# 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 $\sigma_0$ whenever the messages are transmitted without errors, and a copy of a state $\sigma_1$, 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 $\sigma_0$ and $\sigma_1$ 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 $\rho$, with probability $p$ Bob receives such a state together with an ancillary system prepared into the state $\sigma_0$, and, with probability $1 - p$, the completely mixed state together with the ancillary system in $\sigma_1$. The density matrices $\sigma_0$ and $\sigma_1$ 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 $p, \sigma_0, \sigma_1$ 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 $\Lambda$: $\mathcal{L}(\mathcal{H}_A)\to\mathcal{L}(\mathcal{H}_B)$ 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 $\Lambda$ 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 $\Lambda$ 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 $\Lambda$. 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 $\sigma_i$, and that when this happens information carrier gets transformed by a given CPTP transformation $\Lambda_i$. 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 $\rho$ is any input state and the $\sigma_i$s 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 $\Lambda_i$ and Bob receives a quantum flag $\sigma_i$ that encodes in a quantum state the information about which channel is acting. As $\Lambda$ 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.,

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

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

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

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 $\sigma_0$ and $\sigma_1$, by removing the flag states from (6) via partial trace, $\overset{\circ}{\Lambda}{}^d_p$ reduces to a standard DC,

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

Therefore, invoking the monotonicity (5) we can upper bound the rather elusive quantum capacity of $\Lambda_p^d$, with the quantum capacity of $\overset{\circ}{\Lambda}{}^d_p$ 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 $\overset{\circ}{\Lambda}{}^d_p$ is that such channel is covariant under the action of arbitrary unitary transformations $U$ of SU($d$), i.e., $\overset{\circ}{\Lambda}{}^d_p[U\rho U^\dagger] = (U \otimes I)\overset{\circ}{\Lambda}{}^d_p[\rho](U^\dagger \otimes I)$, 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., $S(\overset{\circ}{\Lambda}{}^d_p[|\psi\rangle\langle\psi|]) = t(p, d, \sigma_0, \sigma_1)$. We restrict to the case where $\sigma_1 = |e_1\rangle\langle e_1|$ is a pure state, and $\sigma_0$ 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 $\overset{\circ}{\Lambda}{}^d_p$ and its complementary counterpart can be parametrized by the fidelity between $\sigma_0$ and $\sigma_1$, i.e., via the parameter $c$ [in particular we can write $\overset{\circ}{\Lambda}{}^d_{p,c}(\rho) \coloneqq (1-p)\rho \otimes [c^2|e_1\rangle\langle e_1| + (1-c^2)|e_1^\perp\rangle\langle e_1^\perp|] + p(I^d/d) \otimes |e_1\rangle\langle e_1|$]. In the Supplemental Material [40], using a simple measurement and action channel as a candidate for the degrading channel, we showed that $\overset{\circ}{\Lambda}{}^d_{p,c}$ is degradable for $c$ fulfilling the inequality

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

In this regime due to Eq. (3) the quantum capacity of $\overset{\circ}{\Lambda}{}^d_{p,c}$ can be obtained by maximizing its single shot coherent information $J(\rho, \overset{\circ}{\Lambda}{}^d_{p,c})$. 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 $\overset{\circ}{\Lambda}{}^d_{p,c}$ and a side effect of degradability, i.e., the concavity of the functional $J(\rho, \overset{\circ}{\Lambda}{}^d_{p,c})$ in the input state $\rho$ [44]: these two facts imply that $J(\rho, \overset{\circ}{\Lambda}{}^d_{p,c})$ gets its maximum on the completely mixed input state, i.e.,

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

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 $\Lambda_p^d$ can be upper bounded by the capacity of $\overset{\circ}{\Lambda}{}^d_{p,c}$, 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 $\Lambda_p^d$ 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

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

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

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

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

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

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

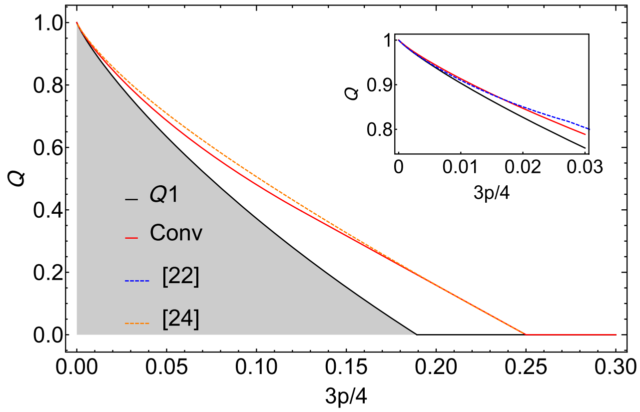


FIG. 1. Quantum capacity upper and lower bounds for $d = 2$: in this case it can be shown that $Q_1 = Q_{\rm low}$, the lower bound from Eq. (17), and the shaded region is excluded by the lower bound; Conv is the convex hull of all the upper bounds defined in Eq. (15), thus the allowed region for $Q$ 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 $\Lambda_p^d$ 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

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

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

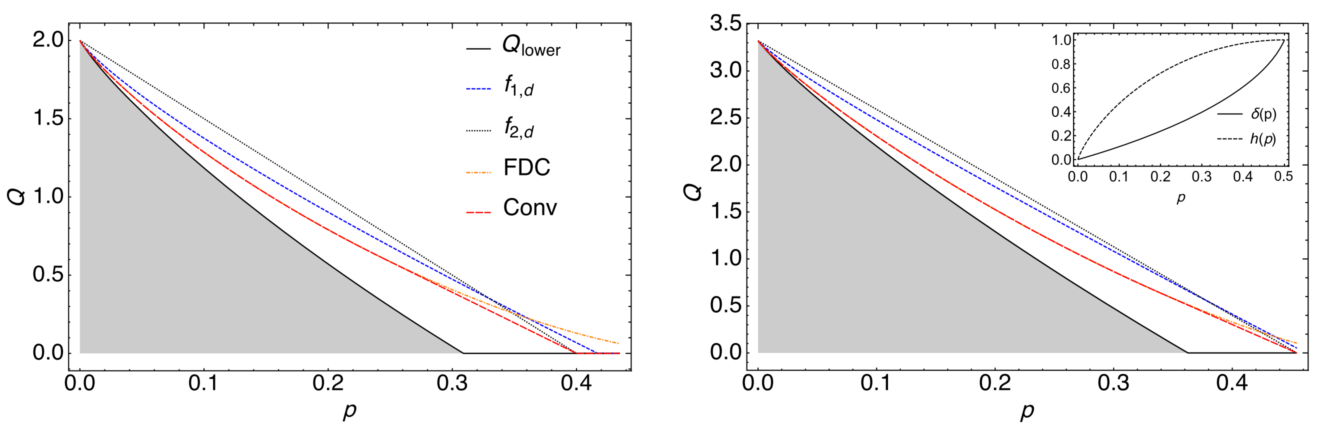


FIG. 2. Quantum capacity bounds for $d = 4$ (left) and $d = 10$ (right): $Q_{\rm low}$ is the lower bound from Eq. (17); $f_{1,d}$ and $f_{2,d}$ are the bounds of Refs. [19,21] [see Eqs. (13) and (14)]; FDC is the bound of Eq. (10); while finally Conv is the convex hull of all the other bounds defined 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 $Q_{\rm low}$ 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).

# SUPPLEMENTAL MATERIAL

## Contents



### Stinespring representation and complementary channel of FDC

In this section we find the Stinespring representation of FDC and its complementary channel according to the definition in Eq. (1) of the main text. FDC can be seen as a quantum channel that with probability $p$ replaces the input state with the maximally mixed state i.e. $\frac{I}{d}$ and adds a flag $\sigma_1$, and with probability $1-p$ acts as the identity channel on the input and adds a flag $\sigma_0$. With this intuition in mind, we can write the environment Hilbert space as $\mathcal{H}_E = \mathcal{H}_F \otimes \mathcal{H}_{\bar{F}} \otimes \mathcal{H}_I \otimes \mathcal{H}_{\bar{I}} \otimes \mathcal{H}_C$. The purifications of flags live in $\mathcal{H}_F \otimes \mathcal{H}_{\bar{F}}$, and the purification of the maximally mixed state i.e. a maximally entangled state live in $\mathcal{H}_I \otimes \mathcal{H}_{\bar{I}}$, and a control qubit lives in $\mathcal{H}_C$. The Stinespring unitary related to FDC is a controlled swap with flags as following

$$U_{AE} \left|\psi\right>_A \left|0\right>_F \left|0\right>_{\bar{F}} \left|\Phi^d\right>_{I,\bar{I}} \left|a\right>_C = \sqrt{1-p}\left|\psi\right>_A |\sigma_0\rangle\!\rangle_{F,\bar{F}} \left|\Phi^d\right>_{I,\bar{I}} \left|0\right>_C + \sqrt{p}\left|\Phi^d\right>_{A,\bar{I}} |\sigma_1\rangle\!\rangle_{F,\bar{F}} \left|\psi\right>_I \left|1\right>_C \ , \tag{1}$$

where $\left|a\right>_C = \sqrt{1-p}\left|0\right>_C + \sqrt{p}\left|1\right>_C$ is a control qubit, and $\left|\Phi^d\right>$ is a $d$-dimensional maximally entangled state – see Fig. 1 for a schematic representation of the quantum circuit that implement the transformation. To prove that (1) provides a Stinespring representation of FDC we first notice that it defines an isometry, since it preserves the scalar product: indeed introducing the compact notation $\left|\psi, e\right>_{AE} := \left|\psi\right>_A \left|0\right>_F \left|0\right>_{\bar{F}} \left|\Phi^d\right>_{I,\bar{I}} \left|a\right>_C$ we have

$${}_{AE}\left<\phi, e\right| U_{AE}^\dagger U_{AE} \left|\psi, e\right>_{AE} = (1-p)\ {}_A\langle\phi|\psi\rangle_A + p\ {}_I\langle\phi|\psi\rangle_I = {}_A\langle\phi|\psi\rangle_A = {}_{AE}\langle\phi, e|\psi, e\rangle_{AE} \ . \tag{2}$$

Next we notice that by tracing over subsystems $\bar{F}, I, \bar{I}, C$ we get

$$\mathrm{tr}^{(A,F)}\left[U_{AE}\left|\psi, e\right>_{AE}\left<\psi, e\right| U_{AE}^\dagger\right] = (1-p)\left|\psi\right>_A\left<\psi\right| \otimes \sigma_0 + p\frac{I^d}{d} \otimes \sigma_1 = \mathring{\Lambda}_p^d(\left|\psi\right>_A\left<\psi\right|) \ , \tag{3}$$

for all possible input state $\left|\psi\right>_A$ (here $\mathrm{tr}^{(A,F)}$ indicates that we are taking the partial trace with respect to all degree of freedom of the system but $A, F$).

From the above definition we now give an expression for the complementary channel of the family of FDC $\mathring{\Lambda}_{p,c}^d$ defined in the main text and parametrized by $c$. The complementary channel of $\mathring{\Lambda}_{p,c}^d$ satisfies

$$\tilde{\mathring{\Lambda}}_{p,c}^d[\rho] = \mathrm{tr}_{A,F}\left[U_{AE}(\rho \otimes \left|e\right>\left<e\right|_E)U_{AE}^\dagger\right] , \tag{4}$$

for all input states $\rho$ (here $\mathrm{tr}_{A,F}$ indicates that the trace is taken on $A$ and $F$). Without loss of generality we can always focus on pure input states. Under this condition the right side of the previous expression yields

$$\begin{aligned} U_{AE}\left(|\psi\rangle_A\langle\psi| \otimes \left|e\right>\left<e\right|_E\right)U_{AE}^\dagger =& (1-p)\left|\psi\right>\left<\psi\right|_A \otimes |\sigma_0\rangle\!\rangle\langle\!\langle\sigma_0|_{F,\bar{F}} \otimes \left|\Phi^d\right>\left<\Phi^d\right|_{I,\bar{I}} \otimes \left|0\right>\left<0\right|_C \\ &+ p\left|\Phi^d\right>\left<\Phi^d\right|_{A,\bar{I}} \otimes \left|e_1\right>\left<e_1\right|_F \otimes \left|e_1\right>\left<e_1\right|_{\bar{F}} \otimes \left|\psi\right>\left<\psi\right|_I \otimes \left|1\right>\left<1\right|_C \\ &+ \sqrt{p(1-p)}\left|\psi\right>_A\left|\Phi^d\right>_{I,\bar{I}}\left<\Phi^d\right|_{A,\bar{I}}\left<\psi\right|_I \otimes |\sigma_0\rangle\!\rangle_{F,\bar{F}}\left<e_1\right|_F\left<e_1\right|_{\bar{F}} \otimes \left|0\right>\left<1\right|_C + h.c. \end{aligned} \tag{5}$$

Given that $|\sigma_0\rangle\!\rangle_{F,\bar{F}} = c\left|e_1\right>\left|e_1\right>_{F,\bar{F}} + \sqrt{1-c^2}\left|e_1^\perp\right>\left|e_1^\perp\right>_{F,\bar{F}}$ we can take trace over $A, F$ and get

$$\begin{aligned} \tilde{\mathring{\Lambda}}_{p,c}^d(\left|\psi\right>\left<\psi\right|) =& (1-p)\sigma_{0\bar{F}} \otimes \left|\Phi^d\right>\left<\Phi^d\right|_{I,\bar{I}} \otimes \left|0\right>\left<0\right|_C + p|e_1\rangle\left<e_1\right|_{\bar{F}} \otimes \left|\psi\right>\left<\psi\right|_I \otimes \frac{I_{\bar{I}}^d}{d} \otimes \left|1\right>\left<1\right|_C \\ &+ \sqrt{p(1-p)}\left[c\ \mathrm{tr}_A(\left|\psi\right>_A\left|\Phi^d\right>_{I,\bar{I}}\left|0\right>_C\left<\Phi^d\right|_{A,\bar{I}}\left<\psi\right|_I\left<1\right|_C) + h.c.\right] \otimes \left|e_1\right>\left<e_1\right|_{\bar{F}} \ . \end{aligned} \tag{6}$$

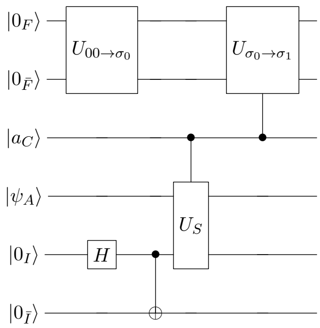


FIG. 1: Schematic of the the quantum circuit implementing the unitary transformation $U_{AE}$ of Eq. (1). In the first step the unitary operator $U_{00\to\sigma_0}$ is applied on the flags. This unitary prepares the state $|\sigma_0\rangle\rangle_{F,\bar F}$ on the flags. In addition, we prepare a maximally entangled state in subspace $\mathcal{H}_I\otimes\mathcal{H}_{\bar I}$ using Hadamard and CNOT gates. Then controlled swap gate $U_S$ acts on $\mathcal{H}_C\otimes\mathcal{H}_A\otimes\mathcal{H}_I$. If the state of the controlled qubit is $|0\rangle$ the controlled swap acts as identity, instead if the state is $|1\rangle$, it swaps $\mathcal{H}_A$ with $\mathcal{H}_I$. In the last step another controlled unitary is performed on the flags.

### Sufficient conditions for degradability for FDC with mixed flags

In this section we analyse the degradability of $\mathring{\Lambda}^d_{p,c}$ and establish the following lemma:

**Lemma.** *$\mathring{\Lambda}^d_{p,c}$ is degradable if $c\le\sqrt{\frac{1-2p}{2-2p}}$*

*Proof.* We look for the existence of a degrading CPTP channel $W_{p,c}$ connecting $\mathring{\Lambda}^d_{p,c}$ and $\tilde{\mathring{\Lambda}}^d_{p,c}$, i.e. satisfying the condition $W_{p,c}\circ\mathring{\Lambda}^d_{p,c}=\tilde{\mathring{\Lambda}}^d_{p,c}$ or explicitly

$$(1-p)W_{p,c}(\rho_A\otimes\sigma_{0F})+\tfrac{p}{d}W_{p,c}(I^d_A\otimes|e_1\rangle\langle e_1|_F)=\tilde{\mathring{\Lambda}}^d_{p,c}(\rho)\,. \tag{7}$$

As a suitable candidate for $W_{p,c}$ we consider a two-step process which first performs a measurement on flags then triggers an action on $A$. Specifically for the measurement we assume an orthogonal projection in the basis $|e_1\rangle$ and $|e_1^\perp\rangle$. For the action on $A$ instead we assume that if the measurement outcome is $|e_1\rangle$ we will prepare whatever state was left on $A$ into the fixed state $|e_1^\perp\rangle\langle e_1^\perp|_{\bar F}\otimes|\Phi^d\rangle\langle\Phi^d|_{I,\bar I}\otimes|0\rangle\langle 0|_C$; on the contrary, if the result is $|e_1^\perp\rangle$ we operate on $A$ with a channel of the form $\tilde{\mathring{\Lambda}}^d_{q,c'}$ with properly selected parameters $q,c'$. With this choice, the resulting mapping $W_{p,c}$ on $\rho_{A,F}$ is hence given by

$$W_{p,c}(\rho_{A,F}):=\langle e_1|\operatorname{tr}_A(\rho_{A,F})|e_1\rangle\,|e_1^\perp\rangle\langle e_1^\perp|_{\bar F}\otimes|\Phi^d\rangle\langle\Phi^d|_{I,\bar I}\otimes|0\rangle\langle 0|_C+\langle e_1^\perp|\operatorname{tr}_A(\rho_{A,F})|e_1^\perp\rangle\,\tilde{\mathring{\Lambda}}^d_{q,c'}(\operatorname{tr}_F(\rho_{A,F}))\,. \tag{8}$$

With this choice the condition (7) becomes

$$(1-p)\big[c^2|e_1^\perp\rangle\langle e_1^\perp|_{\bar F}\otimes|\Phi^d\rangle\langle\Phi^d|_{I,\bar I}\otimes|0\rangle\langle 0|_C+(1-c^2)\tilde{\mathring{\Lambda}}^d_{q,c'}(\rho_A)\big]+p|e_1^\perp\rangle\langle e_1^\perp|_{\bar F}\otimes|\Phi^d\rangle\langle\Phi^d|_{I,\bar I}\otimes|0\rangle\langle 0|_C=\tilde{\mathring{\Lambda}}^d_{p,c}(\rho_A)\,. \tag{9}$$

which can be satisfied if it is possible to find $q,c'\in[0,1]$ such that

$$q=\tfrac{p}{(1-p)(1-c^2)}\,,\qquad c'^2=\tfrac{c^2(1-p)}{1-2p-c^2+pc^2}\,. \tag{10}$$

Implying that $0\le q,c'\le 1$ and doing some simple algebra we get

$$c\le\sqrt{\tfrac{1-2p}{2-2p}}\,. \tag{11}$$

Under this condition the channel $\mathring{\Lambda}^d_{p,c}$ is degradable. $\square$

## FDC with pure flags

In this section we find a sufficient condition for degradability for FDC with pure flags analogous to the case of $\mathring{\Lambda}^{d}_{p,c}[\rho]$, and we use it to obtain another upper bound on the quantum capacity of the depolarizing channel. In this scenario the channel explicitly writes as

$$\mathring{\Lambda}'^{d}_{p,c}[\rho] = (1-p)\rho \otimes |e_0\rangle\langle e_0| + p\tfrac{I^d}{d} \otimes |e_1\rangle\langle e_1| \ , \tag{12}$$

where the parameter $c$ refers now to the overlap $c := \langle e_1|e_0\rangle$.

Note that the phase in $c$ is not important in studying the degradability of $\Lambda'^{d}_{p,c}$ since it can be set to zero by acting with a unitary transformation after the action of the channel (12): accordingly in the following we shall assume $c$ to be real without loss of generality.

To find complementary channel $\mathring{\Lambda}'^{d}_{p,c}$ we should first write the Stinespring form of this transformation. The Hilbert space of the environment is decomposed as $\mathcal{H}_E = \mathcal{H}_F \otimes \mathcal{H}_I \otimes \mathcal{H}_{\bar{I}} \otimes \mathcal{H}_C$. It is similar to the previous case, but we do not need $\mathcal{H}_{\bar{F}}$ because the flags are pure. The unitary interaction between system and environment acts as following

$$U'_{AE}\,|\psi\rangle_A |0\rangle_F \left|\Phi^d\right\rangle_{I,\bar{I}} |0\rangle_C = \sqrt{1-p}\,|\psi\rangle_A |e_0\rangle_F \left|\Phi^d\right\rangle_{I,\bar{I}} |0\rangle_C + \sqrt{p}\left|\Phi^d\right\rangle_{A,\bar{I}} |e_1\rangle_F |\psi\rangle_I |1\rangle_C \ , \tag{13}$$

where $|0\rangle, |1\rangle$ are two orthogonal states, $\left|\Phi^d\right\rangle$ is a maximally entangled states in dimension $d$, and the trace in Eq. (1) of the main text here is on states $I, \bar{I}, C$. Doing simple calculation we can show that this is a Stinespring representation of (12). To find the complementary channel instead of taking trace over states $I, \bar{I}, C$ we should take trace over states $A, F$. Carrying out the calculation we get

$$\begin{aligned}\tilde{\mathring{\Lambda}}'^{d}_{p,c}(|\psi\rangle\langle\psi|) =& (1-p)\left|\Phi^d\right\rangle\left\langle\Phi^d\right|_{I,\bar{I}} \otimes |0\rangle\langle 0|_C + p\,|\psi\rangle\langle\psi|_I \otimes \tfrac{I^d_{\bar{I}}}{2} \otimes |1\rangle\langle 1|_C \\ &+ \sqrt{p(1-p)}\big[c\ \mathrm{tr}_A(|\psi\rangle_A \left|\Phi^d\right\rangle_{I,\bar{I}} |0\rangle_C \left\langle\Phi^d\right|_{A,\bar{I}} \langle\psi|_I \langle 1|_C) + h.c.\big] \ .\end{aligned} \tag{14}$$

Noting that the form of $\tilde{\mathring{\Lambda}}'^{d}_{p,c}$ is exactly the same as $\tilde{\mathring{\Lambda}}^{d}_{p,c}$, and doing some algebra one can find that the regime where $\mathring{\Lambda}'^{d}_{p,c}$ is degradable is

$$c \leq \sqrt{\tfrac{1-2p}{2-2p}} \ . \tag{15}$$

In this regime the quantum capacity of $\mathring{\Lambda}'^{d}_{p,c}$ can be computed as in Eq. (10) of the main text, i.e.

$$Q(\mathring{\Lambda}'^{d}_{p,c}) = \log d + S\big((1-p)\,|e_0\rangle\langle e_0| + p\,|e_1\rangle\langle e_1|\big) - t(p, d^2, |e_0\rangle, |e_1\rangle) \ , \tag{16}$$

which after some algebra can be cast into the expression

$$\begin{aligned}Q(\mathring{\Lambda}'^{d}_{p,c}) =& \log d + \eta[\tfrac{1}{2}\left(1 - \sqrt{-2(p-1)p\cos(\theta) + 2(p-1)p + 1}\right)] + \eta[\tfrac{1}{2}\left(1 + \sqrt{-2(p-1)p\cos(\theta) + 2(p-1)p + 1}\right)] \\ &- \eta[\tfrac{d^2(-p)+d^2-\sqrt{d^4p^2-2d^4p+d^4-2d^2p^2\cos(\theta)+2d^2p\cos(\theta)+p^2}+p}{2d^2}] \\ &- \eta[\tfrac{d^2(-p)+d^2+\sqrt{d^4p^2-2d^4p+d^4-2d^2p^2\cos(\theta)+2d^2p\cos(\theta)+p^2}+p}{2d^2}] + \tfrac{p(d^2-1)}{d^2}\log\big(\tfrac{p}{d^2}\big) \ ,\end{aligned} \tag{17}$$

where $\cos(\theta) = 2c^2 - 1$ and $\eta(z) := -z\log(z)$.

## Combining different bounds from degradable extensions

In this section we present one of the results in Ref. [1]. We call $N$ a degradable extension of $\Lambda$ if $N$ is degradable and there is a second channel $R$ such that $R \circ N = \Lambda$. In Ref. [1] it has been shown that if $N_0$ is a degradable extension of $\Lambda_0$ and $N_1$ is a degradable extension of $\Lambda_1$ then $N = \lambda N_0 \otimes |0\rangle\langle 0| + (1-\lambda)N_1 \otimes |1\rangle\langle 1|$ is a degradable extension of $\Lambda = \lambda\Lambda_0 + (1-\lambda)\Lambda_1$ for every $0 \leq \lambda \leq 1$, and the quantum capacities satisfy the following relation

$$Q(\Lambda) \leq Q_1(N) \leq \lambda Q_1(N_0) + (1-\lambda)Q_1(N_1) \ . \tag{18}$$

This theorem can be used to show if we have upper bounds for the quantum capacity of two channels, all obtained from degradable extensions, the convex combination of the bounds is also an upper bound for the respective convex combination of the channels. We clarify this with an example: Consider the depolarizing channel i.e. $\Lambda_p^d[\rho] = (1-p)\rho + p\,\mathrm{Tr}[\rho]\frac{I^d}{d}$. The set of all values of $p$ for which $\Lambda_p^d$ is a CPTP is $P$, and $N_p$ is a degradable extension of $\Lambda_p^d$ for all $p \in P$. If $p_0, p_1 \in P$, then $N_{p_0}, N_{p_1}$ are degradable extensions of $\Lambda_{p_0}^d, \Lambda_{p_1}^d$ respectively, then

$$Q(\Lambda_{p_0}^d) \leq g(p_0) := Q_1(N_{p_0}), \quad Q(\Lambda_{p_1}^d) \leq g(p_1) := Q_1(N_{p_1}) . \tag{19}$$

Therefore

$$N = \lambda N_{p_0} \otimes |0\rangle\langle 0| + (1-\lambda) N_{p_1} \otimes |1\rangle\langle 1| , \tag{20}$$

is a degradable extension of $\Lambda^d_{\lambda p_0 + (1-\lambda)p_1}$, then using (18) we get $Q(\Lambda^d_{\lambda p_0+(1-\lambda)p_1}) \leq \lambda g(p_0) + (1-\lambda) g(p_1)$. As this holds for all $p_0, p_1 \in P$, $\mathrm{conv}\{g(p)\}$ is also an upper bound for the quantum capacity of $\Lambda_p^d$, where

$$\mathrm{conv}\{g(p)\} := \inf_{\substack{p_0,p_1\in P,\\ 0\leq\lambda\leq 1}} \{\lambda g(p_0) + (1-\lambda) g(p_1) : p = \lambda p_0 + (1-\lambda) p_1\} .$$

In particular, given $g_1(p), ..., g_n(p)$, all upper bounds for the quantum capacity of depolarizing channel all derived from degradable extensions, then $g_{\min}(p) := \min\{g_1(p), ..., g_n(p)\}$ is also an upper bound and therefore $\mathrm{conv}\{g_1(p), ..., g_n(p)\} := \mathrm{conv}\{g_{\min}(p)\}$, is an upper bound too.

### Output entropy of the FDC channel $t(p, d, \sigma_0, \sigma_1)$

The fact that $\mathring{\Lambda}_p^d$ of Eq. (6) of the main text is covariant under $SU(d)$ implies that its output von Neumann entropy associated with a generic input state is a constant $t(p, d, \sigma_0, \sigma_1)$ that explicitly does not depend upon the specific value of $|\psi\rangle$ but only upon the parameters that characterize the map, i.e. $p$, $\sigma_0$, $\sigma_1$ and $d$. Some simple algebra lets us to explicitly determine the value of $t(p, d, \sigma_0, \sigma_1)$ obtaining

$$\begin{aligned} t(p,d,\sigma_0,\sigma_1) &:= S(\mathring{\Lambda}_p^d[|\psi\rangle\langle\psi|]) = S\left((1-p)|\psi\rangle\langle\psi| \otimes \sigma_0 + p\tfrac{I^d}{d} \otimes \sigma_1\right) \\ &= h\left(\tfrac{d(1-p)+p}{d}\right) + \tfrac{p(d-1)}{d}\log(d-1) + \tfrac{d(1-p)+p}{d} S\left(\frac{(1-p)\sigma_0 + \frac{p}{d}\sigma_1}{\frac{d(1-p)+p}{d}}\right) + \tfrac{p(d-1)}{d} S(\sigma_1) , \end{aligned} \tag{21}$$

where $h(x) := -x\log x - (1-x)\log(1-x)$ is the binary entropy.

### Classical capacities of FDC

In this section first we discuss the evaluation of the classical capacity and the entanglement assisted capacity of FDC. The classical capacity $C(\Lambda)$ of $\Lambda$ is the highest achievable rate at which classical data can be faithfully transmitted through such channel. Following [2, 3] it can be computed as

$$C(\Lambda) = \lim_{n\to\infty} C_n(\Lambda) = \lim_{n\to\infty} \tfrac{1}{n}\chi(\Lambda^{\otimes n}) , \tag{22}$$

with $C_n(\Lambda) = \chi(\Lambda) := \max_{\{p_i;\rho_i\}} \chi(\{p_i; \Lambda(\rho_i)\})$ where the Holevo quantity of an ensemble is defined as $\chi(\{p_i;\rho_i\}) := S(\sum_i p_i\rho_i) - \sum_i p_i S(\rho_i)$, and $C_n(\Lambda)$ is the rate achievable with codewords having entanglement across at most $n$ uses of the channel. Similarly the entanglement assisted classical capacity $C_E(\Lambda)$ measures the highest rate at which the classical information can be transmitted through $\Lambda$ when Alice and Bob share unlimited resource of entanglement. From [4] it follows that

$$C_E(\Lambda) = \max_\rho I(\rho, \Lambda) , \tag{23}$$

where the mutual information is defined as $I(\rho,\Lambda) := S(\rho) + S(\Lambda(\rho)) - S(\tilde{\Lambda}(\rho))$ .

By concavity of the Von Neumann entropy, it follows that

$$\min_\rho S(\mathring{\Lambda}_p^d[\rho]) = t(p, d, \sigma_0, \sigma_1) . \tag{24}$$

Using above observation we compute the Holevo capacity of the map $C_1(\mathring{\Lambda}_p^d)$. Notice that for any ensemble $\{p_i;\rho_i\}$, one can create a larger ensemble $\{p_i, dU; U\rho_i U^\dagger\}$, where the state $U\rho_i U^\dagger$ is extracted with probability density $p_i dU$, where $dU$ is the Haar measure of $SU(d)$. By the concavity of Von Neumann entropy we can write

$$\chi(\{p_i;\mathring{\Lambda}_p^d[\rho_i]\}) \leq \chi(\{p_i, dU;\mathring{\Lambda}_p^d[U\rho_i U^\dagger]\}) = \log d + S((1-p)\sigma_0 + p\sigma_1) - \sum_i p_i S(\mathring{\Lambda}_p^d[\rho_i])\;, \tag{25}$$

where we used the depolarizing identity $\int dU\; U\rho U^\dagger = \frac{I^d}{d}$.

We can now invoke (24) to put an upper bound on $\chi(\{p_i, dU;\mathring{\Lambda}_p^d[U\rho_i U^\dagger]\})$ by replacing all the $S(\mathring{\Lambda}_p^d[\rho_i])$ terms with the constant $t(p,d,\sigma_0,\sigma_1)$. The resulting quantity no longer depends on the input of the channel and provide an achievable maximum for the Holevo information of the channel yielding the identity

$$C_1(\mathring{\Lambda}_p^d) = \log d + S((1-p)\sigma_0 + p\sigma_1) - t(p,d,\sigma_0,\sigma_1), \tag{26}$$

(the achievability being granted e.g. by ensembles of the form $\{dU; U|\psi\rangle\langle\psi|U^\dagger\}$, with $|\psi\rangle$ arbitrarily chosen). However, we were not able to find $C_n(\Lambda)$ for $n>1$ and the classical capacity of FDC remains open.

To compute the entanglement assisted capacity of $\mathring{\Lambda}_p^d$, we use the fact that the quantum mutual information of a channel is concave in $\rho$ [5]. Exploiting this and the covariance of $\mathring{\Lambda}_p^d$ under $SU(d)$ we can then write

$$I\left(\tfrac{I^d}{d},\mathring{\Lambda}_p^d\right) = I\left(\int dU\; U\rho U^\dagger,\mathring{\Lambda}_p^d\right) \geq \int dU\; I\left(U\rho U^\dagger,\mathring{\Lambda}_p^d\right) = I(\rho,\mathring{\Lambda}_p^d)\;. \tag{27}$$

Therefore, we can conclude that the state that maximizes the quantum mutual information is $\frac{I}{d}$ and after some algebra we get

$$C_E(\mathring{\Lambda}_p^d) \;=\; I\left(\tfrac{I^d}{d},\mathring{\Lambda}_p^d\right) = 2\log d + S\big((1-p)\sigma_0 + p\sigma_1\big) - t(p,d^2,\sigma_0,\sigma_1). \tag{28}$$


[1] G. Smith and J. A. Smolin, IEEE *Information Theory Workshop*, vol. **54**, 4208 (2008).
[2] A. S. Holevo, Problemy Peredachi Informatsii **9**, 3 (1973).
[3] B. Schumacher and M. D. Westmoreland, Phys. Rev. A **56**, 131 (1997).
[4] C. H. Bennett, *et al.* IEEE Trans. Inf. Th. **48**, 2637 (2002).
[5] A. S. Holevo, *Quantum systems, channels, information: a mathematical introduction* (de Gruyter, 2012).