THE BELL SYSTEM
TECHNICAL JOURNAL

DEVOTED TO THE SCIENTIFIC AND ENGINEERING
ASPECTS OF ELECTRICAL COMMUNICATION

Volume 53 November 1974 Number 9

Source Coding for a Simple Network

By R. M. GRAY and A. D. WYNER

(Manuscript received March 15, 1974)

We consider the problem of source coding subject to a fidelity criterion for a simple network connecting a single source with two receivers via a common channel and two private channels. The region of attainable rates is formulated as an information-theoretic minimization. Several upper and lower bounds are developed and shown to actually yield a portion of the desired region in certain cases.

I. INTRODUCTION

1.1 Informal statement of the problem

To fix ideas, let us consider the following problem. Suppose that we are given a data source whose output is a sequence $U_1, U_2, \cdots$, that appears at the source output at the rate of 1 per second. The $\{U_k\}_1^\infty$ is a sequence of independent copies of the discrete random variable U, with probability distribution $\Pr\{U = u\} = Q(u)$, $u \in \mathfrak{U}$ a finite set. Our task is to transmit this data sequence over a communication channel having a capacity of C bits per second so that it is represented at the output as $\hat{U}_1, \hat{U}_2, \cdots, \in \mathfrak{U}$. We assume that the data are trans-

The work of R. M. Gray was supported in part by NSF Grants GK-5452 and GK-31630 and by the Joint Services Program at Stanford Electronics Laboratory and U.S. Navy Contract N0014-67-A-112-004. Parts of this work were performed while A. D. Wyner was visiting Stanford University with the partial support of the ISL Stanford Affiliates.

mitted over the channel in blocks of length n, and allow processing at both the channel input and output (encoding and decoding). We define the "error rate" as

$$\Delta = E \frac{1}{n} \sum_{k=1}^{n} d_H(U_k, \hat{U}_k), \tag{1a}$$

where

$$d_H(u, \hat{u}) = \begin{cases} 0, & u = \hat{u}, \\ 1, & u \neq \hat{u}, \end{cases} \tag{1b}$$

is the Hamming metric. Thus, Δ is the average fraction of data digits delivered in error.

The question we pose is: What is the smallest capacity C such that (for n sufficiently large) we can transmit the data through the channel and achieve an arbitrarily small Δ? The well-known answer to the question is that the minimum capacity C is the entropy $H(U)$, defined by*

$$H(U) = - \sum_{u \in \mathfrak{U}} Q(u) \log Q(u). \tag{2}$$

Now consider the case where the random variable U is a pair (X, Y) where $x \in \mathfrak{X}$ and $y \in \mathfrak{Y}$. We have

$$Q(u) = Q(x, y) = \Pr \{X = x, Y = y\},$$

and

$$H(U) = H(X, Y) = - \sum_{x,y} Q(x, y) \log Q(x, y).$$

Setting $\hat{U} = (\hat{X}, \hat{Y})$, Δ [as defined in (1)] is the fraction of pairs delivered in error. Thus, we conclude that $H(X, Y)$ is the minimum channel capacity required to transmit the source output $\{(X_k, Y_k)\}$ with the error rate Δ arbitrarily small.

Next, let us assume that, as above, $U = (X, Y)$, but that it is only required to transmit the sequence $\{X_k\}$ through a channel having a capacity C_1, and to deliver it at the channel output as $\{\hat{X}_k\}$. Let

$$\Delta_X = E \frac{1}{n} \sum_{k=1}^{n} d_H(X_k, \hat{X}_k)$$

be the error rate for a system with block coding of block length n. The special assumption here is that the random sequence $\{Y_k\}_{k=1}^{n}$ is available to the encoder and the decoder. See Fig. 1.

* All logarithms in this paper are assumed to be taken to the base 2.

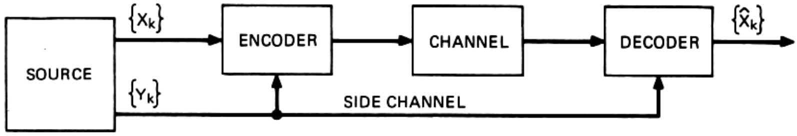

Fig. 1—Source coding with side information.

Again we ask: What is the minimum capacity C_1 required to transmit $\{X_k\}$ with Δ_X arbitrarily small (with n sufficiently large)? The answer[1,2] is that the minimum C_1 is the "conditional entropy," $H(X|Y)$, defined by

$$H(X|Y) = -\sum_{x,y} Q(x, y) \log \frac{Q(x, y)}{Q_Y(y)}$$
$$= -\sum_{y} Q_Y(y)[\sum_{x} Q_{X|Y}(x|y) \log Q_{X|Y}(x|y)], \tag{3a}$$

where

$$Q_Y(y) = \Pr\{Y = y\} = \sum_{x \in \mathfrak{X}} Q(x, y) \tag{3b}$$

and

$$Q_{X|Y}(x|y) = \frac{Q(x, y)}{Q_Y(y)} = \Pr\{X = x | Y = y\}. \tag{3c}$$

Note that $H(X|Y) + H(Y) = H(X, Y)$.

Let us remark that the above still holds if, instead of delivering $\{Y_k\}$ to the decoder, we delivered a sequence $\{\hat{Y}_k\}$, where $\Delta_Y = E(1/n)\sum_{k=1}^{n} d_H(Y_k, \hat{Y}_k)$ can be made arbitrarily small. Thus, the capacity of the "side channel" must be at least $H(Y)$.

Finally, we turn our attention to the problem to which this paper is devoted. Let the source output be $\{(X_k, Y_k)\}_{k=1}^{\infty}$, as above. We assume here, however, that there are *two* receivers. Receiver 1 is interested in obtaining a reproduction $\{\hat{X}_k\}$ of the sequence $\{X_k\}$, and receiver 2 is interested in obtaining a reproduction $\{\hat{Y}_k\}$ of the sequence $\{Y_k\}$. Assume further that a network consisting of three channels is available, as in Fig. 2. The first of these channels is a "common" channel (with capacity C_0) that connects the transmitter to both receivers, and the other two are "private" channels that connect the transmitter to each of the two receivers (with capacities C_1 and C_2). Assuming that we use block coding with block length n, the error rates are

$$\Delta_X = E\frac{1}{n}\sum_{k=1}^{n} d_H(X_k, \hat{X}_k) \tag{4a}$$

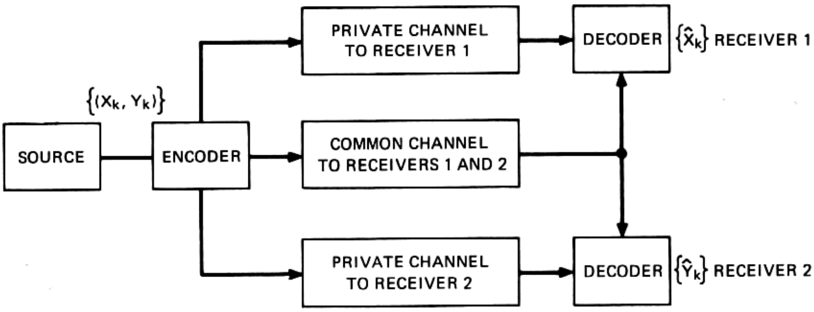

Fig. 2—Source coding for a network.

and

$$\Delta_Y = E\frac{1}{n}\sum_{k=1}^{n} d_H(Y_k, \hat{Y}_k). \tag{4b}$$

We say that a "rate-triple" (R_0, R_1, R_2) is *achievable* if, for any triple of channel capacities (C_0, C_1, C_2) for which $C_i > R_i$ $(i = 0, 1, 2)$ and any $\epsilon > 0$, transmission over the network of Fig. 2 (with these capacities) is possible (with n sufficiently large) with $\Delta_X, \Delta_Y \leqq \epsilon$. Our problem is the determination of the set $\mathcal{R}$ of achievable rate-triples.

Before stating our results, we digress to give a formal and precise statement of the problem as well as some other specialized information. This digression can be omitted by the casual reader.

1.2 Digression—formal statement of the problem

Let $\{(X_k, Y_k)\}_{k=1}^{\infty}$ be a sequence of independent drawings of a pair of random variables (X, Y), $X \in \mathcal{X}$, $Y \in \mathcal{Y}$. $\mathcal{X}$ and $\mathcal{Y}$ are finite sets and $\Pr\{X = x, Y = y\} = Q(x, y)$, $x \in \mathcal{X}$, $y \in \mathcal{Y}$. The marginal distributions are

$$Q_X(x) = \sum_{y \in \mathcal{Y}} Q(x, y) \qquad \text{and} \qquad Q_Y(y) = \sum_{x \in \mathcal{X}} Q(x, y).$$

Often, when the random variables are clear from the context, we write $Q_X(x)$ as $Q(x)$, etc. Define, for $m = 1, 2, \cdots$, the set

$$I_m = \{0, 1, 2, \cdots, m - 1\}. \tag{5}$$

An *encoder* with parameters (n, M_0, M_1, M_2) is a mapping

$$f_E \colon \mathcal{X}^n \times \mathcal{Y}^n \to I_{M_0} \times I_{M_1} \times I_{M_2}. \tag{6}$$

Given an encoder, a *decoder* is a pair of mappings

$$f_D^{(X)} : I_{M_0} \times I_{M_1} \to \mathfrak{X}^n \tag{7a}$$

$$f_D^{(Y)} : I_{M_0} \times I_{M_2} \to \mathcal{Y}^n. \tag{7b}$$

An encoder-decoder with parameters (n, M_0, M_1, M_2) is applied as follows. Let

$$f_E(\mathbf{X}, \mathbf{Y}) = (S_0, S_1, S_2), \tag{8a}$$

where

$$\mathbf{X} = (X_1, \cdots, X_n) \qquad \text{and} \qquad \mathbf{Y} = (Y_1, \cdots, Y_n).$$

Then let

$$\hat{\mathbf{X}} = f_D^{(X)}(S_0, S_1), \tag{8b}$$

$$\hat{\mathbf{Y}} = f_D^{(Y)}(S_0, S_1). \tag{8c}$$

The resulting error rate is

$$\Delta = \max\,(\Delta_X, \Delta_Y), \tag{9a}$$

where

$$\Delta_X = E\frac{1}{n}\sum_{k=1}^{n} d_H(X_k, \hat{X}_k), \tag{9b}$$

$$\Delta_Y = E\frac{1}{n}\sum_{k=1}^{n} d_H(Y_k, \hat{Y}_k), \tag{9c}$$

$d_H(\cdot, \cdot)$ is defined by (1b), and $\hat{X}_k$, $\hat{Y}_k$ are the kth coordinate of $\hat{\mathbf{X}}$ and $\hat{\mathbf{Y}}$, respectively. The Hamming distance $D_H(\mathbf{u}, \mathbf{v})$ between the n-vectors $\mathbf{u}$ and $\mathbf{v}$ is the number of positions in which $\mathbf{u}$ and $\mathbf{v}$ differ. Thus, $\Delta_X = E(1/n)D_H(\mathbf{X}, \mathbf{Y})$ and $\Delta_Y = E(1/n)D_H(\mathbf{Y}, \hat{\mathbf{Y}})$.

The correspondence between the encoder-decoder pair (or "code") as defined here and the communication system of Fig. 2 should be clear. Note that the capacities of the channels in that diagram must be at least $C_i = (1/n)\log_2 M_i$ $(i = 0, 1, 2)$.

A triple (R_0, R_1, R_2) is said to be *achievable* if, for arbitrary $\epsilon > 0$, there exists (for n sufficiently large) a code with parameters (n, M_0, M_1, M_2) with $M_i \leqq 2^{n(R_i+\epsilon)}$, $i = 0, 1, 2$, and error rate $\Delta \leqq \epsilon$. We define $\mathcal{R}$ as the set of achievable rates. Our main problem is to ascertain the region $\mathcal{R}$.

It follows from the definition that $\mathcal{R}$ is a closed subset of Euclidean three-space and the $\mathcal{R}$ has the property that

$$(R_0, R_1, R_2) \in \mathcal{R} \to (R_0 + \epsilon_0, R_1 + \epsilon_1, R_2 + \epsilon_2) \in \mathcal{R}, \tag{10}$$

$\epsilon_i \geqq 0$, $i = 0, 1, 2$. The region $\mathcal{R}$ is therefore completely defined by giving its lower boundary $\overline{\mathcal{R}}$, where

$$\begin{aligned}\overline{\mathcal{R}} \triangleq \{R_0, R_1, R_2) \in \mathcal{R}\colon (\hat{R}_0, \hat{R}_1, \hat{R}_2) \in \mathcal{R},\\ \hat{R}_i \leqq R_i(i = 0, 1, 2) \rightarrow \hat{R}_i = R_i(i = 0, 1, 2)\}.\end{aligned} \tag{11}$$

It follows immediately that $\overline{\mathcal{R}}$ too is closed.

It can also be verified by a simple "time-sharing" argument that $\mathcal{R}$ is convex (see appendix). This leads us to the following equivalent formulation of the problem. Let $\alpha_i \geqq 0$, $i = 0, 1, 2$ be arbitrary. Then define

$$T_1(\alpha_0, \alpha_1, \alpha_2) = \min_{(R_0, R_1, R_2) \in \mathcal{R}} (\alpha_0 R_0 + \alpha_1 R_1 + \alpha_2 R_2).$$

Then it follows from the convexity of $\mathcal{R}$ that the lower boundary $\overline{\mathcal{R}}$ is the upper envelope of the family of planes $\sum_0^2 \alpha_i R_i = T_1(\alpha_0, \alpha_1, \alpha_2)$.

We can think of $T_1(\alpha_0, \alpha_1, \alpha_2)$ as the minimum cost of transmitting, using a code with rate-triple (R_0, R_1, R_2) over the network of Fig. 2, when the cost of transmitting a bit per second over the common channel is α_0 and the costs of transmitting a bit per second over the private channels to receivers 1 and 2 are α_1 and α_2, respectively. Now, since information sent over the common channel (in Fig. 2) can alternatively be sent over *both* private channels, it is never necessary to consider the case where the sum of the costs of a bit per second on the private channels $\alpha_1 + \alpha_2 < \alpha_0$, the cost of a bit per second on the common channel. Similarly, we need never consider the cases where $\alpha_1 > \alpha_0$, or $\alpha_2 > \alpha_0$, since information transmitted over a private channel can alternatively be sent over the common channel. Since we can normalize α_0 as unity, the following theorem should be plausible. A complete proof is given in the appendix.

For $\mathbf{R} = (R_0, R_1, R_2)$ satisfying $R_i \geqq 0$, and $\boldsymbol{\alpha} = (\alpha_1, \alpha_2)$ arbitrary, let the "cost" be defined by

$$C(\boldsymbol{\alpha}, \mathbf{R}) = R_0 + \alpha_1 R_1 + \alpha_2 R_2. \tag{12}$$

With $\boldsymbol{\alpha}$ held fixed, let

$$T(\boldsymbol{\alpha}) = \min_{\mathbf{R} \in \mathcal{R}} C(\boldsymbol{\alpha}, \mathbf{R}). \tag{13}$$

The indicated minimum exists because $\mathcal{R}$ is closed. For $\boldsymbol{\alpha}$ arbitrary, let $\mathcal{S}(\boldsymbol{\alpha})$ be the set of $\mathbf{R} \in \mathcal{R}$ that achieve $T(\boldsymbol{\alpha}) = C(\boldsymbol{\alpha}, \boldsymbol{R})$.

Theorem 1:

$$(i)\,\bar{\mathcal{R}} \subseteq \bigcup_{\boldsymbol{\alpha} \in \mathcal{A}} \mathcal{S}(\boldsymbol{\alpha}),$$

$$(ii)\,\bigcup_{\boldsymbol{\alpha} \in \mathcal{A}'} \mathcal{S}(\boldsymbol{\alpha}) \subseteq \bar{\mathcal{R}},$$

where the boundary $\bar{\mathcal{R}}$ *is defined in (11),* $\mathcal{A}$ *is the set of* $\boldsymbol{\alpha} = (\alpha_1, \alpha_2)$ *that satisfy*

$$0 \leqq \alpha_1, \alpha_2 \leqq 1, \qquad \alpha_1 + \alpha_2 \geqq 1,$$

and $\mathcal{A}'$ *is* $\mathcal{A}$ *with the elements* $(0, 1)$ *and* $(1, 0)$ *deleted.*

Remarks:

(1) $(0, 1)$ and $(1, 0)$ are the only pairs in $\mathcal{A}$ with zero elements. Thus, $\mathcal{A}$ and $\mathcal{A}'$ are nearly identical.

(2) The theorem implies that $\bar{\mathcal{R}}$ is upper envelope in (R_0, R_1, R_2)-space of the family of planes defined by

$$R_0 + \alpha_1 R_1 + \alpha_2 R_2 = T(\boldsymbol{\alpha}),$$

$\boldsymbol{\alpha} \in \mathcal{A}$.

1.3 Upper and lower bounds on $\mathcal{R}$

1.3.1 Lower bounds

We can immediately give some lower bounds to the region $\mathcal{R}$. We state them as

Theorem 2: *If* $(R_0, R_1, R_2) \in \mathcal{R}$, *then*

(a) $R_0 + R_1 + R_2 \geqq H(X, Y)$,

(b) $R_0 + R_1 \geqq H(X)$,

(c) $R_0 + R_2 \geqq H(Y)$.

Proof: Suppose that $(R_0, R_1, R_2) \in \mathcal{R}$. Then, for arbitrary $\epsilon > 0$, we can (for sufficiently large block length n) reproduce $\{X_k\}$, and $\{Y_k\}$ with arbitrarily small Δ_X, Δ_Y, with capacity triple (in Fig. 2)

$$(C_0, C_1, C_2) = (R_0 + \epsilon, R_1 + \epsilon, R_2 + \epsilon).$$

That is, with a code with $M_i = 2^{nC_i}$, $i = 0, 1, 2$.

Since the total capacity of the three channels is $C_0 + C_1 + C_2$, we must have

$$C_0 + C_1 + C_2 = R_0 + R_1 + R_2 + 3\epsilon \geqq H(X, Y).$$

Letting $\epsilon \to 0$, we have established (*a*). Inequality (*b*) follows in an identical way on observing that the common channel (with capacity C_0) and the private channel to receiver 1 (with capacity C_1) together transmit $\{X_k\}$. Inequality (*c*) follows, similarly.

Let us remark that inequality (*a*) is an expression of the fact that a communication system with the constraints imposed in Fig. 2 cannot perform better than in the "best of all possible worlds" situation in which the receivers can collaborate. It is therefore called the "Pangloss bound." The set of triples (R_0, R_1, R_2) that satisfy $\sum_0^2 R_i = H(X, Y)$ are called the "Pangloss plane." Corresponding to rate-triples that lie on the intersection of $\mathcal{R}$ and the Pangloss plane, the approximately $H(X, Y)$ bits per second that characterize $\{X_k, Y_k\}$ can be split up into three parts (corresponding to the information transmitted over the three channels in our network) such that $\{X_k, Y_k\}$ can be essentially perfectly reconstructed by the three receivers in the network. In this situation, the information transmitted over the common channel represents a kind of "core" process. Furthermore, the smallest R_0, such that $(R_0, R_1, R_2) \in \mathcal{R}$ and lies on the Pangloss plane (for some R_1, R_2), can be thought of as a measure of the "common information" of $\{X_k\}$ and $\{Y_k\}$. This point is explored thoroughly in Ref. 3.

1.3.2 Some easily achievable rate-triples

We now assert that certain rate-triples are achievable.

Theorem 3: *The following triples belong to* $\mathcal{R}$:

$$(A) \quad R_0 = H(X, Y), \qquad R_1 = R_2 = 0$$

$$(B) \quad R_0 = 0, \qquad R_1 = H(X), \qquad R_2 = H(Y)$$

$$(C) \quad R_0 = H(Y), \qquad R_1 = H(X|Y), \qquad R_2 = 0$$

$$(D) \quad R_0 = H(X), \qquad R_1 = 0, \qquad R_2 = H(Y|X).$$

Proof: To achieve (*A*), simply transmit $\{(X_k, Y_k)\}$ over the common channel (and do not use the private channels). To achieve (*B*), transmit $\{X_k\}$ and $\{Y_k\}$ over the private channels to receivers 1 and 2, respectively (and do not use the common channel). To achieve (*C*), transmit $\{Y_k\}$ over the common channel (requiring a capacity of about $H(Y)$), and deliver $\{\hat{Y}_k\}$ to receiver 1 to use as side information for transmitting $\{X_k\}$ over the private channel to receiver 1. This will require a capacity of about $H(X|Y)$. We do not use the private channel to receiver 2. Triple (D) can be achieved as in (C) by reversing to roles of X and Y.

Let us remark that points (C) and (D) lie on the Pangloss plane (i.e., they satisfy relation (a) of Theorem 2 with equality), since $H(X) + H(Y|X) = H(Y) + H(X|Y) = H(X, Y)$. Furthermore, because of the convexity of $\mathcal{R}$, all triples that are linear combinations of triples (A) to (D) are also members of $\mathcal{R}$. The situation is summarized in Fig. 3. The plane labeled "(a)" in the figure is the Pangloss plane defined by $R_0 + R_1 + R_2 = H(X, Y)$. Theorem 2$(a)$ states that the region $\mathcal{R}$ (and therefore its lower boundary $\overline{\mathcal{R}}$) lies above this plane. Similarly, Theorem 2(b, c) states that $\mathcal{R}$ and $\overline{\mathcal{R}}$ lie above the planes labeled "(b)" and "(c)" in Fig. 3.

Now the points labeled "A," "B," "C," and "D" in the figure are points (A) to (D) respectively in Theorem 3. As we mentioned previously, points C and D (as well as A) lie on plane a. Thus (from the convexity of $\mathcal{R}$), the triangle ADC lies in $\mathcal{R}$ and must therefore be part of the lower boundary $\overline{\mathcal{R}}$. Further, since points D and B lie on plane b,

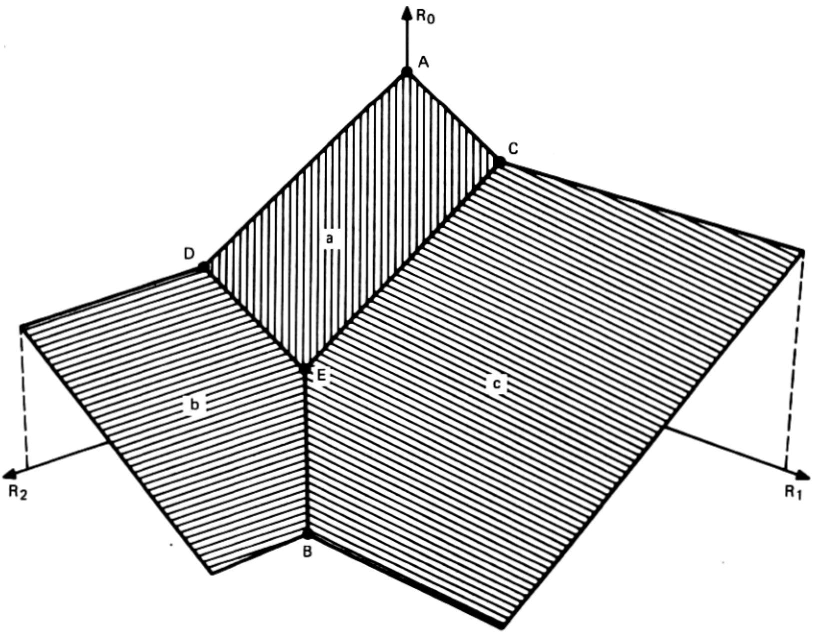

Fig. 3—Estimates of rate-region $\mathcal{R}$.

the line DB is part of $\overline{\mathcal{R}}$. Similarly, line BC is part $\overline{\mathcal{R}}$. Finally, since points B, C, and D are achievable, so are the points on the triangle BCD. Thus, the only unknown part of the lower boundary $\overline{\mathcal{R}}$ lies in the (upside-down) triangular pyramid with base BCD and apex at point E (the intersection of planes a, b, c). The coordinates of point E are easily seen to be $(R_0, R_1, R_2) = [I(X;Y), H(X|Y), H(Y|X)]$.

Let us remark here that there is one special source distribution $Q(x, y)$ for which point E is achievable.† In this case, the entire boundary region $\overline{\mathcal{R}}$ lies on planes a, b, c. This special case is when X, Y can be written $X = (X', V)$, $Y = (Y', V)$, where X' and Y' are conditionally independent given V. Then $I(X, Y) = H(V)$, $H(X|Y) = H(X'|V)$, $H(Y|X) = H(Y'|V)$, so that point E is $R_0 = H(V)$, $R_1 = H(X'|V)$, $R_2 = H(Y'|V)$. Clearly, if, in the system of Fig. 2, we transmit V over the common channel and X' and Y' over the two private channels, we can reconstruct $X = (X', V)$ at receiver 1 and $Y = (Y', V)$ at receiver 2. This requires a capacity triple $C_0 = H(V) + \epsilon$, $C_1 = H(X'|V) + \epsilon$, $C_2 = H(Y'|V) + \epsilon$ ($\epsilon > 0$ arbitrary), so that point is in fact achievable.

We now give a characterization of the region $\mathcal{R}$ (and therefore of $\overline{\mathcal{R}}$) in terms of information theoretic quantities. This characterization is, in fact, the main result.

1.4 Characterization of $\mathcal{R}$—the main result

Suppose we are given $Q(x, y)$, $x \in \mathfrak{X}$, $y \in \mathcal{Y}$, an arbitrary probability function, where $\mathfrak{X}$, $\mathcal{Y}$ are finite. Let $\mathcal{P}$ be the family of probability functions $p(x, y, w)$, where $x \in \mathfrak{X}$, $y \in \mathcal{Y}$, $w \in \mathcal{W}$, and $\mathcal{W}$ is another finite set, for which

$$\sum_{w \in \mathcal{W}} p(x, y, w) = Q(x, y), x \in \mathfrak{X}, y \in \mathcal{Y}. \tag{14}$$

Each $p \in \mathcal{P}$ defines discrete random variables X, Y, W in an obvious way. For each $p \in \mathcal{P}$, define the subset of Euclidean three-space

$$\mathcal{R}^{(p)} = \{(R_0, R_1, R_2): R_0 \geqq I(X, Y; W), \quad R_1 \geqq H(X|W), \quad R_2 \geqq H(Y|W)\}, \tag{15a}$$

and then let

$$\mathcal{R}^* = (\bigcup_{p \in \mathcal{P}} \mathcal{R}^{(p)})^c, \tag{15b}$$

† M. Kaplan has shown that, in fact, this special case is the only one for which point E is achievable.

where $(\ )^c$ denotes set closure. Then our main result (the proof of which is given in Section III) is

Theorem 4: $\mathcal{R} = \mathcal{R}^*$.

Remarks:

(1) Let us define $\mathcal{P}_T$ as the family of "test channel" transition probabilities. That is, $\mathcal{P}_T$ is the family of all $p_t(w|x,y)$ $(x \in \mathcal{X}, y \in \mathcal{Y}, w \in \mathcal{W})$, where $\mathcal{W}$ is a finite set, and for each (x, y), $p_t(w|x, y)$ is a probability function on $\mathcal{W}$. Corresponding to each $p_t \in \mathcal{P}_T$, we have $p(x, y, w) = Q(x, y)p_t(w|x, y) \in \mathcal{P}$. Further, for each $p \in \mathcal{P}$, we have $p_t(w|x, y) = [p(x, y, w)/Q(x, y)] \in \mathcal{P}_T$. Thus $\mathcal{P}$ is in 1-1 correspondence with $\mathcal{P}_T$.

(2) Since $\mathcal{R}$ is convex, Theorem 4 implies that $\mathcal{R}^*$ is convex also.

(3) Theorem 4 can be invoked to show that $T(\boldsymbol{\alpha})$ defined in (13) is also given by

$$T(\boldsymbol{\alpha}) = \inf_{p \in \mathcal{P}} [I(X, Y; W) + \alpha_1 H(X|W) + \alpha_2 H(Y|W)]. \quad (16)$$

Thus, from Theorem 1, the lower boundary $\bar{\mathcal{R}}$, and therefore $\mathcal{R}$, is essentially determined by $T(\boldsymbol{\alpha})$ given by (16).

(4) Theorems 2 and 3 can be verified easily by using Theorem 4. Thus, if $(R_0, R_1, R_2) \in \mathcal{R}$, from Theorem 4 for arbitrary $\epsilon > 0$ we can find a triple of random variables X, Y, W such that

$$\begin{aligned} R_0 + R_1 + R_2 &\geqq I(X, Y; W) + H(X|W) + H(Y|W) - \epsilon \\ &= H(X, Y) + [H(X|W) + H(Y|W) - H(X, Y|W)] - \epsilon \\ &\geqq H(X, Y) - \epsilon \to H(X, Y), \qquad \text{as } \epsilon \to 0. \end{aligned} \quad (17)$$

This is Theorem 2(a). The second inequality in (17) follows from the fact that the entropy of a pair of random variables is less than the sum of the respective entropies. Part (*b*) of Theorem 2 follows from

$$\begin{aligned} I(X, Y; W) + H(X|W) &= I(X; W) + I(Y; W|X) \\ &\quad + H(X|W) \geqq I(X; W) + H(X|W) = H(X). \end{aligned} \quad (18)$$

The first equality in (18) follows from a standard identity [Ref. 4, Eq. (2.2.29)].

Theorem 3 follows from Theorem 4 on taking W as follows: (*A*) $W = (X, Y)$, (*B*) $W = 0$, (*C*) $W = Y$, (*D*) $W = X$.

(5) Although Theorem 4 characterizes $\mathcal{R}$ and $\bar{\mathcal{R}}$ by an information theoretic minimization, it must be emphasized that the minimization is not, in general, easy. In fact, there is no nontrivial case for which we have succeeded in calculating the entire boundary $\bar{\mathcal{R}}$ analytically.

Its major utility at this point has been in finding upper bounds on $\bar{\mathcal{R}}$ by guessing at a p or p_t and calculating the corresponding triple $[I(X, Y; W), H(X|W), H(Y|W)]$, which must lie above $\bar{\mathcal{R}}$. See the example below. The problem of computation of $\bar{\mathcal{R}}$ both analytically and numerically is still open.†

(6) For $p \in \mathcal{P}$, we can define the quantities

$$\beta_{xy}(w) = \Pr\{X = x, Y = y | W = w\}, \qquad x \in \mathfrak{X}, y \in \mathcal{Y}, w \in \mathcal{W},$$

which can be thought of as the transition probabilities of the "backward test channel." For a given (x, y), we can think of $\beta_{xy} = \beta_{x,y}(W)$ as a random variable. Of course, β_{xy} must satisfy

$$\beta_{xy} \geqq 0, \tag{19a}$$

$$\sum_{x,y} \beta_{xy} = 1, \tag{19b}$$

and

$$E\beta_{xy} = Q(x, y), \tag{19c}$$

where the expectation is taken over the distribution for W. Further,

$$I(X, Y; W) = H(X, Y) - H(X, Y|W) = H(X, Y) - E \sum_{x,y} \beta_{xy} \log \frac{1}{\beta_{xy}}, \tag{20a}$$

$$H(X|W) = E \sum_{x} \beta_x^{(1)} \log \frac{1}{\beta_x^{(1)}}, \qquad H(Y|W) = E \sum_{y} \beta_y^{(2)} \log \frac{1}{\beta_y^{(2)}}, \tag{20b}$$

where

$$\beta_x^{(1)} = \sum_y \beta_{xy} = \Pr\{X = x | W\}, \quad \text{and} \quad \beta_y^{(2)} = \sum_x \beta_{xy} = \Pr\{Y = y | W\}, \tag{20c}$$

and the expectation is taken over the distribution for W. Using this idea, it is possible to characterize, for example, $T(\boldsymbol{\alpha})$ as follows (see Ref. 3, for a precise proof of this characterization). Given $Q(x, y)$, $x \in \mathfrak{X}$, $y \in \mathcal{Y}$, define $\mathcal{B}$ as the family of collections of random variables, $\{\beta_{xy}\}$, $x \in \mathfrak{X}$, $y \in \mathcal{Y}$, which satisfy (19). Then

$$T(\boldsymbol{\alpha}) = \min [I(X, Y; W) + \alpha_1 H(X|W) + \alpha_2 H(Y|W)],$$

† One reason for the difficulty is that $I(X, Y; W) + \alpha_1 H(X|W) + \alpha_2 H(Y|W)$ is apparently neither convex nor concave in p_t.

where $I(X, Y; W)$, $H(X|W)$, $H(Y|W)$ are given by (20) and the minimum (which can be shown to exist) is over all sets $\{\beta_{xy}\}$ in $\mathcal{B}$.

This characterization may have value in the computation problem, since the quantities in (20) are *linear* functions of the joint distribution function for the $\{\beta_{xy}\}$ and the constraints of (19) are also linear inequalities in this distribution function. Thus, calculation of $T(\boldsymbol{\alpha})$ is a linear programming problem.

(7) If $p \in \mathcal{P}$ is such that X and Y are conditionally independent given W, then $H(X, Y|W) = H(X|W) + H(Y|W)$. Thus, with $R_0 = I(X, Y; W)$, $R_1 = H(X|W)$, $R_2 = H(Y|W)$,

$$\begin{aligned} R_0 + R_1 + R_2 &= H(X, Y) - H(X, Y|W) + H(X|W) + H(Y|W) \\ &= H(X, Y), \end{aligned}$$

and $(R_0, R_1, R_2) \in \mathcal{R}$ and lies on the Pangloss plane. Reference 3 shows that this class of triples (corresponding to a $p \in \mathcal{P}$, with X, Y conditionally independent given W) completely characterizes the intersection of $\mathcal{R}$ and the Pangloss plane.

1.5 An example

As an example of the preceding, let us consider the special case where the source is the "doubly symmetric binary source" (DSBS), where $\mathcal{X} = \mathcal{Y} = \{0, 1\}$, and

$$Q(x, y) = \tfrac{1}{2}(1 - p_0)\delta_{x,y} + \tfrac{1}{2}p_0(1 - \delta_{x,y}), \quad x, y = 0, 1, \tag{21}$$

and the parameter p_0 satisfies $0 \leqq p_0 \leqq \frac{1}{2}$. We can think of X as being an unbiased binary input into a binary symmetric channel (BSC) with crossover probability p_0, and Y as being the corresponding output, or vice versa. To get a clearer picture of the set of achievable rates $\mathcal{R}$, let us restrict ourselves to the plane in (R_0, R_1, R_2)-space, where $R_1 = R_2$. The intersection of $\mathcal{R}$ and this plane can be plotted in a two-dimensional picture.

Let us first take a look at the implications of Theorems 2 and 3. In this source,

$$H(X) = H(Y) = 1, \qquad H(X|Y) = H(Y|X) = h(p_0)$$

and

$$H(X, Y) = H(X) + H(Y|X) = 1 + h(p_0),$$

where

$$h(\lambda) = -\lambda \log \lambda - (1 - \lambda) \log (1 - \lambda), \qquad 0 \leqq \lambda \leqq 1 \tag{22}$$

is the entropy function. [We take $h(0) = h(1) = 1$.] With $R_1 = R_2$, Theorem 2 yields

$$R_0 + 2\,R_1 \geqq 1 + h(p_0), \tag{23a}$$

$$R_0 + R_1 \geqq 1. \tag{23b}$$

Thus, $\mathcal{R}$ and therefore the lower boundary $\bar{\mathcal{R}}$ must lie above the lines labeled a and b in Fig. 4.

Now Theorem 3 implies that points $A[R_0 = 1 + h(p_0),\ R_1 = 0]$, and $B(R_0 = 0,\ R_1 = 1)$ are achievable, so that any point on the line connecting them is also achievable. But we can do better. Let us drop for a moment the requirement that $R_1 = R_2$. From Theorem 3, C and D, the points $[R_0 = 1, R_1 = h(p_0), R_2 = 0]$ and $[R_0 = 1, R_1 = 0, R_2 = h(p_0)]$

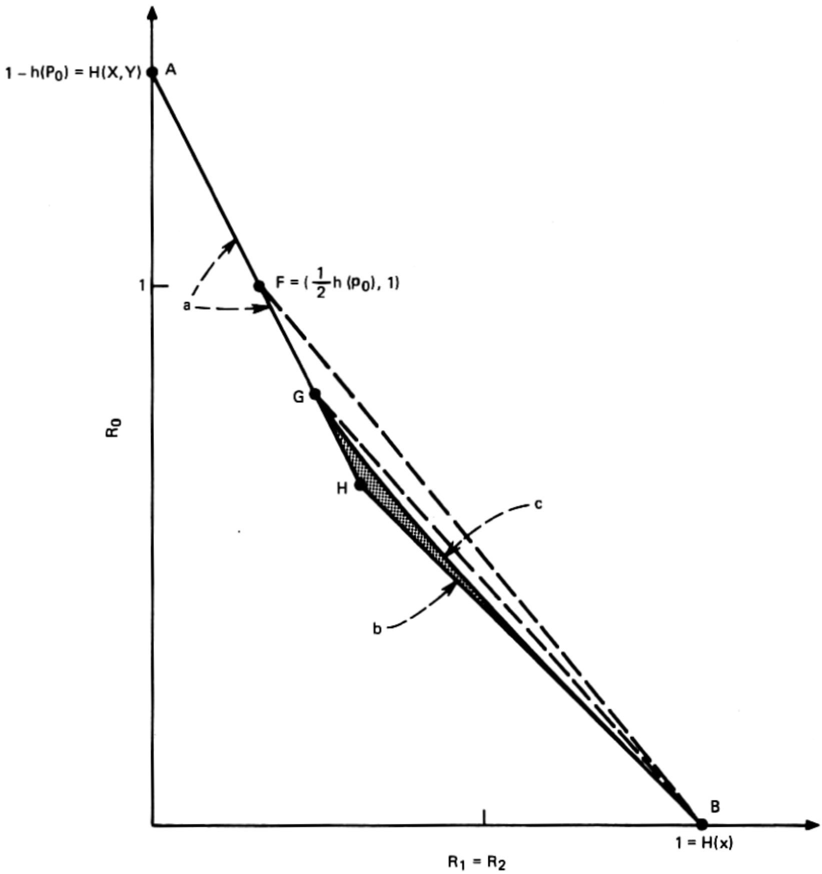

Fig. 4—Estimates of rate-region $\mathcal{R}$ for the DSBS.

are achievable. Thus, the point in (R_0, R_1, R_2)-space halfway between them is also achievable. But this point,

$$[R_0 = 1, R_1 = \tfrac{1}{2}h(p_0), R_2 = \tfrac{1}{2}h(p_0)],$$

satisfies $R_1 = R_2$, and is therefore of interest to us now. Point F in Fig. 4 is therefore achievable, and therefore so are line segments AF and FB. But line segment AF coincides with line a, so that it must be on the boundary $\bar{\mathcal{R}}$. So far, the unknown part of the boundary curve $\bar{\mathcal{R}}$ lies in triangle FHB. We can do better, however, by using Theorem 4.

Theorem 4 asserts that any triple in $\mathcal{R}^{(p)}$, $p \in \mathcal{P}$, is achievable. We therefore guess at a $p \in \mathcal{P}$ that defines random variables X, Y, W, and then assert that the triple $R_0 = I(X, Y; W)$, $R_1 = H(X|W)$, $R_2 = H(Y|W)$ is achievable. Since we choose a $p \in \mathcal{P}$ such that $R_1 = R_2$, this triple is of interest in our present discussion. The $p \in \mathcal{P}$ we have chosen is (with $\mathcal{W} = \{0, 1\}$) given by Table I. The quantity $p_1 = \frac{1}{2}(1 - \sqrt{1 - 2p_0})$. One way of characterizing p is to think of W as an unbiased binary input and X, Y the respective outputs of two independent BSC's, each with crossover probability p_1. Note that these two BSC's in cascade are equivalent to a single BSC with crossover probability, $2p_1(1 - p_1) = p_0$.

With X, Y, W so defined, X, Y are conditionally independent given W, so that (R_0, R_1, R_2) lies on the Pangloss plane. [See remark (6) following Theorem 4.] We have

$$\begin{aligned} R_0 &= I(X, Y; W) = H(X, Y) - H(X, Y|W) \\ &= 1 + h(p_0) - 2h(p_1), \\ R_1 &= R_2 = H(X|W) = h(p_1). \end{aligned} \tag{24}$$

This is point G in Fig. 4. Line segment AG is therefore on the boundary $\bar{\mathcal{R}}$. From these simple arguments, we see that the unknown part of the boundary $\bar{\mathcal{R}}$ lies in the triangle GHB.

To obtain a still tighter bound on $\bar{\mathcal{R}}$, we employ the same technique as above—i.e., "guessing" at a $p \in \mathcal{P}$ and then deducing that $\mathcal{R}^{(p)}$

Table I

$W \backslash XY$	00	01	10	11
0	$\frac{1}{2}(1 - p_1)^2$	$\frac{1}{2}p_1(1 - p_1)$	$\frac{1}{2}p_1(1 - p_1)$	$\frac{1}{2}p_1^2$
1	$\frac{1}{2}p_1^2$	$\frac{1}{2}p_1(1 - p_1)$	$\frac{1}{2}p_1(1 - p_1)$	$\frac{1}{2}(1 - p_1)^2$

Table II

$W \backslash XY$	00	01	10	11
0	$\frac{1}{2}\left(1-\beta-\frac{p_0}{2}\right)$	$p_0/4$	$p_0/4$	$\frac{1}{2}\left(\beta-\frac{p_0}{2}\right)$
1	$\frac{1}{2}\left(\beta-\frac{p_0}{2}\right)$	$p_0/4$	$p_0/4$	$\frac{1}{2}\left(1-\beta-\frac{p_0}{2}\right)$

$\subseteq \mathcal{R}$. Let β be a parameter for which

$$p_1 = \tfrac{1}{2}(1 - \sqrt{1-2p_0}) \leqq \beta \leqq \tfrac{1}{2}. \tag{25}$$

Then let $\mathcal{W} = \{0, 1\}$, and $p(x, y, w)$ be given by Table II. Then the triples $(R_0, R_1, R_2) \in \mathcal{R}$, where

$$\begin{aligned} R_0 &= I(X, Y; W) = H(X, Y) - H(X, Y \mid W) \\ &= 1 + h(p_0) + \frac{1}{2}\left(1-\beta-\frac{p_0}{2}\right)\log\frac{1}{2}\left(1-\beta-\frac{p_0}{2}\right) \\ &\qquad + \frac{p_0}{2}\log\frac{p_0}{4} + \frac{1}{2}\left(\beta-\frac{p_0}{2}\right)\log\frac{1}{2}\left(\beta-\frac{p_0}{2}\right) \end{aligned} \tag{26a}$$

and

$$R_1 = R_2 = H(X \mid W) = H(Y \mid W) = h(\beta). \tag{26b}$$

For $\beta = p_1$, the triple of (26) coincides with that of (25), i.e., point G in Fig. 4. For $\beta = \frac{1}{2}$, the triple of (26) is $R_0 = 0$, $R_1 = R_2 = 1$, i.e., point B of Fig. 4. As β increases from p_1 to $\frac{1}{2}$, the family of rate-triples of (26) generate a curve c, which lies below the line GB and therefore constitutes a tighter upper bound on $\overline{\mathcal{R}}$. We conclude that the unknown portion of $\overline{\mathcal{R}}$ lies in the shaded region in Fig. 4.

In Section 2.5 we give some insight into how we "guessed" at these distributions $p \in \mathcal{P}$.

II. GENERALIZATION TO A FIDELITY CRITERION

In this section we formulate a generalization of the problem of Section I in which we require that the source sequences $\{X_k\}$ and $\{Y_k\}$ be reproduced to within a specified fidelity criterion and not, as in Section I, essentially perfectly. The proofs of the main theorems appear in Section III.

2.1 Definitions and formulation of the problem

Let $\{(X_k, Y_k)\}_{k=1}^{\infty}$ be a sequence of independent drawings of a pair of random variables $X \in \mathcal{X}$, $Y \in \mathcal{Y}$, where the "source alphabets"

$\mathfrak{X}$ and $\mathfrak{Y}$ are either discrete sets, the reals, or arbitrary measurable spaces. We assume that we are given a probability law that defines (X, Y). If $\mathfrak{X}$ and $\mathfrak{Y}$ are discrete, then we write

$$Q(x, y) = \Pr\{X = x, Y = y\}, \qquad x \in \mathfrak{X}, y \in \mathfrak{Y}.$$

If $\mathfrak{X}$, $\mathfrak{Y}$ are the reals, then (X, Y) may be defined by a probability density $Q(x, y)$, $-\infty < x, y < \infty$. For arbitrary measurable $\mathfrak{X}$, $\mathfrak{Y}$, the pair (X, Y) is defined by a probability measure Q on $\mathfrak{X} \times \mathfrak{Y}$. The marginal distribution for X, Y will be defined similarly by Q_X, Q_Y respectively.

As in (5), define the set $I_m = \{0, 1, \cdots, m-1\}$ for $m = 1, 2, \cdots$. An *encoder* with parameters (n, M_0, M_1, M_2) is [as in (6)] a mapping

$$f_E: \mathfrak{X}^n \times \mathfrak{Y}^n \rightarrow I_{M_0} \times I_{M_1} \times I_{M_2}. \tag{27}$$

We assume that the sequences $\{X_k\}$ and $\{Y_k\}$ are to be reproduced as sequences of elements of sets $\hat{\mathfrak{X}}$ and $\hat{\mathfrak{Y}}$, respectively, called "reproducing alphabets." Thus [as in (7)], corresponding to a given encoder, a decoder is a pair of mappings

$$f_D^{(X)}: I_{M_0} \times I_{M_1} \rightarrow \hat{\mathfrak{X}}^n, \tag{28a}$$

$$f_D^{(Y)}: I_{M_0} \times I_{M_1} \rightarrow \hat{\mathfrak{Y}}^n. \tag{28b}$$

Let us adopt the convention of denoting n-vectors with bold-face type (either upper or lower case) and the components as the same subscripted letter in ordinary type. For example, $\mathbf{u} = (u_1, \cdots, u_n)$.

An encoder-decoder with parameters (n, M_0, M_1, M_2) is applied as follows. Say

$$f_E(\mathbf{X}, \mathbf{Y}) = (S_0, S_1, S_2), \tag{29a}$$

where $\mathbf{X} \in \mathfrak{X}^n$, $\mathbf{Y} \in \mathfrak{Y}^n$, and (S_0, S_1, S_2) is a triplet of indices. Then set

$$\hat{\mathbf{X}} = f_D^{(X)}(S_0, S_1), \qquad \hat{\mathbf{Y}} = f_D^{(Y)}(S_0, S_2), \tag{29b}$$

where $\hat{\mathbf{X}} \in \hat{\mathfrak{X}}^n$, $\hat{\mathbf{Y}} \in \hat{\mathfrak{Y}}^n$. The encoder-decoder is said to have *average distortion* (Δ_X, Δ_Y), where

$$\Delta_X = ED_1(\mathbf{X}, \hat{\mathbf{X}}), \qquad \Delta_Y = ED_2(\mathbf{Y}, \hat{\mathbf{Y}}), \tag{30a}$$

and the single-letter distortion functions are defined by

$$D_1(\mathbf{x}, \hat{\mathbf{x}}) = \frac{1}{n} \sum_{k=1}^{n} d_1(x_k, \hat{x}_k), \tag{30b}$$

$$D_2(\mathbf{y}, \hat{\mathbf{y}}) = \frac{1}{n} \sum_{k=1}^{n} d_2(y_k, \hat{y}_k), \tag{30c}$$

$\mathbf{x} \in \mathfrak{X}^n$, $\hat{\mathbf{x}} \in \hat{\mathfrak{X}}^n$, $\mathbf{y} \in \mathcal{Y}^n$, $\hat{\mathbf{y}} \in \mathcal{Y}^n$, and $d_1(\cdot, \cdot)$ is a given nonnegative per-letter distortion function for the X-receiver and $d_2(\cdot, \cdot)$ is a given nonnegative per-letter distortion function for the Y-receiver. An encoder-decoder with parameters (n, M_0, M_1, M_2) with average distortion (Δ_X, Δ_Y) is said to be a *code* $(n, M_0, M_1, M_2, \Delta_X, \Delta_Y)$.

A rate-triple (R_0, R_1, R_2) is said to be (Δ_1, Δ_2)-*achievable* if, for arbitrary $\epsilon > 0$ and n sufficiently large, there exists a code $(n, M_0, M_1, M_2, \Delta_X, \Delta_Y)$ with

$$M_i \leqq 2^{n(R_i+\epsilon)}, \qquad i = 0, 1, 2,$$

and

$$\Delta_X \leqq \Delta_1 + \epsilon, \qquad \Delta_Y \leqq \Delta_2 + \epsilon.$$

The set of all (Δ_1, Δ_2)-achievable rate-triples is called $\mathcal{R}(\Delta_1, \Delta_2)$. Our main problem is to ascertain $\mathcal{R}(\Delta_1, \Delta_2)$, $\Delta_1, \Delta_2 \geqq 0$. Clearly, this generalized problem reduces to the problem of Section I, if $\mathfrak{X} = \hat{\mathfrak{X}}$, $\mathcal{Y} = \hat{\mathcal{Y}}$, $d_1 = d_2 = d_H$, and $\Delta_1 = \Delta_2 = 0$. As in Section I, the region $\mathcal{R}(\Delta_1, \Delta_2)$ is completely defined by the boundary $\overline{\mathcal{R}}(\Delta_1, \Delta_2)$, where $\overline{\mathcal{R}} = \mathcal{R}(\Delta_1, \Delta_2)$ is defined in (11). Further, we show in the appendix that $\mathcal{R}(\Delta_1, \Delta_2)$ is convex and that Theorem 1 holds with $\mathcal{R} = \mathcal{R}(\Delta_1, \Delta_2)$.

2.2 Rate-distortion functions and conditional rate-distortion functions

A major tool in this study is rate-distortion theory. Specifically, joint, marginal, and conditional rate-distortion functions (or simply "rates") are used both in evaluations and bounds. These functions and their properties are dealt with in Refs. 1, 4, and 5. Here we only summarize some pertinent definitions and properties.

The marginal, joint, and conditional rates are defined as follows. Consider first the case where the alphabets $\mathfrak{X}$, $\hat{\mathfrak{X}}$, $\mathcal{Y}$, $\hat{\mathcal{Y}}$, are finite and $Q(x, y)$, $Q_X(x)$, $Q_Y(y)$ are probability functions. Then the (joint) rate-distortion function is defined by

$$R_{XY}(\Delta_1, \Delta_2) = \min I(XY; \hat{X}\hat{Y}), \tag{31}$$

where the random variables $\hat{X}\hat{Y}$ are defined by a "test-channel" $q_t(\hat{x}; \hat{y} | x, y)$—i.e., a probability function on $\hat{\mathfrak{X}} \times \hat{\mathcal{Y}}$ for every $(x, y) \in \mathfrak{X} \times \mathcal{Y}$. The information in (31) is calculated for the joint distribution

$$\Pr \{X = x, Y = y, \hat{X} = \hat{x}, \hat{Y} = \hat{y}\} = Q(x, y) q_t(\hat{x}, \hat{y} | x, y). \tag{32}$$

The minimum in (31) is taken with respect to all test channels q_t such that $Ed_1(X, \hat{X}) \leqq \Delta_1$, $Ed_2(Y, \hat{Y}) \leqq \Delta_2$, where the expectations are taken with respect to the distribution (32). The minimum always exists. Similarly, the marginal rates are defined by

$$R_X(\Delta_1) = \min_{q_t(\hat{x}|x):Ed_1(x,\hat{x}) \leqq \Delta_1} I(X;\hat{X}), \tag{33a}$$

$$R_Y(\Delta_2) = \min_{q_t(\hat{y}|y):Ed(Y,\hat{Y}) \leqq \Delta_2} I(Y;\hat{Y}), \tag{33b}$$

where the expressions in (33) are interpreted analogously to that of (31). Detailed discussions of these quantities and their significance can be found in Refs. 1 and 4.

Another quantity that plays a crucial role in our study is the "conditional rate distortion function." Let $\mathfrak{X}$, $\mathcal{Y}$ be finite, and let $Q(x, y)$ be given. Let $p(x, y, w)$ be a probability function on $\mathfrak{X} \times \mathcal{Y} \times \mathcal{W}$, where $\mathcal{W}$ is a finite set such that $\sum_w p(x, y, w) = Q(x, y)$. Then $p(x, y, w)$ defines a triple of random variables X, Y, W, where the marginal distribution for X, Y is Q. The conditional rate-distortion functions are defined as

$$R_{XY|W}(\Delta_1, \Delta_2) = \min I(X, Y; \hat{X}\hat{Y}|W), \tag{34}$$

where the minimum (which always exists) is taken with respect to all test channels $q_t(\hat{x}, \hat{y}|x, y, w)$ such that $Ed_1(X, \hat{X}) \leqq \Delta_1$, $Ed_2(Y, \hat{Y}) \leqq \Delta_2$. The conditional information in (34) is defined in Ref. 4, p. 21. The conditional rates $R_{X|W}(\Delta_1)$, $R_{Y|W}(\Delta_2)$ are defined analogously. A detailed discussion of conditional rates is given in Ref. 5. Of course, these definitions are meaningful if $X \equiv W$ or $Y \equiv W$. Roughly speaking, $R_{X,Y|W}(\Delta_1, \Delta_2)$ is the channel capacity required to transmit X, Y and to reproduce it as $\hat{X}$, $\hat{Y}$ to within an average distortion (Δ_1, Δ_2) when both the transmitter and receiver know W.

We shall need several properties of the conditional rate-distortion function in the sequel. The first is given in Ref. 5. For $\Delta \geqq 0$,

$$R_{X|W}(\Delta) = \min \sum_w \Pr\{W = w\} R_{X|W=w}(\Delta_w), \tag{35}$$

where $R_{X|W=w}(\cdot)$ is the rate-distortion function calculated for a source with outputs $x \in \mathfrak{X}$ with probability distribution $P_{X|W}(x|w)$ (the conditional probability function for X given $W = w$). The minimum is taken over all sets $\{\Delta_w\}_{w\in\mathcal{W}}$ such that $\sum_w \Pr\{W = w\}\Delta_w \leqq \Delta$.

A second fact of importance is that, say, $R_{X|W}(\Delta)$ is a continuous, convex, nonincreasing function of Δ for $\Delta \geqq 0$. That $R_{X|W}(\Delta)$ is nonincreasing follows from the definition. The proof that it is convex parallels the proof of the convexity of the ordinary rate-distortion function. The continuity of $R_{X|W}(\Delta)$, $\Delta > 0$ follows from its convexity.

Finally, the continuity of $R_{X|W}(\Delta)$ at $\Delta = 0$ follows from (35), and the continuity of $R_{X|W=w}(\Delta)$ at $\Delta = 0$.

A third fact we shall need is that, for any X, W_1, W_2, $\Delta \geqq 0$,

$$R_{X|W_1W_2}(\Delta) \leqq R_{X|W_1}(\Delta). \tag{36}$$

This follows from $R_{X|W_1W_2}(\Delta) = \inf I(X;\hat{X}|W_1W_2)$, where the infimum is with respect to test channels $q_t(\hat{x}|x, w_1, w_2)$ such that $Ed_1(X, \hat{X}) \leqq \Delta$. Included in this class of test channels are those that are independent of w_2, i.e., $q_t(\hat{x}|x, w_1, w_2) = q_t(\hat{x}|x_1, w_1)$. This subclass is exactly the class of test channels in the minimization for computing $R_{X|W_1}(\Delta)$.

The final property of conditional rates is stated as a lemma below. The proof is given in the appendix.

Let $X \in \mathfrak{X}$ be a random variable with probability distribution $Q_X(x) = \Pr(X = x\}$, where $\mathfrak{X}$ is a finite source alphabet. Let $\hat{\mathfrak{X}}$ be a finite reproducing alphabet and let $d(x, \hat{x}) \geqq 0$, $x \in \mathfrak{X}$, $\hat{x} \in \hat{\mathfrak{X}}$ be a distortion function.

Now let $\{\mathcal{W}_k\}_{k=1}^n$ be a family of disjoint finite sets and let $\{p_k(x, w)\}_{k=1}^n$ be a family of probability distributions on $\mathfrak{X} \times \mathcal{W}_k$ such that

$$\sum_{w \in \mathcal{W}_k} p_k(x, w) = Q_X(x).$$

The random pairs (X, W_k) are defined by

$$\Pr\{X = x, W_k = w\} = p_k(x, w), \qquad x \in \mathfrak{X}, w \in \mathcal{W}_k.$$

Let $R_{X|W_k}(\Delta)$, $\Delta \geqq 0$ be the corresponding conditional rate-distortion function.

Next, set $\mathcal{W} = \sum_{k=1}^n \mathcal{W}_k$, where $\sum$ indicates union of disjoint sets. Define the probability distribution on $\mathfrak{X} \times \mathcal{W}$:

$$p^*(x, w) = \frac{1}{n} p_k(x, w), \qquad \text{for } w \in \mathcal{W}_k, 1 \leqq k \leqq n,$$

and let (X, W) be the corresponding random pair with conditional rate-distortion function $R_{X|W}(\Delta)$, $\Delta \geqq 0$. Clearly, $p^*(\cdot)$ is a mixture of the n disjoint probability distributions $\{p_k\}$, with prior probability $1/n$. We now state the lemma.

Lemma 5: For arbitrary $\{\Delta_k\}_{k=1}^n$, $\Delta_k \geqq 0$,

$$R_{X|W}\left(\frac{1}{n}\sum_1^n \Delta_k\right) \leqq \frac{1}{n}\sum_{k=1}^n R_{X|W_k}(\Delta_k).$$

We note here that all the above is meaningful for the case where $Q(x, y)$ is a probability density function or Q is an abstract probability measure. We need only make the obvious correspondences between discrete distributions and more general probability measures and replace "minimum" in (31), (33), (34), and (35) by "infimum."

We conclude this section by taking a look at the specialization of the above to the case where $d_1 = d_2 = d_H$, the Hamming distortion defined in (1b), and $\Delta_1 = \Delta_2 = 0$. Then

$$\begin{aligned} R_{XY}(0, 0) &= H(X, Y), \quad R_X(0) = H(X), \quad R_Y(0) = H(Y), \\ R_{XY|W}(0, 0) &= H(X, Y|W), \quad R_{X|W}(0) = H(X|W), \\ R_{Y|W}(0) &= H(Y|W), \end{aligned}$$

where the entropy $H(\cdot)$ and the conditional entropy $H(\cdot|\cdot)$ are defined in (2) and (3), respectively. Analogous to the relation

$$H(X|Y) + H(Y) = H(X, Y) \leqq H(X) + H(Y), \tag{37a}$$

which holds for this special case, the following is established in Ref. 5 for the general case:

$$R_{X|Y}(\Delta_1) + R_Y(\Delta_2) \leqq R_{XY}(\Delta_1, \Delta_2) \leqq R_X(\Delta_1) + R_Y(\Delta_2). \tag{37b}$$

Further, it is shown in Ref. 5, Corollary 3.2, that the left inequality in (37b) holds with equality in some neighborhood of the origin $\{(\Delta_1, \Delta_2): 0 \leqq \Delta_1, \Delta_2 \leqq \gamma\}$, provided that

$$Q(x, y) > 0, \quad \text{all } x \in \mathfrak{X}, y \in \mathfrak{Y}, \tag{38a}$$

and d_1, d_2 satisfy

$$\begin{aligned} d_1(x, \hat{x}) > d_1(x, x) = 0, \quad x \neq \hat{x}, \\ d_2(y, \hat{y}) > d_2(y, y) = 0, \quad y \neq \hat{y}. \end{aligned} \tag{38b}$$

2.3 Characterization of $\mathfrak{R}(\Delta_1, \Delta_2)$—the main result

We first state two simple theorems that are generalizations of Theorems 2 and 3. The proofs are analogous to the proofs of Section I, and are therefore omitted. Theorem 6(a) is also called the Pangloss bound.

Theorem 6: If $(R_0, R_1, R_2) \in R(\Delta_1, \Delta_2)$, *then*

(*a*) $R_0 + R_1 + R_2 \geqq R_{XY}(\Delta_1, \Delta_2)$.

(*b*) $R_0 + R_1 \geqq R_X(\Delta_1)$.

(*c*) $R_0 + R_2 \geqq R_Y(\Delta_2)$.

Theorem 7: *The following triples belong to* $\mathcal{R}(\Delta_1, \Delta_2)$:

(*A*) $R_0 = R_{XY}(\Delta_1, \Delta_2), \quad R_1 = R_2 = 0.$

(*B*) $R_0 = 0, \quad R_1 = R_X(\Delta_1), \quad R_2 = R_Y(\Delta_2).$

It is also possible to generalize Theorem 3(C) and (D), but this must await presentation of the main result, which we now give.

Consider first the case where $\mathfrak{X}$, $\mathcal{Y}$ are finite. Let $Q(x, y)$, $x \in \mathfrak{X}$, $y \in \mathcal{Y}$ be given. Now let $\mathcal{P}$ be the family of probability functions $p(x, y, w)$, where $x \in \mathfrak{X}$, $y \in \mathcal{Y}$, $w \in \mathcal{W}$, $\mathcal{W}$ is another finite set, and

$$\sum_{w \in \mathcal{W}} p(x, y, w) = Q(x, y), \quad x \in \mathfrak{X}, y \in \mathcal{Y}. \tag{39}$$

Thus, $\mathcal{P}$ is exactly as in Section 1.4. Now each $p \in \mathcal{P}$ defines three discrete random variables X, Y, W in the obvious way. For $p \in \mathcal{P}$ and $\Delta_1, \Delta_2 \geqq 0$, define the subset of Euclidean three-space

$$\mathcal{R}_1^{(p)}(\Delta_1, \Delta_2) = \{(R_0, R_1, R_2): R_0 \geqq I(X, Y; W), \\ R_2 \geqq R_{X|W}(\Delta_1), \qquad R_2 \geqq R_{Y|W}(\Delta_2)\}. \tag{40a}$$

Then let

$$\mathcal{R}^*(\Delta_1, \Delta_2) = [\bigcup_{p \in \mathcal{P}} \mathcal{R}^{(p)}(\Delta_1, \Delta_2)]^c, \tag{40b}$$

where $(\ )^c$ denotes set closure. Since $R_{X|W}(\Delta_1)$ and $R_{Y|W}(\Delta_2)$ are continuous for $\Delta_1, \Delta_2 \geqq 0$, we conclude that $\mathcal{R}^*(\Delta_1, \Delta_2)$ is continuous in (Δ_1, Δ_2) according to the Hausdorff set metric. This metric $\rho(S_1, S_2)$ between two subsets S_1, S_2 of a Euclidean space is defined by

$$\rho(S_1, S_2) = \sup_{r_1 \in S_1} \inf_{r_2 \in S_2} \|r_1 - r_2\| + \sup_{r_2 \in S_2} \inf_{r_1 \in S_1} \|r_1 - r_2\|,$$

where $\|\cdot\|$ denotes Euclidean norm.

If Q is either a density or a probability measure, then $\mathcal{R}^*(\Delta_1, \Delta_2)$ can be defined in an analogous way. In this more general case, we must require that the source has the property that there exists an $\hat{x} \in \hat{\mathfrak{X}}$, $\hat{y} \in \hat{\mathcal{Y}}$ such that

$$Ed_1(X, \hat{x}) < \infty, \quad Ed_2(Y, \hat{y}) < \infty. \tag{41}$$

If $\mathfrak{X}$, $\mathcal{Y}$ are finite, then (41) is always satisfied. We can now state our main result.

Theorem 8: $\mathcal{R}(\Delta_1, \Delta_2) = \mathcal{R}^*(\Delta_1, \Delta_2)$.

Remarks:

(1) Theorem 8 reduces to Theorem 4 when $\mathfrak{X}$, $\mathcal{Y}$ are finite, $d_1 = d_2 = d_H$, and $\Delta_1 = \Delta_2 = 0$.

(2) If we define $\mathcal{P}_T$ as in remark (1) following Theorem 4 as the set of test channels $p_t(w|x, y)$, then $\mathcal{P}_T$ is in 1-1 correspondence with $\mathcal{P}$.

(3) Since $\mathcal{R}(\Delta_1, \Delta_2)$ is convex, Theorem 8 implies that $\mathcal{R}^*(\Delta_1, \Delta_2)$ is convex also.

(4) Since Theorem 1 is valid for $\mathcal{R}(\Delta_1, \Delta_2)$, the present theorem implies that $T(\boldsymbol{\alpha})$, defined in (13) is also given by

$$T(\boldsymbol{\alpha}) = \inf_{p \in \mathcal{P}} [I(X, Y; W) + \alpha_1 R_{X|W}(\Delta_1) + \alpha_2 R_{Y|W}(\Delta_2)]. \quad (42)$$

Thus, from Theorem 1, the lower boundary $\overline{\mathcal{R}}(\Delta_1, \Delta_2)$, and therefore $\mathcal{R}(\Delta_1, \Delta_2)$, is determined by $T(\boldsymbol{\alpha})$ given in (42).

(5) As in remark (4) after Theorem 4, Theorems 6 and 7 can be obtained directly from Theorem 8. The steps parallel those in remark (4) and will be omitted. We will, however, give the generalization of Theorem 3(*C*) and (*D*). We state this as follows. The following triples $(R_0, R_1, R_2) \in \mathcal{R}(\Delta_1, \Delta_2)$:

$$(C)\ R_0 = R_Y(\Delta_2), \quad R_1 = R_{X|\hat{Y}}(\Delta_1), \quad R_2 = 0,$$
$$(D)\ R_0 = R_X(\Delta_1), \quad R_1 = 0, \quad R_2 = R_{Y|\hat{X}}(\Delta_2),$$

where the random variable $\hat{Y}$ is defined by the test channel that achieves the infimum in $R_Y(\Delta_2)$ (assuming that the infimum can be achieved; if not, a simple modification is possible), and $\hat{X}$ is defined by the test channel that achieves $R_X(\Delta_1)$. In the discrete case, we can achieve point (*C*) as follows. Let $p_t^*(\hat{y}|y)$ be the test channel that achieves $I(Y; \hat{Y}) = R_Y(\Delta_2)$. Let $\mathcal{W} = \hat{\mathcal{Y}}$ and let

$$p(x, y, \hat{y}) = Q(x, y) p_t^*(\hat{y}|y) \in \mathcal{P}.$$

The random variables X, Y, $\hat{Y}$ are defined in an obvious way by $p(x, y, \hat{y})$. Further, since X, $\hat{Y}$ are conditionally independent given Y,

$$\begin{aligned} I(X, Y; \hat{Y}) &= I(Y; \hat{Y}) + I(X; \hat{Y}|Y) \\ &= I(Y; \hat{Y}) = R_Y(\Delta_2). \end{aligned}$$

Also, the conditional rate $R_{Y|\hat{Y}}(\Delta_2) = 0$. Thus, from Theorem 8, with $W = \hat{Y}$, we have $(R_0, R_1, R_2) \in \mathcal{R}(\Delta_1, \Delta_2)$ where

$$\begin{aligned} R_0 &= I(X, Y; W) = R_Y(\Delta_2) \\ R_1 &= R_{X|W}(\Delta_1) = R_{X|\hat{Y}}(\Delta_1), \\ R_2 &= R_{Y|W}(\Delta_2) = 0. \end{aligned}$$

This is point (*C*). Point (*D*) is obtained on reversing the roles of X and Y.

Since $\mathcal{R}(\Delta_1, \Delta_2)$ is convex, any linear combination of points (A) and (B) of Theorem 7 and (C) and (D) above also belongs to $\mathcal{R}(\Delta_1, \Delta_2)$. But there is no guarantee in this case that points (C) and (D) will lie on the Pangloss plane. There are cases for which a portion of the Pangloss plane is known to be realizable, as is shown in the example below.

2.4 A technique for overbounding $\bar{\mathcal{R}}(\Delta_1, \Delta_2)$

In this section we present an intuitively sensible ad hoc scheme for choosing probability distributions $p \in \mathcal{P}$ that yield triples $[I(X, Y; W), R_{X|W}(\Delta_1), R_{Y|W}(\Delta_2)]$ which are often close to or actually on the boundary curve $\bar{\mathcal{R}}(\Delta_1, \Delta_2)$. In fact, in many cases this triple will lie on the Pangloss plane.

A natural coding scheme to apply to our network would be to send a "coarse" version of the source output $(\mathbf{X}, \mathbf{Y})$ over the common channel, and then send to each receiver over its private channel only the necessary "fine tuning" it needs to meet its fidelity requirement. This reasoning leads us to the following family of rate triples that belong to $\mathcal{R}(\Delta_1, \Delta_2)$. Assume for simplicity that $\mathcal{X}$, $\mathcal{Y}$, $\hat{\mathcal{X}}$, $\hat{\mathcal{Y}}$ are finite.

Let $\Delta_1, \Delta_2 \geqq 0$ be given. Let β_1, β_2 satisfy

$$\beta_1 \geqq \Delta_1, \quad \beta_2 \geqq \Delta_2.$$

Now let $q_t(\tilde{x}, \tilde{y} | x, y)$ be the test channel that achieves $I(X, Y; \tilde{X}, \tilde{Y}) = R_{XY}(\beta_1, \beta_2)$. Then with $W = (\tilde{X}, \tilde{Y})$ we have that the triple $(R_0, R_1, R_2) \in \mathcal{R}(\Delta_1, \Delta_2)$, where

$$R_0 = I(X, Y; W) = R_{XY}(\beta_1, \beta_2),$$

and

$$R_1 = R_{X|\tilde{X}\tilde{Y}}(\Delta_1), \quad R_2 = R_{Y|\tilde{X}\tilde{Y}}(\Delta_2). \tag{43}$$

Note that the rates corresponding to Theorem 7(A) and (B) and to points (C) and (D) in remark (5) following Theorem 8 can be generated as special cases of the rate in (43). We do this as follows:

A: Let $(\beta_1, \beta_2) = (\Delta_1, \Delta_2)$.

B: Let β_1, β_2 be large enough so that $R_{XY}(\beta_1, \beta_2) = 0$. Then $\tilde{X}$, $\tilde{Y}$ are degenerate.

C: Let β_1 be large enough so that $R_X(\beta_1) = 0$, and let $\beta_2 = \Delta_2$. Then $\tilde{X}$ is degenerate.

D: Let β_2 be large enough so that $R_Y(\beta_2) = 0$, and let $\beta_1 = \Delta_1$.

The power of this technique is illustrated by the following theorem, which asserts that under weak assumption the family of rates given by (43) includes a substantial subfamily that lies on the Pangloss bound and therefore on the boundary.

Theorem 9: Given a source that satisfies

(*i*) $\mathfrak{X} = \hat{\mathfrak{X}}$, $\mathcal{Y} = \hat{\mathcal{Y}}$, $\mathfrak{X}$, $\mathcal{Y}$ *finite,*
(*ii*) $Q(x, y) > 0$, *all* $x \in \mathfrak{X}$, $y \in \mathcal{Y}$,
(*iii*) $d_1(x, \hat{x}) > d_1(x, x) = 0$, *all distinct* x, $\hat{x} \in \mathfrak{X}$, *and* $d_2(y, \hat{y}) > d_2(y, y) = 0$, *all distinct* y, $\hat{y} \in \mathcal{Y}$.

Then there exists two neighborhoods of the origin

$$\eta_1 = \{(\Delta_1, \Delta_2) : 0 \leqq \Delta_1, \Delta_2 \leqq a\}$$
$$\eta_2 = \{(\beta_1, \beta_2) : 0 \leqq \beta_1, \beta_2 \leqq b\},$$

where $0 < a \leqq b$, *such that, if* $(\Delta_1, \Delta_2) \in \eta_1$ *and* $(\beta_1, \beta_2) \in \eta_2$, *then*

$$R_0 + R_1 + R_2 = R_{XY}(\Delta_1, \Delta_2),$$

where (R_0, R_1, R_2) *is given by* (43).

The theorem can be proved using Shannon lower-bound techniques[1,5] and, in particular, the proof is similar to that of Theorem 32 in Ref. 5. Since the proof requires the generation of special machinery that is only tangential to the main ideas in this paper, we have elected to omit it.

2.5 Examples

(*A*) Our first example will be the DSBS considered in the example of Section 1.5. Here $\mathfrak{X} = \mathcal{Y} = \{0, 1\}$, and

$$Q(x, y) = \tfrac{1}{2}(1 - p_0)\delta_{x,y} + \tfrac{1}{2}p_0(1 - \delta_{x,y}), \quad x, y = 0, 1, \tag{44}$$

where the parameter $p_0 \in [0, \frac{1}{2}]$. The distortion function will be the Hamming metric, i.e., $d_1 = d_2 = d_H$, where d_H is defined in (1b). Again, as in Section 1.4, we consider only the plane in (R_0, R_1, R_2)-space where $R_1 = R_2$ and $\Delta_1 = \Delta_2 = \Delta$. We employ the technique of Section 2.4 to obtain an upper bound for $\overline{\mathfrak{R}}(\Delta, \Delta)$.

Making use of Ref. 1, pp. 46–50 (Ex. 2.7.2), we have

$$R_{XY}(\beta, \beta) = \begin{cases} 1 + h(p_0) - 2h(\beta), & 0 \leqq \beta \leqq p_1 \\ L(1 - p_0) - \frac{1}{2}\{L(2\beta - p_0) + L[2(1 - \beta) - p_0]\}, & p_1 \leqq \beta \leqq \frac{1}{2} \end{cases} \tag{45a}$$

where

$$p_1 = \tfrac{1}{2} - \tfrac{1}{2}\sqrt{1 - 2p_0}, \tag{45b}$$

$$h(\lambda) = \lambda \log \lambda - (1 - \lambda) \log (1 - \lambda), \quad 0 \leqq \lambda \leqq 1, \tag{45c}$$

$$L(\lambda) = -\lambda \log \lambda, \qquad 0 \leqq \lambda \leqq 1. \tag{45d}$$

Now, from Ref. 1, the random variables $\tilde{X}$ and $\tilde{Y}$, which satisfy $I(X, Y; \tilde{X}\tilde{Y}) = R_{XY}(\beta, \beta)$ are such that

$$\begin{aligned}\Pr \{X = x | \tilde{X} = \tilde{x}, \tilde{Y} = \tilde{y}\} &= \Pr \{X = x | \tilde{X} = \tilde{x}\} \\ &= (1 - \beta)\delta_{x,\tilde{x}} + \beta(1 - \delta_{x,\tilde{x}}), \quad x, \tilde{x}, \tilde{y} = 0, 1\end{aligned} \tag{46a}$$

and

$$\begin{aligned}\Pr \{Y = y | \tilde{X} = \tilde{x}; \tilde{Y} = \tilde{y}\} &= \Pr \{Y = y | Y = \tilde{y}\} \\ &= (1 - \beta)\delta_{y,\tilde{y}} + \beta(1 - \delta_{y,\tilde{y}}), \quad y, \tilde{y}, \tilde{x} = 0, 1.\end{aligned} \tag{46b}$$

Thus, again from Ref. 1 (p. 46, Ex. 2.7.1), for $0 \leqq \Delta \leqq \beta$,

$$\begin{aligned}R_{X|\tilde{X}\tilde{Y}}(\Delta) &= R_{X|\tilde{X}}(\Delta) = R_{Y|\tilde{X}\tilde{Y}}(\Delta) = R_{Y|\tilde{Y}}(\Delta) \\ &= h(\beta) - h(\Delta).\end{aligned} \tag{47}$$

Thus, we conclude that, for arbitrary $0 \leqq \Delta \leqq \beta \leqq \frac{1}{2}$, the triple $(R_0, R_1, R_2) \in \mathcal{R}(\Delta, \Delta)$, where $R_0 = R_{XY}(\beta, \beta)$ [as in (45)], and $R_1 = R_2 = h(\beta) - h(\Delta)$. Let us note that, for $0 \leqq \Delta \leqq \beta \leqq p_1$, these rate-triples (R_0, R_1, R_2) satisfy

$$R_0 + R_1 + R_2 = 1 + h(p_0) - 2h(\Delta) = R_{XY}(\Delta, \Delta), \tag{48}$$

and therefore lie on the Pangloss plane and $\bar{\mathcal{R}}(\Delta, \Delta)$. One special case occurs when $\Delta = 0$, $\beta = p_1$. This yields the rate-triple of (24)—i.e., point G in Fig. 4. In fact, the distribution $p(x, y, w) \in \mathcal{P}$, which we guessed at in Section 1.5, was obtained by setting $W = (\tilde{X}, \tilde{Y})$, where $\tilde{X}$, $\tilde{Y}$ are as above for $\beta \geqq p_1$.

(*B*) Our second example is a source where $Q(x, y)$ is a density function and $\mathfrak{X}$, $\hat{\mathfrak{X}}$, $\mathfrak{Y}$, $\hat{\mathfrak{Y}}$ are the reals. The ad hoc technique used in the previous example (*A*) will work here with obvious modifications. The random variables X, Y in this case will be jointly gaussian with $EX = EY = 0$, $EX^2 = EY^2 = 1$, and $EXY = r$, $0 \leqq r \leqq 1$. Thus, the density

$$Q(x, y) = \frac{1}{2\pi(1 - r^2)^{\frac{1}{2}}} \exp \left\{ - \frac{(x^2 + y^2 - 2rxy)}{2(1 - r^2)} \right\}. \tag{49}$$

We take the distortion to be $d_1(\cdot, \cdot) = d_2(\cdot, \cdot)$, where

$$d_1(x, \hat{x}) = (x - \hat{x})^2, \quad -\infty < x, \hat{x} < \infty.$$

For $0 < \beta < \infty$, it can be shown[1,4] that

$$R_{XY}(\beta, \beta) = \begin{cases} \frac{1}{2}\log\left(\frac{1-r^2}{\beta^2}\right), & 0 \leqq \beta \leqq 1 - r, \\ \frac{1}{2}\log\left(\frac{1+r}{2\beta - (1-r)}\right), & 1 - r \leqq \beta \leqq 1, \\ 0, & \beta \geqq 1. \end{cases} \tag{50}$$

Further, the random variables $\tilde{X}$, $\tilde{Y}$ which satisfy $I(X, Y; \tilde{X}, \tilde{Y}) = R_{XY}(\beta, \beta)$, $0 < \beta \leqq 1$ are such that, given $\tilde{X} = \tilde{x}$, $\tilde{Y} = \tilde{y}$, X and Y are gaussian with

$$E(X|\tilde{X} = \tilde{x}, \tilde{Y} = \tilde{y}) = \tilde{x},$$
$$E(Y|\tilde{X} = \tilde{x}, \tilde{Y} = \tilde{y}) = \tilde{y},$$
$$\text{var}\,(X|\tilde{X} = \tilde{x}, \tilde{Y} = \tilde{y}) = \text{var}\,(Y|\tilde{X} = \tilde{x}, \tilde{Y} = \tilde{y}) = \beta.$$

Thus,[1,4] for $0 < \Delta \leqq \beta < 1$,

$$R_{X|\tilde{X}\tilde{Y}}(\Delta) = R_{Y|\tilde{X}\tilde{Y}}(\Delta) = \frac{1}{2}\log\frac{\beta}{\Delta}.$$

Thus, we conclude that, for arbitrary $0 < \Delta \leqq \beta \leqq 1$, the triple $(R_0, R_1, R_2) \in \mathcal{R}(\Delta, \Delta)$, where $R_0 = R_{XY}(\beta, \beta)$ [as in (50)] and $R_1 = R_2 = \frac{1}{2}\log \beta/\Delta$. Again, observe that for $0 \leqq \Delta \leqq \beta \leqq 1 - r$,

$$R_0 + R_1 + R_2 = \frac{1}{2}\log\left(\frac{1-r^2}{\Delta}\right) = R_{XY}(\Delta, \Delta), \tag{51}$$

and therefore (R_0, R_1, R_2) lies on the Pangloss plane and therefore on $\bar{\mathcal{R}}(\Delta, \Delta)$.

III. PROOF OF THE MAIN RESULT—THEOREM 8

The proof of Theorem 8 consists of two parts: (*i*) the "converse" part, which asserts that any point in $\mathcal{R}(\Delta_1, \Delta_2)$ belongs to $\mathcal{R}^*(\Delta_1, \Delta_2)$ and (*ii*) the "direct" or "positive" part, which asserts that any point in $\mathcal{R}^*(\Delta_1, \Delta_2)$ belongs to $\mathcal{R}(\Delta_1, \Delta_2)$. We give the proof for the case where $\mathcal{X}$, $\mathcal{Y}$ are finite sets. The proof for arbitrary $\mathcal{X}$, $\mathcal{Y}$ follows in a parallel way with integrals replacing sums, etc., in the standard way. We will begin with the converse.

3.1 The converse

Let $(f_E, f_D^{(X)}, f_D^{(Y)})$ define a code $(n, M_0, M_1, M_2, \Delta_X, \Delta_Y)$. We find a $p^*(x; y, w) \in \mathcal{P}$ for an appropriate set $\mathcal{W}$ such that

$$\left(\frac{1}{n}\log M_0, \frac{1}{n}\log M_1, \frac{1}{n}\log M_2\right) \in \mathcal{R}^{(p^*)}(\Delta_X, \Delta_Y) \subseteq \mathcal{R}^*(\Delta_X, \Delta_Y). \tag{52}$$

The converse follows on applying the definition of (Δ_1, Δ_2)-achievable rates and applying the continuity of $\mathcal{R}^*$ as discussed in Section II.

First, let $f_E(\mathbf{X}, \mathbf{Y}) = (S_0, S_1, S_2)$ where $S_i \in I_{M_i}$ is a random variable ($i = 0, 1, 2$). Then we have

$$\frac{1}{n}\log M_0 \overset{(1)}{\geqq} \frac{1}{n}H(S_0) \overset{(2)}{\geqq} \frac{1}{n}I(\mathbf{X}, \mathbf{Y}; S_0) \overset{(3)}{=} \frac{1}{n}[H(\mathbf{X}, \mathbf{Y}) - H(\mathbf{X}, \mathbf{Y}|S_0)]$$

$$\overset{(4)}{=} \frac{1}{n}\sum_{k=1}^{n}[H(X_k, Y_k) - H(X_k, Y_k|S_0, X_1, \cdots, X_{k-1}, Y_1, \cdots, Y_{k-1})]. \tag{53}$$

These steps are justified as follows:

(1) From $S_0 \in I_{M_0}$.
(2) Standard inequality.
(3) Definition of $I(\mathbf{X}, \mathbf{Y}; S_0)$.
(4) $H(\mathbf{X}, \mathbf{Y}) = \sum_k H(X_k, Y_k)$ follows from the independence of the pairs (X_k, Y_k), $k = 1, 2, \cdots, n$. The rest is also a standard identity.

Now, for $1 \leqq k \leqq n$, let $W_k = (S_0, X_1, \cdots, X_{k-1}, Y_1, \cdots, Y_{k-1})$, a random variable belonging to, say, $\mathcal{W}_k$, a finite set.† Relation (53) is then

$$\frac{1}{n}\log M_0 \geqq \frac{1}{n}\sum_{k=1}^{n} I(X_k, Y_k; W_k). \tag{54}$$

Next, let $\hat{\mathbf{X}} = f_D^{(X)} \circ f_E(\mathbf{X}, \mathbf{Y})$. Let $\Delta_{1k} = Ed_1(X_k, \hat{X}_k)$, $1 \leqq k \leqq n$. Then

$$\Delta_X = ED_1(\mathbf{X}, \hat{\mathbf{X}}) = \frac{1}{n}\sum_{k=1} \Delta_{1k}. \tag{55a}$$

We now write

$$\frac{1}{n}\log M_1 \overset{(1)}{\geqq} \frac{1}{n}H(\hat{\mathbf{X}}|S_0) \overset{(2)}{\geqq} \frac{1}{n}I(\mathbf{X}; \hat{\mathbf{X}}|S_0)$$

$$\overset{(3)}{=} \frac{1}{n}\sum_{k=1}^{n} I(X_k; \hat{\mathbf{X}}|S_0, X_1, \cdots, X_{k-1})$$

$$\overset{(4)}{\geqq} \frac{1}{n}\sum_{k=1}^{n} I(X_k; \hat{X}_k|S_0, X_1, \cdots, X_{k-1})$$

$$\overset{(5)}{\geqq} \frac{1}{n}\sum_{k=1}^{n} R_{X_k|V_k}(\Delta_{1k}) \overset{(6)}{\geqq} \frac{1}{n}\sum_{k=1}^{n} R_{X_k|W_k}(\Delta_{1k}), \tag{55b}$$

where $V_k = (S_0, X_1, \cdots, X_{k-1})$, and $W_k = (V_k, Y_1, \cdots, Y_{k-1})$ as above. These steps are justified as follows:

---

† We can, of course, take $\mathcal{W}_k = I_{M_0} \times \mathcal{X}^{k-1} \times \mathcal{Y}^{k-1}$.

(1) Since $\hat{X}$ is a function of S_0 and S_1, we have that, conditioned on $S_0 = s_0$, $\hat{X}$ can take no more than M_1 values (since $S_1 \in I_{M_1}$). Thus, $H(\hat{\mathbf{X}} | S_0 = s_0) \leqq \log M_1$, all s_0.

(2–4) Standard inequalities and identities.

(5) From the definition of $\mathcal{R}_{X_k|V_k}(\Delta_{1k})$, since $Ed_1(X_k, \hat{X}_k) = \Delta_{1k}$.

(6) Follows from (36).

A similar derivation yields

$$\frac{1}{n} \log M_2 \geqq \frac{1}{n} \sum_{k=1}^{n} R_{Y_k|W_k}(\Delta_{2k}), \tag{56a}$$

where

$$\Delta_Y = \frac{1}{n} \sum_{k=1}^{n} \Delta_{2k}. \tag{56b}$$

We are now in a position to define $p^*(x, y, w)$. With W_k defined as above, let

$$p_k(x, y, w) = \Pr \{X_k = x, Y_k = y, W_k = w\},$$
$$x \in \mathfrak{X}, y \in \mathcal{Y}, w \in \mathcal{W}_k.$$

Let

$$p_{1k}(x, w) = \sum_y p_k(x, y, w),$$
$$p_{2k}(y, w) = \sum_x p_k(x, y, w)$$

be the marginal distributions for (X_k, W_k) and (Y_k, W_k), respectively. The $\{\mathcal{W}_k\}$ can be considered a class of disjoint sets. Let $\mathcal{W} = \Sigma \mathcal{W}_k$, and define the probability function on $\mathfrak{X} \times \mathcal{Y} \times \mathcal{W}$

$$p^*(x, y, w) = \frac{1}{n} p_k(x, y, w), \quad w \in \mathcal{W}_k, \quad 1 \leqq k \leqq n.$$

Since

$$\sum_{w \in \mathcal{W}} p^*(x, y, w) = \sum_{k=1}^{n} \sum_{w \in \mathcal{W}_k} \frac{1}{n} p_k(x, y, w) = Q(x, y),$$

we have $p^* \in \mathcal{P}$. The random variables X, Y, W are defined by p^* in the obvious way. We can think of W as being generated by choosing an integer $K \in [1, n]$ without bias, and setting $W = W_k$ when $K = k$, $1 \leqq k \leqq n$. A straightforward calculation yields

$$I(X,Y; W) = \frac{1}{n} \sum_{k=1}^{n} I(X, Y; W_k). \tag{57a}$$

Furthermore, Lemma 5 can be applied to p_{1k}, p_{2k} to yield

$$R_{X|W}(\Delta_X) \leqq \frac{1}{n} \sum_{k=1}^{n} R_{X_k|W_k}(\Delta_k), \tag{57b}$$

$$R_{Y|W}(\Delta_Y) \leqq \frac{1}{n} \sum_{k=1}^{n} R_{Y_k|W_k}(\Delta_k). \tag{57c}$$

Inequalities (57a, b, c) can be substituted into (54), (55), (56), respectively, to obtain

$$\frac{1}{n} \log M_0 \geqq I(X, Y; W)$$

$$\frac{1}{n} \log M_1 \geqq R_{X|W}(\Delta_X)$$

$$\frac{1}{n} \log M_2 \geqq R_{Y|W}(\Delta_Y),$$

which is (52). This completes the proof of the converse.

3.2 The direct half

We begin the proof by stating a lemma concerning conventional source coding for a single memoryless source. The source is defined by a random variable $X \in \mathfrak{X}$, with probability distributions $Q_X(x)$, and a reproducing alphabet $\hat{\mathfrak{X}}$ with distortion function $d_1(x, \hat{x})$. As above, $\mathbf{X} = (X_1, \cdots, X_n)$ are n independent copies of X. Let $Q_X^{(n)}(\mathbf{x}) = \Pi_{k=1}^{n} Q_X(x_k)$ be the probability distribution for $\mathbf{X}$. Let $R(\Delta)$ be the rate-distortion function.

A source *code* with parameters (n, M) may be thought of as a mapping $f: \mathfrak{X}^n \rightarrow \mathfrak{C} \subseteq \hat{\mathfrak{X}}^n$, where card $\mathfrak{C} \leq M$. Let $\hat{\mathbf{X}} = (\hat{X}_1, \cdots, \hat{X}_n) = f(\mathbf{X})$. Then $D_1(\mathbf{X}, \hat{\mathbf{X}}) = 1/n \sum_{k=1}^{n} d_1(X_k, \hat{X}_k)$ is a random variable. We are interested in the quantity

$$\Gamma(\Delta_1 + \delta) = \Pr\{D_1(\mathbf{X}, \hat{\mathbf{X}}) \geqq \Delta + \delta\} = \sum_{\mathbf{x} \in \mathfrak{X}^n} Q_X^{(n)}(\mathbf{x}) \Phi(\mathbf{x}), \tag{58}$$

where $\Phi(\mathbf{x}) = 1$, if $D[\mathbf{x}, f(\mathbf{x})] \geqq \Delta + \delta$, and $\Phi(\mathbf{x}) = 0$, otherwise. We now state a lemma, which follows immediately from Lemma 9.3.1 and inequality (9.3.31) of Gallager.[4]

Lemma 10: Let $\Delta \geqq 0$, and ϵ, $\delta > 0$ be arbitrary. Then there exist A, $B > 0$ such that for all $n = 1, 2, \cdots$ there exists a code with parameters (n, M) satisfying

$$M \leqq 2^{n[R(\Delta)+\epsilon]},$$

and

$$\Gamma(\Delta + \delta) = \Pr\{D_1(\mathbf{X}, \hat{\mathbf{X}}) \geqq \Delta + \delta\} \leqq Ae^{-Bn}.$$

With the aid of this lemma the standard source coding theorem follows readily (Ref. 4, Theorem 9.3.1).

Next, let us consider a compound source for which the source output in n time units is an n-vector $\mathbf{X} = (X_1, \cdots, X_n) \in \mathfrak{X}^n$. The $\{X_k\}$ are still independent, but the X_k are not identically distributed.

Let $n_1, n_2, \cdots, n_J$ be such that $\sum_{j=1}^{J} n_j = n$, and let $Q_1(\cdot)$, $Q_2(\cdot), \cdots, Q_J(\cdot)$ be J probability distribution functions on X. The source is characterized by the fact that a known subset n_j of the n coordinates of $\mathbf{X}$ are distributed according to $Q_j(\cdot)$, $j = 1, \cdots, J$. Let $R_j(\Delta)$ be the rate-distortion function corresponding to $Q_j(\cdot)$ relative to the distortion function d_1. A code is defined exactly as above, and $\hat{\mathbf{X}} = f(\mathbf{X})$. We now have

Corollary 11: Let $\Delta_j \geqq 0$, $j = 1, 2, \cdots, J$, and ϵ, $\delta > 0$ be arbitrary. Then there exist $A_j, B_j > 0$, $j = 1, \cdots, J$, such that, for all $n = 1, 2, \cdots$ and any set $\{n_j\}_1^J$ such that $\Sigma n_j = n$, there exists a code with parameter M satisfying

$$M \leqq \prod_{j=0}^{J-1} \exp_2 \{n_j[R_j(\Delta_j) + \epsilon]\} = \exp_2 \{\sum n_j[R_j(\Delta_j) + \epsilon]\} \quad (59a)$$

and

$$\Gamma(\Delta + \delta) = \Pr \{D_1(\mathbf{X}, \hat{\mathbf{X}}) \geqq \Delta + \delta\} \leqq \sum_{j=0}^{J-1} A_j 2^{-B_j n_j}, \quad (59b)$$

where $\Delta = n^{-1}\sum n_j\Delta_j$. The (A_j, B_j)'s are the (A, B) of Lemma 10 corresponding to $Q_j(\cdot)$.

The corollary follows immediately from Lemma 10 on noting that, for any random variables $\{U_j\}$ and any set of constants $\{c_j\}$,

$$\Pr \{\sum_j U_j \geqq \sum_j c_j\} \leqq \sum_j \Pr \{U_j \geqq c_j\}.$$

Let us also remark that the $Q^{(n)}(\mathbf{x})$ used to compute $\Gamma(\Delta + \delta)$ in the corollary is of the form

$$Q_X^{(n)}(\mathbf{x}) = \prod_{k=1}^{n_1} Q_1(x_{i_{1k}}) \prod_{k=1}^{n_2} Q_2(x_{i_{2k}}) \cdots \prod_{k=1}^{n_J} Q_J(x_{i_{Jk}}), \quad (60)$$

where the i_{jk}th coordinate of $\mathbf{x}$ has distribution $Q_j(\cdot)$, $1 \leqq k \leqq n_j$, $0 \leqq j \leqq J - 1$.

Let us now turn to our network coding problem. An alternative (though equivalent) way of defining a code for our network with parameters (n, M_0, M_1, M_2) is

(1) A mapping

$$g: \mathcal{X}^n \times \mathcal{Y}^n \rightarrow \mathcal{C}, \tag{61a}$$

where $\mathcal{C}$ is an arbitrary set with cardinality $\leqq M_0$. The mapping g is called a "core code."

(2) For each $\mathbf{w} \in \mathcal{C}$, a mapping

$$g_{\mathbf{w}}^{(X)}: \mathcal{X}^n \times \mathcal{Y}^n \rightarrow \mathcal{C}_{\mathbf{w}}^{(X)} \subseteq \hat{\mathcal{X}}^n, \tag{61b}$$

where card $\mathcal{C}_{\mathbf{w}}^{(X)} \leqq M_1$.

(3) For each $\mathbf{w} \in \mathcal{C}$, a mapping

$$g_{\mathbf{w}}^{(Y)}: \mathcal{X}^n \times \mathcal{Y}^n \rightarrow \mathcal{C}_{\mathbf{w}}^{(Y)} \subseteq \hat{\mathcal{Y}}^n, \tag{61c}$$

where card $\mathcal{C}_{\mathbf{w}}^{(Y)} \leqq M_2$.

The code defined in this way can be used on our network (Fig. 2) as follows. Let $\mathcal{C} = \{\mathbf{w}_i\}_1^{M_0}$. Then, if $g(\mathbf{x}, \mathbf{y}) = \mathbf{w}_i$, the index i is transmitted over the common channel. Let $\mathcal{C}_{\mathbf{w}_i}^{(X)} = \{\hat{\mathbf{x}}_{il}\}_{l=1}^{M_1}$, $1 \leqq i \leqq M_0$. Then, if $g(\mathbf{x}, \mathbf{y}) = \mathbf{w}_i$, and $g_{\mathbf{w}_i}^{(X)}(\mathbf{x}, \mathbf{y}) = \hat{\mathbf{x}}_{il}$, we transmit the index l over the private channel to receiver 1. The decoder at receiver 1, knowing the indices i and l, emits $\hat{\mathbf{x}}_{il}$, and the resulting distortion is $D_1(\mathbf{x}, \hat{\mathbf{x}}_{il})$. Receiver 2 works analogously.

Let us fix our attention on receiver 1, and assume that $g_{\mathbf{w}_i}^{(X)}(\mathbf{x}, \mathbf{y}) = g_{\mathbf{w}_i}^{(X)}(\mathbf{x})$. Then we define the quantity $q(\mathbf{x}, \mathbf{w}_i)(\mathbf{x} \in \mathcal{X}^n, \mathbf{w}_i \in \mathcal{C})$ as the probability that $\mathbf{X} = \mathbf{x} \in \mathcal{X}^n$ and $\mathbf{Y} = \mathbf{y}$ such that $g(\mathbf{x}, \mathbf{y}) = \mathbf{w}_i$. Thus,

$$q(\mathbf{x}, \mathbf{w}_i) = \sum_{\mathbf{y}: g(\mathbf{x}, \mathbf{y}) = \mathbf{w}_i} Q^{(n)}(\mathbf{x}, \mathbf{y}). \tag{62}$$

$Q^{(n)}(\mathbf{x}, \mathbf{y}) \triangleq \prod_{k=1}^{n} Q(x_k, y_k)$ is the probability distribution for $\mathbf{X}$ and $\mathbf{Y}$.

Then, as in (58) with $\hat{X} = g_{\mathbf{W}}^{(X)}(\mathbf{X})$, $\mathbf{W} = g(\mathbf{X}, \mathbf{Y})$,

$$\Gamma(\Delta_1 + \delta) \triangleq \Pr\{D_1(\mathbf{X}, \hat{\mathbf{X}}) > \Delta_1 + \delta\} = \sum_{i=1}^{M_0} \sum_{\mathbf{x}} q(\mathbf{x}, \mathbf{w}_i)\Phi_i(\mathbf{x}), \tag{63a}$$

where

$$\Phi_i(\mathbf{x}) = \begin{cases} 1, \text{ if } D_1[\mathbf{x}, g_{\mathbf{w}_i}^{(X)}(\mathbf{x})] > \Delta_1 + \delta, \\ 0, \text{ otherwise.} \end{cases} \tag{63b}$$

Substituting (62) into (63), we obtain

$$\Gamma(\Delta_1 + \delta) = \sum_{i=1}^{M_0} \{ \sum_{(x,y) \in G_i} Q^{(n)}(\mathbf{x}, \mathbf{y})\Phi_i(\mathbf{x}) \}, \tag{64a}$$

where

$$G_i = \{(\mathbf{x}, \mathbf{y}) : g(\mathbf{x}, \mathbf{y}) = \mathbf{w}_i\}, \quad 1 \leqq i \leqq M_0. \tag{64b}$$

Now our goal is to show that there exists a code for n sufficiently large, with M_0, M_1, M_2 appropriately chosen, and with $\Gamma(\Delta_1 + \delta)$ arbitrarily small.

Let us assume that we are given a probability distribution $p(x, y, w) \in \mathcal{P}$, where $x \in \mathfrak{X}$, $y \in \mathcal{Y}$, $w \in \mathcal{W}$. Let $p_W(w) = \sum_{x,y} p(x, y, w)$, $w \in \mathcal{W}$ be the marginal distribution of W. Assume with no loss of generality that $p_W(w) > 0$. Let

$$p_b(x, y \mid w) = \frac{p(x, y, w)}{p_W(w)}$$

be the "backward test channel." For $\mathbf{x} \in \mathfrak{X}^n$, $\mathbf{y} \in \mathcal{Y}^n$, $\mathbf{w} \in \mathcal{W}^n$, let

$$p^{(n)}(\mathbf{x}, \mathbf{y}, \mathbf{w}) = \Pi_{k=1}^n p(x_k, y_k, w_k)$$

be the probability distribution for $\mathbf{X}$, $\mathbf{Y}$, $\mathbf{W}$ (n independent drawings of X, Y, W). Let $p_W^{(n)}(\mathbf{w}) = \prod_{k=1}^n p_W(w_k)$, and $p_b^{(n)}(\mathbf{x}, \mathbf{y} \mid \mathbf{w}) = \prod_{k=1}^n p_b(x_k, y_k \mid w_k)$. For $(\mathbf{x}, \mathbf{y}, \mathbf{w}) \in \mathfrak{X}^n \times \mathcal{Y}^n \times \mathcal{W}^n$, let

$$i^{(n)}(\mathbf{x}, \mathbf{y}; \mathbf{w}) = \log \frac{p_b^{(n)}(\mathbf{x}, \mathbf{y} \mid \mathbf{w})}{Q^{(n)}(\mathbf{x}, \mathbf{y})} = \sum_{k=1}^n \log \frac{p_b(x_k, y_k \mid w_k)}{Q(x_k, y_k)}, \tag{65}$$

be the information "density." Of course,

$$Ei^{(n)}(\mathbf{X}, \mathbf{Y}; \mathbf{W}) = I\{\mathbf{X}, \mathbf{Y}; \mathbf{W}\} = nI\{X, Y; W\}.$$

Finally, let $\Delta_1 \geqq 0$ be given and let $\{\Delta_w\}_{w \in \mathcal{W}}$ satisfy

$$R_{X|W}(\Delta_1) = \sum_{w \in \mathcal{W}} p_W(w) R_{X|W=w}(\Delta_w), \tag{66a}$$

and

$$\Delta_1 = \sum_{w \in \mathcal{W}} p_W(w) \Delta_w. \tag{66b}$$

See (35). A similar expression can be written for $R_{Y|W}(\Delta_2)$.

We now return to our network coding problem. With $p \in \mathcal{P}$ given, we set out to construct a core code g with certain desirable properties. For any core code g: $\mathfrak{X}^n \times \mathcal{Y}^n \to \mathcal{C} = \{\mathbf{w}_i\}_1^{M_0} \subseteq \mathcal{W}^n$, let N_{iw} = the number of occurrences of symbol w in code vector $\mathbf{w}_i$, $1 \leqq i \leqq M_0$, $w \in \mathcal{W}$. The existence of a desirable core code is assured by

Lemma 12: *Let* $p \in \mathcal{P}$ *and* $\epsilon > 0$ *be arbitrary. Let* $I^* = I(X, Y; W)$ *correspond to* $p \in \mathcal{P}$. *For* n *sufficiently large, there exists a code* g *as in*

(61a) *such that*

$$(i)\ M_0 \leqq 2^{n(I^*+\epsilon)}$$

$$(ii)\ \left|\frac{N_{iw}}{n} - p_W(w)\right| \leqq \epsilon[\min_w p_W(w)], \quad \textit{for all } w \in \mathcal{W},$$

$$(iii)\ \Pr(S_\epsilon^c) = \sum_{(\mathbf{x},\mathbf{y}) \notin S_\epsilon} Q^{(n)}(\mathbf{x}, \mathbf{y}) \leqq \epsilon,$$

where

$$S_\epsilon = \{(\mathbf{x}, \mathbf{y}) : \frac{1}{n} i^{(n)}[\mathbf{x}, \mathbf{y}; g(\mathbf{x}, \mathbf{y})] \geqq I^* - \epsilon\},$$

and $i(\mathbf{x}, \mathbf{y}; \mathbf{w})$ *is defined in* (68).

We defer discussion on the proof of Lemma 12 to the end of this section.

Let us suppose that g is a code that satisfies conditions (i), (ii), and (iii) of Lemma 12. Let $\{g_{\mathbf{w}_i}^{(X)}\}_1^{M_0}$ be a family of encoders as in (61b). Consider expression (64a). The term in braces is

$$\sum_{(\mathbf{x},\mathbf{y}) \in G_i} Q^{(n)}(\mathbf{x}, \mathbf{y})\Phi_i(\mathbf{x})$$

$$= \sum_{(\mathbf{x},\mathbf{y}) \in G_i \cap S_\epsilon} Q^{(n)}(\mathbf{x}, \mathbf{y})\Phi_i(\mathbf{x}) + \sum_{(\mathbf{x},\mathbf{y}) \in G_i \cap S_\epsilon^c} Q^{(n)}(\mathbf{x}, \mathbf{y}). \quad (67)$$

But if $(\mathbf{x}, \mathbf{y}) \in G_i$ [i.e., $g(\mathbf{x}, \mathbf{y}) = \mathbf{w}_i$)] and $(\mathbf{x}, \mathbf{y}) \in S_\epsilon$, then

$$Q^{(n)}(\mathbf{x}, \mathbf{y}) \leqq 2^{-(I^*-\epsilon)n} p_b^{(n)}(\mathbf{x}, \mathbf{y} | \mathbf{w}_i),$$

so that the first summation in the right member of (67) can be overbounded:

$$\begin{aligned} &\leqq 2^{-(I^*-\epsilon)n} \sum_{\substack{(\mathbf{x},\mathbf{y}) : g(\mathbf{x},\mathbf{y}) = w_i \\ (\mathbf{x},\mathbf{y}) \in S_\epsilon}} p_b^{(n)}(\mathbf{x}, \mathbf{y} | \mathbf{w}_i)\Phi_i(\mathbf{x}) \\ &\leqq 2^{-(I^*-\epsilon)n} \sum_{\mathbf{x},\mathbf{y}} p_b^{(n)}(\mathbf{x}, \mathbf{y}) | w_i)\Phi_i(\mathbf{x}) \qquad (68) \\ &\leqq 2^{-(I^*-\epsilon)n} \sum_{\mathbf{x}} p_b^{(n)}(\mathbf{x} | \mathbf{w}_i)\Phi_i(\mathbf{x}). \end{aligned}$$

Combining (68), (67), and (64), we have

$$\Gamma(\Delta_1 + \delta) = \Pr\{D_1(\mathbf{x}, \hat{\mathbf{x}}) > \Delta_1 + \delta\}$$

$$\leqq 2^{-(I^*-\epsilon)n} \sum_{i=1}^{M_0} \sum_{\mathbf{x}} p_b^{(n)}(\mathbf{x} | \mathbf{w}_i)\Phi_i(\mathbf{x}) + \sum_{i=1}^{M_0} \sum_{(\mathbf{x},\mathbf{y}) \in G_i \cap S_\epsilon^c} Q^{(n)}(\mathbf{x}, \mathbf{y})$$

$$\leqq 2^{-(I^*-\epsilon)n} \sum_{i=1}^{M_0} \sum_{\mathbf{x}} p_b^{(n)}(\mathbf{x} | \mathbf{w}_i)\Phi_i(\mathbf{x}) + \Pr(S_\epsilon^c). \quad (69)$$

Now consider

$$p_b^{(n)}(\mathbf{x}|\mathbf{w}_i) = \prod_{k=1}^{n} p_b(x_k|w_{ik}),$$

where $\mathbf{w}_i = (w_{i1}, w_{i2}, \cdots, w_{in})$, $1 \leqq i \leqq M_0$. With N_{iw} as defined just above Lemma 12, we see that (for a given i) $p_b^{(n)}(\mathbf{x}|\mathbf{w}_i)$ is the same form as $Q^{(n)}(\mathbf{x})$ in (60) with $n_j = N_{iw}$. It then follows from Corollary 11 that, with $\mathbf{w}_i$ held fixed, we can find a source code for X—i.e., a mapping $g_{\mathbf{w}_i}^{(X)}$—with parameter $M = M_1$ such that, for arbitrary $\epsilon, \delta > 0$,

$$M_1 \leqq \exp_2 \{ \sum_{w \in \mathcal{W}} N_{iw}[R_{X|W=w}(\Delta_w) + \epsilon]\}, \tag{70a}$$

and

$$\sum_{\mathbf{x}} p_b^{(n)}(\mathbf{x}|\mathbf{w}_i)\Phi_i(\mathbf{x}) \leqq \sum_{w} A_w 2^{-B_w N_{iw}}, \tag{70b}$$

where $\Phi_i(\mathbf{x})$ is defined in (63b) with $\Delta_1 = n^{-1}\sum_w N_{iw}\Delta_w$, and $\{\Delta_w\}$ satisfying (66). The $\{(A_w, B_w)\}$ correspond to $p_b(x|w)$. Further, since the $\{N_{iw}\}$ satisfy condition (*ii*) of Lemma 12, (70) becomes [using (66)]

$$\begin{aligned} M_1 &\leqq \exp_2 \{n \sum_{w \in \mathcal{W}} (p_W[w] + \epsilon)(R_{X|W=w}[\Delta_w] + \epsilon)\} \\ &\leqq \exp_2 \{n(R_{X|W}[\Delta_1] + \epsilon H[X] + \epsilon)\} \end{aligned} \tag{71a}$$

and

$$\sum_{\mathbf{x}} p_b^{(n)}(\mathbf{x}|\mathbf{w}_i)\Phi_i(\mathbf{x}) \leqq \sum_{w \in \mathcal{W}} A_w 2^{-B_w n p_W(w)(1-\epsilon)} \leqq C 2^{-nB(1-\epsilon)}, \tag{71b}$$

where $C = (\text{card } \mathcal{W}) \cdot \max_w A_w$ and $B = \min_w B_w p_w(w)$. Substituting (71b) into (69) and using conditions (*i*) and (*iii*) of Lemma 12, we have

$$\begin{aligned} \Gamma(\Delta_1 + \delta) &\leqq 2^{-n(I^*-\epsilon)} M_0 \cdot C 2^{-nB(1-\epsilon)} + \Pr(S_\epsilon^c) \\ &\leqq 2^{-n(B-B\epsilon-2\epsilon)} + \epsilon \to 0, \quad \text{as } n \to \infty \text{ and then } \epsilon \to 0. \end{aligned}$$

Since we can do an identical construction for Y, we have proved

Lemma 13: *Let* $p \in \mathcal{P}$, *and let the corresponding information be* $I(X, Y; W) = I^*$. *Let* $\Delta_1, \Delta_2 \geqq 0$ *and* $\epsilon, \delta > 0$ *be arbitrary. Then, for* n *sufficiently large, there exists a coding scheme as in* (61) *with parameters* (n, M_0, M_1, M_2) *such that*

(*i*) $M_0 \leqq 2^{n(I^*+\epsilon)}$,
(*ii*) $M_1 \leqq 2^{n(R_{X|W}(\Delta_1)+\epsilon)}$,
(*iii*) $M_2 \leqq 2^{n(R_{Y|W}(\Delta_2)+\epsilon)}$,

and

(*iv*) $\Pr\{D_1(\mathbf{X}, \hat{\mathbf{X}}) > \Delta_1 + \delta\} \leqq \epsilon$,
(*v*) $\Pr\{D_2(\mathbf{Y}, \mathbf{Y}) > \Delta_2 + \delta\} \leqq \epsilon$.

The following corollary follows from Lemma 13 in the usual way (exactly as does Theorem 9.3.1 in Gallager[4]).

Corollary 14: *Let* $p \in \mathcal{P}$, *and let the corresponding information* $I(X, Y; W) = I^*$. *Then, for arbitrary* $\Delta_1, \Delta_2 \geqq 0$, *the rate-triple* $[I^*, R_{X|W}(\Delta_1), R_{Y|W}(\Delta_2)]$ *is* (Δ_1, Δ_2)*-achievable. Thus,* $\mathcal{R}^{(p)}(\Delta_1, \Delta_2) \subseteq \mathcal{R}(\Delta_1, \Delta_2)$, *for all* $p \in \mathcal{P}$.

The direct-half now follows on noting that, if $S_1 \subseteq S_2$ and S_2 is closed, then the closure $S_1^c \subseteq S_2$. Thus, $\mathcal{R}^*(\Delta_1, \Delta_2) = [\bigcup_p \mathcal{R}^{(p)}(\Delta_1, \Delta_2)]^c \subseteq \mathcal{R}(\Delta_1, \Delta_2)$, which is what we had to prove.

It remains to prove Lemma 12. Since the proof is nearly identical to that of Lemma 9.3.1 in Gallager,[4] we will only outline the steps. Let $\epsilon > 0$ be arbitrary. For $\hat{\mathbf{w}} \in \mathcal{W}^n$, let $N_w(\hat{\mathbf{w}})$ = the number of occurrences of symbol $w \in \mathcal{W}$ in the n-vector $\hat{\mathbf{w}}$. Then define

$$T(\epsilon) = \left\{ \hat{\mathbf{w}} \in \mathcal{W}^n : \text{all } w \in \mathcal{W}, \left| \frac{N_w(\hat{\mathbf{w}})}{n} - p_w(w) \right| \leqq \epsilon \right\}.$$

Then, paralleling Gallager, there exists a mapping g [as in (61a)] for which

$$M_0 \leqq 2^{n(I^* + \epsilon)}$$

and

$$\Pr\left\{ \frac{1}{n} i^{(n)}[\mathbf{X}, \mathbf{Y}; g(\mathbf{X}, \mathbf{Y})] \leqq I^* - \epsilon \quad \text{or} \quad g(\mathbf{X}, \mathbf{Y}) \notin T(\epsilon) \right\}$$
$$\leqq P_t(A) + \exp\{-e^{n(\epsilon - \epsilon_2)}\} \triangleq \xi(n),$$

where $\epsilon_2 > 0$ is arbitrary and

$$A = \left\{ (\mathbf{X}, \mathbf{Y}, \mathbf{W}) : \text{either } \frac{1}{n} i^{(n)}(\mathbf{X}, \mathbf{Y}; \mathbf{W}) > I^* + \epsilon_2 \text{ or} \right.$$
$$\left. \frac{1}{n} i^{(n)}(\mathbf{X}, \mathbf{Y}; \mathbf{W}) \leqq I^* - \epsilon, \text{ or } \mathbf{W} \notin T(\epsilon) \right\},$$

and $P_t(\cdot)$ is probability computed with respect to $p(x, y, w) \in \mathcal{P}$. By the weak law of large numbers, if $\epsilon_2 < \epsilon$, then $\xi_n \to 0$, as $n \to \infty$.

Let the code whose existence we have just asserted be $\{\mathbf{w}_i\}_1^{M_0}$. There must be at least one code vector, say, $\mathbf{w}_1$, which belongs to $T(\epsilon)$. Now

$$\Pr\{g(\mathbf{X}, \mathbf{Y}) \notin T(\epsilon)\} \leqq \xi(n).$$

If $\mathbf{x}, \mathbf{y}$ are such that $g(\mathbf{x}, \mathbf{y}) \notin T(\epsilon)$, change $g(\mathbf{x}, \mathbf{y})$ to $\mathbf{w}_1$. The new code has $g(\mathbf{x}, \mathbf{y}) \in T(\epsilon)$ and

$$\Pr\left\{\frac{1}{n} i^{(n)}[\mathbf{X}, \mathbf{Y}; g(\mathbf{X}, \mathbf{Y})] \leqq I^* - \epsilon\right\} \leqq 2\xi(n) \xrightarrow{n} 0.$$

Thus, this new code satisfies conditions (*i*), (*ii*), and (*iii*) of Lemma 12.

IV. ACKNOWLEDGMENT

Useful discussions with T. Cover, J. Wolf, D. Slepian, M. Kaplan, and H. Witsenhausen are acknowledged with thanks.

APPENDIX

A.1 Proof of the convexity of $\mathcal{R}(\Delta_1, \Delta_2)$

Let Δ_1, Δ_2 be given and held fixed. Write $\mathcal{R}(\Delta_1, \Delta_2)$ as $\mathcal{R}$.

Theorem 15: *$\mathcal{R}$ is convex.*

Proof: The theorem follows by a "time-sharing" argument. Let $\mathbf{R}^{(1)}, \mathbf{R}^{(2)} \in \mathcal{R}$, and $0 \leqq \theta \leqq 1$. We must show that

$$\mathbf{R} \in \mathcal{R}, \tag{72a}$$

$$\mathbf{R} = \theta \mathbf{R}^{(1)} + (1 - \theta)\mathbf{R}^{(2)}. \tag{72b}$$

Let $(g_E, g_D^{(X)}, g_D^{(Y)})$ and $(h_E, h_D^{(X)}, h_D^{(Y)})$ be codes with parameters $(n_1, M_0^{(1)}, M_1^{(1)}, M_2^{(1)}, \Delta_X^{(1)}, \Delta_Y^{(1)})$ and $(n_2, M_0^{(2)}, M_1^{(2)}, M_2^{(2)}, \Delta_X^{(2)}, \Delta_Y^{(2)})$, respectively, where $\Delta_X^{(1)}, \Delta_X^{(2)} \leqq \Delta_1$, $\Delta_Y^{(1)}, \Delta_Y^{(2)} \leqq \Delta_2$. Say $\theta = A/B$, where A, B are integers, $0 \leqq A \leqq B \leqq \infty$. We show how to construct a code $(n, M_0, M_1, M_2, \Delta_X, \Delta_Y)$, where

$$\frac{1}{n} \log M_i = \theta\left(\frac{1}{n_1} \log M_i^{(1)}\right) + (1 - \theta)\left(\frac{1}{n_2} \log M_i^{(2)}\right), \tag{73}$$

$(i = 0, 1, 2)$, and $\Delta_X \leqq \Delta_1$, $\Delta_Y \leqq \Delta_2$. This will establish (72) for rational θ. Since the region $\mathcal{R}$ is closed, (72) must hold for all θ, establishing Theorem 15.

We now define a code with block length $n = cn_1 + dn_2$, where $c = An_2$, $d = (B - A)n_1$. Let $(\mathbf{x}, \mathbf{y}) \in \mathcal{X}^n \times \mathcal{Y}^n$ be a sequence of n pairs. Partition this sequence into c blocks of n_1 pairs and d blocks of n_2 pairs. Encode-decode the first c blocks using encoder-decoder $(g_E, g_D^{(X)}, g_D^{(Y)})$, and encode-decode the remaining d blocks using encoder-decoder $(h_E, h_D^{(X)}, h_D^{(Y)})$. Denote this combination encoder-decoder by $(f_E, f_D^{(X)}, f_D^{(Y)})$. Consider $f_E(\mathbf{x}, \mathbf{y}) = (S_0, S_1, S_2)$. The quantity

$S_i (i = 0, 1, 2)$ takes values in a set with

$$(M_i^{(1)})^c \cdot (M_i^{(2)})^d \triangleq M_i$$

members. This set can, of course, be put in 1-1 correspondence with I_{M_i}. Thus, for $i = 0, 1, 2$,

$$\frac{1}{n} \log M_i = \frac{c}{n} \log M_i^{(1)} + \frac{d}{n} \log M_i^{(2)} = \frac{\theta}{n_1} \log M_i^{(1)} + \frac{(1-\theta)}{n_2} \log M_i^{(2)},$$

which is (73). Further, the new code has $\Delta_X \leqq \Delta_1$, and $\Delta_Y \leqq \Delta_2$, so that the lemma follows.

A.2 Proof of Theorem 1

Again let $\Delta_1, \Delta_1 \geqq 0$ be given and fixed, and write $\mathcal{R}(\Delta_1, \Delta_2) = \mathcal{R}$, and $\overline{\mathcal{R}}(\Delta_1, \Delta_2) = \overline{\mathcal{R}}$.

We first establish part (*ii*) of the theorem. Let $\mathbf{R} \in \mathcal{S}(\boldsymbol{\alpha})$, $\boldsymbol{\alpha} \in \mathcal{A}'$. If $\mathbf{R} \notin \overline{\mathcal{R}}$, then there exists an $\hat{\mathbf{R}} = (\hat{R}_0, \hat{R}_1, \hat{R}_2) \in \mathcal{R}$, such that $\hat{R}_i \leqq R_i$, $i = 0, 1, 2$, and at least one of these inequalities holds strictly. Thus,

$$C(\boldsymbol{\alpha}, \hat{\mathbf{R}}) - C(\boldsymbol{\alpha}, \mathbf{R}) = (\hat{R}_0 - R_0) + \alpha_1(\hat{R}_1 - R_1) + \alpha_2(\hat{R}_2 - R_2) < 0. \quad (74)$$

The inequality follows from $\alpha_1, \alpha_2 > 0$. This contradicts $\mathbf{R} \in \mathcal{S}(\boldsymbol{\alpha})$. Thus $\mathbf{R} \in \overline{\mathcal{R}}$, which establishes part (*ii*).

It remains to establish part (*i*). We must first obtain some preliminary facts.

Lemma 16: Let $(R_0, R_1, R_2) \in \mathcal{R}$. Then

(*a*) *for* $a_i \geqq 0$ $(i = 0, 1, 2)$, $(R_0 + a_0, R_1 + a_1, R_2 + a_2) \in \mathcal{R}$,
(*b*) *for* $0 \leqq \theta \leqq 1$, $[(1-\theta)R_0, R_1 + \theta R_0, R_2 + \theta R_0] \in \mathcal{R}$,
(*c*) *for* $0 \leqq \theta_1, \theta_2 \leqq 1$,

$$[R_0 + \theta_1 R_1 + \theta_2 R_2, (1-\theta_1)R_1, (1-\theta_2)R_2] \in \mathcal{R}.$$

Proof:

(*a*) follows immediately from the definition of $\mathcal{R}$.

(*b*) follows on noting that data sent through the common channel can be transmitted instead through *each* private channel.

(*c*) follows on noting that any data transmitted through either private channel can be transmitted instead over the common channel.

Next, for $R_1, R_2 \geqq 0$ write $\mathbf{r} = (R_1, R_2)$, and define the function

$$F(\mathbf{r}) = F(R_1, R_2) = \min\{R_0: (R_0, R_1, R_2) \in \mathcal{R}\}. \tag{75}$$

The minimum exists because $\mathcal{R}$ is closed. Clearly, $(R_0, R_1, R_2) \in \overline{\mathcal{R}}$ only if $R_0 = F(R_1, R_2)$.

Lemma 17: $F(\mathbf{r})$ is convex.

Proof: Let $\mathbf{r}^{(1)}, \mathbf{r}^{(2)}$ be arbitrary. Then $[F(\mathbf{r}^{(1)}), \mathbf{r}^{(1)}], [F(\mathbf{r}^{(2)}), \mathbf{r}^{(2)}] \in \mathcal{R}$. Since $\mathcal{R}$ is convex, for $0 \leqq \theta \leqq 1$,

$$\theta[F(\mathbf{r}^{(1)}), \mathbf{r}^{(1)}] + (1 - \theta)[F(\mathbf{r}^{(2)}), \mathbf{r}^{(2)}]$$
$$= [\theta F(\mathbf{r}^{(1)}) + (1 - \theta)F(\mathbf{r}^{(2)}), \theta\mathbf{r}^{(1)} + (1 - \theta)\mathbf{r}^{(2)}] \in \mathcal{R}.$$

Thus, by the definition of $F(\cdot)$,

$$F[\theta\mathbf{r}_1^{(1)} + (1 - \theta)\mathbf{r}^{(2)}] \leqq \theta F(\mathbf{r}^{(1)}) + (1 - \theta)F(\mathbf{r}^{(2)}),$$

establishing the lemma.

Now it follows from the convexity of $F(\cdot)$ that, for arbitrary $\mathbf{r}^* = (R_1^*, R_2^*)$, $R_1^* R_2^* \geqq 0$, there exists constants $\alpha_i = \alpha_i(\mathbf{r}^*)$, $i = 1, 2$, such that, for all $\mathbf{r}$,

$$F(\mathbf{r}) - F(\mathbf{r}^*) \geqq \sum_{i=1}^{2} \alpha_i(R_i^* - R_i). \tag{76}$$

This is a statement of the well-known fact that any convex curve lies above a plane of support. Here the curve is the locus of points in (R_0, R_1, R_2)-space given by $R_0 = F(\mathbf{r}) = F(R_1, R_2)$, and the plane is the locus of points $R_0 = F(\mathbf{r}^*) + \sum_{i=1}^{2}\alpha_i(R_i^* - R_i)$. Note that the curve and the plane coincide at $\mathbf{r} = \mathbf{r}^*$.

Now let $\mathbf{R}^* = (R_0^*, R_1^*, R_2^*) \in \overline{\mathcal{R}}$. Then $R_0^* = F(R_1^*, R_2^*)$. Let $\mathbf{R} = (R_0, R_1, R_2)$ be any triple in $\mathcal{R}$. Then with $\mathbf{r} = (R_1, R_2)$, (76) yields

$$F(r) + \alpha_1 R_1 + \alpha_2 R_2 \geqq R_0^* + \alpha_1 R_1^* + \alpha_2 R_2^*.$$

Since, by definition of $F(\cdot)$, $F(\mathbf{r}) \leqq R_0$, we have:

$$R_0^* + \alpha_1 R_1^* + \alpha_2 R_2^* = \min_{\mathbf{R}\in\mathcal{R}} (R_0 + \alpha_1 R_1 + \alpha_2 R_2)$$
$$= \min_{\mathbf{R}\in\mathcal{R}} C(\boldsymbol{\alpha}, \mathbf{R}) = T(\boldsymbol{\alpha}),$$

where $C(\boldsymbol{\alpha}, \mathbf{R})$ and $T(\boldsymbol{\alpha})$ are defined by (12) and (13), respectively. Thus, we have shown that, if the triple $\mathbf{R}^* \in \overline{\mathcal{R}}$, then $\mathbf{R}^* \in \mathcal{S}(\boldsymbol{\alpha})$, where $\boldsymbol{\alpha}$ need not necessarily belong to $\mathcal{A}$.

Now suppose that $\mathbf{R}^* = (R_0^*, R_1^*, R_2^*) \in \bar{\mathcal{R}}$, and $\mathbf{R}^* \in \mathcal{S}(\boldsymbol{\alpha})$ where, say, $\alpha_1 < 0$. From Lemma 16(*a*) (with $a > 0$), $\hat{\mathbf{R}} = (R_0, R_1^* + a, R_2^*) \in \mathcal{R}$, and

$$C(\boldsymbol{\alpha}, \hat{\mathbf{R}}) < C(\boldsymbol{\alpha}, \mathbf{R}^*),$$

which implies $\mathbf{R}^* \notin \mathcal{S}(\boldsymbol{\alpha})$, a contradiction. Thus, α_1 (and similarly α_2) $\geqq 0$. Next, suppose $\mathbf{R}^* \in \bar{\mathcal{R}}$, and $\mathbf{R}^* \in \mathcal{S}(\boldsymbol{\alpha})$ where, say, $\alpha_1 > 1$. Then, from Lemma 16(*c*), $\hat{\mathbf{R}} = (R_0^* + R_1^*, 0, R_2^*) \in \mathcal{R}$, and

$$C(\boldsymbol{\alpha}, \hat{\mathbf{R}}) < C(\boldsymbol{\alpha}, \mathbf{R}^*),$$

again a contradiction. Thus, α_1, and similarly $\alpha_2 \leqq 1$. Finally, suppose that $\mathbf{R}^* \in \bar{\mathcal{R}}$ and $\mathbf{R}^* \in \mathcal{S}(\boldsymbol{\alpha})$, where $\alpha_1 + \alpha_2 < 1$. By Lemma 16(*b*), $\hat{\mathbf{R}} = (0, R_1^* + R_0^*, R_2^* + R_0^*) \in \mathcal{R}$, and

$$C(\boldsymbol{\alpha}, \hat{\mathbf{R}}) \leqq C(\boldsymbol{\alpha}, \mathbf{R}^*),$$

a contradiction. Thus, $\alpha_1 + \alpha_2 \geqq 1$. We conclude that

$$\bar{\mathcal{R}} \subseteq \bigcup_{\boldsymbol{\alpha} \in \mathcal{A}} \mathcal{S}(\boldsymbol{\alpha}), \tag{77}$$

where $\mathcal{A}$ is the set of $\boldsymbol{\alpha} = (\alpha_1, \alpha_2)$ that satisfy $0 \leqq \alpha_1, \alpha_2 \leqq 1$, and $\alpha_1 + \alpha_2 \geqq 1$. This is part (*i*). This completes the proof of Theorem 1.

A.3 Proof of Lemma 5

Let $\{\Delta_k\}_1^n$ be given, and, for $k = 1, 2, \cdots, n$, let $q_{tk}(\hat{x}|x, w)$, $\hat{x} \in \hat{\mathfrak{X}}$, $w \in \mathcal{W}_k$ be a test channel for which

$$\sum_{w \in \mathcal{W}_k} \sum_{x,\hat{x}} d(x, \hat{x}) q_{tk}(\hat{x}|x, w) p_k(x, w) \leqq \Delta_k, \tag{78a}$$

and

$$I(X; \hat{X}|W_k) \leqq R_{X|W_k}(\Delta_k) + \epsilon, \tag{78b}$$

where $\epsilon > 0$ is arbitrary. For $w \in \mathcal{W} = \sum_{k=1}^n \mathcal{W}_k$, $x \in \mathfrak{X}$, $\hat{x} \in \hat{\mathfrak{X}}$, define the test channel

$$q_t^*(\hat{x}|x, w) = q_{tk}(\hat{x}|x, w), \quad \text{for } w \in \mathcal{W}_k, 1 \leqq k \leqq n.$$

Then

$$\sum_{x,y,w} d(x, \hat{x}) q_t(\hat{x}|x, w) p^*(x, w)$$

$$= \sum_{k=1}^n \sum_{w \in \mathcal{W}_k} \sum_{x,y} \frac{1}{n} d(x, \hat{x}) q_{t_k}(\hat{x}|x, w) p_k(x, w) \leqq \frac{1}{n} \sum_{k=1} \Delta_k.$$

Thus, corresponding to the distribution $p^*(x, w) \cdot q_t^*(\hat{x} | x, w)$,

$$I(X, \hat{X} | W) \geqq R_{X|W}\left(\frac{1}{n} \sum_k \Delta_k\right). \tag{79}$$

However, by a straightforward calculation,

$$I(X, \hat{X} | W) = \frac{1}{n} \sum_{k=1}^{n} I(X, \hat{X} | W_k) \leqq \frac{1}{n} \sum_{k=1}^{n} R_{X|W_k}(\Delta k) + \epsilon. \tag{80}$$

The inequality follows from (78b). Combining (79) and (80) and letting $\epsilon \to 0$, we have Lemma 5.

REFERENCES

1. T. Berger, *Rate-Distortion Theory*, Englewood Cliffs, N. J.: Prentice-Hall, 1971.
2. D. Slepian and J. K. Wolf, "Noiseless Coding of Correlated Information Sources," IEEE Trans. on Information Theory, *IT-19*, July 1973, pp. 471–480.
3. A. D. Wyner, "The Common Information of Two Dependent Random Variables," to appear in IEEE Trans. on Information Theory.
4. R. G. Gallager, *Information Theory and Reliable Communication*, New York: John Wiley, 1968.
5. R. M. Gray, "A New Class of Lower Bounds to Information Rates of Stationary Sources via Conditional Rate-Distortion Functions," IEEE Trans. on Information Theory, *IT-19*, July 1973, pp. 480–489.