THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 59, No. 8, October 1980
*Printed in U.S.A.*

# Source Coding for Multiple Descriptions

By J. K. WOLF, A. D. WYNER, and J. ZIV

(Manuscript received March 31, 1980)

*This paper discusses an idealization of the situation in which it is required to send information over two separate channels, as in a packet communication network, and it is desired to recover as much as possible of the original information should one of the channels break down. Let* $\{X_k\}_{k=1}^{\infty}$ *be a sequence of independent copies of the binary random variable X, where* $Pr\{X = 0\} = \Pr\{X = 1\} = 1/2$. *Assume that this sequence appears at a rate of one symbol per second as the output of a data source. 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* $D_1, D_2, D_0$ *be the error rates that 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 that gives a necessary condition on the achievable quintuples.*

## I. INTRODUCTION

Let $\{X_k\}_{k=1}^{\infty}$ be a sequence of independent copies of the binary random variable $X$, where $\Pr\{X = 0\} = \Pr\{X = 1\} = 1/2$. Assume that this sequence appears at a rate of one symbol per second as the output of the data source in Fig. 1. The encoder in the figure 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 $D_1, D_2, D_0$ be the error rates which result when the streams at rate $R_1$, rate $R_2$, and both streams are used by a decoder,

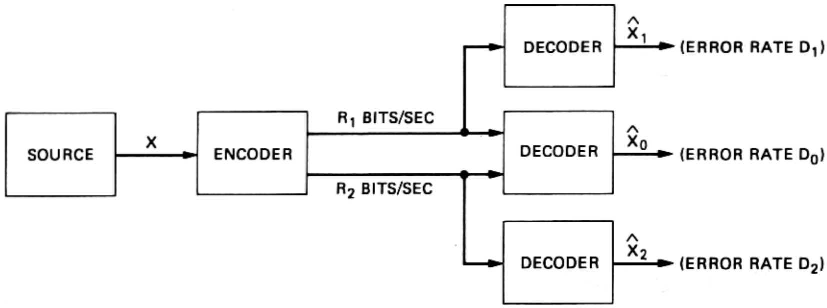


Fig. 1—Source encoder-decoder.

respectively, our problem is to determine (in the usual Shannon sense) the set of achievable $(R_1, R_2, D_0, D_1, D_2)$.

To fix ideas, let us say that $R_1 = R_2 = \frac{1}{2}$, $D_0 = 0$ *and* $D_1 = D_2$. Thus the source sequence at rate 1 is to be encoded into two sequences of rate ½ 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 $D_1 = D_2$?) 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, $D_1 = D_2 = \frac{1}{4}$, 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[1] have looked at this problem and have a theorem that can be used to show that we can make $D_1 = D_2 = (\sqrt{2} - 1)/2 \approx 0.207$. In the present paper, we prove a theorem from which it follows that (with $R_1 = R_2 = \frac{1}{2}$, $D_0 = 0$), $D_1 = D_2 \geq \frac{1}{6}$. An exact determination of the best $D_1 = D_2$ is at present an open problem.*

Let us remark that this problem can be generalized in an obvious way to an arbitrary source $\{X_k\}$, and arbitrary distortion measure. An especially interesting case is where the $\{X_k\}$ are Gaussian and the distortion is the squared-error criterion. For this case, L. H. Ozarow[3] has obtained the complete solution.

## II. FORMAL STATEMENT OF PROBLEM AND RESULTS

Let $\mathscr{B} = \{0, 1\}$, and let $d_H(\mathbf{x}, \mathbf{y})$, $\mathbf{x}, \mathbf{y} \in \mathscr{B}^N$, 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 $(N, M_1, M_2, D_0, D_1, D_2)$ is a quintuple of mappings $(f_1, f_2, g_0, g_1, g_2)$ where

* In Ref. 2, Witsenhausen proved a closely related result which encourages the conjecture that $D_1 = D_2 = 0.207$ is in fact the best possible.

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

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

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

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

$$\hat{\mathbf{X}}_\alpha = g_\alpha \bullet f_\alpha(\mathbf{X}), \qquad \alpha = 1, 2, \tag{2a}$$

and

$$\hat{\mathbf{X}}_0 = g_0[f_1(\mathbf{X}), f_2(\mathbf{X})]. \tag{2b}$$

Then the average *error-rates* are

$$D_\alpha = \frac{1}{N} E d_H(\mathbf{X}, \hat{\mathbf{X}}_\alpha), \qquad \alpha = 1, 2 \tag{3a}$$

$$D_0 = \frac{1}{N} E d_H(\mathbf{X}, \hat{\mathbf{X}}_0). \tag{3b}$$

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 $M_\alpha \leq 2^{(R_\alpha + \epsilon)N}$, $\alpha = 1, 2$, and $D_\alpha \leq d_\alpha + \epsilon$, $\alpha = 0, 1, 2$. 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 the following "converse" theorem.

*Theorem*: *If* $(R_1, R_2, d_0, d_1, d_2)$ *is achievable, then*

$$R_1 + R_2 \geq \begin{cases} 2 - h(d_0) - h(d_1 + 2d_2) \\ 2 - h(d_0) - h(2d_1 + d_2), \end{cases} \tag{4}$$

*where*

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

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

*Discussion*: When $R_1 = R_2 = \tfrac{1}{2}$, $d_0 = 0$, and $d_1 = d_2$, the rate-distortion bound implies that $1 - h(d_1) \leq R_1 = \tfrac{1}{2}$, or $d_1 \geq 0.11$. Our result (4), however, yields $h(3d_1) \geq 1$, or $d_1 \geq \tfrac{1}{6}$.

## III. INFORMAL OUTLINE OF THE PROOF OF THE THEOREM

In this section, we give an informal "$\epsilon$-free" discussion which contains the main ideas behind the proof of our theorem.

Let $(f_1, f_2, g_0, g_1, g_2)$ be a code with parameters $(N, M_1, M_2, D_0, D_1, D_2)$, where $M_\alpha \approx 2^{R_\alpha N} (\alpha = 1, 2)$ and $D_\alpha \approx d_\alpha$ $(\alpha = 0, 1, 2)$. Let

$$A_i = \{\mathbf{x} \in \mathscr{B}^N\colon \ f_1(\mathbf{x}) = i\}, \qquad 1 \le i \le M_1, \tag{5a}$$

$$B_j = \{\mathbf{x} \in \mathscr{B}^N\colon \ f_2(\mathbf{x}) = j\}, \qquad 1 \le j \le M_2, \tag{5b}$$

$$C_{ij} = A_i \cap B_j. \tag{5c}$$

For this informal discussion, let us assume that for all $\mathbf{x} \in \mathscr{B}^N$,

$$d_\alpha \approx \frac{1}{N} d_H(\mathbf{x}, g_\alpha \circ f_\alpha(\mathbf{x})), \qquad \alpha = 1, 2 \tag{6a}$$

$$d_0 \approx \frac{1}{N} d_H(\mathbf{x}, g_0(f_1(x), f_2(\mathbf{x}))). \tag{6b}$$

Also, since $\sum_{i=1}^{M_1} A_i = \mathscr{B}^N$, the average cardinality of the $A_i$ is $2^N/M_1 \approx 2^{N(1-R_1)}$. For our informal discussion, we will assume that all the $A_i$ have approximately equal cardinality, so that

$$|A_i| \approx 2^{N(1-R_1)}, \qquad 1 \le i \le M. \tag{7}$$

where $|A|$ denotes the cardinality of the set $A$.

Now for $1 \le j \le M_2$, define

$$K_j = \sum_{i:C_{ij} \neq \phi} A_i. \tag{8}$$

In other words, $K_j$ is the set of $\mathbf{x} \in \mathscr{B}^N$ such that, for some $\mathbf{x}' \in \mathscr{B}^N$, $f_1(\mathbf{x}) = f_1(\mathbf{x}')$ and $f_2(\mathbf{x}') = j$. The proof proceeds by combining two bounds on $|K_j|$.

We begin by obtaining an upper bound on $|K_j|$. Let

$$\Delta_j = \max_{x_1, x_2 \in K_j} d_H(\mathbf{x}_1, \mathbf{x}_2)$$

be the diameter of $K_j$. Now the subset of $\mathscr{B}^N$ with diameter $\Delta$ with largest cardinality is a sphere with radius $\Delta/2$.

Thus

$$|K_j| \le \sum_{k=0}^{\Delta_j/2} \binom{N}{k} \le 2^{Nh(\Delta_j/2N)},$$

or

$$\frac{1}{N} \log |K_j| \le h(\Delta_j/2N). \tag{9}$$

Next let $\mathbf{x}_1, \mathbf{x}_2, \in K_j$ achieve $d_H(\mathbf{x}_1, \mathbf{x}_2) = \Delta_j$. From the definition of $K_j$ (8), a pair $\mathbf{x}_1', \mathbf{x}_2'$ exists such that

$$f_1(\mathbf{x}_1) = f_1(\mathbf{x}_1'), \qquad f_1(x_2) = f_1(x_2') \tag{10a}$$

and

$$f_2(x_1') = f_1(x_2') = j. \tag{10b}$$

Now (10a), and (6a) with $\alpha = 1$, imply that

$$d_H(\mathbf{x}_1, \mathbf{x}_1') \le 2d_1 N, \qquad d_H(\mathbf{x}_2, \mathbf{x}_2') \le 2d_1 N, \tag{11a}$$

and (10b) and (6a) with $\alpha = 2$ imply that

$$d_H(\mathbf{x}_1', \mathbf{x}_2') \le 2d_2 N. \tag{11b}$$

The triangle inequality and inequalities (11) yield

$$\Delta_j = d_H(\mathbf{x}_1, \mathbf{x}_2) \le d_H(\mathbf{x}_1', \mathbf{x}_2') + d_H(\mathbf{x}_1, \mathbf{x}_1') + d_H(\mathbf{x}_2, \mathbf{x}_2') \le 2d_2 N + 4d_1 N$$

or

$$\frac{\Delta_j}{2N} \le d_2 + 2d_1. \tag{12}$$

Inequalities (12) and (9) yield, for $1 \le j \le M_2$,

$$\frac{1}{N} \log |K_j| \le h(d_2 + 2d_1),$$

and, averaging over $j$, we have

$$\frac{1}{N} \sum_{j=1}^{M_2} \Pr\{f_2(\mathbf{X}) = j\} \log |K_j| \le h(d_2 + 2d_1). \tag{13}$$

This is the first of our bounds on $|K_j|$.

We now obtain a lower bound on $|K_j|$. For $1 \le j \le M_2$, let $m_j$ be the number of $i (1 \le i \le M_1)$ such that $C_{ij} \ne \phi$. Then, using (7) and (8),

$$\log |K_j| \approx \log m_j + N(1 - R_1). \tag{14}$$

Further, from the rate-distortion bound (since $\mathbf{X}$ can be constructed from $f_1(\mathbf{X})$ and $f_2(\mathbf{X})$ with an average distortion $d_0$),

$$\begin{aligned}[1 - h(d_0)]N &\le I(f_1(\mathbf{X}), f_2(\mathbf{X}); \mathbf{X}) \\ &\le H(f_1, f_2) = H(f_2) + H(f_1 \mid f_2) \\ &\le \log M_2 + H(f_1 \mid f_2) = \log M_2 + \sum_{j=1}^{M_2} \Pr\{f_2(\mathbf{X}) = j\} H(f_1 \mid f_2 = j). \end{aligned} \tag{15}$$

Now, given that $f_2(\mathbf{X}) = j$, $f_1(\mathbf{X})$ takes values in a set of cardinality $m_j$. Thus, $H(f_1 \mid f_2 = j)$ can be overbounded by $\log m_j$. Since $M_2 \approx 2^{NR_2}$, we have, from (15),

$$\frac{1}{N} \sum_{j=1}^{M_2} \Pr\{f_2 = j\} \log m_j \ge [1 - h(d_0)] - R_2,$$

and, from (14),

$$\frac{1}{N}\sum_{j=1}^{N} \Pr\{f_2 = j\}\log|K_j| \geq 2 - h(d_0) - R_1 - R_2. \tag{16}$$

This is our second bound on $|K_j|$. Combining (16) and (13), we obtain

$$h(d_2 + 2d_1) \geq 2 - h(d_0) - R_1 - R_2, \tag{17}$$

which is the second part of the theorem, the remainder following on interchanging the roles of $f_1$ and $f_2$.

In the next section, we put the epsilons into the above outline to give a precise proof.

## IV. PROOF OF THE THEOREM

In order to make the ideas in the previous section precise, we must show that in some approximate sense equations (6) and (7) can be assumed to hold. To do this, we begin as follows. As in Section III, let $(f_1, f_2, g_0, g_1, g_2)$ be a code with parameters $(N, M_1, D_0, D_1, D_2)$, and let $A_i$, $B_j$, $C_{ij}$ be as defined in (5). Define the set $S \subseteq \mathscr{B}^N$ as the set of $\mathbf{x}$ such that

$$\frac{1}{N} d(\mathbf{x}, g_\alpha \circ f_\alpha(\mathbf{x})) \leq D_\alpha + \delta, \qquad \alpha = 1, 2, \tag{18a}$$

and

$$\frac{1}{N}\log|A_{f_1(x)}| \geq 1 - R_1 - \delta, \tag{18b}$$

where $\delta > 0$ is an arbitrary fixed parameter. Next, define the set $I \subseteq \{1, \cdots, M_1\}$ as the set of integers $i$ such that

$$P(A_i \cap S) > \delta P(A_i). \tag{19}$$

Note that, if $i \in I$, then $A_i \cap S \neq \phi$. Let $\mathbf{x} \in \mathrm{A}_i \cap \mathrm{S}$, where $i \in I$. Then

$$\begin{aligned} \frac{1}{N}\log|A_i \cap S| &= \frac{1}{N}\log|A_{f_1(x)} \cap S| \\ &\geq \frac{1}{N}\log\delta + \frac{1}{N}\log|A_{f_1(x)}| \\ &\geq \frac{1}{N}\log\delta + (1 - R_1 - \delta). \end{aligned} \tag{20}$$

Finally, paralleling (8), define

$$K'_j = \sum_{i\epsilon I: C_{ij}\cap S\neq\phi} (A_i \cap S), \qquad 1 \leq j \leq M_2. \tag{21}$$

If we imagine for now that $P(S)$ and $\Pr\{f_1(\mathbf{X}) \in \mathrm{I}\}$ are close to 1, then $K'_j$ is an approximation of $K_j$. As in Section III, the theorem is proved by establishing two bounds on $|K'_j|$.

### 4.1 Upper bound on $|K'_j|$

As in Section III, for $1 \le j \le M_2$, let $\Delta'_j$ be the diameter of $K'_j$, so that

$$\frac{1}{N} \log |K'_j| \le h(\Delta'_j/2N). \tag{22}$$

Also, let $\mathbf{x}_1, \mathbf{x}_2 \in K'_j$ achieve $d_H(\mathbf{x}_1, \mathbf{x}_2) = \Delta'_j$. Hence a pair $\mathbf{x}'_1, \mathbf{x}'_2$ exists such that

$$f_1(\mathbf{x}'_1) = f_1(\mathbf{x}_1), \qquad f_1(\mathbf{x}'_2) = f_1(\mathbf{x}_2)$$

and

$$f_2(\mathbf{x}'_1) = f_2(\mathbf{x}'_2) = j.$$

Also, $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}'_1, \mathbf{x}'_2 \in S$. Thus

$$d_H(\mathbf{x}_1, \mathbf{x}'_1),\ d_H(\mathbf{x}_2, \mathbf{x}'_2) \le 2(D_1 + \delta)N$$

and

$$d_H(\mathbf{x}'_1, \mathbf{x}'_2) \le 2(D_2 + \delta)N,$$

so that the triangle inequality yields

$$\frac{\Delta'_j}{2N} \le D_2 + 2D_1 + 3\delta. \tag{23}$$

Inequalities (22) and (23) yield, on averaging over $j$,

$$\frac{1}{N} \sum_{j=1}^{M_2} \Pr\{f_2(x) = j\} \log |K'_j| \le h(D_2 + 2D_1 + 3\delta), \tag{24}$$

which is our upper bound.

### 4.2 Lower bound on $|K'_j|$

Let $j \in \{1, \cdots, M_2\}$, and let $i$ be an integer in the summation set of (21), i.e., $i \in I$, $C_{ij} \cap S \ne \phi$. Thus (20) yields

$$\frac{1}{N} \log |A_i \cap S| \ge \frac{1}{N} \log \delta + (1 - R_1 - \delta).$$

Using this fact, we have from (21), for $1 \le j \le M_2$,

$$\frac{1}{N} \log |K'_j| \ge \frac{1}{N} \log m'_j + \frac{1}{N} \log \delta + (1 - R_1 - \delta), \tag{25}$$

where $m'_j \triangleq |\{i \in I : C_{ij} \cap S \ne \phi\}|$ is the number of terms in the summation of (21). We must now lower bound $m'_j$.

Let $\psi(\mathbf{x})$, $\mathbf{x} \in \mathscr{B}^N$, be defined by

$$\psi(\mathbf{x}) = \begin{cases} 1, & \mathbf{x} \in S,\ f_1(\mathbf{x}) \in I, \\ 0, & \text{otherwise}. \end{cases} \tag{26}$$

Then from the rate-distortion bound (as in (15)),

$$\begin{aligned} N(1 - h(D_0)) &\le I(f_1(\mathbf{X}),\ f_2(\mathbf{X}); \mathbf{X}) \le H(f_1,\ f_2, \psi) \\ &= H(\psi) + H(f_2, | \psi) + H(f_1 | f_2, \psi) \\ &= H(\psi) + H(f_2 | \psi) \\ &\quad + \sum_{j=1}^{M_2} \Pr\{f_2 = j\} \Pr\{\psi = 1 | f_2 = j\} H(f_1 | f_2 = j, \psi = 1) \\ &\quad + \sum_{j=1}^{M_2} \Pr\{\psi = 0\} \Pr\{f_2 = j | \psi = 0\} H(f_1 | f_2 = j, \psi = 0). \end{aligned}$$

Let $\pi = \Pr\{\psi(\mathbf{x}) = 0\}$. Then $H(\psi) \le h(\pi)$. Also $H(f_2 | \psi) \le \log M_2$ and $H(f_1 | f_2 = j, \psi = 0) \le \log M_1$, thus,

$$\begin{aligned} N(1 - h(D_0)) &\le h(\pi) + \log M_2 + \pi \log M_1 \\ &\quad + \sum_{j=1}^{M_2} \Pr\{f_2 = j\} \Pr\{\psi = 1 | f_2 = j\} H(f_1 | f_2 = j, \psi = 1). \end{aligned}$$

Now suppose that $\mathbf{x}$ is such that $f_2(\mathbf{x}) = j$, $\psi(\mathbf{x}) = 1$. Let $f_1(\mathbf{x}) = i_0$. From (26), $i_0 \in I$, and from the definition (5c), $C_{i_0 j} \ne \phi$. Thus $i_0$ is one of the $m'_j$ terms in the summation (21) defining $K'_j$. We can therefore overbound $H(f_1 | f_2 = j, \psi = 1)$ by $\log m'_j$, yielding

$$\begin{aligned} N(1 - h(D_0)) &\le h(\pi) + \log M_2 + \pi \log M_1 \\ &\quad + \sum_{j=1}^{M_2} \Pr\{f_2 = j\} \log m'_j. \end{aligned} \tag{27}$$

Combining (27) with (25), we have

$$\begin{aligned} \frac{1}{N} \sum_{j=1}^{M_2} \Pr\{f_2 = j\} \log |K'_j| &\ge 2 - h(D_0) - R_1 - \delta \\ &\quad - \frac{1}{N} \log M_2 - \frac{h(\pi)}{N} - \frac{\pi}{N} \log M_1 + \frac{1}{N} \log \delta, \end{aligned} \tag{28}$$

which is our lower bound.

### 4.3 Conclusion of the proof

Combining the bounds of (24) and (28), we have

$$\begin{aligned} h(D_2 + 2D_1 + 3\delta) &\ge 2 - h(D_0) - R_1 - \frac{1}{N} \log M_2 \\ &\quad - \frac{h(\pi)}{N} - \delta - \frac{\pi}{N} \log M_1 + \frac{1}{N} \log \delta. \end{aligned} \tag{29}$$

At the conclusion of this section, we prove the following lemma.

*Lemma: If* $(R_1, R_2, d_0, d_1, d_2)$ *is achievable, then for all* $\delta, \epsilon > 0$, *there exists, for N sufficiently large, a code with parameters* $(N, M_1, M_2, D_0, D_1, D_2)$, where $M_\alpha \le 2^{(R_\alpha+\epsilon)N}$ $(\alpha = 1, 2)$, $D_\alpha \le d_\alpha + \epsilon$ $(\alpha = 0, 1, 2)$, *and* $Pr\{\psi(\mathbf{x}) = 0\} = \pi \le \delta + \epsilon$. *(Note that* $\pi$ *depends on S and I, which in turn depend on* $\delta$.)

Now, suppose that $(R_1, R_2, d_0, d_1, d_2)$ is achievable. Pick $\epsilon, \delta > 0$ and let $N$ be large enough so that the code described in the lemma exists. Applying (29) to this code, we have

$$h(d_2 + 2d_1 + 3\delta + 3\epsilon) \ge 2 - h(d_0 + \epsilon) - R_1 - (R_2 + \epsilon) - \delta - \frac{h(\delta + \epsilon)}{N} - (\delta + \epsilon)(R_2 + \epsilon) + \frac{1}{N}\log \delta.$$

Letting $N \to \infty$, and then $\delta, \epsilon \to 0$, we have

$$h(d_2 + 2d_1) \ge 2 - h(d_0) - R_1 - R_2,$$

which is the theorem.

It remains to give a

*Proof of the Lemma:* Let $(x_1, x_2, \cdots)$ be an infinite binary sequence. Define $\mathbf{x}_n^m = (x_n, \cdots x_m)$, $1 \le n < m < \infty$. Now let $(\hat{f}_1, \hat{f}_2, \hat{g}_0, \hat{g}_1, \hat{g}_2)$ be a code with parameters $(\hat{N}, \hat{M}_1, \hat{M}_2, D_0, D_1, D_2)$. For $K = 1, 2, \cdots$, define a "super" code which operates on $K$ successive blocks of $(X_1, X_2, \cdots)$ in the obvious way. That is, letting $q_1, q_2$ be one-to-one mappings

$$q_1\colon \{1, \cdots, \hat{M}_1\}^K \to \{1, \cdots, \hat{M}_1^K\},$$

$$q_2\colon \{1, \cdots, \hat{M}_2\}^K \to \{1, \cdots, \hat{M}_2^K\},$$

then the super code $(f_1, f_2, g_0, g_1, g_2)$ is the code with parameters $(K\hat{N}, \hat{M}_1^K, \hat{M}_2^K, D_0, D_1, D_2)$ defined by

$$f_1(\mathbf{X}_1^{K\hat{N}}) = q_1[\hat{f}_1(\mathbf{X}_1^{\hat{N}}), \hat{f}_1(\mathbf{X}_{\hat{N}+1}^{2\hat{N}}), \cdots, \hat{f}_1(\mathbf{X}_{(K-1)\hat{N}+1}^{K\hat{N}})]$$

and

$$g_1(i) = [\hat{g}_1(i_1), \cdots, \hat{g}_1(i_K)],$$

where $(i_1, \cdots, i_K) = q_1^{-1}(i)$, $1 \le i \le \hat{M}_1^K$, etc.

By the law of large numbers, for $\alpha = 1, 2$, as $K \to \infty$ (with $\hat{N}$ held fixed),

$$\frac{1}{K\hat{N}} d(X_1^{K\hat{N}}, g_\alpha \circ f_\alpha(\mathbf{X}_1^{K\hat{N}})) \to \hat{D}_\alpha,$$

and

$$\frac{1}{K\hat{N}} d(X_1^{K\hat{N}}, g_0(f_1(X_1^{K\hat{N}}), f_2(X_1^{K\hat{N}}))) \to D_0.$$

Also, let $\hat{H}_1 = H(\hat{f}_1(X_1^{\hat{N}}))$. For $\delta > 0$, define

$$I_0 = \left\{ i \in \{1, \cdots, \hat{M}_1^K\} : \frac{1}{K\hat{N}} \log \Pr\{ f_1(\mathbf{X}_1^{K\hat{N}}) = i\} \geq - \frac{\hat{H}_1 + \delta}{\hat{N}} \right\}.$$

$$= \left\{ i : \frac{1}{N} \log |A_i| \geq 1 - \frac{\hat{H}_1}{\hat{N}} - \frac{\delta}{\hat{N}} \right\}.$$

By the AEP, as $K \to \infty$

$$\Pr\{ f_1(\mathbf{X}_1^{K\hat{N}}) \in I_0\} \to 1.$$

Further, since $\hat{H}_1/\hat{N} \leq R_1$, we have, as $K \to \infty$,

$$\Pr\left\{ \frac{1}{N} \log |A_{f_1(X_1^{K\hat{N}})}| \geq 1 - R_1 - \delta \right\} \to 1.$$

We now take a look at the lemma. Let $(R_1, R_2, d_0, d_1, d_2)$ be achievable. Let $\epsilon > 0$ be given, and let $(\hat{f}_1, \hat{f}_2, \hat{g}_0, \hat{g}_1, \hat{g}_2)$ be a code with parameters $(\hat{N}, \hat{M}_1, \hat{M}_2, D_0, D_1, D_2)$ such that $\hat{M}_\alpha \leq 2^{\hat{N}(R_\alpha + \epsilon)}$, $D_\alpha \leq d_\alpha + \epsilon$, $\alpha = 0, 1, 2$. Construct the code $(f_1, f_2, g_0, g_1, g_2)$ with parameters $(N, M_1, M_2, D_0, D_1, D_2)$ where $N = K\hat{N}$, $M_\alpha = \hat{M}_\alpha^K$, $\alpha = 1, 2$, as above. Let $S \subseteq \mathscr{B}^N$ be as defined by (18) with $\delta > 0$ specified. Then, for $K$ sufficiently large, we can make

$$P(S^c) \leq \epsilon/2$$

Then, with $\psi(\cdot)$ defined by (24) and $I$ by (19),

$$\begin{aligned}
\pi &= \Pr\{\psi(\mathbf{X}_1^N) = 0\} = \Pr\{ f_1(\mathbf{X}) \notin I \quad \text{or} \quad \mathbf{X} \notin S\} \\
&\leq \Pr\{ f_1(\mathbf{X}) \notin I\} + \Pr\{\mathbf{X} \notin S\} \\
&= \sum_{i \notin I} P(A_i) + P(S^c) \\
&= \sum_{i \notin I} P(A_i \cap S^c) + \sum_{i \notin I} P(A_i \cap S^c) + P(S^c) \\
&\leq \sum_{i \notin I} \delta P(A_i) + \sum_{i \notin I} P(A_i \cap S^c) + P(S^c) \\
&\leq \delta + 2P(S^c) \leq \epsilon + \delta,
\end{aligned}$$

which establishes the lemma.

## V. ACKNOWLEDGMENT

This problem was formulated by A. Gersho and H. S. Witsenhausen. The authors wish to acknowledge useful discussions with them and also with L. H. Ozarow.

## REFERENCES

1. A. El Gamal, and T. Cover, paper to appear.
2. H. S. Witsenhausen, "On Source Networks with Minimal Breakdown Degradation," B.S.T.J., *59*, No. 6 (July–August 1980), pp. 1083–1087.
3. L. H. Ozarow, "On a Source Coding Problem with Two Channels and Three Receivers," B.S.T.J., *59*, No. 10 (December 1980).