THE ϵ-CAPACITY OF CLASSES OF UNKNOWN CHANNELS

by

L. J. Forys

P. P. Varaiya

Memorandum No. ERL-M 242

7 March 1968

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

INTRODUCTION

In the classical theory of communication, nearly all of the results pertaining to the rate of transmission of information have depended upon accurate characterizations of the channel operator and the statistics of the noise. Recently however, W. L. Root and P. P. Varaiya have investigated the problem of determining the channel capacity in situations where accurate characterizations are not possible. In [1], Root studied the problem of estimating the maximum rate of transmission for the case where the channel operator was known to a large extent, but where it was not possible to characterize the distortions and perturbations of the signal probabilistically in a more than rudimentary manner. Other investigators [2], [6], considered only the limiting behavior, where the perturbations and distortions became vanishingly small.

The case where the noise was known to be Gaussian but where the channel operator was assumed only to lie in some conditionally compact class of linear operators was investigated jointly by Root and Varaiya in [3], [4].

In this paper, we shall consider a combination of both these cases. We will assume that the channel operator is known only to belong to a conditionally compact class of linear operators and that the probability that the noise has average power greater than some number ϵ is vanishingly small. The number ϵ may be arbitraily large and no

It is convenient to denote sequences of x's by $u = (x_1, x_2, \ldots, x_n)$ and sequences of y's by $w = (y_1, y_2, \ldots, y_n)$. We can then characterize the n-extension of a channel A by

$$w = A_\gamma^n u$$

where A_γ^n is the $np \times np$ matrix defined as follows:

$$A_\gamma^n = \begin{bmatrix} A_\gamma & 0 & 0 & 0 & \cdots & 0 \\ 0 & A_\gamma & 0 & 0 & \cdots & 0 \\ 0 & 0 & A_\gamma & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \cdots & A_\gamma \end{bmatrix}$$

In most physical problems we are not interested in transmitting all of R^{np}. We will therefore only consider transmitted sequences u which satisfy an average power constraint i.e., $u \in U_M^p = \{u \mid \|u\|^2 \leq nM, M > 0\}$ where $\|u\|^2 = \Sigma_{j=1}^n \Sigma_{i=1}^p (x_j^i)^2$ and x_j^i is the ith component of x_j. It is also convenient to assume that the class $\mathcal{C}$ is bounded i.e., there exists a real number a, $0 < a < \infty$ such that the operator norm $\|A_\gamma\|$ of any matrix $A_\gamma \in \mathcal{C}$ satisfies $\|A_\gamma\| < a$.

In order to define our problem it is necessary to introduce the concept of the distance between the "total images" of any two input sequences. Given any two input (transmitted) sequences u_i, $u_j \in U_M^p$, we define the number $d_{i,j}$ as follows

$$d_{i,j} = \inf ||z - w||$$

where the infimum is taken over all $z \in \bigcup_{A_\gamma \in \mathcal{C}} A_\gamma^n u_i$, $w \in \bigcup_{A_\gamma \in \mathcal{C}} A_\gamma^n u_j$.

Using this notation we first define what we mean by an attainable ϵ - rate of transmission and then define the ϵ - capacity as the supremum of such rates.

We say that R is an <u>attainable ϵ - rate of transmission</u> for $\mathcal{C}$ if there exists a sequence n_k with $\lim_{k \to \infty} n_k = \infty$ and $G(n_k)$ transmitted signals $u_j \in U_M^p$ of length n_k such that $\hat{d} = \inf_{i \neq j} d_{i,j} \geq \epsilon \sqrt{n_k}$ and $\lim_{k \to \infty} \frac{1}{n_k} \log_e G(n_k) = R$.

The <u>ϵ - capacity</u> for $\mathcal{C}$, denoted by C_ϵ is the supremum of all attainable ϵ - rates of transmission for $\mathcal{C}$.

In Sections II and III, we determine numbers $\overline{C}_\epsilon$ and $\underline{C}_\epsilon$ such that $\overline{C}_\epsilon \geq C_\epsilon \geq \underline{C}_\epsilon$. Although these estimates are not as tight as one would like, nevertheless they are in many cases better than those obtained in [1] even when $\mathcal{C}$ is a single point. This will be discussed in more detail in Section IV.

B. Discrete Infinite Dimensional Model

The infinite dimensional model is related to the finite dimensional model in the sense that the transmitted signals x and the received signals y are related by the matrix transformation

$$y = A_\gamma x$$

In this case however, x and y are infinite dimensional column vectors and A_γ is an infinite dimensional matrix which belongs to a class of such matrices. The n-extension of A_γ, A_γ^n is defined as before so that it carries an n-sequence of transmitted vectors $u = (x_1, x_2, \ldots, x_n)$ into an n-sequence of received vectors $w = (y_1, y_2, \ldots, y$ with

$$y_i = A_\gamma x_i, \quad i = 1, 2, \ldots, n$$

We will also assume the following:

(i) Each matrix $A_\gamma \in \mathcal{C}$, is Hilbert-Schmidt, i.e., if $A_\gamma = \{a_{ij}\}$ then $\Sigma_{ij}\, a_{ij}^2 < \infty$.

(ii) For any two Hilbert-Schmidt matrices $A = \{a_{ij}\}$, $B = \{b_{ij}\}$ we define $||A - B||^2 = \Sigma_{i,j}\, |a_{ij} - b_{ij}|^2$. Then, $||\ \ ||$ defines a metric (in fact, a norm). We further assume that $\mathcal{C}$ is a totally bounded subset of the metric space of all Hilbert-Schmidt matrices.

As before, we impose the average input power constraint M, i.e., we only select transmitted sequences u which satisfy

$$u \in U_M^\infty = \{u \mid ||u||^2 = \Sigma_{i,j}\, (x_j^i)^2 < Mn\}.$$

For any two transmitted sequences u_i, u_j, we again define the distance between the total images of u_i and u_j, $d_{i,j}$ as follows

$$d_{i,j} = \inf ||z - w||$$

where the infimum is taken over all $z \in \bigcup_{A_\gamma \in \mathscr{C}} A_\gamma^n u_i$, $w \in \bigcup_{A_\gamma \in \mathscr{C}} A_\gamma^n u_j$ ($\bigcup_{A_\gamma \in \mathscr{C}} A_\gamma^n u = \{z \,|\, z = (A_\gamma x_1, A_\gamma x_2, \ldots, A_\gamma x_n),\ A_\gamma \in \mathscr{C}\}$). With this notation, the definitions of attainable ϵ - rate of transmission, and ϵ - capacity are identical to those in Section I.A.

C. Continuous Time Model

We also consider classes of channels that can be described as follows. By a transmitted signal over the time interval $[-T, T]$ we mean a real valued function x which is square integrable with respect to Lebesgue measure on $[-T, T]$. If x is the input signal over $[-T, T]$ the received signal y over an interval $[-T, T]$ is to be given by an expression of the form

$$y(t) = \int_{-T}^{T} h_\gamma(t - \tau)\, x(\tau)\, d\tau, \quad -T \leq t \leq T$$

The function h_γ belongs to a class $\mathscr{C}$ of channel operators which has the following properties:

(i) If $h_\gamma \in \mathscr{C}$, then $h_\gamma \in L_2(-\infty, \infty)$ and there exists a finite positive number a such that $\|h_\gamma\|_2 < a$, for all $h_\gamma \in \mathscr{C}$.

(ii) $\mathscr{C}$ is a conditionally compact subset of $L_2(-\infty, \infty)$ i.e., the closure of $\mathscr{C}$ is compact in $L_2(-\infty, \infty)$.

It should be noted that in [1], Root makes the assumption that the convolution operator determined by a function h have finite memory, i.e., there exists a $\delta > 0$ such that $h(t) = 0$ for all $|t| > \delta$. We do not require that $\mathcal{L}$ be composed only of finite memory operators. Root's results however can be extended to include the case where the operators do not have finite memory.

In Section II we approximate the class $\mathcal{L}$ by a class of finite memory operators. The assumptions imposed on $\mathcal{L}$ give this approximating class some useful properties. Let τ_δ denote a "truncation" operator defined as follows;

$$
\begin{aligned}
(\tau_\delta h)(t) &= h(t), \qquad |t| \leq \delta \\
&= 0 \quad , \qquad |t| > \delta
\end{aligned}
$$

Notice that τ_δ is a linear, continuous operator which maps $L_2(-\infty, \infty)$ into $L_2(-\infty, \infty)$. Thus, $\mathcal{L}_\delta$ defined as

$$\mathcal{L}_\delta = \{h_{\gamma,\delta} \mid h_{\gamma,\delta} = \tau_\delta h_\gamma, \; h_\gamma \in \mathcal{L}\}$$

is also a conditionally compact subset of $L_2(-\infty, \infty)$ and $\|h_{\gamma,\delta}\|_2 < a$ for all $h_{\gamma,\delta} \in \mathcal{L}_\delta$.

We can let H_γ denote the integral operator with kernel $h_\gamma(t - \tau)$ and write the relation between x and y as follows

$$y = H_\gamma x$$

We again impose an average power constraint on the transmitted signals by requiring that $x \in \mathcal{U}_M^T$ where

$$\mathcal{U}_M^T = \left\{ x \mid \int_{-T}^{T} x^2(t)\, dt \leq 2MT \right\}$$

For any two transmitted signals x_i, x_j, we define the distance between the "total images" of x_i, x_j as follows

$$d_{i,j} = \inf \| z - w \|_2$$

where the infimum is taken over all $z \in \bigcup_{H_\gamma \in \mathcal{C}} H_\gamma x_i$, $w \in \bigcup_{H_\gamma \in \mathcal{C}} H_\gamma x_j$.

We say that R is an attainable <u>ϵ - rate of transmission</u> for $\mathcal{C}$ if there exists a sequence T_k, $\lim_{k \to \infty} T_k = \infty$ and $G(T_k)$ transmitted signals $u_j \in \mathcal{U}_M^{T_k}$ over $[-T_k, T_k]$ such that $d = \lim_{i \neq j} d_{i,j} \geq \epsilon \sqrt{T_k}$ and $\lim_{k \to \infty} \frac{1}{T_k} \log_e G(T_k) = R$.

The <u>ϵ - capacity</u> for $\mathcal{C}$, denoted by C_ϵ is the supremum of all attainable ϵ - rates of transmission for $\mathcal{C}$.

Having defined the three types of models of interest, we will show in the next section that the lower bounds for C_ϵ obtainable for model A can be used to obtain similar results for models B and C.

II. LOWER BOUNDS ON C_ϵ

In this section we derive lower bounds on C_ϵ, $\underline{C}_\epsilon$, for each of the models given in Section I. The best results to date for such bounds were derived by Root in [1] for the case where the class $\mathcal{C}$ consisted of a single point or a ball. His bounds were obtained by using variations of packing arguments. It appears however that such methods cannot be satisfactorily generalized to include cases where the structure of $\mathcal{C}$ is complex; the total images may, for example, be disconnected and their shapes may depend on the choice of transmitted sequences.

Our approach is fundamentally quite different. Our lower bound for the C_ϵ of the model A (see Theorem 1A) is obtained in three steps:

(1) We imbed our problem in a stochastic problem by artificially adding Gaussian noise. A rate of transmission is chosen so that the maximum probability of error for this rate is less than some judiciously chosen exponential. This is done in Theorem 2A.

(2) We then show that certain exponential bounds on the maximum probability of error imply that the total images of the transmitted sequences yielding these bounds are separated from each other by the required distance. This result is stated as Theorem 3A.

(3) Theorems 2A and 3A are combined to give the bound. The results for models B and C as given in Theorems 1B and 1C are obtained by approximating these models by model A and applying Theorem 1A.

A. The Finite Dimensional Case

Theorem 1A

If the class $\mathcal{C}$ of channels satisfies the assumptions of model A then

$$C_\epsilon \geq \underline{C}_\epsilon = \sup_{\beta>0} \{ \sup_{S \in \mathcal{S}_M^p} \inf_{A_\gamma \in \mathcal{C}} \log_e |I + A_\gamma S A_\gamma' / 2\beta|^{1/2} - \epsilon^2/2\beta \}$$

where $\mathcal{S}_M^p = \{S \mid \text{trace } S \leq M, \text{ S is a } p \times p \text{ covariance matrix}\}$, $|\ |$ denotes the determinant of a $p \times p$ matrix and I is the $p \times p$ identity matrix.

We defer the proof of Theorem 1A until the end of this section. It is a logical consequence of Theorems 2A and 3A which follow. Before we proceed any further, note that we may always assume that $\underline{C}_\epsilon > 0$ If $\underline{C}_\epsilon \leq 0$, there is nothing to prove since $C_\epsilon \geq 0$.

It is necessary at this point to "imbed" model A in a stochastic model. That is, to the received signal y of model A we now add a Gaussian random vector z and call the new received vector $\hat{y}$. To be precise, $\hat{y}$, x, and z are related by

$$\hat{y} = A_\gamma x + z, \quad A_\gamma \in \mathcal{C}$$

where z is a zero mean Gaussian random (column) vector of dimension p and with covariance matrix βI, β a positive scalar, I is the $p \times p$ identity matrix, and where x, A_γ, $\mathcal{C}$ are defined as before.

The n-extension channel A^n_γ of this new model is defined according to the rule

$$\hat{y}_i = A_\gamma x_i + z_i, \qquad i = 1, 2, \ldots, n$$

where the z_i are mutually independent Gaussian random vectors each with covariance matrix βI and zero mean. We denote sequences of x's again by $u = (x_1, x_2, \ldots, x_n)$ and sequences of $\hat{y}$'s by $v = (\hat{y}_1, \hat{y}_2, \ldots, \hat{y}_n)$.

In order to take advantage of random coding arguments we also define a probability distribution for the transmitted (or input) signals x_i. For convenience we assume that the x_i's are zero mean, independent Gaussian random vectors each with covariance matrix S belonging to the set $\mathscr{S}^p_M = \{S \mid \text{trace } S \leq M,\ S \text{ is a } p \times p \text{ covariance matrix}\}$. We let $q(u)$ denote the np-variate probability density function for the transmitted sequences u which are assumed to be statistically independent of the noise z. The conditional probability of a set of received signals B, given a transmitted sequence u depends upon the channel A^n_γ. We denote this probability by $P_\gamma(B|u)$ and define $p_\gamma(v|u)$ to be the np-variate Gaussian density function determined by $P_\gamma(B|u)$. Then,

$$p_\gamma(v) = \int p_\gamma(v|u)\, q(u)\, du$$

defines a probability density for the received (or output) vectors v.

The <u>mutual information</u> $J_\gamma(u,v)$ for the density q(u) and the channel A^n_γ is defined to be

$$J_\gamma(u,v) = \log_e[p_\gamma(v|u)/p_\gamma(v)]$$

To prove Theorem 2A we need a number of facts which involve slight modifications of known results. We state these facts in the series of lemmas which follow.

<u>Lemma 1</u>

Let $\mathscr{L}' = \{A_1, A_2, \ldots, A_{\gamma'}, \ldots, A_L\}$ be a finite class of channels and q(u) be an input (transmitted sequence) probability density function determining $p_\gamma(u,v)$ and $J_\gamma(u,v)$. Let $E \subset R^{pn}$ be a fixed constraint set. For any $\alpha > 0$, $\delta > 0$, $G \geq 1$, if

$$\Delta = LGe^{-\alpha} + L^2 e^{-\delta} + \Delta LP\{E^c\} + \sum_{\gamma'=1}^{L} P_{\gamma'}\{J_{\gamma'}(u,v) \leq \alpha + \delta\}$$

and $0 < \Delta < 1$, then there exists G distinct input sequences $u_1, \ldots, u_G$ all belonging to E, and G disjoint sets $B_1, \ldots, B_G$ of received sequences $\bigcup_{i=1}^{G} B_i = R^{pn}$, such that

$$P_\gamma\{B_i^c|u_i\} \leq \Delta \quad \text{for all} \quad i = 1,2,\ldots,G \quad \text{and} \quad \gamma' = 1,2,\ldots,L$$

The proof is a trivial modification of Lemma 3 in [5].

Lemma 2

For any channel A_γ^n,

$$\text{Expectation of } J_\gamma(u,v) = E\, J_\gamma(u,v) = n \log_e \left| I + \frac{1}{\beta} A_\gamma S A_\gamma' \right|^{1/2}$$

Proof:

This lemma is a special case Lemma 3 of [3]. We reprove this lemma to familiarize ourselves with the notation. It is easily seen that v is a Gaussian random vector with mean zero and $pn \times pn$ covariance matrix

$$E\, vv' = \begin{bmatrix} \Gamma & 0 & 0 & \cdots & \cdots & 0 \\ 0 & \Gamma & 0 & 0 & \cdots & 0 \\ \vdots & & \Gamma & & & \\ \vdots & & & \Gamma & & \\ \vdots & & & & \ddots & \\ 0 & 0 & 0 & 0 & \cdots & \Gamma \end{bmatrix}$$

where $\Gamma = A_\gamma S A_\gamma' + \beta I$. Hence,

$$p_\gamma(v) = \prod_{i=1}^{n} [1/(2\pi)^{p/2} |\Gamma|^{1/2}] \exp[-\tfrac{1}{2} \hat{y}_i' \Gamma^{-1} \hat{y}_i]$$

The Gaussian random vector $(z_1, z_2, \ldots, z_n) = v - A_\gamma^n u$ has covariance matrix βI, I the $np \times np$ identity matrix and mean zero, hence

$$p_\gamma(v|u) = \prod_{i=1}^{n} (2\pi)^{-p/2} \beta^{-p/2} \exp[-\frac{1}{2\beta}(\hat{y}_i - A_\gamma x_i)'(\hat{y}_i - A_\gamma x_i)]$$

Thus $J_\gamma(u,v)$ is given by

$$J_\gamma(u,v) = \sum_{i=1}^{n} \log_e |\Gamma|^{1/2}/\beta^{p/2} + \sum_{i=1}^{n} \{\hat{y}_i' \Gamma^{-1} \hat{y}_i - \frac{1}{2\beta}(\hat{y}_i - A_\gamma x_i)'(\hat{y}_i - A_\gamma x_i)\}$$

The first term on the right is a constant and in fact equals $n \log_e |I + \frac{1}{\beta} A_\gamma S A_\gamma'|^{1/2}$; the expectation of the second term is

$$\sum_{i=1}^{n} E[\hat{y}_i' \Gamma^{-1} \hat{y}_i - \frac{1}{\beta} z_i' z_i] = \sum_{i=1}^{n} \frac{1}{2}(p-p) = 0$$

<u>Lemma 3</u>

Let A_γ^n be a fixed channel, $q(u)$ be the Gaussian distribution for u with covariance matrix S and let $J_\gamma(u,v)$ be defined as above. Then for any $\Theta > 0$, and any t satisfying $1 \geq t > 0$,

$$P_\gamma\{J_\gamma(u,v) \leq E\, J_\gamma(u,v) - n\Theta\}$$

$$\leq \exp\{-nt\Theta - n \log_e[\,|I + (1-t^2)A_\gamma S A_\gamma'/\beta|^{1/2} / |I + A_\gamma S A_\gamma'/\beta|^{1/2}]\}$$

Proof:

The proof of this lemma involves an exact duplication of part of the proof of Lemma 4 in [3].

<u>Lemma 4</u>

Let x_i, $i = 1, 2, \ldots, n$ be independent identically distributed p-dimensional Gaussian random vectors with mean zero and covariance matrix S. Let the trace of S be equal to $M - \xi$, $0 < \xi < M$. Then,

$$P\{||u||^2 = \sum_{i=1}^{n} ||x_i||^2 \geq nM\} \leq \exp\{-(M/(M-\xi) - 1 - \log_e[M/(M-\xi)])\frac{n}{2}\}$$

Proof:

This lemma is just an application of Lemma 5 in [3].

<u>Lemma 5</u>

Let A_γ^n, $A_{\hat{\gamma}}^n$ be the n-extension of two channels and let u be a transmitted sequence of vectors x_i. Let $p_\gamma(v|u)$ be the np-variate probability density for the received sequence v, given u, for A_γ^n, and $p_\gamma(v|u)$ be the corresponding density for $A_{\hat{\gamma}}^n$. Then, for those v satisfying $||v||^2 \leq nC$, $C > 0$,

$$p_\gamma(v|u)/p_{\hat{\gamma}}(v|u) \leq \exp\{n[\sqrt{MC} + aM]\,||A_\gamma - A_{\hat{\gamma}}||/\beta\}$$

Here, M and a are the numbers introduced in Section I.A, i.e.,
$||u||^2 = \sum_{i=1}^{n} ||x_i||^2 \leq nM$, $||A_\gamma||$ and $||A_{\hat{\gamma}}|| \leq a$.

Proof:

This lemma is just a special case of Lemma 7 in [3].

Lemma 6

Let A_γ^n be any channel satisfying the conditions for model A. Let u be any transmitted n-sequence satisfying $||u||^2 \leq nM$. Then, the received sequence v satisfies

$$P_\gamma\{||v||^2 \geq n[2a^2M + 2\beta p + 2k] \mid u\}$$

$$\leq [(1 + k/\beta^p)\exp\{-k/\beta^p\}]^{n/2}$$

for every scalar $k > 0$.

Proof:

The proof of this lemma involves only a slight modification of the proof of Lemma 8 in [3].

We can now state and prove Theorem 2A.

Theorem 2A

Suppose that the class $\mathscr{C}$ satisfies the conditions for model A. If R is any positive number satisfying $R < \underline{C}_\epsilon$, and $G = [e^{Rn}]$ = the greatest integer less than e^{Rn} then for the additive Gaussian noise channel, there exists G disjoint (decoding) sets B_i, $i = 1, 2, \ldots, G$ with $\bigcup_{i=1}^{G} B_i = R^{np}$, G transmitted sequences (code words) u_i satisfying $||u_i||^2 \leq Mn$, and finite positive numbers $\tilde{\beta}$, N, θ such that

$$P_\gamma\{B_i^c | u_i\} \leq \exp\{-(\epsilon^2/8\tilde{\beta} + \theta)n\} \text{ for all } A_\gamma \epsilon \mathscr{b}, \; i = 1, \ldots, G \text{ and } n \geq N.$$

Proof:

Since $R < \underline{C}_\epsilon$, then by the definition in Theorem 1A, we can find positve numbers $\tilde{\beta}$, θ such that

$$R + 21\theta < \sup_{S \epsilon \mathscr{S}_M^p} \inf_{A_\gamma \epsilon \mathscr{b}} \log_e |I + A_\gamma S A_\gamma' / 2\tilde{\beta}|^{1/2} - \epsilon^2/2\tilde{\beta}$$

Let $\hat{R}$ denote the first term on the right hand side of this inequality.

We will first consider the case where $\mathscr{b}$ consists of only a finite number L_η of elements and denote such a class by $\mathscr{b}'$.

With the conditions imposed on $\mathscr{b}'$ and $\mathscr{S}_M^p$, the product set $(S, A_{\gamma'})$ of all $S \epsilon \mathscr{S}_M^p$ and $A_{\gamma'} \epsilon \mathscr{b}'$ is a conditionally compact set with respect to the usual metric for $R^{p^2} \times R^{p^2}$. Because of this fact and the continuity of the determinant and $\log_e$ functions it follows that there exists a scalar ξ, $0 < \xi < M$ and $\hat{S} \epsilon \mathscr{S}_M^p$ with trace of $\hat{S} = M - \xi$ such that

$$\log_e \left|I + \frac{1}{2\tilde{\beta}} A_\gamma \hat{S} A_\gamma'\right|^{1/2} \geq \hat{R} - \theta \qquad \text{for all } A_\gamma \epsilon \mathscr{b}'.$$

We let S be the covariance matrix for the transmitted vectors x_i and $\tilde{\beta} I$ be the covariance matrix of the noise vectors z_i in the stochastic model described earlier in this section. Apply Lemma 1 to this stochastic model by choosing $G = [e^{Rn}]$, $E = U_M^p = \{u \mid ||u||^2 \leq Mn\}$, $\alpha = (R + \epsilon^2/8\tilde{\beta} + 5\theta)n$, $\delta = (\epsilon^2/8\tilde{\beta} + 5\theta)n$. As a consequence of this lemma we know that there exists G transmitted sequences u_i in U_M^p and G disjoint sets B_i, $\bigcup_{i=1}^{G} B_i = R^{np}$ such that

$$P_\gamma\{B_i^c | u_i\} \leq \Delta \quad \text{for all} \quad i = 1, \ldots, G \text{ and all } \gamma = 1, 2, \ldots, L_\eta ,$$

where

$$\begin{aligned}\Delta[1 - L_\eta P\{||u||^2 \geq Mn\}] \leq L_\eta\, & e^{Rn}[\exp\{-(R + \epsilon^2/8\tilde{\beta} + 5\theta)n\}] \\ & + \sum_{\gamma=1}^{L_\eta} P_\gamma\{J_\gamma(u,v) \leq (R + 2\epsilon^2/8\tilde{\beta} + 10\theta)n\} \\ & + L_\eta^2[\exp\{-(\epsilon^2/8\tilde{\beta} + 5\theta)n\}]\end{aligned}$$

We now proceed to show that $\Delta \leq \exp\{-(\epsilon^2/8\tilde{\beta} + 3\theta)n\}$ for n greater than some finite number N_0.

Notice that there exists a number $N_1 < \infty$ such that

$$\begin{aligned}L_\eta[\exp - (\epsilon^2/8\tilde{\beta} + 5\theta)n] &= \frac{1}{3}\exp\{-(\epsilon^2/8\tilde{\beta} + 5\theta)n + \log_e (3L_\eta)\} \\ &\leq \frac{1}{3}\exp\{-(\epsilon^2/8\tilde{\beta} + 4\theta)n\} \text{ for } n \geq N_1\end{aligned}$$

Likewise, there exists a number $N_2 < \infty$ such that

$$L_\eta^2 \exp\{-(\epsilon^2/8\tilde{\beta} + 5\theta)n\} \leq \frac{1}{3}\exp\{-(\epsilon^2/8\tilde{\beta} + 4\theta)n\} \text{ for } n \geq N_2$$

Using Lemma 3 we see that there exists an $N_3 < \infty$ such that for $n \geq N_3$, the following string of inequalities is valid.

$$\sum_{\gamma=1}^{L_\eta} P_\gamma\{J_\gamma(u,v) \leq (R + 2\epsilon^2/8\tilde{\beta} + 10\theta)n\}$$

$$\leq \sum_{\gamma=1}^{L_\eta} P_\gamma\{J_\gamma(u,v) \leq (\hat{R} - \epsilon^2/2\tilde{\beta} - 21\theta + \epsilon^2/4\tilde{\beta} + 10\theta)n\}$$

$$\leq \sum_{\gamma=1}^{L_\eta} P_\gamma\{J_\gamma(u,v) \leq (\log_e |I + A_\gamma \hat{S} A'_\gamma/2\tilde{\beta}|^{1/2} - 10\theta - \epsilon^2/4\tilde{\beta})n\}$$

$$= \sum_{\gamma=1}^{L_\eta} P_\gamma\{J_\gamma(u,v) \leq n \log_e |I + A_\gamma \hat{S} A'_\gamma/\tilde{\beta}|^{1/2}$$

$$- n(\frac{1}{2}\log_e[|I + A_\gamma \hat{S} A'_\gamma/\tilde{\beta}|/|I + A_\gamma \hat{S} A'_\gamma/2\tilde{\beta}|]$$

$$+ 10\theta + \epsilon^2/4\tilde{\beta})\}$$

$$\leq L_\eta \exp\{-n[\frac{t}{2}\log_e(|I + A_\gamma \hat{S} A'_\gamma/\tilde{\beta}| \quad / |I + A_\gamma \hat{S} A'_\gamma/2\tilde{\beta}|)$$

$$+ \frac{1}{2}\log_e(|I + (1-t^2)A_\gamma \hat{S} A'_\gamma/\tilde{\beta}|/|I + A_\gamma \hat{S} A'_\gamma/\tilde{\beta}|)$$

$$+ 10t\theta + t\epsilon^2/4\tilde{\beta}]\} \qquad \text{for all} \quad 0 < t \leq 1$$

$$\leq L_\eta \exp\{-n[\frac{1}{4}\log_e(|I + A_\gamma \hat{S} A'_\gamma/\tilde{\beta}|/|I + A_\gamma \hat{S} A'_\gamma/2\tilde{\beta}|)$$

(cont'd.)

$$+\ \frac{1}{2}\log_e(\,|I+\tfrac{3}{4}A_\gamma\hat{S}A'_\gamma/\tilde{\beta}\,|/\,|I+A_\gamma\hat{S}A'_\gamma/\tilde{\beta}\,|$$

$$+\ 5\theta+\epsilon^2/8\tilde{\beta}]\}$$

$$\Rightarrow L_\eta \exp\{-n[\tfrac{1}{4}\log_e(\,|I+\tfrac{3}{4}A'_\gamma\hat{S}A'_\gamma/\tilde{\beta}\,|^2/\,|I+\tfrac{1}{2}A_\gamma\hat{S}A'_\gamma/\tilde{\beta}\,|\ |I+A_\gamma\hat{S}A'_\gamma/\tilde{\beta}\,|)$$

$$+\ 5\theta+\epsilon^2/8\tilde{\beta}]\}$$

$$\leq L_\eta \exp\{-n(\epsilon^2/8\tilde{\beta}+5\theta\}\leq \tfrac{1}{3}\exp\{-n(\epsilon^2/8\tilde{\beta}+4\theta)\}$$

Thus, for $n \geq \max(N_1, N_2, N_3) = N_4$,

$$\Delta[\,1-L_\eta P\{\|u\|^2\geq Mn\}] < \exp\{-n(\epsilon^2/8\tilde{\beta}+4\theta)\}$$

Applying Lemma 4 we can bound the bracketed term on the left as follows:

$$1-L_\eta P\{\|u\|^2\geq Mn\}\geq 1-L_\eta \exp\{-[\,M/(M-\gamma)-1 - \log_e M/(M-\gamma)]\tfrac{n}{2}\}$$

Since $\log_e x < x-1$ for $x \neq 1$, there exists a constant $\infty > \omega > 0$ such that

$$1-L_\eta P\{\|u\|^2\geq Mn\}\geq 1-L_\eta \exp\{-\omega n\}$$

Hence, there exists a constant $N_5 < \infty$ such that for $n \geq N_5$,

$$1 - \mathbb{P}_{\eta}\{\|u\|^2 \geqslant Mn\} \geqslant 1 - \exp\{-\tfrac{1}{2}\omega n\} > 0$$

We now have that for $n \geqslant \max(N_4, N_5) = N_6$,

$$\Delta < \exp\{-(\epsilon^2/8\tilde{\beta} + 4\theta)n\} \;/[1 - \exp\{-\tfrac{1}{2}\omega n\}]$$

$$= \exp\{-(\epsilon^2/8\tilde{\beta} + 4\theta)n + \log_e[1/(1 - \exp\{-\tfrac{1}{2}\omega n\})]\}$$

$$\leqslant \exp\{-(\epsilon^2/8\tilde{\beta} + 4\theta)n + \exp\{-\tfrac{1}{2}\omega n\}/(1 - \exp\{-\tfrac{1}{2}\omega n\})\}$$

Thus, there exists a finite number $N_7 \geqslant N_6$ such that for $n \geqslant N_7$,

$$\Delta < \exp\{-(\epsilon^2/8\tilde{\beta} + 3\theta)n\}$$

and hence Theorem 1B is proved for the case where $\mathcal{C}$ is a finite set.

If $\mathcal{C}$ is not finite, we can pick a finite subset $\mathcal{C}'$ of $\mathcal{C}$ such that given any number $\eta > 0$, for every $A_\gamma \in \mathcal{C}$ there is an $A_{\gamma'} \in \mathcal{C}'$ with the property that $\|A_\gamma - A_{\gamma'}\| \leqslant \eta$. This follows from the observation that a bounded subset of $\mathbb{R}^{p^2}$ is totally bounded. We now proceed to show that if

$$P_{\gamma'}\{B_i^c \mid u_1\} < \exp\{-n(\epsilon^2/8\tilde{\beta} + 3\theta)\} \quad \text{for all } n \geqslant N_7,$$
$$\text{all } i = 1,\dots,G \text{ and all } A_{\gamma'} \in \mathcal{C}',$$

then there exists a finite number N, $N \geqslant N_7$ such that

$$P_\gamma\{B_i^c \mid u_i\} < \exp\{-n(\epsilon^2/8\tilde{\beta} + \theta)\} \quad \text{for all } n \geq N,$$
$$\text{all } i = 1,\ldots,G \text{ and all } A_\gamma \in \mathcal{C}.$$

Define the set F by $F = \{v \mid \|v\|^2 \leq nC\}$, $C = 2a^2M + 2\tilde{\beta}p + 2k$, where k is a number that will be selected later. We see that for all $k > 0$,

$$P_\gamma\{B_i^c \mid u_i\} = P_\gamma\{(B_i^c \cap F) \cup (B_i^c \cap F^c) \mid u_i\}$$
$$\leq P_\gamma\{B_i^c \cap F \mid u_i\} + P_\gamma\{F^c \mid u_i\} \quad \text{for all } i = 1,\ldots,G$$
$$\text{and all } A_\gamma \in \mathcal{C}$$

Using Lemmas 5 and 6 we see that

$$P_\gamma\{B_i^c \mid u_i\} \leq P_{\gamma'}\{B_i^c \cap F \mid u_i\} \exp\{n(\sqrt{MC} + aM)\|A_{\gamma'} - A_\gamma\|/\tilde{\beta}\}$$
$$+ \left[(1 + k/\tilde{\beta}^p) \exp\{-k/\tilde{\beta}^p\}\right]^{n/2}$$

$$\leq P_{\gamma'}\{B_i^c \mid u_i\} \exp\{n(\sqrt{MC} + aM)\|A_{\gamma'} - A_\gamma\|/\tilde{\beta}\}$$
$$+ \left[(1 + k/\tilde{\beta}^p) \exp\{-k/\tilde{\beta}^p\}\right]^{n/2}$$

Notice that $(1 + k/\tilde{\beta}^p) \exp\{-k/\tilde{\beta}^p\}$ can be made arbitrarily small by choosing k large enough. Choose k so that this term is less than $\frac{1}{2}\exp\{-\epsilon^2/8\tilde{\beta} + \theta\}$ and choose $\eta \geq \|A_{\gamma'} - A_\gamma\|$ to satisfy $(\sqrt{MC} + aM)\eta/\tilde{\beta} < \theta$. Hence there exists an finite number N, such that $N > N_7$ and

$$P\{B_i^c \mid u_i\} \leq \exp\{-n(\epsilon^2/8\tilde{\beta} + 2\theta)\} + (\tfrac{1}{2})^n \exp\{-n(\epsilon^2/8\tilde{\beta} + \theta)\}$$

$$\leq \exp\left\{ - n(\epsilon^2/8\tilde{\beta} + \theta)\right\} \qquad \text{for all } n \geq N,\ i = 1,\ldots,G, \text{ and all } A_\gamma \in \mathcal{C}.$$

This completes the proof of Theorem 2A. We will now relate the maximum probability of error as in Theorem 2A to the minimum distance between the total images of the transmitted sequences through Theorem 3A which follows.

<u>Theorem 3A</u>

If there exists G disjoint (decoding) sets B_i, $\bigcup_{i=1}^{G} B_i = \mathbb{R}^{np}$, G transmitted sequences u_i satisfying $\| u_i \|^2 \leq Mn$, constants $\alpha > 0$ and $0 < N < \infty$, and noise covariance matrix $\tilde{\beta} I$ such that $P_\gamma \{ B_i^c | u_i \} < e^{-\alpha n}$ for all $n \geq N$, $i = 1,2,\ldots,G$ and for all $A_\gamma \in \mathcal{C}$, then given any positive number b there exists a number $N_o = N_o(b)$ satisfying $N \leq N_o < \infty$ such that

$$\hat{d} = \inf_{i \neq j} d_{ij} > \sqrt{8\alpha\tilde{\beta}n/(1 + b)} \qquad \text{for } n \geq N_o$$

(Recall that $d_{ij} = \inf \| z - w \|$ where the infimum is taken over all $z \in \bigcup_{A_\gamma \in \mathcal{C}} A_\gamma^n u_i$, $w \in \bigcup_{A_\gamma \in \mathcal{C}} A^n u_j$.)

Proof:

Let u_k be any one of the transmitted sequences in the statement of the theorem. Let u_c be any other such sequence. Given any positive number δ, we can choose points

$A^n_{\hat{\gamma}} u_c$, $A^n_{\tilde{\gamma}} u_k$ so that $\| A^n_{\hat{\gamma}} u_c - A^n_{\tilde{\gamma}} u_k \| < d_{k,c} + \delta$

Define the set $D_{c,k}$ as follows:

$$D_{c,k} = \{ v \mid \|v - A^n_{\hat{\gamma}} u_c\| \leq \|v - A^n_{\tilde{\gamma}} u_k\| \}$$

See figure 1 for a symbolic sketch of $D_{c,k}$. Notice that if $v \in D_{c,k}$, $p_{\hat{\gamma}}(v \mid u_c) \geq p_{\tilde{\gamma}}(v \mid u_k)$ since

$$p_\gamma(v \mid u) = \exp\{-\tfrac{1}{2\tilde{\beta}}(v - A^n_\gamma u)'(v - A^n_\gamma u)\} \,/(2\pi\tilde{\beta})^{p/2}$$

$$= \exp\{-\tfrac{1}{2\tilde{\beta}}\|v - A^n_\gamma u\|^2\} \,/(2\pi\tilde{\beta})^{p/2}$$

Therefore $P_{\hat{\gamma}}\{F \mid u_c\} \geq P_{\tilde{\gamma}}\{F \mid u_k\}$ if $F \subset D_{c,k}$. Thus we have that

$$P_{\tilde{\gamma}}\{D_{c,k} \mid u_k\} = P_{\tilde{\gamma}}\{D_{c,k} \cap B_k \mid u_k\} + P_{\tilde{\gamma}}\{D_{c,k} \cap B_k^c \mid u_k\}$$

$$\leq P_{\hat{\gamma}}\{D_{c,k} \cap B_k \mid u_c\} + P_{\tilde{\gamma}}\{D_{c,k} \cap B_k^c \mid u_k\}$$

$$\leq P_{\hat{\gamma}}\{B_k \mid u_c\} + P_{\tilde{\gamma}}\{B_k^c \mid u_k\}$$

$$\leq P_{\hat{\gamma}}\{B_c^c \mid u_c\} + P_{\tilde{\gamma}}\{B_k^c \mid u_k\}$$

$$\leq 2 e^{-\alpha n}$$

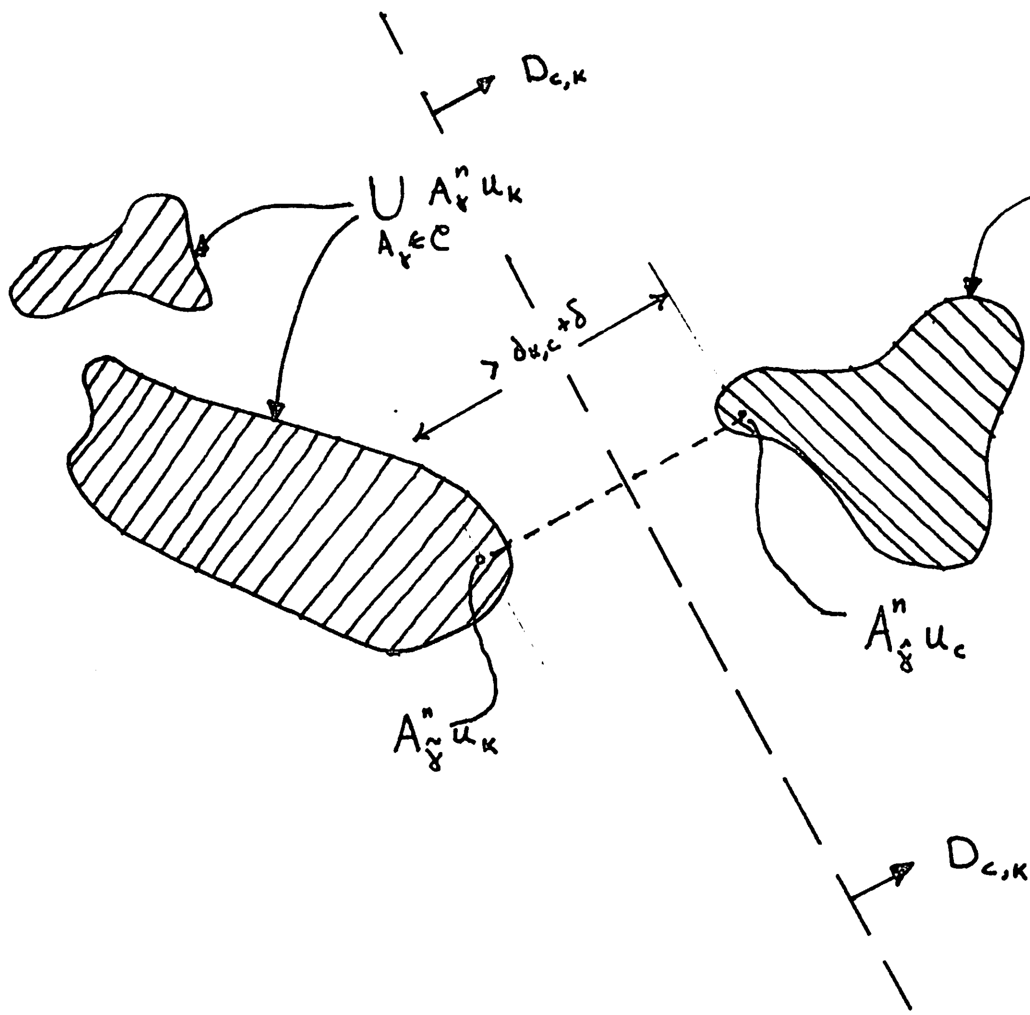

Figure 1 : A symbolic sketch of $D_{c,k}$.

Since the noise is sph erically symmetric, by the definition of $D_{c,k}$ we have that

$$P_{\tilde{\gamma}}\{D_{c,k} \mid u_k\} = \int\limits_{w > (d_{k,c}+\delta)/2} \exp\{-w^2/2\tilde{\beta}\}/(2\pi\tilde{\beta})^{\frac{1}{2}}\, dw$$

$$= \int\limits_{w > (d_{k,c}+\delta)/2\sqrt{\tilde{\beta}}} \exp\{-w^2/2\}/(2\pi)^{\frac{1}{2}}\, dw\,, \quad w \text{ is a scalar variable.}$$

So, we have that

$$\int\limits_{w > (d_{k,c}+\delta)/2\sqrt{\tilde{\beta}}} \exp\{-\tfrac{1}{2}w^2\}/(2\pi)^{\frac{1}{2}}\, dw \leqslant 2\exp\{-\alpha n\}$$

From this expression we can see that as n increases, $d_{k,c}$ must become unbounded. Now, since $\lim\limits_{w \to \infty} w\, e^{-bw} = 0$, for any scalar $b > 0$, we can find a finite positive number $N_o = N_o(b)$, $N_o \geqslant N$ such that for $n \geqslant N_o$,

$$\int\limits_{w > (d_{k,c}+\delta)/2\sqrt{\tilde{\beta}}} \exp\{-\tfrac{1}{2}w^2\}/(2\pi)^{\frac{1}{2}}\, dw \geqslant \int\limits_{w > (d_{k,c}+\delta)/2\sqrt{\tilde{\beta}}} 2\, w(1+b)\exp\{-\tfrac{1}{2}(1+b)w^2\}dw$$

$$= 2\exp\{-(1+b)(d_{k,c}+\delta)^2/8\tilde{\beta}\}$$

Hence,

$$(1+b)(d_{k,c}+\delta)^2/8\tilde{\beta} > \alpha n \quad \text{for } n \geqslant N_o$$

and

$$d_{k,c} > \sqrt{8\alpha\tilde{\beta}n/(1+b)} - \delta$$

Since N_o does not depend on δ, k, c and these numbers are arbitrarily chosen,

$$d_{k,c} > \sqrt{8\alpha\tilde{\beta}\, n/(1 + b)} \quad \text{for all } k \neq c$$

and hence, $\hat{d} > \sqrt{8\alpha\tilde{\beta}\, n/(1 + b)}$

This completes the proof of Theorem 3A.

Proof of Theorem 1A

We now combine Theorems 2A and 3A to prove Theorem 1A. That is, given any number R satisfying $0 < R < \underline{C}_\epsilon$, we show that R is an attainable ϵ- rate of transmission for $\mathcal{C}$. If $0 < R < \underline{C}_\epsilon$, by Theorem 2A there exist $G = [e^{Rn}]$ transmitted sequences $u_i \in U_M^p$ of length n, G disjoint (decoding) sets B_i, $\bigcup_{i=1}^{G} B_i = \mathbb{R}^{np}$ and finite numbers $\tilde{\beta}$, N, θ such that $P_\gamma\{B_i^c \mid u_i\} \leq \exp\{-(\epsilon^2/8\tilde{\beta} + \theta)n\}$ for all $i = 1, \ldots, G$, $n \geq N$ and all $A_\gamma \in \mathcal{C}$.

Using Theorem 3A with these B_i's, u_i's, and $\tilde{\beta}$, and choosing $\alpha = \epsilon^2/8\tilde{\beta} + \theta$, $b = 8\tilde{\beta}\theta/\epsilon^2$ we have that there exists a finite number N_o, $N_o \geq N$, such that

$$\hat{d} = \inf_{i \neq j} d_{i,j} > \sqrt{8\alpha\tilde{\beta}\, n/(1 + b)}$$

$$= \sqrt{8(\epsilon^2/8\tilde{\beta} + \theta)\tilde{\beta}\, n/(1 + 8\tilde{\beta}\theta/\epsilon^2)}$$

$$= \sqrt{\epsilon^2 n} = \epsilon\sqrt{n}$$

Hence, R is an attainable ϵ- rate of transmission for $\mathcal{C}$, and Theorem 1A is proven.

B. The Discrete Infinite Dimensional Case

The results of this section depend critically on Theorem 1A. In fact, we will show that the infinite dimensional model can be approximated arbitrarily closely by the finite dimensional model because of the assumptions imposed on the class $\mathcal{C}$ and the form of the model.

In order to state the results precisely, some further notation must be introduced. Let S denote an infinite dimensional covariance matrix and define the set $\mathcal{S}_M^\infty$ by $\mathcal{S}_M^\infty = \{ S \mid \text{trace } S \leq M, \ S \text{ an infinite dimensional covariance matrix} \}$. Since $S \in \mathcal{S}_M^\infty$, $A_\gamma S A_\gamma'$ is a symmetric positive semi-definite matrix and we will let $\lambda_1^\gamma \geq \lambda_2^\gamma \geq \ldots \geq 0$ denote the eigenvalues of $A_\gamma S A_\gamma'$.

For any matrix $B = \{b_{ij}\}$ and positive integer k, let ${}^kB = \{{}^kb_{ij}\}$ be the matrix given by ${}^kb_{ij} = b_{ij}$ if $i \leq k$, $j \leq k$ and ${}^kb_{ij} = 0$ otherwise. For $S \in \mathcal{S}_M^\infty$ and $A_\gamma \in \mathcal{C}$ we denote the eigenvalues of ${}^kA_\gamma \, {}^kS \, {}^kA_\gamma'$ by $\lambda_1^{\gamma,k} \geq \lambda_2^{\gamma,k} \geq \ldots \geq 0$. Note that ${}^kS \in \mathcal{S}_M^\infty$. To show that the infinite dimensional model can be approximated arbitrarily closely by a finite dimensional model we need the following lemma.

<u>Lemma 7</u>

Let $S \in \mathcal{S}_M^\infty$ be a fixed diagonal matrix i.e. if $S = \{s_{ij}\}$ then $s_{ij} = 0$ for $i \neq j$. Then for each

$\theta > 0$, $\sigma^2 > 0$, there exists $k_o = k_o(\theta, \sigma^2) < \infty$ such that for all $k \geq k_o$ and for all $A_\gamma \in \mathcal{C}$

$$\left| \tfrac{1}{2} \sum_{i=1}^{\infty} \log_e(1 + \lambda_1^\gamma/\sigma^2) - \tfrac{1}{2} \sum_{i=1}^{\infty} \log_e(1 + \lambda_1^{\gamma,k}/\sigma^2) \right| \leq \theta$$

Proof: This is just a restatement of Lemma 11 of [3].

We are now prepared to state and prove the main result of this section, Theorem 1B.

<u>Theorem 1B</u>

For the class $\mathcal{C}$ of channels satisfying the conditions of model B,

$$C_\epsilon \geq \underline{C}_\epsilon = \sup_{\beta > 0} \left\{ \sup_{S \in \mathcal{S}_M^\infty} \inf_{A_\gamma \in \mathcal{C}} \tfrac{1}{2} \sum_{i=1}^{\infty} \log_e[(1 + \lambda_1^\gamma/2\beta) - \epsilon^2/2\beta \right\}$$

Proof:

If $\underline{C}_\epsilon$ is less than zero, there is nothing to prove since $C_\epsilon \geq 0$. Hence, assume $\underline{C}_\epsilon > 0$ in what follows.

Choose any number R satisfying $0 < R < \underline{C}_\epsilon$. By the definition of $\underline{C}_\epsilon$, there exist $\theta > 0$ and $\tilde{\beta} > 0$ such that

$$\sup_{S \in \mathcal{S}_M^\infty} \inf_{A_\gamma \in \mathcal{C}} \tfrac{1}{2} \sum_{i=1}^{\infty} \log_e (1 + \lambda_1^\gamma/2\tilde{\beta}) - \epsilon^2/2\tilde{\beta}$$

$$> R + 3\theta$$

From the definition of sup we see that there exists $\underline{S} \in \mathcal{S}_M^\infty$ such that if $\lambda_1^\gamma \geq \lambda_2^\gamma \geq \ldots \geq 0$ are the eigenvalues of $A_\gamma \underline{S} A_\gamma'$ then

$$\frac{1}{2} \sum_{i=1}^{\infty} \log_e \left[(1 + \lambda_i^\gamma / 2\tilde{\beta}) - \epsilon^2 / 2\tilde{\beta} > R + 2\theta \right.$$

for all $A_\gamma \in \mathcal{C}$.

We can choose an appropriate basis for the transmitted and received spaces so that $\underline{S}$ is diagonal. We note that the matrix representation of a channel relative to this new basis may be different, but this does not change the value of the eigenvalues of $A_\gamma \underline{S} A_\gamma'$. Using Lemma 7 we see that there exists a k_o such that for $k \geq k_o$, the following inequalities are valid.

$$\frac{1}{2} \sum_{i=1} \log_e \; (1 + \lambda_i^{\gamma,\kappa} / 2\tilde{\beta}) - \epsilon^2 / 2\tilde{\beta} > R + \theta \quad \text{for all} \quad A_\gamma \in \mathcal{C}$$

and hence

$$\sup_{S \in \mathcal{S}_M^\infty} \; \inf_{A_\gamma \in \mathcal{C}} \; \frac{1}{2} \sum_{i=1}^{K} \log_e \; (1 + \lambda_i^{\gamma K} / 2\tilde{\beta}) > R + \theta$$

The $\lambda_i^{\gamma,\kappa}$ are the eigenvalues of the matrix ${}^k A_\gamma {}^k S {}^k A_\gamma'$ which is effectively a k - dimensional matrix. If we consider the $k \times k$ submatrices of ${}^{k}A_{\gamma}$, $A_{\gamma} \in \mathcal{C}$ formed from the first k rows and columns, we see that this class of submatrices, call it $\mathcal{C}_{k}$, is bounded. Likewise if we consider the corresponding $k \times k$ submatrices formed from the ${}^{k}S$'s $S \in \mathcal{S}_{M}^{\infty}$ we see that this class of submatrices is just $\mathcal{S}_{M}^{k}$. Thus, from Theorem 1A, we can see that R is an attainable ϵ - rate of transmission for the class $\mathcal{C}_{k}$. This means that for the class of channels $\mathcal{C}_{k}$ there exists a sequence n with $\lim_{\ell \to \infty} n_{\ell} = \infty$, and $G(n_{\ell})$ transmitted sequences $u_j \in U_M^k$ of length n_{ℓ} such that $\hat{d} = \inf_{i \neq j} d_{i,j} \geq \epsilon\sqrt{n_{\ell}}$ and $\lim_{\ell \to \infty} \frac{1}{n_{\ell}} G(n_{\ell}) = R$.

With this fact, let us now show that R is an attainable ϵ- rate of transmission for the infinite dimensional channel $\mathcal{C}$. Suppose that $u_i = (x_{i1}, x_{i2}, \ldots, x_{in_{\ell}})$, $i = 1,2,\ldots,G(n_{\ell})$ is a set of transmitted sequences which yield the ϵ- attainable rate R for $\mathcal{C}_{k}$. Choose $\hat{x}_{ij} = (x_{ij}^{1}, x_{ij}^{2}, x_{ij}^{3}, \ldots, x_{ij}^{k}, 0, 0, \ldots)$ as the transmitted vectors for the class of channels $\mathcal{C}$ and let $\hat{u}_i = (\hat{x}_{i1}, \hat{x}_{i2}, \ldots, \hat{x}_{in})$, $i = 1,2,\ldots G(n_{\ell})$ be the transmitted sequences for the class of channels $\mathcal{C}$. Notice that $\hat{u}_i \in U_M^{\infty}$ and that $A_{\gamma}\hat{x}_{ij} = {}^{k}A_{\gamma}\hat{x}_{ij}$. Thus, the distances $d_{i,j}$ for $\mathcal{C}_{k}$ are the same as the distances $d_{i,j}$ for $\mathcal{C}$ and hence $\hat{d} > \epsilon\sqrt{n_{\ell}}$. It follows that R is an ϵ- attainable rate for $\mathcal{C}$. Since R is arbitrary except that $0 < R < \underline{C}_{\epsilon}$, we see that $C_{\epsilon} \geq \underline{C}_{\epsilon}$.

C. The Continuous Time Case

Just as we obtained results for the discrete infinite dimensional model by approximating it by a finite dimensional model, we can obtain results for the continuous time model by approximating it with a discrete infinite dimensional model. In order to do this, some further notation is necessary.

Let L_p denote $L_p(-\infty,\infty)$ for $1 \leq p \leq \infty$ where $L_p(a,b)$ is the space of complex valued functions such that the p-th power of its magnitude is Lebesque integrable on the interval (a,b). $L_p(T)$ will denote $L_p(-T,T)$, for $1 \leq p \leq \infty$. If $f \in L_p$ or $L_p(T)$, then $\| f \|_p$ denotes the norm of f in that space. If $f,g \in L_2$ or $L_2(T)$, their inner product is written as (f,g). An <u>operator</u> on a space X is a continuous linear transformation of X into itself. P_T is to denote the projection operator on L_p, $1 \leq p \leq \infty$ defined by

$$\begin{aligned}(P_T x)(t) &= x(t) \ , \quad |t| \leq T \\ &= 0 \quad\ \ , \quad |t| > T \quad \text{for all } x \in L_p.\end{aligned}$$

If $f \in L_2(T) \cap L_1(T)$ then for each $T < \infty$, a compact operator F_T on $L_2(T)$ is given by

$$(F_T x)(t) = \int_{-T}^{T} f(t-\tau)x(\tau)\, d\tau \ , \quad -T \leq t \leq T$$

We also have that f defines an operator F on L_2 given by the convolution

$$(Fx)(t) = \int_{-\infty}^{\infty} f(t-\tau)x(\tau)\,d\tau \quad , -\infty < t < \infty \quad (*)$$

With a slight abuse of notation we identify the operators F_T and $P_T F P_T$. If f has finite memory (i.e. $f(t) = 0$ for $|t| \geq \delta$) then

$$P_T F = P_T F P_{T+\delta} \quad \text{and} \quad F P_T = P_{T+\delta} F P_T \quad .$$

If A is an operator on $L_2(T)$ or L_2 then A^* will denote its adjoint which is defined by

$$(x \, , \, A\, y) = (A^* x \, , \, y) \quad \text{for all } x, y \in L_2(T) \text{ or } L_2.$$

If A is a compact symmetric operator, its trace Tr(A), is defined if the sum of the eigenvalues of A converges, and is equal to that sum.

Throughout this section, $\mathcal{C}$ will denote the class of channel operators which satisfy the assumptions of model C. $\mathcal{C}_\delta$ will denote the class of finite memory operators obtained from $\mathcal{C}$ by truncating the kernels of the operators in $\mathcal{C}$, i.e.

$$\mathcal{C}_\delta = \{ h_{\gamma,\delta} \mid h_{\gamma,\delta}(t) = h_\gamma(t) \text{ for } |t| \leq \delta \; , \; h_{\gamma,\delta}(t) = 0 \text{ otherwise}, \; h_\gamma \in \mathcal{C} \}$$

$\mathcal{C}_\delta$ also satisfies the conditions for model C. We will let s denote the covariance function of a stationary stochastic process with the additional property that $s \in L_1$. By known properties of positive semi - definite functions (see e.g. [8], Th. 9) it follows that $\tilde{s} \in L_1 \cap L_\infty$ and hence there exists a number Δ such that $|\tilde{s}(\nu)| < \Delta$[1]. $\mathcal{A}^c_M$ will denote the set of such covariance functions which satisfy

$$\int_{-\infty}^{\infty} \tilde{s}(\nu)\, d\nu = M$$

For each $T<\infty$, $h_{\gamma,\delta} \in \mathcal{C}_\delta$ and s as above, let us define the operators $H_{\gamma,T} = P_T H_\gamma P_T$, $S_T = P_T S P_T$ where H_γ and S are defined in terms of $h_{\gamma,\delta}$ and s as in equation (*). $H_{\gamma,T}$ is then a compact operator and so the positive semidefinite operator

$$W_{\gamma,T} = P_T H_\gamma P_T S P_T H_\gamma^* P_T = H_{\gamma,T} S H^*_{\gamma,T}$$

is also compact. Finally we define

$$Q_{\gamma,T} = P_T H_\gamma S H_\gamma^* P_T = P_T Q_\gamma P_T$$

$Q_{\gamma,T}$ is compact by virtue of the fact that $Q_\gamma = H_\gamma S H_\gamma^*$ is a convolution operator with kernel in L_2. Then, $q_\gamma \in L_1 \cap L_2$ and its Fourier transform is

$$\tilde{q}_\gamma(\nu) = |\tilde{h}_{\gamma\delta}(\nu)|^2\, \tilde{s}(\nu)$$

[1] If $f \in L_1$, its Fourier transform denoted by $\tilde{f}$ is given by $\tilde{f}(\nu) = \int_{-\infty}^{\infty} f(t)\, e^{i\pi\nu t} dt$.

which also belongs to L_1 since $\tilde{s}(\nu)$ is bounded.

We will denote the eigenvalues of $Q_{\gamma,T}$, $W_{\gamma,T}$ by $q_{\gamma 1}(T) \geq q_{\gamma 2}(T) \geq \dots \geq 0$ and $w_{\gamma 1}(T) \geq w_{\gamma 2}(T) \geq \dots \geq 0$ respectively. For convenience we now state three facts which will be essential in proving the main result of this section, Theorem 1C.

Lemma 8

Let g be a continuous monotone-increasing real valued function on the real number which satisfies $g(0) = 0$, $g(x) \geq k_1 x$ in a neighborhood of 0 for some $0 < k_1 < \infty$, and $|g(x) - g(y)| \leq k|x - y|$ for all $x, y \in \mathbb{R}$ and some $0 < k < \infty$. Define the functions $q_T : \mathcal{C}_\delta \to \mathbb{R}$ and $w_T : \mathcal{C}_\delta \to \mathbb{R}$ by

$$q_T(h_{\gamma,\delta}) = \frac{1}{2T} \sum_{i=1}^{\infty} g(q_{\gamma i}(T))$$

and $$w_T(h_{\gamma,\delta}) = \frac{1}{2T} \sum_{i=1}^{\infty} g(w_{\gamma i}(T))$$

Then, i) $$\lim_{T \to \infty} q_T(h_{\gamma,\delta}) = \int_{-\infty}^{\infty} g(|\tilde{h}_{\gamma,\delta}(\nu)|^2 \tilde{s}(\nu)) d\nu$$

uniformly for $h_{\gamma,\delta} \in \mathcal{C}_\delta$

ii) $$\lim_{T \to \infty} (w_T(h_{\gamma,\delta}) - q_T(h_{\gamma,\delta})) = 0$$

uniformly for $h_{\gamma,\delta} \in \mathcal{C}_\delta$

iii) $\lim_{T \to \infty} w_T(h_{\gamma,\delta}) = \int_{-\infty}^{\infty} g(\,|\tilde{h}_{\gamma,\delta}(\nu)|^2\, \tilde{s}(\nu))\,d\nu$

uniformly for $h_{\gamma,\delta} \in \mathcal{C}_\delta$

Proof:

This lemma follows directly from Lemmas 4 and 5 of [4].

Lemma 9

Let $\mathcal{C}_\delta$ be defined as before and let β be any positive number. Then,

$$\lim_{T \to \infty} \frac{1}{2T} \sum_{i=1}^{\infty} \log_e(1 + w_{\gamma i}(T)/\beta)$$

$$= \int_{-\infty}^{\infty} \log_e(1 + |\tilde{h}_{\gamma,\delta}(\nu)|^2 \tilde{s}(\nu)/\beta)\,d\nu$$

uniformly over $\mathcal{C}_\delta$.

Proof:

Note that the function $g(x) = \log_e(1 + x/\beta)$ satisfies the assumptions of Lemma 8.

Lemma 10

Let $\mathcal{C}$ satisfy the condtions for model C. Then given any positive number β, any $\tilde{s}(\nu) \in \mathcal{S}_M^c$ and any positive number θ there exists a finite number δ such that

i) $\left| \int_{-\infty}^{\infty} \log_e(1 + |\tilde{h}_{\gamma,\delta}(\nu)|^2 \tilde{s}(\nu)/2\beta) d\nu \right.$

$$\left. - \int_{-\infty}^{\infty} \log_e(1 + |\tilde{h}_{\gamma}(\nu)|^2 \tilde{s}(\nu)/2\beta) d\nu \right| < \theta$$

for all $h_\gamma \in \mathcal{C}$

ii) $\|\tilde{h}_\gamma - \tilde{h}_{\gamma,\delta}\|_2 = \|h_\gamma - h_{\gamma,\delta}\|_2 < \frac{\theta}{\sqrt{2M}}$ for all $h_\gamma \in \mathcal{C}$

Proof:

Since $s \in L_\infty$, there exists a constant Δ such that $|\tilde{s}(\nu)| < \Delta < \infty$ for all ν. Recall that $\|h_\gamma\|_2 \leq a$ for all $h_\gamma \in \mathcal{C}$ and that $\mathcal{C}$ is conditionally compact. We therefore can find a finite positive number δ such that

$$\|\tilde{h}_\gamma - \tilde{h}_{\gamma,\delta}\|_2 = \|h_\gamma - h_{\gamma,\delta}\|_2 < \min(\theta\beta/3a\Delta \ , \ \theta/\sqrt{2M})$$

For each ν define $\underline{\tilde{h}}_\gamma(\nu) = \max(\tilde{h}_\gamma(\nu), \tilde{h}_{\gamma,\delta}(\nu))$. Thus,

$$\|\underline{\tilde{h}}_\gamma - \tilde{h}_\gamma\|_2 \leq \|\tilde{h}_{\gamma,\delta} - \tilde{h}_\gamma\|_2 \ , \ \|\underline{\tilde{h}}_\gamma - \tilde{h}_{\gamma,\delta}\|_2 \leq \|\tilde{h}_\gamma - \tilde{h}_{\gamma,\delta}\|_2$$

and $\|\underline{\tilde{h}}_\gamma\|_2 < 2a$ for all $h_\gamma \in \mathcal{C}$

Therefore

$$\left| \int_{-\infty}^{\infty} \log_e(1 + |\tilde{h}_{\gamma,\delta}(\nu)|^2 \tilde{s}(\nu)/2\beta)\,d\nu - \int_{-\infty}^{\infty} \log_e(1 + |\tilde{h}_{\gamma}(\nu)|^2 \tilde{s}(\nu)/2\beta)\,d\nu \right|$$

$$\leq \left| \int_{-\infty}^{\infty} \log_e(1 + |\underline{\tilde{h}}_{\gamma}(\nu)|^2 \tilde{s}(\nu)/2\beta)\,d\nu - \int_{-\infty}^{\infty} \log_e(1 + |\tilde{h}_{\gamma,\delta}(\nu)|^2 \tilde{s}(\nu)/2\beta)\,d\nu \right|$$

$$+ \left| \int_{-\infty}^{\infty} \log_e(1 + |\underline{\tilde{h}}_{\gamma}(\nu)|^2 \tilde{s}(\nu)/2\beta)\,d\nu - \int_{-\infty}^{\infty} \log_e(1 + |\tilde{h}_{\gamma}(\nu)|^2 \tilde{s}(\nu)/2\beta)\,d\nu \right|$$

Let us examine the last term on the right.

$$\left| \int_{-\infty}^{\infty} \log_e(1 + |\underline{\tilde{h}}_{\gamma}(\nu)|^2 \tilde{s}(\nu)/2\beta)\,d\nu - \int_{-\infty}^{\infty} \log_e(1 + |\tilde{h}_{\gamma}(\nu)|^2 \tilde{s}(\nu)/2\beta)\,d\nu \right|$$

$$\leq \left| \int_{-\infty}^{\infty} \log_e\left(1 + \left[|\underline{\tilde{h}}_{\gamma}(\nu)|^2 - |\tilde{h}_{\gamma}(\nu)|^2 \right] \tilde{s}(\nu)/\left[2\beta + |\tilde{h}_{\gamma}(\nu)|^2 \tilde{s}(\nu)\right] \right) d\nu \right|$$

$$\leq \left| \int_{-\infty}^{\infty} \left[\tilde{s}(\nu) \left[|\underline{\tilde{h}}_{\gamma}(\nu)|^2 - |\tilde{h}_{\gamma}(\nu)|^2 \right] / \left[2\beta + |\tilde{h}_{\gamma}(\nu)|^2 \tilde{s}(\nu)\right] \right] d\nu \right|$$

$$\leq \frac{\Delta}{2\beta} \left[\|\underline{\tilde{h}}_{\gamma}\|_2^2 - \|\tilde{h}_{\gamma}\|_2^2 \right]$$

$$\leq \frac{\Delta}{2\beta}\left[\|\underline{\tilde{h}}_\gamma\|_2 + \|\tilde{h}_\gamma\|_2\right]\left[\|\underline{\tilde{h}}_\gamma\|_2 - \|\tilde{h}_\gamma\|_2\right]$$

$$\leq \quad 3a\Delta\left[\|\underline{\tilde{h}}_\gamma\|_2 - \|\tilde{h}_\gamma\|_2\right]/2\beta$$

$$\leq \quad \frac{3a\Delta}{2\beta}\|\underline{\tilde{h}}_\gamma - \tilde{h}_\gamma\|_2 \leq \frac{3a\Delta}{2\beta}\|\tilde{h}_{\gamma,\delta} - \tilde{h}_\gamma\|_2 < \theta/2$$

The lemma follows from duplicating such arguments in bounding the other terms.

We can now state and prove Theorem 1C

<u>Theorem 1C</u>

Let $\mathcal{C}$ be a class of channels satisfying the conditions for model C. Then,

$$C_\epsilon \geq \underline{C}_\epsilon = \sup_{\beta>0}\left\{\sup_{s\in\mathcal{S}_M^c}\ \inf_{h_\gamma\in\mathcal{C}}\int_{-\infty}^{\infty}\log_e\left(1+|\tilde{h}_\gamma(\nu)|^2\tilde{s}(\nu)/2\beta\right)d\nu - \epsilon^2/2\beta\right\}$$

Proof:

As before, we need only prove Theorem 1C for $\underline{C}_\epsilon>0$, hence assume that $\underline{C}_\epsilon>0$ in what follows.

For any R, $0<R<\underline{C}_\epsilon$, there exists a $\tilde{\beta}>0$ and a $\theta>0$ such that

$$\sup_{s \in \mathcal{S}_M^c} \inf_{h_\gamma \in \mathcal{C}} \int_{-\infty}^{\infty} \log_e \left(1 + |\tilde{h}_\gamma(\nu)|^2 \tilde{s}(\nu)/2\tilde{\beta}\right) d\nu$$

$$- (\epsilon + 2\theta)^2/2\tilde{\beta} > R + 4\theta$$

Furthermore, we can always find an $\tilde{s}_o \in \mathcal{S}_M^c$ such that

$$\int_{-\infty}^{\infty} \log_e \left(1 + |\tilde{h}_\gamma(\nu)|^2 \tilde{s}_o(\nu)/2\tilde{\beta}\right) d\nu - (\epsilon + 2\theta)^2/2\tilde{\beta}$$

$$> R + 3\theta$$

for all $h_\gamma \in \mathcal{C}$

For this choice of $\tilde{\beta}, \theta$, and $\tilde{s}_o$, we see from Lemma 10, that there exists a finite number δ such that

$\| \tilde{h}_{\gamma,\delta} - \tilde{h}_\gamma \|_2 = \|h_{\gamma,\delta} - h_\gamma\|_2 < \theta/\sqrt{2M}$ for all $h_\gamma \in \mathcal{C}$.

$$\int_{-\infty}^{\infty} \log_e \left(1 + |\tilde{h}_{\gamma,\delta}(\nu)|^2 \tilde{s}_o(\nu)/2\tilde{\beta}\right) d\nu$$

$$- (\epsilon + 2\theta)^2/2\tilde{\beta} > R + 2\theta$$

for all $h_{\gamma,\delta} \in \mathcal{C}_\delta$.

Using Lemma 9 we can see that there exists a $\hat{T}$, $0 < \hat{T} < \infty$ such that

$$\left| \int_{-\infty}^{\infty} \log_e (1 + |\tilde{h}_{\gamma,\delta}(\nu)|^2 \tilde{s}_o(\nu)/\tilde{\beta}\,)\, d\nu \right.$$

$$\left. - \frac{1}{2(\hat{T}+\delta)} \sum_{i=1}^{\infty} \log_e (1 + w_{\gamma,i}(\hat{T})/\tilde{\beta}) \right| < \theta$$

holds for all $h_{\gamma,\delta} \in \mathcal{C}_{\delta}$.

Using both these inequalities, it is obvious that

$$\tfrac{1}{2} \sum_{i=1}^{\infty} \log_e \,(1 + w_{\gamma,i}(\hat{T})/2\tilde{\beta}]$$

$$- \;(\hat{T} + \delta)(\epsilon + 2\theta)^2/2\tilde{\beta}$$

$$> \;(\hat{T} + \delta)R \qquad\qquad \text{for all} \;\; h_{\gamma,\delta} \in \mathcal{C}_{\delta}$$

and therefore that

$$\sup_{s \in \mathcal{S}_M^{\epsilon}} \; \inf_{h_{\gamma,\delta} \in \mathcal{C}_{\delta}} \; \tfrac{1}{2} \sum_{i=1}^{\infty} \log_e [\, 1 + w_{\gamma,i}(\hat{T}))/2\tilde{\beta}\,]$$

$$- \;(\hat{T} + \delta)(\epsilon + 2\theta)^2/2\tilde{\beta} > (\hat{T} + \delta)R \qquad (\dagger)$$

We now show how we can relate the continuous time problem to the infinite dimensional discrete time problem. Let $\{ \varphi_i \mid 1 \leq i < \infty \}$ be a complete orthonormal basis in

$L_2(T)$. Relative to this basis, the operators $P_{\hat{T}} H_\gamma P_{\hat{T}}$, $P_{\hat{T}} S P_{\hat{T}}$ and $P_{\hat{T}} H_\gamma^* P_{\hat{T}}$ have a representation as infinite dimensional matrices. (Recall that H_γ is the convolution operator with kernel $h_{\gamma,\delta}$). We denote these matrices by $H_{\gamma,\hat{T}}$, $S_{\hat{T}}$ and $H^*_{\gamma,\hat{T}}$ respectively.

We note that $H^*_{\gamma,\hat{T}}$ is the transpose of $H_{\gamma,\hat{T}}$ and the collection $\mathcal{C}_{\hat{T}}$ of matrices $H_{\gamma,\hat{T}}$ form a conditionally compact set in the Hilbert-Schmidt norm. Notice also that $w_{\gamma,i}(\hat{T})$ are the eigenvalues of $H_{\gamma,\hat{T}} S_{\hat{T}} H^*_{\gamma,\hat{T}}$. Furthermore, $S_{\hat{T}}$ is an infinite dimensional covariance matrix and the trace of $S_{\hat{T}}$ is less than or equal to $2M\hat{T}$. Hence, $S_{\hat{T}} \in \mathcal{S}^\infty_{2M\hat{T}}$.

From inequality (†) and Theorem 1B we see that $R(\hat{T}+\delta)$ is an $(\epsilon + 2\theta)\sqrt{(\hat{T}+\delta)}$ attainable rate for $\mathcal{C}_{\hat{T}}$. Hence, there exists a sequence n_k , $\lim_{k\to\infty} n_k = \infty$, and $G(n_k)$ transmitted signals $u_i \in U^\infty_{2M\hat{T}}$ of length n_k such that $\hat{d} = \inf_{i\neq j} d_{i,j} \geq (\epsilon + 2\theta)\sqrt{n_k(T+\delta)}$ and

$$\lim_{k\to\infty} \frac{1}{(\hat{T}+\delta)n_k} \log_e G(n_k) = R .$$

We now proceed to show that R is an attainable $(\epsilon + 2\theta)$ - rate of transmission for $\mathcal{C}_\delta$. Let $u_j = (x_{j1}, x_{j2}, \ldots, x_{jn_k}) \in U^\infty_{2M\hat{T}}$ be the transmitted sequences yielding the $R(\hat{T}+\delta)$ rate for $\mathcal{C}_{\hat{T}}$. Corresponding to each vector $x_{ij} = (x^1_{ij}, x^2_{ij}, \ldots, \ldots)$ define the function $x_{ij}(t)$, $-\hat{T} \leq t \leq \hat{T}$ by

$$x_{ij}(t) = \sum_{k=1}^{\infty} x_{ij}^{k} \mathcal{S}_k(t) \qquad i = 1,2,\ldots,G(n_k)$$
$$j \quad 1,2,\ldots,n_k$$

For $i = 1,2,\ldots G(n_k)$ define the function $\hat{u}_i(t)$, $0 \leq t \leq 2n_k(\hat{T} + \delta)$ by

$$\hat{u}_i(t) = x_{ij}(t - \hat{T} - 2(j-1)(\hat{T} + \delta) - \delta)$$

for $2(j-1)(\hat{T} + \delta) + \delta \leq t \leq 2j(\hat{T} + \delta) - \delta$,
$j = 1,2,\ldots,n_k$,

and $\hat{u}_i(t) = 0$ elsewhere.

Next define the functions $u_1(t),\ldots,u_{G(n_k)}(t)$ on the interval $-n_k(\hat{T} + \delta) \leq t \leq n_k(\hat{T} + \delta)$ by

$$u_i(t) = \hat{u}_i(t + n_k(\hat{T} + \delta))$$

Notice from the construction of the functions $u_i(t)$, that $u_i(t) \in \mathcal{U}_M^{n_k(\hat{T}+\delta)}$ for all $i = 1,2,\ldots,G(n_k)$. (See figure 2) Since $\mathcal{C}_\delta$ has memory δ , we see that $\hat{d} > (\epsilon + 2\theta)\sqrt{n_k(\hat{T} + \delta)}$ Thus, R is an $(\epsilon + 2\theta)$ -attainable rate for $\mathcal{C}_\delta$. Recall however that δ was selected so that $\| h_{\gamma,\delta} - h_\gamma \|_2 < \theta/\sqrt{2M}$. If we use the same transmitted signals $u_i(t)$ for $\mathcal{C}$, $\hat{d} = \inf_{i \neq j} d_{i,j} > \epsilon\sqrt{n_k(\hat{T} + \delta)}$. Hence R is an ϵ-rate for $\mathcal{C}$. Since R is arbitrary subject only to $0 < R < \underline{C}_\epsilon$, Theorem 1C is proven.

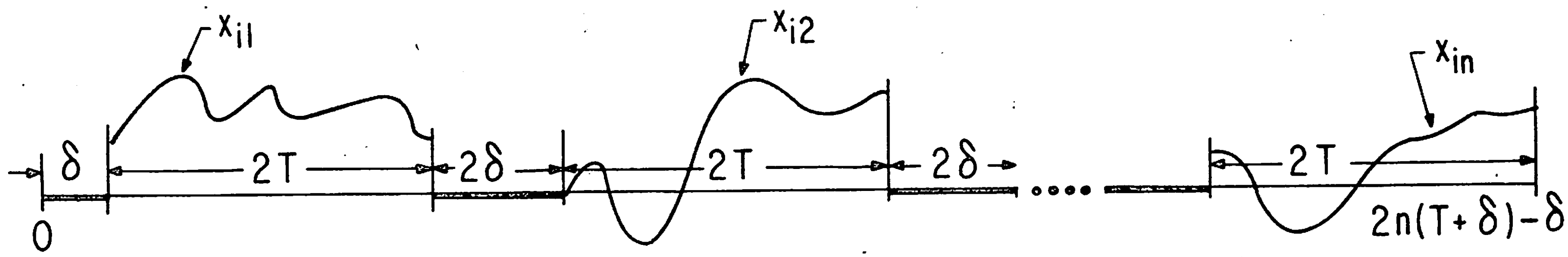

Fig. 2. $\hat{u}_i(t)$.

III. UPPER BOUNDS FOR C_ϵ

Upper bounds for C_ϵ have also been developed by Root [1] for the case where $\mathcal{C}$ consists of a single point. or is a ball. For these upper bounds, Root again used variations of packing arguments. The inapplicability of such arguments when $\mathcal{C}$ has a complex structure has been discussed in Section II. We will use a much simpler technique to obtain upper bounds. We relate the usual capacity theorems for Gaussian channels to the ϵ - capacity of such channels by a judicious choice of noise statistics.

Since the bounds for each of the three cases are derived using the same techniques, we will provide the details for only the continuous time case.

<u>Theorem 4A</u>

If the class $\mathcal{C}$ of channels satisfies the conditions of model A, then

$$C_\epsilon \leq \overline{C}_\epsilon = \sup_{S \in \mathcal{S}_M^p} \inf_{A_\gamma \in C} \inf_{W \in \mathcal{W}} \log_e \left[| W + A_\gamma S A_\gamma' |^{\frac{1}{2}} / | W |^{\frac{1}{2}} \right]$$

where $\mathcal{W} = \{ W \mid W \text{ is a } p \times p \text{ covariance matrix and } \operatorname{trace} W < \frac{1}{4}\epsilon^2 \}$

Theorem 4B

If $\mathcal{C}$ is a class of channels satisfying the conditions of model B, then

$$C_\epsilon \leq \overline{C}_\epsilon = \sup_{S \in \mathcal{S}_M^\infty} \inf_{A_\gamma \in \mathcal{C}} \inf_{\sum_{i=1}^\infty \sigma_i^2 < \epsilon^2/4} \frac{1}{2} \sum_{i=1}^\infty \log_e (1 + \lambda_i^\gamma / \sigma_i^2)$$

Theorem 4C

If $\mathcal{C}$ is a class of channels satisfying the conditions of model C and in addition , there exists a finite positive number δ such that $h_\gamma(t) = 0$ if $|t| > \delta$ for all $h_\gamma \in \mathcal{C}$ then,

$$C_\epsilon \leq \overline{C}_\epsilon = \sup_{\tilde{s} \in \mathcal{S}_M^c} \inf_{h_\gamma \in \mathcal{C}} \inf_{N(\nu) \in \mathcal{N}_\epsilon} \int_{-\infty}^{\infty} \log_e (1 + |\tilde{h}_\gamma(\nu)|^2 \tilde{s}(\nu) / N(\nu)) d\nu$$

where $\mathcal{N}_\epsilon = \{ N(\nu) \mid N(\nu)$ is the spectral density of a stationary Gaussian noise with $\int_{-\infty}^{\infty} N(\nu)\, d\nu < \epsilon^2/8 \}$

Proof:

Assume the contrary, that $\overline{C}_\epsilon < C_\epsilon$. Then there exists real numbers $\theta > 0$, $\Delta > 0$ and an $\underline{N}(\nu) \in \mathcal{N}_\epsilon$ such that

$$\int_{-\infty}^{\infty} \underline{N}(\nu)\, d\nu = \epsilon^2/8 - \Delta \quad \text{and}$$

$$C_\epsilon > 2\theta + \sup_{\tilde{s} \in \mathcal{S}_M^c} \inf_{h_\gamma \in \mathcal{C}} \int_{-\infty}^{\infty} \log_e (1 + |\tilde{h}_\gamma(\nu)|^2 \tilde{s}(\nu) / \underline{N}(\nu))\, d\nu$$

There also exists an $h_{\hat{\gamma}} \in \mathcal{C}$ such that

$$C_\epsilon > \theta + \sup_{\tilde{s} \in \Delta_M^c} \int_{-\infty}^{\infty} \log_e(1 + |\tilde{h}_{\hat{\gamma}}(\nu)|^2 \tilde{s}(\nu)/\underline{N}(\nu)) \, d\nu$$

Let the second term on the right be denoted by C, then $C + \theta$ is an attainable ϵ- rate of transmission for $\mathcal{C}$. That is, there is a sequence T_k, $\lim_{k\to\infty} T_k = \infty$ and $G(T_k)$ code words $u_1(t), u_2(t), \ldots, u_G(t)$ over $(-T_k, T_k)$ such that

$$\lim_{k\to\infty} \frac{1}{T_k} \log_e G(T_k) = C + \theta$$

and $\hat{d} \geq \epsilon\sqrt{T_k}$ (i.e. $\| H_{\hat{\gamma}} u_i - H_{\hat{\gamma}} u_j \|_2 \geq \epsilon\sqrt{T_k}$ for all $i \neq j$)

Now consider the statistical problem gotten by adding Gaussian noise $z(t)$ to $(H_{\hat{\gamma}} x)(t)$, with the spectral density of $z(t)$ being $\underline{N}(\nu)$. Thus we have the following noisy channel :

$$y(t) = (H_{\hat{\gamma}} x)(t) + z(t)$$

Since $\hat{d} \geq \epsilon\sqrt{T_k}$, there exist $G(T_k)$ disjoint sets B_i such that

$$B_i \supset \{ v(t) \mid \| v - H_{\hat{\gamma}} u_i \|_2^2 < \tfrac{1}{4} \epsilon^2 T_k \}$$

where the u_i are the transmitted signals which yield the ϵ - rate of transmission $C + \theta$. Let these signals, u_i , be the code words for the statistical channel, and let the B_i be the corresponding decoding sets. Notice that

$$P_{\hat{\gamma}}\left\{ B_i^c \mid u_i \right\} \leq P_{\hat{\gamma}}\left\{ \int_{-T_k}^{T_k} z^2(t)dt \geq \tfrac{1}{4}\epsilon^2 T_k \right\}$$

$$\leq P_{\hat{\gamma}}\left\{ \frac{1}{2T_k}\int_{-T_k}^{T_k} z^2(t)dt \geq (\epsilon^2/8 - \Delta) + \Delta \right\}$$

$$= P_{\hat{\gamma}}\left\{ \frac{1}{2T_k}\int_{-T_k}^{T_k} z^2(t)dt \geq \int_{-\infty}^{\infty} \underline{N}(\nu)d\nu + \Delta \right\}$$

$$\leq P_{\hat{\gamma}}\left\{ \left| \frac{1}{2T_k}\int_{-T_k}^{T_k} z^2(t)dt - \int_{-\infty}^{\infty} \underline{N}(\nu)d\nu \right| \geq \Delta \right\}$$

Hence, $\lim_{k\to\infty} P_{\hat{\gamma}}\left\{ B_i^c \mid u_i \right\} = 0$ for $i = 1, 2, \ldots, G(T_k)$

and $C + \theta$ is an attainable rate for the statistical channel with spectral density $\underline{N}(\nu)$ and channel operator $h_{\hat{\gamma}}$. Notice however that C is the channel capacity for such a channel (se e.g. [7]). Thus $C + \theta$ cannot be an attainable rate. We therefore have that $C_\epsilon \leq \overline{C}_\epsilon$.

IV APPLICATIONS AND CONCLUSIONS

The most obvious application of the estimates of sections II and III is to the situation where only rudimentary knowledge of the additive noise and channel operator of a communication channel is available. Suppose for example that we have a channel model which relates transmitted signals $x(t) \in \mathcal{U}_M^T$ to received signals $y(t)$ in the following way

$$y(t) = (H_\gamma x)(t) + z(t) \quad , \quad -T \leq t \leq T$$

where H_γ is a convolution type of operator known only to belong to a class $\mathcal{C}$ which satisfies the conditions of model C and $z(t)$ is additive noise which satisfies the following condition:

$$\lim_{T \to \infty} P \left\{ \frac{1}{2T} \int_{-T}^{T} z^2(t)\, dt > \epsilon \right\} = 0$$

Nothing else is known about $z(t)$. Then, C_ϵ will be a lower bound on the channel capacity for this model.

The ϵ- capacity can also be used to provide lower bounds on the channel capacity for certain classes of nonlinear and time varying channels. Suppose N is a nonlinear, time varying operator which maps $L_2(T)$ into $L_2(T)$ for all $T > 0$. Suppose there exists a class $\mathcal{C}$ of linear convolution operators which satisfies the assumptions of model C $_{\epsilon}$ such that for each $x \in \mathcal{U}_M^T$ there exists an $H_x \in \mathcal{C}$ for which

$$\int_{-T}^{T} |(H_x x)(t) - (Nx)(t)|^2 \, dt < \epsilon^2 T/4 \qquad \text{for all } T > 0$$

Then, the 2ϵ - capacity for this class $\mathcal{C}$ is obviously a lower bound for the ϵ - capacity of the nonlinear time varying channel N .

Although we have derived bounds for the ϵ - capacity of quite general classes of channels, in many situations our bounds are better than those of Root even when $\mathcal{C}$ consists of only a single operator. We will verify this by means of an example.

Consider a continuous time channel with $\tilde{h}(\nu)$ defined as follows

$$\tilde{h}(\nu) = e^{-|\nu|}$$

We will let $M = \frac{1}{2}$ and will assume, for convenience, that $\epsilon < 1$.

Root's lower bound for C_ϵ which we denote by $\underline{C}^R_\epsilon$ is

$$\underline{C}^R_\epsilon = 2 \int_{|\tilde{h}(\nu)| > \epsilon/2} \log_e 2|\tilde{h}(\nu)|/\epsilon \, d\nu - 2 \log_e 2 \int_{|\tilde{h}(\nu)| > \epsilon/2} d\nu$$

and his upper bound $\overline{C}^R_\epsilon$ is

$$\overline{C}_\epsilon^R = 2\int_{|\tilde{h}(\nu)| > \epsilon/2} \log_e 2|\tilde{h}(\nu)|/\epsilon \, d\nu + 5\log_e 2 \int_{|\tilde{h}(\nu)| > \epsilon/2} d\nu$$

By making the appropriate calculations we see that

$$\underline{C}_\epsilon^R = 2\log_e 2/\epsilon \,[\log_e 2/\epsilon - 2\log_e 2]$$

$$\overline{C}_\epsilon^R = 2\log_e 2/\epsilon \,[\log_e 2/\epsilon + 5\log_e 2]$$

By applying Theorem 1 C with $\tilde{s}(\nu)\ \beta/2\epsilon^2 \ -\beta e^{2|\nu|}$

for $|\nu| < 1/2 \log_e 1/2\epsilon^2$, $\tilde{s}(\nu) = 0$ otherwise and

$\beta = 1/2[\ 1/2\epsilon^2 \log 1/2\epsilon^2 - 1/2\epsilon^2 + 1]$, we see that

$$\underline{C}_\epsilon \geq 1/2 \quad \log_e^2 1/2\epsilon^2 - \log_e 1/2\epsilon^2 + 1 - 2\epsilon^2$$

For $\epsilon = 1/2$, $\underline{C}_\epsilon^R = 0$ but $\underline{C}_\epsilon \geq .05$. Thus our lower bound is better than Root's in some instances.

By applying Theorem 4C with $N(\nu) = \epsilon^2/16 \Big/ 4(1 - \epsilon^2/16) \log 2/\epsilon$

for $|\nu| < \log 2/\epsilon$, $N(\nu) = e^{-2|\nu|} \Big/ 4(1 - \epsilon^2/16) \log 2/\epsilon$

for $|\nu| \geq \log 2/\epsilon$ and $\tilde{s}(\nu) = \dfrac{1}{4(1-\epsilon^2/16)\log 2\epsilon} - e^{2|\nu|}N(\nu)$

for $|\nu| < \log 2/\epsilon$, $\tilde{s}(\nu) = 0$ otherwise, we see that

$$\overline{C}_\epsilon \leq 2\log_e 2/\epsilon \,[\log_e 2/\epsilon + 2\log_e 2]$$

which is less than $\overline{C}_\epsilon^R$ for all ϵ.

BIBLIOGRAPHY

1 "Estimation Of ϵ- Capacity For Certain Linear Communication Channels", W. L. Root, Willow Run Laboratories (Technical Report), U. of Michigan, January, 1967.

2 "The ϵ- Entropy And ϵ- Capacity Of Certain Time-Varying Channels", Reese T. Prosser, <u>Journal Of Mathematical Analysis And Applications</u> Vol. 16, Number 3, December, 1966.

3 "Capacity Of Classes Of Gaussian Channels, Part I : Discrete - Time", W. L. Root and P. P. Varaiya, Memorandum No. ERL-208, Electronics Research Laboratory, U. of California, March, 1967.

4 "Capacity Of Classes Of Gaussian Channels, Part II : Continuous Time", W. L. Root and P.P. Varaiya, Memorandum No. ERL-M212, Electronics Research Laboratory, U. of California, May, 1967.

5 "The Capacity Of A Class Of Channels", D. Blackwell, L. Breiman, A. J. Thomasian, <u>Annals of Mathematical Statistics</u>, 30, pp. 1229-1241.

6 "ϵ- Entropy and ϵ- Capacity Of Sets In Function Spaces", A.M. Kolmogorov and V.M. Tihomirov, Uspeki Mat. Nauk , 14 (1959)

7 Information Theory And Reliable Communication , Robert G. Gallager, to be published by Wiley in Sept. 1968.

8 Fourier Transforms , Annals Of Mathematical Studies No. 19, Princeton University Press (1949), S. Bochner and K. Chandrasekharan.