Non-signaling reliability of classical–quantum channels with full-rank outputs

Abstract

We prove that non-signaling assistance attains the sphere-packing error exponent for every finite classical–quantum channel with full-rank output states, at every strictly positive rate below the Holevo capacity. Thus, in this setting, the auxiliary noiseless bit in the activated reliability theorem of Oufkir, Tomamichel, and Berta (Lett. Math. Phys., 2026) is unnecessary. The main result is a dimension-independent one-shot comparison: for every finite classical–quantum channel V and integer $M \geq 2$, $\varepsilon_{\mathrm{MC}}(M, V) \leq \varepsilon_{\mathrm{NS}}(M, V) \leq \varepsilon_{\mathrm{MC}}(M + 6, V)$, where $\varepsilon_{\mathrm{NS}}$ is the optimal non-signaling error and $\varepsilon_{\mathrm{MC}}$ its meta-converse relaxation. We obtain the upper bound by completing the operator normalization of a relaxed decoding family while preserving its input distribution and increasing every effect in the positive-semidefinite order. The correction is determined by a linear map close to the identity; the message-count slack gives spectral margins that keep the solution between zero and the identity. Unlike an additive error estimate, this comparison preserves exponentially small errors. The one-shot result also holds for singular output states. For full-rank channels, continuity of the sphere-packing function transfers the known meta-converse exponent to non-signaling coding and yields ordinary limits for both assisted models with exactly $\lceil 2^{nR} \rceil$ messages.

Note. This paper was generated entirely by AI, including MiMo, using an automated research pipeline developed by Chenghua Liu and Hanyu Li.

1 Introduction

A finite classical–quantum (cq) channel maps each letter of a finite input alphabet to a density operator on a finite-dimensional output space. Its reliability function describes the exponential decay of the optimal decoding error at a fixed communication rate below capacity. To determine this exponent from a converse relaxation, an achievability argument must preserve errors that are already exponentially small. An additive discrepancy that vanishes with the block length need not suffice. For non-signaling-assisted cq coding, the coding program and its meta-converse relaxation differ only by an operator normalization constraint. We show that this constraint can be imposed without increasing the error, at the cost of a fixed reduction in the number of messages.

For classical channels, Matthews [1] identified the optimum over non-signaling codes with the finite-blocklength converse of Polyanskiy, Poor, and Verdú. Wang, Xie, and Duan [2] developed semidefinite formulations for non-signaling-assisted coding over quantum channels. Wang, Fang, and Tomamichel [3] showed that a single auxiliary noiseless classical bit suffices to attain the meta-converse after accounting for the communication supplied by that bit. For cq channels, Oufkir, Tomamichel, and Berta [4] determined the resulting activated reliability function and proved that it equals the sphere-packing exponent. They also removed activation for channels with positive non-signaling zero-error capacity and, more generally, above a specified rate threshold [4, Propositions 5.1 and 5.2]. These results do not cover all low positive rates for general full-rank output states.

The sphere-packing exponent was established as an unassisted converse bound by Dalai [5]. Renes [6] and Li and Yang [7] proved unassisted achievability of the random-coding exponent, which matches the sphere-packing bound above the critical rate. Cheng and Liu [8] subsequently obtained a one-shot random-coding bound through an operator layer cake theorem. These results concern unassisted coding. Here the additional power of non-signaling assistance allows us to attain the sphere-packing exponent also at low positive rates for full-rank channels.

A different comparison between non-signaling coding and the meta-converse was proved by Oufkir and Berta [9, Proposition 8]. In terms of error probabilities, their success-probability bound gives

$$\varepsilon_{\mathrm{NS}}(M,V) \leq \varepsilon_{\mathrm{MC}}(M,V) + \frac{1-\varepsilon_{\mathrm{MC}}(M,V)}{M} \leq \varepsilon_{\mathrm{MC}}(M,V) + \frac{1}{M}. \tag{1}$$

This estimate is useful for strong-converse exponents, where the success probability is small. In the reliability regime, however, the additive term $1/M$ has exponent equal to the rate and need not preserve a larger error exponent. We instead compare errors at message counts differing by an absolute constant.

Write $\varepsilon_{\mathrm{NS}}(M,V)$ for the minimum average error in transmitting one of M equiprobable messages through V with non-signaling assistance, and $\varepsilon_{\mathrm{MC}}(M,V)$ for its meta-converse relaxation. Their optimization programs are given in Section 2.

Theorem 1 (One-shot comparison)**.** *For every finite cq channel V and integer $M \geq 2$,*

$$\varepsilon_{\mathrm{MC}}(M,V) \leq \varepsilon_{\mathrm{NS}}(M,V) \leq \varepsilon_{\mathrm{MC}}(M+6,V). \tag{2}$$

The bound is uniform in the input alphabet and output dimension and imposes no rank assumption on the output states. The constant six is sufficient for the spectral estimates in the proof; we do not claim that it is optimal.

To state the asymptotic consequence, let $W : x \mapsto \rho_x$ have output space $\mathcal{H}$. Write $\mathcal{P}(\mathcal{X})$ for the input probability distributions and $\mathcal{S}(\mathcal{H})$ for the density operators on $\mathcal{H}$. All logarithms are base two. With $S(\rho) = -\operatorname{Tr}\rho\log_2\rho$, the Holevo capacity is

$$C(W) = \max_{p\in\mathcal{P}(\mathcal{X})} \left\{ S\left(\sum_x p(x)\rho_x\right) - \sum_x p(x)S(\rho_x) \right\}. \tag{3}$$

For $0 < \alpha < 1$, define the Petz Rényi divergence and the order-α channel capacity by

$$D_\alpha(\rho\|\sigma) = \frac{1}{\alpha-1}\log_2 \operatorname{Tr}\rho^\alpha\sigma^{1-\alpha}, \tag{4}$$

$$C_\alpha(W) = \max_{p\in\mathcal{P}(\mathcal{X})} \inf_{\sigma\in\mathcal{S}(\mathcal{H})} \sum_{x:p(x)>0} p(x)D_\alpha(\rho_x\|\sigma). \tag{5}$$

The divergence is $+\infty$ when the trace in (4) vanishes. The continuous extension at $\alpha = 1$ is $C_1(W) = C(W)$. The sphere-packing function is

$$E_{\mathrm{sp}}(R) = \sup_{0<\alpha\leq 1} \frac{1-\alpha}{\alpha}(C_\alpha(W) - R), \tag{6}$$

where the $\alpha = 1$ term is defined to be zero.

Set $M_n(R) = \lceil 2^{nR} \rceil$ and let I_2 denote a noiseless classical bit channel. When the limits exist, put

$$E_{\mathrm{NS}}(R) = \lim_{n\to\infty} -\frac{1}{n}\log_2 \varepsilon_{\mathrm{NS}}(M_n(R), W^{\otimes n}), \tag{7}$$

$$E_{\mathrm{ANS}}(R) = \lim_{n\to\infty} -\frac{1}{n}\log_2 \varepsilon_{\mathrm{NS}}(M_n(R), W^{\otimes n}\otimes I_2). \tag{8}$$

Here I_2 is used once for the entire block, not once per channel use, and we use the convention $-\log_2 0 = +\infty$.

Theorem 2 (Reliability without activation)**.** *Let W be a finite cq channel whose output states are full rank. For every $0 < R < C(W)$, the limits in* (7) *and* (8) *exist and satisfy*

$$E_{\mathrm{NS}}(R) = E_{\mathrm{ANS}}(R) = E_{\mathrm{sp}}(R). \tag{9}$$

The new ingredient is the one-shot comparison, rather than the meta-converse exponent, which is supplied by [4]. Its proof is an effect-completion argument. Adding the same positive normalization defect to every effect can violate the upper bound by the identity. Instead, we add a correction of the form $(I - T_x)^{1/2} X (I - T_x)^{1/2}$. This preserves the effect bounds whenever $0 \le X \le I$ and turns the normalization constraint into a linear equation. A norm estimate and the six-message slack place its solution inside that interval. Full rank is used only in the asymptotic argument, to ensure that the sphere-packing function is finite and continuous at every positive rate.

2 Coding programs and effect completion

Let $V : x \mapsto \omega_x$ be a finite cq channel. An effect on its output space is an operator T satisfying $0 \le T \le I$, where inequalities between Hermitian operators are in the positive-semidefinite order. For a real parameter $L \ge 1$, the meta-converse program is

$$\begin{aligned} \varepsilon_{\mathrm{MC}}(L, V) = \min_{q,T} \quad & \sum_x q_x \operatorname{Tr} \omega_x (I - T_x) \\ \text{subject to} \quad & q \in \mathcal{P}(\mathcal{X}), \qquad 0 \le T_x \le I, \\ & \sum_x q_x T_x \le I/L. \end{aligned} \tag{10}$$

For integer L, the non-signaling program has the same objective and replaces the last inequality by equality. These are the cq specializations of the assisted coding programs in [2, 3]; in the notation of [4, Eq. (29) and Eq. (35)], the change of variables is $p_x = L q_x$ and $\Lambda_x = L q_x T_x$. At indices with $q_x = 0$, the choice of T_x is immaterial. The feasible sets in (q, T) are compact, so the minima are attained. Equivalently, substituting $S_x = q_x T_x$ gives semidefinite programs with constraints $0 \le S_x \le q_x I$. The operational interpretation of the equality constraint is recalled in Appendix A.

Both error functions are nondecreasing in the message count and vanish at one message. For the meta-converse, monotonicity follows by inclusion of the feasible sets. For the non-signaling program, let $1 \le L \le N$ be integers with $N > 1$, and let (q, T) be feasible at N. The effects $T'_x = T_x + t(I - T_x)$, where

$$t = \frac{1/L - 1/N}{1 - 1/N} \in [0, 1],$$

satisfy $T'_x \ge T_x$ and $\sum_x q_x T'_x = I/L$. They are therefore feasible at L and have no larger error, proving the claimed monotonicity.

The next lemma supplies a completion with the stronger requirement that every original effect be increased. It is independent of the channel states.

Lemma 1 (Effect completion)**.** *Let $M \ge 2$ be an integer and put $K = M + 6$. Suppose q is a probability distribution on a finite set and T_x are effects on a finite-dimensional Hilbert space such that*

$$A := \sum_x q_x T_x \le I/K. \tag{11}$$

Then there are effects Q_x satisfying

$$T_x \le Q_x \le I, \qquad \sum_x q_x Q_x = I/M. \tag{12}$$

Proof. Set $C_x = (I - T_x)^{1/2}$, $D_x = I - C_x$, and $B = \sum_x q_x D_x$. For $0 \le t \le 1$, we have $0 \le 1 - \sqrt{1-t} \le t \le 1$. Spectral calculus therefore gives $0 \le D_x^2 \le D_x \le T_x$, and hence

$$0 \le B \le A \le I/K, \qquad \sum_x q_x D_x^2 \le A \le I/K. \tag{13}$$

On the real vector space of Hermitian operators, equipped with the operator norm $\|\cdot\|_\infty$, define

$$\Phi(X) = \sum_x q_x C_x X C_x, \qquad \Psi(X) = \sum_x q_x D_x X D_x.$$

Expanding $C_x = I - D_x$ gives

$$\Phi(X) = X - BX - XB + \Psi(X). \tag{14}$$

For Hermitian X, the inequalities $-\|X\|_\infty I \le X \le \|X\|_\infty I$ imply

$$-\|X\|_\infty \sum_x q_x D_x^2 \le \Psi(X) \le \|X\|_\infty \sum_x q_x D_x^2.$$

Consequently, $\|\Psi(X)\|_\infty \le \|X\|_\infty / K$. Together with $\|B\|_\infty \le 1/K$, equation (14) yields the induced-norm bound

$$\|\mathrm{id} - \Phi\|_{\infty\to\infty} \le 3/K < 1. \tag{15}$$

To obtain the required normalization, we solve $\Phi(X) = H$ with $H = I/M - A$. By (15), the Neumann series

$$X = \Phi^{-1}(H) = \sum_{j=0}^{\infty} (\mathrm{id} - \Phi)^j (H) \tag{16}$$

converges in operator norm to a Hermitian solution. Since $0 \le H \le I/M$, its distance from H satisfies

$$\|X - H\|_\infty \le \frac{3/K}{1 - 3/K} \frac{1}{M} = \frac{3}{M(K-3)}. \tag{17}$$

Using $H \ge (1/M - 1/K) I$ and $K = M + 6$, we obtain

$$X \ge \left(\frac{1}{M} - \frac{1}{M+6} - \frac{3}{M(M+3)} \right) I = \frac{3}{(M+6)(M+3)} I > 0, \tag{18}$$

$$X \le \left(\frac{1}{M} + \frac{3}{M(M+3)} \right) I = \frac{M+6}{M(M+3)} I \le \frac{4}{5} I < I. \tag{19}$$

Thus X lies in the effect interval. This conclusion uses the spectral margins above, not positivity of the inverse map Φ^{-1}.

Define

$$Q_x = T_x + C_x X C_x. \tag{20}$$

Then $Q_x - T_x = C_x X C_x \ge 0$ and $I - Q_x = C_x (I - X) C_x \ge 0$. Finally,

$$\sum_x q_x Q_x = A + \Phi(X) = A + H = I/M,$$

which proves (12). □

Proof of Theorem 1. Apply Lemma 1 to a minimizing pair (q, T) in (10) with $L = M + 6$. The completed pair (q, Q) is feasible for the non-signaling program at M. Since $Q_x \ge T_x$ and $\omega_x \ge 0$, we have

$$\mathrm{Tr}\, \omega_x (I - Q_x) \le \mathrm{Tr}\, \omega_x (I - T_x)$$

for every x. Averaging with the unchanged distribution q proves $\varepsilon_{\mathrm{NS}}(M, V) \le \varepsilon_{\mathrm{MC}}(M + 6, V)$. The other inequality in (2) follows by relaxing the non-signaling equality constraint. No rank or commutation assumption on the channel outputs is used. □

3 Passage to reliability exponents

We use the meta-converse exponent established by Oufkir, Tomamichel, and Berta [4, Theorem 4.1 and Eq. (34)]. The remaining issue is to justify that the fixed message-count change in Theorem 1, and the rounding required for actual codes, do not change the limit.

Proof of Theorem 2. Let $d = \dim \mathcal{H}$ and $\lambda = \min_x \lambda_{\min}(\rho_x) > 0$. Evaluating the infimum in (5) at $\sigma = I/d$ gives

$$C_\alpha(W) \le -\frac{\alpha}{1-\alpha} \log_2(d\lambda), \qquad 0 < \alpha < 1. \tag{21}$$

Indeed, $\operatorname{Tr} \rho_x^\alpha \ge d\lambda^\alpha$, so $\operatorname{Tr} \rho_x^\alpha (I/d)^{1-\alpha} \ge (d\lambda)^\alpha$; division by $\alpha - 1 < 0$ yields (21). It follows that, for every $r > 0$,

$$0 \le E_{\mathrm{sp}}(r) \le -\log_2(d\lambda) < \infty. \tag{22}$$

As a supremum of affine functions of r, E_{sp} is convex and therefore continuous on $(0, \infty)$. Moreover, $C_0(W) := \lim_{\alpha \downarrow 0} C_\alpha(W) = 0$, by (21) and nonnegativity of the Rényi divergence.

The achievability and converse bounds in [4, Propositions 4.2 and 4.3], together with its meta-converse identification, therefore apply throughout $0 < r < C(W)$. In base-two units, they give

$$\lim_{n\to\infty} -\frac{1}{n} \log_2 \varepsilon_{\mathrm{MC}}(2^{nr}, W^{\otimes n}) = E_{\mathrm{sp}}(r), \qquad 0 < r < C(W). \tag{23}$$

Here the real-valued message parameter is interpreted through (10). The limit is ordinary: the two cited propositions provide matching upper and lower asymptotic bounds. The errors in (23) are positive. More generally, for $L > 1$, any feasible pair in the block-channel program satisfies

$$\sum_{x^n} q_{x^n} \operatorname{Tr} \rho_{x^n}(I - T_{x^n}) \ge \lambda^n \operatorname{Tr}\left(I - \sum_{x^n} q_{x^n} T_{x^n}\right) \ge (d\lambda)^n (1 - 1/L) > 0,$$

where $\rho_{x^n} = \rho_{x_1} \otimes \cdots \otimes \rho_{x_n}$.

We first extend (23) to arbitrary message-count sequences. Suppose $L_n \ge 1$ and $n^{-1} \log_2 L_n \to r \in (0, C(W))$. For every $0 < \delta < \min\{r, C(W) - r\}$, eventually

$$2^{n(r-\delta)} \le L_n \le 2^{n(r+\delta)}.$$

Monotonicity of $\varepsilon_{\mathrm{MC}}$ and (23) imply

$$E_{\mathrm{sp}}(r+\delta) \le \liminf_{n\to\infty} -\frac{1}{n} \log_2 \varepsilon_{\mathrm{MC}}(L_n, W^{\otimes n}),$$
$$\limsup_{n\to\infty} -\frac{1}{n} \log_2 \varepsilon_{\mathrm{MC}}(L_n, W^{\otimes n}) \le E_{\mathrm{sp}}(r-\delta).$$

Letting $\delta \downarrow 0$ and using continuity gives

$$\lim_{n\to\infty} -\frac{1}{n} \log_2 \varepsilon_{\mathrm{MC}}(L_n, W^{\otimes n}) = E_{\mathrm{sp}}(r). \tag{24}$$

Now fix $0 < R < C(W)$ and set $V_n = W^{\otimes n}$. Theorem 1 gives

$$\varepsilon_{\mathrm{MC}}(M_n(R), V_n) \le \varepsilon_{\mathrm{NS}}(M_n(R), V_n) \le \varepsilon_{\mathrm{MC}}(M_n(R) + 6, V_n). \tag{25}$$

Both bounding message counts have normalized logarithm tending to R. Applying (24) to the two bounds proves the existence of $E_{\mathrm{NS}}(R)$ and the equality $E_{\mathrm{NS}}(R) = E_{\mathrm{sp}}(R)$.

For the activated model, the one-bit activation identity [3, 4] states that, for every integer $m \geq 1$,

$$\varepsilon_{\mathrm{NS}}(2m, V_n \otimes I_2) = \varepsilon_{\mathrm{MC}}(m, V_n); \tag{26}$$

see [4, Eq. (32)]. Thus, for every integer $L \geq 2$, monotonicity of the non-signaling error gives

$$\varepsilon_{\mathrm{MC}}(\lfloor L/2 \rfloor, V_n) \leq \varepsilon_{\mathrm{NS}}(L, V_n \otimes I_2) \leq \varepsilon_{\mathrm{MC}}(\lceil L/2 \rceil, V_n). \tag{27}$$

With $L = M_n(R)$, both bounding message counts again have normalized logarithm tending to R. Equation (24) proves that $E_{\mathrm{ANS}}(R)$ exists and equals $E_{\mathrm{sp}}(R)$, completing the proof. $\square$

References

[1] William Matthews. A linear program for the finite block length converse of Polyanskiy–Poor–Verdú via nonsignaling codes. *IEEE Trans. Inf. Theory* **58**(12), 7036–7044 (2012). doi:10.1109/TIT.2012.2210695.

[2] Xin Wang, Wei Xie, and Runyao Duan. Semidefinite programming strong converse bounds for classical capacity. *IEEE Trans. Inf. Theory* **64**(1), 640–653 (2018). doi:10.1109/TIT.2017.2741101.

[3] Xin Wang, Kun Fang, and Marco Tomamichel. On converse bounds for classical communication over quantum channels. *IEEE Trans. Inf. Theory* **65**(7), 4609–4619 (2019). doi:10.1109/TIT.2019.2898656.

[4] Aadil Oufkir, Marco Tomamichel, and Mario Berta. Error exponent of activated non-signaling-assisted classical-quantum channel coding. *Lett. Math. Phys.* **116**, article 33 (2026). doi:10.1007/s11005-026-02061-z. Author version: arXiv:2410.01084v2 (2024), https://arxiv.org/abs/2410.01084v2. The theorem, proposition, and equation numbers cited here refer to this author version.

[5] Marco Dalai. Lower bounds on the probability of error for classical and classical-quantum channels. *IEEE Trans. Inf. Theory* **59**(12), 8027–8056 (2013). doi:10.1109/TIT.2013.2283794.

[6] Joseph M. Renes. Tight lower bound on the error exponent of classical-quantum channels. *IEEE Trans. Inf. Theory* **71**(1), 530–538 (2025). doi:10.1109/TIT.2024.3500578.

[7] Ke Li and Dong Yang. Reliability function of classical-quantum channels. *Phys. Rev. Lett.* **134**, 010802 (2025). doi:10.1103/PhysRevLett.134.010802.

[8] Hao-Chung Cheng and Po-Chieh Liu. Error exponents for quantum packing problems via an operator layer cake theorem. arXiv:2507.06232v4 (2026). https://arxiv.org/abs/2507.06232v4.

[9] Aadil Oufkir and Mario Berta. Quantum channel coding: Approximation algorithms and strong converse exponents. *Quantum* **9**, 1877 (2025). doi:10.22331/q-2025-10-06-1877.

A Operational form of the non-signaling program

For completeness, we describe the non-signaling operation associated with the equality constraint in Section 2. Alice receives a classical message $m \in \{1, \ldots, M\}$ and outputs a channel input x; Bob receives the quantum output of the channel and outputs an estimate $\widehat{m}$. Such an operation is specified by positive operators $F_{x,\widehat{m}|m}$ on Bob's input space. For each m, they sum to I over $(x, \widehat{m})$. No signaling from Bob to Alice requires $\sum_{\widehat{m}} F_{x,\widehat{m}|m} = q(x \mid m)I$ for a probability distribution $q(\cdot \mid m)$, and no signaling from Alice to Bob requires $\sum_x F_{x,\widehat{m}|m}$ to be independent of m.

Given a feasible pair (q, Q) with $M \geq 2$, define

$$F_{x,\widehat{m}|m} = \begin{cases} q_x Q_x, & \widehat{m} = m, \\ \dfrac{q_x(I - Q_x)}{M - 1}, & \widehat{m} \neq m. \end{cases} \tag{28}$$

These operators are positive and satisfy

$$\sum_{\widehat{m}} F_{x,\widehat{m}|m} = q_x I, \qquad \sum_x F_{x,\widehat{m}|m} = I/M.$$

They therefore define a non-signaling operation. Connecting Alice's output x to the channel $V : x \mapsto \omega_x$ gives average success probability

$$\frac{1}{M} \sum_{m=1}^{M} \sum_x \operatorname{Tr} \omega_x F_{x,m|m} = \sum_x q_x \operatorname{Tr} \omega_x Q_x.$$

Conversely, average any non-signaling operation over simultaneous permutations of the message and estimate labels. This preserves its average success probability and both non-signaling conditions. The resulting operators have the form G_x when $\widehat{m} = m$ and H_x otherwise. The Bob-to-Alice condition gives $G_x + (M-1)H_x = q_x I$ for a distribution q. The Alice-to-Bob condition and normalization give $\sum_x G_x = \sum_x H_x = I/M$. Thus $0 \leq G_x \leq q_x I$, and setting $Q_x = G_x/q_x$ for $q_x > 0$, with $Q_x = 0$ otherwise, yields $0 \leq Q_x \leq I$ and $\sum_x q_x Q_x = I/M$. This recovers the equality program and its objective.