Relay with Side Information

Ryoulhee Kwak, Wooyul Lee, Abbas El Gamal and John M. Cioffi
Department of Electrical Engineering
Stanford University
Stanford, CA 94305, USA
E-mail: rhkwak@stanford.edu, wylee@stanford.edu, abbas@isl.stanford.edu, cioffi@stanford.edu

***Abstract*— This paper establishes necessary and sufficient conditions for reliable transmission of a source over a relay channel when source side information is available non-causally (a) only at the receiver, (b) only at the relay, or (c) at both the relay and the receiver. For the cases of side information only at the receiver and at both the relay and the receiver, we establish tight necessary and sufficient conditions that apply to any relay channel and show that source-channel separation is optimal. When side information is available only at the relay, we establish a necessary condition for reliable transmission and show that it is tight for the class of degraded relay channels and that source-channel separation is optimal in this case.**

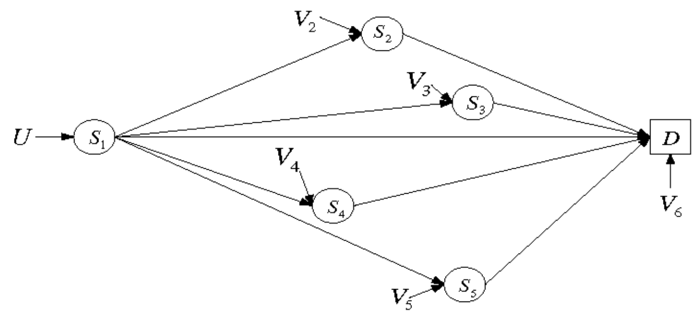

Fig. 1. Sensor network with correlated measurements.

I. Introduction

The general problem of multiple user information theory is that of reliable transmission of correlated sources over noisy channels. In [1], Shannon showed that source-channel separation is optimal for sending a source over a channel. This is not in general the case for multiple user scenarios. In [2], it was shown that source-channel separation is not in general optimal for sending correlated sources over a multiple access channel. The paper also proposed a joint source-channel coding scheme, which was later shown not to be optimal in general [3]. In [4], the problem of sending a source over a broadcast channel with receiver side information is studied, where a sufficient condition is proved using a joint source-channel approach [4].

In this paper we study the problem of sending a source over a relay channel with source side information available non-causally at the relay and/or at the receiver. This problem is motivated by the sensor network setting depicted in Fig. 1 in which a sensor S_1 is queried about its measurement by a data collection center D. To send S_1's measurement to D, other nodes in the network can be used as relays. In many applications the nodes, including the collection center, may have measurements $V_2, \dots, V_6$ that are correlated with that of S_1. How should these measurements be used to help in transmitting S_1's measurement to D?

The problem of sending a source over a relay channel with side information available only at the relay was studied in [5]. The paper established a sufficient condition for reliable transmission of a source that uses block-Markov and list-decoding and showed that the correlated source at the relay can increase the transmission rate. This sufficient condition was proved using a joint source-channel scheme.

The remainder of this paper is organized as follows. In Section II, we provide the needed definitions and review the main results in the paper. In Section III, we study the case when side information is available only at the receiver. In Section IV, we study the case when side information is available at both the relay and the receiver and in Section V, we study the case when the side information is available only at the relay.

II. Definitions and Main Results

Let $(U, V) \sim p(u, v)$ be a pair of correlated i.i.d. sources and $(\mathcal{X}, \mathcal{X}_1, p(y, y_1|x, x_1), \mathcal{Y}, \mathcal{Y}_1)$ be a discrete-memoryless relay channel. We wish to send U reliably from X to Y at a rate of one symbol per transmission in the presence of side information V available non-causally (a) only at the receiver Y, (b) only at the relay (X_1, Y_1), or (c) at both the receiver and the relay as illustrated in Fig. 2.

More precisely, we define a $(2^{nR}, n)$ code for the *relay-with-side-information* to consist of:

1) An encoder that maps each sequence $u^n \in \mathcal{U}^n$ into a codeword $x^n(u^n)$.
2) A set of relaying functions defined as $x_{1i} = f_i(y_1^{i-1})$, $i = 1, 2, 3, \dots$, for scenario (a), and as $x_{1i} = f_i(v^n, y_1^{i-1})$, $i = 1, 2, 3, \dots$, for scenarios (b) and (c).
3) A decoder that maps each received sequence pair (y^n, v^n) into an estimate $\hat{u}^n$ in scenarios (a) and (c) and maps each received sequence y^n to an estimate $\hat{u}^n$ in scenario (b).

The average probability of decoding error is defined as

$$P_e^{(n)} = P\{\hat{U}^n \neq U^n\}.$$

The source U can be sent reliably over the relay channel if there exists a sequence of $(2^{nR}, n)$ codes with $P_e^{(n)} \to 0$ as $n \to \infty$. We seek to find necessary and sufficient conditions

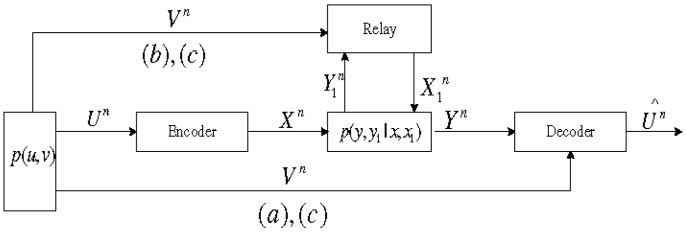

Fig. 2. Relay channels with source side information V^n.

for reliable transmission of U over the relay channel for each of the aforementioned side information scenarios. We also seek to answer the question of whether source-channel separation is optimal for each of these scenarios.

In the following, we shall denote the capacity of the relay channel $(\mathcal{X}, \mathcal{X}_1, p(y, y_1|x, x_1), \mathcal{Y}, \mathcal{Y}_1)$ by C and recall the multi-letter expression for C in [6]

$$C = \lim_{n\to\infty} C_n,$$

where

$$C_n = \frac{1}{n} \max_{p(x^n),\ \{f_i(y_1^{i-1})\}} I(X^n; Y^n).$$

Note that without side information, U can be reliably transmitted over the relay channel iff $H(U) \le C$, i.e., if $H(U) < C$, then there exists a sequence of $(2^{nR}, n)$ codes with $P_e^{(n)} \to 0$ as $n \to \infty$, and $H(U) \le C$ for any sequence of $(2^{nR}, n)$ codes with $P_e^{(n)} \to 0$ as $n \to \infty$. The proof of this result is straightforward and therefore will not be included.

Main Results

The main results of this paper are as follows:

- For scenario (a) discussed in Section III, we show that a source can be transmitted over a general relay channel iff $H(U|V) \le C$. This shows that source-channel separation is optimal in general for this case.
- For scenario (c) discussed in Section IV, we again show that a source can be sent over a general relay channel iff $H(U|V) \le C$ and that source-channel separation is optimal in general.
- For scenario (b) discussed in Section V, we show that for a source U to be reliably transmitted over a general relay channel, it is necessary that

$$H(U) \le \max_{p(x,x_1)} \min\{I(X, X_1; Y),\\ I(X; Y_1|X_1) + I(U, V)\}. \quad (1)$$

We show that this condition is sufficient for the class of degraded relay channels. The proof of sufficiency follows from the result in [5]. We provide a proof, which uses separate source and channel coding that shows the optimality of source-channel separation for this class.

III. Side Information only at the Receiver

We establish a necessary and sufficient condition for scenario (a), that is, when the side information V^n is available only at the receiver Y.

Theorem 1: A necessary and sufficient condition for U to be reliably sent over the discrete-memoryless relay channel $(\mathcal{X}, \mathcal{X}_1, p(y, y_1|x, x_1), \mathcal{Y}, \mathcal{Y}_1)$ with capacity C is given by $H(U|V) \le C$.

Proof: To establish sufficiency we use separate source and channel coding. The encoder performs Slepian-Wolf encoding followed by optimal channel encoding of the bin index, and the relay uses its optimal set of functions. The decoder first decodes the bin index, then performs Slepian-Wolf decoding. Clearly this can be performed reliably if $H(U|V) < C$.

To show the necessity of the condition, we first apply Fano's inequality to obtain

$$H(U^n|\hat{U}^n) \le n\epsilon_n,$$

for some $\epsilon_n \to 0$ as $n \to \infty$. Next consider

$$\begin{aligned} nH(U) &= I(U^n; \hat{U}^n) + H(U^n|\hat{U}^n) \\ &\le I(U^n; \hat{U}^n) + n\epsilon_n \\ &\le I(U^n; Y^n, V^n) + n\epsilon_n \\ &= I(U^n; Y^n) + I(U^n; V^n|Y^n) + n\epsilon_n \\ &= I(U^n; Y^n) + H(V^n) - H(V^n|Y^n, U^n) + n\epsilon_n \\ &\overset{(a)}{=} I(U^n; Y^n) + H(V^n) - H(V^n|U^n) + n\epsilon_n \\ &\overset{(b)}{\le} I(X^n; Y^n) + nI(U; V) + n\epsilon_n \\ &\overset{(c)}{\le} n(C + I(U; V) + \epsilon_n). \end{aligned}$$

Step (a) follows from the fact that $V^n \to U^n \to Y^n$ form a Markov chain, step (b) follows from the fact that $U^n \to X^n \to Y^n$ form a Markov chain, and (c) follows by the multi-letter definition of relay channel capacity in Section II. Thus as $n \to \infty$, $H(U|V) \le C$. This completes the proof of the theorem. ■

Note that proof of the above theorem also shows that source-channel separation is optimal for this scenario.

IV. Side Information at both the Relay and the Receiver

We now consider scenario (c), where the side information is available at both the relay and the receiver.

Theorem 2: A necessary and sufficient condition for sending U reliably over a discrete-memoryless relay channel when the side information V is available at both the relay and the receiver is given by

$$H(U|V) \le C. \quad (2)$$

Proof: To establish sufficiency we use the same scheme as in scenario (a). Because the condition is same as that of scenario (a), $H(U|V) < C$ is achievable.

To show the necessity of the condition, we first apply Fano's inequality to obtain inequality

$$H(U^n|V^n, Y^n) \le n\epsilon_n,$$

where $\epsilon_n \to 0$ as $n \to \infty$. Now consider

$$\begin{aligned} nH(U|V) &\leq I(U^n; Y^n|V^n) + n\epsilon_n \\ &\leq I(X^n; Y^n|V^n) + n\epsilon_n \\ &= \sum_{v^n} p(v^n) I(X^n; Y^n|v^n) + n\epsilon_n \\ &\stackrel{(a)}{\leq} \sum_{v^n} p(v^n) \max_{p(x^n|v^n),\ \{f_i(y_1^{i-1}, v^n)\}} I(X^n; Y^n|v^n) \\ &+ n\epsilon_n \\ &= nC_n + n\epsilon_n, \end{aligned}$$

where the last step follows by the fact that the max term in the previous step is the same as $\max_{p(x^n),\ \{f_i(y_1^{i-1})\}} I(X^n; Y^n) = nC_n$. To see this note that the mutual information term in (a) is a function of the conditional pmf

$$p(x^n|v^n) \prod_{i=1}^n p(x_{1i}|y_1^{i-1}, v^n) \prod_{i=1}^n p(y_i, y_{1i}|x_i, x_{1i}).$$

Thus v^n is simply a label and does not change the range of $p(x^n|v^n),\ \{f_i(y_1^{i-1}, v^n)\}$ in the maximization. Hence as $n \to \infty$, $H(U|V) \leq C$. ■

The above theorem shows that source-channel separation is also optimal for scenario (c).

V. Side Information only at the Relay

Here we consider scenario (b), where side information V is available only at the relay. In the following subsection we establish a general necessary condition for reliable transmission of U over the relay channel. In Subsection V-B, we show that this necessary condition is also sufficient for degraded channels and that source-channel separation is optimal in this case.

A. Necessary condition

We establish the following necessary condition for scenario (b).

Theorem 3: A necessary condition for the source U to be sent reliably over a discrete-memoryless relay channel $(\mathcal{X}, \mathcal{X}_1, p(y, y_1|x, x_1), \mathcal{Y}, \mathcal{Y}_1)$ with side information V available non-causally only at the relay is given by

$$H(U) \leq \max_{p(x,x_1)} \min\{I(X, X_1; Y), I(X; Y_1, Y|X_1) + I(U, V)\}. \quad (3)$$

Proof: By Fano's inequality

$$H(U^n|\hat{U}^n) \leq n\epsilon_n,$$

for some $\epsilon_n \to 0$ as $n \to \infty$. To show the first inequality, consider

$$\begin{aligned} nH(U) &\leq I(U^n; \hat{U}^n) + n\epsilon_n \\ &\leq I(U^n; Y^n) + n\epsilon_n \\ &= \sum_{i=1}^n I(U^n; Y_i|Y^{i-1}) + n\epsilon_n \\ &= \sum_{i=1}^n (H(Y_i) - H(Y_i|U^n, Y^{i-1})) + n\epsilon_n \\ &\leq \sum_{i=1}^n (H(Y_i) - H(Y_i|X_i, X_{1i}, U^n, Y^{i-1})) + n\epsilon_n \\ &= \sum_{i=1}^n (H(Y_i) - H(Y_i|X_i, X_{1i})) + n\epsilon_n \\ &= \sum_{i=1}^n I(X_i, X_{1i}; Y_i) + n\epsilon_n. \end{aligned}$$

To show the second inequality, consider

$$\begin{aligned} nH(U|V) &\leq I(U^n; \hat{U}^n|V^n) + n\epsilon_n \\ &\leq I(X^n; Y^n, Y_1^n|V^n) + n\epsilon_n \\ &= \sum_{i=1}^n I(X^n; Y_i, Y_{1i}|V^n, Y_1^{i-1}, Y^{i-1}) + n\epsilon_n \\ &\stackrel{(a)}{=} \sum_{i=1}^n I(X^n; Y_i, Y_{1i}|V^n, Y_1^{i-1}, X_{1i}, Y^{i-1}) + n\epsilon_n \\ &\leq \sum_{i=1}^n (H(Y_i, Y_{1i}|X_{1i}) - H(Y_i, Y_{1i}|X_{1i}, X_i)) + n\epsilon_n \\ &= \sum_{i=1}^n I(X_i; Y_i, Y_{1i}|X_{1i}) + n\epsilon_n, \end{aligned}$$

where (a) follows from the fact that $X_{1i} = f_i(V^n, Y_1^{i-1})$. ■

B. Degraded Relay

We show that the necessary condition in the previous subsection is tight for the degraded relay channel, that is, when the channel conditional pmf is given by

$$p(y, y_1|x, x_1) = p(y_1|x, x_1)p(y|y_1, x_1),$$

for all (x, x_1, y, y_1). Note that for the degraded relay, the necessary condition in Theorem 3 reduces to

$$H(U) \leq \max_{p(x,x_1)} \min\{I(X, X_1; Y), I(X; Y_1|X_1) + I(U, V)\}. \quad (4)$$

We now show that this condition is also sufficient.

Proposition 1: The necessary condition in Theorem 3 is tight for the degraded relay channel.

Proof: The achievability of the above necessary condition was proved using a joint source-channel scheme in [5]. We show that it can also be achieved using separate source and channel coding scheme. This is acheived by a combination of Slepian-Wolf source coding and decode-and-forward [7]. The set of u^n sequences is partitioned into roughly $2^{n(H(U|V)+\epsilon)}$

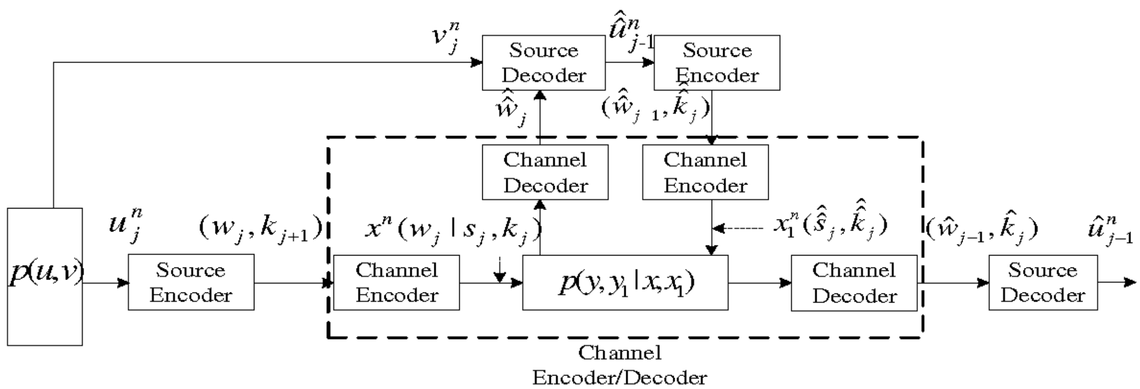

Fig. 3. Source-channel separation for the degraded relay channels with source side information V^n at the relay.

bins. Each typical sequence $u^n \in A_\epsilon^{(n)}$, where $A_\epsilon^{(n)}$ refers to the set of weakly ϵ-jointly typical sequences as defined in [8] is then assigned a unique pair of indices, its bin number w and its index number k in the bin (non-typical u^n sequences are assigned $k = 1$). In block j, the sender sends the bin number of the sequence u_j^n and additional partial information about the bin number of u_{j-1}^n to cooperate with the relay to resolve the receiver's uncertainty about u_{j-1}^n. The relay first decodes the bin number of u_j^n, then finds u_j^n by using the side information v_j^n. Simultaneously, the relay cooperates with the sender to resolve receiver's uncertainty about u_{j-1}^n in block j. This scheme is illustrated in Fig. 3.

We now provide a more detailed outline of the proof.

Generation of codebooks: For the source codebook, randomly partition the set u^n into $2^{n(H(U|V)+\epsilon)}$ bins. In each bin, there are almost $2^{n(I(U;V)+\epsilon)}$ u^n typical sequences. Assign a unique index $w \in [1, 2^{n(H(U|V)+\epsilon)}]$ to each bin and assign a unique index $k \in [1, 2^{n(I(U;V)+\epsilon)}]$ to each typical u^n sequence in each bin. Assign the index $k = 1$ to non-typical u^n sequences.

For the channel codebook, fix $p(x_1)p(x|x_1)$. Generate at random $2^{n(R_0+I(U;V)+\epsilon)}$ i.i.d $x_1^n(s,k)$, $s \in [1, 2^{nR_0}]$ according to $\prod_{i=1}^n p(x_{1i})$. For each $x_1^n(s,k)$, generate $2^{n(H(U|V)+\epsilon)}$ conditionally independent $x^n(w|s,k)$ sequences according to $\prod_{i=1}^n p(x_i|x_{1i})$. For each bin number w, assign an index $s(w)$ at random from $[1, 2^{nR_0}]$. The set of bin numbers with the same index s forms a bin $S_s \subseteq [1, 2^{n(H(U|V)+\epsilon)}]$.

Encoding: In block j, the source encoder at the sender finds the bin number for the new source sequence u_j^n and the index of u_j^n in that bin. Let w_j be the source message bin number to be sent in block j, and k_j be the index of u_{j-1}^n in block $j-1$. The sender source encoder outputs (w_j, k_{j+1}) to the channel encoder. Assume that $w_{j-1} \in S_{s_j}$. The channel encoder at the sender sends $x^n(w_j|s_j, k_j)$.

The source decoder at the relay has an estimate of u_{j-1}^n and outputs $\hat{\hat{u}}_{j-1}^n$ to the source encoder at the relay. The source encoder at the relay estimates $(\hat{\hat{w}}_{j-1}, \hat{\hat{k}}_j)$ of (w_{j-1}, k_j) and sends $(\hat{\hat{w}}_{j-1}, \hat{\hat{k}}_j)$ to the channel encoder at the relay. Assume that $\hat{\hat{w}}_{j-1} \in S_{\hat{\hat{s}}_j}$, then the channel encoder at the relay sends $x_1^n(\hat{\hat{s}}_j, \hat{\hat{k}}_j)$ in block j.

Decoding : Assume that at the end of block $j-1$, the channel decoder at the receiver knows $(w_1, w_2, ..., w_{j-2})$, $(s_2, ..., s_{j-1})$, and $(k_2, ..., k_{j-1})$, and the channel decoder at the relay knows $(w_1, w_2, ..., w_{j-1})$, $(s_2, ..., s_j)$, and $(k_2, ..., k_j)$. The decoding procedures at the end of block j are as follows.

1) Upon receiving $y_1^n(j)$, the channel decoder at the relay declares that $\hat{\hat{w}}_j$ was sent if it is the unique index such that $(x^n(\hat{\hat{w}}_j|\hat{\hat{s}}_j, \hat{\hat{k}}_j), x_1^n(\hat{\hat{s}}_j, \hat{\hat{k}}_j), y_1^n(j)) \in A_\epsilon^{(n)}$, otherwise an error is declared. The source decoder at the relay then finds the unique index $\hat{\hat{k}}_{j+1}$ such $(u^n(\hat{\hat{w}}_j, \hat{\hat{k}}_{j+1}), v^n) \in A_\epsilon^{(n)}$. It can be shown that $\hat{\hat{w}}_j = w_j$ and a unique index k_{j+1} is found with an arbitrarily small probability of error if n is sufficiently large and
$$H(U|V) < I(X;Y_1|X_1).$$
Adding $I(U;V)$ to both sides gives
$$H(U) < I(X;Y_1|X_1) + I(U;V).$$

2) The channel decoder at the receiver declares that $(\hat{s}_j, \hat{k}_j)$ was sent if it is the unique index such that $(x_1^n(\hat{s}_j, \hat{k}_j), y^n(j)) \in A_\epsilon^{(n)}$, otherwise an error is declared. The indices (s_j, k_j) can be decoded with arbitrarily small probability of error if n is sufficiently large and
$$R_0 + I(U;V) < I(X_1;Y).$$

3) Assuming that s_j is decoded correctly at the receiver, it constructs the list $L(y^n(j-1))$ of message indices whose codewords are jointly typical with $y^n(j-1)$ in block $j-1$. The channel decoder at the receiver then declares that $\hat{w}_{j-1}$ is sent in block $j-1$ if it is the unique bin number in $S_{s_j} \cap L(y^n(j-1))$, otherwise an error is declared. It can be shown that $\hat{w}_{j-1} = w_{j-1}$ with an arbitrarily small probability of error if n is sufficiently large and
$$\begin{aligned} H(U|V) &< I(X;Y|X_1) + R_0 \\ &< I(X;Y|X_1) + I(X_1;Y) - I(U;V). \end{aligned}$$

Thus

$$H(U) < I(X, X_1; Y).$$

Now, given k_j and w_{j-1}, the u_{j-1}^n sequence can be decoded with arbitrarily small probability of error at the source decoder of the receiver.

■

The above result shows that source-channel separation is optimal for the degraded relay channel as shown in Fig. 3.

VI. Conclusion

We showed that a source U can be reliably transmitted over a discrete-memoryless channel with capacity C in the presence of side information V only at the receiver or at both the relay and the receiver iff $H(U|V) \leq C$. This also shows that source-channel separation is optimal for these cases. When the side information is available only at the relay, we established a necessary condition for reliable transmission of U. We showed that it is tight when the channel is degraded by using a combination of Slepian-Wolf coding and decode-and-forward. A tighter sufficient condition can be found by using a combination of lossless source coding with side information [9] and partial decode-and-forward. We do not believe, however, that separate source-channel coding is in general optimal when side information is available only to the relay.

References

[1] C. E. Shannon, "A Mathematical Theory of Communication," *Bell Sys. Tech. Journal*, vol. 27, pp. 379-423, pp. 623-656, Jul. Oct. 1948.

[2] T. M. Cover, A. A. El Gamal, and M. Salehi, "Multiple Access Channels with Arbitrarily Correlated Sources," *IEEE Trans. Info. Theory*, vol. IT-26, pp. 648-657, Nov. 1980.

[3] G. Dueck, "A Note on the Multiple Access Channel with Correlated Sources," *IEEE Trans. Info. Theory*, vol. IT-27, pp. 232-235, Mar. 1981.

[4] E. Tuncel, "Lossless Joint Source-Channel Coding Across Broadcast Channels with Decoder Side Information," *Proc. IEEE ISIT*, pp. 641-645, Sep. 2005.

[5] B. Smith, and S. Vishwanath, "Cooperative Communication in Sensor Networks: Relay Channel with Correlated Sources," *Proc. 42nd Annu. Allerton Conf. on Commun., Control and Computing*, Sep. 2004.

[6] E. C. van der Meulen, "Three-terminal Communication Channels," *Adv. Appl. Prob.*, vol. 3, pp. 120-154, 1971.

[7] A. A. El Gamal, *Lecture notes for EE478: Network Information Thery,* Stanford University.

[8] T. M. Cover, and J. A. Thomas, *Elements of Information Theory,* Wiley, Newyork, 1991.

[9] A. Wyner, "On Source Coding with Side Information at the Decoder," *IEEE Trans. Info. Theory,* vol. IT-21, pp. 163-179, 1975.