

# Cooperative Communication in Sensor Networks: Relay Channels with Correlated Sources

Brian Smith and Sriram Vishwanath

Department of ECE, University of Texas at Austin, Austin, TX

*E-mail* : {bsmith,sriram}@ece.utexas.edu

## Abstract

This paper emphasizes the importance of relays channels that possess correlated information at the transmitter and the relay as basic building blocks for sensor networks. This work characterizes achievable rates, and in some cases, the capacity of these channels. This characterization is obtained for two distinct scenarios. The first scenario involves the transmission of a single source to the receiver using the relay as an intermediate node. The second involves communicating two sources of information simultaneously (available at the transmitter and at the relay respectively) to the receiver.

In this paper, explicit conditions are found under which the transmitter and relay can achieve full data cooperation, as if they were co-located and acted as a multiple antenna transmitter. Finally, the MIMO Gaussian relay channel is used as an illustrative example to show that the set of relay channels with correlated information in which full data cooperation is possible, and thus *capacity* can be characterized, is non-trivial and is a fairly realistic set.

## 1 Introduction

The memoryless relay channel, introduced by van der Meulen [2,3] (as shown in Figure 1) is a channel consisting of an input  $X$ , a relay output  $Y_1$  and a relay sender  $X_1$  (which can depend on previously received  $Y_1$ ), a channel output  $Y$ , and a conditional distribution  $p(y, y_1|x, x_1)$ . The capacity region of the relay channel is an open problem. However, many interesting cases such as the physically degraded relay [1], semi-deterministic relay [4], and relay with feedback have been characterized. Most of these characterizations are based on the block-Markov coding scheme, introduced by Cover and El Gamal in [1] to obtain the best known achievable region for this channel.

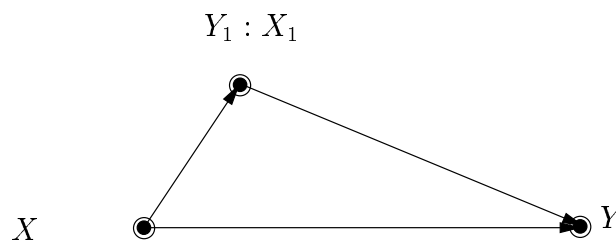


Figure 1: The Relay Channel

Interest in relay channels and their capacity has seen a resurgence in recent years [6, 9, 11, 16–19]. Multi-hop communication has become the most critical enabling technology for sensor networks, thus increasing the importance of understanding the fundamental limits of the relay channel [10]. However, the relay channel embedded in a sensor network has one major difference from that defined in [3]: the relay in a sensor network has side information on the source.

It is highly likely that sensors in the proximity of one another possess correlated data, as illustrated in Figure 2. Since relaying involves transmission between two neighboring nodes, the more appropriate relay channel model to be studied is one with correlated sources  $U$  and  $V$ , as shown in Figure 3. Common randomness amongst sources has shown to aid in interactive communication in sensor networks [7, 8], and we show explicitly how this is so in the relay channel.

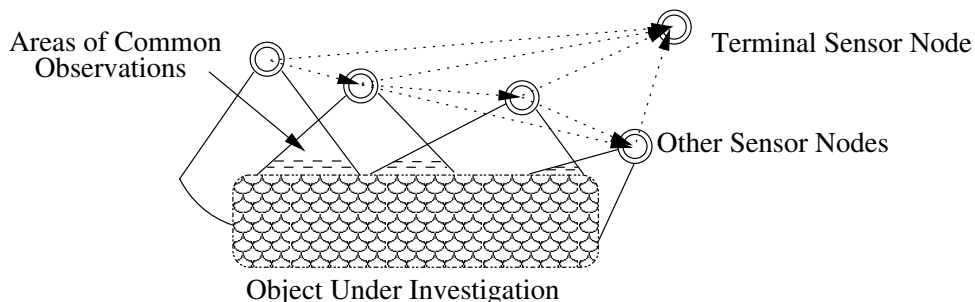


Figure 2: Nodes in a Sensor Network with Correlated Observations of an Object

Transmitting correlated sources [20, 21] over a noisy multi-user channel remains a complex problem, as the source-channel separation theorem does not hold in this domain. Cover et al. [13] discuss the transmission of correlated sources over a multiple-access channel (MAC) and obtain an achievable region. However, as Dueck and others noticed, [5, 14, 15], the difficulty coordinating communication in a MAC, unless the two sources are completely correlated, is in isolating the common randomness between the two sources. Ahlswede and Csiszàr [12] have shown that determining one bit of common information is as complex as isolating the entire common information between two spatially separated correlated random variables. In the correlated relay channel, however, the communication link between the transmitter and the relay aids in the process of determining that information which is common between the transmitter and the relay.

In this paper, we combine the block-Markov coding strategy [1] with list-decoding at both the relay and the receiver to characterize an achievable rate for the relay channel with correlated sources. Specifically, we show that, unlike the MAC with correlated sources, the degree of cooperative communication between the source and the relay is directly related to the correlation between their sources. Thus, a higher throughput is achieved by the overall sensor network, or alternatively, less power is required to relay the same amount of data through a sensor network using our scheme.

In the next section, we review the relay channel. In Section 3, we obtain an achievability region for the transmission of a source given side information at the relay, using block-Markov and list-decoding. In Section 4 we address the transmission of two sources simultaneously (the source and the side information) to the receiver. We then provide an example in the context of the additive Gaussian MIMO channel to illustrate that the set of cases where the achievable rate equals capacity is a fairly large set.

## 2 The Relay Channel

The conventional relay channel without side information is shown in Figure 1. For this channel, a cut-set upper bound (using max-flow min-cut principles) was obtained in [1], and is given by:

$$\mathcal{C} \leq \max_{p(x, x_1)} \min \{I(X, X_1; Y), I(X; Y, Y_1|X_1)\} \quad (1)$$

where  $\mathcal{C}$  denotes the capacity of the channel.

Achievable rate for this channel are also obtained in [1, Theorems 1 & 7]. Here, we focus on the simpler block-Markov coding scheme utilized to obtain the rate described in [1, Theorem 1]. This rate differs from the cut-set upper bound (1) by the omission of  $Y$  in the second term, and is given by:

$$\mathcal{C} \geq \max_{p(x, x_1)} \min \{I(X, X_1; Y), I(X; Y_1|X_1)\} \quad (2)$$

A brief review of the block-Markov coding scheme is given below. Details of this scheme can be found in [1].

*Codebook Generation:* Fix  $p(x_1)p(x|x_1)$ . Generate  $2^{nR_0}$  i.i.d.  $x_1^n$  sequences as  $\prod_n p(x_1)$  and index them as  $x_1^n(s)$ ,  $s \in \mathcal{S} = \{1, 2, \dots, 2^{nR_0}\}$ . For each  $x_1^n$ , generate  $2^{nR}$  conditionally independent  $x^n$  sequences as  $\prod_n p(x|x_1)$  and index them as  $x^n(w|s)$ ,  $w \in \mathcal{W} = \{1, 2, \dots, 2^{nR}\}$ . This defines the joint codebook  $\mathcal{C} = \{x^n(w|s), x_1^n(s)\}$ . For each message  $w \in \mathcal{W}$  assign an index  $s(w)$  at random from  $\mathcal{S}$  to form the bins  $S_s \subseteq \mathcal{W}$ .

*Encoding:* In each block  $b$  the transmitter sends the codeword  $x^n(w_b|s_{b-1})$ , depending on the current message  $w_b$  and the bin index  $s_{b-1}$  such that  $w_{b-1} \in S_{s_{b-1}}$ . The relay is assumed to have a message estimate  $\hat{w}_{b-1}$  and can therefore generate a bin estimate  $\hat{s}_{b-1}$ . The relay transmits  $x_1^n(\hat{s}_{b-1})$  in the same block  $b$ .

*Decoding:* The decoding proceeds in three steps as detailed below:

1. At the end of block  $b$ , the relay declares that  $\hat{w}_b$  was sent if it is the unique index such that  $(x^n(\hat{w}_b|\hat{s}_{b-1}), x_1^n(\hat{s}_{b-1}), y_1^n(b)) \in A_\epsilon^{(n)}$ . This decoding is correct with an arbitrarily small probability of error for  $n$  large and  $R < I(X; Y_1|X_1)$ .
2. The receiver declares  $\hat{s}_{b-1}$  was sent by the relay if there exists a unique index  $\hat{s}_{b-1}$  such that  $(x_1^n(\hat{s}_{b-1}), y^n(b)) \in A_\epsilon^{(n)}$ . This is correct with an arbitrarily small probability of error if  $n$  is large and  $R_0 < I(X_1; Y)$ .
3. Finally, the receiver constructs the list  $\mathcal{L}(y^n(b-1))$  of message indices  $w \in \mathcal{W}$  whose codewords are jointly typical with  $y^n(b-1)$ . The receiver then declares  $\hat{w}_{b-1} \in \mathcal{W}$  was received if it is the unique message in  $S_{\hat{s}_{b-1}} \cap \mathcal{L}(y^n(b-1))$ . The receiver is correct with arbitrarily small probability of error for  $n$  large if  $R - R_0 < I(X; Y|X_1)$ , or equivalently when  $R < I(X, X_1; Y)$  [1]. These two bounds on  $R$  lead to the achievable lower-bound on capacity given by Equation (2).

## 3 Side Information at the Relay

Consider the relay channel shown in Figure 3. The random variable  $U$  describes the source at the transmitter, and  $V$  represents an alternative data source available to the relay that is correlated with  $U$ . We will show that, when  $I(X; Y|X_1) > H(U|V)$  for a particular  $p(x, x_1) =$

$p^\dagger(x, x_1)$ , the cooperative rate of  $R < I(X, X_1; Y)$  for  $p^\dagger(x, x_1)$  is achievable. In other words, the rates given by:

$$\begin{aligned} H(U) &< I(X, X_1; Y) \\ H(U|V) &< I(X; Y_1|X_1) \end{aligned}$$

for a joint  $p(x, x_1)$  are achievable. We use block-Markov coding and list-decoding as the basis of our analysis.

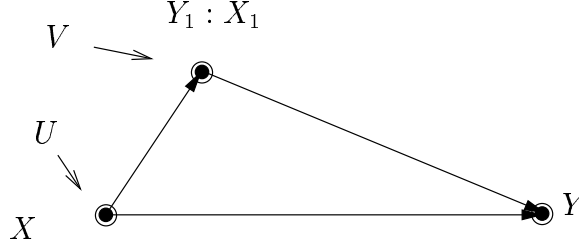


Figure 3: The Relay Channel with Side Information

The strategy is as follows:

*Codebook Generation:* Identical to block-Markov strategy described in Section 2.

*Encoding:* Generate  $2^{nR}$  sequences  $U^n$  and index them by  $w \in \mathcal{W}$ . The remainder of the encoding steps are identical to those in Section 2.

*Decoding:* Assume the relay has an estimate  $\hat{s}_{b-1}$  at the end of block  $b$ . Upon receiving  $y_1^n(b)$ , the relay forms two lists described below.

1. The list  $\mathcal{L}_1(y_1^n(b))$  of all message indices  $w \in \mathcal{W}$  such that  $(x^n(w|\hat{s}_{b-1}), x_1^n(\hat{s}_{b-1}), y_1^n(b)) \in A_\epsilon^{(n)}$ .
2. The list  $\mathcal{L}_2(V_b^n)$  of indices  $w \in \mathcal{W}$  such that  $(U^n(w), V_b^n) \in A_\epsilon^{(n)}$ .

The decoder at the relay declares  $\hat{w}_b$  as the message if  $\hat{w}_b$  is the unique message in  $\mathcal{L}_1 \cap \mathcal{L}_2$ . This can be done with arbitrarily small probability of error for  $n$  large if  $R < I(X; Y_1|X_1) + I(U; V)$ . A detailed derivation of probability of error is included in the appendix.

Decoding at the receiver follows the strategy described in Section 2: the receiver declares  $\hat{s}_{b-1}$  was sent by the relay and then declares  $\hat{w}_{b-1}$  was sent in the previous block. This can be done with arbitrarily small error if  $R_0 < I(X_1; Y)$ ,  $R - R_0 < I(X; Y|X_1)$  for  $n$  sufficiently large. Thus, the achievable rate of the system is given by

$$\begin{aligned} &\max_{p(x, x_1)} I(X, X_1; Y) \\ \text{such that} & \\ &H(U|V) < I(X; Y_1|X_1). \end{aligned} \tag{3}$$

In conclusion, a rate of  $I(X, X_1; Y)$  can be achieved, where the set of joint distributions is constrained to that set of distributions that satisfy  $H(U|V) < I(X; Y_1|X_1)$ . In other words, when  $H(U|V) < I(X; Y_1|X_1)$ , a cooperative rate of  $I(X, X_1; Y)$  can be achieved on this channel. Let  $p^*(x, x_1)$  be the probability distribution that maximizes  $I(X, X_1; Y)$ . Assume that, for  $p^*(x, x_1)$ , we have  $H(U|V) < I(X; Y_1|X_1)$ . This implies that complete cooperation

between the transmitter and relay is possible. In other words, the same capacity that could be achieved if the transmitter and the relay were co-located and cooperating completely to send the single source  $U$ . Even when the inequality is not fulfilled, correlated side information can substantially increase the allowable transmission rate.

## 4 Two Correlated Sources

In this section, we desire to transmit information from two nodes, namely from both the transmitter and the relay, to the receiver. Based on the system described in Figure 3, we communicate both  $U$  and  $V$  to the receiver. This system is analogous to a multiple access channel with correlated sources, where one of the sources has the means of communicating with the other. For this case, for a  $p(z)p(x|z)p(x_1|z)$ , we will show that the rates given by

$$H(U, V) < I(X, X_1; Y), \quad (4)$$

$$H(U) < I(X, Z; Y), \quad (5)$$

$$H(U|V) < I(X; Y_1|Z), \quad (6)$$

$$H(V|U) < I(X_1; Y|Z). \quad (7)$$

are achievable.

The coding and decoding arguments for this problem differ from the case when we only wish to transmit  $U$  (in Section 3) because the codeword  $x_1^n$  has the added role of carrying information about  $V^n$ . In the coding arguments presented in Section 3, the transmitter knows the codeword transmitted from the relay a-priori (i.e.  $x_1^n$ ) at the beginning of each block. In this case, the transmitter has no knowledge of  $V$ , and hence has only partial information about the codeword  $x_1^n$ , namely of the part representing the bin index  $s$ . Hence, this necessitates the introduction of an auxiliary random variable  $Z$  to represent the common information  $s$  between the transmitter and the relay. Using  $Z$ , and the strategies of block-Markov and list-decoding, we obtain the following coding strategy:

*Codeword Generation:* Fix  $p(z)p(x|z)p(x_1|z)$ . Generate  $2^{nR_0}$  i.i.d.  $z^n$  sequences as  $\prod_n p(z)$  and index them as  $z^n(s)$ ,  $s \in \mathcal{S} = \{1, 2, \dots, 2^{nR_0}\}$ . For each  $z^n$ , generate  $2^{nR_u}$ ,  $R_u \geq R_0$  conditionally independent  $x^n$  sequences as  $\prod_n p(x|z)$  and index them as  $x^n(w|s)$ ,  $w \in \mathcal{W} = \{1, 2, \dots, 2^{nR_u}\}$ . Also, for each  $z^n$ , generate  $2^{nR_1}$  conditionally independent  $x_1^n$  sequences as  $\prod_n p(x_1|z)$  and index them as  $x_1^n(k, s)$ ,  $k \in \mathcal{K} = \{1, 2, \dots, 2^{nR_1}\}$ . For each message  $w \in \mathcal{W}$ , assign an index  $s(w) \in \mathcal{S}$  randomly from  $\{1, 2, \dots, 2^{nR_0}\}$  to form the bins  $S_s \subseteq \mathcal{W}$ . This defines a joint codebook  $C(x^n(w|s), x_1^n(k, s))$ .

*Encoding:* Generate  $2^{nR_u}$  i.i.d  $U^n \sim \prod_n p(u)$  indexed by  $w \in \mathcal{W}$ . Also generate  $2^{nR_v}$  i.i.d.  $V^n$  and assign them randomly to  $2^{nR_1}$  bins indexed by  $k \in \mathcal{K}$ . In block  $b$ , the transmitter coding strategy is identical to that in Section 3. It transmits the codeword  $x^n(w_b|s_{b-1})$ , where  $w_b \in \mathcal{W}$  is the message to be transmitted in block  $b$  and  $s_{b-1}$  is the bin index such that  $w_{b-1} \in S_{s_{b-1}}$ . If  $V_b^n$  is the message at the relay in block  $b$ , the relay determines the corresponding bin index  $k_b \in \mathcal{K}$ . Assuming the relay has an estimate  $\hat{s}_{b-1}$  of  $s_{b-1}$ , the relay transmits  $x_1^n(\hat{s}_{b-1}, k_b)$ .

*Decoding:* The decoding follows the following steps:

1. The decoding at the relay is identical to that described for the case described in Section 3. It is easy to show that this decoding can be performed with arbitrarily small probability of error when  $R_u < I(X; Y_1|Z) + I(U; V)$  for  $n$  sufficiently large.

2. The receiver declares  $\hat{s}_{b-1}$  was sent if there exists  $\hat{s}_{b-1}$  as the unique index such that  $z^n(\hat{s}_{b-1})$  is jointly typical with  $y^n(b)$ . This is possible with arbitrarily small probability of error if  $R_0 < I(Y; Z)$  for  $n$  large enough.
3. The receiver then declares  $\hat{w}_{b-1}$  was sent in block  $b-1$  if there exists a unique  $\hat{w}_{b-1} \in \mathcal{W}$  in the set  $\mathcal{L}(y^n(b-1) \cap S_{\hat{s}_{b-1}})$  where  $\mathcal{L}$  is obtained as in Section 3. Again, this is possible with arbitrarily small probability of error when  $R_u - R_0 < I(X; Y|Z)$  and  $n$  sufficiently large. From  $\hat{w}_{b-1}$  the receiver determines  $\hat{U}_{b-1}^n$ .
4. The receiver declares  $\hat{k}_{b-1}$  was sent in block  $b-1$  if there exists  $\hat{k}_{b-1} \in \mathcal{K}$  as the unique index such that  $(x^n(\hat{w}_{b-1}|\hat{s}_{b-2}), x_1^n(\hat{s}_{b-2}, \hat{k}_{b-1}), y^n(b-1)) \in A_\epsilon^{(n)}$ . It is correct with arbitrarily low probability of error if  $n$  is sufficiently large and  $R_1 < I(X_1; Y|X, Z)$ . Finally, the receiver determines  $\hat{V}_{b-1}^n$  that is jointly typical with  $\hat{U}_{b-1}^n$  in bin  $\hat{k}_{b-1}$ .

These inequalities combine to create the conditions in Equations (4-7). They are derived in detail in the Appendix.

## 5 Example

A class of channels for which the gains of joint source-channel coding are substantial is that of the MIMO relay channel with correlated sources at the transmitter and relay nodes. For example, consider a Gaussian relay channel with two receive antennas and a single antenna each at the transmitter and relay. This channel can be represented as:

$$\begin{aligned} Y_1 &= hX + n_1 \\ Y &= H[X \ X_1]^T + N. \end{aligned}$$

where  $H$  is a  $2 \times 2$  matrix and  $Y$  &  $N$  are  $2 \times 1$  vectors. The noise at both the relay and receiver is assumed to be Gaussian  $\sim \mathcal{N}(0, 1)$ . Also, the transmitter and relay possess separate average power constraints of  $P$  and  $P_1$  respectively.

We desire to communicate the source  $U$  to the receiver by using the relay. Full data cooperation and hence capacity is achieved on this channel if the correlation between  $U$  and  $V$  exceeds a certain threshold value. In this section, we find this threshold value for an example MIMO relay channel, showing that the capacity of a large set of channels can be characterized using our strategy.

Consider the Gaussian MIMO relay channel, illustrated in Figure 4.

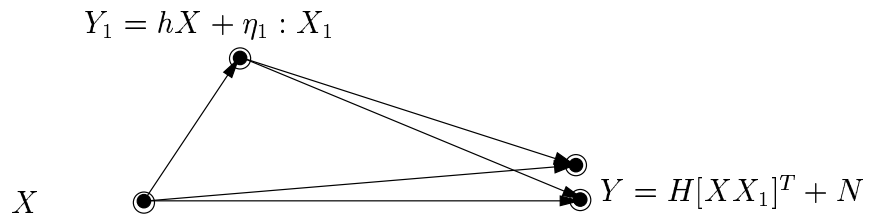


Figure 4: The MIMO Gaussian Relay Channel

The achievable rates for a discrete memoryless relay channel with correlated sources are described by Equation (3). For the MIMO Gaussian case, this expression is modified to be written as a supremization over all cumulative distribution functions (c.d.f.s)  $F(x, x_1)$  that

fulfill the power constraints on the transmitter and relay. Thus for a joint c.d.f.  $F(x, x_1)$ , the achievable rates on this channel are:

$$\begin{aligned} H(U) &< I(X, X_1; Y), \\ H(U|V) &< I(X; Y_1|X_1). \end{aligned} \quad (8)$$

The maximization problem in Equation (8) is complicated, as it contains the term  $H(U|V)$  which is dependent upon  $H(U) = R$ , the quantity we desire to maximize. A specific model for the relationship between  $U$  and  $V$  is required for further analysis. Assuming  $U$  and  $V$  to belong to a discrete  $M$ -ary alphabet, we model  $U$  to be a noised version of  $V$ , given by  $U = V + N$ . Here,  $N$  is assumed to be independent of  $V$ , and to have an entropy of  $(1 - \alpha)H(U)$ . An easy calculation shows that an  $\alpha$  of *zero* implies no side information at the relay, and  $\alpha = 1$  implies complete correlation between  $U$  and  $V$ . Thus, intermediate values of  $\alpha$  characterize the degree of correlation between  $U$  and  $V$ , resulting in  $I(U; V) = \alpha H(U)$ . Given the parameter  $\alpha$ , Equation (8) is equivalent to

$$R \leq \sup_{F(x, x_1)} \min \left\{ I(X, X_1; Y), \frac{I(X; Y_1|X_1)}{1 - \alpha} \right\} \quad (9)$$

Consider an example MIMO channel given by:

$$\begin{aligned} Y_1 &= 3X + n_1 \\ Y &= \begin{bmatrix} 1 & 1.5 \\ .5 & 2.5 \end{bmatrix} [X \ X_1]^T + N \end{aligned}$$

and a power constraint of  $P = P_1 = 2$  on each of the transmitters. This example was picked to reflect the scenario where the transmitter is closer to the relay than the receiver. For this scenario, we compute the maximum value attainable by permitting full data cooperation between  $X$  and  $X_1$ , i.e., we obtain that  $p^*(x, x_1)$  that maximizes

$$I(X, X_1; Y) \quad (10)$$

This maximization problem corresponds to that of a  $2 \times 2$  point-to-point MIMO channel with *separate* power constraints at each antenna of the transmitter. It is well known that Gaussian inputs maximize the capacity of these channels, and for our particular example channel, the covariance

$$\Sigma_{X, X_1} = \begin{bmatrix} 2 & 0.898 \\ 0.898 & 2 \end{bmatrix} \quad (11)$$

between  $X$  and  $X_1$  maximizes  $I(X, X_1; Y)$ , leading to a value of 2.569 bits/use. For this choice of the joint distribution between  $X$  and  $X_1$  (given by 11),

$$I(X; Y_1|X_1) = 1.971.$$

Thus, if  $I(U; V)$  exceeds 0.6, then **the capacity** of this relay channel with correlated sources equals 2.569 bits/use.

Using the model that  $I(U; V) = \alpha H(U)$ , we find that  $\alpha = 0.233$  is the *threshold value*. For all correlations over 25%, the capacity of this channel equals the full data cooperative capacity of 2.569 bits/use.

Note that a correlation of 25% between the data streams at the transmitter and relay is a highly realistic scenario, and that full data cooperation is possible with this correlation. This is unlike the multiple access channel with correlated sources, where 99% of correlation is insufficient, in general, to enable any cooperation between the two sources.

## 6 Conclusions and Further Work

We have shown that a large degree of cooperation is possible between the transmitter and relay if side information is available at the relay. Specifically, the capacity of the relay channel with correlated sources is known if the joint distribution between  $X$  and  $X_1$  that maximizes  $I(X, X_1; Y)$  is such that  $I(X; Y_1|X_1) > H(U|V)$ . In addition, cooperation with isolation of common information is also possible when we desire to transmit two correlated sources on a relay channel to the receiver. Finally, we characterized the *capacity* of a MIMO Gaussian relay channel with sources correlated beyond a threshold value.

While this simple case will have application to the sensor network cooperative communications problem, it is desirable to extend the result to see how correlated data could provide a rate advantage in more complex, multiple relay situations. Such an extension could help lead to results which would be directly applied for more general sensor network configurations.

## Appendix

**Probability of Error Calculations, Single Source to be Transmitted** A decoding error occurs at the relay if either 1) the transmitted and received sequences  $x^n$  and  $y_1^n$  are not jointly typical or 2) a message  $w$  other than the correct message has *both* a transmitted sequence  $x^n$  jointly typical with the received sequence  $y_1^n$  and an associated source sequence  $U^n$  which is jointly typical with the side information  $V^n$ .

The first probability, by the AEP, is less than  $\epsilon$ . Given the sequence  $x_1^n$ , there are  $2^{nR} - 1$  possible messages  $w$  that could be incorrectly identified as the true message. The probability of an incorrect codeword being jointly typical with the received sequence  $y_1^n$  is less than  $2^{-n(I(X; Y_1|X_1) - 3\epsilon)}$ . Also, the probability that the  $U^n$  corresponding to that incorrect message  $w$  is jointly typical with the known  $V^n$  is less than  $2^{-n(I(U; V) - \epsilon)}$ . The total probability of error along this leg, then, is less than

$$\epsilon + \sum_{i=2}^{2^{nR}} 2^{-n(I(X; Y_1|X_1) - 3\epsilon)} 2^{-n(I(U; V) - \epsilon)} \quad (12)$$

that is, less than the sum of all possible error events. When  $R = H(U)$ , this probability of error is driven  $\epsilon$ -small by increasing  $n$  when  $H(U) - I(X; Y_1|X_1) - I(U; V) < 0$  which is equivalent to the statement

$$I(X; Y_1|X_1) > H(U|V). \quad (13)$$

The probability of error analysis for every other step in the decoding process is exactly the same as in the block-Markov scheme described in Theorem 2 of [1].

**Probability of Error Calculations, Two Correlated Sources to be Transmitted** Four decoding steps in the two-sources problem were described in Section 4.

The first step is the analogue to the relay decoding in the single source problem. Given  $z^n$ , there are  $2^{nR_u} - 1$  incorrect messages  $w$  that each have probability  $2^{-n(I(X; Y_1|Z) - 3\epsilon)} 2^{-n(I(U; V) - \epsilon)}$  to be jointly typical with both  $y_1^n$  and  $V^n$ . When  $R_u < H(U)$ , the probability of error is driven arbitrarily small for large  $n$  under the condition  $I(X; Y_1|Z) > H(U|V)$ , which is Equation (6).

To determine the condition for low probability of error in decoding the bin number  $s$  at the receiver, examine the density  $p(y|z) = \sum_{x_1} p(y|x_1)p(x_1|z)$ . Thus, the probability of receiving a  $y^n$  that is jointly typical with a specific  $z^n$  other than the one at the relay is less than  $2^{-n(I(Y; Z) - 3\epsilon)}$ . The union bound on error  $2^{nR_0} 2^{-n(I(Y; Z) - 3\epsilon)}$  goes arbitrarily small as  $n$

increases if

$$R_0 < I(Y; Z). \quad (14)$$

In the third decoding step, the receiver has the estimated bin index  $\hat{s}$ , and has to choose among less than  $2^{n(R-R_0+\epsilon)}$  messages in that bin. The probability for each of the incorrect messages to be jointly typical with the received  $y^n$  is less than  $2^{-n(I(X;Y|Z)-3\epsilon)}$ , leading to the constraint of

$$R - R_0 < I(X; Y|Z). \quad (15)$$

Adding Equations (14) and (15) gives the condition in Equation (5).

Given the bin estimate  $\hat{s}$ , the receiver has  $2^{nR_1}$  codewords  $x_1^n$  to choose among. The probability that one of the incorrect codewords is jointly typical with  $y^n$  is less than  $2^{-n(I(X_1;Y|X,Z)-3\epsilon)}$ , which creates the condition

$$R_1 < I(X_1; Y|X, Z). \quad (16)$$

Since  $X$  and  $X_1$  are conditionally independent given  $Z$ ,  $I(X_1; Y|X, Z) = I(X_1; Y|Z)$ .

Finally, the probability that an incorrect  $V^n$  is jointly typical with the known  $U^n$  and in the bin  $k$  is less than the sum of the total number of such incorrect jointly typical sequences  $2^{n(H(V|U)+\epsilon)}$  multiplied by the probability that each is in this particular  $k$  bin, or  $2^{-nR_1}$ . This sum of  $2^{-nR_1}2^{n(H(V|U)+\epsilon)}$  goes to zero when  $R_1 > H(V|U) + \epsilon$ . Combining this condition with Equation(16), yields Equation (7). Summing Equations (5) and (7) lead to the total capacity equation  $H(U, V) < I(X, X_1; Y)$ , which can be maximized over the set of probability distributions considered to obtain Equation (4).

## References

- [1] T. M. Cover and A. A. El Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inform. Theory*, vol. IT-25, No.5, pp. 572-584, Sep 1979.
- [2] E. C. van der Meulen, "A survey of multi-way channels in information theory: 1961-1976," *IEEE Trans. Inform. Theory*, vol. IT-23, No. 1, pp. 1-37, Jan 1977.
- [3] E. C. van der Meulen, "Three-terminal communication channels," *Adv. Appl. Prob.*, vol. 3, pp. 120-154, 1971.
- [4] A. A. El Gamal and M. Aref, "The capacity of the semi-deterministic relay channel," *IEEE Trans. Inform. Theory*, vol. IT-28, No. 3, pg. 536, May 1982.
- [5] P. Gacs and J. Körner, "Common information is much less than mutual information," *Problems of Control and Information Theory*, vol. 2, pp. 149-162, 1973.
- [6] G. Kramer, M. Gastpar and P. Gupta, "Capacity theorems for wireless relay channels," *Proc. 41st Allerton Conf. on Commun., Control and Computing*, Oct 1-3, 2003.
- [7] R. Ahlswede and N. Cai and Z. Zhang, "On interactive communication," *IEEE Trans. Inform. Theory*, vol. 43, No. 1, pp. 22-37, Jan 1997.
- [8] J. Barros and S. D. Servetto, "The sensor reachback problem," submitted to *IEEE Trans. Inform. Theory*, Nov 2003, available at <http://cn.ece.cornell.edu/publications/papers/20031112/>.
- [9] G. Kramer, M. Gastpar and P. Gupta, "Information theoretic multi-hopping for relay networks," *2004 International Zurich Seminar*, pp. 192-195, Zurich, Switzerland, Feb. 18-20, 2004.
- [10] M. Gastpar and M. Vetterli, "On the capacity of wireless networks: the relay case," In *Proc IEEE InfoComm 2002*, New York, pp. 1577-1586, Jun. 23-27, 2002.

- [11] B. Wang and J. Zhang and A. Host-Madsen, "On capacity bounds of MIMO relay channel," submitted for publication, available at <http://www.eas.asu.edu/junshan/publications.html>.
- [12] R. Ahlswede and I. Csiszàr, "To get a bit of information may be as hard as to get full information," *IEEE Trans. Inform. Theory*, vol. IT-27, No 4, pp. 398-408, Jul. 1981.
- [13] T. Cover and A. A. El Gamal and M. Salehi, "Multiple access channels with arbitrarily correlated sources," *IEEE Trans. Inform. Theory*, vol. IT-26, No. 6, pp. 648-657, Nov. 1980.
- [14] G. Dueck, "A note on the multiple access channel with correlated sources," *IEEE Trans. Inform. Theory*, vol. IT-27, No. 2, pp. 232-235, Mar. 1981.
- [15] R. Ahlswede and T. S. Han, "On source coding with side information via a multiple-access channel and related problems in Multi-user information theory," *IEEE Trans. Inform. Theory*, vol. IT-29, No. 3, pp. 396-412, May 1983.
- [16] B. E. Schein, *Distributed Coordination in Network Information Theory*, Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA, Oct. 2001.
- [17] P. Gupta and P. R. Kumar, "The capacity of wireless networks," *IEEE Trans. Inform. Theory*, vol. 46, no. 3, pp. 388-404, Mar. 2000.
- [18] P. Gupta and P. R. Kumar, "Towards an information theory of large networks: An achievable rate region," *IEEE Trans. Inform. Theory*, vol. 49, pp. 1877-1894, Aug. 2003.
- [19] L.-L. Xie and P. R. Kumar, "A network information theory for wireless communication: scaling laws and optimal operation," *IEEE Trans. Inform. Theory*, vol. 50, no. 5, pp. 748-767, May 2004.
- [20] D. Slepian and J. K. Wolf, "A coding theorem for multiple access channels with correlated sources," *Bell Sys. Tech. J.*, vol. 52, pp. 1037-1076, Sept. 1973.
- [21] T. Cover, "A proof of the data compression theorem of Slepian and Wolf for ergodic sources," *IEEE Trans. Inform. Theory*, vol. 21, pp. 226-228, Mar. 1975.