

Review on Capacity of Gaussian Channel with or without Feedback

Kenjiro Yanagi (Yamaguchi University)

1 Probability Measure on Banach Space

Let X be a real separable Banach space and X^* be its dual space. Let $\mathcal{B}(X)$ be Borel σ -field of X . For finite dimensional subspace F of X^* we define the cylinder set C based on F as follows

$$C = \{x \in X; (\langle x, f_1 \rangle, \langle x, f_2 \rangle, \dots, \langle x, f_n \rangle) \in D\}.$$

where $n \geq 1$, $\{f_1, f_2, \dots, f_n\} \subset F$, $D \in \mathcal{B}(\mathbb{R}^n)$. We denote all of cylinder sets based on F by $\mathcal{C}_F$. Then we put

$$\mathcal{C}(X, X^*) = \bigcup \{\mathcal{C}_F; F \text{ is finite dimensional subspaces of } X^*\}.$$

It is easy to show that $\mathcal{C}(X, X^*)$ is a field. Let $\bar{\mathcal{C}}(X, X^*)$ be the σ -field generated by $\mathcal{C}(X, X^*)$. Then $\bar{\mathcal{C}}(X, X^*) = \mathcal{B}(X)$. If μ is a probability measure on $(X, \mathcal{B}(X))$ satisfying $\int_X \|x\|^2 d\mu(x) < \infty$, then there exist a vector $m \in X$ and an operator $R : X^* \rightarrow X^*$ such that

$$\langle m, x^* \rangle = \int_X \langle x, x^* \rangle d\mu(x),$$

$$\langle Rx^*, y^* \rangle = \int_X \langle x - m, x^* \rangle \langle x - m, y^* \rangle d\mu(x),$$

for any $x^* \in X^*$, $y^* \in Y^*$. m is a mean vector of μ and R is a covariance operator of μ which is a bounded linear operator. We remark that R is symmetric in the following sense.

$$\langle Rx^*, y^* \rangle = \langle Ry^*, x^* \rangle, \text{ for any } x^*, y^* \in X^*.$$

And also R is positive in the following sense.

$$\langle Rx^*, x^* \rangle \geq 0, \text{ for any } x^* \in X^*.$$

When $\mu_f = \mu \circ f^{-1}$ is a Gaussian measure on $\mathbb{R}$ for any $f \in X^*$, we call μ a Gaussian measure on $(X, \mathcal{B}(X))$. For any $f \in X^*$, the characteristic function $\bar{\mu}(f)$ is represented by

$$\bar{\mu}(f) = \exp\{i\langle m, f \rangle - \frac{1}{2}\langle Rf, f \rangle\}, \quad (1.1)$$

where $m \in X$ is mean vector of μ and $R : X^* \rightarrow X$ is covariance operator of μ . Conversely when the characteristic function of a probability measure μ on $(X, \mathcal{B}(X))$ is given by (1.1), μ is Gaussian measure whose mean vector is $m \in X$ and covariance operator is $R : X^* \rightarrow X$. Then we can represent $\mu = [m, R]$ as Gaussian measure with mean vector μ and covariance operator R .

2 Reproducing Kernel Hilbert Space and Mutual Information

For any symmetric positive operator $R : X^* \rightarrow X$, there exists a Hilbertian subspace $H (\subset X)$ and a continuous embedding $j : H \rightarrow X$ such that $R = jj^*$. H is isomorphic to the reproducing kernel Hilbert space (RKHS) $\mathcal{H}(k_R)$ which is defined by positive definite kernel k_R satisfying $k_R(x^*, y^*) = \langle Rx^*, y^* \rangle$. Then we call H itself a reproducing kernel Hilbert space. Now we can define mutual information as follows. Let X, Y be real Banach spaces. Let μ_X, μ_Y be probability measures on $(X, \mathcal{B}(X)), (Y, \mathcal{B}(Y))$, respectively, and let μ_{XY} be joint probability measure on $(X \times Y, \mathcal{B}(X) \times \mathcal{B}(Y))$ with marginal distributions μ_X, μ_Y , respectively. That is

$$\begin{aligned} \mu_X(A) &= \mu_{XY}(A \times Y), \quad A \in \mathcal{B}(X), \\ \mu_Y(B) &= \mu_{XY}(X \times B), \quad B \in \mathcal{B}(Y), \end{aligned}$$

If we assume

$$\int_X \|x\|^2 d\mu_X(x) < \infty, \quad \int_Y \|y\|^2 d\mu_Y(y) < \infty,$$

then there exists $m = (m_1, m_2) \in X \times Y$ such that for any $(x^*, y^*) \in X^* \times Y^*$

$$\langle (m_1, m_2), (x^*, y^*) \rangle = \int_{X \times Y} \langle (x, y), (x^*, y^*) \rangle d\mu_{XY}(x, y),$$

where m_1, m_2 are mean vectors of μ_X, μ_Y , respectively, and there exists $\mathcal{R}$ such that

$$\mathcal{R} = \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix} : X^* \times Y^* \rightarrow X \times Y$$

satisfies the following relation: for any $(x^*, y^*), (z^*, w^*) \in X^* \times Y^*$

$$\begin{aligned} &\left\langle \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix} \begin{pmatrix} x^* \\ y^* \end{pmatrix}, \begin{pmatrix} z^* \\ w^* \end{pmatrix} \right\rangle = \\ &\int_{X \times Y} \langle (x, y) - (m_1, m_2), (x^*, y^*) \rangle \langle (x, y) - (m_1, m_2), (z^*, w^*) \rangle d\mu_{XY}(x, y), \end{aligned}$$

where $R_{11} : X^* \rightarrow X$ is covariance operator of μ_X , $R_{22} : Y^* \rightarrow Y$ is covariance operator of μ_Y , and $R_{12} = R_{21}^* : Y^* \rightarrow X$ is cross covariance operator defined by

$$\langle R_{12}y^*, x^* \rangle = \int_{X \times Y} \langle x - m_1, x^* \rangle \langle y - m_2, y^* \rangle d\mu_{XY}(x, y)$$

for any $(x^*, y^*) \in Y^* \times X^*$.

When we put $\mu_{XY} = \left[(0, 0), \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix} \right]$, we obtain $\mu_X = [0, R_X]$, $\mu_Y = [0, R_Y]$.

And there exist RKHSs $H_X \subset X$ of R_X , $H_Y \subset Y$ of R_Y with continuous embeddings $j_X : H_X \rightarrow X$, $j_Y : H_Y \rightarrow Y$ satisfying $R_X = j_X j_X^*$, $R_Y = j_Y j_Y^*$, respectively. Furthermore if we assume RKHS H_X is dense in X and RKHS H_Y is dense in Y , then there exist $V_{XY} : H_Y \rightarrow H_X$ such that

$$R_{XY} = j_X V_{XY} j_Y^*, \quad \|V_{XY}\| \leq 1.$$

Then the following theorem holds.

Theorem 2.1 $\mu_{XY} \sim \mu_X \otimes \mu_Y$ if and only if V_{XY} is Hilbert-Schmidt operatorsatisfying $\|V_{XY}\| < 1$.

Next we define mutual information of μ_{XY} in the following. We put

$$\mathcal{F} = \{(\{A_j\}, \{B_j\}); \{A_j\} \text{ is finite measurable partitions of } X \text{ with } \mu_X(A_j) > 0 \text{ and } \{B_j\} \text{ is finite measurable partitions of } Y \text{ with } \mu_Y(B_j) > 0\}.$$

Then

$$I(\mu_{XY}) = \sup \sum_{i,j} \mu_{XY}(A_i \times B_j) \log \frac{\mu_{XY}(A_i \times B_j)}{\mu_X(A_i) \mu_Y(B_j)}.$$

where the supremum is taken by all $(\{A_i\}, \{B_j\}) \in \mathcal{F}$.

It is easy to show that if $\mu_{XY} \ll \mu_X \otimes \mu_Y$, then

$$I(\mu_{XY}) = \int_{X \times Y} \log \frac{d\mu_{XY}}{d\mu_X \otimes \mu_Y}(x, y) d\mu_{XY}(x, y)$$

and if otherwise, we put $I(\mu_{XY}) = \infty$.

We introduce several properties without proofs in order to state the exact representation of mutual information. Let X be real separable Banach space and $\mu_X = [0, R_X]$, H_X be RKHS of R_X . Let $L_X \equiv \overline{X^*}^{\|\cdot\|_2^{\mu_X}}$ be the completion by norm of $L_2(X, \mathcal{B}(X), \mu_X)$. Then L_X is a Hilbert space with the inner product

$$\langle f, g \rangle_{L_X} = \int_X \langle x, f \rangle \langle x, g \rangle d\mu_X(x)$$

For any embedding $j_X : H_X \rightarrow X$, there exists an unitary operator $U_X : L_X \rightarrow H_X$ such that $U_X f = j_X^* f$, $f \in X^*$.

We give the following important properties of Radon-Nykodym derivatives.

Lemma 2.1 (Pan [17]) *Let X be a real separable Banach space and let $\mu_X = [0, R_X]$, $\mu_Y = [m, R_Y]$. Then $\mu_X \sim \mu_Y$ if and only if the following (1), (2), (3) are satisfied.*

(1) $H_X = H_Y$,

(2) $m \in H_X$,

(3) $JJ^* - I_X$: Hilbert Schmidt operator,

where H_X, H_Y are RKHS of R_X, R_Y , respectively, $J : H_Y \rightarrow H_X$ is continuous injection and $I_X : H_X \rightarrow H_X$ is an identity operator.

And When (1), (2), (3) hold, we assume $\{\lambda_n\}$ is eigenvalues ($\neq 1$) of JJ^* , $\{v_n\}$ is normalized eigenvectors with respect to $\{\lambda_n\}$. Then

$$\begin{aligned} \frac{d\mu_Y}{d\mu_X}(x) &= \exp\{U_X^{-1}[(JJ^*)^{-1/2}m](x) - \frac{1}{2} \langle m, (JJ^*)^{-1}m \rangle_{H_X} \\ &\quad - \frac{1}{2} \sum_{n=1}^{\infty} [(U_X^{-1}v_n)^2(x) (\frac{1}{\lambda_n} - 1) + \log \lambda_n]\}, \end{aligned}$$

where $U_X : L_X \rightarrow H_X$ is an unitary operator.

And when at least one of (1), (2), (3) does not hold, $\mu_X \perp \mu_Y$.

Lemma 2.2 *Let $R_X : X^* \rightarrow X$, $R_Y : Y^* \rightarrow Y$ and*

$$\mathcal{R}_{X \otimes Y} \equiv \begin{pmatrix} R_X & 0 \\ 0 & R_Y \end{pmatrix}.$$

Then $\mathcal{R}_{X \otimes Y} : X^ \times Y^* \rightarrow X \times Y$ is symmetric, positive. And let $H_X, H_Y, H_{X \otimes Y}$ be RKHS of $R_X, R_Y, \mathcal{R}_{X \otimes Y}$, respectively. Then $H_{X \otimes Y} \cong H_X \times H_Y$.*

We obtain the exact representation of mutual information.

Theorem 2.2 *If $\mu_{XY} \sim \mu_X \otimes \mu_Y$, then $I(\mu_{XY}) < \infty$ and*

$$I(\mu_{XY}) = -\frac{1}{2} \sum_{n=1}^{\infty} \log(1 - \gamma_n),$$

where $\{\gamma_n\}$ are eigenvalues of $V_{XY}^* V_{XY}$.

3 Gaussian Channel

We define Gaussian channel without feedback as follows.

Let X be a real separable Banach space representing input space, Y be a real separable Banach space representing output space, respectively. We assume that $\lambda : X \times \mathcal{B}(Y) \rightarrow [0, 1]$ satisfies the following (1), (2).

- (1) For any $x \in X$, $\lambda(x, \cdot) = \lambda_x$ is Gaussian measure on $(Y, \mathcal{B}(Y))$.
- (2) For any $B \in \mathcal{B}(Y)$, $\lambda(\cdot, B)$ is Borel measurable function on $(X, \mathcal{B}(X))$.

We call a triple $[X, \lambda, Y]$ Gaussian channel. When an input source μ_X is given, we can define corresponding output source μ_Y and compound source μ_{XY} as follows.

For any $B \in \mathcal{B}(Y)$

$$\mu_Y(B) = \int_X \lambda(x, B) d\mu_X(x),$$

For any $C \in \mathcal{B}(X) \times \mathcal{B}(Y)$

$$\mu_{XY}(C) = \int_X \lambda(x, C_x) d\mu_X(x),$$

where $C_x = \{y \in Y; (x, y) \in X \times Y\}$.

Capacity of Gaussian channel is defined as the supremum of mutual information $I(\mu_{XY})$ under appropriate constraint on input sources. We put $X = Y$ and $\lambda(x, B) = \mu_Z(B - x)$, $\mu_Z = [0, R_Z]$ for the simplicity. When the constraint is given by

$$\int_X \|x\|_Z^2 d\mu_X(x) \leq P,$$

it is called matched Gaussian channel. The capacity is well known to be $P/2$. On the other hand when the constraint is given by

$$\int_X \|x\|_W^2 d\mu_X(x) \leq P,$$

where μ_W is different from μ_Z , it is called mismatched Gaussian channel. The capacity is given by Baker [4] in the case of X and Y are the same real separable Hilbert space H . Yanagi [21] considered the case of channel distribution $\lambda_x = [0, R_x]$ and showed this channel corresponds to the change of density operator ρ after the measurement.

4 Discrete Time Gaussian Channel with Feedback

The model of discrete time Gaussian channel with feedback is defined as follows.

$$Y_n = S_n + Z_n, \quad n = 1, 2, \dots,$$

where $Z = \{Z_n; n = 1, 2, \dots\}$ is nondegenerate zero mean Gaussian process representing noise, $S = \{S_n; n = 1, 2, \dots\}$ is stochastic process representing input signal and $Y = \{Y_n; n = 1, 2, \dots\}$ is stochastic process representing output signal. The input signal S_n at time n can be represented by some function of message W and output signal $Y_1, Y_2, \dots, Y_{n-1}$. The error probability for code word $x^n(W, Y^{n-1})$, $W \in \{1, 2, \dots, 2^{nR}\}$ with rate R and length n and the decoding function $g_n : \mathbb{R}^n \rightarrow \{1, 2, \dots, 2^{nR}\}$ is defined by

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

where W is uniform distribution which is independent with the noise $Z^n = (Z_1, Z_2, \dots, Z_n)$. The input signals is assumed average power constraint. That is

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

The feedback is causal. That is $S_i (i = 1, 2, \dots, n)$ is dependent with $Z_1, Z_2, \dots, Z_{i-1}$. In the nonfeedback case $S_i (i = 1, 2, \dots, n)$ is independent with $Z^n = (Z_1, Z_2, \dots, Z_n)$. Since the input signals can be assumed Gaussian, we can represent as follows.

$$C_{n,FB}(P) = \max \frac{1}{2n} \log \frac{|R_X^{(n)} + R_Z^{(n)}|}{|R_Z^{(n)}|},$$

where $|\cdot|$ is determinant and the maximum is taken under strictly lower triangle matrix B and nonnegative symmetric matrix $R_X^{(n)}$ satisfying

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

The nonfeedback capacity is given by the condition $B = 0$. The feedback capacity can be represented by the different form.

$$C_{n,FB}(P) = \max \frac{1}{2n} \log \frac{|R_{S+Z}^{(n)}|}{|R_Z^{(n)}|},$$

where the maximum is taken under nonnegative symmetric matrix $R_S^{(n)}$. Cover and Pombra [9] obtained the following.

Proposition 4.1 (Cover and Pombra [9]) *For any $\epsilon > 0$ there exists $2^{n(C_{n,FB}(P) - \epsilon)}$ code words with block length n such that $Pe^{(n)} \rightarrow 0$ for $n \rightarrow \infty$. Conversely For any $\epsilon > 0$ and any $2^{n(C_{n,FB}(P) + \epsilon)}$ code words with block length n , $Pe^{(n)} \rightarrow 0$ ($n \rightarrow \infty$) does not hold.*

$C_n(P)$ is given exactly.

Proposition 4.2 (Gallager [10])

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

where $0 < r_1 \leq r_2 \leq \cdots \leq r_n$ are eigenvalues of $R_Z^{(n)}$, $k(\leq n)$ is the largest integer satisfying $nP + r_1 + r_2 + \cdots + r_k > kr_k$.

4.1 Necessary and sufficient condition for increase of feedback capacity

We give the following definition for $R_Z^{(n)}$.

Definition 4.1 (Yanagi [23]) Let $R_Z^{(n)} = \{z_{ij}\}$ and $L_k = \{\ell(\neq k); z_{k\ell} \neq 0\}$. Then

- (a) $R_Z^{(n)}$ is called *white* if $L_k = \emptyset$ for any k .
- (b) $R_Z^{(n)}$ is called *completely non-white* if $L_k \neq \emptyset$ for any k .
- (c) $R_Z^{(n)}$ is *blockwise white* if there exists k, ℓ such that $L_k = \emptyset$ and $L_\ell \neq \emptyset$.
We denote by $\tilde{R}_Z$ the submatrix of $R_Z^{(n)}$ generated by k with $L_k \neq \emptyset$.

Theorem 4.1 (Ihara and Yanagi [12], Yanagi [23]) The following (1), (2) and (3) hold.

- (1) If $R_Z^{(n)}$ is white, then $C_n(P) = C_{n,FB}(P)$ for any $P > 0$.
- (2) If $R_Z^{(n)}$ is completely non-white, then $C_n(P) < C_{n,FB}(P)$ for any $P > 0$.
- (3) If $R_Z^{(n)}$ is blockwise white, then we have two cases in the following.
Let r_m is the minimum eigenvalue of $\tilde{R}_Z$ and $nP_0 = mr_m - (r_1 + r_2 + \cdots + r_m)$.
 - (a) If $P > P_0$, then $C_n(P) < C_{n,FB}(P)$.
 - (b) If $P \leq P_0$, then $C_n(P) = C_{n,FB}(P)$.

4.2 Upper bound of $C_{n,FB}(P)$

Since we can't obtain the exact value of $C_{n,FB}(P)$ generally, the upper bound of $C_{n,FB}(P)$ is important. The following theorem has a kind of beautiful expression.

Theorem 4.2 (Cover and Pombra [9])

$$C_{n,FB}(P) \leq \min\{2C_n(P), C_n(P) + \frac{1}{2} \log 2\}.$$

Proof. We use $R_S, R_Z, \dots$ for a simplification of $R_S^{(n)}, R_Z^{(n)}, \dots$. We obtain the following relation by using properties of covariance matrices.

$$\frac{1}{2}R_{S+Z} + \frac{1}{2}R_{S-Z} = R_S + R_Z. \quad (4.1)$$

By operator concavity of $\log x$

$$\frac{1}{2} \log R_{S+Z} + \frac{1}{2} \log R_{S-Z} \leq \log\left\{\frac{1}{2}R_{S+Z} + \frac{1}{2}R_{S-Z}\right\} = \log\{R_S + R_Z\}.$$

We take Tr and get

$$\frac{1}{2} \log |R_{S+Z}| + \frac{1}{2} \log |R_{S-Z}| \leq \log |R_S + R_Z|.$$

Then

$$\frac{1}{2} \frac{1}{2n} \log \frac{|R_{S+Z}|}{|R_Z|} + \frac{1}{2} \frac{1}{2n} \log \frac{|R_{S-Z}|}{|R_Z|} \leq \frac{1}{2n} \log \frac{|R_S + R_Z|}{|R_Z|}.$$

Now since

$$\frac{1}{2n} \log \frac{|R_{S-Z}|}{|R_Z|} \geq 0,$$

we have

$$\frac{1}{2} \frac{1}{2n} \log \frac{|R_{S+Z}|}{|R_Z|} \leq \frac{1}{2n} \log \frac{|R_S + R_Z|}{|R_Z|}.$$

By maximizing under the condition $Tr[R_S] \leq nP$

$$C_{n,FB}(P) \leq 2C_n(P).$$

By (4.1)

$$R_{S+Z} \leq 2(R_S + R_Z).$$

Then

$$\frac{1}{2n} \log \frac{|R_{S+Z}|}{|R_Z|} \leq \frac{1}{2n} \log \frac{|R_S + R_Z|}{|R_Z|} + \frac{1}{2} \log 2.$$

By maximizing under the condition $Tr[R_S] \leq nP$

$$C_{n,FB}(P) \leq C_n(P) + \frac{1}{2} \log 2.$$

□

4.3 Cover's conjecture

Cover gave the following conjecture.

Conjecture 4.1 (Cover [8])

$$C_n(P) \leq C_{n,FB}(P) \leq C_n(2P).$$

We remark the following.

Proposition 4.3 (Chen and Yanagi [5])

$$C_n(2P) \leq \min\{2C_n(P), C_n(P) + \frac{1}{2} \log 2\}.$$

Then if we can prove Conjecture 4.1, we obtain Theorem 4.2 as its colrollary. On the other hand we proved conjecture for $n = 2$. But conjecture is not solved in the case of $n \geq 3$ still now.

Theorem 4.3 (Chen and Yanagi [5])

$$C_2(P) \leq C_{2,FB}(P) \leq C_2(2P).$$

4.4 Concavity of $C_{n,FB}(\cdot)$

Concavity of non-feedback capacity $C_n(\cdot)$ is clear, but concavity of feedback capacity $C_{n,FB}(\cdot)$ is also given.

Theorem 4.4 (Chen and Yanagi [7], Yanagi, Chen and Yu [26]) *For any $P, Q \geq 0$ and any for $\alpha, \beta \geq 0 (\alpha + \beta = 1)$*

$$C_{n,FB}(\alpha P + \beta Q) \geq \alpha C_{n,FB}(P) + \beta C_{n,FB}(Q).$$

5 Mixed Gaussian channel with feedback

Let Z_1, Z_2 be Gaussian processes with mean 0 and covariance operator $R_{Z_1}^{(n)}, R_{Z_2}^{(n)}$, respectively. Let $\tilde{Z}$ be Gaussian process with mean 0 and covariance operator

$$R_{\tilde{Z}}^{(n)} = \alpha R_{Z_1}^{(n)} + \beta R_{Z_2}^{(n)},$$

where $\alpha, \beta \geq 0 (\alpha + \beta = 1)$. We define the mixed Gaussian channel by additive Gaussian channel with $\tilde{Z}$ as noise. $C_{n, \tilde{Z}}(P)$ is called capacity of mixed Gaussian channel without feedback. And $C_{n, FB, \tilde{Z}}(P)$ is called capacity of mixed Gaussian channel with feedback. Now we gave concavity of $C_{n, \tilde{Z}}(P)$ in the following sence.

Theorem 5.1 (Yanagi, Chen and Yu [26], Yanagi, Yu and Chao [27]) *For any $P > 0$*

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

Theorem 5.2 (Yanagi, Chen and Yu [26], Yanagi, Yu and Chao [27]) *For any $P > 0$ there exit $P_1, P_2 \geq 0 (P = \alpha P_1 + \beta P_2)$ such that*

$$C_{n, FB, \tilde{Z}}(P) \leq \alpha C_{n, FB, Z_1}(P_1) + \beta C_{n, FB, Z_2}(P_2).$$

The proof is given by the operator convexity of $\log(1 + t^{-1})$ essentially. But the following conjecture is not solved still now.

Conjecture 5.1 *For $P > 0$*

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

Conjecture is partially solved under some condition.

Theorem 5.3 (Yanagi, Yu and Chao [27]) *If one of the following conditions is satisfied, the corollay holds.*

(a) $R_{Z_1}^{(n-1)} = R_{Z_2}^{(n-1)}$.

(b) $R_{\tilde{Z}}$ is white.

We also give the following conjecture.

Conjecture 5.2 *For any $Z_1, Z_2, P_1, P_2 \geq 0, \alpha, \beta \geq 0 (\alpha + \beta = 1)$,*

$$\begin{aligned} & \alpha C_{n, FB, Z_1}(P_1) + \beta C_{n, FB, Z_2}(P_2) \\ & \leq C_{n, FB, \tilde{Z}}(\alpha P_1 + \beta P_2) + \frac{1}{2n} \log \frac{|R_{\tilde{Z}}|}{|R_{Z_1}|^\alpha |R_{Z_2}|^\beta}. \end{aligned}$$

6 Kim's result

Definition 6.1 $Z = \{Z_i; i = 1, 2, \dots\}$ is first order moving average Gaussian process if the following equivalent three conditions.

(1) $Z_i = \alpha U_{i-1} + U_i$, $i = 1, 2, \dots$, where $U_i \sim N(0, 1)$ is i.i.d.

(2) Spectral density function (SDF) $f(\lambda)$ is given by

$$f(\lambda) = \frac{1}{2\pi} |1 + \alpha e^{-i\lambda}|^2 = \frac{1}{2\pi} (1 + \alpha^2 + 2\alpha \cos \lambda).$$

(3) $Z_n = (Z_1, \dots, Z_n) \sim N_n(0, K_Z)$, $n \in \mathbb{N}$, where covariance matrix K_Z is given by

$$K_Z = \begin{pmatrix} 1 + \alpha^2 & \alpha & 0 & \cdots & 0 \\ \alpha & 1 + \alpha^2 & \alpha & \cdots & 0 \\ 0 & \alpha & 1 + \alpha^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \alpha \\ 0 & 0 & 0 & \cdots & 1 + \alpha^2 \end{pmatrix}.$$

Then entropy rate of Z is given by

$$\begin{aligned} h(Z) &= \frac{1}{4\pi} \int_{-\pi}^{\pi} \log\{4\pi^2 e f(\lambda)\} d\lambda \\ &= \frac{1}{4\pi} \int_{-\pi}^{\pi} \log\{2\pi e |1 + \alpha e^{-i\lambda}|^2\} d\lambda \\ &= \frac{1}{2} \log(2\pi e) \quad \text{if } |\alpha| \leq 1 \\ &= \frac{1}{2} \log(2\pi e \alpha^2) \quad \text{if } |\alpha| > 1, \end{aligned}$$

where the last term is used by the following Poisson's integral formula.

$$\begin{aligned} \frac{1}{2\pi} \int_{-\pi}^{\pi} \log |e^{i\lambda} - \alpha| d\lambda &= 0 \quad \text{if } |\alpha| \leq 1, \\ &= \log |\alpha| \quad \text{if } |\alpha| > 1. \end{aligned}$$

Capacity of Gaussian channel with $MA(1)$ Gaussian noise is given by

$$C_{Z,FB}(P) = \lim_{n \rightarrow \infty} C_{n,Z,FB}(P).$$

Recently Kim obtained capacity of Gaussian channel with feedback for the first time.

Theorem 6.1 (Kim [15])

$$C_{Z,FB}(P) = -\log x_0,$$

where x_0 is only one positive solution of the following equation;

$$Px^2 = (1 - x^2)(1 - |\alpha|x)^2.$$

7 Counter example of Conjecture 4.1

Kim [16] gave the counter example of Conjecture 4.1. When

$$f_Z(\lambda) = \frac{1}{4\pi} |1 + e^{i\lambda}|^2 = \frac{1 + \cos \lambda}{2\pi},$$

input is known to be taken by

$$f_X(\lambda) = \frac{1 - \cos \lambda}{2\pi}.$$

Then output is given by

$$f_Y(\lambda) = f_X(\lambda) + f_Z(\lambda) = \frac{1}{\pi}.$$

Then nonfeedback capacity is given by

$$\begin{aligned} C_Z(2) &= \frac{1}{4\pi} \int_{-\pi}^{\pi} \log \frac{f_Y(\lambda)}{f_Z(\lambda)} d\lambda \\ &= \frac{1}{4\pi} \int_{-\pi}^{\pi} \log \frac{4}{|1 + e^{i\lambda}|^2} d\lambda \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \log \frac{2}{|1 + e^{i\lambda}|} d\lambda \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \log 2 d\lambda - \frac{1}{2\pi} \int_{-\pi}^{\pi} \log |1 + e^{i\lambda}| d\lambda \\ &= \frac{1}{2\pi} 2\pi \log 2 - 0 \\ &= \log 2. \end{aligned}$$

On the other hand feedback capacity is given by

$$C_{Z,FB}(1) = -\log x_0,$$

where x_0 is only one positive solution of equation

$$x^2 = (1+x)(1-x)^3.$$

Since $x_0 < \frac{1}{2}$ is assumed, we have the following

$$C_{Z,FB}(1) = -\log x_0 > \log 2 = C_Z(2).$$

This is a counter example of Conjecture 4.1. And we can show that there exists $n_0 \in \mathbb{N}$ such that

$$C_{n_0,Z,FB}(1) > C_{n_0,Z}(2).$$

References

- [1] N.Aronszajn, Theory of reproducing kernels, Trans. Amer. Math. Soc., vol 68, pp 337-404, 1950.
- [2] C.R.Baker, Joint measures and cross-covariance operators, Trans. Amer. Math. Soc., vol 186, pp 273-289, 1973.
- [3] C.R.Baker, Capacity of the Gaussian channel without feedback, Information and Control, vol 37, pp 70-89, 1978.
- [4] C.R.Baker, Capacity of mismatched Gaussian channel, IEEE Trans. Information Theory, vol IT-33, pp 802-812, 1987.
- [5] 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, 1997.
- [6] H.W.Chen and K.Yanagi, Refinements of the half-bit and factor-of-two bounds for capacity in Gaussian channel with feedback, IEEE Trans. Information Theory, vol IT-45, no 1, pp 319-325, 1999.
- [7] H.W.Chen and K.Yanagi, Upper bounds on the capacity of discrete-time block-wise white Gaussian channels with feedback, IEEE Trans. Information Theory, vol IT-46, no 3, pp 1125-1131, 2000.
- [8] T.M.Cover and B.Gopinath, Open problems in communication and computation, Springer-Verlag, New York, 1987.
- [9] T.M.Cover and S.Pombra, Gaussian feedback capacity, IEEE Trans. Information Theory, vol IT-35, no 1, pp 37-43, 1989.
- [10] R.G.Gallager, Information Theory and reliable communication, John Wiley and Sons, New York, 1968.
- [11] F.Hiai and K.Yanagi, Hilbert spaces and linear operators, Makino-Syoten, 1995, (in Japanese).
- [12] 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.
- [13] S. Ihara, Information theory for continuous systems, World Scientific, 1993.
- [14] K.Kunisawa and H.Umegaki, Progress of information theory, Iwanami-Syoten, 1965, (in Japanese).

- [15] Y.H.Kim, Feedback capacity of the first-order moving average Gaussian channel, *IEEE Trans. Information Theory*, vol IT-52, no 7, pp 3063-3079, 2006.
- [16] Y.H.Kim, A counter example to Cover's 2P conjecture on Gaussian feedback capacity, *IEEE Trans. Information Theory*, vol IT-52, no 8, pp 3792-3793, 2006.
- [17] Y.M.Pan, Simple proofs of equivalence conditions for measures induced by Gaussian processes, *Selected Transl. Math. Stat. Prob.*, vol 12, pp 109-118, 1973.
- [18] H.Umegaki, Conditional expectations in an operator algebra IV (entropy and information), *Kodai Math. Sem. Rep.*, vol 14, pp 59-85, 1962.
- [19] H.Umegaki, Basic of information mathematics, Science-Sya, 1993, (in Japanese).
- [20] N.N.Vakhania, V.I.Tarieladze and S.A.Chobanyan, Probability distribution on Banach spaces, D.Reidel Publishing Company, 1987.
- [21] K.Yanagi, Covariance operators and von Neumann's theory of measurements, *Kodai Math. J.*, vol 5, no 3, pp 435-445, 1982.
- [22] 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.
- [23] 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, 1992.
- [24] 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, 1994.
- [25] 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.
- [26] K.Yanagi, H.W.Chen and J.W.Yu, Operator inequality and its application to capacity of Gaussian channel, *Taiwanese J. Math.*, vol 4, no 3, pp 407-416, 2000.
- [27] K.Yanagi, J.W.Yu and I.F.Chao, On some inequalities for capacity in mixed Gaussian channels with feedback, *Arch. Inequalities Appl.*, vol 2, pp 13-24, 2004.