On Error Exponents in Quantum Hypothesis Testing

In the simple quantum hypothesis testing problem, upper bounds on the error probabilities are shown based on a key operator inequality between a density operator and its pinching. Concerning the error exponents, the upper bounds lead to a non-commutative analogue of the Hoeffding bound, which is identical with the classical counterpart if the hypotheses, composed of two density operators, are mutually commutative. The upper bounds also provide a simple proof of the direct part of the quantum Stein's lemma.

1 Introduction

Quantum hypothesis testing is a fundamental problem in quantum information theory, because it is one of the most simple problems where the difficulty derived from non-commutativity of operators appears. It is also closely related to other topics in quantum information theory, as in classical information theory. Actually, its relation with quantum channel coding is discussed in [7, 15].

Let us outline briefly significant results in classical hypothesis testing for probability distributions $p^n(\cdot)$ versus $q^n(\cdot)$, where $p^n(\cdot)$ and $q^n(\cdot)$ are i.i.d. extensions of some probability distributions $p(\cdot)$ and $q(\cdot)$ on a finite set $\mathcal{X}$. In the classical case, the asymptotic behaviors of the first kind error probability α_n and the second kind error probability β_n for the optimal test were studied thoroughly as follows.

First, when α_n satisfies the constant constraint $\alpha_n \leq \varepsilon$ ($\varepsilon > 0$), the error exponent of βn for the optimal test, say $\beta_n^*(\varepsilon)$, is written asymptotically as

$$\limsup_{n\to\infty} \frac{1}{n} \log \beta_n^* = -D(p||q) \tag{1}$$

R. Ahlswede, *Identification and Other Probabilistic Models*,
Foundations in Signal Processing, Communications and Networking 16,
https://doi.org/10.1007/978-3-030-65072-8_25

for any ε, where $D(p||q)$ is the relative entropy. The equality (1) is called Stein's lemma (see e.g. [4, p.115]), and the quantum analogue of (1) was established recently [8, 14].

Next, when α_n satisfies the exponential constraint $\alpha_n \leq e^{-nr}$ $(r > 0)$, the error exponent of β_n for the optimal test is asymptotically determined by

$$\limsup_{n\to\infty} \frac{1}{n} \log \beta_n^{\dagger}(r) = - \min_{p':D(p'||p)\leq r} D(p'||q) \tag{2}$$

$$= - \max_{0<s\leq 1} \frac{\Psi(s) - (1-s)r}{s} \tag{3}$$

where the function $\Psi(s)$ is defined as

$$\Psi(s) \triangleq - \log \sum_{x\in\mathcal{X}} p(x)^{1-s} q(x)^s. \tag{4}$$

Historically speaking, (2) and the test achieving it were shown in [9], followed by another expression (3) (see [3]), which we call the Hoeffding bound here. In quantum hypothesis testing, the error exponent of $1 - \beta_n$ was studied in [14] to obtain a similar result to (3), which led to the strong converse property in quantum hypothesis testing. Concerning quantum fixed-length pure state source coding, the error exponent of erroneously decoded probability was determined in [5], where the optimality of the error exponent similar to (3) was discussed.

In this lecture (see [13]), a quantum analogue of the Hoeffding bound (3), (4) is introduced to derive a bound on the error exponent in quantum hypothesis testing. As a by-product of the process to derive the exponent, a simple proof of the quantum Stein's lemma is also given.

2 Definition and Main Results

Let $\mathcal{H}$ be a Hilbert space which represents a physical system in interest. We assume $\dim\mathcal{H} < \infty$ for mathematical simplicity. Let us denote the set of linear operators on $\mathcal{H}$ as $\mathcal{L}(\mathcal{H})$ and define the set of density operators on $\mathcal{H}$ by

$$\mathcal{S}(\mathcal{H}) \triangleq \{\rho \in \mathcal{L}(\mathcal{H}) : \rho = \rho^* \geq 0, \mathrm{Tr}[\rho] = 1\}. \tag{5}$$

We study the hypothesis testing problem for the null hypothesis

$$H_0 : \rho_n \triangleq \rho^{\otimes n} \in \mathcal{S}(\mathcal{H}^{\otimes n})$$

versus the alternative hypothesis

$$H_1 : \sigma_n \triangleq \sigma^{\otimes n} \in \mathcal{S}(\mathcal{H}^{\otimes n})$$

where $\rho^{\otimes n}$ and $\sigma^{\otimes n}$ are the nth tensor powers of arbitrarily given density operators ρ and σ in $\mathcal{S}(\mathcal{H})$.

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 $\sum_i M_{n,i} = I_n$ by non-negative 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, an operator $A_n \in \mathcal{L}(\mathcal{H}^{\otimes n})$ satisfying inequalities $0 \leq A_n \leq I_n$ is called a test in the sequel, since A_n is identified with the POVM $\{A_n, I_n - A_n\}$. For a test A_n, the error probabilities of the first kind and the second kind are, respectively, defined by

$$\alpha_n(A_n) \triangleq \mathrm{Tr}[\rho_n(I_n - A_n)]$$

$$\beta_n(A_n) \triangleq \mathrm{Tr}[\sigma_n A_n].$$

Let us define the optimal value for $\beta_n(A_n)$ under the constant constraint on $\alpha_n(A_n)$

$$\beta_n^*(\varepsilon) \triangleq \min\{\beta_n(A_n) : \ A_n : \ \text{test}\ , \alpha_n(A_n) \leq \varepsilon\} \tag{6}$$

and let

$$D(\rho||\sigma) \triangleq \mathrm{Tr}[\rho(\log\rho - \log\sigma)] \tag{7}$$

which is called the quantum relative entropy. Then we have the following theorem, which is one of the most essential theorems in quantum information theory.

Proposition 277 (The Quantum Stein's Lemma) *For all* $0 < \varepsilon < 1$*, it holds that*

$$\lim_{n\to\infty} \frac{1}{n} \log \beta_n^*(\varepsilon) = -D(\rho||\sigma). \tag{8}$$

The first proof of (8) was composed of two inequalities, the direct part and the converse part. The direct part, concerned with existence of good tests, claims that

$$\forall\, 0 < \varepsilon \leq 1, \qquad \limsup_{n\to\infty} \frac{1}{n} \log \beta_n^*(\varepsilon) \leq -D(\rho||\sigma) \tag{9}$$

and it was given by Hiai and Petz [8]. In this lecture, the main focus is on the direct part. Note that the direct part (9) is equivalent to the existence of a sequence of tests

$\{A_n\}$ such that

$$\lim_{n\to\infty} \alpha_n(A_n) = 0 \quad \text{and} \quad \limsup_{n\to\infty} \frac{1}{n} \log \beta_n(A_n) \le -D(\rho||\sigma) \tag{10}$$

(see [14]). On the other hand, the converse part, concerned with nonexistence of too good tests, asserts that

$$\forall\, 0 < \varepsilon < 1, \qquad \liminf_{n\to\infty} \frac{1}{n} \log \beta_n^*(\varepsilon) \ge -D(\rho||\sigma) \tag{11}$$

which was given by Ogawa and Nagaoka [14]. A direct proof of the equality (8) was also given by Hayashi [6] using the information spectrum approach in quantum setting [10, 12], and a considerably simple proof of the converse part (11) was given in [11].

In this lecture, the asymptotic behavior of the error exponent $\frac{1}{n} \log \beta_n(A_n)$ under the exponential constraint

$$\alpha_n(A_n) \le e^{-nr}, \qquad r > 0$$

is studied, and a non-commutative analogue of the Hoeffding bound [9] similar to (3) is given as follows.

Theorem 278 (Ogawa and Hayashi 2004, [13]) *For all $r > 0$, there exists a sequence of tests $\{A_n\}$ which satisfies*

$$\limsup_{n\to\infty} \frac{1}{n} \log \alpha_n(A_n) \le -r, \tag{12}$$

$$\limsup_{n\to\infty} \frac{1}{n} \log \beta_n(A_n) \le -\max_{0<s\le 1} \frac{\overline{\psi}(s) - (1-s)r}{s} \tag{13}$$

where

$$\overline{\psi}(s) \triangleq -\log \mathrm{Tr}\left[\rho\sigma^{\frac{s}{2}}\rho^{-s}\sigma^{\frac{s}{2}}\right]. \tag{14}$$

We will prove the theorem in 4. If ρ and σ commutate, $\overline{\psi}(s)$ is identical with the classical counterpart $\Psi(s)$ defined in (4), and (13) coincides with the Hoeffding bound (3), which is optimal in classical hypothesis testing.

This lecture is organized as follows. In 3, upper bounds on the error probabilities are shown based on a key operator inequality [6]. Using the upper bounds, we will prove Theorem 278 in 4. In 5, we will make some remarks toward further investigations.

Section 7 is devoted to the definition of pinching (see, e.g., [2], p. 50), which is known as a special notion of the conditional expectation in literature on the operator

algebra and is used effectively in 3. In 8, the key operator inequality used in 3 is summarized along with another proof of it for readers' convenience.

3 Bounds on Error Probabilities

In the sequel, let $\mathcal{E}_{\sigma_n}(\rho_n)$ be the conditional expectation of ρ_n to the commutant of the $*$-subalgebra generated by σ_n, which we call pinching (see 7) and denote it as $\overline{\rho_n}$ for simplicity. Let $v(\sigma_n)$ be the number of eigenvalues of σ_n mutually different from others as defined in 7. Then a key operator inequality[1] follows from Lemma 285 in 8, which originally appeared in [6]

$$\rho_n \leq v(\sigma_n)\overline{\rho}n. \tag{15}$$

Note that the type counting argument provides

$$v(\sigma_n) \leq (n+1)^d \tag{16}$$

where $d \triangleq \dim \mathcal{H}$. Following [6], let us apply the operator monotonicity of the function $x \mapsto -x^{-s}$, $0 \leq s \leq 1$ (see, e.g, [2, Sec. V.1]) to (15) so that we have

$$\overline{\rho_n}^{-s} \leq v(\sigma_n)^s \rho_n^{-s} \leq (n+1)^{sd} \rho_n^{-s}. \tag{17}$$

Following the notation used in [10, 12], let us define the projection $\{X > 0\}$ for a Hermitian operator $X = \sum_i x_i E_i$ as

$$\{X > 0\} \triangleq \sum_{i:x_i>0} E_i \tag{18}$$

where E_i is the projection onto the eigenspace corresponding to an eigenvalue x_i. In the sequel, we will focus on a test defined by

$$\overline{S}_n(a) \triangleq \{\overline{\rho_n} - e^{na}\sigma_n > 0\} \tag{19}$$

where a is a real parameter, and derive the upper bounds on the error probabilities for the test $\overline{S}_n(a)$ as follows.

Theorem 279 (Ogawa and Hayashi 2004, [13])

$$\alpha_n\left(\overline{S}_n(a)\right) \leq (n+1)^d e^{-n\overline{\varphi}(a)}, \tag{20}$$

$$\beta_n\left(\overline{S}_n(a)\right) \leq (n+1)^d e^{-n[\overline{\varphi}(a)+a]} \tag{21}$$

[1] Although the way to derive the operator inequality and the definition of $v(\sigma_n)$ are different from those of [6], it results in the same one as [6] in the case that both of ρ_n and σ_n are tensored states.

where $\overline{\varphi}(a)$ is defined by $\overline{\psi}(s)$ given in (14) as

$$\overline{\varphi}(a) \stackrel{def}{=} \max_{0\le s\le 1} \left\{\overline{\psi}(s) - as\right\}. \tag{22}$$

Proof The definition of $\overline{S}_n(a)$ and commutativity of operators $\overline{\rho_n}$ and σ_n lead to

$$\left(\overline{\rho_n}^{1-s} - e^{na(1-s)}\sigma_n^{1-s}\right)\overline{S}_n(a) \ge 0 \tag{23}$$

$$\left(\overline{\rho_n} - e^{nas}\sigma_n^s\right)\left(I_n - \overline{S}_n(a)\right) \le 0 \tag{24}$$

for all $0 \le s \le 1$. Note that $\overline{S}_n(a)$ also commutes with σ_n. Therefore, the inequality (24), with the property of pinching (63) in 7, provides

$$\begin{aligned}
\alpha_n(\overline{S}_n(a)) &= \mathrm{Tr}[\rho_n(I_n - \overline{S}_n(a))] \\
&= \mathrm{Tr}[\overline{\rho_n}(I_n - \overline{S}_n(a))] \\
&= \mathrm{Tr}[\overline{\rho_n}^{1-s}\overline{\rho_n}^{s}(I_n - \overline{S}_n(a))] \\
&\le e^{nas}\,\mathrm{Tr}[\overline{\rho_n}^{1-s}\sigma_n^s(I_n - \overline{S}_n(a))] \\
&\le e^{nas}\,\mathrm{Tr}[\overline{\rho_n}^{1-s}\sigma_n^s].
\end{aligned} \tag{25}$$

In the same way, (23) yields

$$\begin{aligned}
\beta_n(\overline{S}_n(a)) &= \mathrm{Tr}[\sigma_n\overline{S}_n(a)] \\
&= \mathrm{Tr}[\sigma_n^s\sigma_n^1\overline{S}_n(a)] \\
&\le e^{-na(1-s)}\,\mathrm{Tr}[\sigma_n^s\overline{\rho_n}^{1-s}\overline{S}_n(a)] \\
&\le e^{-na}e^{nas}\,\mathrm{Tr}[\overline{\rho_n}^{1-s}\sigma_n^s].
\end{aligned} \tag{26}$$

It follows from (63) and (17) that

$$\begin{aligned}
\mathrm{Tr}[\overline{\rho_n}^{1-s}\sigma_n^s] &= \mathrm{Tr}\left[\overline{\rho_n}\sigma_n^{\frac{s}{2}}\overline{\rho_n}^{-s}\sigma_n^{\frac{s}{2}}\right] \\
&= \mathrm{Tr}\left[\rho_n\sigma_n^{\frac{s}{2}}\overline{\rho_n}^{-s}\sigma_n^{\frac{s}{2}}\right] \\
&\le (n+1)^{sd}\,\mathrm{Tr}\left[\rho_n\sigma_n^{\frac{s}{2}}\rho_n^{-s}\sigma_n^{\frac{s}{2}}\right]
\end{aligned}$$

$$= (n+1)^{sd} \left(\mathrm{Tr} \left[\rho \sigma^{\frac{s}{2}} \rho^{-s} \sigma^{\frac{s}{2}} \right] \right)^n$$

$$= (n+1)^{sd} e^{-n\overline{\psi}(s)} \tag{27}$$

for all $0 \le s \le 1$. Combining (25)–(27), we have

$$\alpha_n \left(\overline{S}_n(a) \right) \le (n+1)^{sd} e^{-n\left[\overline{\psi}(s) - as\right]}$$

$$\le (n+1)^{d} e^{-n\left[\overline{\psi}(s) - as\right]}, \tag{28}$$

$$\beta_n \left(\overline{S}_n(a) \right) \le (n+1)^{sd} e^{-n\left[\overline{\psi}(s) - as + a\right]}$$

$$\le (n+1)^{d} e^{-n\left[\overline{\psi}(s) - as + a\right]}, \tag{29}$$

which lead to (20) and (21) by taking the maximum in the exponents. □

4 Proof of Theorem 278 and the Quantum Stein's Lemma

In this section, we will prove Theorem 278 by using Theorem 279. To this end, the behavior of $\overline{\varphi}(a)$ in the error exponents (20) and (21) is investigated in the following lemmas. We will also show that Theorem 279 provides a simple proof of the direct part of the quantum Stein's lemma (10).

Lemma 280 *$\overline{\varphi}(a)$ is convex and monotonically nonincreasing.*

Proof The assertion immediately follows from the definition of $\overline{\varphi}(a)$. Actually, we have for all $0 \le t \le 1$

$$\overline{\varphi}(ta + (1-t)b) = \max_{0 \le s \le 1} \{\overline{\psi}(s) - (ta + (1-t)b)s\}$$

$$\le t \max_{0 \le s \le 1} \{\overline{\psi}(s) - as\} + (1-t) \max_{0 \le s \le 1} \{\overline{\psi}(s) - bs\}$$

$$= t\overline{\varphi}(a) + (1-t)\overline{\varphi}(b). \tag{30}$$

Next, let $a \le b$ and $s_b \triangleq \arg\max_{0 \le s \le 1} \{\overline{\psi}(s) - bs\}$. Then we have

$$\overline{\varphi}(b) = \overline{\psi}(s_b) - bs_b$$

$$\le \overline{\psi}(s_b) - as_b$$

$$\leq \max_{0\leq s\leq 1}\{\overline{\psi}(s) - as\}$$

$$= \overline{\varphi}(a). \tag{31}$$

□

Lemma 281 *$\overline{\varphi}(a)$ ranges from 0 to infinity.*

Proof Since we can calculate the derivative of $\overline{\psi}(s)$ explicitly, $\overline{\psi}(s)$ is continuous and differentiable. Therefore, it follows from the mean value theorem that for $s > 0$ there exists $0 \leq t \leq s$ such that

$$\overline{\psi}(t) = \frac{\overline{\psi}(s) - \overline{\psi}(0)}{s - 0}. \tag{32}$$

Let $a \leq \max_{0\leq t\leq 1} \overline{\psi}'(t)$, then we have

$$a \geq \frac{\overline{\psi}(s) - \overline{\psi}(0)}{s - 0}. \tag{33}$$

and hence,

$$\overline{\psi}(0) \geq \overline{\psi}(s) - as \tag{34}$$

which yields

$$0 = \overline{\psi}(0) = \max_{0\leq s\leq 1}\{\overline{\psi}(s) - as\} = \overline{\varphi}(a). \tag{35}$$

On the other hand, it is obvious that

$$\lim_{a\to-\infty} \overline{\varphi}(a) = \infty. \tag{36}$$

Since $\overline{\varphi}(a)$ is continuous, which follows from convexity by Lemma 280, the assertion follows from (35) and (36). □

Combined with the above lemma, Theorem 279 leads to Theorem 278 as follows.

Proof of Theorem 278 For all $r > 0$, there exists $a_r \in \mathbb{R}$ such that $r = \overline{\varphi}(a_r)$ from Lemma 281. Let $\overline{u}(r) \triangleq \overline{\varphi}(a_r) + a_r$, then it follows from Theorem 279 that

$$\limsup_{n\to\infty} \frac{1}{n} \log \alpha_n(\overline{S}n(a_r)) \leq -r \tag{37}$$

$$\limsup_{n\to\infty} \frac{1}{n} \log \beta_n \overline{S}_n(a_r)) \leq -\overline{u}(r). \tag{38}$$

Therefore, it suffices to show that

$$\overline{u}(r) = \max_{0 \le s \le 1} \frac{\overline{\psi} - (1-s)r}{s} \tag{39}$$

For all $0 \le s \le 1$, we have from the definition of $\overline{\varphi}(a)$

$$r = \overline{\varphi}(a_r) \ge \overline{\psi}(s) - a_r s \tag{40}$$

and there exists a number s_0, $0 < s_0 \le 1$, achieving the equality since $r = \overline{\varphi}(a_r) > 0$. On the other hand, the definitions of $\overline{u}(r)$ and a_r lead to

$$\overline{u}(r) = \overline{\varphi}(a_r) + a_r = r + a_r. \tag{41}$$

Eliminating a_r from (40) and (41), we have

$$\overline{u}(r) \ge \frac{\overline{\psi}(s) - (1-s)r}{s}) \tag{42}$$

and s_0 achieves the equality in (42) as well. Thus, we have shown (39), and Theorem 278 has been proved. □

Next, observing that $\overline{\psi}(0) = 0$ and $\overline{\psi}'(0) = D(\rho||\sigma)$, we have

$$\overline{\varphi}(a) > 0 \qquad \text{for all } a < D(\rho||\sigma) \tag{43}$$

which leads to the following theorem combined with Theorem 279.

Theorem 282 (Ogawa and Hayashi 2004, [13]) *For all $a < D(\rho||\sigma)$, we have*

$$\lim_{n\to\infty} \alpha_n(\overline{S}_n(a)) = 0 \tag{44}$$

$$\limsup_{n\to\infty} \frac{1}{n} \log \beta_n(\overline{S}_n(a)) \le -a. \tag{45}$$

Since $a < D(\rho||\sigma)$ can be arbitrarily near $D(\rho||\sigma)$, we have shown the direct part of the quantum Stein's lemma (10).

5 Toward Further Investigations

The error exponents derived here do not seem to be natural, since $\overline{\psi}(s)$ lacks symmetry between ρ and σ that the original hypothesis testing problem has. We need further investigation to determine the error exponents in quantum hypothesis testing. In this section, we make a few remarks on some candidates for the

alternative to $\overline{\psi}(s)$ in the expectation that the error exponents would be written in the form of Theorem 278.

Among many candidates, let us consider the following functions:

$$\psi_1(s) \triangleq \max\left\{\overline{\psi}(s), \tilde{\psi}(s)\right\} \tag{46}$$

$$\psi_2(s) \triangleq -\log \mathrm{Tr}\left[\rho^{1-s}\sigma^s\right] \tag{47}$$

$$\psi_3(s) \triangleq -\log \mathrm{Tr}\left[e^{(1-s)\log\rho + s\log\sigma}\right] \tag{48}$$

where

$$\tilde{\psi}(s) \triangleq -\log \mathrm{Tr}\left[\sigma\rho^{\frac{1-s}{2}}\sigma^{-(1-s)}\rho^{1-s}2\right] \tag{49}$$

and define the corresponding functions

$$u_i(r) \triangleq \max_{0<2\le 1} \frac{\psi_i(s) - (1-s)r}{s} \qquad i = 1, 2, 3. \tag{50}$$

The reason to consider these functions is as follows. First $\psi_1(s)$ is a symmetrized version of $\overline{\psi}(s)$, and Theorem 278 still holds with $\overline{\psi}(s)$ replaced by $\psi_1(s)$, since similar upper bounds to Theorem 279 using $\tilde{\psi}(s)$ are valid by exchanging ρ and σ and replacing s with $1-s$. On the other hand, $\psi_2(s)$ for $-1 \le s \le 0$ appeared in [14] to show the strong converse property in quantum hypothesis testing. Concerning $\psi_3(s)$, $u_3(r)$ is a quantum analogue of (2). Actually, we can show that

$$u_3(r) = \min_{\rho' : D(\rho'||\rho) \le r} D(\rho'||\rho) \tag{51}$$

by the same way as [14, Sec. VI]. At present it is not clear whether $u_2(r)$ and $u_3(r)$ are achievable exponents in quantum hypothesis testing. It should be noted, however, that $\psi_i(s)$, $i = 1, 2, 3$, are reduced to the classical one (4) if ρ and σ commute, and they have desirable properties

$$\begin{aligned} &\psi_i(0) = \psi_i(1) = 0 \\ &\psi_i'(0) = D(\rho||\sigma), \\ &\psi_i'(1) = D(\rho||\sigma) \qquad i = 1, 2, 3 \end{aligned} \tag{52}$$

which are consistent with the quantum Stein's lemma. The above properties of $\psi_2(s)$ and $\psi_3(s)$ are verified by the direct calculations while those of $\psi_1(s)$ follow from

the following fact:

$$\psi_1(s) = \overline{\psi}(s) \geq \tilde{\psi}(s), \qquad \text{if } s \text{ is sufficiently near 0} \tag{53}$$

$$\psi_1(s) = \tilde{\psi}(s) \geq \overline{\psi}(s), \qquad \text{if } s \text{ is sufficiently near 1} \tag{54}$$

which is a consequence of $\overline{\psi}(0) = \psi_2(0)$, $\tilde{\psi}(1) = \psi_2(1)$, and the following lemma.

Lemma 283 *For all* $0 \leq s \leq 1$*, we have*

$$\overline{\psi}(s) \leq \psi_2(s) \tag{55}$$

$$\tilde{\psi}(s) \leq \psi_2(s) \tag{56}$$

Proof Let us apply the monotonicity property of the quantum quasi-entropy [17, 18] to $\mathrm{Tr}[\rho^{1-s}\sigma^s]$, $0 \leq s \leq 1$,[2] so that we have

$$\begin{aligned} e^{-n\psi_2(s)} &= \left(\mathrm{Tr}[\rho^{1-s}\sigma^s]\right)^n \\ &= \mathrm{Tr}[\rho_n^{1-s}\sigma_n^s] \\ &\leq \mathrm{Tr}[\overline{\rho_n}^{1-s}\sigma_n^s] \\ &\leq (n+1)^{sd} e^{-n\overline{\psi}(s)} \end{aligned} \tag{57}$$

where we used (27) in the last inequality. Thus, we obtain

$$\overline{\psi}(s) \leq \psi_2(s) + \frac{sd}{n}\log(n+1) \tag{58}$$

for any natural number n, and we have (55) by letting n go to infinity. Exchanging ρ and σ and replacing s with $1-s$ in (55), we obtain (56). □

It follows immediately from Lemma 283 that $\psi_1(s) \leq \psi_2(s)$, and it was pointed out in [14] that we have $\psi_2(s) \leq \psi_3(s)$ as a consequence of the Golden-Thompson inequality (see, e.g., [16, p. 128])

$$\mathrm{Tr}\left[e^{A+B}\right] \leq \mathrm{Tr}\left[e^A e^B\right] \tag{59}$$

[2] Comprehensible explanations of the monotonicity property are found in [1, Sec. 7.2] and [14].

for Hermitian operators A and B with the equality if and only if A and B commute. These facts are stated as the following proposition

Proposition 284 *It holds that*

$$\psi_1(s) \leq \psi_2(s) \leq \psi_3(s) \qquad \forall\, 0 \leq s \leq 1 \tag{60}$$

$$u_1(r) \leq u_2(r) \leq u_3(r) \qquad \forall\, r > 0 \tag{61}$$

Especially, if ρ and σ do not commute, we have $\psi_2(s) < \psi_2(s)$ and $u_2(r) < u_3(r)$.

As mentioned above, $u_1(r)$ is an achievable exponent in quantum hypothesis testing, while it is not known whether $u_2(r)$ and $u_3(r)$ are achievable or not. It is interesting to study the achievability of these functions, especially that of $u_2(r)$, and the problem is left open.

6 Concluding Remarks

In the quantum hypothesis problem, we have presented upper bounds on the error probabilities of the first and the second kind, based on a key operator inequality satisfied by a density operator and pinching of it. The upper bounds are regarded as a noncommutative analogue of the Hoeffding bound [9], which is the optimal bound in classical hypothesis testing, and the upper bounds provide a simple proof of the direct part of the quantum Stein's lemma. Compared with [6], the proof is considerably simple and leads to the exponential convergence of the error probability of the first kind.

7 Definition of Pinching

In this section, we summarize the definition of pinching (see, e.g., [2, p. 50]) for readers' convenience. Pinching is known as a special notion of the conditional expectation in the field of operator algebra.

Given a Hermitian operator $A \in \mathcal{L}(\mathcal{H})$, let $A = \sum_{i=1}^{v(A)} a_i E_i$ be its spectral decomposition, where $v(A)$ is the number of eigenvalues of A mutually different from others, and each E_i is the projection corresponding to an eigenvalue a_i. The following map defined by using the PVM $E = \{Ei\}_{i=1}^{v(A)}$ is called pinching:

$$\mathcal{E}_A : B \in \mathcal{L}(\mathcal{H}) \to \mathcal{E}_A(B) = \triangleq \sum_{i=1}^{v(A)} E_i B E_i \in \mathcal{L}(\mathcal{H}). \tag{62}$$

The operator $\mathcal{E}_A(B)$ is also called pinching when no confusion is likely to arise, and it is sometimes denoted as $\mathcal{E}_E(B)$. It should be noted here that pinching is the conditional expectation (with respect to the tracial state) to the commutant of the $*$-subalgebra generated by A or PVM E, since $\mathcal{E}_A(B)$ is the one and only operator which satisfies

$$\mathrm{Tr}[BC] = \mathrm{Tr}[\mathcal{E}_A(B)C] \tag{63}$$

for any operator $C \in \mathcal{L}(\mathcal{H})$ commuting with A.

8 Key Operator Inequality

The following lemma has played an important role in this lecture. Although the lemma for a two-valued PVM has been widely used, it appeared in [6] for the general case. Here, we will show another proof of it for readers' convenience.

Lemma 285 (Hayashi 2002, [6]) *Given a PVM $M = \{M_i\}_{i=1}^{v(M)}$ on $\mathcal{H}$, we have for all $\rho \in \mathcal{S}(\mathcal{H})$*

$$\rho \le v(M)\mathcal{E}_M(\rho) \tag{64}$$

where $\mathcal{E}_M(\rho)$ is the pinching defined in 7.

Proof First, note that the following map, defined with respect to a non-negative operator $A \in \mathcal{L}(\mathcal{H})$, is operator convex

$$f_A : X \in \mathcal{L}(\mathcal{H}) \to X^*AX \in \mathcal{L}(\mathcal{H}) \tag{65}$$

which is shown by a direct calculation

$$tf_A(X)+(1-t)f_A(Y)-f_A(tX+(1-t)Y) = t(1-t)(X-Y)^*A(X-Y) \ge 0 \tag{66}$$

for $0 \le t \le 1$. Using the convexity, the lemma is verified as follows:

$$\begin{aligned}\frac{1}{v(M)^2}\rho &= \left(\frac{1}{v(M)}\sum_{i=1}^{v(M)} M_i\right)\rho\left(\frac{1}{v(M)}\sum_{i=1}^{v(M)} M_i\right)\\ &\le \frac{1}{v(M)}\sum_{i=1}^{v(M)} M_i\rho M_i\\ &= \frac{1}{v(M)}\mathcal{E}_M(\rho).\end{aligned} \tag{67}$$

□

References

1. S. Amari, H. Nagaoka, *Methods of Information Geometry* (AMS/Oxford University, Oxford, 1993)
2. R. Bhatia, *Matrix Analysis* (Springer, New York, 1997)
3. R.E. Blahut, Hypothesis testing and information theory. IEEE Trans. Inf. Theory **IT-20**, 405–417 (1974)
4. R.E. Blahut, *Principles and Practice of Information Theory* (Addison-Wesley, Reading, 1991)
5. M. Hayashi, Exponents of quantums fixed-length pure state source coding. Phys. Rev. A **66**(3), 032321 (2002)
6. M. Hayashi, Optimal sequence of POVM's in the sense of Stein's lemma in quantum hypothesis testing. J. Phys. A Math. Gen. **35**, 10759–10773 (2002)
7. M. Hayashi, H. Nagaoka, General formulas for capacity of classical-quantum channels. IEEE Trans. Inf. Theory **49**(7), 1753–1768 (2003)
8. F. Hiai, D. Petz, The proper formula for relative entropy and its asymptotics in quantum probability. Commun. Math. Phys. **143**, 99–114 (1991)
9. W. Hoeffding, On probabilities of large deviations, in *Proceedings of the 5th Berkeley Symposium Mathematical Statistics and Probability, Berkeley, CA* (1965), pp. 203–219
10. H. Nagaoka, On asymptotic theory of quantum hypothesis testing, in *Proceedings of the Symposium Statistical Inference Theory and its Information Theoretical Aspect* (1998), pp. 49–52
11. H. Nagaoka, Strong converse theorems in quantum information theory, in *Proceedings of the ERATO Workshop in Quantum Information Science* (2001)
12. H. Nagaoka, M. Hayashi, *An Information-Spectrum Approach to Classical and Quantum Hypothesis Testing for Simple Hypotheses* (2002)
13. T. Ogawa, M. Hayashi, On error exponents in quantum hypothesis testing. IEEE Trans. Inf. Theory **50**(6), 1368–1372 (2004)
14. T. Ogawa, H. Nagaoka, Strong converse and Stein's lemma in quantum hypothesis testing. IEEE Trans. Inf. Theory **46**, 2428–2433 (2000)
15. T. Ogawa, H. Nagaoka, A new proof of the channel coding theorem via hypothesis testing in quantum information theory, in *Proceedings of the 2002 IEEE International Symposium Information Theory, Lausanne, Switzerland* (2002)
16. M. Ohya, D. Petz, *Quantum Entropy and its Use, Berlin/Heidelberg* (Springer, Germany, 1993)
17. D. Petz, *Quasi-entropies for States of a von Neumann Algebra* (RIMS, Kyoto University, Kyoto, 1985), pp. 787–800
18. D. Petz, Quasi-entropies for finite quantum systems. Rep. Math. Phys. **23**, 57–65 (1986)