

Low Complexity and Fractional Coded Cooperation for Wireless Networks

Andrew W. Eckford, *Member, IEEE*, Josephine P. K. Chu, *Student Member, IEEE*,
and Raviraj S. Adve, *Senior Member, IEEE*

Abstract—Wireless networks, and especially wireless sensor networks, have complexity and energy constraints, within which they must confront the challenging wireless fading environment. In this paper, fractional cooperation is introduced, which is shown to provide energy-efficient and low-complexity diversity gains for constant energy costs per bit throughout the network. To minimize complexity, cooperation is based on demodulate-and-forward, wherein the relay nodes encode demodulated, not decoded symbols. A scheme is presented for cooperative error-control coding in complexity-constrained networks, using low-density generator-matrix codes and repeat-accumulate codes, both chosen for being simple to encode, as well as for their easily adaptable rates. It is shown that these codes, coupled with fractional cooperation, are robust to system parameters and conditions, and introduce little added complexity at the receiver, while providing excellent performance.

Index Terms—Wireless networks, error-control coding, relay networks, cooperative communication.

I. INTRODUCTION

FOR many contemporary applications, wireless networks are required to be unobtrusive, with numerous nodes that are dependent on a battery power source. For instance, the present paper is motivated by *sensor networks* [1], in which distributed sensors, embedded in *nodes*, take local measurements of a phenomenon, and form a wireless network to share their information amongst themselves, or transmit it to some central authority, known as the *data sink*. Such networks have a wide variety of potential applications, from wildlife monitoring [2] to load monitoring in structures [3]. In this type of wireless network, nodes must be as small, inexpensive, and as efficient as possible, which places stringent constraints on their computational and energy resources. On the other hand, the data sink is assumed to have access to substantial energy and computational resources, within the limits of reasonable expense and contemporary technology.

Manuscript received February 23, 2007; revised September 12, 2007; accepted November 30, 2007. The associate editor coordinating the review of this paper and approving it for publication was A. Nosratinia. Material in this paper was presented in part at the 40th Annual Conference on Information Sciences and Systems, Princeton, NJ, March 2006; and at the 2006 IEEE International Conference on Communications, Istanbul, Turkey, June 2006.

A. W. Eckford is with the Department of Computer Science and Engineering, York University, 4700 Keele Street, Toronto, Ontario, Canada M3J 1P3 (e-mail: aeckford@yorku.ca).

J. P. K. Chu and R. S. Adve are with The Edward S. Rogers Sr. Department of Electrical and Computer Engineering, University of Toronto, 10 King's College Road, Toronto, Ontario, Canada M5S 3G4 (e-mail: {chuj, rsadve}@comm.utoronto.ca).

Digital Object Identifier 10.1109/TWC.2008.070221.

The challenging nature of the wireless channel poses several problems for ubiquitous wireless networks, in that their low complexity nodes must ensure reliable communication in the presence of fading. A useful technique is *spatial diversity*, which exploits the large number of antennas available in the wireless network – at least one of the nodes likely has a good link to the data sink, and that node can be used as a *relay* for its neighbors [4]. This idea can be generalized to cooperative diversity [5], [6], [7], in which two (or more) transmitters assist each other in sending their messages to a common receiver.

To implement cooperative diversity, it is generally necessary to use error-control coding across multiple nodes in the network – a scheme known as *cooperative coding*. This concept has attracted much attention in the literature. In [8], cooperative diversity was combined with error-control coding as a more flexible strategy than merely repetition by the partner. In [9], a simple and practical ARQ-based scheme was introduced to improve the performance of relays, while in [10], adaptive Raptor codes (described in [11]) were used to guarantee decoding over a source-relay link. There have been a variety of alternative approaches presented in the literature. In [12], a scheme using punctured convolutional codes was presented. Efforts using more powerful error-correcting codes included [13], where a capacity-approaching method using Turbo codes was presented; and [14], [15], [16], where various optimized schemes using LDPC codes were presented.

The results in this paper are inspired by scenarios such as the ones described in [9], [10], [17]. However, in the spirit of our focus on low complexity, low-energy devices, we are more concerned with relay schemes that achieve both diversity and reliable transmission in a robust and low-cost manner, rather than in maximizing data rate. To this end, we make two key assumptions throughout this paper, which are distinct from other scenarios in the available literature. First, we place stringent constraints on the nodes' computational abilities, which we formalize in Section II-C (e.g., they are incapable of decoding each others' error-control-coded transmissions). Second, we assume that it is more efficient to vary the number of transmitted symbols, at constant transmitter power, rather than varying the transmitter power, for a constant number of transmitted symbols. Furthermore, our results do not rely on any external knowledge or intervention, such as knowledge of the geographic positions of the nodes, or any form of ARQ. With these assumptions in mind, the main contributions of this paper are as follows. First, suppose that, due to its

own constraints, relay nodes can only offer a fraction of their transmission window to relay their neighbor's data. This is significantly different from the widely used "all-or-nothing" model where a relay either gives up all its power/bandwidth resources to cooperation or does not cooperate at all. For this scenario, we provide an information-theoretic analysis to show that, *once a minimum number of relays are in use*, each additional relay contributes a full order of diversity, *regardless of the fraction of the source's symbols that are relayed*. Indeed, fractional cooperation allows us to minimize the number of symbols transmitted by each relay to achieve a given diversity order. Second, we develop the use of error-correcting codes which are simple to encode, particularly low-density generator-matrix (LDGM) and repeat-accumulate (RA) codes, in relay systems. Furthermore, we investigate these codes under the constraint that relays *cannot decode* received codewords – they merely *demodulate* the signal before re-encoding it. Demodulate-and-forward was proposed in different contexts in [18], [19], although our process differs by allowing the relay to encode the demodulated data using flexible codes. In spite of the simplicity and adaptability of this scenario, we show that excellent and robust performance can be achieved. Although these results are applicable to general wireless networks, they are most likely to be applicable to networks with low complexity nodes, such as sensor networks.

The remainder of the paper is organized as follows. In Section II, we introduce the system model used throughout this paper. In Section III, we extend our framework to multiple relays by introducing *fractional cooperation*, and develop the theoretical properties of our approach. In Section IV, we provide a framework for coded cooperation under complexity constraints, which is particularly useful for fractional cooperation, and which is based on a demodulate-and-forward scheme. Finally, in Section V, we present simulation results demonstrating the efficacy of our method.

II. SYSTEM MODEL

Our relaying system is assumed to operate in an environment with the following conditions. Nodes may not transmit and receive at the same time (this half-duplex assumption has previously been suggested for "cheap" hardware in [20]; performance limits for half-duplex relaying were considered in [7]). Each node has an orthogonal channel in which it transmits, so that there is no co-channel interference on any link. Signals are observed coherently at each node. The communication links use binary phase shift keying (BPSK) for data modulation. The channel amplitudes follow a Rayleigh fading model. Finally, the fading amplitudes are quasi-static, remaining constant throughout the transmission of a single codeword. The orthogonal channel and coherency assumptions are made for ease of exposition, but in Section V-C we briefly discuss how these assumptions may be relaxed. Similarly, using BPSK simplifies the exposition; coupled with coherent reception, we can focus on real-valued signals. However, our results can be easily generalized to other modulation schemes.

Throughout this paper, we disregard the problem of assigning relay resources to assist a particular source. We assume that the destination is aware of the identities of all the

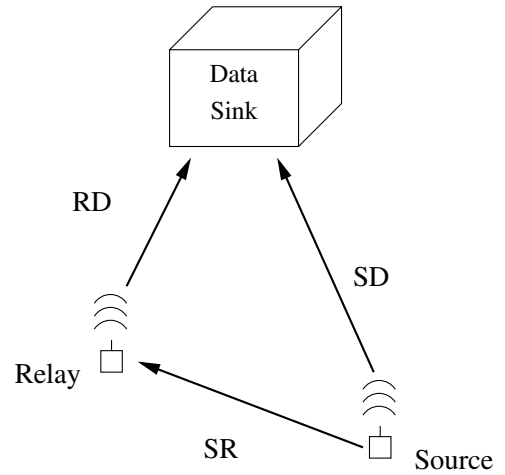


Fig. 1. A single-relay system.

relays, and of all parameters relevant to the relaying, including channel state information.

A. One-relay model

We start by describing a three-node model, having a source, a single relay, and a destination. As shown in Figure 1, there are three radio links: source-to-relay (SR), source-to-destination (SD), and relay-to-destination (RD). A particular realization of the channel is parameterized by the triple (a_{SR}, a_{SD}, a_{RD}) , representing the amplitude on the three links.

The overall communication occurs in two phases. In the first phase, the source encodes its data sequence as a codeword of an error-correcting code, and transmits the length- n sequence $\mathbf{x}^{(S)} = [x_1^{(S)}, x_2^{(S)}, \dots, x_n^{(S)}] \in \mathcal{X}^n$, where \mathcal{X} represents the set of symbols available to the transmitter. The relay and destination simultaneously observe this sequence as real-valued vectors $\mathbf{y}^{(R)}$ and $\mathbf{y}^{(SD)}$, respectively, where

$$\mathbf{y}^{(R)} = a_{SR}\mathbf{x}^{(S)} + \mathbf{n}^{(R)}, \quad (1)$$

$$\mathbf{y}^{(SD)} = a_{SD}\mathbf{x}^{(S)} + \mathbf{n}^{(SD)}, \quad (2)$$

where $\mathbf{n}^{(R)}$ and $\mathbf{n}^{(SD)}$ represent unit-variance additive white Gaussian noise (AWGN) vectors at the relay and destination, respectively.

The signal $\mathbf{y}^{(R)}$ is incident to the relay. Let the function $\phi: \mathbb{R}^n \rightarrow \mathcal{X}^m$, represent the processing at the relay, so that

$$\mathbf{x}^{(R)} = \phi(\mathbf{y}^{(R)}) \quad (3)$$

where $\mathbf{x}^{(R)}$ is the the sequence transmitted by the relay in the second phase, which is also the codeword of an error-correcting code. (Note that $\mathbf{x}^{(R)}$ is an m -dimensional vector, while $\mathbf{x}^{(S)}$ is an n -dimensional vector, which admits the possibility that the transmissions of the source and relay are of different lengths.) The signal received by the destination in this second phase is given by

$$\mathbf{y}^{(RD)} = a_{RD}\mathbf{x}^{(R)} + \mathbf{n}^{(RD)}. \quad (4)$$

The Rayleigh fading model is parameterized by $\bar{\gamma}$, the average signal-to-noise ratio, assuming unit noise power. Thus, in a fading channel, the amplitude triple (a_{SR}, a_{SD}, a_{RD}) is

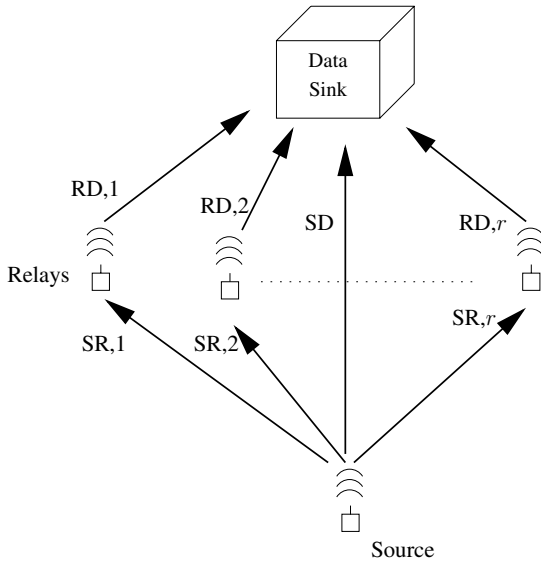


Fig. 2. A multiple-relay system.

a three-dimensional vector of three independent Rayleigh-distributed random variables, parameterized by the triple $(\bar{\gamma}_{\text{SR}}, \bar{\gamma}_{\text{SD}}, \bar{\gamma}_{\text{RD}})$.

B. Multiple-relay model

A system with r relays is depicted in Figure 2. The notation is similar to the one-relay case, with the following generalizations:

- 1) There is an index set $\mathcal{I} = \{1, 2, \dots, r\}$ containing a unique index for each relay.
- 2) The channel is parameterized by $(\mathbf{a}^{(\text{SR})}, a^{(\text{SD})}, \mathbf{a}^{(\text{RD})})$, where $\mathbf{a}^{(\text{SR})} = [a_1^{(\text{SR})}, a_2^{(\text{SR})}, \dots, a_r^{(\text{SR})}]$ is the vector of source-to-relay amplitudes for each relay in \mathcal{I} , and $\mathbf{a}^{(\text{RD})} = [a_1^{(\text{RD})}, a_2^{(\text{RD})}, \dots, a_r^{(\text{RD})}]$ is the vector of source-to-relay amplitudes for each relay in \mathcal{I} . (Note that $a^{(\text{SD})}$ is still a scalar.)
- 3) The source-destination link (2) does not change, but Equations (1) and (4) must be specified for each relay. The notation can be made compact by writing

$$\mathbf{Y}^{(\text{R})} = \mathbf{A}^{(\text{SR})} \mathbf{X}^{(\text{S})} + \mathbf{N}^{(\text{R})}, \quad \text{and} \quad (5)$$

$$\mathbf{Y}^{(\text{RD})} = \mathbf{A}^{(\text{RD})} \mathbf{X}^{(\text{R})} + \mathbf{N}^{(\text{RD})}. \quad (6)$$

For (5), we define

$$\mathbf{Y}^{(\text{R})} := \begin{bmatrix} \mathbf{y}^{(\text{R},1)} \\ \mathbf{y}^{(\text{R},2)} \\ \vdots \\ \mathbf{y}^{(\text{R},r)} \end{bmatrix}, \quad \mathbf{A}^{(\text{SR})} := \text{diag}(\mathbf{a}^{(\text{SR})}),$$

$$\mathbf{X}^{(\text{S})} := \begin{bmatrix} \mathbf{x}^{(\text{S})} \\ \mathbf{x}^{(\text{S})} \\ \vdots \\ \mathbf{x}^{(\text{S})} \end{bmatrix}, \quad \mathbf{N}^{(\text{R})} := \begin{bmatrix} \mathbf{n}^{(\text{R},1)} \\ \mathbf{n}^{(\text{R},2)} \\ \vdots \\ \mathbf{n}^{(\text{R},r)} \end{bmatrix}.$$

The superscript (R, i) for $i \in \mathcal{I}$ refers to processes at

the i th relay. Similarly, for (6), we define

$$\mathbf{Y}^{(\text{RD})} := \begin{bmatrix} \mathbf{y}^{(\text{RD},1)} \\ \mathbf{y}^{(\text{RD},2)} \\ \vdots \\ \mathbf{y}^{(\text{RD},r)} \end{bmatrix}, \quad \mathbf{A}^{(\text{RD})} := \text{diag}(\mathbf{a}^{(\text{RD})}),$$

$$\mathbf{X}^{(\text{R})} := \begin{bmatrix} \mathbf{x}^{(\text{R},1)} \\ \mathbf{x}^{(\text{R},2)} \\ \vdots \\ \mathbf{x}^{(\text{R},r)} \end{bmatrix}, \quad \mathbf{N}^{(\text{RD})} := \begin{bmatrix} \mathbf{n}^{(\text{RD},1)} \\ \mathbf{n}^{(\text{RD},2)} \\ \vdots \\ \mathbf{n}^{(\text{RD},r)} \end{bmatrix}.$$

The superscript (RD, i) for $i \in \mathcal{I}$ refers to processes taking place at the destination, corresponding to the i th relay. Note that, as indicated in $\mathbf{x}^{(\text{R})}$, with the superscripts (R, i) , each relay is allowed to send a *different* transmission to the data sink.

- 4) For all $i \in \mathcal{I}$, $\phi_i(\cdot)$ represents the processing function at the i th relay (thus, each relay may process the signal differently).
- 5) In a Rayleigh fading scenario, the amplitudes $(\mathbf{a}^{(\text{SR})}, a^{(\text{SD})}, \mathbf{a}^{(\text{RD})})$ are independent and Rayleigh distributed, with $(\bar{\gamma}^{(\text{SR})}, \bar{\gamma}^{(\text{SD})}, \bar{\gamma}^{(\text{RD})})$ representing respective vectors of average SNRs on each link.

In (6), it seems like the relay sequences $\mathbf{x}^{(\text{R},i)}$ must all be of the same length. However, this is only done for convenience in specifying the notation. For instance, the sequences $\mathbf{x}^{(\text{R},i)}$ can be zero-padded so that they are all the same length.

C. Relay Processing Function

The previous section introduced the function $\phi_i(\cdot)$ to represent the processing performed by the relay. As mentioned in the introduction, the relay operates under complexity constraints, so in this section we define $\phi_i(\cdot)$, bearing this in mind. In particular, since high-performance error-correcting codes require computationally complex decoding algorithms, we assume that these algorithms are beyond the reach of small, low complexity wireless nodes. As a result, our focus here is on demodulate-and-forward, rather than decode-and-forward. Our relay processing function may be summarized as follows: demodulate by making hard decisions on the observed symbols, and encode these as information symbols using another error-correcting code. If we choose not to forward all the hard decisions, the ones to be forwarded are chosen at random (with the identities of the forwarded decisions known to the data sink).

Since all transmitters in the system use BPSK modulation, each transmitter's symbol alphabet is $\mathcal{X} = \{+1, -1\}$. For convenience, we define $\sigma : \{0, 1\} \rightarrow \{+1, -1\}$ as the function for translating between binary alphabets, with $\sigma^{-1}(\cdot)$ its inverse. With a slight abuse of notation, if the argument of $\sigma(\cdot)$ is a vector, then the function $\sigma(\cdot)$ is applied to each element of the vector, and similarly for $\sigma^{-1}(\cdot)$.

The function $\phi_i(\cdot)$ takes $\mathbf{y}^{(\text{SR},i)}$ as its argument. It is difficult to describe $\phi_i(\mathbf{y}^{(\text{SR},i)})$ in closed form, so we will describe the sequence of operations that are carried out to calculate it. From (5), the i th relay observes $\mathbf{y}^{(\text{SR},i)}$, and makes hard decisions on it, forming $\text{sign}(\mathbf{y}^{(\text{SR},i)})$, where $\text{sign}(\cdot)$ returns $+1$ if the argument is positive, and -1 if the

argument is negative, with the same modified definition for a vector argument. The bit values in $\{0, 1\}$ are recovered from this sequence, resulting in

$$\mathbf{b} := \sigma^{-1} \left(\text{sign} \left(\mathbf{y}^{(R,i)} \right) \right).$$

The i th relay then selects k of the n received bits in \mathbf{b} to relay, where $k \leq n$, and where k could be different from relay to relay. Let $\mathbf{q} \in \{0, 1\}^n$ represent the indicator sequence which indicates the bits from \mathbf{b} to be relayed; clearly, the Hamming weight of \mathbf{q} is k . The relay then forms the information sequence \mathbf{w} from \mathbf{b} by letting h_j represent the indices of these bits, given formally by

$$h_j := \min \left\{ h : \sum_{t=1}^h q_t = j \right\}, \quad (7)$$

and then letting $w_j := b_{h_j}$. That is, \mathbf{w} contains those elements, and only those elements, of \mathbf{b} where $q_t = 1$. This length- k information sequence may be encoded with an error-correcting code, resulting in the length- m codeword \mathbf{c} (specifics are given for two types of codes in Section IV). Finally, the function $\phi_i(\mathbf{y}^{(R,i)})$ returns the modulated version of the codeword, $\sigma(\mathbf{c})$.

The relays, therefore, demodulate the received signal, select a subset of the received *coded bits* as the information bits to be transmitted, and encode these into a codeword for transmission – hence the terminology *demodulate-and-forward*. Note that the indicator sequence \mathbf{q} and k may differ from relay to relay. We assume that the destination has complete knowledge of the function $\phi_i(\cdot)$ used by each relay.

III. FRACTIONAL COOPERATION

In this section, we present our main theoretical contribution: the notion of *fractional cooperation*, in which a node enlists a large number of its neighbors to each relay a *small fraction* of its transmission (i.e., each relay uses an indicator sequence \mathbf{q} with weight much less than n). This fraction may be fixed in advance or determined dynamically by the bandwidth/energy resources available at each relay.

For this approach, our results indicate that a fractional cooperation system has a critical number of relays, r_c , depending on its system parameters. If the number of relays is greater than or equal to r_c , then each additional relay provides a full order of diversity gain (Theorem 1). Furthermore, the cost of complexity-reducing assumptions, such as a lack of centralized coordination, and no intermediate decoding, is at most a *finite number* of orders of diversity; thus, the above property is not affected by the relaxation of these assumptions (Corollary to Theorem 1).

The first point illustrates why this technique is powerful: each relay could be relaying an arbitrarily small fraction of the source's transmission, but each additional relay nonetheless adds a full order of diversity. Furthermore, the second point indicates that the strategy is highly robust and thus appropriate for a distributed, complexity-constrained network. Fractional cooperation is therefore particularly useful in systems where greater power savings are achieved by varying the number of symbols transmitted, rather than varying the transmitter amplitude – for example, where a constant amount of energy

is expended in observing or processing each symbol to be relayed, in addition to the power used by the radio. Our scheme minimizes these costs by minimizing the number of symbols required from each potential relay.

A. Key assumptions and definitions

We adopt all the modeling assumptions given in Section II. We assume that the relays choose the symbols they relay at random and uniformly from all the symbols transmitted by the source. Rather than assigning a particular number of symbols to relay, we assume for convenience that each symbol transmitted by the source has some probability ν of being relayed by a given relay, so the expected value of bits relayed by a neighbor is νn . We will also assume for convenience that *the average SNR $\bar{\gamma}$ on every link is the same* (but we will show in Section V-C that relaxing the assumption makes no difference to our results). As before, let r represent the number of relays in the system.

We define a *system outage* as the event where the overall probability of error between the source and destination fails to achieve a given frame error rate criterion. To describe asymptotic outage probabilities, designated P_{out} , we will be using the Θ order notation, where $g(x) = \Theta(f(x))$ means that there exists a constant c such that $\lim_{x \rightarrow \infty} f(x)/g(x) = c$. In Rayleigh fading, a system with diversity order d has probability of system outage $P_{\text{out}} = \Theta(\bar{\gamma}^{-d})$.

For a single link in Rayleigh fading, and given a minimum SNR γ_{min} to avoid system outage, it is easy to show that P_{out} is given by

$$P_{\text{out}} = \Pr(\gamma < \gamma_{\text{min}}) = 1 - e^{-\alpha \gamma_{\text{min}} / \bar{\gamma}} \quad (8)$$

where α is a positive constant. Given the Taylor series expansion of e^x , it is easy to show that $1 - e^{-\alpha \gamma_{\text{min}} / \bar{\gamma}} = \Theta(\bar{\gamma}^{-1})$.

When using error-correcting codes, we assume that there exists an SNR $\gamma > 0$ so that the code (and its decoding algorithm) is equal or superior to using no coding for every SNR greater than γ . In other words, for sufficiently high SNR, the use of the code does not result in a higher probability of error than using no code.

With these definitions and assumptions in mind, we are now ready to present the main theoretical results of the paper.

B. Theoretical results

Our main result is stated in Theorem 1. However, we first give some preliminary results that are useful in the proof of the theorem.

Lemma 1: Let $r_c(\gamma^{(\text{SD})})$ represent the number of relays such that whenever $r < r_c(\gamma^{(\text{SD})})$, a system outage always occurs at that value of $\gamma^{(\text{SD})}$ for all values of $\gamma^{(\text{SR},i)}$, $\gamma^{(\text{RD},i)}$. Then $r_c(\gamma^{(\text{SD})})$ exists and $r_c(\gamma^{(\text{SD})}) < \infty$ for all $\gamma^{(\text{SD})} \geq 0$.

Proof: To simplify the proof, we first assume that $\gamma^{(\text{SD})} = 0$, and $\gamma^{(\text{RD},i)} = \infty$ for all i , and relax these assumptions later. Suppose any transmission received by a relay below a minimum SNR, γ_{min} , is discarded. The probability of a given symbol being selected by a given relay was defined as ν , making the probability that a given symbol is selected by a relay *not* in outage is $\nu \Pr(\gamma^{(\text{SR})} \geq \gamma_{\text{min}})$, since these are independent events. Given r relays, the probability that a given

symbol is *not* selected by *any* relay with $\Pr(\gamma^{(\text{SR})} \geq \gamma_{\min})$ is p_{nr} , where

$$p_{nr} = \left(1 - \nu \Pr(\gamma^{(\text{SR})} \geq \gamma_{\min})\right)^r. \quad (9)$$

Thus, p_{nr} is the probability that a symbol is not selected for relaying at all, or is not selected by any relays with high enough $\gamma^{(\text{SR})}$. The *equivalent relay channel* is thus (at worst) a binary symmetric channel with crossover probability corresponding to γ_{\min} (ignoring the possibility that symbols are chosen by more than one relay) concatenated with an erasure channel with erasure probability p_{nr} . The overall bit probability of error is then, at worst,

$$p_{nr} + \frac{1 - p_{nr}}{2} \operatorname{erfc}\left(\sqrt{\frac{\gamma_{\min}}{2}}\right).$$

Ignoring the error-correcting code, for any finite block length n , and assuming that the SNR is at least γ_{\min} and constant over the n symbols, the frame error rate is then at worst

$$1 - \left(1 - p_{nr} - \frac{1 - p_{nr}}{2} \operatorname{erfc}\left(\sqrt{\frac{\gamma_{\min}}{2}}\right)\right)^n. \quad (10)$$

There exists γ_{\min} large enough so that $\operatorname{erfc}(\gamma_{\min}/2) \rightarrow 0$, and from (9), for any γ_{\min} , there exists r large enough so that $p_{nr} \rightarrow 0$. So any frame error rate criterion for outage probability can be satisfied, with r_c being the satisfying value of r .

In the case $\gamma^{(\text{RD})} < \infty$, we have a system with two binary symmetric channels concatenated with an erasure channel with erasure probability p_{nr} . By a similar argument, there exists a sufficiently high γ_{\min} so that any frame error rate criterion can be satisfied. Furthermore, the value of $r_c(\gamma^{(\text{SD})})$ for any $\gamma^{(\text{SD})} > 0$ must be less than the case where $\gamma^{(\text{SD})} = 0$. Clearly, $r_c(\gamma^{(\text{SD})})$ exists, since it is finite and lower bounded by zero for all $\gamma^{(\text{SD})}$. ■

For simplicity, the proof above ignores the error correcting code, instead using a minimum SNR threshold of γ_{\min} for the “protection” of the data. In practice, for any γ_{\min} , there exists some code length (block length) n such that the data is protected. This is true as long as an adequate fraction of the bits are relayed. Thus, the lemma emphasizes the importance of *coded cooperation* to reduce r_c to a realistic value. This will be explored further in Section IV.

Let $r_c = \lim_{\gamma^{(\text{SD})} \rightarrow 0} r_c(\gamma^{(\text{SD})})$. Furthermore, since r_c is an integer, there must be some $\gamma^{(\text{SD})} > 0$ so that $r_c = r_c(\gamma^{(\text{SD})})$; so let $\gamma_0^{(\text{SD})} = \max\{\gamma^{(\text{SD})} : r_c = r_c(\gamma^{(\text{SD})})\}$ be the largest such value (thus, we have that r_c applies to every $\gamma^{(\text{SD})}$ from 0 to $\gamma_0^{(\text{SD})}$). The following lemma is a straightforward consequence of these definitions:

Lemma 2: If $r < r_c$, then by including the source-destination channel, the system diversity order is 1, and if $r = r_c$, then the system diversity order is 2.

Proof: From Lemma 1, r_c exists and is less than ∞ . Suppose $r < r_c$. Then, by the definition of r_c , the system is in outage if $\gamma^{(\text{SD})} \leq \gamma_0^{(\text{SD})}$. The fact that a minimum SNR must be maintained on the SD link implies that the diversity order is 1.

Suppose $r = r_c$. From the definition of r_c , there must exist $\gamma_{\text{suf}} > 0$, which is the minimum $\gamma^{(\text{SR},i)}$ or $\gamma^{(\text{RD},i)}$ for any

i , such that the system is not in outage when $\gamma^{(\text{SD})} < \gamma_0^{(\text{SD})}$. Then a sufficient condition for system outage is that $\gamma^{(\text{SD})} \leq \gamma_0^{(\text{SD})}$, and $\gamma^{(\text{SR},i)} < \gamma_{\text{suf}}$ or $\gamma^{(\text{RD},i)} < \gamma_{\text{suf}}$ for any i . Thus, letting $\zeta^{(\text{SR},i)} := 1 - \Pr(\gamma^{(\text{SR},i)} < \gamma_{\text{suf}})$ and $\zeta^{(\text{RD},i)} := (1 - \Pr(\gamma^{(\text{RD},i)} < \gamma_{\text{suf}}))$, we have that

$$P_{\text{out}} \geq \Pr(\gamma^{(\text{SD})} < \gamma_0^{(\text{SD})}) \left(1 - \prod_{i=1}^r \zeta^{(\text{SR},i)} \zeta^{(\text{RD},i)}\right), \quad (11)$$

where the inequality follows from the sufficient condition. Since, by assumption, the average SNR of every channel (SD, SR, and RD) is equal to $\bar{\gamma}$, the probabilities of an SNR less than γ_{suf} are all equal to (8). Substituting into (11), we have

$$\begin{aligned} P_{\text{out}} &\geq (1 - e^{-\alpha\gamma_0^{(\text{SD})}/\bar{\gamma}}) \left(1 - \prod_{i=1}^r (e^{-\alpha\gamma_{\text{suf}}/\bar{\gamma}})(e^{-\alpha\gamma_{\text{suf}}/\bar{\gamma}})\right) \\ &= (1 - e^{-\alpha\gamma_0^{(\text{SD})}/\bar{\gamma}}) (1 - e^{-2r\alpha\gamma_{\text{suf}}/\bar{\gamma}}) \\ &= \Theta(\bar{\gamma}^{-1})\Theta(\bar{\gamma}^{-1}) = \Theta(\bar{\gamma}^{-2}). \end{aligned}$$

Similarly, from the definition of r_c , there must exist $\gamma_{\text{neq}} < \infty$, which is the maximum $\gamma^{(\text{SR},i)}$ or $\gamma^{(\text{RD},i)}$ for any i , such that the system is in outage when $\gamma^{(\text{SD})} \leq \gamma_0^{(\text{SD})}$. Thus, a necessary condition for system outage is that $\gamma^{(\text{SD})} \leq \gamma_0^{(\text{SD})}$, and $\gamma^{(\text{SR},i)} < \gamma_{\text{neq}}$ or $\gamma^{(\text{RD},i)} < \gamma_{\text{neq}}$ for any i . The analysis of the necessary case follows similarly, with \leq in place of \geq in (11), which is enough to show that the diversity order is 2. ■

Since this proof depends on the definitions given in Lemma 1, and since that result applied to both demodulate-and-forward as well as decode-and-forward, this result applies to both methods as well.

We are now ready to state our main result.

Theorem 1: For a relay system with a number of relays r , where $r_c \leq r < \infty$, diversity order is given by $r - r_c + 2$. That is, each additional relay adds an order of diversity, regardless of system parameters.

Proof: We follow similarly to the proof of Lemma 2. From Lemma 1, r_c exists and is less than ∞ . From the definition of r_c , note that for each number of relays r , there exists $\gamma_{\text{suf},r} > 0$ such that if $\gamma^{(\text{SD})} < \gamma_{\text{suf},r}$ and at least $r - r_c + 1$ relays have $\gamma^{(\text{SR},i)} \leq \gamma_{\text{suf},r}$ or $\gamma^{(\text{RD},i)} \leq \gamma_{\text{suf},r}$, then an outage always occurs. Since this is a sufficient condition for system outage, and letting

$$p_{o,r} := \Pr(\gamma^{(\text{SR},i)} < \gamma_{\text{suf},r} \cup \gamma^{(\text{RD},i)} < \gamma_{\text{suf},r}),$$

we have that

$$P_{\text{out}} \geq \Pr(\gamma^{(\text{SD})} < \gamma_{\text{suf},r}) \sum_{j=r-r_c+1}^r \binom{r}{j} p_{o,r}^j (1 - p_{o,r})^{r-j}.$$

Since $\Pr(\gamma^{(\text{SD})} < \gamma_{\text{suf},r}) = \Theta(\bar{\gamma}^{-1})$, $p_{o,r} = \Theta(\bar{\gamma}^{-1})$, $(1 - p_{o,r}) = \Theta(1)$, and the binomial coefficient is constant with respect to SNR, we have that $P_{\text{out}} \geq \Theta(\bar{\gamma}^{-(r-r_c+2)})$. Similarly to the proof of Lemma 2, the necessary condition for system outage leads to $P_{\text{out}} \leq \Theta(\bar{\gamma}^{-(r-r_c+2)})$. Thus, the diversity order is $r - r_c + 2$, and the theorem follows. ■

In the statement of the theorem, we say that the diversity gains are achieved “regardless of system parameters”. Of course ν is a system parameter, and r_c is a function of ν . However,

once the system parameters are fixed, then r_c is fixed, so an additional order of diversity is gained for every relay above r_c , regardless of how small ν was fixed in the beginning.

From Theorem 1 and Lemmas 1 and 2, the diversity order depends on the system parameters *only through* r_c . Since Lemma 1 made no assumptions concerning the error-correcting code in use, it is sufficiently broad to include any code and any SNR on the source-relay link, including an SNR of ∞ , which corresponds to the case of intermediate decoding. Furthermore, it is easy to make minor modifications to Lemma 1 to include any method of choosing symbols for relaying, so long as there exists sufficiently large (but finite) r so that $p_{nr} < \epsilon$ for any $\epsilon > 0$. As well, by assumption, using the error-correcting code does not increase the probability of error on any link. Thus, Lemma 1 applies to decode-and-forward as well. Furthermore, Lemma 2 and Theorem 1 applies to decode-and-forward, since these results depend on γ_{nec} and γ_{suf} , which would simply have smaller values under decode-and-forward. Thus, we have the following:

Corollary (to Theorem 1): Consider a relaying system of the type used in Theorem 1. Consider any of the following system parameters: the strength of the code used by the transmitter, or decoder used by the receiver; the use or exclusion of intermediate decoding at the relay (due to the generality of Lemmas 1 and 2, and Theorem 1); or the method of choosing symbols for relaying (so long as there exists sufficiently large $r < \infty$ to achieve $p_{nr} < \epsilon$ for any $\epsilon > 0$). Regardless of these system design choices, Theorem 1 holds, and the impact of these three choices is reflected in the value of r_c .

Owing to this corollary, for any simplifying assumption that we can make in designing a fractional relay system, the “cost” is at most a *finite number of orders of diversity*, through the changing value of r_c ; for $r > r_c$, the property that each additional relay leads to a full order of diversity still holds. In other words, fractional relaying is a tremendously robust strategy, which is highly amenable to distributed wireless networking hardware.

IV. LOW COMPLEXITY CODED COOPERATION VIA DEMODULATE-AND-FORWARD

From Lemma 1, there is a clear motivation for implementing fractional cooperation with good error-correcting codes. From (9)-(10), to achieve any practical frame error rate, the value of r_c needs to be enormous in the absence of error-correcting codes. However, from an information-theoretic perspective, recalling that r is the number of relays and ν is the probability of any relay selecting a source symbol for retransmission, if symbols not selected for relaying are treated as erasures, then from (9), the erasure probability is $(1-\nu)^r$, and so the capacity of the equivalent relay channel is at most $1 - (1-\nu)^r$. Thus, letting R represent the rate of the SD codeword, r_c is lower-bounded by the smallest r such that $R \leq 1 - (1-\nu)^r$, which is a comparatively small number. A very good error-correcting code could come close to achieving this bound.

Again, bearing our computational constraints in mind, we seek error-correcting codes that are *powerful*, but that are also *adaptive* and have *low encoding complexity*. In this section, we specify the transformation from the information sequence \mathbf{w} to

the codeword \mathbf{c} , first using systematic low-density generator-matrix (LDGM) codes, and subsequently using punctured systematic repeat-accumulate (PSRA) codes. Although we have chosen to implement systematic codes in this paper, which have the desirable property that the encoded symbols are repeated in the transmitted codeword, any code with a sparse parity check matrix may be used. For instance, one may use low-density parity-check (LDPC) codes (which are non-systematic), though these codes have fixed block lengths and higher (though still reasonable) encoding complexity than the codes we discuss in this section [21]. However, although we emphasize the low-complexity requirements at the relay, in our system the same codes will also be used by the source, which is itself a low-complexity device.

A. Coded cooperation with LDGM codes

LDGM codes [22] are parallel concatenations of many simultaneous single parity check codes, resulting in a very sparse generator matrix. This makes encoding very simple, which is our motivation for choosing this type of code. Furthermore, these codes are closely related to LT codes [23], which have the property of being “rateless”; thus, it is also easy to adapt the rate of an LDGM code to the channel conditions. However, a negative consequence of the sparse generator matrix is a low minimum distance (and hence error floors at high SNR). To keep the encoder’s complexity low, we choose not to implement strategies for improving the error floor, such as proposed in [24].

As we mentioned, our work considers systematic LDGM codes, where the information sequence forms part of the codeword. Let $\mathbf{w} = [w_1, w_2, \dots, w_k] \in \{0, 1\}^k$ represent a k -bit binary information vector. Recall that the relay’s transmission is of length m . An arbitrary linear systematic code of length m has a $k \times m$ generator matrix \mathbf{G} of the form

$$\mathbf{G} = \left[\mathbf{I}^{(k)} \mid \mathbf{P} \right], \quad (12)$$

where $\mathbf{I}^{(k)}$ represents a $k \times k$ identity matrix, and \mathbf{P} is some $k \times (m - k)$ binary matrix. A codeword \mathbf{c} is formed by calculating

$$\mathbf{c} = \mathbf{w}\mathbf{G} \quad (13)$$

in modulo-2 arithmetic. In a systematic LDGM code, \mathbf{G} must be of low density (as the name implies), which means \mathbf{P} must also be sparse.

For the purposes of this paper, given a column weight d_j , the j th column of \mathbf{P} is formed by placing d_j ones in the column at random, uniformly distributed over all possible arrangements of d_j ones, and independently of any other column. Letting \mathbf{p}_j represent the j th column of \mathbf{P} , from (12) and (13) we can write

$$c_{k+j} = \mathbf{w}\mathbf{p}_j. \quad (14)$$

Furthermore, if $\psi_j(h)$ represents the location of the h th one in \mathbf{p}_j (for $1 \leq h \leq d_j$), we can write (14) as

$$c_{k+j} = w_{\psi_j(1)} \oplus w_{\psi_j(2)} \oplus \dots \oplus w_{\psi_j(d_j)},$$

where \oplus represents mod-2 addition. Thus, the following simple scheme generates an LDGM codeword \mathbf{c} :

- 1) For $1 \leq j \leq k$, let $c_j = w_j$.
- 2) For $k < j \leq m$, select d_{j-k} symbols at random from \mathbf{w} , and take their sum (modulo 2), yielding c_j . The identities of the selected symbols are known in advance by the decoder.

The algorithm does not change at all for different values of m , so the value of m can be chosen dynamically in response to channel conditions (unlike many other kinds of block codes which must be redesigned in order to change their length). Thus, LDGM codes satisfy both our *low complexity* and *adaptivity* criteria.

B. Coded cooperation with PSRA codes

Repeat-accumulate (RA) codes [25] are Turbo-like codes with a very simple encoding method, and have excellent performance, especially in low-SNR channels. Code rates are made adaptive by *puncturing*, i.e., by transmitting a subset of the bits from a codeword, and omitting the remainder. It is known that to retain good performance under puncturing, the codes must be *systematic* [26]. As a result, we propose to use punctured systematic repeat accumulate (PSRA) codes as part of our coded cooperation scheme. Since they rely on puncturing to change their rate, PSRA codes are not as flexible as LDGM codes, but they avoid the problem of error floors.

Once again, let $\mathbf{w} = [w_1, w_2, \dots, w_k] \in \{0, 1\}^k$ represent a k -bit binary information vector. A PSRA codeword is formed in the following four steps.

- 1) **Repeat:** For some integer $\rho > 0$, form the vector \mathbf{v} of length $k\rho$, containing ρ concatenated versions of \mathbf{w} . For example, if $\rho = 3$, then $\mathbf{v} = [\mathbf{w}, \mathbf{w}, \mathbf{w}]$.
- 2) **Permute:** Form the vector $\mathbf{v}^{(\Pi)}$ by permuting \mathbf{v} with respect to the permutation Π on $k\rho$ letters. The permutation Π is selected uniformly at random from the set of all possible such permutations, and is known by the decoder.
- 3) **Accumulate:** Form the vector of *accumulator* bits \mathbf{a} , whose elements are the running modulo-2 sum of the elements of $\mathbf{v}^{(\Pi)}$. In other words,

$$\begin{aligned} a_j &= \sum_{h=1}^j v_h^{(\Pi)} \pmod{2} \\ &= \begin{cases} v_1^{(\Pi)}, & j = 1; \\ a_{j-1} + v_j^{(\Pi)} \pmod{2}, & 1 < j \leq k\rho. \end{cases} \end{aligned}$$

From the second line, the accumulator bits satisfy $(a_j + a_{j-1} + v_j^{(\Pi)}) \pmod{2} = 0$, which is a valid *parity check*; thus, these codes are very similar to LDPC codes.

- 4) **Puncture:** This is carried out in much the same manner as the selection of symbols from \mathbf{b} , described in Section II-C. Let $\mathbf{q} \in \{0, 1\}^{k\rho}$ represent a *puncturing sequence*. Form \mathbf{a}' from \mathbf{a} by letting $a'_j := a_{h_j}$, where h_j is as defined in (7).
- 5) **Form the systematic codeword:** The codeword \mathbf{c} is then given by

$$\mathbf{c} := [\mathbf{w}, \mathbf{a}'].$$

which is of length $m = k + \sum_{t=1}^{k\rho} q_t$.

Notice that the information sequence is not punctured in forming the codeword \mathbf{c} . Given a random number generator, there exist algorithms with very low complexity for producing a random permutation (e.g., [27, p. 145]), and all the other encoding operations are simple. Thus, PSRA codes satisfy our criterion of low encoding complexity.

Furthermore, since the information rate is given by

$$R = \frac{k}{m} = \frac{k}{k + \sum_{t=1}^{k\rho} q_t},$$

the rate can be altered by changing \mathbf{u} , although once ρ is fixed, it is easy to see that the minimum rate is $1/(1 + \rho)$, unlike the LDGM code in which there is no minimum rate. Nevertheless, ρ can be changed dynamically from codeword to codeword.

C. Decoding

Codes with sparse parity check matrices, including LDGM and PSRA codes (as well as LDPC codes), have the property that their parity check matrices may be expressed using factor graphs, and the code may be decoded iteratively as message-passing over the factor graph using the sum-product algorithm [28]. We rely heavily on this property in describing the decoding algorithms for our system.

Since there is no intermediate decoding at the relays in our setup, the data sink must combine all of the relays' transmissions, and decode the corresponding codes (as well as the code on the SD link), in order to recover the data originally transmitted by the source. This could be accomplished in two ways: with *serial decoding*, in which the relays' transmissions are decoded first, followed by the source's transmissions; or by *parallel decoding*, in which the transmissions are jointly decoded at the same time.

First consider serial decoding. For ease of exposition, suppose there is only one relay, so that the observed signal $\mathbf{Y}^{(\text{RD})}$ from the relay is decoded first. This is merely a single LDGM or PSRA codeword transmitted through a Gaussian channel, which is decoded using the sum-product algorithm. Assuming the relay's transmission is decoded successfully, the destination has $\mathbf{x}^{(\text{R})}$ available, which is a noisy observation of the source transmission $\mathbf{x}^{(\text{S})}$. From (1) and (3), and recalling the definition of the relay processing function, $\mathbf{x}^{(\text{R})}$ is an observation of certain symbols of $\mathbf{x}^{(\text{S})}$ at the i th relay, with a certain probability that each symbol is inverted. For example, say the u -th symbol $x_u^{(\text{S})}$ from the source is relayed as the v -th symbol $x_v^{(\text{R})}$ (with different subscripts due to our description of the relay processing function). Due to the hard decision at the relay, there is some probability p that $x_v^{(\text{R})} \neq x_u^{(\text{S})}$, and hence

$$p \left(x_v^{(\text{R})} \mid x_u^{(\text{S})} \right) = \begin{cases} p, & x_v^{(\text{R})} \neq x_u^{(\text{S})}; \\ 1 - p, & x_v^{(\text{R})} = x_u^{(\text{S})}. \end{cases} \quad (15)$$

Now, the code used to protect the SD link is decoded with respect to the simultaneous observations $\mathbf{y}^{(\text{SD})}$, equivalent to a Gaussian channel, and $\mathbf{x}^{(\text{R})}$, equivalent to a binary symmetric channel with erasures. This is easily done under the sum-product algorithm, by calculating the channel message $\mu_{x_j^{(\text{S})}}$, corresponding to the symbol $x_j^{(\text{S})}$. This is the log-likelihood

ratio of $x_j^{(S)}$, taking into account all observations of $x_j^{(S)}$, but excluding any information from the code, and is obtained as follows:

$$\mu_{x_j^{(S)}} = \frac{a_{SD} y_j^{(SD)}}{N_0} + q_j x_{h_j}^{(R)} \log \frac{1-p}{p}, \quad (16)$$

recalling the notation for the relay processing function defined in Section II-C. To generalize this to the multiple relay case, a new term of the form of the second term in (16) can be added for each additional relay.

If the decoder fails to correctly decode any relay's codeword, that error propagates and can cause the overall decoding task to fail. This problem can be avoided using parallel decoding, in which the codes are decoded *simultaneously* at the destination on a joint factor graph, as in Figure 3. Again assuming a one-relay case, the sequences $\mathbf{x}^{(S)}$ and $\mathbf{x}^{(R)}$ are each codewords of separate codes; these codes are represented using separate factor graphs. Furthermore, the relationship between $\mathbf{x}^{(S)}$ and $\mathbf{x}^{(R)}$ is given in (15). Since, for example, $x_u^{(S)}$ and $x_v^{(R)}$ are probabilistically related, they are connected on a factor graph, and thus these pairs form connections between the factor graphs for the two codes. Furthermore, through these connections, messages are exchanged between the two code graphs. A message $\mu_{x_u^{(S)} \rightarrow x_v^{(R)}}$ is passed from $x_u^{(S)}$ to $x_v^{(R)}$, and is given (in log-likelihood ratio form) by

$$\mu_{x_u^{(S)} \rightarrow x_v^{(R)}} = \log \frac{\sum_{x_u^{(S)} \in \{+1, -1\}} p_e(x_u^{(S)}) p(x_v^{(R)} = +1 | x_u^{(S)})}{\sum_{x_u^{(S)} \in \{+1, -1\}} p_e(x_u^{(S)}) p(x_v^{(R)} = -1 | x_u^{(S)})}, \quad (17)$$

where $p_e(x_u^{(S)})$ is a message representing all the summarized information about $x_u^{(S)}$, passed from the factor graph representing the source codeword, and where $p(x_v^{(R)} | x_u^{(S)})$ is given in (16). Similarly, the message $\mu_{x_u^{(S)} \rightarrow x_v^{(R)}}$ from $x_v^{(R)}$ to $x_u^{(S)}$ is given by

$$\mu_{x_v^{(R)} \rightarrow x_u^{(S)}} = \frac{\sum_{x_v^{(R)} \in \{+1, -1\}} p_e(x_v^{(R)}) p(x_v^{(R,i)} | x_u^{(S)} = +1)}{\sum_{x_v^{(R)} \in \{+1, -1\}} p_e(x_v^{(R)}) p(x_v^{(R,i)} | x_u^{(S)} = -1)}, \quad (18)$$

where $p_e(x_v^{(R)})$ is a message representing all the summarized information about $x_v^{(R)}$, passed from the factor graph representing the relay codeword. These messages are exchanged for every relayed symbol, and the messages are combined with the channel messages at each symbol node, where the sum of all incident messages is taken. Using the factor graph, it is straightforward to include more relays in this scenario, as each relay has its own code factor graph, which is connected to the source graph in the same manner. From (18), when $x_v^{(R)}$ is perfectly known, the equation is either equal to $\log(1-p)/p$ when $x_v^{(R)} = +1$, or $\log p/(1-p)$ when $x_v^{(R)} = -1$, equivalent to the second term in (16).

As we have mentioned throughout, the decoding operations are expected to take place at the data sink, which increases the destination's computational load. However, this is assumed to be fair, in that the data sink is likely to have far more

resources than any other node. For codes decoded on graphs, it is well known that $\Theta(n)$ operations are required to decode a code of length n ; furthermore, the decoder is called upon to decode the r codes transmitted by each relay, each with length proportional to the original code, for a total computational load of $\Theta(rn)$. This level of complexity is the same as in, for example, decode-and-forward.

D. Relationship to systems in the literature

The hard decision, and selection of a subset of the hard decisions for relaying, can be thought of as a crude quantization. Quantization and compression are a strategy described in [4] for communication over the relay channel. Although this strategy is normally performed when the relay is *incapable* of decoding the source's transmission (i.e., if the source-to-relay channel's capacity is below the source's transmission rate), it represents a bound on the performance of any relay channel. In its most general form, this scheme is sometimes referred to as "estimate-and-forward" or "compress-and-forward." Bounds on achievable rates these schemes were given for full-duplex channels in [29], and for half-duplex channels in [30]. Practical work to implement such schemes has been done as an extension of the Slepian-Wolf problem [31] and the Wyner-Ziv problem [32], [33].

We also wish to emphasize the difference of our system from that of [6], [7] where a system called "decode-and-forward" is described which assumes error control *decoding* at the relay. Such a decoding is also explicitly specified by other coded cooperation schemes such as in [8] are related works; indeed, [15] is implemented similarly to our work, though without fractional cooperation, and with the requirement of decode-and-forward. The important difference from these works is our restriction of not allowing relays to decode the transmitted code word. Our relays, therefore, often transmit erroneous data; however, the reliability of these is taken into account in the factor graph as described above. A demodulate-and-forward scheme similar to ours, was proposed in [18] and then used in [34] to discuss the capacity of relay systems. Independently, demodulate-and-forward was proposed to implement cooperative diversity in sensor networks in [19]. However, our approach of re-encoding the demodulated data is not suggested in these references, nor is fractional cooperation.

V. SIMULATION RESULTS AND DISCUSSION

This paper has developed the two variants of cooperative diversity, for the specific aim of minimizing complexity in a distributed wireless network of complexity-constrained nodes. In Section III studied a network setting with r potential relays, introducing the notion of fractional cooperation wherein any relay only forwards a fraction of the source's codeword to the destination. The main theorem in Section III shows that there is a single parameter, r_c , that determines the overall diversity order of the fractionally cooperative system. In Section IV we developed the notion of cooperation via demodulate-and-forward within the context of error correcting codes (specifically LDGM and RA codes), which is useful both on its own and in the context of fractional cooperation. This section presents results of simulations undertaken to illustrate the

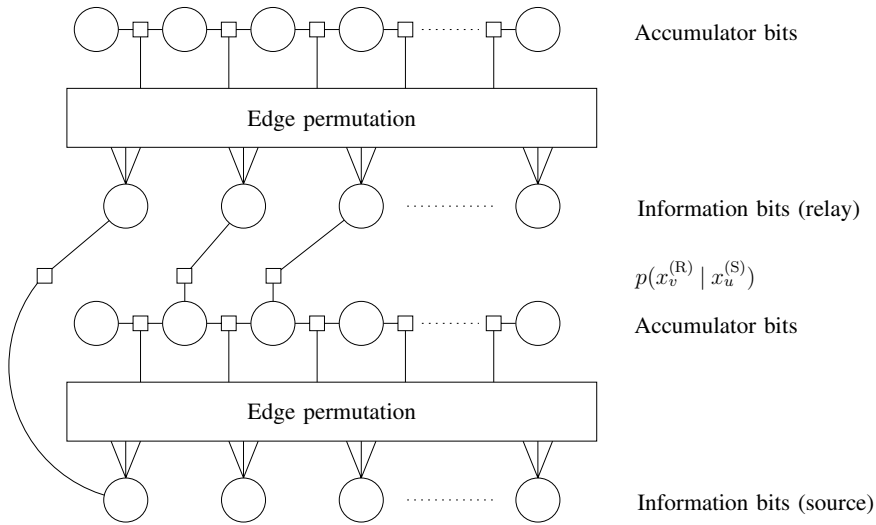


Fig. 3. Joint factor graph with RA codes. Notice that some information bits and some accumulator bits are relayed by the second code, and that not all code symbols are relayed.

efficacy of our proposals. The first set focuses on coded cooperation via demodulate-and-forward. The second set focuses on the extension to fractional cooperation, illustrating the flexibility and robustness of the proposed schemes.

Throughout this section, we use E_b to represent energy per information bit. We also use R_S to represent the code rate at the source, and R_R to represent the code rate at the relays (for convenience, the same rate is generally used at every relay). The number of information bits to be transmitted by the source is given by k_S .

A. Coded Cooperation via Demodulate-and-Forward

Our first results concern the use of single relays, to prove the effectiveness of coded cooperation with a range of possible rates and channel conditions. To evaluate the effectiveness in relaying systematic bits as opposed to parity bits, we define the new parameter ϵ_p as the fraction of source *parity* bits that are relayed. In these simulations we consider the relay forward none of the parity bits ($\epsilon_p = 0$) or all of the parity bits it demodulates ($\epsilon_p = 1$). We assume that the SNR (or average SNR, for fading channels) of every link is the same.

Figures 4 and 5 demonstrate the performance of cooperation using LDGM codes, where $k_S = 2000$ and $R_R = 1/2$, with other parameters given in the respective figures. For convenience we assume that the encoder uses a constant check degree of 10 (the robust soliton distribution [23] was attempted, but the results were not significantly better). Figure 4 plots the frame error rate (FER) for Gaussian channels for LDGM codes with different code rates and the two choices of ϵ_p . The figure shows that for higher data rates, the LDGM code results in a shallow “waterfall”. However, it is interesting that with only systematic bits relayed and with *minimal complexity* at the relay, reliable communications is possible. Figure 5 presents the FER in a Rayleigh fading channel, with similar variations in the parameters. The key observation is that, as expected, a diversity order of two is achieved, although the gain for using a low-rate code is generally small. At the cost of additional energy expended by the relay, for medium code

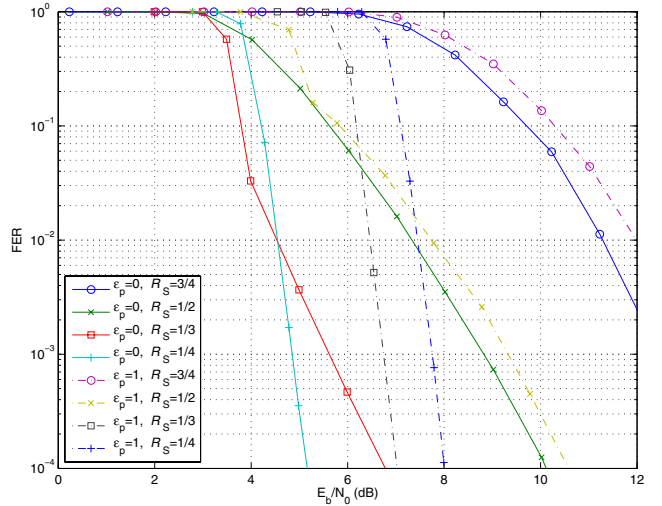


Fig. 4. Frame error rate for Gaussian channels with LDGM codes and a single relay.

rates, significant gains are achieved by forwarding both data and parity bits ($\epsilon_p = 1$ as opposed to data only, $\epsilon_p = 0$). All in all, the advantages of LDGM codes such as ease of encoding and flexibility in changing rate, are observed over the relay channels. The rate of an LDGM code can be changed easily, meaning that all of these performance curves could be achieved by a single code, where additional parity bits are generated as required.

Figures 6 and 7 demonstrate the performance of cooperation using PSRA codes, again with $k_S = 2000$ and $R_R = 1/2$. For R_S and R_R greater than or equal to $1/4$, the PSRA code was configured with a repetition of $\rho = 3$ and, if necessary, puncturing to increase the rate. Figure 6 illustrates the performance for fixed Gaussian channels, demonstrating that the PSRA code yields better performance at low SNR and has sharp waterfalls. Furthermore, for most cases a clear-cut SNR threshold is seen, above which the error rate is practically zero. Figure 7 illustrates the performance of PSRA

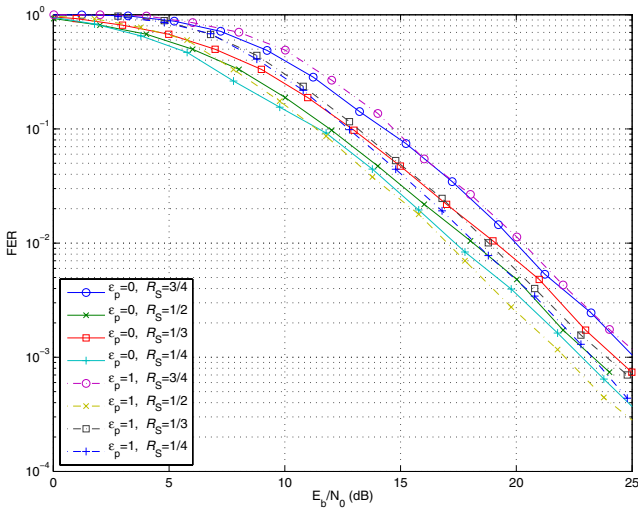


Fig. 5. Frame error rate for Rayleigh fading channels with LDGM codes and a single relay.

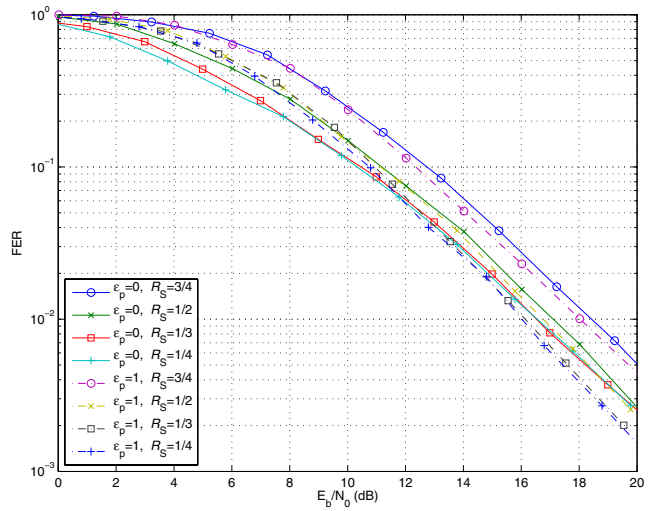


Fig. 7. Frame error rate for Rayleigh fading channels with RA codes and a single relay.

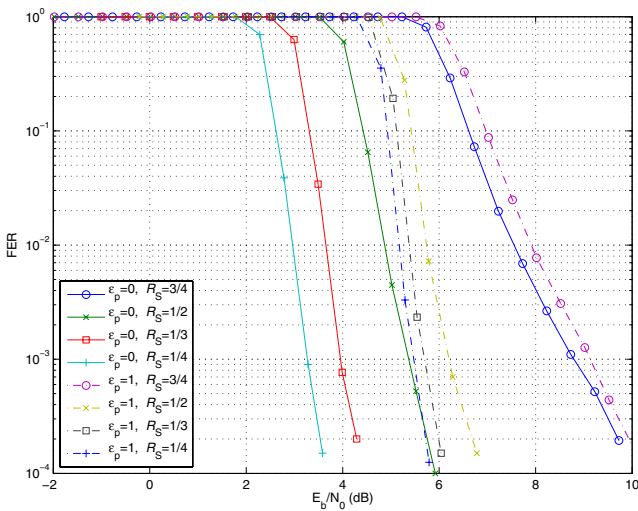


Fig. 6. Frame error rate for Gaussian channels with RA codes and a single relay.

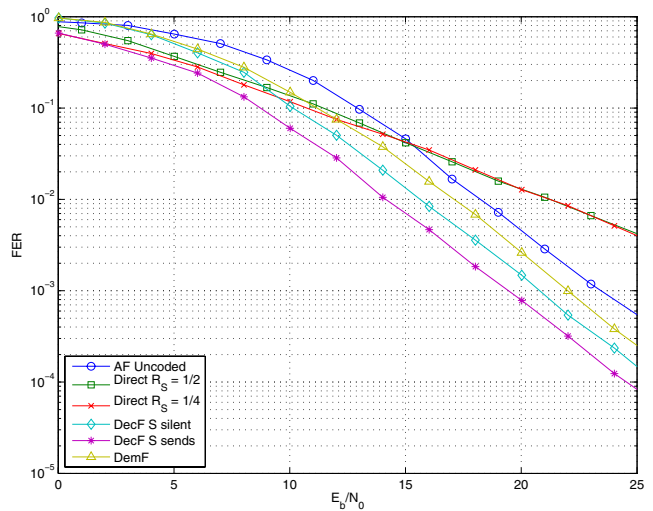


Fig. 8. Comparison of no relaying, amplify-and-forward, decode-and-forward, and demodulate-and-forward in Rayleigh fading.

codes in Rayleigh fading channels. Again, a diversity order two is observed. In both these figures, the gains in transmitting the parity bits is significantly larger than with LDGM codes. As with the LDGM case, all of these performance curves could again be achieved by a single code, where symbols were punctured or transmitted as required.

The results in Figure 8 compare our method against uncoded amplify-and-forward (AF) as well as decode-and-forward (DF) in Rayleigh fading. In these curves, all codes are PSRA codes with rate 1/2 (whether R_S or R_R), and $k_S = 2000$. We give one implementation of AF and two implementations of DF. For AF, we consider the case of no encoding at the source, where the relay merely amplifies the transmission. For DF, we consider the cases where the source is silent after its initial transmission (S silent), and where it transmits again after a decoding failure at the relay (S sends). Furthermore, we add results comparing all these methods to the case of no relaying (direct). As expected, demodulate-and-forward is better than AF and worse than DF, but the loss compared to the “S

silent” case is only around 1 dB. Furthermore, excluding DF, our method outperforms every other method at realistic frame error rates. A system designer may wish to use coded AF, in which the source employs an error-correcting code prior to transmitting its information. Under various circumstances this may be either better or worse than demodulate-and-forward; we leave the general question of which method is better for future work. Nonetheless, regardless of performance, a system designer might find AF unattractive in a practical system for implementational reasons (such as the half-duplex assumption, which may require the storage of analog information).

Figures 4-8 illustrate the ability to use simple, flexible error control codes to achieve cooperation with complexity-constrained relays using coded cooperation with demodulate-and-forward. Even though the relays do not individually decode the codeword, but only demodulate coded bits, a diversity order of two is achieved. This is because the destination is able to account for the reliability of relay bits within the decoding process. This assumes some knowledge of the source-relay

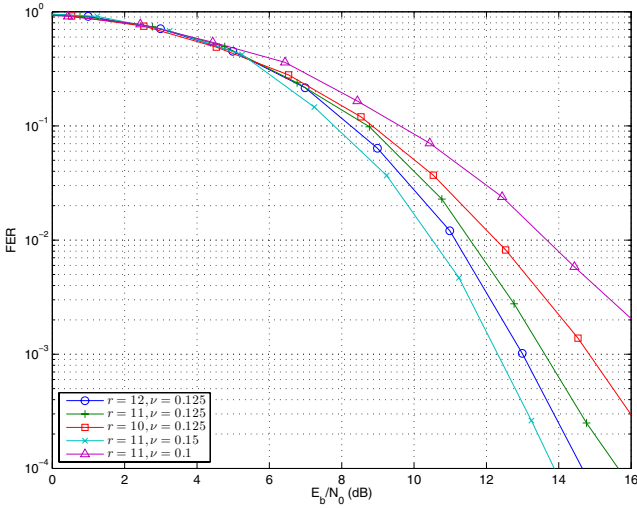


Fig. 9. Frame error rate with fractional cooperation.

channel at the destination. How the destination gets this information is an issue of importance that is acknowledged, but not addressed here.

B. Fractional Cooperation in Network Settings

We now illustrate the performance of fractional cooperation as proposed in Section III. In this extension of the notion of coded cooperation to networks with multiple relays, each relay forwards only a fraction ν of the n coded bits it demodulates; recalling our definitions from Section IV, the value of k at each relay would be νn . A critical parameter is r_c , the number of relays required to achieve successful communication without a source-destination link (Lemmas 1 and 2). The results below show that fractional cooperation is a flexible and remarkably robust scheme. These curves were generated using PSRA codes with $R_S = 1/2$ and $k_S = 2000$, where every channel has the same average SNR. All relays forward a fraction of the received coded bits. For convenience of simulation, we have assumed that successful decoding is always achieved along the RD channel, which (as we have noted) does not affect the diversity results (as a result, R_R is not specified).

Figure 9 presents results for a fractional cooperation scheme. Consider the three middle curves in the figure with $\nu = 0.125$, i.e., 500 of the 4000 coded bits are forwarded by each relay. From top to bottom, these curves exhibit diversity orders of 4, 5, and 6; with $r = 10$, $r = 11$, and $r = 12$, respectively. Thus, we conclude that $r_c = 8$. The increasing diversity order is an effective demonstration of Theorem 1. The figure also shows the case of $r = 11$ and $\nu = 0.1$ (400 out of 4000 bits forwarded) with a diversity order of 3 (implying $r_c = 10$), while with $r = 11$ and $\nu = 0.15$ (600 out of 4000 bits forwarded) a diversity order of roughly 6 is observed.

A most interesting suggestion from these curves is that the value of r_c is roughly $1/\nu$. This suggests that a minimal system using fractional cooperation should be designed such that, on average, the total number of forwarded symbols is n . Considering the erasure channel bound on r_c given at the beginning of Section IV, we should have for a good code that $r_c \simeq \log(1 - R)/\log(1 - \nu)$. Furthermore, the leading

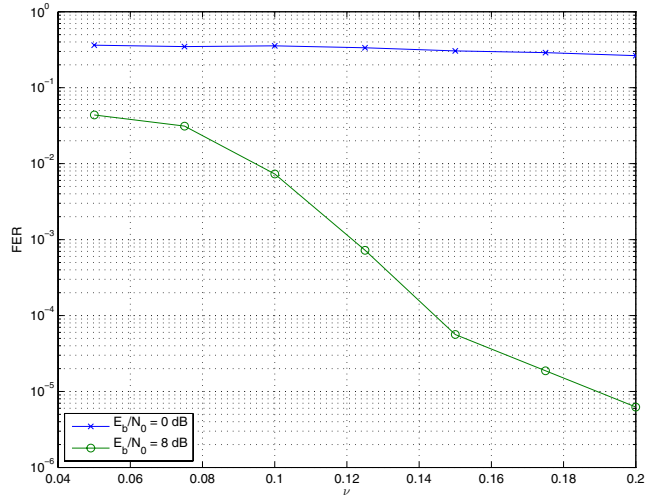


Fig. 10. Frame error rate for fractional cooperation as a function of ν .

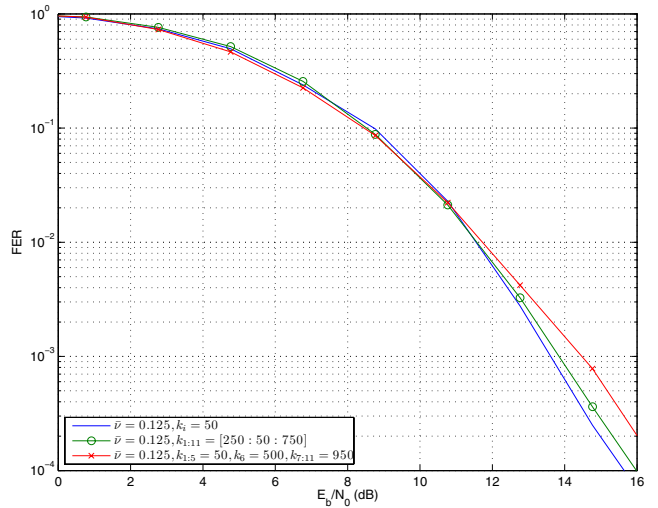


Fig. 11. Frame error rate for fractional cooperation with each relay choosing a differing ν .

term in the Taylor series expansion of $\log(1 - \nu)$ is $-\nu$ (and $\log(1 - R)$ is negative), so we should expect for a good code that, roughly speaking, $r_c \simeq -\log(1 - R)/\nu$. The fact that the results are close to this value clearly validates the use of coding in fractional cooperation, as well as our specific choice of code. However, there exists some room for code optimization.

Figure 10 plots the FER of a $r = 11$ system as a function of ν for SNR levels of 0 dB and 8 dB. At low SNR, the FER is not significantly effected by ν , but at higher SNR, where the diversity order is a factor, increasing ν (thereby decreasing r_c) significantly improves the FER. This is clearly at the expense of increased energy consumed by the relays, because larger ν implies more symbols relayed.

Figure 11 illustrates the robustness of the system to each relay choosing a different ν . This could happen due to each relay having differing energy resources, and hence be willing to give up less or more of its energy to cooperation. For a fair comparison, each curve in the figure has the same number of bits forwarded over all the relays. On average

each relay forwards 500 bits ($\nu = 0.125$). Three different cases are presented: (a) each relay forwards 500 bits; (b) the number of bits relayed by each relay ranges from 250 to 750 bits, in 50-bit increments; and (c) 5 relays forward 50 bits, 1 relay forwards 500 bits, and 5 relays forward 950 bits. The performance loss is surprisingly small, indicating the fact that fractional cooperation is very robust to the level of cooperation. The similar performance curves also illustrate that every relay (above r_c) contributes a full order of diversity, independent of how many bits it relays.

C. Discussion

We address two points of discussion raised by our work: the importance of coded cooperation to fractional cooperation, and the relaxation of assumptions.

1) *Coded fractional cooperation*: As we showed in the examples, an important result is that a *full order of diversity* is gained by going from r_c to $r_c + 1$ relays, or from $r_c + 1$ to $r_c + 2$ relays, even though each additional relay is forwarding only a small fraction of source's information. This indicates that fractional cooperation is a distributed and highly energy efficient method for achieving diversity gains.

Pairing the fractional cooperation technique with simple, but powerful, error-correcting codes, as we have done in this paper, is important to achieving good performance and strong diversity gains. From Lemma 1, very high SNRs and an enormous value of r_c is required for a fractional cooperation scheme in the absence of coding. However, from our single-relay results, we see that good performance is achieved using our coded cooperation method even at very low SNRs. Meanwhile, our result in Figure 9 indicates that coded cooperation leads to a reasonable value of r_c , $r_c \simeq 1/\nu$. Furthermore, since we have required that our coded cooperation scheme operate at varying rates, it is straightforward to change the value of ν used by each relay node, with little added complexity.

2) *Relaxation of assumptions*: For the sake of convenience, we have made three assumptions throughout this work, and here we describe how these assumptions may be relaxed without impacting the results.

Orthogonal channels. We assumed that the users suffer no co-channel interference from other nodes. There are several ways to relax this assumption (such as multiuser detection), but the most practical option would be to treat the interference as noise. Our coded cooperation scheme is designed to function even for very strong noise, which includes the worst case scenario where co-channel interference might decrease the average SNR on each link. These circumstances fall within the scope of the system we have proposed.

Coherent reception. We could use a standard technique to eliminate the ambiguity in the signal phase, such as the transmission of pilot symbols. Such techniques can be performed independently of what we suggest.

Same average SNRs. Consider the case that each link has a different average SNR, $\bar{\gamma}_i$ (e.g., based on the geographical locations of the nodes). Thus, the order of the system outage

probability is given by

$$\begin{aligned} P_{\text{out}} &= \prod_{i=1}^{r-r_c+2} \Theta(\bar{\gamma}_i) \\ &= \Theta\left(\prod_{i=1}^{r-r_c+2} \bar{\gamma}_i\right) \\ &= \Theta\left(\prod_{i=1}^{r-r_c+2} \tilde{\gamma}\right) = \Theta(\tilde{\gamma}^{r-r_c+2}), \end{aligned}$$

where $\tilde{\gamma}$ is the geometric mean of the link average SNRs, which maintains the system diversity order.

VI. CONCLUSION AND FUTURE WORK

This paper has introduced a new and practical framework for coded cooperation in wireless networks. Using fractional cooperation, we have shown that energy-efficient diversity gain can be achieved by a system that is flexible enough to be used in a distributed network. Furthermore, the derivation of fractional cooperation has made practical assumptions about the capabilities of wireless networking hardware; specifically we do not assume that a relay node can decode the source's transmission. Our strategy, of letting relays contribute only as much to the communication link as they can and splitting up the relaying task over the available relays, is particularly amenable to powerful error-correcting codes such as LDGM and PSRA codes. Considering these features, fractional cooperation based on demodulate-and-forward presents a robust and flexible option for designers of wireless networks.

Our theoretical and practical results suggest several interesting avenues for future and continuing work, such as exploring the relationship between r_c and the strength of the error-correcting code employed in fractional cooperation. Furthermore, future work will investigate *relay selection* to achieve higher energy efficiencies, as well as the practical problem of devising protocols to implement fractional cooperation.

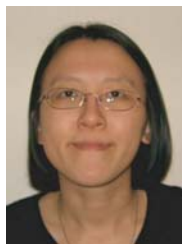
REFERENCES

- [1] C.-Y. Chong and S. P. Kumar, "Sensor networks: evolution, opportunities and challenges," *Proc. IEEE*, vol. 91, pp. 1247–1256, Aug. 2003.
- [2] P. Zhang, C. M. Sadler, S. Lyon, and M. Martonosi, "Hardware design experiences in ZebraNet," in *Proc. 2nd International Conference on Embedded Network Sensor Systems*, Baltimore, MD, pp. 227–238, Nov. 2004.
- [3] N. Xu, S. Rangwala, K. K. Chintalapudi, D. Ganesan, A. Broad, R. Govindan, and D. Estrin, "A wireless sensor network for structural monitoring," in *Proc. 2nd International Conference on Embedded Network Sensor Systems*, Baltimore, MD, pp. 13–24, Nov. 2004.
- [4] T. Cover and A. E. Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inform. Theory*, vol. 25, pp. 572–584, Sept. 1979.
- [5] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity—part I: system description," *IEEE Trans. Commun.*, vol. 51, pp. 1927–1938, Nov. 2003.
- [6] J. N. Laneman and G. W. Wornell, "Distributed space-time-coded protocols for exploiting cooperative diversity in wireless networks," *IEEE Trans. Inform. Theory*, vol. 49, pp. 2415–2425, Oct. 2003.
- [7] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: efficient protocols and outage behavior," *IEEE Trans. Inform. Theory*, vol. 50, pp. 3062–3080, Dec. 2004.
- [8] M. Janani, A. Hedayat, T. E. Hunter, and A. Nosratinia, "Coded cooperation in wireless communications: space-time transmission and iterative decoding," *IEEE Trans. Signal Processing*, vol. 52, pp. 362–371, Feb. 2004.

- [9] B. Zhao and M. C. Valenti, "Practical relay networks: a generalization of Hybrid-ARQ," *IEEE J. Select. Areas Commun.*, vol. 23, pp. 7–18, Jan. 2005.
- [10] J. Castura and Y. Mao, "Rateless coding over fading channels," *IEEE Commun. Lett.*, vol. 10, pp. 46–48, Jan. 2006.
- [11] O. Etesami and M. A. Shokrollahi, "Raptor codes on binary memoryless symmetric channels," *IEEE Trans. Inform. Theory*, vol. 52, no. 5, pp. 2033–2051, May 2006.
- [12] A. Stefanov and E. Erkip, "Cooperative coding for wireless networks," *IEEE Trans. Commun.*, vol. 52, no. 9, pp. 1470–1476, Sept. 2004.
- [13] Z. Zhang and T. M. Duman, "Capacity-approaching Turbo coding and iterative decoding for relay channels," *IEEE Trans. Commun.*, vol. 53, no. 11, pp. 1895–1905, Nov. 2005.
- [14] G. Kramer, "Distributed and layered codes for relaying," in *Proc. 39th Asilomar Conf. on Signals, Systems and Computers*, Pacific Grove, CA, pp. 1752–1756, Nov. 2005.
- [15] P. Razaghi and W. Yu, "Bilayer low-density parity-check codes for decode-and-forward in relay channels," accepted, to appear in *IEEE Trans. Inform. Theory*.
- [16] J. Hu and T. M. Duman, "Low density parity check codes over half-duplex relay channels," in *Proc. IEEE Int. Symp. Info. Theory*, Seattle, WA, pp. 972–976, July 2006.
- [17] K. Azarian, H. E. Gamal, and P. Schniter, "On the achievable diversity-multiplexing tradeoff in half-duplex cooperative channels," *IEEE Trans. Inform. Theory*, vol. 51, no. 12, pp. 4152–4172, Dec. 2005.
- [18] D. Chen and J. N. Laneman, "Modulation and demodulation for cooperative diversity in wireless systems," *IEEE Trans. Wireles Commun.*, vol. 5, no. 7, pp. 1785–1794, July 2006.
- [19] J. P. K. Chu and R. S. Adve, "Implementation of co-operative diversity using message-passing in wireless sensor networks," in *Proc. IEEE Global Communications Conference*, St. Louis, MO, pp. 1167–1171, Dec. 2005.
- [20] M. Khojastepour, A. Sabharwal, and B. Aazhang, "On the capacity of 'cheap' relay networks," in *Proc. 37th Annual Conference on Information Sciences and Systems*, Baltimore, MD, Mar. 2003.
- [21] T. J. Richardson and R. L. Urbanke, "Efficient encoding of low-density parity-check codes," *IEEE Trans. Inform. Theory*, vol. 47, no. 2, pp. 638–656, Feb. 2001.
- [22] T. R. Oenning and J. Moon, "A low density generator matrix interpretation of parallel concatenated single bit parity codes," *IEEE Trans. Magnetics*, vol. 37, no. 2, pp. 737–741, Mar. 2001.
- [23] M. Luby, "LT codes," in *Proc. 43rd Annual IEEE Symposium on Foundations of Computer Science*, Vancouver, BC, pp. 271–280, Nov. 2002.
- [24] J. Garcia-Frias and W. Zhong, "Approaching Shannon performance by iterative decoding of linear codes with low-density generator matrix," *IEEE Commun. Lett.*, vol. 7, no. 7, pp. 266–268, June 2003.
- [25] D. Divsalar, H. Jin, and R. J. McEliece, "Coding theorems for 'Turbo-like' codes," in *Proc. 36th Annual Allerton Conference on Communications, Control, and Computing*, Monticello, IL, USA, pp. 201–210, Sept. 1998.
- [26] A. Abbasfar, D. Divsalar, and K. Yao, "Accumulate repeat accumulate codes," in *Proc. IEEE International Symposium on Information Theory*, Chicago, IL, p. 505, July 2004.
- [27] D. E. Knuth, *The Art of Computer Programming, Volume 2*. Addison-Wesley, 1997.
- [28] F. R. Kschischang, B. J. Frey, and H.-A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Trans. Inform. Theory*, vol. 47, no. 2, pp. 498–519, Feb. 2001.
- [29] R. Dabora and S. D. Servetto, "On the role of estimate-and-forward with time-sharing in cooperative communication," submitted to *IEEE Trans. Inform. Theory*.
- [30] A. Chakrabarti, A. de Baynast, A. Sabharwal, and B. Aazhang, "Half-duplex estimate-and-forward relaying: bounds and code design," in *Proc. IEEE Int. Symp. Info. Theory*, Seattle, WA, pp. 1239–1243, July 2006.
- [31] R. Hu and J. Li, "Exploiting Slepian-Wolf coding in wireless relay channels," in *Proc. IEEE International Workshop on Signal Processing Advances for Wireless Communications (SPAWC)*, New York, NY, pp. 293–297, May 2005.
- [32] R. Hu and J. Li, "Practical compress-forward in user cooperation: Wyner-Ziv cooperation," in *Proc. IEEE Int. Symp. Info. Theory*, Seattle, WA, pp. 489–493, July 2006.
- [33] Z. Liu, V. Stanković, and Z. Xiong, "Practical compress-and-forward code design for the half-duplex relay channel," in *Proc. Conf. on Inf. Sci. and Systems*, Baltimore, MD, Mar. 2005.
- [34] K. S. Gomadam and S. A. Jafar, "On the capacity of memoryless relay networks," in *Proc. IEEE International Conf. on Communications*, Istanbul, Turkey, pp. 1580–1585, June 2006.



Andrew W. Eckford (M'96–S'97–M'04) is originally from Edmonton, AB, Canada. He received the B.Eng. degree in electrical engineering from the Royal Military College of Canada, Kingston, ON, Canada, in 1996, and the M.A.Sc. as well as Ph.D. degrees in electrical engineering from the University of Toronto, Toronto, ON, Canada, in 1999 and 2004, respectively. Dr. Eckford is currently an Assistant Professor of Computer Science and Engineering at York University, Toronto, ON, Canada, a post he has held since July 2006.



Josephine P. K. Chu (S'98–M'99–S'01) received the B.Sc. degree in Electrical and Computer Engineering and the M.Sc. degree in Electrical Engineering from Queens University, Kingston, Ontario, Canada in 1999 and 2003 respectively. She is currently pursuing her Ph.D. at the Department of Electrical and Computer Engineering at the University of Toronto, Toronto, Ontario, Canada. Her research interests include communications and coding, with special emphasis on low-complexity relaying techniques.



Raviraj S. Adve (S'88–M'97–SM'03) received his B. Tech. degree in Electrical Engineering from IIT, Bombay in 1990 and his Ph.D. from Syracuse University in 1996, where his dissertation received a Doctoral Prize. Between 1997 and August 2000, he worked for Research Associates for Defense Conversion Inc. on contract with the Air Force Research Laboratory at Rome, NY. He joined the faculty at the University of Toronto in August 2000. Dr. Adve's research interests include practical signal processing algorithms for smart antennas with applications in wireless communications and radar. He is currently focused on linear precoding in wireless communications, cooperative communications and augmenting space-time adaptive processing with a waveform dimension.