

# BOUND ON DISTORTION DUE TO ERROR FOR A MINIMUM DISTORTION DECODING SCHEME

by BHU DEV SHARMA and GURDIAL\*, *Faculty of Mathematics,  
University of Delhi, Delhi 110007*

(Received 13 March 1975)

Gallager has obtained upper bound on the probability of error under maximum likelihood decoding scheme. When distortion is taken into account the decoding should depend on the distortion and accordingly the aim would shift from upper bounding the probability of error to upper bounding the distortion due to error. This is done in this paper. The upper bound so obtained exponentially tends to zero as block length increases provided that rate distortion function is less than the capacity of an auxiliary channel.

## 1. INTRODUCTION

Shannon (1948) proved the coding theorem which states that for a broad class of communication channels there is a maximum rate, called capacity, at which information can be transmitted with arbitrarily low probability of error. Later in 1959, Shannon (1959) proved the coding theorem under distortion. This is a generalization of earlier studies in information theory as it allows data quality to be taken into account rather than requiring the probability of error to be small.

Consider a discrete channel with input alphabet  $X = (x_1, \dots, x_K)$  and output  $Y = (y_1, \dots, y_J)$  together with a transition probability matrix  $\{P(y_j/x_k)\}$ ,  $k = 1, \dots, K$ ;  $j = 1, \dots, J$ . Also let single letter distortion when  $x_k$  is sent and  $y_j$  received be denoted by  $d(x_k, y_j)$ .

Fano (1961), Gallager (1968) and Shannon *et al.* (1967) have obtained bounds on the probability of error  $P_e$  for codes of block length  $N$ . For discrete memoryless channels the simplest and most elegant proof to this theorem was given by Gallager (1968), together with an upper bound on the probability of decoding error, expressed in the form

$$P_e \leq \exp [-N \{ -\rho R + \max_{\mathbf{p}} E_0(\rho, \mathbf{p}) \}], \quad 0 \leq \rho \leq 1 \quad \dots(1.1)$$

where  $R$  is the rate of transmission and

$$E_0(\rho, \mathbf{p}) = - \ln \sum_{j=1}^J \left( \sum_{k=1}^K p_k P_{j/k}^{1/(1+\rho)} \right)^{1+\rho}, \quad 0 \leq \rho \leq 1. \quad \dots(1.2)$$

---

\*Present address : Departamento de Estatística & Matemática Aplicada, Fortaleza, Brazil.

Recently Arimoto (1973) proved a dual of Gallager's theorem for rates above capacity, together with a lower bound on the probability of decoding error expressed in the form

$$P_e \geq 1 - \exp [- N\{-\rho R + \min_{\mathbf{p}} E_0(\rho, \mathbf{p})\}], \quad 0 > \rho \geq -1. \quad \dots(1.3)$$

We shall denote the set of all sequences of length  $N$  formed of input and output alphabets by  $\mathbf{X}_N$  and  $\mathbf{Y}_N$  respectively. A code with  $M$  code words of length  $N$  is a mapping from a set of  $M$  source messages denoted by the integers  $1, \dots, M$  into a set of  $M$  code words,  $\mathbf{x}_1, \dots, \mathbf{x}_M$ , where  $\mathbf{x}_m \in \mathbf{X}_N, 1 \leq m \leq M$ . A decoder is a mapping of the set of output sequences  $\mathbf{Y}_N$  into the integers  $1, \dots, M$ .

The average distortion per letter is given by

$$D = \sum_{k=1}^K \sum_{j=1}^J p(x_k) P(y_j/x_k) d(x_k, y_j)$$

As is well known (Gallager 1968) the mutual information between  $X$  and  $Y$  is given by

$$I(X; Y) = \sum_{k=1}^K \sum_{j=1}^J p(x_k) P(y_j/x_k) \log \frac{P(y_j/x_k)}{q(y_j)}$$

where  $q(y_j)$  is the probability of the output letter  $y_j$ .

The rate distortion function  $R(D^*)$  is defined (Gallager 1968) as

$$R(D^*) = \inf I(X; Y)$$

where infimum is taken over all transition probabilities  $P(y_j/x_k)$  for which  $D \leq D^*$ .

### Minimum Distortion Decoding Scheme

Let  $d(\mathbf{x}_m, \mathbf{y})$  be the distortion when  $\mathbf{x}_m \in \mathbf{X}_N$  is sent and  $\mathbf{y} \in \mathbf{Y}_N$  is received. We will use the minimum distortion decoding scheme, i.e. the output sequence  $\mathbf{y}$  will be decoded into integer  $m$  if

$$d(\mathbf{x}_m, \mathbf{y}) < d(\mathbf{x}_{m'}, \mathbf{y}) \quad 1 \leq m' \leq M, m' \neq m$$

Depending upon the minimum distortion decoding scheme, we shall obtain upper bounds on 'Distortion due to error'. Sharma and Gurdial (1974) have earlier considered a decoding scheme depending upon the ratio of probabilities and distortion and have obtained bounds on the distortion due to error under this scheme.

## 2. DERIVATION OF THE UPPER BOUND

Let  $D_{em}$  denote the distortion due to error when  $\mathbf{x}_m \in \mathbf{X}_N$  is transmitted, then (cf. Gallager 1965)

$$D_{em} = \sum_{\mathbf{y} \in \mathbf{Y}_N} P(\mathbf{y}/\mathbf{x}_m) d(\mathbf{x}_m, \mathbf{y}) \phi_m(\mathbf{y}) \tag{2.1}$$

where

$$\phi_m(\mathbf{y}) = 1, \text{ if } d(\mathbf{x}_m, \mathbf{y}) \geq d(\mathbf{x}_{m'}, \mathbf{y}) \text{ for some } m' \neq m, \tag{2.2}$$

$$= 0, \text{ otherwise.} \tag{2.3}$$

In this section we shall upper bound average of  $D_{em}$  over all  $m$ , by suitably upperbounding the function  $\phi_m(\mathbf{y})$ .

*Theorem 1* — For any  $D^* \geq 0$ , and for sufficiently large block length  $N$ , there exists a code with  $M$  code words where  $M \leq \exp NR(D^*)$  for which average distortion due to error is such that

$$D_e \leq \exp - N[-\rho R(D^*) + E_0(\rho, \mathbf{p}, \mathbf{d}) - \rho] \tag{2.4}$$

where

$$E_0(\rho, \mathbf{p}, \mathbf{d}) = - \left( \frac{1 + \rho}{N} \right) \ln \sum_{k=1}^K \sum_{j=1}^J \{p(x_k) q(y_j) d(x_k, y_j)\}^{1/(1+\rho)}, \tag{2.5}$$

$$0 \leq \rho \leq 1.$$

PROOF : We take

$$\phi_m(\mathbf{y}) \leq \left[ \sum_{m'=1}^M \frac{d(\mathbf{x}_{m'}, \mathbf{y})^{1/(1+\rho)}}{d(\mathbf{x}_m, \mathbf{y})^{1/(1+\rho)}} \right]^\rho, \rho \geq 0. \tag{2.6}$$

Substituting this value of  $\phi_m(\mathbf{y})$  in (2.1), we get

$$D_{em} \leq \sum_{\mathbf{y} \in \mathbf{Y}_N} P(\mathbf{y}/\mathbf{x}_m) d^{1/(1+\rho)}(\mathbf{x}_m, \mathbf{y}) \left[ \sum_{m'=1}^M d^{1/(1+\rho)}(\mathbf{x}_{m'}, \mathbf{y}) \right]^\rho. \tag{2.7}$$

Equation (2.7) yields a bound on  $D_{em}$  for a particular code. We shall simplify the bound on  $D_{em}$  by averaging over an appropriately chosen ensemble of codes. Clearly at least one code in the ensemble will have a distortion due to error that is as small as the ensemble average distortion due to error. Now following the same steps as in Gallager (1965), we get

$$\bar{D}_{em} \leq \sum_{\mathbf{y} \in \mathbf{Y}_N} \overline{P(\mathbf{y}/\mathbf{x}_m)} \overline{d^{1/(1+\rho)}(\mathbf{x}_m, \mathbf{y})} \left[ \sum_{m'=1}^M \overline{d^{1/(1+\rho)}(\mathbf{x}_{m'}, \mathbf{y})} \right]^\rho \tag{2.8}$$

where an upper bar denotes average.

Since the code words are chosen with probability  $p(\mathbf{x})$ ,

we have

$$\overline{d^{1/(1+\rho)}(\mathbf{x}_m, \mathbf{y})} = \sum_{\mathbf{x} \in \mathbf{X}_N} p(\mathbf{x}) d^{1/(1+\rho)}(\mathbf{x}, \mathbf{y}) \quad \dots(2.9)$$

and

$$\overline{P(\mathbf{y}/\mathbf{x}_m)} = \sum_{\mathbf{x} \in \mathbf{X}_N} p(\mathbf{x}) P(\mathbf{y}/\mathbf{x}) = q(\mathbf{y}). \quad \dots(2.10)$$

Using (2.9) and (2.10), (2.8) becomes

$$\overline{D}_{em} \leq M^\rho \sum_{\mathbf{y} \in \mathbf{Y}_N} q(\mathbf{y}) \left[ \sum_{\mathbf{x} \in \mathbf{X}_N} p(\mathbf{x}) d^{1/(1+\rho)}(\mathbf{x}, \mathbf{y}) \right]^{1+\rho}. \quad \dots(2.11)$$

Let us now denote upper bound  $M$  by  $\exp NR(D^*)$ , where  $R(D^*)$  is the rate-distortion function. Thus

$$\overline{D}_{em} \leq \exp [\rho NR(D^*)] \sum_{\mathbf{y} \in \mathbf{Y}_N} q(\mathbf{y}) \left( \sum_{\mathbf{x} \in \mathbf{X}_N} p(\mathbf{x}) d^{1/(1+\rho)}(\mathbf{x}, \mathbf{y}) \right)^{1+\rho}. \quad \dots(2.12)$$

Since the words  $\mathbf{x}$  and  $\mathbf{y}$  are  $N$ -tuples, let  $\mathbf{x} = (x_1, \dots, x_N)$ ,  $\mathbf{y} = (y_1, \dots, y_N)$  and

$$d(\mathbf{x}, \mathbf{y}) = \sum_{n=1}^N \frac{d(x_n, y_n)}{N}.$$

Using the following inequality (Gallager 1965)

$$\sum_k \left( \sum_j Q_j a_{jk}^{1/r} \right)^r \leq \left[ \sum_j Q_j \left( \sum_k a_{jk} \right)^{1/r} \right]^r, \quad r > 1 \quad \dots(2.13)$$

we get

$$\overline{D}_{em} \leq \exp [\rho NR(D^*)] \left[ \sum_{\mathbf{x}} p(\mathbf{x}) \left( \sum_{\mathbf{y}} q(\mathbf{y}) d^{1/(1+\rho)}(\mathbf{x}, \mathbf{y}) \right) \right]^{1+\rho} \quad \dots(2.14)$$

$$\begin{aligned} &= \exp [\rho NR(D^*)] \left[ \sum_{x_1} \dots \sum_{x_N} \prod_{n=1}^N p(x_n) \left( \sum_{y_1} \dots \sum_{y_N} \prod_{n=1}^N q(y_n) \right. \right. \\ &\quad \left. \left. \times \sum_{n=1}^N \frac{d(x_n, y_n)}{N} \right)^{1/(1+\rho)} \right]^{1+\rho} \quad \dots(2.15) \end{aligned}$$

$$\begin{aligned} &= \exp [\rho NR(D^*)] \left[ \sum_{x_1} \dots \sum_{x_N} \prod_{n=1}^N p(x_n) \left( \sum_{y_1} \frac{q(y_1) d(x_1, y_1)}{N} \right. \right. \\ &\quad \left. \left. + \dots + \sum_{y_N} \frac{q(y_N) d(x_N, y_N)}{N} \right)^{1/(1+\rho)} \right]^{1+\rho} \quad \dots(2.16) \end{aligned}$$

$$\begin{aligned} &\leq \frac{1}{N} \exp [\rho NR(D^*)] \left[ \sum_{x_1} \dots \sum_{x_N} \prod_{n=1}^N p(x_n) \left( \left( \sum_{y_1} q(y_1) d(x_1, y_1) \right)^{1/(1+\rho)} \right. \right. \\ &\quad \left. \left. + \dots + \left( \sum_{y_N} q(y_N) d(x_N, y_N) \right)^{1/(1+\rho)} \right) \right]^{1+\rho} \quad \dots(2.17) \end{aligned}$$

where we have used the inequality

$$\left( \sum_i a_i \right)^r \leq \sum_i a_i^r. \quad \dots(2.18)$$

Using again this inequality, (2.17) becomes

$$\begin{aligned} \bar{D}_{em} &\leq \frac{1}{N} \exp [\rho NR(D^*)] \left[ \sum_{x_1} \dots \sum_{x_N} \prod_{n=1}^N p(x_n) \left( \left( \sum_{y_1} q(y_1) \right. \right. \right. \\ &\quad \left. \left. \times d(x_1, y_1) \right)^{1/(1+\rho)} + \dots + \left( \sum_{y_N} q(y_N) d(x_N, y_N) \right)^{1/(1+\rho)} \right) \right]^{1+\rho} \quad \dots(2.19) \end{aligned}$$

$$\begin{aligned} &= \frac{1}{N} \exp [\rho NR(D^*)] \left[ \sum_{x_1} p(x_1) \left( \sum_{y_1} (q(y_1) d(x_1, y_1))^{1/(1+\rho)} \right) \right. \\ &\quad \left. + \dots + \sum_{x_N} p(x_N) \left( \sum_{y_N} q(y_N) d(x_N, y_N) \right)^{1/(1+\rho)} \right]^{1+\rho}. \quad \dots(2.20) \end{aligned}$$

Using  $p(x_1) \leq p(x_1)^{1/(1+\rho)}$ , we get

$$\begin{aligned} \bar{D}_{em} &\leq \frac{1}{N} \exp [\rho NR(D^*)] \left[ \sum_{n=1}^N \sum_{x_n} \sum_{y_n} (p(x_n) q(y_n) \right. \\ &\quad \left. \times d(x_n, y_n))^{1/(1+\rho)} \right]^{1+\rho} \quad \dots(2.21) \end{aligned}$$

$$\leq N^\rho \exp [\rho NR(D^*)] \left[ \sum_{k=1}^K \sum_{j=1}^J (p(x_k) q(y_j) d(x_k, y_j))^{1/(1+\rho)} \right]^{1+\rho} \quad \dots(2.22)$$

$$\begin{aligned} &= \exp \left[ \rho NR(D^*) + \rho \ln N + \ln \left[ \sum_{k=1}^K \sum_{j=1}^J (p(x_k) q(y_j) d(x_k, y_j))^{1/(1+\rho)} \right]^{1+\rho} \right]. \quad \dots(2.23) \end{aligned}$$

Using the inequality  $\ln x \leq x - 1 < x$ , we get

$$\bar{D}_{em} \leq \exp - N [ -\rho R(D^*) - \rho + E_0(\rho, \mathbf{p}, \mathbf{d}) ]. \quad \dots(2.24)$$

Since the right-hand side of (2.24) is independent of  $m$ , it is a bound on the ensemble distortion due to error and is independent of the probabilities with which the code words are used. Since at least one code in the ensemble must have a distortion due to error as small as the average, the theorem is proved.

For simplicity we shall use the following notation

$$E(R(D^*), \mathbf{p}, \rho) = E_0(\rho, \mathbf{p}, \mathbf{d}) - \rho R(D^*) - \rho. \tag{2.25}$$

In next section, we study some properties of this function.

### 3. PROPERTIES OF $E(R(D^*), \mathbf{p}, \rho)$

The properties are given in the following two theorems.

Given the discrete memoryless channel with input distribution  $\mathbf{p} = (p(1), \dots, p(K))$  and transition probability matrix  $\{P(y_j/x_k)\}, j = 1, \dots, J, k = 1, \dots, K$ , we define an auxiliary channel with input probability given by

$$p'_k = \sum_{j=1}^J \frac{p_k q_j d_{kj}}{\sum_j \sum_k p_k q_j d_{kj}}, \quad k = 1, \dots, K$$

and joint probability given by

$$A(x_k, y_j) = \frac{p_k q_j d_{kj}}{\sum_i \sum_k p_k q_j d_{ki}}, \quad k = 1, \dots, K, j = 1, \dots, J.$$

*Theorem 2* — Let  $I(\mathbf{p}'; A) = \sum_{k=1}^K \sum_{j=1}^J A_{jk} \ln A_{jk} \neq 0$ . ... (3.1)

$$\begin{aligned} \text{Then } E_0(0, \mathbf{p}, \mathbf{d}) &= \frac{1}{N} \ln \sum_j \sum_k p_k q_j d_{kj} \\ &= - \frac{1}{N} \ln D \end{aligned} \tag{3.2}$$

$$> 0 \text{ if } \frac{1}{D} > 1$$

$$E_0(\rho, \mathbf{p}, \mathbf{d}) > 0, \rho > 0 \tag{3.3}$$

$$\frac{1}{N} I(\mathbf{p}') \geq \frac{\partial E_0(\rho, \mathbf{p}, \mathbf{d})}{\partial \rho} > 0, \rho \geq 0 \tag{3.4}$$

and

$$\frac{\partial^2 E_0(\rho, \mathbf{p}, \mathbf{d})}{\partial \rho^2} \leq 0, \rho \geq 0 \tag{3.5}$$

PROOF: Taking partial derivative with respect to  $\rho$  of  $E_0(\rho, \mathbf{p}, \mathbf{d})$  we get,

$$\begin{aligned} \frac{\partial E_0(\rho, \mathbf{p}, \mathbf{d})}{\partial \rho} &= \frac{1}{N} \sum_k \sum_j \frac{(p_k q_j d_{kj})^{1/(1+\rho)}}{\sum_k \sum_j (p_k q_j d_{kj})^{1/(1+\rho)}} \ln \frac{(p_k q_j d_{kj})^{1/(1+\rho)}}{\sum_k \sum_j (p_k q_j d_{kj})^{1/(1+\rho)}} \\ \frac{\partial E_0(\rho, \mathbf{p}, \mathbf{d})}{\partial \rho} \Big|_{\rho=0} &= \frac{1}{N} \sum_j \sum_k \frac{p_k q_j d_{kj}}{\sum_j \sum_k p_k q_j d_{kj}} \ln \frac{p_k q_j d_{kj}}{\sum_j \sum_k p_k q_j d_{kj}} \\ &= \frac{1}{N} I(\mathbf{p}'). \end{aligned} \quad \dots(3.6)$$

Taking the 2nd partial derivative with respect to  $\rho$  of  $E_0(\rho, \mathbf{p}, \mathbf{d})$ , we get

$$\begin{aligned} \frac{\partial^2 E_0(\rho, \mathbf{p}, \mathbf{d})}{\partial \rho^2} &= \frac{1}{N} \frac{1}{1+\rho} \sum_j \sum_k \frac{(p_k q_j d_{kj})^{1/(1+\rho)}}{\sum_j \sum_k (p_k q_j d_{kj})^{1/(1+\rho)}} [\ln (p_k q_j d_{kj})^{1/(1+\rho)}]^2 \\ &\quad + \frac{1}{N} \left[ \sum_j \sum_k \frac{(p_k q_j d_{kj})^{1/(1+\rho)}}{\sum_j \sum_k (p_k q_j d_{kj})^{1/(1+\rho)}} \ln (p_k q_j d_{kj})^{1/(1+\rho)} \right]^2. \end{aligned}$$

If we consider  $\ln (p_k q_j d_{kj})^{1/(1+\rho)}$  as a random variable then

$$\begin{aligned} &\sum_j \sum_k \frac{(p_k q_j d_{kj})^{1/(1+\rho)}}{\sum_j \sum_k (p_k q_j d_{kj})^{1/(1+\rho)}} [\ln (p_k q_j d_{kj})^{1/(1+\rho)}]^2 \\ &\quad - \left[ \sum_j \sum_k \frac{(p_k q_j d_{kj})^{1/(1+\rho)}}{\sum_j \sum_k (p_k q_j d_{kj})^{1/(1+\rho)}} \ln (p_k q_j d_{kj}) \right]^2 \end{aligned}$$

is the variance of the random variable and is positive.

Therefore

$$\frac{\partial^2 E_0(\rho, \mathbf{p}, \mathbf{d})}{\partial \rho^2} \leq 0. \quad \dots(3.7)$$

Using Theorem 2, we can maximize  $E_0(R(D^*), \mathbf{p}, \rho)$  with respect to  $\rho$ . Let

$$E(R(D^*), \mathbf{p}) = \max_{0 \leq \rho \leq 1} [-\rho R(D^*) + E_0(\rho, \mathbf{p}, \mathbf{d})] - \rho. \quad \dots(3.8)$$

$$\text{Then } R(D^*) + 1 = \frac{\partial E_0(\rho, \mathbf{p}, \mathbf{d})}{\partial \rho} \quad \dots(3.9)$$

Now if some  $\rho$  in the range  $0 \leq \rho \leq 1$  satisfies (3.9) then that  $\rho$  must maximize (3.8). Again from (3.7)  $\frac{\partial E_0(\rho, \mathbf{p}, \mathbf{d})}{\partial \rho}$  is non-increasing with  $\rho$ , so that a solution to (3.9) exists if  $R(D^*)$  lies in the following range

$$\frac{\partial E_0(\rho, \mathbf{p}, \mathbf{d})}{\partial \rho} \Big|_{\rho=1} \leq R(D^*) + 1 \leq I(\mathbf{p}'). \quad \dots(3.10)$$

Thus for the range specified above,

$$E(R(D^*), \mathbf{p}) = E_0(\rho, \mathbf{p}, \mathbf{d}) - \rho \frac{\partial E_0(\rho, \mathbf{p}, \mathbf{d})}{\partial \rho} \quad \dots(3.11)$$

and

$$R(D^*) + 1 = \frac{\partial E_0(\rho, \mathbf{p}, \mathbf{d})}{\partial \rho}, \quad 0 \leq \rho \leq 1. \quad \dots(3.12)$$

For  $R(D^*) + 1 < \frac{\partial E_0(\rho, \mathbf{p}, \mathbf{d})}{\partial \rho}$  (3.11) and (3.12) are not valid.

In this case, the function  $E(R(D^*), \mathbf{p})$  increases with  $\rho$  in the range  $0 \leq \rho \leq 1$  and thus the maximum occurs at  $\rho = 1$ . Hence

$$E(R(D^*), \mathbf{p}) = E_0(1, \mathbf{p}, \mathbf{d}) - R(D^*) - 1, \text{ for } R(D^*) + 1 < \frac{\partial E_0(\rho, \mathbf{p}, \mathbf{d})}{\partial \rho} \Big|_{\rho=1}. \quad \dots(3.13)$$

Differentiating (3.11) with respect to  $\rho$ , we get  $-\rho \frac{\partial^2 E_0(\rho, \mathbf{p}, \mathbf{d})}{\partial \rho^2}$ . Thus  $E(R(D^*), \mathbf{p})$  is increasing with for  $\rho \geq 0$  and is equal to  $\frac{1}{N} \ln \frac{1}{D}$  for  $\rho = 0$ .

From (3.11) and (3.12), we have

$$\frac{\partial E(R(D^*), \mathbf{p})}{\partial R} = -\rho.$$

Further let  $E(R(D^*)) = \max_{\mathbf{p}} E(R(D^*), \mathbf{p})$ . Then  $E(R(D^*))$  is the upper envelope of all of the  $E(R(D^*), \mathbf{p})$  curves and we have the following theorem :

*Theorem 3* — For auxiliary memoryless channel the random coding exponent  $E(R(D^*))$  is a convex  $\cup$ , decreasing, positive function of  $R(D^*)$  for  $0 \leq R(D^*) < C'$ .

**PROOF :** Since  $E(R(D^*))$  is a maximum over a set of functions that are convex  $\cup$  and decreasing in  $R(D^*)$ . Therefore the maximizing function is also convex  $\cup$  and decreasing in  $R(D^*)$ . Also for the  $\mathbf{p}$  that yields capacity on the auxiliary channel,  $E(R(D^*), \mathbf{p})$  is positive for  $R(D^*) < C'$  and, therefore,  $E(R(D^*))$  is positive for  $R(D^*) < C'$ .

REFERENCES

Arimoto, S. (1973). On the converse to the coding theorem for discrete memoryless channels. *IEEE Trans. Inform. Theory*, IT-19, 357-59.



- Fano, R. M. (1961). *Transmission of Information*. M. I. T. Press, Cambridge, and Wiley, New York.
- Gallager, R. G. (1965). A simple derivation of the coding theorem and some applications. *IEEE Trans. Inform. Theory*, IT-11, 3-18.
- (1968). *Information Theory and Reliable Communication*. Wiley, New York.
- Shannon, C. E. (1948). A mathematical theory of communication. *BSTJ*, 27, 379-423.
- (1959). Coding theorems for a discrete source with a fidelity criterion. *IRE Nat. Conv. Rec.*, Part 4, 142-63.
- Shannon, C. E., Gallager, R. C., and Berlekamp, E. R. (1967). Lower bounds to error probability for coding on discrete memoryless channels I. *Inform. Control*, 10, 65-103.
- Sharma, B. D., and Gurdial (1974). Bounds on distortion due to error for rates below and above channel capacity. *Inform. Control*, 26, No. 3, 272-79