Strong converse exponent for classical-quantum channel coding

Milán Mosonyi[1, 2, *] and Tomohiro Ogawa[3, †]

[1] *Física Teòrica: Informació i Fenomens Quàntics, Universitat Autònoma de Barcelona, ES-08193 Bellaterra (Barcelona), Spain.*

[2] *Mathematical Institute, Budapest University of Technology and Economics, Egry József u 1., Budapest, 1111 Hungary.*

[3] *Graduate School of Information Systems, University of Electro-Communications, 1-5-1 Chofugaoka, Chofu-shi, Tokyo, 182-8585, Japan.*

I. OVERVIEW

We determine the exact strong converse exponent of classical-quantum channel coding, for every rate above the Holevo capacity. Our form of the exponent is an exact analogue of Arimoto's exponent, given as a transform of the Rényi capacities of the channel with parameters $\alpha > 1$.

It is important to note that, unlike in the classical case, there are (infinitely) many inequivalent ways to define the Rényi divergence of quantum states, and hence the Rényi capacities of channels. Our exponent is in terms of the Rényi capacities corresponding to a version of the Rényi divergences (denoted by D_α^* here) that has been introduced recently in [13] and [19]. These Rényi divergences have been shown to be the natural quantifiers of the strong converse trade-off relations in various binary hypothesis testing problems [4, 8, 11]. Our result proves that this distinguished role of the D_α^*-divergences is not restricted to hypothesis testing problems, and supports the expectation that it might hold in any information theoretic problem with two competing operational quantities and a well-defined strong converse region.

It is known that, at least in the problem of binary state discrimination, a different notion of Rényi divergence (denoted by D_α) is needed to quantify the trade-off relations in the direct domain [3, 6, 15], and it is expected that these two versions, D_α and D_α^*, are sufficient to describe the full trade-off curve in a large variety of coding problems. In this work, however, we show that there is at least one more quantum Rényi divergence (denoted by $D_\alpha^\flat$) that is worth considering when extending classical information theoretic results to the quantum domain. Its importance stems from the fact that classical divergence sphere-optimization forms of optimal exponents translate naturally to expressions in terms of the $D_\alpha^\flat$-divergence instead of the correct divergence D_α or D_α^*. Although the resulting exponents are suboptimal in the quantum setting, they may be asymptotically convertible to the right exponents, as we demonstrate on the present example of classical-quantum channel coding.

This submission is based on [12].

II. MAIN RESULT

A classical-quantum channel W is defined by a map $W : \mathcal{X} \to \mathcal{S}(\mathcal{H})$, where $\mathcal{X}$ is an arbitrary set and $\mathcal{S}(\mathcal{H})$ is the set of density operators on a Hilbert space $\mathcal{H}$. n uses of the channel is described by $W^{\otimes n}(\underline{x}) := W(x_1) \otimes \ldots \otimes W(x_n)$, $\underline{x} = x_1 \ldots x_n \in \mathcal{X}^n$. A code $\mathcal{C}_n$ for n uses of the channel consists of an encoding $\phi_n : \{1, 2, \ldots, M_n\} \to \mathcal{X}^n$ of messages into sequences of input signals, and a decoding POVM $D_n = \{D_n(k)\}_{k=1}^{M_n}$ on $\mathcal{H}^{\otimes n}$. The size of the code is $|\mathcal{C}_n| := M_n$. The average success and error probabilities of the code are

$$P_s(W^{\otimes n}, \mathcal{C}_n) = \frac{1}{M_n} \sum_{k=1}^{M_n} \operatorname{Tr} W^{\otimes n}(\phi_n(k)) M_n(k) \qquad \text{and} \qquad P_e(W^{\otimes n}, \mathcal{C}_n) = 1 - P_s(\Phi_n, W^{\otimes n}).$$

* milan.mosonyi@gmail.com

† ogawa@is.uec.ac.jp

The capacity $C(W)$ of the channel is the largest rate $\liminf_{n\to+\infty}\frac{1}{n}\log|\mathcal{C}_n|$ that can be achieved by a sequence of codes $\{\mathcal{C}_n\}_n$ such that $\lim_{n\to+\infty}P_e(W^{\otimes n},\mathcal{C}_n)=0$. According to the Holevo-Schumacher-Westmoreland (HSW) theorem [9, 18], $C(W)=\chi(W):=\max_{P\in\mathcal{P}(\mathcal{X})}\chi(W,P)$, where $\mathcal{P}(\mathcal{X})$ is the set of finitely supported probability distributions on $\mathcal{X}$, and $\chi(W,P):=\min_{\sigma\in\mathcal{S}(\mathcal{H})}D(\mathbb{W}(P)\|\hat{P}\otimes\sigma)$ is the Holevo quantity. Here, $\{|x\rangle\}_{x\in\mathcal{X}}$ is an orthonormal system in some auxiliary Hilbert space $\mathcal{H}_{\mathcal{X}}$, $\mathbb{W}(P):=\sum_{x\in\mathcal{X}}P(x)|x\rangle\langle x|\otimes W(x)$ is a classical-quantum state between the input and the output of the channel, and $\hat{P}:=\sum_x P(x)|x\rangle\langle x|$.

It is known that for any rate R above the capacity, the success probability goes to 0 with an exponential speed [16, 20]; this is called the strong converse property. The strong converse exponent $R_c(R,W)$ is the optimal rate of this exponential decay, i.e.,

$$R_c(R,W):=\inf\left\{r\;\middle|\;\exists\{\mathcal{C}_n\}_{n\in\mathbb{N}}\ \text{s.t.}\ \liminf_{n\to\infty}\frac{1}{n}\log|\mathcal{C}_n|\geq R\quad\text{and}\quad\liminf_{n\to\infty}\frac{1}{n}\log P_s(\mathcal{C}_n,W^{\otimes n})\geq -r\right\}.$$

Our main result is the following expression for the strong converse exponent:

Theorem II.1 *For every $R>0$,*

$$R_c(R,W)=\sup_{\alpha>1}\frac{\alpha-1}{\alpha}\left\{R-\sup_{P\in\mathcal{P}(\mathcal{X})}\chi^*_\alpha(W,P)\right\},\tag{1}$$

where $\chi^*_\alpha(W,P)$ is a generalized Holevo quantity defined below.

III. QUANTUM RÉNYI DIVERGENCES AND CAPACITIES

For non-commuting operators on a Hilbert space $\mathcal{H}$, various inequivalent generalizations of the Rényi divergences have been proposed. Here we will use the Rényi divergences built on the quantities

$$Q_\alpha(\rho\|\sigma):=\operatorname{Tr}\rho^\alpha\sigma^{1-\alpha},\qquad Q^*_\alpha(\rho\|\sigma):=\operatorname{Tr}\left(\rho^{\frac{1}{2}}\sigma^{\frac{1-\alpha}{\alpha}}\rho^{\frac{1}{2}}\right)^\alpha,\qquad Q^\flat_\alpha(\rho\|\sigma):=\operatorname{Tr}e^{\alpha\log\rho+(1-\alpha)\log\sigma},\tag{2}$$

defined for every positive definite ρ,σ, and every $\alpha\in(0,+\infty)\setminus\{1\}$. We can extend these definitions for semidefinite operators ρ and σ by $Q^{\{x\}}_\alpha(\rho\|\sigma):=\lim_{\varepsilon\searrow 0}Q^{\{x\}}_\alpha(\rho+\varepsilon I\|\sigma+\varepsilon I)$, where $\{x\}=\{\ \}$, $\{x\}=\{*\}$ or $\{x\}=\{\flat\}$. The corresponding *quantum Rényi divergences* are then defined as

$$D^{\{x\}}_\alpha(\rho\|\sigma):=\frac{1}{\alpha-1}\log Q^{\{x\}}_\alpha(\rho\|\sigma)-\frac{1}{\alpha-1}\log\operatorname{Tr}\rho.$$

The Araki-Lieb-Thirring inequality [2, 10] yields that $D^*_\alpha\leq D_\alpha$, and here we prove that

$$D_\alpha\leq D^\flat_\alpha,\quad \alpha\in[0,1),\qquad\text{and}\qquad D^\flat_\alpha\leq D^*_\alpha,\quad \alpha>1.$$

Note that D_α is the traditional notion of quantum Rényi divergence that features in the Hoeffding bound theorem [3, 6, 15]. D^*_α has been introduced recently in [13, 19], and has found operational interpretation in the strong converse part of various hypothesis testing problems [4, 8, 11]. $D^\flat_\alpha$ doesn't seem to have been much studied in information theory so far, although it is relevant in information geometry [1], and is also related to the free energy in some problems in statistical physics [17]. We prove the following variational representation, which is the main reason for the relevance of $D^\flat_\alpha$ for our purposes, and from which many important properties (e.g., convexity in σ) follow immediately:

Theorem III.1 *For every $\rho,\sigma\in\mathcal{L}(\mathcal{H})_+$ with non-orthogonal supports, and every $\alpha\in(0,+\infty)\setminus\{1\}$,*

$$D^\flat_\alpha(\rho\|\sigma)=\sup_{\tau\in\mathcal{S}(\mathcal{H})}\left\{D(\tau\|\sigma)-\frac{\alpha}{\alpha-1}D(\tau\|\rho)\right\}.\tag{3}$$

For a quantum channel $W : \mathcal{X} \to \mathcal{S}(\mathcal{H})$ and a probability distribution $P \in \mathcal{P}(\mathcal{X})$, we define the *generalized Holevo quantities*, corresponding to each Rényi α-divergence $D_\alpha^{\{x\}}$, as

$$\chi_{\alpha,1}^{\{x\}}(W,P) := \min_{\sigma\in\mathcal{S}(\mathcal{H})} D_\alpha^{\{x\}}(\mathbb{W}(P)\|\hat{P}\otimes\sigma) = \frac{1}{\alpha-1}\log\min_{\sigma\in\mathcal{S}(\mathcal{H})}\sum_x P(x) Q_\alpha^{\{x\}}(W(x)\|\sigma). \tag{4}$$

IV. MAIN STEPS OF THE PROOF

It is fairly easy, following Nagaoka's method [14], to show that

$$R_c(R,W) \geq \sup_{\alpha>1}\frac{\alpha-1}{\alpha}\left\{R - \sup_{P\in\mathcal{P}(\mathcal{X})}\chi_\alpha^*(W,P)\right\}. \tag{5}$$

Hence, the real challenge lies in proving the converse inequality. We first extend a result of Dueck and Körner [5] to classical-quantum channels, and show that

$$R_c(R,W) \leq F(P,R,W) := \inf_V\left\{D(\mathbb{V}(P)\|\mathbb{W}(P)) + \max\{0, R-\chi(V,P)\}\right\},$$

where the infimum is taken over all channels $V : \mathcal{X} \to \mathcal{S}(\mathcal{H})$. Next, we show that for any $P \in \mathcal{P}(\mathcal{X})$,

$$F(P,R,W) = \sup_{\alpha>1}\frac{\alpha-1}{\alpha}\left\{R - \inf_{\sigma\in\mathcal{S}(\mathcal{H})}\sum_{x\in\mathcal{X}} P(x) D_\alpha^\flat(W(x)\|\sigma)\right\},$$

where $\inf_{\sigma\in\mathcal{S}(\mathcal{H})}\sum_{x\in\mathcal{X}} P(x)D_\alpha^\flat(W(x)\|\sigma)$ is a variant of the generalized Holevo quantity $\chi_\alpha^\flat$. This can be changed to $\chi_\alpha^\flat$ after taking the infimum in P:

$$\inf_{P\in\mathcal{P}(\mathcal{X})} F(P,R,W) = \sup_{\alpha>1}\frac{\alpha-1}{\alpha}\left\{R - \inf_{\sigma\in\mathcal{S}(\mathcal{H})}\chi_\alpha^\flat(W,P)\right\}.$$

This, however, is a suboptimal bound. To obtain the desired bound, we apply the above argument to $\mathcal{E}_m W^{\otimes m}$, where $\mathcal{E}_m$ is the pinching of $W^{\otimes m}$ with respect to a universal symmetric state [7]. We show that

$$\chi_\alpha^\flat(\mathcal{E}_m W^{\otimes m}, P^{\otimes m}) \geq m\chi_\alpha^*(W,P) - 3\log v_m$$

for every $P \in \mathcal{P}(\mathcal{X})$, where $\lim_{m\to+\infty}\frac{1}{m}\log v_m = 0$. Taking the block size m to infinity then yields the converse to the inequality in (5).

---

[1] Shun-ichi Amari and Hiroshi Nagaoka. *Methods of information geometry*, volume 191 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 2000.

[2] H. Araki. On an inequality of Lieb and Thirring. *Letters in Mathematical Physics*, 19:167–170, 1990.

[3] K. M. R. Audenaert, M. Nussbaum, A. Szkola, and F. Verstraete. Asymptotic error rates in quantum hypothesis testing. *Communications in Mathematical Physics*, 279:251–283, 2008. arXiv:0708.4282.

[4] Tom Cooney, Milán Mosonyi, and Mark M. Wilde. Strong converse exponents for a quantum channel discrimination problem and quantum-feedback-assisted communication. August 2014. arXiv:1408.3373.

[5] G. Dueck and J. Körner. Reliability function of a discrete memoryless channel at rates above capacity. *IEEE Transactions on Information Theory*, 25:82–85, 1979.

[6] Masahito Hayashi. Error exponent in asymmetric quantum hypothesis testing and its application to classical-quantum channel coding. *Physical Review A*, 76(6):062301, December 2007. arXiv:quant-ph/0611013.

[7] Masahito Hayashi. Universal coding for classical-quantum channel. *Commun. Math. Phys.*, 289(3):1087–1098, May 2009.

[8] Masahito Hayashi and Marco Tomamichel. Composite hypothesis testing and an operational interpretation of the rényi mutual information. *arXiv:1408.6894*, 2014.

[9] Alexander S. Holevo. The capacity of the quantum channel with general signal states. *IEEE Transactions on Information Theory*, 44(1):269–273, January 1998.

[10] E.H. Lieb and W. Thirring. *Studies in mathematical physics*. University Press, Princeton, 1976.

[11] Milán Mosonyi and Tomohiro Ogawa. Quantum hypothesis testing and the operational interpretation of the quantum Rényi relative entropies. *arXiv:1309.3228; to appear in Commun. Math. Phys.*, September 2013.

[12] Milán Mosonyi and Tomohiro Ogawa. Strong converse exponent for classical-quantum channel coding. *arXiv:1409.3562*, September 2014.

[13] Martin Müller-Lennert, Frédéric Dupuis, Oleg Szehr, Serge Fehr, and Marco Tomamichel. On quantum Rényi entropies: A new generalization and some properties. *Journal of Mathematical Physics*, 54(12):122203, December 2013. arXiv:1306.3142.

[14] Hiroshi Nagaoka. Strong converse theorems in quantum information theory. *Proceedings of ERATO Workshop on Quantum Information Science*, page 33, 2001. Also appeared in Asymptotic Theory of Quantum Statistical Inference, ed. M. Hayashi, World Scientific, 2005.

[15] Hiroshi Nagaoka. The converse part of the theorem for quantum Hoeffding bound. November 2006. arXiv:quant-ph/0611289.

[16] Tomohiro Ogawa and Hiroshi Nagaoka. Strong converse to the quantum channel coding theorem. *IEEE Transactions on Information Theory*, 45(7):2486–2489, November 1999. arXiv:quant-ph/9808063.

[17] M. Ohya and D. Petz. *Quantum Entropy and its Use*. Springer, 1993.

[18] Benjamin Schumacher and Michael Westmoreland. Sending classical information via noisy quantum channels. *Physical Review A*, 56(1):131–138, July 1997.

[19] Mark M. Wilde, Andreas Winter, and Dong Yang. Strong converse for the classical capacity of entanglement-breaking and Hadamard channels via a sandwiched Rényi relative entropy. *Communications in Mathematical Physics*, 331(2):593–622, October 2014. arXiv:1306.1586.

[20] Andreas Winter. Coding theorem and strong converse for quantum channels. *IEEE Transactions on Information Theory*, 45(7):2481–2485, 1999.