This is due to the fact that the *AF cooperative* link provided by the relay can not support multiplexing gains greater than 0.5—a consequence of the half-duplex constraint. Hence, for multiplexing gains larger than 0.5, there is only one link from the source to the destination, and, thus, the tradeoff curve is identical to that of a point-to-point system with one transmit and one receive antenna. Later, we will show that the proposed DDF strategy avoids this drawback.

3. As shown in the proof of Theorem 2, the achievability of the optimal tradeoff is not very sensitive to the choice of the repetition gain "b" (i.e., for a wide range of choices, the NAF protocol achieves the optimal tradeoff). In practice, one should optimize the repetition gain, experimentally if needed, to minimize the outage probability at the target rate and signal-to-noise ratio.

The NAF protocol can be extended to the case of arbitrary number of relays (i.e., $N \geq 2$) as follows. First, we define a super-frame as a concatenation of $N-1$ consecutive cooperation frames. Within each super-frame, the relays take turns repeating the signals they previously observed as they did in the case of a single relay. Thus, the destination's received signals during a super-frame will be

$$\begin{aligned}
y_{1,1} &= g_1 x_{1,1} + v_{1,1} \\
y_{2,1} &= g_1 x_{2,1} + g_2 b_2 (h_2 x_{1,1} + w_{1,1}) + v_{2,1} \\
y_{1,2} &= g_1 x_{1,2} + v_{1,2} \\
y_{2,2} &= g_1 x_{2,2} + g_3 b_3 (h_3 x_{1,2} + w_{1,2}) + v_{2,2} \\
&\vdots \\
y_{1,N-1} &= g_1 x_{1,N-1} + v_{1,N-1} \\
y_{2,N-1} &= g_1 x_{2,N-1} + g_N b_N (h_N x_{1,N-1} + w_{1,N-1}) + v_{2,N-1},
\end{aligned}$$

where the source-relay channel gain, relay-destination channel gain, relay repetition gain, and relay-observed noise for relay $i \in \{1, ..., N-1\}$ are denoted by h_{i+1}, g_{i+1}, b_{i+1} and $w_{1,i}$, respectively. As before, g_1 represents the source-destination channel gain. The quantities $y_{j,k}$, $v_{j,k}$, and $x_{j,k}$ represent the received signal, noise sample, and source symbol, respectively, during the j^{th} symbol-interval of the k^{th} cooperation frame. Note that there is nothing to be gained by having more than one relay transmitting the same symbol simultaneously. The following Theorem characterizes the diversity-multiplexing tradeoff achieved by this protocol.

Theorem 3 *The diversity-multiplexing tradeoff achieved by the NAF protocol with $N-1$ relays is characterized by*

$$d(r) = (1-r) + (N-1)(1-2r)^+.$$

Proof: The proof is virtually identical to that of Theorem 2, and hence, is omitted for brevity.

It is interesting to note that the generalized NAF protocol uniformly dominates the space-time coded protocol proposed by Laneman and Wornell [4]. The generalized NAF protocol offers the additional advantage of low complexity since it does not require decoding/encoding at the relays.

3.2 Decode and Forward Protocols

In this class of protocols, we allow for the possibility of decoding/encoding at the different relays. In [3], Laneman-Tse-Wornell presented a particular variant of DF protocols where the source transmits in the first half of the codeword. Based on its received signal in this interval, the relay attempts to decode. If successful, the relay re-encodes and transmits the encoded stream in the second half of the codeword. One can easily see that the diversity-multiplexing tradeoff of this scheme is upper bounded by that of the LTW-AF protocol in Fig. 1 (i.e., $d(r) \leq 2(1-2r)$) [3]. Here, we propose a Dynamic Decode and Forward (DDF) protocol and characterize its tradeoff curve. This characterization will reveal the uniform dominance of this protocol over all known cooperation protocols for the half-duplex relay channel and, furthermore, establish its optimality over a certain range of multiplexing gains. For clarity, we first describe and analyze the protocol for the case of a single relay. Generalization to $N-1$ relays will then follow.

Similar to the previous section, we assume that a codeword consists of l consecutive symbol-intervals, during which all the channel gains remain unchanged. In the DDF protocol, the source transmits during every symbol-interval in the codeword, while the relay listens to the source until it collects sufficient energy to decode the message error-free. It then encodes the message using an *independent* Gaussian code-book and transmits it during the rest of the codeword. The dynamic nature of the protocol is manifested in the fact that we allow the relay to listen for a time duration that depends on the instantaneous channel realization to maximize the probability of successful decoding. We denote the signals transmitted by the source and relay as $\{x_k\}_{k=1}^{l}$ and $\{\tilde{x}_k\}_{k=l'+1}^{l}$, respectively, where l' is the number of symbol-intervals the relay waits before starting transmission. Using this notation, the received signals (at the destination) can be written as:

$$
\begin{aligned}
y_1 &= g_1 x_1 + v_1 \\
y_2 &= g_1 x_2 + v_2 \\
&\vdots \\
y_{l'} &= g_1 x_{l'} + v_{l'} \\
y_{l'+1} &= g_1 x_{l'+1} + g_2 \tilde{x}_{l'+1} + v_{l'+1} \\
y_{l'+2} &= g_1 x_{l'+2} + g_2 \tilde{x}_{l'+2} + v_{l'+2} \\
&\vdots \\
y_l &= g_1 x_l + g_2 \tilde{x}_l + v_l.
\end{aligned}
$$

The number of symbols where the relay listens should be chosen as:

$$l' = \min \left\{ l, \left\lceil \frac{lR}{\log_2 (1 + |h|^2 c\rho)} \right\rceil \right\}, \tag{13}$$

where R is the source data rate in BPCU, h is the source-relay channel gain, and $c = \sigma_v^2 / \sigma_w^2$. One can now see the dependence of this choice of l' on the instantaneous channel realization and that this choice, together with the asymptotically large l, means that the relay average probability of error with a Gaussian code ensemble is arbitrarily small when $l' < l$. Clearly, when $l' = l$ the relay does not contribute to the transmission of the message, and hence, incorrect decoding at the relay in this case does not affect performance. The following Theorem describes the diversity-multiplexing tradeoff achievable with this cooperation protocol.

Theorem 4 *The diversity-multiplexing tradeoff achieved by the single relay DDF protocol is given by*

$$d(r) = \begin{cases} 2(1-r) & \text{if } \frac{1}{2} \geq r \geq 0 \\ (1-r)/r & \text{if } 1 \geq r \geq \frac{1}{2} \end{cases}. \tag{14}$$

Proof: Please refer to the Appendix.

The diversity-multiplexing tradeoff of (14) is shown in Fig. 2. It is now clear that the DDF protocol is optimal for $0 \leq r \leq 0.5$ since it achieves the genie aided diversity (where the relay is assumed to know the information message *a-priori*). For $r > 0.5$, the DDF protocol suffers from a loss, compared to the genie aided strategy, since, on the average, the relay will only be able to help during a small fraction of the codeword. It is easy to see that, the performance for this range of multiplexing gains can not be improved through employing a mixed AF and DF strategy. In fact, the DDF strategy dominates all such strategies[3]. It remains to be seen whether there exists a strategy that closes the gap to the genie aided strategy when $r > 0.5$ or the DDF protocol is indeed optimal. Finally, we note that the gain offered by the DDF protocol, compared to AF protocols, can be attributed to the ability of this strategy to transmit independent Gaussian symbols after successful decoding. In AF strategies, on the other hand, the relay is limited to repeating the noisy Gaussian symbols it receives from the source.

Next, we describe the generalization of the DDF protocol to the case of multiple relays. In this case, the source and relays cooperate in nearly the same manner as in the

[3]The proof for this is rather straightforward, and hence, is omitted here for brevity.

single relay case. Specifically, the source transmits during the whole codeword while each relay listens until it collects sufficient energy for error-free decoding. Once a relay decodes the message, it uses an independent code-book to re-encode the message, which it then transmits for the rest of the codeword. Note that, since the source-relay channel gains may differ, the relays may require different wait times for decoding. This complicates the protocol, since a given relay ability to decode the message requires precise knowledge of the times at which every other relay begins its transmission. To address this problem, the codeword is divided into a number of segments, and relays are allowed to start transmitting only at the beginning of a segment. In between the segments, every relay is allowed to broadcast a (well protected) beacon, informing all other relays whether or not it will start transmission. Judicious choice of the segment length, relative to the codeword length, results in only a small loss compared to the genie-aided case, whereby all relays know all decoding times *a-priori*. Here, we assume that the number of segments is sufficiently large and the length of the beacon signals is much smaller than the segment length. Therefore, in characterizing the diversity-multiplexing tradeoff achieved by this protocol, we ignore the losses associated with the beacon and the quantization of the starting times for the different delays.

Theorem 5 *The diversity-multiplexing tradeoff achieved by the DDF protocol with $N-1$ relays is characterized by:*

$$d(r) = \begin{cases} N(1-r), & \frac{1}{N} \geq r \geq 0, \\ 1 + \frac{(N-1)(1-2r)}{1-r}, & \frac{1}{2} \geq r \geq \frac{1}{N}, \\ \frac{1-r}{r}, & 1 \geq r \geq \frac{1}{2}. \end{cases} \qquad (15)$$

Proof: Please refer to the Appendix.

The diversity-multiplexing tradeoff (15) is shown in Fig. 3 and Fig. 4 for different values of N. While the loss of the DDF protocol compared to the genie-aided protocol increases with N, it is not clear at the moment if this loss is due to the half-duplex constraint or due to the sub-optimality of the DDF strategy.

4 The Half-Duplex Cooperative Broadcast Channel

We now consider the cooperative broadcast (CB) scenario, where a single source broadcasts to N destinations. We assume that $R_c = r_c \log(\rho)$ BPCU is used to encode a common portion of the message (intended for all destinations) while $R_j = r_j \log(\rho)$ BPCU is used to encode the portion of the message intended for the j^{th} destination, where $j \in \{1, \ldots, N\}$. The total rate is then $R = R_c + \sum_{j=1}^{N} R_j$ and the multiplexing gain tuple is given by $\mathbf{r} = (r_c, r_1, ..., r_N)$. We define the overall diversity gain d based on the performance of the worst case receiver as

$$d = \min_{1 \leq j \leq N} \{d_j\},$$

where we require all the receivers to decode the common information[4]. Now, as a first step, one can see that if $r_c = 0$ then the techniques developed for the relay channel can be *exported* to this setting through a proportional time sharing strategy. With this assumption, all the properties of the NAF and DDF protocols, established for the relay

[4]Clearly this definition does not allow for different Quality of Service (QOS) constraints.

channel, carry over to this scenario. The problem becomes slightly more challenging when $r_c > 0$. In fact, it is easy to see that, for a fixed total rate, the highest probability of error corresponds to the case where all destinations are required to decode all the streams. This translates to the following condition (that applies to any cooperation scheme)

$$d(r_c, r_1, r_2, ..., r_N) \leq d(r_c + r_1 + ... + r_N, 0, 0, .., 0). \tag{16}$$

So, we will focus the following discussion on this worst case scenario, i.e.,

$$\mathbf{r} = (r_c, 0, 0, ..., 0)\,, 0 \leq r_c \leq 1. \tag{17}$$

The first observation is that, in this scenario, the only AF strategy that achieves the full rate extreme point ($r = 1, d = 0$) is the non-cooperative protocol. Any other AF strategy will require some of the nodes to re-transmit, and therefore not to *listen* during parts of the codeword[5], which prevents it from achieving full rate. Fortunately, this drawback can be avoided in the DDF protocol. The reason is that, in this protocol, any node will start helping only after it has successfully decoded the message. We now propose a protocol for the CB scenario that is a direct extension of the DDF relay protocol. This will be referred to as the CB-DDF protocol in the sequel. The only modification needed, compared to the relay channel case, is that now every node can act as a relay for the other nodes, based on its instantaneous channel gain. Similar to the relay channel, the protocol must include a mechanism that keeps every destination informed of the re-transmission starting times of all the other nodes. Again, in deriving the following result, we ignore the associated cost of this mechanism, relying on the asymptotic assumptions.

Theorem 6 *The diversity-multiplexing tradeoff achieved by the CB-DDF protocol with N destinations is given by:*

$$d(r_c) = \begin{cases} N(1 - r_c), & \frac{1}{N} \geq r_c \geq 0, \\ 1 + \frac{(N-1)(1-2r_c)}{1-r_c}, & \frac{1}{2} \geq r_c \geq \frac{1}{N}, \\ \frac{1-r_c}{r_c}, & 1 \geq r_c \geq \frac{1}{2}. \end{cases} \tag{18}$$

Proof: Please refer to the Appendix.

It is interesting to note that this is exactly the same tradeoff obtained in the relay channel. This implies that requiring all nodes to decode the same message does not entail a price in terms of the achievable tradeoff.

5 The Half-Duplex Cooperative Multiple-Access Channel

In this section, we consider the cooperative multiple-access (CMA) scenario, where N sources transmit their independent messages to a common destination. We assume symmetry so that all sources transmit information at the same rate and are limited by the same power constraint. The basic idea of the proposed protocol, which we refer to as the CMA-NAF protocol, is to create an artificial ISI channel. Towards this end, each of the N

[5] This follows from the half-duplex constraint.

sources transmits once per cooperation frame, where a cooperation frame is defined as N consecutive symbol-intervals. Each source is assigned unique *transmission and reception* symbol-intervals within the cooperation frame. During its transmission symbol-interval, a source transmits a linear combination of its own symbol and the signal it observed during its most recent reception symbol-interval. In other words, every source is assigned a "helper" which repeats a noisy version of its symbol sometime within the next $N-1$ symbol-intervals. Without loss of generality, we set the j^{th} source transmission symbol-interval equal to j.

We now provide an illustrative example for the $N = 3$ case. Here we assume that sources 1, 2, and 3 help sources 3, 1, and 2, respectively. For the j^{th} source and the k^{th} cooperation frame, $t_{j,k}$ denotes the transmission, $r_{j,k}$ the (assigned) reception, and $x_{j,k}$ the originating symbol. Using a_j and b_j to denote the broadcast and repetition gains of the j^{th} source, respectively, the signals transmitted during the first two cooperation frames would be (in chronological order)

$$\begin{aligned}
t_{1,1} &= a_1 x_{1,1} \\
t_{2,1} &= a_2 x_{2,1} + b_2 r_{2,1} \\
t_{3,1} &= a_3 x_{3,1} + b_3 r_{3,1} \\
t_{1,2} &= a_1 x_{1,2} + b_1 r_{1,1} \\
t_{2,2} &= a_2 x_{2,2} + b_2 r_{2,2} \\
t_{3,2} &= a_3 x_{3,2} + b_3 r_{3,2}.
\end{aligned}$$

Using h_{ji} to denote the i^{th}-source-to-j^{th}-source channel gain, and $w_{j,k}$ to denote the noise observed by the j^{th} source during its k^{th}-frame reception symbol-interval, the assigned receptions become

$$\begin{aligned}
r_{2,1} &= h_{21} t_{1,1} + w_{2,1} \\
r_{3,1} &= h_{32} t_{2,1} + w_{3,1} \\
r_{1,1} &= h_{13} t_{3,1} + w_{1,1} \\
r_{2,2} &= h_{21} t_{1,2} + w_{2,2} \\
r_{3,2} &= h_{32} t_{2,2} + w_{3,2}.
\end{aligned}$$

Using g_j to denote the j^{th}-source-to-destination channel gain, and $v_{j,k}$ to denote the noise observed by the destination during the j^{th} symbol-interval of the k^{th} frame, the signals observed at the destination would be (in chronological order)

$$\begin{aligned}
y_{1,1} &= g_1 t_{1,1} + v_{1,1} \\
y_{2,1} &= g_2 t_{2,1} + v_{2,1} \\
y_{3,1} &= g_3 t_{3,1} + v_{3,1} \\
y_{1,2} &= g_1 t_{1,2} + v_{1,2} \\
y_{2,2} &= g_2 t_{2,2} + v_{2,2} \\
y_{3,2} &= g_3 t_{3,2} + v_{3,2}.
\end{aligned}$$

The source-observed noises $\{w_{j,k}\}$ have variance σ_w^2 for all j, k, and the destination-observed noises $\{v_{j,k}\}$ have variance σ_v^2 for all j, k. Note that, as mandated by our half-duplex constraint, no source transmits and receives simultaneously. The broadcast and repetition gains $\{a_j, b_j\}$ should be chosen to satisfy the average power constraint

$$E\{|t_{j,k}|^2\} \leq E. \tag{19}$$

Let us now define L consecutive cooperation frames as a super-frame. We will assume that helper assignments are fixed within a super-frame but are scheduled to change across super-frames. We impose the following requirements on helper scheduling.

1. In each super-frame, every source is helped by a different source.

2. Across super-frames, every source is helped equally by every other source.

Among the many scheduling rules that satisfy these requirements, we choose the following circular rule. In super-frame i, sources with indices $(1, \ldots, N)$ are assigned helpers with indices given by the j^{th} right circular shift of $(1, \ldots, N)$, where $j = \langle i-1 \rangle_{N-1} + 1$. For example, when $N = 4$, the helper configurations are given by the following table.

Super-frame index	Helper assigned to 1	2	3	4
1	4	1	2	3
2	3	4	1	2
3	2	3	4	1
4	4	1	2	3

Since this scheduling algorithm generates $N-1$ distinct helper configurations, the length of the super-frames is chosen such that a coherence-interval consists of $N-1$ consecutive super-frames. To achieve maximal diversity for a given multiplexing gain, it is required that all codewords span the entire coherence-interval. For this reason, we choose codes of length l given by

$$l = (N-1)L. \tag{20}$$

Similar to the broadcast channel, defining the multiplexing gain r and diversity gain d for the cooperative multiple-access channel requires some care. Note that, using (3), the pair (r_j, d_j) can be defined for communication between the j^{th} source and the destination. However, since we assumed a symmetric CMA setup, all multiplexing gains are equal, i.e., $r = r_j$ for all j. Furthermore, since CMA-NAF mandates that only one source transmits in any symbol-interval, the destination's multiplexing gain is also equal to r. That is, the destination receives information at rate R given by

$$R = r \log(\rho). \tag{21}$$

We define the overall diversity gain d based on the worst case probability of error for the N information streams, i.e.,

$$d = \min_{1 \leq j \leq N} \{d_j\}.$$

With these definitions, Theorem 7 establishes the optimality of the CMA-NAF in the symmetric scenario with N sources.

Theorem 7 *The CMA-NAF protocol achieves the optimal (genie-aided) diversity-multiplexing tradeoff for the symmetric scenario with N sources, given by*

$$d^*(r) = N(1-r). \tag{22}$$

Proof: Please refer to the Appendix.

Theorem 7 not only establishes the optimality of the CMA-NAF protocol, but also it shows that the half-duplex constraint does not entail any cost, in terms of the tradeoff, in the symmetric CMA channel. One can now attribute the sub-optimality of previously proposed CMA schemes to the use of orthogonal subspaces. It is interesting to observe that one can achieve the optimal tradeoff in the symmetric CMA channel with a simple AF strategy. In fact, by comparing Theorems 1 and 7, one can see the fundamental difference between the half duplex CMA and relay channels.

6 Numerical Results

In this section, we report numerical results that quantify the performance gains offered by the proposed protocols. These numerical results correspond to outage probabilities and are meant to show that the superiority of the proposed protocols in terms of diversity-multiplexing tradeoff translates into significant SNR gains. In Fig. 5, Fig. 6, and Fig. 7, we compare the proposed protocols with the non-cooperative (direct transmission) and the LTW-AF protocols. To ensure fairness, we have imposed more strict power constraints on the NAF and the DDF relay protocols; specifically, we lowered the average transmission energy of the source and the relay from E to $E/2$ during the interval when both are transmitting. This way, the total average energy per symbol-interval, spent by any of the protocols considered here is E[6]. While one may find other energy allocation strategies that offer performance improvement (in terms of the outage probability), any such optimization will not affect the achievable diversity-multiplexing tradeoff, and hence, will not be pursued here. To obtain lower bounds on the gains offered by the DDF and CMA-NAF protocols, we assume a noiseless source-relay channel for the LTW-AF and NAF relay protocols. For the DDF relay and the CMA-NAF protocols, the SNR of the link between the two cooperating partners was assumed to be only 3 dB better than that of the relay-destination or source-destination channels. We optimized the broadcast and repetition gains for the CMA-NAF protocol experimentally. In all the considered cases, the outage probabilities are computed through Monte-Carlo simulations.

Fig. 5 shows the performance gain offered by the NAF relay protocol over both the non-cooperative protocol and the LTW-AF protocol at high SNRs and two different data rates. The same comparison is repeated in Fig. 6 with the DDF protocol where, as expected, the gains are shown to be larger. The CMA channel is considered in Fig. 7 where the optimality of the CMA-NAF protocol is shown to translate into significant SNR gains. It is also interesting to note that the gap between CMA-NAF performance and genie-aided strategy is less than 3 dB when the date rate is equal to 2 BPCU. We can also observe that the gains offered by the DDF and CMA-NAF protocols compared with the LTW-AF protocol increase with the date rate. This is a direct consequence of the higher multiplexing gains achievable with our newly proposed protocols. Overall, these results re-emphasize the fact that the full diversity criterion alone[7] is a rather weak design tool.

We conclude this section with a brief comment on our choice for the diversity-multiplexing tradeoff as our design tool. This choice is inspired by the convenient tradeoff, between analytical tractability and accuracy, that this tool offers. Ideally, one should seek

[6] In the CMA-NAF protocol, the constant average energy per symbol interval is automatically implied.

[7] Full diversity corresponds to the point $(d = 2, r = 0)$ on the tradeoff curve.

cooperation schemes that minimize the outage probability at the target rate and SNR. Unfortunately, it is easy to see that such an approach would lead to an intractable problem even in very simplified scenarios. Our results, on the other hand, demonstrate that one can use the diversity-multiplexing tradeoff to analytically guide the design in many relevant scenarios. From the accuracy point of view, our simulation results validate that schemes with better tradeoff characteristics *always* offer significant SNR gains at sufficiently high SNRs. In this context, the main drawback of the diversity-multiplexing tradeoff is that it fails to predict at which SNR the promised gains will start to appear. For example, from the figures, one can see that the DDF and CMA-NAF schemes yield performance gains at relatively moderate SNRs whereas the NAF protocol only offers gain at larger SNRs.

7 Conclusions

In this paper, we considered the design of cooperative protocols for a system consisting of half-duplex nodes. In particular, we differentiated between three scenarios. For the relay channel, we investigated the AF and DF protocols. We established the uniform dominance of the proposed DDF protocol compared to all known cooperation strategies and its optimality in a certain range of multiplexing gains. We then proceeded to the cooperative broadcast channel where the gain offered by the DDF strategy was argued to be more significant, as compared to the relay channel. For the multiple-access scenario, we proposed a novel AF cooperative protocol where an *artificial* ISI channel is created. We proved the optimality (in the sense of diversity-multiplexing tradeoff) of this protocol by showing that it achieves the same tradeoff curve as the genie-aided $N \times 1$ point-to-point system.

Our results reveal interesting insights on the structure of optimal cooperation strategies with half-duplex partners. First, we observe that, without the half-duplex constraint, achieving the optimal tradeoff in the three channels considered here is rather straightforward (i.e., one can easily construct a simple AF strategy that results in an N-tap ISI channel, and hence, the optimal tradeoff). With the half-duplex constraint, more care is necessary in constructing the cooperation strategies, but, as shown, one can still achieve the optimal tradeoff in many relevant scenarios. One of the important insights is that one should strive to transmit *independent* symbols as frequently as possible. Indeed, the optimality of the proposed CMA-NAF protocol stems from exploiting the distributed nature of the information to enable transmission of an independent symbol in every symbol interval. It is now easy to see that the use of orthogonal subspaces to enable cooperation, as in [3] for example, entails a significant loss in the achievable tradeoff.

This work poses many interesting questions. For example, proving (or disproving) the optimality of the DDF protocol for the single relay channel and $r > 0.5$ is an open problem (currently under investigation). Generalizations of the proposed schemes to multi-antenna nodes, scenarios with different QOS constraints, and asymmetric CMA channels are of definite interest. Finally, the design of practical coding/decoding strategies that approach the fundamental limits achievable with Gaussian codes and maximum likelihood decoding is an important venue to pursue.

8 Appendix

In this section, we collect all the proofs.

8.1 Proof of Theorem 1

Due to the source average energy constraint, setting A_1 and A_2 to anything other than the identity matrix will reduce the mutual information between $\mathbf{x}$ and $\mathbf{y}$. Since we are interested in obtaining an upper bound, we will choose $A_1 = I_{l'}$ and $A_2 = I_{l-l'}$, in which case (8) reduces to

$$\mathbf{y} = \begin{bmatrix} g_1 I_{l'} & 0 \\ g_2 h B & g_1 I_{l-l'} \end{bmatrix} \mathbf{x} + \begin{bmatrix} 0 \\ g_2 B \end{bmatrix} \mathbf{w} + \mathbf{v}. \tag{23}$$

Using singular value decomposition (SVD), the matrix B can be factored as

$$B = UDV^H,$$

where $U \in \mathbb{C}^{(l-l')\times(l-l')}$ and $V \in \mathbb{C}^{l'\times l'}$ are unitary and where $D \in \mathbb{C}^{(l-l')\times l'}$ is non-negative diagonal with the diagonal elements in decreasing order. Using these matrices, we define $\tilde{\mathbf{y}} \triangleq T\mathbf{y}$, $\tilde{\mathbf{x}} \triangleq T\mathbf{x}$, $\tilde{\mathbf{v}} \triangleq T\mathbf{v}$, and $\tilde{\mathbf{w}} \triangleq T\mathbf{w}$, for unitary transformation

$$T \triangleq \begin{bmatrix} V^H & 0 \\ 0 & U^H \end{bmatrix}.$$

The unitary property of T implies that $\Sigma_{\tilde{\mathbf{w}}} = \sigma_w^2 I_l$ and $\Sigma_{\tilde{\mathbf{v}}} = \sigma_v^2 I_l$, as well as

$$I(\mathbf{x};\mathbf{y}) = I(\tilde{\mathbf{x}};\tilde{\mathbf{y}}). \tag{24}$$

In terms of the new variables, (23) becomes

$$\begin{aligned} \tilde{\mathbf{y}} &= \begin{bmatrix} g_1 I_{l'} & 0 \\ g_2 h D & g_1 I_{l-l'} \end{bmatrix} \tilde{\mathbf{x}} + \begin{bmatrix} 0 \\ g_2 D \end{bmatrix} \tilde{\mathbf{w}} + \tilde{\mathbf{v}} \\ &= \begin{bmatrix} g_1 I_{l'} & 0 \\ g_2 h D & g_1 I_{l-l'} \end{bmatrix} \tilde{\mathbf{x}} + \tilde{\mathbf{n}} \end{aligned} \tag{25}$$

with

$$\Sigma_{\tilde{\mathbf{n}}} = \begin{bmatrix} \sigma_v^2 I_{l'} & 0 \\ 0 & \sigma_v^2 I_{l-l'} + |g_2|^2 \sigma_w^2 D D^H \end{bmatrix}. \tag{26}$$

If we denote the non-zero diagonal elements of D as $\{d_i\}_{i=1}^m$, then (26) can be written as

$$\begin{aligned} \tilde{\mathbf{y}}_i &= G_i \tilde{\mathbf{x}}_i + \tilde{\mathbf{n}}_i, \quad i = 1,\ldots,m \\ \tilde{y}_i &= g_i \tilde{x}_i + \tilde{n}_i, \quad i = m+1,\ldots,l' \text{ and } i = l'+m+1,\ldots,l, \end{aligned}$$

where $\tilde{y}_i$, $\tilde{x}_i$ and $\tilde{n}_i$ represent the i^{th} element of $\tilde{\mathbf{y}}$, $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{n}}$, respectively, and where $\tilde{\mathbf{y}}_i \triangleq [\tilde{y}_i, \tilde{y}_{l'+i}]^t$, $\tilde{\mathbf{x}}_i \triangleq [\tilde{x}_i, \tilde{x}_{l'+i}]^t$, $\tilde{\mathbf{n}}_i \triangleq [\tilde{n}_i, \tilde{n}_{l'+i}]^t$, and

$$\begin{aligned} G_i &\triangleq \begin{bmatrix} g_1 & 0 \\ g_2 h d_i & g_1 \end{bmatrix} \\ \Sigma_{\tilde{\mathbf{n}}_i} &= \begin{bmatrix} \sigma_v^2 & 0 \\ 0 & \sigma_v^2 + |g_2|^2 d_i^2 \sigma_w^2 \end{bmatrix}. \end{aligned}$$

Note that, according to the SVD theorem,

$$m \leq \min\{l', l - l'\}. \tag{27}$$

Because $\Sigma_{\tilde{\mathbf{n}}}$ is diagonal, $I(\tilde{\mathbf{x}}; \tilde{\mathbf{y}})$ (and therefore $I(\mathbf{x}; \mathbf{y})$) is maximized when $\{\tilde{\mathbf{x}}_i\}_{i=1}^{m} \cup \{\tilde{x}_i\}_{i=m+1}^{l'} \cup \{\tilde{x}_i\}_{i=l'+m+1}^{l}$ are mutually independent, in which case we would have

$$I(\tilde{\mathbf{x}}; \tilde{\mathbf{y}}) = \sum_{i=1}^{m} I(\tilde{\mathbf{x}}_i; \tilde{\mathbf{y}}_i) + \sum_{i=m+1}^{l'} I(\tilde{x}_i; \tilde{y}_i) + \sum_{i=l'+m+1}^{l} I(\tilde{x}_i; \tilde{y}_i). \tag{28}$$

The mutual information between $\tilde{\mathbf{x}}_i$ and $\tilde{\mathbf{y}}_i$ is given by

$$I(\tilde{\mathbf{x}}_i; \tilde{\mathbf{y}}_i) = \log(\det{(I_2 + \Sigma_{\tilde{\mathbf{n}}_i}^{-\frac{1}{2}} G_i \Sigma_{\tilde{\mathbf{x}}_i} G_i^H \Sigma_{\tilde{\mathbf{n}}_i}^{-\frac{1}{2}})}). \tag{29}$$

A lower-bound on $I(\tilde{\mathbf{x}}_i; \tilde{\mathbf{y}}_i)$ is easily obtained by replacing $\Sigma_{\tilde{\mathbf{x}}_i}$ by EI_2:

$$\log(\det{(I_2 + EG_i G_i^H \Sigma_{\tilde{\mathbf{n}}_i}^{-1})}) \leq I(\tilde{\mathbf{x}}_i; \tilde{\mathbf{y}}_i).$$

Since $\log(\det{(.)})$ is an increasing function on the cone of positive-definite Hermitian matrices and since $\lambda_{\max} I_2 - \Sigma_{\tilde{\mathbf{x}}_i} \geq 0$ (where $\lambda_{\max}$ represents the largest eigenvalue of $\Sigma_{\tilde{\mathbf{x}}}$), we get the following upper-bound on $I(\tilde{\mathbf{x}}_i; \tilde{\mathbf{y}}_i)$:

$$I(\tilde{\mathbf{x}}_i; \tilde{\mathbf{y}}_i) \leq \log(\det{(I_2 + \lambda_{\max} G_i G_i^H \Sigma_{\tilde{\mathbf{n}}_i}^{-1})}).$$

Now, since $\lambda_{\max}$ is of the same exponential order as E, the bounds converge as ρ grows to infinity. Then, it is straightforward to show

$$\lim_{\rho \to \infty} \max_{\Sigma_{\tilde{\mathbf{x}}_i}, d_i} I(\tilde{\mathbf{x}}_i; \tilde{\mathbf{y}}_i) = (\max\{2(1 - v_1), 1 - u\})^+ \log(\rho),$$

where v_1, v_2 and u are the exponential orders of $1/|g_1|^2$, $1/|g_2|^2$ and $1/|h|^2$, respectively. In deriving this expression, we have assumed that $(v_1, v_2, u) \in \mathbb{R}^{3+}$; as explained earlier, we do not need to consider realizations in which v_1, v_2 or u are negative. Similarly,

$$\lim_{\rho \to \infty} \max I(\tilde{x}_i; \tilde{y}_i) = (1 - v_1)^+ \log(\rho),$$

which, together with (24) and (28), results in:

$$\lim_{\rho \to \infty} \max_{\Sigma_{\mathbf{x}}} I(\mathbf{x}; \mathbf{y}) = [(l - 2m)(1 - v_1)^+ + m(\max\{2(1 - v_1), 1 - u\})^+] \log(\rho). \tag{30}$$

For the quasi-static fading setup, the outage event is defined as the set of channel realizations for which the instantaneous capacity falls below the target data rate. Thus, our outage event O becomes

$$O = \{(v_1, v_2, u) | \max_{\Sigma_{\mathbf{x}}} I(\mathbf{x}, \mathbf{y}) < lR\}.$$

Letting R grow with ρ according to

$$R = r \log(\rho),$$

and using (30), we conclude that, for large ρ,

$$O^+ = \{(v_1, v_2, u) \in \mathbb{R}^{3+} | (l - 2m)(1 - v_1)^+ + m(\max\{2(1 - v_1), 1 - u\})^+ < rl\}, \tag{31}$$

and thus

$$P_O(R)\dot{=}\rho^{-d_o(r)} \quad \text{for} \quad d_o(r) = \inf_{(v_1,v_2,u)\in O^+}(v_1+v_2+u). \tag{32}$$

As Zheng and Tse have shown in Lemma 5 of [8], $d_o(r)$ provides an upper-bound on $d^*(r)$ (i.e., the optimal diversity gain at multiplexing gain r):

$$d^*(r) \le d_o(r). \tag{33}$$

From (31) and (32), it is easy to see that the right hand side of (33) is maximized when m is set to its maximum, which, according to (27), is $\min\{l', l-l'\}$. This is the case when B is full-rank. On the other hand, $\min\{l', l-l'\}$ itself is maximized when $l' = l/2$ (assuming an even codeword length l), which corresponds to B being a square matrix. For this B, $d_o(r)$ can be shown to take the value of the right hand side of (10). This completes the proof.

8.2 Proof of Theorem 2

The proof closely follows that for the MIMO point-to point communication system in [8]. In particular, we assume that the source uses a Gaussian random code-book of codeword length l, where l is taken to be even, and data rate R, where R increases with ρ according to

$$R = r\log(\rho).$$

The error probability of the ML decoder, $P_e(\rho)$, can be upper bounded using Bayes' rule:

$$\begin{aligned} P_e(\rho) &= P_O(R)P_{e|O} + P_{e,O_c} \\ P_e(\rho) &\le P_O(R) + P_{e,O_c}, \end{aligned}$$

where O and O_c denote the outage event and its complement, respectively. The outage event O is chosen such that $P_O(R)$ dominates P_{e,O_c}, i.e.,

$$P_{e,O_c}\dot{\le}P_O(R), \tag{34}$$

in which case

$$P_e(\rho)\dot{\le}P_O(R). \tag{35}$$

In order to characterize O, we note that, since the destination observations during different frames are independent, the upper-bound on the ML conditional PEP [recalling (7)] changes to

$$P_{pe|g_1,g_2,h} \le \det\left(I_2 + \frac{1}{2}\Sigma_s\Sigma_n^{-1}\right)^{-l/2}, \tag{36}$$

where here $\Sigma_\mathbf{s}$ and $\Sigma_\mathbf{n}$ denote the covariance matrices of the destination's observation vector and its noise component during a single frame:

$$\Sigma_\mathbf{s} = \begin{bmatrix} |g_1|^2 & g_1 g_2^* b^* h^* \\ g_1^* g_2 bh & |g_1|^2 + |g_2|^2|bh|^2 \end{bmatrix} E \tag{37}$$

$$\Sigma_\mathbf{n} = \begin{bmatrix} \sigma_v^2 & 0 \\ 0 & \sigma_v^2 + |g_2|^2|b|^2\sigma_w^2 \end{bmatrix}. \tag{38}$$

Let us define v_1, v_2, u, and w as the exponential orders of $1/|g_1|^2$, $1/|g_2|^2$, $1/|h|^2$, and $|b|^2$, respectively. Then the constraint on b given in (11) implies the following constraint on w:

$$w \leq \min\{u, 1\} \tag{39}$$

We assume b is chosen such that the exponential order w becomes

$$w \triangleq (u)^{-}.$$

which satisfies the constraint given by (39). Interestingly, if we consider $(v_1, v_2, u) \in \mathbb{R}^{3+}$, then w becomes zero and vanishes in the expressions. Plugging (37)-(38) into (36), we obtain

$$P_{pe|v_1,v_2,u} \dot{\leq} \rho^{-\frac{l}{2}(\max\{2(1-v_1),1-(v_2+u)\})^{+}} \quad \text{for} \quad (v_1, v_2, u) \in \mathbb{R}^{3+}.$$

With rate $R = r \log \rho$ BPCU and codeword length l, we have a total of ρ^{rl} codewords. Thus,

$$P_{e|v_1,v_2,u} \dot{\leq} \rho^{-\frac{l}{2}[(\max\{2(1-v_1),1-(v_2+u)\})^{+}-2r]} \quad \text{for} \quad (v_1, v_2, u) \in \mathbb{R}^{3+}.$$

P_{e,O_c} is the average of $P_{e|v_1,v_2,u}$ over the set of channel realizations that do not cause an outage (i.e., O_c). Using (5), one can see that

$$P_{e,O_c} \dot{\leq} \int_{O_c^+} \rho^{-d_e(r,v_1,v_2,u)} dv_1 dv_2 du.$$

for

$$d_e(r, v_1, v_2, u) = \frac{l}{2}[(\max\{2(1-v_1), 1-(v_2+u)\})^{+} - 2r] + (v_1 + v_2 + u).$$

Now, P_{e,O_c} is dominated by the term corresponding to the minimum value of $d_e(r, v_1, v_2, u)$ over O_c^+, where $O_c^+ \triangleq O_c \cap \mathbb{R}^{3+}$:

$$P_{e,O_c} \dot{\leq} \rho^{-d_e(r)} \quad \text{for} \quad d_e(r) = \inf_{v_1,v_2,u \in O_c^+} d_e(r, v_1, v_2, u). \tag{40}$$

Using (6), $P_O(R)$ can be expressed

$$P_O \doteq \rho^{-d_o(r)} \quad \text{for} \quad d_o(r) = \inf_{(v_1,v_2,u) \in O^+} (v_1 + v_2 + u). \tag{41}$$

Comparing (40) and (41), we realize that for (34) to be met, O^+ should be defined as

$$O^+ = \{(v_1, v_2, u) \in \mathbb{R}^{3+} | (\max\{2(1-v_1), 1-(v_2+u)\})^{+} \leq 2r\}.$$

Then, for any $(v_1, v_2, u) \in O_c^+$, it is possible to choose l to make $d_e(r, v_1, v_2, u)$ arbitrarily large, ensuring (34). Note that, because of (35), $d_o(r)$ provides a lower-bound on the diversity gain achieved by the protocol. But $d_o(r)$, as given by (41), turns out to be identical to right hand side of (10) (refer to Fig. 8). Thus the optimal diversity-multiplexing tradeoff for this scenario is indeed given by (12) and the NAF protocol achieves it.

8.3 Proof of Theorem 4

For the single relay DDF protocol, the average conditional pairwise error probability (PEP) is given by[8]

$$P_{pe|g_1,g_2,h} \leq \left(1+|g_1|^2\frac{E}{2\sigma_v^2}\right)^{-l'}\left(1+\left(|g_1|^2+|g_2|^2\right)\frac{E}{2\sigma_v^2}\right)^{-(l-l')}.$$

Defining v_1, v_2, and u as the exponential orders of g_1, g_2, and h, respectively, gives

$$P_{pe|v_1,v_2,u} \dot{\leq} \rho^{-l[f(1-v_1)+(1-f)(1-\min\{v_1,v_2\})]} \quad \text{for} \quad (v_1,v_2,u) \in \mathbb{R}^{3+},$$

where $f \triangleq l'/l$. At a rate of $R = r\log\rho$ BPCU and a codeword length of l, there are a total of ρ^{rl} codewords. Thus,

$$P_{e,O_c} \dot{\leq} \rho^{-d_e(r)}$$

for

$$d_e(r) = \inf_{(v_1,v_2,u)\in O_c^+} l[f(1-v_1)+(1-f)(1-\min\{v_1,v_2\})-r] + (v_1+v_2+u) \qquad (42)$$

Examining (42), we realize that for (34) to hold, O^+ should be defined as

$$O^+ = \{(v_1,v_2,u) \in \mathbb{R}^{3+} | f(1-v_1)+(1-f)(1-\min\{v_1,v_2\}) \leq r\}$$

so that it is possible to choose l to make $d_e(r)$ arbitrarily large, ensuring (34). As before, $P_O(R)$ is given by (41), which turns out to be identical to $d(r)$ given by (14). To see this, one needs to consider two cases. The first case, when $f \leq 0.5$, is very easy. Referring to Fig. 9 reveals that, in this case, $\inf_{(v_1,v_2)\in O^+} v_1+v_2$ and therefore $d_o(r)$ is equal to $2(1-r)$ (the genie aided tradeoff). The second case, when $f > 0.5$, is a little bit more difficult. As can be seen from Fig. 10, in this case

$$\inf_{(v_1,v_2)\in O^+} = \frac{1-r}{f}. \qquad (43)$$

On the other hand, from (13), one can show that

$$u = 1 - \frac{r}{f}, \qquad (44)$$

which implies that $f > \max\{r, 0.5\}$ (u is nonnegative and we had assumed $f > 0.5$). From (44) and (43), we conclude that

$$d_o(r) = \inf_{f>\max\{r,0.5\}} 1 + \frac{1-2r}{f}, \qquad (45)$$

which gives (14). Again, according to (35), $d_o(r)$ provides a lower-bound on the diversity gain achieved by the protocol. On the other hand, $d_o(r)$ is also an upper bound on the diversity since: 1) for $0 \leq r \leq 0.5$ $d_o(r)$ is the genie-aided diversity and 2) for $0.5 \leq r \leq 1$ it is easy to see that $v_1 = \frac{1-r}{r} + \epsilon$, $v_2 = 0$ and $u = 0$ correspond to a *channel* outage for any $\epsilon > 0$. Thus (14) is the diversity achieved by the DDF protocol and the proof is complete.

[8] Note that the probability of error at the relay is arbitrarily small with the asymptotically large block length considered here, and hence, is ignored in the proof.

8.4 Proof of Theorem 5

In this scenario, the PEP of the ML decoder conditioned on a certain channel realization and averaged over the ensemble of Gaussian random codes, is upper-bounded as

$$P_{pe|g_j,h_j} \leq \prod_{j=1}^{N} \left[1 + \left(\sum_{i=1}^{j} |g_j|^2\right) \frac{E}{2\sigma_v^2}\right]^{-l_j},$$

where the nodes are labeled according to the order in which they start transmission. That is, the source is labeled as node 1, the first relay that starts transmission as node 2, and so on. As before, the gain of the channel that connects the j^{th} node to the destination is denoted by g_j, while the gain of the channel that connects nodes i and j is denoted by h_{ji}. We use l_j to denote the number of symbol-intervals in the codeword during which a total of j nodes are transmitting, so that $\sum_{j=1}^{N} l_j = l$, with l denoting the total codeword length. Note that $\sum_{j=1}^{p} l_j$ is the number of symbol-intervals that relay $p+1$ has to wait, before it collects sufficient energy for error-free decoding of the message. Thus

$$\sum_{j=1}^{p} l_j \leq \min\{l, \lceil \frac{lR}{\log(1+|h_{p+1,1}|^2 c\rho)} \rceil\}, \quad \text{for } N-1 \geq p \geq 1. \tag{46}$$

Defining v_j and u_{ji} as the exponential orders of g_j and h_{ji}, respectively, we have

$$P_{pe|v_j,u_{ji}} \dot{\leq} \rho^{-\sum_{j=1}^{N} l_j (1-\min\{v_1,\ldots,v_j\})^+}.$$

Choosing $R = r\log(\rho)$ for a total of ρ^{rl} codewords, the following expression for the conditional error probability can be derived.

$$P_{e|v_j,u_{ji}} \dot{\leq} \rho^{-l\left[\sum_{j=1}^{N} \frac{l_j}{l}(1-\min\{v_1,\ldots,v_j\})^+ - r\right]}.$$

Thus, O^+ is the set of channel realizations that satisfy

$$\sum_{j=1}^{N} \frac{l_j}{l}(1-\min\{v_1,\ldots,v_j\})^+ \leq r,$$

which can be simplified to

$$1 - r \leq \sum_{j=1}^{N} \frac{l_j}{l} \min\{1, v_1, \ldots, v_j\}. \tag{47}$$

As before, $P_O(R)$ is characterized by

$$P_O(R) \dot{=} \rho^{-d_o(r)} \quad \text{for} \quad d_o(r) = \inf_{O^+} \sum_{j=1}^{N} \left(v_j + \sum_{i<j} u_{ji}\right). \tag{48}$$

Defining $\tilde{v}_j \triangleq \min\{v_1, \cdots, v_j\}, j = 1, \cdots, N$ lets us simplify (47) and (48) to

$$1 - r \leq \sum_{j=1}^{N} \frac{l_j}{l} \min\{1, \tilde{v}_j\} \tag{49}$$

$$d_o(r) = \inf_{O^+} \sum_{j=1}^{N} \left(\tilde{v}_j + \sum_{i<j} u_{ji}\right). \tag{50}$$

From the definition of $\tilde{v}_j$, it follows that

$$\tilde{v}_1 \geq \tilde{v}_2 \geq \cdots \geq \tilde{v}_N \geq 0.$$

Note that (46) can also be simplified to

$$\sum_{j=1}^{p} \frac{l_j}{l} \leq \min\{1, \lceil \frac{r}{(1-u_{p+1,1})^+} \rceil\}, \quad \text{for} \quad N-1 \geq p \geq 1,$$

or

$$1 - \frac{r}{\sum_{k=1}^{p} \frac{l_k}{l}} \leq u_{j1}, \qquad \text{for} \quad j > p. \tag{51}$$

In order to characterize $d_o(r)$, we need to consider three cases. The first case is when $1 \geq \tilde{v}_1$. In this case, (49) simplifies to

$$1 - r \leq \sum_{j=1}^{N} \frac{l_j}{l} \tilde{v}_j.$$

Let us define $x_j \triangleq j(\tilde{v}_j - \tilde{v}_{j+1}), j = 1, \cdots, N-1$ and $x_N \triangleq N\tilde{v}_N$. It immediately follows that $x_j \geq 0, j = 1, \ldots, N$. It is also easy to verify that

$$\sum_{j=1}^{N} \tilde{v}_j = \sum_{j=1}^{N} x_j \quad \text{and} \quad 1 - r \leq \sum_{j=1}^{N} \frac{f_j}{j} x_j, \tag{52}$$

where $f_j \triangleq \sum_{k=1}^{j} l_k / l$. From (52), it can be seen that

$$\inf_{\substack{O^+ \\ 1 \geq \tilde{v}_1}} \sum_{j=1}^{N} \tilde{v}_j = p(\frac{1-r}{f_p}), \quad \text{where} \quad p = \arg \max_{N \geq j \geq 1} \{\frac{f_j}{j}\}. \tag{53}$$

The infimum value corresponds to $x_p = p(1-r)/f_p$ and $x_j = 0, j \neq p$ or $\tilde{v}_j = (1 - r)/f_p, p \geq j \geq 1$ and $\tilde{v}_j = 0, j > p$. But we assumed $1 \geq \tilde{v}_1$, so

$$f_p \geq 1 - r. \tag{54}$$

From (51), it follows that,

$$\inf_{\substack{O^+ \\ 1 \geq \tilde{v}_1}} \sum_{j>p} u_{j1} = (N-p)(1 - \frac{r}{f_p}), \quad 1 \geq f_p \geq r. \tag{55}$$

Now, from (53) and (55) we conclude that

$$\inf_{\substack{O^+, \\ 1 \geq \tilde{v}_1}} \sum_{j=1}^{N} \left(\tilde{v}_j + \sum_{i<j} u_{ji} \right) \geq \inf_{\substack{N \geq p \geq 1, \\ 1 \geq f_p \geq max\{r, 1-r\}}} d_o(r, p, f_p), \tag{56}$$

where,

$$d_o(r, p, f_p) \triangleq p(\frac{1-r}{f_p}) + (N-p)(1 - \frac{r}{f_p}). \tag{57}$$

It turns out that, (57) is an increasing function of p. Therefore, its infimum corresponds to $p = 1$. Now, examining $d_o(r, 1, f_1)$, i.e.,

$$d_o(r, 1, f_1) = (\frac{1-r}{f_1}) + (N-1)(1 - \frac{r}{f_1}),$$

we realize that, for $1/N \geq r \geq 0$, it decreases with f_1, thus its infimum corresponds to $f_1 = 1$. On the other hand, for $1 \geq r \geq 1/N$, $d_o(r, 1, f_1)$ becomes an increasing function of f_1, which means that its infimum corresponds to $f_1 = \max\{r, 1-r\}$, i.e.

$$\inf_{\substack{O^+, \\ 1 \geq \tilde{v}_1}} \sum_{j=1}^{N} \left(\tilde{v}_j + \sum_{i<j} u_{ji} \right) \geq \begin{cases} N(1-r), & \frac{1}{N} \geq r \geq 0, \\ 1 + \frac{(N-1)(1-2r)}{1-r}, & \frac{1}{2} \geq r \geq \frac{1}{N}, \\ \frac{1-r}{r}, & 1 \geq r \geq \frac{1}{2}. \end{cases} \tag{58}$$

The second case to be considered is when $\tilde{v}_i > 1 \geq \tilde{v}_{i+1}$, $N-1 \geq i \geq 1$. It immediately follows that

$$\inf_{\substack{O^+ \\ \tilde{v}_i > 1 \geq \tilde{v}_{i+1}}} \sum_{j=1}^{i} \tilde{v}_j = i. \tag{59}$$

In this case, (49) can be written as

$$1 - r - f_i \leq \sum_{j=i+1}^{N} \frac{l_j}{l} \tilde{v}_j. \tag{60}$$

If $f_i \geq 1-r$, then from (60), we get

$$\inf \sum_{j=i+1}^{N} \tilde{v}_j = 0. \tag{61}$$

On the other hand, from (51), it follows that,

$$\inf \sum_{j=i+1}^{N} u_{j1} = (N-i)(1 - \frac{r}{f_i}), \quad 1 \geq f_i \geq r. \tag{62}$$

Now, from (59), (61) and (62) one can see that

$$\inf_{\substack{O^+, \\ \tilde{v}_i > 1 \geq \tilde{v}_{i+1}, \\ f_i \geq 1-r}} \sum_{j=1}^{N} \left(\tilde{v}_j + \sum_{i<j} u_{ji} \right) \geq \inf_{\substack{N-1 \geq i \geq 1, \\ 1 \geq f_i \geq max\{r, 1-r\}}} d_o(r, i, f_i),$$

with

$$d_o(r, i, f_i) \triangleq i + (N-i)(1 - \frac{r}{f_i}).$$

The infimum of $d_o(r, i, f_i)$ corresponds to $i = 1$ and $f_i = \max\{r, 1-r\}$, i.e.,

$$\inf_{\substack{O^+, \\ \tilde{v}_i > 1 \geq \tilde{v}_{i+1}, \\ f_i \geq 1-r}} \sum_{j=1}^{N} \left(\tilde{v}_j + \sum_{i<j} u_{ji} \right) \geq \begin{cases} 1 + \frac{(N-1)(1-2r)}{1-r}, & \frac{1}{2} \geq r \geq 0, \\ 1, & 1 \geq r \geq \frac{1}{2}. \end{cases} \tag{63}$$

If $f_i < 1-r$, then the problem of finding $\inf \sum_{j=i+1}^{N} \tilde{v}_j$ reduces to the first case (i.e., $1 \geq \tilde{v}_1$). Specifically, $\inf \sum_{j=i+1}^{N} \tilde{v}_j$ is given by (53), with $N-i$, $f_p - f_i$, $r + f_i$ and $p - i$ substituting N, f_p, r and p. Thus,

$$\inf(\sum_{j=i+1}^{N} \tilde{v}_j) = (p-i)(\frac{1-r-f_i}{f_p - f_i}), \qquad \text{where } \; p = \arg \max_{N \geq j \geq i+1} \{\frac{f_j - f_i}{j-i}\}. \tag{64}$$

Note that (54) still holds. Derivation of $\inf \sum_{j=1}^{N} \sum_{i<j} u_{ji}$ follows from (51),

$$\inf \sum_{j=1}^{N} \sum_{i<j} u_{ji} = (p-i)(1-\frac{r}{f_i}) + (N-p)(1-\frac{r}{f_p}), \;\; \text{with } \; f_p > f_i \geq r. \tag{65}$$

From (59), (64) and (65), we conclude that

$$\inf_{\substack{O^+, \\ \tilde{v}_i > 1 \geq \tilde{v}_{i+1}, \\ 1-r > f_i}} \sum_{j=1}^{N} \left(\tilde{v}_j + \sum_{i<j} u_{ji} \right) \geq \inf_{\substack{N \geq p > i \geq 1, \\ 1 \geq f_p \geq 1-r > f_i \geq r}} d_o(r, i, p, f_i, f_p), \tag{66}$$

where,

$$d_o(r, i, p, f_i, f_p) \triangleq i + (p-i)(\frac{1-r-f_i}{f_p - f_i}) + (p-i)(1-\frac{r}{f_i}) + (N-p)(1-\frac{r}{f_p}). \tag{67}$$

As can be seen from (67), $d_o(r, i, p, f_i, f_p)$ is a linear, and therefore monotonic, function of p. Thus, its infimum corresponds to either $p = i+1$ or $p = N$. Now if the infimum indeed corresponds to $p = i+1$, by plugging in $p = i$ into (67), we derive a lower-bound on it. That is,

$$\inf_{\substack{N \geq p > i \geq 1, \\ 1 \geq f_p \geq 1-r > f_i \geq r}} d_o(r, i, p, f_i, f_p) \geq \inf_{\substack{N > i \geq 1 \\ 1 \geq f_p \geq 1-r}} i + (N-i)(1-\frac{r}{f_p}).$$

or

$$\inf_{\substack{N \geq p > i \geq 1, \\ 1 \geq f_p \geq 1-r > f_i \geq r}} d_o(r, i, p, f_i, f_p) \geq 1 + (N-1)\frac{1-2r}{1-r}, \;\; \text{for } \; \frac{1}{2} > r \geq 0. \tag{68}$$

Choosing $p = N$, on the other hand, gives

$$d_o(r, i, N, f_i, 1) = i + (N-i)(2 - \frac{r}{1-f_i} - \frac{r}{f_i}),$$

which has an infimum value, corresponding to $i = 1$ and $f_i = r$ or $f_i = 1-r$, identical to the right-hand side of (68). This means that

$$\inf_{\substack{N \geq p > i \geq 1, \\ 1 \geq f_p \geq 1-r > f_i \geq r}} d_o(r, i, p, f_i, f_p) = 1 + (N-1)\frac{1-r}{1-2r}, \;\; \text{for } \; \frac{1}{2} > r \geq 0. \tag{69}$$

Now, from (69), (66) and (63), we conclude that

$$\inf_{\substack{O^+, \\ \tilde{v}_i > 1 \geq \tilde{v}_{i+1}}} \sum_{j=1}^{N} \left(\tilde{v}_j + \sum_{i<j} u_{ji} \right) \geq \begin{cases} 1 + \frac{(N-1)(1-2r)}{1-r}, & \frac{1}{2} \geq r \geq 0, \\ 1, & 1 \geq r \geq \frac{1}{2}. \end{cases} \tag{70}$$

The third case (i.e., $\tilde{v}_N > 1$), is trivial

$$\inf_{\substack{O^+,\\ \tilde{v}_N > 1}} \sum_{j=1}^{N} \left(\tilde{v}_j + \sum_{i<j} u_{ji} \right) \geq N. \tag{71}$$

From (58), (70) and (71) we conclude that (15) provides a lower-bound on the diversity gain achieved by the protocol. On the other hand, $d_o(r)$ is also an upper bound on the diversity since: 1) for $1/N \geq r \geq 0$, $d_o(r)$ is the genie-aided diversity, 2) for $0.5 \geq r \geq 1/N$, it can be shown that the realization, where $v_1 = 1+\epsilon$, $\{v_j\}_{j=2}^{N} = 0$, $\{u_{j1}\}_{j=2}^{N} = \frac{1-2r}{1-r}$ and $\{u_{ji}\}_{i \neq j} = 0$ corresponds to a *channel* outage for any $\epsilon > 0$, and 3) for $1 \geq r \geq 0.5$, realization $v_1 = \frac{1-r}{r} + \epsilon$, $\{v_j\}_{j=2}^{N} = 0$ and $\{u_{ji}\} = 0$ also corresponds to a channel outage for any $\epsilon > 0$. Thus (15) is the diversity achieved by the $N-1$ relay DDF protocol and the proof is complete.

8.5 Proof of Theorem 6

Note that the error probability of destination j can be written as

$$P_e^j(\rho) = P_{e|S^j}^j(\rho) P_{S^j}^j + P_{e|S_c^j}^j(\rho)(1 - P_{S^j}^j), \tag{72}$$

where S^j denotes the event that destination j decodes the message and starts re-transmission before the end of the codeword, and S_c^j denotes its complement. Note $P_{e|S^j}^j(\rho)$ equals zero, since re-transmission occurs only if the message was decoded error-free. On the other hand, $0 \leq 1 - P_{S^j}^j \leq 1$, thus (72) gives

$$P_e^j(\rho) \leq P_{e|S_c^j}^j(\rho). \tag{73}$$

Now, the case in which the j^{th} destination (out of N total) spends the entire codeword listening is identical to that of the DDF relay protocol with $N-1$ relays. Thus, from (73), we see that communication to the j^{th} destination achieves the same diversity order as does the DDF relay protocol with $N-1$ relays, namely, (18). This completes the proof.

8.6 Proof of Theorem 7

Realizing that (22) also corresponds to the optimal diversity-multiplexing tradeoff for a MIMO point-to-point communication system with N transmit antennas and a single receive antenna (i.e., the case of "genie-aided" cooperation between N sources), we only need to show that the CMA-NAF protocol achieves this tradeoff. To achieve this goal, we assume that each of the sources uses a Gaussian random code with codeword length l and data rate R, where l is chosen as in (20) and R grows with ρ according to (21). We then characterize the joint ML decoder's error probability $P_e(\rho)$. Note that the error probability of the joint ML decoder upper-bounds the error probabilities of the source-specific ML decoders and thus provides a lower-bound on the achievable overall diversity gain (as a function of r). In characterizing $P_e(\rho)$, we follow the approach of Tse *et al.* [9] by partitioning the error event e into the set of partial error events $\{e^I\}$, i.e.,

$$e = \bigcup_I e^I,$$

where I denotes any *nonempty* subset of $\{1, ..., N\}$ and e^I (referred to as a "type-I error") is the event that the joint ML decoder incorrectly decodes the messages from sources whose indices belong to I while correctly decoding all other messages. Because the partial error events are mutually exclusive,

$$P_e(\rho) = \sum_I P_{e^I}(\rho). \tag{74}$$

Using Bayes' rule, one can upper-bound $P_{e^I}(\rho)$ as

$$P_{e^I}(\rho) = P_O(R)P_{e^I|O} + P_{e^I,O_c}$$
$$P_{e^I}(\rho) \leq P_O(R) + P_{e^I,O_c},$$

where, as before, O and O_c denote the outage event and its complement, respectively. The outage event is defined such that $P_O(R)$ dominates P_{e^I,O_c} for all I:

$$P_{e^I,O_c} \dot{\leq} P_O(R). \tag{75}$$

Thus,

$$P_{e^I}(\rho) \dot{\leq} P_O(R),$$

which, together with (74), results in

$$P_e(\rho) \dot{\leq} P_O(R). \tag{76}$$

This means that $P_O(R)$, as defined by (75), provides an upper-bound to the joint ML decoder's error probability and therefore a lower-bound to the achievable diversity gain $d^*(r)$. The derivation of $P_O(R)$, however, requires the characterization of $P_{pe^I|g_j,h_{ji}}$ (i.e., the joint ML decoder's type-I PEP, conditioned on a particular channel realization and averaged over the ensemble of Gaussian random codes). Here, we upper-bound $P_{pe^I|g_j,h_{ji}}$, for each I, by the PEP of a suboptimal joint ML decoder that uses only a subset of the destination's observations (referred to as the *type-I decoder*):

$$P_{pe^I|g_j,h_{ji}} \leq \det(I_m + \frac{1}{2}\Sigma_{\mathbf{s}^I}\Sigma_{\mathbf{n}^I}^{-1})^{-1} \tag{77}$$

In (77), $\Sigma_{\mathbf{s}^I}$ and $\Sigma_{\mathbf{n}^I}$ represent the $m \times m$ covariance matrices corresponding to the signal and noise components, respectively, of the *partial* observation vector used by the type-I decoder, provided that the symbols of the sources that are not in set I are set to zero. The size m will be characterized in the sequel.

Before going into more detail on the type-I decoder, we note that, since $\Sigma_{\mathbf{s}^I}$ and $\Sigma_{\mathbf{n}^I}$ are both positive definite matrices, the right-hand side of (77) can be upper-bounded as

$$P_{pe^I|g_j,h_{ji}} \dot{\leq} \det(\Sigma_{\mathbf{s}^I})^{-1}\det(\Sigma_{\mathbf{n}^I}). \tag{78}$$

The discussion is simplified if we define v_j and u_{ji} as the exponential orders of $1/|g_j|^2$ and $1/|h_{ji}|^2$, respectively. Note that the exponential orders of $\{|b_j|^2\}_{j=1}^N$ do not appear in the following expressions for the reasons outlined in the proof of Theorem 2. Also, recalling (5), the PDFs of negative v_j and u_{ji} are effectively zero for large values of ρ, allowing us to concern ourselves only with their non-negative realizations. With this ideas in mind, we return to (78) and claim that

$$\det(\Sigma_{\mathbf{n}^I}) \dot{\leq} 1. \tag{79}$$

To understand (79), recall that the noise component of the destination observation is a linear combination of the noise originating at the sources (i.e., $\{w_{j,k}\}_{j=1}^N$) and the noise originating at the destination (i.e., $v_{j,k}$). Furthermore, the coefficients of this linear combination are the products of some channel, broadcast, and repetition gains. Then, because these noise variances and magnitude-squared gains can be written as non-positive powers of ρ, equation (79) must hold. Combining (79) and (78) yields

$$P_{pe^I|v_j,u_{ji}} \dot{\leq} \det(\Sigma_{\mathbf{s}^I})^{-1} \quad \text{for} \quad v_j \geq 0, u_{ji} \geq 0. \tag{80}$$

As mentioned earlier, $\Sigma_{\mathbf{s}^I}$ represents the covariance matrix of the signal component of the partial observation used by the type-I decoder, provided that the symbols of the sources that are not in I are set to zero. To fully characterize $\Sigma_{\mathbf{s}^I}$, though, we must know which observations are used by the type-I decoder and which are discarded. The type-I decoder picks *one* observation for every source in set I, for a total of $m = |I|$ observations per frame (where $|I|$ denotes the size of I and therefore $1 \leq |I| \leq N$). Provided that frame k is not the last frame in its super-frame and assuming that during this super-frame, source i is helping source $j \in I$, the destination observation component corresponding to source j will be either the $y_{j,k}$ that corresponds to source j's broadcast of $x_{j,k}$ or the $y_{i,k'}$ that corresponds to helper i's re-broadcast of $x_{j,k}$ (where $k' \in \{k, k+1\}$). As an example, consider the case when $N = 4$ and assume that during a certain super-frame, source 3 is helping source $2 \in I$ (i.e., $j = 2$, $i = 3$). In this case, the type-I decoder picks either $y_{2,k}$ or $y_{3,k}$ in correspondence to $x_{2,k}$. However, if instead of source 3, source 1 is helping source 2 (i.e., $j = 2$, $i = 1$), then the type-I decoder has to choose between $y_{2,k}$ or $y_{1,k+1}$. Back to our description of the type-I decoder, if $i \in I$, then the decoder always picks $y_{j,k}$ over $y_{i,k'}$. On the other hand, if $i \notin I$, then the decoder chooses $y_{j,k}$ when $|g_j|^2 \geq |g_i|^2$ or $y_{i,k'}$ when $|g_j|^2 < |g_i|^2$ (i.e., the observation received through the better channel). The preceding discussion focused on the case where frame k is not the last frame of the super-frame. If frame k is indeed last, then the decoder always chooses $y_{j,k}$ over $y_{i,k'}$.

We define $\mathbf{s}_{j,k}^I$, where $j \in I$, as the vector (of dimension $ml \times 1$) of contributions of symbol $x_{j,k}$ to the destination observations picked by the type-I decoder. Clearly,

$$\mathbf{s}^I = \sum_{k=1}^{l} \sum_{j \in I} \mathbf{s}_{j,k}^I.$$

Taking into account the independence of the transmitted symbols (i.e., $x_{j,k}$), we have

$$\Sigma_{\mathbf{s}^I} = \sum_{k=1}^{l} \sum_{j \in I} \mathrm{E}\{\mathbf{s}_{j,k}^I (\mathbf{s}_{j,k}^I)^H\}. \tag{81}$$

In order to illuminate some of the properties of $\mathbf{s}_{j,k}^I$, assume that we sort the chosen observations in chronological order. From the description given, it is apparent that, associated with each chosen observation (i.e., $y_{j,k}$ or $y_{i,k'}$) there is *one* symbol $x_{j,k}$ (with $j \in I$) which has contributions only from this observation forward. This means that if we define S^I as

$$S^I \triangleq [\mathbf{s}_{j_1,k_1}^I \mathbf{s}_{j_2,k_2}^I \ldots \mathbf{s}_{j_{ml},k_{ml}}^I]_{ml \times ml},$$

where $j_p \in I$ and $k_p \in \{1, \ldots, l\}$ are chosen such that the first non-zero elements of $\mathbf{s}_{j_p,k_p}^I, p = 1, \ldots, ml$ are sorted in chronological order, then S^I will be lower-triangular and consequently $(S^I)^H$ will be upper-triangular. Furthermore, based on the choice between $y_{j,k}$ or $y_{i,k'}$ (corresponding to $x_{j,k}$), the first non-zero element of $\mathbf{s}^I_{j_p,k_p}$ (i.e., the p^{th} diagonal element of S^I) will be $g_j x_{j,k}$ or $g_i b_i h_{ij} x_{j,k}$, respectively. Next, we define $\psi^I_{j,k}$ as the signature of $x_{j,k}$, i.e.,

$$\psi^I_{j,k} \triangleq \frac{1}{x_{j,k}} \mathbf{s}^I_{j,k} \quad j \in I,$$

and Ψ^I as

$$\Psi^I \triangleq [\psi^I_{j_1,k_1} \psi^I_{j_2,k_2} \dots \psi^I_{j_{ml},k_{ml}}]_{ml \times ml}.$$

It follows then, that Ψ^I is also lower-triangular with the p^{th} diagonal element being equal to g_j or $g_i b_i h_{ij}$. Using these definitions, (81) can be written as

$$\Sigma_{\mathbf{s}^I} = E \sum_{k=1}^{l} \sum_{j \in I} \psi^I_{j,k} (\psi^I_{j,k})^H. \tag{82}$$

The significance of Ψ^I can now be seen from the fact that (82) can be written as

$$\Sigma_{\mathbf{s}^I} = E \Psi^I (\Psi^I)^H.$$

Now, as the determinant of triangular matrices is simply the product of their diagonal elements, from (80) we conclude that

$$P_{pe^I|v_j,u_{ji}} \dot{\leq} \rho^{-m(N-1)L + \sum_{j \in I} \left[(m-1)L v_j + \sum_{i \notin I} \left(\min\{v_j, u_{ji}+v_i\}(L-1)+v_j\right)\right]}, \qquad v_j \geq 0, u_{ji} \geq 0.$$

It is obvious that for large L's, the previous inequality can be rewritten as

$$P_{pe^I|v_j,u_{ji}} \dot{\leq} \rho^{-\left[-\sum_{j \in I}\left((m-1)v_j + \sum_{i \notin I} \min\{v_j, u_{ji}+v_i\}\right) + m(N-1)\right] L}, \qquad v_j \geq 0, u_{ji} \geq 0. \tag{83}$$

At rate $R = r\log(\rho)$ and codeword length l, and when the symbols of the sources that are not in I are set to zero, there are a total of $\rho^{m(N-1)Lr}$ unique codewords. Thus,

$$P_{e^I|v_j,u_{ji}} \dot{\leq} \rho^{-\left[-\sum_{j \in I}\left((m-1)v_j + \sum_{i \notin I} \min\{v_j, u_{ji}+v_i\}\right) + m(N-1)(1-r)\right] L}, \qquad v_j \geq 0, u_{ji} \geq 0. \tag{84}$$

This conditional type-I error probability leads to

$$P_{e^I,O_c} \dot{\leq} \rho^{-d_{e^I}(r)},$$

where

$$\begin{aligned} d_{e^I}(r) \triangleq \min_{O_c^+} \sum_j \left(v_j + \sum_i u_{ji}\right) + \cdots \\ \left[-\sum_{j \in I}\left((m-1)v_j + \sum_{i \notin I} \min\{v_j, u_{ji}+v_i\}\right) + m(N-1)(1-r)\right] L \end{aligned} \tag{85}$$

Examining (85), we realize that for (75) to be met, O^+ should be defined as the set of all real $\frac{N(N+1)}{2}$-tuples with nonnegative elements that satisfy the following condition for at least one nonempty $I \subseteq \{1, \dots, N\}$:

$$\sum_{j \in I}\left((m-1)v_j + \sum_{i \notin I} \min\{v_j, v_i + u_{ji}\}\right) \geq m(N-1)(1-r) \tag{86}$$

This way, by choosing large enough l, $d_{e^I}(r)$ can be made arbitrary large and thus (75) is always met. From (86), it follows that

$$\sum_{j\in I}\left((m-1)v_j + \sum_{i\notin I}\min\{v_j, v_i + \max_{j\neq i}\{u_{ji}\}\}\right) \geq m(N-1)(1-r). \tag{87}$$

Substituting $\min\{v_j, v_i + \max_{j\neq i}\{u_{ji}\}\}$ in this expression by v_j gives

$$\sum_{j\in I} v_j \geq m(1-r). \tag{88}$$

On the other hand, replacing $\min\{v_j, v_i + \max_{j\neq i}\{u_{ji}\}\}$ in (87) by $v_i + \max_{j\neq i}\{u_{ji}\}$ results in

$$(m-1)\sum_{j\in I} v_j + m\sum_{i\notin I}(v_i + \max_{j\neq i}\{u_{ji}\}) \geq m(N-1)(1-r). \tag{89}$$

Under the constraints given by (88) and (89), it is easy to see that

$$\inf_{O^+}\left(\sum_{j\in I} v_j + \sum_{i\notin I}(v_i + \max_{j\neq i}\{u_{ji}\})\right) \geq N(1-r). \tag{90}$$

Now, from (90) and (48), it follows that

$$d_o(r) \geq N(1-r).$$

Again, according to (76), $d_o(r)$ provides a lower-bound on the diversity gain achieved by the protocol. Thus the protocol achieves the diversity gain given by (22) and the proof is complete.

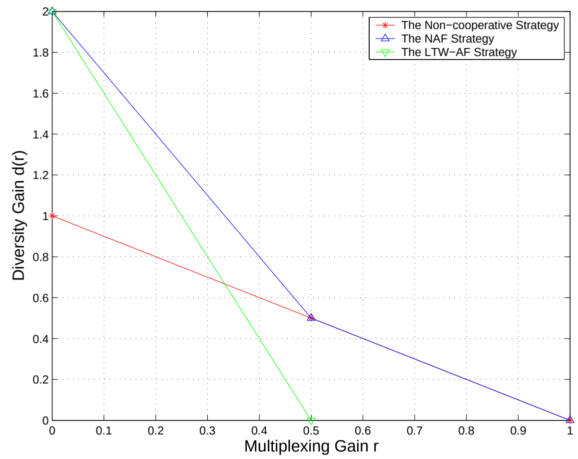

Figure 1: Optimal diversity-multiplexing tradeoff for a single AF relay.

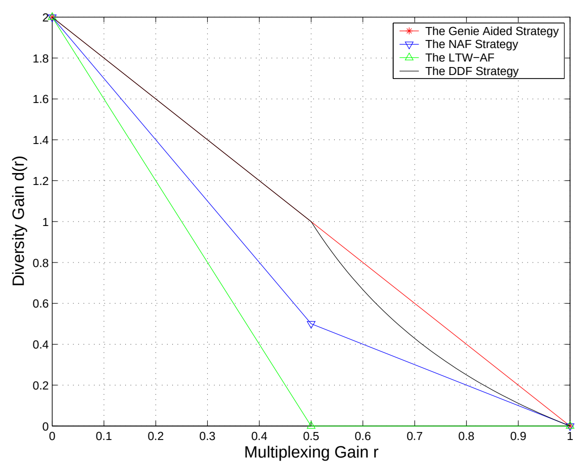

Figure 2: Diversity-multiplexing tradeoff for the DDF protocol with one relay.

References

[1] A. Sendonaris, E. Erkip and B. Aazhang, "User Cooperation Diversity. Part I. System Description," *IEEE Transactions on On Communications*, page(s): 1927- 1938 Volume: 51, Issue: 11, Nov. 2003

[2] A. Sendonaris, E. Erkip and B. Aazhang, "User Cooperation Diversity. Part II. Implementation Aspects and Performance Analysis," *IEEE Transactions on On Communications*, page(s): 1939- 1948 Volume: 51, Issue: 11, Nov. 2003

[3] J. N. Laneman, D. N. C. Tse and G. W. Wornell, "Cooperative Diversity in Wireless Networks: Efficient Protocols and Outage Behavior," *IEEE Transactions on Information Theory*, 2002, submitted.

[4] J. N. Laneman and G. W. Wornell, " Distributed Space-Time-Coded Protocols for Exploiting Cooperative Diversity in Wireless Networks," *IEEE Transactions on Information Theory*, Volume: 49 , Issue: 10 , Oct. 2003 Pages:2415 - 2425

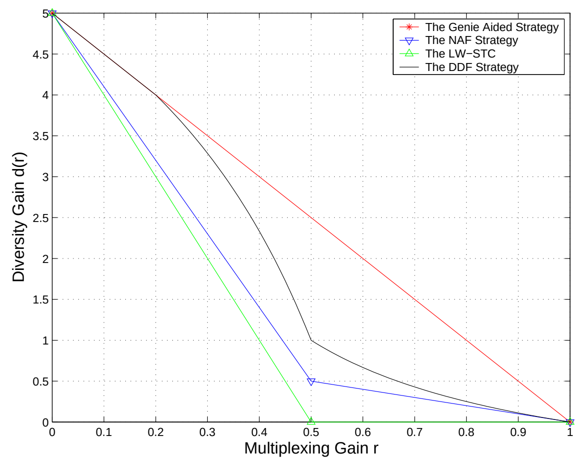

Figure 3: Diversity-multiplexing tradeoff for the NAF, DDF, Laneman-Wornell space-time-coded (LW-STC), and non-cooperative protocols with 4 relays.

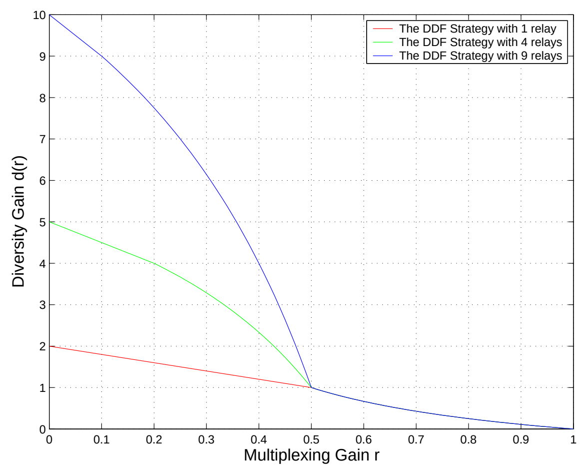

Figure 4: Diversity-multiplexing tradeoff for the DDF protocol with different number of relays.

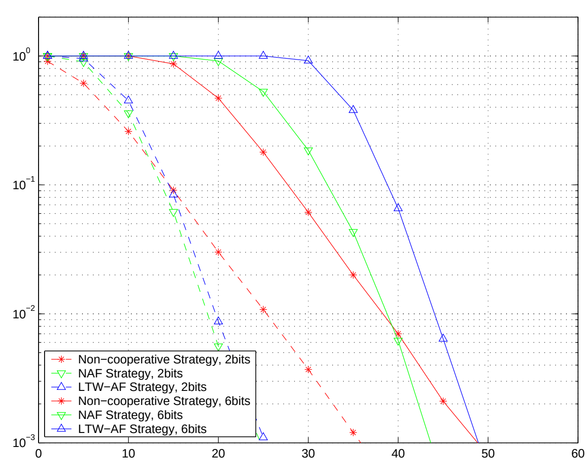

Figure 5: Comparison of the outage probability for the NAF relay, LTW-AF, and non-cooperative 1×1 protocols ($N = 2$).

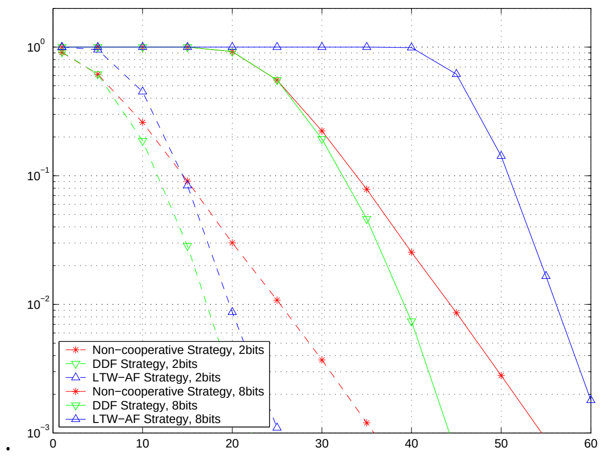

Figure 6: Comparison of the outage probability for the DDF relay, LTW-AF and non-cooperative 1×1 protocols ($N = 2$).

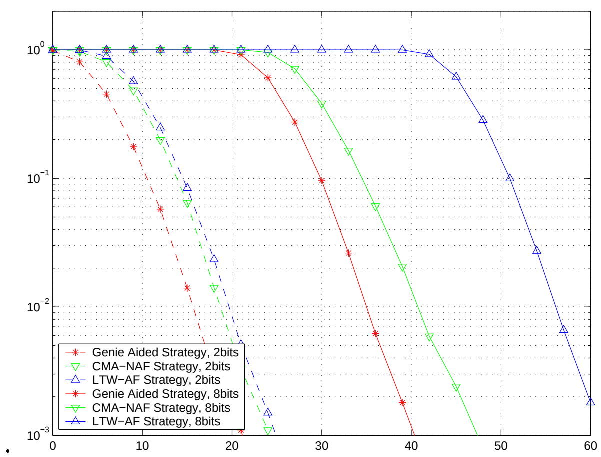

Figure 7: Comparison of the outage probability for the CMA-NAF, LTW-AF and genie-aided 2×1 protocols ($N = 2$).

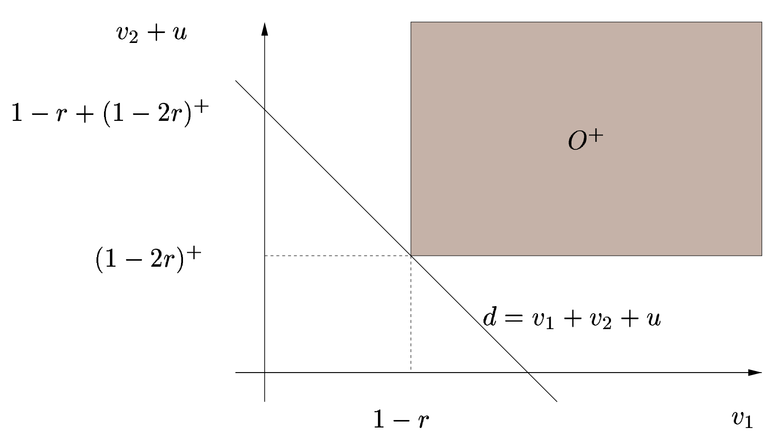

Figure 8: Outage Region for the NAF protocol with a single relay.

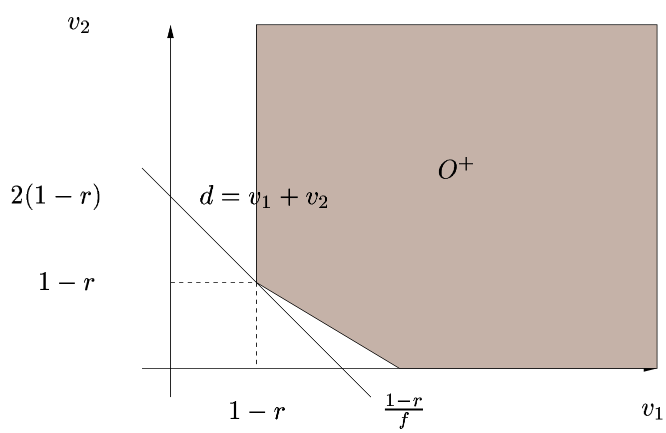

Figure 9: Outage Region for the DDF protocol with a single relay ($f \leq 0.5$).

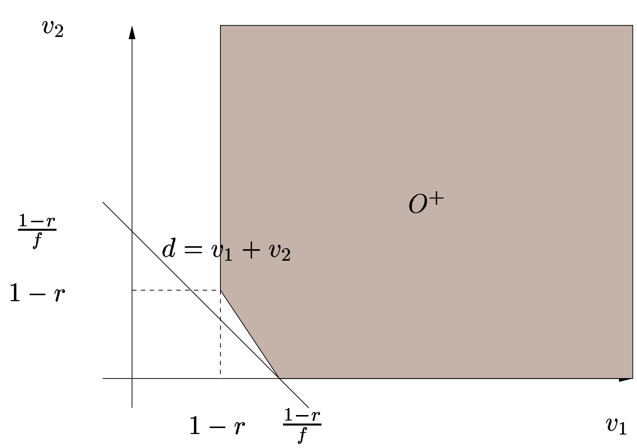

Figure 10: Outage Region for the DDF protocol with a single relay ($f > 0.5$).

[5] G. Kramer, M. Gastpar P. Gupta, "Cooperative Strategies and Capacity Theorems for Relay Networks," *submitted to the IEEE Transactions on Information Theory*, Feb. 2004

[6] A. Stefanov and E. Erkip, "Cooperative Coding for Wireless Networks," *IEEE Transactions on Communications*, accepted for publication

[7] M. Janani, A. Hedayat, T. Hunter, and A. Nosratinia, ""Coded Cooperation in Wireless Communications: Space-time transmission and iterative decoding," *IEEE Transactions on Signal Processing*, Volume: 52 , Issue: 2 , Feb. 2004 Pages:362-371

[8] L. Zheng and D. N. C. Tse. Diversity and Multiplexing: A Fundamental Tradeoff in Multiple Antenna Channels. *IEEE Trans. Info. Theory*, 49:1073 –1096, May 2003.

[9] D. N. C. Tse, P. Viswanath and L. Zheng Diversity-Multiplexing Tradeoff in Multiple Access Channels. *IEEE Trans. on Info. Theory*, 2003, submitted.

[10] K. Azarian Yazdi, H. El Gamal, and P. Schniter, "On the Design of Cooperative Transmission Schemes," *Proc. Allerton Conf. on Communication, Control, and Computing, (Monticello, IL)*, Oct. 2003.

[11] K. Azarian, H. El Gamal, and P. Schniter, "On the Achievable Diversity-vs-Multiplexing Tradeoff in Cooperative Channels," *Proc. Conference on Information Sciences and Systems, (Princeton, NJ)*, Mar. 2004.

[12] H. El Gamal, "The Diversity-Multiplexing Tradeoff in Half-Duplex Cooperative Channels: Achievable Curves and Optimal Strategies," *LIDS Colloquia, EECS Dept., MIT*, May 2004.

[13] T. M. Cover and A. A.El Gamal, "Capacity Theorems for the Relay Channel," *IEEE Transactions on Information Theory*, Volume: 25, pp. 572-584, Sept. 1979.