Error exponent in asymmetric quantum hypothesis testing and its application to classical-quantum channel coding

Masahito Hayashi
ERATO-SORST Quantum Computation and Information Project, Japan Science and Technology Agency, 201 Daini Hongo White Bldg. 5-28-3, Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
and Graduate School of Information Sciences, Tohoku University, Aoba-ku, Sendai, 980-8579, Japan
(Received 12 November 2006; published 5 December 2007)

An upper bound on simple quantum hypothesis testing in the asymmetric setting is shown using a useful inequality by Audenaert *et al.* [Phys. Rev. Lett. **98**, 160501 (2007)] which was originally invented for symmetric setting. Using this upper bound, we obtain the Hoeffding bound, which is identical with the classical counterpart if the hypotheses, composed of two density operators, are mutually commutative. Its attainability has been a long-standing open problem. Further, using this bound, we obtain a better exponential upper bound of the average error probability of classical-quantum channel coding.

DOI: 10.1103/PhysRevA.76.062301 PACS number(s): 03.67.Hk, 03.65.Ta, 03.65.Wj

I. INTRODUCTION

One of the main theoretical difficulties appearing in quantum information lies in the noncommutativity. Hence, for further development of quantum information, it is needed to accumulate the methods to resolve such difficulties. Simple quantum hypothesis testing is the simplest problem describing this kind of difficulty because this problem is discriminating two quantum states (the null hypothesis and the alternative hypothesis) as the candidates of the true state. (In statistics, hypothesis testing is called simple when both hypotheses consist of one element.) This problem is also the fundamental tool for other problems in quantum information. For example, sending classical information via quantum channel (classical-quantum channel coding) is known as an important topic in quantum information, and is closely related to asymmetric simple quantum hypothesis testing, which was shown in Hayashi-Nagaoka [1]. Recently, Nussbaum-Szkoła [2] and Audenaert *et al.* [3] have derived the optimal exponential rate of error probability in symmetric quantum hypothesis testing, which is called the quantum Chernoff bound [4]. However, its symmetric framework is not directly linked to classical-quantum channel coding, but only the asymmetric framework is applicable to classical-quantum channel coding. Further, it was clarified through information spectrum method [5] that other topics in classical information theory, wire-tap channel, source coding, and identification code, are related to the asymmetric framework. Hence, asymmetric simple quantum hypothesis testing can be expected to be a useful tool for quantum information.

In simple quantum hypothesis testing, we usually focus on the two types of error probabilities, i.e., the error probability of the first kind (the null hypothesis ρ is rejected despite of being correct) and the error probability of the second kind (the alternative hypothesis σ is rejected despite of being correct). While we minimize the sum of two kinds of error probabilities in the symmetric hypothesis testing, we do not treat both error probabilities equally in the asymmetric hypothesis testing. For example, we minimize the error probability of the second kind under the constraint of the error probability of the first kind. In this paper, we will treat the asymmetric hypothesis testing in the asymptotic framework, in which, we usually adopt independent and identical condition, i.e., the true state is assumed to be the tensor product state $\rho^{\otimes n}$ or $\sigma^{\otimes n}$. In the symmetric case with the asymptotic framework, the exponential rate of the sum of two error probabilities is equal to $\max_{0\leqslant s\leqslant 1}-\phi(s|\rho\|\sigma)$, which is called quantum Chernoff bound, where $\phi(s|\rho\|\sigma) := \log \mathrm{Tr}\rho^s\sigma^{1-s}$. In this paper, we choose the base of the logarithm to be e. The asymmetric case is more complicated than the symmetric case even with the asymptotic framework. In the asymmetric case, we often assume the constant constraint for the error probability of the first kind. Then, the optimal decreasing rate of the error probability of the second kind is equal to the quantum relative entropy $D(\rho\|\sigma) := \mathrm{Tr}\rho(\log\rho - \log\sigma)$, which is known as quantum Stein's lemma [6,7]. As is described in Hayashi-Nagaoka [1] and Ogawa-Nagaoka [7], the capacity theorem of classical-quantum channel coding can be shown via quantum Stein's lemma. For a more precise analysis, we often treat the same optimal decreasing rate under the exponential constraint for the error probability of the first kind. In the classical case (i.e., the commutative case), this optimal rate is equal to $\max_{0<s\leqslant 1}\frac{-sr-\phi(s|\rho\|\sigma)}{1-s}$, and is called the Hoeffding bound [8], where r is the exponent of the error probability of the first kind. As was pointed by Han, the exponential decreasing rate of average error probability of channel coding can be characterized by the optimal exponential decreasing rate of error probability in asymmetric simple hypothesis testing in the classical case. Also, Han [5] implicitly indicated that asymmetric simple classical hypothesis testing is closely related to other topics in information theory via information spectrum method.

In this paper, we focus on Hoeffding bound in the quantum case, and derive an upper bound of the optimal decreasing rate of the error probability of the second kind under the exponential constraint for the error probability of the first kind. Applying this bound, we derive an exponential upper bound of the average error probability of the optimal code in classical-quantum channel coding. While the capacity theorem of classical-quantum channel coding is quite familiar, but exponential evaluation is more important from applied viewpoint. This is because in order to evaluate the average

error probability with a finite length code, we need not the capacity but an exponential upper bound. Further, as was pointed out by Nagaoka [9], since asymmetric hypothesis testing can be regarded as a meeting point among statistics, information theory, and large deviation theory, this research area is quite important for interdisciplinary research. Therefore, it can be expected that the study of its quantum case yields many fruitful products. Further, Hoeffding bound itself is more meaningful from statistical viewpoint. Since the meanings of two kinds of errors are different in the classical statistical hypothesis testing, it is usual to treat two kinds of error probabilities asymmetrically. Thus, it is suitable to apply not Stein's lemma but Hoeffding bound to an exponentially small error probability of the first kind because Stein's lemma treats the constant constraint for the error probability of the first kind.

Now, we trace the history of the research of quantum simple hypothesis testing. First, the quantum extension of Stein's lemma has been solved by Hiai and Petz [6] and Ogawa and Nagaoka [7]. That is, Hiai and Petz [6] proved the existence of a sequence of test attaining the required property by combining the classical case of Stein's lemma and the following fact: When a suitable measurement is chosen, the classical relative entropy concerning the measurement outcome approaches quantum relative entropy. Ogawa and Nagaoka proved the negative part: if the exponential rate of the error probability of the second kind is greater than the quantum relative entropy, the probability correctly accepting the null hypothesis goes to 0. Concerning the symmetric setting, Hayashi obtained quantum Chernoff bound in Chap. 3 of Ref. [10] when $\phi(s|\rho\|\sigma)$ is symmetric. Nussbaum and Szkoła [2] obtained its lower bound. Audenaert *et al.* [3] showed that the bound by Nussbaum and Szkoła [2] can be attained. In their proof, they derived a quite useful inequality (Lemma 2 in this paper).

However, concerning the quantum extension of Hoeffding bound, only a lower bound has been obtained by Ogawa and Hayashi [11]. Their approach is valid only in the finite dimensional case, and their bound does not work effectively in the pure states case. They also suggested the existence of a tighter lower bound. Hence, a tighter lower bound of these problems has been desired. In this paper, we obtain a tighter lower bound of Hoeffding bound by using a powerful inequality by Audenaert *et al.* [3]. This method is valid even in the infinite-dimensional case. After the first version of this paper, Nagaoka [12] proved that our exponential upper bound is tight.

Next, we trace the history of the research of classical-quantum channel coding. The capacity theorem has been established by combining the achievable part shown by Holevo [13] and Schumacher-Westmoreland [14] with the impossibility part that goes back to the 1970's works of Holevo [15,16]. Brunashev and Holevo [17] derived an exponential upper bound of average error probability when all of the sent states are pure. Ogawa-Nagaoka [18] pointed out that quantum channel coding can be treated by using quantum hypothesis testing. Hayashi and Nagaoka [4] derived a good relation between this problem and the asymmetric treatment of hypothesis testing. Applying this relation to Ogawa-Hayashi's result, they derived an exponential upper bound of average error probability for the general case. However, since the Ogawa-Hayashi bound in quantum hypothesis testing is not optimal, there is a possibility to improve the Hayashi-Nagaoka [1] exponential upper bound in classical-quantum channel coding.

II. FORMULATION AND MAIN RESULTS

We study the simple hypothesis testing problem for the null hypothesis $H_0:\rho^{\otimes n}$ versus the alternative hypothesis $H_1:\sigma^{\otimes n}$, where $\rho^{\otimes n}$ and $\sigma^{\otimes n}$ are the nth tensor powers of arbitrarily given density operators ρ and σ on a Hilbert space $\mathcal{H}$ which represents a physical system in interest. The problem is to decide which hypothesis is true based on the data drawn from a quantum measurement, which is described by a positive operator valued measure (POVM) on $\mathcal{H}^{\otimes n}$, i.e., a resolution of identity $\Sigma_i M_{n,i}=I_n$ by nonnegative operators $M_n=\{M_{n,i}\}$ on $\mathcal{H}^{\otimes n}$. If a POVM consists of projections on $\mathcal{H}^{\otimes n}$, it is called a projection valued measure (PVM). In the hypothesis testing problem, however, it is sufficient to treat a two-valued POVM $\{M_0,M_1\}$, where the subscripts 0 and 1 indicate the acceptance of H_0 and H_1, respectively. Thus, a hermitian matrix T_n satisfying inequalities $0\leqq T_n\leqq I$ is called a test in the sequel, since T_n is identified with the POVM $\{T_n^c,T_n\}$. For a test T_n, the error probabilities of the first kind and the second kind are, respectively, given by $\mathrm{Tr}[\rho^{\otimes n}T_n]$ and $\mathrm{Tr}[\sigma^{\otimes n}T_n^c]$, where $T^c:=I-T$.

This problem is considered in an asymmetric framework. Let us define the optimal value for the error probability of the second kind $\mathrm{Tr}[\sigma^{\otimes n}T_n^c]$ under the constant constraint on the error probability of the first kind $\mathrm{Tr}[\rho^{\otimes n}T_n]$:

$$\beta_n^*(\epsilon)\stackrel{\mathrm{def}}{=}\min\{\mathrm{Tr}[\sigma^{\otimes n}T_n^c]|T_n:\mathrm{test},\mathrm{Tr}[\rho^{\otimes n}T_n]\leqq\epsilon\}.$$

Then, we have the quantum Stein's lemma, which was obtained by Hiai-Petz [6] and Ogawa-Nagaoka [7]: For $0<\forall\epsilon<1$, it holds that

$$\lim_{n\to\infty}\frac{1}{n}\log\beta_n^*(\epsilon)=-D(\rho\|\sigma). \tag{1}$$

For a further analysis, we focus on the decreasing exponent of the error probability of the second kind under an exponential constraint for the error probability of the first kind. For this purpose, we define

$$B(r|\rho\|\sigma)\stackrel{\mathrm{def}}{=}\sup_{\{T_n\}}\left\{\lim_{n\to\infty}\frac{-\log\mathrm{Tr}\sigma^{\otimes n}T_n^c}{n}\middle|\varliminf_{n\to\infty}\frac{-\log\mathrm{Tr}\rho^{\otimes n}T_n}{n}\geqq r\right\}.$$

Then, we obtain the following theorem:

Theorem 1. The inequality

$$B(r|\rho\|\sigma)\geqq\sup_{0\leqq s\leqq 1}\frac{-sr-\phi(s|\rho\|\sigma)}{1-s} \tag{2}$$

holds.

After the first version of this paper, the opposite inequality was proved by Nagaoka [12]. That is, Nagaoka's result implies that the inequality (2) is the final, tight quantum Hoef-

fding bound. In fact, Ogawa and Hayashi [11] obtained a weaker inequality

$$B(r|\rho\|\sigma) \geqslant \max_{0\leqslant s\leqslant 1} \frac{-sr-\tilde{\phi}(s|\rho\|\sigma)}{1-s} \tag{3}$$

and treated the inequality (2) as an open problem, where $\tilde{\phi}(s|\rho\|\sigma) := \mathrm{Tr}\sigma\rho^{s/2}\sigma^{-s}\rho^{s/2}$.

III. PROOF OF MAIN THEOREMS

In the following we abbreviate $\phi(s|\rho\|\sigma)$ to $\phi(s)$. In order to prove theorem 1, we use the following lemma.

Lemma 1. For any two positive-semidefinite operators X, Y and a real number $0\leqslant s\leqslant 1/2$, we obtain

$$\mathrm{Tr}X^sY^{1-s} \geqslant \mathrm{Tr}\{X^{1-s} \geqslant Y^{1-s}\}Y + \mathrm{Tr}\{X^{1-s} < Y^{1-s}\}X,$$

where for any Hermitian matrixes C and D we denote the projection $\Sigma_{c_i\geqslant 0}E_i$ ($\Sigma_{c_i<0}E_i$) by $\{C\geqslant D\}$ ($\{C<D\}$) with the spectral decomposition $C-D=\Sigma_i c_i E_i$. Only the case of $s=1/2$ has been proved in Chap. 3 of Hayashi [10].

Substituting $\sigma^{\otimes n}$ and $\rho^{\otimes n}e^{-na}$ to Y and X in this lemma, the projection $T_{n,s} := \{(\rho^{\otimes n}e^{-na})^{1-s} < (\sigma^{\otimes n})^{1-s}\}$ satisfies

$$\mathrm{Tr}\rho^{\otimes n}T_{n,s} = \mathrm{Tr}XT_{n,s}e^{na} \leqslant \mathrm{Tr}X^sY^{1-s}e^{na} = e^{n(1-s)a}e^{n\phi(s)} \tag{4}$$

$$\mathrm{Tr}\sigma^{\otimes n}(I-T_{n,s}) = \mathrm{Tr}Y(I-T_{n,s})$$

$$\leqslant \mathrm{Tr}X^sY^{1-s} = e^{-nsa}e^{n\phi(s)} \tag{5}$$

for $0\leqslant s\leqslant 1/2$. For $1/2\leqslant t\leqslant 1$, the projection $T_{n,t} := \{(\rho^{\otimes n}e^{-na})^t < (\sigma^{\otimes n})^t\}$ satisfies

$$\mathrm{Tr}\rho^{\otimes n}T_{n,t} \leqslant e^{n(1-t)a}e^{n\phi(t)}, \tag{6}$$

$$\mathrm{Tr}\sigma^{\otimes n}(I-T_{n,t}) \leqslant e^{-nta}e^{n\phi(t)}, \tag{7}$$

where we substitute $1-t$, $\sigma^{\otimes n}$, and $\rho^{\otimes n}e^{-na}$ into s, X, and Y.

We choose $s_r \overset{\mathrm{def}}{=} \arg\max_{0\leqslant s\leqslant 1}\frac{-sr-\phi(s|\rho\|\sigma)}{1-s}$. Then, we have

$$r = (s_r-1)\phi'(s_r) - \phi(s_r),$$

$$\max_{1\geqslant s'\geqslant 0}\frac{-s'r-\phi(s')}{1-s'} = s_r\phi'(s_r) - \phi(s_r).$$

Thus, choosing a to be $\phi'(s_r)$, Eqs. (4)–(7) imply

$$\mathrm{Tr}\rho^{\otimes n}T_{n,s} \leqslant e^{-nr},$$

$$\mathrm{Tr}\sigma^{\otimes n}(I-T_{n,s}) \leqslant e^{-n\max_{0\leqslant s\leqslant 1}[-sr-\phi(s|\rho\|\sigma)/1-s]}.$$

Therefore, we obtain Eq. (2).

Let us now move to prove lemma 1. Note that the proof that we present here goes through in infinite dimensions. The proof relies on the following quite powerful lemma.

Lemma 2 (Audenaert et al. [3]). For any two positive-semidefinite operators A, B and a real number $0\leqslant t\leqslant 1$, we obtain

$$\mathrm{Tr}\{A\geqslant B\}B(A^t-B^t) \geqslant 0.$$

Proof of Lemma 1. We apply lemma 2 to the case $t=s/(1-s)$, $A=X^{1-s}$ and $B=Y^{1-s}$, where $0\leqslant s\leqslant 1/2$ is the inequality. Then

$$\mathrm{Tr}\{X^{1-s}\geqslant Y^{1-s}\}Y^{1-s}(X^s-Y^s) \geqslant 0$$

holds. Subtracting both sides from $\mathrm{Tr}\{X^{1-s}\geqslant Y^{1-s}\}(X-Y)$ then yields

$$\mathrm{Tr}X^s\{X^{1-s}\geqslant Y^{1-s}\}(X^{1-s}-Y^{1-s}) \leqslant \mathrm{Tr}\{X^{1-s}\geqslant Y^{1-s}\}(X-Y).$$

Since $\{X^{1-s}\geqslant Y^{1-s}\}(X^{1-s}-Y^{1-s}) \geqslant (X^{1-s}-Y^{1-s})$, we have

$$\mathrm{Tr}X - \mathrm{Tr}X^sY^{1-s} = \mathrm{Tr}X^s(X^{1-s}-Y^{1-s}) \leqslant \mathrm{Tr}X^s\{X^{1-s}\geqslant Y^{1-s}\} \times (X^{1-s}-Y^{1-s}) \leqslant \mathrm{Tr}\{X^{1-s}\geqslant Y^{1-s}\}(X-Y).$$

Thus, the relation $I-\{X^{1-s}\geqslant Y^{1-s}\}=\{X^{1-s}<Y^{1-s}\}$ yields

$$\mathrm{Tr}(I-\{X^{1-s}\geqslant Y^{1-s}\})X + \mathrm{Tr}\{X^{1-s}\geqslant Y^{1-s}\}Y \leqslant \mathrm{Tr}X^sY^{1-s}.$$

IV. APPLICATION TO CLASSICAL-QUANTUM CHANNEL CODING

As is mentioned in Ref. [1], the error exponent in classical-quantum channel coding are derived from the error exponent in simple quantum hypothesis testing. Now, we consider the nth stationary memoryless channel of the classical-quantum channel $x\mapsto\rho_x$. Define the densities R, S_p and σ_p for a distribution p,

$$R \overset{\mathrm{def}}{=} \begin{pmatrix} p(x_1)\rho_{x_1} & & 0 \\ & \ddots & \\ 0 & & p(x_k)\rho_{x_k} \end{pmatrix},$$

$$S_p \overset{\mathrm{def}}{=} \begin{pmatrix} p(x_1)\sigma_p & & 0 \\ & \ddots & \\ 0 & & p(x_k)\sigma_p \end{pmatrix}, \quad \sigma_p \overset{\mathrm{def}}{=} \sum_x p(x)\rho_x.$$

In the channel coding, we usually treat the trade-off between the average error probability $P_e(\Phi^{(n)})$ and the number N of transmitted massages. That is, the receiver should choose the recovered message among N elements via the received quantum state. This number is called the size.

Then, the inequality (44) in Ref. [1] implies that for any distribution p and any test $T^{(n)}$, there exists a code $\Phi^{(n)}$ with the size N whose average error probability $P_e(\Phi^{(n)})$ satisfies

$$P_e(\Phi^{(n)}) \leqslant 2(1-\mathrm{Tr}R^{\otimes n}T^{(n)}) + 4N\mathrm{Tr}S^{\otimes n}T^{(n)}. \tag{8}$$

This kind of relation between hypothesis testing and channel coding was obtained by Verdú and Han [19], and it was researched by Han further in Ref. [20].

When $N=e^{na}$, applying Lemma 1 to the two cases: $X=S_pN, Y=R$ and $Y=S_pN, X=R$, we obtain

$$P_e(\Phi^{(n)}) \leqslant 4e^{-n[sa-\varphi_p(s)]} \tag{9}$$

for $0\leqslant s\leqslant 1$, where

$$\varphi_p(s) \overset{\mathrm{def}}{=} \log \mathrm{Tr}R^{1-s}S_p^s = \log\sum_x p_x\mathrm{Tr}\rho_x^{1-s}\sigma_p^s. \tag{10}$$

This gives the exponential decreasing rate of error probability. This upper bound improves the bound given in Hayashi-

Nagaoka [1], which was obtained by using the Ogawa-Hayashi [11] Hoeffding bound. Also, it can be regarded as a generalization of the Brunashev-Holevo [17] result, which gives the exponential decreasing rate of error probability in the pure states case.

V. DISCUSSIONS

In this paper, we treated the asymmetric setting of quantum hypothesis testing, and obtained a quantum extension of Hoeffding bound $\max_{0\leqslant s\leqslant 1}\frac{-sr-\phi(s|\rho\|\sigma)}{1-s}$, which improves the Ogawa-Hayashi [11] bound. Since Nagaoka [12] proved the opposite inequality, our exponential rate is tight. Further, we applied this result to classical-quantum channel coding and obtained a better error exponent. As is discussed in the beginning part of this paper, the asymmetric setting of simple hypothesis testing has wider connection with information theory. So, we can expect that our results will be applied to many other topics in quantum information.

ACKNOWLEDGMENTS

This work was partially supported by ERATO-SORST Quantum Computation, Information Project and Special Coordination Funds for Promoting Science and Technology, and a MEXT Grant-in-Aid for Scientific Research on Priority, Deepening and Expansion of Statistical Mechanical Informatics (DEX-SMI), Grant No. 18079014.

[1] M. Hayashi and H. Nagaoka, IEEE Trans. Inf. Theory **49**, 1753 (2003).
[2] M. Nussbaum and A. Szkoła, e-print arXiv:quant-ph/0607216.
[3] K. M. R. Audenaert, J. Calsamiglia, R. Munoz-Tapia, E. Bagan, L. Masanes, A. Acin, and F. Verstraete, Phys. Rev. Lett. **98**, 160501 (2007).
[4] H. Chernoff, Ann. Math. Stat. **23**, 493 (1952).
[5] T. S. Han, *Information-Spectrum Methods in Information Theory* (Springer, Berlin, 2002).
[6] F. Hiai and D. Petz, Commun. Math. Phys. **143**, 99 (1991).
[7] T. Ogawa and H. Nagaoka, IEEE Trans. Inf. Theory **46**, 2428 (2000).
[8] W. Hoeffding, in *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability* (University of California Press, Berkeley, 1965), pp. 203–219.
[9] H. Nagaoka (private communication).
[10] M. Hayashi, *Quantum Information: An Introduction* (Springer, Berlin, 2006).
[11] T. Ogawa and M. Hayashi, IEEE Trans. Inf. Theory **50**, 1368 (2004).
[12] H. Nagaoka, e-print arXiv:quant-ph/0611289.
[13] A. S. Holevo, IEEE Trans. Inf. Theory **44**, 269 (1998).
[14] B. Schumacher and M. D. Westmoreland, Phys. Rev. A **56**, 131 (1997).
[15] A. S. Holevo, Probl. Peredachi Inf. **9**, 3 (1973). [Probl. Inf. Transm. **9**, 177 (1975)].
[16] A. S. Holevo, Probl. Peredachi Inf. **15**, 3 (1979). [Probl. Inf. Transm. **15**, 247 (1979)].
[17] M. V. Burnashev and A. S. Holevo, Probl. Inf. Transm. **34**, 97 (1998).
[18] T. Ogawa and H. Nagaoka, IEEE Trans. Inf. Theory **53**, 2261 (2007).
[19] S. Verdú and T. S. Han, IEEE Trans. Inf. Theory **40**, 1147 (1994).
[20] Han [5] treated the exponential error rate of the channel coding in the original Japanese version. However, he did not treat this topic in the English translation.