Chapter 1

COOPERATIVE DIVERSITY

Models, Algorithms, and Architectures

J. Nicholas Laneman
Department of Electrical Engineering
University of Notre Dame
jnl@ieee.org

Abstract Cooperative diveristy allows a collection of radio terminals that relay signals for each other to emulate an antenna array and exploit spatial diversity in wireless fading channels. For a variety of processing algorithms and transmission protocols, performance improvements in terms of transmission rate and reliability have been demonstrated. This chapter summarizes some low-complexity processing algorithms and transmission protocols for cooperative diversity, summarizes performance predictions in terms of information-theoretic outage probability, and how such constructs can be integrated into wireless network architectures, both existing and new.

Keywords: Relay channel, outage probability, spatial diversity, layered architectures, cross-layer design.

1. Introduction

In wireless communication systems, the physics of electromagnetic waves lead to multipath propagation of wireless signals and, in turn, variations in received signal strength as a function of transciever location and frequency. Combined with transciever motion, these effects produce wireless channel variations, generally called *fading*, in space, frequency, and time. Diversity techniques for mitigating, and even exploiting, multipath fading are central to improving the performance of wireless communication systems and networks. As such they are prevalent, in one form or another, in all modern wireless systems. Indeed, numerous forms of time diversity, frequency diversity, and spatial diversity are leveraged in modern systems Proakis, 2001; Rappaport, 2002.

Cooperative diversity Laneman et al., 2004; Laneman and Wornell, 2003 is a relatively new class of spatial diversity techniques that is enabled by relaying van der Meulen, 1971; Cover and El Gamal, 1979 and cooperative communications Sendonaris et al., 2003a; Sendonaris et al., 2003b more generally. As illustrated by other chapters of this book, cooperative communications refers to scenarios in which distributed radios interact to jointly transmit information in wireless environments. Cooperative diversity results when cooperative communications is used primarily to leverage the spatial diversity available among distributed radios. The main motivation here is to improve the reliability of communications in terms of, for example, outage probability, or symbol- or bit-error probability, for a given transmission rate. By contrast, as discussed in other chapters, cooperative communications can also be used primarily to increase the transmission rate. In both cases, cooperation allows for tradeoffs between target performance and required transmitted power, and thus provides additional design options for energy-efficient wireless networks.

The remainder of this chapter is organized as follows. Sec. 2 summarizes the simplest model within which cooperative diversity is well motivated, namely, wireless network environments with limited time and frequency diversity. Some low-complexity algorithms are defined and evaluated in terms of information-theoretic outage probability. Sec. 3 describes, at a high level, how such algorithms can be integrated into existing wireless network architectures such as infrastructure (cellular and wireless local area network (WLAN)) and ad hoc networks with cluters. Finally, Sec. 4 concludes the chapter with some discussion and future directions for research and development.

2. Elements of Cooperative Diversity

In this section, we summarize some of the main elements of cooperative diversity protocols and illustrate some of their performance advantages. To illustrate the issues associated with cooperative communications, we consider a single source, two relays, and a single destination as shown in Fig. 1.1. Generalizations to multi-source, and multi-stage cooperation have also been considered by a variety of authors; see, for example, Gupta and Kumar, 2003; Xie and Kumar, 2004; Kramer et al., 2005; Bölcskei et al., 2004.

Cooperative communications exploits the broadcast nature of the wireless medium and allows radios to jointly transmit information through relaying. As illustrated in Fig. 1.1(a), the two relays can receive signals resulting from the source transmission, suitably process those received signals, and transmit signals of their own so as to increase the capacity and/or improve reliability of end-to-end transmissions between the source and destination radios. Fig. 1.1(b) illustrates that relaying can be performed in multiple stages so that relays as well as the destination benefit from spatial diversity. As we will see, among other

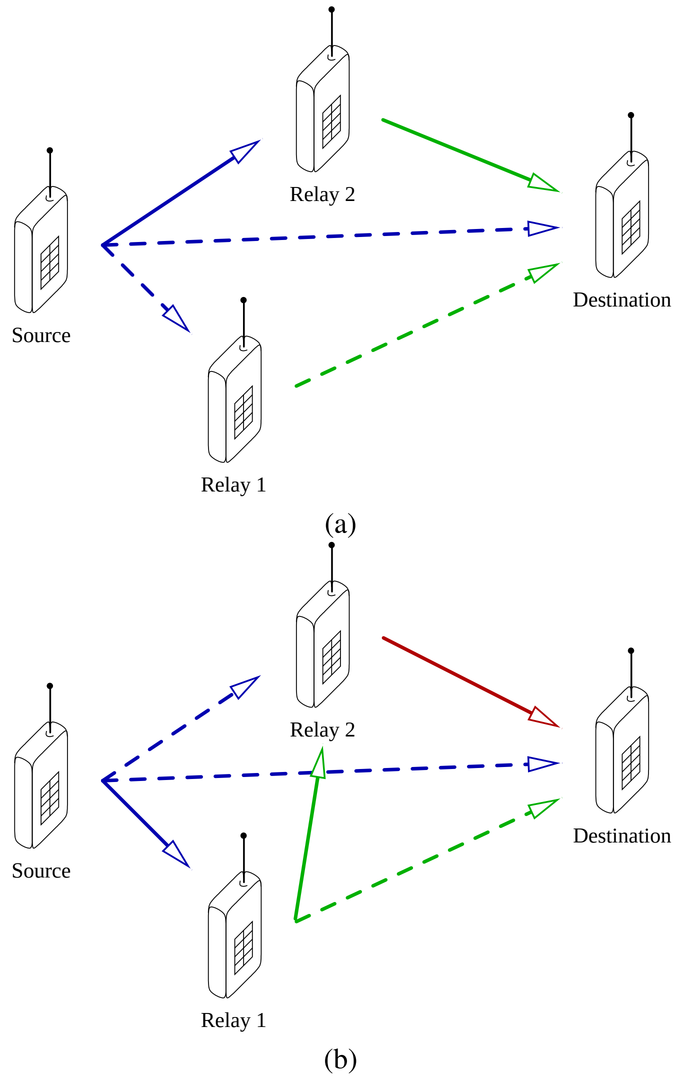


Figure 1.1. Illustration of cooperative wireless transmission with (a) parallel relays and (b) serial relays. Colors indicate transmissions that occur in different time slots or frequency bands. Solid arrows indicate transmissions that are utilized in traditional multihop transmission. Cooperative communications utilizes transmissions corresponding to both solid and dashed arrows by having the appropriate receivers perform some form of combining of their respective incoming signals.

potential benefits, cooperative communications leverages the spatial diversity available when multiple transmissions experience fading and/or shadowing that is essentially independent. For example, if the source signal experiences a deep fade at the destination, there remains a significant chance that it can be effectively communicated to the destination via one of the relays.

Because cooperative communications is inherently a network problem, issues of protocol layering and cross-layer architectures naturally arise. Starting as low as the physical layer, encoding and signal processing algorithms are required in at the source(s) and relay(s), and signal processing and decoding algorithms are required at the destination(s). However, such issues can be addressed as part of link layer coding and retransmissions, such as automatic repeat request (ARQ). Organizing the schedule for transmissions in time and frequency must be addressed by protocols in the link layer and medium-access control sublayer in coordination with the physical layer. Synchronization of signals in terms of carrier, symbol, and frame synchronization is also particularly important at the physical and link layers. Finally, collecting sets of radios into cooperating groups is inherently a cross-layer issue that can involve the physical, medium-access control, link, and even network layers. Designing an effective cooperative communication system requires insights about all of these issues. As we will further elaborate, the right combination of architecture (what logical components are identified and how they can interact) and algorithms (specific signal encoding, processing, and decoding techniques) can depend upon the application context, radio hardware, and complexity of the system.

Background

Early formulations of general relaying problems appeared in the information theory community van der Meulen, 1968; van der Meulen, 1971; Cover and El Gamal, 1979 and were inspired by the concurrent development of the ALOHA system at the University of Hawaii. The relay channel model is comprised of three terminals: a source that transmits information, a destination that receives information, and a relay that both receives and transmits information in order to enhance communication between the source and destination. More recently, models with multiple relays have been examined Schein and Gallager, 2000; Schein, 2001; Gupta and Kumar, 2003; Xie and Kumar, 2004; Khojastepour et al., 2004; Kramer et al., 2005. Cooperative communications Sendonaris et al., 2003a; Sendonaris et al., 2003b is a generalization of the relay channel to multiple sources with information to transmit that also serve as relays for each other. Combinations of relaying and cooperation are also possible, and are often referred to generically as "cooperative communications". Less well know is the fact that all of these models fall within the broader class of channels with "generalized feedback" King, 1978; Carleial, 1982; Willems, 1982; Willems et al., 1983.

Although problems of relaying and cooperation have been examined in the information theory community for years, the fundamental performance limits, in terms of the Shannon capacity or capacity region, are not known in general. Nevertheless, useful bounds on capacity have been obtained for various approaches. When applied to wireless channel models in particular, relaying and cooperation can be shown to offer significant performance enhancements in terms of various performance metrics, including: increased capacity (or larger capacity region) Sendonaris et al., 2003a; Kramer et al., 2005; Host-Madsen and Zhang, 2005; Host-Madsen, 2004; improved reliability in terms of diversity gain Laneman et al., 2004; Laneman and Wornell, 2003; Nabar et al., 2003; Mitran et al., 2005, diversity-multiplexing tradeoff Laneman et al., 2004; Laneman and Wornell, 2003; Azarian et al., 2005, and bit- or symbol-error probabilities Laneman and Wornell, 2000; Sendonaris et al., 2003b; Ribeiro et al., 2005; Chen and Laneman, 2005. These modern perspectives on and applications of relaying and cooperation have generated considerable research activity on relaying and cooperation within the communications, signal processing, and networking communities, and renewed interest within the information theory community.

System Model

This section summarizes a simple model for cooperative diversity within which later sections describe algorithms, characterize performance, and suggest network architectures. More details on system modeling are available in, for example, Laneman, 2002; Laneman et al., 2004; Laneman and Wornell, 2003; Kramer, 2004.

Cooperative diversity is motivated by a need to mitigate wireless channel effects resulting from slowly-time varying, frequency non-selective multipath fading, large-scale shadowing, and path-loss. One cost of employing relays in practical systems is that current radios cannot transmit and receive simultaneously in the same frequency band, *i.e.*, they must operate in *half-duplex* mode. In addition to power constraints, half-duplex constraints must be an integral part of the model.

More specifically, consider a network with $t \geq 2$ radios. Each radio has a baseband-equivalent, discrete-time transmit signal $X_i[k]$, with average power constraint $\sum_{k=1}^{n} |X_i[k]|^2 \leq nP_i$, and receive signal $Y_i[k]$, $i = 1, 2, \ldots, t$. Incorporating the half-duplex constraints, we model the receive signal at radio i and time sample k as Kramer, 2004

$$Y_i[k] = \begin{cases} \sum_{j \neq i}^{t} A_{i,j} X_j[k] + Z_i[k] \,, & \text{if radio } i \text{ receives at time } k \\ 0 \,, & \text{if radio } i \text{ transmits at time } k \end{cases} , \quad (1.1)$$

where $A_{i,j}$ captures the combined effects of frequency-nonselective, quasi-static multipath fading, shadowing, and path-loss between radios i and j, and $Z_i[k]$ captures the thermal noise and other interference received at radio i. Note that $A_{i,j}$ is assumed to be fixed throughout the transmission blocklength. These effects are captured in the simplest settings possible to isolate the benefits of spatial diversity.

Motivated by quasi-static conditions, we consider the scenario in which the coefficients $A_{i,j}$ are known to, *i.e.*, accurately measured by, the appropriate receivers, but not fully known to, or not exploited by, the transmitters. That is, radio i knows the realized $A_{i,j}$ but not $A_{i',j}$, for $i' \neq i$, and $j = 1, 2, \ldots, t$. Statistically, we model $A_{i,j}$ as independent complex-valued random variables, which is reasonable for scenarios in which the radios are separated by a number of carrier wavelengths. Furthermore, we model $Z_i[n]$ as zero-mean mutually independent, white circularly-symmetric, complex Gaussian random sequences with common variance N_0.

We expect that some level of synchronization between the terminals is required for cooperative diversity to be effective. For purposes of exposition, we consider the scenario in which the terminals are block, carrier, and symbol synchronous. Given some form of network block synchronization, carrier and symbol synchronization for the network can build upon the same between the individual transmitters and receivers. Although a discussion of how synchronization is achieved is beyond our scope, we note that much of the performance benefits of cooperative diversity appear to be robust to small symbol synchronization errors Wei et al., 2005 and lack of carrier phase synchronizatin Chen and Laneman, 2005.

Example Relaying Algorithms

General relaying allows sophisticated joint encoding in the transmit signals of the source(s) and relay(s) as well as intricate processing and decoding of the received signals at the relay(s) and destination(s). Since a growing number of relaying algorithms are appearing in the literature, we summarize only a few simple algorithms for illustration.

Amplify-and-Forward. For amplify-and-forward, relays simply amplify what they receive subject to their power constraint. Amplifying corresponds to a linear transformation at the relay.

Consider first the case of a single relay. The simplest algorithm described below divides transmissions into two blocks of equal duration, one block for the source transmission and one block for the relay transmission; more elaborate amplify-and-forward algorithms, as well as more general linear relaying schemes, have been considered in Nabar et al., 2003; Jing and Hassibi, 2004.

For the simplest algorithm, the source transmits $X_s[k]$ for $k = 1, 2, \ldots, n$. The relay processes its corresponding received signal $Y_r[k]$ for $k = 1, 2, \ldots, n$, and relays the information by transmitting

$$X_r[k] = \beta_r Y_r[k-n] \ , \quad k = n+1, n+2, \ldots, 2n \ . \tag{1.2}$$

To remain within its power constraint, an amplifying relay must use gain

$$\beta_r \leq \sqrt{\frac{P_s}{|A_{r,s}|^2 \, P_r + N_0}} \ , \tag{1.3}$$

where the gain is allowed to depend upon the fading coefficient $A_{r,s}$ between the source and relay. The destination processes its received signal $Y_d[k]$ for $k = 1, 2, \ldots, 2n$ by some form of diversity combining of the two subblocks of length n.

When multiple relays are active, they can each relay in their own block of channel uses so that their transmissions do not interfere at the destination, or they can relay simultaneously so that their transmissions interfere at the destination. The former approach offers better diversity benefits, but decreases bandwidth efficiency.

Decode-and-Forward. For decode-and-forward, relays apply some form of detection and/or decoding algorithms to their received signals and re-encode the information into their transmit signals. This decoding and re-encoding process often corresponds to a non-linear transformation of the received signals. Although decoding at the relays has the advantages of reducing the impact of receiver noise, as we will see, it can limit performance because of the incoming fading effects.

Again, consider first the case of a single relay. The simplest algorithm described below again divides transmissions into two blocks of equal duration, one block for the source transmission and one block for the relay transmission; more elaborate decode-and-forward algorithms have been considered in Azarian et al., 2005; Mitran et al., 2005.

For the simplest algorithm, the source transmits $X_s[k]$ for $k = 1, 2, \ldots, n$. The relay forms an estimate $\hat{X}_s[k]$ by decoding its corresponding received signal $Y_r[k]$ for $k = 1, 2, \ldots, n$, and relays a re-encoded version of $\hat{X}_s[k]$. For example, the relay can implement repetition coding by transmitting the signal

$$X_r[k] = \sqrt{\frac{P_r}{P_s}} \hat{X}_s[k-n] \ , \quad k = n+1, n+2, \ldots, 2n \ . \tag{1.4}$$

Again, the destination processes its received signal $Y_d[k]$ for $k = 1, 2, \ldots, 2n$ by some form of diversity combining of the two subblocks of length n.

Instead of repetition coding, the relay can encode the source message using a codeword that is generally correlated with, by not necessarily identical to, the source codeword. Within the context of the simple algorithms, this corresponds to a form of parallel channel coding. When multiple relays are involved, they can all employ repetition coding or a more general space-time code to transmit information jointly with the source to the destination. Like amplify-and-forward, repetition coding in separate blocks has the advantage of low complexity, but the disadvantages of scheduling and low spectral efficiency.

Selection and Dynamic Relaying. As we might expect, fixed decode-and-forward is limited by direct transmission between the source and relay. However, since the fading coefficients are known to the appropriate receivers, $A_{r,s}$ can be measured to high accuracy by the cooperating terminals; thus, they can adapt their transmission format according to the realized value of $A_{r,s}$.

This observation suggests the following class of selection relaying algorithms. If the measured $|A_{r,s}|^2$ falls below a certain threshold, the source simply continues its transmission to the destination, in the form of repetition or more powerful codes. If the measured $|A_{r,s}|^2$ lies above the threshold, the relay forwards what it received from the source, using either amplify-and-forward or decode-and-forward, in an attempt to achieve diversity gain.

Informally speaking, selection relaying of this form should offer diversity because, in either case, two of the fading coefficients must be small in order for the information to be lost. Specifically, if $|A_{r,s}|^2$ is small, then $|A_{d,s}|^2$ must also be small for the information to be lost when the source continues its transmission. Similarly, if $|A_{r,s}|^2$ is large, then both $|A_{d,s}|^2$ and $|A_{d,r}|^2$ must be small for the information to be lost when the relay employs amplify-and-forward or decode-and-forward. We formalize this notion when we consider outage performance of selection relaying in Sec. 2.0.

A further improvement of decode-and-forward is dynamic decode-and-forward Azarian et al., 2005; Mitran et al., 2005. In dynamic decode-and-forward, the relay starts by receiving from the source and does not begin transmitting until it is sure it has correctly received the source transmission. Because of quasi-static conditions, the reception time at the relay can be modeled as a random variable, and the coding scheme must take this into account.

Incremental Relaying. Fixed and selection relaying can make inefficient use of the degrees of freedom of the channel, especially for high rates, because the relays repeat all the time. In this section, we describe incremental relaying protocols that exploit limited feedback from the destination terminal, *e.g.*, a single bit indicating the success or failure of the direct transmission. These incremental relaying protocols can be viewed as extensions of incremental redundancy, or hybrid automatic-repeat-request (ARQ), to the relay context. In

ARQ, the source retransmits if the destination provides a negative acknowledgment via feedback; in incremental relaying, the relay retransmits in an attempt to exploit spatial diversity.

As one example, consider the following protocol utilizing feedback and amplify-and-forward transmission. First, the source transmits its information to the destination. The destination indicates success or failure by broadcasting a single bit of feedback to the source and relay, which we assume is detected reliably by at least the relay. If the source-destination signal-to-noise ratio (SNR) is sufficiently high, the feedback indicates success of the direct transmission, and the relay does nothing. If the source-destination SNR is not sufficiently high for successful direct transmission, the feedback requests that the relay amplify-and-forward what it received from the source. In the latter case, the destination tries to combine the two transmissions. Protocols of this form make more efficient use of the degrees of freedom of the channel, because they repeat rarely, and only when necessary.

Performance Benefits

Having described some basic relaying algorithms, we now turn to illustrating their performance. To evaluate algorithms, we utilize outage probability Ozarow et al., 1994 as a metric throughout. Because the channels are quasi-static, channel mutual informations become random variables as functions of the fading coefficients. The outage probability is then the probability that a mutual information random variable falls below some fixed rate chosen *a priori*. There are several advantages to such an information-theoretic treatment of the problem, including abstracting away many of the details of the channel coding and decoding algorithms as well as accounting for the decreased spectral efficiency required by half-duplex operation in the relays. We note that results similar to those obtained for outage probability can also be obtained for symbol- and bit-error rates of uncoded cooperative modulation and demodulation; see, *e.g.*, Laneman and Wornell, 2000; Ribeiro et al., 2005; Chen and Laneman, 2005.

Non-Cooperative Transmission. To be more precise, and for comparison with the results to follow, let us compute the outage probability of a system without cooperative diversity in the model (1.1).

Consider non-cooperative transmission from radio s to radio d. In this case, the mutual information random variable, in bits per channel use, viewed as a function of the fading coefficient $A_{d,s}$, satisfies Cover and Thomas, 1991; Telatar, 1999

$$I_{\mathrm{NC}} \leq \log\left(1 + \frac{|A_{d,s}|^2 P_s}{N_0}\right) , \tag{1.5}$$

with equality achieved for independent, identically distributed complex circular Gaussian inputs with zero-mean and variance P_s. The outage probability for rate R, in bits per channel use, is then given by Ozarow et al., 1994

$$\begin{aligned} p_{\text{out}}^{\text{NC}} &:= \Pr[I_{\text{NC}} \leq R] \\ &= \Pr\left[|A_{d,s}|^2 \leq \frac{2^R - 1}{(P_s/N_0)}\right] . \end{aligned}$$

Note that if radios s and r transmit and receive, respectively, in only k out of the n channel uses, the mutual information random variable becomes

$$I_{\text{NC}} = \frac{k}{n} \log\left(1 + \frac{n}{k}\frac{|A_{d,s}|^2 P_s}{N_0}\right) .$$

Because of the reduced number of channel uses, radio s can increase its transmitted power per channel use and remain within its average power constraint for the entire block.

Cooperative Transmission. Outage results for cooperative transmission can be obtained by extending similar results for multiple-input, multiple-output (MIMO) systems Telatar, 1999.

The simplest amplify-and-forward algorithm for a single source and relay produces an equivalent one-input, two-output conditionall complex Gaussian noise channel with different noise levels in the outputs. As Laneman et al., 2004 details, the mutual information random variable between the input and the two outputs is

$$I_{\text{AF}} = \frac{1}{2} \log\left(1 + 2\frac{|A_{d,s}|^2 P_s}{N_0} + f\left(2\frac{|A_{r,s}|^2 P_s}{N_0}, 2\frac{|A_{d,r}|^2 P_r}{N_0}\right)\right) \tag{1.6}$$

as a function of the fading coefficients, where

$$f(x, y) := \frac{xy}{x + y + 1} . \tag{1.7}$$

We note that (1.6) is achieved by i.i.d. complex Gaussian inputs, and that the amplifier gain β_r does not appear in (1.6), because the constraint (1.3) is met with equality.

For the simplest decode-and-forward algorithm with repetition coding, the mutual information random variable is

$$I_{\text{RDF}} = \frac{1}{2} \min\left\{\log\left(1 + 2\frac{|A_{r,s}|^2 P_s}{N_0}\right), \log\left(1 + 2\frac{|A_{d,s}|^2 P_s}{N_0} + 2\frac{|A_{d,r}|^2 P_r}{N_0}\right)\right\} , \tag{1.8}$$

and is achieved by i.i.d. zero-mean complex Gaussian inputs. If parallel channel coding is used instead of repetition coding, then the mutual information random variable for independent source and relay codebooks is

$$I_{\mathrm{PDF}} = \frac{1}{2} \min \left\{ \log\left(1 + 2\frac{|A_{r,s}|^2 P_s}{N_0}\right), \log\left(1 + 2\frac{|A_{d,s}|^2 P_s}{N_0}\right) + \log\left(1 + 2\frac{|A_{d,r}|^2 P_r}{N_0}\right) \right\}. \tag{1.9}$$

It is the sum of the signal-to-noise ratio random variables $|A_{i,j}|^2 P_j / N_0$ in (1.6) and (1.8) that can lead to diversity gains when compared to (1.5). Since parallel channel coding is superior to repetition coding, (1.9) can also offer diversity gains. However, in either case, the source-relay link can limit performance.

By contrast, selection decode-and-forward can be shown to offer *full diversity* regardless of the quality of the source-relay link. For the simplest selection decode-and-forward algorithm with repetition coding, the mutual information random variable is

$$I_{\mathrm{SRDF}} = \begin{cases} \frac{1}{2}\log\left(1 + 2\frac{|A_{d,s}|^2 P_s}{N_0}\right), & \text{if } \frac{1}{2}\log\left(1 + 2\frac{|A_{r,s}|^2 P_s}{N_0}\right) \leq R \\ \frac{1}{2}\log\left(1 + 2\frac{|A_{d,s}|^2 P_s}{N_0} + 2\frac{|A_{d,r}|^2 P_r}{N_0}\right), & \text{if } \frac{1}{2}\log\left(1 + 2\frac{|A_{r,s}|^2 P_s}{N_0}\right) > R \end{cases} \tag{1.10}$$

If parallel channel coding is used instead of repetition coding, then the mutual information random variable for independent source and relay codebooks is

$$I_{\mathrm{SPDF}} = \begin{cases} \frac{1}{2}\log\left(1 + 2\frac{|A_{d,s}|^2 P_s}{N_0}\right), & \text{if } \frac{1}{2}\log\left(1 + 2\frac{|A_{r,s}|^2 P_s}{N_0}\right) \leq R \\ \frac{1}{2}\log\left(1 + 2\frac{|A_{d,s}|^2 P_s}{N_0}\right) + \log(\left(1 + 2\frac{|A_{d,r}|^2 P_r}{N_0}\right), & \text{if } \frac{1}{2}\log\left(1 + 2\frac{|A_{r,s}|^2 P_s}{N_0}\right) > R \end{cases} \tag{1.11}$$

Fig. 1.2 illustrates example outage performance for non-cooperative transmission and cooperative transmission with up to two relays. The outage probabilities corresponding to the mutual informations for no cooperation (1.5), amplify-and-forward (1.6), repetition decode-and-forward (1.10), and parallel/space-time decode-and-forward (1.11) are shown. We observe from Fig. 1.2 that cooperation increases the diversity order, i.e., the negative slope of a plot of log-outage vs. SNR in dB, and provides full spatial diversity in the number of cooperating nodes, i.e., the source plus the number of relays. Although the two forms of decode-and-forward have similar performance for the case of one relay for the particular network geometry, path-loss exponent, and spectral efficiency considered, for two relays the advantages of parallel/space-time decode-and-forward are apparent in Fig. 1.2. More general analytical results in this direction are available in, e.g., Laneman et al., 2004; Laneman and Wornell, 2003.

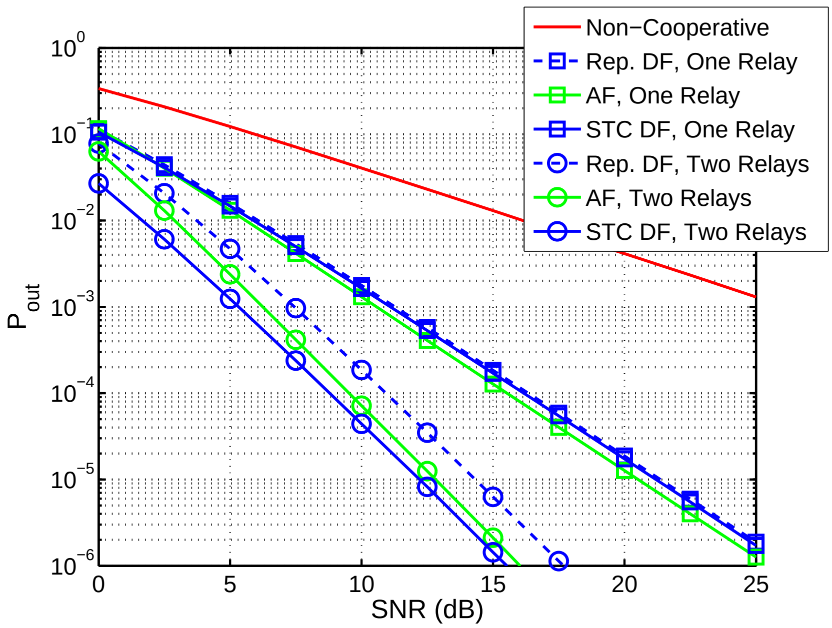


Figure 1.2. Example outage performance of non-cooperative and cooperative transmission computed via Monte Carlo simulations. The model has: path-loss with exponent $\alpha = 3$, independent Rayleigh fading, network geometry with relays located at the midpoint between the source and destination, spectral efficiency $R = 1/2$, and uniform power allocation.

3. Cooperative Diversity in Existing Network Architectures

In this section, we suggest some simple ways for integrating cooperative diversity into existing network architectures such as infrastructure networks and clustered ad hoc networks.

Centralized Partitioning for Infrastructure Networks

Our focus in this section is on infrastructure networks, in which all terminals communicate through an access point (AP). In such scenarios, the AP can gather information about the state of the network, *e.g.*, the path-losses among terminals, select a cooperative mode based upon some network performance criterion, and feed back its decision on the appropriate control channels. Here cooperative diversity lives across the medium-access control, and physical layers; routing is not considered.

Matching Algorithms. In this section, we consider grouping terminals into cooperating *pairs*. Additional studies of grouping algorithms appear in Hunter and Nosratinia, 2006; Lin et al., 2006. As we will see, choosing pairs of cooperating terminals is an instance of a more general set of problems known as *matching* problems on graphs Rosen, 2000. To outline the general matching framework, let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a graph, with $\mathcal{V}$ a set of vertices and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ a set of edges between vertices. A subset $\mathcal{M}$ of $\mathcal{E}$ is called a *matching* if edges in $\mathcal{M}$ are pairwise disjoint, *i.e.*, no two edges in $\mathcal{M}$ are incident on the same vertex. Note that

$$|\mathcal{M}| \leq \lfloor |\mathcal{V}|/2 \rfloor \ , \tag{1.12}$$

where $|\mathcal{M}|$ is the cardinality of the set $\mathcal{M}$ and $\lfloor x \rfloor$ denotes the usual floor function. When the bound (1.12) is achieved with equality, the matching is called a *perfect matching*. Since we will be working with *complete* graphs, *i.e.*, there is an edge between each pair of vertices, there will always be a perfect matching for $|\mathcal{V}|$ even. As a result, we will not be concerned with so-called *maximal matching* problems.

Instead, we focus on *weighted matching* problems. Given an edge e in $\mathcal{E}$, the *weight* of the edge is some real number $w(e)$. Given a subset $\mathcal{S}$ of $\mathcal{E}$, we denote its sum weight by

$$w(\mathcal{S}) = \sum_{e \in \mathcal{S}} w(e) \ . \tag{1.13}$$

The *minimal weighted matching* problem is to find a matching $\mathcal{M}$ of minimal weight Rosen, 2000. We also consider two other matching algorithms, both based upon randomization, that approximate minimal weighted matching and offer lower complexity.

Specifically, we consider the following algorithms:

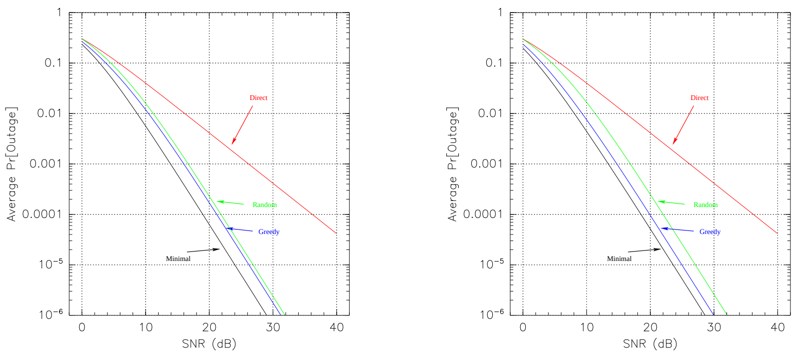
- **Minimal Weighted Matching:** Since algorithms for implementing minimal weighted matching are well-studied and readily available Ahuja et al., 1993; Rosen, 2000, we do not go into their details. We note, however, that more recent algorithms for minimal weighted matching have complexity $O(|\mathcal{V}|^3)$ Rosen, 2000.

- **Greedy Matching:** To reduce complexity and approximate minimal weighted matching, we consider a greedy algorithm in which we randomly select a free vertex v and match it with another free vertex v' such that the edge $e = (v, v')$ has minimal weight. The process continues until all of the vertices have been matched. Since each step of the algorithm takes at most $|\mathcal{V}|$ comparisons, and there are $|\mathcal{V}|/2$ steps, the complexity of this algorithm is $O(|\mathcal{V}|^2)$. We note that this greedy algorithm need not be optimal for this order of complexity.

- **Random Matching:** To reduce complexity still further, we consider a random matching algorithm where we pair vertices randomly. The complexity of this algorithm is $O(|\mathcal{V}|)$.

In addition to the algorithms outlined above for matching cooperating terminals, there are a variety of other possibilities. Instead of the general weighted matching approach, we can randomly partition the terminals into two sets and utilize *bipartite weighted matching* algorithms, which have slightly lower complexity (in terms of their coefficients, not order) and are conceptually simpler to implement than general weighted matching algorithms Rosen, 2000. Another possibility is to again randomly partition the terminals into two sets and utilize *stable marriage* algorithms with still lower complexity $O(|\mathcal{V}|^2)$. Such algorithms may be suitable for decentralized implementation Ahuja et al., 1993.

Example Performance. Fig. 1.3 shows a set of example results from the various matching algorithms described above. Terminals are independently and uniformly distributed in a square of side 2000 m, with the basestation/access point located in the center of the square. Variances for Rayliegh fading are computed using a $d^{-\alpha}$ path-loss model, with $\alpha = 3$. The weight of an edge $e = (v, v')$ is the average of the outage probabilities for terminal v using v' as a relay, and vice versa. In particular, we utilize the amplify-and-forward result (1.6) for this example; more generally, we can employ any of the outage probability expressions for a pair of cooperating terminals. Each set of results is averaged over 100 trial networks with the various matching algorithms applied. The results are normalized so that the performance of non-cooperative transmission is the same in each trial, *i.e.*, the received SNR for direct transmission averaged



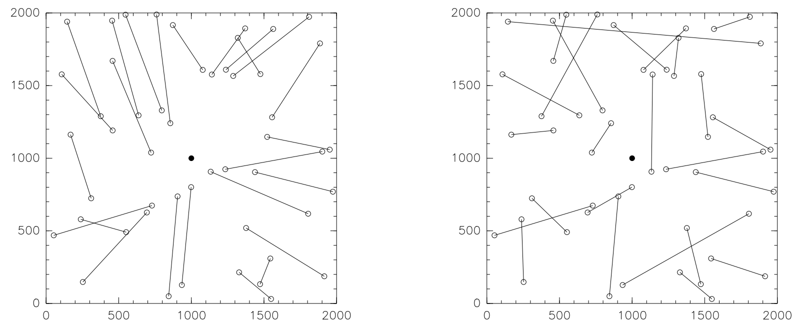
(a) $n = 10$ (b) $n = 50$

Figure 1.3. Matching algorithm performance in terms of average outage probability vs. received SNR (normalized for direct transmission).

over all the terminals in the network is normalized to be the horizontal axis in Fig. 1.3.

We note several features of the results in Fig. 1.3. First, all the matching algorithms exhibit full diversity gain of order two with respect to non-cooperative transmission. As we would expect, random, greedy, and minimal matching perform increasingly better, but only in terms of SNR gain. Although diversity gain remains constant because we only group terminals into cooperating pairs, the relative SNR gain does improve slightly with increasing network size. This effect appears most pronounced in the case of greedy matching. This observation suggests that optimal matching is more crucial to good performance in small networks, because there are fewer choices among a small number of terminals. In general, the SNR gains of the more computationally demanding matching algorithms are most beneficial in low to moderate SNR regimes where the benefits of the diversity gains are smallest. As the diversity gains increase for higher SNR, its becomes less crucial to utilize complex matching algorithms.

Fig. 1.4 compares the results of minimal and greedy matching for a sample network with 50 terminals. We see that the minimal matching tends to have pairs such that one of the terminals is almost on the line connecting the basestation and the other terminal. By comparison, the greedy matching algorithm exhibits much more randomness.



(a) Minimal Matching (b) Greedy Matching

Figure 1.4. Matching algorithm results for an example network: (a) minimal matching, (b) greedy matching. Terminals are indicated by circles, and matched terminals are connected with lines.

Clustering in Ad-Hoc Networks

Our focus in this section is on clustering in ad-hoc networks, and how cooperative diversity can be integrated into such architectures. Clustering algorithms partition a large ad-hoc network into a set of clusters, each centered around a clusterhead. Terminals communicate directly to their associated clusterhead, and routing is usually performed between clusterheads. In this sense, clustering mimics some of the features of infrastructure networks: clusters correspond to cells and clusterheads correspond to basestations. However, in ad-hoc settings the clusters and clusterheads may be varying as the network operates, the clusterheads themselves can have information to transmit, and the clusterhead network must share the wireless bandwidth.

There are many tradeoffs in the design of clustering algorithms, too many to fully address here. For example, clustering algorithms can be designed in order to reduce the complexity and overhead of routing through the network Das and Bharghavan, 1997; they can be designed in coordination with turning radios on and off in order to reduce power consumption in the network Chen et al., 2002; and they can be designed to facilitate fusion of measurements in sensor networks Heinzelman et al., 2000; Heinzelman et al., 2002.

Instead our objective in this section is to suggest how cooperative diversity can be integrated into an existing clustered ad-hoc network. To this end, we

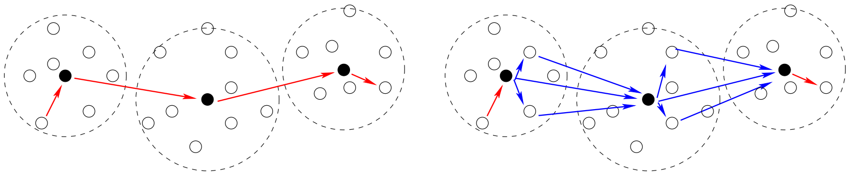


Figure 1.5. Clustering with (a) direct transmission and (b) cooperative diversity transmission.

consider three clusters as in Fig. 1.5. Fig. 1.5(a) illustrates how direct transmission can be utilized to communicate information between terminals in different clusters. A terminal transmits to its clusterhead, clusterheads route the transmission to the destination cluster, and finally the destination clusterhead transmits to the destination terminal. Fig. 1.5(b) illustrates how each of the inter-cluster direct transmissions can be converted into cooperative diversity transmissions. If the average inter-terminal spacing is $\overline{r}$, we see that inter-cluster transmissions occur roughly over a distance $2\overline{r}$ on average. There are likely to be useful relays between clusterheads, and in order to coordinate the transmissions we utilize relays in the originating cluster.

As we increase the cluster size, the number of clusters in the network decreases. Thus the complexity of routing across the entire network decreases as well; however, the utility and complexity of routing within a cluster increases. From a diversity standpoint, the inter-cluster direct transmissions are over longer and longer distances, but there are more and more potential relays to exploit. The complexity and benefits of cooperative diversity between clusters thus increases as well. Again, at least from a diversity standpoint because of reduced bandwidth efficiency and diminishing returns of diversity gains, there is no reason to grow the cluster size too large. On the other hand, growing the cluster size allows more and more terminals to be asleep when they are not transmitting and still conserve the transport of the network Chen et al., 2002. The point is, there are a variety of issues to explore here, and cooperative diversity can be one of them.

4. Discussion and Future Directions

Although it seems safe to suggest that some form of cooperative communications will emerge in future wireless networks, the final form of the network architectures and the resulting impact on performance for various applications remain unclear. In this section, we discuss some important directions for clarifying these issues.

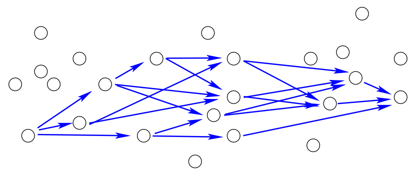


Figure 1.6. Illustration of multi-stage cooperative transmission. Downstream receivers can combine signals from all upstream transmitters. Only one complicated “link” is presented to the network layer.

Rethinking the Link Abstraction

Cooperative communications raises interesting questions about the right “link” abstraction for wireless networks. Much of the work on cooperation focuses on a relatively small number of cooperating terminals relaying in parallel, or in a single stage. As we saw in Sec. 2 and Sec. 3, such approaches can lead to performance improvements and can be readily integrated into existing network architectures by replacing individual links with single-stage cooperative links.

For multi-hop networks, one rather provacative idea Scaglione and Hong, 2003; Scaglione et al., 2006 is to replace a multi-hop route consisting of many links by a single, multi-stage cooperative link such as the one shown in Fig. 1.6. Typical multi-hop architectures for ad hoc networks present the network layer with a large number of relatively simple links, and routing over all those links becomes non-trivial. By contrast, multi-stage cooperation presents the network layer with a smaller number of more complicated links. Between these two extremes, one can imagine shifting complexity among the network, link and multi-access, and physical layers depending upon the network size and terminal processing capabilities. Although some analysis of individual multi-stage cooperative links has appeared Boyer et al., 2004; Ribeiro et al., 2005, more work should focus on how such complicated links interact in the context of a network architecture.

Hardware Testbeds

Because cooperative communications involves more than a single transmitter and a single receiver, accurate modeling and performance prediction can become challenging without many simplifying assumptions such as those considered in Sec. 2. Verifying that such assumptions are reasonable, or developing better models, requires hardware testbeds for cooperative communications.

The first, and perhaps only, public demonstration of cooperative diversity in radio hardware appears in Bletsas, 2005 as part of work performed at the MIT

Media Laboratory. Custom terminals were built on a printed circuit board consisting of a 916 MHz on-off keying radio from RF Monolithics and a low-cost Cygnal 8051 microcontroller running at 22 MHz. Software for transmission, synchronization, and reception was written from scratch.

The protocols developed within this testbed reflect the design objective of simple radio hardware. Among multiple potential relays between a source and destination, the "best" relay is selected in a distributed fashion using a request-to-send (RTS) followed by clear-to-send (CTS). The destination performs simple selection combining of the direct and relayed transmissions. Tags in the packet headers are used to display the selected relay at the destination visually with different colors. During a demonstration indoors, one can often see the colors change, indicating selection of different relays, as people and other objects move about the room, making the benefits of relaying much more tangible than a plot of outage probability against signal-to-noise ratio.

Dumb Cooperation, Smart Routing, and Distributed Beamforming

Varying amounts of channel state information (CSI) at the source and relay terminals for their channels to the destination leads to a spectrum of transmission strategies. We briefly summarize some classes of strategies here.

One class of strategies involves no transmit CSI at the source or relay terminals. That is, the source and relays do not obtain CSI on their outgoing channels, but they can obtain receive CSI on their incoming channels. This scenario is the one considered in the model in Sec. 2, which is based upon the original work on cooperative diversity in Laneman et al., 2004; Laneman and Wornell, 2003. Relative to more elaborate schemes to be discussed next, the transmission strategies can be viewed as "dumb" cooperation since they spend little overhead to obtain CSI. On the other hand, these strategies effectively waste transmit energy by having multiple relays transmit in order to achieve diversity gain at the destination.

A more informed class of strategies involves some transmit CSI at the relays, and perhaps the source. However, the key distinction between this class and the one to follow is that the CSI is used only to select a single relay. Such "opportunistic relaying" protocols, as developed within the context of the hardware testbed described above, are in some sense simpler than dumb cooperation protocols, but are more power and bandwidth efficient than dumb cooperation due to the additional CSI Bletsas et al., 2006. From a higher perspective, opportunistic relaying can be viewed as a form of "smart routing" in a local area with frequent routing table updates.

The most informed class of strategies involves complete transmit CSI at the source and relays, including both amplitude and phase information, and allows for “distributed beamforming” Sendonaris et al., 2003a; Host-Madsen and Zhang, 2005; Barriac et al., 2004; Ochiai et al., 2005. Beamforming exploits coherent combination of signals at the receiver, and can be very power efficient and bandwidth. Effective distributed beamforming requires accurate carrier synchronization among the source and relays Brown et al., 2005, but much simpler codes than those designed for dumb cooperation can be employed.

Having quickly summarized a broader spectrum of cooperative strategies beyond those of Sec. 2, we conclude by saying that it is unclear which of these approaches will have the largest impact in practical networks. For simplicity, analysis suppresses many interactions among signal processing for channel estimation, beamforming, and synchronization; channel coding for error control; and network protocols for data transfer. Overhead, particularly for dissemination of CSI and network protocols, is often not taken into account. It seems unlikely that a model general enough to evaluate all these interactions and tradeoffs will be tractable. Instead, extensive simulations and implementations within wireless testbeds will have to be pursued in order to complete the story.

Appendix: CV of Author

J. Nicholas Laneman received B.S. degrees (summa cum laude) in Electrical Engineering and in Computer Science from Washington University, St. Louis, MO, in 1995. At the Massachusetts Institute of Technology (MIT), Cambridge, MA, he earned the S.M. and Ph.D. degrees in Electrical Engineering in 1997 and 2002, respectively.

Since 2002, Dr. Laneman has been on the faculty of the Department of Electrical Engineering, University of Notre Dame, where his current research interest lie in wireless communications and networking, information theory, and detection & estimation theory. From 1995 to 2002, he was affiliated with the Department of Electrical Engineering and Computer Science and the Research Laboratory of Electronics, MIT, where he held a National Science Foundation Graduate Research Fellowship and served as both a Research and Teaching Assistant. During 1998 and 1999 he was also with Lucent Technologies, Bell Laboratories, Murray Hill, NJ, both as a Member of the Technical Staff and as a Consultant, where he developed robust source and channel coding methods for digital audio broadcasting. His industrial interactions have led to five U.S. patents.

Dr. Laneman received the MIT EECS Harold L. Hazen Teaching Award in 2001 and the ORAU Ralph E. Powe Junior Faculty Enhancement Award in 2003. He is a member of IEEE, ASEE, and Sigma Xi.

References

Ahuja, Ravindra K., Magnanti, Thomas L., and Orlin, James B. (1993). *Network Flows: Theory, Algorithms, and Applications*. Prentice-Hall, Englewood Cliffs, NJ.

Azarian, Kambiz, El Gamal, Hesham, and Schniter, Philip (2005). On the Achievable Diversity-Multiplexing Tradeoff in Half-Duplex Cooperative Channels. *IEEE Trans. Inform. Theory*, 51(12):4152–4172.

Barriac, G., Mudumbai, R., and Madhow, U. (2004). Distributed Beamforming for Information Transfer in Sensor Networks. In *Proc. International Symposium on Information Processing in Sensor Networks (IPSN)*, Berkeley, CA.

Bletsas, Aggelos (2005). *Intelligent Antenna Sharing in Cooperative Diversity Wireless Networks*. PhD thesis, Massachusetts Institute of Technology, Cambridge, MA.

Bletsas, Aggelos, Khisti, Ashish, Reed, D. P., and Lippman, A. (2006). A Simple Cooperative Diversity Method Based on Network Path Selection. *IEEE J. Select. Areas Commun.* To appear.

Bölcskei, Helmut, Nabar, Rohit U., Oyman, Özgür, and Paulraj, Arogyaswami J. (2004). Capacity Scaling Laws in MIMO Relay Networks. *IEEE Trans. Wireless Commun.* Submitted for publication.

Boyer, John, Falconer, David, and Yanikomeroglu, Halim (2004). Multihop Diversity in Wireless Relaying Channels. *IEEE Trans. Commun.*, 52(10):1820–1830.

Brown, III, D. Richard, Prince, Gregary B., and McNeill, John A. (2005). A Method for Carrier Frequency and Phase Synchronization of Two Autonomous Cooperative Transmitters. In *Proc. IEEE Workshop on Sig. Proc. Adv. in Wireless Comm. (SPAWC)*, New York, NY.

Carleial, Aydano B. (1982). Multiple-Access Channels with Different Generalized Feedback Signals. *IEEE Trans. Inform. Theory*, 28(6):841–850.

Chen, Benjie, Jamieson, Kyle, Balakrishnan, Hari, and Morris, Robert (2002). Span: An Energy-Efficient Coordination Algorithm for Topology Maintenance in Ad Hoc Wireless Networks. *ACM Wireless Networks Journal*, 8(5).

Chen, Deqiang and Laneman, J. Nicholas (2005). Modulation and Demodulation for Cooperative Diversity in Wireless Systems. *IEEE Trans. Wireless Commun.* To appear.

Cover, Thomas M. and El Gamal, Abbas A. (1979). Capacity Theorems for the Relay Channel. *IEEE Trans. Inform. Theory*, 25(5):572–584.

Cover, Thomas M. and Thomas, Joy A. (1991). *Elements of Information Theory*. John Wiley & Sons, Inc., New York.

Das, Bevan and Bharghavan, Vaduver (1997). Routing in Ad-Hoc Networks Using Minimum Connected Dominating Sets. In *Proc. IEEE Int. Conf. Communications (ICC)*, volume 1, pages 376–380, Montreal, Canada.

Gupta, Piyush and Kumar, P. R. (2003). Towards an Information Theory of Large Networks: An Achievable Rate Region. *IEEE Trans. Inform. Theory*, 49(8):1877–1894.

Heinzelman, Wendi, Chandrakasan, Anantha, and Balakrishnan, Hari (2000). Energy-Efficient Communication Protocols for Wireless Microsensor Networks. In *Proc. of the Hawaii Int. Conf. on System Sciences*, pages 3005–3014, Maui, HW.

Heinzelman, Wendi B., Chandrakasan, Anantha P., and Balakrishnan, Hari (2002). An Application-Specific Protocol Architecture for Wireless Microsensor Networks. *IEEE Trans. Wireless Commun.*, 1(4):660–670.

Host-Madsen, Anders (2004). Capacity Bounds for Cooperative Diversity. *IEEE Trans. Inform. Theory*. Submitted for publication.

Host-Madsen, Anders and Zhang, Junshan (2005). Capacity Bounds and Power Allocation in Wireless Relay Channel. *IEEE Trans. Inform. Theory*, 51(6):2020–2040.

Hunter, Todd E. and Nosratinia, Aria (2006). Coded Cooperation in Multi-User Wireless Networks. *IEEE Trans. Wireless Commun.* To appear.

Jing, Yindi and Hassibi, Babak (2004). Wireless Networks, Diversity, and Space-Time Codes. In *Proc. IEEE Information Theory Workshop (ITW)*, San Antonio, TX.

Khojastepour, M. A., Sabharwal, A., and Aazhang, B. (2004). Improved Achievable Rates for User Cooperation and Relay Channels. In *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Chicago, IL.

King, Roger C. (1978). *Multiple Access Channels with Generalized Feedback*. PhD thesis, Stanford University, Palo Alto, CA.

Kramer, Gerhard (2004). Models and Theory for Relay Channels with Receive Constraints. In *Proc. Allerton Conf. Communications, Control, and Computing*, Monticello, IL.

Kramer, Gerhard, Gastpar, Michael, and Gupta, Piyush (2005). Cooperative Strategies and Capacity Theorems for Relay Networks. *IEEE Trans. Inform. Theory*, 51(9):3037–3063.

Laneman, J. Nicholas (2002). *Cooperative Diversity in Wireless Networks: Algorithms and Architectures*. PhD thesis, Massachusetts Institute of Technology, Cambridge, MA.

Laneman, J. Nicholas, Tse, David N. C., and Wornell, Gregory W. (2004). Cooperative Diversity in Wireless Networks: Efficient Protocols and Outage Behavior. *IEEE Trans. Inform. Theory*, 50(12):3062–3080.

Laneman, J. Nicholas and Wornell, Gregory W. (2000). Energy-Efficient Antenna Sharing and Relaying for Wireless Networks. In *Proc. IEEE Wireless Comm. and Networking Conf. (WCNC)*, Chicago, IL.

Laneman, J. Nicholas and Wornell, Gregory W. (2003). Distributed Space-Time Coded Protocols for Exploiting Cooperative Diversity in Wireless Networks. *IEEE Trans. Inform. Theory*, 49(10):2415–2525.

Lin, Zinan, Erkip, Elza, and Stefanov, Andrej (2006). Cooperative Regions and Partner Choice in Coded Cooperative Systems. *IEEE Trans. Commun.* To appear.

Mitran, P., Ochia, H., and Tarokh, V. (2005). Space-Time Diversity Enhancements using Cooperative Communications. *IEEE Trans. Inform. Theory*, 51(6):2041–2057.

Nabar, R. U., Bölcskei, Helmut, and Kneubuhler, F. W. (2003). Fading Relay Channels: Performance Limits and Space-Time Signal Design. *IEEE J. Select. Areas Commun.*, 22(6):1099–1109.

Ochiai, H., Mitran, P., Poor, H. V., and Tarokh, V. (2005). Collaborative Beamforming for Distributed Wireless Ad Hoc Sensor Networks. *IEEE Trans. Signal Processing*, 53(11):4110–4124.

Ozarow, Lawrence H., Shamai (Shitz), Shlomo, and Wyner, Aaron D. (1994). Information Theoretic Considerations for Cellular Mobile Radio. *IEEE Trans. Veh. Technol.*, 43(5):359–378.

Proakis, John G. (2001). *Digital Communications*. McGraw-Hill, Inc., New York, fourth edition.

Rappaport, Theodore S. (2002). *Wireless Communications: Principles and Practice*. Prentice-Hall, Inc., Upper Saddle River, New Jersey, second edition.

Ribeiro, Alejandro, Cai, Xiaodong, and Giannakis, Georgios B. (2005). Symbol Error Probabilities for General Cooperative Links. *IEEE Trans. Wireless Commun.*, 4(3):1264–1273.

Rosen, Kenneth H., editor (2000). *Handbook of Discrete and Combinatorial Mathematics*. CRC Press, Boca Raton, FL.

Scaglione, Anna, Goeckel, Dennis, and Laneman, J. Nicholas (2006). Cooperative Communications in Mobile Ad-Hoc Networks: Rethinking the Link Abstraction. *IEEE Signal Processing Mag.* Submitted for publication.

Scaglione, Anna and Hong, Yao-Win (2003). Opportunistic Large Arrays: Cooperative Transmission in Wireless Multihop Ad Hoc Networks to Reach Far Distances. *IEEE Trans. Signal Processing*, 51(8):2082–2092.

Schein, Brett (2001). *Distributed Coordination in Network Information Theory*. PhD thesis, Massachusetts Institute of Technology, Cambridge, MA.

Schein, Brett and Gallager, Robert G. (2000). The Gaussian Parallel Relay Network. In *Proc. IEEE Int. Symp. Information Theory (ISIT)*, page 22, Sorrento, Italy.

Sendonaris, Andrew, Erkip, Elza, and Aazhang, Behnaam (2003a). User Cooperation Diversity, Part I: System Description. *IEEE Trans. Commun.*, 15(11):1927–1938.

Sendonaris, Andrew, Erkip, Elza, and Aazhang, Behnaam (2003b). User Cooperation Diversity, Part II: Implementation Aspects and Performance Analysis. *IEEE Trans. Commun.*, 51(11):1939–1948.

Telatar, I. Emre (1999). Capacity of Multi-Antenna Gaussian Channels. *European Trans. on Telecomm.*, 10(6):585–596.

van der Meulen, Edward C. (1968). *Transmission of Information in a T-Terminal Discrete Memoryless Channel*. Department of Statistics, University of California, Berkeley, CA.

van der Meulen, Edward C. (1971). Three-Terminal Communication Channels. *Adv. Appl. Prob.*, 3:120–154.

Wei, Shuangqing, Goeckel, Dennis, and Valenti, Matthew (2005). Asynchronous Cooperative Diversity. *IEEE Trans. Wireless Commun.* Submitted for publication.

Willems, Frans M. J. (1982). *Informationtheoretical Results for the Discrete Memoryless Multiple Access Channel*. PhD thesis, Katholieke Universiteit Leuven, Leuven, Belgium.

Willems, Frans M. J., van der Meulen, Edward C., and Schalkwijk, J. Pieter M. (1983). An Achievable Rate Region for the Multiple Access Channel with Generalized Feedback. In *Proc. Allerton Conf. Communications, Control, and Computing*, pages 284–292, Monticello, IL.

Xie, Liang-Liang and Kumar, P. R. (2004). An Achievable Rate for the Multiple Level Relay Channel. In *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Chicago, IL.