Quantum Conditional Mutual Information, Reconstructed States, and State Redistribution

Fernando G. S. L. Brandão,[1] Aram W. Harrow,[2] Jonathan Oppenheim,[3] and Sergii Strelchuk[4]
[1]*Department of Computer Science, University College London, Gower St, London WC1E 6BT, UK and Quantum Architectures and Computation Group, Microsoft Research, Redmond, Washington 98052-6399, USA*
[2]*Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*
[3]*Department of Physics, University College London, Gower St, London WC1E 6BT, UK*
[4]*Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Wilberforce Road, Cambridge CB3 0WA, UK*
(Received 29 December 2014; published 29 July 2015)

We give two strengthenings of an inequality for the quantum conditional mutual information of a tripartite quantum state recently proved by Fawzi and Renner, connecting it with the ability to reconstruct the state from its bipartite reductions. Namely, we show that the conditional mutual information is an upper bound on the regularized relative entropy distance between the quantum state and its reconstructed version. It is also an upper bound for the measured relative entropy distance of the state to its reconstructed version. The main ingredient of the proof is the fact that the conditional mutual information is the optimal quantum communication rate in the task of state redistribution.

DOI: 10.1103/PhysRevLett.115.050501 PACS numbers: 03.67.Mn

Quantum information theory is the successful framework describing the transmission and storage of information. It not only generalized all of the classical information-theoretic results, but also developed a wealth of tools to analyze a number of scenarios beyond the reach of the latter, such as entanglement processing. One of the central quantities of the classical information theory that directly generalizes to quantum information is the conditional mutual information (CMI). For a tripartite state ρ_{BCR} it is defined as

$$I(C:R|B)_\rho := S(BC)_\rho + S(BR)_\rho - S(BCR)_\rho - S(B)_\rho, \qquad (1)$$

with $S(X)_\rho := -tr(\rho_X \log \rho_X)$ as the von Neumann entropy. It measures the correlations of subsystems C and R relative to subsystem B. The fact the classical CMI is non-negative is a simple consequence of the properties of the probability distributions; the same fact for the quantum CMI is equivalent to a deep result of quantum information theory—strong subadditivity of the von Neumann entropy [1]. Naturally, this led to a variety of applications in different areas, ranging from quantum information theory [2–4] to condensed matter physics [5–7].

In the classical case, for every tripartite probability distribution p_{XYZ},

$$I(X:Z|Y) = \min_{q \in \mathrm{MC}} S(p\|q), \qquad (2)$$

where $S(p\|q) := \sum_i p_i \log(p_i/q_i)$ is the relative entropy and the minimum is taken over the set MC of all distributions q such that $X - Y - Z$ form a Markov chain. Equivalently, the minimization in the right-hand side of Eq. (2) could be taken over $\Lambda \otimes \mathrm{id}_Z(p_{YZ})$, for reconstruction channels $\Lambda: Y \to YX$. In particular, $I(X:Z|Y) = 0$ if, and only if, $X - Y - Z$ form a Markov chain [which is equivalent to the existence of a channel $\Lambda: Y \to YX$ such that $p_{XYZ} = \Lambda \otimes \mathrm{id}_Z(p_{YZ})$].

The class of tripartite quantum states ρ_{BCR} satisfying $I(C:R|B)_\rho = 0$ has also been similarly characterized [8]: The B subsystem can be decomposed as $B = \bigoplus_k B_{L,k} \otimes B_{R,k}$ (with orthogonal vector spaces $B_{L,k} \otimes B_{R,k}$) and the state written as

$$\rho_{BCR} = \bigoplus_k p_k \rho_{CB_{L,k}} \otimes \rho_{B_{R,k}R} \qquad (3)$$

for a probability distribution $\{p_k\}$ and states $\rho_{CB_{L,k}} \in C \otimes B_{L,k}$ and $\rho_{B_{R,k}R} \in B_{R,k} \otimes R$. States of this form are called quantum Markov because in analogy to Markov chains, conditioned on the outcome of the measurement onto $\{B_{L,k} \otimes B_{R,k}\}$, the resulting state on C and R is product.

Paralleling the classical case, ρ_{BCR} is a quantum Markov state if, and only if, there exists a reconstruction channel $\Lambda: B \to BC$ such that $\Lambda \otimes \mathrm{id}_R(\rho_{BR}) = \rho_{BCR}$ [9]. Having generalized the definition of CMI, can we also retain the above equivalence, with the set of quantum Markov states taking the role of Markov chains? Surprisingly, it turns out that this is not the case, [10] and it seems not to be possible to connect states that are close to Markov states with states of small CMI in a meaningful way (see, however, [2,11]). Nonetheless, it might be possible to relate states with small CMI with those that can be approximately reconstructed from their bipartite reductions, i.e., such that $\Lambda \otimes \mathrm{id}_R(\rho_{BR}) \approx \rho_{BCR}$. Indeed, several conjectures appeared recently to this respect [5,12–14].

A recent breakthrough result from Fawzi and Renner gives the first such connection. They proved the following inequality [15]:

$$I(C:R|B)_\rho \geq \min_{\Lambda:B\to BC} S_{1/2}[\rho_{BCR}\|\Lambda\otimes \mathrm{id}_R(\rho_{BR})] \tag{4}$$

with $S_{1/2}(\rho\|\sigma) \coloneqq -2\log F(\rho,\sigma)$ the order-1/2 Rényi relative entropy, where $F(\rho,\sigma) = \mathrm{tr}[(\sigma^{1/2}\rho\sigma^{1/2})^{1/2}]$ is the fidelity [16]. It implies that if the CMI of ρ_{BCR} is small, there exists a reconstructing channel $\Lambda: B\to BC$ such that $\Lambda\otimes\mathrm{id}_R(\rho_{BR})$ has high fidelity with ρ_{BCR}.

In this Letter, we prove a strengthened version of the Fawzi-Renner inequality. We also give a simpler proof of the inequality based on the task of state redistribution [4], which gives an operational interpretation to the CMI.

Result.—Let $S(\rho\|\sigma)\coloneqq \mathrm{tr}[\rho(\log\rho-\log\sigma)]$ be the quantum relative entropy of ρ and σ. Define the measured relative entropy as

$$\mathbb{M}S(\rho\|\sigma) = \max_{M\in\mathcal{M}} S[M(\rho)\|M(\sigma)], \tag{5}$$

where $\mathcal{M}$ is the set of all quantum-classical channels $M(\rho)=\sum_k \mathrm{tr}(M_k\rho)|k\rangle\langle k|$, with $\{M_k\}$ a positive operator-valued measure and $\{|k\rangle\}$ an orthonormal basis.

The main result of this Letter is the following.

Theorem 1: For every state ρ_{BCR},

$$I(C:R|B)_\rho, \tag{6a}$$

$$\geq \lim_{n\to\infty}\min_{\Lambda_n:B^n\to B^nC^n}\frac{1}{n}S[\rho_{BCR}^{\otimes n}\|\Lambda_n\otimes\mathrm{id}_{R^n}(\rho_{BR}^{\otimes n})], \tag{6b}$$

$$\geq \min_{\Lambda:B\to BC}\mathbb{M}S[\rho_{BCR}\|\Lambda\otimes\mathrm{id}_R(\rho_{BR})], \tag{6c}$$

$$\geq \min_{\Lambda:B\to BC}S_{1/2}[\rho_{BCR}\|\Lambda\otimes\mathrm{id}_R(\rho_{BR})]. \tag{6d}$$

Equation (6d) is the Fawzi-Renner inequality [Eq. (4)] and follows from Eq. (6c) using the bound $S(\pi\|\sigma)\geq S_{1/2}(\pi\|\sigma)$ [17] and the fact that $\min_{M\in\mathcal{M}}F(M(\pi),M(\sigma)) = F(\pi,\sigma)$ [18]. Equation (6c) also generalizes one side of Eq. (2) to quantum states, implying that it is optimal at least for classical states ρ.

Our lower bound provides a substantial improvement over the original Fawzi-Renner bound even for classical states. To see this, consider the classically correlated state $\rho_{CBR}=\rho_{CR}\otimes\mathbb{1}_B/d_B$ with $d\coloneqq d_C=d_R$ and $\rho_{CR}=(1-\epsilon)|00\rangle\langle 00|_{CR}+(\epsilon/d-1)\sum_{k=1}^{d-1}|kk\rangle\langle kk|_{CR}$. Then Eq. (6c) becomes $\mathbb{M}S(\rho_{BCR}\|\sigma_{BC}\otimes\rho_R)$, where σ_{BC} depends on the channel Λ that minimizes Eq. (6c). The measured relative entropy is equal to the ordinary classical relative entropy between the distribution p_Bp_{CR} (generated from ρ_{BCR}) and the product distribution $q_{BC}p_R$ (generated from $\sigma_{BC}\otimes\rho_R$) optimized over all quantum-classical channels. Observing that p_{CR} is maximally correlated, whereas q_Cp_B is the product distribution irrespective of Λ, Eq. (6c) equals to $I(C:R)\approx\epsilon\log(d-1)$. The corresponding Fawzi-Renner bound (6d) becomes $-\log F(\rho_{CR},\rho_C\otimes\rho_R)\leq -\log(1-\epsilon)\approx\epsilon$. Thus, the lower bound (6c) is optimal for classical states.

Another application of our result is the well-known problem of classification of the short-range entangled states studied by Kitaev [19]. Defining such a class of states is nontrivial, and one of the natural properties to be required is the ability generate them locally: There must exist a $O(1)$ quantum circuit that generates the designated state from a product state. In particular, one sees that states with low CMI can be generated from the product states according to the Fawzi-Renner bound. Our result improves the lower bound when we quantify the distance between the states using measured relative entropy.

Li and Winter conjectured in [11] that Eq. (6c) could be strengthened to have the relative entropy in the right-hand side (instead of the measured relative entropy). We leave this as an open question, but we note that Eq. (6b) shows that an asymptotic version of the conjectured inequality does hold true.

Proof of Theorem 1: The main tool in the proof will be the state redistribution protocol of Devetak and Yard [4,20,21], which gives an operational meaning for the CMI as twice the optimal quantum communication cost of the protocol. Consider the state $|\psi\rangle_{ABCR}^{\otimes n}$ shared by two parties (Alice and Bob) and the environment (or reference system). Alice has A^nC^n (where we denote n copies of A by A^n and likewise for C, B, and R), Bob has B^n, and R^n is the reference system. In state redistribution, Alice wants to redistribute the C^n subsystem to Bob using preshared entanglement and quantum communication.

It was shown in [4,21] that using preshared entanglement Alice can send the C^n part of her state to Bob, transmitting approximately $(n/2)I(C:R|B)$ qubits in the limit of a large number of copies n. More precisely:

Lemma 2 (State redistribution protocol [4,21]): For every $|\psi\rangle_{ABCR}$ there exist completely positive trace-preserving encoding maps $\mathcal{E}_n: A^nC^nX_n\to A^nG_n$ and decoding maps $\mathcal{D}_n: B^nG_nY_n\to B^nC^n$ such that

$$\lim_{n\to\infty}\|\mathcal{D}_n\circ\mathcal{E}_n(|\psi\rangle\langle\psi|_{ABCR}^{\otimes n}\otimes\Phi_{X_nY_n}) - |\psi\rangle\langle\psi|_{ABCR}^{\otimes n}\|_1 = 0 \tag{7}$$

and

$$\lim_{n\to\infty}\frac{\log\ \dim(G_n)}{n} = \frac{1}{2}I(C:R|B)_\rho, \tag{8}$$

where $\rho_{BCR}\coloneqq\mathrm{tr}_A(|\psi\rangle\langle\psi|_{ABCR})$ and $\Phi_{X_nY_n}$ is a maximally entangled state shared by Alice (who has X_n) and Bob (who has Y_n), and $\|.\|_1$ denotes the trace norm.

We split the proof of Theorem 1 into the proof of Proposition 3 and Eq. (17), below.

Proposition 3 follows from the state redistribution protocol outlined above. The main idea is the following:

Suppose that in the state redistribution protocol Bob does not receive any quantum communication from Alice, but instead he "mocks" the communication (locally preparing the maximally mixed state in G_n) and applies the decoding map $\mathcal{D}_n$. It will follow that even though the output state might be very far from the target one, the relative entropy per copy of the output state and the original one cannot be larger than twice the amount of communication of the protocol (which is given by the conditional mutual information).

Proposition 3: For every state ρ_{BCR},

$$I(C:R|B)_\rho \geq \lim_{n\to\infty} \min_{\Lambda: B^n\to B^nC^n} \frac{1}{n} S[\rho_{BCR}^{\otimes n} \| \Lambda \otimes \mathrm{id}_{R^n}(\rho_{BR}^{\otimes n})]. \tag{9}$$

Proof. Let $|\psi\rangle_{ABCR}$ be a purification of ρ_{BCR}. Consider the state redistribution protocol for sending C from Alice (who has AC) to Bob (who has B). Let $\phi_{G_nY_nA^nB^nR^n} \coloneqq \mathcal{E}_n \otimes \mathrm{id}_{B^nR^nY_n}(|\psi\rangle\langle\psi|_{ABCR}^{\otimes n} \otimes \Phi_{X_nY_n})$ be the state after the encoding operation.

Using the operator inequality $\pi_{MN} \leq \dim(M) I_M \otimes \pi_N$, valid for every state π_{MN}, we find

$$\phi_{G_nY_nB^nR^n} \leq \dim(G_n)^2 \tau_{G_n} \otimes \tau_{Y_n} \otimes \rho_{BR}^{\otimes n} \tag{10}$$

with τ_{Y_n}, τ_{G_n} as the maximally mixed state on Y_n and G_n, respectively. We used $\phi_{Y_nB^nR^n} = \tau_{Y_n} \otimes \rho_{BR}^{\otimes n}$, which holds true since $\mathcal{E}_n$ only acts nontrivially on $A^nC^nX_n$.

Let $\mathcal{D}_n: G_nY_nB^n \to B^nC^n$ be the decoding operation of Bob in state redistribution (see Lemma 2) and define $\tilde{\mathcal{D}}_n \coloneqq (1-2^{-n})\mathcal{D}_n + 2^{-n}\Lambda_{\mathrm{dep}}$, with Λ_{dep} the depolarizing channel mapping all states to the maximally mixed. Since $\tilde{\mathcal{D}}_n$ is completely positive, using Eq. (10) we get

$$\begin{aligned}(\tilde{\mathcal{D}}_n \otimes \mathrm{id}_{R^n})(\tau_{G_n} \otimes \tau_{Y_n} \otimes \rho_{BR}^{\otimes n})\\ \geq \dim(G_n)^{-2}(\tilde{\mathcal{D}}_n \otimes \mathrm{id}_{R^n})(\phi_{G_nY_nB^nR^n}).\end{aligned} \tag{11}$$

From the operator monotonicity of the log (see Lemma 1 in the Supplemental Material [22]),

$$\begin{aligned}&S[\rho_{BCR}^{\otimes n} \| (\tilde{\mathcal{D}}_n \otimes \mathrm{id}_{R^n})(\tau_{G_n} \otimes \tau_{Y_n} \otimes \rho_{BR}^{\otimes n})]\\ &\quad\leq S[\rho_{BCR}^{\otimes n} \| (\tilde{\mathcal{D}}_n \otimes \mathrm{id}_{R^n})(\phi_{G_nY_nB^nR^n})] + 2\log[\dim(G_n)].\end{aligned} \tag{12}$$

Equation (7) gives

$$\lim_{n\to\infty} \| \rho_{BCR}^{\otimes n} - (\tilde{\mathcal{D}}_n \otimes \mathrm{id}_{R^n})(\phi_{G_nY_nB^nR^n}) \|_1 = 0. \tag{13}$$

Because $(\tilde{\mathcal{D}}_n \otimes \mathrm{id}_{R^n})(\phi_{G_nY_nB^nR^n}) = (1-2^{-n})(\mathcal{D}_n \otimes \mathrm{id}_{R^n})(\phi_{G_nY_nB^nR^n}) + 2^{-n}\tau_{BC}^{\otimes n} \otimes \rho_R^{\otimes n}$ (with τ_{BC} the maximally mixed state on BC), Lemma 2 in the Supplemental Material [22] gives

$$\lim_{n\to\infty} \frac{1}{n} S[\rho_{BCR}^{\otimes n} \| (\tilde{\mathcal{D}}_n \otimes \mathrm{id}_{R^n})(\phi_{G_nY_nB^nR^n})] = 0, \tag{14}$$

and so

$$\begin{aligned}I(C:R|B)_\rho &= 2\lim_{n\to\infty} \frac{\log[\dim(G_n)]}{n}\\ &\geq \lim_{n\to\infty} \min_{\Lambda_n: B^n\to B^nC^n} \frac{1}{n} S[\rho_{BCR}^{\otimes n} \| (\Lambda_n \otimes \mathrm{id}_{R^n})(\rho_{BR}^{\otimes n})].\end{aligned} \tag{15}$$

■

Even though we do not know whether

$$\begin{aligned}&\lim_{n\to\infty} \min_{\Lambda: B^n\to B^nC^n} \frac{1}{n} S[\rho_{BCR}^{\otimes n} \| \Lambda \otimes \mathrm{id}_{R^n}(\rho_{BR}^{\otimes n})]\\ &\quad\overset{?}{\geq} \min_{\Lambda: B\to BC} S[\rho_{BCR} \| \Lambda \otimes \mathrm{id}_R(\rho_{BR})],\end{aligned} \tag{16}$$

it turns out that a similar inequality holds true if we replace the relative entropy by its measured variant (see Sec. B in the Supplemental Material [22]): For every state ρ_{BCR} one has

$$\begin{aligned}&\lim_{n\to\infty} \min_{\Lambda: B^n\to B^nC^n} \frac{1}{n} S[\rho_{BCR}^{\otimes n} \| \Lambda \otimes \mathrm{id}_{R^n}(\rho_{BR}^{\otimes n})]\\ &\quad\geq \min_{\Lambda: B\to BC} \mathbb{M}S[\rho_{BCR} \| \Lambda \otimes \mathrm{id}_R(\rho_{BR})].\end{aligned} \tag{17}$$

Discussion and open problems.—The main result of this Letter, on one hand, and Theorem 4 of Ref. [10], on the other hand, give

$$\begin{aligned}\min_{\sigma\in \mathrm{QMS}} S(\rho_{BCR} \| \sigma_{BCR}) &\geq I(C:R|B)\\ &\geq \min_{\Lambda: B\to BR} \mathbb{M}[\rho_{BCR} \| \Lambda \otimes \mathrm{id}_R(\rho_{BR})],\end{aligned} \tag{18}$$

with the set of quantum Markov states (QMS) given by Eq. (3). For probability distributions the lower and upper bounds in Eq. (18) coincide, giving Eq. (2). However, in the quantum case, the two can be very far from each other.

An interesting question is whether we can also have equality in the quantum case when minimizing over the set of reconstructed states. In particular, we can ask whether Eq. (9) holds with equality. It turns out that this is false and can be disproved using pure states of dimension $2\times 2\times 2$ and the transpose channel, defined for a tripartite state ρ_{BCR} as

$$T(\pi) \coloneqq \sqrt{\rho_{BC}}(\rho_B^{-1/2}\pi\rho_B^{-1/2} \otimes \mathrm{id}_C)\sqrt{\rho_{BC}}. \tag{19}$$

In Fig. 1 we plot the CMI against the reconstructed relative entropy using the transpose channel, i.e., $S[\rho_{BCR} \| T_B \otimes \mathrm{id}_R(\rho_{BR})]$, for 10 000 randomly chosen pure states of dimension $2\times 2\times 2$. We see that for roughly 73%

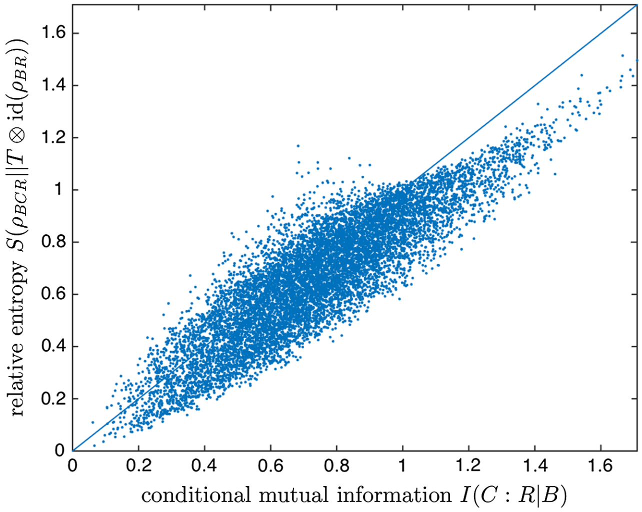

FIG. 1 (color online). Counterexamples for the case of equality in Eq. (9): Conditional mutual information against the reconstructed relative entropy using the transpose channel. The sample consists of 10 000 random pure states of dimension $2\times2\times2$.

of the points, the relative entropy is strictly smaller than the CMI when using the transpose channel. Since any particular reconstruction map also puts an upper bound on the minimum relative entropy, Eq. (9) must sometimes be a strict inequality. Similar numerical results were found in an unpublished early version of [11].

In the proof of Theorem 1 we were not able to give an explicit optimal reconstruction map. In the context of approximate recovery for pure states, the transpose channel is optimal up to a square factor [23] (using the fidelity as a figure of merit). We could ask whether the same holds for mixed states.

Another interesting open problem is whether we can improve the lower bound in Eq. (18) to have the relative entropy, instead of the measured relative entropy. Proposition 3 and Lemma 5 in the Supplemental Material [22] show that the result would follow from the following conjectured inequality: Given a state ρ, a convex closed set of states $\mathcal{S}$, and a measure μ with support only on $\mathcal{S}$,

$$\lim_{n\to\infty}\inf\frac{1}{n}S\left(\rho^{\otimes n}\Big\|\int\mu(d\sigma)\sigma^{\otimes n}\right)\overset{?}{\geq}\min_{\sigma\in\mathcal{S}}S(\rho\|\sigma). \qquad (20)$$

The case when $\rho_{BR}=\rho_B\otimes\rho_R$ was recently proven in [24]. We can also easily prove the inequality classically using hypothesis testing, which is universal for the alternative hypothesis. However, since there is no quantum hypothesis test universal for the alternative hypothesis [25] for general sets $\mathcal{S}$, we leave the inequality in the quantum case as an open problem for future work.

F. G. S. L. B. and J. O. thank EPSRC for financial support. A. W. H. was funded by NSF Grants No. CCF-1111382 and CCF-1452616, ARO Contract No. W911NF-12-1-0486, and a grant from the Leverhulme Trust. S. S. acknowledges the support of Sidney Sussex College. We thank Omar Fawzi, Ke Li, Marco Tomamichel, Mark Wilde, and Lin Zhang for helpful comments and the Newton Institute for their hospitality while some of this research was conducted.

[1] E. Lieb and M. Ruskai, Proof of strong subadditivity of quantum-mechanical entropy, J. Math. Phys. (N.Y.) **14**, 1938 (1973).

[2] F. G. S. L. Brandão, M. Christandl, and J. Yard, Faithful squashed entanglement, Commun. Math. Phys. **306**, 805 (2011).

[3] M. Christandl and A. Winter, 'Squashed entanglement'—an additive entanglement measure, J. Math. Phys. (N.Y.) **45**, 829 (2004).

[4] I. Devetak and J. Yard, Exact Cost of Redistributing Multipartite Quantum States, Phys. Rev. Lett. **100**, 230501 (2008).

[5] I. H. Kim, Conditional independence in quantum many-body systems, Ph.D. thesis, California Institute of Technology, 2013.

[6] A. Kitaev (private communication).

[7] D. Poulin and M. Hastings, Markov Entropy Decomposition: A Variational Dual for Quantum Belief Propagation, Phys. Rev. Lett. **106**, 080403 (2011).

[8] P. Hayden, R. Jozsa, D. Petz, and A. Winter, Structure of states which satisfy strong subadditivity of quantum entropy with equality, Commun. Math. Phys. **246**, 359 (2004).

[9] Reference [8] shows that Markov states have perfect reconstruction maps; see Eq. (19) for an explicit formula. The converse is straightforward. If $\Lambda\otimes\mathrm{id}_R(\rho_{BR})=\rho_{BCR}$, then $\Lambda(\rho_B)=\rho_{BC}$. Using the recovery condition, and the monotonicity of relative entropy after applying first $(\Lambda\otimes\mathrm{id}_R)$ and then tr_C, we obtain $S(\rho_{BCR}\|\rho_R\otimes\rho_{BC})=S[(\Lambda\otimes\mathrm{id}_R)\rho_{BR}\|(\Lambda\otimes\mathrm{id}_R)(\rho_B\otimes\rho_R)]\leq S(\rho_{BR}\|\rho_B\otimes\rho_R)\leq S(\rho_{BCR}\|\rho_{BC}\otimes\rho_R)$. It then follows that $I(C:R|B)=S(\rho_{BCR}\|\rho_{BC}\otimes\rho_R)-S(\rho_{BR}\|\rho_B\otimes\rho_R)=0$.

[10] B. Ibinson, N. Linden, and A. Winter, Robustness of quantum Markov chains, Commun. Math. Phys. **277**, 289 (2008).

[11] K. Li and A. Winter, Relative entropy and squashed entanglement, Commun. Math. Phys. **326**, 63 (2014).

[12] J. Oppenheim and S. Strelchuk, Robust quantum Markov states (unpublished).

[13] M. Berta, K. P. Seshadreesan, and M. M. Wilde, Rényi generalizations of the conditional quantum mutual information, J. Math. Phys. (N.Y.) **56**, 022205 (2015).

[14] L. Zhang, Conditional mutual information and commutator, Int. J. Theor. Phys. **52**, 2112 (2013).

[15] O. Fawzi and R. Renner, Quantum conditional mutual information and approximate Markov chains, arXiv:1410.0664.

[16] R. Jozsa, Fidelity for mixed quantum states, J. Mod. Opt. **41**, 2315 (1994).

[17] M. Müller-Lennert, F. Dupuis, O. Szehr, S. Fehr, and M. Tomamichel, On quantum Rényi entropies: A new generalization and some properties, J. Math. Phys. (N.Y.) **54**, 122203 (2013).

[18] C. A. Fuchs and C. M. Caves, Ensemble-Dependent Bounds for Accessible Information in Quantum Mechanics, Phys. Rev. Lett. **73**, 3047 (1994).

[19] A. Kitaev, Topological Phases of Matter Program, Caltech, SCGP, http://scgp.stonybrook.edu/archives/7874.

[20] J. Oppenheim, State redistribution as merging: Introducing the coherent relay, arXiv:0805.1065.

[21] J. Yard and I. Devetak, Optimal quantum source coding with quantum information at the encoder and decoder, IEEE Trans. Inf. Theory **55**, 5339 (2009).

[22] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.115.050501 for auxiliary lemmas and the proof of the measured entropy lower bound.

[23] H. Barnum and E. Knill, Reversing quantum dynamics with near-optimal quantum and classical fidelity, J. Math. Phys. (N.Y.) **43**, 2097 (2002).

[24] M. Hayashi and M. Tomamichel, Correlation detection and an operational interpretation of the Rényi mutual information, arXiv:1408.6894.

[25] I. Bjelaković, J.-D. Deuschel, T. Krüger, R. Seiler, R. Siegmund-Schultze, and A. Szkoła, A quantum version of Sanov's theorem, Commun. Math. Phys. **260**, 659 (2005).