Quantum Flags and New Bounds on the Quantum Capacity of the Depolarizing Channel

Marco Fanizza,* Farzad Kianvash,† and Vittorio Giovannetti
NEST, Scuola Normale Superiore and Istituto Nanoscienze-CNR, I-56126 Pisa, Italy

(Received 4 December 2019; accepted 4 May 2020; published 8 July 2020)

A new upper bound for the quantum capacity of the d-dimensional depolarizing channels is presented. Our derivation makes use of a flagged extension of the map where the receiver obtains a copy of a state σ_0 whenever the messages are transmitted without errors, and a copy of a state σ_1, when instead the original state gets fully depolarized. By varying the overlap between the flag states, the resulting transformation nicely interpolates between the depolarizing map (when $\sigma_0 = \sigma_1$), and the d-dimensional erasure channel (when σ_0 and σ_1 have orthogonal support). We find sufficient conditions for degradability of the flagged channel, which let us calculate its quantum capacity in a suitable parameter region. From this last result we get the upper bound for the depolarizing channel, which by a direct comparison appears to be tighter than previous available results for $d > 2$, and for $d = 2$ it is tighter in an intermediate regime of noise. In particular, in the limit of large d values, our findings present a previously unnoticed $\mathcal{O}(1)$ correction.

DOI: 10.1103/PhysRevLett.125.020503

Introduction.—Quantum Shannon theory [1,2] provides a characterization of the maximum transmission rates (capacities) achievable in sending classical or quantum data through a quantum channel. Unfortunately, at variance with the classical case [3], for most of the models the evaluation of these quantities cannot be performed algorithmically, the computation being so hard that the identification of good bounds is already considered as an important achievement. The difficulty of the task originates on one hand from the possibility of sending entangled messages across successive uses of the transmission line, and, on the other hand, from the superadditivity properties of the information-theoretic quantities involved in the computation. Instances of superadditivity have been shown for the classical capacity [4], the quantum capacity [5–11], the classical private capacity [12], and for the trade-off capacity region [13,14]. A striking consequence of these effects is superactivation: two channels with zero quantum capacity show nonzero quantum capacity if used together [15,16]. However, it is crucial to stress that, rather than being an exotic phenomenon, superadditivity manifests itself even in the simplest cases. Indeed, it holds for the coherent information of the depolarizing channel (DC) [5,6], which is the simplest and most symmetric nonunitary quantum channel [17]. Still, despite the considerable efforts that have been spent on this issue [18–30], its quantum capacity [31,32] is not known.

DC has a peculiar position in the theory which makes it an important error model for finite dimensional systems, like qubits in a quantum computer. Indeed by pre- and postprocessing and classical communication via twirling [33], any other channel can be mapped into a DC whose quantum capacity is lower than or equal to the quantum capacity of the original channel [34]. Accordingly the value of the quantum capacity of DC can be used to lower bound the minimum number of physical qubits needed to preserve quantum information in quantum processors and memories. In the view of these facts it is clear that the DC quantum capacity problem is of primary importance in quantum information theory: solving it would likely help in understanding the peculiar difficulties of quantum communication and error correction.

The main result of this Letter is a new analytic upper bound to the quantum capacity of the DC valid for any finite dimension, which outperforms previous results in many different regimes. To achieve this goal we rely on flagged extensions of quantum channels, a construction which, in other contexts, proved to be a powerful tool, see, e.g., Ref. [10]. In our case we define the flagged depolarizing channel (FDC) assuming that if Alice sends the density matrix ρ, with probability p Bob receives such a state together with an ancillary system prepared into the state σ_0, and, with probability $1 - p$, the completely mixed state together with the ancillary system in σ_1. The density matrices σ_0 and σ_1 behave as flags that encode information about what happened to the input and, at variance with previous approaches [19,21,24], are not assumed to be necessarily orthogonal—when this happens Bob can know exactly if he received the original message or an error, and our FDC is equivalent to the erasure channel [35]. By tracing out the flags, Bob effectively receives the output of a DC. This means that the FDC is a better communication line than its associated DC, therefore every capacity of the former is larger than or equal to the corresponding value of the latter. Most importantly it is possible to find p, σ_0, σ_1 such that the FDC becomes degradable [36,37]. Degradable maps are a special set of quantum channels which have the peculiar property of admitting a nonsuperadditive coherent information [18], hence allowing for a quantum capacity formula that needs not to be regularized over infinite many channel uses—see Eqs. (3) and (9) below. Exploiting this fact and the special symmetries of the model we can produce an analytical expression for the quantum capacity of the FDC which in turns provides an analytic bound (throughout the Letter "bound" refers to upper bound, unless explicitly stated otherwise) for the quantum capacity of the associated DC.

Furthermore, exploiting a convexity argument given in [19], we also show how to merge our new inequality with those obtained in [19,21] to get an extra bound. The resulting constraint is strictly better than the one obtainable by [19,21] alone and yields the best analytic limit on the quantum capacity of the DC for all choices of d and p. For $d = 2$, the bounds in [22,23] perform better at low noise, while for higher noise our expression is better, surpassing also the one in [24] in an intermediate region. Most notably the improvement increases in the large d limit: the gap between the best upper bound and lower bound of the quantum capacity is given by a $\mathcal{O}(1)$ function of p, which is differentiable in $p = 0$, in contrast with previous bounds for which the $\mathcal{O}(1)$ term of the gap is the binary entropy $h(p)$.

Preliminaries.—Given a finite dimensional Hilbert space $\mathcal{H}$, we write the space of linear operators on $\mathcal{H}$ as $\mathcal{L}(\mathcal{H})$ and the set of density operators as $\mathfrak{S}(\mathcal{H})$. The action of a quantum channel Λ: $\mathcal{L}(\mathcal{H}_A) \to \mathcal{L}(\mathcal{H}_B)$ connecting two systems described by the Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$, is a completely positive trace preserving (CPTP) map [2] on $\mathcal{L}(\mathcal{H}_A)$ which can always be cast in the Stinespring representation form,

$$\Lambda(\theta) = \mathrm{tr}_{E'}(U_{AE}\theta_A \otimes |e\rangle\langle e|_E U_{AE}^\dagger), \tag{1}$$

where $|e\rangle_E$ is the state of environment interacting with the system A, and U_{AE} is an unitary interaction acting on $\mathcal{H}_A \otimes \mathcal{H}_E \cong \mathcal{H}_B \otimes \mathcal{H}_{E'}$. In this setting the complementary channel $\tilde{\Lambda}$: $\mathcal{L}(\mathcal{H}_A) \to \mathcal{L}(\mathcal{H}_{E'})$ is defined as the CPTP mapping

$$\tilde{\Lambda}(\theta) \coloneqq \mathrm{tr}_B(U_{AE}\theta_A \otimes |e\rangle\langle e|_E U_{AE}^\dagger). \tag{2}$$

The channel Λ is said to be degradable if there exists a third CPTP channel W: $\mathcal{L}(\mathcal{H}_B) \to \mathcal{L}(\mathcal{H}_{E'})$ (dubbed degrading channel) such that $W \circ \Lambda = \tilde{\Lambda}$. Similarly, it is said to be antidegradable if instead there exists a CPTP channel V: $\mathcal{L}(\mathcal{H}_{E'}) \to \mathcal{L}(\mathcal{H}_B)$ such that $V \circ \tilde{\Lambda} = \Lambda$. Finally, we call N a degradable extension of Λ if N is degradable and there is a second channel R such that $R \circ N = \Lambda$.

The quantum capacity $Q(\Lambda)$ gives the highest rate at which quantum information can be transmitted over many uses of Λ. In this case from [32,38] we get $Q(\Lambda) = \lim_{n\to\infty} Q_n(\Lambda) = \lim_{n\to\infty} \max_{\rho\in\mathfrak{S}(\mathcal{H}_A^{\otimes n})}(1/n)J(\rho, \Lambda^{\otimes n})$, with $J(\rho, \Lambda^{\otimes n}) \coloneqq S[\Lambda^{\otimes n}(\rho)] - S[\tilde{\Lambda}^{\otimes n}(\rho)]$, and $S(\rho)$ being the Von Neumann entropy. Due to the no-cloning theorem [39], the quantum capacity of an antidegradable channel is zero. On the contrary, as anticipated in the introduction, for a degradable channel the regularization limit on n is not needed and the expression for $Q(\Lambda)$ reduces to the following single-letter formula:

$$Q(\Lambda) = Q_1(\Lambda) \coloneqq \max_{\rho\in\mathfrak{S}(\mathcal{H}_A)} J(\rho, \Lambda), \tag{3}$$

an identity which, while not having a simple physical interpretation, mathematically originates from the monotonicity of the relative entropy under CPTP transformations [37].

The FDC model.—In a standard approach to quantum communication the interaction between the quantum carriers of the information and their environment, the associated interaction time, as well as the state of environment are assumed to be known. However, it is possible to think about scenarios where the state of environment is changing in time and it can be monitored with quantum measurements. In this setting, suppose that with probability p_i the state of environment is the state σ_i, and that when this happens information carrier gets transformed by a given CPTP transformation Λ_i. If there was no other information except the probability distribution of the environment, the resulting channel would be just the weighted sum of each individual map, i.e., $\Lambda \coloneqq \sum_i p_i \Lambda_i$. Instead, we assume that in our case Bob collects a copy of the environment describing the channel as

$$\mathring{\Lambda}[\rho] \coloneqq \sum_i p_i \Lambda_i[\rho] \otimes \sigma_i, \tag{4}$$

where ρ is any input state and the σ_is live on an auxiliary space $\mathcal{H}_1$ on which Bob has complete access. More abstractly, this model can be also seen as a quantum channel with *quantum flags*, where with probability p_i the channel acts as Λ_i and Bob receives a quantum flag σ_i that encodes in a quantum state the information about which channel is acting. As Λ can be obtained from $\mathring{\Lambda}$ by simply tracing away the flags, it turns out that the capacities of the latter provide natural upper bounds for the corresponding ones of the former, i.e.,

$$Q(\Lambda) \leq Q(\mathring{\Lambda}). \tag{5}$$

A special example of a channel of the form (4) was considered in [19,21] where the σ_i were assumed to be orthogonal pure states. Here, on the contrary, we allow the σ_i to be mixed and not necessarily orthogonal and focus on the case where the resulting mapping has the form

$$\mathring{\Lambda}_p^d[\rho] = (1-p)\rho \otimes \sigma_0 + p\mathrm{Tr}[\rho]\frac{I^d}{d} \otimes \sigma_1. \tag{6}$$

This channel acts on a d dimensional Hilbert space and it can be expressed as in (4) with two components, the first associated with the identity channel and the second associated with a completely depolarizing transformation that replaces every input with the completely mixed state I^d/d. Notice however that Eq. (6) describes a proper CPTP mapping also for values of p larger than 1—indeed its Choi state [1,2] can be easily shown to be positive for any $p > 0$ such that $(1-p)\sigma_0 + (p/d^2)\sigma_1 \geq 0$. Most importantly, irrespectively from the value of σ_0 and σ_1, by removing the flag states from (6) via partial trace, $\overset{\circ}{\Lambda}{}_p^d$ reduces to a standard DC,

$$\Lambda_p^d[\rho] \coloneqq (1-p)\rho + p\mathrm{Tr}[\rho]\frac{I^d}{d}. \tag{7}$$

Therefore, invoking the monotonicity (5) we can upper bound the rather elusive quantum capacity of Λ_p^d, with the quantum capacity of $\overset{\circ}{\Lambda}{}_p^d$ that, as we shall see in the following section, is relatively easy to characterize.

The quantum capacity of FDC.—A fundamental ingredient in studying the capacities of $\overset{\circ}{\Lambda}{}_p^d$ is that such channel is covariant under the action of arbitrary unitary transformations U of SU(d), i.e., $\overset{\circ}{\Lambda}{}_p^d[U\rho U^\dagger] = (U\otimes I)\overset{\circ}{\Lambda}{}_p^d[\rho](U^\dagger \otimes I)$, the operators I being the identity on the flags. This implies that the output Von Neumann entropy associated with a generic pure input state is a constant quantity $t(p,d,\sigma_0,\sigma_1)$, which does not explicitly depend upon the specific value of $|\psi\rangle$, but only upon the parameters that characterize the map, i.e., $S(\overset{\circ}{\Lambda}{}_p^d[|\psi\rangle\langle\psi|]) = t(p,d,\sigma_0,\sigma_1)$. We restrict to the case where $\sigma_1 = |e_1\rangle\langle e_1|$ is a pure state, and σ_0 is diagonalizable in that basis, i.e., $\sigma_0 = c^2|e_1\rangle\langle e_1| + (1-c^2)|e_1^\perp\rangle\langle e_1^\perp|$. For this case both $\overset{\circ}{\Lambda}{}_p^d$ and its complementary counterpart can be parametrized by the fidelity between σ_0 and σ_1, i.e., via the parameter c [in particular we can write $\overset{\circ}{\Lambda}{}_{p,c}^d(\rho) \coloneqq (1-p)\rho\otimes[c^2|e_1\rangle\langle e_1| + (1-c^2)|e_1^\perp\rangle\langle e_1^\perp|] + p(I^d/d)\otimes |e_1\rangle\langle e_1|]$. In the Supplemental Material [40], using a simple measurement and action channel as a candidate for the degrading channel, we showed that $\overset{\circ}{\Lambda}{}_{p,c}^d$ is degradable for c fulfilling the inequality

$$c \leq c(p) \coloneqq \sqrt{(1-2p)/(2-2p)}. \tag{8}$$

In this regime due to Eq. (3) the quantum capacity of $\overset{\circ}{\Lambda}{}_{p,c}^d$ can be obtained by maximizing its single shot coherent information $J(\rho, \overset{\circ}{\Lambda}{}_{p,c}^d)$. While in general the maximum of such quantity does not allow for a close analytical expression, in our case the problem gets further simplified when putting together the covariance of $\overset{\circ}{\Lambda}{}_{p,c}^d$ and a side effect of degradability, i.e., the concavity of the functional $J(\rho, \overset{\circ}{\Lambda}{}_{p,c}^d)$ in the input state ρ [44]: these two facts imply that $J(\rho, \overset{\circ}{\Lambda}{}_{p,c}^d)$ gets its maximum on the completely mixed input state, i.e.,

$$\begin{aligned} Q(\overset{\circ}{\Lambda}{}_{p,c}^d) &= Q_1(\overset{\circ}{\Lambda}{}_{p,c}^d) = \max_\rho J(\rho, \overset{\circ}{\Lambda}{}_{p,c}^d) = J\left(\frac{I^d}{d}, \overset{\circ}{\Lambda}{}_{p,c}^d\right) \\ &= \log d + S[(1-p)\sigma_0 + p\sigma_1] - t(p,d^2,\sigma_0,\sigma_1). \end{aligned} \tag{9}$$

For the interested reader we point out that in the Supplemental Material [40] we also report other capacities of the FDC, specifically the entanglement assisted capacity and product state classical capacity.

Upper bounds for the DC quantum capacity.—According to Eq. (5), the quantum capacity of the DC Λ_p^d can be upper bounded by the capacity of $\overset{\circ}{\Lambda}{}_{p,c}^d$, irrespectively from the choice we make on the parameter c, as long as the degradability constraint (8) holds true. Intuitively however, as c gets larger, the bound gets better. To get the best upper bound for the quantum capacity of Λ_p^d we hence set $c = c(p)$. Accordingly, using the expression for $t(p,d^2,\sigma_0,\sigma_1)$ computed in the Supplemental Material [40], we can write

$$\begin{aligned} Q(\Lambda_p^d) \leq Q(\overset{\circ}{\Lambda}{}_{p,c(p)}^d) &= \log d + \eta\left(\frac{1}{2}\right) \\ &\quad - \eta\left(\frac{1}{2} - \frac{(d^2-1)p}{d^2}\right) - (d^2-1)\eta\left(\frac{p}{d^2}\right), \end{aligned} \tag{10}$$

where $\eta(z) \coloneqq -z\log z$ [an alternative inequality can be obtained by choosing the flag states to be pure: as discussed in the Supplemental Material [40] the resulting expression is however much more involved than (10) and a numerical check reveals that it is worse than the latter]. In order to test the quality of our findings we now proceed with a comparison with the limits previously proposed in the literature. We start considering first the low noise regime ($p \ll 1$) where (10) gives

$$\begin{aligned} Q(\overset{\circ}{\Lambda}{}_{p,c(p)}^d) &= \log d + \frac{d^2-1}{d^2}\left[\log\left(\frac{p}{d^2}\right) - \log e + 1\right]p \\ &\quad + O(p^2). \end{aligned} \tag{11}$$

It turns out that for $d = 2$, the above expression is less tight if compared with the numerical bounds given in Refs. [22,24] (see Fig. 1), and with the analytic bound of Ref. [23] which for this special regime implies

$$Q(\Lambda_p^2) \leq Q(\overset{\circ}{\Lambda}{}_{p,c(p)}^d) - \frac{3}{4}p + O(p^2 \log p). \tag{12}$$

Things however change when we move out from the $d = 2$, low noise regime for which to our knowledge the best

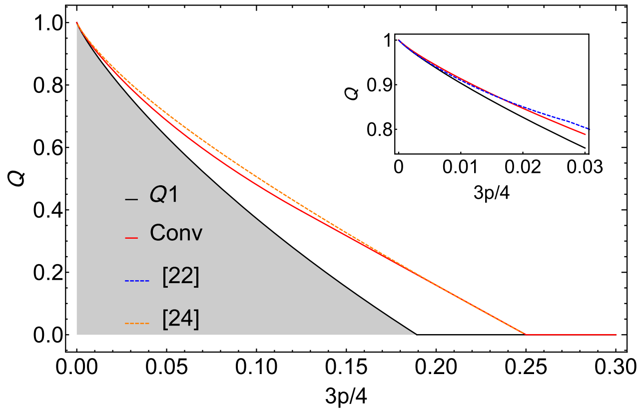

FIG. 1. Quantum capacity upper and lower bounds for $d = 2$: in this case it can be shown that $Q_1 = Q_{\mathrm{low}}$, the lower bound from Eq. (17), and the shaded region is excluded by the lower bound; Conv is the convex hull of all the upper bounds defined in Eq. (15), thus the allowed region for Q is the below Conv and above Q_1. Finally, the dashed lines represent the numerical upper bounds of Refs. [22,24].

performances up to date are provide by the results presented in Refs. [19,21]. The first one consists in the following inequality [21]

$$Q(\Lambda_p^d) \leq f_{1,d}(p) \coloneqq \eta\left(\frac{1+(d-1)\gamma}{d}\right) + (d-1)\eta\left(\frac{1-\gamma}{d}\right) - \eta\left(1-\frac{(d-1)\gamma}{d}\right) - (d-1)\eta\left(\frac{\gamma}{d}\right), \tag{13}$$

with $\gamma \coloneqq 2d/(d^2-1)\{\sqrt{1-p[(d^2-1)/d^2]} - (1-p[(d^2-1)/2])\}$. The second one instead relays on the fact that Λ_p^d is anti-degradable when $p = d/[2(d+1)]$ [19,21,45]; it implies that

$$Q(\Lambda_p^d) \leq f_{2,d}(p) \coloneqq \left(1 - \frac{2p(d+1)}{d}\right)\log d. \tag{14}$$

A direct comparison reveals that our inequality (10) beats both (13) and (14) in most of the parameter space, see e.g., Fig. 2 where we plot the relative functions for two values of d. We further notice that both $f_{1,d}(p)$ and $f_{2,d}(p)$, as well as our bound $Q(\overset{\circ}{\Lambda}{}^d_{p,c(p)})$, originate from degradable extensions of DCs. We can hence invoke the convexity of upper bounds obtained from degradable extensions [19], to derive the following improved inequality (see the Supplemental Material [40] for the detailed proof)

$$Q(\Lambda_p^d) \leq \mathrm{conv}\{Q(\overset{\circ}{\Lambda}{}^d_{p,c(p)}), f_{1,d}(p), f_{2,d}(p)\}, \tag{15}$$

where the convex hull $\mathrm{conv}\{g_1(p), g_2(p), \ldots\}$ is defined as the maximal convex function that is less than or equal to all the $g_i(p)$ s. Equation (15) is our ultimate result which, outside the special $d = 2$ low noise regime, clearly overcomes all the others results reported so far—see Fig. 2.

As a final observation we now focus on the asymptotic expansions of the various bounds for large d. Defining $\delta(p) \coloneqq \eta(\frac{1}{2}) - \eta(\frac{1}{2} - p) + \eta(1-p)$ from Eqs. (10), (13), and (14) we get

$$Q(\overset{\circ}{\Lambda}{}^d_{p,c(p)}) = (1-2p)\log d - h(p) + \delta(p) + \mathcal{O}\left(\frac{1}{\log d}\right),$$
$$f_{1,d}(p) = (1-2p)\log d + \mathcal{O}\left(\frac{\log d}{d}\right),$$
$$f_{2,d}(p) = (1-2p)\log d + \mathcal{O}\left(\frac{\log d}{d}\right), \tag{16}$$

with $h(p) \coloneqq -p\log p - (1-p)\log(1-p)$ the binary entropy functional [3]. Due to the fact that for $p < 1/2$

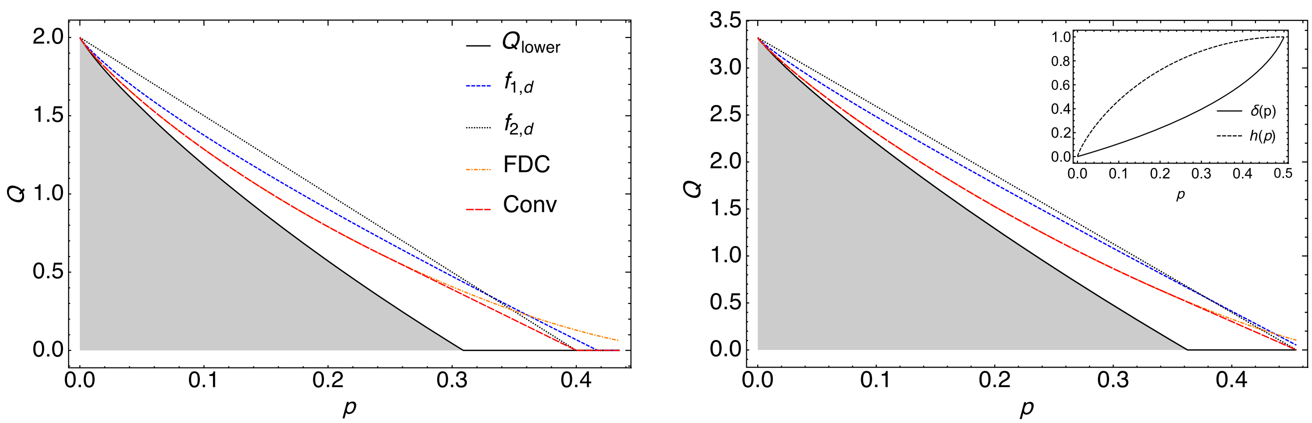

FIG. 2. Quantum capacity bounds for $d = 4$ (left) and $d = 10$ (right): Q_{low} is the lower bound from Eq. (17); $f_{1,d}$ and $f_{2,d}$ are the bounds of Refs. [19,21] [see Eqs. (13) and (14)]; FDC is the bound of Eq. (10); while finally Conv is the convex hull of all the other bounds defined in Eq. (15). The shaded region is excluded by the lower bound, thus the allowed region for Q is below Conv and above Q_1. Inset: comparison for the $\mathcal{O}(1)$ gaps for large d between the upper bounds and the lower bound Q_{low} as a function of p: for the bounds of Refs. [19,21] the gap is given by the binary entropy function $h(p)$, for ours it is instead given by the function $\delta(p)$ of Eq. (18).

(where the quantum capacity of the DC is not zero) one has $h(p) \geq \delta(p)$, Eq. (16) makes it clear that our bound is the only one that shows an $\mathcal{O}(1)$ term which is not zero (and negative)—see inset in the right panel of Fig. 2. A deeper insight on this can be gained by considering the lower bound of $Q(\Lambda_p^d)$ one gets by evaluating $J(\rho, \Lambda_p^d)$ on the completely mixed state, i.e.,

$$\begin{aligned} Q(\Lambda_p^d) \geq Q_{\mathrm{low}}(\Lambda_p^d) &\coloneqq J\left(\frac{I^d}{d}, \Lambda_p^d\right), \\ &= \log d - \eta\left(1 - p + \frac{p}{d^2}\right) - (d^2 - 1)\eta\left(\frac{p}{d^2}\right), \\ &= (1 - 2p)\log d - h(p) + \mathcal{O}\left(\frac{1}{\log d}\right). \end{aligned} \tag{17}$$

From Eq. (16) it then follows that the gap between our bound and $Q_{\mathrm{low}}(\Lambda_p^d)$ scales as

$$Q(\overset{\circ}{\Lambda}{}^d_{p,c(p)}) - Q_{\mathrm{low}}(\Lambda_p^d) = \delta(p) + \mathcal{O}\left(\frac{1}{\log d}\right), \tag{18}$$

while the differences between the other upper bounds and the lower bound exhibit a $\mathcal{O}(1)$ gap equal to $h(p)$ which as already noticed is always larger than $\delta(p)$ for the relevant values of p. In particular, it appears that our inequality gives a much better bound for low p, since $h(p)$ has derivative that diverges as $-\log p$ when $p \to 0$, while $\delta(p)$ scales linearly in p.

Discussion.—We introduced a specific flagged version of DC that, for a certain values of the parameter, is degradable. This allows us to compute an analytic bound for the quantum capacity of the original map. Our result works in any dimension, and it is the tightest available analytical upper bound. Unlike other degradable extensions of depolarizing channel [19,21], the introduced flags are not orthogonal. However, considering a general form for the flags and finding the degradability conditions is an open question. The idea we used is of general applicability and could give new good bounds for many other channels.

We thank Felix Leditzky, Andreas Winter, and Mark Wilde for helpful feedbacks. We acknowledge support by MIUR via Grant No. PRIN 2017 (Progetto di Ricerca di Intresse Nazionale): Project QUSHIP (Taming Complexity with Quantum Strategies: A Hybrid Integrated Photonics Approach) (Grant No. 2017SRN- BRK).

*marco.fanizza@sns.it
†farzad.kianvash@sns.it

[1] M. Wilde, *Quantum Information Theory* (Cambridge University Press, Berlin-Boston, 2013).
[2] A. S. Holevo, *Quantum Systems, Channels, Information: A Mathematical Introduction* (de Gruyter, New York, 2012).
[3] C. Shannon, Bell Syst. Tech. J. **27**, 379 (1948).
[4] M. Hastings, Nat. Phys. **5**, 255 (2009).
[5] P. W. Shor and J. A. Smolin, arXiv:quant-ph/9604006.
[6] D. P. DiVincenzo, P. W. Shor, and J. A. Smolin, Phys. Rev. A 1998) 830 ,**57**).
[7] G. Smith and J. A. Smolin, Phys. Rev. Lett. **98**, 030501 (2007).
[8] J. Fern and K. B. Whaley, Phys. Rev. A **78**, 062335 (2008).
[9] T. Cubitt, D. Elkouss, W. Matthews, M. Ozols, D. Pérez-García, and S. Strelchuk, Nat. Commun. **6**, 6739 (2015).
[10] F. Leditzky, D. Leung, and G. Smith, Phys. Rev. Lett. **121**, 160501 (2018).
[11] J. Bausch and F. Leditzky, arXiv:1910.00471.
[12] K. Li, A. Winter, X. B. Zou, and G. C. Guo, Phys. Rev. Lett. **103**, 120501 (2009).
[13] E. Y. Zhu, Q. Zhuang, and P. W. Shor, Phys. Rev. Lett. **119**, 040503 (2017).
[14] E. Y. Zhu, Q. Zhang, M-H Hsieh, and P. W. Shor, IEEE Trans. Inf. Theory **65**, 3973 (2018).
[15] G. Smith and J. Yard, Science **321**, 1812 (2008).
[16] G. Smith, J. Smolin, and J. Yard, Nat. Photonics **5**, 624 (2011).
[17] C. King, IEEE Trans. Inf. Theory **49**, 221 (2003).
[18] C. Adami and N. J. Cerf, Phys. Rev. A **56**, 3470 (1997).
[19] G. Smith and J. A. Smolin, IEEE Inf. Theory Workshop **54**, 4208 (2008).
[20] G. Smith, J. A. Smolin, and A. Winter, IEEE Trans. Inf. Theory **54**, 4208 (2008).
[21] Y. Ouyang, Quantum Inf. Comput. **14**, 917 (2014).
[22] D. Sutter, V. B. Scholz, A. Winter, and R. Renner, IEEE Trans. Inf. Theory **63**, 7832 (2017).
[23] F. Leditzky, D. Leung, and G. Smith, Phys. Rev. Lett. **120**, 160503 (2018).
[24] F. Leditzky, N. Datta, and G. Smith, IEEE Trans. Inf. Theory **64**, 4689 (2018).
[25] N. J. Cerf, Phys. Rev. Lett. **84**, 4497 (2000).
[26] E. M. Rains, Phys. Rev. A **60**, 179 (1999).
[27] E. M. Rains, IEEE Trans. Inf. Theory **47**, 2921 (2001).
[28] D. P. DiVincenzo, P. W. Shor, and J. A. Smolin, Phys. Rev. A **57**, 830 (1998).
[29] G. Smith and J. A. Smolin, Phys. Rev. Lett. **98**, 030501 (2007).
[30] J. Fern and K. B. Whaley, Phys. Rev. A **78**, 062335 (2008).
[31] S. Lloyd, Phys. Rev. A **55**, 1613 (1997).
[32] I. Devetak, IEEE Trans. Inf. Theory **51**, 44 (2005).
[33] M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Rev. A **60**, 1888 (1999).
[34] H. Barnum, E. Knill, and M. A. Nielsen, IEEE Trans. Inf. Theory **46**, 1317 (2000).
[35] C. H. Bennett, D. P. DiVincenzo, and J. A. Smolin, Phys. Rev. Lett. **78**, 3217 (1997).
[36] I. Devetak and P. W. Shor, Commun. Math. Phys. **256**, 287 (2005).
[37] T. S. Cubitt, M. B. Ruskai, and G. Smith, J. Math. Phys. (N.Y.) **49**, 102104 (2008).
[38] H. Barnum, E. Knill, and M. A. Nielsen, IEEE Trans. Inf. Theory **46**, 1317 (2000).
[39] W. K. Wootters and W. H. Zurek, Nature (London) **299**, 802 (1982).
[40] See the Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.125.020503 for the de-

tailed calculations of results of the main text, which includes Refs. [2,41–43].

[41] A. S. Holevo, Prob. Peredachi Inf. **9**, 3 (1973), http://www.mathnet.ru/php/archive.phtml?wshow=paper&jrnid=ppi&paperid=903&option_lang=eng.

[42] B. Schumacher and M. D. Westmoreland, Phys. Rev. A **56**, 131 (1997).

[43] C. H. Bennett, P. W. Shor. J. A. Smolin, and A. V. Thapliyal, IEEE Trans. Inf. Theory **48**, 2637 (2002).

[44] J. Yard, P. Hayden, and I. Devetak, IEEE Trans. Inf. Theory **54**, 3091 (2008).

[45] D. Bruss, D. P. DiVincenzo, A. Ekert, C. A. Fuchs, C. Macchiavello, and J. A. Smolin, Phys. Rev. A **57**, 2368 (1998).