Strong converse theorems in quantum information theory

Hiroshi Nagaoka *

Abstract: There are many situations in information theory where it is required to find the threshold value of some relevant quantity for which a kind of error probability can asymptotically be made arbitrarily small. The strong converse property means that when the quantity exceeds the threashold value the error probability inevitably goes to one asymptotically. In quantum information theory, the strong converse property is known for (classical) capacity of quantum memoryless channel ([1, 2]) and for hypothesis testing problem of two quantum i.i.d. states ([3]). In this article we provide new proofs, which are based on a simple and unified argument, with these results.

1 Key inequality

Let ρ and σ be density operators on a common Hilbert space $\mathcal{H}$. For any $s \in [0,1]$ and any POVM $\{X_i\}_{i=1}^k$ it holds that $\sum_{i=1}^k (\mathrm{Tr}\,[\rho X_i]))^{1+s}\ (\mathrm{Tr}\,[\sigma X_i]))^{-s} \leq \mathrm{Tr}\,[\rho^{1+s}\sigma^{-s}]$, which follows from a result in [4] and was used to prove the strong converse theorem for quantum hypothesis testing in [3]. In particular, for any self-adjoint operator A satisfying $0 \leq A \leq I$, which we call a *test* in general, we have

$$(\mathrm{Tr}\,[\rho A])^{1+s}\ (\mathrm{Tr}\,[\sigma A])^{-s} \leq \mathrm{Tr}\,[\rho^{1+s}\sigma^{-s}]. \qquad (1)$$

This is the key inequality for the arguments below.

2 Hypothesis testing

Given a test A_n on $\mathcal{H}^{\otimes n}$, let $\alpha_n(A_n) \stackrel{\mathrm{def}}{=} 1 - \mathrm{Tr}\,[\rho^{\otimes n}A_n]$ and $\beta_n(A_n) \stackrel{\mathrm{def}}{=} \mathrm{Tr}\,[\sigma^{\otimes n}A_n]$, which are interpreted as the error probabilities of the first and second kinds, and apply the above inequality to $\rho^{\otimes n}$, $\sigma^{\otimes n}$ and A_n to yield

$$\frac{1+s}{n}\log(1-\alpha_n(A_n)) - \frac{s}{n}\log\beta_n(A_n) \leq \psi(s),$$

where $\psi(s) \stackrel{\mathrm{def}}{=} \log \mathrm{Tr}\,[\rho^{1+s}\sigma^{-s}]$. From $\lim_{s\downarrow 0}\psi(s)/s = \mathrm{Tr}\,[\rho(\log\rho - \log\sigma)] \stackrel{\mathrm{def}}{=} D(\rho\,||\,\sigma)$, we see that if $r \stackrel{\mathrm{def}}{=} -\limsup_{n\to\infty}\frac{1}{n}\log\beta_n(A_n) > D(\rho\,||\,\sigma)$ then $\alpha_n(A_n)$ neccessarily goes to one with exponential rate

$$\limsup_{n\to\infty}\frac{1}{n}\log(1-\alpha_n(A_n)) \leq -\max_{0\leq s\leq 1}\left\{\frac{s}{1+s}r - \frac{\psi(s)}{1+s}\right\}.$$

3 Channel coding

Let Δ be a compact subset of $\mathcal{S}(\mathcal{H})$ (:the set of density operators on $\mathcal{H}$), which is supposed to be the set of possible received states of a quantum communication channel. For $n, M \in \mathbb{N}$ and $\varepsilon > 0$, an (n, M, ε)-code consists of a family of product states $\rho_i^{(n)} = \rho_{i,1}\otimes\cdots\otimes\rho_{i,n}$ with $\rho_{i,j} \in \Delta$, $i = 1, \ldots, M$, and a POVM $X^{(n)} = \{X_i^{(n)}\}_{i=1}^M$ on $\mathcal{H}^{\otimes n}$ satisfying $\frac{1}{M}\sum_i \mathrm{Tr}\,[\rho_i^{(n)}X_i^{(n)}] \geq 1-\varepsilon$. Setting ρ, σ and A in (1) to be the block diagonal matrices $\frac{1}{M}\mathrm{diag}\{\rho_1^{(n)}, \ldots, \rho_M^{(n)}\}$, $\frac{1}{M}\mathrm{diag}\{\sigma^{\otimes n}, \ldots, \sigma^{\otimes n}\}$ and $\mathrm{diag}\{X_1^{(n)}, \ldots, X_M^{(n)}\}$, where σ is an arbitrary state in $\mathcal{S}(\mathcal{H})$, we have

$$\begin{aligned}(1-\varepsilon)^{1+s}M^s &\leq \frac{1}{M}\sum_{i=1}^M\prod_{j=1}^n \mathrm{Tr}\,[\rho_{i,j}^{1+s}\sigma^{-s}] \\ &\leq \max_{\rho\in\Delta}\left(\mathrm{Tr}\,[\rho^{1+s}\sigma^{-s}]\right)^n.\end{aligned}$$

It follows that if a sequence of (n, M_n, ε_n) codes for $n = 1, 2, \ldots$ satisfies

$$\begin{aligned}R &\stackrel{\mathrm{def}}{=} \liminf_{n\to\infty}\frac{1}{n}\log M_n > \min_{\sigma\in\mathcal{S}(\mathcal{H})}\max_{\rho\in\Delta} D(\rho\,||\,\sigma) \\ &= \max_{\lambda,\{\rho_i\}\subset\Delta}\sum_i \lambda_i D(\rho_i\,||\,\sum_j\lambda_j\rho_j),\end{aligned}$$

then ε_n neccessarily goes to one with exponential rate

$$\limsup_{n\to\infty}\frac{1}{n}\log(1-\varepsilon_n) \leq -\max_{0\leq s\leq 1}\left\{\frac{s}{1+s}R - \min_{\sigma\in\mathcal{S}(\mathcal{H})}\max_{\rho\in\Delta}\frac{\log\mathrm{Tr}\,[\rho^{1+s}\sigma^{-s}]}{1+s}\right\}.$$

References

[1] T. Ogawa and H. Nagaoka, *IEEE Trans. Inform. Theory*, vol.45, 2486-2489, 1999.

[2] A. Winter, *IEEE Trans. Inform. Theory*, vol.45, 2481-2485, 1999.

[3] T. Ogawa and H. Nagaoka, *IEEE Trans. Inform. Theory*, vol.46, 2428-2433, 2000.

[4] D. Petz, *Rep. Math. Phys.*, vol.23, 57–65, 1986, *Publ. RIMS, Kyoto Univ.*, vol. 21, 787-800, 1985.

*Graduate School of Information Systems, University of Electro-Communications, 1-5-1, Chofugaoka, Chofu-shi, Tokyo 182-8585, Japan. Email: nagaoka@is.uec.ac.jp