Journal of Physics Communications

OPEN ACCESS

RECEIVED
2 April 2024

REVISED
1 August 2024

ACCEPTED FOR PUBLICATION
14 August 2024

PUBLISHED
27 August 2024

PAPER

The *n*-shot classical capacity of the quantum erasure channel

Matteo Rosati

Dipartimento di Ingegneria Civile, Informatica e delle Tecnologie Aeronautiche, Universitá degli Studi Roma Tre, Via Vito Volterra 62, Roma, I-00146, Italy

E-mail: matteo.rosati@uniroma3.it

Keywords: quantum channel, classical capacity, *n*-shot capacity, finite-size capacity

Abstract

We compute the *n*-shot classical capacity of the quantum erasure channel, providing upper bounds and almost-matching lower bounds for it, the latter achievable via large-minimum-distance classical linear codes for any *n*. The protocols are in full product form, i.e. no entanglement is needed either at the encoder or decoder to attain the capacity, and they explicitly adapt to the transition between different error regimes as the erasure probability increases. Finally, we show that our upper and lower bounds on the capacity are tighter than those obtainable from the general theory of finite-size capacity via generalized divergences.

1. Introduction

The information transmission capabilities of quantum channels have been studied since the birth of quantum information theory [1], with the motivation that information carriers are ultimately described by quantum-mechanical laws and hence might exhibit counter-intuitive effects with respect to classical channels. Among such effects is the possibility of using entanglement between different channel uses, either at the transmitter or at the receiver, to enhance the bit-transmission rate, achieving the celebrated Holevo capacity [1, 2]. Interestingly, even those quantum channels for which entanglement at the encoder is not needed to achieve the Holevo capacity, exhibit a phenomenon of super-additivity at the receiver end [3, 4]. Indeed, the decoder has to perform a collective, non-product and thus potentially entangling, measurement on the joint Hilbert space of the entire received quantum codeword, e.g. called pretty-good or square-root measurement (PGM) [5, 6], for which an explicit design remains an open problem to date [7–16].

In the last decade, the attention has shifted to the quantification of non-asymptotic information-processing tasks, with the hope of gathering a better understanding of the problem in the finite-size regime [17, 18]. For the finite-size classical capacity, i.e. the maximum bit-transmission rate attainable by using a quantum channel a finite number of times, upper and lower bounds have been derived by Wang and Renner [19] in terms of generalized divergences (GD). These bounds are valid for any number *n* of finite uses of the channel, hence they are sometimes referred to as *n*-shot capacity.

While GD bounds offer a general method to bound the *n*-shot capacity, and approach the Holevo capacity for large *n*, they leave ample margin for improvement, at least in two respects: (i) even with the simplification of a finite number of uses of the channel, *n*-shot bounds based on GD still rely on a communication protocol that makes use of the abstract PGM; (ii) upper- and lower-bounds obtained via GD are asymptotically matching, i.e. they differ only by channel-independent terms that tend to zero for large *n*, however their calculation can be difficult and can exhibit a significant gap for finite *n*, effectively allowing a wide range of possible values for the *n*-shot capacity.

In this paper we introduce explicit protocols for the transmission of bits on the *d*-dimensional quantum erasure channel, obtaining narrow bounds for its *n*-shot classical capacity, which are not based on GD.

Firstly, our method provides significantly tighter bounds than divergence-based methods, showing, for the first time to our knowledge, that the literature bounds on the finite-size classical capacity of a quantum channel can be determined significantly more precisely than previously expected.

Secondly, the tighter bounds we obtain are based on a communication protocol tailored for the channel under study, and they show that sharp transitions in the capacity's behaviour happen when varying the erasure probability. This behaviour is not apparent when using divergence-based bounds and it is caused by sharp changes in the capacity-achieving protocol, which needs to protect against an increasing number of errors as the erasure probability increases.

Thirdly, using the channel's simple structure, we show that entanglement is not needed at the encoder nor at the decoder, a fact previously known only for the asymptotic channel capacity [20]. Indeed, our protocol shows that the n-shot classical capacity of the erasure channel can be attained without the PGM, but rather using a simple product measurement, which is amenable to practical realization.

The article is structured as follows: in section 2 we review some necessary concepts and quantities about finite-size classical capacities of quantum channels; in section 3 we compute the 1-shot capacity exactly (Theorem 3); in section 4 we generalize our analysis to the n-shot capacity, obtaining nearly matching bounds for finite n (Theorems 4, 5) and compare them with GD-based bounds.

2. Non-asymptotic classical capacity of a quantum channel

The quantum erasure channel is a completely positive and trace preserving linear map $\mathcal{E}_q$: $\mathcal{D}(\mathcal{H}_d) \to \mathcal{D}(\mathcal{H}_{d+1})$ from the space of density operators on a d-dimensional Hilbert space $\mathcal{H}_d$, to that of a Hilbert space with one more dimension $\mathcal{H}_{d+1} = \mathcal{H}_d \cup \{|e\rangle\}$, where $|e\rangle$ represents a default error vector, orthogonal to the input Hilbert space. The channel acts on an input state $\rho \in \mathcal{D}(\mathcal{H}_d)$ as follows:

$$\mathcal{E}_q(\rho) = (1 - q)\rho + q|e\rangle\langle e|, \tag{1}$$

where q is the erasure probability.

For a general quantum channel $\mathcal{N}$, the 1-shot classical capacity is defined as the maximum bit transmission rate attainable by encoding a random message $m = 1, \cdots, K$ into a quantum state $\rho_m \in \mathcal{H}$, and decoding it after transmission on the channel via a positive-operator-valued measurement (POVM) $\{D_m\}_{m=0}^K$, where the operator D_0 corresponds to an error outcome upon which no guess on the message is made. Since in a single round the decoding error might be strictly larger than zero, one allows for an error margin ϵ, obtaining the following definition[1]:

Definition 1. The ϵ-error 1-shot classical capacity of a quantum channel $\mathcal{N}$: $\mathcal{D}(\mathcal{H}) \mapsto \mathcal{D}(\mathcal{H}')$ is

$$C_1^\epsilon(\mathcal{N}) := \max_{\{\rho_m\},\{D_m\}} \log K \text{ s.t.} 1 - p_{\text{succ}} \leqslant \epsilon, \tag{2}$$

where

$$p_{\text{succ}} = \frac{1}{K}\sum_{m=1}^{K} \operatorname{Tr} \mathcal{N}(\rho_m) D_m \tag{3}$$

is the average decoding success probability and the maximization is over all input state ensembles $\{\rho_m\}_{m=1}^K$ and decoding POVMs $\{D_m\}_{m=0}^K$. The n-shot capacity is

$$C_n^\epsilon(\mathcal{N}) := \frac{1}{n} C_1^\epsilon(\mathcal{N}^{\otimes n}), \tag{4}$$

where the optimization over n channel uses allows to employ non-product input states on $\mathcal{D}(\mathcal{H}^{\otimes n})$ and measurements on $\mathcal{D}(\mathcal{H}'^{\otimes n})$.

In [19], it was shown that the 1-shot classical capacity of a quantum channel $\mathcal{N}$ can be bounded in terms of generalized divergences (GD). For our purposes, we are going to need the following family:

Definition 2. For any $\alpha \in (0, 1) \cup (\infty)$, the α-Petz-Rényi relative entropy between two quantum states $\rho, \sigma \in \mathcal{D}(\mathcal{H})$ is

$$D_\alpha(\rho||\sigma) := \frac{1}{\alpha - 1} \log \operatorname{Tr} \rho^\alpha \sigma^{1-\alpha}. \tag{5}$$

We then have the following characterization of the classical n-shot capacity:

Theorem 1. *([17, 19]) The ϵ-error* n*-shot capacity of a quantum channel admits the following upper- and lower-bounds:*

[1] Henceforth we use the base-2 logarithm.

$$C_n^{\epsilon}(\mathcal{N}) \leqslant \frac{1}{n}\left(\max_{\{p_\ell,\rho_\ell\}} D_\beta(\omega_{\mathrm{AB}}||\omega_{\mathrm{A}}\otimes\omega_{\mathrm{B}}) + \frac{\beta}{\beta-1}\log\frac{1}{1-\epsilon}\right), \tag{6}$$

for all $\beta \in (1, \infty)$ *and*

$$C_n^{\epsilon}(\mathcal{N}) \geqslant \frac{1}{n}\left(\max_{\{p_\ell,\rho_\ell\}} D_\alpha(\omega_{\mathrm{AB}}||\omega_{\mathrm{A}}\otimes\omega_{\mathrm{B}}) - \frac{1}{1-\alpha}\log\frac{4}{\epsilon} - 4\right), \tag{7}$$

for all $\alpha \in (0, 1)$, *where*

$$\omega_{\mathrm{AB}} = \sum_{\ell=1}^{K} p_\ell |\ell\rangle\langle\ell| \otimes \mathcal{N}^{\otimes n}(\rho_\ell). \tag{8}$$

Finally, in the asymptotic limit, these GD bounds match and the n-shot capacity approaches the Holevo limit:
Theorem 2. [20] *The Holevo capacity of the erasure channel is*

$$\chi(\mathcal{E}_q) = (1-q)\log d. \tag{9}$$

3. Exact calculation of the 1-shot classical capacity

Intuitively, one can expect the 1-shot capacity of the erasure channel to exhibit two different regimes, with a threshold around the value $q \sim \epsilon$. Indeed, when the erasure error probability is smaller than the allowed error threshold $q \lesssim \epsilon$, we have an additional error margin that permits us to increase the number of messages sent beyond the space dimension, as for the identity channel. On the other hand, when the erasure probability is larger than the allowed error threshold, we will be forced to reduce the number of messages in order to increase their average distinguishability. Finally, the threshold between the two regimes is not exactly at $q = \epsilon$ because, even for full erasure, i.e. a constant channel, it is always possible to reduce a little bit the error by making a random guess. Harnessing these observations, we can prove the following theorem:

Theorem 3. *The 1-shot classical capacity of the quantum erasure channel has the closed-form expression:*

$$C_\epsilon(\mathcal{E}_q) = \begin{cases} \log\left\lfloor \dfrac{(1-q)d+q}{1-\epsilon} \right\rfloor & q \leqslant q(\epsilon), \\ \log\left\lfloor \dfrac{q}{q-\epsilon} \right\rfloor & q > q(\epsilon), \end{cases} \tag{10}$$

where

$$q(\epsilon) := \epsilon\frac{d}{d-1}. \tag{11}$$

Proof. We start by deriving matching upper and lower bounds for the region $q \lesssim \epsilon$. Consider a generic code $\{\rho_m\}_{m=1}^K$ and decoding POVM $\{D_m\}_{m=0}^K$; its average success probability can be bounded as follows:

$$p_{\mathrm{succ}} = \frac{1}{K}\sum_{m=1}^{K} \mathrm{Tr}\,\mathcal{E}_q(\rho_m)D_m = \frac{1}{K}\sum_{m=1}^{K}[(1-q)\mathrm{Tr}\,\rho_m D_m + q\langle e|D_m|e\rangle] \tag{12}$$

$$\leqslant \frac{1}{K}[(1-q)\mathrm{Tr}\,1_d\cdot 1_{d+1} + q\cdot 1] \leqslant \frac{(1-q)d+q}{K}, \tag{13}$$

where in the first inequality we used that $\rho_m \leqslant 1_d$ for all $m > 0$ and $\sum_{m=0}^K D_m = 1_{d+1}$. We conclude that, if $p_{\mathrm{succ}} \geqslant 1-\epsilon$, the number of messages can be at most

$$K \leqslant \frac{(1-q)d+q}{1-\epsilon}. \tag{14}$$

This upper bound is achievable by the following strategy:

$$\rho_m := |m \bmod d\rangle\langle m \bmod d|,\ \forall\, m = 1,\cdots, K \tag{15}$$

$$D_m := \begin{cases} \dfrac{1}{Q+1}|m \bmod d\rangle\langle m \bmod d| \oplus \dfrac{1}{K}|e\rangle\langle e| & 0 < m \bmod d \leqslant R, \\ \dfrac{1}{Q}|m \bmod d\rangle\langle m \bmod d| \oplus \dfrac{1}{K}|e\rangle\langle e| & m \bmod d > R \text{ or } m \bmod d = 0, \end{cases} \tag{16}$$

$$D_0 := 0, \tag{17}$$

where $\{|i\rangle\}_{i=0}^{d-1}$ is an orthonormal basis of the input Hilbert space and we have defined $Q,\ R$ as the quotient and the remainder of the quotient between integers K and d, such that $K = Qd + R$. The intuition behind this construction is that, for each computational basis state $|m'\rangle$, the number of m values that are encoded in that state can be either $Q+1$, if $m' \leqslant R$, or Q, if $m' > R$ or $m' = 0$.

It is straightforward to show that the success probability corresponding to this scheme equals the upper bound of (14), since

$$\sum_{m=1}^{K} \mathrm{Tr}\, \rho_m D_m = \sum_{m \bmod d \leqslant R} \frac{1}{Q+1} + \sum_{m \bmod d > R} \frac{1}{Q} = \frac{R(Q+1)}{Q+1} + \frac{(d-R)Q}{Q} = d, \tag{18}$$

implying that the optimal number of messages is

$$K = \left\lfloor \frac{(1-q)d+q}{1-\epsilon} \right\rfloor. \tag{19}$$

However, note that the communication strategy detailed above relies on encoding in each state $|i\rangle$ roughly K/d equiprobable messages, which can be distinguished only with average success probability d/K. It is clear then that $K \geqslant d$ for this strategy to be valid, which means that the upper bound (13) is not achievable when the number of messages it predicts is smaller than the input space dimension. We thus expect this to be tight in the region $q \lesssim \epsilon$.

Let us then introduce another upper bound, for which a matching lower bound can be found in the large-noise parameter region $q \gtrsim \epsilon$:

$$p_{\mathrm{succ}} \leqslant \frac{1}{K}\left[(1-q)\sum_m \mathrm{Tr}\, \rho_m \cdot 1_{d+1} + q \cdot 1\right] \leqslant 1 - q + \frac{q}{K}, \tag{20}$$

where we have used the fact that $D_m \leqslant 1_{d+1}$. In order for this to be larger than $1-\epsilon$, the number of messages must be bounded by

$$K \leqslant \frac{q}{q-\epsilon}. \tag{21}$$

The following explicit strategy achieves the same error probability:

$$\rho_m \coloneqq |m\rangle\langle m|, \quad \forall\, m = 1, \cdots, K \tag{22}$$

$$D_m \coloneqq |m\rangle\langle m| \oplus \frac{1}{K}|e\rangle\langle e| \quad m = 1, \cdots, K \tag{23}$$

$$D_0 \coloneqq \sum_{m=K+1}^{d} |m\rangle\langle m|, \tag{24}$$

so that the maximum number of messages is

$$K = \left\lfloor \frac{q}{q-\epsilon} \right\rfloor. \tag{25}$$

Note that this strategy is valid only for $K \leqslant d$, conversely to the first strategy. The transition between the two regimes thus happens at

$$\frac{(1-q)d+q}{1-\epsilon} = d = \frac{q}{q-\epsilon}, \tag{26}$$

which is satisfied for $q = q(\epsilon)$. □

In figure 1 we plot (10) as a function of q for ϵ fixed. We observe a discontinuity of the first derivative exactly at $q(\epsilon)$, which constitutes a turning point where the two bounds cross each other. Note also that the threshold is slightly larger than ϵ, because the optimal strategies for both regimes employ a decoding POVM that tries to decode the message also when the constant error $|e\rangle$ has happened, by making a random guess on all possible messages. A suboptimal rate can be obtained by removing this component and taking D_m that act only on A; the resulting threshold in such a case will be then exactly at $q = \epsilon$.

4. Near-optimal bounds for the *n*-shot classical capacity

The strategy used for the 1-shot case gives us insight on the behaviour of the n-shot capacity and allows us to derive almost-matching upper- and lower-bounds for this general case. As a result, we obtain a family of almost-capacity-achieving protocols for the erasure channel that work for any number of channel uses n.

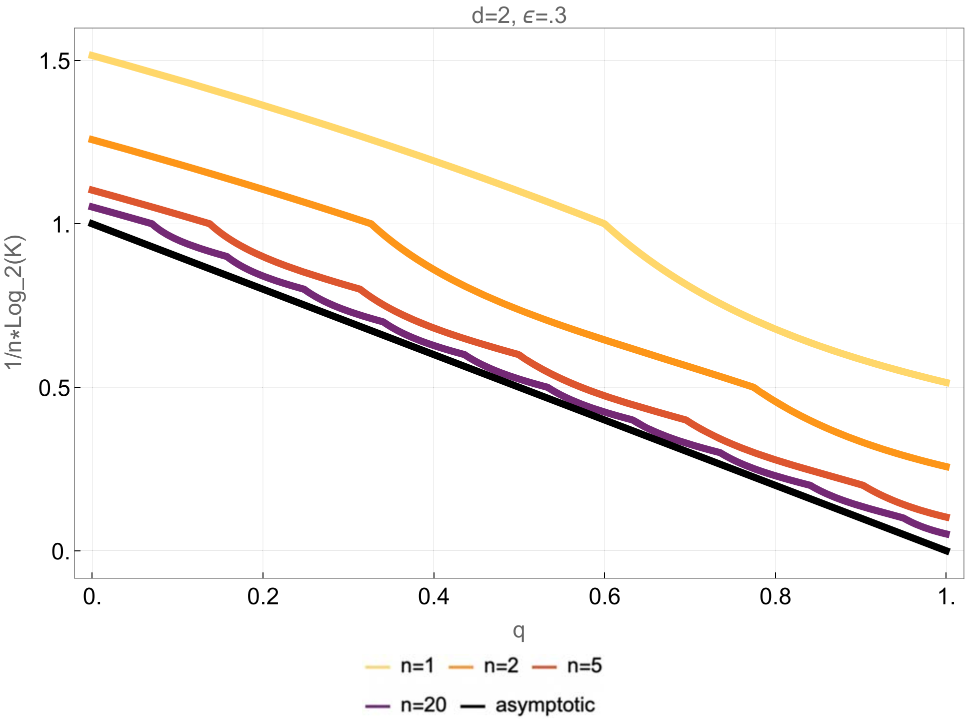

Figure 1. Plot of the n-shot capacity upper bound for the quantum erasure channel $\mathcal{E}_q$ versus the erasure probability q for $d = 2$ and $\epsilon = 0.3$. For $n = 1$, there is a matching lower bound, hence we plot the exact expression (10). For $n > 1$ we plot the only the upper bound (27). We see that, as n increases, the upper bound approaches the asymptotic Holevo capacity of the channel (9). The discontinuous derivative is explained by the transition to different communication regimes as the erasure probability increases. For $n = 1$ the bound is attainable and the turning point is exactly at $q(0.3) = 0.6$.

4.1. Upper bound

In the 1-shot case, we had two distinct upper bounds, each tight in a different region, respectively $K \geqslant d$ or $K < d$. By close inspection, we see that the first upper bound (13) corresponds to a situation where one is able to decode at least d messages, i.e. the maximum allowed by the Hilbert space dimension and a bit more for finite error threshold; clearly, this is possible only if the erasure probability is sufficiently small. On the other hand, the second upper bound (20) corresponds to being able to decode only $K < d$ messages, due to the erasure probability being large.

In the n-shot case, one is allowed to use the channel n times, hence the total number of erasures will follow a binomial distribution, giving rise to multiple error thresholds. The analysis is more complex with respect to the 1-shot case and does not give rise to a closed formula, however we can state the folliwing theorem:

Theorem 4. *The n-shot capacity of the quantum erasure channel can be upper-bounded as*

$$C_n^\epsilon(\mathcal{E}_q) \leqslant \frac{S_i^n(d)}{S_i^n(1) - \epsilon} \quad \text{for } q(\epsilon, i+1) \leqslant q \leqslant q(\epsilon, i), \tag{27}$$

where

$$S_a(d) := \binom{n}{a}((1-q)d)^{n-a}q^a, \quad S_a^b(d) := \sum_{i'=a}^{b} S_{i'}(d), \tag{28}$$

and $q(\epsilon, i)$ *is the solution to the equation*

$$\frac{S_i^n(d)}{S_i^n(1) - \epsilon} = d^{n-i} = \frac{S_{i+1}^n(d)}{S_{i+1}^n(1) - \epsilon}. \tag{29}$$

Proof. In this case, the success probability can be written as

$$p_{\mathrm{succ}} = \frac{1}{K}\sum_{m=1}^{K}\mathrm{Tr}\,\mathcal{E}_q^{\otimes n}(\rho_m)D_m \tag{30}$$

$$= \sum_{i=0}^{n}(1-q)^{n-i}q^i \sum_{J\subseteq[n]:|J|=i}\frac{1}{K}\sum_m \mathrm{Tr}\left((\rho_m)_{J^c}\otimes(|e\rangle\langle e|^{\otimes i})_J\right)D_m, \tag{31}$$

where J (respectively J^c) identifies the subsystems where an erasure has happened (respectively not happend)), $[n]=\{1,\cdots,n\}$ is the full index set and $|J|$ the cardinality of J. Note that for each i there are $\binom{n}{i}$ such sets. Furthermore, $(\rho_m)_{J^c}$ is the reduced state of ρ_m on subsystems J^c.

Each term with $i < n$ in (31) can be bounded in the two different ways adopted for the 1-shot case:

$$\frac{1}{K}\sum_m \mathrm{Tr}\left((\rho_m)_{J^c}\otimes(|e\rangle\langle e|^{\otimes i})_J\right)D_m \leqslant \frac{1}{K}\mathrm{Tr}(1_{d^{n-i}}\otimes|e\rangle\langle e|^{\otimes i})1_{(d+1)^n} = \frac{d^{n-i}}{K}, \tag{32}$$

which corresponds to a high-error region where only $d^{n-i}\leqslant K$ messages can be decoded upon erasure of i parties, or

$$\frac{1}{K}\sum_m \mathrm{Tr}\left((\rho_m)_{J^c}\otimes(|e\rangle\langle e|^{\otimes i})_J\right)D_m \leqslant \frac{1}{K}\sum_m \mathrm{Tr}\left((\rho_m)_{J^c}\otimes(|e\rangle\langle e|^{\otimes i})_J\right)1_{(d+1)^n} = 1. \tag{33}$$

which corresponds to a low-error region where all K messages can be decoded upon erasure of i parties. In turn, this gives rise to $n+1$ different bounds, depending on how many erasures the communication protocol is able to correct: for $i=0,\cdots,n$ and $d^{n-i}\leqslant K\leqslant d^{n-i+1}$ the protocol can recover each message perfectly, provided that at most i erasures took place. In this region we can bound the terms $i'\leqslant i$ with few erasures via (33) and the terms $i'>i$ with many erasures via (32). We conclude that the success probability in the i-th region is upper-bounded as

$$p_{\mathrm{succ}} \leqslant p_{\mathrm{succ}}^{(i)} \coloneqq \sum_{i'=0}^{i-1}\binom{n}{i'}(1-q)^{n-i'}q^{i'} + \frac{1}{K}\sum_{i'=i}^{n}\binom{n}{i'}((1-q)d)^{n-i'}q^{i'}, \tag{34}$$

where we have defined $p_{\mathrm{succ}}^{(i)}$ as the i-th region bound.

From (34), setting $p_{\mathrm{succ}} = 1-\epsilon$ we obtain the following bound on the maximum number of messages transmittable in each region i:

$$K \leqslant \frac{S_i^n(d)}{S_i^n(1)-\epsilon}, \tag{35}$$

where we have defined

$$S_a(d) \coloneqq \binom{n}{a}((1-q)d)^{n-a}q^a, \qquad S_a^b(d) \coloneqq \sum_{i'=a}^{b} S_{i'}(d), \tag{36}$$

and used the fact that $S_0^{i-1}(1) = S_0^n(1) - S_i^n(1) = 1 - S_i^n(1)$. In particular, we expect the capacity to be continuous, hence at the boundary of each couple of adjacent regions, labelled by $i=0,\cdots,n-1$ it must hold that

$$\frac{S_i^n(d)}{S_i^n(1)-\epsilon} = d^{n-i} = \frac{S_{i+1}^n(d)}{S_{i+1}^n(1)-\epsilon}. \tag{37}$$

□

We observe that the limiting cases of (34) have a simple interpretation. When the erasure probability is so small that all (but n) erasures can be corrected on average, $i=0$, and one can even increase the number of messages beyond the Hilbert space dimension if allowing a certain margin of error, i.e. $d^n\leqslant K$; this case has high success probability:

$$p_{\mathrm{succ}}^{(0)} = \frac{1}{K}\sum_{i'=0}^{n}\binom{n}{i'}((1-q)d)^{n-i'}q^{i'} = \frac{((1-q)d+q)^n}{K}. \tag{38}$$

The opposite case is when the erasure probability is so large that no erasure can be corrected on average, $i=n$, then one can use only $K\leqslant d$ messages; this case has low success probability:

$$p_{\mathrm{succ}}^{(n)} = \sum_{i'=0}^{n-1}\binom{n}{i'}(1-q)^{n-i'}q^{i'} + \frac{q^n}{K} = 1 - q^n + \frac{q^n}{K}. \tag{39}$$

Furthermore, note that, unlike the $n=1$ case, this equation cannot be solved explicitly to find the transition value $q(\epsilon, i)$ for each ϵ, i. Nevertheless, in figure 1 we plot (27) as a function of q for several values of n, choosing the minimum among all values over i. The plot confirms that there are $n+1$ turning points with discontinuous derivative, where the $i+1$-th upper bound becomes tighter than the i-th one.

4.2. Nearly-matching lower bound

The upper bound (35), identifies n distinct regions that are characterized by being able to correct, on average, $n-i$ errors. In order to show that this bound is tight, one needs to find a protocol with average success probability equal or close to (34). This can be done via a code that has $d^{n-i} \leqslant K \leqslant d^{n-i+1}$ messages and is robust against up to i erasures, i.e. any two codewords differ in at least $i+1$ parties so that, if i erasures happen, it is still possible to distinguish any two codewords in at least one position. We then obtain the following theorem:

Theorem 5. *For any $D = 0, \cdots, i$ such that there exists a $[n, n-i, i-D+1]_d$ classical error-correcting code on* d *symbols, with* n *message bits,* n − i *information bits and minimum distance $i - D + 1$, the* n*-shot capacity of the quantum erasure channel can be lower-bounded as*

$$C_n^{\epsilon}(\mathcal{E}_q) \geqslant \frac{d^{n-i}\left(S_0^{i-D}(1) + \sum_{i'=i-D+1}^{n} S_{i'}(1)\cdot \tilde{d}_{i'}^{-1}\right)}{1-\epsilon} \quad \text{for } d^{n-i} \leqslant C_n^{\epsilon}(\mathcal{E}_q) \leqslant d^{n-i+1}, \tag{40}$$

with the same notation of Theorem 4, and $\tilde{d}_{i'} := \min\{d^{n-i}, d^{i'}\}$. The bound is decreasing in D.

Proof. Suppose that, for all i, a $[n, n-i, i+1]_d$ code $\mathcal{C}$ on a d-ary field exists with length n, information bits $n-i$ and minimum Hamming distance between codewords $i+1$; this is called a maximum-distance-separable (MDS) code, as it saturates the Singleton bound [21]. Then this code is able to perfectly withstand up to $i' = i$ erasures in random positions. If more than i erasures happen, then multiple codewords could be compatible with the received message and we have to make a random guess among them. The number of compatible codewords after $i' > i$ erasures is, in the worst case, $\tilde{d}_{i'}$.[2]

Let us label the codewords of the code via an integer index, such that $\mathbf{b}_c \in \mathcal{C} \subset \{0, \cdots, d-1\}^n$ for all $c = 0, \cdots, d^{n-i}$. Consider then a protocol with $d^{n-i} \leqslant K < d^{n-i+1}$ messages

$$\rho_m = |\mathbf{b}(m)\rangle\langle\mathbf{b}(m)| \quad \forall\, m = 1, \cdots, K, \tag{41}$$

where we have defined the function

$$\mathbf{b}\colon m \mapsto \mathbf{b}_{m \bmod d^{n-i}} \tag{42}$$

mapping the message index to a codeword of $\mathcal{C}$. Note that, if $K = d^{n-i} + R$ with $R > 0$, there will be $R-1$ codewords that encode two messages. The corresponding quantum decoding operators are of the form

$$D_m = \begin{cases} \dfrac{1}{2}\displaystyle\sum_{i'=0}^{i}\left\langle \bigotimes_{\ell=1}^{i'} |e\rangle\langle e|_\ell \otimes \bigotimes_{\ell=i'+1}^{n} |b_\ell(m)\rangle\langle b_\ell(m)|_\ell \right\rangle & \\ + \displaystyle\sum_{i'=i+1}^{n} \frac{1}{2\tilde{d}_{i'}}\left\langle \bigotimes_{\ell=1}^{i'} |e\rangle\langle e|_\ell \otimes \bigotimes_{\ell=i'+1}^{n} |b_\ell(m)\rangle\langle b_\ell(m)|_\ell \right\rangle & 0 < m \bmod d^{n-i} \leqslant R, \\ \displaystyle\sum_{i'=0}^{i}\left\langle \bigotimes_{\ell=1}^{i'} |e\rangle\langle e|_\ell \otimes \bigotimes_{\ell=i'+1}^{n} |b_\ell(m)\rangle\langle b_\ell(m)|_\ell \right\rangle & \\ + \displaystyle\sum_{i'=i+1}^{n} \frac{1}{\tilde{d}_{i'}}\left\langle \bigotimes_{\ell=1}^{i'} |e\rangle\langle e|_\ell \otimes \bigotimes_{\ell=i'+1}^{n} |b_\ell(m)\rangle\langle b_\ell(m)|_\ell \right\rangle & m \bmod d^{n-i} > R \\ & \text{or } m \bmod d^{n-i} = 0. \end{cases} \tag{43}$$

where $\langle\cdot\rangle$ represents the sum over all different permutations of the error positions and $b_\ell(m)$ is the ℓ-th entry of codeword $\mathbf{b}$(m). We note that these are separable operators, proving that no coherent processing is required at the decoder [22]. For a fixed order of $i' \leqslant i$ erasures, we have that only a single codeword is compatible with the non-erased parties, therefore

$$\sum_m \sum_{i'=0}^{i} \bigotimes_{\ell=1}^{i'} |e\rangle\langle e|_\ell \otimes \bigotimes_{\ell=i'+1}^{n} |b_\ell(m)\rangle\langle b_\ell(m)|_\ell \leqslant \sum_{i'=0}^{i} \bigotimes_{\ell=1}^{i'} |e\rangle\langle e|_\ell \otimes 1_d^{\otimes n-i'-1}, \tag{44}$$

if the decoded codeword corresponds to a single value of m, and analogously if it corresponds to two values of m, thanks to the normalization introduced in case $m \bmod d^{n-i} \leqslant R$. Instead, for $i' > i$ erasures, at most $\tilde{d}_{i'}$ different codewords could be compatible with the non-erased parties and, distinguishing the compatible m in two groups based on $m \bmod d^{n-i}$, we have

[2] A more refined estimate is $d^{i'-i}\binom{i'}{i'-i}$, though it does not seem to beat the bounds obtained with the rougher estimates above. In order to obtain it observe that, by correcting at most $i'-i$ out of the i' erased bits, one can uniquely identify a codeword compatible with the received one. There are $\binom{i'}{i'-i}$ ways to choose the positions of such bits and at most $d^{i'-i}$ different codewords for each sequence of positions.

$$\sum_{\substack{m \bmod d^{n-i}>0 \\ m \bmod d^{n-i}\leqslant R}} \sum_{i'=i+1}^{n} \frac{1}{2\tilde{d}_{i'}} \bigotimes_{\ell=1}^{i'} |e\rangle\langle e|_\ell \otimes \bigotimes_{\ell=i'+1}^{n} |b_\ell(m)\rangle\langle b_\ell(m)|_\ell \tag{45}$$

$$+ \sum_{\substack{m \bmod d^{n-i}>R \\ \text{or} m \bmod d^{n-i}=0}} \sum_{i'=i+1}^{n} \frac{1}{\tilde{d}_{i'}} \bigotimes_{\ell=1}^{i'} |e\rangle\langle e|_\ell \otimes \bigotimes_{\ell=i'+1}^{n} |b_\ell(m)\rangle\langle b_\ell(m)|_\ell \tag{46}$$

$$= \sum_{b_{i'+1},\cdots,\, b_n} \sum_{\substack{c:b_{c,\ell}=b_\ell \\ \forall \ell=i'+1,\cdots,\, n}} \sum_{i'=i+1}^{n} \frac{1}{\tilde{d}_{i'}} \bigotimes_{\ell=1}^{i'} |e\rangle\langle e|_\ell \otimes \bigotimes_{\ell=i'+1}^{n} |b_\ell\rangle\langle b_\ell|_\ell \tag{47}$$

$$\leqslant \sum_{i'=i+1}^{n} \frac{1}{\tilde{d}_{i'}} \bigotimes_{\ell=1}^{i'} |e\rangle\langle e|_\ell \otimes \tilde{d}_{i'}\, 1_d^{\otimes n-i'-1}, \tag{48}$$

where $b_{c,\ell}$ is the ℓ-th entry of codeword $\mathbf{b}_c$. Therefore, we conclude that $\sum_{\mathbf{m}} D_{\mathbf{m}} \leqslant 1_{d+1}^{\otimes n}$, completing this to a measurement via the additional error-outcome operator $D_0 = 1 - \sum_{\mathbf{m}\in\mathcal{C}} D_{\mathbf{m}}$.

Defining $(\rho_m)_{i'+1,\cdots,\, n}$ as the reduced state on $\mathcal{D}(\mathcal{H})^{\otimes n-i'}$ after erasures in the first i' positions took place, we have that

$$\sum_m \mathrm{Tr}\,(\rho_m)_{i'+1,\cdots,\, n} D_m = \sum_{m \bmod d^{n-i}\leqslant R} \frac{1}{2} + \sum_{m \bmod d^{n-i}>R} 1 \tag{49}$$

$$= \frac{2R}{2} + d^{n-i} - R = d^{n-i} \tag{50}$$

for $i' \leqslant i$, while

$$\sum_m \mathrm{Tr}\,(\rho_m)_{i'+1,\cdots,\, n} D_m = \frac{d^{n-i}}{\tilde{d}_{i'}} \tag{51}$$

for $i' > i$. The resulting lower-bound to the success probability in the i-th region then is

$$p_{\text{succ}} \geqslant \tilde{p}_{\text{succ}}^{(i)} \coloneqq \sum_{i'=0}^{i} \binom{n}{i'} (1-q)^{n-i'} q^{i'} \cdot \frac{d^{n-i}}{K} + \sum_{i'=i+1}^{n} \binom{n}{i'} (1-q)^{n-i'} q^{i'} \cdot \frac{d^{n-i}}{\tilde{d}_{i'}\cdot K}, \tag{52}$$

which has a similar form to the upper-bound (34). Setting $p_{\text{succ}} = 1 - \epsilon$ and solving for K, we obtain the lower-bound (40) on the capacity.

The question is then whether $[n, n-i, i+1]_d$ MDS codes can be constructed for any d, n and i, which can be answered in the negative [21]; for example, for $d = 2$ only trivial codes of this kind exist. On the other hand, there are known classes of MDS codes for certain parameter values: Reed-Solomon MDS codes exist for dimension $d = u^m \geqslant n$ with u prime and m integer. Therefore, in general one can only employ lesser-performing $[n, n-i, i-D+1]_d$ codes with $D \leqslant i$, e.g. from the family of BCH codes [21], obtaining a sub-optimal lower-bound to the success probability that corresponds to correcting only $i - D$ erasures on average. Repeating the previous steps by setting the threshold at $i' = i - D$ instead of i, we obtain in the i-th region

$$p_{\text{succ}} \geqslant S_0^{i-D}(1)\frac{d^{n-i}}{K} + \sum_{i'=i-D+1}^{n} S_{i'}(1)\frac{d^{n-i}}{\tilde{d}_{i'}K}. \tag{53}$$

and the corresponding lower-bound on the number of messages. □

Note that the lower-bound (52) can be further simplified by fixing $\tilde{d}_{i'}$ as one of the two possible choices, obtaining

$$p_{\text{succ}} \geqslant \max\left\{\frac{S_0^{i-D}(1)d^{n-i} + S_{i-D+1}^{n}(1)}{K}, \frac{S_0^{i-D}(1)d^{n} + S_{i-D+1}^{n}(d)}{d^{i}\cdot K}\right\}. \tag{54}$$

The upper and lower bounds (27), (40) are plotted in figure 2 for several values of D. It can be seen that these bounds are very tight, determining the capacity up to a narrow interval, contrarily to what happens for the GD-based bounds discussed in the next section.

4.3. Bounds from hypothesis testing

We can compare the bounds found in the previous sections with those coming from the general theory of n-shot classical capacity of a quantum channel [19]. In order to apply 1, one needs to compute or approximate the GD for the erasure channel under consideration, obtaining the following theorem:

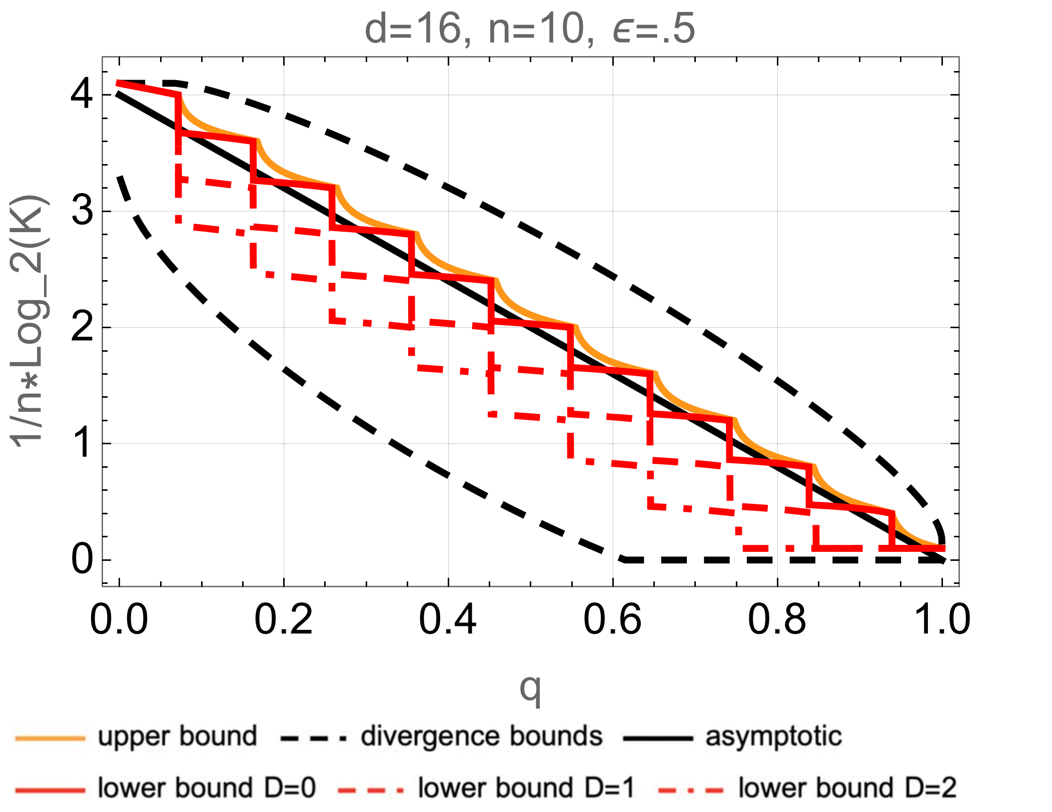

Figure 2. Plot of upper- (27) and lower-bounds (40) on the 10-shot channel capacity of a 16-dimensional system for $\epsilon = 0.5$ versus the erasure probability q. We compare it with the asymptotic channel capacity (9), as well as the divergence-based upper- (55) and lower-bounds (56) (the latter plotted only when it is positive), optimized over the admissible α and β range. We observe that our explicit analysis provides quite accurate results with respect to the general GD-based theory for finite size.

Theorem 6. *The* n*-shot classical capacity of the quantum erasure channel is upper- and lower-bounded via GD as*

$$C_n^{\epsilon}(\mathcal{E}_q) \leqslant \log d + \frac{1}{\beta - 1}\log(1 - q + q \cdot d^{1-\beta}) + \frac{\beta}{n(\beta - 1)}\log\frac{1}{1 - \epsilon}, \tag{55}$$

for any $\beta \in (1, \infty)$ *and*

$$C_n^{\epsilon}(\mathcal{E}_q) \geqslant \log d - \frac{1}{1 - \alpha}\log(1 - q + q \cdot d^{1-\alpha}) + \frac{\alpha}{n(\alpha - 1)}\log\frac{2}{\epsilon} - \frac{1}{n}\left(\log\frac{2}{\epsilon} + 4\right), \tag{56}$$

for any $\alpha \in (0, 1)$.

Proof. Let us first write the classical-quantum state of the input and output subsystems for n channel uses as

$$\omega_{\mathrm{AB}} = \sum_{\ell=1}^{K} p_\ell |\ell\rangle\langle\ell| \otimes \mathcal{E}_q^{\otimes n}(\rho_\ell) \tag{57}$$

$$= \sum_{\ell=1}^{K} p_\ell |\ell\rangle\langle\ell| \otimes \sum_{i=0}^{n}(1 - q)^{n-i}q^{i} \sum_{J\subseteq[n]:|J|=i} (|e\rangle\langle e|^{\otimes i})_J \otimes (\rho_\ell)_{J^c}, \tag{58}$$

with $\{|\ell\rangle\}_{\ell=1}^{K}$ an orthonormal basis of A. Clearly, this state is block-diagonal and we have that

$$\omega_{\mathrm{AB}}^{\alpha} = \sum_{\ell=1}^{K} p_\ell^{\alpha} |\ell\rangle\langle\ell| \otimes \sum_{i=0}^{n}(1 - q)^{(n-i)\alpha}q^{i\alpha} \sum_{J\subseteq[n]:|J|=i} (|e\rangle\langle e|^{\otimes i})_J \otimes (\rho_\ell)_{J^c}^{\alpha}, \tag{59}$$

and similarly

$$(\omega_{\mathrm{A}} \otimes \omega_{\mathrm{B}})^{1-\alpha} = \sum_{\ell'=1}^{K} p_{\ell'}^{1-\alpha} |\ell'\rangle\langle\ell'| \otimes \sum_{i=0}^{n}(1 - q)^{(n-i)(1-\alpha)}q^{i(1-\alpha)} \sum_{J\subseteq[n]:|J|=i} (|e\rangle\langle e|^{\otimes i})_J \otimes (\bar{\rho})_{J^c}^{(1-\alpha)}, \tag{60}$$

where $\bar{\rho} = \sum_\ell p_\ell \rho_\ell$ is the average output state of the ensemble.

We can then observe that

$$\mathrm{Tr}\,\omega_{\mathrm{AB}}^{\alpha}(\omega_{\mathrm{A}}\otimes\omega_{\mathrm{B}})^{1-\alpha}=\sum_{\ell}p_{\ell}\sum_{i=0}^{n}(1-q)^{n-i}q^{i}\sum_{J\subseteq[n]:|J|=i}\mathrm{Tr}\,(\rho_{\ell})_{J^c}^{\alpha}(\bar{\rho})_{J^c}^{1-\alpha}. \tag{61}$$

Therefore, we can lower-bound the GD by providing a completely mixed input ensemble $\left(\{\frac{1}{d},\,|\ell\rangle\langle\ell|\}_{\ell=0}^{d-1}\right)^{\otimes n}$ in product form, such that $p_{\ell}=1/d^{n}$, $\rho_{\ell}=|\ell\rangle\langle\ell|$ for each $\ell=(\ell_1,\cdots,\ell_n)$ and $\bar{\rho}=\frac{1}{d^n}$, obtaining

$$\max_{\{p_{\ell},\rho_{\ell}\}}D_{\alpha}(\omega_{\mathrm{AB}}||\omega_{\mathrm{A}}\otimes\omega_{\mathrm{B}})\geqslant\frac{1}{\alpha-1}\log\sum_{\ell}\frac{1}{d^{n}}\sum_{i=0}^{n}(1-q)^{n-i}q^{i}\sum_{J\subseteq[n]:|J|=i}\frac{\prod_{j\in J^c}\langle\ell_j|1_j|\ell_j\rangle}{d^{(n-i)(1-\alpha)}} \tag{62}$$

$$=\frac{1}{\alpha-1}\log\frac{1}{d^{n(1-\alpha)}}\sum_{i=0}^{n}\binom{n}{i}(1-q)^{n-i}(q\cdot d^{1-\alpha})^{i} \tag{63}$$

$$=\frac{1}{\alpha-1}\log\frac{(1-q+q\cdot d^{1-\alpha})^{n}}{d^{n(1-\alpha)}} \tag{64}$$

$$=n\log d-\frac{n}{1-\alpha}\log(1-q+q\cdot d^{1-\alpha}), \tag{65}$$

where in the inequality we have used that $(\bar{\rho})_{J^c}=1_{J^c}/d^{n-i}$.

For the upper bound instead, observe that for $\beta>1$ and any state σ,τ it holds $\sigma^{\beta}\leqslant\sigma$ and $\mathrm{Tr}\,\sigma^{\beta}\tau\leqslant\mathrm{Tr}\,\sigma\tau$. Therefore

$$\sum_{\ell}p_{\ell}\,\mathrm{Tr}\,(\rho_{\ell})_{J^c}^{\beta}(\bar{\rho})_{J^c}^{1-\beta}\leqslant\sum_{\ell}p_{\ell}\,\mathrm{Tr}\,(\rho_{\ell})_{J^c}(\bar{\rho})_{J^c}^{1-\beta}=\mathrm{Tr}\,(\bar{\rho})_{J^c}^{2-\beta}. \tag{66}$$

Since $x^{2-\beta}$ is concave for $\beta\in(1,\,2]$, we have that

$$\sigma\succ\rho\Rightarrow\mathrm{Tr}\,\sigma^{2-\beta}\leqslant\mathrm{Tr}\,\rho^{2-\beta}, \tag{67}$$

where $\sigma\succ\rho$ means that the vector of eigenvalues of σ majorizes that of ρ. We conclude that (66) is maximized by the maximally mixed state in $(\bar{\rho})_{J^c}=1_{J^c}/d^{n-i}$, and hence for $\beta\in(1,\,2]$ it holds

$$\max_{\omega}D_{\beta}(\omega_{\mathrm{AB}^{\otimes n}}||\omega_{\mathrm{A}}\otimes\omega_{\mathrm{B}^{\otimes n}})\leqslant\frac{1}{\beta-1}\log\sum_{i=0}^{n}(1-q)^{n-i}q^{i}\sum_{J\subseteq[n]:|J|=i}\frac{d^{n-i}}{d^{(n-i)(2-\beta)}} \tag{68}$$

$$=\frac{1}{\beta-1}\log\frac{1}{d^{n(1-\beta)}}\sum_{i=0}^{n}\binom{n}{i}(1-q)^{n-i}(q\cdot d^{1-\beta})^{i} \tag{69}$$

$$=n\log d+\frac{n}{\beta-1}\log(1-q+q\cdot d^{1-\beta}), \tag{70}$$

which has a similar form to the expression of the lower bound, albeit for different values of β (65). Combining these together with theorem (1) we obtain the result. □

In figure 2 we plot the near-optimal bound of section 4.1 together with the asymptotic capacity and the divergence-based bounds found above, optimized with respect to α and β. We observe that these bounds are quite loose with respect to the one found by an explicit error-probability analysis and protocol design.

5. Conclusions

We have computed the n-shot classical capacity of the quantum erasure channels from first principles, providing upper bounds and nearly matching lower bounds. Specifically, our upper bounds are nearly achieved by our lower bounds, based on the existence of classical codes with large minimum distance.

At variance with the general theory based on GD, we show that the explicit analysis of the error probability provides an understanding on the optimal encoder and decoder's behaviour as a function of the erasure probability, directly hinting at a practical communication protocol that explicitly adapts the number of codewords depending on the noise level. Furthermore, the capacity bounds obtained in this way are nearly optimal and highlight interesting features of the coding and decoding problem at finite size. Indeed, given the erasure channel's simple structure, the capacity-achieving protocol is fully separable and does not make use of the PGM.

In light of these reasons, we suggest that a direct coding approach might determine the n-shot capacity of quantum channels more precisely than the GD approach. On one hand, this advantage comes at the cost of generality, as our bounds are obtained making use of the channel's specific properties, hence they need to be adapated on a case-by-case basis. On the other hand, by harnessing the channel's properties, our approach offers a way to devise practical communication protocols that are close to optimal even in the finite-size regime.

Acknowledgments

M R acknowledges support from the project PNRR—Finanziato dall'Unione europea—MISSIONE 4 COMPONENTE 2 INVESTIMENTO 1.2 - 'Finanziamento di progetti presentati da giovani ricercatori' - Id MSCA_0000011-SQUID—CUP F83C22002390007 (Young Researchers) - Finanziato dallUnione europea—NextGenerationEU.

Data availability statement

All data that support the findings of this study are included within the article (and any supplementary files).

ORCID iDs

Matteo Rosati https://orcid.org/0000-0002-8972-2936

References

[1] Holevo A S 2012 *Quantum Systems, Channels, Information* (De Gruyter) (https://doi.org/10.1515/9783110273403)

[2] Hastings M B 2009 Superadditivity of communication capacity using entangled inputs. *Nat. Phys.* **5** 255–7

[3] Holevo A S and Giovannetti V 2012 Quantum channels and their entropic characteristics *Reports Prog. Phys.* **75** 046001

[4] Giovannetti V, García-Patrón R, Cerf N J and Holevo A S 2014 Ultimate classical communication rates of quantum optical channels *Nat. Photonics* **8** 796–800

[5] Hausladen P, Jozsa R, Schumacher B, Westmoreland M and Wootters W K 1996 Classical information capacity of a quantum channel *Phys. Rev. A—At. Mol. Opt. Phys.* **54** 1869–76

[6] Schumacher B and Westmoreland M D 1997 Sending classical information via noisy quantum channels *Phys. Rev. A—At. Mol. Opt. Phys.* **56** 131–8

[7] Giovannetti V, Lloyd S and Maccone L 2012 Achieving the Holevo bound via sequential measurements *Phys. Rev.* A **85** 012302

[8] Wilde M M and Guha S 2013 Polar codes for classical-quantum channels *IEEE Trans. Inf. Theory* **59** 1175–87

[9] Takeoka M, Krovi H and Guha S 2013 Achieving the Holevo capacity of a pure state classical-quantum channel via unambiguous state discrimination *2013 IEEE Int. Symp. Inf. Theory* pp 166–70

[10] Rosati M and Giovannetti V 2016 Achieving the Holevo bound via a bisection decoding protocol *J. Math. Phys.* **57** 062204

[11] Rosati M, Mari A and Giovannetti V 2017 Capacity of coherent-state adaptive decoders with interferometry and single-mode detectors *Phys. Rev.* A **96** 012317

[12] Rosati M, De Palma G, Mari A and Giovannetti V 2017 Optimal quantum state discrimination via nested binary measurements *Phys. Rev. A—At. Mol. Opt. Phys.* **95** 1–10

[13] Rosati M 2021 Performance of coherent frequency-shift keying for classical communication on quantum channels *2021 IEEE Int. Symp. Inf. Theory* (https://doi.org/10.1109/ISIT45174.2021.9517959)

[14] Holevo A S 2021 Accessible information of a general quantum Gaussian ensemble **62** 092201

[15] Mishra H K, Lami L, Mandayam P and Wilde M M 2023 Pretty good measurement for bosonic Gaussian ensembles *International Journal of Quantum Information* 2440010

[16] Rosati M and Solana A 2023 Optical decoder learning for fiber communication at the quantum limit *Quantum Physics* arXiv preprint arXiv:2312.13693

[17] Khatri S and Wilde M M 2020 Principles of quantum communication theory: a modern approach *Quantum Physics* arXiv preprint arXiv:2011.04672

[18] Tomamichel M 2016 *Quantum Information Processing with Finite Resources vol 5 of SpringerBriefs in Mathematical Physics* (Springer International Publishing) (https://doi.org/10.1007/978-3-319-21891-5)

[19] Wang L and Renner R 2012 One-shot classical-quantum capacity and hypothesis testing *Phys. Rev. Lett.* **108** 200501

[20] Bennett C H, DiVincenzo D P and Smolin J A 1997 Capacities of quantum erasure channels *Phys. Rev. Lett.* **78** 3217–20

[21] Ling S and Chaoping X 2004 *Coding Theory: A First Course* (Cambridge University Press)

[22] Diaz M G, Desef B, Rosati M, Egloff D, Calsamiglia J, Smirne A, Skotiniotis M and Huelga S F 2020 Accessible coherence in open quantum system dynamics *Quantum* **4** 249