Chapter 1

COOPERATIVE COMMUNICATIONS

Fundamental Limits and Practical Implementation

Arnab Chakrabarti
Rice University
arnychak@rice.edu

Ashutosh Sabharwal
Rice University
ashu@rice.edu

Behnaam Aazhang
Rice University
aaz@rice.edu

Abstract This chapter summarizes theoretically achievable gains and the construction of practical codes for user-cooperation. Most of these results relate to the *relay* channel, which is a three-terminal channel that captures the essence of user-cooperation and serves as one of the primary building blocks for cooperation on a larger scale. In investigating the fundamental limits of relaying, we present information-theoretic results on the achievable throughput of relay channel in mutual-information terms. We also include results on Gaussian channels, and for the practically important case of half-duplex relaying. In the domain of relay coding, we specifically discuss pragmatic code constructions for half as well as full-duplex relaying, using LDPC codes as components.

Keywords: wireless communication, user cooperation, relay, broadcast, multiple-access, decode-and-forward, estimate-and-forward, amplify-and-forward, information theory, coding, LDPC, max-flow min-cut

Introduction

Cooperative communication is one of the fastest growing areas of research, and it is likely to be a key enabling technology for efficient spectrum use in future. [1] The key idea in user-cooperation is that of resource-sharing among multiple nodes in a network. The reason behind the exploration of user-cooperation is that willingness to share power and computation with neighboring nodes can lead to savings of overall network resources. Mesh networks provide an enormous application space for user-cooperation strateges to be implemented. In traditional communication networks, the physical layer is only responsible for communicating information from one node to another. In contrast, user-cooperation implies a paradigm shift, where the channel is not just one link but the network itself. The current chapter summarizes the fundamental limits achievable by cooperative communication, and also discusses practical code constructions that carry the potential to reach these limits.

Cooperation is possible whenever the number of communicating terminals exceeds two. Therefore, a three-terminal network is a fundamental unit in user-cooperation. Indeed, a vast portion of the literature, especially in the realm of information theory, has been devoted to a special three-terminal channel, labeled the *relay* channel. The focus of our discussion will be the relay channel, and its various extensions. In contrast, there is also a prominent portion of literature devoted to cooperation as viewed from a network-wide perspective, which we will only briefly allude to.

Our emphasis is on user-cooperation in the domain of wireless communication, and the fundamental limits that we discuss are information theoretic in nature. In this regard, we first bound the achievable rates of relaying using mutual information expressions involving inputs and outputs of the cooperating nodes. We then investigate relaying in the context of Gaussian channels, and summarize known results for well-known relaying protocols. In recent years, half-duplex relaying has been accepted as a practical form of relaying that has potential for implementation in near future. Therefore, we devote a section to the derivation of the fundamental limits of half-duplex relaying. Last, we consider a scenario where the source and the relay exchange roles, which is a departure from the conventional relay channel. This departure, however, captures the essence of user-cooperation where both nodes stand to gain from sharing their resources, which is why this model is a prominent candidate for future implementation.

As regards the coding strategies, we will discuss practical code constructions that emulate random coding strategies used in information theoretic achievability proofs. The component codes of choice are LDPC (Low Density Parity Check) codes, because of their simple factor graph representations, and low-complexity belief propagation decoding. We present code constructions for both half and full-duplex Gaussian relay channels. Many practical challenges encountered in relay coding are exposed in the course of our treatment.

This chapter is organized as follows. First, we present a historical summary of important contributions in the field of relaying. Following that, we include a section to introduce preliminary concepts and terminology for the reader who is unfamiliar with the literature. The next section is devoted to a discussion of information-theoretic limits on the throughput achievable by relaying. In this regard, we pay special attention to Gaussian links; discuss limits of half-duplex relay communication; and finally investigate a scenario where two nodes cooperate with each other without any separate notion of one node being a source and another a relay. The next section is devoted to explicit code constructions for the relay channel, and here we discuss LDPC codes for both full and half-duplex relaying. The final section concludes with a few closing remarks.

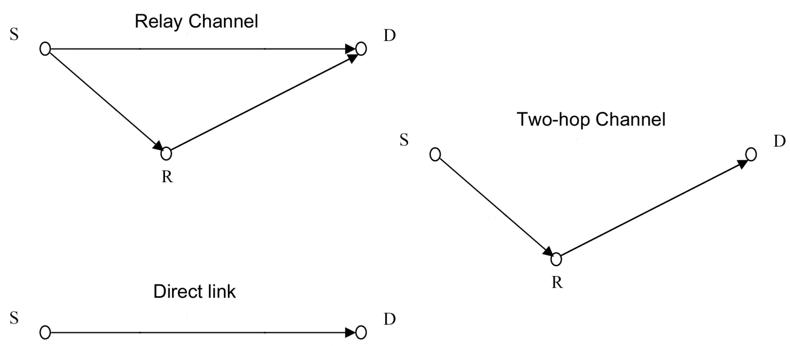

Figure 1.1. Direct, two-hop and relay communications.

A Brief History of Relaying

We summarize prominent contributions in the area of user-cooperation. Our list of contributions is by no means exhaustive, but we attempt to touch upon the many approaches to user-cooperation over the years.

Communication from a single source to a single destination without the help of any other communicating terminal is called *direct* , *single-user* or *point-to-point* communication (Figure 1.1). User-cooperation is possible whenever there is at least one additional node willing to aid in communication. The simplest and oldest form of user-cooperation is perhaps multi-hopping , which is nothing but a chain of point-to-point links from the source to the destination (Figure 1.1 shows two-hop communication). No matter what the channel, there is some attenuation of the signal with distance, which makes long-range point-to-point communication impractical. This problem is overcome by replacing a single long-range link with a chain of short-range links, where at each intermediate node there is a *booster* or *repeater* to enhance signal quality. Multi-hopping was conceived about the same time as smoke and drum signals, therefore we do not attempt to put a time stamp on it.

More recently, the three-terminal relay channel (depicted in Figure 1.1) was introduced by van der Meulen, 1968; van der Meulen, 1971. In his original work, van der Meulen discovered upper and lower bounds on the capacity of the relay channel, and made several observations that led to improvement of his results in later years. The capacity of the general relay channel is still unknown, but the bounds discovered by van der Meulen were improved significantly by Cover and El Gamal, 1979. In the interim, Sato, 1976 also looked at the relay channel in the context of the Aloha system. Notably, an extensive review of results on several channels that are important to network information theory was published in van der Meulen, 1977. The review summarized the state-of-the-art at that time, but our understanding of relaying has improved considerably since then. Other important contributions of the era which contributed to the understanding of user-cooperation include Slepian and Wolf, 1973; Gaarder and Wolf, 1975; Cover and Leung, 1981; Willems, 1982; Cover, 1972; Cover, 1975; Bergmans and Cover, 1974; Marton, 1979; Gel'fand and Pinsker, 1980; Han, 1981; El Gamal and van der Meulen, 1981; Cover et al., 1980; Wyner, 1978; Wyner and Ziv, 1976.

Undoubtedly, the most prominent work on relaying to date is Cover and El Gamal, 1979. Most of the results in this work have still not been superseded. In the years following Cover and El Gamal, 1979, there was some interest in the relay channel, as is evident from the literature. In El Gamal and Aref, 1982, the authors discovered the capacity of the semideterministic relay channel, where the received signal at the relay is a deterministic function of the source and relay transmissions. There was an effort to generalize the results of Cover and El Gamal, 1979 to networks with multiple relays in Aref, 1980; El Gamal, 1981. These works also investigated deterministic relay networks with no interference, and deterministic broadcast relay networks .

In parallel with the effort on relaying, there was a prominent body of research on the capacity of the multiple-access channel with generalized feedback (MACGF) . This channel was studied in King, 1978 with a model where two transmitters transmit to a common destination, and these transmitters also receive a common feedback from the destination. In Carleial, 1982, this model was generalized to include different feedback to the two transmitters. It is easy to see that the relay channel is a special case of Carleial's model. Remarkably, as discussed in Kramer et al., 2005, Carleial introduced a coding scheme that is different from, and in some respects preferable to the superposition block-Markov encoding introduced by Cover and El Gamal, 1979.

Perhaps due to the difficulty of finding new and better information-theoretic results, and the technological challenges of implementing user-cooperation, the interest in relaying and user cooperation diminished after the early 80's. Until the turn of the century, there were sporadic contributions on relaying, broadcast, and multiple-access channels as evidenced in Zhang, 1988; Zeng et al., 1989; Thomas, 1987. Efforts on relay coding continued until the late 80's and early 90's as evidenced in Ahlswede and Kaspi, 1987; Kobayashi, 1987; Vanroose and van der Meulen, 1992. On the other hand, some truly remarkable strides were made during this period in the general area of digital and wireless communications, such as discovering the capacity of multi-antenna systems by Foschini and Gans, 1998; Telatar, 1999, a great deal of advancement in our understanding of fading channels (summarized in Biglieri et al., 1998), and remarkable progress in channel coding including the discovery of Turbo codes in Berrou et al., 1993, space-time codes in Tarokh et al., 1998, and the rediscovery of LDPC codes of Gallager, 1963 in MacKay, 1999; Luby et al., 2001; Richardson and Urbanke, 2001. These advances set the stage for a second wave of research on relaying by providing a whole new context and new tools to attack the problem.

One of the prominent works that helped to draw attention to user-cooperation in recent years is Sendonaris et al., 2003a; Sendonaris et al., 2003b. In this work, the authors propose user-cooperation as a form of diversity in a mobile uplink scenario, and show its benefits using various metrics. Also noteworthy are the contributions of Laneman, 2002; Laneman and Wornell, 2003; Laneman et al., 2004 for studying the performance of important relaying protocols in fading environments. Yet another important set of contributions came in the form of novel information theoretic results and new insights into information theoretic coding in Kramer et al., 2005 (also more recently in Chong et al., 2005). In Schein and Gallager, 2000; Schein, 2001 the authors considered a variation of the relay channel where there is no direct source-destination link, but there are two relays to aid communication. In Schein and Gallager, 2000; Schein, 2001 the authors considered a variation of the relay channel where there is no direct source-destination link, but there are two relays to aid communication. New information theoretic results and results on power control were also discovered in Wang et al., 2005; Høst-Madsen and Zhang, 2005. A variety of contributions to relaying including new bounds, cut-set theorems , power control strategies, LDPC relay code designs, and some of the earliest results on half-duplex relaying inxxrelay,half-duplex were proposed in Khojastepour, 2004. Researchers realized that relaying can mimic multiple-antenna systems even when the communicating entities were incapable of supporting multiple antennas. Prominent literature on the use of space-time codes with relays includes Laneman and Wornell, 2003; Nabar et al., 2004; Mitran et al., 2005. Other noteworthy recent contributions are by El Gamal et al., 2004; Reznik et al., 2004; Hasna and

Alouini, 2003; Boyer et al., 2004; Toumpis and Goldsmith, 2003; Liang and Veeravalli, 2005.

In a different direction, Gupta and Kumar, 2000 proposed a new approach towards finding network information carrying capacity, which led to research on finding scaling laws for wireless networks in a variety of settings. Numerous works were published on studies of networks with large numbers of nodes, as contrasted to the simple few-node channels studied by traditional information theory. In Gupta and Kumar, 2003, the authors showed that the use of advanced multi-user schemes can improve network transport capacity significantly. Subsequently, Xie and Kumar, 2004 discovered an achievable rate expression for a degraded Gaussian channel with multiple relays and established bounds on its transport capacity. The results of Reznik et al., 2004 also treated the case of multiple Gaussian degraded relay stages with a total average power constraint. In another direction, Gastpar and Vetterli, 2005 showed that the upper-bounds on relay capacity obtained from cut-set theorems coincide with known lower bounds as the number of relays becomes large. Other prominent contributions in this area include Xue et al., 2005; Xie and Kumar, 2005; Grossglauser and Tse, 2002.

With significant advances in technology over the last two decades, the promise of relaying is very real. A large body of research is currently geared towards developing practical user-cooperation schemes to harvest the gains predicted by information theory. Solutions in this direction include Sendonaris et al., 2003b; Stefanov and Erkip, 2004; Janani et al., 2004; Hunter et al., 2004; Khojastepour et al., 2004a; Chakrabarti et al., 2005a; Zhang et al., 2004; Zhang and Duman, 2005; Castura and Mao, 2005; Zhao and Valenti, 2003.

Yet another area of user cooperation where recent years have seen an explosion of publications is that of *network coding* . The field grew largely after the publication of Ahlswede et al., 2000, although an earlier publication Yeung and Zhang, 1999 also contained seeds of the idea. Subsequently, several important advancements to the field have been made, of which some of the fundamental ones are in Li et al., 2003; Koetter and Medard, 2003; Chou et al., 2003; Jaggi et al., 2005; Ho et al., 2003; Li and Li, 2004; Yeung et al., 2005.

Preliminaries of Relaying

The relay channel is the three-terminal communication channel shown in Figure 1.2. The terminals are labeled the source (S), the relay (R), and the destination (D). All information originates at S, and must travel to D. The relay aids in communicating information from S to D without actually being an information source or sink. The signal being transmitted from the source is labeled X. The signal received by the relay is V. The transmitted signal from the relay is W, and the received signal at the destination is Y. Several notions

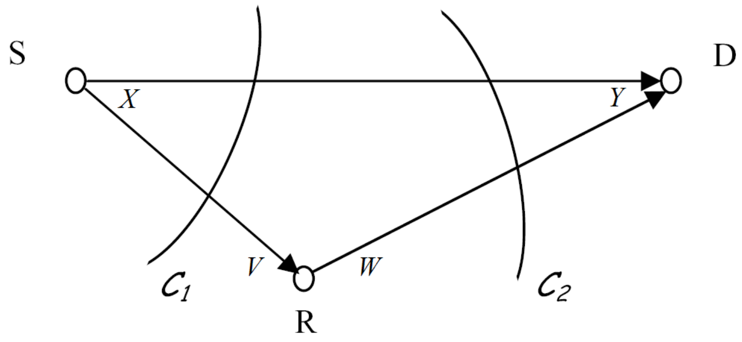

Figure 1.2. The relay channel with three nodes: the source S, the relay R, and the destination D. These three nodes are conceptually divided into two subsets by two cuts of interest: C_1 or the broadcast cut which separates S from $\{R, D\}$, and C_2 or the multiple-access cut, which separates $\{S, R\}$ from D. The channel input at S is given by X, the input at R is W, and the outputs at R and D are V and Y respectively.

of relaying exist in the literature. We will list the prominent ones in this section. Conceptually, information is relayed in two phases or modes : first, when S transmits and (R, D) receive, commonly called the broadcast (BC) mode ; and second when (S, R) transmit and D receive, also known as the multiple-access (MAC) mode . Note that this differentiation is only conceptual since it is possible for communication in both modes to take place simultaneously. We will elaborate on this a little later, but first we will enumerate four different models of relaying that can be classified based on the above two modes.

1 $S \rightarrow (R, D)$; $(S, R) \rightarrow D$ (most general form of relaying);

2 $S \rightarrow R$; $(S, R) \rightarrow D$ (D ignores signal from S in first mode);

3 $S \rightarrow (R, D)$; $R \rightarrow D$ (S does not transmit in second mode);

4 $S \rightarrow R$; $R \rightarrow D$ (multi-hop communication) .

Of these, the first model is the most general, and most early results on relaying were based on the first model. The second and the third are simplified models introduced mainly for analytical tractability. For example, they simplify the analysis of outage probabilities and the design of space-time codes for fading relay channels in Laneman et al., 2004; Nabar et al., 2004.

The last model of relaying is much older as well as simpler than the other three, and is commonly known as multi-hop communication. Unlike the other three models, multi-hop communication does not yield diversity benefits, and it is primarily used to combat signal attenuation in long-range communication. In wireless communication, usually there is severe attenuation of signal power with distance. This attenuation is characterized by a channel attenuation exponent γ. In other words, if the transmitted power is P, then the received power at a distance d is $\frac{P}{d^\gamma}$. The value of γ lies in the range of 2 to 6 for most wireless channels. This attenuation makes long-range communication virtually impossible. The simplest solution to this problem is to replace a single long-range link with a chain of short-range links by placing a series of nodes in between the source and the destination. A distinguishing feature of multi-hopping is that each node in this chain communicates only with the one before and the one after in the chain, or nodes that are one "hop" away. In a wireless environment, it may be possible for a node to receive or transmit its signal to other nodes that are several hops away, but such capability is ignored in multi-hopping, making it a simple and extremely popular, but suboptimal mode of user-cooperation. Of all the modes of user-cooperation discussed in this chapter, multi-hopping is the only one that is widely implemented today.

Half-duplex versus Full-duplex Relaying. A relay is said to be half-duplex (or 'cheap' as in Khojastepour et al., 2003) when it cannot simultaneously transmit and receive in the same band. In other words, the transmission and reception channels must be orthogonal. Orthogonality between transmitted and received signals can be in time-domain, in frequency domain, or using any set of signals that are orthogonal over the time-frequency plane. If a relay tries to transmit and receive simultaneously in the same band, then the transmitted signal interferes with the received signal. In theory, it is possible for the relay to cancel out interference due to the transmitted signal because it knows the transmitted signal. In practice, however, any error in interference cancellation (due to inaccurate knowledge of device characteristics or due to the effects of quantization and finite-precision processing) can be catastrophic because the transmitted signal is typically 100-150dB stronger than the received signal as noted in Laneman et al., 2004. Due to the difficulty of accurate interference cancellation, full-duplex radios are not commonly used; however, dvances in analog processing could potentially enable full-duplex relaying.

Although early literature on information theoretic relaying was based almost entirely on full-duplex relaying (eg. van der Meulen, 1971; Cover and El Gamal, 1979), in recent years a lot of research, and especially research directed towards practical protocols, has been based on the premise of half-duplex relaying(eg. Khojastepour et al., 2003; Liang and Veeravalli, 2005; Janani et al., 2004; Laneman et al., 2004; Nabar et al., 2004).

Relay Protocols. The capacity of the general relay channel of Figure 1.2 is not known even today, over thirty years after the channel was first proposed.

Moreover, there is no single cooperation strategy known that works best for the general relay channel. As we will discuss in a subsequent section on fundamental limits, there are at least two fundamental ideas (and a third that is practically less important) based on which the source and relay nodes can share their resources to achieve the highest throughput possible for any known coding scheme. The cooperation strategies based on these different ideas have come to be known as *relay protocols* .

The first idea involves decoding of the source transmission at the relay. The relay then retransmits the decoded signal after possibly compressing or adding redundancy. This strategy is known as the *decode-and-forward* protocol , named after the fact that the relay can and does decode the source transmission. The decode-and-forward protocol is close to optimal when the source-relay channel is excellent, which practically happens when the source are relay are physically near each other. When the source-relay channel becomes perfect, the relay channel becomes a 2×1 multiple-antenna system. Following the naming convention of Cover and El Gamal, 1979, some authors use the term *cooperation* to strictly mean the decode-and-forward type of cooperation.

The second idea, sometimes called *observation* , is important when the source-relay and the source-destination channels are comparable, and the relay-destination link is good. In this situation, the relay may not be able to decode the source signal, but nonetheless it has an independent observation of the source signal that can aid in decoding at the destination. Therefore, the relay sends an estimate of the source transmission to the destination. This strategy is known as the *estimate-and-forward* (also known as *compress-and-forward* or *quantize-and-forward*) protocol .

The *amplify-and-forward* (also sometimes called *scale-and-forward*) protocol is a special case of the above strategy where the estimate of the source transmission is simply the signal received by the relay, scaled up or down before retransmission. A 1×2 multi-antenna system is a relay channel where amplify-and-forward is the optimal strategy, and the amplification factor is dictated by the relative strengths of the source-relay and source-destination links.

The third idea, known as *facilitation* , is mostly of theoretical interest. When the relay is not able to contribute any new information to the destination, then it simply tries to stay out of the way by transmitting the signal that would be least harmful to source-destination communication.

The names for the protocols that we have described above are generally accepted by the relaying community. However, the reader is cautioned that some authors refer to the aforementioned protocols differently. For example, in Khojastepour, 2004, *scale-and-forward* and *amplify-and-forward* refer to different schemes. Therefore, it is always a good idea to check the authors' definitions of scientific terms used in a document.

Relaying : Fundamental Limits

The following is a brief outline of this section. First, we will summarize well-known information theoretic results on the full-duplex relay channel stated in terms of mutual information expressions. Second, we will present the achievable rates of Gaussian channels for various relay protocols. Following that, we will discuss results on half-duplex relaying. Finally, we will briefly summarize the results of Sendonaris et al., 2003a; Sendonaris et al., 2003b for a three-terminal network where each node acts as both source, and (full-duplex) relay for the other node. A discussion on fundamental limits of relaying cannot be complete without results on fading channels; however, they will be treated in a separate chapter by Nicholas Laneman.

The relay channel is shown in Figure 1.2. We will assume that the channel is discrete-time and memoryless . The signals X,V,W, and Y are chosen from finite sets $\mathcal{X}$,$\mathcal{V}$,$\mathcal{W}$, and $\mathcal{Y}$ respectively, and the channel is described by the conditional probability densities $p(v, y|X = x, W = w)$ on $(\mathcal{V} \times \mathcal{Y})$ for all $(x \times w) \in (\mathcal{X} \times \mathcal{W})$. In what follows, we briefly summarize important results known for the relay channel in terms of mutual information expressions. Many of these results are due to Cover and El Gamal, 1979.

For the special case of a degraded relay channel, i.e., when $X \rightarrow (V, W) \rightarrow Y$ is a Markov chain , the following theorem from Cover and El Gamal, 1979 gives the capacity of the relay channel.

THEOREM 1.1 *[Cover and El Gamal, 1979] The capacity C_d of the degraded relay channel is given by*

$$C_d = \sup_{p(x,w)} \min \left(I(X, W; Y), I(X; V|W) \right). \tag{1.1}$$

The rate of Theorem 1.1 is achievable for any relay channel (not necessarily degraded). The premise of degradedness is used only to prove that this rate cannot be surpassed, and is therefore the capacity when the channel is degraded. The practical utility of Theorem 1.1 stems from the fact that it provides an achievable rate (a lower bound on the capacity) for the general relay channel. This lower bound is fairly tight if the source-relay (SR) channel is better than the relay-destination (RD) channel, which may physically correspond to a scenario where R is closer to S than to D. This can be attributed to the fact that the relay channel resembles a degraded channel more and more closely as the SR link improves relative to the RD link. A 2×1 multi-antenna system can be thought of as a degraded relay channel, where the two transmitter antennas corresponding to the source and the relay have a perfect communication channel in between them.

Achieving the rate of Theorem 1.1 requires a coding scheme where R decodes the signal it receives from S before passing it on to D. Therefore, Theorem 1.1

corresponds to the achievable rate of the *decode-and-forward* relaying protocol. We will discuss practical code designs for this protocol in a subsequent section on relay coding.

For a reversely degraded relay channel i.e., when $X \rightarrow (Y, W) \rightarrow V$ is a Markov chain, the capacity of the relay channel is given by the following theorem in Cover and El Gamal, 1979.

THEOREM 1.2 *[Cover and El Gamal, 1979] The capacity C_{rd} of the reversely degraded relay channel is given by*

$$C_{rd} = \max_{w \in \mathcal{W}} \max_{p(x)} I(X; Y|w). \tag{1.2}$$

Theorem 1.2 advocates *facilitation* , where the relay transmits a signal that will maximize the capacity of the SD link. Since R receives a signal that is a noisy version of what D receives, R knows nothing that D does not already know - "thus w is set constantly at the symbol that "opens" the channel for the transmission of x, directly to y at rate $I(X; Y|w)$"as quoted from Cover and El Gamal, 1979. As in the case of Theorem 1.1, the achievability of the rate in Theorem 1.2 does not require the reverse degradedness assumption, and is true for all relay channels. This theorem is usually not important in practice because practical relay channels that resemble reversely degraded channels have small capacity. A channel where the SR distance is much larger than the SD distance would mimic reverse degradedness. Geometric constraints dictate that R will be far from both S and D in such a scenario, and it is therefore intuitive that such a relay will not be of much use.

For the general relay channel of Figure 1.2, the following upper bound is due to Cover and El Gamal, 1979.

THEOREM 1.3 *[Cover and El Gamal, 1979] The capacity C of the general relay channel is bounded above as follows*

$$C \leq \sup_{p(x,w)} \min \left(I(X, W; Y), I(X; V, Y|W) \right). \tag{1.3}$$

The above upper bound is a consequence of a general cut-set theorem for information flow in networks. The reader should refer to Page 444 of Cover and Thomas, 1991 for an extended discussion of this theorem. Here, we will briefly define notation, and state this elegant and powerful theorem without proof.

A network with multiple terminals is shown in Figure 1.3. We closely follow the conventions in Cover and Thomas, 1991. The network consists of N nodes. Node i is characterized by the input-output pair $(X^{(i)}, Y^{(i)})$, and sends information at a rate $R^(ij)$ to node j. The nodes are divided in two sets S

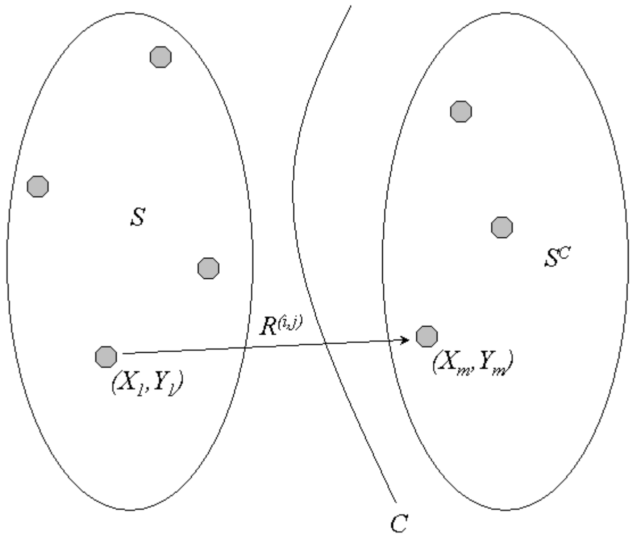

Figure 1.3. A network with multiple nodes divided in two sets S and S^C separated by a cut C.

and S^C (the complement of S), and a cut C conceptually separates the nodes in these two sets. The channel is represented by the conditional probability mass function $p(y^{(1)}, y^{(2)}, ..., y^{(N)}|x^{(1)}, x^{(2)}, ..., x^{(N)})$ of the outputs given the inputs. The following theorem bounds the rate of information transfer from nodes in S to those in S^C by the conditional mutual information.

THEOREM 1.4 *[Cover and Thomas, 1991] If the information rates* $\{R^{(ij)}\}$ *are achievable, then there exists a joint input probability distribution* $p(x^{(1)}, x^{(2)}, ..., x^{(N)})$ *such that*

$$\sum_{i \in S, j \in S^C} R^{(ij)} \leq I\left(X^{(S)}; Y^{(S^C)} | X^{(S^C)}\right) \tag{1.4}$$

for all $S \subset 1, 2, ..., N$.

The above theorem is analogous to the well-known *max-flow min-cut* theorem Ford and Fulkerson, 1962. It says that the rate of information flow across any cut (boundary) dividing the set of nodes in a network in two parts (the transmitter and receiver sides) cannot exceed the mutual information between the channel inputs on the transmitter side and the channel outputs on the receiver side conditioned on the knowledge of inputs on the receiver side.

The aforementioned cut-set theorem is a powerful tool for bounding the capacity region from above for several important multi-terminal communication channels. Capacity results in network information theory are usually difficult to prove, and in many important cases such as the multiple-access channel and the degraded relay channel, the capacity of a channel is known because the upper bound given by the max-flow min-cut theorem is achievable. Unfortunately, the upper bound given by the cut-set theorem is not always achievable. There are examples where the capacity has been shown to be smaller than what is indicated by the cut-set bound. One of the prominent cases is that of the Gaussian vector broadcast channel (GVBC) (see Caire and Shamai, 2003; Vishwanath et al., 2003; Yu and Cioffi, 2004; Viswanath and Tse, 2003), where the upper bound on sum capacity is derived based on the work of Sato, 1978, and achievability has been shown using Costa, 1983; Marton, 1979. Nevertheless, the cut-set theorem remains one of the most powerful tools in network information theory.

Notably, the above cut-set theorem has recently been extended to networks with multiple states by Khojastepour, 2004. We will discuss the extension in the context of half-duplex relaying in a subsequent section.

Returning to our discussion of the relay channel, the reader will observe that Theorem 1.3 follows immediately from Theorem 1.4. For the general relay channel, the upper bound of Theorem 1.3 is not known to be achievable. However, it *is* achievable when feedback is available, as shown in Theorem 3 of Cover and El Gamal, 1979. In a relay channel with feedback, S as well as R know the signal received by D (with unit delay, which does not reduce capacity). In fact, a relay channel with feedback is a degraded relay channel where V is replaced with (Y, V), therefore using Theorem 1.1, the rate of (1.3) is achievable. Here, the authors would like to caution the reader that results based on feedback in information theory can be misleading because this feedback does not carry a cost. In practice, feedback, like all other information, must be communicated, and will therefore consume a fraction of the channel capacity.

Apart from decode-and-forward (Theorem 1.1) and facilitation (Theorem 1.2), there is at least one other idea in cooperation that yields useful achievable rates, exceeding those of Theorem 1.1 in some scenarios. The idea corresponds to the estimate-and-forward protocol . Cover and El Gamal realized that in between the two cases of a degraded channel, where the relay can perfectly decode the source signal, and the reversely degraded channel, where the relay can contribute no new information, there exists a regime where the relay can contribute partial information to improve decoding at the destination. This partial information is in the form of an estimate of the source signal received by the relay. As an example, suppose that both SR and SD channels are AWGN links having the same noise variance. The channel is neither degraded nor reversely degraded, but since the relay and the destination have independent views of the noise, the destination would benefit from knowing the received signal at the relay, or an estimate of it. The achievable rate for this strategy is given by the following theorem

THEOREM 1.5 *[Cover and El Gamal, 1979] The rate R_{cf} is achievable for a general relay channel where*

$$R_{cf} = \sup I(X; Y, \widehat{V}|W) \tag{1.5}$$

subject to the constraint

$$I(W; Y) \geq I(V; \widehat{V}|W, Y) \tag{1.6}$$

where the supremum is over all joint distributions of the form $p(x, w, y, v, \widehat{v}) = p(x)p(w)p(y, v|x, w)p(\widehat{v}|v, w)$ on $\mathcal{X} \times \mathcal{W} \times \mathcal{Y} \times \mathcal{V} \times \widehat{\mathcal{V}}$ and $\widehat{V}$ has a finite range.

The above theorem introduces the idea of an estimate $\widehat{V}$ of V that is communicated from R to D when decoding is not possible. The final theorem (Theorem 7) in Cover and El Gamal, 1979 superimposes the decode-and-forward and compress-and-forward strategies to yield a composite achievable rate that reduces to the achievable rates of the two aforementioned strategies in special cases. We do not present the theorem here, but the interested reader can refer directly to Cover and El Gamal, 1979.

In addition to the above results, all of which are from Cover and El Gamal, 1979, there are few other results known for special cases of the relay channel. The capacity of the semideterministic relay channel, where the channel output at the relay is a deterministic function of both source and relay inputs, was derived in El Gamal and Aref, 1982 using Theorem 7 of Cover and El Gamal, 1979 and Theorem 1.3 presented above. In addition, two new results were discovered in Khojastepour, 2004 using a new set of tools developed in the context of half-duplex relay channels. These tools include a new cut-set theorem (presented as Theorem 1.11 in this chapter) as well as a novel coding scheme introduced in Khojastepour et al., 2002b; Khojastepour, 2004. We will present these results next.

THEOREM 1.6 *[Khojastepour, 2004] If the relay channel transition function can be written in the form*

- $p(y, v|x, w) = p((y_1, y_2), v_1|(x_1, x_2), w_2) = p(y_1|v_1)p(v_1|x_1)p(y_2|x_2, w_2)$ *then the capacity of the relay channel is*

$$C_1 = \sup_{p(x_1)p(x_2,w_2)} \min \Big(I(X_1; V_1) + I(X_2; Y_2|W_2), \\ I(X_1; Y_1) + I(X_2, W_2; Y_2) \Big); \tag{1.7}$$

- $p(y, v|x, w) = p((y_1, y_2), v_1|(x_1, x_2), w_2) = p(y_1, v_1|x_1)p(y_2|x_2, w_2)$ *then the capacity of the channel is*

$$C_2 = \sup_{p(x_1)p(x_2,w_2)} I(X_1; Y_1) + I(X_2, W_2; Y_2). \tag{1.8}$$

subject to the constraint

$$I(X_1; V_1) + I(X_2; Y_2|W_2) \geq I(X_1; Y_1) + I(X_2, W_2; Y_2). \tag{1.9}$$

The above two results were derived in the context of the half-duplex relay channel, therefore the subscripts 1 and 2 for broadcast and multiple-access modes respectively. However, the results are equally applicable to the full-duplex relay channel.

In addition to the above results, new and better achievable rates channel have been recently reported in Chong et al., 2005 based on two new coding strategies that combine ideas of decode-and-forward and compress-and-forward. The encoding and decoding differ from that of Cover and El Gamal, 1979 in that regular block-Markov sperposition encoding and backward decoding Willems, 1982 are used.

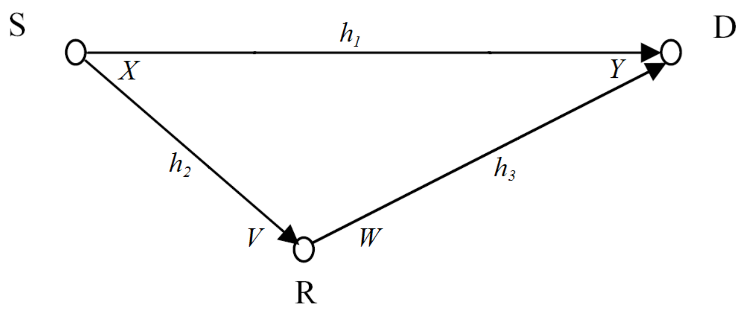

Figure 1.4. Gaussian relay channel

Capacity Results for the Gaussian Relay Channel. In this section, we discuss information theoretic results for Gaussian relay channels . We first introduce our notation and state the well-known upper bound that can be derived from Theorem 1.3. Then we discuss several lower bounds including some recently discovered ones. The lower bounds are of practical interest because every lower bound is a consequence of a constructive coding scheme, and carries the potential for implementation in future.

A full-duplex Gaussian relay channel is shown in Figure 1.4. For this discussion, we will use the notation of Khojastepour, 2004. The dependence of channel outputs on inputs is as follows: the output at the relay is $v = h_2x + n_1$, and the output at the destination is $y = h_1x + h_3w + n_2$. Here, h_1, h_2 and h_3 are the SD, SR and RD channel gains respectively, and n_1 and n_2 are Gaussian noise with variance N_1 and N_2 respectively. The three variables $\gamma_1 = \frac{|h_1|^2}{N_2}$, $\gamma_2 = \frac{|h_2|^2}{N_1}$, and $\gamma_3 = \frac{|h_3|^2}{N_2}$ denote the SD, SR, and RD channel SNRs respectively as shown in Figure 1.4. The input power constraints are given by $E[X^2] \leq P_1$ and $E[W^2] \leq P_2$.

The following upper bound on the Gaussian relay capacity can be obtained from Theorem 1.3.

THEOREM 1.7 *[Cover and El Gamal, 1979]*

$$C_{AWGN} \leq \max_{\rho:0\leq\rho\leq 1} \min \Big(C((1-\rho^2)(\gamma_1+\gamma_2)P_1), C(\gamma_1 P_1 + \gamma_3 P_2 + 2\rho\sqrt{\gamma_1\gamma_3 P_1 P_2}) \Big) \quad (1.10)$$

$$\textit{where} \quad C(x) = \frac{1}{2}\log(1+x). \quad (1.11)$$

The parameter ρ in the above equation has the physical interpretation of the correlation between source and relay signals. Increasing ρ in equation (1.9) increases the mutual information term corresponding to the multiple-access cut, which is the second argument of the $\min$ function above. At the same time, increasing ρ decreases the mutual information term corresponding to the broadcast cut, which is the first argument of the $\min$ function above. This interpretation of ρ is also true for several of the lower bounds that we present below.

Perhaps the most prominent lower bound on the capacity of the Gaussian relay channel is a consequence of Theorem 1.1. This lower bound was actually derived in Theorem 5 of Cover and El Gamal, 1979 as the capacity of the Gaussian degraded relay channel . Using our notation, we present the lower bound in the following theorem

THEOREM 1.8 *[Cover and El Gamal, 1979]*

$$C_{AWGN} \geq R_{DF} = \max_{\rho:0\leq\rho\leq 1} \min \Big(C((1-\rho^2)\gamma_2 P_1), C(\gamma_1 P_1 + \gamma_3 P_2 + 2\rho\sqrt{\gamma_1\gamma_3 P_1 P_2}) \Big) \quad (1.12)$$

where $C(x)$ is defined in (1.11).

Since the above lower bound coincides with the capacity of the degraded AWGN relay channel, we do not expect it to be tight when the channel is far from being degraded, for example when the SD link received SNR exceeds that of the SR link. In this scenario, even the capacity of the direct link $C(\gamma_1 P_1)$ exceeds that of decode-and-forward relaying. This is true even if unbounded power is available at the relay. This is certainly not the best that we can do with the relay. For instance, consider the case where the SR and SD channels have identical SNRs, and the RD channel is noiseless. Here, the relay channel becomes a 1×2 MIMO system, where the best strategy is for the relay to simply forward its analog received signal with appropriate power, so that the destination effectively receives a signal that is the result of maximal-ratio combining. In general, depending on the amount of noise in the RD channel, the relay may spend variable amounts of resources on sending an estimate of its received signal to the destination. Estimate-and-forward relaying, as proposed in Cover and El Gamal, 1979, does not explicitly state the nature of the estimate to be forwarded. Achievable rates of the estimate-and-forward protocol have been derived for the Gaussian channel in Khojastepour et al., 2004b; Kramer et al., 2005 using ideas of source coding at the relay and decoding with side information at the destination Wyner and Ziv, 1976. The following theorem gives the achievable rate of estimate-and-forward relaying for the Gaussian relay channel of Figure 1.4.

THEOREM 1.9 *[Khojastepour et al., 2004b] The achievable rate of estimate-and-forward relaying is given by*

$$C_{AWGN} \geq R_{EF} = C\left(\gamma_1 P_1 + \frac{\gamma_2 P_1 \gamma_3 P_2}{1 + \gamma_2 P_1 + \gamma_1 P_1 + \gamma_3 P_2}\right) \quad (1.13)$$

where $C(x)$ is defined in (1.11).

The above theorem shows that the rate achievable by estimate-and-forward relaying can always surpass that of the direct link, although not always exceed that of the decode-and-forward protocol.

An achievable rate for the estimate-and-forward protocol has also been derived in Sabharwal and Mitra, 2005. Yet another rate is known to be achievable using retransmission of a quantized version of the received signal at the relay. The rate is presented as the achievable rate of the quantize-and-forward protocol in Chapter 3 of Khojastepour, 2004.

The following theorem of Khojastepour, 2004 gives the rate of amplify-and-forward relaying under the name of scale-and-forward relaying. Conceptually, amplify-and-forward is a simple (but not always the best) way of implementing the principle of Theorem 1.5. Here, the estimate being forwarded is simply a scaled version of the received signal. We divide the transmission into $L \to \infty$ consecutive sub-blocks of length M. Furthermore, assume that the relay builds

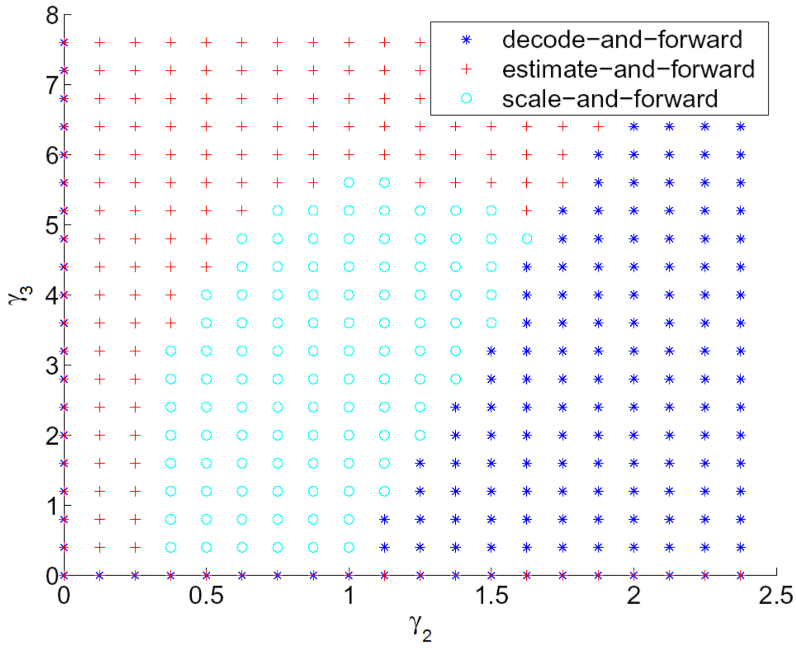

Figure 1.5. Regions where each protocol outperforms all others ($P_1 = P_2 = -10dB, \gamma_1 = 1$).

its input signal based on all previously received signals in each sub-block. For the simple case of $M = 2$, the following theorem gives the achievable rate

THEOREM 1.10 *[Khojastepour, 2004] The optimal achievable rate of the scale-and-forward scheme with the sub-block length M = 2 for the Gaussian relay channel is given by*

$$C_{AWGN} \geq R_{SF(M=2)} = \max_{Q_1,Q_2:Q_1+Q_2\leq 2P_1;\beta\leq\frac{2P_2}{h_2^2Q_1+N_1}} \frac{1}{2}C\left(\frac{h_1^2Q_1}{N_2} + \frac{h_1^2Q_2+\beta h_2^2h_3^2Q_1}{N_2+\beta h_3^2N_1} + \frac{h_1^2Q_1Q_2+\beta h_2^2h_3^2N_2^2}{N_2(N_2+\beta h_3^2N_1)}\right) \quad (1.14)$$

where $C(x)$ is defined in (1.11).

Figure 1.5 shows the region of channel conditions where each protocol performs optimally. In this figure, only the SR and RD links change whereas the SD link is fixed. Therefore, in a way, the plot shows relative performance of these protocols as a function of relay position.

The reader is reminded that several notions coexist in the domain of relaying, and for each notion there may be several achievable rates depending on the assumptions made during derivation. Often, finding the achievable rate corresponding to a given protocol under a given channel model is not analytically feasible. This may be true, for instance, if the throughput maximizing input distribution does not have a familiar form. For this reason, it is important to view each achievable rate expression in the light of its context. We have only summarized some of the best achievable rates known for three prominent protocols. The interested reader can find additional results on Gaussian channels in Gupta and Kumar, 2003; Khojastepour et al., 2004b; Laneman et al., 2004; Sabharwal and Mitra, 2005; Gastpar and Vetterli, 2005; Schein and Gallager, 2000.

Fundamental Limits for Half-duplex Relays. In this section, we will first present a new cut-set theorem for networks with multiple states due to Khojastepour, 2004. This cut-set theorem yields an upper bound on the half-duplex relay capacity. Following that, we will discuss well-known upper and lower bounds for the half-duplex relay channel.

The following theorem from Khojastepour, 2004 can be thought of as a generalization of Theorem 1.4. For the convenience of the reader, we define our notation again. The network consists of N nodes. Node i is characterized by the input-output pair $(X^{(i)}, Y^{(i)})$, and sends information at a rate $R^{(ij)}$ to node j. The nodes are divided in two sets S and S^C (the complement of S), and a cut C conceptually separates the nodes in these two sets. The channel is represented by a collection of m conditional probability mass functions $p(y^{(1)}, y^{(2)}, ..., y^{(N)}|x^{(1)}, x^{(2)}, ..., x^{(N)}|m)$ of the outputs given the inputs, where m is the state of the network which takes its values from a set of possible states $\mathcal{M}$, with finite cardinality $M = |\mathcal{M}|$. We denote the state of the channel in the k^{th} network use as m_k. For any state m define $n_m(k)$ as the number of times that the network is used in state m in the first k network uses. Let

$$t_m = \lim_{k\to\infty} \frac{n_m(k)}{k} \tag{1.15}$$

denote the portion of the time that the network has been used in state m as the total number of network uses goes to infinity. The following theorem bounds the rate of information transfer from nodes in S to those in S^C by the conditional mutual information.

THEOREM 1.11 *[Khojastepour, 2004] Consider a general network with M states, where M is finite. If the information rates $\{R^{(ij)}\}$ are achievable, then the sum rate of information transfer from a node set S_1 to a disjoint node set S_2, where $S_1, S_2 \subset \{1, 2, ..., N\}$ and for any choice of network state sequence*

$(m_k)_{k=1}^{\infty}$, is bounded by:

$$\sum_{i\in S_1, j\in S_2} R^{(ij)} \leq \sup_{t_m} \min_{S} \sum_{m=1}^{M} t_m I\left(X_{(m)}^{(S)}; Y_{(m)}^{(S^C)} | X_{(m)}^{(S^C)}\right) \tag{1.16}$$

for some joint probability distributions $p(x^{(1)}, x^{(2)}, ..., x^{(N)}|m), m = 1, 2, ..., M$ when the minimization is taken over all sets $S \subset \{1, 2, ..., N\}$ subject to $S \cap S_1 = S_1$, $S \cap S_2 = \phi$ and the supremum is over all non-negative t_m subject to $\sum_{i=1}^{M} t_m = 1$.

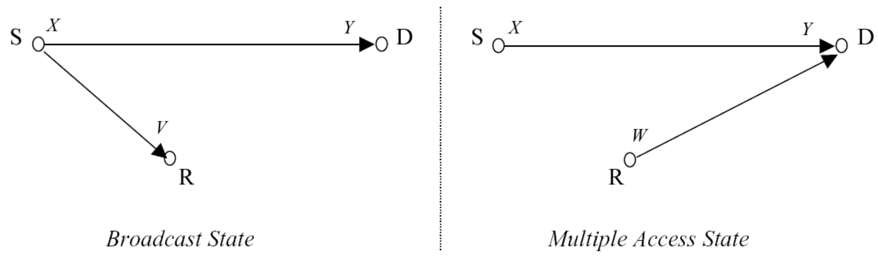

Figure 1.6. Two states of the half-duplex relay channel.

Using the above theorem, an upper bound for the capacity of the half-duplex relay channel can be found. In a half-duplex relay, the network can be in one of two states, based on whether the relay is transmitting or receiving. These states are called the *broadcast* and the *multiple-access* states and are shown in Figure 1.6. The two cuts of interest, which are also called the *broadcast* and *multiple-access* cuts, are shown in Figure 1.2. A simple application of Theorem 1.11 now gives us the following result.

THEOREM 1.12 *[Khojastepour, 2004] The capacity of a general half-duplex relay channel is upper bounded as follows.*

$$C_{hd} \leq \sup_{t:0\leq t\leq 1} \min \Big(tI(X_1; Y_1, V_1) + (1-t)I(X_2; Y_2|W_2), \\ tI(X_1; Y_1) + (1-t)I(X_2, W_2; Y_2)\Big) \tag{1.17}$$

where the subscript 1 stands for the broadcast state, and 2 stands for the multiple-access state.

The most well-known lower bound for half-duplex decode-and-forward relaying is given by the following theorem.

THEOREM 1.13 *[Khojastepour, 2004] The capacity of a general half-duplex relay channel is upper bounded as follows.*

$$C_{hd} \leq \sup_{t, 0 \leq t \leq 1} \min \Big(tI(X_1; V_1) + (1-t)I(X_2; Y_2 | W_2), \\ tI(X_1; Y_1) + (1-t)I(X_2, W_2; Y_2) \Big) \quad (1.18)$$

where the subscript 1 *stands for the broadcast state, and* 2 *stands for the multiple-access state.*

The similarity between Theorem 1.13 and Theorem 1.12 is easy to see. The relationship between the rates in the half-duplex and full-duplex cases is also visible or closer inspection. Results on half-duplex relaying have been discussed in various papers, including Liang and Veeravalli, 2005; Khojastepour et al., 2002a; Høst-Madsen and Zhang, 2005; Zahedi et al., 2004. A comprehensive treatment of this channel can be found in Khojastepour, 2004, including results on Gaussian half-duplex relay channels, and the development of a comprehensive framework that encompasses the cases of time-division and frequency-division half-duplex relaying as special cases of a channel with multiple states.

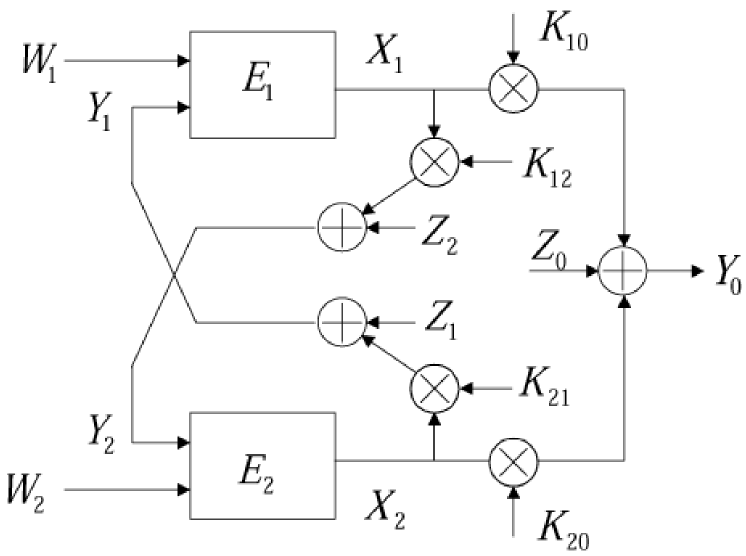

Figure 1.7. Channel model for cooperation when source and relay exchange roles

Limits of Two-user Cooperation. Here, we consider the channel model of Figure 1.7. There are two sources and a single destination . Both transmitters can overhear each other, and are willing to cooperate by forwarding information from the other. Transmitters are capable of full-duplex communication. The mathematical model for this channel is given by the following equations

$$\begin{aligned} Y_0 &= K_{10}X_1 + K_{20}X_2 + Z_0 && (1.19)\\ Y_1 &= K_{21}X_2 + Z_1 && (1.20)\\ Y_2 &= K_{12}X_1 + Z_2, && (1.21)\end{aligned}$$

with $Z_0 \sim \mathcal{N}(0, \Xi_0)$, $Z_1 \sim \mathcal{N}(0, \Xi_1)$ and $Z_2 \sim \mathcal{N}(0, \Xi_2)$. In general, we assume that $\Xi_1 = \Xi_2$. The system is causal and transmission is done for B blocks of length n, therefore the signal of Source 1 at time j, $j = 1, \ldots, n$, can be expressed as $X_1(W_1, Y_1(j-1), Y_1(j-2), \ldots, Y_1(1))$, where W_1 is the message that Source 1 wants to transmit to the destination in that block. Similarly, for Source 2 we have $X_2(W_2, Y_2(j-1), Y_2(j-2), \ldots, Y_2(1))$.

We assume that Source 1 divides its information W_1 into two parts: W_{10}, which is sent directly to the destination, and W_{12}, which is sent to Source 2 and then forwarded by Source 2 to the destination. Source 1 structures its transmit signal so that it is able to send the above information as well as some additional cooperative information to the destination. This is done as follows

$$X_1 = X_{10} + X_{12} + U_1 \tag{1.22}$$

where the power is divided as

$$P_1 = P_{10} + P_{12} + P_{U1}. \tag{1.23}$$

Here, U_1 refers to the part of the signal that carries cooperative information. Thus, X_{10} uses power P_{10} to send W_{10} at rate R_{10} directly to the destination, X_{12} uses power P_{12} to send W_{12} to Source 2 at rate R_{12}, and U_1 uses power P_{U1} to send cooperative information to the destination. Forwarding is based on the principle of decode-and-forward, therefore the transmission rate of W_{12}, i.e. R_{12}, and the power allocated to W_{12}, i.e. P_{12}, should be such that W_{12} can be perfectly decoded by Source 2. Source 2 similarly structures its transmit signal X_2 and divides its total power P_2. In the above setup, the following theorem from Sendonaris et al., 2003a gives an achievable rate region .

THEOREM 1.14 *[Sendonaris et al., 2003a] An achievable rate region for the system given in (1.19)-(1.21) is the closure of the convex hull of all rate pairs* (R_1, R_2) *such that* $R_1 = R_{10} + R_{12}$ *and* $R_2 = R_{20} + R_{21}$ *where*

$$R_{12} < E\left\{C\left(\frac{K_{12}^2 P_{12}}{K_{12}^2 P_{10} + \Xi_1}\right)\right\} \tag{1.24}$$

$$R_{21} < E\left\{C\left(\frac{K_{21}^2 P_{21}}{K_{21}^2 P_{20}+\Xi_2}\right)\right\} \tag{1.25}$$

$$R_{10} < E\left\{C\left(\frac{K_{10}^2 P_{10}}{\Xi_0}\right)\right\} \tag{1.26}$$

$$R_{20} < E\left\{C\left(\frac{K_{20}^2 P_{20}}{\Xi_0}\right)\right\} \tag{1.27}$$

$$R_{10}+R_{20} < E\left\{C\left(\frac{K_{10}^2 P_{10}+K_{20}^2 P_{20}}{\Xi_0}\right)\right\} \tag{1.28}$$

$$R_{10}+R_{20}+R_{12}+R_{21} < E\left\{C\left(\frac{K_{10}^2 P_1+K_{20}^2 P_2+2K_{10}K_{20}\sqrt{P_{U1}P_{U2}}}{\Xi_0}\right)\right\} \tag{1.29}$$

for some power assignment satisfying $P_1 = P_{10} + P_{12} + P_{U1}$, and $P_2 = P_{20} + P_{21} + P_{U2}$. The function $C(x)$ is defined in (1.11) and E denotes expectation with respect to the fading parameters K_{ij}.

Practical Strategies for Relaying Information

There are several parallel efforts going on to harness the gains of relay coding in practice, of which we mentioned some in an earlier section on the history of relaying. In this section, we will focus only on the results of Sendonaris et al., 2003a; Sendonaris et al., 2003b; Khojastepour et al., 2004a; Chakrabarti et al., 2005a.

We first describe a CDMA based user-cooperation strategy that was proposed in Sendonaris et al., 2003a; Sendonaris et al., 2003b. It was one of the first implementations of user-cooperation to have been proposed, and it was designed keeping in mind the realities of cellular communication. After describing the aforementioned scheme, we present relay code designs using LDPC component codes for both full-duplex (Khojastepour et al., 2004a) and half-duplex (Chakrabarti et al., 2005a) relays.

CDMA Implementation for User-cooperation. To begin with, let us assume that each user has a single spreading code , which is orthogonal to the spreading codes of all other users. We further assume that the coherence time of the channel equals L symbol periods, i.e. the channel does not change for L symbol periods. For the simple case of $L = 3$, we describe the transmitted signals. If the sources were not cooperating, they would transmit

$$\begin{aligned} X_1(t) &= a_1\, b_1^{(1)}\, c_1(t)\,, \quad a_1\, b_1^{(2)}\, c_1(t)\,, \quad a_1\, b_1^{(3)}\, c_1(t) \\ X_2(t) &= \underbrace{a_2\, b_2^{(1)}\, c_2(t)}_{\text{Period 1}}, \quad \underbrace{a_2\, b_2^{(2)}\, c_2(t)}_{\text{Period 2}}, \quad \underbrace{a_2\, b_2^{(3)}\, c_2(t)}_{\text{Period 3}} \end{aligned} \tag{1.30}$$

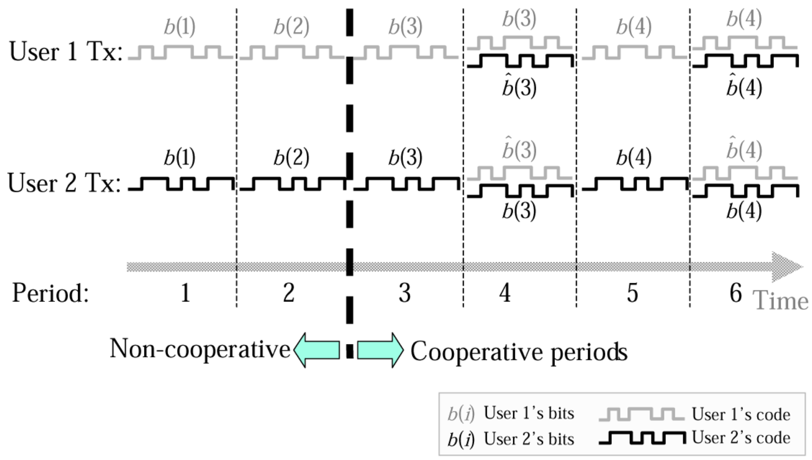

Figure 1.8. User cooperation using spreading codes

where $b_j^{(i)}$ is the i^{th} bit from user j, $c_j(t)$ is the spreading code used by user j, and $a_j = \sqrt{\frac{P_j}{T_s}}$ where P_j is the power used by user j and T_s is the symbol period.

For fairness, any cooperative scheme developed in the same framework must satisfy some basic criteria. The total number of spreading codes used by the two users must remain the same as in the non-cooperative scheme, and the cooperative strategy should be of comparable complexity to the non-cooperative scheme. Under the proposed cooperative scheme, the users transmit

$$\begin{array}{rcllll} X_1(t) & = & a_{11}\, b_1^{(1)}\, c_1(t)\,, & a_{12}\, b_1^{(2)}\, c_1(t)\,, & a_{13}\, b_1^{(2)}\, c_1(t) & \\ & & & & +a_{14}\, \hat{b}_2^{(2)}\, c_2(t) & \\ X_2(t) & = & a_{21}\, b_2^{(1)}\, c_2(t)\,, & a_{22}\, b_2^{(2)}\, c_2(t)\,, & a_{23}\, \hat{b}_1^{(2)}\, c_1(t) & (1.31) \\ & & & & +a_{24}\, b_2^{(2)}\, c_2(t) & \\ & & \underbrace{\qquad\qquad}_{\text{Period 1}} & \underbrace{\qquad\qquad}_{\text{Period 2}} & \underbrace{\qquad\qquad}_{\text{Period 3}} & \end{array}$$

where $\hat{b}_j^{(i)}$ is the partner's estimate of user j's i^{th} bit. The parameters $\{a_{ji}\}$ control the amount of power allocated to a user's own bits versus the bits of the partner, while maintaining an average power constraint of P_j for user j, over L periods.

The way to interpret the above is as follows. In Period 1, each user sends data to the base station only. In period 2, users send data to each other in addition to the base station. After this, each user estimates its partner's data and constructs a cooperative signal that is sent to the destination in Period 3. This cooperative signal is a superposition of spreading codes of both users.

To generalize the above scheme to arbitrary number of symbol periods L, we define another parameter L_c. The two users cooperate for $2L_c$ periods, whereas the remaining $L - 2L_c$ periods are used for sending their own information. For example, in (1.31), $L = 3$ and $L_c = 1$, whereas in (1.30), $L = 3$ and $L_c = 0$. By changing the value of L_c over time, it is possible to achieve different points on the capacity region. The $\{a_{ij}\}$ are chosen to satisfy the power constraints

$$\begin{aligned} \tfrac{1}{L}(L_n a_{11}^2 + L_c(a_{12}^2 + a_{13}^2 + a_{14}^2)) &= P_1 \\ \tfrac{1}{L}(L_n a_{21}^2 + L_c(a_{22}^2 + a_{23}^2 + a_{24}^2)) &= P_2. \end{aligned} \quad (1.32)$$

This cooperative scheme is depicted in Figure 1.8 for the case of $L = 6$, $L_c = 2$. The performance of the above scheme and the design of optimal receivers for this type of user-cooperation is discussed in Sendonaris et al., 2003b.

LDPC Codes for Full-duplex Relaying. One of the first attempts on practical full-duplex relay code design was due to Khojastepour et al., 2004a. Although the designs of Khojastepour et al., 2004a are not optimal in an information-theoretic sense, they perform well, and they incorporate most of the essential components of practical relay LDPC code design, namely - capacity analysis for finite alphabet (eg. BPSK), protocol design, power allocation, factor graph construction, and decoding algorithms. In order to design capacity-approaching relay codes, each of these must be done optimally, and in addition there is often the additional step of code-profile optimization.

Here we will briefly describe two protocols proposed in Khojastepour et al., 2004a, and the factor graph construction. Decoding is performed using belief propagation , and code profiles are optimized using density evolution . The interested reader can refer to Khojastepour et al., 2004a for additional details.

Two protocols are proposed in Khojastepour et al., 2004a. The first is called the *simple protocol*, where transmission from the source occurs in B blocks of length N. A pair of consecutive blocks uses a pair of jointly designed *constituent codes*. Odd blocks use one of the constituent codes, and even blocks use the other. The source sends new information in each block. At the end of each block, the relay finds the codeword that is closest to its received signal, and retransmits it without re-encoding.

The second protocol, which is called the *DF protocol* is inspired by the decode-and-forward scheme, and is somewhat similar to the simple protocol. Again, transmission from the source occurs in B blocks of length N. In each block, the source sends a superposition of a new codeword and a repetition of the previous codeword with an appropriate power ratio. In the first and last blocks, only one codeword is sent. At the end of each block, the relay decodes the new codeword from the received signal and retransmits it the same way as in the simple protocol. The constituent codes used in the above protocols are irregular LDPC codes proposed in Luby et al., 2001; Richardson et al., 2001, chosen for their capacity-approaching performance.

Before proceeding, we present a very brief introduction to factor graphs in the context of LDPC codes. The interested reader will find extensive reading material on factor graphs in Kschischang et al., 2001; Richardson and Urbanke, 2004.

Any block code (LDPC codes are block codes) can be represented completely by its parity check matrix H. This binary matrix, in turn, can be uniquely represented by a bipartite graph. For a discussion of full-duplex LDPC codes only, we follow the following conventions. A variable node is represented by a circle in the graph and corresponds to a column of the parity check matrix H. A check node is represented by a square and corresponds to a row of the same matrix. Last, there is a connection between a check node and variable node if and only if the parity check matrix H has a 1 in the corresponding row and column (we confine ourselves to binary LDPC codes). For example, the following is the parity check matrix of a rate $\frac{1}{2}$ LDPC code.

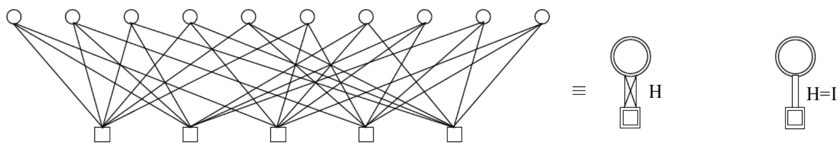

Figure 1.9. Factor graph for parity check matrix H and shorthand noations.

$$H = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 1 \end{bmatrix}.$$

Its factor graph is depicted in Figure 1.9. When $H = I$, the factor graph is a series of parallel connections between the check nodes and the variable nodes. Two shorthand notations are introduced in Figure 1.9 for a general parity check matrix and for the special case of $H = I$.

The signal received by the destination in each block is a superposition of two codewords. Corresponding to this fact, Figure 1.10 shows the factor graph structure for optimal decoding at the destination based on the entire set of B

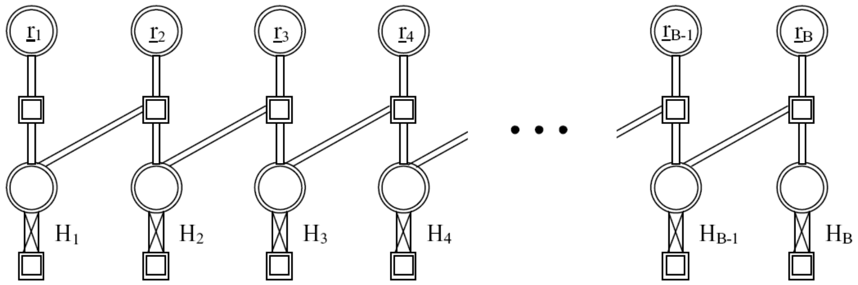

Figure 1.10. Factor graph for optimal decoding at the destination.

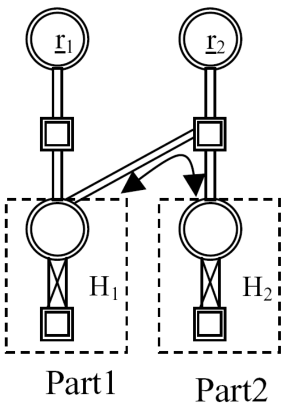

Figure 1.11. Factor graph of a pair of consecutive codes.

blocks. It is extremely complex to find optimized LDPC code profiles for the entire factor graph since it requires joint optimization of B matrices. Therefore, as a practical alternative, only pairs of codes, as depicted in Figure 1.11 are optimized at a time. The two codes in a pair are then alternately used over consecutive channel uses. It is to be noted that a set of LDPC codes optimized for the entire factor graph of Figure 1.10 would perform optimally only when the decoding is performed jointly over all B blocks, which is infeasible. For a block-by-block successive decoding scheme (see Figure 1.12), the optimization of successive code pairs is actually the optimal strategy.

Two algorithms were proposed in Khojastepour, 2004 for decoding the received signals at the destination, called the forward and the backward decoding

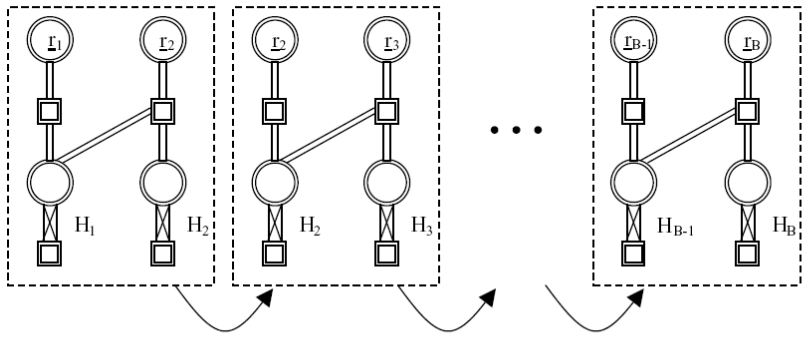

Figure 1.12. Factor graph for block-by-block successive decoding at the destination.

algorithms. Note that the first and the last transmissions in the above coding scheme use only a single code, whereas any intermediate received signal is a superposition of a pair of codes. Therfore, decoding may either commence from the first or the last received codeword, corresponding to forward and backward decoding respectively. Forward decoding has a minimal latency of two blocks, and also performs better when the relay is near the destination. Backward decoding, in contrast, is better when source and relay are close to each other; however, it has a decoding latency of B blocks. The performance of the proposed LDPC codes is ahown in Figure 1.13.

LDPC Codes for Half-duplex Relaying. In this section, we discuss LDPC code designs proposed in Chakrabarti et al., 2005a for the half-duplex relay channel . The code designs are based on the information theoretic random-coding scheme for half-duplex decode-and-forward relaying. Although the relay channel is commonly visualized as a combination of a broadcast and a multiple-access channel, it is shown that the achievable rate of decode-and-forward relaying can be approached by using single-user codes decoded with single-user receivers. The single-user decodability of these codes supports the practicality of half-duplex relaying.

It is shown in Høst-Madsen, 2004; Sabharwal, 2004; Chakrabarti et al., 2005b that the gains of relaying are significant only in low to medium SNRs. At high SNRs, the throughput of relaying is not a significant fraction larger than that of a direct link. And in the low to medium SNR range, binary modulation on each channel dimension (QPSK) approaches the capacity of the AWGN channel. This justifies the use of binary codes. Another challenge in code construction is that the implementation of source-relay correlation in multiple-access mode

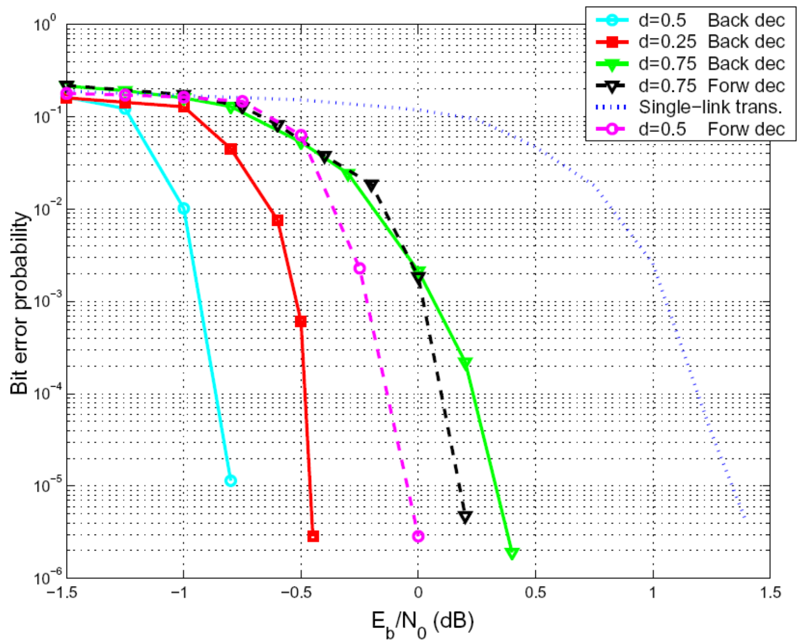

Figure 1.13. Performance of full-duplex relay coding scheme for the simple protocol. Performance of a single-user code using the same constituent parity check matrix is shown for comparison. Here source and relay are unit distance apart and the relay is on the line joining them at a distance d from the source.

introduces an added level of complexity. In contrast it is simple to devise coding schemes where this correlation is either 0 or 1. Empirical results in Chakrabarti et al., 2005b show that the loss in throughput is negligible when the better of $\rho = 0, 1$ is chosen instead of the optimal correlation.

We assume that the two rate terms in the achievable rate expression (1.17) are equal at the point where the achievable rate is maximized, i.e.

$$\begin{aligned} & tI(X_1; V_1) + (1-t)I(X_2; Y_2|W_2) \\ = \; & tI(X_1; Y_1) + (1-t)I(X_2, W_2; Y_2). \end{aligned} \tag{1.33}$$

The above is easy to prove for Gaussian codebooks when the source and the relay have separate power constraints. Even if this is not true in general, it is easy to see that we can find rates $R_1 \leq I(X_1; V_1)$, $R_2 \leq I(X_2; Y_2|W_2)$, $R_3 \leq I(X_1; Y_1)$, and $R_4 \leq I(X_2, W_2; Y_2)$ satisfying

$$\text{Achievable rate} = tR_1 + (1-t)R_2 = tR_3 + (1-t)R_4. \tag{1.34}$$

The proposed coding scheme then remains the same with the mutual information terms replaced by the corresponding rates.

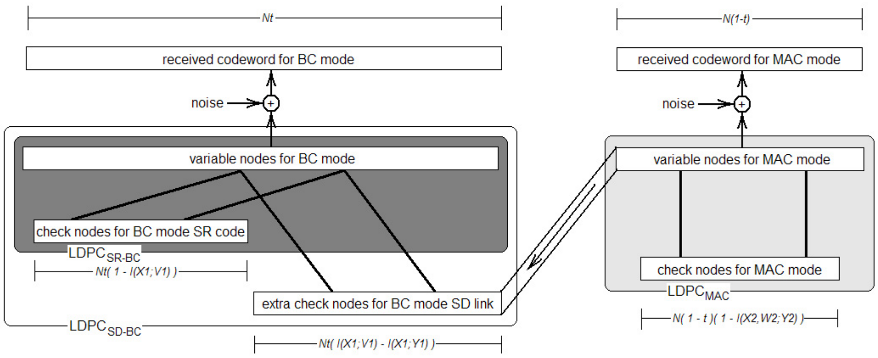

Figure 1.14. LDPC code structure for $\rho = 1$.

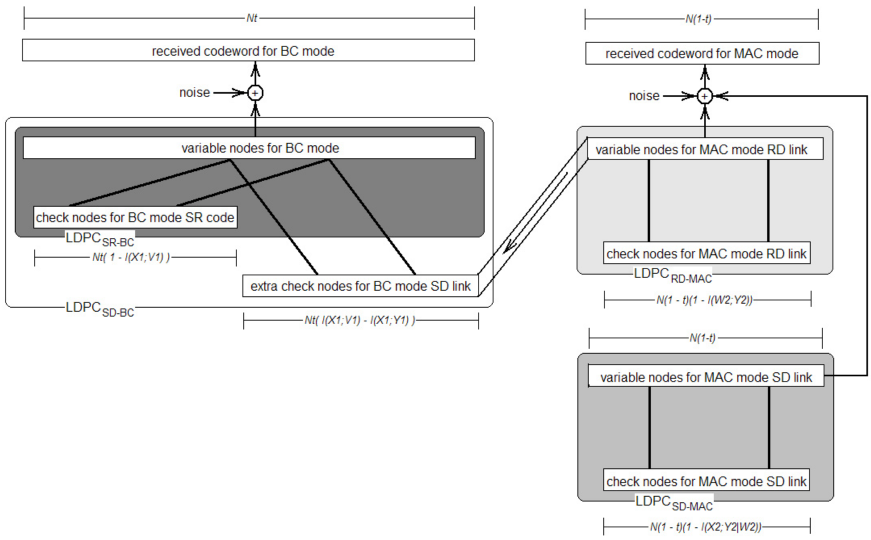

Figure 1.15. LDPC code structure for $\rho = 0$.

The factor-graph structures of the codes for the two cases corresponding to $\rho = 0$ and $\rho = 1$ are shown in Figures 1.14 and 1.15 respectively. The factor graphs in this section are represented as a pair of rectangular blocks representing the variable and check nodes, and connected by a pair of parallel lines denoting the edges of the graph. In contrast to the convention of the previous section, the parallel lines do not indicate that the parity check matrix is identity. Several single-user codes combine to form the factor graphs of Figure 1.14 and Figure 1.15. Any set of variable nodes and associated check nodes that are enclosed in a rectangular box with rounded edges indicates that the enclosed nodes correspond to a single-user LDPC code. We now explain the encoding and decoding for both values of ρ. We assume that the sum length of BC and MAC mode codewords is N bits.

When $\rho = 1$, S and R transmit identical signals in MAC mode. For this case, the following scheme is used. In the beginning of BC mode, S encodes information bits using a code $LDPC_{SR-BC}$ to yield a codeword of length tN bits. This codeword is transmitted by S. At the end of BC mode (which is also the beginning of MAC mode), both R and D receive the BC mode source signal. This signal is successfully decoded by R. However, D cannot decode the received signal, and stores a copy of it. In the beginning of MAC mode, the tN variable bits from BC mode are compressed. Compression is done at both S and R, by multiplying with the same parity matrix. These compressed bits, acting as parity together with the parity bits of $LDPC_{SR-BC}$ form a composite code $LDPC_{SD-BC}$ that can be decoded at D at the end of MAC mode. In order to communicate the compressed bits to D reliably, S, and R treat them as information bits for MAC mode, and re-encode them using a code $LDPC_{MAC}$ to yield a codeword of length $(1-t)N$, which is then transmitted synchronously from S and R with appropriate powers. The structure of the code is shown in Figure 1.14.

For $\rho = 1$, decoding is performed as follows. R decodes $LDPC_{SR-BC}$ at the end of BC mode using belief propagation like any single-user LDPC code. D waits for both BC and MAC mode signals to arrive before it commences decoding. $LDPC_{MAC}$ is decoded like a single-user LDPC code, from which side information in the form of additional parity bits is obtained about the BC mode signal. Using knowledge of the single-user BC mode source-relay code, and with the help of these additional parity bits, $LDPC_{SD-BC}$ is decoded. This final decoding also is performed using belief propagation.

For $\rho = 0$, the BC mode is the same as before. In MAC mode, however, S and R transmit independent (therefore uncorrelated) information using codes $LDPC_{SD-MAC}$ and $LDPC_{RD-MAC}$ respectively. As before, R compresses the information bits received in BC mode to produce additional parity bits, which serve as relay information bits in MAC mode. These bits are re-encoded by R using $LDPC_{RD-MAC}$ to yield $(1-t)N$ coded bits. The source, in MAC mode, sends bits of new information encoded using $LDPC_{SD-MAC}$ to yield another set of $(1-t)N$ coded bits. Thus, $(1-t)N$ coded bits each from S and R are transmitted simultaneously with appropriate power allocation, so that the two (uncorrelated) signals appear superimposed at D. The structure of the code is shown in Figure 1.15.

For $\rho = 0$, decoding proceeds as follows. R decodes the BC mode signal like a single-user LDPC code. D waits for both BC and MAC mode signals. In MAC mode, the rates for SD and RD channels correspond to one of the corner points of the MAC capacity region, for which it is well known Cover and Thomas, 1991 that capacity can be achieved by successive decoding (onion-peeling). The MAC signal is successively decoded to first reveal the relay codeword, treating both noise and interference from S as noise. Next, the relay codeword is subtracted out to reveal the source codeword in the presence of noise alone, which is then decoded. The MAC mode source information is new information, whereas the relay information provides additional parity bits to aid in decoding the BC mode codeword.

The main challenge is the design of codes $LDPC_{SD-BC}$ and $LDPC_{SR-BC}$, which must be jointly optimized, since the factor graph of the latter is a subgraph of the factor graph of the former. The reader should note that these codes are of different rates, and although the received codeword is same for both R and D, the received SNRs are different. To avoid confusion, we would like to mention that neither S, nor R actually uses $LDPC_{SD-BC}$ to encode information. It is merely a convenience to visualize the side information received by D in MAC mode as extra parity bits in addition to the actual parity bits of $LDPC_{SR-BC}$, and call the composite a code $LDPC_{SD-BC}$. The optimization of code profiles is performed using a modification of the density evolution algorithm. In the implementation of density evolution, the messages have been approximated as Gaussians to speed up the optimization, the cost being usually small inaccuracy in threshold determination. We omit details of code profile optimization using modified density evolution; the interested reader can find these in Chakrabarti et al., 2006.

Figure 1.16 shows how far the coding schemes for direct, two-hop and relay coding schemes are from their respective theoretical bounds. The thresholds for direct and two-hop channels are calculated using code ensembles with maximum variable node degrees comparable to the relay codes. In this figure, we fix the rates of the codes, and calculate E_b/N_0 from the numerically calculated thresholds of the LDPC codes. For the relay coding scheme, the achievable E_b/N_0 values are less than 0.4 dB away from the theoretical minimum. Note that the thresholds for the relay channel are not thresholds of any single code, but a function of the thresholds of all component codes.

Figure 1.17 plots the BER performance of the overall relay LDPC coding scheme for $\rho = 1$, taking into account the individual BER performances of all three component codes. Some modifications have been incorporated into the coding scheme to improve performance when the number of decoding iterations is small. The code designs correspond to a total power of -1dB ($P = 10^{-0.1}$)

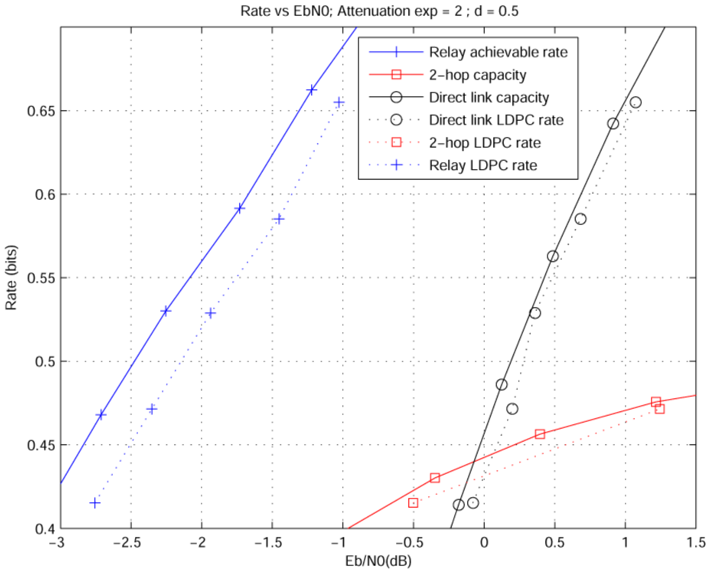

Figure 1.16. Rate vs. E_b/N_0. Theoretical limits and LDPC performance based on thresholds.

and correlation $\rho = 1$ in this case. For this case, the SR code has rate 0.9, the SD code has rate 0.47 and the MAC code has rate 0.8. The results are for a blocklength of 100K. The codes are randomly constructed and there is no cycle removal with the exception of removing double edges.

Conclusion

Problems in cooperative communication continue to intrigue researchers by their difficulty and the potential for faster and more reliable communication. There is a wide open space for the implementation of principles of user-cooperation in mesh and sensor networks. User-cooperation also has potential for implementation in mobile handheld devices, but here fair sharing of resources must be ensured by a suitable protocol. We anticipate that our current understanding of the principles of user-cooperation, together with advances in technology will enable cooperative communication networks in future.

Glossary

Diversity Order The slope of the bit error rate (BER) vs. the signal-to-noise ratio (SNR) at high SNR. Diversity order is a measure of the number of

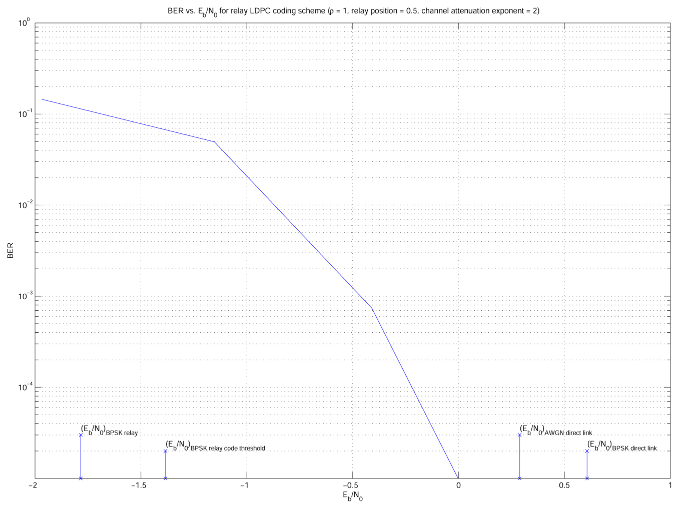

Figure 1.17. BER vs. E_b/N_0 for relay LDPC coding scheme ($\rho = 1$).

independent datapaths from the source to the destination in a communication system. <—- slope is diversity order, diversity by itself is not a number

Amplify-and-forward A relay protocol where the relay retransmits a scaled version of its received analog signal.

Broadcast Channel A communication system where a single transmitter sends potentially different information to multiple users.

CDMA(Code-division Multiple-access) A technology where each user modulates its transmission symbols with a spreading code before transmission. The spreading codes of different users are often orthogonal.

Cut-set theorems A class of theorems that give upper bounds, and sometimes achievable rates, for flow in networks. This flow may correspond to information flow, or flow of a fluid through a network of pipes, or any other physical quantity.

Decode-and-forward A relay protocol where the relay first decodes its received signal, then transmits a signal that is derived from the decoded information.

Degraded relay channel A channel where the signal received at the destination is a corrupted version of the signal received at the relay.

Diversity Order The slope of the bit error rate (BER) vs. the signal-to-noise ratio (SNR) at high SNR. Diversity order is a measure of the number of independent data-paths from the source to the destination in a communication system. Common forms of communication diversity include temporal, spatial, spectral and multiuser diversity.

Estimate-and-forward A relay protocol where the relay transmits an estimate of its received analog signal without decoding but potentially compressed.

Full-duplex A communication node is said to operate in full-duplex mode when it can simultaneously transmit and receive in the same frequency band.

Half-duplex A communication node is said to operate in half-duplex mode when it can simultaneously transmit and receive only if the transmitted and received signals are orthogonal in time.

LDPC (Low-density Parity-check) Code A class of block codes characterized by sparse parity-check matrices.

Multi-hop Communication Communication from a source to a destination through a chain of intermediate nodes, where each intermediate node communicates only with the node immediately preceding it and immediately following it in the chain.

Multiple-access Channel A shared channel where multiple sources transmit to a single destination in the same frequency band.

Network-coding A network information processing paradigm where intermediate nodes combine and encode received information before forwarding it, as contrasted to traditional communication where nodes are restricted to the passive role of forwarding information without processing.

Power Control Adjusting transmission power based on channel condition to achieve a given target. For example, transmission at high power when the channel is good, and low power when the channel is bad, achieves the target of increasing overall rate for a given average power constraint.

Relay Channel A three-terminal communication channel where communication from a source to a destination is aided by a third terminal called the relay.

Relay Protocol The information-processing strategy employed at the relay for retransmitting received information in a relay channel. For example, the relay may choose between retransmitting the received analog signal without decoding, or it may decode the received signal and re-encode it before retransmission.

Space-time Code A channel coding technique in multi-antenna systems, where the antennas as well as time are treated as dimensions of the channel. Space-time codes can be used to yield higher diversity or to achieve higher communication rates than in possible with single-antenna communication.

Spreading Code A noise-like sequence with which a transmission symbol is modulated (multiplied) with the goal of spreading the information in the symbol among different dimensions in the time-frequency plane.

Notes

1. The authors are grateful to Andrew Sendonaris, Elza Erkip, and Mohammad Ali Khojastepour for permission to use figures and various other content from their publications in this chapter.

References

Ahlswede, R., Cai, N., Li, S.-Y. R., and Yeung, R. W. (2000). Network information flow. *IEEE Trans. Inform. Theory*, 46(4):1204–1216.

Ahlswede, R. and Kaspi, A. H. (1987). Optimal coding strategies for certain permuting channels. *IEEE Trans. Inform. Theory*, 33(3):310–314.

Aref, M. R. (1980). *Information flow in relay networks*. PhD thesis, Stanford University.

Bergmans, P. and Cover, T. M. (1974). Cooperative broadcasting. *IEEE Trans. Inform. Theory*, 20:317–324.

Berrou, C., Glavieux, A., and Thitimajshima, P. (1993). Near shannon limit error-correcting coding and decoding: Turbo-codes. In *Proc. of ICC*, volume 2, pages 1064–1070.

Biglieri, E., Proakis, J., and Shamai, S. (1998). Fading channels: Information-theoretic and communications aspects. *IEEE Trans. Inform. Theory*, 44(6):2619–2692.

Boyer, J., Falconer, D. D., and Yanikomeroglu, H. (2004). Multihop diversity in wireless relaying channels. *IEEE Trans. Commun.*, 52(10):1820–1830.

Caire, G. and Shamai, S. (2003). On the achievable throughput of a multiantenna Gaussian broadcast channel. *IEEE Trans. Inform. Theory*, 49:1691–1706.

Carleial, A. B. (1982). Multiple-access channels with different generalized feedback signals. *IEEE Trans. Inform. Theory*, 28(6):841–850.

Castura, J. and Mao, Yongyi (2005). Rateless coding for wireless relay channels. In *Proc. of ISIT*, pages 810–814.

Chakrabarti, A., de Baynast, A., Sabharwal, A., and Aazhang, B. (2005a). LDPC code design for half-duplex decode-and-forward relaying. In *Proc. of the Allerton Conference*, Monticello, IL.

Chakrabarti, A., de Baynast, A., Sabharwal, A., and Aazhang, B. (2006). LDPC codes for decode-and-forward half-duplex relaying. in preparation for IEEE J. Select. Areas Commun.

Chakrabarti, A., Sabharwal, A., and Aazhang, B. (2005b). Sensitivity of achievable rates for half-duplex relay channel. In *Proc. of SPAWC*.

Chong, H. F., Motani, M., and Garg, H. K. (2005). New coding strategies for the relay channel. In *Proc. of ISIT*, pages 1086–1090.

Chou, P. A., Wu, Y., and Jain, K. (2003). Practical network coding. In *Proc. of the Allerton conference*.

Costa, M. H. M. (1983). Writing on dirty paper. *IEEE Trans. Inform. Theory*, 29:439–441.

Cover, T., El Gamal, A., and Salehi, M. (1980). Multiple access channels with arbitrarily correlated sources. *IEEE Trans. Inform. Theory*, 26(6):648–657.

Cover, T. and Thomas, J. (1991). *Elements of Information Theory*. John Wiley and Sons.

Cover, T. M. (1972). Broadcast channels. *IEEE Trans. Inform. Theory*, 18:2–14.

Cover, T. M. (1975). An achievable rate region for the broadcast channel. *IEEE Trans. Inform. Theory*, 21(4):399–404.

Cover, T. M. and El Gamal, A. A. (1979). Capacity theorems for the relay channel. *IEEE Trans. Inform. Theory*, 25:572–584.

Cover, T. M. and Leung, C. S. K. (1981). An achievable rate region for the multiple-access channel with feedback. *IEEE Trans. Inform. Theory*, 27(3):292–298.

El Gamal, A., Mohseni, M., and Zahedi, S. (2004). On reliable communication over additive white Gaussian noise relay channels. *submitted to IEEE Trans. Inform. Theory*.

El Gamal, A. and van der Meulen, E. C. (1981). A proof of Marton's coding theorem for the discrete memoryless broadcast channel. *IEEE Trans. Inform. Theory*, 27(1):120–122.

El Gamal, A. A. (1981). On information flow in relay networks. In *Proc. of IEEE National Telecommunications Conference*, volume 2, pages D4.1.1–D4.1.4.

El Gamal, A. A. and Aref, M. (1982). The capacity of the semideterministic relay channel. *IEEE Trans. Inform. Theory*, 28(3):536.

Ford, L. and Fulkerson, D. (1962). *Flows in Networks*. Princeton University Press.

Foschini, G. J. and Gans, M. J. (1998). On limits of wireless communications in a fading environment when using multiple antennas. *Wireless Personal Communications*, 6(3):311–335.

Gaarder, N. T. and Wolf, J. K. (1975). The capacity region of a multiple-access discrete memoryless channel can increase with feedback. *IEEE Trans. Inform. Theory*, 21(1):100–102.

Gallager, R. G. (1963). *Low Density Parity Check Codes*. PhD thesis, MIT.

Gastpar, M. and Vetterli, M. (2005). On the capacity of large Gaussian relay networks. *IEEE Trans. Inform. Theory*, 51(3):765–779.

Gel'fand, S. I. and Pinsker, M. S. (1980). Capacity of a broadcast channel with one deterministic component. *Probl. Pered. Inform.*, 16(1):24–34.

Grossglauser, M. and Tse, D. N. C. (2002). Mobility increases the capacity of ad hoc wireless networks. *IEEE/ACM Trans. Networking*, 10(4):477–486.

Gupta, P. and Kumar, P. R. (2000). The capacity of wireless networks. *IEEE Trans. Inform. Theory*, pages 388–404.

Gupta, P. and Kumar, P. R. (2003). Towards an information theory of large networks: An achievable rate region. *IEEE Transactions on Information Theory*, 49(8):1877–1894.

Han, T. S. (1981). The capacity region for the deterministic broadcast channel with a common message. *IEEE Trans. Inform. Theory*, 27(1):122–125.

Hasna, M. O. and Alouini, M.-S. (2003). End-to-end performance of transmission systems with relays over Rayleigh-fading channels. *IEEE Trans. Wireless Commun.*, 2(6):1126–1131.

Ho, T., Koetter, R., Medard, M., Karger, D., and Effros, M. (2003). The benefits of coding over routing in a randomized setting. In *Proc. of ISIT*.

Hunter, T. E., Sanayei, S., and Nosratinia, A. (2004). The outage behavior of coded cooperation. In *Proc. of ISIT*, page 270.

Høst-Madsen, A. (2004). Cooperation in the low power regime. In *Proc. of the Allerton Conference*.

Høst-Madsen, A. and Zhang, J. (2005). Capacity bounds and power allocation for the wireless relay channel. *IEEE Trans. Inform. Theory*, 51(6):2020–2040.

Jaggi, S., Sanders, P., Chou, P. A., Effros, M., Egner, S., Jain, K., and Tolhuizen, L. (2005). Polynomial time algorithms for multicast network code construction. *IEEE Trans. Inform. Theory*, 51:1973–1982.

Janani, M., Hedayat, A., Hunter, T. E., and Nosratinia, A. (2004). Coded cooperation in wireless communications: space-time transmission and iterative decoding. *IEEE Trans. Signal Processing*, 52:362 – 371.

Khojastepour, M. A. (2004). *Distributed Cooperative Communications in Wireless Networks*. PhD thesis, Dept. of Electrical and Computer Engg., Rice University.

Khojastepour, M. A., Ahmed, N., and Aazhang, B. (2004a). Code design for the relay channel and factor graph decoding. In *Proc. of Asilomar Conference*, volume 2, pages 2000–2004.

Khojastepour, M. A., Sabharwal, A., and Aazhang, B. (2002a). Bounds on achievable rates for general multi-terminal networks with practical constraints. In *Proc. of IPSN*.

Khojastepour, M. A., Sabharwal, A., and Aazhang, B. (2002b). On the capacity of ’cheap’ relay networks. In *Proc. of CISS*.

Khojastepour, M. A., Sabharwal, A., and Aazhang, B. (2003). On capacity of Gaussian ’cheap’ relay channel. *GLOBECOM*, pages 1776 – 1780.

Khojastepour, M. A., Sabharwal, A., and Aazhang, B. (2004b). Improved achievable rates for user cooperation and relay channels. In *Proc. of ISIT*.

King, R. C. (1978). *Multiple access channels with generalized feedback*. PhD thesis, Stanford University.

Kobayashi, K. (1987). Combinatorial structure and capacity of the permuting relay channel. *IEEE Trans. Inform. Theory*, 33(6):813–826.

Koetter, R. and Medard, M. (2003). An algebraic approach to network coding. *IEEE/ACM Transactions on network coding*, 11:782–795.

Kramer, G., Gastpar, M., and Gupta, P. (2005). Cooperative strategies and capacity theorems for relay networks. *IEEE Trans. Inform. Theory*, 51(9):3037–3063.

Kschischang, F. R., Frey, B. J., and Loeliger, H.-A. (2001). Factor graphs and the sum-product algorithm. *IEEE Trans. Inform. Theory*, 47:498–519.

Laneman, J. N. (2002). *Cooperative diversity in wireless networks: Algorithms and Architectures*. PhD thesis, Massachusetts Institute of Technology, Cambridge, MA.

Laneman, J. N., Tse, D. N. C., and Wornell, G. W. (2004). Cooperative diversity in wireless networks: Efficient protocols and outage behavior. *IEEE Trans. Inform. Theory*, 50:3062 – 3080.

Laneman, J. N. and Wornell, G. W. (2003). Distributed space-time coded protocols for exploiting cooperative diversity in wireless networks. *IEEE Trans. Inform. Theory*, 49(10):2415–2425.

Li, S.-Y. R., Yeung, R. W., and Cai, N. (2003). Linear network coding. *IEEE Trans. Inform. Theory*, 49:371–381.

Li, Z. and Li, B. (2004). Network coding in undirected networks. *Proc. of CISS*.

Liang, Y. and Veeravalli, V. V. (2005). Gaussian orthogonal relay channels: Optimal resource allocation and capacity. *IEEE Trans. Inform. Theory*, 51:3284–3289.

Luby, M. G., Mitzenmacher, M., Shokrollahi, M. A., and Spielman, D. (2001). Improved low-density parity-check codes using irregular graphs. *IEEE Trans. Inform. Theory*, 47:585 – 598.

MacKay, D. J. C. (1999). Good error-correcting codes based on very sparse matrices. *IEEE Trans. Inform. Theory*, 45:399–431.

Marton, K. (1979). A coding theorem for the discrete memoryless broadcast channel. *IEEE Trans. Inform. Theory*, 25(3):306–311.

Mitran, P., Ochiai, H., and Tarokh, V. (2005). Space-time diversity enhancements using collaborative communications. *IEEE Trans. Inform. Theory*, 51(6):2041–2057.

Nabar, R. U., Bolcskei, H., and Kneubuhler, F. W. (2004). Fading relay channels: Performance limits and space-time signal design. *IEEE J. Select. Areas Commun.*, 22:1099–1109.

Reznik, A., Kulkarni, S. R., and Verdú, S. (2004). Degraded Gaussian multiple relay channel: capacity and optimal power allocation. *IEEE Trans. Inform. Theory*, 50(12):3037–3046.

Richardson, T. J., Shokrollahi, M. A., and Urbanke, R. L. (2001). Design of capacity-approaching irregular low-density parity-check codes. *IEEE Trans. Inform. Theory*, 47:619 – 637.

Richardson, T. J. and Urbanke, R. L. (2001). The capacity of low-density parity-check codes under message-passing decoding. *IEEE Trans. Inform. Theory*, 47:599 – 618.

Richardson, T. J. and Urbanke, R. L. (2004). Modern coding theory (draft of book).

Sabharwal, A. (2004). Impact of half-duplex radios and decoding latencies on mimo relay channels. In *Proc. of the Allerton Conference*, Monticello, IL.

Sabharwal, A. and Mitra, U. (2005). Rate-constrained relaying: A model for cooperation with limited relay resources. submitted to IEEE Trans. Inform. Theory.

Sato, H. (1976). Information transmission through a channel with relay. The Aloha System, University of Hawaii, Honolulu, Tech. Rep.

Sato, H. (1978). An outer bound to the capacity region of broadcast channels. *IEEE Trans. Inform. Theory*, 24(3):374–377.

Schein, B. (2001). *Distributed coordination in network information theory*. PhD thesis, Massachusetts Institute of Technology.

Schein, B. and Gallager, R. (2000). The Gaussian parallel relay network. In *Proc. of ISIT*, page 22, Sorrento, Italy.

Sendonaris, A., Erkip, E., and Aazhang, B. (2003a). User cooperation diversity. Part I. System description. *IEEE Trans. Commun.*, 51:1927 – 1938.

Sendonaris, A., Erkip, E., and Aazhang, B. (2003b). User cooperation diversity. Part II. Implementation aspects and performance analysis. *IEEE Trans. Commun.*, 51:1939 – 1948.

Slepian, D. and Wolf, J. (1973). Noiseless coding of correlated information sources. *IEEE Trans. Inform. Theory*, 19:471–480.

Stefanov, A. and Erkip, E. (2004). Cooperative coding for wireless networks. *IEEE Trans. Commun.*, 52(9):1470–1476.

Tarokh, V., Seshadri, N., and Calderbank, A. R. (1998). Space-time codes for high data rate wireless communication: Performance criterion and code construction. *IEEE Trans. Inform. Theory*, 44:744–765.

Telatar, I. E. (1999). Capacity of multiple-antenna Gaussian channels. *Eur. Trans. Tel.*, 10(6):585–595.

Thomas, J. A. (1987). Feedback can at most double Gaussian multiple access channel capacity. *IEEE Trans. Inform. Theory*, 33(5):711–716.

Toumpis, S. and Goldsmith, A. J. (2003). Capacity regions for wireless ad hoc networks. *IEEE Trans. Wireless Commun.*, 2(4):736–748.

van der Meulen, E. C. (1968). *Transmission of information in a T-terminal discrete memoryless channel*. PhD thesis, Dept. of Statistics, University of California, Berkeley.

van der Meulen, E. C. (1971). Three-terminal communication channels. *Advanced Applied Probability*, 3:120–154.

van der Meulen, E.C. (1977). A survey of multi-way channels in information theory: 1961-1976. *IEEE Trans. Inform. Theory*, 23:1–37.

Vanroose, P. and van der Meulen, E. C. (1992). Uniquely decodable codes for deterministic relay channels. *IEEE Trans. Inform. Theory*, 38(4):1203–1212.

Vishwanath, S., Jindal, N., and Goldsmith, A. J. (2003). Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels. *IEEE Trans. Inform. Theory*, 49(10):2658–2668.

Viswanath, P. and Tse, D. N. C. (2003). Sum capacity of the vector Gaussian broadcast channel and uplink-downlink duality. *IEEE Trans. Inform. Theory*, 49:1912–1921.

Wang, Bo, Zhang, Junshan, and Høst-Madsen, A. (2005). On the capacity of MIMO relay channels. *IEEE Trans. Inform. Theory*, 51:29 – 43.

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

Wyner, A. D. (1978). The rate-distortion function for source coding with side information at the receiver—II: General sources. *Inform. Contr.*, 38:60–80.

Wyner, A. D. and Ziv, J. (1976). The rate-distortion function for source coding with side information at the receiver. *IEEE Trans. Inform. Theory*, 22(1):1–11.

Xie, L. L. and Kumar, P. R. (2005). An achievable rate for the multiple-level relay channel. *IEEE Trans. Inform. Theory*, 51(4):1348–1358.

Xie, Liang-Liang and Kumar, P. R. (2004). A network information theory for wireless communication: Scaling laws and optimal operation. *IEEE Trans. Inform. Theory*, 50(5):748–767.

Xue, F., Xie, Liang-Liang, and Kumar, P. R. (2005). The transport capacity of wireless networks over fading channels. *IEEE Trans. Inform. Theory*, 51(3):834–847.

Yeung, R. W., Li, S.-Y. R., Cai, N., and Zhang, Z. (2005). Theory of network coding.

Yeung, R. W. and Zhang, Z. (1999). Distributed source coding for satellite communications. *IEEE Trans. Inform. Theory*, 45:1111–1120.

Yu, Wei and Cioffi, J. M. (2004). Sum capacity of Gaussian vector broadcast channels. *IEEE Trans. Inform. Theory*, 50:1875–1892.

Zahedi, S., Mohseni, M., and El Gamal, A. (2004). On the capacity of AWGN relay channels with linear relaying functions. In *Proc. of ISIT*, page 399, Chicago, IL.

Zeng, C. M., Kuhlmann, F., and Buzo, A. (1989). Achievability proof of some multiuser channel coding theorems using backward decoding. *IEEE Trans. Inform. Theory*, 35(6):1160–1165.

Zhang, Z. (1988). Partial converse for a relay channel. *IEEE Trans. Inform. Theory*, 34(5):1106–1110.

Zhang, Z., Bahceci, I., and Duman, T. M. (2004). Capacity approaching codes for relay channels. In *Proc. of ISIT*.

Zhang, Zheng and Duman, T. M. (2005). Capacity approaching turbo coding for half duplex relaying. In *Proc. of ISIT*.

Zhao, B. and Valenti, M. C. (2003). Distributed turbo coded diversity for relay channel. *Electronics letters*, 39:786–787.