International Symposium on Information Theory and its Applications, ISITA2006
Seoul, Korea, October 29–November 1, 2006

Counter-examples of the trace inequalities related to the auxiliary function of the quantum reliability function

Shigeru Furuichi† and Kenjiro Yanagi‡

† Department of Electronics and Computer Science
Tokyo University of Science in Yamaguchi
1-1-1 Daigakudori Sanyou Onoda City, Yamaguchi, 756-0884, Japan
E-mail: furuichi@ed.yama.tus.ac.jp

‡ Division of Applied Mathematical Science
Graduate School of Science and Engineering, Yamaguchi University
2-16-1 Tokiwadai, Ube City, Yamaguchi, 755-0811, Japan
E-mail: yanagi@yamaguchi-u.ac.jp

Abstract

We study the open problem given by Holevo and Ogawa-Nagaoka on the concavity of the auxiliary function of the quantum reliability function. Firstly we review the previous results on this problem in the case that the parameter s is positive. Secondly we consider the problem in the case that the parameter s is negative.

1. INTRODUCTION

In classical information theory [1], the random coding exponent $E_r^c(R)$, the lower bound of the reliability function, is defined by

$$E_r^c(R) = \max_{p,s}\left[E_c(p,s) - sR\right].$$

As for the classical auxiliary function $E_c(p,s)$, it is well-known the following properties [1].

(a) $E_c(p,0) = 0$.

(b) $\frac{\partial E_c(p,s)}{\partial s}|_{s=0} = I(X;Y)$, where $I(X;Y)$ presents the classical mutual information.

(c) $E_c(p,s) > 0$ $(0 \le s \le 1)$. $E_c(p,s) < 0$ $(-1 < s < 0)$.

(d) $\frac{\partial E_c(p,s)}{\partial s} > 0$, $(-1 < s \le 1)$.

(e) $\frac{\partial^2 E_c(p,s)}{\partial s^2} \le 0$, $(-1 < s \le 1)$.

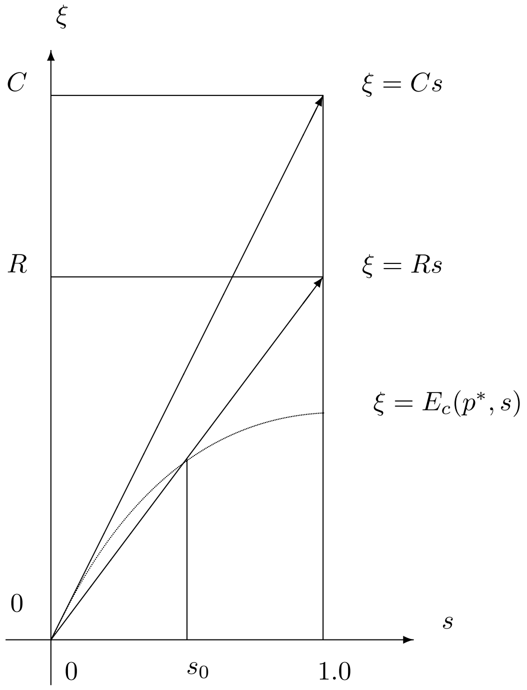

Figure 1: The sketch of the auxiliary function $\xi = E_c(p^*,s)$ in $0 \le s \le 1$.

In figure 1, we suppose that p^* is *a priori* probability which attains the maximum of the classical mutual information. We then find that there exists a code satisfying $E_r^c(R) > 0$ by the above properties. Thus the upper bound [1] of the error probability P_e due to the random coding and the maximum likelihood decoding

$$P_e \le \exp\left[-nE_r^c(R)\right], \quad (0 \le s \le 1)$$

goes to 0 as the code length $n \to \infty$. The parameter $s \in (-1, 0)$ (*resp.*$s \in [0, 1]$) corresponds to the converse (*resp.* direct) part of the channel coding theorem.

In quantum information theory, it is also important to study the properties of the auxiliary function $E_q(\pi, s)$, which will be defined in the below, appearing in the lower bound with respect to the random coding in the reliability function for general quantum states. The corresponding properties to (a),(b),(c) and (d) in quantum system have been shown in [4, 6]. Also the concavity of the auxiliary function $E_q(\pi, s)$ is shown in the case when the signal states are pure [5], and when the expurgation method is adopted [6]. However, for general signal states, the concavity of the auxiliary function $E_q(\pi, s)$ which corresponds to (e) in the above has remained as an open question [4, 6].

The reliability function of classical-quantum channel is defined by

$$E(R) \equiv -\liminf_{n\to\infty} \frac{1}{n} \log P_e(2^{nR}, n), \quad 0 < R < C, \quad (1)$$

where C is a classical-quantum capacity, R is a transmission rate $R = \frac{\log_2 M}{n}$ (n and M represent the number of the code words and the messages, respectively), $P_e(M, n)$ can be taken any minimal error probabilities of $\min_{\mathcal{W},\mathcal{X}} \bar{P}(\mathcal{W}, \mathcal{X})$ or $\min_{\mathcal{W},\mathcal{X}} P_{\max}(\mathcal{W}, \mathcal{X})$. These error probabilities are defined by

$$\begin{aligned} \bar{P}(\mathcal{W}, \mathcal{X}) &= \frac{1}{M} \sum_{j=1}^{M} P_j(\mathcal{W}, \mathcal{X}), \\ P_{\max}(\mathcal{W}, \mathcal{X}) &= \max_{1\le j\le M} P_j(\mathcal{W}, \mathcal{X}), \end{aligned}$$

where

$$P_j(\mathcal{W}, \mathcal{X}) = 1 - \mathrm{Tr} S_{w^j} X_j$$

is the usual error probability associated with the positive operator valued measurement $\mathcal{X} = \{X_j\}$ satisfying $\sum_{j=1}^{M} X_j \le I$. Here we note S_{w^j} represents the density operator corresponding to the code word w^j chosen from the code(blook) $\mathcal{W} = \left\{w^1, w^2, \cdots, w^M\right\}$. For details, see [3, 4, 6].

The lower bound for the quantum reliability function defined in Eq.(1), when we use random coding, was conjectured [5, 6] by

$$E(R) \ge E_r^q(R) \equiv \max_{\pi} \sup_{0<s\le 1} \left[E_q(\pi, s) - sR\right],$$

where $\pi = \{\pi_1, \pi_2, \cdots, \pi_a\}$ is *a priori* probability distribution satisfying $\sum_{i=1}^{a} \pi_i = 1$ and

$$\begin{aligned} E_q(\pi, s) &= -\log G(s), \\ G(s) &= \mathrm{Tr}\left[A(s)^{1+s}\right], \\ A(s) &= \sum_{i=1}^{a} \pi_i S_i^{\frac{1}{1+s}}, \end{aligned}$$

where each S_i is density operator which corresponds to the output state of the classical-quantum channel $i \to S_i$ from the set of the input alphabet $A = \{1, 2, \cdots, a\}$ to the set of the output quantum states in the Hilbert space $\mathcal{H}$.

2. A sufficient condition on concavity of the auxiliary function $E_q(\pi, s)$

Proposition 2.1 ([8]) For any real number s $(-1 < s \le 1)$, density operators $S_i (i = 1, \cdots, a)$ and *a priori* probability $\pi = \{\pi_i\}_{i=1}^{a}$, if the trace inequality

$$\mathrm{Tr}\left[A(s)^s \sum_{j=1}^{a} \pi_j S_j^{\frac{1}{1+s}} (\log S_j^{\frac{1}{1+s}})^2\right] - \mathrm{Tr}\left[A(s)^{-s+1} \left(\sum_{i=1}^{a} \pi_i H(S_i^{\frac{1}{1+s}})\right)^2\right] \ge 0 \quad (2)$$

holds, then the auxiliary function

$$E_q(\pi, s) = -\log\left[\mathrm{Tr}\left\{\left(\sum_{i=1}^{a} \pi_i S_i^{\frac{1}{1+s}}\right)^{1+s}\right\}\right] \quad (3)$$

is concave in s. Where $H(x) = -x \log x$ is the operator entropy [7].

The condition (2) can be weakened by

$$\mathrm{Tr}\left[A(s)^s \sum_{j=1}^{a} \pi_j A_j (\log A_j)^2\right] - \mathrm{Tr}\left[A(s)^{-s+1} \left(\sum_{i=1}^{a} \pi_i H(A_j)\right)^2\right] \ge 0 \quad (4)$$

for $0 \le A_j \le I$.

3. Previous results

In this section, we review the previous results on the present problem limited the parameter $s \in [0, 1]$. In the previous section, we found that in order to prove the concavity of the auxiliary function Eq.(3), we have only to prove the sufficient condtion (2) for any a, s, $(0 \le s \le 1)$ and any density matrices S_i. If the condition (4) is proven, we see the condition (2) holds. Thus we considered the simple case $a = 2$ and then we put $A = S_1^{\frac{1}{1+s}}$, $B = S_2^{\frac{1}{1+s}}$ and $\pi_1 = \pi_2 = \frac{1}{2}$ for simplicity. Thus our problem could be deformed as follows:

Problem 3.1 Prove

$$\mathrm{Tr}[(A+B)^s\left\{A(\log A)^2+B(\log B)^2\right\}]$$
$$-\mathrm{Tr}[(A+B)^{-1+s}(A\log A+B\log B)^2]\geq 0 \ (5)$$

for any $s, (0\leq s\leq 1)$ and two positive matrices $A\leq I$ and $B\leq I$.

For this problem, we obtained the following results.

Theorem 3.2 ([12]) For two positive matrices $A\leq I$ and $B\leq I$, Eq.(5) holds in the case of $s=1$:

$$\mathrm{Tr}[(A+B)\left\{A(\log A)^2+B(\log B)^2\right\}]$$
$$-\mathrm{Tr}[(A\log A+B\log B)^2]\geq 0.$$

Theorem 3.3 ([12]) For two positive matrices $A\leq I$ and $B\leq I$, Eq.(5) holds in the case of $s=0$:

$$\mathrm{Tr}[\left\{A(\log A)^2+B(\log B)^2\right\}]$$
$$-\mathrm{Tr}[(A+B)^{-1}(A\log A+B\log B)^2]\geq 0.$$

To prove the above theorem, we used the Jensen's operator inequality:

Lemma 3.4 ([10, 11]) For the continuous function $f:[0,\alpha)\to\mathbf{R}$, $(0<\alpha\leq\infty)$, the following statements are equivalent.

(i) f is operator convex and $f(0)\leq 0$.

(ii) For the bounded linear operators K_i, $(i=1,2,\cdots,n)$ satisfying $\sigma(K_i)\subset[0,\alpha)$, where $\sigma(Z)$ represents the set of all spectrums of the bounded linear operator Z, and the bounded linear operators C_i, $(i=1,2,\cdots,n)$ satisfying $\sum_{i=1}^n C_i^*C_i\leq I$, we have

$$f(\sum_{i=1}^n C_i^*K_iC_i)\leq\sum_{i=1}^n C_i^*f(K_i)C_i.$$

With the help of the following lemma, we could obtained the following Theorem 3.6 as a kind of the interpolation between Theorem 3.2 and Theorem 3.3.

Lemma 3.5 Suppose the positive numbers t_1,t_2,a_1,a_2,b_1 and b_2 satisfy the following two conditions.

(i) $t_1a_1+t_2a_2\geq b_1+b_2$

(ii) $a_1+a_2\geq t_1^{-1}b_1+t_2^{-1}b_2$

Then for any $0\leq s\leq 1$ we have

$$t_1^sa_1+t_2^sa_2\geq t_1^{-1+s}b_1+t_2^{-1+s}b_2.$$

Theorem 3.6 Suppose A and B are 2×2 positive matrices. Then for any $0\leq s\leq 1$ we have

$$\mathrm{Tr}[(A+B)^s\left\{A(\log A)^2+B(\log B)^2\right\}]$$
$$-\mathrm{Tr}[(A+B)^{-1+s}(A\log A+B\log B)^2]\geq 0.$$

Remark 3.7 In the process of the proof of Theorem 3.3, we found the operator inequality holds in the case of $s=0$. However, we did not know whether the following matrix inequalities

$$(A+B)^{1/2}\left\{A(\log A)^2+B(\log B)^2\right\}(A+B)^{1/2}$$
$$\geq(A\log A+B\log B)^2 \quad (6)$$

or

$$\left\{A(\log A)^2+B(\log B)^2\right\}^{1/2}(A+B)$$
$$\times\left\{A(\log A)^2+B(\log B)^2\right\}^{1/2}$$
$$\geq(A\log A+B\log B)^2 \quad (7)$$

corresponding to the case of $s=1$ for any two positive matrices $A\leq I$ and $B\leq I$ hold or not. We have not yet found any counter-examples, namely the examples that the matrix ineqalities both Eq.(6) and Eq.(7) are not satisfied simultaneously, for some positive matrices $A\leq I$ and $B\leq I$. For this question, T.Furuta give the answer by finding the counter example [13].

We expected that our Lemma 3.5 can be extended to the general $n\geq 3$, where n represents the number of the eigenvalues given in the Schatten decomposition of $A+B$ in Theorem 3.6.

$$A+B=\sum_n t_n|\phi_n\rangle\langle\phi_n|, \quad (8)$$

where $\{t_n\}$ are the eigenvalues of $A+B$, $\{|\phi_n\rangle\}$ are the corresponding eigenvectors. However it is impossible to prove it, because we have a counter-example for such a generalization. This means that our Lemma 3.5 can not be extended to the general case of $n\geq 3$. Therefore one must produce an another method to prove Theorem 3.6 for any $n\times n$ positive matrices A and B. In such a situation, J.I.Fujii solved this problem by proving the remarkable trace inequality [14, 15, 16]. Using this method, the open problem given in [6, 4] was completely solved in the case of $s\in[0,1]$ by J.I.Fujii, R.Nakamoto and K.Yanagi [17] in the following way.

Definition 3.8 ([15, 16]) Let f, g be real valued continuous functions. Then (f, g) is called a monotone (*resp.* antimonotone) pair of functions on the domain $D \subset \mathbb{R}$ if

$$(f(a) - f(b))(g(a) - g(b)) \geq 0 \ \ (resp. \leq)$$

for any $a, b \in D$.

Proposition 3.9 ([15, 16, 14]) If (f, g) is a monotone (resp. antimonotone) pair, then

$$\mathrm{Tr}[f(A)Xg(A)X] \leq \mathrm{Tr}[f(A)g(A)X^2] \ \ (resp. \geq)$$

for selfadjoint matrices A and X whose spectra are included in D.

Theorem 3.10 ([14]) For $0 \leq A, B \leq I$ and $s \geq 0$, we have

$$\begin{aligned} &\mathrm{Tr}[(A+B)^s(A(\log A)^2 + B(\log B)^2)] \\ &-\mathrm{Tr}[(A+B)^{s-1}(A\log A + B\log B)^2] \geq 0. \end{aligned}$$

Theorem 3.11 ([17]) For the operators $0 \leq A_i \leq I$, the probability distribution π_i, $(i = 1, \ldots, a)$, and $s \geq 0$ we have

$$\begin{aligned} &\mathrm{Tr}\left[\left(\sum_{k=1}^a \pi_k A_k\right)^s \sum_{i=1}^a \pi_i A_i \left(\log A_i\right)^2\right] \\ &-\mathrm{Tr}\left[\left(\sum_{k=1}^a \pi_k A_k\right)^{s-1} \left(\sum_{i=1}^a \pi_i A_i \log A_i\right)^2\right] \geq 0, \end{aligned}$$

4. Study on the case of $s \in (-1, 0)$

Our remained problem is the following.

Problem 4.1 Prove the trace inequality

$$\begin{aligned} &\mathrm{Tr}\left[A(s)^s \left\{\sum_{j=1}^a \pi_j S_j^{\frac{1}{1+s}} \left(\log S_j^{\frac{1}{1+s}}\right)^2\right\}\right] \\ &-\mathrm{Tr}\left[A(s)^{-1+s} \left\{\sum_{j=1}^a \pi_j H\left(S_j^{\frac{1}{1+s}}\right)\right\}^2\right] \geq 0 \quad (9) \end{aligned}$$

for any real number s $(-1 < s < 0)$, any density matrices $S_i (i = 1, \cdots, a)$ and any probability distributions $\pi = \{\pi_i\}_{i=1}^a$, under the assumption that $A(s) \equiv \sum_{i=1}^a \pi_i S_i^{\frac{1}{1+s}}$ is invertible. Or find the counter example of the inequality (9).

As similar way of the previous section, this problem can be weakened in the following.

Problem 4.2 Prove

$$\begin{aligned} &\mathrm{Tr}\left[(A+B)^s \left\{A(\log A)^2 + B(\log B)^2\right\}\right] \\ &-\mathrm{Tr}\left[(A+B)^{-1+s}(A\log A + B\log B)^2\right] \geq 0 \quad (10) \end{aligned}$$

for any s, $(-1 < s < 0)$ and two positive matrices $A \leq I$ and $B \leq I$. Or find the counter example of the inequality (10).

Here we give a counter-example of the inequality (10) for $s \in (-1, 0)$. Putting $A = e^{-X}$ and $B = e^{-Y}$ for $X, Y > 0$, the inequality (10) is equivalent to

$$\begin{aligned} &\mathrm{Tr}\left[\left(e^{-X} + e^{-Y}\right)^s \left(e^{-X}X^2 + e^{-Y}Y^2\right)\right] \\ &-\mathrm{Tr}\left[\left(e^{-X} + e^{-Y}\right)^{s-1} \left(e^{-X}X + e^{-Y}Y\right)^2\right] \geq 0. \quad (11) \end{aligned}$$

If we take

$$X = \begin{pmatrix} 5 & 3 \\ 3 & 2 \end{pmatrix}, Y = \begin{pmatrix} 4 & 0 \\ 0 & 25 \end{pmatrix}, s = -1/2,$$

then the left hand side of the inequality (11) takes -0.441722.

However this counter example does not necessarily assure that the concavity of the auxiliary function of the quantum reliability function does not hold. In order to show that the concavity of the auxiliary function of the quantum reliability function does not hold, we must find the counter example of the original trace inequality (9) for $a = 2$. However we have not found such counter examples yet.

Acknowledgments

S.F. was partially supported by the Japanese Ministry of Education, Science, Sports and Culture, Grant-in-Aid for Encouragement of Young Scientists (B), 17740068.

References

[1] R.G.Gallager, Information theory and reliable communication, John Wiley and Sons,1968.

[2] S.Arimoto, On the converse to the coding theorem for discrete memoryless channels, IEEE Trans.IT, Vol.19, pp.357-359, 1973.

[3] A.S.Holevo, The capacity of quantum channel with general signal states, IEEE Trans.IT, Vol.44, pp.269-273, 1998.

[4] T.Ogawa and H.Nagaoka, Strong converse to the quantum channel coding theorem, IEEE Trans.IT, Vol.45, pp.2486-2489, 1999.

[5] M.V.Burnashev and A.S.Holevo, On the reliability function for a quantum communication channel, Problems of Information Transmission, Vol.34,pp.97-107, 1998.

[6] A.S.Holevo, Reliability function of general classical-quantum channel, IEEE Trans.IT, Vol.46, pp.2256-2261, 2000.

[7] M.Nakamura and H.Umegaki, A note on the entropy for operator algebras, Proc. Jap. Acad., Vol.37,pp.149-154, 1961.

[8] S.Furuichi, K.Yanagi and K.Kuriyama, A sufficient condition on concavity of the auxiliary function appearing in quantum reliability function, INFROMATION,Vol.6,pp.71-76, 2003.

[9] J.I.Fujii and E.Kamei, Relative operator entropy in noncommutative information theory, Math. Japonica, Vol.34,pp.341-348, 1989.

[10] T. Ando, Topics on operator inequalities, Lecture Notes (mimeographed), Hokkaido Univ., Sapporo, 1978.

[11] F. Hansen and G. K. Pedersen, Jensen's inequality for operators and Lowner's theorem, Math. Ann., Vol.258, pp.229-241, 1982.

[12] K.Yanagi, S.Furuichi, and K.Kuriyama, On trace inequalities and their applications to noncommutative communication theory, Linear Alg.Appl., Vol.395, pp.351-359, 2005.

[13] T.Furuta, A counterexample to the question proposed by Yanagi-Furuichi-Kuriyama on matrix inequalities and related counterexamples,Linear Alg.Appl., Vol.397,pp.345-353, 2005.

[14] J.I.Fujii, A trace inequality arising from quantum information theory, Linear Alg.Appl.,Vol.400,pp.141-146, 2005.

[15] J.C.Bourin, Some inequalities for norms on matrices and operators, Linear Alg.Appl., Vol.292, pp.139-154, 1999.

[16] J.C.Bourin, Compressions, Dilations and matrix inequalities, RGMIA Monographs, Victoria University, 2004.

[17] J.I.Fujii, R.Nakamoto and K.Yanagi, Concavity of the auxiliary function appearing in quantum reliability function, IEEE Trans. IT, Vol.52, pp.3310-3313, 2006.