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

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

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

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 $Z^n$. 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

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

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 [6])** 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 $P_e^{(n)} \to 0$, 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 $n$, $P_e^{(n)}$ 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 [11])**

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

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 $nP + r_1 + \cdots + r_k > kr_k$.

In this paper, we first show that the Gaussian feedback capacity $C_{n,FB,Z}(P)$ is a concave function of $P$ by using the operator concavity of $\log t$. 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 $C_{n,FB,\cdot}(P)$. For the sake of simplicity, we use the notations $R_S, R_Z, R_{S+Z}, \ldots$ instead of $R_S^{(n)}, R_Z^{(n)}, R_{S+Z}^{(n)}, \ldots$ from the next section.
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 some known results. We denote $\text{ran} A, \ker A$ as the range of $A$, the kernel of $A$, respectively.

**Proposition 3 (Douglas [8])** Let $\mathcal{H}$ be a real Hilbert space and let $\mathcal{B}(\mathcal{H})$ be the all of bounded linear operators on $\mathcal{H}$. And let $A, B \in \mathcal{B}(\mathcal{H})$. Then the following assertions are equivalent.

(i) $\text{ran} A \subset \text{ran} B$

(ii) There exists $\alpha \geq 0$ such that $AA^* \leq \alpha BB^*$.

(iii) There exists $C \in \mathcal{B}(\mathcal{H})$ such that $A = BC$.

Furthermore when the above condition (iii) holds, $C$ is uniquely determined such that the following three conditions are satisfied:

(1) $\|C\|^2 = \inf \{\alpha : AA^* \leq \alpha BB^*\}$,

(2) $\ker A = \ker C$,

(3) $\overline{\text{ran} C} \subset (\ker B)^\perp$.

**Proposition 4 (Baker [1])** Let $\mathcal{H}_1$ (resp. $\mathcal{H}_2$) be a real and separable Hilbert space with Borel $\sigma$-field $\Gamma_1$ (resp. $\Gamma_2$). Let $\mu_X$ (resp. $\mu_Y$) be a probability measure on $(\mathcal{H}_1, \Gamma_1)$ (resp. $(\mathcal{H}_2, \Gamma_2)$) satisfying

$$\int_{\mathcal{H}_1} \|x\|^2_1 d\mu_X(x) < \infty \quad (\text{resp.} \quad \int_{\mathcal{H}_2} \|y\|^2_2 d\mu_Y(y) < \infty)$$

Let $R_X$ and $m_X$ (resp. $R_Y$ and $m_Y$) denote the covariance operator and mean element of $\mu_X$ (resp. $\mu_Y$). Let $(\mathcal{H}_1 \times \mathcal{H}_2, \Gamma_1 \times \Gamma_2)$ be the product measurable space generated by the measurable rectangles. Let $\mu_{XY}$, having $\mathcal{R}$ as covariance and $m$ as mean element, be a joint measure on $(\mathcal{H}_1 \times \mathcal{H}_2, \Gamma_1 \times \Gamma_2)$ with projections $\mu_X$ and $\mu_Y$. Then the cross-covariance operator $R_{XY}$ of the $\mu_{XY}$ has a decomposition as

$$R_{XY} = R_X^\frac{1}{2} V R_Y^\frac{1}{2},$$

where $V$ is a unique bounded linear operator such that $V : \mathcal{H}_2 \rightarrow \mathcal{H}_1$, $\|V\| \leq 1$, $\ker R_Y \subset \ker V$ and $\overline{\text{ran} V} \subset \overline{\text{ran} R_X}$.

The following is given by the operator concavity of $\log t$. 

3
Proposition 5 (Cover and Pombra [6]) Let $A$ and $B$ be nonnegative definite matrices. For any $\alpha, \beta \geq 0$ satisfying $\alpha + \beta = 1$, we have

$$|\alpha A + \beta B| \geq |A|^\alpha |B|^\beta.$$ 

Lemma 1 Let $R_S$ be the covariance matrix of a random vector $S$. For any $\alpha, \beta \geq 0$ satisfying $\alpha + \beta = 1$, the following formulas hold.

(i) $\alpha R_{S_1} + \beta R_{S_2} = R_{\alpha S_1 + \beta S_2} = R_{\alpha S_1} + \beta R_{S_2} - \alpha \beta R_{S_1 S_2}.$

(ii) $\alpha R_{S_1} + \beta R_{S_2} \geq R_{\alpha S_1 + \beta S_2}$, where, if $0 < \alpha < 1$, then the equality holds if and only if $S_1 = S_2$.

(iii) $\alpha R_{S_1 + Z} + \beta R_{S_2 + Z} = R_{\alpha S_1 + \beta S_2 + Z} = \alpha R_{S_1} + \beta R_{S_2}.$

Proof of Lemma 1.

(i) It is easy to obtain the following relations by the properties of non-negative definite matrices.

$$R_{\alpha S_1 + \beta S_2} + \alpha \beta R_{S_1 - S_2} = \alpha \beta R_{S_1 - S_2} = \alpha (\alpha + \beta) R_{S_1} + \beta (\alpha + \beta) R_{S_2}.$$

Then we have the result (i).

(ii) We can directly get the result (ii) from (i), because $R_{S_1 - S_2}$ is non-negative definite matrix.

(iii) It is easy to see from (i). Let $S_1 = S_1^* + Z$ and $S_2 = S_2^* + Z$, then

$$\alpha S_1 + \beta S_2 = \alpha (S_1^* + Z) + \beta (S_2^* + Z) = \alpha S_1^* + \beta S_2^* + Z.$$

Therefore

$$S_1 - S_2 = S_1^* + Z - S_2^* - Z = S_1^* - S_2^*.$$

Then we have the result (iii).
Theorem 1 $C_{n,FB,Z}(P)$ is a concave function with respect to $P$. That is, for any $P_1, P_2 \geq 0$ and for any $\alpha, \beta \geq 0$ satisfying $\alpha + \beta = 1$,

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

Proof of Theorem 1. We can define $C_{n,FB,Z}(P)$ as the following

$$C_{n,FB,Z}(P) = \max\{\frac{1}{2n} \ln \frac{|R_{S+Z}|}{|R_Z|}; S \in \Gamma(P)\},$$

where $\Gamma(P) = \{S; Tr[R_S] \leq nP\}$. By Lemma 1 (i) we have

$$\alpha R_{S_1+Z} + \beta R_{S_2+Z} = R_{\alpha S_1+\beta S_2+Z} + \alpha \beta R_{S_1-S_2}$$
$$= R_{\alpha S_1+\beta S_2} + R_{S_1+\beta S_2, Z} + R_{Z, S_1, \alpha S_1+\beta S_2} + R_{Z} + \alpha \beta R_{S_1-S_2}$$
$$= \alpha R_{S_1} + \beta R_{S_2} + R_{\alpha S_1+\beta S_2} V R_{Z}^\frac{1}{2} + R_{Z} V^t R_{\alpha S_1+\beta S_2}^\frac{1}{2} + R_{Z}$$
$$= \alpha R_{S_1} + \beta R_{S_2} + (\alpha R_{S_1} + \beta R_{S_2})^\frac{1}{2} W V R_{Z}^\frac{1}{2} + R_{Z}$$

Here (a) follows from the Lemma 1 (i), and (b) follows from Proposition 4, where $\|V\| \leq 1$, and (c) follows from the fact that we can gain $R_{\alpha S_1+\beta S_2} \leq \alpha R_{S_1} + \beta R_{S_2}$ by Lemma 1 (ii) and $(R_{\alpha S_1+\beta S_2})^\frac{1}{2} = (\alpha R_{S_1} + \beta R_{S_2})^\frac{1}{2} W$ by Proposition 3 (iii), where $\|W\| \leq 1$. By getting determinants on the both sides of the equation above, we have

$$|\alpha R_{S_1} + \beta R_{S_2} + (\alpha R_{S_1} + \beta R_{S_2})^\frac{1}{2} W V R_{Z}^\frac{1}{2} + R_{Z} V^t (\alpha R_{S_1} + \beta R_{S_2})^\frac{1}{2} + R_{Z}|$$
$$= |\alpha R_{S_1+Z} + \beta R_{S_2+Z}| \geq (e) |R_{S_1+Z}|^{\alpha} \left|R_{S_2+Z}\right|^{\beta}.$$ 

Here (e) follows from Proposition 5. Therefore

$$\frac{1}{2n} \ln \frac{|\alpha R_{S_1} + \beta R_{S_2} + (\alpha R_{S_1} + \beta R_{S_2})^\frac{1}{2} W V R_{Z}^\frac{1}{2} + R_{Z} V^t (\alpha R_{S_1} + \beta R_{S_2})^\frac{1}{2} + R_{Z}|}{|R_Z|}$$
$$\geq \frac{1}{2n} \ln \frac{|R_{S_1+Z}|^{\alpha} \left|R_{S_2+Z}\right|^{\beta}}{|R_Z|}$$
$$= \frac{\alpha}{2n} \ln \frac{|R_{S_1+Z}|}{|R_Z|} + \frac{\beta}{2n} \ln \frac{|R_{S_2+Z}|}{|R_Z|}. \quad (1)$$
We can get both the \( S_1 \in \Gamma(P_1) \) attaining to \( C_{n,FB,Z}(P_1) \) and the \( S_2 \in \Gamma(P_2) \) attaining to \( C_{n,FB,Z}(P_2) \). Then the right hand side of (1) has
\[
\text{RHS} = \alpha C_{n,FB,Z}(P_1) + \beta C_{n,FB,Z}(P_2).
\]

Since
\[
\text{Tr}[\alpha R_{S_1} + \beta R_{S_2}] = \alpha \text{Tr}[R_{S_1}] + \beta \text{Tr}[R_{S_2}] \leq \alpha n P_1 + \beta n P_2 = n(\alpha P_1 + \beta P_2)
\]

and \( \|WV\| \leq \|W\|\|V\| \leq 1 \), we maximize the left hand side of (1) over \( \Gamma(\alpha P_1 + \beta P_2) \) and we get
\[
C_{n,FB,Z}(\alpha P_1 + \beta P_2) \geq \text{LHS}.
\]
Then we have
\[
C_{n,FB,Z}(\alpha P_1 + \beta P_2) \geq \alpha C_{n,FB,Z}(P_1) + \beta C_{n,FB,Z}(P_2).
\]

\( \square \)

### 3 OPERATOR INEQUALITY

Let \( \mathcal{H} \) be a Hilbert space. Let \( \mathcal{B}(\mathcal{H}) \) be all bounded linear operators on \( \mathcal{H} \) and \( \mathcal{B}(\mathcal{H})_+ = \{ A \in \mathcal{B}(\mathcal{H}); A \geq 0 \} \). And let \( J \) be any interval of \( \mathbb{R} \) and \( \sigma(A) \) be spectrum of \( A \in \mathcal{B}(\mathcal{H}) \).

**Definition 1** Let \( f : J \to \mathbb{R} \) be continuous.

1. \( f \) is called operator monotone if for any self-adjoint \( A, B \in \mathcal{B}(\mathcal{H}) \) satisfying \( \sigma(A), \sigma(B) \subset J \),
\[
A \leq B \implies f(A) \leq f(B).
\]

2. \( f \) is called operator convex if for any self-adjoint \( A, B \in \mathcal{B}(\mathcal{H}) \) satisfying \( \sigma(A), \sigma(B) \subset J \),
\[
f\left(\frac{A + B}{2}\right) \leq \frac{f(A) + f(B)}{2}.
\]

By the continuity of \( f \), it is equivalent to
\[
f(\lambda A + (1 - \lambda)B) \leq \lambda f(A) + (1 - \lambda) f(B)
\]
for any \( 0 \leq \lambda \leq 1 \).

3. \( f \) is called operator concave if \(-f \) is operator convex.
Proposition 6 Let \( f \) be nonnegative continuous function on \([0, \infty)\). Then \( f \) is operator monotone if and only if \( f \) is operator concave.

Proposition 7 \( f(t) = t^{-1} \) is operator convex on \((0, \infty)\).

Definition 2 (Kubo and Ando [14]) \( \sigma \) is called operator connection if \( \sigma \) is binary operation on \( \mathcal{B}(\mathcal{H})_+ \) satisfying the following axioms;

1. (Monotonicity) \( A \leq C, B \leq D \implies A\sigma B \leq C\sigma D \).

2. (Transform Inequality) \( C(A\sigma B)C \leq (CAC)\sigma(CBC) \).

3. (Upper Continuity) \( A_n \downarrow A, B_n \downarrow B \implies A_n\sigma B_n \downarrow A\sigma B, \)  
where \( A_n \downarrow A \) represents \( A_1 \geq A_2 \geq \cdots \)  
and \( A_n \to A \) (strong operator topology).

\( \sigma \) is called operator mean if \( \sigma \) is operator connection satisfying \( I\sigma I = I \).

Proposition 8 For any operator connection \( \sigma \), there exists a unique nonnegative operator monotone function \( f \) on \([0, \infty)\) such that

\[ f(t)I = I\sigma(tI), \quad t \geq 0. \]

Then we have the followings;

1. \( \sigma \to f \) is an affine order isomorphism between the class of connections and the class of nonnegative operator monotone functions on \([0, \infty)\).

2. For invertible \( A \in \mathcal{B}(\mathcal{H})_+ \),

\[ A\sigma B = A^{1/2}f(A^{-1/2}BA^{-1/2})A^{1/2}. \]

3. \( \sigma \) is operator mean if and only if \( f(1) = 1 \).

Proposition 9 Let \( \sigma \) be operator connection and \( A, B, C \in \mathcal{B}(\mathcal{H})_+ \).
(1) For any invertible \( C \),
\[ C(A\sigma B)C = (CAC)\sigma(CBC). \]

(2) For any \( \alpha \geq 0 \),
\[ \alpha(A\sigma B) = (\alpha A)\sigma(\alpha B). \]

**Definition 3 (Kubo and Ando [14])** For invertible \( A, B \in \mathcal{B}(\mathcal{H})_+ \), parallel sum is defined by
\[ A : B = (A^{-1} + B^{-1})^{-1}. \]

In general for \( A, B \in \mathcal{B}(\mathcal{H})_+ \), it is defined by
\[ A : B = s - \lim_{\epsilon \downarrow 0}(A + \epsilon I) : (B + \epsilon I). \]

Harmonic mean is defined by
\[ A!B = 2(A : B). \]

**Proposition 10** Let \( \sigma \) be operator connection and \( A, B, C, D \in \mathcal{B}(\mathcal{H})_+ \). Then
\[ (A\sigma B) : (C\sigma D) \geq (A : C)\sigma(B : D). \]

**Proposition 11** Let \( f \) be nonnegative continuous function on \([0, \infty)\). If \( f \) is operator monotone, then for any \( A, B \in \mathcal{B}(\mathcal{H})_+ \)
\[ f(A!B) \leq f(A)!f(B). \]

**Proof of Proposition 11.** By Proposition 8, there exists a unique operator connection \( \sigma \) such that \( f(t)I = I\sigma(tI) \).
\[
\begin{align*}
(I\sigma A) : (I\sigma B) & \geq (I : I)\sigma(A : B) \quad \text{(by Proposition 10)} \\
n & = \frac{1}{2}I\sigma(A : B) \\
n & = \frac{1}{2}(I\sigma(\frac{1}{2}(2(A : B)))) \\
n & = \frac{1}{2}(I\sigma(2(A : B))) \quad \text{(by Proposition 9(2))} \\
n & = \frac{1}{2}(I\sigma(A!B)).
\end{align*}
\]
Then we have

\[ 2((I\sigma A) : (I\sigma B)) \geq I\sigma(A!B). \]

Hence

\[ (I\sigma A)!(I\sigma B) \geq I\sigma(A!B). \]

Thus

\[ f(A)!f(B) \geq f(A!B). \]

\[ \square \]

**Proposition 12** Let \( f \) be positive continuous function on \((0, \infty)\). If \( f \) is operator monotone, then \( f(t^{-1}) \) is operator convex.

**Proof of Proposition 12.** For any invertible \( A, B \in B(H)_+ \), we have

\[
\begin{align*}
 f((A + B)^{-1}) &= f(A^{-1}B^{-1}) \\
 &\leq f(A^{-1})!f(B^{-1}) \quad \text{(by Proposition 11)} \\
 &= \left\{ \frac{(f(A^{-1}))^{-1} + (f(B^{-1}))^{-1}}{2} \right\}^{-1} \\
 &\leq \frac{1}{2} f(A^{-1}) + \frac{1}{2} f(B^{-1}) \quad \text{(by Proposition 7)}
\end{align*}
\]

\[ \square \]

### 4 CONVEXITY OF \( C_{n,\tilde{Z}}(P), C_{n,FB,\tilde{Z}}(P) \)

Let \( \tilde{R}_Z = \alpha R_{Z_1} + \beta R_{Z_2} \), where \( \alpha, \beta \geq 0 \) \((\alpha + \beta = 1)\). Then we have the following convexity of \( C_{n,\tilde{Z}}(P) \).

**Theorem 2** For any \( Z_1, Z_2 \), for any \( P \geq 0 \) and for any \( \alpha, \beta \geq 0 \) \((\alpha + \beta = 1)\),

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

**Proof of Theorem 2.** By Proposition 12, \( f(t) = \log(1 + t^{-1}) \) is operator convex on \((0, \infty)\). Then
We can get the \( S \in \Gamma(P) \) attaining to \( C_{n,\tilde{Z}}(P) \), where \( \Gamma(P) = \{ S; Tr[R_S] \leq nP \} \).
And we take the maximization of the right hand side of (2). Then we have the result.

\[ \frac{1}{2n} \log \frac{|R_S + R_{\tilde{Z}}|}{|R_{\tilde{Z}}|} \leq \alpha \frac{1}{2n} \log \frac{|R_S + R_{Z_1}|}{|R_{Z_1}|} + \beta \frac{1}{2n} \log \frac{|R_S + R_{Z_2}|}{|R_{Z_2}|}. \] (2)

\[ \frac{1}{2n} \log \frac{|R_S + R_{\tilde{Z}}|}{|R_{\tilde{Z}}|} \leq \alpha \frac{1}{2n} \log \frac{|R_S + R_{Z_1}|}{|R_{Z_1}|} + \beta \frac{1}{2n} \log \frac{|R_S + R_{Z_2}|}{|R_{Z_2}|}. \] (3)

Theorem 3 For any \( Z_1, Z_2 \), for any \( P \geq 0 \) and for any \( \alpha, \beta \geq 0 (\alpha + \beta = 1) \), there exist \( P_1, P_2 \geq 0 (P = \alpha P_1 + \beta P_2) \) such that
\[ C_{n,FB,\tilde{Z}}(P) \leq \alpha C_{n,FB,Z_1}(P_1) + \beta C_{n,FB,Z_2}(P_2). \]

Proof of Theorem 3. By Proposition 12, \( f(t) = \log(1 + t^{-1}) \) is operator convex on \((0, \infty)\). Then
\[ \frac{1}{2n} \log \frac{|R_S + R_{\tilde{Z}}|}{|R_{\tilde{Z}}|} \leq \alpha \frac{1}{2n} \log \frac{|R_S + R_{Z_1}|}{|R_{Z_1}|} + \beta \frac{1}{2n} \log \frac{|R_S + R_{Z_2}|}{|R_{Z_2}|}. \]

We can get the \((X^*, B^*) \in \Delta(P)\) attaining to \( C_{n,FB,\tilde{Z}}(P) \), where
\[ \Delta(P) = \{(X, B); Tr[(I + B)R_X(I + B^t) + BR_{\tilde{Z}}B^t] \leq nP \}. \]

Since
\[ Tr[(I + B^*)R_X(I + (B^*)^t) + B^*R_{\tilde{Z}}(B^*)^t] = \alpha Tr[(I + B^*)R_X(I + (B^*)^t) + B^*R_{Z_1}(B^*)^t] + \beta Tr[(I + B^*)R_X(I + (B^*)^t) + B^*R_{Z_2}(B^*)^t], \]
we have \( \alpha P_1 + \beta P_2 = P \), where
\[ Tr[(I + B^*)R_X(I + (B^*)^t) + B^*R_{Z_1}(B^*)^t] = nP_1 \]
and
\[ Tr[(I + B^*)R_X(I + (B^*)^t) + B^*R_{Z_2}(B^*)^t] = nP_2. \]

We take the maximization of the right hand side of (3). Then we have the result.

Finally we have the following open problem.

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


