Private Capacity of Quantum Channels is Not Additive

Ke Li,[1,*] Andreas Winter,[2,3,†] XuBo Zou,[1,‡] and GuangCan Guo[1,§]

[1]*Key Laboratory of Quantum Information, University of Science and Technology of China, Chinese Academy of Sciences, Hefei, Anhui 230026, China*

[2]*Department of Mathematics, University of Bristol, University Walk, Bristol BS8 1TW, United Kingdom*

[3]*Centre for Quantum Technologies, National University of Singapore, 2 Science Drive 3, Singapore 117542*

(Received 1 May 2009; published 15 September 2009)

Recently there has been considerable activity on the subject of the additivity of various quantum channel capacities. Here, we construct a family of channels with a sharply bounded classical and, hence, private capacity. On the other hand, their quantum capacity when combined with a zero private (and zero quantum) capacity erasure channel becomes larger than the previous classical capacity. As a consequence, we can conclude for the first time that the classical private capacity is nonadditive. In fact, in our construction even the quantum capacity of the tensor product of two channels can be greater than the sum of their individual classical private capacities. We show that this violation occurs quite generically: every channel can be embedded into our construction, and a violation occurs whenever the given channel has a larger entanglement-assisted quantum capacity than (unassisted) classical capacity.

DOI: 10.1103/PhysRevLett.103.120501 PACS numbers: 03.67.Hk, 03.67.Dd

Information theory, established by Claude Shannon in the 1940s as a ‘‘Mathematical Theory of Communication’’ [1], is the theoretical foundation of today’s communication technologies. The main problem it is concerned with is how much information can be transmitted down a noisy channel asymptotically, i.e., the capacity of the channel. Shannon provided a beautifully simple formula for the capacity of a discrete memoryless channel, which only involves an entropic expression of a single channel use. Subsequent research revealed that this simple capacity formula fully characterizes the information-carrying capability of a channel under a large range of circumstances [2], serving as a very robust measure. E.g., the ability of two channels together to transmit information is quantified by the sum of their individual capacities.

Our world however is not the classical one of Shannon’s noisy channels, but is at a basic level described by quantum theory. To understand the ultimate limit the laws of physics impose on our ability to communicate, the underlying quantum behavior of the channels should be considered. A quantum channel $\mathcal{N}$ is mathematically described by an isometric map V from the input Hilbert space A to the combined Hilbert space of the output B and the so-called environment system E. Then the channel and its natural complement $\tilde{\mathcal{N}}$ act as $\mathcal{N}(\rho) = \mathrm{Tr}_E V\rho V^\dagger$ and $\tilde{\mathcal{N}}(\rho) = \mathrm{Tr}_B V\rho V^\dagger$. It can in general not only convey classical messages, but also quantum data, i.e., a Hilbert space of quantum states. It can also carry classical private information, inaccessible to the environment, enabling the classically impossible, provably unconditionally secure key distribution [3]. Naturally, deriving capacity formulas of a quantum channel for transmitting various kinds of information is a central task of quantum information theory.

The classical capacity, $C(\mathcal{N})$, is the maximal rate of classical information that the quantum channel $\mathcal{N}$ can asymptotically transmit with vanishing errors. In contrast to the classical capacity, the definition of classical private capacity $P(\mathcal{N})$ further requires that the classical information conveyed is secret from the environment. Finally, the quantum capacity $Q(\mathcal{N})$ quantifies how large a Hilbert space of states the channel $\mathcal{N}$ can transmit asymptotically and with the error approaching zero. Operationally, quantum information transmission implies classical transmission, which in turn implies classical communication. I.e.,

$$C(\mathcal{N}) \geq P(\mathcal{N}) \geq Q(\mathcal{N}). \tag{1}$$

Despite considerable progress, tractable formulae for the quantum, private, and unrestricted classical capacities are still out of reach. The HSW theorem [4], Devetak [5], and the LSD theorem [5–7] give the classical, private, and quantum capacities, respectively, as the regularization of single-letter quantities:

$$\chi(\mathcal{N}) \leq C(\mathcal{N}) = \lim_{n\to\infty} \frac{1}{n} \chi(\mathcal{N}^{\otimes n}), \tag{2}$$

$$P^{(1)}(\mathcal{N}) \leq P(\mathcal{N}) = \lim_{n\to\infty} \frac{1}{n} P^{(1)}(\mathcal{N}^{\otimes n}), \tag{3}$$

$$Q^{(1)}(\mathcal{N}) \leq Q(\mathcal{N}) = \lim_{n\to\infty} \frac{1}{n} Q^{(1)}(\mathcal{N}^{\otimes n}). \tag{4}$$

All three single-letter quantities are obtained via finite optimizations of entropic expressions: the Holevo capacity $\chi(\mathcal{N})$ is the maximum over all ensembles $\{p_i, \rho_i\}$ of states on A of the Holevo information

$$\chi_{\{p_i,\rho_i\}}(\mathcal{N}) = H\left(\left(\mathcal{N}\left(\sum_i p_i\rho_i\right)\right)\right) - \sum_i p_i H(\mathcal{N}(\rho_i)), \quad (5)$$

where $H(\rho) = -\mathrm{Tr}\rho\log\rho$ is the von Neumann entropy (log is always the binary logarithm). Similarly, $P^{(1)}(\mathcal{N}) = \max_{\{p_i,\rho_i\}}(\chi_{\{p_i,\rho_i\}}(\mathcal{N}) - \chi_{\{p_i,\rho_i\}}(\tilde{\mathcal{N}}))$, and $Q^{(1)}(\mathcal{N}) = \max_\rho I_c(\rho, \mathcal{N})$, with the coherent information [8]

$$I_c(\rho, \mathcal{N}) = H(\mathcal{N}(\rho)) - H(\tilde{\mathcal{N}}(\rho)). \quad (6)$$

Neither $\chi(\mathcal{N})$, nor $Q^{(1)}(\mathcal{N})$, nor $P^{(1)}(\mathcal{N})$ are additive; in fact, $\chi \neq C$ [9], $P^{(1)} \neq P$ [10], $Q^{(1)} \neq Q$ [11]. However, for certain classes of channels it is known that $C(\mathcal{N}) = \chi(\mathcal{N})$ [12–14], and for other classes $P(\mathcal{N}) = P^{(1)}(\mathcal{N})$, $Q(\mathcal{N}) = Q^{(1)}(\mathcal{N})$ [15].

As measures of a channel's information transmitting capability, the above three capacity quantities might be expected to be robust, i.e., just like Shannon's capacity for classical channels, to be applicable under a large range of settings. While this is no longer true when various auxiliary resources (e.g. free entanglement or classical communications) are available [16], another weird feature of the quantum capacity was discovered recently. Smith and Yard [17] show that, as a function of channels, $Q(\mathcal{N})$ is not additive. Specifically, for the two channels $\mathcal{N}_1$ and $\mathcal{N}_2$ with $\mathcal{N}_1$ satisfying $Q(\mathcal{N}_1) = 0$ and $P(\mathcal{N}_1) > 0$, and $\mathcal{N}_2$ the (zero quantum and zero private capacity) 50% erasure channel, $Q(\mathcal{N}_1 \otimes \mathcal{N}_2) \geq \frac{1}{2}P(\mathcal{N}_1) > 0$. One might attribute this superactivation of quantum capacity to the ability to transmit privacy [18], recalling the close relationship between $Q(\mathcal{N})$ and $P(\mathcal{N})$ [19]. But surprisingly again, Smith and Smolin [20] have found two channels such that either they have large joint quantum capacity but negligible individual private classical capacities, or one of them exhibits a large nonadditivity of χ.

In this Letter, we present quantum channels $\mathcal{T}^k_{\mathcal{N}}$ for given channel $\mathcal{N}$ with finite environment dimension (this includes all channels with finite dimensional input and output), and integer k; it inherits input and output from $\mathcal{N}$, but has also auxiliary registers. We can show that $C(\mathcal{N}) \leq C(\mathcal{T}^k_{\mathcal{N}}) \leq C(\mathcal{N}) + \delta(k)$, where $\delta(k)$ goes to zero as $k \to \infty$. Regarding the capability of the channel $\mathcal{T}^k_{\mathcal{N}}$, together with a 50% erasure channel $\mathcal{A}$, for quantum communication, we find that the quantum capacity of the combined channel $\mathcal{T}^k_{\mathcal{N}} \otimes \mathcal{A}$ is lower bounded by $Q_E(\mathcal{N})$, the entanglement-assisted quantum capacity of $\mathcal{N}$ [21]. So, for channels $\mathcal{N}$ such that $Q_E(\mathcal{N}) > C(\mathcal{N})$, $\mathcal{T}^k_{\mathcal{N}}$—when combined with the above erasure channel—can transmit more quantum information than its classical capacity $C(\mathcal{T}^k_{\mathcal{N}})$. Referring to Eq. (1), we conclusively prove that the classical private capacity, in fact even the quantum capacity, of two channels can be greater than the sum of their individual classical private capacities. Our findings not only demonstrate that the classical private capacity of a quantum channel is generally not additive, but also yield another counterexample to the additivity of quantum capacity, of which the underlying reasoning is different from that of Smith and Yard's [17].

The channel construction.—In the Stinespring representation $\mathcal{N}(\rho) = \mathrm{Tr}_E V\rho V^\dagger$, the partial trace embodies all the noise of the channel as loss of information; if Bob got E as well, there would be no noise at all as he can undo the isometry. However, a well-known way of giving him E anyway, is to completely randomize it: denoting the discrete Weyl operators on E by W_j ($j = 1, \ldots, |E|^2$), if the channel internally picks j uniformly at random and applies W_j to E, it creates a new channel with output $\mathcal{N}(\rho)^B \otimes \frac{1}{|E|}\mathbb{1}^E$. The extra register is always constant, so the new channel has the same information properties as $\mathcal{N}$. The idea of the following channel construction is to add another "gadget" on top of this, which outputs some randomness approximating the uniform distribution above—see Fig. 1; so, intuitively, on its own it does not alter too much the classical capacity of the channel, but if paired with the right resources can increase the quantum capacity.

A comment on why we need the rather large register A_1, most of which is discarded anyway. In fact, the size (parametrized by k) has a double purpose: on the one hand, we need A_{11} to be close to maximally mixed for most inputs. But more importantly, to make it very "costly", though not impossible, to use entanglement with another system to access the index J_1 (see the proof of Theorem 1).

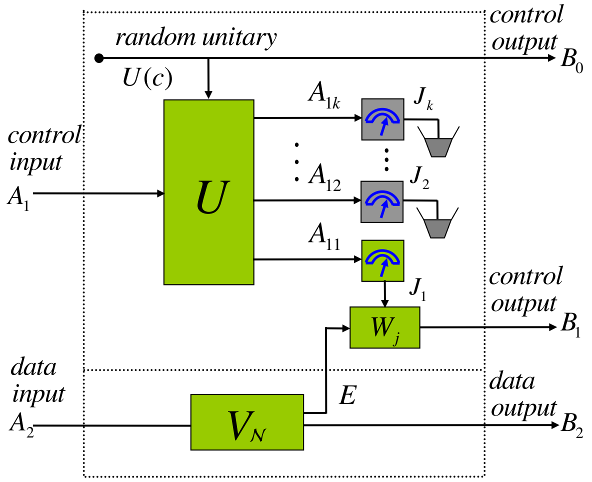

FIG. 1 (color online). The channel $\mathcal{T}^k_{\mathcal{N}}$: in the lower part it contains $\mathcal{N}$ (with input register A_2, output register B_2, and environment E). It also has another input register A_1 of dimension $c = |E|^{2k}$, which we view in a fixed way as a tensor product of k $|E|^2$-dimensional systems $A_{11}, \ldots, A_{1k}$, each coming with a fixed computational basis $\{|j\rangle\}_{j=1,\ldots,|E|^2}$. This big register is subjected to a random unitary rotation U, where U is chosen from the Haar measure and subsequently output (a classical description of it) in register B_0. All registers $A_{12}, \ldots, A_{1k}$ are discarded, only A_{11} is measured in the computational basis, and the result j used to control a unitary transformation (Weyl operator) W_j on the environment E, which is then output in the register B_1. A formal definition can be found in the supplementary material [22].

The additivity violation.—Now, if we knew that the Holevo quantity $\chi(\mathcal{T}^k_{\mathcal{N}})$ were additive for $\mathcal{T}^k_{\mathcal{N}}$, we would have $C(\mathcal{T}^k_{\mathcal{N}}) = \chi(\mathcal{T}^k_{\mathcal{N}})$. Since it is possible to show that $\chi(\mathcal{T}^k_{\mathcal{N}}) \leq \chi(\mathcal{N}) + o(1)$—this is a special case of Theorem 1 below—, we would have an upper bound

$$P(\mathcal{T}^k_{\mathcal{N}}) \leq C(\mathcal{T}^k_{\mathcal{N}}) \leq C(\mathcal{N}) + o(1). \tag{7}$$

While we are not able to show additivity for the channels $\mathcal{T}^k_{\mathcal{N}}$, the above relation is nevertheless true. In fact, we have the following general theorem, proved in full in the supplementary material section [22].

Theorem 1.—For any channel $\mathcal{N}$ with input space A, output space B and environment E, and any integer k, let $\delta(k) = \frac{1}{k}(5 + 4\log|E|)$. Then, for arbitrary channel $\mathcal{E}$,

$$\chi(\mathcal{N} \otimes \mathcal{E}) \leq \chi(\mathcal{T}^k_{\mathcal{N}} \otimes \mathcal{E}) \leq \chi(\mathcal{N} \otimes \mathcal{E}) + \delta(k). \tag{8}$$

As a consequence,

$$C(\mathcal{N}) \leq C(\mathcal{T}^k_{\mathcal{N}}) \leq C(\mathcal{N}) + \delta(k).$$

On the other hand, we can get a lower bound on $P(\mathcal{T}^k_{\mathcal{N}} \otimes \mathcal{A})$, where $\mathcal{A}$ is the 50% erasure channel of input dimension c; note that by the no-cloning principle, $P(\mathcal{A}) = Q(\mathcal{A}) = 0$. Since the private classical capacity is not smaller than the quantum capacity, which is in turn lower bounded by the coherent information, we evaluate the coherent information of the channel $\mathcal{T}^k_{\mathcal{N}} \otimes \mathcal{A}$. Let us look at an input state σ as follows: Alice prepares a maximally entangled state Φ^{A_1C} of dimension $c \times c$ and feeds the two halves into the control input (A_1) and the 50% erasure channel (C; its quantum output we also denote C, and the erasure flag D). She feeds another arbitrary state ρ^{A_2}, whose purification is denoted as $|\varphi\rangle^{AA_2}$, into the data input and keeps the system A. So the final state after the channel action is

$$\omega^{AB_0B_1B_2CD} = (\mathrm{id}_A \otimes \mathcal{T}^k_{\mathcal{N}} \otimes \mathcal{A}_C)(\Phi^{A_1C} \otimes \varphi^{AA_2}).$$

The coherent information, with respect to this state, is

$$\begin{aligned} I_c(\sigma^{A_1A_2C}, \mathcal{T}^k_{\mathcal{N}} \otimes \mathcal{A}) = {} & H(B_1B_2CD|B_0) \\ & - H(AB_1B_2CD|B_0). \end{aligned}$$

By an argument similar to that in [20], we divide the computation into two cases: the information sent into $\mathcal{A}$ is erased or not erased,

$$I_c(\sigma^{A_1A_2C}, \mathcal{T}^k_{\mathcal{N}} \otimes \mathcal{A}) = \tfrac{1}{2}(I_c^{\text{erased}} + I_c^{\text{not erased}}).$$

In the erased case, the state is decoupled between AB_2 and B_0B_1C, so the coherent information simplifies to

$$\begin{aligned} I_c^{\text{erased}} &= H(B_2) - H(AB_2) \\ &= H(\mathcal{N}(\rho)) - H((\mathrm{id} \otimes \mathcal{N})|\varphi\rangle\langle\varphi|^{AA_2}). \end{aligned}$$

When the transmitted information is not erased, Bob will be able to correct the errors encountered by the noisy channel $\mathcal{N}$ as follows. Bob reads the output B_0, learning which unitary transformation is applied by the channel $\mathcal{T}^k_{\mathcal{N}}$. Then he can measure C in the proper basis to get j, and then apply $W_j^\dagger$ to B_1, recovering the environment E. As a result, Bob possesses the output and the environment of $\mathcal{N}$ simultaneously, effectively obtaining the quantum information input into $\mathcal{N}$ completely. In this case, the system B_0C is decoupled from AB_1B_2, which is in the pure state $(\mathbb{1} \otimes V)|\varphi\rangle^{AA_2}$. So,

$$I_c^{\text{not erased}} = H(B_1B_2) - H(AB_1B_2) = H(\rho^{A_2}).$$

Adding these two cases together, we have

$$\begin{aligned} I_c(\sigma, \mathcal{T}^k_{\mathcal{N}} \otimes \mathcal{A}) = {} & \tfrac{1}{2}[H(\rho) + H(\mathcal{N}(\rho)) \\ & - H((\mathrm{id} \otimes \mathcal{N})\varphi)]. \end{aligned}$$

The term in brackets on the right-hand side is called quantum mutual information (between input and output of $\mathcal{N}$). In [21], it is proved that the maximum over ρ of the right-hand side is the entanglement-assisted quantum capacity $Q_E(\mathcal{N})$ of the channel $\mathcal{N}$. I.e.,

$$Q_E(\mathcal{N}) = \max_\rho I_c(\Phi \otimes \rho, \mathcal{T}^k_{\mathcal{N}} \otimes \mathcal{A}),$$

and hence

$$P(\mathcal{T}^k_{\mathcal{N}} \otimes \mathcal{A}) \geq Q(\mathcal{T}^k_{\mathcal{N}} \otimes \mathcal{A}) \geq Q_E(\mathcal{N}). \tag{9}$$

Now, comparing Eqs. (7) and (9), also making use of Eq. (1), we see that for all channels $\mathcal{N}$ such that

$$Q_E(\mathcal{N}) > C(\mathcal{N}), \tag{10}$$

we have, for sufficiently large k,

$$P(\mathcal{T}^k_{\mathcal{N}} \otimes \mathcal{A}) \geq Q(\mathcal{T}^k_{\mathcal{N}} \otimes \mathcal{A}) > P(\mathcal{T}^k_{\mathcal{N}}) \geq Q(\mathcal{T}^k_{\mathcal{N}}).$$

Note that the channel $\mathcal{A}$ has zero private classical capacity and zero quantum capacity, so this exhibits the violations of the additivity of private classical capacity and the quantum capacity at the same time.

All we need now is to find quantum channels that satisfy Eq. (10). One example is the depolarizing channel of arbitrary dimension d, for which both capacities are known [13,21], $\mathcal{D}_q(\rho) = (1-q)\rho + q\frac{1}{d}\mathbb{1}$. For large d, the gap becomes asymptotically $\frac{1}{2}H(q, 1-q)$ [22].

There also exist large additivity violations: In [[23], theorem V.1] it is proven that in sufficiently large dimension d, there exist $n = \lfloor(\log d)^4\rfloor$ orthogonal bases $\mathcal{B}_\nu = (|b_1^{(\nu)}\rangle, \ldots, |b_d^{(\nu)}\rangle)$ such that for all states ρ,

$$\frac{1}{n}\sum_{\nu=1}^{n} H(\mathcal{B}_\nu|\rho) \geq \log d - 4,$$

where $H(\mathcal{B}_\nu|\rho) = -\sum_{i=1}^{d}\langle b_i^{(\nu)}|\rho|b_i^{(\nu)}\rangle \log\langle b_i^{(\nu)}|\rho|b_i^{(\nu)}\rangle$ is the Shannon entropy of the outcome distribution when measuring the state ρ in basis $\mathcal{B}_\nu$. What this means is that the channel $\mathcal{N}$ from d to dn dimensions, defined as

$$\mathcal{N}(\rho) = \sum_{\nu=1}^{n}\sum_{i=1}^{d}\frac{1}{n}\langle b_i^{(\nu)}|\rho|b_i^{(\nu)}\rangle|i\rangle\langle i|^B \otimes |\nu\rangle\langle\nu|^{B'},$$

satisfies $\chi(\mathcal{N}) \leq 4$. Since the channel is entanglement-breaking, the additivity result of [14] applies, so $C(\mathcal{N}) = \chi(\mathcal{N}) \leq 4$. On the other hand, it is straightforward to see that $Q_E(\mathcal{N}) = \frac{1}{2}\log d$. Thus we find that almost the entire bandwidth of $\mathcal{N}$ can be activated by the presence of entanglement. Now, to construct the example for activation of the secret capacity by a 50% erasure channel, we observe that $|E| = dn = d(\log d)^4$. We choose $k = \log|E|$, and get from Theorem 1 that $C(\mathcal{T}^k_{\mathcal{N}}) \leq O(1)$, while at the same time $Q(\mathcal{T}^k_{\mathcal{N}} \otimes \mathcal{A}) \geq \frac{1}{2}\log d$.

Conclusion.—We showed a way of converting any gap between classical capacity and entanglement-assisted quantum capacity of a channel into a violation of the additivity of the private capacity of the channel tensored with a 50% erasure channel. In fact, the quantum capacity of the tensor product channel is larger than the classical capacity of the single channel.

The construction is based on a certain embedding of the given channel into a version of the echo-correctable channels from [16]. That the pairing with the erasure channel gives larger quantum capacity follows from the echo-correctable reasoning of the benefit of sharing entanglement. On the other hand, the upper bound on the classical capacity relies on showing that the additional "gadgets" built around the given channel increase the capacity by an arbitrarily small amount. The argument is different from proving additivity of χ of the channel (which we cannot do for $\mathcal{T}^k_{\mathcal{N}}$), and also from the use of the recent continuity bound [24] (which cannot be applied as $\mathcal{T}^k_{\mathcal{N}}$ is at finite distance from any channel for which we know the capacity).

Thus, we even get a new type of example for the nonadditivity of the quantum capacity Q, which is different from that of Smith and Yard [17] as our channel is not PPT entanglement binding. Furthermore, while in [17] the lower bound of half the private capacity on the quantum capacity of the tensor product was enough, here we experience even a large gap between these two quantities. However, we also note a conceptual analogy in the constructions: The PPT entanglement binding channel used in [17] derives from a so-called pbit state [25]. It provides Alice and Bob with shared randomness—which is made private by distributing the purification among Alice and Bob, but in a scrambled way that makes it impossible for them to recover much of the entanglement. Our channel randomizes the environment and hence gives it to Bob in an encrypted way, limiting the receiver's knowledge about the noise encountered by the channel. In the construction of [17] as in the present one, the availability of additional resources allows Alice and Bob to break the encryption and access the entanglement.

K. L. thanks the University of Bristol and CQT at NUS, where part of this work was done, for their hospitality. K. L., X. B. Z. and G. C. G. were supported by the National Fundamental Research Program (Grant No. 2009CB929601), the National Natural Science Foundation of China (Grant Nos. 10674128 and 60121503), and the Innovation Funds and "Hundreds of Talents" program of the Chinese Academy of Sciences. A. W. was supported by the U.K. EPSRC, by the Royal Society, the Philip Leverhulme Trust, and by the European Commission. The Centre for Quantum Technologies is funded by the Singapore Ministry of Education and the National Research Foundation as part of the Research Centres of Excellence programme.

*leeke@mail.ustc.edu.cn
†a.j.winter@bris.ac.uk
‡xbz@ustc.edu.cn
§gcg@ustc.edu.cn

[1] C. E. Shannon, Bell Syst. Tech. J. **27**, 379 (1948).
[2] T. M. Cover and J. A. Thomas, *Elements of Information Theory* (Wiley, New York, 1991).
[3] C. H. Bennett and G. Brassard, *Proceedings of IEEE International Conference Computers, System and Signal Processing* (IEEE, New York, 1984), p. 175.
[4] A. S. Holevo, IEEE Trans. Inf. Theory **44**, 269 (1998); B. Schumacher and M. D. Westmoreland, Phys. Rev. A **56**, 131 (1997).
[5] I. Devetak, IEEE Trans. Inf. Theory **51**, 44 (2005).
[6] S. Lloyd, Phys. Rev. A **55**, 1613 (1997).
[7] P. W. Shor, http://www.msri.org/publications/ln/msri/2002/quantumcrypto/shor/1/.
[8] B. Schumacher and M. Nielsen, Phys. Rev. A **54**, 2629 (1996).
[9] M. Hastings, Nature Phys. **5**, 255 (2009).
[10] G. Smith, J. Renes, and J. A. Smolin, Phys. Rev. Lett. **100**, 170502 (2008).
[11] D. P. DiVincenzo, P. W. Shor, and J. A. Smolin, Phys. Rev. A **57**, 830 (1998).
[12] C. King, J. Math. Phys. (N.Y.) **43**, 4641 (2002).
[13] C. King, IEEE Trans. Inf. Theory **49**, 221 (2003).
[14] P. W. Shor, J. Math. Phys. (N.Y.) **43**, 4334 (2002).
[15] I. Devetak and P. W. Shor, Commun. Math. Phys. **256**, 287 (2005).
[16] C. H. Bennett, I. Devetak, P. W. Shor, and J. A. Smolin, Phys. Rev. Lett. **96**, 150502 (2006).
[17] G. Smith and J. Yard, Science **321**, 1812 (2008).
[18] J. Oppenheim, Science **321**, 1783 (2008).
[19] I. Devetak and A. Winter, Phys. Rev. Lett. **93**, 080501 (2004).
[20] G. Smith and J. A. Smolin, Phys. Rev. Lett. **102**, 010501 (2009).
[21] C. H. Bennett, P. W. Shor, J. A. Smolin, and A. V. Thapliyal, Phys. Rev. Lett. **83**, 3081 (1999); IEEE Trans. Inf. Theory **48**, 2637 (2002).
[22] See EPAPS Document No. E-PRLTAO-103-012940 for the mathematical details. For more information on EPAPS, see http://www.aip.org/pubservs/epaps.html.
[23] P. Hayden, D. Leung, P. W. Shor, and A. Winter, Commun. Math. Phys. **250**, 371 (2004).
[24] D. Leung and G. Smith, arXiv:0810.4931.
[25] K. Horodecki, M. Horodecki, P. Horodecki, and J. Oppenheim, Phys. Rev. Lett. **94**, 160502 (2005).