Concavity of the Auxiliary Function Appearing in Quantum Reliability Function

Jun Ichi Fujii, Ritsuo Nakamoto, and Kenjiro Yanagi, *Member, IEEE*

***Abstract*—Reliability functions characterize the asymptotic behavior of the error probability for transmission of data on a channel. Holevo introduced the quantum channel, and gave an expression for a random-coding lower bound involving an auxiliary function. Holevo, Ogawa, and Nagaoka conjectured that this auxiliary function is concave. Here we give a proof of this conjecture.**

***Index Terms*—Quantum information theory, quantum reliability function, random coding exponent.**

I. Introduction

In classical information theory, there are two commonly used methods for proving the channel coding theorem. One uses typical sequences (see [4]), the other is a direct method involving reliability functions (see [8]). In quantum information theory, the channel coding theorem for classical-quantum channels was obtained by Holevo in [10] by means of a generalization of the typical sequence method. So far, there is no proof based on quantum reliability functions. In a classical-quantum channel, each symbol i of our alphabet $\{1, 2 \ldots, a\}$ is transmitted in the form of a density operator S_i. The receiver can infer which word of a code (a set of words) is transmitted by making a joint quantum measurement on the channel outputs. For such a channel, the quantum reliability function is defined by

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

where C is the classical-quantum capacity obtained by Holevo, R is the transmission rate $R = \frac{\log_2 M}{n}$ with n the length of the code and M the number of code words, and $P_e(M, n)$ is the minimum average error probability $\bar{P}(\mathcal{W}, \mathcal{X})$ or the worst-case error probability $P_{\max}(\mathcal{W}, \mathcal{X})$. These are defined by

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

where $\mathcal{W} = \{w^1, w^2, \ldots, w^M\}$ ranges over codes, $\mathcal{X} = \{X_i\}(\sum_i X_i \le I)$ ranges over (partial) positive operator valued measurements, and

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

is the usual error probability associated with $\mathcal{X} = \{X_j\}$, where S_{w^j} is the density operator corresponding to w^j. Let $\mathcal{E}$ denote expectations with respect to codes $\mathcal{W}$ whose codewords are chosen i.i.d with probability $\mathcal{P}(w^i = (i_1, i_2, \ldots, i_n)) = \pi_{i_1} \ldots \pi_{i_n}$ for an a priori probability distribution $\pi = \{\pi_i\}$. In [3], it was conjectured that the random coding bound on the channel capacity is determined by the following:

$$\mathcal{E} \min_X \bar{P}(\mathcal{W}, \mathcal{X}) \le c \inf_{0<s\le 1} (M-1)^s \left[\operatorname{Tr}\left(\sum_{i=1}^{a} \pi_i S_i^{\frac{1}{1+s}}\right)^{1+s}\right]^n \tag{2}$$

This bound holds for pure states S_i, in which case $S_i^{\frac{1}{1+s}} = S_i$ and $c = 2$. For commuting S_i it reduces to the classical bound of Theorem 5.6.2 in [8] with $c = 1$. By setting $M = 2^{nR}$, it implies a lower bound on the quantum reliability function defined in (1). In particular

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

where

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

with π ranging over probability distributions. In analogy to the classical case, we expect that E_q satisfies the following properties

a) $E_q(\pi, 0) = 0$.
b) $\frac{\partial E_q(\pi,s)}{\partial s}|_{s=0} = I(X;Y)$, where $I(X;Y)$ presents the mutual information.
c) $E_q(\pi, s) > 0 (0 < s \le 1)$. $E_q(\pi, s) < 0 (-1 < s < 0)$.
d) $\frac{\partial E_q(\pi,s)}{\partial s} > 0, (-1 < s \le 1)$.
e) $\frac{\partial^2 E_q(\pi,s)}{\partial s^2} \le 0, (-1 < s \le 1)$.

Of these properties, (a), (b), (c), and (d) are proven in [11][12]. Property (e) was conjectured in [11][12] and implies concavity of the auxiliary function $E_q(\pi, s)$ in s. Before this work, (e) was shown to be true for the case where the S_i are pure [3], and where instead of random coding, one uses the expurgation method [11].

II. Concavity of $E_q(\pi, a)$

We state the main theorem.

Theorem 2.1: $E_q(\pi, s)$ is concave in s for $s \in [0, 1]$.

However we still have the conjecture that $E_q(\pi, s)$ is concave in s for $s \in (-1, 0]$. A sufficient condition on concavity of the auxiliary function $E_q(\pi, s)$ is the following proposition proven in [7]. Here $H(x) = -x \log x$ is the matrix entropy.

Proposition 2.2 ([7]): Let $S_i \quad (i = 1, \ldots, a)$ be density matrices and $\pi = \{\pi_i\}_{i=1}^a$ a probability distribution such that $A(s) = \sum_{i=1}^{a} \pi_i S_i^{1/(1+s)}$ is invertible. If the trace inequality

$$\operatorname{Tr}\left[A(s)^s \left\{\sum_{j=1}^{a} \pi_j S_j^{\frac{1}{1+s}} \left(\log S_j^{\frac{1}{1+s}}\right)^2\right\} - A(s)^{-1+s}\left\{\sum_{j=1}^{a} \pi_j H\left(S_j^{\frac{1}{1+s}}\right)\right\}^2\right] \ge 0. \tag{3}$$

holds for s with $-1 < s \le 1$, then the auxiliary function $E_q(\pi, s)$ is concave at s.

We note that our assumption that $A(s)$ is invertible is generic, because $A(s)$ becomes invertible if we have at least one invertible S_i. Moreover, $A(s)$ may be invertible even if none of the S_i are invertible. It suffices that the span of the support of the $\pi_i S_i$ is the full space. In

Manuscript received January 21, 2005; revised March 14, 2006. The material in this correspondence was presented at the IEEE International Symposium on Information Theory, Adelaide, Australia, September 2005.

J. I. Fujii is with the Department of Arts and Sciences (Information Science), Osaka Kyoiku University, Osaka 582-8582, Japan; (e-mail: fujii@cc.osaka-kyoiku.ac.jp).

R. Nakamoto is with the Faculty of Engineering, Ibaraki University, Hitachi City 316-8511, Japan; (e-mail: nakamoto@base.ibaraki.ac.jp).

K. Yanagi is with the Division of Applied Mathematical Science, Graduate School of Science and Engineering, Yamaguchi University, Ube City 755-8611, Japan (e-mail: yanagi@yamaguchi-u.ac.jp).

Communicated by E. Knill, Associate Editor for Quantum Information Theory.

Digital Object Identifier 10.1109/TIT.2006.876248

[13], Yanagi, Furuichi, and Kuriyama proved the concavity of $E_q(\pi, s)$ in the special case $a = 2$ with $\pi_1 = \pi_2 = \frac{1}{2}$ under the assumption that the dimension of $\mathcal{H}$ is two by proving the trace inequality (3). And recently in [6], Fujii proved (3) in the case $a = 2$ with $\pi_1 = \pi_2 = \frac{1}{2}$ for any dimension of $\mathcal{H}$. In this paper we prove (3) for all a, any dimension of $\mathcal{H}$ and $0 \le s \le 1$, which, according to Proposition 2.2, implies that $E_q(\pi, \cdot)$ is concave on $[0, 1]$.

Definition 2.3 ([1], [2]): 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

$$(f(a) - f(b))(g(a) - g(b)) \ge 0 \quad (\text{resp.} \le)$$

for all $a, b \in D$.

Proposition 2.4 ([1], [2], [6]): If (f, g) is a monotone (resp. antimonotone) pair, then

$$\mathrm{Tr}\,[f(A)Xg(A)X] \le \mathrm{Tr}\,[f(A)g(A)X^2] \quad (\text{resp.} \ge)$$

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

Proofs of Proposition 2.2 and 2.4 are given in Appendix for the reader's convenience.

Proof of Theorem 2.1: We recall the following operator Jensen's inequality (e.g., [5], [9]):

If $\sum_{i=1}^{a} C_i^* C_i = I$, then

$$\sum_{i=1}^{a} C_i^* X_i^2 C_i \ge \left(\sum_{i=1}^{a} C_i^* X_i C_i\right)^2$$

holds for all Hermitian operators X_i, since $f(x) = x^2$ is operator convex on any interval. We put

$$X_i = \log A_i, \; C_i = (\pi_i A_i)^{1/2} \left(\sum_{k=1}^{a} \pi_k A_k\right)^{-1/2}$$

for $i = 1, 2, \ldots, a$. Since $\sum_{i=1}^{a} C_i^* C_i = I$, we have

$$\begin{aligned} &\sum_{i=1}^{a} \left(\sum_{k=1}^{a} \pi_k A_k\right)^{-1/2} (\pi_i A_i)^{1/2} (\log A_i)^2 \\ &\times (\pi_i A_i)^{1/2} \left(\sum_{k=1}^{a} \pi_k A_k\right)^{-1/2} \\ &\quad \ge \left(\sum_{i=1}^{a} \left(\sum_{k=1}^{a} \pi_k A_k\right)^{-1/2} (\pi_i A_i)^{1/2} \log A_i \right. \\ &\quad\quad \left. \times (\pi_i A_i)^{1/2} \left(\sum_{k=1}^{a} \pi_k A_k\right)^{-1/2}\right)^2. \end{aligned}$$

And so we have

$$\begin{aligned} &\left(\sum_{k=1}^{a} \pi_k A_k\right)^{-1/2} \sum_{i=1}^{a} (\pi_i A_i)^{1/2} (\log A_i)^2 \\ &\times (\pi_i A_i)^{1/2} \left(\sum_{k=1}^{a} \pi_k A_k\right)^{-1/2} \\ &\quad \ge \left(\left(\sum_{k=1}^{a} \pi_k A_k\right)^{-1/2} \left(\sum_{i=1}^{a} \pi_i A_i \log A_i\right) \right. \\ &\quad\quad \left. \times \left(\sum_{k=1}^{a} \pi_k A_k\right)^{-1/2}\right)^2. \end{aligned}$$

Hence it follows that

$$\begin{aligned} &\sum_{i=1}^{a} (\pi_i A_i)^{1/2} (\log A_i)^2 (\pi_i A_i)^{1/2} \\ &\quad \ge \left(\sum_{i=1}^{a} \pi_i A_i \log A_i\right) \left(\sum_{k=1}^{a} \pi_k A_k\right)^{-1} \\ &\quad\quad \times \left(\sum_{i=1}^{a} \pi_i A_i \log A_i\right). \end{aligned}$$

Since $(\pi_i A_i)^{1/2} (\log A_i)^2 (\pi_i A_i)^{1/2} = \pi_i A_i (\log A_i)^2$, we have

$$\begin{aligned} &\left(\sum_{k=1}^{a} \pi_k A_k\right)^{s/2} \sum_{i=1}^{a} \pi_i A_i (\log A_i)^2 \\ &\times \left(\sum_{k=1}^{a} \pi_k A_k\right)^{s/2} \\ &\quad \ge \left(\sum_{k=1}^{a} \pi_k A_k\right)^{s/2} \left(\sum_{i=1}^{a} \pi_i A_i \log A_i\right) \\ &\quad\quad \times \left(\sum_{k=1}^{a} \pi_k A_k\right)^{-1} \left(\sum_{i=1}^{a} \pi_i A_i \log A_i\right) \\ &\quad\quad \times \left(\sum_{k=1}^{a} \pi_k A_k\right)^{s/2}. \end{aligned}$$

Thus

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

Since $f(x) = x^s \; (s \ge 0)$ and $g(x) = x^{-1}$, it is clear that (f, g) is an antimonotone pair. By Proposition 2.4,

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

Q.E.D.

Appendix A

Proof of Proposition 2.2 from [7].: This is a copy of the proof in [7]. We put

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

Since

$$\frac{\partial E_q(\pi,s)}{\partial s} = -G(s)^{-1}G'(s)$$

we have

$$\frac{\partial^2 E_q(\pi,s)}{\partial s^2} = G(s)^{-2}\left(G'(s)^2 - G(s)G''(s)\right).$$

By the use of the formula [11] for the operator valued function $A(s)$ w.r.t. the real number s

$$\frac{d}{ds}\mathrm{Tr} f(s,A(s)) = \mathrm{Tr} f'_s(s,A(s)) + \mathrm{Tr} f'_A(s,A(s))A'(s)$$

we have

$$\begin{aligned} G'(s) &= \mathrm{Tr}\left[A(s)^s\left(A(s)\log A(s) + (1+s)A'(s)\right)\right] \\ &= -\mathrm{Tr}\left[A(s)^s \Delta H(\pi,s)\right], \end{aligned}$$

where

$$\Delta H(\pi,s) = H(A(s)) - \sum_{i=1}^{a}\pi_i H(S_i^{\frac{1}{1+s}}).$$

By some simple calculations, we have

$$\begin{aligned} &G''(s) \\ &\quad = \mathrm{Tr}\left[A(s)^{s-1}\left\{A(s)^2(\log A(s))^2 + s(1+s)A'(s)^2\right\}\right] \\ &\qquad + \mathrm{Tr}\left[A(s)^{s-1}\left\{A(s)(2(1+(1+s)\log A(s))A'(s)\right.\right. \\ &\qquad \left.\left. +\ (1+s)A''(s))\right\}\right] \end{aligned} \tag{4}$$

where

$$A'(s) = -\frac{1}{(1+s)^2}\sum_{i=1}^{a}\pi_i S_i^{\frac{1}{1+s}}\log S_i \tag{5}$$

$$A''(s) = \frac{1}{(1+s)^4}\sum_{i=1}^{a}\pi_i S_i^{\frac{1}{1+s}}(2(1+s)\log S_i + (\log S_i)^2). \tag{6}$$

Substituting (5) and (6) into (4), we have

$$\begin{aligned} &G''(s) \\ &= \mathrm{Tr}\left[A(s)^{s-1}\left\{H(A(s))^2 + \frac{s}{1+s}\left(\sum_{i=1}^{a}\pi_i H(S_i^{\frac{1}{1+s}})\right)^2\right.\right. \\ &\quad -\ 2H(A(s))\sum_{i=1}^{a}\pi_i H(S_i^{\frac{1}{1+s}}) \\ &\quad \left.\left. +\ \frac{1}{1+s}\sum_{i=1}^{a}\pi_i S_i^{\frac{1}{1+s}}\sum_{j=1}^{a}\pi_j S_j^{\frac{1}{1+s}}(\log S_j^{\frac{1}{1+s}})^2\right\}\right] \\ &= \mathrm{Tr}\left[A(s)^{s-1}\left\{H(A(s))^2 - 2H(A(s))\sum_{i=1}^{a}\pi_i H(S_i^{\frac{1}{1+s}})\right.\right. \\ &\quad +\ \left(\sum_{i=1}^{a}\pi_i H(S_i^{\frac{1}{1+s}})\right)^2 \\ &\quad +\ \frac{1}{1+s}\sum_{i=1}^{a}\pi_i S_i^{\frac{1}{1+s}}\sum_{j=1}^{a}\pi_j S_j^{\frac{1}{1+s}}(\log S_j^{\frac{1}{1+s}})^2 \\ &\quad \left.\left. -\ \frac{1}{1+s}\left(\sum_{i=1}^{a}\pi_i H(S_i^{\frac{1}{1+s}})\right)^2\right\}\right] \end{aligned} \tag{7}$$

By the Cauchy–Schwarz inequality, we have

$$G'(s)^2 - G(s)\widetilde{G''}(s) \le 0$$

where

$$\widetilde{G''}(s) = \mathrm{Tr}[A(s)^{-1+s}\Delta H(\pi,s)^2]. \tag{8}$$

Therefore if we have

$$G'(s)^2 - G(s)G''(s) \le G'(s)^2 - G(s)\widetilde{G''}(s)$$

that is,

$$\widetilde{G''}(s) \le G''(s) \tag{9}$$

then the theorem holds. From (7) and (8), (9) can be deformed,

$$\begin{aligned} &\mathrm{Tr}\left[A(s)^{s-1}\left\{-H(A(s))\sum_{i=1}^{a}\pi_i H(S_i^{\frac{1}{1+s}})\right.\right. \\ &\quad \left.\left. +\ \sum_{i=1}^{a}\pi_i H(S_i^{\frac{1}{1+s}})H(A(s))\right\}\right] \\ &\quad +\ \frac{1}{1+s}\mathrm{Tr}\left[A(s)^{s-1}\left\{\sum_{i=1}^{a}\pi_i S_i^{\frac{1}{1+s}}\right.\right. \\ &\quad \times\sum_{j=1}^{a}\pi_j S_j^{\frac{1}{1+s}}(\log S_j^{\frac{1}{1+s}})^2 \\ &\quad \left.\left. -\ \left(\sum_{i=1}^{a}\pi_i H(S_i^{\frac{1}{1+s}})\right)^2\right\}\right] \ge 0. \end{aligned} \tag{10}$$

Since $H(A(s))$ commutes with $A(s)^{-1+s}$, the first term of (10) equal to 0 so that (10) can be rewritten in the following:

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

which implies the proposition. Q.E.D.

Proof of Proposition 2.4: We may assume that A is diagonal. Let $A = diag(t_1, t_2, \ldots, t_n)$ and $X = (x_{ij})$.

If (f,g) is a monotone pair, then

$$f(a)g(b) + f(b)g(a) \le f(a)g(a) + f(b)g(b)$$

for any $a, b \in D$. Then we have the following;

$$\begin{aligned} &\mathrm{Tr}\left[f(A)Xg(A)X\right] \\ &\quad = \sum_{k=1}^{n} f(t_k)g(t_k)x_{kk}^2 + \sum_{k<j}\{f(t_k)g(t_j) + f(t_j)g(t_k)\}x_{kj}^2 \\ &\quad \le \sum_{k=1}^{n} f(t_k)g(t_k)x_{kk}^2 + \sum_{k<j}\{f(t_k)g(t_k) + f(t_j)g(t_j)\}x_{kj}^2 \\ &\quad = \mathrm{Tr}\left[f(A)g(A)X^2\right]. \end{aligned}$$

If (f,g) is an antimonotone pair, then we obtain the result by the same method. Q.E.D.

REFERENCES

[1] J. C. Bourin, "Some inequalities for norms on matrices and operators," *Linear Alg. Appl.*, vol. 292, pp. 139–154, 1999.

[2] J. C. Bourin, *Compressions, Dilations and matrix inequalities, RGMIA Monographs*. Victoria, Australia: Victoria University, 2004.

[3] M. V. Burnashev and A. S. Holevo, "On the reliability function for a quantum communication channel," *Prob. Inf. Trans.*, vol. 34, no. 2, pp. 97–107, 1998.

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

[5] J. I. Fujii and M. Fujii, "Jensen's Inequalities on any interval for operators," in *Proc. 3rd Int. Conf. Nonlinear Anal. Convex Anal.*, 2004, pp. 29–39.

[6] J. I. Fujii, "A trace inequality arising from quantum information theory," *Linear Alg. Appl.*, vol. 400, pp. 141–146, 2005.

[7] S. Furuichi, K. Yanagi, and K. Kuriyama, "A sufficient condition on concavity of the auxiliary function appearing in quantum reliability function," *INFORMATION*, vol. 6, no. 1, pp. 71–76, 2003.

[8] R. G. Gallager, *Information theory and reliable communication*. New York: Wiley, 1968.

[9] F. Hansen and G. K. Pedersen, "Jensen's operator inequality," *Bull. London Math. Soc.*, vol. 35, pp. 553–564, 2003.

[10] A. S. Holevo, "The capacity of quantum channel with general signal states," *IEEE. Trans. Inf. Theory*, vol. 44, no. 1, pp. 269–273, 1998.

[11] A. S. Holevo, "Reliability function of general classical-quantum channel," *IEEE. Trans. Inf. Theory*, vol. 46, no. 6, pp. 2256–2261, 2000.

[12] T. Ogawa and H. Nagaoka, "Strong converse to the quantum channel coding theorem," *IEEE. Trans. Inf. Theory*, vol. 45, no. 7, pp. 2486–2489, 1999.

[13] 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.