

Capacity Analysis of the Relay Channel 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

Paper Type: Correspondence

Abstract—This paper emphasizes the importance of relay channels that possess correlated information at the transmitter and the relay as basic building blocks for sensor networks. This work characterizes sources which can be reliably transmitted over a relay channel. 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 scalar and MIMO Gaussian relay channels are used as illustrative examples. These cases 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.

Keywords— relay channel, correlated sources, sensor networks, block-Markov

I. INTRODUCTION

The memoryless relay channel, introduced by van der Meulen [1, 2] (as shown in Figure 1) is a channel consisting of an input X , a relay input 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 [3], semi-deterministic relay [4], and relay with feedback [3] have been characterized. Most of these characterizations are based on the block-Markov coding scheme, introduced by Cover and El Gamal in [3] to obtain the best known achievable region for this channel.

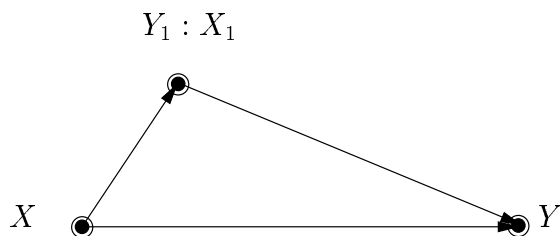


Fig. 1. The Relay Channel

Interest in relay channels and their capacity has seen a resurgence in recent years [5–11]. 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 [12]. However, the relay channel embedded in a sensor network has one major difference from that defined in [2]: 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 [13]. 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 2. Common randomness amongst sources has shown to aid in interactive communication in sensor networks [14, 15], and we show explicitly how this is so in the relay channel.

Transmitting correlated sources [16, 17] 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. [18] discuss the transmission of correlated sources over a multiple-access channel (MAC) and obtain an achievable region. However, the difficulty coordinating communication in a MAC, unless the two sources are completely correlated, is in isolating the common randomness between the two sources [19–21]. Ahlswede and Csiszár [22] 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 [3] with list-decoding at both the relay and the receiver to characterize a set of correlated sources which can be reliably communicated via the relay channel. 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 III, 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 IV we address the transmission of two sources simultaneously (the source and the side information) to the receiver. We then provide two examples of the relay channel with correlated sources in Section V. First, we study the scalar additive white Gaussian noise (AWGN) relay channel. Then, we demonstrate, for the AWGN MIMO channel, a case where the entropy of the source that we communicate over the relay channel achieves the cut-set upper bound on rates, and thus is the capacity of the channel.

II. THE RELAY CHANNEL

The conventional relay channel without side information is shown in Figure 1. For this channel, a cut-set upper bound on the rate which data can be reliably communicated from the transmitter to the receiver (using max-flow min-cut principles) was obtained in [3], and is given by:

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

Here \mathcal{C} denotes the channel capacity, or highest rate at which information can be sent with arbitrarily low probability of error [23].

Achievable rates for this channel are also obtained in [3, Theorems 1 & 7]. Here, we focus on the simpler block-Markov coding scheme utilized to obtain the rate described in [3, 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 [3].

Codebook Generation: Fix $p(x_1)p(x|x_1)$. Generate 2^{nR_0} i.i.d. x_1^n sequences as $\prod_{i=1}^n p(x_{1i})$ 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 $\prod_{i=1}^n p(x_i|x_{1i})$ 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 2^{nR_0} 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 $w_b \in S_{s_{b-1}}$. 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: Let $y_1^n(b)$ and $y^n(b)$ be the channel outputs received by the relay and receiver, respectively, in the b^{th} block. 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 $(x^n(\hat{w}_b|\hat{s}_{b-1}), x_1^n(\hat{s}_{b-1}), y_1^n(b)) \in A_\epsilon^{(n)}$, where $A_\epsilon^{(n)}$ is the set of jointly ϵ -typical sequences as defined in [23]. This decoding is correct with an arbitrarily small positive probability of error for n large and $R < I(X; Y_1|X_1)$ [23].
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 $R_0 < I(X_1; Y)$ [23].
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 $R < I(X, X_1; Y)$

[3]. These two bounds on R lead to the achievable lower-bound on capacity given by Equation (2).

III. SIDE INFORMATION AT THE RELAY

Consider the relay channel shown in Figure 2. 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 . Both U and V are considered to be discrete. We will use block-Markov coding and list-decoding as the basis of our analysis.

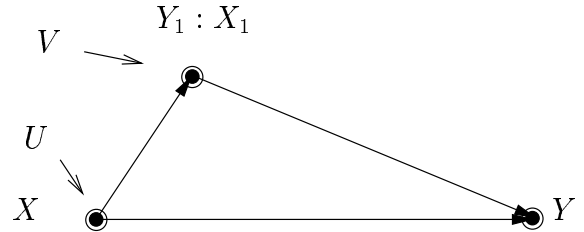


Fig. 2. The Relay Channel with Side Information

Theorem 1: The discrete memoryless source U , with (U, V) distributed as $p(u, v)$, can be sent with arbitrarily low probability of error over the relay channel if

$$H(U) < I(X, X_1; Y) \quad (3)$$

$$H(U|V) < I(X; Y_1|X_1) \quad (4)$$

are satisfied for a joint $p(x, x_1)$ and the source V is known at the relay.

Proof: The strategy is as follows:

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

Encoding: Generate 2^{nR} i.i.d. sequences U^n as $\prod_{i=1}^n p(u_i)$ and index them by $w \in \mathcal{W}$. The remainder of the encoding steps are identical to those in Section II.

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 II: 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. ■

Remark 1: From 1, we see that when $H(U|V) < I(X; Y_1|X_1)$, the completely 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, we see the same rate, or 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, as will be shown in Section V.

IV. 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 2, we communicate both U and V to the receiver. This system is analogous to a multiple access channel with correlated sources [18], with the additional aspect that one of the sources has the means of communicating with the other.

Theorem 2: A discrete memoryless source (U, V) distributed as $p(u, v)$ can be reliably decoded at the receiver of a relay channel if U is known at the transmitter, V is known at the relay, and for some $p(z)p(x|z)p(x_1|z)$, the following relations are satisfied:

$$H(U, V) < I(X, X_1; Y), \quad (5)$$

$$H(U) < I(X, Z; Y), \quad (6)$$

$$H(U|V) < I(X; Y_1|Z), \quad (7)$$

$$H(V|U) < I(X_1; Y|Z). \quad (8)$$

The coding and decoding arguments for this problem differ from the case when we only wish to transmit U (in Section III) because the codeword x_1^n has the added role of carrying information about V^n . In the coding arguments presented in Section III, 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.

Proof: 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 $\prod_{i=1}^n p(z)$ 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 $\prod_{i=1}^n p(x_i|z_i)$ 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 $\prod_{i=1}^n p(x_{1i}|z_i)$ 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 as $\prod_{i=1}^n p(u_i)$ indexed by $w \in \mathcal{W}$. Also generate 2^{nR_v} i.i.d. V^n as $\prod_{i=1}^n p(v_i)$ 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 III. 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 III. 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 III. 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 (5-8). They are derived in detail in the appendix. ■

V. EXAMPLES

We consider two examples which each apply the joint-source channel coding scheme described in this paper: the scalar and MIMO Gaussian relay channels with correlated sources at the transmitter and relay nodes. In both examples, we demonstrate achievable rates when we desire to transmit U to the receiver given side information V at the relay (as described in Section III). While the gains of joint-source channel coding are substantial in either case, there is a fundamental difference in the results. For the scalar Gaussian channel, joint source-channel coding clearly improves rate, but our method provides the explicit calculation of capacity only in the trivial case where the two sources are completely correlated. In the class of MIMO Gaussian relay channels, however, the capacity can be found if the correlation between the two sources exceeds a computable threshold.

The first example is the scalar Gaussian relay channel, as illustrated in Figure 3. Without loss of generality, this channel can be described by the relations

$$y = x + x_1 + \eta \quad \text{and} \quad y_1 = \beta x + \eta_1$$

with power constraints

$$E[X^2] \leq P \text{ and } E[X_1^2] \leq P_1,$$

where η and η_1 are realizations of independent zero-mean unit-variance Gaussian random variables \mathcal{N} and \mathcal{N}_1 .

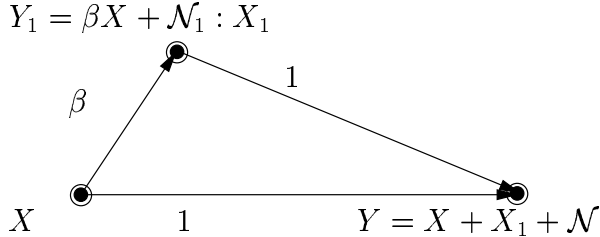


Fig. 3. The Gaussian Relay Channel

The achievable rates for a discrete memoryless relay channel with correlated sources are described by Equation (3). For the case of continuous channels (i.e. Gaussian), 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:

$$\begin{aligned} H(U) &< I(X, X_1; Y), \\ H(U|V) &< I(X; Y_1|X_1). \end{aligned} \quad (9)$$

The maximizing of the expression in Equation (9) 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 $U = V \oplus N$. The symbol \oplus indicates addition modulo M . Here, N is assumed to be independent of V . The mutual information between U and V is thus $I(U; V) = I(V + N; V) = H(V + N) - H(V + N|V) = H(U) - H(N)$, and for any distribution of the random variable N we can define α as $1 - H(N)/H(U)$. When N is identically zero U and V are perfectly correlated and α is 1; when the distribution of N is such that $H(U) = H(N)$ then α is zero and U and V are independent. Intermediate values of α monotonically characterize the degree of correlation between U and V for any set of distributions on N , resulting in $I(U; V) = \alpha H(U)$. Given the parameter α , we can write $H(U|V) = R(1 - \alpha)$ and thus Equation (9) is equivalent to

$$R \leq \sup_{F(x, x_1)} \min \left\{ I(X, X_1; Y), \frac{I(X; Y_1|X_1)}{1 - \alpha} \right\} \quad (10)$$

and is valid for both the scalar Gaussian case and the MIMO example investigated later.

Examine the first of two terms that appear inside the minimization in Equation (10): $I(X, X_1; Y) = H(Y) - H(\eta)$. For a given covariance, $H(Y)$ is maximized by a joint Gaussian distribution between X and X_1 . Likewise, the second term in Equation (10), $I(X; Y_1|X_1) = H(Y_1|X_1) - H(\eta_1)$ is also maximized by a joint Gaussian distribution. The jointly Gaussian

random variables X and X_1 are completely described by their covariance, and all that remains is to find the covariance (subject to the constraints of the channel) that fulfills the maximization in Equation (10).

Since we can only decrease the entropies by using less than the maximum available power in X and X_1 , distribute X_1 normally with zero mean and variance P_1 . Let γ describe the correlation between X and X_1 such that $p(x|x_1)$ is distributed normally with mean γx_1 and variance $P - \gamma^2 P_1$ where $0 \leq \gamma \leq \sqrt{P/P_1}$. According to the channel model, we get $I(X, X_1; Y) = \frac{1}{2} \log(1 + P + P_1 + 2\gamma P_1)$ and $I(X; Y_1|X_1) = \frac{1}{2} \log(1 + \beta^2(P - \gamma^2 P_1))$. Since the first term in Equation (10) is uniformly increasing in γ and the second is uniformly decreasing, the maximum occurs either at an endpoint of the γ range or where the two terms are equal. We can solve numerically to find the maximizing γ for given values of the other parameters.

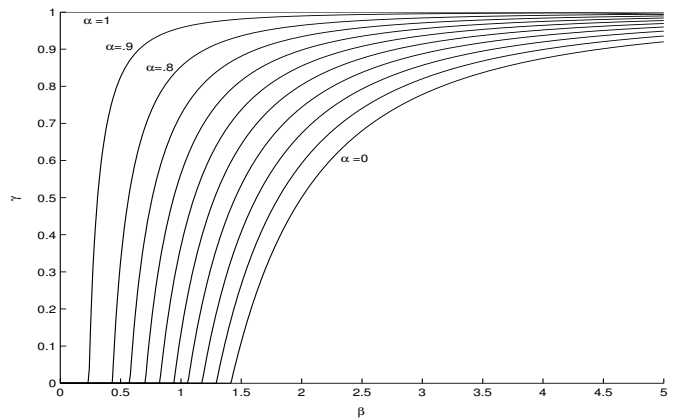


Fig. 4. Cooperation γ which Maximizes Rate $H(U)$ for Different Correlation Levels α versus Channel Gain β

A larger γ implies a larger correlation between X and X_1 , and allows a higher degree of cooperation between the transmitter and the relay. As the channel gain β between the transmitter and the relay increases, a larger extent of cooperation is possible, as it becomes easier to communicate U from the transmitter to the relay and ultimately to the receiver. For any particular value of β , the availability of the side information V at the relay allows a larger class of γ and a higher degree of cooperation between the transmitter and relay, as illustrated in Figure 4. Hence, a significantly higher data rate is possible to the receiver. In this figure equal unit powers are assumed at the transmitter and relay ($P = P_1 = 1$). As the side information V becomes more and more correlated with the source (as represented by the set of constant α curves), the amount of cooperation increases until X and X_1 can send exactly the same codewords, when $V = U$. Note that $\alpha = 1$ (complete correlation between sources) is the only case in the scalar Gaussian example where the relay and transmitter can cooperate completely (i.e. $\gamma = 1$), since for all other values $I(X; Y_1|X_1) < H(U|V)$ for the maximizing value of γ .

In contrast, consider a Gaussian relay channel with two receive antennas and a single antenna each at the transmitter and relay, illustrated in Figure 5. Full data cooperation and hence capacity is achieved on this channel if the correlation between

the sources U and V exceeds a certain threshold value. In this section, we find this threshold value for a particular example MIMO relay channel, showing that the capacity of a non-trivial set of channels can be characterized using our strategy.

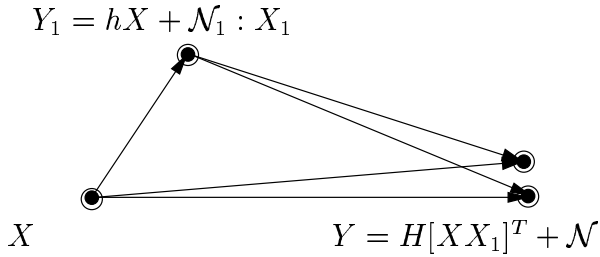


Fig. 5. The MIMO Gaussian Relay Channel

This channel can be represented as:

$$\begin{aligned} Y_1 &= hX + \mathcal{N}_1 \\ Y &= H[X \ X_1]^T + \mathcal{N}. \end{aligned}$$

where H is a 2×2 matrix and Y & \mathcal{N} are 2×1 vectors. The noise at both the relay and receiver is assumed to be Gaussian with zero mean and unit variance. Also, the transmitter and relay possess separate average power constraints of P and P_1

Consider a specific example MIMO channel given by:

$$\begin{aligned} Y_1 &= 3X + \mathcal{N}_1 \\ Y &= \begin{bmatrix} 1 & 1.5 \\ .5 & 2.5 \end{bmatrix} [X \ X_1]^T + \mathcal{N} \end{aligned}$$

and a power constraint of $P = P_1 = 2$ on each of the transmitters. This example was picked to reflect the scenario where the relay is closer to each of the receiver and the transmitter than they are to each other. 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

$$I(X, X_1; Y). \quad (11)$$

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

$$\Sigma_{X, X_1} = \begin{bmatrix} 2 & 0.898 \\ 0.898 & 2 \end{bmatrix} \quad (12)$$

between X and X_1 maximizes $I(X, X_1; Y)$, leading to a value of 2.569 bits/use. The covariance matrix in Equation 12) was found numerically, as $I(X, X_1; Y)$ is here a convex function of the cross-covariance term. For this choice of the joint distribution between X and X_1 (given by 12),

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

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. It is also unlike the scalar Gaussian relay channel, where correlation enables some, but never (in a non-trivial sense) *full*, data cooperation.

VI. 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. This cooperation takes the form of correlation between the relay and transmitter codebooks. 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 [23], is less than ϵ . 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

$$\epsilon + \sum_{i=2}^{2^{nR}} 2^{-n(I(X; Y_1 | X_1) - 3\epsilon)} 2^{-n(I(U; V) - \epsilon)} \quad (13)$$

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

$$I(X; Y_1 | X_1) > H(U|V). \quad (14)$$

The probability of error analysis for every other step in the decoding process, including the derivation of the second requirement for achievability in Equation (3), $H(U) < I(X, X_1; Y)$, is exactly the same as in the block-Markov scheme described in Theorem 2 of [3].

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

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 $2^{-n(I(X; Y_1|Z) - 3\epsilon)} \times 2^{-n(I(U; V) - \epsilon)}$ 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 (7).

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

$$R_0 < I(Y; Z). \quad (15)$$

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

$$R - R_0 < I(X; Y|Z). \quad (16)$$

Adding Equations (15) and (16) gives the condition in Equation (6).

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

$$R_1 < I(X_1; Y|X, Z). \quad (17)$$

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

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(17), yields Equation (8). Summing Equations (6) and (8) 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 (5).

REFERENCES

[1] 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.
 [2] E. C. van der Meulen, "Three-terminal communication channels," *Adv. Appl. Prob.*, vol. 3, pp. 120-154, 1971.

[3] 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.
 [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] 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.
 [6] 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.
 [7] 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>.
 [8] B. E. Schein, *Distributed Coordination in Network Information Theory*, Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA, Oct. 2001.
 [9] P. Gupta and P. R. Kumar, "The capacity of wireless networks," *IEEE Trans. Inform. Theory*, vol. 46, no. 3, pp. 388-404, Mar. 2000.
 [10] 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.
 [11] 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.
 [12] 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.
 [13] G. Pottie and W. Kaiser, "Wireless integrated sensor networks," *Communications of the ACM*, no. 43, pp. 51-58, 2000.
 [14] R. Ahlswede and N. Cai and Z. Zhang, "On interactive communication," *IEEE Trans. Inform. Theory*, vol. 43, No. 1, pp. 22-37, Jan 1997.
 [15] 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/>.
 [16] 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.
 [17] 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.
 [18] 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.
 [19] 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.
 [20] 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.
 [21] 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.
 [22] R. Ahlswede and I. Csiszár, "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.
 [23] T. Cover and J. Thomas, *Elements of Information Theory*, Wiley, New York 1991.