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

Kambiz Azarian, Hesham El Gamal, *Senior Member, IEEE*, and Philip Schniter, *Member, IEEE*

***Abstract*—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 (CMA) (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 Wornell 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 CMA 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 suboptimality of previously proposed protocols stems from their use of orthogonal subspaces rather than the half-duplex constraint.**

***Index Terms*—Cooperative diversity, diversity–multiplexing tradeoff, dynamic decode and forward (DDF), half-duplex node, multiple-access channel, nonorthogonal amplify and forward (NAF), relay channel.**

I. Introduction

RECENTLY, there has been a growing interest in the design and analysis of wireless cooperative transmission protocols (e.g., [1]–[17]). These works consider several interesting scenarios (e.g., fading versus additive white Gaussian noise (AWGN) channels, ergodic versus quasi-static channels, and full-duplex versus half-duplex transmission) and devise appropriate transmission techniques and analysis tools, based on the settings. Here, we focus on the delay-limited coherent channel and adopt 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. In this setup, the basic idea is to leverage the antennas available at the other nodes in the network as a source of *virtual* spatial diversity. 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) codebook. 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 followup works have focused on developing practical coding schemes that attempt to exploit the promised information-theoretic gains (e.g., [5], [6]).

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 multiple-input multiple-output (MIMO) channels in [18] and later used by Tse, Viswanath, and Zheng to study the (noncooperative) multiple-access channel in [19].

In the following, we summarize the main results of this paper, some of which were initially reported in [20]–[24].

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

Manuscript received July 16, 2004; revised June 30, 2005. This work was supported in part by the National Science Foundation under Grant ITR-0219892. The material in this paper was presented in part at the 38th Annual Conference on Information Sciences and Systems, Princeton, NJ, March 2004, the 42nd Allerton Conference on Communication, Control, and Computing, Monticello, IL, October 2004, and the IEEE Information Theory Workshop, San Antonio, TX, October 2004.

The authors are with the Department of Electrical and Computer Engineering, The Ohio State University, Columbus, OH 43210-1272 USA (e-mail: azariany@ece.osu.edu; helgamal@ece.osu.edu; schniter@ece.osu.edu).

Communicated by R. R. Müller, Associate Editor for Communications.
Digital Object Identifier 10.1109/TIT.2005.858920

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

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 (LW-STC) [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 protocol where an *artificial* intersymbol-interference (ISI) channel is created. We prove the optimality (in the sense of diversity–multiplexing tradeoff) of this protocol by showing that, for all multiplexing gains (i.e., $0 \leq r \leq 1$), 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 suboptimality of the schemes proposed in [3] 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 CMA channels.

Before proceeding further, a brief remark regarding two independent prior works [8], [9] is in order. In [7], [8], Nabar, Bölcskei, and Kneubuhler considered the half-duplex single-relay channel, under *almost* the same assumptions as in [3] (i.e., the only difference is that, for diversity analysis, the relay-destination channel was assumed to be nonfading) and proposed a set of AF and DF protocols. In one of their AF protocols (NBK-AF), Nabar *et al.* allowed the source to continue transmission over the whole duration of the codeword, while the relay listened to the source for the first half of the codeword and relayed the received signal over the second half. This makes the NBK-AF protocol identical to the NAF protocol [7], [21]. Here, we characterize the diversity–multiplexing tradeoff achieved by this protocol while relaxing the assumption of nonfading relay-destination channel which is invoked in the analysis reported in [8]. Using this analysis, we establish the optimality of this scheme within the class of linear AF protocols. Furthermore, we generalize the NAF protocol to the case of arbitrary number of relays and characterize its achieved tradeoff curve. In [9], Prasad and Varanasi derived upper bounds on the diversity–multiplexing tradeoffs achieved by the DF protocols proposed in [8]. In the sequel, we establish the gain offered by the proposed DDF protocol by comparing its diversity–multiplexing tradeoff with the upper bounds in [9]. Finally, we emphasize that, except for the single-relay NAF protocol, all the other protocols proposed in this paper are novel.

In this paper, we use $(x)^+$ to mean $\max\{x,0\}$, $(x)^-$ to mean $\min\{x,0\}$, and $\lceil x \rceil$ to mean nearest integer to x toward 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 nonnegative N-tuples. We denote the complement of set $O \subseteq \mathbb{R}^N$, in $\mathbb{R}^N$, by O^c, while O^+ means $O \cap \mathbb{R}^{N+}$. I_N denotes the $N \times N$ identity matrix, $\Sigma_{\boldsymbol{x}}$ denotes the autocovariance matrix of vector $\boldsymbol{x}$, and $\log(\cdot)$ denotes the base-2 logarithm.

The rest of the paper is organized as follows. In Section II, 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 III where we describe the NAF and DDF protocols and derive their tradeoff curves. In Section IV, we extend the DDF protocol to the cooperative broadcast channel. Section V is devoted to the CMA 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 VI, we present numerical results that show the signal-to-noise (SNR) gains offered by the proposed schemes in certain representative scenarios. Finally, we offer some concluding remarks in Section VII. To enhance the flow of the paper, we collect all the proofs in the Appendix.

II. 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, i.e., 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' s 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] cannot, 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 codebooks, where a codeword spans the entire coherence interval of the channel. Furthermore, we assume asymptotically large code lengthes. This implies that the diversity–multiplexing tradeoffs derived in this paper serve as upper bounds for the performance of the proposed protocols with finite code lengths. 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 [18]. 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 $k=0,\ldots,\min\{M,N\}$, 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.,

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

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

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

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

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

provided that O^+ is not empty. In other words, the exponential order of P_O only depends on O^+. This is due to the fact that the probability of any set, consisting of N-tuples $(v_1,\ldots,v_N)$ with at least one negative element, decreases exponentially with SNR and therefore can be neglected compared to P_{O^+} which decreases polynomially with SNR.

6. Consider a coherent linear Gaussian channel, i.e.,

$$\boldsymbol{y} = \boldsymbol{s} + \boldsymbol{n}$$

where $\boldsymbol{s} \in \mathbb{C}^N$ and $\boldsymbol{n} \in \mathbb{C}^N$ denote the signal and noise components of the observed vector, respectively. For this channel, 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

$$P_{\text{PE}} \leq \det\left(I_N + \frac{1}{2}\Sigma_{\boldsymbol{s}}\Sigma_{\boldsymbol{n}}^{-1}\right)^{-1}. \tag{7}$$

7. The following lemma will be used in characterizing the diversity–multiplexing tradeoff of the DDF protocols.

Lemma 1: Consider a coherent linear Gaussian channel of data rate R and codeword length l. The error probability of the ML decoder which utilizes a fraction of the codeword such that the mutual information between the received and transmitted signals exceeds lR, averaged over the ensemble of random Gaussian codes, can be made arbitrarily small provided that the codeword length l is sufficiently large.

Proof: Please refer to the Appendix.

III. 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 subproblems have been studied (for example, see [25]). In this work, we focus on two important classes of relay protocols. The first is the class of AF protocols, where a relaying node can only process the observed signal linearly before retransmitting it. The second is the class of DF protocols, where the relays are allowed to decode and re-encode the message using (a possibly different) codebook. Here we emphasize that, *a priori*, it is not clear which class (i.e., AF or DF) offers a better performance (e.g., [3]).

A. Amplify and Forward (AF) 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 [7], [21], 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:

$$\boldsymbol{y} = \begin{bmatrix} g_1 A_1 & 0 \\ g_2 h B A_1 & g_1 A_2 \end{bmatrix} \boldsymbol{x} + \begin{bmatrix} 0 \\ g_2 B \end{bmatrix} \boldsymbol{w} + \boldsymbol{v}. \quad (8)$$

In (8), $\boldsymbol{y} \in \mathbb{C}^l$ represents the vector of observations at the destination, $\boldsymbol{x} \in \mathbb{C}^l$ the vector of source symbols, $\boldsymbol{w} \in \mathbb{C}^{l'}$ the vector of noise samples (of variance σ_w^2) observed by the relay, and $\boldsymbol{v} \in \mathbb{C}^l$ 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 $B = bI_{l'}$ (with $b \le \sqrt{E/(|h|^2 E + \sigma_w^2)}$ denoting the relay repetition gain), we obtain Laneman–Tse–Wornell AF (LTW-AF) protocol [3]. Finally, we note that when the source symbols are independent, the average energy constraint translates to

$$|h|^2 E \sum_{i=1}^{l'} |b_{ji}|^2 |a_i|^2 + \sigma_w^2 \sum_{i=1}^{l'} |b_{ji}|^2 \le E, \qquad j = 1, \ldots, l - l' \quad (9)$$

where $B = [b_{ji}]$ and $A_1 = \operatorname{diag}(a_1, \ldots, a_{l'})$.

Theorem 2: The optimal diversity gain for the cooperative relay scenario with a single AF relay is upper-bounded by

$$d^*(r) \le (1 - r) + (1 - 2r)^+. \quad (10)$$

Proof: Please refer to the Appendix.

The upper bound on $d^*(r)$, as given by (10), is shown in Fig. 1. Having Theorem 2 at hand, it now suffices to identify an AF protocol that achieves this upper bound in order to establish its optimality. Toward this end, we observe that, in the proof of Theorem 2, 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

$$A_1 = I_{l/2} \quad A_2 = I_{l/2} \quad B = bI_{l/2}, \qquad \text{for } b \le \sqrt{\frac{E}{|h|^2 E + \sigma_w^2}} \quad (11)$$

inspire the NAF protocol [7], [21]. 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,

$$\begin{aligned} y_{1,k} &= g_1 x_{1,k} + v_{1,k} \\ y_{2,k} &= g_1 x_{2,k} + g_2 b(h x_{1,k} + w_{1,k}) + v_{2,k} \end{aligned}$$

where the repetition gain b must satisfy (11). Note that, in order to decode the message, the destination needs to know the relay repetition gain b, the source-relay channel gain h, the source-destination channel gain g_1, and the relay-destination channel gain g_2. Now, we are ready to establish the optimality of the NAF protocol with respect to the diversity–multiplexing tradeoff.

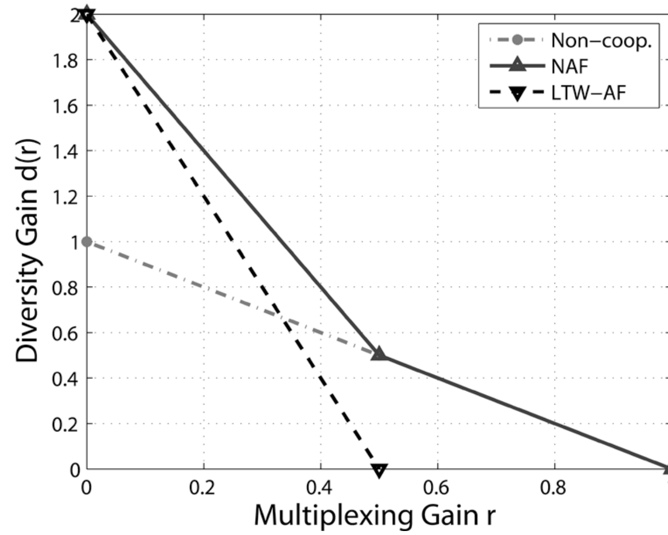

Fig. 1. Optimal diversity–multiplexing tradeoff for a single-relay AF protocol.

Theorem 3: The NAF protocol achieves the optimal diversity–multiplexing tradeoff for the AF single-relay scenario, which is

$$d^*(r) = (1 - r) + (1 - 2r)^+. \quad (12)$$

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 NAF relay protocol is identical to that of the noncooperative

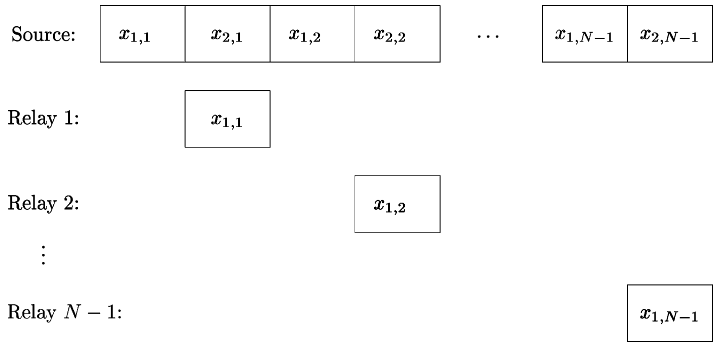

Fig. 2. The super-frame in the NAF protocol with $N-1$ relays.

protocol. This is due to the fact that the *AF cooperative* link provided by the relay cannot 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 3, 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 SNR.

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 (refer to Fig. 2). 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 $i \in \{1, \ldots, N-1\}$ 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 jth symbol interval of the kth cooperation frame. Note that there is nothing to be gained by having more than one relay transmitting the same symbol simultaneously. Also, similar to the single-relay NAF scenario, the destination needs to know all relay repetition gains $\{b_i\}_{i=2}^N$ as well as all channel gains $\{g_i\}_{i=1}^N$ and $\{h_i\}_{i=2}^N$. The following theorem characterizes the diversity–multiplexing tradeoff achieved by this protocol.

Theorem 4: 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 3, and hence, is omitted for brevity.

It is interesting to note that the generalized NAF protocol uniformly dominates the LW-STC. This can be attributed to the fact that in the generalized NAF protocol, in contrast to the LW-STC protocol, the source transmits over the whole duration of the codeword. The generalized NAF protocol offers the additional advantage of low complexity since it does not require decoding/encoding at the relays.

B. Decode and Forward (DF) Protocols

In this class of protocols, we allow for the possibility of decoding/encoding at the different relays. In [3], Laneman, Tse, and Wornell presented a particular variant of DF protocols (LTW-DF) where the source transmits in the first half of the codeword. Based on its received signal in this interval, the relay attempts to decode the message. It then re-encodes and transmits the encoded stream in the second half of the codeword. In [4], Laneman and Wornell derived the diversity–multiplexing tradeoff achieved by this scheme (i.e., $d(r) = 2(1-2r)$), which is depicted in Fig. 3. Here, we propose a DDF protocol and characterize its tradeoff curve. This characterization reveals the uniform dominance of this protocol over all known *full-diversity* (i.e., $d(0) = 2$) protocols proposed for the half-duplex single-relay channel and furthermore establishes its optimality, over a certain range of multiplexing gains (i.e., $0 \leq r \leq 1/2$). 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 data at a rate of R BPCU during every symbol interval in the codeword. The relay, on the other hand, listens to the source until the mutual information between its received signal and source signal exceeds lR. It then decodes and re-encodes the message using an *independent* Gaussian codebook and transmits it during the rest of the codeword. The

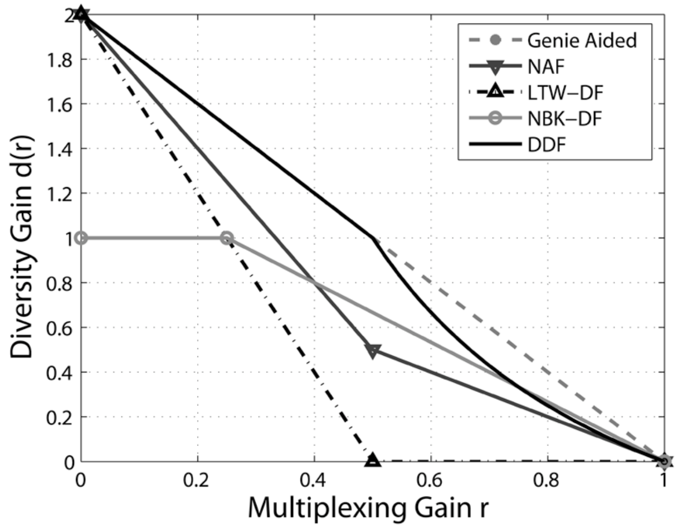

Fig. 3. Diversity–multiplexing tradeoff for the DDF protocol with one relay.

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

$$\begin{aligned} y_k &= g_1 x_k + v_k, && \text{for } l' \geq k \geq 1 \\ y_k &= g_1 x_k + g_2 \tilde{x}_k + v_k, && \text{for } l \geq k > l'. \end{aligned}$$

From the protocol description, it is clear that 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 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, guarantees that when $l' < l$, the relay average probability of error with a Gaussian code ensemble is arbitrarily small.[2] 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. Here, we observe that, in contrast to the NAF protocol, the destination does not need to know the source-relay channel gain. It does, however, need to know the relay waiting time l', along, with the source-destination and relay-destination channel gains. The following theorem describes the diversity–multiplexing tradeoff achievable with this cooperation protocol.

Theorem 5: 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.

[2]This point will be established rigorously in the proofs of Theorem 5 and Lemma 1.

The diversity–multiplexing tradeoff of (14) is shown in Fig. 3. 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 cannot 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 not. Note also 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, in contrast, the relay is limited to repeating the noisy Gaussian symbols it receives from the source. Fig. 3 also compares the DDF protocol with the DF protocol proposed in [8], which we refer to as NBK-DF. In this comparison, we utilize the upper bound derived by Prasad and Varanasi on the diversity–multiplexing tradeoff of the NBK-DF, which was reported in [9]. One can see from Fig. 3 that the NBK-DF protocol does not achieve any diversity gain greater than one. This can be attributed to the fact that in this protocol, the message is split up into two parts, out of which, only one is retransmitted by the relay. Fig. 3 also shows that for multiplexing gains close to one, the NBK-DF upper bound outperforms the DDF protocol. Therefore, in this range, the comparison between the two protocols depends on the tightness of the NBK-DF upper bound which was not discussed in [9].

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 single-relay case. Specifically, the source transmits during the whole codeword while each relay listens until the mutual information between its received signal and the signals transmitted by the source and other relays exceeds lR. It is assumed that every relay knows the codebooks used by the source and other relays. Once a relay decodes the message, it uses an independent codebook 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's 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 transmission only at the beginning of a segment. In between the segments, every relay is allowedto 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

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

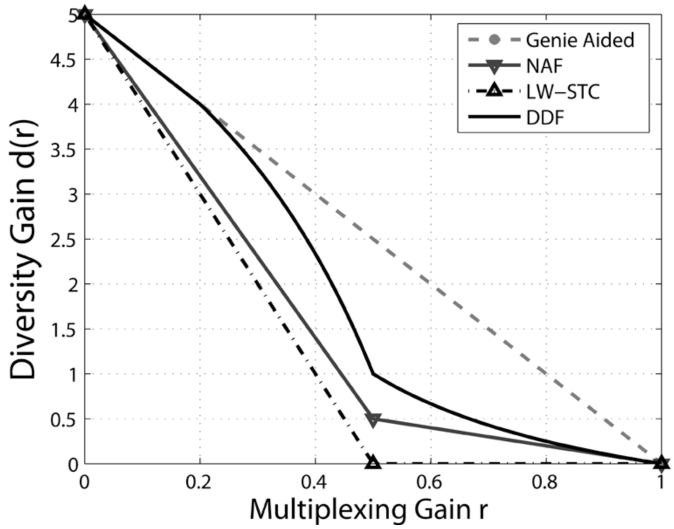

Fig. 4. Diversity–multiplexing tradeoff for the NAF, DDF, LW-STC, and genie-aided protocols with four relays.

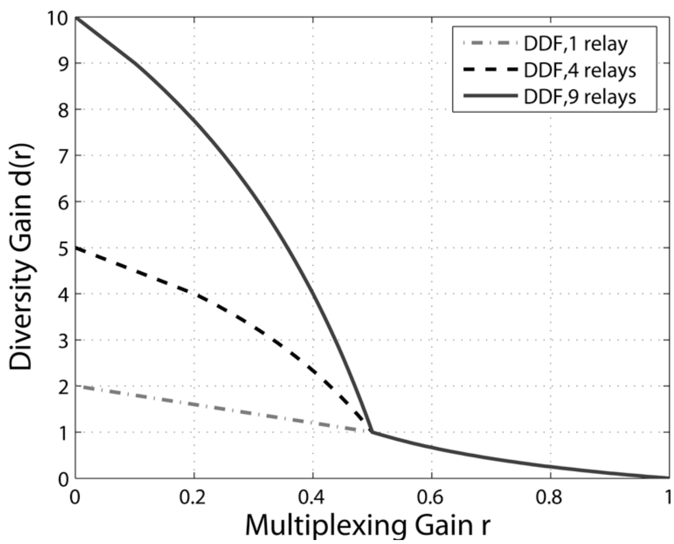

Fig. 5. Diversity–multiplexing tradeoff for the DDF protocol with different number of relays.

length. Therefore, in characterizing the diversity–multiplexing tradeoff achieved by this protocol, we ignore the losses associated with the beacons and the quantization of the starting times for the different relays.

Theorem 6: The diversity–multiplexing tradeoff achieved by the DDF protocol with $N-1$ relays is characterized by

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

Proof: Please refer to the Appendix.

The diversity-multiplexing tradeoff (15) is shown in Figs. 4 and 5 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 suboptimality of the DDF strategy.

IV. The Half-Duplex Cooperative Broadcast (CB) Channel

We now consider the CB scenario, where a single source broadcasts to N destinations. This model corresponds to a generalized version of a single-cell downlink where the destinations are allowed to cooperate through helping one another in receiving their messages. Similar to the relay channel scenario, we adopt the quasi-static flat Raleigh-fading model for all the source-destination and inter-destination channels. We assume that the message intended for destination $j \in \{1, \ldots, N\}$ consists of two parts. A common part of rate $R_c = r_c \log(\rho)$ BPCU, which is intended for all of the destinations and an individual part of rate $R_j = r_j \log(\rho)$ BPCU, which is specific to the jth destination. The total rate is then $R = R_c + \sum_{j=1}^{N} R_j$ and the multiplexing gain tuple is given by $\boldsymbol{r} = (r_c, r_1, \ldots, r_N)$. We define the overall diversity gain d based on the performance of the worst 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 $r_c = 0$, i.e., if there is no common message, then the techniques developed for the relay channel can be *exported* to this setting through a proportional time sharing strategy. With this assumption, all of the properties of the NAF and DDF protocols, established for the relay 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 R, the highest probability of error corresponds to the case where all destinations are required to decode all the messages. This translates to the following condition (that applies to any cooperation scheme):

$$d(r_c, r_1, r_2, \ldots, r_N) \geq d(r_c + r_1 + \cdots + r_N, 0, 0, \ldots, 0). \quad (16)$$

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

$$\boldsymbol{r} = (r_c, 0, 0, \ldots, 0), \qquad 0 \leq r_c \leq 1. \quad (17)$$

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 noncooperative protocol. Any other AF strategy will require some of the nodes to retransmit, 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 destination can act as a relay for the other destinations, based on its instantaneous channel gain. Specifically, the source transmits during the whole codeword while each destination listens until the mutual information between its received signal and the signals transmitted by the source and other destinations exceeds lR. Once a

[4]Clearly, this definition does not allow for different quality of service (QoS) constraints.

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

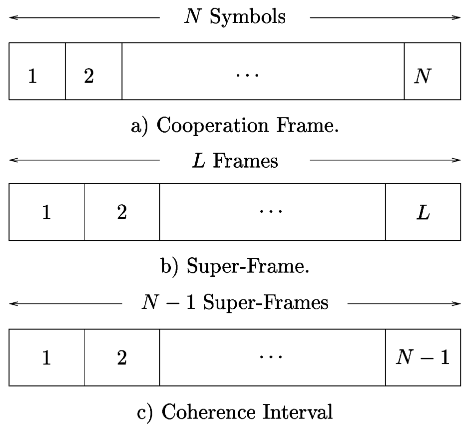

Fig. 6. The cooperation frame, super-frame, and coherence interval in the CMA-NAF protocol with N sources.

destination decodes the message, it uses an independent codebook to re-encode the message, which it then transmits for the rest of the codeword. Similar to the relay channel, it is assumed that every destination knows the codebooks used by the source and other destinations. Also, the protocol must include a mechanism that keeps every destination informed of the retransmission starting times of all the other destinations. Again, in deriving the following result, we ignore the associated cost of this mechanism, relying on the asymptotic assumptions.

Theorem 7: The diversity–multiplexing tradeoff achieved by the CB-DDF protocol with N destinations is given by

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

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 message does not entail a price in terms of the achievable tradeoff.

V. The Half-Duplex Cooperative Multiple-Access (CMA) Channel

In this section, we consider the CMA scenario, where N sources transmit their independent messages to a common destination. This model corresponds to a generalized version of a single-cell uplink where the sources are allowed to cooperate. The CMA scenario was previously considered in [1], [3] (and references therein). Again, the same quasi-static flat Rayleigh-fading model is invoked for all the channels. 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. Toward this end, each of the N sources transmits once per cooperation frame, where a cooperation frame is defined as N consecutive symbol-intervals (refer to part a) of Fig. 6). 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, in addition to sending its own symbol, *helps* another source by repeating the (noisy) signal it last received from it. Without loss of generality, we set the jth 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 jth source and the kth 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 jth 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 ith-source-to-jth-source channel gain, and $w_{j,k}$ to denote the noise observed by the jth source during its kth-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 jth-source-to-destination channel gain, and $v_{j,k}$ to denote the noise observed by the destination during the jth symbol interval of the kth frame, the signals observed at the destination would be

$$y_{j,k} = g_j t_{j,k} + v_{j,k}.$$

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

$$E\{|t_{j,k}|^2\} \leq E. \quad (19)$$

Let us now define L consecutive cooperation frames as a super-frame (refer to part b) of Fig. 6). 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.

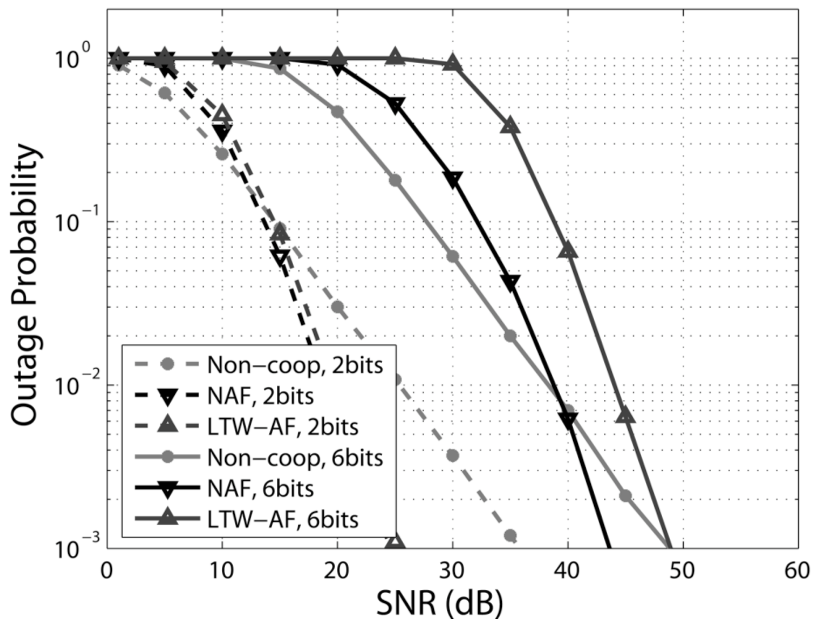

Fig. 7. Comparison of the outage probability for the NAF relay, LTW-AF, and noncooperative 1×1 protocols ($N = 2$).

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 jth 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 L is chosen such that a coherence-interval consists of $N-1$ consecutive super-frames (refer to part c) of Fig. 6). 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 CMA channel requires some care. Note that, using (3), the pair (r_j, d_j) can be defined for communication between the jth 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.,

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

With these definitions, Theorem 8 establishes the optimality of the CMA-NAF in the symmetric scenario with N sources.

Theorem 8: 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 8 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 diversity–multiplexing tradeoff, in the symmetric CMA channel. One can now attribute the suboptimality of the CMA schemes reported in [3], [4] 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 2 and 8, one can see the fundamental difference between the half duplex CMA and relay channels.

VI. 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 Figs. 7–9, we compare the proposed protocols with the noncooperative (direct transmission) and

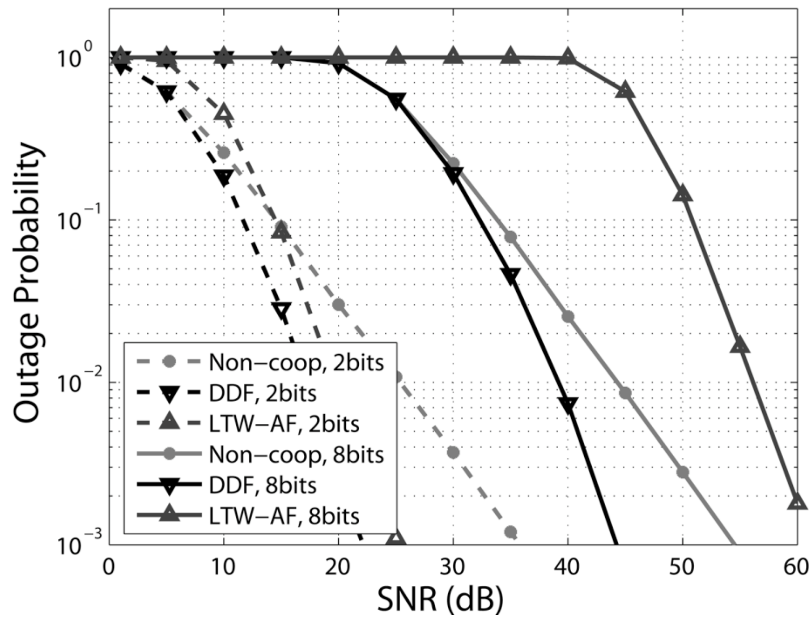

Fig. 8. Comparison of the outage probability for the DDF relay, LTW-AF, and noncooperative 1×1 protocols ($N = 2$).

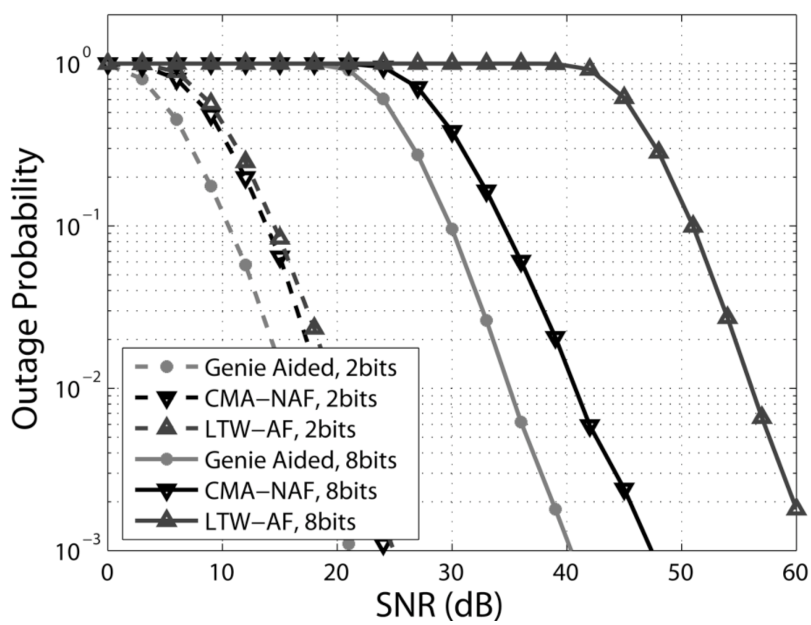

Fig. 9. Comparison of the outage probability for the CMA-NAF, LTW-AF, and genie-aided 2×1 protocols ($N = 2$).

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[6] is E. 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. 7 shows the performance gain offered by the NAF relay protocol over both the noncooperative protocol and the LTW-AF protocol at high SNRs and two different data rates. The same comparison is repeated in Fig. 8 with the DDF protocol where, as expected, the gains are shown to be larger. The CMA channel is considered in Fig. 9, where the

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

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 data 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 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.

VII. Conclusion

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 full diversity 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 was 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 (as observed in [7], [21] for the relay channel case). 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. Generalizations of the proposed schemes to multiple-antenna nodes, cooperative automatic retransmission request (ARQ) channels [23], 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 ML decoding is an important venue to pursue.

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

Appendix

In this appendix, we collect all the proofs.

A. *Proof of Lemma 1*

For simplicity, we consider the real AWGN channel (the complex channel is a straightforward extension), i.e.,

$$y_k = hx_k + n_k, \qquad \text{for } l \geq k \geq 1$$

where y_k is the received signal over the kth symbol interval, h is the channel gain, and n_k is the independent and identically distributed (i.i.d.) Gaussian noise sample during symbol interval k. In our analysis, we derive the average probability of error, where averaging is invoked with respected to the ensemble of Gaussian codes, conditioned on a particular h, which is assumed to be known at the decoder. We denote this error probability by $P_{E|h}$. Toward this end, we divide a codeword into N segments, each consisting of L symbol intervals, i.e., $l = NL$. In our analysis, we consider the asymptotic scenario where N and L grow to infinity. The ML decoder waits for N' segments before it starts decoding, where N' is given by

$$N' = \left\lfloor \frac{NR}{I(x_k; y_k)} \right\rfloor + 1. \tag{23}$$

In (23), $I(x_k; y_k)$ denotes the mutual information between x_k and y_k. Notice that assuming the data rate to be less than the channel capacity guarantees $N' \leq N$. Also, observe that the fraction of codeword that the decoder has to wait before decoding is given by

$$\lim_{N\to\infty} \frac{N'}{N} = \frac{R}{I(x_k; y_k)}.$$

Our analysis is based on *exactly* the same techniques used to establish the achievability of the AWGN channel capacity, i.e., we first upper-bound the ML error probability with that of a typical set decoder and then take the average of the latter error probability, over the ensemble of random Gaussian codebooks.

In particular, following *exactly* the same argument as in [26, Theorem 10.1.1], we get

$$P_{E|h} \leq 2\epsilon + 2^{3LN'\epsilon}2^{-L\left(N'I(x_k;y_k)-NR\right)}, \\ \leq 3\epsilon$$

for sufficiently large N and L. This follows from the choice of N' in (23). It is important to note that the transmission rate R is constant, i.e., independent of h, and N' is *not* known to the transmitter.

B. Proof of Theorem 2

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 $\boldsymbol{x}$ and $\boldsymbol{y}$. 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

$$\boldsymbol{y} = \begin{bmatrix} g_1 I_{l'} & 0 \\ g_2 h B & g_1 I_{l-l'} \end{bmatrix} \boldsymbol{x} + \begin{bmatrix} 0 \\ g_2 B \end{bmatrix} \boldsymbol{w} + \boldsymbol{v}. \tag{24}$$

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 nonnegative diagonal with the diagonal elements in decreasing order. Using these matrices, we define $\tilde{\boldsymbol{y}} \triangleq T\boldsymbol{y}$, $\tilde{\boldsymbol{x}} \triangleq T\boldsymbol{x}$, $\tilde{\boldsymbol{v}} \triangleq T\boldsymbol{v}$, and $\tilde{\boldsymbol{w}} \triangleq V^H\boldsymbol{w}$, for unitary transformation

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

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

$$I(\boldsymbol{x};\boldsymbol{y}) = I(\tilde{\boldsymbol{x}};\tilde{\boldsymbol{y}}). \tag{25}$$

In terms of the new variables, (24) becomes

$$\begin{aligned}\tilde{\boldsymbol{y}} &= \begin{bmatrix} g_1 I_{l'} & 0 \\ g_2 h D & g_1 I_{l-l'} \end{bmatrix} \tilde{\boldsymbol{x}} + \begin{bmatrix} 0 \\ g_2 D \end{bmatrix} \tilde{\boldsymbol{w}} + \tilde{\boldsymbol{v}} \\ &= \begin{bmatrix} g_1 I_{l'} & 0 \\ g_2 h D & g_1 I_{l-l'} \end{bmatrix} \tilde{\boldsymbol{x}} + \tilde{\boldsymbol{n}} \end{aligned} \tag{26}$$

with

$$\Sigma_{\tilde{\boldsymbol{n}}} = \begin{bmatrix} \sigma_v^2 I_{l'} & 0 \\ 0 & \sigma_v^2 I_{l-l'} + |g_2|^2\sigma_w^2 DD^H \end{bmatrix}. \tag{27}$$

If we denote the nonzero diagonal elements of D as $\{d_i\}_{i=1}^m$, then (27) can be written as

$$\begin{aligned}\tilde{\boldsymbol{y}}_i &= G_i\tilde{\boldsymbol{x}}_i + \tilde{\boldsymbol{n}}_i, \;\; i = 1,\ldots,m \\ \tilde{y}_i &= g_1\tilde{x}_i + \tilde{n}_i, \;\; i = m+1,\ldots,l' \text{ and } i = l'+m+1,\ldots,l\end{aligned}$$

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

$$G_i \triangleq \begin{bmatrix} g_1 & 0 \\ g_2 h d_i & g_1 \end{bmatrix} \tag{28}$$

$$\Sigma_{\tilde{\boldsymbol{n}}_i} = \begin{bmatrix} \sigma_v^2 & 0 \\ 0 & \sigma_v^2 + |g_2|^2 d_i^2 \sigma_w^2 \end{bmatrix}. \tag{29}$$

Note that, according to the SVD theorem

$$m \leq \min\{l', l-l'\}. \tag{30}$$

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

$$\max_{\Sigma_{\tilde{\boldsymbol{x}}}} I(\tilde{\boldsymbol{x}};\tilde{\boldsymbol{y}}) = \sum_{i=1}^{m} \max_{\Sigma_{\tilde{\boldsymbol{x}}_i}} I(\tilde{\boldsymbol{x}}_i;\tilde{\boldsymbol{y}}_i) + \sum_{i=m+1}^{l'} \max I(\tilde{x}_i;\tilde{y}_i) + \sum_{i=l'+m+1}^{l} \max I(\tilde{x}_i;\tilde{y}_i). \tag{31}$$

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

$$I(\tilde{\boldsymbol{x}}_i;\tilde{\boldsymbol{y}}_i) = \log\left(\det\left(I_2 + \Sigma_{\tilde{\boldsymbol{n}}_i}^{-\frac{1}{2}} G_i \Sigma_{\tilde{\boldsymbol{x}}_i} G_i^H \Sigma_{\tilde{\boldsymbol{n}}_i}^{-\frac{1}{2}}\right)\right). \tag{32}$$

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

$$\log\left(\det\left(I_2 + EG_iG_i^H\Sigma_{\tilde{\boldsymbol{n}}_i}^{-1}\right)\right) \leq \max_{\Sigma_{\tilde{\boldsymbol{x}}_i}} I\left(\tilde{\boldsymbol{x}}_i;\tilde{\boldsymbol{y}}_i\right). \tag{33}$$

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

$$\max_{\Sigma_{\tilde{\boldsymbol{x}}_i}} I(\tilde{\boldsymbol{x}}_i;\tilde{\boldsymbol{y}}_i) \leq \log(\det(I_2 + \lambda_{\max}G_iG_i^H\Sigma_{\tilde{\boldsymbol{n}}_i}^{-1})). \tag{34}$$

From (33) and (34), we conclude that

$$\frac{\log\left(\det\left(I_2 + EG_iG_i^H\Sigma_{\tilde{\boldsymbol{n}}_i}^{-1}\right)\right)}{\log(\rho)} \leq \frac{\max_{\Sigma_{\tilde{\boldsymbol{x}}_i}} I\left(\tilde{\boldsymbol{x}}_i;\tilde{\boldsymbol{y}}_i\right)}{\log(\rho)} \leq \frac{\log\left(\det\left(I_2 + \lambda_{\max}G_iG_i^H\Sigma_{\tilde{\boldsymbol{n}}_i}^{-1}\right)\right)}{\log(\rho)}.$$

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

$$\lim_{\rho\to\infty} \frac{\max_{\Sigma_{\tilde{\boldsymbol{x}}_i},d_i} I(\tilde{\boldsymbol{x}}_i;\tilde{\boldsymbol{y}}_i)}{\log(\rho)} = \lim_{\rho\to\infty} \frac{\log\left(\det\left(I_2 + EG_iG_i^H\Sigma_{\tilde{\boldsymbol{n}}_i}^{-1}\right)\right)}{\log(\rho)}.$$

Plugging in for G_i and $\Sigma_{\tilde{\boldsymbol{n}}_i}$ from (28) and (29), respectively, we get

$$\lim_{\rho\to\infty} \frac{\max_{\Sigma_{\tilde{\boldsymbol{x}}_i},d_i} I(\tilde{\boldsymbol{x}}_i;\tilde{\boldsymbol{y}}_i)}{\log(\rho)} = \lim_{\rho\to\infty} \frac{1}{\log(\rho)} \log\Bigg(1+\frac{|g_1|^2E}{\sigma_v^2} + \\ +\frac{(|g_1|^2+|g_2|^2|h|^2|d_i|^2)E}{\sigma_v^2+|g_2|^2d_i^2\sigma_w^2} + \frac{|g_1|^4E^2}{\sigma_v^2(\sigma_v^2+|g_2|^2d_i^2\sigma_w^2)}\Bigg).$$

It is then straightforward to see that

$$\lim_{\rho\to\infty} \frac{\max_{\Sigma_{\tilde{\boldsymbol{x}}_i},d_i} I(\tilde{\boldsymbol{x}}_i;\tilde{\boldsymbol{y}}_i)}{\log(\rho)} = (\max\{2(1-v_1), 1-(v_2+u)\})^+ \quad (35)$$

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

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

which, together with (35), (25), and (31), results in

$$\lim_{\rho\to\infty} \frac{\max_{\Sigma_{\boldsymbol{x}}} I(\boldsymbol{x};\boldsymbol{y})}{\log(\rho)} = (l-2m)(1-v_1)^+ \\ +m(\max\{2(1-v_1), 1-(v_2+u)\})^+. \quad (36)$$

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

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

Letting R grow with ρ according to

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

and using (36), we conclude that, for large ρ

$$O^+ = \big\{(v_1,v_2,u)\in\mathbb{R}^{3+} | (l-2m)(1-v_1)^+ \\ +m(\max\{2(1-v_1), 1-(v_2+u)\})^+ < rl\big\} \quad (37)$$

and thus,

$$P_O(R) \doteq \rho^{-d_o(r)} \text{ for } d_o(r) = \inf_{(v_1,v_2,u)\in O^+}(v_1+v_2+u). \quad (38)$$

As Zheng and Tse have shown in [18, Lemma 5], $d_o(r)$ provides an upper bound on $d^*(r)$ (i.e., the optimal diversity gain at multiplexing gain r)

$$d^*(r) \le d_o(r). \quad (39)$$

From (37) and (38), it is easy to see that the right-hand side of (39) is maximized when m is set to its maximum, which, according to (30), 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.

C. Proof of Theorem 3

The proof closely follows that for the MIMO point-to-point communication system in [18]. In particular, we assume that the source uses a Gaussian random codebook 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

$$P_E(\rho) = P_O(R)P_{E|O} + P_{E,O^c}$$
$$P_E(\rho) \le P_O(R) + P_{E,O^c}$$

where O denotes the outage event. The outage event O is chosen such that $P_O(R)$ dominates P_{E,O^c}, i.e.,

$$P_{E,O^c} \dot{\le} P_O(R) \quad (40)$$

in which case

$$P_E(\rho) \dot{\le} P_O(R). \quad (41)$$

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)), assuming l to be even, changes to

$$P_{\mathrm{PE}|g_1,g_2,h} \le \det\left(I_2 + \frac{1}{2}\Sigma_s\Sigma_n^{-1}\right)^{-l/2} \quad (42)$$

where $\Sigma_{\boldsymbol{s}}$ and $\Sigma_{\boldsymbol{n}}$ denote the covariance matrices of destination observation's signal and noise components during a single frame

$$\Sigma_{\boldsymbol{s}} = \begin{bmatrix} |g_1|^2 & g_1g_2^*b^*h^* \\ g_1^*g_2bh & |g_1|^2+|g_2|^2|bh|^2 \end{bmatrix} E \quad (43)$$

$$\Sigma_{\boldsymbol{n}} = \begin{bmatrix} \sigma_v^2 & 0 \\ 0 & \sigma_v^2+|g_2|^2|b|^2\sigma_w^2 \end{bmatrix}. \quad (44)$$

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:

$$w \le \min\{u, 1\}. \quad (45)$$

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

$$w \triangleq (u)^-$$

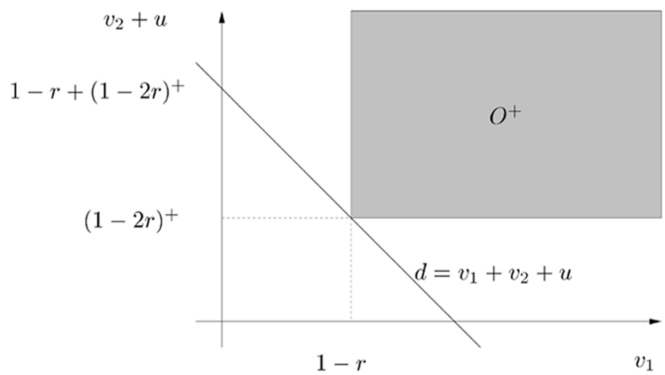

Fig. 10. Outage region for the NAF protocol with a single relay.

which satisfies the constraint given by (45). Interestingly, if we consider $(v_1, v_2, u) \in \mathbb{R}^{3+}$, then w becomes zero and vanishes in the expressions. Plugging (43) and (44) into (42), we obtain

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

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

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

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+}

$$P_{E,O^c} \dot{\leq} \rho^{-d_e(r)} \text{ for } d_e(r) = \inf_{v_1,v_2,u \in O^{c+}} d_e(r,v_1,v_2,u). \quad (46)$$

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

$$P_O \doteq \rho^{-d_o(r)} \text{ for } d_o(r) = \inf_{(v_1,v_2,u)\in O^+} (v_1+v_2+u). \quad (47)$$

Comparing (46) and (47), we realize that for (40) 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 (40). Note that, because of (41), $d_o(r)$ provides a lower bound on the diversity gain achieved by the protocol. But $d_o(r)$, as given by (47), turns out to be identical to right-hand side of (10) (refer to Fig. 10). Thus, the optimal diversity–multiplexing tradeoff for this scenario is indeed given by (12) and the NAF protocol achieves it.

D. Proof of Theorem 5

Instead of considering specific codes, in the following we upper-bound the average probability of error over random Gaussian ensemble of codebooks (employed by both the source and relay). Therefore, averaging is invoked with respect to both the fading channel distribution and the random codebooks. It is then straightforward to see that there is at least one codebook in this ensemble whose average performance, now with respect only to the fading channel distribution, is better than the predictions of our upper bound. For the single-relay DDF protocol, the error probability of the ML decoder, averaged over the ensemble of Gaussian codebooks and conditioned on a certain channel realization, can be upper-bounded using Bayes' rule to give

$$P_{E|g_1,g_2,h} = P_{E,E_r^c|g_1,g_2,h} + P_{E,E_r|g_1,g_2,h}$$
$$P_{E|g_1,g_2,h} \leq P_{E|E_r^c,g_1,g_2,h} + P_{E_r|g_1,g_2,h}$$

where E_r and E_r^c denote the events that the relay decodes source's message erroneously and its complement, respectively. The first step in the proof follows from Lemma 1 by observing that if (13) is met, i.e., if the mutual information between the signal transmitted by the source and the signal received by the relay exceeds lR, then $P_{E_r|g_1,g_2,h}$ can be made arbitrarily small, provided that the code length is sufficiently large. This means that for any $\epsilon > 0$ and for a sufficiently large code length,

$$P_{E|g_1,g_2,h} < P_{E|E_r^c,g_1,g_2,h} + \epsilon.$$

Taking the average over the ensemble of channel realizations gives

$$P_E < P_{E|E_r^c} + \epsilon$$
$$P_E \dot{\leq} P_{E|E_r^c}.$$

This means that the exponential order of $P_{E|E_r^c}$, i.e., destination's ML error probability assuming *error-free* decoding at the relay, provides a lower bound on the diversity gain achieved by the protocol. Therefore, we only need to characterize $P_{E|E_r^c}$, which for the sake of notational simplicity, we will denote by P_E in the sequel. To characterize P_E, we note that the corresponding PEP (recalling (7)) is given by

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

Defining v_1, v_2, and u as the exponential orders of $1/|g_1|^2$, $1/|g_2|^2$, and $1/|h|^2$, respectively, gives

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

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

$$d_e(r) = \inf_{(v_1,v_2,u)\in O^{c+}} l\big[f(1-v_1)^+ + (1-f)(1-\min\{v_1,v_2\})^+ - r\big] + (v_1+v_2+u). \quad (48)$$

Examining (48), we realize that for (40) to hold, O^+ should be defined as

$$O^+ = \{(v_1, v_2, u) \in \mathbb{R}^{3+} | f(1-v_1)^+ + (1-f)(1-\min\{v_1, v_2\})^+ \leq r\} \quad (49)$$

so that it is possible to choose l to make $d_e(r)$ arbitrarily large, ensuring (40). As before, $P_O(R)$ is given by (47), which turns out to be identical to $d(r)$ given by (14). To see this, one needs to consider four different categories of channel realizations. The first category is when both v_1 and v_2 are greater than one. For this category

$$\inf_{\substack{(v_1,v_2,u)\in O^+,\\ v_1>1, v_2>1}} (v_1 + v_2 + u) = 2. \quad (50)$$

The second category is when $1 \geq v_1 \geq 0$ and $v_2 > 1$. It is easy to see from (49) that for this category

$$\inf_{\substack{(v_1,v_2,u)\in O^+,\\ 1\geq v_1\geq 0, v_2>1}} (v_1 + v_2 + u) = 2 - r. \quad (51)$$

The third category to be considered is when $v_1 > 1$ and $1 \geq v_2 \geq 0$. Before proceeding further, note that from (13), one can show that

$$u = 1 - \frac{r}{f}. \quad (52)$$

This implies that $f \geq r$, since u is nonnegative. Returning back to (49), it is easy to verify that for this category

$$v_2 \geq 1 - \frac{r}{1-f}. \quad (53)$$

Now, if $f \geq \max\{r, 1-r\}$, then from (53) and (52) we get

$$\inf_{\substack{(v_1,v_2,u)\in O^+,\\ v_1>1, 1\geq v_2\geq 0,\\ f\geq\max\{r,1-r\}}} (v_1 + v_2 + u) = \inf_{f\geq\max\{r,1-r\}} 2 - \frac{r}{f}$$

or

$$\inf_{\substack{(v_1,v_2,u)\in O^+,\\ v_1>1, 1\geq v_2\geq 0,\\ f\geq\max\{r,1-r\}}} (v_1 + v_2 + u) = \begin{cases} 1 + \frac{1-2r}{1-r}, & \frac{1}{2} \geq r \geq 0 \\ 1, & 1 \geq r \geq \frac{1}{2}. \end{cases} \quad (54)$$

On the other hand, if $1 - r > f \geq r$, then

$$\inf_{\substack{(v_1,v_2,u)\in O^+,\\ v_1>1, 1\geq v_2\geq 0,\\ 1-r>f\geq r}} (v_1 + v_2 + u) = \inf_{1-r>f\geq r} 3 - \frac{r}{1-f} - \frac{r}{f}$$

or

$$\inf_{\substack{(v_1,v_2,u)\in O^+,\\ v_1>1, 1\geq v_2\geq 0,\\ 1-r>f\geq r}} (v_1 + v_2 + u) = 1 + \frac{1-2r}{1-r}, \quad \text{for } \frac{1}{2} > r \geq 0.$$

This means that $\inf_{O^+}(v_1 + v_2 + u)$, for the third category, is indeed given by (54). It is noteworthy that the tradeoff curves given by (50), (51), and (54), are all better than the genie-aided tradeoff. In other words, the diversity gain achieved by this protocol is determined by the fourth category, where both v_1 and v_2 are less than or equal to one. For this category, one needs to consider two cases (note that (52) is still valid, implying $f \geq r$). The first case, when $0.5 \geq f \geq r$, is very easy. Referring to Fig. 11 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 $f > \max\{r, 0.5\}$, is a little bit more difficult. As can be seen from Fig. 12, in this case

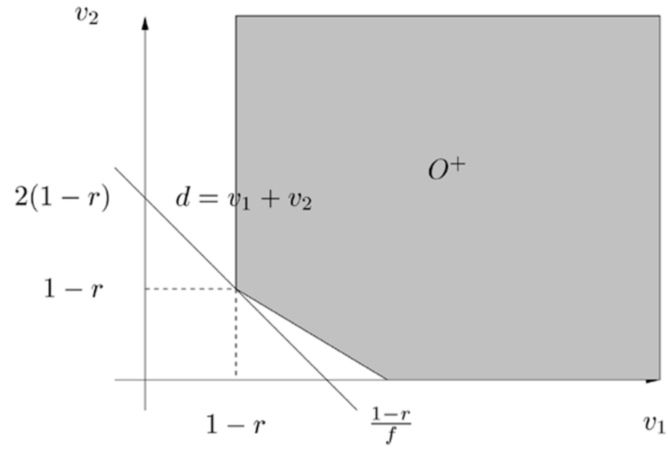

Fig. 11. Outage region for the DDF protocol with a single relay ($f \leq 0.5$).

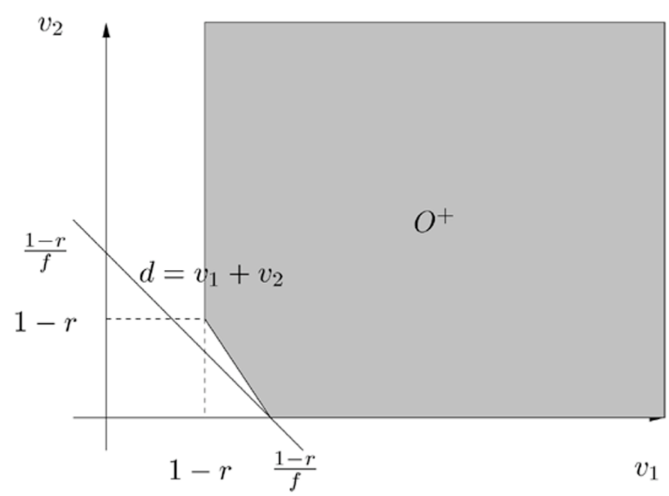

Fig. 12. Outage region for the DDF protocol with a single relay ($f > 0.5$).

$$\inf_{(v_1,v_2)\in O^+} (v_1 + v_2) = \frac{1-r}{f}. \quad (55)$$

From (52) and (55), we conclude that

$$d_o(r) = \inf_{f>\max\{r,0.5\}} 1 + \frac{1-2r}{f} \quad (56)$$

which gives (14). Again, according to (41), $d_o(r)$ 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 achieved 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.

E. Proof of Theorem 6

Inspired by the single-relay case, we use ensembles of Gaussian codebooks at the source and all the relays. To characterize the diversity–multiplexing tradeoff achieved by the DDF protocol with $N-1$ relays, we first label the nodes 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. We then use Bayes' rule to upper-bound the error probability of the ML decoder, averaged over the ensemble of Gaussian codebooks and conditioned on a certain channel realization, to get

$$P_{E|g_j,h_{ji}} \leq P_{E|\{E_p^c\}_{p=2}^N, g_j, h_{ji}} + \sum_{n=2}^{N} P_{E_n|\{E_p^c\}_{p<n}, g_j, h_{ji}} \quad (57)$$

where E_n, $n \in \{2, \ldots, N\}$, denotes the event that node n decodes the source message in error, while E_n^c denotes its complement. Let us now examine $P_{E_n|\{E_p^c\}_{p<n}, g_j, h_{ji}}$, i.e., the probability that node $n \in \{2, \ldots, N\}$ makes an error in decoding the source message, assuming error-free decoding at all previous nodes. It follows from Lemma 1, that if the mutual information between the signals transmitted by the source and active relays and the signal received by node n exceeds lR, then $P_{E_n|\{E_p^c\}_{p<n}, g_j, h_{ji}}$ can be made arbitrarily small, provided that the code length is sufficiently large. This means that for any $\epsilon > 0$ and for sufficiently large code lengths

$$P_{E_n|\{E_p^c\}_{p<n}, g_j, h_{ji}} < \epsilon, \qquad n \in \{2, \ldots, N\}. \quad (58)$$

Using (58), (57) can be written as

$$P_{E|g_j,h_{ji}} \leq P_{E|\{E_p^c\}_{p=2}^N, g_j, h_{ji}} + (N-1)\epsilon.$$

Taking the average over the ensemble of channel realizations gives

$$P_E < P_{E|\{E_p^c\}_{p=2}^N} + (N-1)\epsilon$$
$$P_E \dot{\leq} P_{E|\{E_p^c\}_{p=2}^N}.$$

This means that the exponential order of $P_{E|\{E_p^c\}_{p=2}^N}$, i.e., destination's ML error probability assuming *error-free* decoding at all of the relays, provides a lower bound on the diversity gain achieved by the protocol. Therefore, we only need to characterize $P_{E|\{E_p^c\}_{p=2}^N}$, which for the sake of notational simplicity, we will denote by P_E in the sequel. To characterize P_E, we note that the corresponding PEP is upper-bounded by

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

As before, the gain of the channel that connects the jth 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 the mutual information between its received signal and the signals that the source and other relays transmit exceeds lR. Thus,

$$\sum_{j=1}^{p} l_j \leq \min\left\{l, \left\lceil \frac{lR}{\log(1+|h_{p+1,1}|^2 c\rho)} \right\rceil \right\}, \quad \text{for } N-1 \geq p \geq 1. \quad (59)$$

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

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

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

$$1 - r \leq \sum_{j=1}^{N} \frac{l_j}{l} \min\{1, v_1, \ldots, v_j\}. \quad (60)$$

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

$$P_O(R) \doteq \rho^{-d_o(r)} \text{ for } d_o(r) = \inf_{O^+} \sum_{j=1}^{N} \left(v_j + \sum_{i<j} u_{ji}\right). \quad (61)$$

Defining $\tilde{v}_j \triangleq \min\{v_1, \ldots, v_j\}$, $j = 1, \ldots, N$, lets us simplify (60) and (61) to

$$1 - r \leq \sum_{j=1}^{N} \frac{l_j}{l} \min\{1, \tilde{v}_j\} \quad (62)$$

$$d_o(r) \geq \inf_{O^+} \sum_{j=1}^{N} \left(\tilde{v}_j + \sum_{i<j} u_{ji}\right). \quad (63)$$

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 (59) can also be simplified to

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

or

$$1-\frac{r}{\sum_{k=1}^{p}\frac{l_k}{l}} \le u_{j1}, \qquad \text{for } j>p. \tag{64}$$

In order to characterize $d_o(r)$, we need to consider three cases. The first case is when $1\ge \tilde{v}_1$. In this case, (62) simplifies to

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

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

$$\sum_{j=1}^{N}\tilde{v}_j=\sum_{j=1}^{N}x_j \quad \text{and} \quad 1-r\le \sum_{j=1}^{N}\frac{f_j}{j}x_j \tag{65}$$

where $f_j \triangleq \sum_{k=1}^{j} l_k/l$. From (65), it can be seen that

$$\inf_{\substack{O^+\\ 1\ge\tilde{v}_1}}\sum_{j=1}^{N}\tilde{v}_j = p\left(\frac{1-r}{f_p}\right), \quad \text{where } p=\arg\max_{N\ge j\ge 1}\left\{\frac{f_j}{j}\right\}. \tag{66}$$

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

$$f_p \ge 1-r. \tag{67}$$

From (64), it follows that

$$\inf_{\substack{O^+\\ 1\ge\tilde{v}_1}}\sum_{j>p}u_{j1}=(N-p)\left(1-\frac{r}{f_p}\right), \qquad 1\ge f_p\ge r. \tag{68}$$

Now, from (66) and (68) we conclude that

$$\inf_{\substack{O^+,\\ 1\ge\tilde{v}_1}}\sum_{j=1}^{N}\left(\tilde{v}_j+\sum_{i<j}u_{ji}\right) \ge \inf_{\substack{N\ge p\ge 1,\\ 1\ge f_p\ge max\{r,1-r\}}} d_o(r,p,f_p) \tag{69}$$

where

$$d_o(r,p,f_p)\triangleq p\left(\frac{1-r}{f_p}\right)+(N-p)\left(1-\frac{r}{f_p}\right). \tag{70}$$

It turns out that (70) is an increasing function of p. Therefore, its infimum corresponds to $p=1$. Now, examining $d_o(r,1,f_1)$, i.e.,

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

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

$$\inf_{\substack{O^+,\\ 1\ge\tilde{v}_1}}\sum_{j=1}^{N}\left(\tilde{v}_j+\sum_{i<j}u_{ji}\right) \ge \begin{cases} N(1-r), & \frac{1}{N}\ge r\ge 0\\ 1+\frac{(N-1)(1-2r)}{1-r}, & \frac{1}{2}\ge r\ge\frac{1}{N}\\ \frac{1-r}{r}, & 1\ge r\ge\frac{1}{2}. \end{cases} \tag{71}$$

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

$$\inf_{\substack{O^+\\ \tilde{v}_i>1\ge\tilde{v}_{i+1}}}\sum_{j=1}^{i}\tilde{v}_j=i. \tag{72}$$

In this case, (62) can be written as

$$1-r-f_i\le\sum_{j=i+1}^{N}\frac{l_j}{l}\tilde{v}_j. \tag{73}$$

If $f_i\ge 1-r$, then from (73), we get

$$\inf_{\substack{O^+\\ \tilde{v}_i>1\ge\tilde{v}_{i+1},\\ f_i\ge\max\{r,1-r\}}}\sum_{j=i+1}^{N}\tilde{v}_j=0. \tag{74}$$

On the other hand, from (64), it follows that

$$\inf_{u_{j1}\ge 1-\frac{r}{f_i},\, j>i}\sum_{j=i+1}^{N}u_{j1}=(N-i)\left(1-\frac{r}{f_i}\right), \quad 1\ge f_i\ge r. \tag{75}$$

Now, from (72), (74), and (75) one can see that

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

with

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

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

$$\inf_{\substack{O^+,\\ \tilde{v}_i>1\ge\tilde{v}_{i+1},\\ f_i\ge\max\{r,1-r\}}}\sum_{j=1}^{N}\left(\tilde{v}_j+\sum_{i<j}u_{ji}\right) \ge \begin{cases} 1+\frac{(N-1)(1-2r)}{1-r}, & \frac{1}{2}\ge r\ge 0\\ 1, & 1\ge r\ge\frac{1}{2}. \end{cases} \tag{76}$$

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., $1\ge\tilde{v}_1$). Specifically, $\inf\sum_{j=i+1}^{N}\tilde{v}_j$ is given by (66), with $N-i$, f_p-f_i, $r+f_i$, and $p-i$ substituting N, f_p, r, and p. Thus,

$$\inf_{\substack{O^+,\\ \tilde{v}_i>1\ge\tilde{v}_{i+1},\\ 1-r>f_i\ge r}}\left(\sum_{j=i+1}^{N}\tilde{v}_j\right)=(p-i)\left(\frac{1-r-f_i}{f_p-f_i}\right),$$

$$\text{where } p=\arg\max_{N\ge j\ge i+1}\left\{\frac{f_j-f_i}{j-i}\right\}. \tag{77}$$

Note that (67) still holds. Derivation of $\inf\sum_{j=1}^{N}\sum_{i<j}u_{ji}$ follows from (64)

$$\inf_{u_{j1}\ge 1-\frac{r}{f_k},\, j>k}\sum_{j=1}^{N}\sum_{i<j}u_{ji}\ge(p-i)\left(1-\frac{r}{f_i}\right) + (N-p)\left(1-\frac{r}{f_p}\right), \qquad \text{with } f_p>f_i\ge r. \tag{78}$$

From (72), (77), and (78), we conclude that

$$\inf_{\substack{O^+,\\ \tilde{v}_i>1\geq\tilde{v}_{i+1},\\ 1-r>f_i\geq r}} \sum_{j=1}^{N}\left(\tilde{v}_j+\sum_{i<j}u_{ji}\right) \geq \inf_{\substack{N\geq p>i\geq 1,\\ 1\geq f_p\geq 1-r>f_i\geq r}} d_o(r,i,p,f_i,f_p) \quad (79)$$

where

$$d_o(r,i,p,f_i,f_p) \triangleq i+(p-i)\left(\frac{1-r-f_i}{f_p-f_i}\right) + (p-i)\left(1-\frac{r}{f_i}\right)+(N-p)\left(1-\frac{r}{f_p}\right). \quad (80)$$

As can be seen from (80), $d_o(r,i,p,f_i,f_p)$ 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 (80), we derive a lower bound on it. That is,

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

or

$$\inf_{\substack{N\geq p>i\geq 1,\\ 1\geq f_p\geq 1-r>f_i\geq r}} d_o(r,i,p,f_i,f_p) \geq 1+(N-1)\frac{1-2r}{1-r}, \quad \text{for } \frac{1}{2}>r\geq 0. \quad (81)$$

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

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

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 (81). This means that

$$\inf_{\substack{N\geq p>i\geq 1,\\ 1\geq f_p\geq 1-r>f_i\geq r}} d_o(r,i,p,f_i,f_p) = 1+(N-1)\frac{1-2r}{1-r}, \quad \text{for } \frac{1}{2}>r\geq 0. \quad (82)$$

Now, from (82), (79), and (76), we conclude that

$$\inf_{\substack{O^+,\\ \tilde{v}_i>1\geq\tilde{v}_{i+1}}} \sum_{j=1}^{N}\left(\tilde{v}_j+\sum_{i<j}u_{ji}\right) \geq \begin{cases} 1+\frac{(N-1)(1-2r)}{1-r}, & \frac{1}{2}\geq r\geq 0\\ 1, & 1\geq r\geq \frac{1}{2}. \end{cases} \quad (83)$$

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

$$\inf_{\substack{O^+,\\ \tilde{v}_N>1}} \sum_{j=1}^{N}\left(\tilde{v}_j+\sum_{i<j}u_{ji}\right)\geq N. \quad (84)$$

From (71), (83), and (84) 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.

F. Proof of Theorem 7

To characterize the diversity–multiplexing tradeoff achieved by the CB-DDF protocol, we first label the N destinations according to the order in which they start transmission. That is, the first destination that starts transmission is denoted as destination 1, the next destination as destination 2, and so on. Note that the error probability of destination j can be written as

$$P_{E_j}=P_{E_j|S_j^c}P_{S_j^c}+P_{E_j|S_j}P_{S_j} \quad (85)$$

where S_j denotes the event that destination j decodes the message and starts retransmission before the end of the codeword and S_j^c is its complement. Now, since both P_{S_j} and $P_{S_j^c}$ are less than one, (85) gives

$$P_{E_j}\leq P_{E_j|S_j^c}+P_{E_j|S_j}. \quad (86)$$

In order to characterize $P_{E_j|S_j}$, we need to characterize $P_{E_j|S_j,g,h}$, i.e., destination j's ML error probability, averaged over the ensemble of Gaussian codebooks and conditioned on a certain channel realization, under the assumption that it started transmission before the end of the codeword. Toward this end and through using Bayes' rule, one can upper-bound $P_{E_j|S_j,g,h}$ to get

$$P_{E_j|S_j,g,h}\leq\sum_{i=1}^{j}P_{E_i|\left\{E_p^c\right\}_{p<i},S_j,g,h}. \quad (87)$$

Now, let us examine $P_{E_i|\{E_p^c\}_{p<i},S_j,g,h}$, i.e., the probability that destination i $(i\leq j)$, makes an error in decoding the source message, conditioned on S_j (which ensures that destination i has indeed started retransmission) and assuming *error-free* decoding at all of the active destinations. It follows from Lemma 1, that if the mutual information between the signals transmitted by the source and active destinations and the signal received by destination i exceeds lR (which is implied by S_j), then $P_{E_i|\left\{E_p^c\right\}_{p<i},S_j,g,h}$ can be made arbitrarily small, provided that the code length is sufficiently large. This means that for any $\epsilon>0$ and for sufficiently large code lengths

$$P_{E_i|\left\{E_p^c\right\}_{p<i},S_j,g,h}<\epsilon, \qquad i\leq j. \quad (88)$$

Using (88), (87) can be written as

$$P_{E_j|S_j,g,h}\leq j\epsilon.$$

Taking the average over the ensemble of channel realizations gives

$$P_{E_j|S_j}<j\epsilon.$$

This together with (86), yields

$$P_{E_j}<P_{E_j|S_j^c}+j\epsilon$$
$$P_{E_j}\dot{\leq}P_{E_j|S_j^c}. \quad (89)$$

This means that the exponential order of $P_{E_j|S_j^c}$ provides a lower bound on the diversity gain achieved by the protocol. Now, examining $P_{E_j|S_j^c}$, it is easy to realize that the event in which the jth destination (out of N destinations), spends the entire codeword listening, i.e., S_j^c, is identical to the DDF relay protocol with the rest of the destinations taking the role of the $N-1$ relays. Thus, from (89), we see that communication to the jth destination achieves the same diversity order as does the DDF relay protocol with $N-1$ relays, namely, (18). This completes the proof.

G. Proof of Theorem 8

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.* [19] 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 $\{1,\dots,N\}$ 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

$$P_E(\rho) = \sum_I P_{E^I}(\rho). \tag{90}$$

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

$$P_{E^I}(\rho) = P_O(R)P_{E^I|O} + P_{E^I,O^c}$$
$$P_{E^I}(\rho) \le P_O(R) + P_{E^I,O^c}$$

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

$$P_{E^I,O^c} \dot{\le} P_O(R). \tag{91}$$

Thus,

$$P_{E^I}(\rho) \dot{\le} P_O(R),$$

which, together with (90), results in

$$P_E(\rho) \dot{\le} P_O(R). \tag{92}$$

This means that $P_O(R)$, as defined by (91), 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 $P_{\mathrm{PE}^I|g_j,h_{ji}}$ (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 $P_{\mathrm{PE}^I|g_j,h_{ji}}$, 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*)

$$P_{\mathrm{PE}^I|g_j,h_{ji}} \le \det\left(I_m + \frac{1}{2}\Sigma_{\boldsymbol{s}^I}\Sigma_{\boldsymbol{n}^I}^{-1}\right)^{-1}. \tag{93}$$

In (93), $\Sigma_{\boldsymbol{s}^I}$ and $\Sigma_{\boldsymbol{n}^I}$ 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 $\Sigma_{\boldsymbol{s}^I}$ and $\Sigma_{\boldsymbol{n}^I}$ are both positive-definite matrices, the right-hand side of (93) can be upper-bounded as

$$P_{\mathrm{PE}^I|g_j,h_{ji}} \dot{\le} \det(\Sigma_{\boldsymbol{s}^I})^{-1}\det(\Sigma_{\boldsymbol{n}^I}). \tag{94}$$

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 3. We also note that the exponential orders of the broadcast gains $\{|a_j|^2\}_{j=1}^N$ are zero. Furthermore, 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 nonnegative realizations. With this ideas in mind, we return to (94) and claim that

$$\det(\Sigma_{\boldsymbol{n}^I}) \dot{\le} 1. \tag{95}$$

To understand (95), 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 nonpositive powers of ρ, (95) must hold. Combining (95) and (94) yields

$$P_{\mathrm{PE}^I|v_j,u_{ji}} \dot{\le} \det(\Sigma_{\boldsymbol{s}^I})^{-1}, \quad \text{for } v_j \ge 0,\ u_{ji} \ge 0. \tag{96}$$

As mentioned earlier, $\Sigma_{\boldsymbol{s}^I}$ 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 $\Sigma_{\boldsymbol{s}^I}$, 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 $1 \le |I| \le N$). 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 rebroadcast of $x_{j,k}$ (where $k' \in \{k, k+1\}$). As an example, consider the case when $N = 4$ and assume that during

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

$$P_{\mathrm{PE}^I|v_j,u_{ji}} \dot{\leq} \rho^{-\left[-\sum_{j\in I}\left((m-1)v_j+\sum_{i\notin I}\min\{v_j,u_{ji}+v_i\}\right)+m(N-1)\right]L}, \qquad v_j \geq 0,\ u_{ji}\geq 0. \tag{99}$$

$$P_{E^I|v_j,u_{ji}} \dot{\leq} \rho^{-\left[-\sum_{j\in I}\left((m-1)v_j+\sum_{i\notin I}\min\{v_j,u_{ji}+v_i\}\right)+m(N-1)(1-r)\right]L}, \qquad v_j \geq 0,\ u_{ji}\geq 0. \tag{100}$$

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 $\boldsymbol{s}^I_{j,k}$, 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

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

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

$$\Sigma_{\boldsymbol{s}^I} = \sum_{k=1}^{l}\sum_{j\in I}\mathrm{E}\left\{\boldsymbol{s}^I_{j,k}\left(\boldsymbol{s}^I_{j,k}\right)^H\right\}. \tag{97}$$

In order to illuminate some of the properties of $\boldsymbol{s}^I_{j,k}$, 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

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

where $j_p\in I$ and $k_p\in\{1,\ldots,l\}$ are chosen such that the first nonzero elements of $\boldsymbol{s}^I_{j_p,k_p}$, $p=1,\ldots,ml$ 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 nonzero element of $\boldsymbol{s}^I_{j_p,k_p}$ (i.e., the pth diagonal element of S^I) will be $g_ja_jx_{j,k}$ or $g_ib_ih_{ij}a_jx_{j,k}$, respectively. Next, we define $\boldsymbol{\psi}^I_{j,k}$ as the signature of $x_{j,k}$, i.e.,

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

and Ψ^I as

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

It follows then, that Ψ^I is also lower-triangular with the pth diagonal element being equal to g_ja_j or $g_ib_ih_{ij}a_j$. Using these definitions, (97) can be written as

$$\Sigma_{\boldsymbol{s}^I} = E\sum_{k=1}^{l}\sum_{j\in I}\boldsymbol{\psi}^I_{j,k}\left(\boldsymbol{\psi}^I_{j,k}\right)^H. \tag{98}$$

The significance of Ψ^I can now be seen from the fact that (98) can be written as

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

Now, as the determinant of triangular matrices is simply the product of their diagonal elements, from (96) we get the first equation at the top of the page. It is obvious that for large L's, the previous inequality can be rewritten as (99) at the top of the page. 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, we have (100), also at the top of the page. This conditional type-I error probability leads to

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

where

$$\begin{aligned} d_{e^I}(r) \triangleq \min_{O_c^+}\sum_j\left(v_j+\sum_i u_{ji}\right) &\\ +\Bigg[-\sum_{j\in I}\left((m-1)v_j+\sum_{i\notin I}\min\{v_j,u_{ji}+v_i\}\right) &\\ +m(N-1)(1-r)\Bigg]L. & \end{aligned} \tag{101}$$

Examining (101), we realize that for (91) 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 $I\subseteq\{1,\ldots,N\}$:

$$\sum_{j\in I}\left((m-1)v_j+\sum_{i\notin I}\min\{v_j,v_i+u_{ji}\}\right) \geq m(N-1)(1-r). \tag{102}$$

This way, by choosing large enough l, $d_{e^I}(r)$ can be made arbitrary large and thus (91) is always met. From (102), it follows that

$$\sum_{j\in I}\left((m-1)v_j+\sum_{i\notin I}\min\{v_j, v_i+\max_{j\neq i}\{u_{ji}\}\}\right) \geq m(N-1)(1-r). \quad (103)$$

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

$$\sum_{j\in I} v_j \geq m(1-r). \quad (104)$$

On the other hand, replacing $\min\{v_j, v_i+\max_{j\neq i}\{u_{ji}\}\}$ in (103) by $v_i+\max_{j\neq i}\{u_{ji}\}$ results in

$$(m-1)\sum_{j\in I} v_j + m\sum_{i\notin I}\left(v_i+\max_{j\neq i}\{u_{ji}\}\right) \geq m(N-1)(1-r). \quad (105)$$

Under the constraints given by (104) and (105), it is easy to see that

$$\inf_{O^+}\left(\sum_{j\in I} v_j+\sum_{i\notin I}\left(v_i+\max_{j\neq i}\{u_{ji}\}\right)\right) \geq N(1-r). \quad (106)$$

Now, from (106) and (61), it follows that

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

Again, according to (92), $d_o(r)$ 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.

References

[1] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity. Part I. System description," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1927–1938, Nov. 2003.

[2] ——, "User cooperation diversity. Part II. Implementation aspects and performance analysis," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1939–1948, 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 Trans. Inf. Theory*, submitted for publication.

[4] J. N. Laneman and G. W. Wornell, "Distributed space-time-coded protocols for exploiting cooperative diversity in wireless networks," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2415–2425, Oct. 2003.

[5] A. Stefanov and E. Erkip, "Cooperative coding for wireless networks," *IEEE Trans. Commun.*, to be published.

[6] M. Janani, A. Hedayat, T. Hunter, and A. Nosratinia, "Coded cooperation in wireless communications: Space-time transmission and iterative decoding," *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 362–371, Feb. 2004.

[7] R. U. Nabar, F. W. Kneubuhler, and H. Bölcskei, "Performance limits of amplify-and-forward based fading relay channels," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing*, vol. 4, Montreal, QC, Canada, May 2004, pp. 565–568.

[8] R. U. Nabar, H. Bölcskei, and F. W. Kneubuhler, "Fading relay channels: Performance limits and space-time signal design," *IEEE J. Sel. Areas Commun.*, vol. 22, no. 6, pp. 1099–1109, Aug. 2004.

[9] N. Prasad and M. K. Varanasi, "Diversity and multiplexing tradeoff bounds for cooperative diversity protocols," in *Proc. IEEE Int. Symp. Information Theory*, Chicago, IL, Jun./Jul. 2004, p. 268.

[10] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Trans. Inf. Theory*, vol. 51, no. 9, pp. 3037–3063, Sep. 2005.

[11] A. Host-Madsen, "On the capacity of wireless relaying," in *Proc. 2002 IEEE 56th Vehicular Technology Conf.*, vol. 3, Sep. 2002, pp. 1333–1337.

[12] ——, "On the capacity of cooperative diversity in slow fading channels," in *Proc. 40th Allerton Conf. Communications, Control and Computing*, Monticello, IL, Oct. 2002.

[13] ——, "A new achievable rate for cooperative diversity based on generalized writing on dirty paper," in *Proc. IEEE Int. Symp. Information Theory*, Yokohama, Japan, Jun./Jul. 2003.

[14] M. A. Khojastepour, A. Sabharwal, and B. Aazhang, "Lower bounds on the capacity of Gaussian relay channel," in *Proc. 38th Annu. Conf. Information Sciences and Systems*, Princeton, NJ, Mar. 2004.

[15] ——, "On the capacity of Gaussian cheap relay channel," in *Proc IEEE 2003 Global Communications Conf.*, San Francisco, CA, Dec. 2003.

[16] U. Mitra and A. Sabharwal, "On achievable rates for complexity constrained relay channels," in *Proc 41st Allerton Conf. Communication, Control and Computing*, Monticello, IL, Oct. 2003.

[17] S. Zahedi, M. Mohseni, and A. El Gamal, "On the capacity of AWGN relay channel with linear relaying functions," in *Proc. IEEE Int. Symp. Information Theory*, Chicago, IL, Jun. 2004, p. 399.

[18] L. Zheng and D. N. C. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple antenna channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1073–1096, May 2003.

[19] D. N. C. Tse, P. Viswanath, and L. Zheng, "Diversity–multiplexing tradeoff in multiple-access channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 9, pp. 1859–1874, Sep. 2004.

[20] K. A. Yazdi, H. El Gamal, and P. Schniter, "On the design of cooperative transmission schemes," in *Proc. 41st Allerton Conf. Communication, Control, and Computing*, Monticello, IL, Oct. 2003.

[21] K. Azarian, H. El Gamal, and P. Schniter, "On the achievable diversity-vs-multiplexing tradeoff in cooperative channels," in *Proc. Conf. Information Sciences and Systems*, Princeton, NJ, Mar. 2004.

[22] H. El Gamal, "The diversity-multiplexing tradeoff in half-duplex cooperative channels: Achievable curves and optimal strategies," in *LIDS Colloquia*. Cambridge, MA: MIT Elec. Eng. Comp. Sci. Dept., May 2004.

[23] K. Azarian, H. El Gamal, and P. Schniter, "On the achievable diversity-multiplexing tradeoff in half duplex cooperative channels," in *Proc. 42nd Allerton Conf. Communication, Control, and Computing*, Monticello, IL, Oct. 2004.

[24] ——, "Achievable diversity-vs-multiplexing tradeoffs in half-duplex cooperative channels," in *Proc. IEEE Information Theory Workshop*, San Antonio, TX, Oct. 2004.

[25] T. M. Cover and A. A. El Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory*, vol. IT-25, no. 5, pp. 572–584, Sep. 1979.

[26] T. M. Cover and J. A. Thomas, *Elements of Information Theory*: Wiley, 1991.