Second Order Asymptotics for Quantum Hypothesis Testing

Ke Li*
Centre for Quantum Technologies, National University of Singapore, 3 Science Drive 2, Singapore 117543
(Dated: November 20, 2012)

(Full version available at arXiv[quant-ph]:1208.1400)

Abstract In the asymptotic theory of quantum hypothesis testing, the error probability of the first kind jumps sharply from zero to one when the error exponent of the second kind passes by the point of the relative entropy of the two states, in an increasing way. This is well known as the direct part and strong converse of quantum Stein's lemma.

Here we look into the behavior of this sudden change and have make it clear how the error of first kind grows according to a lower order of the error exponent of the second kind, and hence, we obtain the second order asymptotics for quantum hypothesis testing. Meanwhile, our analysis also yields tight bounds for the case of finite sample size. These results have nice applications in quantum information theory. Our method is elementary, based on basic linear algebra and probability theory. It deals with the achievability part and the converse part in a unified framework, with a clear geometric picture.

Introduction. In the problem of asymptotic hypothesis testing, there are two hypotheses, each being many copies of independent and identically-distributed instances, occurring according to some given statistical description ρ or σ, respectively. Here the statistical descriptions are probability distributions in classical setting and quantum states in quantum mechanics. The task is to distinguish these two hypotheses with minimal error probabilities.

Classically, this problem has been well understood [1]. Moving to the quantum case, it becomes much more difficult due to the non-commutativity of the quantum states ρ and σ, and the more complicated mechanics for observing the physical systems of interest(i.e., quantum measurement). Substantial achievements have already been made in the asymptotic theory of quantum hypothesis testing. Most notably, these include the establishment of the quantum Stein's lemma with a strong converse [2, 3], the quantum Chernoff bound [4, 5], and the quantum Hoeffding bound [6–8].

Given a large number n of identical quantum systems, which are either of the state $\rho^{\otimes n}$ (the null hypothesis) or of the state $\sigma^{\otimes n}$ (the alternative hypothesis), we want to identify which state the systems belong to. This is achieved by doing a two outcome measurement $\{A_n, \mathbb{1} - A_n\}$. We define two types of errors. Type I error (or the error of the first kind) is the probability that we falsely conclude that the state is σ while it is actually ρ, given by $\alpha_n(A_n) := \operatorname{Tr} \rho^{\otimes n}(\mathbb{1} - A_n)$; type II error (the error of the second kind) instead is the probability that we mistake σ for ρ, given by $\beta_n(A_n) := \operatorname{Tr} \sigma^{\otimes n} A_n$. In an asymmetric situation, we want to minimize the type II error while only simply requiring that the type I error converges to 0. The quantum Stein's lemma states that the maximal error exponent of type II is the relative entropy $D(\rho\|\sigma)$:
(direct part [2]) For arbitrary $R < D(\rho\|\sigma)$, there exists a test $\{A_n, \mathbb{1} - A_n\}$ satisfying

$$\liminf_{n\to\infty} \frac{-1}{n} \log \beta_n(A_n) \geq R \quad \text{and} \quad \lim_{n\to\infty} \alpha_n(A_n) = 0;$$

(strong converse [3]) If a test $\{A_n, \mathbb{1} - A_n\}$ is such that

$$\liminf_{n\to\infty} \frac{-1}{n} \log \beta_n(A_n) > D(\rho\|\sigma),$$

then $\lim_{n\to\infty} \alpha_n(A_n) = 1$.

Second order asymptotics. From the quantum Stein's lemma, we see that the error probability of the first kind jumps sharply from zero to one when the error exponent of the second kind passes by the relative entropy $D(\rho\|\sigma)$ from the smaller side to the larger side. In this paper,

we resort to a smaller order of the type II error exponent and have understood how this sudden change happens. After defining the error-dependency function, we present our result in Theorem 2 as follows.

Definition 1 *We define the function $\alpha_n(E_1, E_2|f)$, which reflects the dependency of the optimal error probability of the first kind on the error exponent of the second kind, up to the order n and $\sqrt{n}$, as*

$$\alpha_n(E_1, E_2|f) := \min_{A_n} \left\{ \alpha_n(A_n) \middle| \beta_n(A_n) \leq \exp\big(-(E_1 n + E_2\sqrt{n} + f(n))\big) \right\}, \tag{1}$$

where $f(n)$ is a function of some order other than n and $\sqrt{n}$, which is to be specified when necessary.

Theorem 2 *Let $\Phi(x)$ be the cumulative distribution function of the standard normal distribution, i.e., $\Phi(x) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\,\mathrm{d}t$, we have*

$$\lim_{n\to\infty} \alpha_n(E_1, E_2|f) = \begin{cases} 0 & \text{if } E_1 < D(\rho\|\sigma), f \in o(n) \\ \Phi\left(\frac{E_2}{\sqrt{V(\rho\|\sigma)}}\right) & \text{if } E_1 = D(\rho\|\sigma), f \in o(\sqrt{n}) \\ 1 & \text{if } E_1 > D(\rho\|\sigma), f \in o(n), \end{cases} \tag{2}$$

where $V(\rho\|\sigma)$, which we name the quantum relative variance of ρ and σ, is defined as

$$V(\rho\|\sigma) := \operatorname{Tr} \rho(\log\rho - \log\sigma)^2 - D^2(\rho\|\sigma).$$

Our result provides a fundamental deepening to quantum Stein's lemma, in a way just like the central limit theorem does to the law of large numbers. The first and third cases of Eq. (2) are nothing else but the direct part and strong converse of quantum Stein's lemma, respectively. The second case is the second order asymptotics. In fact, it is easy to show that the second case implies the first and third cases. (We include them in the theorem such that one easily gets the full information at first sight.) *So, our second order asymptotics is not only about the point "$E_1 = D(\rho\|\sigma)$"; instead, it characterizes the whole interval "$E_1 \geq 0$".*

The proof is described shortly as follows. At first, we divide Theorem 2 into the achievability part and the optimality part, and express them equivalently in a more technical way as in Theorem 3. Then, we prove these two parts separately (but in the same framework) using a direct and elementary method, based on basic linear algebra and probability theory. For the achievability part, we construct explicitly a projective measurement A_n as the test, specifying a basis of its supporting space. For the optimality part, we evaluate the type I error $\alpha_n(A_n)$ in a geometrically intuitive way.

Theorem 3 *For quantum hypothesis testing with the null hypothesis $\rho^{\otimes n}$ and the alternative hypothesis $\sigma^{\otimes n}$, and the error probabilities of the first and second kinds denoted as $\alpha_n(A_n)$ and $\beta_n(A_n)$, respectively, we have*
(Achievability): For any $E_2 \in \mathbb{R}$ and $f(n) \in o(\sqrt{n})$, there exists a sequence of measurements $\{A_n, \mathbb{1} - A_n\}_n$, such that

$$\beta_n(A_n) \leq \exp\left\{-\big(nD(\rho\|\sigma) + E_2\sqrt{n} + f(n)\big)\right\}, \tag{3}$$

$$\limsup_{n\to\infty} \alpha_n(A_n) \leq \Phi\left(\frac{E_2}{\sqrt{V(\rho\|\sigma)}}\right); \tag{4}$$

(Optimality): If there is a sequence of measurements $\{A_n, \mathbb{1} - A_n\}_n$ such that

$$\beta_n(A_n) \leq \exp\left\{-\big(nD(\rho\|\sigma) + E_2\sqrt{n} + f(n)\big)\right\} \tag{5}$$

holds for given $E_2 \in \mathbb{R}$ and $f(n) \in o(\sqrt{n})$, then

$$\liminf_{n\to\infty} \alpha_n(A_n) \geq \Phi\left(\frac{E_2}{\sqrt{V(\rho\|\sigma)}}\right). \tag{6}$$

One technical tool we used is Lemma 4 below, which is similar to the quantum-to-classical mapping first appeared in [4]. Without loss of generality, we suppose that σ is of full rank (see full version). Write $\rho = \sum_x \lambda(x)|a_x\rangle\langle a_x|$ and $\sigma = \sum_y \mu(y)|b_y\rangle\langle b_y|$ in their diagonal form, where $\{|a_x\rangle\langle a_x|\}_x$ and $\{|b_y\rangle\langle b_y|\}_y$ are two orthonormal bases. Then, we write $|a_x\rangle = \sum_y \gamma_{xy}|b_y\rangle$, with $\gamma_{xy} = \langle b_y|a_x\rangle \in \mathbb{C}$ and $\sum_x |\gamma_{xy}|^2 = \sum_y |\gamma_{xy}|^2 = 1$. Define a pair of random variables (X, Y), with alphabet $\{(x, y)\}_{x,y=1}^{|\mathcal{H}|}$ and joint distribution $P_{X,Y}(x, y) = \lambda(x)|\gamma_{xy}|^2$.

Lemma 4 *The quantum relative entropy and relative variance can be expressed as statistical quantities of classical random variables:*

$$D(\rho\|\sigma) = \mathrm{E}_{(X,Y)} \log \frac{\lambda(X)}{\mu(Y)}, \tag{7}$$

$$V(\rho\|\sigma) = \mathrm{Var}_{(X,Y)} \log \frac{\lambda(X)}{\mu(Y)}. \tag{8}$$

Finite sample-size analysis. (This part is not appeared in the arXiv full version yet, but can be derived from that straightforwardly.) During the derivation of the second order asymptotics, if we use the Berry-Esseen theorem [9] instead of the central limit theorem, and reserve all the $o(1)$ terms in the valuation of $\alpha_n(A_n)$, and then make optimizations over some parameters, we obtain for finite n the following bounds, which is relatively tighter and cleaner compared to that obtained in [13].

$$\begin{aligned}
&\exp\left\{-\left(nD(\rho\|\sigma) + \sqrt{n}\sqrt{V(\rho\|\sigma)}\Phi^{-1}\Big(\varepsilon + \frac{1}{\sqrt{n}}\big(\frac{CT}{\sqrt{V(\rho\|\sigma)}^3} + 1\big)\Big) + \log(2^9 n^2)\right)\right\} \\
&\leq \beta_n(\varepsilon) := \min_{A_n}\big\{\beta_n(A_n)\big|\alpha_n(A_n) \leq \varepsilon\big\} \leq \\
&\exp\left\{-\left(nD(\rho\|\sigma) + \sqrt{n}\sqrt{V(\rho\|\sigma)}\Phi^{-1}\Big(\varepsilon - \frac{1}{\sqrt{n}}\frac{CT}{\sqrt{V(\rho\|\sigma)}^3}\Big)\right)\right\},
\end{aligned} \tag{9}$$

where $0.40973 \leq C \leq 0.4784$ is a constant and $T = \mathrm{E}_{(X,Y)}\left|\log\frac{\lambda(X)}{\mu(Y)} - D(\rho\|\sigma)\right|^3$. This means

$$-\log\beta_n(\varepsilon) \sim nD(\rho\|\sigma) + \sqrt{n}\sqrt{V(\rho\|\sigma)}\Phi^{-1}(\varepsilon) + O(\log n). \tag{10}$$

Especially, the leading order in $O(\log n)$ of Eq. (10) (this is the third order asymptotics) must be positive unless it is a constant term. This is because, from the upper bound in Eq. (9), we see that it is achievable with $O(\log n)$ being a constant term.

Applications. There is a deep connection between hypothesis testing and other information processing tasks, both in the classical regime [10, 11] and in the quantum regime [12]. Hence the result obtained for quantum hypothesis testing here, can be used to study the second order asymptotics and finite block-length analysis for other quantum tasks. For example, data compression with quantum side information and randomness extraction against quantum side information [13], and classical information transmission over quantum channels [14]. Besides, we hope that the method employed in this work, will find applications in other problems of non-commutative probability and statistics.

Note added: The result of this work was independently and concurrently obtained in [13], using different method. In [13], they also have done something else. The bounds for finite sample size in these two works are slightly different. This information will be added in an update of my full version arXiv[quant-ph]:1208.1400.

---

* Electronic address: carl.ke.lee@gmail.com

[1] T. M. Cover, J. A. Thomas, *Elements of Information Theory* (New York: Wiley, 1991).
[2] F. Hiai, D. Petz, Comm. Math. Phys. **143**, 99 (1991).
[3] T. Ogawa, H. Nagaoka, IEEE Trans. Inf. Theory **46**, 2428 (2000).
[4] M. Nussbaum, A. Szkoła, Ann. Stat. **37**, 1040 (2009).
[5] K. M. R. Audenaert, J. Calsamiglia, Ll. Masanes, R. Munoz-Tapia, A. Acin, E. Bagan, F. Verstraete, Phys. Rev. Lett. **98**, 160501 (2007).
[6] M. Hayashi, Phys. Rev. A **76**, 062301 (2007).
[7] H. Nagaoka, Arxiv preprint quant-ph/0611289 (2006).
[8] K. M. R. Audenaert, M. Nussbaum, A. Szkoła, F. Verstraete, Comm. Math. Phys. **279**, 251 (2008).
[9] V. Korolev, I. Shevtsova, Scandinavian Actuarial Journal **2012**(2):81–105 (2012).
[10] S. Verdú, T. S. Han, IEEE Trans. Inf. Theory **40**, 1147 (1994).
[11] T. S. Han, *Information-Spectrum Methods in Information Theory* (Berlin: Springer, 2003).
[12] M. Hayashi, H. Nagaoka, IEEE Trans. Inf. Theory **49**, 1753 (2003).
[13] M. Tomamichel, M. Hayashi, arXiv[quant-ph]:1208.1478.
[14] M. Hayashi, Ke Li, in preparation (2012).