**Coding Capacity of Generalized Additive Channels**

**Personal Author(s)**
C.R. Baker

**Title**
Coding Capacity of Generalized Additive Channels

**Type of Report**
TECHNICAL

**Date of Report**
October, 1987

**Page Count**
13

**Subject Terms**
Channel capacity; Shannon theory; Information Theory.
CODING CAPACITY OF GENERALIZED ADDITIVE CHANNELS

LISS 20
October, 1987

Charles R. Baker
Department of Statistics
University of North Carolina
Chapel Hill, N.C. 27514

This research was supported by ONR Contract N00014-86-K-0039.
Introduction

The generalized additive channel was introduced in [1]. It is described by an additive noise process with sample functions inducing a measure on a linear topological vector space, and by a constraint which includes dimensionality. The coding capacity of the matched channel was analyzed in [1], with an exact value obtained for the Gaussian channel and an upper bound for a class of non-Gaussian channels. Bounds on the coding capacity for the mismatched Gaussian generalized channel were obtained in [2].

In this paper, the exact coding capacity of the mismatched Gaussian generalized channel is determined, along with an upper bound for a class of non-Gaussian mismatched channels. The set of admissible constraints is also greatly increased over that considered in [2]. Although the treatment here is restricted to noise measures induced on a separable Hilbert space, it can readily be seen that the results extend immediately to the class of linear topological vector spaces considered in [1]. The results of the present paper are partly based on the Hilbert space results on information capacity given in [3]; for the extension to linear topological vector spaces, one would use the corresponding results given in [4]. The focus on Hilbert space is useful for application of the results given here to the discrete-time or continuous-time additive channel.

The basic path followed here is well-known to information theorists, appearing in the analysis of much simpler channels. A generalization of Feinstein's Fundamental Lemma is used to obtain a lower bound on capacity, and Fano's inequality is used to obtain an upper bound. However, the generality of the model requires a development considerably different from that of the classical treatment; central to these results is the spectral representation of unbounded self-adjoint operators.
determining bounds on coding capacity of the continuous-time channel. These bounds will be given elsewhere.

\( \mu_G \) is defined as the zero-mean Gaussian cylinder set measure on \( H \) having the same covariance operator as \( \mu_N \). The entropy \( H_G(N) \) of \( \mu_N \) with respect to \( \mu_G \) is defined as follows. Let \( H_n \) be any finite-dimensional subspace of \( H \), with \( \mu_N^n \) and \( \mu_G^n \) the measures induced on \( H_n \) by the projection operator \( P_n : H \rightarrow H_n \). Let \( H_G(N|H_n) \) be the entropy of \( \mu_N^n \) with respect to \( \mu_G^n \):

\[
H_G(N|H_n) = \infty \text{ if it is false that } \mu_N^n \ll \mu_G^n, \text{ while otherwise}
\]

\[
H_G(N|H_n) = \int_{H_n} \left[ \log \frac{d\mu_N^n}{d\mu_G^n} \right] d\mu_G^n. \text{ Define } H_G(N) \text{ by } H_G(N) = \sup_{H_n \subset H, \dim n \geq 1} H_G(N|H_n).
\]

The induced measures \( \mu_G^n \) and \( \mu_N^n \) are always countably additive for any finite-dimensional subspace \( H_n \), while the measure \( \mu_G \) will be countably additive if and only if \( R_N \) is trace-class.

Since \( R_N^{-1} \) exists and \( \text{range}(R_N^2) \subset \text{range}(\frac{1}{2} I + S) \), \( R_N = \frac{1}{2} (I + S)R_N^2 \) for a self-adjoint linear operator \( S \), with \( (I + S)^{-1} \) existing and bounded [5]. \( \Theta \) is the smallest limit point of the spectrum of \( S \). A limit point of the spectrum is either the limit of a sequence of distinct eigenvalues, or an eigenvalue of infinite multiplicity, or a point of the continuous spectrum [6].

Coding Capacity

Theorem 1: (1) If \( H_G(N) < \infty \), then

\[
C_W^\infty(P) \leq \frac{1}{2} \log \left[ 1 + \frac{P}{1 + \Theta} \right].
\]

(2) If \( H_G(N) < \infty \) and \( \dim(H) < \infty \), then \( C^\infty_W(P) = 0 \).

(3) If \( \mu_N \) is Gaussian and \( \dim(H) = \infty \), then \( C^\infty_W(P) = \frac{1}{2} \log \left[ 1 + \frac{P}{1 + \Theta} \right] \).

Proof: The complete theorem will first be proved under the assumption that \( \Theta < \infty \).
Suppose that \( \mu_N \) is Gaussian, with \( \theta < \infty \). We will show that
\[
C_\infty^p(P) \geq \frac{1}{2} \log \left[ 1 + \frac{P}{1+\theta} \right].
\]

Fix any \( \delta > 0 \). Since \( 1 + \theta \) is the smallest limit point of the spectrum of the self-adjoint operator \( I + S \), there exists an infinite o.n. set \( \{ v_n, n > 1 \} \) in the range of the projection operator \( P_{1+\theta+\delta} \), where \( \{ P_t, t \in \mathbb{R} \} \) is the left-continuous resolution of the identity for \( I + S \) such that \( x \in \mathcal{D}(I+S) \) if and only if \( \int_0^\infty \lambda^2 \| P_\lambda x \|^2 \, d\lambda < \infty \), and then \( (I+S)x = \int_0^\infty \lambda dP_\lambda x \) where the integral exists as a limit in the strong operator topology [6].

If \( x \) is any element in \( \text{span}\{ v_n, n \geq 1 \} \), then \( P_t x = x \) for \( t \geq 1+\theta+\delta \), since
\[
\langle x, v_i \rangle = \langle x, v_i \rangle P_t v_i = \langle x, v_i \rangle v_i. \quad \text{Thus, if } x \text{ is in span}\{ v_1, \ldots, v_n \}, \text{then}
\]
\[
\int_0^\infty \lambda^2 \| P_\lambda x \|^2 < \infty \text{, and then } (I+S)x = \int_0^\infty \lambda dP_\lambda x \text{ where the integral exists as a limit in the strong operator topology [6].}
\]

This also shows that \( \text{span}\{ v_n, n \geq 1 \} \) is contained in \( \mathcal{D}(I+S) \), and that
\[
\| (I+S)x \|^2 \leq (1+\theta+\delta)^2 \| x \|^2 \text{ for all } x \in \text{span}\{ v_n, n \geq 1 \}. \text{Similarly,}
\]
\[
\| (I+S)^{\frac{1}{2}} x \|^2 \leq (1+\theta+\delta) \| x \|^2 \text{ if } x \in \text{span}\{ v_n, n \geq 1 \}.
\]

Let \( U \) be the unitary operator in \( H \) which satisfies \( R^\frac{1}{2}_W (I+S)^{\frac{1}{2}} U^* = R^\frac{1}{2}_N \) [5].

For each \( v_n \), define \( u_n = U v_n \), so that \( (I+S)^{\frac{1}{2}} u_n = (I+S)^{\frac{1}{2}} v_n \).

Choose \( Q \in (0, P) \). For \( n \geq 1 \), define \( \mu_X^n \) to be the zero-mean Gaussian measure with covariance operator
\[
\frac{Q}{1+\theta+\delta} \sum_{i=1}^n R^\frac{1}{2}_N u_i \otimes R^\frac{1}{2}_N u_i. \text{ Let}
\]
\[
H_n = \text{span}\{ R^\frac{1}{2}_W u_1, \ldots, R^\frac{1}{2}_W u_n \}. \text{ Note that } H_n \subset \text{range}(R^\frac{1}{2}_W), \text{ because } R^\frac{1}{2}_W u_1 =
\]
\[
R^\frac{1}{2}_W (I+S)^{\frac{1}{2}} U v_i = R^\frac{1}{2}_W (I+S)^{\frac{1}{2}} v_i; \text{ since } \mu_X^n[H_n] = 1, \text{ this shows that}
\]
\[
\mu_X^n[\text{range}(R^\frac{1}{2}_W)] = 1. \text{ Let } \mu_X^n \text{ and } \mu_X^n \otimes \mu_Y^n \text{ be the joint cylinder set measures.
defined by \(\mu_X^n\) and \(\mu_N^n\). Since \(\mu_X^n\) gives full measure to \(H_n\), we can replace \(\mu_N^n\) by the measure \(\mu_N^n \circ P_n^{-1}\), where \(P_n\) is the projection operator with range equal to \(H_n\). Thus the joint measure of interest is concentrated on \(H_n \times H_n\), and if \(B_1\) and \(B_2\) are Borel sets in \(H_n\), then \(\mu_{XY}^{n}(B_1 \times B_2) = \mu_X^n \otimes \mu_Y^n((x, y) : (x, x + P_n y) \in B_1 \times B_2)\). Similarly, \(\mu_Y^n(B_2) = \mu_X^n \otimes \mu_Y^n((x, y) : x + P_n y \in B_2)\). Since both \(\mu_{XY}^n\) and \(\mu_X^n \otimes \mu_Y^n\) are countably additive measures on \(H_n \times H_n\), the results of [3] can be applied. Set \(F_n = \{x \in \text{range}(R^2_n) : \|x\| \leq n P\}\).

It will now be shown that \(\mu_X^n[F_n^c] \rightarrow 0\) as \(n \rightarrow \infty\). Note that \(\mu_X^n = \mu_X^n \circ (R^2_n)^{-1}\), where \(\mu_X^n\) is the zero-mean Gaussian measure with covariance operator
\[
\frac{Q}{1 + \Theta + \delta} \sum_{i=1}^{n} u_i u_i^\star, \text{ so that } x = \sum_{i=1}^{n} \langle x, u_i \rangle u_i \text{ a.e. } d\mu_T^n(x). \text{ Thus}
\]
\[
\mu_X^n[F_n^c] = \mu_X^n(x : \|R_w^{-\frac{1}{2}} R_n^\frac{1}{2} x\|^2 > n P) = \mu_T^n(x : \|I + S\| x\|^2 > n P)
= \mu_T^n(x : (I + S)^{\frac{1}{2}} U_n x\|^2 > n P)
\leq \mu_T^n(x : (I + \Theta + \delta) \sum_{i=1}^{n} \langle u_i, x \rangle^2 > n P).
\]
The random variables \(\{\langle u_i, \cdot \rangle, i \leq n\}\) are i.i.d. Gaussian random variables with respect to \(\mu_T^n\), mean zero and variance \(Q/[1 + \Theta + \delta]\). Applying Chebyshev's inequality, one has \(\mu_X^n[F_n^c] \leq \frac{2n^2}{[n P - n Q]^2}\), so that \(\mu_X^n[F_n^c] \rightarrow 0\) as \(n \rightarrow \infty\).

From the proof of Prop. 2 of [7],
\[
\frac{d\mu_X^n}{d\mu_X^n}(x, y) = \frac{1}{2} \sum_{i=1}^{n} (a_1^2(x, y) - b_1^2(x, y)) + \frac{1}{2} n \log(1 + \frac{Q}{1 + \Theta + \delta})
\]
where \(\{a_1, ..., a_n, b_1, ..., b_n\}\) is a family of i.i.d. Gaussian random variables with respect to \(\mu_X^n\), each having zero mean and variance
\[
\frac{Q/[1 + \Theta + \delta]}{Q/(1 + \Theta + \delta + Q)} = \left[\frac{Q}{1 + \Theta + \delta + Q}\right]^{\frac{1}{2}}. \text{ Take } \gamma > 0, \text{ and define}
\]

JMVA - 11/12/87 - 6
\[ \alpha_n = \frac{1}{2} n \log \left[ 1 + \frac{Q}{1 + \Theta + \delta} \right] - n \gamma, \]

\[ A_n = \{(x,y) : \log \frac{d\mu_{X^n}^n}{d\mu_{X^n}^n}(x,y) > \alpha_n \}, \]

so that \( A_n^c = \{(x,y) : \frac{1}{2} \sum_{i=1}^n (a_i^2 - b_i^2) \leq -n \gamma \}. \) Since the sequence of random variables \( (a_i^2 - b_i^2) \) are independent and have zero mean w.r.t. \( \mu_{XY} \), Chebyshev's inequality gives \( \mu_{XY}^n[A_n^c] \leq \frac{1}{n^2} - 4n \left[ \frac{Q}{1 + \Theta + \delta + Q} \right]^2 \to 0. \)

Let \( R < \frac{1}{2} \log \left[ 1 + \frac{Q}{1 + \Theta + \delta} \right] \) and set \( k_n = [e^{nR}] \). Then,

\[ k_n e^{-\alpha_n nR + n \gamma - \frac{1}{2} n \log [1 + \frac{Q}{1 + \Theta + \delta}]} \leq e. \]

By the Thomasian-Kadota generalization of Feinstein's Fundamental Lemma (see, e.g., [1, p. 165]), there exists a code \( (k_n, F_n, \varepsilon_n) \) with \( \varepsilon_n \leq k_n e^{-\alpha_n} + \mu_{XY}^n(A_n^c) + \mu_X^n(F_n^c) \). From above, both \( \mu_{XY}^n(A_n^c) \) and \( \mu_X^n(F_n^c) \) tend to zero as \( n \to \infty \). Considering \( k_n e^{-\alpha_n} \), choose \( \gamma \) so that

\[ R + \gamma < \frac{1}{2} \log \left[ 1 + \frac{Q}{1 + \Theta + \delta} \right], \]

Then \( k_n e^{-\alpha_n} \to 0 \) also.

This shows that any rate less than \( \frac{1}{2} \log \left[ 1 + \frac{Q}{1 + \Theta + \delta} \right] \) is admissible, for all \( Q < P \) and for all \( \delta > 0 \). Hence, the supremum over all admissible rates must be at least \( \frac{1}{2} \log \left[ 1 + \frac{P}{1 + \Theta} \right] \), so that \( C_f^\infty(P) \geq \frac{1}{2} \log \left[ 1 + \frac{P}{1 + \Theta} \right] \) when \( \mu_N \) is Gaussian.

Now consider the case of a possibly nonGaussian \( \mu_N \), not necessarily countably additive, with \( \Theta < \infty \) and \( H_{GN}(N) < \infty \). Proceeding exactly as in the proof of this result for the matched channel [1, pp. 167-168], it is found that any admissible \( R \) must satisfy \( R \leq \limsup_n \frac{1}{n} C_n^\infty(P) \). \( C_n^\infty(P) \) is the information capacity of the additive Gaussian channel with noise covariance operator \( R_N \), subject to the constraints that \( \text{support}(\mu_X) \) has linear dimension \( \leq n \) and \( \int_{H_{GN}} ||x||_{\infty}^2 d\mu_X(x) \leq np^2. \)
It now remains only to verify that \( \lim_{n \to \infty} \frac{C_{W}^{n}(p)}{n} = \frac{1}{2} \log \left[ 1 + \frac{p}{1+\theta} \right] \).

To show this, one can apply Theorem 2 of [3]. If the operator \( S \) has no eigenvalues less than \( \theta \), then \( C_{W}^{n}(n^p) = \frac{n^p}{2} \log \left[ 1 + \frac{n^p}{(1+\theta)} \right] \) for all \( n \geq 1 \), so \( \lim_{n \to \infty} \frac{1}{n} C_{W}^{n}(n^p) \) exists and equals \( \frac{1}{2} \log \left[ 1 + \frac{p}{1+\theta} \right] \).

If the operator \( S \) has a finite set of eigenvalues less than \( \theta \), \( \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_K < \theta \), then \( \sum_{i=1}^{K} \lambda_i + n^p > K\theta \) for sufficiently large \( n \), so that applying Theorem 2(c) of [3],

\[
\frac{1}{n} C_{W}^{n}(n^p) = \frac{1}{2n} \sum_{i=1}^{K} \lambda_i + \frac{1}{2} \log \left[ 1 + \frac{n^p + \sum_{i=1}^{K} \lambda_i - \theta}{n(1+\theta)} \right]
\]

and this again converges to the limit \( \frac{1}{2} \log \left[ 1 + \frac{p}{1+\theta} \right] \).

Finally, suppose that \( S \) has an infinite sequence of eigenvalues \( \lambda_n \) strictly less than \( \theta \). Since \( \theta \) is the smallest limit point of the spectrum, \( \lambda_n \uparrow \theta \). This means that for any fixed \( \theta \), \( K\lambda_K + \sum_{i=1}^{K} \lambda_i > K\lambda_K \) for all sufficiently large \( K \). To see this, one notes that for any \( \Delta > 0 \), there exists \( M_0 \) such that \( \theta - \lambda_i < \Delta \) for \( i > M_0 \). Thus, for \( K > M_0 \),

\[
K\lambda_K - \sum_{i=1}^{M_0} \lambda_i \leq \sum_{i=1}^{M_0} (\lambda_K - \lambda_i) + (K-M_0)\Delta \leq \sum_{i=1}^{M_0} (\theta - \lambda_i) + (K-M_0)\Delta.
\]

so that

\[
\frac{1}{K} \left[ K\lambda_K - \sum_{i=1}^{K} \lambda_i \right] \leq \frac{1}{K} \left[ \sum_{i=1}^{M_0} (\theta - \lambda_i) + (K-M_0)\Delta \right].
\]

with the right side converging to \( \Delta \) as \( K \to \infty \). Thus, choosing \( \Delta < p \),

\[
K P + \sum_{i=1}^{K} \lambda_i > K\lambda_K \] for \( K \) sufficiently large. One can thus apply part (c) of Theorem 2 of [3], giving
Since \( \log \frac{1+\theta}{1+\lambda_n} \to 0 \), \( \frac{1}{n} \sum_{i=1}^{n} \log \left[ \frac{1+\theta}{1+\lambda_n} \right] \to 0 \). Similarly, \( \frac{1}{n} \sum_{i=1}^{n} (\lambda_i - \theta) \to 0 \).

Thus, one again has \( \lim_{n \to \infty} \frac{1}{n} C_W(nP) = \frac{1}{2} \log \left[ 1 + \frac{P}{1+\theta} \right] \); part (1) is proved, and this also completes the proof of part (3).

If \( \dim \text{range}(R_N) = M < \infty \), then in the immediately preceding result one has for \( n \) sufficiently large,

\[
C_W(nP) = \frac{1}{2} \sum_{i=1}^{M} \log \left[ \frac{M + nP + \sum_{j=1}^{M} \beta_j}{M(1+\beta_1)} \right]
\]

where \( \beta_1 \leq \beta_2 \leq \ldots \leq \beta_M \) are the eigenvalues of \( S \). In this case, \( \lim_{n \to \infty} \frac{1}{n} C_W(nP) = 0 \), so that \( R > 0 \) is not permissible.

The theorem is now proved when \( \theta < \infty \). If \( \theta = \infty \), then obviously \( C_W(P) = \frac{1}{2} \log \left[ 1 + \frac{P}{1+\theta} \right] = 0 \). Part (2) of the theorem can be ignored, since \( \theta = \infty \) cannot occur unless \( \dim \text{range}(R_N) \) is infinite-dimensional. Thus, it only remains to prove part (1), and this is equivalent to showing that

\[
\lim_{n \to \infty} \frac{1}{n} C_W(nP) = 0 \text{ when } \theta = \infty .
\]

If there exists an integer \( M \) such that \( \lambda_{n+1} > P + \frac{1}{n} \sum_{i=1}^{M} \lambda_i \) for all \( n \geq M \), then

\[
\lim_{n \to \infty} \frac{1}{n} C_W(nP) = \lim_{n \to \infty} \frac{1}{2n} \sum_{j=1}^{M} \log \left[ \frac{1 + \lambda_j}{1 + \lambda_j} \right] = 0 .
\]

Suppose that there exists a subsequence \( (n_k) \) of the integers such that

for all \( k \geq 1 \), \( \lambda_{n_k+1} - \frac{1}{n_k} \sum_{i=1}^{n_k} \lambda_i \leq P \). This gives
\[ \lim_{n \to \infty} \frac{1}{n} C^n_W(nP) = \lim_{k} \frac{1}{2n_k} \sum_{k=1}^{n_k} \left[ \frac{P - \left( \frac{1}{n_k+1} \right) \sum_{j=1}^{n_k} \lambda_j + 1 + \lambda_k+1}{1 + \lambda_i} \right] \]

\[ \leq \lim_{k} \frac{1}{2n_k} \sum_{k=1}^{M} \left[ \frac{P + 1 + \lambda_k+1}{1 + \lambda_i} \right] + \lim_{k} \frac{1}{2n_k} \sum_{k=M+1}^{n_k} \left[ \frac{P + 1 + \lambda_k+1}{1 + \lambda_i} \right] \]

for any fixed integer \( M \). Now, since \( \frac{1}{n_k} \sum_{k=1}^{n_k} \lambda_k+1 - \lambda_i \leq P \), and since

\[ \frac{1}{n_k} \sum_{k=1}^{n_k} \lambda_k+1 \to 1 \] as \( k \to \infty \), we must have that \( \frac{n_k}{1+\lambda_i} \) is bounded, so that

\[ \frac{\lambda_k+1}{n_k} \leq C_0 \] for some \( C_0 < \omega \) and all \( k \geq 1 \). The first term on RHS(\( \gamma \)) above is then

\[ \leq \lim_{k} \frac{M}{2n_k} \log \left[ \frac{P + 1 + C_0 n_k}{1 + \lambda_i} \right] = 0. \]

We now have, for any \( M \geq 1 \),

\[ \lim_{n \to \infty} \frac{1}{n} C^n_W(nP) \leq \lim_{k} \frac{1}{2n_k} \sum_{k=M+1}^{n_k} \left[ \frac{P + 1 + \lambda_k+1}{1 + \lambda_i} \right] \]

\[ \leq \lim_{k} \frac{1}{2n_k} \sum_{k=M+1}^{n_k} \left[ \frac{P + \lambda_k+1 - \lambda_i}{1 + \lambda_i} \right] \]

\[ \leq \lim_{k} \frac{1}{2n_k} \sum_{k=M+1}^{n_k} \left[ \frac{\lambda_k+1 - \lambda_i}{1 + \lambda_i} \right] + \frac{P}{2(1+\lambda_{M+1})} \]

\[ \leq \frac{P}{1 + \lambda_{M+1}}. \]

Since \( M \) is arbitrary and \( \lambda \to \infty \), \( \lim_{n \to \infty} \frac{1}{n} C^n_W(nP) = 0 \), and thus \( C_W^\infty(P) = 0 \) when \( \theta = \omega \).
Bounds on Coding Capacity of the Discrete-Time Gaussian Channel

We now consider the following situation. A zero-mean Gaussian stochastic process \( \{N_t, t = 1, 2, \ldots \} \) is represented by a bounded, non-negative, self-adjoint operator \( R_N \) in \( \ell_2 \); \( R_N \) is an infinite matrix with \( R_N(i, j) = E N_i N_j \). The constraint is given in terms of a second such operator \( R_W \) in \( \ell_2 \). The basic assumption to be made is that \( \text{range}(R_N) \) contains \( \text{range}(R_W^2) \).

A simple example of such a channel and constraint is the memoryless Gaussian channel with \( R_W = I \) (leading to an average power constraint) and \( R_N \) given by \( R_N(i, j) = \alpha_j \delta_{ij} \), with \( \alpha_j \geq \epsilon \) for all \( j \geq 1 \), some \( \epsilon > 0 \).

In the discrete-time channel, a code \( (k, n, \epsilon) \) is a set of \( k \) code words \( x_1, \ldots, x_k \) and corresponding decoding sets \( C_1, \ldots, C_k \), satisfying the constraints given below, with the requirement that each \( x_i \) belong to \( \mathbb{R}^n \). The decoding sets are thus Borel sets in \( \mathbb{R}^n \). The constraints on the code words are that

\[
\|x_i\|_{W,n}^2 \leq np, \quad \|x_i\|_{W,n}^2 = \|R_{W,n}^{-2} x_i\|_{n}^2; \quad \|\cdot\|_{n} \text{ is the } n\text{-dimensional Euclidean norm, and } R_{W,n} \text{ is the restriction of } R_W \text{ to } \{1, 2, \ldots, n\} \times \{1, 2, \ldots, n\}. \quad \text{As before, we require that } \mu_n^W(y: y + x_i \in C_i) \geq 1 - \epsilon \text{ for } i \leq k, \text{ where } \mu_n^W \text{ is the measure on } \mathbb{B}[\mathbb{R}^n] \text{ induced from } \mu_n \text{ by the map } q_n: x \mapsto (x_1, x_2, \ldots, x_n). \quad R \geq 0 \text{ is an admissible rate if there exists a sequence of codes } (([e_1^R], n_1, e_1^R) \text{ with } e_1^R \to 0 \text{ as } n_1 \to \infty. \quad \text{The capacity } C_W^\infty(P) \text{ is the supremum over the set of admissible rates.}

An exact expression for the coding capacity of the discrete-time Gaussian channel is given in [8]. In some applications, the value of the coding capacity will be difficult to determine, as it involves rather detailed knowledge of the spectrum of the operator \( S \), defined above. In such cases it is useful to have bounds on coding capacity. For example, a lower bound enables one to strive toward communicating at a rate that is certain to be admissible. We give here upper and lower bounds on coding capacity.
Theorem 2: Suppose that $N$ is Gaussian. Let $\theta_1$ be the smallest and $\theta_K$ the largest limit point of the spectrum of the operator $S$. Then

$$\log \left[ 1 + \frac{P}{(1+\theta_K^*)} \right] \leq C_w^\infty(P) \leq \frac{1}{2} \log \left[ 1 + \frac{P}{(1+\theta_1^*)} \right].$$

If $N$ is not Gaussian, and $H_{GN}(N) < \infty$, then

$$C_w^\infty(P) \leq \frac{1}{2} \log \left[ 1 + \frac{P}{(1+\theta_1^*)} \right].$$

Proof: The upper bound can be obtained from part (1) of Theorem 1. That is, we can identify $\mathbb{R}^n$ with $H_n$, the subspace of $\ell_2$ consisting of all elements $x$ such that $(x)_i = 0$ for $i > n$. The constraint that any admissible code word belong to $H_n$ thus imposes an additional constraint beyond those imposed in proving the theorem; this gives $C_w^\infty(P) \leq \frac{1}{2} \log \left[ 1 + \frac{P}{(1+\theta_1^*)} \right].$

To prove the lower bound, we can of course assume that $\theta_K^* < \infty$. We then simply mimic the proof of part (3) of Theorem 1, but now defining $\mu_X^n$ to be the Gaussian measure with zero mean and covariance matrix

$$R_X^n = \frac{Q_{n^\delta}}{1+\theta_1^*+\delta} \sum_{i=1}^{M_n^\delta} R_{n+i}^{\frac{1}{2}+\frac{1}{2}} R_{n+i}^{\frac{1}{2}-\frac{1}{2}}$$

where the $\{u_{i}^{\delta}, i \leq M_n^\delta\}$ are determined as follows. $\{v_i, i \leq M_n^\delta\}$ are o.n. elements in $\mathbb{R}^n$ such that $\| (I_n + S_n)^{\frac{1}{2}} v_i \|_n^2 \leq 1 + \theta_K^* + \delta$; such elements always exist [3]. $\{u_i^{\delta}, i \leq M_n^\delta\}$ are then defined by $u_i^{\delta} = U_n v_i$, where $U_n$ is the unitary operator in $\mathbb{R}^n$ satisfying $R_{n,n}^{\frac{1}{2}} = \frac{1}{2} (I_n + S_n)^{\frac{1}{2}} U_n^* U_n$. \qed
References


