Operator Inequality and its Application to Capacity of Gaussian Channel

Kenjiro YANAGI [*], Han Wu CHEN [†] and Ji Wen YU [‡]

Key Words: Operator Inequality, Gaussian Channel, Capacity, Feedback

1991 Mathematical Classification: 94A

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 $g_n : \mathbb{R}^n \to \{1, \ldots, 2^{nR}\}$, the probability of error is

$$Pe^{(n)} = Pr\{g_n(Y^n) \neq W; Y^n = x^n(W, Y^{n-1}) + Z^n\},$$

*Department of Applied Science, Faculty of Engineering, Yamaguchi University, Ube 755-8611, Japan. This paper was partially supported by Grant-in-Aid for Scientific Research (C), No.11640169, Japan Society for the Promotion of Science

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

‡Graduate School of Science and Engineering, Yamaguchi University, Ube 755-8611, Japan

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 $2^{n(C_{n,FB,Z}(P)-\epsilon)}$ codewords, $n = 1,2,\ldots$, such that $Pe^{(n)} \to 0$, as $n \to \infty$. Conversely, for every $\epsilon > 0$ and any sequence of codes with $2^{n(C_{n,FB,Z}(P)+\epsilon)}$ codewords and block length n, $Pe^{(n)}$ 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])

$$C_{n,Z}(P) = \frac{1}{2n}\sum_{i=1}^{k} \ln \frac{nP + r_1 + \cdots + r_k}{kr_i},$$

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* $D = R_Z^{(n)} > 0$. *Then*

$$C_{n,FB,Z}(P) = \max \frac{1}{2n} \log \frac{|T + BD + DB^t + D|}{|D|}, \tag{1}$$

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 trianglar B such that

$$Tr[(I + B)A(I + B^t) + BDB^t] \leq nP \tag{2}$$

and

$$C_{n,FB,Z}(P) = \frac{1}{2n} \log \frac{|A + D|}{|D|}. \tag{3}$$

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),

$$C_{n,FB,Z}(P) \leq \textbf{RHS} \text{ of } (1).$$

Conversely there is $T > 0$ and strictly lower triangular B such that $T - BDB^t > 0$ and

$$\textbf{RHS} \text{ of } (1) = \frac{1}{2n} \log \frac{|T + BD + DB^t + D|}{|D|}. \tag{4}$$

Let

$$A = (I + B)^{-1}(T - BDB^t)(I + B^t)^{-1}.$$

Then since $T - BDB^t > 0$, we have $A > 0$ and

$$(I + B)A(I + B^t) + BDB^t = T$$

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

$$\textbf{RHS} \text{ of } (1) \leq C_{n,FB,Z}(P).$$

This completes the proof. □

In this paper, we first show that the Gaussian feedback capacity $C_{n,FB,Z}(P)$ is a concave function of P. And 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 $C_{n,FB,\cdot}(P)$.

2 CONCAVITY OF $C_{n,FB,Z}(\cdot)$

Before proving the concavity of $C_{n,FB,Z}(P)$ as the function of P, we need two lemmas.

Lemma 1 *For $D \geq 0$, and B_1, B_2 and $\alpha, \beta \geq 0$ with $\alpha+\beta=1$*

$$\alpha B_1 D B_1^t + \beta B_2 D B_2^t \geq (\alpha B_1 + \beta B_2) D (\alpha B_1^t + \beta B_2^t).$$

Proof. This is known and easy to prove. In fact,

$$\begin{aligned} & \{\alpha B_1 D B_1^t + \beta B_2 D B_2^t\} - (\alpha B_1 + \beta B_2) D (\alpha B_1^t + \beta B_2^t) \\ = \; & \alpha\beta (B_1 - B_2) D (B_1^t - B_2^t) \geq 0. \end{aligned}$$

□

Lemma 2 *The function $\log t$ is operator concave on $(0,\infty)$, that is, for $T_1, T_2 > 0$ and $\alpha, \beta \geq 0$ with $\alpha+\beta=1$*

$$\log(\alpha T_1 + \beta T_2) \geq \alpha \log(T_1) + \beta \log(T_2).$$

Proof. This is a well known fact. By Lemma 1 we have first

$$(\alpha T_1 + \beta T_2) \geq (\alpha T_1^{1/2} + \beta T_2^{1/2})^2,$$

which implies by Löwner theorem

$$(\alpha T_1 + \beta T_2)^{1/2} \geq \alpha T_1^{1/2} + \beta T_2^{1/2}.$$

Repeating this argument we can conclude

$$(\alpha T_1 + \beta T_2)^{1/(2^k)} \geq \alpha T_1^{1/(2^k)} + \beta T_2^{1/(2^k)} \;\; (k = 1, 2 \ldots).$$

Now the operator concavity of the function $\log t$ can be derived as

$$\begin{aligned} \log(\alpha T_1 + \beta T_2) &= \lim_{k\to\infty} 2^k \{(\alpha T_1 + \beta T_2)^{1/(2^k)} - I\} \\ &\geq \alpha \lim_{k\to\infty} 2^k (T_1^{1/(2^k)} - I) + \beta \lim_{k\to\infty} 2^k (T_2^{1/(2^k)} - I) \\ &= \alpha \log(T_1) + \beta \log(T_2). \end{aligned}$$

□

Now we can prove the convacity of $C_{n,FB,Z}(\cdot)$.

Theorem 1 *Fix Z. Then $C_{n,FB,Z}(P)$ is a concave function of P, that is, for any $P_1, P_2 \geq 0$ and for any $\alpha, \beta \geq 0$ with $\alpha + \beta = 1$*

$$C_{n,FB,Z}(\alpha P_1 + \beta P_2) \geq \alpha C_{n,FB,Z}(P_1) + \beta C_{n,FB,Z}(P_2).$$

Proof. By Proposition 3 there are $T_1, T_2 > 0$ and strictly lower triangular B_1, B_2 such that

$$C_{n,FB,Z}(P_i) = \frac{1}{2n} \log \frac{|T_i + B_i D + D B_i^t + D|}{|D|} \;\; (i = 1, 2),$$

and

$$T_i - B_i D B_i^t > 0, \quad \text{and} \quad Tr(T_i) \leq n P_i \;\; (i = 1, 2).$$

Let

$$T = \alpha T_1 + \beta T_2, \quad \text{and} \quad B = \alpha B_1 + \beta B_2.$$

Then clearly $Tr(T) \leq n(\alpha P_1 + \beta P_2)$ and B is strictly lower triangular.
Since by Lemma 1

$$BDB^t = (\alpha B_1 + \beta B_2) D (\alpha B_1^t + \beta B_2^t) \leq \alpha B_1 D B_1^t + \beta B_2 D B_2^t,$$

we have

$$T - BDB^t \geq \alpha(T_1 - B_1 D B_1^t) + \beta(T_2 - B_2 D B_2^t) > 0.$$

Then again by Proposition 2 we have

$$C_{n,FB,Z}(\alpha P_1 + \beta P_2) \geq \frac{1}{2n} \log \frac{|T + BD + DB^t + D|}{|D|}.$$

Since

$$T + BD + DB^t + D = \alpha(T_1 + B_1 D + DB_1^t + D) + \beta(T_2 + B_2 D + DB_2^t + D),$$

we have by Lemma 2

$$\log(T+BD+DB^t+D) \geq \alpha \log(T_1+B_1D+DB_1^t+D)+\beta \log(T_2+B_2D+DB_2^t+D),$$

which implies

$$Tr[\log(T+BD+DB^t+D)] \geq \alpha Tr[\log(T_1+B_1D+DB_1^t+D)]+\beta Tr[\log(T_2+B_2D+DB_2^t+D)].$$

The inequality

$$C_{n,FB,Z}(\alpha P_1 + \beta P_2) \geq \alpha C_{n,FB,Z}(P_1) + \beta C_{n,FB,Z}(P_2)$$

follows from the relation

$$\log |A| = Tr[\log(A)] \ \ (A > 0).$$

This completes the proof. $\square$

3 CONVEXITY OF $C_{n,\cdot}(P), C_{n,FB,\cdot}(P)$

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

Lemma 3 *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* $T_1, T_1 > 0$

$$\log(I + (\alpha T_1 + \beta T_2)^{-1}) \leq \alpha \log(I + T_1^{-1}) + \beta \log(I + T_2^{-1}). \tag{5}$$

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

$$f(t) = \log(1+t) - \log t = \int_0^1 \frac{1}{\lambda + t} d\lambda = \int_0^1 f_\lambda(t) d\lambda,$$

(5) follows from (6). □

Now we can prove the convexity of $C_{n,\cdot}(P)$.

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

$$C_{n,Z}(P) \leq \alpha C_{n,Z_1}(P) + \beta C_{n,Z_2}(P).$$

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

$$C_{n,Z_i}(P) = \max\{\frac{1}{2n} \log \frac{|A + D_i|}{|D_i|}; A > 0, Tr(A) \leq nP\} \quad (i = 1, 2)$$

and

$$C_{n,Z}(P) = \max\{\frac{1}{2n} \log \frac{|A + D|}{|D|}; A > 0, Tr(A) \leq nP\}.$$

Remark that

$$\begin{aligned} \log \frac{|A + D|}{|D|} &= \log |AD^{-1} + I| \\ &= \log |A^{1/2} D^{-1} A^{1/2} + I| \\ &= \log |I + (A^{-1/2} D A^{-1/2})^{-1}|. \end{aligned}$$

Since by Lemma 3

$$\begin{aligned}
\log \frac{|A+D|}{|D|} &= Tr[\log\{I + (\alpha(A^{-1/2}D_1A^{-1/2}) + \beta(A^{-1/2}D_2A^{-1/2}))^{-1}\}] \\
&\leq \alpha Tr[\log\{I + (A^{-1/2}D_1A^{-1/2})^{-1}\}] + \beta Tr[\log\{I + (A^{-1/2}D_2A^{-1/2})^{-1}\}] \\
&\leq \alpha \log \frac{|A+D_1|}{|D_1|} + \beta \log \frac{|A+D_2|}{|D_2|}.
\end{aligned}$$

This completes the proof. □

Theorem 3 *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 3. 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

$$\begin{aligned}
&Tr[(I+B)A(I+B^t) + BDB^t] \\
= \; &\alpha Tr[(I+B)A(I+B^t) + BD_1B^t] + \beta Tr[(I+B)A(I+B^t) + BD_2B^t],
\end{aligned}$$

we have $\alpha P_1 + \beta P_2 = P$, where

$$P_i = \frac{1}{n} Tr[(I+B)A(I+B^t) + BD_iB^t] \;\; (i = 1, 2).$$

Since, as in the proof of Theorem 2,

$$\log \frac{|A+D|}{|D|} \leq \alpha \log \frac{|A+D_1|}{|D_1|} + \beta \log \frac{|A+D_2|}{|D_2|},$$

we can conclude

$$\begin{aligned}
C_{n,FB,Z}(P) &\leq \frac{\alpha}{2n} \log \frac{|A+D_1|}{|D_1|} + \frac{\beta}{2n} \log \frac{|A+D_2|}{|D_2|} \\
&\leq \alpha C_{n,FB,Z_1}(P_1) + \beta C_{n,FB,Z_2}(P_2).
\end{aligned}$$

This completes the proof. □

Finally we have the following open problem.

Open Problem. For any Z_1, Z_2, for any $P \geq 0$ and for any $\alpha, \beta \geq 0$ $(\alpha + \beta = 1)$,

$$C_{n,FB,Z}(P) \leq \alpha C_{n,FB,Z_1}(P) + \beta C_{n,FB,Z_2}(P).$$

ACKNOWLEDGEMENT. The authors wish to thank the reviewer for his very valuable advice and constructive criticism.

References

[1] H.W.Chen and K.Yanagi, On the Cover's conjecture on capacity of Gaussian channel with feedback, *IEICE Trans. Fundamentals*, vol E80-A, no 11, pp 2272-2275, November 1997.

[2] H.W.Chen and K.Yanagi, Refinements of the half-bit and factor-of-two bounds for capacity in Gaussian channels with feedback, *IEEE Trans. Information Theory*, vol IT-45, no 1, pp 319-325, January 1999.

[3] H.W.Chen and K.Yanagi, Upper bounds on the capacity of discrete time blockwise white Gaussian channels with feedback, *IEEE Trans. Information Theory*, vol IT-46, no 3, pp 1125-1131, May 2000.

[4] T.M.Cover, Conjecture: Feedback does not help much, in *Open problems in communication and computation, T.Cover and B.Gopinath (Ed.)*, pp 70-71, Springer-Verlag, New York, 1987.

[5] T.M.Cover and S.Pombra, Gaussian feedback capacity, *IEEE Trans. Information Theory*, vol IT-35, no 1, pp 37-43, January 1989.

[6] T.M.Cover and J.A.Thomas, *Elements of Information Theory*, New York, Wiley, 1991.

[7] A.Dembo, On Gaussian feedback capacity, *IEEE Trans. Information Theory*, vol IT-35, no 5, pp 1072-1089, September 1989.

[8] P.Ebert, The capacity of the Gaussian channel with feedback, *Bell. Syst. Tech. J.*, vol 49, pp 1705-1712, 1970.

[9] R.G.Gallager, *Information Theory and Reliable Communication*, John Wiley and Sons, New York, 1968.

[10] S.Ihara and K.Yanagi, Capacity of discrete time Gaussian channel with and without feedback, II, *Japan J. Appl. Math.*, vol 6, pp 245-258, 1989.

[11] M.Pinsker, talk delivered at the Soviet Information Theory Meeting, (no abstract published), 1969.

[12] K.Yanagi, An upper bound to the capacity of discrete time Gaussian channel with feedback, *Lecture Notes in Math.*, vol 1299, pp 565-570, 1988.

[13] K.Yanagi, Necessary and sufficient condition for capacity of the discrete time Gaussian channel to be increased by feedback, *IEEE Trans. Information Theory*, vol IT-38, no 6, pp 1788-1791, November 1992

[14] K.Yanagi, An upper bound to the capacity of discrete time Gaussian channel with feedback, II, *IEEE Trans. Information Theory*, vol IT-40, no 2, pp 588-593, March 1994.

[15] K.Yanagi, An upper bound to the capacity of discrete time Gaussian channel with feedback, III, *Bull. Kyushu Inst. Tech., Pure and Applied Mathematics*, vol 45, pp 1-8, 1998.