THE BELL SYSTEM TECHNICAL JOURNAL

DEVOTED TO THE SCIENTIFIC AND ENGINEERING
ASPECTS OF ELECTRICAL COMMUNICATION

Volume 60 | December 1981 | Number 10

Source Coding for Multiple Descriptions II: A Binary Source

By H. S. WITSENHAUSEN and A. D. WYNER

(Manuscript received March 19, 1981)

A uniformly distributed (iid) binary source is encoded into two binary data streams at rates R_1 and R_2, respectively. These sequences are such that by observing either one separately, a decoder can recover a good approximation of the source (at average error rates D_1, D_2, respectively), and by observing both sequences, a decoder can obtain a better approximation of the source (at average error rate D_0). In this paper a "converse" theorem is established on the set of achievable quintuples (R_1, R_2, D_0, D_1, D_2). For the special case $R_1 = R_2 = 1/2$, $D_0 = 0$, and $D_1 = D_2 = D$, our result implies that $D \geq 1/5$.

I. INTRODUCTION

Let $\{X_k\}_{k=1}^{\infty}$ be a sequence of independent drawings of the binary random variable X, where $\Pr\{X = 0\} = \Pr\{X = 1\} = 1/2$. Assume that this sequence appears at a rate of 1 symbol per second as the output of a data source. (Refer to Fig. 1.) An encoder observes this sequence and emits two binary sequences at rates $R_1, R_2 \leq 1$. These sequences are such that by observing either one, a decoder can recover a good approximation to the source output, and by observing both sequences, a decoder can obtain a better approximation to the source output. Letting D_1, D_2, and D_0 be the error rates which result when the streams at rate R_1, rate R_2, and both streams are used by a decoder, respectively, our problem is to determine (in the usual Shannon sense) the set of achievable quintuples (R_1, R_2, D_0, D_1, D_2). Our main result is a "converse" theorem which gives a necessary condition on the achiev-

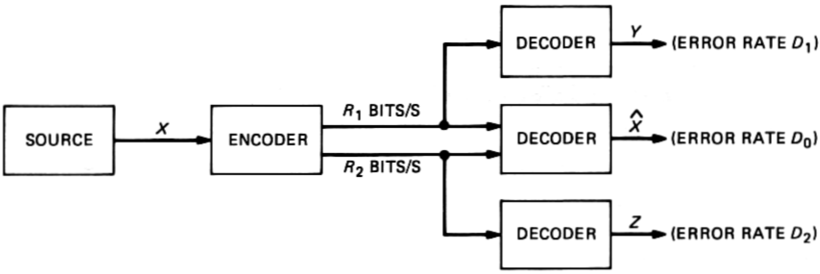

Fig. 1—Communication system.

able quintuples. This paper extends a previous one on the same subject.[1] This paper, however, is self-contained.

This problem is an idealization of the situation in which it is desired to

(*i*) send information over two separate channels, as in a packet communication network, and

(*ii*) recover as much of the original information as possible, should one of the channels break down.

To fix ideas, let us say that $R_1 = R_2 = 1/2$, $D_0 = 0$, and $D_1 = D_2 = D$. Thus, the source sequence at rate 1 is to be encoded into two sequences of rate 1/2 each, such that the original sequence can be recovered from these two encoded sequences with approximately zero error rate (i.e. $D_0 = 0$). Our question then becomes: How well can we reconstruct the source sequence from one of the encoded streams—that is, what is the minimum D? A simple-minded approach would be to let the encoded streams consist of alternate source symbols, which will allow $D_0 = 0$. In this case, $D = 1/4$, since by observing every other source symbol a decoder will make an error half the time on the missing symbol. Is it possible to do better? El Gamal and Cover[2] have looked at this problem and have a theorem which can be used to show that we can make $D = (\sqrt{2} - 1)/2 \approx 0.207$. In a previous paper[1] it was shown that (with $R_1 = R_2 = 1/2$, $D_0 = 0$) $D \geq 1/6$. The new result given here specializes to $D \geq 1/5 = 0.200$. The exact determination of the best D remains an open problem.*

II. FORMAL STATEMENT OF PROBLEM AND RESULTS

Let $\mathbf{B} = \{0, 1\}$, and let $d_H(\mathbf{x}, \mathbf{y})$, $\mathbf{x}, \mathbf{y} \in \mathbf{B}^N$, be the Hamming distance between the binary N-vectors $\mathbf{x}$, $\mathbf{y}$; i.e., $d_H(\mathbf{x}, \mathbf{y})$ is the number of positions in which $\mathbf{x}$ and $\mathbf{y}$ do not agree. A code with parameters

* In Ref. 3, Witsenhausen proved a closely related result which encourages the conjecture that $D = 0.207$ is, in fact, the best possible.

$(N, M_1, M_2, D_0, D_1, D_2)$ is a quintuple of mappings $(f_1, f_2, g_0, g_1, g_2)$ where,

$$f_\alpha: \mathbf{B}^N \to \{1, \cdots, M_\alpha\}, \alpha = 1, 2 \tag{1a}$$

$$g_\alpha: \{1, 2, \cdots, M_\alpha\} \to \mathbf{B}^N, \alpha = 1, 2 \tag{1b}$$

$$g_0: \{1, 2, \cdots, M_1\} \times \{1, 2, \cdots, M_2\} \to \mathbf{B}^N. \tag{1c}$$

The source output is a random vector $\mathbf{X}$ uniformly distributed on $\mathbf{B}^N$. Define

$$\mathbf{Y} = g_1 \circ f_1(\mathbf{X}), \tag{2a}$$

$$\mathbf{Z} = g_2 \circ f_2(\mathbf{X}), \tag{2b}$$

and

$$\hat{\mathbf{X}} = g_0[f_1(\mathbf{X}), f_2(\mathbf{X})]. \tag{2c}$$

Then the average error rates are

$$D_1 = \frac{1}{N} E d_H(\mathbf{X}, \mathbf{Y}), \tag{3a}$$

$$D_2 = \frac{1}{N} E d_H(\mathbf{X}, \mathbf{Z}), \tag{3b}$$

$$D_0 = \frac{1}{N} E d_H(\mathbf{X}, \hat{\mathbf{X}}). \tag{3c}$$

We say that a quintuple $(R_1, R_2, d_0, d_1, d_2)$ is *achievable* if, for arbitrary $\epsilon > 0$, there exists, for N sufficiently large, a code with parameters $(N, M_1, M_2, D_0, D_1, D_2)$, where $M_\alpha \le 2^{(R_\alpha+\epsilon)N}$, $\alpha = 1, 2$, and $D_\alpha \le d_\alpha + \epsilon$, $\alpha = 0, 1, 2$. The relationship of this formalism to the system of Fig. 1 should be clear. Our problem is the determination of the set of achievable quintuples, and our main result is a converse theorem.

Before stating our result, let us take a moment to state a positive theorem by El Gamal and Cover[2] as it specializes to our problem.

Theorem 1: *The quintuple* $(R_1, R_2, d_0, d_1, d_2)$ *is achievable if there exists a quadruple of random variables* $X, \hat{X}, Y, Z$, *which take values in* $\mathbf{B}$, *such that* $\Pr\{X = 0\} = \Pr\{X = 1\} = 1/2$, *and*

$$E d_H(X, \hat{X}) \le d_0, \tag{4a}$$

$$E d_H(X, Y) \le d_1, \tag{4b}$$

$$E d_H(X, Z) \le d_2, \tag{4c}$$

and

$$R_1 \ge I(X; Y), \tag{5a}$$

$$R_2 \geq I(X; Z), \tag{5b}$$

$$R_1 + R_2 \geq I(X; \hat{X}, Y, Z) + I(Y; Z), \tag{5c}$$

where $I(\cdot;\cdot)$ is the usual Shannon information.

For the special case of $R_1 = R_2 = 1/2$, $d_0 = 0$, it can be shown that $d_1 = d_2 = (\sqrt{2} - 1)/2 \approx 0.207$ is achievable.

We now state our converse result.

Theorem 2: If $(R_1, R_2, d_0, d_1, d_2)$ is achievable, then in all cases

$$R_1 + R_2 \geq 1 - h(d_0), \tag{6a}$$

furthermore, if $2d_1 + d_2 \leq 1$, then

$$R_1 + R_2 \geq 2 - h(d_0) - h\left(2d_1 + d_2 - \frac{2d_1^2}{1 - d_2}\right), \tag{6b}$$

and if $d_1 + 2d_2 \leq 1$, then

$$R_1 + R_2 \geq 2 - h(d_0) - h\left(d_1 + 2d_2 - \frac{2d_2^2}{1 - d_1}\right), \tag{6c}$$

where

$$h(\lambda) = \begin{cases} 0, & \lambda = 0, \\ -\lambda \log_2 \lambda - (1 - \lambda)\log_2(1 - \lambda), & 0 < \lambda \leq 1/2. \\ 1, & \lambda \geq 1/2. \end{cases}$$

All logarithms in this paper are taken to the base 2. As (6a) is obvious, and (6c) follows from (6b) by symmetry, we need only prove (6b).

In the special case of $R_1 = R_2 = 1/2$, $d_0 = 0$, and $d_1 = d_2 = d$, inequality (6b) implies that

$$h\left(3d - \frac{2d^2}{1 - d}\right) \geq 1,$$

or

$$3d - \frac{2d^2}{(1 - d)} = \frac{3d(1 - d) - 2d^2}{(1 - d)} \geq \frac{1}{2},$$

which implies that $d \geq 1/5 = 0.200$.

III. PROOF OF THEOREM 2

We start from the standard identity

$$I(U_1; U_2 U_3) = I(U_1; U_2) + I(U_1; U_3 | U_2), \tag{7}$$

for arbitrary random variables U_1, U_2, U_3. We say that U_1, U_2, U_3 is a "Markov chain" if U_1, U_3 are conditionally independent given U_2; i.e., U_3 depends on U_1, U_2 only through U_2. If U_1, U_2, U_3 is a Markov chain then $I(U_1; U_3 | U_2) = 0$, and from (7)

$$I(U_1; U_3) \le I(U_1; U_2\, U_3) = I(U_1; U_2). \tag{8}$$

Note that the hypothesis for (8) holds when U_3 is a function of U_2. A sequence $\{U_n\}$ is a Markov chain if, for all n,

$$(\cdots U_{n-2}, U_{n-1}), U_n, (U_{n+1}, U_{n+2}, \cdots)$$

is a Markov chain.

Let us now suppose that we are given a code $(f_1, f_2, g_0, g_1, g_2)$ with parameters $(N, M_1, M_2, D_0, D_1, D_2)$. We can write

$$\log M_1 + \log M_2 \ge H(f_1(\mathbf{X})) + H(f_2(\mathbf{X})) \tag{9}$$

$$= I(f_1(\mathbf{X}); f_2(\mathbf{X})) + H(f_1(\mathbf{X})\, f_2(\mathbf{X}))$$

$$= I(f_1(\mathbf{X}); f_2(\mathbf{X})) + I(\mathbf{X}; f_1(\mathbf{X})\, f_2(\mathbf{X})) \tag{10}$$

$$\ge I(f_1(\mathbf{X}); f_2(\mathbf{X})) + I(\mathbf{X}; \hat{\mathbf{X}}, \mathbf{Y}, \mathbf{Z}), \tag{11}$$

where (9) follows from the fact that $f_i(\mathbf{X})$ takes its values in a set of cardinality M_i, (10) holds because the pair $f_1(\mathbf{X})\, f_2(\mathbf{X})$ is determined by $\mathbf{X}$ and (11) holds because $\hat{\mathbf{X}}$, $\mathbf{Y}$, and $\mathbf{Z}$ depend on $\mathbf{X}$ only through $f_1(\mathbf{X})$ $f_2(\mathbf{X})$ so that (8) applies.

Now (11) is getting close to (5c) in the direct theorem. In fact, using (8) twice, we can underbound $I[f_1(\mathbf{X}); f_2(\mathbf{X})]$ by $I(\mathbf{Y}; \mathbf{Z})$. Now the components of the source vector $\mathbf{X}$ are independent, and if the components of either $\mathbf{Y}$ or $\mathbf{Z}$ were also independent, we could make use of standard techniques to establish the necessity of (5c). But alas, we cannot assume that either the $\{Y_n\}$ nor the $\{Z_n\}$ are independent, so that another tactic is required. The key idea is the definition of another random vector $\mathbf{V} = (V_1, \cdots, V_n)$ the components of which are in fact independent.

For $1 \le k \le M_1$, define the set

$$A_k = \{\mathbf{x}: f_1(\mathbf{x}) = k\} = f_1^{-1}(k). \tag{12}$$

Let the cardinality of A_k be μ_k. Let the random vector $\mathbf{V}$ be defined by its conditional distribution given $\mathbf{X}$:

$$\Pr\{\mathbf{V} = \mathbf{v} | \mathbf{X} = \mathbf{x}\} = \begin{cases} [\mu_{f_1(\mathbf{x})}]^{-1}, & \mathbf{v} \in A_{f_1(\mathbf{x})}, \\ 0, & \text{otherwise}. \end{cases} \tag{13}$$

Thus, given $\mathbf{X} \in A_k$, $\mathbf{V}$ is uniformly distributed on A_k. It follows that the unconditional distribution for $\mathbf{V}$ is

$$\Pr\{\mathbf{V} = \mathbf{v}\} = 2^{-n}, \qquad \mathbf{v} \in \mathbf{B}^n,$$

and the components of $\mathbf{V}$ are independent, as desired.* Furthermore, $\mathbf{Z}$, $f_2(\mathbf{X})$, $\mathbf{X}$, $f_1(\mathbf{X})$, $\mathbf{V}$ is a Markov chain, so that, using (8),

* In effect, $\mathbf{V}$ is obtained from $f_1(X)$ by a channel with transition probabilities $\Pr\{\mathbf{X} = \mathbf{x} | f_1(\mathbf{X}) = k\}$ so that the distribution of $\mathbf{V}$ is the same as that of $\mathbf{X}$, hence, iid. This is valid for any distribution of X.

$$I(f_1(\mathbf{X}); f_2(\mathbf{X})) = I(\mathbf{V}, f_1(\mathbf{X}); f_2(\mathbf{X}), \mathbf{Z}) \geq I(\mathbf{V}; \mathbf{Z}). \tag{14}$$

Combining (11) and (14), we obtain

$$\begin{aligned}
\frac{1}{N}\log M_1 + \frac{1}{N}\log M_2 &\geq \frac{1}{N} I(\mathbf{V}; \mathbf{Z}) + \frac{1}{N} I(\mathbf{X}; \hat{\mathbf{X}}, \mathbf{Y}, \mathbf{Z}) \\
&\geq \frac{1}{N} I(\mathbf{V}; \mathbf{Z}) + \frac{1}{N} I(\mathbf{X}; \hat{\mathbf{X}}) \\
&\overset{(1)}{\geq} 2 - h\left[\frac{Ed_H(\mathbf{V}; \mathbf{Z})}{N}\right] - h\left[\frac{Ed_H(\mathbf{X}, \hat{\mathbf{X}})}{N}\right] \\
&= 2 - h(\Delta) - h(D_0), \qquad (15a)
\end{aligned}$$

where

$$\Delta = \frac{Ed_H(\mathbf{V}; \mathbf{Z})}{N}. \tag{15b}$$

Step (1) follows from the "rate-distortion bound" which states that if $\mathbf{U}$ is a random vector uniformly distributed in $\mathbf{B}^N$ (as are $\mathbf{V}$ and $\mathbf{X}$), and $\hat{\mathbf{U}}$ is an arbitrary binary random vector, then $I(\mathbf{U}; \hat{\mathbf{U}}) \geq 1 - h\left[\frac{1}{N} Ed_H(\mathbf{U}, \hat{\mathbf{U}})\right]$. (See Ref. 4.)

We will now obtain an upper bound on Δ in terms of D_1 and D_2. As a "warm up," let us observe that from the triangle inequality,

$$\Delta = \frac{1}{N} Ed_H(\mathbf{V}, \mathbf{Z}) \leq \frac{1}{N}[Ed_H(\mathbf{V}, \mathbf{Y}) + Ed_H(\mathbf{Y}, \mathbf{X}) + Ed_H(\mathbf{X}, \mathbf{Z})].$$

Now

$$Ed_H(\mathbf{Y}, \mathbf{X}) = D_1 N, \qquad Ed_H(\mathbf{Z}, \mathbf{X}) = D_2 N, \tag{16}$$

Furthermore,

$$Ed_H(\mathbf{V}, \mathbf{Y}) = \sum_{\mathbf{v}} E[d_H(\mathbf{v}, \mathbf{Y}) \,|\, \mathbf{V} = \mathbf{v}] \Pr\{\mathbf{V} = \mathbf{v}\}.$$

Now suppose that we are given $\mathbf{V} = \mathbf{v} \in A_k$. Then, $\mathbf{Y} = g_1(k)$. Since $\Pr\{\mathbf{V} = \mathbf{v}\} = 2^{-N}$,

$$\begin{aligned}
Ed_H(\mathbf{V}, \mathbf{Y}) &= \sum_{k=1}^{M_1} \sum_{\mathbf{v} \in A_k} 2^{-N} d_H[\mathbf{v}, g_1(k)] \\
&= \sum_{k=1}^{M_1} \sum_{\mathbf{x} \in A_k} \Pr\{\mathbf{X} = \mathbf{x}\} d_H[\mathbf{x}, g_1(k)] = ND_1. \qquad (17)
\end{aligned}$$

Thus,

$$\Delta \leq 2D_1 + D_2. \tag{18}$$

Substitution of (18) into (15a) yields that for achievable $(R_1, R_2, d_0, d_1, d_2)$

$$R_1 + R_2 \geq 2 - h(2d_1 + d_2) - h(d_0), \tag{19}$$

which is the result reported in Ref. 1.

We will now establish a tighter bound on Δ, namely, for $D_2 + 2D_1 \leq 1$,

$$\Delta = \frac{1}{N} Ed_H(\mathbf{V}, \mathbf{Z}) \leq D_2 + 2D_1 - \frac{2D_1^2}{(1 - D_2)}, \tag{20}$$

so that (15) yields that for achievable $(R_1, R_2, d_0, d_1, d_2)$,

$$R_1 + R_2 \geq 2 - h\left(d_2 + 2d_1 - \frac{2d_1^2}{1 - d_2}\right) - h(d_0), \tag{21}$$

which is (6b), the inequality required for Theorem 2.

Upper bound on Δ: We establish inequality (20) as follows. Let k, $1 \leq k \leq M_1$ be fixed. Let A_k be as defined in (12), and let its cardinality $\mu_k = \mu$. Let the members of A_k be the N-vectors $\{\mathbf{x}_m\}_{m=1}^{\mu}$. Let $\mathbf{y} = g_1(k)$. Thus, when $\mathbf{X} = \mathbf{x} \in A_k$, then $\mathbf{Y} = \mathbf{y}$. Now, form a $\mu \times N$ array, $\mathbf{A}$, with mth row

$$\mathbf{a}_m = (a_{m_1}, a_{m_2}, \cdots, a_{mN}) = \mathbf{x}_m \oplus \mathbf{y}, \tag{22}$$

where $\oplus$ denotes modulo 2 vector addition. Thus, $a_{mn} = 1$, when the nth coordinates of $\mathbf{x}_m$ and $\mathbf{y}$ are different, and $a_{mn} = 0$, otherwise. Note that

$$\begin{aligned} &\frac{1}{N} E[d_H(\mathbf{X}, \mathbf{Y}) \,|\, f_1(\mathbf{X}) = k] \\ &= \frac{1}{N}\frac{1}{\mu} \sum_{m=1}^{\mu} d_H(\mathbf{x}_m, \mathbf{y}) = \frac{1}{N}\frac{1}{\mu} \sum_{m=1}^{\mu} \sum_{n=1}^{N} a_{mn} \\ &= \frac{1}{N} \sum_{n=1}^{N} s_n, \end{aligned} \tag{23a}$$

where for $1 \leq n \leq N$,

$$s_n = \frac{1}{\mu} \sum_{m=1}^{\mu} a_{mn} \tag{23b}$$

is the fraction of 1's in column n of $\mathbf{A}$.

Next, for $1 \leq m \leq \mu$, let $\mathbf{z}_m = g_2 \circ f_2(\mathbf{x}_m)$ be the value of $\mathbf{Z}$ which results when $\mathbf{X} = \mathbf{x}_m$. Let $\mathbf{B}$ be the $\mu \times N$ array with mth row

$$\mathbf{b}_m = (b_{m1}, b_{m2}, \cdots, b_{mN}) = \mathbf{z}_m \oplus \mathbf{y}. \tag{24}$$

Then,

$$\frac{1}{N} E[d_H(\mathbf{X}, \mathbf{Z}) \,|\, f_1(\mathbf{X}) = k]$$

$$= \frac{1}{N} \frac{1}{\mu} \sum_{m=1}^{\mu} d_H(\mathbf{x}_m, \mathbf{z}_m)$$

$$= \frac{1}{N\mu} \sum_{m} d_H(\mathbf{a}_m, \mathbf{b}_m)$$

$$= \frac{1}{N} \sum_{n=1}^{N} \left\{ \frac{1}{\mu} \sum_{m=1}^{\mu} [a_{mn}(1 - b_{mn}) + b_{mn}(1 - a_{mn})] \right\}, \tag{25}$$

where the last equality follows from the fact that for $a, b \epsilon \{0,1\}$, $a(1 - b) + b(1 - a) = 0$ or 1 according as $a = b$ or $a \neq b$. Let τ_n be the expression in braces in (25). Then,

$$\tau_n = \frac{1}{\mu} \sum_{m=1}^{\mu} [a_{mn} + b_{mn} - 2a_{mn}b_{mn}]$$

$$= \begin{cases} \frac{1}{\mu} \sum_{m} [a_{mn} - b_{mn} + 2b_{mn}(1 - a_{mn})] \\ \qquad \geq \frac{1}{\mu} \sum_{m} (a_{mn} - b_{mn}) = s_n - t_n \\ \frac{1}{\mu} \sum_{m} [b_{mn} - a_{mn} + 2a_{mn}(1 - b_{mn})] \\ \qquad \geq \frac{1}{\mu} \sum_{m} (b_{mn} - a_{mn}) = t_n - s_n, \end{cases} \tag{26a}$$

where

$$t_n = \frac{1}{\mu} \sum_{m=1}^{\mu} b_{mn}, \qquad 1 \leq n \leq N, \tag{26b}$$

and s_n is given by (23b). We conclude that $\tau_n \geq |t_n - s_n|$, so that (25) yields

$$\frac{1}{N} E[d_H(\mathbf{X}, \mathbf{Z}) \,|\, f_1(\mathbf{X}) = k] \geq \frac{1}{N} \sum_{n=1}^{N} |t_n - s_n|. \tag{27}$$

Finally, consider

$$\frac{1}{N} E[d_H(\mathbf{V}, \mathbf{Z}) \,|\, f_1(\mathbf{X}) = k]$$

$$= \frac{1}{N} \sum_{m=1}^{\mu} \frac{1}{\mu} E[d_H(\mathbf{V}, \mathbf{Z}) \,|\, \mathbf{X} = \mathbf{x}_m]$$

$$= \frac{1}{N\mu} \sum_{m=1}^{\mu} E[d_H(\mathbf{V}, \mathbf{z}_m) \,|\, \mathbf{X} = \mathbf{x}_m]$$

$$= \frac{1}{N\mu} \sum_{m=1}^{\mu} \sum_{\mathbf{v}} d_H(\mathbf{v}, \mathbf{z}_m) \Pr\{\mathbf{V} = \mathbf{v} \,|\, \mathbf{X} = \mathbf{x}_m\}. \tag{28}$$

Now, from (13) which defines $\mathbf{V}$,

$$\Pr\{\mathbf{V} = \mathbf{v} \mid \mathbf{X} = \mathbf{x}_m\} = \begin{cases} \mu^{-1}, & \mathbf{v} \in A_k, \\ 0, & \text{otherwise.} \end{cases}$$

Thus, (28) yields,

$$\begin{aligned} \frac{1}{N} E[d_H(\mathbf{V}, \mathbf{Z}) \mid f_1(\mathbf{X}) = k] \\ &= \frac{1}{N\mu^2} \sum_{m=1}^{\mu} \sum_{m'=1}^{\mu} d_H(\mathbf{x}_{m'}, \mathbf{z}_m) \\ &= \frac{1}{N\mu^2} \sum_{m,m'} d_H(\mathbf{a}_{m'}, \mathbf{b}_m) \\ &= \frac{1}{N\mu^2} \sum_{m,m'} \sum_{n} [a_{m'n}(1 - b_{mn}) + (1 - a_{m'n})b_{mn}] \\ &= \frac{1}{N} \sum_{n=1}^{N} [s_n(1 - t_n) + t_n(1 - s_n)]. \end{aligned} \tag{29}$$

Now make the dependence of s_n and t_n on k explicit by writing s_{nk} and t_{nk}, respectively. Then, on averaging over k, (23b) becomes

$$\begin{aligned} D_1 = \frac{1}{N} E\{d_H(\mathbf{X}, \mathbf{Y})\} &= \sum_{k=1}^{M_1} \Pr\{f_1(\mathbf{X}) = k\} \frac{1}{N} \sum_{n=1}^{N} s_{nk} \\ &= \sum_{k=1}^{M_1} \sum_{n=1}^{N} P(n, k) s_{nk}, \end{aligned} \tag{30a}$$

where $P(n, k) = \Pr\{f_1(\mathbf{X}) = k\} \times \frac{1}{N}$. Similarly, (27) becomes

$$D_2 = \frac{1}{N} E\{d_H(\mathbf{X}, \mathbf{Z})\} \geq \sum_{k=1}^{M_1} \sum_{n=1}^{N} P(n, k) \mid t_{nk} - s_{nk} \mid. \tag{30b}$$

Finally, (29) becomes

$$\begin{aligned} \Delta &= \frac{1}{N} E\{d_H(\mathbf{V}, \mathbf{Z})\} \\ &= \sum_{k=1}^{M_1} \sum_{n=1}^{N} P(n, k)[s_{nk}(1 - t_{nk}) + t_{nk}(1 - s_{nk})]. \end{aligned} \tag{30c}$$

We now apply the following inequality, the proof of which is given in Section IV:

Let S, T be random variables such that $0 \leq S, T \leq 1$, and $E\{S\} \leq D_1$, $E\{\mid T - S \mid\} \leq D_2$, with $2D_1 + D_2 \leq 1$. Then,

$$E\{S(1 - T) + T(1 - S)\} \leq D_2 + 2D_1 - \frac{2D_1^2}{1 - D_2}. \tag{31}$$

Let S, T be the random variables which take the value s_{nk}, t_{nk}, respectively, with probability $P(n, k)$, then (30) and (31) imply that, for $2D_1 + D_2 \le 1$,

$$\Delta \le D_2 + 2D_1 - \frac{2D_1^2}{1 - D_2}, \tag{32}$$

which, when substituted in (15) gives (6b), proving Theorem 2.

IV. PROOF OF THE INEQUALITY

Define $Q(D_1, D_2)$ as the supremum of $E\{S(1 - T) + T(1 - S)\}$ over all distributions of (S, T) on the unit square $[0, 1]^2$ for which $E\{S\} \le D_1$, $E\{|T - S|\} \le D_2$.

Theorem 3: (a) *For* $2D_1 + D_2 \le 1$ *one has*

$$Q(D_1, D_2) = 2D_1 + D_2 - \frac{2D_1^2}{1 - D_2},$$

with $Q(0, 1) = 1$. (b) *For* $2D_1 + D_2 \ge 1$ *one has*

$$Q(D_1, D_2) \ge \frac{1}{2}.$$

To establish this, introduce for S, T, x, y in $[0, 1]$, $y \ne 1$, the function*

$$F(S, T, x, y) = S(1 - T) + T(1 - S) + \left(\frac{4x}{1 - y} - 2\right)(S - x) + \left[\frac{2x^2}{(1 - y)^2} - 1\right](|S - T| - y). \tag{33}$$

Lemma: *For* $2x + y \le 1$, $y < 1$ *the maximum of* $F(S, T, x, y)$ *over all* (S, T) *in* $[0, 1]^2$ *is* $2x + y - 2x^2/(1 - y)$.

Proof: For fixed S, the maximum of F over T must, by piecewise linearity, be at either $T = 0$, S, or 1.

(i) If $T = 0$, then

$$F = S + \left(\frac{4x}{1 - y} - 2\right)(S - x) + \left[\frac{2x^2}{(1 - y)^2} - 1\right](S - y),$$

and this is maximized over S at either $S = 0$ or $S = 1$.

(a) For $S = 0$

$$F = 2x - \frac{4x^2}{1 - y} + y - \frac{2x^2 y}{(1 - y)^2}$$
$$= 2x + y - \frac{2x^2}{1 - y} - \frac{2x^2}{1 - y} - \frac{2x^2 y}{(1 - y)^2} \le 2x + y - \frac{2x^2}{1 - y}.$$

* This choice of F comes from the duality theory of convex hull formation, as described, e.g., in Section VA of Ref. 5.

(b) For $S = 1$

$$F = 1 - 2 + 2x + \frac{4x(1-x)}{1-y} + \frac{2x^2}{1-y} - 1 + y$$

$$= 2x + y - \frac{2x^2}{1-y} - 2 + \frac{4x}{1-y} \leq 2x + y - \frac{2x^2}{1-y}, \text{ using } 2x + y \leq 1.$$

(*ii*) If $T = S$, then

$$F = 2S - 2S^2 + \left(\frac{4x}{1-y} - 2\right)(S - x) + y - \frac{2x^2 y}{(1-y)^2}.$$

The maximum over S of this quadratic is at $S = x/(1-y)$ which is in $[0, 1]$ if $x \leq 1 - y$ and this holds as $x \leq (1-y)/2 \leq 1 - y$. Hence, the maximum is

$$\frac{2x}{1-y} - \frac{2x^2}{(1-y)^2} + \left(\frac{4x}{1-y} - 2\right)\left(\frac{x}{1-y} - x\right) + y - \frac{2x^2 y}{(1-y)^2} = 2x + y - \frac{2x^2}{1-y}.$$

(*iii*) If $T = 1$, then

$$F = 1 - S + \left(\frac{4x}{1-y} - 2\right)(S - x) + \left[\frac{2x^2}{(1-y)^2} - 1\right](1 - S - y),$$

which is maximized by taking S either 0 or 1.

(a) For $S = 0$

$$F = 2x + y - \frac{4x^2}{1-y} + \frac{2x^2}{(1-y)^2} - \frac{2x^2 y}{(1-y)^2}$$

$$= 2x + y - \frac{2x^2}{1-y}$$

(b) For $S = 1$,

$$F = 2x + y + \frac{4x}{1-y} - \frac{4x^2}{1-y} - 2 - \frac{2x^2 y}{(1-y)^2}$$

$$= 2x + y - \frac{2x^2}{1-y} - \frac{2x^2}{1-y} + \frac{4x}{1-y} - 2 - \frac{2x^2 y}{(1-y)^2}$$

$$\leq 2x + y - \frac{2x^2}{1-y},$$

since $4x/(1-y) \leq 2$.

Thus, the maximum is as stated and is attained for $T = S = x/(1-y)$ and for $T = 1$, $S = 0$. This completes the proof of the lemma.

Turning to the proof of Theorem 3, consider any distribution of (S, T) on the unit square for which

$$E\{S\} = x, \qquad E\{|T - S|\} = y, \qquad 2x + y \le 1. \tag{34}$$

If $y = 1$, then $x = 0$ from which it follows that $S = 0$, $T = 1$ almost surely, giving $E\{S(1 - T) + T(1 - S)\} = 1$.

If $y < 1$, one has, by the lemma,

$$E\{F(S, T, x, y)\} = E\{S(1 - T) + T(1 - S)\} \le 2x + y - \frac{2x^2}{1 - y}.$$

If one chooses the distribution $T = 1$, $S = 0$ with probability y and $T = S = x/(1 - y)$ with probability $1 - y$, equality is attained. This determines the maximum of $E\{S(1 - T) + T(1 - S)\}$ subject to (34). As $2x + y - 2x^2/(1 - y)$ is monotone increasing (for $2x + y \le 1$) in both x and y, the maximum is unchanged if one allows all x, y with $0 \le x \le D_1$, $0 \le y \le D_2$. This establishes part (a) of Theorem 3.

For part (b), it suffices to observe that Q is monotone nondecreasing in both arguments by its definition and that on the boundary, where $2D_1 + D_2 = 1$, one has $Q(D_1, D_2) = (1 + D_2)/2$. This establishes part (b). (It could easily be shown that $Q(D_1, D_2) = (1 + D_2)/2$ for all (D_1, D_2) in the unit square satisfying $2D_1 + D_2 \ge 1$.)

REFERENCES

1. J. K. Wolf, A. D. Wyner, and J. Ziv, "Source Coding for Multiple Descriptions," B.S.T.J., *59,* No. 8 (October 1980), pp. 1417–26.
2. A. A. El Gamal and T. M. Cover, "Information Theory of Multiple Descriptions," Technical Report No. 43, Department of Statistics, Stanford University, 1980.
3. H. S. Witsenhausen, "On Source Networks with Minimal Breakdown Degradation," B.S.T.J., *59,* No. 6 (July–August 1980), pp. 1038–87.
4. R. G. Gallager, *Information Theory and Reliable Communication*, New York: John Wiley, 1968, Theorem 4.3.2, p. 79.
5. H. S. Witsenhausen, "Some Aspects of Convexity Useful in Information Theory," IEEE Trans. Inform. Theory, *IT-26* (May 1980), pp. 265–71.