Strong converse for the classical capacity of entanglement-breaking channels

Mark M. Wilde,[1] Andreas Winter,[2,3,4] and Dong Yang[3,5]

[1] Department of Physics and Astronomy, Center for Computation and Technology,
Louisiana State University, Baton Rouge, Louisiana 70803, USA
[2] ICREA – Institució Catalana de Recerca i Estudis Avançats,
Pg. Lluis Companys 23, ES-08010 Barcelona, Spain
[3] Física Teòrica: Informació i Fenomens Quàntics,
Universitat Autònoma de Barcelona, ES-08193 Bellaterra (Barcelona), Spain
[4] School of Mathematics, University of Bristol, Bristol BS8 1TW, United Kingdom
[5] Laboratory for Quantum Information, China Jiliang University, Hangzhou, Zhejiang 310018, China

Introduction.—One of the most fundamental tasks in quantum information theory is the transmission of classical data over many independent uses of a quantum channel, such that, for a fixed rate of communication, the error probability of the transmission decreases to zero in the limit of many channel uses. The maximum rate at which this is possible for a given channel is known as the classical capacity of the channel. Holevo, Schumacher, and Westmoreland (HSW) [11, 27] characterized the classical capacity of a quantum channel $\mathcal{N}$ in terms of the following formula:

$$\chi(\mathcal{N}) \equiv \max_{\{p_X(x),\rho_x\}} I(X;B)_\rho, \tag{1}$$

where $\{p_X(x), \rho_x\}$ is an ensemble of quantum states, $I(X;B)_\rho \equiv H(X)_\rho + H(B)_\rho - H(XB)_\rho$ is the quantum mutual information, and $H(\sigma) \equiv -\text{Tr}\{\sigma \log \sigma\}$ is the von Neumann entropy. In the above formula, the quantum mutual information $I(X;B)$ is computed with respect to the following classical-quantum state:

$$\rho_{XB} \equiv \sum_x p_X(x) |x\rangle\langle x|_X \otimes \mathcal{N}_{A\to B}(\rho_x), \tag{2}$$

for some orthonormal basis $\{|x\rangle\}$, and the notation $\mathcal{N}_{A\to B}$ indicates that the channel accepts an input on the system A and outputs to the system B.

For certain quantum channels, the HSW formula is equal to the classical capacity of the channel [2, 4, 6, 8, 15, 17, 19, 30]. These results follow because the Holevo formula was shown to be additive for these channels, in the sense that the following relation holds for these channels for any positive integer n:

$$\chi(\mathcal{N}^{\otimes n}) = n\,\chi(\mathcal{N}).$$

However, in general, if one cannot show that the HSW formula is additive for a given channel, then our best characterization of the classical capacity is given by a regularized formula:

$$C(\mathcal{N}) = \chi_{\text{reg}}(\mathcal{N}) \equiv \lim_{n\to\infty} \frac{1}{n}\chi(\mathcal{N}^{\otimes n}).$$

The work of Hastings [9] suggests that the regularized limit is necessary unless we are able to find some better characterization of the classical capacity, other than the above one given by HSW. Also, an important implication of the Hastings result, which demonstrates a strong separation between the classical and quantum theories of information, is that using entangled quantum codewords between multiple channel uses can enhance the classical capacity of certain quantum channels, whereas it is known that classically correlated codewords do not [11, 25, 27, 33].

Given the above results, one worthwhile direction is to refine our understanding of the classical capacity of channels for which the HSW formula is additive. Indeed, the achievability part of the HSW coding theorem states that as long as the rate of communication is below the classical capacity of the channel, then there exists a coding scheme such that the error probability of the scheme decreases exponentially fast to zero. The converse part of the capacity theorem makes use of the well known Holevo bound [10], and it states that if the rate of communication exceeds the capacity, then the error probability of any coding scheme is bounded away from zero in the limit of many channel uses.

Such a converse statement as given above might suggest that there is room for a trade-off between error probability and communication rate. That is, such a "weak" converse suggests that it might be possible for one to increase communication rates by allowing for an increased error probability. A *strong converse theorem* leaves no such room for a trade-off—it states that if the rate of communication exceeds the capacity, then the error probability of any coding scheme converges to one in the limit of many channel uses. Importantly, a strong converse theorem establishes the capacity of a channel as a very sharp dividing line between which communication rates are possible or impossible in the limit of many channel uses.

Strong converse theorems hold for all discrete memoryless classical channels [3, 35]. Wolfowitz employed a combinatorial approach based on the theory of types in order to prove the strong converse theorem [34, 35]. Arimoto used Rényi entropies to bound the success probability of any communication scheme [3]. Both the Wolfowitz and Arimoto approaches demonstrate that the success probability converges exponentially fast to zero if the rate of communication exceeds the capacity. Much later, Polyanskiy and Verdú generalized the Arimoto approach in a very useful way, by showing how to obtain a bound on the success probability in terms of any relative-entropy-like quantity satisfying several natural properties [26].

Less is known about strong converses for quantum channels. However, Winter [33] and Ogawa and Nagaoka [25] independently developed a strong converse theorem for channels with classical inputs and quantum outputs. For such channels, the HSW formula in (1) is equal to the classical capacity. The proof in Ref. [33] of the strong converse was based on a combinatorial approach in the spirit of Wolfowitz. Ogawa and Nagaoka's proof [25] was in the spirit of Arimoto.

After this initial work, Koenig and Wehner proved that the strong converse holds for the classical capacity of particular covariant quantum channels [20]. Their proof is in the spirit of Arimoto—they considered a Holevo-like quantity derived from the Rényi relative entropy and then showed that this quantity is additive for particular covariant channels. This reduction of the strong converse question to the additivity of an information quantity is similar to the approach of Arimoto, but the situation becomes more interesting for the case of quantum channels since entanglement between channel uses might lead to the quantity being non-additive.

Summary of results.—We prove that a strong converse theorem holds for the classical capacity of all entanglement-breaking channels [12, 14, 30]. Such channels can be modeled as the following process:

1. The channel performs a quantum measurement on the incoming state.

2. The channel then prepares a particular quantum state at the output depending on the result of the measurement.

The channels are said to be entanglement-breaking because if one applies a channel in this class to a share of an entangled state, then the resulting bipartite state is a separable state, having no entanglement. As important subclasses of the entanglement-breaking channels occur the classical-quantum channels mentioned above and quantum measurement channels, in which only the first step above occurs and the output is classical. Our result thus sharpens our understanding of the classical capacity for the entanglement-breaking channels, as motivated in the introduction.

We now give a brief sketch of the proof of our main result, full details can be found in Ref. [1].

1. First, we recall the argument of Sharma and Warsi [29] (which in turn is based on Ref. [26]), in which they showed that any relative-entropy-like quantity that satisfies some natural requirements gives a bound on the success probability of any coding scheme. Let $\mathcal{D}(\rho||\sigma)$ denote any generalized divergence that satisfies monotonicity (data processing), invariance under tensoring with the same quantum state, and reduces to a classical divergence when evaluated on commuting states. From this generalized divergence, one can define a Holevo-like quantity for a classical-quantum state of the form in (2):

$$\chi_{\mathcal{D}}(\mathcal{N}) \equiv \max_{\{p_X(x),\rho_x\}} I_{\mathcal{D}}(X;B), \tag{3}$$

where $I_{\mathcal{D}}(X;B) \equiv \min_{\sigma_B} \mathcal{D}(\rho_{XB}||\rho_X \otimes \sigma_B)$. Such a quantity itself satisfies a data processing inequality, which we can then exploit to obtain a bound on the success probability for any code that uses the channel n times at a fixed rate R.

2. We then introduce a "sandwiched" Rényi relative entropy (cf. Refs. [7, 23, 24, 32]), based on a parameter α and defined for quantum states ρ and σ as

$$\widetilde{D}_\alpha(\rho||\sigma) \equiv \frac{1}{\alpha-1}\log \mathrm{Tr}\Big\{\Big(\sigma^{\frac{1-\alpha}{2\alpha}}\ \rho\ \sigma^{\frac{1-\alpha}{2\alpha}}\Big)^{\alpha}\Big\}. \tag{4}$$

This definition of the Rényi relative entropy is different from the traditional one employed in quantum information theory (see Ref. [22], for example). Recall that the Rényi relative entropy is defined as

$$D_\alpha(\rho||\sigma) \equiv \frac{1}{\alpha-1}\log \mathrm{Tr}\{\rho^{\alpha}\ \sigma^{1-\alpha}\}.$$

However, it follows from the Lieb-Thirring trace inequality [21] that $\widetilde{D}_\alpha(\rho||\sigma) \leq D_\alpha(\rho||\sigma)$ for all $\alpha > 1$. Also, one can easily see that the two quantities are equal to each other whenever ρ and σ commute (when the states are effectively classical).

We prove that $\widetilde{D}_\alpha(\rho||\sigma)$ is monotone under quantum operations for all $\alpha \in (1, 2]$ and that it reduces to the von Neumann relative entropy in the limit as $\alpha \to 1$. These properties establish $\widetilde{D}_\alpha(\rho||\sigma)$ as a relevant information quantity to consider in quantum information theory. In particular, it will be useful for us in establishing the strong converse for entanglement-breaking channels. We then define a Holevo-like quantity $\widetilde{\chi}_\alpha(\mathcal{N})$ via the recipe in (3).

3. Combining the above two results, we establish the following upper bound on the success probability of any rate R classical communication scheme that uses a channel n times:

$$p_{\text{succ}} \leq 2^{-n\left(\frac{\alpha-1}{\alpha}\right)\left(R-\frac{1}{n}\widetilde{\chi}_\alpha(\mathcal{N}^{\otimes n})\right)}.$$

4. One can realize by inspecting the above formula that subadditivity of $\widetilde{\chi}_\alpha$ would be helpful in proving the strong converse, i.e., if

$$\widetilde{\chi}_\alpha(\mathcal{N}^{\otimes n}) \leq n\widetilde{\chi}_\alpha(\mathcal{N}). \tag{5}$$

We prove this for entanglement-breaking channels by first showing that $\widetilde{\chi}_\alpha$ is equal to an "α-information radius"[5, 22, 31],

$$\widetilde{\chi}_\alpha(\mathcal{N}) = \widetilde{K}_\alpha(\mathcal{N}) \equiv \min_\sigma \max_\rho \widetilde{D}_\alpha(\mathcal{N}(\rho)||\sigma), \tag{6}$$

building upon prior work in Refs. [20, 28].
This allows us to invoke King's result that the maximum output α-norm of an entanglement-breaking channel and any other channel is multiplicative [18] for $\alpha \geq 1$; subsequently Holevo observed that King's proof extends more generally to a completely positive entanglement-breaking map and any other completely positive map [13].
It is crucial for this step that conjugating a completely positive (entanglement-breaking) map by a positive operator does not take it out of this class—i.e., if $\mathcal{M}_{\text{EB}}$ is a completely positive (entanglement-breaking) map, then so is $(\mathcal{X} \circ \mathcal{M}_{\text{EB}})(\rho) = X\mathcal{M}_{\text{EB}}(\rho)X$, for a positive operator X.

5. The bound on the success probability for any coding scheme of rate R when using an entanglement-breaking channel then becomes

$$p_{\text{succ}} \leq 2^{-n\left(\frac{\alpha-1}{\alpha}\right)\left(R-\widetilde{\chi}_\alpha(\mathcal{N}_{\text{EB}})\right)}.$$

Finally, by a standard argument [25, 29], we can choose $\varepsilon > 0$ such that $\widetilde{\chi}_\alpha(\mathcal{N}_{\text{EB}}) < \chi(\mathcal{N}_{\text{EB}}) + \varepsilon$ for all $\alpha \geq 1$ in some neighborhood of 1, so that the success probability decays exponentially fast to zero with n if $R > \chi(\mathcal{N}_{\text{EB}})$. The strong converse theorem for all entanglement-breaking channels then follows.

Conclusion.—We have proven a strong converse theorem for the classical capacity of all entanglement-breaking channels, building on tighter bounds on the success probability in terms of a "sandwiched" Rényi relative entropy. Our approach also allows us to recover the earlier results of Koenig and Wehner [20]. This information measure should find other applications in quantum information theory, given that many other information measures can be obtained from a relative entropy.

An important open question going forward from here is to determine if a strong converse theorem holds for the classical capacity of quantum Hadamard channels [16, 19]. The most general definition of a Hadamard channel is one whose complementary channel is entanglement-breaking. We can already determine that our approach developed here does not seem to be useful for this class of channels. Namely, denoting the Hadamard (i.e., entry-wise) product by $*$, it is not necessarily the case that $X(C * \rho)X = D * \rho$ for some positive operator X and some other operator D, so conjugating a Hadamard channel by a positive operator X can take the channel outside of the Hadamard class.

Acknowledgements: After completing the work in this paper, we discovered that other authors had already defined [7, 32] and proved [23, 24] some of the properties of the sandwiched Rényi relative entropy. However, only the definition of the sandwiched Rényi relative entropy was publicly available at the time when we completed this work. We are grateful to Min-Hsiu Hsieh and Joe Renes for useful discussions. We thank Frédéric Dupuis, Marco Tomamichel, and Serge Fehr for passing along Refs. [7, 23, 24, 32].

[1] **Mark M. Wilde, Andreas Winter, and Dong Yang. Strong converse for the classical capacity of entanglement-breaking channels. June 2013. arXiv:1306.1586.**

[2] Grigori G. Amosov, Alexander S. Holevo, and R. F. Werner. On some additivity problems in quantum information theory. *Problems of Information Transmission*, 36(4):25, 2000.

[3] Suguru Arimoto. On the converse to the coding theorem for discrete memoryless channels. *IEEE Transactions on Information Theory*, 19:357–359, May 1973.

[4] Charles H. Bennett, David P. DiVincenzo, and John A. Smolin. Capacities of quantum erasure channels. *Physical Review Letters*, 78(16):3217–3220, April 1997.

[5] Imre Csiszár. Generalized cutoff rates and Rényi's information measures. *IEEE Transactions on Information Theory*, 41(1):26–34, January 1995.

[6] Nilanjana Datta, Alexander S. Holevo, and Yuri Suhov. Additivity for transpose depolarizing channels. *International Journal of Quantum Information*, 4(1):85–98, 2006.

[7] Serge Fehr. On the conditional Rényi entropy. Lecture at the Beyond IID Workshop at the University of Cambridge, January 2013.

[8] Motohisa Fukuda. Extending additivity from symmetric to asymmetric channels. *Journal of Physics A: Mathematical and General*, 38(45):L753–L758, 2005.

[9] Matthew B. Hastings. Superadditivity of communication capacity using entangled inputs. *Nature Physics*, 5:255–257, April 2009. arXiv:0809.3972.

[10] Alexander S. Holevo. Bounds for the quantity of information transmitted by a quantum communication channel. *Problems of Information Transmission*, 9:177–183, 1973.

[11] Alexander S. Holevo. The capacity of the quantum channel with general signal states. *IEEE Transactions on Information Theory*, 44(1):269–273, January 1998. arXiv:quant-ph/9611023.

[12] Alexander S. Holevo. Quantum coding theorems. *Russian Mathematical Surveys*, 53:1295–1331, 1999.

[13] Alexander S. Holevo. Multiplicativity of p-norms of completely positive maps and the additivity problem in quantum information theory. *Russian Mathematical Surveys*, 61(2):301–339, 2006.

[14] Michał Horodecki, Peter W. Shor, and Mary Beth Ruskai. Entanglement breaking channels. *Reviews in Mathematical Physics*, 15(6):629–641, 2003. arXiv:quant-ph/0302031.

[15] Christopher King. Additivity for unital qubit channels. *Journal of Mathematical Physics*, 43(10):4641–4653, 2002. arXiv:quant-ph/0103156.

[16] Christopher King. An application of the Lieb-Thirring inequality in quantum information theory. *Fourteenth International Congress on Mathematical Physics*, pages 486–490, 2003. arXiv:quant-ph/0412046.

[17] Christopher King. The capacity of the quantum depolarizing channel. *IEEE Transactions on Information Theory*, 49(1):221–229, January 2003. arXiv:quant-ph/0204172.

[18] Christopher King. Maximal p-norms of entanglement breaking channels. *Quantum Information and Computation*, 3(2):186–190, 2003. arXiv:quant-ph/0212057.

[19] Christopher King, Keiji Matsumoto, Michael Nathanson, and Mary Beth Ruskai. Properties of conjugate channels with applications to additivity and multiplicativity. *Markov Processes and Related Fields*, 13(2):391–423, 2007. J. T. Lewis memorial issue, arXiv:quant-ph/0509126.

[20] Robert Koenig and Stephanie Wehner. A strong converse for classical channel coding using entangled inputs. *Physical Review Letters*, 103:070504, August 2009. arXiv:0903.2838.

[21] Elliott H. Lieb and Walter Thirring. *Studies in mathematical physics*, chapter Inequalities for the moments of the eigenvalues of the Schroedinger Hamiltonian and their relation to Sobolev inequalities, pages 269–297. Princeton University Press, Princeton, 1976.

[22] Milán Mosonyi and Fumio Hiai. On the quantum Rényi relative entropies and related capacity formulas. *IEEE Transactions on Information Theory*, 57(4):2474–2487, April 2011. arXiv:0912.1286.

[23] Martin Müller-Lennert. Quantum relative Rényi entropies. Master's thesis, ETH Zurich, April 2013.

[24] Martin Müller-Lennert, Frédéric Dupuis, Serge Fehr, Oleg Szehr, and Marco Tomamichel. On quantum Rényi entropies: a new definition, some properties and several conjectures. June 2013. arXiv:1306.XXXX.

[25] 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.

[26] Yury Polyanskiy and Sergio Verdú. Arimoto channel coding converse and Rényi divergence. In *Proceedings of the 48th Annual Allerton Conference on Communication, Control, and Computation*, pages 1327–1333, September 2010.

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

[28] Benjamin Schumacher and Michael D. Westmoreland. Optimal signal ensembles. *Physical Review A*, 63:022308, January 2001.

[29] Naresh Sharma and Naqueeb Ahmad Warsi. On the strong converses for the quantum channel capacity theorems. June 2012. arXiv:1205.1712.

[30] Peter W. Shor. Additivity of the classical capacity of entanglement-breaking quantum channels. *Journal of Mathematical Physics*, 43(9):4334–4340, 2002. arXiv:quant-ph/0201149.

[31] Robin Sibson. Information radius. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 14(2):149160, 1969.

[32] Marco Tomamichel. Smooth entropies—a tutorial: With focus on applications in cryptography. Tutorial at QCRYPT 2012, slides available at `http://2012.qcrypt.net/docs/slides/Marco.pdf`, September 2012.

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

[34] Jacob Wolfowitz. The coding of messages subject to chance errors. *Illinois Journal of Mathematics*, 1:591–606, 1957.

[35] Jacob Wolfowitz. *Coding Theorems of Information Theory.* Prentice-Hall, Englewood Cliffs, NJ, USA, 1962.