Properties of Capacity in Gaussian Channels
with or without Feedback

Han Wu Chen†, Ji Wen Yu† and Kenjiro Yanagi‡

†Graduate School of Science and Engineering
Yamaguchi University
Ube, 755-8611 JAPAN

‡Department of Applied Science
Faculty of Engineering
Yamaguchi University
Ube, 755-8611 JAPAN
Email: yanagi@po.cc.yamaguchi-u.ac.jp

Abstract
We give some inequalities of capacity in Gaussian channel with or without feedback. The nonfeedback capacity \( C_{n,Z}(P) \) and the feedback capacity \( C_{n,FB,Z}(P) \) are both concave functions of \( P \). Though it is shown that \( C_{n,Z}(P) \) is a convex function of \( Z \) in some sense, \( C_{n,FB,Z}(P) \) is a convex like function of \( Z \).

1. Introduction

The following model for the discrete time Gaussian channel with feedback is considered:

\[ Y_n = S_n + Z_n, \quad n = 1, 2, \ldots \]

where \( Z = \{Z_n; n = 1, 2, \ldots \} \) is a non-degenerate, zero mean Gaussian process representing the noise and \( S = \{S_n; n = 1, 2, \ldots \} \) and \( Y = \{Y_n; n = 1, 2, \ldots \} \) are stochastic processes representing input signals and output signals, respectively. The channel is with noiseless feedback, so \( S_n \) is a function of a message to be transmitted and the output signals \( Y_1, \ldots, Y_{n-1} \). For a code of rate \( R \) and length \( n \), with code words \( x^n(W, Y^n-1), W \in \{1, \ldots, 2^{nR}\} \), and a decoding function \( g_n : \mathcal{R}^n \rightarrow \{1, \ldots, 2^{nR}\} \), the probability of error is

\[ P_e^{(n)} = P_r(g_n(Y^n) \neq W; Y^n = x^n(W, Y^n-1) + Z^n) \]

where \( W \) is uniformly distributed over \( \{1, \ldots, 2^{nR}\} \) and independent of \( Z^n \). The signal is subject to an expected power constraint

\[ \frac{1}{n} \sum_{i=1}^{n} E[S_i^2] \leq P \]

and the feedback is causal, i.e., \( S_i \) is dependent of \( Z_1, \ldots, Z_{i-1} \) for \( i = 1, 2, \ldots, n \). Similarly, when there is no feedback, \( S_i \) is independent of \( Z^n \). We denote by \( R_X^{(n)}, R_Z^{(n)} \) the covariance matrices of \( X, Z \), respectively. It is well known that a finite block length capacity is given by

\[ C_{n,FB,Z}(P) = \max \frac{1}{2n} \ln \frac{|R_X^{(n)} + R_Z^{(n)}|}{|R_Z^{(n)}|} \]

where the maximum is on \( R_X^{(n)} \) symmetric, nonnegative definite and \( B \) strictly lower triangular, such that

\[ Tr[(I + B)R_X^{(n)}(I + B^t) + BR_Z^{(n)}B^t] \leq nP. \]

Similarly, let \( C_{n,Z}(P) \) be the maximal value when \( B = 0 \), i.e. when there is no feedback. Under these conditions, Cover and Pombra proved the following.

Proposition 1 (Cover and Pombra [5]) For every \( \epsilon > 0 \) there exist codes, with block length \( n \) and \( 2^n(C_{n,Z}(P) - \epsilon) \) codewords, \( n = 1, 2, \ldots \), such that \( P_e^{(n)} \to 0 \), as \( n \to \infty \). Conversely, for every \( \epsilon > 0 \) and any sequence of codes with \( 2^n(C_{n,Z}(P) + \epsilon) \) codewords and block length \( n \), \( P_e^{(n)} \) is bounded away from zero for all \( n \). The same theorem holds in the special case without feedback upon replacing \( C_{n,FB,Z}(P) \) by \( C_{n,Z}(P) \).

When block length \( n \) is fixed, \( C_{n,Z}(P) \) is given exactly.

Proposition 2 (Gallager [9])

\[ C_{n,Z}(P) = \frac{1}{2n} \ln \frac{nP + r_1 + \cdots + r_k}{kr_k}, \]

where \( 0 < r_1 \leq r_2 \leq \cdots \leq r_n \) are eigenvalues of \( R_Z^{(n)} \) and \( k(\leq n) \) is the largest integer satisfying \( nP + r_1 + \cdots + r_k > kr_k \).
We can also represent \( C_{n,FB,Z}(P) \) by the different formula.

**Proposition 3** Let \( D = R_{\beta}^{(n)} > 0 \). Then
\[
C_{n,FB,Z}(P) = \max_{2n} \frac{1}{2^n} \log \left| \frac{T + BD + DB^t + D}{|D|} \right|, \tag{1}
\]
where the maximum is on \( T \geq 0 \) and \( B \) strictly lower triangular, such that
\[
T - BDB^t > 0, \quad \text{and} \quad Tr(T) \leq np.
\]

**Proof.** By definition there is \( A > 0 \) and strictly lower triangular \( B \) such that
\[
Tr[(I + B)A(I + B^t) + BDB^t] \leq np \tag{2}
\]
and
\[
C_{n,FB,Z}(P) = \frac{1}{2^n} \log \left| A + D \right|. \tag{3}
\]
Let
\[
T = (I + B)A(I + B^t) + BDB^t.
\]
Then by (2) we have \( Tr(T) \leq np \) and
\[
T - BDB^t = (I + B)A(I + B^t) > 0.
\]
Since
\[
|I + B| = |I + B^t| = 1,
\]
we have
\[
|A + D| = |(I + B)A(I + B^t) + (I + B)D(I + B^t)| = |T + BD + DB^t + D|.
\]
This consideration shows, by (3),
\[
C_{n,FB,Z}(P) \leq \text{RHS of (1)}.
\]
Conversely there is \( T > 0 \) and strictly lower triangular \( B \) such that \( T - BDB^t > 0 \) and
\[
\text{RHS of (1)} = \frac{1}{2^n} \log \left| \frac{T + BD + DB^t + D}{|D|} \right|. \tag{4}
\]
Let
\[
A = (I + B)^{-1}(T - BDB^t)(I + B^t)^{-1}.
\]
Then since \( T - BDB^t > 0 \), we have \( A > 0 \) and
\[
(A + B)A(I + B^t) + BDB^t = T
\]
so that
\[
Tr[(I + B)A(I + B^t) + BDB^t] \leq np.
\]
Just as in the foregoing arguments
\[
|T + BD + DB^t + D| = |A + D|.
\]
By (4) this consideration shows
\[
\text{RHS of (1)} \leq C_{n,FB,Z}(P).
\]
This completes the proof.

We proved in [3] that the Gaussian feedback capacity \( C_{n,FB,Z}(P) \) is a concave function of \( P \). In this paper we also show that \( C_{n,FB,Z}(P) \) is a convexlike function of \( Z \) by using the operator convexity of \( \log(1 + t^{-1}) \).

At last we have an open problem about convexity of \( C_{n,FB,Z}(P) \).

### 2. Convexity of \( C_{n,}(P), C_{n,FB,}(P) \)

Before proving the convexity of \( C_{n,}(P) \) and the convexikness of \( C_{n,FB,Z}(P) \) as the function of \( Z \), we need the following lemma.

**Lemma 1** The function
\[
f(t) = \log(1 + t^{-1}) = \log(1 + t) - \log(t)
\]
is operator convex on \( (0, \infty) \), that is, for any \( \alpha, \beta \geq 0 \) with \( \alpha + \beta = 1 \) and for \( T_1, T_2 > 0 \)
\[
\log(I + (\alpha T_1 + \beta T_2)^{-1}) \leq \alpha \log(I + T_1^{-1}) + \beta \log(I + T_2^{-1}) \tag{5}
\]

**Proof.** It is well known that for any \( \lambda > 0 \) the function
\[
f_\lambda(t) = \frac{1}{\lambda + t}
\]
is operator convex on \( (0, \infty) \), that is, for \( \alpha, \beta \geq 0 \) with \( \alpha + \beta = 1 \) and for \( T_1, T_2 \geq 0 \)
\[
\{\lambda I + (\alpha T_1 + \beta T_2)^{-1}\} \leq \alpha(\lambda I + T_1)^{-1} + \beta(\lambda I + T_2)^{-1} \tag{6}
\]
Then, since
\[
f(t) = \log(1 + t) - \log(t) = \int_0^t \frac{1}{\lambda + t} d\lambda = \int_0^t f_\lambda(t) d\lambda,
\]
(5) follows from (6).

Now we can prove the convexity of \( C_{n,}(P) \).
Theorem 1 Given $Z_1, Z_2$ and $\alpha, \beta \geq 0$ with $\alpha + \beta = 1$, define $Z$ by
\[ R_Z^{(n)} = \alpha R_{Z_1}^{(n)} + \beta R_{Z_2}^{(n)}. \]
Then
\[ C_{n,Z}(P) \leq \alpha C_{n,Z_1}(P) + \beta C_{n,Z_2}(P). \]

Proof. Let
\[ D_i = R_{Z_i}^{(n)} \quad (i = 1, 2), \quad \text{and} \quad D = R_Z^{(n)}. \]
Then by definition
\[ D = \alpha D_1 + \beta D_2 \]
and for $i = 1, 2$
\[ C_{n,Z}(P) = \max \left\{ \frac{1}{2n} \log \frac{|A + D_i|}{|D_i|} \mid A > 0, Tr(A) \leq nP \right\}. \]

Remark that
\[ \log \frac{|A + D|}{|D|} = \log |AD^{-1} + I| = \log |A^{1/2}D^{-1}A^{1/2} + I| = \log |I + (A^{-1/2}DA^{-1/2})^{-1}|. \]

Since by Lemma 1
\[ \log \frac{|A + D|}{|D|} = Tr[\log \{I + (\alpha(A^{-1/2}D_1A^{-1/2})
+ \beta(A^{-1/2}D_2A^{-1/2}))^{-1}\}] \leq \alpha \log [A + D_1] + \beta \log [A + D_2] \]
we can conclude
\[ C_{n,FB,Z}(P) \leq \frac{\alpha}{2n} \log \frac{|A + D_1|}{|D_1|} + \frac{\beta}{2n} \log \frac{|A + D_2|}{|D_2|} \leq \alpha C_{n,FB,Z_1}(P_1) + \beta C_{n,FB,Z_2}(P_2). \]

This completes the proof.

Next we can prove the convex-likeness of $C_{n,FB}(P)$.

Theorem 2 Given $Z_1, Z_2$ and $\alpha, \beta \geq 0$ with $\alpha + \beta = 1$, define $Z$ by
\[ R_Z^{(n)} = \alpha R_{Z_1}^{(n)} + \beta R_{Z_2}^{(n)}. \]
Then there exist $P_1, P_2 \geq 0$ with $\alpha P_1 + \beta P_2 = P$ such that
\[ C_{n,FB,Z}(P) \leq \alpha C_{n,FB,Z_1}(P_1) + \beta C_{n,FB,Z_2}(P_2). \]

Proof. Let us use the notations in the proof of Theorem 1. Take $A > 0$ and strictly triangular $B$ such that
\[ Tr[(I + B)A(I + B^t) + BDB^t] = nP \]
and
\[ \frac{1}{2n} \log \frac{|A + D|}{|D|} = C_{n,FB,Z}(P). \]
Since
\[ Tr[(I + B)A(I + B^t) + BDB^t] = \alpha Tr[(I + B)A(I + B^t) + BDB^t] + \beta Tr[(I + B)A(I + B^t) + BDB^t], \]
we have $\alpha P_1 + \beta P_2 = P$, where
\[ P_i = \frac{1}{n} Tr[(I + B)A(I + B^t) + BDB^t] \quad (i = 1, 2). \]

Since, as in the proof of Theorem 1,
\[ \log \frac{|A + D|}{|D|} \leq \alpha \log \frac{|A + D_1|}{|D_1|} + \beta \log \frac{|A + D_2|}{|D_2|}, \]
we can conclude
\[ C_{n,FB,Z}(P) \leq \frac{\alpha}{2n} \log \frac{|A + D_1|}{|D_1|} + \frac{\beta}{2n} \log \frac{|A + D_2|}{|D_2|} \leq \alpha C_{n,FB,Z_1}(P_1) + \beta C_{n,FB,Z_2}(P_2). \]

This completes the proof.

Now we have the following open problem.

Open Problem. For any $Z_1, Z_2$, for any $P \geq 0$ and for any $\alpha, \beta \geq 0$ ($\alpha + \beta = 1$),
\[ C_{n,FB,Z}(P) \leq \alpha C_{n,FB,Z_1}(P_1) + \beta C_{n,FB,Z_2}(P_2). \]

Finally we give the sufficient condition.

Theorem 3 For any $Z_1, Z_2$ and for any $\alpha, \beta \geq 0$ ($\alpha + \beta = 1$), we assume the following condition (a) or (b).
(a) $R_{Z_1}^{(n-1)} = R_{Z_2}^{(n-1)}$.
(b) $\alpha R_{Z_1}^{(n)} + \beta R_{Z_2}^{(n)}$ is white.
Then for any $P \geq 0$, 
\[ C_{n,FB,Z}(P) \leq \alpha C_{n,FB,Z_1}(P) + \beta C_{n,FB,Z_2}(P). \]

**Proof.** If (a) is satisfied, then 
\[ \text{Tr}[B R_{Z_1}^{(n)} B^t] = \text{Tr}[B R_{Z_2}^{(n)} B^t] \]
for any $B$. Thus we can show $P_1 = P_2$.

And if (b) is satisfied, then it follows from Theorem 1 that 
\[ C_{n,FB,Z}(P) = C_{n,Z}(P) \leq \alpha C_{n,Z_1}(P) + \beta C_{n,Z_2}(P) \leq \alpha C_{n,FB,Z_1}(P) + \beta C_{n,FB,Z_2}(P). \]

This completes the proof.

**References**


