Available online at www.sciencedirect.com
SCIENCE DIRECT®
Linear Algebra and its Applications 395 (2005) 351–359
LINEAR ALGEBRA AND ITS APPLICATIONS
www.elsevier.com/locate/laa

On trace inequalities and their applications to noncommutative communication theory

Kenjiro Yanagi [a], Shigeru Furuichi [b,*], Ken Kuriyama [a]

[a] *Yamaguchi University, Ube city, Yamaguchi 755-8611, Japan*
[b] *Tokyo University of Science, Yamaguchi 1-1-1, Daigakudori, Onoda 756-0884, Japan*

Received 15 March 2004; accepted 25 August 2004

Submitted by R.A. Brualdi

Abstract

Certain trace inequalities related to matrix logarithm are shown. These results enable us to give a partial answer of the open problem conjectured by A.S. Holevo. That is, concavity of the auxiliary function which appears in the random coding exponent as the lower bound of the quantum reliability function for general quantum states is proven in the case of $0 \leqslant s \leqslant 1$.

AMS classification: 15A42; 47A63; 94A05

Keywords: Trace inequalities; Concavity; Quantum reliability function; Noncommutative communication theory

1. Introduction

In noncommutative (quantum) communication theory, the concavity of the auxiliary function of the quantum reliability function has remained as an open question [6] and unsolved conjecture [8]. The auxiliary function $E(s),\ (-1 < s \leqslant 1)$ is defined by

* Corresponding author.
E-mail addresses: yanagi@yamaguchi-u.ac.jp (K. Yanagi), furuichi@ed.yama.tus.ac.jp (S. Furuichi), kuriyama@yamaguchi-u.ac.jp (K. Kuriyama).

0024-3795/$ - see front matter
doi:10.1016/j.laa.2004.08.027

$$E(s) \equiv -\log\left\{\mathrm{Tr}\left[\left(\sum_{i=1}^{a}\pi_i S_i^{\frac{1}{1+s}}\right)^{1+s}\right]\right\}, \tag{1}$$

where each S_i is the density matrix and each π_i is nonnegative number satisfying $\sum_{i=1}^{a}\pi_i = 1$. See [2,6] for details on quantum reliability function theory. For the above problem, we gave the sufficient condition on concavity of the auxiliary function in the previous paper [4].

Proposition 1.1 [4]. *If the trace inequality*

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

holds for any real number s $(-1 < s \leqslant 1)$, *any density matrices* S_i $(i = 1, \ldots, a)$ *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, then the auxiliary function* $E(s)$ *defined by Eq.* (1) *is concave for all* s $(-1 < s \leqslant 1)$. *Where* $H(x) = -x\log x$ *is the matrix entropy introduced in* [7].

We note that our assumption "$A(s)$ is invertible" is not so special condition, because $A(s)$ becomes invertible if we have one invertible S_i at least. Moreover, we have the possibility such that $A(s)$ becomes invertible even if all S_i is not invertible for all $\pi_i \neq 0$.

In the present paper, we show some trace inequalities related to matrix logarithm, and then give a partial solution of the open problem in noncommutative communication theory as an application of them.

2. Main results

In the previous section, we found that in order to prove the concavity of the auxiliary function Eq. (1), we have only to prove the sufficient condition Eq. (2) for any a, s $(-1 < s \leqslant 1)$ and any density matrices S_i. For this purpose, we consider 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 can be deformed as follows:

Problem 2.1. Prove

$$\mathrm{Tr}\left[(A+B)^s\{A(\log A)^2 + B(\log B)^2\} - (A+B)^{-1+s}(A\log A + B\log B)^2\right] \geqslant 0 \tag{3}$$

for any s, $(-1 < s \leqslant 1)$ and two positive matrices $A \leqslant I$ and $B \leqslant I$.

Theorem 2.2. *For two positive matrices $A \leqslant I$ and $B \leqslant I$, Eq. (3) holds in the case of $s = 1$:*

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

Proof of Theorem 2.2. Eq. (3) can be directly calculated by

$$\begin{aligned}
&\mathrm{Tr}\left[(A+B)^s\{A(\log A)^2 + B(\log B)^2\}\right] \\
&\quad - \mathrm{Tr}\left[(A+B)^{-1+s}(A\log A + B\log B)^2\right] \\
&\quad\quad = \mathrm{Tr}\left[(A+B)^{-1+s}(A+B)\{A(\log A)^2 + B(\log B)^2\}\right] \\
&\quad\quad\quad - \mathrm{Tr}\left[(A+B)^{-1+s}(A\log A + B\log B)^2\right] \\
&\quad\quad = \mathrm{Tr}\left[(A+B)^{-1+s}\{A^2(\log A)^2 + AB(\log B)^2\right. \\
&\quad\quad\quad \left. + BA(\log A)^2 + B^2(\log B)^2\}\right] - \mathrm{Tr}\left[(A+B)^{-1+s}\{A^2(\log A)^2\right. \\
&\quad\quad\quad \left. + A\log AB\log B + B\log BA\log A + B^2(\log B)^2\}\right] \\
&\quad\quad = \mathrm{Tr}\left[(A+B)^{-1+s}\{AB(\log B)^2 + BA(\log A)^2\}\right] \\
&\quad\quad\quad - \mathrm{Tr}\left[(A+B)^{-1+s}A\log AB\log B\right] \\
&\quad\quad\quad - \mathrm{Tr}\left[(A+B)^{-1+s}B\log BA\log A\right] \\
&\quad\quad = \mathrm{Tr}\left[(A+B)^{-1+s}AB(\log B)^2\right] + \mathrm{Tr}\left[(A+B)^{-1+s}BA(\log A)^2\right] \\
&\quad\quad\quad - 2\mathrm{Re}\,\mathrm{Tr}\left[A\log A(A+B)^{-1+s}B\log B\right].
\end{aligned} \tag{4}$$

Eq. (4) is further calculated for $s = 1$ such as

$$\begin{aligned}
&\mathrm{Tr}[AB(\log B)^2] + \mathrm{Tr}[BA(\log A)^2] - 2\mathrm{Re}\,\mathrm{Tr}[A\log AB\log B] \\
&\quad = \mathrm{Tr}[AB(\log B)^2] + \mathrm{Tr}[BA(\log A)^2] \\
&\quad\quad - 2\mathrm{Re}\,\mathrm{Tr}[B^{1/2}A^{1/2}\log AA^{1/2}B^{1/2}\log B] \\
&\quad \geqslant \mathrm{Tr}[AB(\log B)^2] + \mathrm{Tr}[BA(\log A)^2] \\
&\quad\quad - 2\left(\mathrm{Tr}[BA(\log A)^2]\right)^{1/2}\left(\mathrm{Tr}[AB(\log B)^2]\right)^{1/2} \\
&\quad = \left\{(\mathrm{Tr}[BA(\log A)^2])^{1/2} - (\mathrm{Tr}[AB(\log B)^2])^{1/2}\right\}^2 \geqslant 0.
\end{aligned}$$

Cuachy–Schwarz inequality:

$$|\mathrm{Tr}[X^*Y]|^2 \leqslant \mathrm{Tr}[X^*X]\mathrm{Tr}[Y^*Y]$$

for the matrices X and Y, has been applied in the above calculation. □

Remark 2.3. After the manner of Theorem 2.2, we can prove Eq. (2) in the case of $s = 1$ for any density matrices S_i and any probability distributions $\pi = \{\pi_i\}$, $(i = 1, 2, \ldots, a)$, since the left hand side of Eq. (2) can be directly calculated in the following

$$\sum_{i<j}\pi_i\pi_j\left\{\mathrm{Tr}\left[S_i^{\frac{1}{2}}S_j^{\frac{1}{2}}\left(\log S_j^{\frac{1}{2}}\right)^2\right]+\mathrm{Tr}\left[S_j^{\frac{1}{2}}S_i^{\frac{1}{2}}\left(\log S_i^{\frac{1}{2}}\right)^2\right]\right.$$
$$\left.-2\mathrm{Re}\,\mathrm{Tr}\left[S_i^{\frac{1}{2}}\log S_i^{\frac{1}{2}}S_j^{\frac{1}{2}}\log S_j^{\frac{1}{2}}\right]\right\}.$$

That is, the extended version of Theorem 2.2 holds, by applying Cuachy–Schwarz inequality to the third term in the brace of the above, after we slightly performed changes as similar as the proof of Theorem 2.2.

Theorem 2.4. *For two positive matrices $A \leqslant I$ and $B \leqslant I$, Eq.* (3) *holds in the case of $s = 0$:*

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

To prove Theorem 2.4 we require the following lemma.

Lemma 2.5 [1,5]. *For the continuous function $f : [0,\alpha) \to \mathbf{R}$, $(0 < \alpha \leqslant \infty)$, the following statements are equivalent.*

(i) *f is operator convex and $f(0) \leqslant 0$.*
(ii) *For the bounded linear operators K_i, $(i = 1, 2, \ldots, n)$ 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, $(i = 1, 2, \ldots, n)$ satisfying $\sum_{i=1}^n C_i^* C_i \leqslant I$, we have*

$$f\left(\sum_{i=1}^n C_i^* K_i C_i\right) \leqslant \sum_{i=1}^n C_i^* f(K_i) C_i.$$

Proof of Theorem 2.4. For $C_1 = A^{1/2}(A+B)^{-1/2}$ and $C_2 = B^{1/2}(A+B)^{-1/2}$, we have $C_1^*C_1 + C_2^*C_2 = I$. Note that $A \leqslant I$ and $B \leqslant I$. Then we set $f(t) = t^2$, $K_1 = -\log A$ and $K_2 = -\log B$ and then apply Lemma 2.5. Thus we have

$$\begin{aligned}&\{(A+B)^{-1/2}A^{1/2}(-\log A)A^{1/2}(A+B)^{-1/2}\\ &\quad + (A+B)^{-1/2}B^{1/2}(-\log B)B^{1/2}(A+B)^{-1/2}\}^2\\ &\quad \leqslant (A+B)^{-1/2}A^{1/2}(-\log A)^2A^{1/2}(A+B)^{-1/2}\\ &\qquad + (A+B)^{-1/2}B^{1/2}(-\log B)^2B^{1/2}(A+B)^{-1/2}.\end{aligned}$$

Since $[A^{1/2}, \log A] = 0$ and $[B^{1/2}, \log B] = 0$, we have

$$\begin{aligned}&\{(A+B)^{-1/2}(-A\log A - B\log B)(A+B)^{-1/2}\}^2\\ &\quad \leqslant (A+B)^{-1/2}\{A(-\log A)^2 + B(-\log B)^2\}(A+B)^{-1/2}.\end{aligned}$$

That is,

$$(A+B)^{-1/2}(A\log A+B\log B)(A+B)^{-1}(A\log A+B\log B)(A+B)^{-1/2} \leqslant (A+B)^{-1/2}\{A(\log A)^2+B(\log B)^2\}(A+B)^{-1/2}.$$

Thus we have

$$(A\log A+B\log B)(A+B)^{-1}(A\log A+B\log B) \leqslant A(\log A)^2+B(\log B)^2. \tag{5}$$

Therefore, if we take the trace in the both sides, then the proof is completed. □

Remark 2.6. After the manner of Theorem 2.4, we can prove Eq. (2) in the case of $s=0$ for any density matrices S_i and any probability distributions $\pi=\{\pi_i\}$, $(i=1,2,\ldots,a)$, since Lemma 2.5 is available for any finite number n. Indeed, we can apply Lemma 2.5 by putting $K_i=-\log S_i$, $C_i=\pi_i^{1/2}S_i{}^{1/2}\left(\sum_{k=1}^{a}\pi_k S_k\right)^{-1/2}$ for $i=1,2,\ldots,a$ and $f(t)=t^2$.

Question 2.7. From Eq. (5), the matrix inequality holds in the case of $s=0$. However, we do not know whether the following matrix inequalities:

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

or

$$\left\{A(\log A)^2+B(\log B)^2\right\}^{1/2}(A+B)\left\{A(\log A)^2+B(\log B)^2\right\}^{1/2} \geqslant (A\log A+B\log B)^2 \tag{7}$$

corresponding to the case of $s=1$ for any two positive matrices $A\leqslant I$ and $B\leqslant I$ hold or not. We have not yet found any counter-examples, namely the examples that the matrix inequalities both Eqs. (6) and (7) are not satisfied simultaneously, for some positive matrices $A\leqslant I$ and $B\leqslant I$.

Theorem 2.8. *Suppose A and B are 2×2 positive matrices. Then for any $0\leqslant s\leqslant 1$ we have*

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

Proof of Theorem 2.8. We consider the Schatten decomposition of $A+B$ as follows:

$$A+B=\sum_n t_n|\phi_n\rangle\langle\phi_n|, \tag{8}$$

where $\{t_n\}$ are the eigenvalues of $A+B$, $\{|\phi_n\rangle\}$ are the corresponding eigenvectors. Then we have

$$\begin{aligned}\mathrm{Tr}\big[(A+B)^s\big\{A(\log A)^2+B(\log B)^2\big\}\big]\\ &=\sum_n\langle\phi_n|(A+B)^{s/2}\{A(\log A)^2+B(\log B)^2\}(A+B)^{s/2}|\phi_n\rangle\\ &=\sum_n\langle\phi_n(A+B)^{s/2}|\{A(\log A)^2+B(\log B)^2\}|(A+B)^{s/2}\phi_n\rangle\\ &=\sum_n t_n^s\langle\phi_n|\{A(\log A)^2+B(\log B)^2\}|\phi_n\rangle\\ &=\sum_n t_n^s a_n.\end{aligned}$$

As similarly, we have

$$\begin{aligned}\mathrm{Tr}\big[(A+B)^{-1+s}(A\log A+B\log B)^2\big]\\ &=\sum_n t_n^{-1+s}\langle\phi_n|(A\log A+B\log B)^2|\phi_n\rangle\\ &=\sum_n t_n^{-1+s}b_n,\end{aligned}$$

where we put $a_n=\langle\phi_n|\{A(\log A)^2+B(\log B)^2\}|\phi_n\rangle$ and $b_n=\langle\phi_n|(A\log A+B\log B)^2|\phi_n\rangle$. The proof is completed by using the following lemma. □

Lemma 2.9. *Suppose the positive numbers t_1,t_2,a_1,a_2,b_1 and b_2 satisfy the following two conditions*:

(i) $t_1a_1+t_2a_2\geqslant b_1+b_2$
(ii) $a_1+a_2\geqslant t_1^{-1}b_1+t_2^{-1}b_2$

Then for any $0\leqslant s\leqslant 1$ *we have*

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

Proof of Lemma 2.9. It is trivial for $t_1=t_2$ so that we can suppose $t_1>t_2$ without loss of generality. From the condition (i), we then have the following:

$$\begin{aligned}t_1^sa_1+t_2^sa_2-t_1^{-1+s}b_1-t_2^{-1+s}b_2&=t_1^sa_1-t_1^{-1+s}b_1+t_2^sa_2-t_2^{-1+s}b_2\\ &=t_1^{-1+s}(t_1a_1-b_1)+t_2^{-1+s}(t_2a_2-b_2)\\ &\geqslant t_1^{-1+s}(b_2-t_2a_2)+t_2^{-1+s}(t_2a_2-b_2)\\ &=(t_2^{-1+s}-t_1^{-1+s})(t_2a_2-b_2).\end{aligned}$$

Since $t_2^{-1+s} - t_1^{-1+s} \geqslant 0$, if $t_2a_2 - b_2 \geqslant 0$, then the lemma follows. On the other hand, if $t_2a_2 - b_2 < 0$, from the condition (ii) we then have

$$\begin{aligned} t_1^s a_1 + t_2^s a_2 - t_1^{-1+s} b_1 - t_2^{-1+s} b_2 &= t_1^s a_1 - t_1^{-1+s} b_1 + t_2^s a_2 - t_2^{-1+s} b_2 \\ &= t_1^s (a_1 - t_1^{-1} b_1) + t_2^s (a_2 - t_2^{-1} b_2) \\ &\geqslant t_1^s (t_2^{-1} b_2 - a_2) + t_2^s (a_2 - t_2^{-1} b_2) \\ &= (t_1^s - t_2^s)(t_2^{-1} b_2 - a_2) \geqslant 0. \qquad \square \end{aligned}$$

Remark 2.10. After the manner of Theorem 2.8, we can prove Eq. (2) for any 2×2 density matrices S_i and any probability distributions $\pi = \{\pi_i\}$, $(i = 1, 2, \ldots, a)$, by considering the Schatten decomposition of the 2×2 positive matrix $\sum_{k=1}^{a} \pi_k S_k^{\frac{1}{1+s}}$ as follows:

$$\sum_{k=1}^{a} \pi_k S_k^{\frac{1}{1+s}} = \sum_n \lambda_n |\phi_n\rangle\langle\phi_n|,$$

where λ_1 and λ_2 are the eigenvalues of $\sum_{k=1}^{a} \pi_k S_k^{\frac{1}{1+s}}$, $\{|\phi_1\rangle\}$ and $\{|\phi_2\rangle\}$ are corresponding eigenvectors, respectively. Therefore it was shown the concavity of the auxiliary function $E(s)$ of the quantum reliability function for any 2×2 density matrices S_i and $0 \leqslant s \leqslant 1$. Thus we gave a partial solution for the open problem given in [6]. However, we still have the unsolved problems that $E(s)$ is concave for $0 \leqslant s \leqslant 1$ and $n \geqslant 3$, and also that $E(s)$ is concave for $-1 < s < 0$.

Remark 2.11. We expect that our Lemma 2.9 can be extended to the general $n \geqslant 3$, where n represents the number of the eigenvalues given in Eq. (8). However it is impossible to prove it, because we have a counter-example for such a generalization. For example, we take

$$s = \frac{1}{2}, \quad t_1 = 3, \quad t_2 = 2, \quad t_3 = 1, \quad a_1 = \frac{2}{3},$$
$$a_2 = 1, \quad a_3 = \frac{3}{2}, \quad b_1 = \frac{1}{2}, \quad b_2 = 4, \quad b_3 = 1.$$

Although it holds two conditions corresponding to the generalization of two conditions (i) and (ii) in Lemma 2.9:

$$t_1a_1 + t_2a_2 + t_3a_3 = b_1 + b_2 + b_3 = \frac{11}{2}$$

and

$$a_1 + a_2 + a_3 = t_1^{-1}b_1 + t_2^{-1}b_2 + t_3^{-1}b_3 = \frac{19}{6},$$

the following calculations:

$$t_1^s a_1 + t_2^s a_2 + t_3^s a_3 = \frac{2\sqrt{3}}{3} + \sqrt{2} + \frac{3}{2} \simeq 4.068914$$

and

$$t_1^{-1+s} b_1 + t_2^{-1+s} b_2 + t_3^{-1+s} b_3 = \frac{\sqrt{3}}{6} + 2\sqrt{2} + 1 \simeq 4.1171021,$$

show that

$$t_1^s a_1 + t_2^s a_2 + t_3^s a_3 \geqslant t_1^{-1+s} b_1 + t_2^{-1+s} b_2 + t_3^{-1+s} b_3$$

does not hold. This means that our Lemma 2.9 cannot be extended to the general case of $n \geqslant 3$. Therefore we must produce an another method to prove Theorem 2.8 for any $n \times n$ positive matrices A and B. Our Theorem 2.8 is constructed by a kind of the interpolation between two conditions generated by Theorems 2.2 and 2.4. If we extend this method to the case of $n \geqslant 3$, we may require the further conditions.

3. The related inequalities

We introduce the following symbol in the relation to quantum relative entropy. For the positive matrices A and B, we define

$$D(A\|B) = A(\log A - \log B).$$

Then we have the next theorem.

Theorem 3.1.
(1) $\mathrm{Tr}[D(A\|B)D(B\|A)] \leqslant 0$.
(2) $\mathrm{Tr}[(A+B)^{-1}D(A\|B)D(B\|A)^*] \leqslant 0$.

Remark 3.2. The quantum relative entropy is defined by $H(A\|B) = \mathrm{Tr}[D(A\|B)]$ for any density matrices A and B. The relative matrix entropy [3] is defined by

$$S(A\|B) = A^{1/2}(\log A^{-1/2}BA^{-1/2})A^{1/2}$$

for any invertible positive matrices A and B. Moreover, if A and B are commutative, then we have $D(A\|B) = -S(A\|B)$.

Acknowledgments

The authors thank referees for valuable comments on this paper.

References

[1] T. Ando, Topics on operator inequalities, Lecture Notes (mimeographed), Hokkaido University, Sapporo, 1978.
[2] M.V. Burnashev, A.S. Holevo, On reliability function of quantum communication channel, Probl. Inf. Transm. 34 (2) (1998) 97–107.
[3] J.I. Fujii, E. Kamei, Relative operator entropy in noncommutative information theory, Math. Japon. 34 (1989) 341–348.
[4] S. Furuichi, K. Yanagi, K. Kuriyama, A sufficient condition on concavity of the auxiliary function appearing in quantum reliability function, Information 6 (1) (2003) 71–76.
[5] F. Hansen, G.K. Pedersen, Jensen's inequality for operators and Lowner's theorem, Math. Ann. 258 (1982) 229–241.
[6] A.S. Holevo, Reliability function of general classical-quantum channel, IEEE Trans. Inform. Theory 46 (2000) 2256–2261.
[7] M. Nakamura, H. Umegaki, A note on the entropy for operator algebras, Proc. Japan Acad. 37 (1961) 149–154.
[8] T. Ogawa, H. Nagaoka, Strong converse to the quantum channel coding theorem, IEEE Trans. Inform. Theory 45 (1999) 149–154.