When Does Noise Increase the Quantum Capacity?

Fernando G. S. L. Brandão,[1] Jonathan Oppenheim,[2] and Sergii Strelchuk[2]

[1]Departamento de Física, Universidade Federal de Minas Gerais, Belo Horizonte, Caixa Postal 702, 30123-970, MG, Brazil
[2]Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge CB3 0WA, United Kingdom

(Received 31 July 2011; published 25 January 2012)

Superactivation is the property that two channels with zero quantum capacity can be used together to yield a positive capacity. Here we demonstrate that this effect exists for a wide class of inequivalent channels, none of which can simulate each other. We also consider the case where one of two zero-capacity channels is applied, but the sender is ignorant of which one is applied. We find examples where the greater the entropy of mixing of the channels, the greater the lower bound for the capacity. Finally, we show that the effect of superactivation is rather generic by providing an example of superactivation using the depolarizing channel.

DOI: 10.1103/PhysRevLett.108.040501 PACS numbers: 03.67.Hk, 05.40.Ca

A quantum channel is any physical process which can be applied to a quantum system. There is an input to the channel, and we are interested in how much information remains at the output. Some channels are so noisy that no quantum information can be reliably transmitted through them—error correction becomes impossible and one cannot send a quantum system through the channel faithfully. We say that such channels have zero capacity. In classical information theory, zero-capacity channels are not interesting, because they only include the case where there is no correlation between the input and output. However, some zero-capacity quantum channels have surprising properties: for example, they can be used to share a private key [1,2], and two zero-capacity channels can be combined in parallel to reliably send quantum states, a situation that is impossible classically [3].

The ability to send quantum information down two channels which have zero capacity is called superactivation, and it is an important phenomenon which suggests that quantum channels are radically different from classical ones. For classical channels, we can quantify a channel by its capacity, while the phenomena of superactivation means that for a quantum channel the capacity does not adequately characterize the channel, since the utility of the channel depends on what other channels are also available. One hopes that a greater understanding of superactivation will enable progress to be made in understanding the quantum capacity, something made difficult because we still do not have an adequate formula for it. Additionally, there appear to be strong links between superactivation and privacy [4,5], and these are not yet properly understood.

Despite the importance of superactivation, only one example is known [3]: one of the channels is a symmetric channel, meaning that the quantum state of the output and the environment is symmetric under exchange. This channel cannot be used for quantum communication because its symmetry implies that if this channel had positive quantum capacity it would violate the no-cloning theorem [6]. An example is the 50% erasure channel, denoted as $\mathcal{N}_e^{0.5}$, which faithfully transmits the input state half of the time and outputs an erasure flag in the rest of the cases. The only known protocol for superactivation involved using the 50% erasure channel. The second channel is one which produces a private key, but cannot be used to send quantum information [1]. Such a channel is known to have zero capacity because it has a positive partial transpose (PPT) [7], which implies that it has zero capacity [8].

It was also shown in [3] that a convex combination of "flagged" channels,

$$\mathcal{N} = \kappa \mathcal{N}_{\gamma^{(d)}} \otimes |0\rangle\langle 0|_B + (1-\kappa)\mathcal{N}_e^{0.5} \otimes |1\rangle\langle 1|_B, \quad (1)$$

has positive quantum capacity for a particular private channel $\mathcal{N}_{\gamma^{(d)}}$ and for a very small amount of mixing ($\kappa = 0.0041$).

It is natural to ask about the generality of this phenomenon. First, do there exist communication protocols that allow for strong nonconvexity of quantum capacity, in the sense that κ can have a large range? Indeed, we will find here that one can achieve positive capacity for any $0 < \kappa < 1$. This surprising result implies that a generic mixing of the zero-capacity channels during the transmission will, nevertheless, increase the quantum capacity. In fact, we find situations where, counterintuitively, the more noise, the greater the lower bound for the capacity given by the so-called coherent information. A second question we address is, what types of channels can be superactivated? Since there are very limited techniques to show a channel has zero capacity, this is a difficult problem. It was not presently known whether this startling effect can be generalized to any channels other than $\mathcal{N}_e^{0.5}$. Here we find that superactivation is possible for a large class of inequivalent and generic channels (in the sense that they cannot simulate each other). This includes erasure channels with any probability $p \in [\frac{1}{2}, 1)$ of erasure, as well as the common depolarizing channel [9]. Third, we are interested in

whether superactivation is robust against noise or can only be demonstrated using perfectly noiseless resources. This is particularly important in lieu of proposed experiments to test this effect [10]. We answer this question affirmatively.

It is of course a basic question in quantum information theory to quantify the ability of quantum channels to transmit quantum states faithfully. The former is described mathematically as a completely positive trace preserving map $\mathcal{N}:A \to B$ from density matrices on input system A to density matrices on an output system B. The performance of a quantum channel for noiseless quantum communication is characterized by its quantum capacity $\mathcal{Q}(\mathcal{N})$, which is the maximum achievable rate for quantum communication. Analogously, $\mathcal{Q}(\mathcal{N})$ quantifies the amount of pure state entanglement that can be transmitted through $\mathcal{N}$.

The quantum capacity is known to be lower bounded by the coherent information [11–13]:

$$\mathcal{Q}(\mathcal{N}) \geq I_c(A\rangle B) := \max_{\rho}[S(B)_\sigma - S(E)_\sigma], \quad (2)$$

where the von Neumann entropies are evaluated on $\sigma_{BE} = U\rho U^\dagger$, with $U:A \mapsto BE$ the isometry associated to the channel $\mathcal{N}$ as follows: $\mathcal{N}(\rho) = \mathrm{tr}_E(U\rho U^\dagger)$. The first family of zero-capacity channels we will consider, denoted as $\mathcal{N}_{\gamma^{(d)}}$, produce bound entangled states—states that need pure state entanglement to create them, but from which no pure state entanglement can be extracted [14]. Such states, despite being useless for transmission of quantum information, may contain secrecy [1]. Here we take $\mathcal{N}_{\gamma^{(d)}}$ to be such a channel which produces bound entangled states that contain secrecy and, in particular, "private bits."

Private bits and coherent information.—Quantum states that contain d bits of secrecy are called private dits, pdits, or twisted ebits [1,15] and have the generic form

$$\gamma^{(d)} = UP^+_{AB} \otimes \sigma_{A'B'}U^\dagger, \quad (3)$$

where $U = \sum_{i,j=0}^{d-1} |ij\rangle\langle ij|_{AB} \otimes U_{ij}$ is a controlled unitary operation termed twisting (with arbitrary unitaries U_{ij}), P^+_{AB} is the projector onto a d dimensional maximally entangled state $\Phi^+_{AB} = \frac{1}{\sqrt{d}}\sum_{i=0}^{d-1}|ii\rangle_{AB}$, and $\sigma_{A'B'}$ is an arbitrary state called the "shield" subsystem of dimension d', for its presence protects private correlations. In the case when $d = 2$ we will call it a pbit. Parties that have A and B subsystems of a pdit (known as the "key") can extract $\log_2 d$ ebits by performing $U^\dagger$ if one of them possesses the shield $A'B'$ in its entirety. However, when the shield is split between the two parties, it can be impossible to perform the untwisting using only local operations, and there exist states which are arbitrarily close to pdits, yet no ebits can be produced from them. The main idea we will be exploiting here is that superactivation can occur by one zero-capacity channel being used to share pdits, and then by Alice using a second zero-capacity channel to send her part of the shield A' to Bob some of the time so that he can perform the untwisting operation, giving them shared ebits [16] on these occasions.

We will thus consider using $\mathcal{N}_{\gamma^{(d)}}$ in conjunction with a number of different channels: first, erasure channels $\mathcal{N}^p_e$, which output an erasure flag with probability $p \in [\frac{1}{2}, 1)$, and faithfully transmit the input state otherwise. These are all inequivalent channels, in the sense that for $p \in \{1 - \frac{1}{n}|n \in \mathbb{N}\backslash\{1\}\}$ no such channel with probability p can simulate one with probability of erasure smaller than p [17]. Moreover, it is known that $\mathcal{N}^p_e$ retains zero capacity in this range since a higher erasure probability can only decrease the capacity. Our results hold for all $p \in [\frac{1}{2}, 1)$.

Strong nonconvexity of quantum capacity.—Consider the convex combination of two channels as in Eq. (1), where $\mathcal{N}_{\gamma^{(d)}}$ is the PPT channel that generates noisy pdits, which can be made arbitrarily close to perfect pdits at the expense of increasing the dimension of the shield, and the erasure probability of the latter channel is in the range $p \in [\frac{1}{2}, 1)$. We take the input dimension of both channels to be equal. For clarity of presentation, we will consider the limiting case, when the dimension of the shield goes to infinity, and take the key part to be perfect. Both the PPT pdit channel and the erasure channel have zero quantum capacity. The quantum capacity of the resulting mixture of the two channels can be strictly positive when $p = 0.5$, and $\kappa \in (0; 0.0041)$ [3]. We now show that this is much more generic, and will employ the protocol described below to show that for the PPT pdit channel and 50% erasure channel in the convex mixture we can surprisingly achieve positive quantum capacity for all $\kappa \in (0, 1)$.

More formally, consider a channel $\mathcal{N}$ in the form of Eq. (1) together with initial state $\rho_{ABA'B'} = (\Phi^+_{AB})^{\otimes \log d} \otimes (\Phi^+_{A'B'})^{\otimes \log d'}$ on Alice with A of dimension d and A' of dimension d' and consider the following protocol.

(1) Alice initially feeds subsystems BB' of $\rho_{ABA'B'}$ through $\mathcal{N}$, keeping AA'. If this is repeated n times, then at the end of this step they share n instances of $AA'BB'$, where each of the instances has the form of a convex mixture of the pdit and the state, which experienced the action of the erasure channel.

(2) Alice feeds her instances of A' into the channel, and pads her input with $\log d$ fresh qubits which will not play any role in this round of the protocol and are discarded by Bob. After the transmission Alice and Bob have n instances of subsystems A and $BB'\tilde{A}'$, respectively.

At the end of the protocol we get

$$I_c(A\rangle B)_{\gamma^{(d)}} = \tfrac{1}{2}(1 - \kappa)[\kappa - p(\kappa + 2) + 1]\log(d). \quad (4)$$

When $p = 0.5$ the expression for the coherent information simplifies to

$$I_c(A\rangle B)_{\gamma^{(d)}} = \tfrac{1}{4}(1 - \kappa)\kappa \log(d). \quad (5)$$

See Appendix A in the Supplemental Material [18] for the calculation of the coherent information. Figure 1

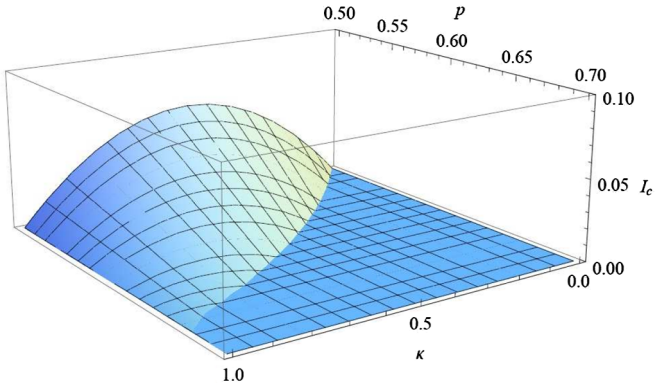

FIG. 1 (color online). Nonconvexity of quantum capacity for $I_c(A\rangle B) = \frac{1}{2}(1-\kappa)[\kappa - p(\kappa+2)+1]\log d$ when $d = 2$ when the dimension of the shield subsystem tends to infinity.

demonstrates the full range of pairs (κ, p) for which the violation of the convexity of quantum capacity is achieved.

The full nonconvexity of the coherent information for the convex combination (1) holds when $p = 0.5$, when the dimension of the shield subsystem tends to infinity, and is not true for larger p. This is also where the greater the entropy of mixing of the two channels, the greater the lower bound for the capacity given by the coherent information.

Inequivalent classes of superactivating channels with noisy resources.—We next address the question of generalizing the superactivation example to the class of erasure channels with $p > \frac{1}{2}$, and we will simultaneously tackle the question of robustness of superactivation to noise. We do so by establishing the region of pairs (p, ϵ), where $p \in [\frac{1}{2}, 1)$ is the erasure probability and ϵ denotes the amount of tolerable noise in the PPT pbit channel, for which we can demonstrate superactivation.

In the two-step protocol [16] that achieves superactivation for an arbitrary pbit channel $\mathcal{N}_{\gamma^{(d)}}$, Alice and Bob first use the pbit channel to share states of Eq. (3), then in the second step Alice sends her part of the shield (subsystem A') through the erasure channel $\mathcal{N}_e^{0.5}$. Half of the time, when the erasure does not take place, Bob is able to perform the $U^\dagger$ of Eq. (3) and they end up sharing an ebit. When erasure occurs, they are left with a classically correlated state and an erasure flag. We now show that this protocol works for other values of p. Since the case of erasure (nonerasure) is distinguishable on Bob's site, the lower bound for the capacity of the joint channel $\mathcal{N}_\gamma \otimes \mathcal{N}_e^p$ is just the coherent information averaged over the two cases:

$$\mathcal{Q}(\mathcal{N}_{\gamma^{(d)}} \otimes \mathcal{N}_e^p) \geq pI_c(A\rangle B)_{\gamma_{\mathrm{er}}^{(d)}} + (1-p)I_c(A\rangle B)_{\gamma_{\mathrm{uner}}^{(d)}}, \tag{6}$$

where $p = 0.5$ in the original example [3], and the first term is evaluated on the state $\gamma_{\mathrm{er}}^{(d)}$ that corresponds to the case when Bob received the erasure flag while the latter is evaluated on $\gamma_{\mathrm{uner}}^{(d)}$, when Alice's share of the shield was successfully transmitted to Bob. If the erasure event takes place, and the shield does not get through, Bob will not be able to undo the unitary U, so $I_c(A\rangle B)_{\gamma_{\mathrm{er}}^{(d)}} = 0$. If the shield gets through, assuming operations are perfect, $I_c(A\rangle B)_{\gamma_{\mathrm{uner}}^{(d)}} = \log d$. In the case of many copies, Alice and Bob will share $m = (1-p)n$ pdits on average and

$$I_c(A\rangle B)_{(\gamma^{(d)})^{\otimes m}} = mI_c(A\rangle B)_{\gamma_{\mathrm{uner}}^{(d)}} = (1-p)n\log(d). \tag{7}$$

This is under the assumption that the pbits are perfectly private, and so to investigate what happens when this restriction is lifted, we consider channels which produce approximate pbits.

Definition: The state $\tilde{\gamma}^{(d)}$ is called an ϵ-approximate pdit if there exists a set of local measurement operators on the key subsystem of Alice and Bob $\{P_i^A \otimes P_j^B\}_{i,j=1}^d$ such that

$$\left| \mathrm{Tr}_{A'B'}\left(\sum_{ij} P_i^A \otimes P_j^B \tilde{\gamma}^{(d)} P_j^B \otimes P_i^A\right) - K_{AB} \otimes M_E \right| \leq \epsilon, \tag{8}$$

where K_{AB} represents the perfect key and M_E represents the environment.

An approximate pdit satisfies the following property: For every $\tilde{\gamma}^{(d)}$ there exists a unitary $U = \{U_{ij}\}$ on the system such that

$$|U^\dagger \tilde{\gamma}^{(d)} U - \Phi_{AB}^+ \otimes \sigma_{A'B'}| \leq \epsilon. \tag{9}$$

This follows directly from Theorem 2 in [15]. From now on, we will limit the set of all approximate pdits to the subset of the approximate pdits which have PPT. The existence of good PPT approximations of pdits is shown in [1]. It is known that using Choi-Jamiolkowski isomorphism from each such state we can construct a channel that produces it in the same way as it was done in [8].

Following the same protocol as in [3,16], consider a pair of channels

$$\tilde{\mathcal{N}}_{\gamma^{(d)}} \otimes \mathcal{N}_e^p, \tag{10}$$

with $p \in [\frac{1}{2}, 1)$, where using $\tilde{\mathcal{N}}_{\gamma^{(d)}}$ results in Alice and Bob sharing an ϵ-approximate pdit $\tilde{\gamma}^{(d)}$. Then Alice sends her share of the shield to Bob using $\mathcal{N}_e^p$ as above. After many independent uses of $\tilde{\mathcal{N}}_{\gamma^{(d)}}$ they share $m = (1-p)n$ copies of $\tilde{\gamma}^{(d)}$. The question of interest is whether given a large number n of $\tilde{\gamma}^{(d)}$ Alice and Bob could superactivate them with an erasure channel of probability p, i.e., whether there exist pairs (p, ϵ) which will make the lower bound on the quantum capacity given by Eq. (10) strictly positive. The following lemma will make use of Eq. (6) and relation (8) to derive a lower bound on the joint channel of Eq. (10).

Lemma 1: Consider independent uses of $\tilde{\mathcal{N}}_{\gamma^{(d)}} \otimes \mathcal{N}_e^p$, $p \in [\frac{1}{2}, 1)$. Then

$$\mathcal{Q}(\tilde{\mathcal{N}}_{\gamma^{(d)}} \otimes \mathcal{N}_e^p) \geq (1 - p - 4\epsilon)\log(d) - 2h(\epsilon), \quad (11)$$

where d is the dimension of the key part, and $h(\cdot)$ is a binary entropy.

See Appendix B in the Supplemental Material [18] for the proof and graphical illustration.

Superactivation using depolarizing channel.—It turns out that the erasure channel and its variants are not the only channels that can be used in conjunction with the PPT pbit channel for superactivation. Here we also consider $\tilde{\mathcal{N}}_{\gamma^{(d)}} \otimes \mathcal{N}_{\text{dep}}$, with $\mathcal{N}_{\text{dep}}$ the commonplace depolarizing channel [9] given by

$$\mathcal{N}_{\text{dep}} = p\mathcal{N}_{\text{id}} + (1 - p)\mathcal{N}_{\text{mix}}. \quad (12)$$

The first channel in this mixture is the identity channel acting as $\mathcal{N}_{\text{id}}(\rho) = \rho$, and the second one is the completely randomizing channel acting as $\mathcal{N}_{\text{mix}}(\rho) = \frac{\mathbb{1}}{r}$. The depolarizing channel is so ubiquitous in part because all quantum channels can be twirled to this form by applying some randomly chosen bilateral unitary to the system left at the sender's site and output of the channel [9]. It follows that $\mathcal{N}_{\text{dep}}$, for arbitrary input dimension r, is antidegradable [19] and thus has zero capacity in the range $p \in [0; \frac{1}{2}]$. This follows from the fact that the Jamiolkowski state associated with the channel $1/2(P_{AB}^+ + \mathbb{1}_{AB}/r^2)$ has a two-symmetric extension, namely, $1/2(P_{AB}^+ \otimes \mathbb{1}_{\mathbb{B}'}/r + P_{AB'}^+ \otimes \mathbb{1}_B/r)$. Remarkably, we will find that this channel can be used for superactivation, even as the amount of noise is made arbitrarily large.

The superactivation protocol is as before—after creating approximate pbits using $\tilde{\mathcal{N}}_{\gamma^{(d)}}$, Alice sends the shield A' to Bob through the depolarizing channel. Unlike the previous examples of erasure channels, there are no flags attached to the output, so Bob does not know which channel was applied. After the transmission, Alice and Bob are left with the mixture of two states: with probability p, after Bob performs the untwisting operation $U^\dagger$, they share the noisy maximally entangled state $\Phi_{AB,\epsilon}^+$ such that $\|\Phi_{AB,\epsilon}^+ - \Phi_{AB}^+\| \leq \epsilon$, and with probability $(1 - p)$ the ebits cannot be untwisted and they share the state $\sigma_{AB,\epsilon}$ that approximates the classically correlated state $\sigma_{AB} := 1/d\sum_k |k, k\rangle\langle k, k|$; i.e., they share the state

$$\omega_{AB} = p\Phi_{AB,\epsilon}^+ + (1 - p)\sigma_{AB,\epsilon}. \quad (13)$$

The fact that we only get an approximation $\sigma_{AB,\epsilon}$ of the classically correlated state is due to the fact that the channel $\tilde{\mathcal{N}}_{\gamma^{(d)}} \otimes \mathcal{N}_{\text{dep}}$ only created approximate pbits. For any $\epsilon > 0$ we can choose the dimension of the shield state and of the depolarizing channel sufficiently large so that $\|\sigma_{AB,\epsilon} - \sigma_{AB}\|_1 \leq \epsilon$. The coherent information, evaluated on ω_{AB} for $d = 2$, can be lower bounded as follows:

$$I_c(A\rangle B)_{\omega_{AB}} \geq 1 + \frac{1 - p}{2}\log\left(\frac{1 - p}{2}\right) \quad (14)$$

$$+ \frac{1 + p}{2}\log\left(\frac{1 + p}{2}\right) - 4\epsilon\log(d) + 2h(\epsilon). \quad (15)$$

This follows by computing the coherent information for $p\Phi_{AB}^+ + (1 - p)\sigma_{AB}$ and using Fannes inequality and the relation $\|\sigma_{AB,\epsilon} - \sigma_{AB}\|_1 \leq \epsilon$. For any fixed p we can take the dimension of the depolarizing channel and of the shield part of $\mathcal{N}_{\gamma^{(2)}}$ sufficiently large so that ϵ is as small as we wish. In this regime we find superactivation for a large region of values of p in the range $(0, \frac{1}{2}]$, which constitute new examples of superactivation using the depolarizing channel [see Appendix C in the Supplemental Material [18] for the plot of the region for $[p, \epsilon(p)]$].

We have seen that superactivation does not only occur for the two special channels considered in the initial discovery of the effect. Rather, there are classes of generic and common channels, as well as inequivalent ones, which can be used for superactivation and, likewise, for the curious effect where adding noise (by increasing the entropy of mixing of two channels) can increase the quantum capacity. Here too, we find that it is not a tiny mixture of noise which increases the capacity, but rather, there are cases where the more the noise, the greater the capacity, and generally any amount of mixing can result in positive capacity. Although we have found superactivation to be more generic than previously thought, we have only considered cases where one channel has zero capacity because it is PPT, and the other channel has zero capacity because of the no-cloning bound. The big question of whether superactivation exists for channels which do not each belong to these classes remains unanswered. This is a challenging question since at the moment we have no other way of showing a channel has zero capacity. We hope the considerations here provide some clues to the answer.

[1] K. Horodecki, M. Horodecki, P. Horodecki, and J. Oppenheim, Phys. Rev. Lett. **94**, 160502 (2005).
[2] K. Horodecki, M. Horodecki, P. Horodecki, D. Leung, and J. Oppenheim, IEEE Trans. Inf. Theory **54**, 2604 (2008).
[3] G. Smith and J. Yard, Science **321**, 1812 (2008).
[4] G. Smith and J. A. Smolin, Phys. Rev. Lett. **103**, 120503 (2009).
[5] K. Li, A. Winter, X. Zou, and G. Guo, Phys. Rev. Lett. **103**, 120501 (2009).
[6] W. K. Wootters and W. H. Zurek, Nature (London) **299**, 802 (1982).
[7] A. Peres, Phys. Rev. Lett. **77**, 1413 (1996).
[8] P. Horodecki, M. Horodecki, and R. Horodecki, J. Mod. Opt. **47**, 347 (2000).

[9] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, Phys. Rev. A **54**, 3824 (1996).
[10] G. Smith, J. A. Smolin, and J. Yard, arXiv:1102.4580.
[11] S. Lloyd, Phys. Rev. A **55**, 1613 (1997).
[12] P. Shor, in Proceedings of the MSRI Workshop on Quantum Computation, 2002 (unpublished) [http://www.msri.org/publications/ln/msri/2002/quantumcrypto/shor/1/meta/aux/shor.pdf].
[13] I. Devetak, arXiv:quant-ph/0304127.
[14] M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Rev. Lett. **80**, 5239 (1998).
[15] K. Horodecki, M. Horodecki, P. Horodecki, and J. Oppenheim, IEEE Trans. Inf. Theory **55**, 1898 (2009).
[16] J. Oppenheim, Science **321**, 1783 (2008).
[17] M. Hastings (private communication).
[18] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.108.040501 for proofs and plots.
[19] T. S. Cubitt, M. B. Ruskai, and G. Smith, J. Math. Phys. (N.Y.) **49**, 102104 (2008).