

# Retrocausal Capacity of a Quantum Channel

Kaiyuan Ji,<sup>1,2,\*</sup> Seth Lloyd,<sup>2,†</sup> and Mark M. Wilde<sup>1,‡</sup>

<sup>1</sup>*School of Electrical and Computer Engineering, Cornell University, Ithaca, New York 14850, USA*

<sup>2</sup>*Department of Mechanical Engineering, Massachusetts Institute of Technology,  
77 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA*

(Dated: April 17, 2026)

We study the capacity of a quantum channel for retrocausal communication, where messages are transmitted backward in time, from a sender in the future to a receiver in the past, through a noisy postselected closed timelike curve (P-CTC) mathematically represented by the channel. We completely characterize the one-shot retrocausal quantum and classical capacities, and we show that the corresponding asymptotic capacities are equal to the average and sum, respectively, of the channel’s max-information and its regularized Doebelin information. This endows these information measures with a novel operational interpretation. Furthermore, our characterization can be generalized beyond quantum channels to all completely positive maps. This imposes information-theoretic limits on transmitting messages via postselected-teleportation-like mechanisms with arbitrary initial- and final-state boundary conditions, including those considered in various black-hole final-state models.

**Introduction.**—The theory of general relativity admits of the existence of closed timelike curves (CTCs) [1–4]. Such a curve sends a “chronology-violating” system backward in time, which can then interact with surroundings of its own past. The existence of CTCs could lead to perplexing consequences such as violation of causality and escape from a black hole [5–8], and it is challenged by grandfather-like paradoxes. To formalize the discussion of these consequences and mitigate paradoxes, abstract quantum-mechanical models of CTCs have been developed, with two notable examples being the Deutschian model [9] based on a self-consistency condition and the postselected model [10–13] based on a postselected version of quantum teleportation [14]. Access to a noiseless CTC of one kind or the other has been shown to unleash stunning information-processing power, such as the ability to distinguish nonorthogonal states perfectly [15, 16], to clone [17], and to speed up computations [11, 18, 19] (although its interpretation has been the subject of debate [20–22]). Recent works have also demonstrated advantages brought by resources simulable using noiseless postselected CTCs (P-CTCs) in other tasks such as metrology [23–25], resource manipulation [26], hypothesis testing [27], and communication [28–30].

In this paper, we investigate fundamental limits on the efficiency of communication [31] through *noisy* P-CTCs. A noisy P-CTC can mathematically be represented by a quantum channel, in the sense that information traveling through the noisy P-CTC effectively evolves as if passing through the channel, except for ending up in the past. Therefore, another way of phrasing the question is: how does the capacity of a quantum channel change, if it were to be used for transmitting messages backward in time (as opposed to forward in time as in conventional communication settings [31–42])? We term this unconventional communication setting *retrocausal communication*. A distinctive feature of retrocausal communication is the possible formation of causal loops, which is counterintuitive, yet perfectly reasonable granted the existence of the noisy P-CTC. In fact, the communication power of the noisy P-CTC can only be fully unlocked by exploiting such an exotic structure of spacetime. Our main technical result is to estab-

lish the Shannon-theoretic limit of retrocausal communication, accompanied by an explicit construction of an optimal strategy in the one-shot regime.

Our use of the postselected model of CTCs is motivated by several considerations. P-CTCs, based on postselected teleportation, can be simulated experimentally with reasonable overhead [11], and they have been shown to be consistent [12] with path-integral approaches [43–48] to CTCs. Most importantly, a noiseless P-CTC amounts to an identity channel from the future to the past, suggesting that the model is consistent with the formalism of quantum Shannon theory [42, 49–51] and suitable for studying CTCs as media for communication. This contrasts with the Deutschian model of CTCs [9], in which the chronology-violating system becomes decorrelated from the environment after traveling backward in time, so that partial information (such as one part of an entangled state) can never be faithfully transmitted from the future to the past. Moreover, P-CTCs are closely related to the Horowitz-Maldacena final-state model of black-hole evaporation [5], which was proposed to reconcile unitarity with Hawking’s semiclassical arguments [52] and inspired by a time-symmetric modification of quantum mechanics [53–55]. In variations of the model [6–8] that allow for arbitrary initial- or final-state boundary conditions, nonunitary postselected-teleportation-like mechanisms resembling noisy P-CTCs could emerge.

**Noisy P-CTCs.**—A chronology-violating system traveling through a noiseless P-CTC [10–13] can be thought of as being teleported to its past under postselecting the event that the Bell measurement in the teleportation protocol [14] yields the desirable outcome that requires no further action to complete the protocol. This time-travel mechanism is “noiseless,” as the past and future versions of the system are identical. Now consider a noisy mechanism of time travel, which turns a future system  $A$  into a past system  $B$  under the effect of a quantum channel  $\mathcal{N}_{A \rightarrow B}$  [56]. A natural such mechanism is to let the future system travel through a noiseless P-CTC and then be acted upon ordinarily by the channel. We refer to this mechanism as a “noisy” P-CTC and simply represent it by the channel  $\mathcal{N}_{A \rightarrow B}$ , with the implicit assumption that  $B$  precedes  $A$  in time.

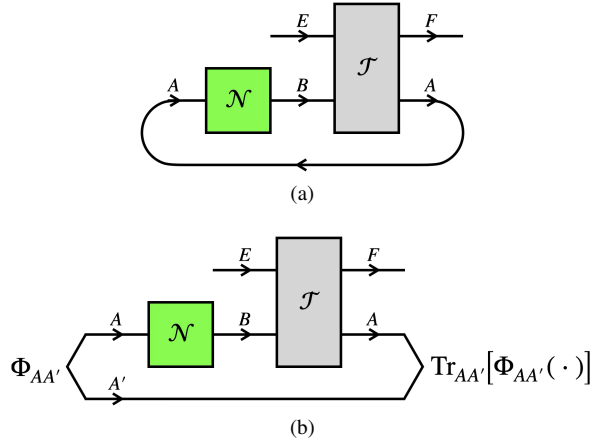


FIG. 1. Enclosing a bipartite channel  $\mathcal{T}_{EB \rightarrow FA}$  with a noisy P-CTC represented by a channel  $\mathcal{N}_{A \rightarrow B}$  (Fig. 1a) is equivalent to preprocessing  $\mathcal{T}_{EB \rightarrow FA}$  with  $\mathcal{N}_{A \rightarrow B}$  and performing postselected teleportation from the output  $A$  of  $\mathcal{T}_{EB \rightarrow FA}$  to the input  $A$  of  $\mathcal{N}_{A \rightarrow B}$  (Fig. 1b with renormalization).

P-CTCs can be combined with other “chronology-respecting” processes. For instance, given a noisy P-CTC represented by  $\mathcal{N}_{A \rightarrow B}$ , one can feed  $B$  into a bipartite channel  $\mathcal{T}_{EB \rightarrow FA}$  as part of the latter’s input and then feed the output  $A$  of  $\mathcal{T}_{EB \rightarrow FA}$  into the noisy P-CTC’s entrance (see Fig. 1a). This might result in a causal loop  $B \rightarrow A \rightarrow B$ , and yet it is perfectly reasonable granted the existence of the noisy P-CTC, because  $B$  can influence  $A$  via  $\mathcal{T}_{EB \rightarrow FA}$  without violating chronology. Potential paradoxes arising from such a loop can be addressed within the model of P-CTCs [11]. The resulting transformation from  $E$  to  $F$  is then represented by the following nonlinear map (see Fig. 1b) [11, 12]:

$$\rho_{RE} \mapsto \frac{\text{Tr}_{AA'}[\Phi_{AA'} \mathcal{T}_{EB \rightarrow FA}[\rho_{RE} \otimes \mathcal{N}_{A \rightarrow B}[\Phi_{AA'}]]]}{\text{Tr}[\Phi_{AA'} \mathcal{T}_{EB \rightarrow FA}[\rho_{RE} \otimes \mathcal{N}_{A \rightarrow B}[\Phi_{AA'}]]]}, \quad (1)$$

where  $R$  is an arbitrary reference system and  $\Phi_{AA'} \equiv (1/d_A) \sum_{i,j} |i\rangle\langle j|_A \otimes |i\rangle\langle j|_{A'}$  is the standard maximally entangled state between  $A$  and  $A'$ .

**Retrocausal communication.**—The setting of retrocausal communication can perhaps be best explained in reference to the film *Interstellar* [57] (see Fig. 2). In the film, a father in the future discovers that he can communicate to his young daughter in the past via some noisy mechanism. Due to its chronology-violating nature, we model this mechanism as a noisy P-CTC represented by a quantum channel  $\mathcal{N}_{A \rightarrow B}$ . We are concerned with the maximum possible efficiency at which messages can be reliably transmitted from the father to the daughter through the noisy P-CTC under an encoding and decoding of their choice. This efficiency is measured by the *retrocausal capacity* of  $\mathcal{N}_{A \rightarrow B}$ , which we will define shortly [58]. One may wonder how this exotic scenario differs from a conventional communication setting in which messages are transmitted forward in time. The difference becomes manifest when taking into account the following possibility (as is also visible in the film [57]): the father, who is in the future, may retrieve his

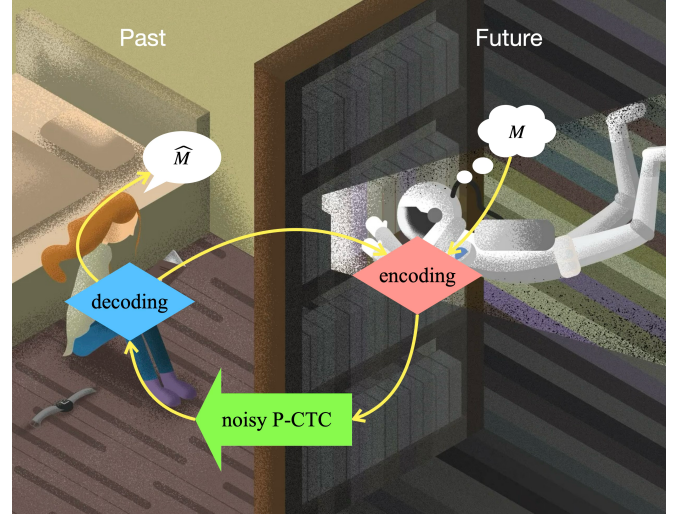


FIG. 2. In the film *Interstellar* [57], a father, who is in the future, encodes his message to his daughter, who is in the past, in a system traveling through a noisy P-CTC. The father can retrieve and consult his memory of past events he has witnessed, so that his encoding could be influenced by what results from the daughter’s decoding, forming a causal loop. Background artwork: *Bookshelf Scene* (2018), ©Alexey Nefed (motion.nalex), reproduced with permission [59].

memory of past events he has witnessed, even including the daughter’s decoding of the message which he is about to send! It would thus not be surprising that he will consult his memory of the daughter’s decoding when encoding his message, so as to maximize the efficiency of the communication.

Evidently, this allows what results from the daughter’s decoding to influence the father’s encoding, which in turn influences the daughter’s decoding via the noisy P-CTC, thus forming a causal loop. However, to stress again, granted the existence of the noisy P-CTC, the formation of this loop is physically justified: the father and the daughter operate in a perfectly chronology-respecting manner, including the father’s retrieving and consulting his memory of past events; the only chronology-violating element here is the noisy P-CTC, which models a given mechanism beyond their control. Note that, if the father is not allowed to retrieve or consult his memory, then the setting under consideration would not be different (at least mathematically) from conventional (forward-in-time) communication settings, since everything would anyway abide by a definite causal order and could thus be realized without necessarily violating chronology.

Following the intuition explained above, the most general form of strategy that the father and the daughter can adopt, without violating chronology (when not invoking the noisy P-CTC), can be described as follows (see Fig. 3a): (i) in the past, the daughter receives  $B$  from the exit of the noisy P-CTC, decodes the message by applying a channel  $\mathcal{D}_{B \rightarrow \hat{M}L}$  with  $\hat{M}$  storing her recovery of the message, and leaves  $L$  in a noiseless quantum memory for the father to retrieve in the future; (ii) in the future, the father retrieves  $L$ , encodes his message stored in  $M$  by applying a channel  $\mathcal{E}_{ML \rightarrow A}$ , and feeds  $A$  into the

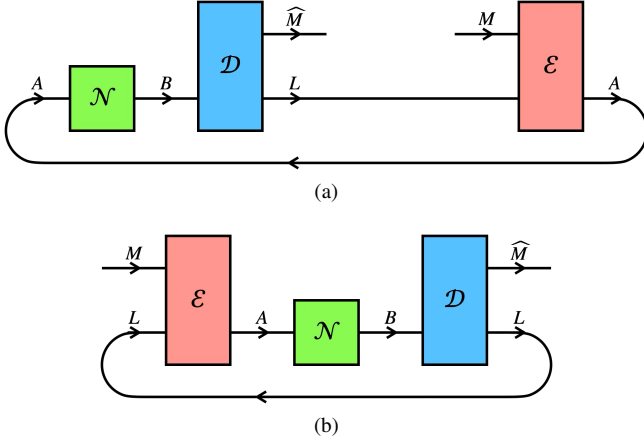


FIG. 3. Retrocausal communication through a noisy P-CTC represented by a channel  $\mathcal{N}_{A \rightarrow B}$  (Fig. 3a) is mathematically equivalent to forward-in-time communication through the channel  $\mathcal{N}_{A \rightarrow B}$  with the assistance of a noiseless P-CTC (Fig. 3b).

entrance of the noisy P-CTC. The resulting transformation from the father's message system to the daughter's recovery system is represented by a nonlinear map  $\mathcal{M}_{M \rightarrow \widehat{M}}$ , whose action on an arbitrary pure state  $\psi_{RM}$ , with  $R$  an arbitrary reference system, is given by

$$\mathcal{M}_{M \rightarrow \widehat{M}}[\psi_{RM}] \equiv \frac{\text{Tr}_{AA'}[\Phi_{AA'} \mathcal{E}_{ML \rightarrow A}[\psi_{RM} \otimes \mathcal{D}_{B \rightarrow L\widehat{M}} \circ \mathcal{N}_{A \rightarrow B}[\Phi_{AA'}]]]}{\text{Tr}[\Phi_{AA'} \mathcal{E}_{ML \rightarrow A}[\psi_{RM} \otimes \mathcal{D}_{B \rightarrow L\widehat{M}} \circ \mathcal{N}_{A \rightarrow B}[\Phi_{AA'}]]]}, \quad (2)$$

following the machinery of Eq. (1).

Another observation is that, in the presence of a causal loop, the notion of past and future becomes elusive and can be treated flexibly. For instance, if one imagines the “first-hand experience” of a message traveling through the noisy P-CTC, then the conclusion would be that the father resides in the “past” where it sets out and the daughter resides in the “future” where it arrives. Judging from this experience, the noisy P-CTC would just appear as an ordinary channel, whereas the memory  $L$  travels through a noiseless P-CTC. This suggests that, mathematically, retrocausal communication (see Fig. 3a) is equivalent to forward-in-time communication assisted by a noiseless P-CTC (see Fig. 3b). For coherence of presentation, we will continue with our original formulation of retrocausal communication (see Fig. 3a), where the father is in the future and the daughter is in the past.

**Retrocausal capacity.**—To quantify the quality of transmitting quantum messages via the nonlinear map  $\mathcal{M}_{M \rightarrow \widehat{M}}$  [see Eq. (2)], we define its worst-case quantum infidelity as

$$q_{\text{infid}}(\mathcal{M}) := 1 - \inf_{\psi_{RM}} \text{Tr}[\psi_{RM} \widehat{\mathcal{M}}_{M \rightarrow \widehat{M}}[\psi_{RM}]], \quad (3)$$

where the infimization is over all pure states with  $R$  of an arbitrary dimension. We then define the *one-shot retrocausal quantum capacity* of the quantum channel  $\mathcal{N}_{A \rightarrow B}$  as the maximum

number of qubits that can be transmitted within an infidelity  $\varepsilon \in [0, 1]$  through the noisy P-CTC it represents:

$$Q_{\text{retro}}^{\varepsilon}(\mathcal{N}) := \sup_{\mathcal{E}_{ML \rightarrow A}, \mathcal{D}_{B \rightarrow \widehat{M}L}} \{\log_2 d_M : q_{\text{infid}}(\mathcal{M}) \leq \varepsilon\}, \quad (4)$$

where  $\mathcal{M}_{M \rightarrow \widehat{M}}$  is given by Eq. (2) and the supremization is over all choices of encoding and decoding (which are quantum channels) with  $M$  and  $L$  of arbitrary dimensions. We define the *asymptotic retrocausal quantum capacity* of  $\mathcal{N}_{A \rightarrow B}$  as

$$Q_{\text{retro}}(\mathcal{N}) := \inf_{\varepsilon \in (0, 1)} \liminf_{n \rightarrow \infty} \frac{1}{n} Q_{\text{retro}}^{\varepsilon}(\mathcal{N}^{\otimes n}). \quad (5)$$

If the messages being transmitted are classical, then one can disregard the reference system  $R$  and only consider all (perfectly distinguishable) symbols in an alphabet of size  $d_M$ . The one-shot retrocausal classical capacity  $C_{\text{retro}}^{\varepsilon}(\mathcal{N})$  is then defined similarly to its quantum counterpart but with the worst-case classical error probability

$$p_{\text{error}}(\mathcal{M}) := 1 - \min_m \text{Tr}[|m\rangle\langle m|_{\widehat{M}} \mathcal{M}_{M \rightarrow \widehat{M}}[|m\rangle\langle m|_M]] \quad (6)$$

in place of  $q_{\text{infid}}(\mathcal{M})$ , where the minimization is over all symbols in the alphabet. The asymptotic retrocausal classical capacity  $C_{\text{retro}}(\mathcal{N})$  is defined accordingly.

To reiterate, the retrocausal capacities defined above assume that the channel  $\mathcal{N}_{A \rightarrow B}$  represents a noisy P-CTC, with  $B$  preceding  $A$  in time. They do not apply to other models of CTCs [9], nor do they suggest backward-in-time signaling in standard quantum theory, where chronology-violating resources are forbidden.

**Main result.**—For a quantum channel  $\mathcal{N}_{A \rightarrow B}$  and a real number  $\varepsilon \in (0, 1)$ , the one-shot retrocausal quantum and classical capacities are equal to, respectively,

$$Q_{\text{retro}}^{\varepsilon}(\mathcal{N}) = \log_2 \left\lfloor \sqrt{\frac{\varepsilon}{1-\varepsilon}} 2^{I_{\text{max}}(\mathcal{N}) + I_{\text{doe}}(\mathcal{N})} + 1 \right\rfloor, \quad (7)$$

$$C_{\text{retro}}^{\varepsilon}(\mathcal{N}) = \log_2 \left\lfloor \frac{\varepsilon}{1-\varepsilon} 2^{I_{\text{max}}(\mathcal{N}) + I_{\text{doe}}(\mathcal{N})} + 1 \right\rfloor, \quad (8)$$

where  $\lfloor \cdot \rfloor$  denotes the floor function. The max-information [60] is defined as

$$I_{\text{max}}(\mathcal{N}) := \inf_{\lambda, \sigma_B} \{\log_2 \lambda : \mathcal{N}_{A \rightarrow B}[\Phi_{A'A}] \leq \lambda \pi_{A'} \otimes \sigma_B\}, \quad (9)$$

where  $\pi_{A'} \equiv \mathbb{1}_{A'}/d_{A'}$  and the infimization is over all real numbers and all quantum states of  $B$ , and the Doeblin information [61] is defined as

$$I_{\text{doe}}(\mathcal{N}) := \inf_{\lambda, \tau_B} \{\log_2 \lambda : \pi_{A'} \otimes \tau_B \leq \lambda \mathcal{N}_{A \rightarrow B}[\Phi_{A'A}]\}, \quad (10)$$

where the infimization is over all real numbers and all unit-trace Hermitian operators in  $B$ .

The asymptotic retrocausal quantum and classical capacities are given by

$$Q_{\text{retro}}(\mathcal{N}) = \frac{1}{2} C_{\text{retro}}(\mathcal{N}) = \frac{1}{2} (I_{\text{max}}(\mathcal{N}) + I_{\text{doe}}^{\infty}(\mathcal{N})), \quad (11)$$

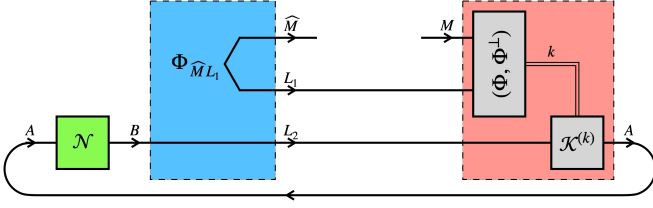


FIG. 4. Amplified probabilistic teleportation is an optimal strategy for retrocausal communication. The blue and red boxes encapsulate the daughter's decoding and the father's encoding, respectively.

where  $I_{\text{doe}}^{\infty}(\mathcal{N}) := \lim_{n \rightarrow \infty} (1/n) I_{\text{doe}}(\mathcal{N}^{\otimes n})$  is the regularized Doebelin information.

Several implications of our result are in order. First, our result provides exact closed-form expressions for the one-shot retrocausal quantum and classical capacities. This is notable considering that having a one-shot characterization up to this precision has been extremely rare in quantum Shannon theory. Second, our result implies a complete characterization of the error exponent and the strong-converse exponent of retrocausal communication (see Supplemental Material [62, Secs. VD and VID] for formal statements). Third, our result provides a novel operational interpretation for the max- and Doebelin informations in the context of communication, contrasting, for instance, the previously known interpretation of the max-information in channel simulation (which is a setting reverse to communication) [41] (see Supplemental Material [62, Sec. II] for more discussion). Fourth, for several classes of channels (including measurement channels and certain covariant channels), the Doebelin information is additive [61], implying that  $I_{\text{doe}}^{\infty}(\mathcal{N}) = I_{\text{doe}}(\mathcal{N})$ . This means that, for such channels, the asymptotic retrocausal capacities have a single-letter expression and are efficiently computable via semidefinite programs (SDPs) (see Supplemental Material [62, Sec. VII] for numerical examples). For general quantum channels, the Doebelin information is not always additive, and the efficient computability of its regularization is unknown. In this case, we provide single-letter and SDP-computable upper and lower bounds on the asymptotic retrocausal capacities in Supplemental Material [62, Remark 6].

**Optimal strategy.**—To provide some intuition behind the derivation of our result, we outline a construction of an optimal strategy, which achieves the one-shot retrocausal quantum capacity precisely. For a rigorous analysis of its performance and a proof of its optimality, see Supplemental Material [62, Sec. V]. The strategy operates as follows (see Fig. 4).

In the past, the daughter receives  $B$  from the exit of the noisy P-CTC. She prepares the standard maximally entangled state between  $\hat{M}$  and a noiseless quantum memory  $L_1$ , keeps  $B$  in another noiseless quantum memory  $L_2$ , and leaves both  $L_1$  and  $L_2$  for the father to retrieve in the future. This concludes the daughter's decoding. In the future, the father retrieves  $L_1$  and  $L_2$  and performs a Bell measurement on  $M$  and  $L_1$ . If the measurement ends up being a projection onto the standard maximally entangled state, then he applies a channel  $\mathcal{K}_{L_2 \rightarrow A}^{(0)}$

to  $L_2$ ; otherwise he applies  $\mathcal{K}_{L_2 \rightarrow A}^{(1)}$ . Finally, he feeds  $A$  into the entrance of the noisy P-CTC. This concludes the father's encoding.

We call the above strategy *amplified probabilistic teleportation*, and its underlying functioning can be understood as follows. Instead of trying to code his message into  $A$  which enters the noisy P-CTC, the father intends to “teleport” [14] the message probabilistically to the daughter via  $L_1$ . Indeed, if the father's Bell measurement ends up being a projection onto the standard maximally entangled state, then what the daughter ends up with in  $\hat{M}$  is a perfect recovery of the message in  $M$ . However, this “desirable” outcome occurs only with probability  $1/d_M^2$ , or at least so without invoking the noisy P-CTC. Yet, as illustrated in Fig. 4, the remaining part of the strategy engineers a causal loop solely consisting of the noisy P-CTC, represented by  $\mathcal{N}_{A \rightarrow B}$ , and the channel  $\mathcal{K}_{L_2 \rightarrow A}^{(k)}$ , conditioned on  $k \in \{0, 1\}$ . Due to the nonlinearity of such a loop (specifically, its capability of renormalizing the probabilities of certain events), the father can choose the channels  $\mathcal{K}_{L_2 \rightarrow A}^{(0)}$  and  $\mathcal{K}_{L_2 \rightarrow A}^{(1)}$  in such a way that the renormalized probability of the desirable outcome is maximally amplified. If the renormalized probability of the desirable outcome can be amplified to at least  $1 - \varepsilon$  for a message system  $M$  of dimension  $d_M$ , then  $\log_2 d_M$  qubits can indeed be “teleported” from  $M$  to  $\hat{M}$  via  $L_1$  subject to a failure probability at most  $\varepsilon$ , which upper bounds the overall infidelity of communication. The max- and Doebelin informations of  $\mathcal{N}_{A \rightarrow B}$  come to relevance in quantifying the noisy P-CTC's maximal amplification of the probability of the desirable outcome, due to their interpretations in terms of the maximum and minimum achievable singlet fractions, respectively [60, 61].

**Discussion.**—We have proposed the setting of retrocausal communication based on the model of P-CTCs, demonstrating that exchanging the temporal locations of the sender and the receiver and “reversing” the chronological direction of the communication channel shifts the paradigm of quantum Shannon theory significantly. In the one-shot regime, the fundamental limit on its efficiency can be determined precisely and achieved with a simple strategy contrasting with conventional coding schemes. In the asymptotic regime, the limit can be incarnated as a resource inequality [63] involving communication resources in both chronological directions:

$$\langle \tilde{\mathcal{N}} \rangle + \infty[q \rightarrow q] \geq \frac{1}{2} (I_{\text{max}}(\mathcal{N}) + I_{\text{doe}}^{\infty}(\mathcal{N})) [q \leftarrow q], \quad (12)$$

where  $\langle \tilde{\mathcal{N}} \rangle$  denotes a noisy P-CTC represented by  $\mathcal{N}_{A \rightarrow B}^{\otimes n}$ , and  $[q \rightarrow q]$  and  $[q \leftarrow q]$  denote a noiseless quantum memory and a noiseless P-CTC, respectively, on  $n$  qubits. The noiseless quantum memory left by the receiver in the past and retrieved by the sender in the future plays an essential role in achieving this limit. An interesting open question is how this limit is affected if the memory is noisy, say, represented by a channel  $\mathcal{L}_{C \rightarrow D}$ . In other words, one may try to characterize the tradeoff between the rates  $R$  and  $Q$  in the following resource inequality:

$$\langle \tilde{\mathcal{N}} \rangle + R \langle \tilde{\mathcal{L}} \rangle \geq Q[q \leftarrow q]. \quad (13)$$

Another future direction is to explore the relation between our result and final-state models of black-hole evaporation [5–8]. The original final-state model proposed by Horowitz and Maldacena [5] postulates that black-hole evaporation can be described by postselected teleportation, which reconciles unitarity and Hawking’s semiclassical arguments [52]. It was subsequently pointed out [6–8] that, even assuming a final-state boundary condition, unitarity could still be spoiled if the initial- or final-state boundary condition can be arbitrary, and the evaporation would then be described by a postselected-teleportation-like mechanism represented by a completely positive map that is not necessarily trace preserving (see Supplemental Material [62, Sec. III B] for details). Notably, our characterization of the retrocausal capacity can be shown to apply to all completely positive maps and not just quantum channels (see Supplemental Material [62, Secs. V and VI] for proofs), and it is therefore general enough to account for final-state models with arbitrary initial- or final-state boundary conditions. Further implications of our result for final-state models will be discussed in a future work [64].

We are grateful to Alexey Nefed (motion.nalex) for permission to reproduce his artwork *Bookshelf Scene* (2018) as the background in Fig. 2. We thank Bartosz Regula, Eric Chitambar, and Theshani Nuradha for helpful discussions and comments. KJ and MMW acknowledge support from the National Science Foundation under grant no. 2329662 and from the Cornell School of Electrical and Computer Engineering. This material is based upon work supported by, or in part by, the U. S. Army Research Laboratory and the U. S. Army Research Office under contract/grant number W911NF2310255, and by DoE under contract, DE-SC0012704.

---

\* kj264@cornell.edu

† slloyd@mit.edu

‡ wilde@cornell.edu

- [1] W. J. van Stockum, *The gravitational field of a distribution of particles rotating about an axis of symmetry*, *Proceedings of the Royal Society of Edinburgh* **57**, 135 (1938).
- [2] K. Gödel, *An example of a new type of cosmological solutions of Einstein’s field equations of gravitation*, *Reviews of Modern Physics* **21**, 447 (1949).
- [3] W. B. Bonnor, *The rigidly rotating relativistic dust cylinder*, *Journal of Physics A: Mathematical and General* **13**, 2121 (1980).
- [4] J. R. Gott, *Closed timelike curves produced by pairs of moving cosmic strings: Exact solutions*, *Physical Review Letters* **66**, 1126 (1991).
- [5] G. T. Horowitz and J. Maldacena, *The black hole final state*, *Journal of High Energy Physics* **2004**, 008 (2004).
- [6] D. Gottesman and J. Preskill, *Comment on “The black hole final state”*, *Journal of High Energy Physics* **2004**, 026 (2004).
- [7] S. Lloyd, *Almost certain escape from black holes in final state projection models*, *Physical Review Letters*, **96** (2006).
- [8] S. Lloyd and J. Preskill, *Unitarity of black hole evaporation in final-state projection models*, *Journal of High Energy Physics* **2014**, 126 (2014).
- [9] D. Deutsch, *Quantum mechanics near closed timelike lines*, *Physical Review D* **44**, 3197 (1991).
- [10] C. H. Bennett and B. Schumacher, *Lecture notes*, (2002), slides available at <https://web.archive.org/web/20030809140213/http://qip-server.tcs.tifr.res.in/%7Eqip/HTML/Courses/Bennett/TIFR5.pdf>.
- [11] S. Lloyd, L. Maccone, R. Garcia-Patron, V. Giovannetti, Y. Shikano, S. Pirandola, L. A. Rozema, A. Darabi, Y. Soudagar, L. K. Shalm, and A. M. Steinberg, *Closed timelike curves via postselection: Theory and experimental test of consistency*, *Physical Review Letters* **106**, 040403 (2011).
- [12] S. Lloyd, L. Maccone, R. Garcia-Patron, V. Giovannetti, and Y. Shikano, *Quantum mechanics of time travel through post-selected teleportation*, *Physical Review D*, **84** (2011).
- [13] G. Svetlichny, *Effective quantum time travel*, (2009), arXiv:0902.4898 [quant-ph].
- [14] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, *Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels*, *Physical Review Letters* **70**, 1895 (1993).
- [15] T. A. Brun, J. Harrington, and M. M. Wilde, *Localized closed timelike curves can perfectly distinguish quantum states*, *Physical Review Letters* **102**, 210402 (2009).
- [16] T. A. Brun and M. M. Wilde, *Perfect state distinguishability and computational speedups with postselected closed timelike curves*, *Foundations of Physics* **42**, 341 (2012).
- [17] T. A. Brun, M. M. Wilde, and A. Winter, *Quantum state cloning using Deutschian closed timelike curves*, *Physical Review Letters* **111**, 190401 (2013).
- [18] S. Aaronson, *Quantum computing, postselection, and probabilistic polynomial-time*, *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* **461**, 3473 (2005).
- [19] S. Aaronson and J. Watrous, *Closed timelike curves make quantum and classical computing equivalent*, *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* **465**, 631 (2009).
- [20] C. H. Bennett, D. Leung, G. Smith, and J. A. Smolin, *Can closed timelike curves or nonlinear quantum mechanics improve quantum state discrimination or help solve hard problems?* *Physical Review Letters* **103**, 170502 (2009).
- [21] T. C. Ralph and C. R. Myers, *Information flow of quantum states interacting with closed timelike curves*, *Physical Review A* **82**, 062330 (2010).
- [22] E. G. Cavalcanti, N. C. Menicucci, and J. L. Pienaar, *The preparation problem in nonlinear extensions of quantum theory*, (2012), arXiv:1206.2725 [quant-ph].
- [23] D. R. M. Arvidsson-Shukur, N. Yunger Halpern, H. V. Lepage, A. A. Lasek, C. H. W. Barnes, and S. Lloyd, *Quantum advantage in postselected metrology*, *Nature Communications* **11**, 3775 (2020).
- [24] D. R. M. Arvidsson-Shukur, A. G. McConnell, and N. Yunger Halpern, *Nonclassical advantage in metrology established via quantum simulations of hypothetical closed timelike curves*, *Physical Review Letters* **131**, 150202 (2023).
- [25] X. Song, F. Salvati, C. Gaikwad, N. Yunger Halpern, D. R. M. Arvidsson-Shukur, and K. Murch, *Agnostic phase estimation*, *Physical Review Letters* **132**, 260801 (2024).
- [26] B. Regula, *Probabilistic transformations of quantum resources*, *Physical Review Letters* **128**, 110505 (2022).
- [27] B. Regula, L. Lami, and M. M. Wilde, *Postselected quantum hypothesis testing*, *IEEE Transactions on Information Theory* **70**, 3453 (2024).
- [28] D. Ebler, S. Salek, and G. Chiribella, *Enhanced communication with the assistance of indefinite causal order*, *Physical Review*

- Letters **120**, 120502 (2018).
- [29] G. Chiribella, M. Banik, S. S. Bhattacharya, T. Guha, M. Alimuddin, A. Roy, S. Saha, S. Agrawal, and G. Kar, *Indefinite causal order enables perfect quantum communication with zero capacity channels*, *New Journal of Physics* **23**, 033039 (2021).
  - [30] K. Ji, B. Regula, and M. M. Wilde, *Postselected communication over quantum channels*, *International Journal of Quantum Information* **22**, 2440012 (2024).
  - [31] C. E. Shannon, *A mathematical theory of communication*, *Bell System Technical Journal* **27**, 379 (1948).
  - [32] A. S. Holevo, *Bounds for the quantity of information transmitted by a quantum communication channel*, *Probl. Peredachi Inf.* **9**, 3 (1973).
  - [33] B. Schumacher, *Sending entanglement through noisy quantum channels*, *Physical Review A* **54**, 2614 (1996).
  - [34] B. Schumacher and M. D. Westmoreland, *Sending classical information via noisy quantum channels*, *Physical Review A* **56**, 131 (1997).
  - [35] S. Lloyd, *Capacity of the noisy quantum channel*, *Physical Review A* **55**, 1613 (1997).
  - [36] A. S. Holevo, *The capacity of the quantum channel with general signal states*, *IEEE Transactions on Information Theory* **44**, 269 (1998).
  - [37] C. H. Bennett, P. W. Shor, J. A. Smolin, and A. V. Thapliyal, *Entanglement-assisted classical capacity of noisy quantum channels*, *Physical Review Letters* **83**, 3081 (1999).
  - [38] P. Shor, *The quantum channel capacity and coherent information*, Lecture Notes, MSRI Workshop on Quantum Computation (2002).
  - [39] I. Devetak, *The private classical capacity and quantum capacity of a quantum channel*, *IEEE Transactions on Information Theory* **51**, 44 (2005).
  - [40] T. S. Cubitt, D. Leung, W. Matthews, and A. Winter, *Zero-error channel capacity and simulation assisted by non-local correlations*, *IEEE Transactions on Information Theory* **57**, 5509 (2011).
  - [41] R. Duan and A. Winter, *No-signalling-assisted zero-error capacity of quantum channels and an information theoretic interpretation of the Lovász number*, *IEEE Transactions on Information Theory* **62**, 891 (2016).
  - [42] S. Khatri and M. M. Wilde, *Principles of Quantum Communication Theory: A Modern Approach* (arXiv, 2024) arXiv:2011.04672v2 [cond-mat, physics:hep-th, physics:math-ph, physics:quant-ph].
  - [43] M. S. Morris, K. S. Thorne, and U. Yurtsever, *Wormholes, time machines, and the weak energy condition*, *Physical Review Letters* **61**, 1446 (1988).
  - [44] S.-W. Kim and K. S. Thorne, *Do vacuum fluctuations prevent the creation of closed timelike curves?* *Physical Review D* **43**, 3929 (1991).
  - [45] D. G. Boulware, *Quantum field theory in spaces with closed timelike curves*, *Physical Review D* **46**, 4421 (1992).
  - [46] H. D. Politzer, *Simple quantum systems in spacetimes with closed timelike curves*, *Physical Review D* **46**, 4470 (1992).
  - [47] H. D. Politzer, *Path integrals, density matrices, and information flow with closed timelike curves*, *Physical Review D* **49**, 3981 (1994).
  - [48] J. B. Hartle, *Unitarity and causality in generalized quantum mechanics for nonchronal spacetimes*, *Physical Review D* **49**, 6543 (1994).
  - [49] M. M. Wilde, *Quantum Information Theory*, 2nd ed. (Cambridge University Press, 2017).
  - [50] M. Hayashi, *Quantum Information Theory: Mathematical Foundation*, Graduate Texts in Physics (Springer Berlin Heidelberg, Berlin, Heidelberg, 2017).
  - [51] J. Watrous, *The Theory of Quantum Information*, 1st ed. (Cambridge University Press, 2018).
  - [52] S. W. Hawking, *Breakdown of predictability in gravitational collapse*, *Physical Review D* **14**, 2460 (1976).
  - [53] Y. Aharonov, P. G. Bergmann, and J. L. Lebowitz, *Time symmetry in the quantum process of measurement*, *Physical Review* **134**, B1410 (1964).
  - [54] R. B. Griffiths, *Consistent histories and the interpretation of quantum mechanics*, *Journal of Statistical Physics* **36**, 219 (1984).
  - [55] M. Gell-Mann and J. B. Hartle, *Time symmetry and asymmetry in quantum mechanics and quantum cosmology*, in *Physical Origins of Time Asymmetry*, edited by J. J. Halliwell, J. Pérez-Mercader, and W. H. Zurek (Cambridge University Press, Cambridge, 1994) pp. 311–345.
  - [56] D. Genkina, G. Chiribella, and L. Hardy, *Optimal probabilistic simulation of quantum channels from the future to the past*, *Physical Review A* **85**, 022330 (2012).
  - [57] K. S. Thorne, *The Science of Interstellar*, 1st ed. (W.W. Norton & Company, New York, 2014).
  - [58] The retrocausal capacity of a channel is also referred to as its *flux capacity*.
  - [59] A. Nefed, *Bookshelf Scene*, (2018), Dribbble, <https://dribbble.com/shots/5696261-Bookshelf-Scene>.
  - [60] R. König, R. Renner, and C. Schaffner, *The operational meaning of min- and max-entropy*, *IEEE Transactions on Information Theory* **55**, 4337 (2009).
  - [61] I. George, C. Hirche, T. Nuradha, and M. M. Wilde, *Quantum Doeblin coefficients: Interpretations and applications*, (2025), arXiv:2503.22823 [quant-ph].
  - [62] See Supplemental Material for technical background, detailed proofs of our results, numerical examples, and additional discussions, which includes Refs. [65–96].
  - [63] I. Devetak, A. W. Harrow, and A. J. Winter, *A resource framework for quantum Shannon theory*, *IEEE Transactions on Information Theory* **54**, 4587 (2008).
  - [64] K. Ji, E. L. Brown, S. Lloyd, M. McFly, and M. M. Wilde, October 26, 1985.
  - [65] N. Datta, *Min- and max-relative entropies and a new entanglement monotone*, *IEEE Transactions on Information Theory* **55**, 2816 (2009).
  - [66] M. K. Gupta and M. M. Wilde, *Multiplicativity of completely bounded  $p$ -norms implies a strong converse for entanglement-assisted capacity*, *Communications in Mathematical Physics* **334**, 867 (2015).
  - [67] M. G. Díaz, K. Fang, X. Wang, M. Rosati, M. Skotiniotis, J. Calsamiglia, and A. Winter, *Using and reusing coherence to realize quantum processes*, *Quantum* **2**, 100 (2018). arXiv:1805.04045 [quant-ph].
  - [68] T. Nuradha, H. K. Mishra, F. Leditzky, and M. M. Wilde, *Multivariate fidelities*, *Journal of Physics A: Mathematical and Theoretical* **58**, 165304 (2025).
  - [69] K. Ji, H. K. Mishra, M. Mosonyi, and M. M. Wilde, *Barycentric bounds on the error exponents of quantum hypothesis exclusion*, (2025), arXiv:2407.13728 [quant-ph].
  - [70] F. Girardi, A. Oufkir, B. Regula, M. Tomamichel, M. Berta, and L. Lami, *Quantum umlaut information*, (2025), arXiv:2503.21479 [quant-ph].
  - [71] K. Fang, X. Wang, M. Tomamichel, and M. Berta, *Quantum channel simulation and the channel’s smooth max-information*, *IEEE Transactions on Information Theory* **66**, 2129 (2020).
  - [72] M. M. Wilde, M. Berta, C. Hirche, and E. Kaur, *Amortized channel divergence for asymptotic quantum channel discrimination*,

- Letters in Mathematical Physics **110**, 2277 (2020).
- [73] E. Chitambar, I. George, B. Doolittle, and M. Junge, *The communication value of a quantum channel*, [IEEE Transactions on Information Theory](#) **69**, 1660 (2023).
  - [74] E. Chitambar and G. Gour, *Quantum resource theories*, [Reviews of Modern Physics](#) **91**, 025001 (2019).
  - [75] G. Gour, *Quantum Resource Theories*, 1st ed. (Cambridge University Press, 2025).
  - [76] G. Vidal and R. Tarrach, *Robustness of entanglement*, [Physical Review A](#) **59**, 141 (1999).
  - [77] A. W. Harrow and M. A. Nielsen, *Robustness of quantum gates in the presence of noise*, [Physical Review A](#), **68** (2003).
  - [78] A. C. Elitzur, S. Popescu, and D. Rohrlich, *Quantum nonlocality for each pair in an ensemble*, [Physics Letters A](#) **162**, 25 (1992).
  - [79] M. Lewenstein and A. Sanpera, *Separability and entanglement of composite quantum systems*, [Physical Review Letters](#) **80**, 2261 (1998).
  - [80] N. Datta, *Max-relative entropy of entanglement, alias log robustness*, [International Journal of Quantum Information](#) **07**, 475 (2009).
  - [81] M. Fekete, *Über die Verteilung der Wurzeln bei gewissen algebraischen Gleichungen mit ganzzahligen Koeffizienten*, [Mathematische Zeitschrift](#) **17**, 228 (1923).
  - [82] G. Chiribella, G. M. D'Ariano, and P. Perinotti, *Transforming quantum operations: Quantum supermaps*, [EPL \(Europhysics Letters\)](#) **83**, 30004 (2008).
  - [83] J. Burniston, M. Grabowecky, C. M. Scandolo, G. Chiribella, and G. Gour, *Necessary and sufficient conditions on measurements of quantum channels*, [Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences](#) **476**, 20190832 (2020).
  - [84] P. W. Shor, *Scheme for reducing decoherence in quantum computer memory*, [Physical Review A](#) **52**, R2493 (1995).
  - [85] B. Schumacher and M. A. Nielsen, *Quantum data processing and error correction*, [Physical Review A](#) **54**, 2629 (1996).
  - [86] H. Barnum, M. A. Nielsen, and B. Schumacher, *Information transmission through a noisy quantum channel*, [Physical Review A](#) **57**, 4153 (1998).
  - [87] H. Barnum, E. Knill, and M. Nielsen, *On quantum fidelities and channel capacities*, [IEEE Transactions on Information Theory](#) **46**, 1317 (2000).
  - [88] C. Bennett, P. Shor, J. Smolin, and A. Thapliyal, *Entanglement-assisted capacity of a quantum channel and the reverse Shannon theorem*, [IEEE Transactions on Information Theory](#) **48**, 2637 (2002).
  - [89] A. S. Holevo, *On entanglement-assisted classical capacity*, [Journal of Mathematical Physics](#) **43**, 4326 (2002).
  - [90] W. Matthews, *A linear program for the finite block length converse of Polyanskiy–Poor–Verdú via nonsignaling codes*, [IEEE Transactions on Information Theory](#) **58**, 7036 (2012).
  - [91] R. Takagi, K. Wang, and M. Hayashi, *Application of the resource theory of channels to communication scenarios*, [Physical Review Letters](#) **124**, 120502 (2020).
  - [92] G. Gour, *Comparison of quantum channels by superchannels*, [IEEE Transactions on Information Theory](#) **65**, 5880 (2019).
  - [93] T. Eggeling, D. Schlingemann, and R. F. Werner, *Semicausal operations are semilocalizable*, [Europhysics Letters \(EPL\)](#) **57**, 782 (2002).
  - [94] C. King, *The capacity of the quantum depolarizing channel*, [IEEE Transactions on Information Theory](#) **49**, 221 (2003).
  - [95] C. H. Bennett, D. P. DiVincenzo, and J. A. Smolin, *Capacities of quantum erasure channels*, [Physical Review Letters](#) **78**, 3217 (1997).
  - [96] V. Giovannetti and R. Fazio, *Information-capacity description of spin-chain correlations*, [Physical Review A](#) **71**, 032314 (2005).

# Supplemental Material: Retrocausal Capacity of a Quantum Channel

## CONTENTS

|                                                   |    |
|---------------------------------------------------|----|
| I. Notation                                       | 1  |
| II. Max-information and Doeblin information       | 2  |
| III. Postselected closed timelike curves (P-CTCs) | 4  |
| A. Noiseless P-CTCs                               | 4  |
| B. Noisy P-CTCs                                   | 6  |
| IV. Retrocausal communication                     | 7  |
| A. General framework                              | 7  |
| B. Retrocausal quantum communication              | 9  |
| C. Retrocausal classical communication            | 10 |
| V. Retrocausal quantum capacity                   | 11 |
| A. Converse bound                                 | 12 |
| B. Achievability                                  | 14 |
| C. Asymptotics                                    | 17 |
| D. Error exponent and strong-converse exponent    | 18 |
| VI. Retrocausal classical capacity                | 20 |
| A. Converse bound                                 | 20 |
| B. Achievability                                  | 22 |
| C. Asymptotics                                    | 24 |
| D. Error exponent and strong-converse exponent    | 25 |
| VII. Examples                                     | 27 |
| A. Depolarizing channels                          | 27 |
| B. Erasure channels                               | 28 |
| C. Amplitude-damping channels                     | 30 |

## I. NOTATION

Let  $\mathbb{R}$  denote the set of real numbers,  $\mathbb{C}$  the set of complex numbers, and  $\mathbb{N}$  the set of nonnegative integers. For a nonnegative real number  $x$ , let  $\lfloor x \rfloor \equiv \max_{d \in \mathbb{N}} \{d : d \leq x\}$  denote the largest integer not exceeding  $x$ .

Let  $d_A \equiv \dim(\mathbb{H}_A)$  denote the dimension of a quantum system  $A$  associated with a Hilbert space  $\mathbb{H}_A$ . Let  $\mathbb{B}_A$  denote the space of bounded linear operators acting on  $\mathbb{H}_A$ . Let  $\mathcal{D}_A$  denote the set of quantum states of  $A$  (i.e., the set of positive semidefinite unit-trace operators in  $\mathbb{B}_A$ ), and let  $\text{aff}(\mathcal{D}_A)$  denote the set of unit-trace Hermitian operators in  $\mathbb{B}_A$  (i.e., the affine hull of  $\mathcal{D}_A$ ). Let  $\pi_A \equiv (1/d_A) \sum_{i=0}^{d_A-1} |i\rangle\langle i|_A$  denote the uniform state of  $A$ . Let  $\Phi_{AA'} \equiv (1/d_A) \sum_{i,j=0}^{d_A-1} |i\rangle\langle j|_A \otimes |i\rangle\langle j|_{A'}$  denote the standard maximally entangled state between two systems  $A$  and  $A'$  (with  $A \cong A'$ ). Let  $\Upsilon_{M'M} \equiv (1/d_M) \sum_{m=0}^{d_M-1} |m\rangle\langle m|_{M'} \otimes |m\rangle\langle m|_M$  denote the standard classically maximally correlated state between two systems  $M'$  and  $M$  (with  $M' \cong M$ ). Let  $\Omega_{M'\hat{M}} \equiv \sum_{m=0}^{d_M-1} |m\rangle\langle m|_{M'} \otimes |m\rangle\langle m|_{\hat{M}}$  denote the classical comparator between  $M'$  and  $\hat{M}$  (with  $M' \cong \hat{M}$ ).

Let  $\mathbb{L}_{A \rightarrow B}$  denote the space of linear maps from  $\mathbb{B}_A$  to  $\mathbb{B}_B$ . For a linear map  $\mathcal{N}_{A \rightarrow B} \in \mathbb{L}_{A \rightarrow B}$ , its Choi state is defined as  $\Phi_{A'B}^{\mathcal{N}} \equiv \mathcal{N}_{A \rightarrow B}[\Phi_{A'A}]$ . Let  $\mathcal{C}_{A \rightarrow B}$  denote the set of quantum channels from  $A$  to  $B$  (i.e., the set of completely positive trace-preserving linear maps in  $\mathbb{L}_{A \rightarrow B}$ ), and let  $\mathcal{L}_{A \rightarrow B}$  denote the set of completely positive maps in  $\mathbb{L}_{A \rightarrow B}$ . For a unit-trace Hermitian operator  $\sigma_B \in \text{aff}(\mathcal{D}_B)$ , a corresponding replacement map  $\mathcal{R}_{A \rightarrow B}^{\sigma}$  is defined by  $\mathcal{R}_{A \rightarrow B}^{\sigma}[\cdot] \equiv \text{Tr}_A[\cdot] \sigma_B$ , whose Choi state is given by  $\Phi_{A'B}^{\mathcal{R}^{\sigma}} = \pi_{A'} \otimes \sigma_B$ .

## II. MAX-INFORMATION AND DOEBLIN INFORMATION

The *max-relative entropy* between two states  $\rho_A, \sigma_A \in \mathcal{D}_A$  is defined as [65]

$$D_{\max}(\rho \parallel \sigma) := \inf_{\lambda \in \mathbb{R}} \{ \log_2 \lambda : \rho \leq \lambda \sigma \}. \quad (\text{S1})$$

We furthermore use the same definitions as above for the max-relative entropy when the arguments  $\rho_A, \sigma_A \in \text{aff}(\mathcal{D}_A)$  are Hermitian operators that are not necessarily positive semidefinite.

The *max-information* of a channel  $\mathcal{N}_{A \rightarrow B} \in \mathcal{C}_{A \rightarrow B}$  is defined as [60, 66, 67]

$$I_{\max}(\mathcal{N}) := \inf_{\sigma_B \in \mathcal{D}_B} D_{\max}(\Phi^{\mathcal{N}} \parallel \Phi^{\mathcal{R}^{\sigma}}) \quad (\text{S2a})$$

$$= \inf_{\substack{\lambda \in \mathbb{R}, \\ \sigma_B \in \mathcal{D}_B}} \{ \log_2 \lambda : \Phi_{A'B}^{\mathcal{N}} \leq \lambda \pi_{A'} \otimes \sigma_B \}. \quad (\text{S2b})$$

The *Doeblin information* of a channel  $\mathcal{N}_{A \rightarrow B} \in \mathcal{C}_{A \rightarrow B}$  is defined as [61]

$$I_{\text{doe}}(\mathcal{N}) := \inf_{\tau_B \in \text{aff}(\mathcal{D}_B)} D_{\max}(\Phi^{\mathcal{R}^{\tau}} \parallel \Phi^{\mathcal{N}}) \quad (\text{S3a})$$

$$= \inf_{\substack{\lambda \in \mathbb{R}, \\ \tau_B \in \text{aff}(\mathcal{D}_B)}} \{ \log_2 \lambda : \pi_{A'} \otimes \tau_B \leq \lambda \Phi_{A'B}^{\mathcal{N}} \}. \quad (\text{S3b})$$

Note that the infimization in  $\tau_B$  above is over all unit-trace Hermitian operators. We furthermore use the same definitions as above, for the max- and Doeblin informations, when the argument  $\mathcal{N}_{A \rightarrow B} \in \mathcal{L}_{A \rightarrow B}$  is a completely positive map that is not necessarily trace preserving.

**Remark 1.** For a function  $\mathbf{D}: \bigcup_A \text{aff}(\mathcal{D}_A) \times \text{aff}(\mathcal{D}_A) \rightarrow \mathbb{R} \cup \{\infty\}$  satisfying the data-processing inequality (DPI) on pairs of unit-trace Hermitian operators, one can define a corresponding mutual information  $\mathbf{I}$  of a channel  $\mathcal{N}_{A \rightarrow B} \in \mathcal{C}_{A \rightarrow B}$  as follows (see, e.g., Ref. [42, Definition 7.8.5] for when  $\mathbf{D}$  satisfies the DPI on pairs of states):

$$\mathbf{I}(\mathcal{N}) := \sup_{\rho_{RA} \in \mathcal{D}_{RA}} \inf_{\sigma_B \in \text{aff}(\mathcal{D}_B)} \mathbf{D}(\mathcal{N}_{A \rightarrow B}[\rho_{RA}] \parallel \rho_R \otimes \sigma_B) \quad (\text{S4a})$$

$$= \sup_{\rho_{A'} \in \mathcal{D}_{A'}} \inf_{\sigma_B \in \text{aff}(\mathcal{D}_B)} \mathbf{D}\left(\rho_{A'}^{\frac{1}{2}} \Phi_{A'B}^{\mathcal{N}} \rho_{A'}^{\frac{1}{2}} \parallel \rho_{A'}^{\frac{1}{2}} \Phi_{A'B}^{\mathcal{R}^{\sigma}} \rho_{A'}^{\frac{1}{2}}\right). \quad (\text{S4b})$$

One can also define a corresponding “reverse” mutual information  $\mathbf{I}^{\text{rev}}$  of a channel as the mutual information corresponding to the reverse function  $\mathbf{D}^{\text{rev}}: (\rho, \sigma) \mapsto \mathbf{D}(\sigma \parallel \rho)$  (see, e.g., Refs. [68–70] for specific instances):

$$\mathbf{I}^{\text{rev}}(\mathcal{N}) := \sup_{\rho_{RA} \in \mathcal{D}_{RA}} \inf_{\tau_B \in \text{aff}(\mathcal{D}_B)} \mathbf{D}(\rho_R \otimes \tau_B \parallel \mathcal{N}_{A \rightarrow B}[\rho_{RA}]) \quad (\text{S5a})$$

$$= \sup_{\rho_{A'} \in \mathcal{D}_{A'}} \inf_{\tau_B \in \text{aff}(\mathcal{D}_B)} \mathbf{D}\left(\rho_{A'}^{\frac{1}{2}} \Phi_{A'B}^{\mathcal{R}^{\tau}} \rho_{A'}^{\frac{1}{2}} \parallel \rho_{A'}^{\frac{1}{2}} \Phi_{A'B}^{\mathcal{N}} \rho_{A'}^{\frac{1}{2}}\right), \quad (\text{S5b})$$

where Eq. (S5b) follows from applying Eq. (S4b) to  $\mathbf{D}^{\text{rev}}$ . Then following a similar argument to Ref. [71, Remark 2] (also see Ref. [72, Lemma 12]), it can be inferred that the max- and Doeblin informations in Eqs. (S2) and (S3) are equivalent, respectively, to the mutual and reverse mutual information in Eqs. (S4) and (S5) corresponding to the max-relative entropy defined on pairs of unit-trace Hermitian operators.

The max-information has several operational interpretations in quantum Shannon theory, most notably as the one-shot zero-error nonsignaling-assisted classical simulation cost [41] and as the entanglement-assisted logarithmic communication value [73]. The Doeblin information has found applications in bounding contraction coefficients and consequently a variety of other contexts such as quantum machine learning, quantum error mitigation, and sample complexity of noisy quantum hypothesis testing [61]. From a resource-theoretic perspective [74, 75], the max- and Doeblin informations can also be interpreted, respectively, as the logarithmic resource robustness [76, 77] and the logarithmic resource weight [78, 79] against the affine hull of replacement channels [61, 80].

For our goal of characterizing the retrocausal capacity of a channel, it is of particular relevance that the max- and Doeblin informations have operational interpretations in terms of the maximum and minimum achievable singlet fractions, respectively [60, 61], as reviewed in the following lemma.

**Lemma 1** ([60, 61]). *The max- and Doeblin informations can be interpreted in terms of the maximum and minimum achievable singlet fractions, respectively: for a completely positive map  $\mathcal{N}_{A \rightarrow B} \in \mathcal{L}_{A \rightarrow B}$ ,*

$$I_{\max}(\mathcal{N}) = 2 \log_2 d_A + \log_2 \max_{\mathcal{K}_{B \rightarrow A} \in \mathcal{C}_{B \rightarrow A}} \text{Tr}[\Phi_{AA'}(\mathcal{K}_{B \rightarrow A} \circ \mathcal{N}_{A \rightarrow B})[\Phi_{AA'}]], \quad (\text{S6})$$

$$I_{\text{doe}}(\mathcal{N}) = -2 \log_2 d_A - \log_2 \min_{\mathcal{K}_{B \rightarrow A} \in \mathcal{C}_{B \rightarrow A}} \text{Tr}[\Phi_{AA'}(\mathcal{K}_{B \rightarrow A} \circ \mathcal{N}_{A \rightarrow B})[\Phi_{AA'}]]. \quad (\text{S7})$$

The max-information is additive in the sense that  $I_{\max}(\mathcal{N}_1 \otimes \mathcal{N}_2) = I_{\max}(\mathcal{N}_1) + I_{\max}(\mathcal{N}_2)$  for two arbitrary completely positive maps  $\mathcal{N}_1$  and  $\mathcal{N}_2$  [60, p. 4342] (also see Ref. [66, Lemma 6]). This implies that, for every  $n \in \mathbb{N}$ ,

$$\frac{1}{n} I_{\max}(\mathcal{N}^{\otimes n}) = I_{\max}(\mathcal{N}). \quad (\text{S8})$$

On the other hand, the Doeblin information is nonadditive in general, despite being superadditive [61, Proposition 20]:  $I_{\text{doe}}(\mathcal{N}_1 \otimes \mathcal{N}_2) \geq I_{\text{doe}}(\mathcal{N}_1) + I_{\text{doe}}(\mathcal{N}_2)$ . By Fekete's lemma [81], the following limit exists, which we term the *regularized Doeblin information* of  $\mathcal{N}_{A \rightarrow B}$ :

$$I_{\text{doe}}^{\infty}(\mathcal{N}) := \lim_{n \rightarrow \infty} \frac{1}{n} I_{\text{doe}}(\mathcal{N}^{\otimes n}). \quad (\text{S9})$$

A slight generalization of Ref. [61, Theorems 3 and 4], shows that, for quantum-to-classical completely positive maps (which include measurement channels) and mictodiactive covariant completely positive maps (which include depolarizing channels), the Doeblin information is additive and thus equal to its regularization. For general completely positive maps, the Doeblin information is upper bounded by the following quantity [61]:

$$I_{\text{doe}}(\mathcal{N}) \leq \inf_{\substack{\lambda \in \mathbb{R}_+, \\ \tau_B \in \text{aff}(\mathcal{D}_B)}} \{ \log_2 \lambda : -\lambda \Phi_{A'B}^{\mathcal{N}} \leq \pi_{A'} \otimes \tau_B \leq \lambda \Phi_{A'B}^{\mathcal{N}} \} \quad (\text{S10})$$

$$=: I_{\pm}(\mathcal{N}), \quad (\text{S11})$$

which can be shown to be additive by slightly generalizing Ref. [61, Proposition 24]. Therefore, since the Doeblin information is superadditive, we can bound the regularized Doeblin information from both sides in terms of single-letter quantities that are efficiently computable via semidefinite programs (SDPs):

$$I_{\text{doe}}(\mathcal{N}) \leq I_{\text{doe}}^{\infty}(\mathcal{N}) \leq I_{\pm}(\mathcal{N}). \quad (\text{S12})$$

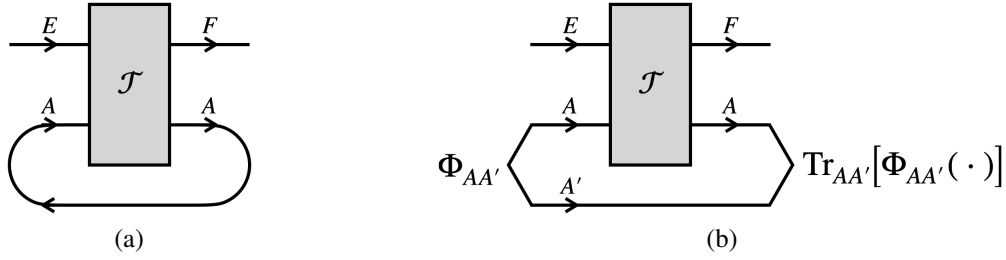


FIG. S1. Enclosing a linear map  $\mathcal{T}_{EA \rightarrow FA}$  with a noiseless P-CTC on the system  $A$  (Fig. S1a) is equivalent to performing postselected teleportation from its output  $A$  to its input  $A$  (Fig. S1b with renormalization).

### III. POSTSELECTED CLOSED TIMELIKE CURVES (P-CTCs)

#### A. Noiseless P-CTCs

As proposed in Ref. [10] and formalized in Refs. [11–13], enclosing a linear map  $\mathcal{T}_{EA \rightarrow FA} \in \mathbb{L}_{EA \rightarrow FA}$  with a noiseless P-CTC on the system  $A$  can be simulated, up to renormalization, by:

1. preparing a standard maximally entangled state  $\Phi_{AA'}$  between  $A$  and  $A'$ ;
2. letting the map  $\mathcal{T}_{EA \rightarrow FA}$  act on  $A$  along with  $E$ , receiving  $A$  back along with  $F$ ;
3. projecting  $A$  and  $A'$  onto  $\Phi_{AA'}$  and discarding them.

The above simulation can be represented by a linear supermap (i.e., a linear map on linear maps [82])  $\Gamma_A$  from  $\mathbb{L}_{A \rightarrow A}$  to  $\mathbb{C}$ , which we term the *loop supermap*, defined as follows:

$$\Gamma_A: \mathcal{T}_{EA \rightarrow FA} \mapsto \Gamma_A\{\mathcal{T}_{EA \rightarrow FA}\}, \quad (\text{S13a})$$

$$\Gamma_A\{\mathcal{T}_{EA \rightarrow FA}\}: \rho_{RE} \mapsto \text{Tr}_{AA'}[\Phi_{AA'} \mathcal{T}_{EA \rightarrow FA}[\rho_{RE} \otimes \Phi_{AA'}]], \quad (\text{S13b})$$

where we use the curly bracket to indicate the action of a supermap on a map. Then the actual transformation from  $E$  to  $F$ , resulting from enclosing  $\mathcal{T}_{EA \rightarrow FA}$  with the noiseless P-CTC, is represented by the following nonlinear map from  $\mathbb{B}_{RE}$  to  $\mathbb{B}_{RF}$ , with  $R$  an arbitrary reference system, which renormalizes  $\Gamma_A\{\mathcal{T}_{EA \rightarrow FA}\}$ :

$$\rho_{RE} \mapsto \frac{\Gamma_A\{\mathcal{T}_{EA \rightarrow FA}\}[\rho_{RE}]}{\text{Tr}[\Gamma_A\{\mathcal{T}_{EA \rightarrow FA}\}[\rho_{RE}]]}. \quad (\text{S14})$$

Note that, if  $\mathcal{T}_{EA \rightarrow FA}$  is completely positive, then so is  $\Gamma_A\{\mathcal{T}_{EA \rightarrow FA}\}$  [83]. It is thus immediate that the nonlinear map in Eq. (S14) preserves the set of states in the presence of an arbitrary reference system  $R$ .

**Fact 1** (Noiselessness). *Surrounding a swap channel between two systems with a noiseless P-CTC on one system results in an identity channel on the other: for two systems  $A$  and  $B$  with  $A \cong B$  and a linear operator  $\rho_{RA} \in \mathbb{B}_{RA}$  with  $R$  a reference system,*

$$\frac{\Gamma_B\{\mathcal{P}_{AB \rightarrow AB}\}[\rho_{RA}]}{\text{Tr}[\Gamma_B\{\mathcal{P}_{AB \rightarrow AB}\}[\rho_{RA}]]} = \rho_{RA}, \quad (\text{S15})$$

where  $\mathcal{P}_{A_I B_I \rightarrow A_O B_O} \equiv \text{id}_{A_I \rightarrow B_O} \otimes \text{id}_{B_I \rightarrow A_O}$  is the swap channel between  $A$  and  $B$ .

*Proof.* It follows from the definition of the loop supermap [Eq. (S13)] that

$$\Gamma_B\{\mathcal{P}_{AB \rightarrow AB}\}[\rho_{RA}] = \text{Tr}_{BB'}[\Phi_{BB'} \mathcal{P}_{AB \rightarrow AB}[\rho_{RA} \otimes \Phi_{BB'}]] \quad (\text{S16})$$

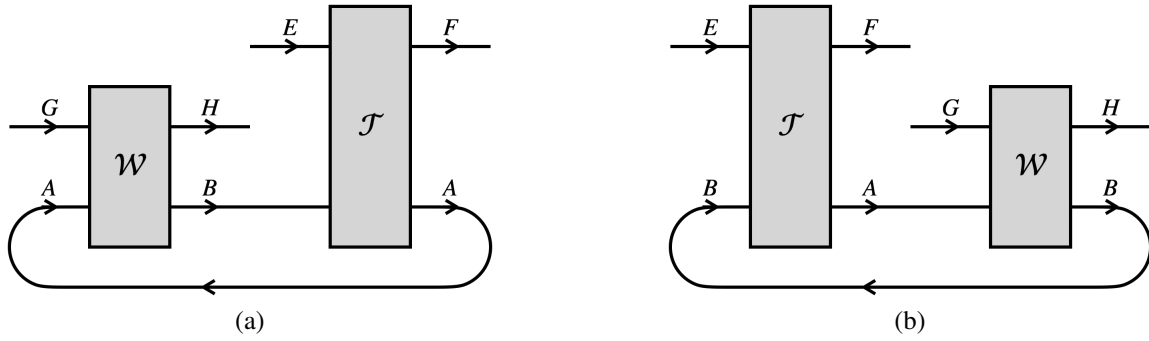


FIG. S2. The cyclicity of noiseless P-CTCs dictates that Figs. S2a and S2b are equivalent.

$$= \text{Tr}_{BB'} [\Phi_{BB'} (\rho_{RB} \otimes \Phi_{AB'})] \quad (\text{S17})$$

$$= \frac{1}{d_A^2} \rho_{RA}, \quad (\text{S18})$$

where Eq. (S18) follows from probabilistic teleportation [14]. The desired statement is then immediate.  $\square$

**Fact 2** (Composition). *Parallel composition of noiseless P-CTCs is represented by tensor product: for two systems  $A$  and  $B$ , we have that  $\Gamma_{AB} = \Gamma_A \otimes \Gamma_B$ .*

*Proof.* The desired statement follows immediately from the fact that  $\Phi_{ABA'B'} = \Phi_{AA'} \otimes \Phi_{BB'}$ .  $\square$

**Lemma 2** (Cyclicity). *Noiseless P-CTCs have the following cyclicity property (see Fig. S2): for two linear maps  $\mathcal{T}_{EB \rightarrow FA} \in \mathbb{L}_{EB \rightarrow FA}$  and  $\mathcal{W}_{GA \rightarrow HB} \in \mathbb{L}_{GA \rightarrow HB}$  and a linear operator  $\rho_{\text{REG}} \in \mathbb{B}_{\text{REG}}$  with  $R$  a reference system, we have that*

$$\frac{\Gamma_A \{\mathcal{T}_{EB \rightarrow FA} \circ \mathcal{W}_{GA \rightarrow HB}\} [\rho_{\text{REG}}]}{\text{Tr}[\Gamma_A \{\mathcal{T}_{EB \rightarrow FA} \circ \mathcal{W}_{GA \rightarrow HB}\} [\rho_{\text{REG}}]]} = \frac{\Gamma_B \{\mathcal{W}_{GA \rightarrow HB} \circ \mathcal{T}_{EB \rightarrow FA}\} [\rho_{\text{REG}}]}{\text{Tr}[\Gamma_B \{\mathcal{W}_{GA \rightarrow HB} \circ \mathcal{T}_{EB \rightarrow FA}\} [\rho_{\text{REG}}]]}. \quad (\text{S19})$$

*Proof.* Since Eq. (S18) implies that  $\Gamma_{B'} \{\mathcal{P}_{BB' \rightarrow BB'}\} = \frac{1}{d_B^2} \text{id}_{B \rightarrow B}$  for  $B' \cong B$ , where  $\mathcal{P}_{BB' \rightarrow BB'}$  is the swap channel between  $B$  and  $B'$ , it follows that

$$\Gamma_A \{\mathcal{T}_{EB \rightarrow FA} \circ \mathcal{W}_{GA \rightarrow HB}\} = d_B^2 \Gamma_A \{\mathcal{T}_{EB \rightarrow FA} \circ \Gamma_{B'} \{\mathcal{P}_{BB'}\} \circ \mathcal{W}_{GA \rightarrow HB}\} \quad (\text{S20})$$

$$= d_B^2 (\Gamma_A \otimes \Gamma_{B'}) \{\mathcal{T}_{EB \rightarrow FA} \circ \mathcal{P}_{BB' \rightarrow BB'} \circ \mathcal{W}_{GA \rightarrow HB}\} \quad (\text{S21})$$

$$= d_B^2 \Gamma_{AB'} \{\mathcal{T}_{EB \rightarrow FA} \circ \mathcal{P}_{BB' \rightarrow BB'} \circ \mathcal{W}_{GA \rightarrow HB}\}. \quad (\text{S22})$$

Considering that

$$\mathcal{T}_{EB_O \rightarrow FA_O} \circ \mathcal{P}_{B_I B'_I \rightarrow B_O B'_O} \circ \mathcal{W}_{GA_I \rightarrow HB_I} = \mathcal{T}_{EB_O \rightarrow FA_O} \circ (\text{id}_{B_I \rightarrow B'_O} \otimes \text{id}_{B'_I \rightarrow B_O}) \circ \mathcal{W}_{GA_I \rightarrow HB_I} \quad (\text{S23})$$

$$= \mathcal{T}_{EB'_I \rightarrow FA_O} \circ \mathcal{W}_{GA_I \rightarrow HB'_O} \quad (\text{S24})$$

$$= \mathcal{T}_{EB'_I \rightarrow FA_O} \otimes \mathcal{W}_{GA_I \rightarrow HB'_O}, \quad (\text{S25})$$

it follows from Eq. (S22) that

$$\Gamma_A \{\mathcal{T}_{EB \rightarrow FA} \circ \mathcal{W}_{GA \rightarrow HB}\} = d_B^2 \Gamma_{AB'} \{\mathcal{T}_{EB' \rightarrow FA} \otimes \mathcal{W}_{GA \rightarrow HB'}\}. \quad (\text{S26})$$

Following the same reasoning, we can also show for  $A' \cong A$  that

$$\Gamma_B \{\mathcal{W}_{GA \rightarrow HB} \circ \mathcal{T}_{EB \rightarrow FA}\} = d_A^2 \Gamma_{A'B} \{\mathcal{W}_{GA' \rightarrow HB} \otimes \mathcal{T}_{EB \rightarrow FA'}\}. \quad (\text{S27})$$

Combining Eqs. (S26) and (S27) yields the desired statement.  $\square$

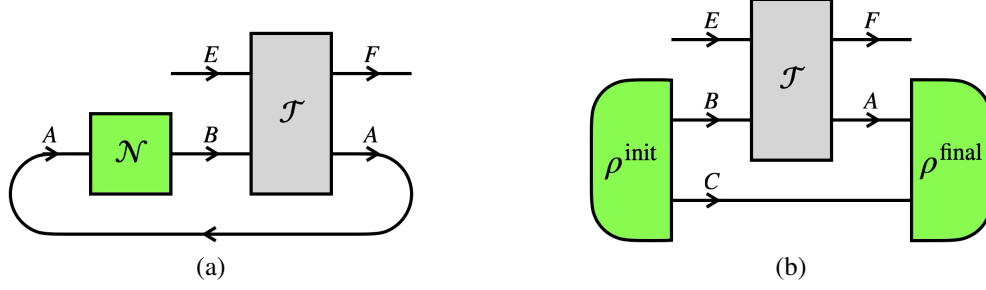


FIG. S3. Enclosing a linear map  $\mathcal{T}_{EB \rightarrow FA}$  with a noisy P-CTC represented by a completely positive map  $\mathcal{N}_{A \rightarrow B}$  is equivalent to preprocessing  $\mathcal{T}_{EB \rightarrow FA}$  with  $\mathcal{N}_{A \rightarrow B}$  and enclosing  $\mathcal{T}_{EB \rightarrow FA} \circ \mathcal{N}_{A \rightarrow B}$  with a noiseless P-CTC on the system  $A$  (Fig. S3a). It is also equivalent to imposing on  $\mathcal{T}_{EB \rightarrow FA}$  initial- and final-state boundary conditions satisfying Eq. (S29) (Fig. S3b with renormalization).

### B. Noisy P-CTCs

Compared to the main text, here we consider a slightly more generalized notion of a noisy P-CTC, which can be represented by a completely positive (but not necessarily trace-preserving) map. As illustrated in Fig. S3a, a noisy P-CTC from  $A$  to  $B$ , represented by a completely positive map  $\mathcal{N}_{A \rightarrow B} \in \mathcal{L}_{A \rightarrow B}$ , can be understood as the concatenation of a noiseless P-CTC on  $A$  and the map  $\mathcal{N}_{A \rightarrow B}$ . In other words, enclosing a linear map  $\mathcal{T}_{EB \rightarrow FA} \in \mathbb{L}_{EB \rightarrow FA}$  with the noisy P-CTC amounts to preprocessing  $\mathcal{T}_{EB \rightarrow FA}$  with  $\mathcal{N}_{A \rightarrow B}$  and enclosing  $\mathcal{T}_{EB \rightarrow FA} \circ \mathcal{N}_{A \rightarrow B}$  with a noiseless P-CTC on  $A$ . Formally, the resulting transformation from  $E$  to  $F$  is represented by the following nonlinear map from  $\mathbb{B}_{RE}$  to  $\mathbb{B}_{RF}$ , with  $R$  an arbitrary reference system:

$$\rho_{RE} \mapsto \frac{\Gamma_A\{\mathcal{T}_{EB \rightarrow FA} \circ \mathcal{N}_{A \rightarrow B}\}[\rho_{RE}]}{\text{Tr}[\Gamma_A\{\mathcal{T}_{EB \rightarrow FA} \circ \mathcal{N}_{A \rightarrow B}\}[\rho_{RE}]]}. \quad (\text{S28})$$

Note that the noisy P-CTC can equivalently be realized by postprocessing  $\mathcal{T}_{EB \rightarrow FA}$  with  $\mathcal{N}_{A \rightarrow B}$  and enclosing  $\mathcal{N}_{A \rightarrow B} \circ \mathcal{T}_{EB \rightarrow FA}$  with a noiseless P-CTC on  $B$ , due to the cyclicity of noiseless P-CTCs (Lemma 2).

A major reason for our interest in noisy P-CTCs represented by completely positive maps, and not just by channels, is that they naturally arise in a time-symmetric modification of quantum mechanics [53–55], where arbitrary initial- and final-state boundary conditions can be specified. A general configuration to consider is one where a linear map  $\mathcal{T}_{EB \rightarrow FA} \in \mathbb{L}_{EB \rightarrow FA}$  is subject to an initial state  $\rho_{BC}^{\text{init}} \in \mathcal{D}_{BC}$  and a final state  $\rho_{AC}^{\text{final}} \in \mathcal{D}_{AC}$ , where  $C$  is a side quantum memory (see Fig. S3b). The resulting transformation from  $E$  to  $F$  is then represented by the following nonlinear map, with  $R$  an arbitrary reference system:

$$\rho_{RE} \mapsto \frac{\text{Tr}_{AC}[\rho_{AC}^{\text{final}} \mathcal{T}_{EB \rightarrow FA}[\rho_{RE} \otimes \rho_{BC}^{\text{init}}]]}{\text{Tr}[\rho_{AC}^{\text{final}} \mathcal{T}_{EB \rightarrow FA}[\rho_{RE} \otimes \rho_{BC}^{\text{init}}]]}. \quad (\text{S29})$$

It is easy to verify that Eq. (S29) recovers Eq. (S28) upon defining the following completely positive but not necessarily trace-preserving map:

$$\mathcal{N}_{A \rightarrow B}[\cdot] := \text{Tr}_{AC}[\rho_{AC}^{\text{final}} ((\cdot)_A \otimes \rho_{BC}^{\text{init}})], \quad (\text{S30})$$

which represents a noisy P-CTC from  $A$  to  $B$ . Conversely, every noisy P-CTC represented by a completely positive map can be realized with some initial- and final-state boundary condition. To see this, for a completely positive map  $\mathcal{N}_{A \rightarrow B} \in \mathcal{L}_{A \rightarrow B}$ , by setting

$$\rho_{BC}^{\text{init}} := \frac{\mathcal{N}_{A \rightarrow B}[\Phi_{AC}]}{\text{Tr}[\mathcal{N}_{A \rightarrow B}[\Phi_{AC}]]}, \quad (\text{S31})$$

$$\rho_{AC}^{\text{final}} := \Phi_{AC}, \quad (\text{S32})$$

one realizes a noisy P-CTC represented by  $\mathcal{N}_{A \rightarrow B} / \text{Tr}[\mathcal{N}_{A \rightarrow B}[\Phi_{AC}]]$ , and this noisy P-CTC is equivalent to the one represented by  $\mathcal{N}_{A \rightarrow B}$ , due to renormalization.

## IV. RETROCAUSAL COMMUNICATION

### A. General framework

The setting of retrocausal communication involves two parties, namely a sender in the future and a receiver in the past. Following the convention of the main text, we refer to the sender as a father and the receiver as his daughter. The father is connected to the daughter by a noisy P-CTC represented by a given completely positive map  $\mathcal{N}_{A \rightarrow B} \in \mathcal{L}_{A \rightarrow B}$ , and they aim to transmit quantum or classical messages backward in time from the father to the daughter through this noisy P-CTC. They can adopt any strategy for communication as long as they do not violate chronology when not invoking the noisy P-CTC.

The freedom for the father and the daughter to adopt any strategy without violating chronology brings forth the following possibility. Since the daughter precedes the father in time, she is naturally able to signal to him, most generally by storing certain quantum information in a noiseless quantum memory of an arbitrary dimension and leaving the memory for him to retrieve in the future. Note that such a memory, which enables signaling from the daughter to the father, combined with the noisy P-CTC from the father to daughter, could incur a causal loop. However, grandfather-type paradoxes can be effectively circumvented within the model of P-CTCs [11], guaranteeing that the setting we consider is well defined. In fact, the potential formation of such a causal loop is what fundamentally distinguishes retrocausal communication from conventional (i.e., forward-in-time) communication settings (such as unassisted classical [31, 32, 34, 36] and quantum [33, 35, 38, 39, 84–87] communication, entanglement-assisted communication [37, 88, 89], and nonsignaling-assisted communication [40, 41, 90, 91]), and it is the reason why a channel's retrocausal capacity can exceed its capacity in any conventional communication setting.

We now put the discussion in formal terms. As illustrated in Fig. S4, the father and the daughter's strategy is, in general, represented by a pair of channels,  $(\mathcal{E}_{ML \rightarrow A}, \mathcal{D}_{B \rightarrow \hat{M}L})$  (with  $M \cong \hat{M}$ ), corresponding to the father's encoding and the daughter's decoding, respectively, where:

- $M$  is the system storing the father's message;
- $\hat{M}$  is the system storing the daughter's recovery of the message;
- $L$  is the noiseless quantum memory that the daughter leaves to the father.

The strategy operates as follows.

- **Encoding.** In the future, the father retrieves  $L$  (which was left by the daughter in the past), encodes his message with the assistance of  $L$  by applying the channel  $\mathcal{E}_{ML \rightarrow A}$ , and feeds  $A$  into the entrance of the noisy P-CTC.
- **Decoding.** In the past, the daughter receives  $B$  from the exit of the noisy P-CTC, decodes the message by applying the channel  $\mathcal{D}_{B \rightarrow \hat{M}L}$ , and leaves  $L$  for the father to retrieve in the future.

Combining the strategy  $(\mathcal{E}_{ML \rightarrow A}, \mathcal{D}_{B \rightarrow \hat{M}L})$  as described above with the noisy P-CTC represented by the given completely positive map  $\mathcal{N}_{A \rightarrow B}$ , a transformation from  $M$  to  $\hat{M}$  is realized, represented by a nonlinear map  $\mathcal{M}_{M \rightarrow \hat{M}}: \mathbb{B}_{RM} \rightarrow \mathbb{B}_{R\hat{M}}$ , with  $R$  an arbitrary reference system, defined as follows:

$$\mathcal{M}_{M \rightarrow \hat{M}}: \rho_{RM} \mapsto \frac{\Gamma_A \{ \mathcal{E}_{ML \rightarrow A} \circ \mathcal{D}_{B \rightarrow \hat{M}L} \circ \mathcal{N}_{A \rightarrow B} \} [\rho_{RM}]}{\text{Tr}[\Gamma_A \{ \mathcal{E}_{ML \rightarrow A} \circ \mathcal{D}_{B \rightarrow \hat{M}L} \circ \mathcal{N}_{A \rightarrow B} \} [\rho_{RM}]]}, \quad (\text{S33})$$

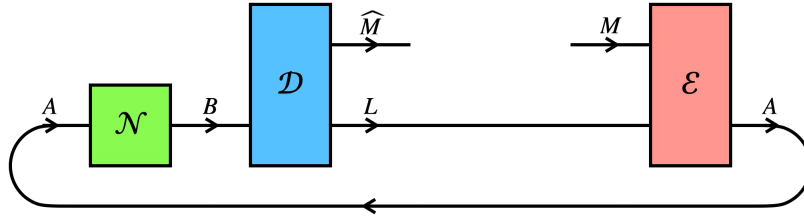


FIG. S4. In retrocausal communication, a father in the future (i.e., the sender) is connected to his daughter in the past (i.e., the receiver) by a noisy P-CTC represented by a completely positive map  $\mathcal{N}_{A \rightarrow B}$  (green). The father applies a channel  $\mathcal{E}_{ML \rightarrow A}$  as his encoding (red), and the daughter applies a channel  $\mathcal{D}_{B \rightarrow \hat{M}L}$  as her decoding (blue). As part of the strategy, the daughter can signal to the father via a noiseless quantum memory  $L$ .

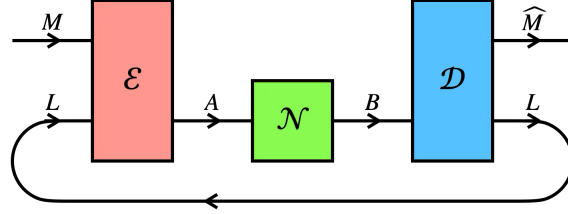


FIG. S5. In forward-in-time communication assisted by a noiseless P-CTC, the sender in the past is connected to the receiver in the future by a completely positive map  $\mathcal{N}_{A \rightarrow B}$ , and the receiver is connected to the sender by a noiseless P-CTC. This setting is mathematically equivalent to retrocausal communication (see Fig. S4).

where  $\Gamma_A$  is the loop supermap defined in Eq. (S13).

**Remark 2** (Forward-in-time communication assisted by a noiseless P-CTC). The setting of retrocausal communication, where communication is conducted *backward* in time (from a sender in the future to a receiver in the past), is equivalent to the setting where communication is conducted *forward* in time (from a sender in the past to a receiver in the future) but assisted by a noiseless P-CTC from the receiver to the sender (see Fig. S5). To see this, it follows from the cyclicity of noiseless P-CTCs (Lemma 2) that, for every linear operator  $\rho_{RM} \in \mathbb{B}_{RM}$  with  $R$  a reference system,

$$\frac{\Gamma_A \{ \mathcal{E}_{ML \rightarrow A} \circ \mathcal{D}_{B \rightarrow \hat{M}L} \circ \mathcal{N}_{A \rightarrow B} \} [\rho_{RM}]}{\text{Tr} [\Gamma_A \{ \mathcal{E}_{ML \rightarrow A} \circ \mathcal{D}_{B \rightarrow \hat{M}L} \circ \mathcal{N}_{A \rightarrow B} \} [\rho_{RM}]]} = \frac{\Gamma_L \{ \mathcal{D}_{B \rightarrow \hat{M}L} \circ \mathcal{N}_{A \rightarrow B} \circ \mathcal{E}_{ML \rightarrow A} \} [\rho_{RM}]}{\text{Tr} [\Gamma_L \{ \mathcal{D}_{B \rightarrow \hat{M}L} \circ \mathcal{N}_{A \rightarrow B} \circ \mathcal{E}_{ML \rightarrow A} \} [\rho_{RM}]]}. \quad (\text{S34})$$

Note that the map from  $\rho_{RM}$  to the right-hand side of Eq. (S34) precisely represents the transformation from  $M$  to  $\hat{M}$  resulting from adopting the strategy  $(\mathcal{E}_{ML \rightarrow A}, \mathcal{D}_{B \rightarrow \hat{M}L})$  in a setting where communication is conducted forward in time through  $\mathcal{N}_{A \rightarrow B}$  and assisted by a noiseless P-CTC on  $L$ . Due to this equivalence, it can clearly be recognized that, mathematically, retrocausal communication reduces to entanglement-assisted communication [37, 88, 89] when  $L$  is only used for entanglement distribution, and it reduces to unassisted communication [31–36, 38, 39, 84–87] when  $L$  is not utilized at all.

**Remark 3** (Axiomatic interpretation). The setting of retrocausal communication can also be understood from an axiomatic perspective. Specifically, in any communication setting, the sender and the receiver's strategy can be interpreted as a bipartite channel  $\Theta_{MB \rightarrow A\hat{M}}$ , where  $M$  and  $A$  are held by the sender and  $B$  and  $\hat{M}$  are held by the receiver (see, e.g., Ref. [41, Fig. 1] for an illustration). Communication through  $\mathcal{N}_{A \rightarrow B}$  by adopting this strategy can then be understood as connecting the output  $A$  to the input  $B$  of  $\Theta_{MB \rightarrow A\hat{M}}$  with  $\mathcal{N}_{A \rightarrow B}$ , resulting in a transformation from the sender's message system  $M$  to the receiver's recovery system  $\hat{M}$ . Then the following can be said.

- In standard (i.e., chronology-respecting) quantum theory, to avoid the formation of a causal loop when combining  $\Theta_{MB \rightarrow A\hat{M}}$  with  $\mathcal{N}_{A \rightarrow B}$  for all  $\mathcal{N}_{A \rightarrow B}$ , a minimal constraint on the bipartite channel  $\Theta_{MB \rightarrow A\hat{M}}$  is that it should be nonsignaling from  $B$  to  $A$ . Under this constraint, combining  $\Theta_{MB \rightarrow A\hat{M}}$  with  $\mathcal{N}_{A \rightarrow B}$  amounts to applying a superchannel to  $\mathcal{N}_{A \rightarrow B}$  [82, 92].
- In addition to the above, in conventional communication settings, to ensure that the strategy itself does not introduce any communication resource from the sender to the receiver, it is also necessary to require that the bipartite channel  $\Theta_{MB \rightarrow A\hat{M}}$  be nonsignaling from  $M$  to  $\hat{M}$ . This renders  $\Theta_{MB \rightarrow A\hat{M}}$  a two-way nonsignaling correlation and, if without further constraints, corresponds to the setting of nonsignaling-assisted communication [40, 41, 90, 91]. As more restrictive settings than nonsignaling-assisted communication, entanglement-assisted communication [37, 88, 89] corresponds to when  $\Theta_{MB \rightarrow A\hat{M}}$  is furthermore required to be realized by local operations and shared entanglement, and unassisted communication [31–36, 38, 39, 84–87] corresponds to when  $\Theta_{MB \rightarrow A\hat{M}}$  is realized by just local operations.
- In the setting of retrocausal communication (or equivalently, communication assisted by a noiseless P-CTC; see Remark 2 and Fig. S5) that we consider, the nonsignaling constraint from  $B$  to  $A$  on  $\Theta_{MB \rightarrow A\hat{M}}$  is released, which enables the possible formation of a causal loop and renders the strategy indescribable by a superchannel; however, the nonsignaling constraint from  $M$  to  $\hat{M}$  must be kept, so that the strategy, despite its inclusiveness, does not introduce any communication resource from the sender to the receiver. Note that the sole requirement of  $\Theta_{MB \rightarrow A\hat{M}}$  being nonsignaling from  $M$  to  $\hat{M}$  is equivalent to our assumption that the strategy can be decomposed into a pair of channels  $(\mathcal{E}_{ML \rightarrow A}, \mathcal{D}_{B \rightarrow \hat{M}L})$  with  $L$  a quantum memory, due to Ref. [93].

## B. Retrocausal quantum communication

We first consider the case where the messages being transmitted are quantum. A quantum message can be represented by a pure state  $\psi_{RM}$ , where  $R$  is an inaccessible reference system that purifies the message system  $M$ . For a strategy  $(\mathcal{E}_{ML \rightarrow A}, \mathcal{D}_{B \rightarrow \hat{M}L})$ , the fidelity of the daughter's recovery of the message is given by the overlap between  $\psi_{R\hat{M}}$  and the actual state of  $R\hat{M}$  that she recovers, which by Eq. (S33) equals

$$F(\psi_{R\hat{M}}|\psi_{RM}; (\mathcal{E}_{ML \rightarrow A}, \mathcal{D}_{B \rightarrow \hat{M}L}), \mathcal{N}_{A \rightarrow B}) = \frac{\text{Tr}[\psi_{R\hat{M}} \Gamma_A \{ \mathcal{E}_{ML \rightarrow A} \circ \mathcal{D}_{B \rightarrow \hat{M}L} \circ \mathcal{N}_{A \rightarrow B} \} [\psi_{RM}]]}{\text{Tr}[\Gamma_A \{ \mathcal{E}_{ML \rightarrow A} \circ \mathcal{D}_{B \rightarrow \hat{M}L} \circ \mathcal{N}_{A \rightarrow B} \} [\psi_{RM}]]}. \quad (\text{S35})$$

The *worst-case quantum infidelity* associated with the strategy is then given by

$$q_{\text{infid}}((\mathcal{E}_{ML \rightarrow A}, \mathcal{D}_{B \rightarrow \hat{M}L}), \mathcal{N}_{A \rightarrow B}) := 1 - \inf_{\substack{d_R \in \mathbb{N}, \\ \psi_{RM} \in \mathcal{D}_{RM}: \\ \text{rank}(\psi_{RM})=1}} F(\psi_{R\hat{M}}|\psi_{RM}; (\mathcal{E}_{ML \rightarrow A}, \mathcal{D}_{B \rightarrow \hat{M}L}), \mathcal{N}_{A \rightarrow B}), \quad (\text{S36})$$

which upper bounds the infidelity for all possible quantum messages.

The *one-shot retrocausal quantum capacity* of a completely positive map  $\mathcal{N}_{A \rightarrow B} \in \mathcal{L}_{A \rightarrow B}$  within infidelity  $\varepsilon \in [0, 1]$  is defined as

$$Q_{\text{retro}}^{\varepsilon}(\mathcal{N}) := \sup_{\substack{d_M, d_L \in \mathbb{N}, \\ \mathcal{E}_{ML \rightarrow A} \in \mathcal{C}_{ML \rightarrow A}, \\ \mathcal{D}_{B \rightarrow \hat{M}L} \in \mathcal{C}_{B \rightarrow \hat{M}L}}} \{ \log_2 d_M : q_{\text{infid}}((\mathcal{E}_{ML \rightarrow A}, \mathcal{D}_{B \rightarrow \hat{M}L}), \mathcal{N}_{A \rightarrow B}) \leq \varepsilon \}, \quad (\text{S37})$$

which quantifies the maximum number of qubits that can be transmitted within infidelity  $\varepsilon$  from the father to the daughter through the noisy P-CTC represented by  $\mathcal{N}_{A \rightarrow B}$  under an encoding and decoding of their choice.

The *asymptotic retrocausal quantum capacity* of a completely positive map  $\mathcal{N}_{A \rightarrow B} \in \mathcal{L}_{A \rightarrow B}$  is defined as

$$Q_{\text{retro}}(\mathcal{N}) := \inf_{\varepsilon \in (0,1)} \liminf_{n \rightarrow \infty} \frac{1}{n} Q_{\text{retro}}^\varepsilon(\mathcal{N}^{\otimes n}), \quad (\text{S38})$$

which quantifies the maximum rate at which the number of reliably transmitted qubits increases with respect to the number of parallel uses of the noisy P-CTC.

The *error exponent of retrocausal quantum communication* for a completely positive map  $\mathcal{N}_{A \rightarrow B} \in \mathcal{L}_{A \rightarrow B}$  at a transmission rate  $r \in [0, \infty)$  is defined as

$$E_{\text{retro,Q}}^r(\mathcal{N}) := \liminf_{n \rightarrow \infty} \sup_{\substack{d_M, d_L \in \mathbb{N}, \\ \mathcal{E}_{ML \rightarrow A^n} \in \mathcal{C}_{ML \rightarrow A^n}, \\ \mathcal{D}_{B^n \rightarrow \widehat{M}_L} \in \mathcal{C}_{B^n \rightarrow \widehat{M}_L}}} \left\{ -\frac{1}{n} \log_2 q_{\text{infid}}((\mathcal{E}_{ML \rightarrow A^n}, \mathcal{D}_{B^n \rightarrow \widehat{M}_L}), (\mathcal{N}^{\otimes n})_{A^n \rightarrow B^n}) : d_M \geq 2^{nr} \right\}, \quad (\text{S39})$$

which quantifies the maximum rate at which the infidelity exponentially decays as the number of transmitted qubits increases at rate  $r$ .

The *strong-converse exponent of retrocausal quantum communication* for a completely positive map  $\mathcal{N}_{A \rightarrow B} \in \mathcal{L}_{A \rightarrow B}$  at a transmission rate  $r \in [0, \infty)$  is defined as

$$S_{\text{retro,Q}}^r(\mathcal{N}) := \limsup_{n \rightarrow \infty} \inf_{\substack{d_M, d_L \in \mathbb{N}, \\ \mathcal{E}_{ML \rightarrow A^n} \in \mathcal{C}_{ML \rightarrow A^n}, \\ \mathcal{D}_{B^n \rightarrow \widehat{M}_L} \in \mathcal{C}_{B^n \rightarrow \widehat{M}_L}}} \left\{ -\frac{1}{n} \log_2 (1 - q_{\text{infid}}((\mathcal{E}_{ML \rightarrow A^n}, \mathcal{D}_{B^n \rightarrow \widehat{M}_L}), (\mathcal{N}^{\otimes n})_{A^n \rightarrow B^n})) : d_M \geq 2^{nr} \right\}, \quad (\text{S40})$$

which quantifies the minimum rate at which the infidelity exponentially approaches 1 as the number of transmitted qubits increases at rate  $r$ .

### C. Retrocausal classical communication

We now consider the case where the messages being transmitted are classical. A classical message  $m$  from an alphabet of size  $d_M$  is represented by the classical state  $|m\rangle\langle m|_M$ , with  $\{|m\rangle\}_{m=0}^{d_M-1}$  an orthonormal basis of  $\mathbb{H}_M$ . For a strategy  $(\mathcal{E}_{ML \rightarrow A}, \mathcal{D}_{B \rightarrow \widehat{M}_L})$ , the probability of the daughter correctly recovering the message is given by the overlap between  $|m\rangle\langle m|_{\widehat{M}}$  and the actual state of  $\widehat{M}$  that she recovers, which by Eq. (S33) equals

$$\Pr\{m|m; (\mathcal{E}_{ML \rightarrow A}, \mathcal{D}_{B \rightarrow \widehat{M}_L}), \mathcal{N}_{A \rightarrow B}\} = \frac{\text{Tr}[|m\rangle\langle m|_{\widehat{M}} \Gamma_A \{\mathcal{E}_{ML \rightarrow A} \circ \mathcal{D}_{B \rightarrow \widehat{M}_L} \circ \mathcal{N}_{A \rightarrow B}\} [|m\rangle\langle m|_M]]}{\text{Tr}[\Gamma_A \{\mathcal{E}_{ML \rightarrow A} \circ \mathcal{D}_{B \rightarrow \widehat{M}_L} \circ \mathcal{N}_{A \rightarrow B}\} [|m\rangle\langle m|_M]]}. \quad (\text{S41})$$

The *worst-case classical error probability* associated with the strategy is then given by

$$p_{\text{error}}((\mathcal{E}_{ML \rightarrow A}, \mathcal{D}_{B \rightarrow \widehat{M}_L}), \mathcal{N}_{A \rightarrow B}) := 1 - \min_m \Pr\{m|m; (\mathcal{E}_{ML \rightarrow A}, \mathcal{D}_{B \rightarrow \widehat{M}_L}), \mathcal{N}_{A \rightarrow B}\}, \quad (\text{S42})$$

where the minimization is over all symbols in the alphabet. It upper bounds the error probability for all possible classical messages.

The *one-shot retrocausal classical capacity* of a completely positive map  $\mathcal{N}_{A \rightarrow B} \in \mathcal{L}_{A \rightarrow B}$  within error  $\varepsilon \in [0, 1]$  is defined as

$$C_{\text{retro}}^\varepsilon(\mathcal{N}) := \sup_{\substack{d_M, d_L \in \mathbb{N}, \\ \mathcal{E}_{ML \rightarrow A} \in \mathcal{C}_{ML \rightarrow A}, \\ \mathcal{D}_{B \rightarrow \widehat{M}_L} \in \mathcal{C}_{B \rightarrow \widehat{M}_L}}} \left\{ \log_2 d_M : p_{\text{error}}((\mathcal{E}_{ML \rightarrow A}, \mathcal{D}_{B \rightarrow \widehat{M}_L}), \mathcal{N}_{A \rightarrow B}) \leq \varepsilon \right\}, \quad (\text{S43})$$

which quantifies the maximum number of qubits that can be transmitted within error  $\varepsilon$  from the father to the daughter through the noisy P-CTC represented by  $\mathcal{N}_{A \rightarrow B}$  under an encoding and decoding of their choice.

The *asymptotic retrocausal classical capacity* of a completely positive map  $\mathcal{N}_{A \rightarrow B} \in \mathcal{L}_{A \rightarrow B}$  is defined as

$$C_{\text{retro}}(\mathcal{N}) := \inf_{\varepsilon \in (0,1)} \liminf_{n \rightarrow \infty} \frac{1}{n} C_{\text{retro}}^\varepsilon(\mathcal{N}^{\otimes n}), \quad (\text{S44})$$

which quantifies the maximum rate at which the number of reliably transmitted qubits increases with respect to the number of parallel uses of the noisy P-CTC. It quantifies the maximum rate at which the number of reliably transmitted classical bits increases with respect to the number of parallel uses of the noisy P-CTC.

The *error exponent of retrocausal classical communication* for a completely positive map  $\mathcal{N}_{A \rightarrow B} \in \mathcal{L}_{A \rightarrow B}$  at a transmission rate  $r \in [0, \infty)$  is defined as

$$E_{\text{retro,C}}^r(\mathcal{N}) := \liminf_{n \rightarrow \infty} \sup_{\substack{d_M, d_L \in \mathbb{N}, \\ \mathcal{E}_{ML \rightarrow A^n} \in \mathcal{C}_{ML \rightarrow A^n}, \\ \mathcal{D}_{B^n \rightarrow \widehat{ML}} \in \mathcal{C}_{B^n \rightarrow \widehat{ML}}} \left\{ -\frac{1}{n} \log_2 p_{\text{error}}((\mathcal{E}_{ML \rightarrow A^n}, \mathcal{D}_{B^n \rightarrow \widehat{ML}}), (\mathcal{N}^{\otimes n})_{A^n \rightarrow B^n}) : d_M \geq 2^{nr} \right\}, \quad (\text{S45})$$

which quantifies the maximum rate at which the error probability exponentially decays as the number of transmitted classical bits increases at rate  $r$ .

The *strong-converse exponent of retrocausal classical communication* for a completely positive map  $\mathcal{N}_{A \rightarrow B} \in \mathcal{L}_{A \rightarrow B}$  at a transmission rate  $r \in [0, \infty)$  is defined as

$$S_{\text{retro,C}}^r(\mathcal{N}) := \limsup_{n \rightarrow \infty} \inf_{\substack{d_M, d_L \in \mathbb{N}, \\ \mathcal{E}_{ML \rightarrow A^n} \in \mathcal{C}_{ML \rightarrow A^n}, \\ \mathcal{D}_{B^n \rightarrow \widehat{ML}} \in \mathcal{C}_{B^n \rightarrow \widehat{ML}}} \left\{ -\frac{1}{n} \log_2 (1 - p_{\text{error}}((\mathcal{E}_{ML \rightarrow A^n}, \mathcal{D}_{B^n \rightarrow \widehat{ML}}), (\mathcal{N}^{\otimes n})_{A^n \rightarrow B^n})) : d_M \geq 2^{nr} \right\}, \quad (\text{S46})$$

which quantifies the minimum rate at which the error probability exponentially approaches 1 as the number of transmitted classical bits increases at rate  $r$ .

## V. RETROCAUSAL QUANTUM CAPACITY

In this section, we precisely determine the one-shot retrocausal quantum capacity.

**Theorem 1** (One-shot retrocausal quantum capacity). *The one-shot retrocausal quantum capacity has a closed-form expression in terms of the max- and Doeblin informations: for a completely positive map  $\mathcal{N}_{A \rightarrow B} \in \mathcal{L}_{A \rightarrow B}$  and a real number  $\varepsilon \in (0, 1)$ ,*

$$Q_{\text{retro}}^\varepsilon(\mathcal{N}) = \log_2 \left[ \sqrt{\frac{\varepsilon}{1 - \varepsilon} 2^{I_{\text{max}}(\mathcal{N}) + I_{\text{doc}}(\mathcal{N})} + 1} \right], \quad (\text{S47})$$

where the max- and Doeblin informations are defined in Eqs. (S2) and (S3), respectively.

**Remark 4** (Efficient computability). Since the max- and Doeblin informations are both defined via semidefinite programs (SDPs) [see Eqs. (S2b) and (S3b)], the one-shot retrocausal quantum capacity is efficiently computable as a closed-form formula thereof.

In what follows, we prove Theorem 1 by first establishing the desired converse bound on the one-shot retrocausal quantum capacity and then showing that it is achievable by construction. The basic idea of the proof is similar to that of the results in Ref. [30].

### A. Converse bound

As shown in the following lemma, if the father and the daughter were connected by a noisy P-CTC represented by a replacement channel, then the resulting transformation from  $M$  to  $\hat{M}$  is necessarily also a replacement channel and thus useless for communication, regardless of the strategy they adopt. A similar statement can be made for affine combinations of replacement channels as well.

**Lemma 3** (Useless noisy P-CTC). *For a strategy  $(\mathcal{E}_{ML \rightarrow A}, \mathcal{D}_{B \rightarrow \hat{M}L})$  and a unit-trace Hermitian operator  $\tau_B \in \text{aff}(\mathcal{D}_B)$ , there exists another unit-trace Hermitian operator  $\eta_{\hat{M}} \in \text{aff}(\mathcal{D}_{\hat{M}})$  such that*

$$\Gamma_A \{ \mathcal{E}_{ML \rightarrow A} \circ \mathcal{D}_{B \rightarrow \hat{M}L} \circ \mathcal{R}_{A \rightarrow B}^\tau \} = \frac{1}{d_A^2} \mathcal{R}_{M \rightarrow \hat{M}}^\eta. \quad (\text{S48})$$

Furthermore, if  $\tau_B$  is a state, then so is  $\eta_{\hat{M}}$ .

*Proof.* Define a unit-trace Hermitian operator  $\eta_{\hat{M}} := \mathcal{D}_{B \rightarrow \hat{M}}[\tau_B]$ . Note that  $\eta_{\hat{M}}$  is a state if  $\tau_B$  is. Then it follows from the definition of the loop supermap [Eq. (S13)] that, for every linear operator  $\rho_{RM} \in \mathbb{B}_{RM}$ ,

$$\Gamma_A \{ \mathcal{E}_{ML \rightarrow A} \circ \mathcal{D}_{B \rightarrow \hat{M}L} \circ \mathcal{R}_{A \rightarrow B}^\tau \} [\rho_{RM}] = \text{Tr}_{AA'} [\Phi_{AA'} (\mathcal{E}_{ML \rightarrow A} \circ \mathcal{D}_{B \rightarrow \hat{M}L} \circ \mathcal{R}_{A \rightarrow B}^\tau) [\rho_{RM} \otimes \Phi_{AA'}]] \quad (\text{S49})$$

$$= \text{Tr}_{AA'} [\Phi_{AA'} (\mathcal{E}_{ML \rightarrow A} \circ \mathcal{D}_{B \rightarrow \hat{M}L}) [\rho_{RM} \otimes \pi_{A'} \otimes \tau_B]] \quad (\text{S50})$$

$$= \frac{1}{d_A} \text{Tr}_A [\pi_A (\mathcal{E}_{ML \rightarrow A} \circ \mathcal{D}_{B \rightarrow \hat{M}L}) [\rho_{RM} \otimes \tau_B]] \quad (\text{S51})$$

$$= \frac{1}{d_A^2} \text{Tr}_{ML} [\mathcal{D}_{B \rightarrow \hat{M}L} [\rho_{RM} \otimes \tau_B]] \quad (\text{S52})$$

$$= \frac{1}{d_A^2} \rho_R \otimes \mathcal{D}_{B \rightarrow \hat{M}}[\tau_B] \quad (\text{S53})$$

$$= \frac{1}{d_A^2} \mathcal{R}_{M \rightarrow \hat{M}}^\eta [\rho_{RM}], \quad (\text{S54})$$

which implies the desired statement.  $\square$

We now prove the desired converse bound on the one-shot retrocausal quantum capacity.

**Proposition 1** (Converse bound on retrocausal quantum capacity). *For a completely positive map  $\mathcal{N}_{A \rightarrow B} \in \mathcal{L}_{A \rightarrow B}$  and a real number  $\varepsilon \in (0, 1)$ ,*

$$Q_{\text{retro}}^\varepsilon(\mathcal{N}) \leq \log_2 \left[ \sqrt{\frac{\varepsilon}{1-\varepsilon}} 2^{I_{\max}(\mathcal{N}) + I_{\text{doe}}(\mathcal{N})} + 1 \right]. \quad (\text{S55})$$

*Proof.* Let  $(\mathcal{E}_{ML \rightarrow A}, \mathcal{D}_{B \rightarrow \hat{M}L})$  be a strategy such that

$$q_{\text{infd}}((\mathcal{E}_{ML \rightarrow A}, \mathcal{D}_{B \rightarrow \hat{M}L}), \mathcal{N}_{A \rightarrow B}) \leq \varepsilon. \quad (\text{S56})$$

Let  $\lambda, \mu \in \mathbb{R}$  be two real numbers,  $\sigma_B \in \mathcal{D}_B$  a state, and  $\tau_B \in \text{aff}(\mathcal{D}_B)$  a unit-trace Hermitian operator, such that

$$\Phi_{A'B}^\mathcal{N} \leq \lambda \Phi_{A'B}^{\mathcal{R}^\sigma}, \quad (\text{S57})$$

$$\Phi_{A'B}^{\mathcal{R}^\tau} \leq \mu \Phi_{A'B}^\mathcal{N}. \quad (\text{S58})$$

Equation (S57) means that  $\lambda \mathcal{R}_{A \rightarrow B}^\sigma - \mathcal{N}_{A \rightarrow B}$  is a completely positive map, and therefore so is  $\lambda \Gamma_A \{ \mathcal{E}_{ML \rightarrow A} \circ \mathcal{D}_{B \rightarrow \hat{M}L} \circ \mathcal{R}_{A \rightarrow B}^\sigma \} - \Gamma_A \{ \mathcal{E}_{ML \rightarrow A} \circ \mathcal{D}_{B \rightarrow \hat{M}L} \circ \mathcal{N}_{A \rightarrow B} \}$ . This implies that

$$\Gamma_A \{ \mathcal{E}_{ML \rightarrow A} \circ \mathcal{D}_{B \rightarrow \hat{M}L} \circ \mathcal{N}_{A \rightarrow B} \} [\Phi_{M'M}] \leq \lambda \Gamma_A \{ \mathcal{E}_{ML \rightarrow A} \circ \mathcal{D}_{B \rightarrow \hat{M}L} \circ \mathcal{R}_{A \rightarrow B}^\sigma \} [\Phi_{M'M}] \quad (\text{S59})$$

$$= \frac{\lambda}{d_A^2} \mathcal{R}_{M \rightarrow \hat{M}}^\omega [\Phi_{M'M}] \quad (\text{S60})$$

$$= \frac{\lambda}{d_A^2} \pi_{M'} \otimes \omega_{\hat{M}}, \quad (\text{S61})$$

where  $\omega_{\hat{M}} \in \mathcal{D}_{\hat{M}}$  is a state, whose existence is guaranteed by Lemma 3. Likewise, it follows from Eq. (S58) that

$$\mu \Gamma_A \{ \mathcal{E}_{ML \rightarrow A} \circ \mathcal{D}_{B \rightarrow \hat{M}L} \circ \mathcal{N}_{A \rightarrow B} \} [\Phi_{M'M}] \geq \Gamma_A \{ \mathcal{E}_{ML \rightarrow A} \circ \mathcal{D}_{B \rightarrow \hat{M}L} \circ \mathcal{R}_{A \rightarrow B}^\tau \} [\Phi_{M'M}] \quad (\text{S62})$$

$$= \frac{1}{d_A^2} \mathcal{R}_{M \rightarrow \hat{M}}^\eta [\Phi_{M'M}] \quad (\text{S63})$$

$$= \frac{1}{d_A^2} \pi_{M'} \otimes \eta_{\hat{M}}, \quad (\text{S64})$$

where  $\eta_{\hat{M}} \in \text{aff}(\mathcal{D}_{\hat{M}})$  is a unit-trace Hermitian operator, whose existence is guaranteed by Lemma 3. Then it follows from Eq. (S56) that

$$1 - \varepsilon \leq 1 - q_{\text{infid}}((\mathcal{E}_{ML \rightarrow A}, \mathcal{D}_{B \rightarrow \hat{M}L}), \mathcal{N}_{A \rightarrow B}) \quad (\text{S65})$$

$$\leq \frac{\text{Tr}[\Phi_{M'\hat{M}} \Gamma_A \{ \mathcal{E}_{ML \rightarrow A} \circ \mathcal{D}_{B \rightarrow \hat{M}L} \circ \mathcal{N}_{A \rightarrow B} \} [\Phi_{M'M}]]}{\text{Tr}[\Gamma_A \{ \mathcal{E}_{ML \rightarrow A} \circ \mathcal{D}_{B \rightarrow \hat{M}L} \circ \mathcal{N}_{A \rightarrow B} \} [\Phi_{M'M}]]} \quad (\text{S66})$$

$$\leq \frac{\lambda \text{Tr}[\Phi_{M'\hat{M}} (\pi_{M'} \otimes \omega_{\hat{M}})]}{d_A^2 \text{Tr}[\Gamma_A \{ \mathcal{E}_{ML \rightarrow A} \circ \mathcal{D}_{B \rightarrow \hat{M}L} \circ \mathcal{N}_{A \rightarrow B} \} [\Phi_{M'M}]]} \quad (\text{S67})$$

$$= \frac{\lambda}{d_M^2 d_A^2 \text{Tr}[\Gamma_A \{ \mathcal{E}_{ML \rightarrow A} \circ \mathcal{D}_{B \rightarrow \hat{M}L} \circ \mathcal{N}_{A \rightarrow B} \} [\Phi_{M'M}]]}, \quad (\text{S68})$$

where Eq. (S66) follows from Eqs. (S36) and (S35) and choosing  $R = M'$  and  $\psi_{RM} = \Phi_{M'M}$ ; Eq. (S67) follows from Eq. (S61). Similarly,

$$\mu \varepsilon \geq \mu q_{\text{infid}}((\mathcal{E}_{ML \rightarrow A}, \mathcal{D}_{B \rightarrow \hat{M}L}), \mathcal{N}_{A \rightarrow B}) \quad (\text{S69})$$

$$\geq \frac{\mu \text{Tr}[(\mathbb{1}_{M'\hat{M}} - \Phi_{M'\hat{M}}) \Gamma_A \{ \mathcal{E}_{ML \rightarrow A} \circ \mathcal{D}_{B \rightarrow \hat{M}L} \circ \mathcal{N}_{A \rightarrow B} \} [\Phi_{M'M}]]}{\text{Tr}[\Gamma_A \{ \mathcal{E}_{ML \rightarrow A} \circ \mathcal{D}_{B \rightarrow \hat{M}L} \circ \mathcal{N}_{A \rightarrow B} \} [\Phi_{M'M}]]} \quad (\text{S70})$$

$$\geq \frac{\text{Tr}[(\mathbb{1}_{M'\hat{M}} - \Phi_{M'\hat{M}}) (\pi_{M'} \otimes \eta_{\hat{M}})]}{d_A^2 \text{Tr}[\Gamma_A \{ \mathcal{E}_{ML \rightarrow A} \circ \mathcal{D}_{B \rightarrow \hat{M}L} \circ \mathcal{N}_{A \rightarrow B} \} [\Phi_{M'M}]]} \quad (\text{S71})$$

$$= \left(1 - \frac{1}{d_M^2}\right) \frac{1}{d_A^2 \text{Tr}[\Gamma_A \{ \mathcal{E}_{ML \rightarrow A} \circ \mathcal{D}_{B \rightarrow \hat{M}L} \circ \mathcal{N}_{A \rightarrow B} \} [\Phi_{M'M}]]}, \quad (\text{S72})$$

where Eq. (S71) follows from Eq. (S64). Note that combining Eqs. (S68) and (S72) implies that

$$d_M^2 \leq \frac{\varepsilon \lambda \mu}{1 - \varepsilon} + 1, \quad (\text{S73})$$

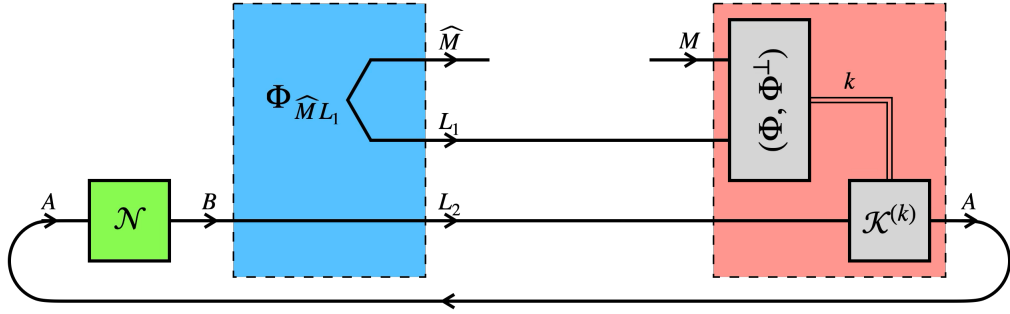


FIG. S6. The basic idea of amplified probabilistic quantum teleportation is to first perform probabilistic teleportation from the father to the daughter, and then engineer a causal loop involving the noisy P-CTC such that it maximally amplifies the renormalized probability of the teleportation being successful. The red and blue boxes encapsulate the father's encoding and the daughter's decoding, respectively.

which, due to  $d_M$  being an integer, further implies that

$$\log_2 d_M \leq \log_2 \left[ \sqrt{\frac{\varepsilon \lambda \mu}{1 - \varepsilon}} + 1 \right]. \quad (\text{S74})$$

Supremizing the left-hand side over every  $d_M, d_L \in \mathbb{N}$  and every strategy  $(\mathcal{E}_{ML \rightarrow A}, \mathcal{D}_{B \rightarrow \hat{M}L})$  satisfying Eq. (S56), and infimizing the right-hand side over every  $\lambda, \mu \in \mathbb{R}$ ,  $\sigma_B \in \mathcal{D}_B$ , and  $\tau_B \in \text{aff}(\mathcal{D}_B)$  satisfying Eqs. (S57) and (S58), the desired statement thus follows from the definitions of the one-shot retrocausal quantum capacity and the max- and Doeblin informations [Eqs. (S37), (S2), and (S3)].  $\square$

## B. Achievability

We now explicitly construct a strategy that saturates the converse bound derived in Proposition 1.

**Strategy 1** (Amplified probabilistic quantum teleportation). The strategy operates as follows (see Fig. S6).

- **Encoding.** In the future, the father retrieves two quantum memories  $L_1$  and  $L_2$  (which were left by Bella in the past) with  $L_1 \cong M$  and  $L_2 \cong B$ . He performs a measurement on  $M$  and  $L_1$  according to the POVM  $(\Phi_{ML_1}, \mathbb{1}_{ML_1} - \Phi_{ML_1})$ , obtaining one of the two possible outcomes,  $k \in \{0, 1\}$ . Conditioned on  $k$ , he applies a channel  $\mathcal{K}_{L_2 \rightarrow A}^{(k)} \in \mathcal{C}_{L_2 \rightarrow A}$  to  $L_2$ . He then feeds  $A$  into the entrance of the noisy P-CTC.
- **Decoding.** In the past, the daughter receives  $B$  from the exit of the noisy P-CTC. She prepares a standard maximally entangled state  $\Phi_{\hat{M}L_1}$  between  $\hat{M}$  and  $L_1$ , renames  $B$  as  $L_2$ , and leaves both  $L_1$  and  $L_2$  for the father to retrieve in the future. She then declares that  $\hat{M}$  stores her recovery of the message.

The intuition behind Strategy 1 is as follows. Instead of trying to code his message into  $A$  which enters the noisy P-CTC, the father intends to “teleport” [14] the message probabilistically to the daughter via  $L_1$ . Indeed, by probabilistic teleportation, whenever the father obtains the outcome  $k = 0$  after performing the measurement on  $ML_1$ , what the daughter ends up with in  $\hat{M}$  is a perfect recovery of the message in  $M$ . However, the outcome  $k = 0$  is not guaranteed to occur; in fact, it occurs only with probability  $1/d_M^2$ , or at least so without invoking the noisy P-CTC. Yet, as illustrated in Fig. S6, the remaining part of the strategy engineers a causal loop solely consisting of the noisy P-CTC, represented by  $\mathcal{N}_{A \rightarrow B}$ , and the channel  $\mathcal{K}_{L_2 \rightarrow A}^{(k)}$ , conditioned on  $k \in \{0, 1\}$ . Due to the nonlinearity of such a loop (specifically, its capability of renormalizing the probabilities of certain events), the father can choose the two channels  $\mathcal{K}_{L_2 \rightarrow A}^{(0)}$  and  $\mathcal{K}_{L_2 \rightarrow A}^{(1)}$

in such a way that the renormalized probability of the event  $k = 0$  is maximally amplified (or equivalently, the renormalized probability of  $k = 1$  is maximally diminished). If the renormalized probability of  $k = 0$  can be amplified to at least  $1 - \varepsilon$  for a message system  $M$  of dimension  $d_M$ , then  $\log_2 d_M$  qubits can indeed be “teleported” from  $M$  to  $\widehat{M}$  via  $L_1$  subject to a failure probability at most  $\varepsilon$ , which upper bounds the overall infidelity of communication.

We now formalize this analysis and show that Strategy 1 can indeed achieve the maximum possible efficiency allowed by the converse bound established in Proposition 1.

**Proposition 2** (Achievability of retrocausal quantum capacity). *For a completely positive map  $\mathcal{N}_{A \rightarrow B} \in \mathcal{L}_{A \rightarrow B}$  and a real number  $\varepsilon \in (0, 1)$ ,*

$$Q_{\text{retro}}^\varepsilon(\mathcal{N}) \geq \log_2 \left\lfloor \sqrt{\frac{\varepsilon}{1 - \varepsilon}} 2^{I_{\max}(\mathcal{N}) + I_{\text{doc}}(\mathcal{N})} + 1 \right\rfloor. \quad (\text{S75})$$

Furthermore, the efficiency on the right-hand side of Eq. (S75) is achievable by Strategy 1 for some channels  $\mathcal{K}_{L_2 \rightarrow A}^{(0)}$  and  $\mathcal{K}_{L_2 \rightarrow A}^{(1)}$ .

*Proof.* The encoding and decoding of Strategy 1 are represented by the following channels, respectively:

$$\mathcal{E}_{ML_1 L_2 \rightarrow A}[\cdot] := \text{Tr}_{ML_1} \left[ \Phi_{ML_1} \mathcal{K}_{L_2 \rightarrow A}^{(0)}[\cdot] \right] + \text{Tr}_{ML_1} \left[ (\mathbb{1}_{ML_1} - \Phi_{ML_1}) \mathcal{K}_{L_2 \rightarrow A}^{(1)}[\cdot] \right], \quad (\text{S76})$$

$$\mathcal{D}_{B \rightarrow \widehat{M} L_1 L_2}[\cdot] := \Phi_{\widehat{M} L_1} \otimes \text{id}_{B \rightarrow L_2}[\cdot]. \quad (\text{S77})$$

Let  $\psi_{RM}$  be a pure state. Then it follows from the definition of the loop supermap [Eq. (S13)] that

$$\begin{aligned} & \Gamma_A \left\{ \mathcal{E}_{ML_1 L_2 \rightarrow A} \circ \mathcal{D}_{B \rightarrow \widehat{M} L_1 L_2} \circ \mathcal{N}_{A \rightarrow B} \right\} [\psi_{RM}] \\ &= \text{Tr}_{AA'} \left[ \Phi_{AA'} \left( \mathcal{E}_{ML_1 L_2 \rightarrow A} \circ \mathcal{D}_{B \rightarrow \widehat{M} L_1 L_2} \circ \mathcal{N}_{A \rightarrow B} \right) [\psi_{RM} \otimes \Phi_{AA'}] \right] \end{aligned} \quad (\text{S78})$$

$$\begin{aligned} &= \text{Tr}_{ML_1 AA'} \left[ (\Phi_{ML_1} \otimes \Phi_{AA'}) \left( \mathcal{K}_{B \rightarrow A}^{(0)} \circ \mathcal{N}_{A \rightarrow B} \right) [\psi_{RM} \otimes \Phi_{\widehat{M} L_1} \otimes \Phi_{AA'}] \right] \\ &\quad + \text{Tr}_{ML_1 AA'} \left[ ((\mathbb{1}_{ML_1} - \Phi_{ML_1}) \otimes \Phi_{AA'}) \left( \mathcal{K}_{B \rightarrow A}^{(1)} \circ \mathcal{N}_{A \rightarrow B} \right) [\psi_{RM} \otimes \Phi_{\widehat{M} L_1} \otimes \Phi_{AA'}] \right] \end{aligned} \quad (\text{S79})$$

$$\begin{aligned} &= \frac{1}{d_M^2} \text{Tr} \left[ \Phi_{AA'} \left( \mathcal{K}_{B \rightarrow A}^{(0)} \circ \mathcal{N}_{A \rightarrow B} \right) [\Phi_{AA'}] \right] \psi_{R\widehat{M}} \\ &\quad + \frac{1}{d_M^2} \text{Tr} \left[ \Phi_{AA'} \left( \mathcal{K}_{B \rightarrow A}^{(1)} \circ \mathcal{N}_{A \rightarrow B} \right) [\Phi_{AA'}] \right] \left( d_M^2 \text{Tr}_M [\psi_{RM}] \otimes \pi_{\widehat{M}} - \psi_{R\widehat{M}} \right). \end{aligned} \quad (\text{S80})$$

Considering the fact that  $d_M^2 \text{Tr}_M [\psi_{RM}] \otimes \pi_{\widehat{M}} - \psi_{R\widehat{M}} \geq 0$ , it follows from Eq. (S35) that

$$\begin{aligned} & 1 - F(\psi_{R\widehat{M}} | \psi_{RM}; (\mathcal{E}_{ML_1 L_2 \rightarrow A}, \mathcal{D}_{B \rightarrow \widehat{M} L_1 L_2}, \mathcal{N}_{A \rightarrow B})) \\ &= 1 - \frac{\text{Tr} \left[ \psi_{R\widehat{M}} \Gamma_A \left\{ \mathcal{E}_{ML_1 L_2 \rightarrow A} \circ \mathcal{D}_{B \rightarrow \widehat{M} L_1 L_2} \circ \mathcal{N}_{A \rightarrow B} \right\} [\psi_{RM}] \right]}{\text{Tr} \left[ \Gamma_A \left\{ \mathcal{E}_{ML_1 L_2 \rightarrow A} \circ \mathcal{D}_{B \rightarrow \widehat{M} L_1 L_2} \circ \mathcal{N}_{A \rightarrow B} \right\} [\psi_{RM}] \right]} \end{aligned} \quad (\text{S81})$$

$$\begin{aligned} &\leq 1 - \frac{\frac{1}{d_M^2} \text{Tr} \left[ \Phi_{AA'} \left( \mathcal{K}_{B \rightarrow A}^{(0)} \circ \mathcal{N}_{A \rightarrow B} \right) [\Phi_{AA'}] \right]}{\frac{1}{d_M^2} \text{Tr} \left[ \Phi_{AA'} \left( \mathcal{K}_{B \rightarrow A}^{(0)} \circ \mathcal{N}_{A \rightarrow B} \right) [\Phi_{AA'}] \right] + \frac{d_M^2 - 1}{d_M^2} \text{Tr} \left[ \Phi_{AA'} \left( \mathcal{K}_{B \rightarrow A}^{(1)} \circ \mathcal{N}_{A \rightarrow B} \right) [\Phi_{AA'}] \right]} \end{aligned} \quad (\text{S82})$$

$$= \left( \frac{\text{Tr} \left[ \Phi_{AA'} \left( \mathcal{K}_{B \rightarrow A}^{(0)} \circ \mathcal{N}_{A \rightarrow B} \right) [\Phi_{AA'}] \right]}{(d_M^2 - 1) \text{Tr} \left[ \Phi_{AA'} \left( \mathcal{K}_{B \rightarrow A}^{(1)} \circ \mathcal{N}_{A \rightarrow B} \right) [\Phi_{AA'}] \right]} + 1 \right)^{-1}, \quad (\text{S83})$$

where Eq. (S82) follows from Eq. (S80). Since this holds for every pure state  $\psi_{RM}$ , it follows from Eq. (S36) that, as long as

$$\left( \frac{\text{Tr} \left[ \Phi_{AA'} \left( \mathcal{K}_{B \rightarrow A}^{(0)} \circ \mathcal{N}_{A \rightarrow B} \right) [\Phi_{AA'}] \right]}{(d_M^2 - 1) \text{Tr} \left[ \Phi_{AA'} \left( \mathcal{K}_{B \rightarrow A}^{(1)} \circ \mathcal{N}_{A \rightarrow B} \right) [\Phi_{AA'}] \right]} + 1 \right)^{-1} \leq \varepsilon, \quad (\text{S84})$$

we have that

$$q_{\text{infid}} \left( \left( \mathcal{E}_{ML_1 L_2 \rightarrow A}, \mathcal{D}_{B \rightarrow \widehat{M} L_1 L_2} \right), \mathcal{N}_{A \rightarrow B} \right) = 1 - \inf_{\substack{\psi_{RM} \in \mathcal{D}_{RM}: \\ \text{rank}(\psi_{RM})=1}} F \left( \psi_{RM} \middle| \psi_{RM}; \left( \mathcal{E}_{ML_1 L_2 \rightarrow A}, \mathcal{D}_{B \rightarrow \widehat{M} L_1 L_2} \right), \mathcal{N}_{A \rightarrow B} \right) \quad (\text{S85})$$

$$\leq \varepsilon. \quad (\text{S86})$$

Since Eq. (S84) is equivalent to

$$d_M^2 \leq \frac{\varepsilon \text{Tr} \left[ \Phi_{AA'} \left( \mathcal{K}_{B \rightarrow A}^{(0)} \circ \mathcal{N}_{A \rightarrow B} \right) [\Phi_{AA'}] \right]}{(1 - \varepsilon) \text{Tr} \left[ \Phi_{AA'} \left( \mathcal{K}_{B \rightarrow A}^{(1)} \circ \mathcal{N}_{A \rightarrow B} \right) [\Phi_{AA'}] \right]} + 1, \quad (\text{S87})$$

by choosing

$$\mathcal{K}_{B \rightarrow A}^{(0)} := \arg \max_{\mathcal{K}_{B \rightarrow A} \in \mathcal{C}_{B \rightarrow A}} \text{Tr} \left[ \Phi_{AA'} \left( \mathcal{K}_{B \rightarrow A} \circ \mathcal{N}_{A \rightarrow B} \right) [\Phi_{AA'}] \right], \quad (\text{S88})$$

$$\mathcal{K}_{B \rightarrow A}^{(1)} := \arg \min_{\mathcal{K}_{B \rightarrow A} \in \mathcal{C}_{B \rightarrow A}} \text{Tr} \left[ \Phi_{AA'} \left( \mathcal{K}_{B \rightarrow A} \circ \mathcal{N}_{A \rightarrow B} \right) [\Phi_{AA'}] \right], \quad (\text{S89})$$

which maximize the right-hand side of Eq. (S87), the following efficiency can be achieved within infidelity  $\varepsilon$  (i.e., subject to Eq. (S86)) and is thus a lower bound on the one-shot retrocausal quantum capacity [see Eq. (S37)]:

$$\log_2 d_M = \log_2 \left[ \sqrt{\frac{\varepsilon \text{Tr} \left[ \Phi_{AA'} \left( \mathcal{K}_{B \rightarrow A}^{(0)} \circ \mathcal{N}_{A \rightarrow B} \right) [\Phi_{AA'}] \right]}{(1 - \varepsilon) \text{Tr} \left[ \Phi_{AA'} \left( \mathcal{K}_{B \rightarrow A}^{(1)} \circ \mathcal{N}_{A \rightarrow B} \right) [\Phi_{AA'}] \right]} + 1} \right] \quad (\text{S90})$$

$$= \log_2 \left[ \sqrt{\frac{\varepsilon \max_{\mathcal{K}_{B \rightarrow A} \in \mathcal{C}_{B \rightarrow A}} \text{Tr} \left[ \Phi_{AA'} \left( \mathcal{K}_{B \rightarrow A} \circ \mathcal{N}_{A \rightarrow B} \right) [\Phi_{AA'}] \right]}{(1 - \varepsilon) \min_{\mathcal{K}_{B \rightarrow A} \in \mathcal{C}_{B \rightarrow A}} \text{Tr} \left[ \Phi_{AA'} \left( \mathcal{K}_{B \rightarrow A} \circ \mathcal{N}_{A \rightarrow B} \right) [\Phi_{AA'}] \right]} + 1} \right] \quad (\text{S91})$$

$$= \log_2 \left[ \sqrt{\frac{\varepsilon}{1 - \varepsilon} 2^{I_{\max}(\mathcal{N}) + I_{\text{doe}}(\mathcal{N})} + 1} \right], \quad (\text{S92})$$

where Eq. (S92) follows from the interpretation of the max- and Doeblin informations in terms of the maximum and minimum achievable singlet fractions, respectively (Lemma 1), implying the desired statement.  $\square$

**Remark 5** (On infinite retrocausal quantum capacity). While the max-information of a completely positive map is always finite, the Doeblin information can be infinite. As can be inferred from the proof of Proposition 2, for a completely positive map whose Doeblin information is equal to infinity, its one-shot retrocausal quantum capacity is equal to infinity for all  $\varepsilon \in [0, 1]$ , including the zero-error case.

Combining Propositions 1 and 2 completes the proof of Theorem 1.

### C. Asymptotics

With the one-shot retrocausal quantum capacity determined, the asymptotic retrocausal quantum capacity immediately follows.

**Theorem 2** (Asymptotic retrocausal quantum capacity). *The asymptotic retrocausal quantum capacity is equal to one half of the sum of the max- and regularized Doeblin informations: for a completely positive map  $\mathcal{N}_{A \rightarrow B} \in \mathcal{L}_{A \rightarrow B}$  and a real number  $\varepsilon \in (0, 1)$ ,*

$$Q_{\text{retro}}(\mathcal{N}) = \lim_{n \rightarrow \infty} \frac{1}{n} Q_{\text{retro}}^\varepsilon(\mathcal{N}^{\otimes n}) = \frac{1}{2} (I_{\text{max}}(\mathcal{N}) + I_{\text{doe}}^\infty(\mathcal{N})), \quad (\text{S93})$$

where the regularized Doeblin information is defined in Eq. (S9).

*Proof.* Since the max- and Doeblin informations are nonnegative, it follows from Theorem 1 that

$$Q_{\text{retro}}^\varepsilon(\mathcal{N}) = \log_2 \left[ \sqrt{\frac{\varepsilon}{1-\varepsilon} 2^{I_{\text{max}}(\mathcal{N}) + I_{\text{doe}}(\mathcal{N})} + 1} \right] \quad (\text{S94})$$

$$\leq \frac{1}{2} \log_2 \left( \frac{\varepsilon}{1-\varepsilon} 2^{I_{\text{max}}(\mathcal{N}) + I_{\text{doe}}(\mathcal{N})} + 1 \right) \quad (\text{S95})$$

$$\leq \frac{1}{2} \log_2 \left( \left( \frac{\varepsilon}{1-\varepsilon} + 1 \right) 2^{I_{\text{max}}(\mathcal{N}) + I_{\text{doe}}(\mathcal{N})} \right) \quad (\text{S96})$$

$$= \frac{1}{2} \left( I_{\text{max}}(\mathcal{N}) + I_{\text{doe}}(\mathcal{N}) + \log_2 \frac{1}{1-\varepsilon} \right). \quad (\text{S97})$$

Due to the fact that  $\lfloor \sqrt{x+1} \rfloor \geq \frac{1}{2} \sqrt{x}$  for every nonnegative real number  $x \in [0, \infty)$ , we also have that

$$Q_{\text{retro}}^\varepsilon(\mathcal{N}) = \log_2 \left[ \sqrt{\frac{\varepsilon}{1-\varepsilon} 2^{I_{\text{max}}(\mathcal{N}) + I_{\text{doe}}(\mathcal{N})} + 1} \right] \quad (\text{S98})$$

$$\geq \log_2 \left( \frac{1}{2} \sqrt{\frac{\varepsilon}{1-\varepsilon} 2^{I_{\text{max}}(\mathcal{N}) + I_{\text{doe}}(\mathcal{N})}} \right) \quad (\text{S99})$$

$$= \frac{1}{2} \left( I_{\text{max}}(\mathcal{N}) + I_{\text{doe}}(\mathcal{N}) + \log_2 \frac{\varepsilon}{4(1-\varepsilon)} \right). \quad (\text{S100})$$

Combining Eqs. (S97) and (S100), we have that, for every  $n \in \mathbb{N}$ ,

$$\begin{aligned} \frac{1}{2n} \left( I_{\text{max}}(\mathcal{N}^{\otimes n}) + I_{\text{doe}}(\mathcal{N}^{\otimes n}) + \log_2 \frac{1}{1-\varepsilon} \right) &\leq \frac{1}{n} Q_{\text{retro}}^\varepsilon(\mathcal{N}^{\otimes n}) \\ &\leq \frac{1}{2n} \left( I_{\text{max}}(\mathcal{N}^{\otimes n}) + I_{\text{doe}}(\mathcal{N}^{\otimes n}) + \log_2 \frac{\varepsilon}{4(1-\varepsilon)} \right). \end{aligned} \quad (\text{S101})$$

Taking the limit  $n \rightarrow \infty$ , it follows from the additivity of the max-information [Eq. (S8)] and the definition of the regularized Doeblin information [Eq. (S9)] that

$$\lim_{n \rightarrow \infty} \frac{1}{n} Q_{\text{retro}}^\varepsilon(\mathcal{N}^{\otimes n}) = \frac{1}{2} (I_{\text{max}}(\mathcal{N}) + I_{\text{doe}}^\infty(\mathcal{N})), \quad (\text{S102})$$

from which and the definition of the asymptotic retrocausal quantum capacity [Eq. (S38)] the desired statement follows.  $\square$

**Remark 6** (Efficiently computable bounds). For quantum-to-classical completely positive maps (which include measurement channels) and mictodiactive covariant completely positive maps (which include depolarizing channels), the Doeblin information is invariant under regularization [61, Theorem 3], and thus the asymptotic retrocausal quantum capacity has a single-letter expression and is efficiently computable via SDPs for such maps. While it is unknown whether the asymptotic retrocausal quantum capacity has a single-letter expression for general completely positive maps, single-letter and SDP-computable bounds on it from both sides can be established. Specifically, it follows from Theorem 2 and Eq. (S12) that

$$\frac{1}{2} (I_{\max}(\mathcal{N}) + I_{\text{doe}}(\mathcal{N})) \leq Q_{\text{retro}}(\mathcal{N}) \leq \frac{1}{2} (I_{\max}(\mathcal{N}) + I_{\Xi}(\mathcal{N})), \quad (\text{S103})$$

where  $I_{\Xi}(\mathcal{N})$  is defined in Eq. (S11).

**Remark 7** (Second-order asymptotics). Equation (S101) implies that, for a completely positive map whose regularized Doeblin information is equal to its Doeblin information (e.g., quantum-to-classical completely positive maps and mictodiactive covariant completely positive maps), the second-order asymptotics of retrocausal quantum communication becomes trivial. However, it is unclear to us whether nontrivial second-order asymptotics exists for general completely positive maps.

#### D. Error exponent and strong-converse exponent

We now show that the error exponent and the strong-converse exponent of retrocausal quantum communication can also be determined based on the one-shot retrocausal quantum capacity.

**Theorem 3** (Error exponent and strong-converse exponent of retrocausal quantum communication). *For a completely positive map  $\mathcal{N}_{A \rightarrow B} \in \mathcal{L}_{A \rightarrow B}$  and a real number  $r \in [0, \infty)$ ,*

$$E_{\text{retro},Q}^r(\mathcal{N}) = \max \{0, I_{\max}(\mathcal{N}) + I_{\text{doe}}^{\infty}(\mathcal{N}) - 2r\}, \quad (\text{S104})$$

$$S_{\text{retro},Q}^r(\mathcal{N}) = \max \{0, 2r - I_{\max}(\mathcal{N}) - I_{\text{doe}}^{\infty}(\mathcal{N})\}. \quad (\text{S105})$$

*Proof.* Let  $(\mathcal{E}_{ML \rightarrow A}, \mathcal{D}_{B \rightarrow \widehat{ML}})$  be a strategy. Let  $\lambda, \mu \in \mathbb{R}$  be two real numbers,  $\sigma_B \in \mathcal{D}_B$  a state, and  $\tau_B \in \text{aff}(\mathcal{D}_B)$  a unit-trace Hermitian operator satisfying Eqs. (S57) and (S58). Comparing Eqs. (S65) with (S68) and Eqs. (S69) with (S72), we have that

$$1 - q_{\text{infid}}((\mathcal{E}_{ML \rightarrow A}, \mathcal{D}_{B \rightarrow \widehat{ML}}), \mathcal{N}_{A \rightarrow B}) \leq \frac{\lambda}{d_M^2 d_A^2 \text{Tr}[\Gamma_A \{\mathcal{E}_{ML \rightarrow A} \circ \mathcal{D}_{B \rightarrow \widehat{ML}} \circ \mathcal{N}_{A \rightarrow B}\} [\Phi_{M'M}]]}, \quad (\text{S106})$$

$$\mu q_{\text{infid}}((\mathcal{E}_{ML \rightarrow A}, \mathcal{D}_{B \rightarrow \widehat{ML}}), \mathcal{N}_{A \rightarrow B}) \geq \left(1 - \frac{1}{d_M^2}\right) \frac{1}{d_A^2 \text{Tr}[\Gamma_A \{\mathcal{E}_{ML \rightarrow A} \circ \mathcal{D}_{B \rightarrow \widehat{ML}} \circ \mathcal{N}_{A \rightarrow B}\} [\Phi_{M'M}]]}. \quad (\text{S107})$$

Note that combining Eqs. (S106) and (S107) implies that

$$q_{\text{infid}}((\mathcal{E}_{ML \rightarrow A}, \mathcal{D}_{B \rightarrow \widehat{ML}}), \mathcal{N}_{A \rightarrow B}) \geq \frac{d_M^2 - 1}{\lambda \mu + d_M^2 - 1}. \quad (\text{S108})$$

For a fixed  $d_M$ , infimizing the left-hand side over every  $d_L \in \mathbb{N}$  and every strategy  $(\mathcal{E}_{ML \rightarrow A}, \mathcal{D}_{B \rightarrow \widehat{ML}})$ , and supremizing the right-hand side over every  $\lambda, \mu \in \mathbb{R}$ ,  $\sigma_B \in \mathcal{D}_B$ , and  $\tau_B \in \text{aff}(\mathcal{D}_B)$  satisfying Eqs. (S57) and (S58), we have that

$$\inf_{\substack{d_L \in \mathbb{N}, \\ \mathcal{E}_{ML \rightarrow A} \in \mathcal{C}_{ML \rightarrow A}, \\ \mathcal{D}_{B \rightarrow \widehat{ML}} \in \mathcal{C}_{B \rightarrow \widehat{ML}}}} q_{\text{infid}}((\mathcal{E}_{ML \rightarrow A}, \mathcal{D}_{B \rightarrow \widehat{ML}}), \mathcal{N}_{A \rightarrow B}) \geq \frac{d_M^2 - 1}{2I_{\max}(\mathcal{N}) + I_{\text{doe}}(\mathcal{N}) + d_M^2 - 1}. \quad (\text{S109})$$

Now consider the strategy  $(\mathcal{E}_{ML_1L_2 \rightarrow A}, \mathcal{D}_{B \rightarrow \widehat{M}L_1L_2})$  defined in Eqs. (S76) and (S77). It follows from Eq. (S83) that, for every pure state  $\psi_{RM}$ ,

$$\begin{aligned} & 1 - F\left(\psi_{R\widehat{M}} \middle| \psi_{RM}; \left(\mathcal{E}_{ML_1L_2 \rightarrow A}, \mathcal{D}_{B \rightarrow \widehat{M}L_1L_2}\right), \mathcal{N}_{A \rightarrow B}\right) \\ & \leq \left( \frac{\text{Tr}\left[\Phi_{AA'} \left(\mathcal{K}_{B \rightarrow A}^{(0)} \circ \mathcal{N}_{A \rightarrow B}\right) [\Phi_{AA'}]\right]}{(d_M^2 - 1) \text{Tr}\left[\Phi_{AA'} \left(\mathcal{K}_{B \rightarrow A}^{(1)} \circ \mathcal{N}_{A \rightarrow B}\right) [\Phi_{AA'}]\right]} + 1 \right)^{-1}. \end{aligned} \quad (\text{S110})$$

Supremizing the left-hand side over every  $d_R \in \mathbb{N}$  and every pure state  $\psi_{RM}$ , and choosing the channels  $\mathcal{K}_{B \rightarrow A}^{(0)}$  and  $\mathcal{K}_{B \rightarrow A}^{(1)}$  according to Eqs. (S88) and (S89), it follows from Eq. (S36) that

$$\begin{aligned} & q_{\text{infid}}\left(\left(\mathcal{E}_{ML_1L_2 \rightarrow A}, \mathcal{D}_{B \rightarrow \widehat{M}L_1L_2}\right), \mathcal{N}_{A \rightarrow B}\right) \\ & \leq \left( \frac{\text{Tr}\left[\Phi_{AA'} \left(\mathcal{K}_{B \rightarrow A}^{(0)} \circ \mathcal{N}_{A \rightarrow B}\right) [\Phi_{AA'}]\right]}{(d_M^2 - 1) \text{Tr}\left[\Phi_{AA'} \left(\mathcal{K}_{B \rightarrow A}^{(1)} \circ \mathcal{N}_{A \rightarrow B}\right) [\Phi_{AA'}]\right]} + 1 \right)^{-1} \end{aligned} \quad (\text{S111})$$

$$= \left( \frac{\max_{\mathcal{K}_{B \rightarrow A} \in \mathcal{C}_{B \rightarrow A}} \text{Tr}\left[\Phi_{AA'} \left(\mathcal{K}_{B \rightarrow A} \circ \mathcal{N}_{A \rightarrow B}\right) [\Phi_{AA'}]\right]}{(d_M^2 - 1) \min_{\mathcal{K}_{B \rightarrow A} \in \mathcal{C}_{B \rightarrow A}} \text{Tr}\left[\Phi_{AA'} \left(\mathcal{K}_{B \rightarrow A} \circ \mathcal{N}_{A \rightarrow B}\right) [\Phi_{AA'}]\right]} + 1 \right)^{-1} \quad (\text{S112})$$

$$= \left( \frac{1}{d_M^2 - 1} 2^{I_{\max}(\mathcal{N}) + I_{\text{doe}}(\mathcal{N})} + 1 \right)^{-1} \quad (\text{S113})$$

$$= \frac{d_M^2 - 1}{2^{I_{\max}(\mathcal{N}) + I_{\text{doe}}(\mathcal{N})} + d_M^2 - 1}, \quad (\text{S114})$$

where Eq. (S113) follows from the interpretation of the max- and Doeblin informations in terms of the maximum and minimum achievable singlet fractions, respectively (Lemma 1). Combining Eqs. (S109) and (S114), we have that

$$\inf_{\substack{d_L \in \mathbb{N}, \\ \mathcal{E}_{ML \rightarrow A} \in \mathcal{C}_{ML \rightarrow A}, \\ \mathcal{D}_{B \rightarrow \widehat{M}L} \in \mathcal{C}_{B \rightarrow \widehat{M}L}}} q_{\text{infid}}\left(\left(\mathcal{E}_{ML \rightarrow A}, \mathcal{D}_{B \rightarrow \widehat{M}L}\right), \mathcal{N}_{A \rightarrow B}\right) = \frac{d_M^2 - 1}{2^{I_{\max}(\mathcal{N}) + I_{\text{doe}}(\mathcal{N})} + d_M^2 - 1}. \quad (\text{S115})$$

Then it follows from the definition of the error exponent of retrocausal quantum communication [Eq. (S39)] that

$$E_{\text{retro,Q}}^r(\mathcal{N}) = \lim_{n \rightarrow \infty} \sup_{d_M \in \mathbb{N}} \left\{ -\frac{1}{n} \log_2 \frac{d_M^2 - 1}{2^{I_{\max}(\mathcal{N}^{\otimes n}) + I_{\text{doe}}(\mathcal{N}^{\otimes n})} + d_M^2 - 1} : d_M \geq 2^{nr} \right\} \quad (\text{S116})$$

$$= \lim_{n \rightarrow \infty} -\frac{1}{n} \log_2 \frac{2^{2nr} - 1}{2^{I_{\max}(\mathcal{N}^{\otimes n}) + I_{\text{doe}}(\mathcal{N}^{\otimes n})} + 2^{2nr} - 1} \quad (\text{S117})$$

$$= \max \{0, I_{\max}(\mathcal{N}) + I_{\text{doe}}(\mathcal{N}) - 2r\}. \quad (\text{S118})$$

where Eq. (S118) follows from the additivity of the max-information [Eq. (S8)] and the definition of the regularized Doeblin information [Eq. (S9)]. Likewise, it follows from the definition of the strong-converse exponent of retrocausal quantum communication [Eq. (S40)] that

$$S_{\text{retro,Q}}^r(\mathcal{N}) = \lim_{n \rightarrow \infty} \inf_{d_M \in \mathbb{N}} \left\{ -\frac{1}{n} \log_2 \left( 1 - \frac{d_M^2 - 1}{2^{I_{\max}(\mathcal{N}^{\otimes n}) + I_{\text{doe}}(\mathcal{N}^{\otimes n})} + d_M^2 - 1} \right) : d_M \geq 2^{nr} \right\} \quad (\text{S119})$$

$$= \lim_{n \rightarrow \infty} -\frac{1}{n} \log_2 \frac{2^{I_{\max}(\mathcal{N}^{\otimes n}) + I_{\text{doe}}(\mathcal{N}^{\otimes n})}}{2^{I_{\max}(\mathcal{N}^{\otimes n}) + I_{\text{doe}}(\mathcal{N}^{\otimes n})} + 2^{2nr} - 1} \quad (\text{S120})$$

$$= \max \{0, 2r - I_{\max}(\mathcal{N}) - I_{\text{doe}}^{\infty}(\mathcal{N})\}, \quad (\text{S121})$$

which completes the proof of the desired statement.  $\square$

## VI. RETROCAUSAL CLASSICAL CAPACITY

In this section, we precisely determine the one-shot retrocausal classical capacity. All proofs follow a similar idea to their quantum counterparts.

**Theorem 4** (One-shot retrocausal classical capacity). *The one-shot retrocausal classical capacity has a closed-form expression in terms of the max- and Doeblin informations: for a completely positive map  $\mathcal{N}_{A \rightarrow B} \in \mathcal{L}_{A \rightarrow B}$  and a real number  $\varepsilon \in (0, 1)$ ,*

$$C_{\text{retro}}^{\varepsilon}(\mathcal{N}) = \log_2 \left[ \frac{\varepsilon}{1 - \varepsilon} 2^{I_{\max}(\mathcal{N}) + I_{\text{doe}}(\mathcal{N})} + 1 \right], \quad (\text{S122})$$

where the max- and Doeblin informations are defined in Eqs. (S2) and (S3), respectively.

As in the quantum case, we prove Theorem 4 by first establishing the desired converse bound on the one-shot retrocausal classical capacity and then showing that it is achievable by construction.

### A. Converse bound

We now prove the desired converse bound on the one-shot retrocausal classical capacity.

**Proposition 3** (Converse bound on retrocausal classical capacity). *For a completely positive map  $\mathcal{N}_{A \rightarrow B} \in \mathcal{L}_{A \rightarrow B}$  and a real number  $\varepsilon \in (0, 1)$ ,*

$$C_{\text{retro}}^{\varepsilon}(\mathcal{N}) \leq \log_2 \left[ \frac{\varepsilon}{1 - \varepsilon} 2^{I_{\max}(\mathcal{N}) + I_{\text{doe}}(\mathcal{N})} + 1 \right]. \quad (\text{S123})$$

*Proof.* Let  $(\mathcal{E}_{ML \rightarrow A}, \mathcal{D}_{B \rightarrow \hat{M}L})$  be a strategy such that

$$p_{\text{error}}((\mathcal{E}_{ML \rightarrow A}, \mathcal{D}_{B \rightarrow \hat{M}L}), \mathcal{N}_{A \rightarrow B}) \leq \varepsilon. \quad (\text{S124})$$

It follows from Eqs. (S42) and (S41) that, for every symbol  $m$  in the alphabet,

$$1 - \varepsilon \leq 1 - p_{\text{error}}((\mathcal{E}_{ML \rightarrow A}, \mathcal{D}_{B \rightarrow \hat{M}L}), \mathcal{N}_{A \rightarrow B}) \quad (\text{S125})$$

$$\leq \frac{\text{Tr}[|m\rangle\langle m|_{\hat{M}} \Gamma_A \{ \mathcal{E}_{ML \rightarrow A} \circ \mathcal{D}_{B \rightarrow \hat{M}L} \circ \mathcal{N}_{A \rightarrow B} \} [|m\rangle\langle m|_M]]}{\text{Tr}[\Gamma_A \{ \mathcal{E}_{ML \rightarrow A} \circ \mathcal{D}_{B \rightarrow \hat{M}L} \circ \mathcal{N}_{A \rightarrow B} \} [|m\rangle\langle m|_M]]}, \quad (\text{S126})$$

which implies that

$$1 - \varepsilon \leq 1 - p_{\text{error}}((\mathcal{E}_{ML \rightarrow A}, \mathcal{D}_{B \rightarrow \hat{M}L}), \mathcal{N}_{A \rightarrow B}) \quad (\text{S127})$$

$$\leq \frac{\sum_{m=0}^{d_M-1} \frac{1}{d_M} \text{Tr}[|m\rangle\langle m|_{\hat{M}} \Gamma_A \{ \mathcal{E}_{ML \rightarrow A} \circ \mathcal{D}_{B \rightarrow \hat{M}L} \circ \mathcal{N}_{A \rightarrow B} \} [|m\rangle\langle m|_M]]}{\sum_{m=0}^{d_M-1} \frac{1}{d_M} \text{Tr}[\Gamma_A \{ \mathcal{E}_{ML \rightarrow A} \circ \mathcal{D}_{B \rightarrow \hat{M}L} \circ \mathcal{N}_{A \rightarrow B} \} [|m\rangle\langle m|_M]]} \quad (\text{S128})$$

$$= \frac{\text{Tr}[\Omega_{M'\hat{M}} \Gamma_A \{ \mathcal{E}_{ML \rightarrow A} \circ \mathcal{D}_{B \rightarrow \hat{M}L} \circ \mathcal{N}_{A \rightarrow B} \} [\Upsilon_{M'M}]]}{\text{Tr}[\Gamma_A \{ \mathcal{E}_{ML \rightarrow A} \circ \mathcal{D}_{B \rightarrow \hat{M}L} \circ \mathcal{N}_{A \rightarrow B} \} [\Upsilon_{M'M}]]}, \quad (\text{S129})$$

where the standard classically maximally correlated state  $\Upsilon_{M'M}$  and the classical comparator  $\Omega_{M'\hat{M}}$  are defined in Sec. I. Let  $\lambda, \mu \in \mathbb{R}$  be two real numbers,  $\sigma_B \in \mathcal{D}_B$  a state, and  $\tau_B \in \text{aff}(\mathcal{D}_B)$  a unit-trace Hermitian operator, such that

$$\Phi_{A'B}^{\mathcal{N}} \leq \lambda \Phi_{A'B}^{\mathcal{R}^\sigma}, \quad (\text{S130})$$

$$\Phi_{A'B}^{\mathcal{R}^\tau} \leq \mu \Phi_{A'B}^{\mathcal{N}}. \quad (\text{S131})$$

Equation (S130) means that  $\lambda \mathcal{R}_{A \rightarrow B}^\sigma - \mathcal{N}_{A \rightarrow B}$  is a completely positive map, and therefore so is  $\lambda \Gamma_A \{\mathcal{E}_{ML \rightarrow A} \circ \mathcal{D}_{B \rightarrow \hat{M}L} \circ \mathcal{R}_{A \rightarrow B}^\sigma\} - \Gamma_A \{\mathcal{E}_{ML \rightarrow A} \circ \mathcal{D}_{B \rightarrow \hat{M}L} \circ \mathcal{N}_{A \rightarrow B}\}$ . This implies that

$$\Gamma_A \{\mathcal{E}_{ML \rightarrow A} \circ \mathcal{D}_{B \rightarrow \hat{M}L} \circ \mathcal{N}_{A \rightarrow B}\} [\Upsilon_{M'M}] \leq \lambda \Gamma_A \{\mathcal{E}_{ML \rightarrow A} \circ \mathcal{D}_{B \rightarrow \hat{M}L} \circ \mathcal{R}_{A \rightarrow B}^\sigma\} [\Upsilon_{M'M}] \quad (\text{S132})$$

$$= \frac{\lambda}{d_A^2} \mathcal{R}_{M \rightarrow \hat{M}}^\omega [\Upsilon_{M'M}] \quad (\text{S133})$$

$$= \frac{\lambda}{d_A^2} \pi_{M'} \otimes \omega_{\hat{M}}, \quad (\text{S134})$$

where  $\omega_{\hat{M}} \in \mathcal{D}_{\hat{M}}$  is a state, whose existence is guaranteed by Lemma 3. Likewise, it follows from Eq. (S131) that

$$\mu \Gamma_A \{\mathcal{E}_{ML \rightarrow A} \circ \mathcal{D}_{B \rightarrow \hat{M}L} \circ \mathcal{N}_{A \rightarrow B}\} [\Upsilon_{M'M}] \geq \Gamma_A \{\mathcal{E}_{ML \rightarrow A} \circ \mathcal{D}_{B \rightarrow \hat{M}L} \circ \mathcal{R}_{A \rightarrow B}^\tau\} [\Upsilon_{M'M}] \quad (\text{S135})$$

$$= \frac{1}{d_A^2} \mathcal{R}_{M \rightarrow \hat{M}}^\eta [\Upsilon_{M'M}] \quad (\text{S136})$$

$$= \frac{1}{d_A^2} \pi_{M'} \otimes \eta_{\hat{M}}, \quad (\text{S137})$$

where  $\eta_{\hat{M}} \in \text{aff}(\mathcal{D}_{\hat{M}})$  is a unit-trace Hermitian operator, whose existence is guaranteed by Lemma 3. Then it follows from Eq. (S129) that

$$1 - \varepsilon \leq 1 - p_{\text{error}}((\mathcal{E}_{ML \rightarrow A}, \mathcal{D}_{B \rightarrow \hat{M}L}), \mathcal{N}_{A \rightarrow B}) \quad (\text{S138})$$

$$\leq \frac{\text{Tr}[\Omega_{M'\hat{M}} \Gamma_A \{\mathcal{E}_{ML \rightarrow A} \circ \mathcal{D}_{B \rightarrow \hat{M}L} \circ \mathcal{N}_{A \rightarrow B}\} [\Upsilon_{M'M}]]}{\text{Tr}[\Gamma_A \{\mathcal{E}_{ML \rightarrow A} \circ \mathcal{D}_{B \rightarrow \hat{M}L} \circ \mathcal{N}_{A \rightarrow B}\} [\Upsilon_{M'M}]]} \quad (\text{S139})$$

$$\leq \frac{\lambda \text{Tr}[\Omega_{M'\hat{M}} (\pi_{M'} \otimes \omega_{\hat{M}})]}{d_A^2 \text{Tr}[\Gamma_A \{\mathcal{E}_{ML \rightarrow A} \circ \mathcal{D}_{B \rightarrow \hat{M}L} \circ \mathcal{N}_{A \rightarrow B}\} [\Upsilon_{M'M}]]} \quad (\text{S140})$$

$$= \frac{\lambda}{d_M d_A^2 \text{Tr}[\Gamma_A \{\mathcal{E}_{ML \rightarrow A} \circ \mathcal{D}_{B \rightarrow \hat{M}L} \circ \mathcal{N}_{A \rightarrow B}\} [\Upsilon_{M'M}]]}, \quad (\text{S141})$$

where Eq. (S140) follows from Eq. (S134). Similarly, it follows from Eq. (S129) that

$$\mu \varepsilon \geq \mu p_{\text{error}}((\mathcal{E}_{ML \rightarrow A}, \mathcal{D}_{B \rightarrow \hat{M}L}), \mathcal{N}_{A \rightarrow B}) \quad (\text{S142})$$

$$\geq \frac{\mu \text{Tr}[(\mathbb{1}_{M'\hat{M}} - \Omega_{M'\hat{M}}) \Gamma_A \{\mathcal{E}_{ML \rightarrow A} \circ \mathcal{D}_{B \rightarrow \hat{M}L} \circ \mathcal{N}_{A \rightarrow B}\} [\Upsilon_{M'M}]]}{\text{Tr}[\Gamma_A \{\mathcal{E}_{ML \rightarrow A} \circ \mathcal{D}_{B \rightarrow \hat{M}L} \circ \mathcal{N}_{A \rightarrow B}\} [\Upsilon_{M'M}]]} \quad (\text{S143})$$

$$\geq \frac{\text{Tr}[(\mathbb{1}_{M'\hat{M}} - \Omega_{M'\hat{M}}) (\pi_{M'} \otimes \eta_{\hat{M}})]}{d_A^2 \text{Tr}[\Gamma_A \{\mathcal{E}_{ML \rightarrow A} \circ \mathcal{D}_{B \rightarrow \hat{M}L} \circ \mathcal{N}_{A \rightarrow B}\} [\Upsilon_{M'M}]]} \quad (\text{S144})$$

$$= \left(1 - \frac{1}{d_M}\right) \frac{1}{d_A^2 \text{Tr}[\Gamma_A \{\mathcal{E}_{ML \rightarrow A} \circ \mathcal{D}_{B \rightarrow \hat{M}L} \circ \mathcal{N}_{A \rightarrow B}\} [\Upsilon_{M'M}]]}, \quad (\text{S145})$$

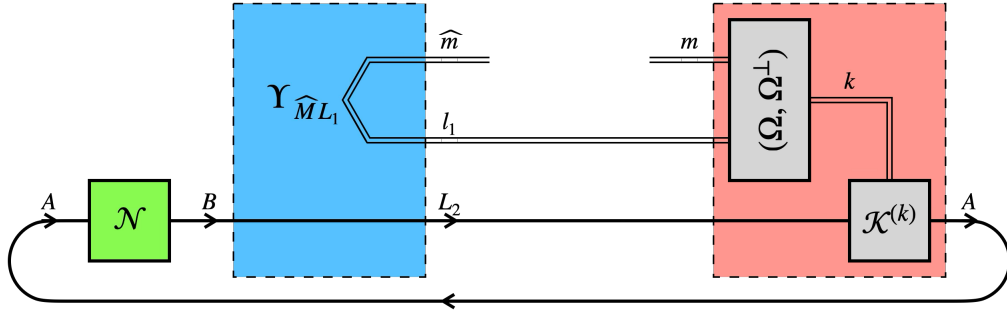


FIG. S7. The basic idea of amplified probabilistic classical teleportation is to first perform a classical version of probabilistic teleportation from the father to the daughter, and then engineer a causal loop involving the noisy P-CTC such that it maximally amplifies the renormalized probability of the teleportation being successful. The red and blue boxes encapsulate the father's encoding and the daughter's decoding, respectively.

where Eq. (S144) follows from Eq. (S137). Note that combining Eqs. (S141) and (S145) implies that

$$d_M \leq \frac{\varepsilon\lambda\mu}{1-\varepsilon} + 1, \quad (\text{S146})$$

which, due to  $d_M$  being an integer, further implies that

$$\log_2 d_M \leq \log_2 \left\lceil \frac{\varepsilon\lambda\mu}{1-\varepsilon} + 1 \right\rceil. \quad (\text{S147})$$

Supremizing the left-hand side over every  $d_M, d_L \in \mathbb{N}$  and every strategy  $(\mathcal{E}_{ML \rightarrow A}, \mathcal{D}_{B \rightarrow \hat{M}L})$  satisfying Eq. (S124), and infimizing the right-hand side over every  $\lambda, \mu \in \mathbb{R}$ ,  $\sigma_B \in \mathcal{D}_B$ , and  $\tau_B \in \text{aff}(\mathcal{D}_B)$  satisfying Eqs. (S130) and (S131), the desired statement thus follows from the definitions of the one-shot retrocausal classical capacity and the max- and Doeblin informations [Eqs. (S43), (S2), and (S3)].  $\square$

## B. Achievability

We now construct a strategy that saturates the converse bound derived in Proposition 3 by slightly adapting Strategy 1.

**Strategy 2** (Amplified probabilistic classical teleportation). The strategy operates as follows (see Fig. S7).

- **Encoding.** In the future, the father retrieves a classical memory  $L_1$  and a quantum memory  $L_2$  (both of which were left by Bella in the past) with  $L_1 \cong M$  and  $L_2 \cong B$ . He reads the data  $l_1$  in  $L_1$  and compares it with his message  $m$ . He sets  $k := 0$  if  $m = l_1$  and sets  $k := 1$  otherwise. Conditioned on  $k$ , he applies a channel  $\mathcal{K}_{L_2 \rightarrow A}^{(k)} \in \mathcal{C}_{L_2 \rightarrow A}$  to  $L_2$ . He then feeds  $A$  into the entrance of the noisy P-CTC.
- **Decoding.** In the past, the daughter receives  $B$  from the exit of the noisy P-CTC. She generates a symbol  $\hat{m}$  from the alphabet uniformly at random, writes a copy of  $l_1 := \hat{m}$  into  $L_1$ , renames  $B$  as  $L_2$ , and leaves both  $L_1$  and  $L_2$  for the father to retrieve in the future. She then declares that  $\hat{m}$  is her recovery of the message.

The intuition behind Strategy 2 is the same as that behind its quantum counterpart, Strategy 1. Instead of trying to code his message into  $A$  which enters the noisy P-CTC, the father intends to “teleport” the message probabilistically to the daughter via  $L_1$ , using shared randomness. Indeed, whenever the father sets  $k = 0$ , his message  $m$  has been found to be identical to the data  $l_1$  in  $L_1$ , which is in turn identical to the daughter's

recovery  $\hat{m}$ . This occurs with probability  $1/d_M$  without invoking the noisy P-CTC. The noisy P-CTC and the channels  $\mathcal{K}_{L_2 \rightarrow A}^{(0)}$  and  $\mathcal{K}_{L_2 \rightarrow A}^{(1)}$  are then utilized in the same manner as in Strategy 1, to maximally amplify the renormalized probability of the event  $k = 0$ .

We now formalize this analysis and show that Strategy 2 can indeed achieve the maximum possible efficiency allowed by the converse bound established in Proposition 3.

**Proposition 4** (Achievability of retrocausal classical capacity). *For a completely positive map  $\mathcal{N}_{A \rightarrow B} \in \mathcal{L}_{A \rightarrow B}$  and a real number  $\varepsilon \in (0, 1)$ ,*

$$C_{\text{retro}}^\varepsilon(\mathcal{N}) \geq \log_2 \left[ \frac{\varepsilon}{1 - \varepsilon} 2^{I_{\max}(\mathcal{N}) + I_{\text{doe}}(\mathcal{N})} + 1 \right]. \quad (\text{S148})$$

Furthermore, the efficiency on the right-hand side of Eq. (S148) is achievable by Strategy 2 for some channels  $\mathcal{K}_{L_2 \rightarrow A}^{(0)}$  and  $\mathcal{K}_{L_2 \rightarrow A}^{(1)}$ .

*Proof.* The encoding and decoding of Strategy 2 are represented by the following channels, respectively:

$$\mathcal{E}_{ML_1 L_2 \rightarrow A}[\cdot] := \text{Tr}_{ML_1} \left[ \Omega_{ML_1} \mathcal{K}_{L_2 \rightarrow A}^{(0)}[\cdot] \right] + \text{Tr}_{ML_1} \left[ (\mathbb{1}_{ML_1} - \Omega_{ML_1}) \mathcal{K}_{L_2 \rightarrow A}^{(1)}[\cdot] \right], \quad (\text{S149})$$

$$\mathcal{D}_{B \rightarrow \hat{M} L_1 L_2}[\cdot] := \Upsilon_{\hat{M} L_1} \otimes \text{id}_{B \rightarrow L_2}[\cdot]. \quad (\text{S150})$$

Let  $m$  be a symbol in the alphabet. Then it follows from the definition of the loop supermap [Eq. (S13)] that

$$\begin{aligned} \Gamma_A \left\{ \mathcal{E}_{ML_1 L_2 \rightarrow A} \circ \mathcal{D}_{B \rightarrow \hat{M} L_1 L_2} \circ \mathcal{N}_{A \rightarrow B} \right\} [|m\rangle\langle m|_M] \\ = \text{Tr}_{AA'} \left[ \Phi_{AA'} \left( \mathcal{E}_{ML_1 L_2 \rightarrow A} \circ \mathcal{D}_{B \rightarrow \hat{M} L_1 L_2} \circ \mathcal{N}_{A \rightarrow B} \right) [|m\rangle\langle m|_M \otimes \Phi_{AA'}] \right] \end{aligned} \quad (\text{S151})$$

$$\begin{aligned} = \text{Tr}_{ML_1 AA'} \left[ (\Omega_{ML_1} \otimes \Phi_{AA'}) \left( \mathcal{K}_{B \rightarrow A}^{(0)} \circ \mathcal{N}_{A \rightarrow B} \right) [|m\rangle\langle m|_M \otimes \Upsilon_{\hat{M} L_1} \otimes \Phi_{AA'}] \right] \\ + \text{Tr}_{ML_1 AA'} \left[ ((\mathbb{1}_{ML_1} - \Omega_{ML_1}) \otimes \Phi_{AA'}) \left( \mathcal{K}_{B \rightarrow A}^{(1)} \circ \mathcal{N}_{A \rightarrow B} \right) [|m\rangle\langle m|_M \otimes \Upsilon_{\hat{M} L_1} \otimes \Phi_{AA'}] \right] \end{aligned} \quad (\text{S152})$$

$$\begin{aligned} = \frac{1}{d_M} \text{Tr} \left[ \Phi_{AA'} \left( \mathcal{K}_{B \rightarrow A}^{(0)} \circ \mathcal{N}_{A \rightarrow B} \right) [\Phi_{AA'}] \right] |m\rangle\langle m|_{\hat{M}} \\ + \frac{1}{d_M} \text{Tr} \left[ \Phi_{AA'} \left( \mathcal{K}_{B \rightarrow A}^{(1)} \circ \mathcal{N}_{A \rightarrow B} \right) [\Phi_{AA'}] \right] (\mathbb{1}_{\hat{M}} - |m\rangle\langle m|_{\hat{M}}). \end{aligned} \quad (\text{S153})$$

It follows from Eq. (S41) that

$$\begin{aligned} 1 - \Pr \left\{ m \middle| m; \left( \mathcal{E}_{ML_1 L_2 \rightarrow A}, \mathcal{D}_{B \rightarrow \hat{M} L_1 L_2} \right), \mathcal{N}_{A \rightarrow B} \right\} \\ = 1 - \frac{\text{Tr} \left[ |m\rangle\langle m|_{\hat{M}} \Gamma_A \left\{ \mathcal{E}_{ML_1 L_2 \rightarrow A} \circ \mathcal{D}_{B \rightarrow \hat{M} L_1 L_2} \circ \mathcal{N}_{A \rightarrow B} \right\} [|m\rangle\langle m|_M] \right]}{\text{Tr} \left[ \Gamma_A \left\{ \mathcal{E}_{ML_1 L_2 \rightarrow A} \circ \mathcal{D}_{B \rightarrow \hat{M} L_1 L_2} \circ \mathcal{N}_{A \rightarrow B} \right\} [|m\rangle\langle m|_M] \right]} \end{aligned} \quad (\text{S154})$$

$$\leq 1 - \frac{\frac{1}{d_M} \text{Tr} \left[ \Phi_{AA'} \left( \mathcal{K}_{B \rightarrow A}^{(0)} \circ \mathcal{N}_{A \rightarrow B} \right) [\Phi_{AA'}] \right]}{\frac{1}{d_M} \text{Tr} \left[ \Phi_{AA'} \left( \mathcal{K}_{B \rightarrow A}^{(0)} \circ \mathcal{N}_{A \rightarrow B} \right) [\Phi_{AA'}] \right] + \frac{d_M - 1}{d_M} \text{Tr} \left[ \Phi_{AA'} \left( \mathcal{K}_{B \rightarrow A}^{(1)} \circ \mathcal{N}_{A \rightarrow B} \right) [\Phi_{AA'}] \right]} \quad (\text{S155})$$

$$= \left( \frac{\text{Tr} \left[ \Phi_{AA'} \left( \mathcal{K}_{B \rightarrow A}^{(0)} \circ \mathcal{N}_{A \rightarrow B} \right) [\Phi_{AA'}] \right]}{(d_M - 1) \text{Tr} \left[ \Phi_{AA'} \left( \mathcal{K}_{B \rightarrow A}^{(1)} \circ \mathcal{N}_{A \rightarrow B} \right) [\Phi_{AA'}] \right] + 1} \right)^{-1}, \quad (\text{S156})$$

where Eq. (S155) follows from Eq. (S153). Since this holds for every symbol  $m$ , it follows from Eq. (S42) that, as long as

$$\left( \frac{\text{Tr}[\Phi_{AA'} (\mathcal{K}_{B \rightarrow A}^{(0)} \circ \mathcal{N}_{A \rightarrow B}) [\Phi_{AA'}]]}{(d_M - 1) \text{Tr}[\Phi_{AA'} (\mathcal{K}_{B \rightarrow A}^{(1)} \circ \mathcal{N}_{A \rightarrow B}) [\Phi_{AA'}]]} + 1 \right)^{-1} \leq \varepsilon, \quad (\text{S157})$$

we have that

$$p_{\text{error}}\left(\left(\mathcal{E}_{ML_1 L_2 \rightarrow A}, \mathcal{D}_{B \rightarrow \hat{M} L_1 L_2}\right), \mathcal{N}_{A \rightarrow B}\right) = 1 - \min_m \Pr\left\{m \mid m; \left(\mathcal{E}_{ML_1 L_2 \rightarrow A}, \mathcal{D}_{B \rightarrow \hat{M} L_1 L_2}\right), \mathcal{N}_{A \rightarrow B}\right\} \quad (\text{S158})$$

$$\leq \varepsilon. \quad (\text{S159})$$

Since Eq. (S157) is equivalent to

$$d_M \leq \frac{\varepsilon \text{Tr}[\Phi_{AA'} (\mathcal{K}_{B \rightarrow A}^{(0)} \circ \mathcal{N}_{A \rightarrow B}) [\Phi_{AA'}]]}{(1 - \varepsilon) \text{Tr}[\Phi_{AA'} (\mathcal{K}_{B \rightarrow A}^{(1)} \circ \mathcal{N}_{A \rightarrow B}) [\Phi_{AA'}]]} + 1, \quad (\text{S160})$$

by choosing

$$\mathcal{K}_{B \rightarrow A}^{(0)} := \arg \max_{\mathcal{K}_{B \rightarrow A} \in \mathcal{C}_{B \rightarrow A}} \text{Tr}[\Phi_{AA'} (\mathcal{K}_{B \rightarrow A} \circ \mathcal{N}_{A \rightarrow B}) [\Phi_{AA'}]], \quad (\text{S161})$$

$$\mathcal{K}_{B \rightarrow A}^{(1)} := \arg \min_{\mathcal{K}_{B \rightarrow A} \in \mathcal{C}_{B \rightarrow A}} \text{Tr}[\Phi_{AA'} (\mathcal{K}_{B \rightarrow A} \circ \mathcal{N}_{A \rightarrow B}) [\Phi_{AA'}]], \quad (\text{S162})$$

which maximize the right-hand side of Eq. (S160), the following efficiency can be achieved within error  $\varepsilon$  (i.e., subject to Eq. (S159)) and is thus a lower bound on the one-shot retrocausal classical capacity [see Eq. (S43)]:

$$\log_2 d_M = \log_2 \left[ \frac{\varepsilon \text{Tr}[\Phi_{AA'} (\mathcal{K}_{B \rightarrow A}^{(0)} \circ \mathcal{N}_{A \rightarrow B}) [\Phi_{AA'}]]}{(1 - \varepsilon) \text{Tr}[\Phi_{AA'} (\mathcal{K}_{B \rightarrow A}^{(1)} \circ \mathcal{N}_{A \rightarrow B}) [\Phi_{AA'}]]} + 1 \right] \quad (\text{S163})$$

$$= \log_2 \left[ \frac{\varepsilon \max_{\mathcal{K}_{B \rightarrow A} \in \mathcal{C}_{B \rightarrow A}} \text{Tr}[\Phi_{AA'} (\mathcal{K}_{B \rightarrow A} \circ \mathcal{N}_{A \rightarrow B}) [\Phi_{AA'}]]}{(1 - \varepsilon) \min_{\mathcal{K}_{B \rightarrow A} \in \mathcal{C}_{B \rightarrow A}} \text{Tr}[\Phi_{AA'} (\mathcal{K}_{B \rightarrow A} \circ \mathcal{N}_{A \rightarrow B}) [\Phi_{AA'}]]} + 1 \right] \quad (\text{S164})$$

$$= \log_2 \left[ \frac{\varepsilon}{1 - \varepsilon} 2^{I_{\text{max}}(\mathcal{N}) + I_{\text{doc}}(\mathcal{N})} + 1 \right], \quad (\text{S165})$$

where Eq. (S165) follows from the interpretation of the max- and Doeblin informations in terms of the maximum and minimum achievable singlet fractions, respectively (Lemma 1), implying the desired statement.  $\square$

Combining Propositions 3 and 4 completes the proof of Theorem 4.

### C. Asymptotics

With the one-shot retrocausal classical capacity determined, the asymptotic retrocausal classical capacity immediately follows.

**Theorem 5** (Asymptotic retrocausal classical capacity). *The asymptotic retrocausal classical capacity is equal to the sum of the max- and regularized Doeblin informations: for a completely positive map  $\mathcal{N}_{A \rightarrow B} \in \mathcal{L}_{A \rightarrow B}$  and a real number  $\varepsilon \in (0, 1)$ ,*

$$C_{\text{retro}}(\mathcal{N}) = \lim_{n \rightarrow \infty} \frac{1}{n} Q_{\text{retro}}^\varepsilon(\mathcal{N}^{\otimes n}) = I_{\text{max}}(\mathcal{N}) + I_{\text{doe}}^\infty(\mathcal{N}), \quad (\text{S166})$$

where the regularized Doeblin information is defined in Eq. (S9).

*Proof.* Since the max- and Doeblin informations are nonnegative, it follows from Theorem 4 that

$$C_{\text{retro}}^\varepsilon(\mathcal{N}) = \log_2 \left[ \frac{\varepsilon}{1 - \varepsilon} 2^{I_{\text{max}}(\mathcal{N}) + I_{\text{doe}}(\mathcal{N})} + 1 \right] \quad (\text{S167})$$

$$\leq \log_2 \left( \frac{\varepsilon}{1 - \varepsilon} 2^{I_{\text{max}}(\mathcal{N}) + I_{\text{doe}}(\mathcal{N})} + 1 \right) \quad (\text{S168})$$

$$\leq \log_2 \left( \left( \frac{\varepsilon}{1 - \varepsilon} + 1 \right) 2^{I_{\text{max}}(\mathcal{N}) + I_{\text{doe}}(\mathcal{N})} \right) \quad (\text{S169})$$

$$= I_{\text{max}}(\mathcal{N}) + I_{\text{doe}}(\mathcal{N}) + \log_2 \frac{1}{1 - \varepsilon}. \quad (\text{S170})$$

We also have that

$$C_{\text{retro}}^\varepsilon(\mathcal{N}) = \log_2 \left[ \frac{\varepsilon}{1 - \varepsilon} 2^{I_{\text{max}}(\mathcal{N}) + I_{\text{doe}}(\mathcal{N})} + 1 \right] \quad (\text{S171})$$

$$\geq \log_2 \left( \frac{\varepsilon}{1 - \varepsilon} 2^{I_{\text{max}}(\mathcal{N}) + I_{\text{doe}}(\mathcal{N})} \right) \quad (\text{S172})$$

$$= I_{\text{max}}(\mathcal{N}) + I_{\text{doe}}(\mathcal{N}) + \log_2 \frac{\varepsilon}{1 - \varepsilon}. \quad (\text{S173})$$

Combining Eqs. (S170) and (S173), we have that, for every  $n \in \mathbb{N}$ ,

$$\frac{1}{n} \left( I_{\text{max}}(\mathcal{N}^{\otimes n}) + I_{\text{doe}}(\mathcal{N}^{\otimes n}) + \log_2 \frac{1}{1 - \varepsilon} \right) \leq \frac{1}{n} C_{\text{retro}}^\varepsilon(\mathcal{N}^{\otimes n}) \quad (\text{S174})$$

$$\leq \frac{1}{n} \left( I_{\text{max}}(\mathcal{N}^{\otimes n}) + I_{\text{doe}}(\mathcal{N}^{\otimes n}) + \log_2 \frac{\varepsilon}{1 - \varepsilon} \right). \quad (\text{S175})$$

Taking the limit  $n \rightarrow \infty$ , it follows from the additivity of the max-information [Eq. (S8)] and the definition of the regularized Doeblin information [Eq. (S9)] that

$$\lim_{n \rightarrow \infty} \frac{1}{n} C_{\text{retro}}^\varepsilon(\mathcal{N}^{\otimes n}) = I_{\text{max}}(\mathcal{N}) + I_{\text{doe}}^\infty(\mathcal{N}), \quad (\text{S176})$$

from which and the definition of the asymptotic retrocausal classical capacity [Eq. (S44)] the desired statement follows.  $\square$

## D. Error exponent and strong-converse exponent

We now show that the error exponent and the strong-converse exponent of retrocausal classical communication can also be determined based on the one-shot retrocausal classical capacity.

**Theorem 6** (Error exponent and strong-converse exponent of retrocausal classical communication). *For a completely positive map  $\mathcal{N}_{A \rightarrow B} \in \mathcal{L}_{A \rightarrow B}$  and a real number  $r \in [0, \infty)$ ,*

$$E_{\text{retro,C}}^r(\mathcal{N}) = \max \{0, I_{\text{max}}(\mathcal{N}) + I_{\text{doe}}^\infty(\mathcal{N}) - r\}, \quad (\text{S177})$$

$$S_{\text{retro,C}}^r(\mathcal{N}) = \max \{0, r - I_{\text{max}}(\mathcal{N}) - I_{\text{doe}}^\infty(\mathcal{N})\}. \quad (\text{S178})$$

*Proof.* Let  $(\mathcal{E}_{ML \rightarrow A}, \mathcal{D}_{B \rightarrow \hat{M}L})$  be a strategy. Let  $\lambda, \mu \in \mathbb{R}$  be two real numbers,  $\sigma_B \in \mathcal{D}_B$  a state, and  $\tau_B \in \text{aff}(\mathcal{D}_B)$  a unit-trace Hermitian operator satisfying Eqs. (S130) and (S131). Comparing Eqs. (S138) with (S141) and Eqs. (S142) with (S145), we have that

$$1 - p_{\text{error}}((\mathcal{E}_{ML \rightarrow A}, \mathcal{D}_{B \rightarrow \hat{M}L}), \mathcal{N}_{A \rightarrow B}) \leq \frac{\lambda}{d_M d_A^2 \text{Tr}[\Gamma_A \{\mathcal{E}_{ML \rightarrow A} \circ \mathcal{D}_{B \rightarrow \hat{M}L} \circ \mathcal{N}_{A \rightarrow B}\} [\Upsilon_{M'M}]]}, \quad (\text{S179})$$

$$\mu p_{\text{error}}((\mathcal{E}_{ML \rightarrow A}, \mathcal{D}_{B \rightarrow \hat{M}L}), \mathcal{N}_{A \rightarrow B}) \geq \left(1 - \frac{1}{d_M}\right) \frac{1}{d_A^2 \text{Tr}[\Gamma_A \{\mathcal{E}_{ML \rightarrow A} \circ \mathcal{D}_{B \rightarrow \hat{M}L} \circ \mathcal{N}_{A \rightarrow B}\} [\Upsilon_{M'M}]]}. \quad (\text{S180})$$

Note that combining Eqs. (S179) and (S180) implies that

$$p_{\text{error}}((\mathcal{E}_{ML \rightarrow A}, \mathcal{D}_{B \rightarrow \hat{M}L}), \mathcal{N}_{A \rightarrow B}) \geq \frac{d_M - 1}{\lambda \mu + d_M - 1}. \quad (\text{S181})$$

For a fixed  $d_M$ , infimizing the left-hand side over every  $d_L \in \mathbb{N}$  and every strategy  $(\mathcal{E}_{ML \rightarrow A}, \mathcal{D}_{B \rightarrow \hat{M}L})$ , and supremizing the right-hand side over every  $\lambda, \mu \in \mathbb{R}$ ,  $\sigma_B \in \mathcal{D}_B$ , and  $\tau_B \in \text{aff}(\mathcal{D}_B)$  satisfying Eqs. (S130) and (S131), we have that

$$\inf_{\substack{d_L \in \mathbb{N}, \\ \mathcal{E}_{ML \rightarrow A} \in \mathcal{C}_{ML \rightarrow A}, \\ \mathcal{D}_{B \rightarrow \hat{M}L} \in \mathcal{C}_{B \rightarrow \hat{M}L}}} p_{\text{error}}((\mathcal{E}_{ML \rightarrow A}, \mathcal{D}_{B \rightarrow \hat{M}L}), \mathcal{N}_{A \rightarrow B}) \geq \frac{d_M - 1}{2^{I_{\max}(\mathcal{N}) + I_{\text{doe}}(\mathcal{N})} + d_M - 1}. \quad (\text{S182})$$

Now consider the strategy  $(\mathcal{E}_{ML_1 L_2 \rightarrow A}, \mathcal{D}_{B \rightarrow \hat{M}L_1 L_2})$  defined in Eqs. (S149) and (S150). It follows from Eq. (S156) that, for every symbol  $m$  in the alphabet,

$$1 - \Pr\{m|m; (\mathcal{E}_{ML_1 L_2 \rightarrow A}, \mathcal{D}_{B \rightarrow \hat{M}L_1 L_2}), \mathcal{N}_{A \rightarrow B}\} \leq \left( \frac{\text{Tr}[\Phi_{AA'} (\mathcal{K}_{B \rightarrow A}^{(0)} \circ \mathcal{N}_{A \rightarrow B}) [\Phi_{AA'}]]}{(d_M^2 - 1) \text{Tr}[\Phi_{AA'} (\mathcal{K}_{B \rightarrow A}^{(1)} \circ \mathcal{N}_{A \rightarrow B}) [\Phi_{AA'}]]} + 1 \right)^{-1}. \quad (\text{S183})$$

Supremizing the left-hand side over every  $d_R \in \mathbb{N}$  and every symbol  $m$  in the alphabet, and choosing the channels  $\mathcal{K}_{B \rightarrow A}^{(0)}$  and  $\mathcal{K}_{B \rightarrow A}^{(1)}$  according to Eqs. (S161) and (S162), it follows from Eq. (S42) that

$$\begin{aligned} & p_{\text{error}}((\mathcal{E}_{ML_1 L_2 \rightarrow A}, \mathcal{D}_{B \rightarrow \hat{M}L_1 L_2}), \mathcal{N}_{A \rightarrow B}) \\ &= \left( \frac{\text{Tr}[\Phi_{AA'} (\mathcal{K}_{B \rightarrow A}^{(0)} \circ \mathcal{N}_{A \rightarrow B}) [\Phi_{AA'}]]}{(d_M - 1) \text{Tr}[\Phi_{AA'} (\mathcal{K}_{B \rightarrow A}^{(1)} \circ \mathcal{N}_{A \rightarrow B}) [\Phi_{AA'}]]} + 1 \right)^{-1} \end{aligned} \quad (\text{S184})$$

$$= \left( \frac{\max_{\mathcal{K}_{B \rightarrow A} \in \mathcal{C}_{B \rightarrow A}} \text{Tr}[\Phi_{AA'} (\mathcal{K}_{B \rightarrow A} \circ \mathcal{N}_{A \rightarrow B}) [\Phi_{AA'}]]}{(d_M - 1) \min_{\mathcal{K}_{B \rightarrow A} \in \mathcal{C}_{B \rightarrow A}} \text{Tr}[\Phi_{AA'} (\mathcal{K}_{B \rightarrow A} \circ \mathcal{N}_{A \rightarrow B}) [\Phi_{AA'}]]} + 1 \right)^{-1} \quad (\text{S185})$$

$$= \left( \frac{1}{d_M - 1} 2^{I_{\max}(\mathcal{N}) + I_{\text{doe}}(\mathcal{N})} + 1 \right)^{-1} \quad (\text{S186})$$

$$= \frac{d_M - 1}{2^{I_{\max}(\mathcal{N}) + I_{\text{doe}}(\mathcal{N})} + d_M - 1}, \quad (\text{S187})$$

where Eq. (S186) follows from the interpretation of the max- and Doeblin informations in terms of the maximum and minimum achievable singlet fractions, respectively (Lemma 1). Combining Eqs. (S182) and

(S187), we have that

$$\inf_{\substack{d_L \in \mathbb{N}, \\ \mathcal{E}_{ML \rightarrow A} \in \mathcal{C}_{ML \rightarrow A}, \\ \mathcal{D}_{B \rightarrow \widehat{ML}} \in \mathcal{C}_{B \rightarrow \widehat{ML}}} p_{\text{error}}((\mathcal{E}_{ML \rightarrow A}, \mathcal{D}_{B \rightarrow \widehat{ML}}), \mathcal{N}_{A \rightarrow B}) = \frac{d_M - 1}{2^{I_{\max}(\mathcal{N}) + I_{\text{doe}}(\mathcal{N})} + d_M - 1}. \quad (\text{S188})$$

Then it follows from the definition of the error exponent of retrocausal classical communication [Eq. (S45)] that

$$E_{\text{retro,C}}^r(\mathcal{N}) = \lim_{n \rightarrow \infty} \sup_{d_M \in \mathbb{N}} \left\{ -\frac{1}{n} \log_2 \frac{d_M - 1}{2^{I_{\max}(\mathcal{N}^{\otimes n}) + I_{\text{doe}}(\mathcal{N}^{\otimes n})} + d_M - 1} : d_M \geq 2^{nr} \right\} \quad (\text{S189})$$

$$= \lim_{n \rightarrow \infty} -\frac{1}{n} \log_2 \frac{2^{nr} - 1}{2^{I_{\max}(\mathcal{N}^{\otimes n}) + I_{\text{doe}}(\mathcal{N}^{\otimes n})} + 2^{nr} - 1} \quad (\text{S190})$$

$$= \max \{0, I_{\max}(\mathcal{N}) + I_{\text{doe}}^\infty(\mathcal{N}) - r\}. \quad (\text{S191})$$

where Eq. (S191) follows from the additivity of the max-information [Eq. (S8)] and the definition of the regularized Doeblin information [Eq. (S9)]. Likewise, it follows from the definition of the strong-converse exponent of retrocausal classical communication [Eq. (S46)] that

$$S_{\text{retro,C}}^r(\mathcal{N}) = \lim_{n \rightarrow \infty} \inf_{d_M \in \mathbb{N}} \left\{ -\frac{1}{n} \log_2 \left( 1 - \frac{d_M - 1}{2^{I_{\max}(\mathcal{N}^{\otimes n}) + I_{\text{doe}}(\mathcal{N}^{\otimes n})} + d_M - 1} \right) : d_M \geq 2^{nr} \right\} \quad (\text{S192})$$

$$= \lim_{n \rightarrow \infty} -\frac{1}{n} \log_2 \frac{2^{I_{\max}(\mathcal{N}^{\otimes n}) + I_{\text{doe}}(\mathcal{N}^{\otimes n})}}{2^{I_{\max}(\mathcal{N}^{\otimes n}) + I_{\text{doe}}(\mathcal{N}^{\otimes n})} + 2^{nr} - 1} \quad (\text{S193})$$

$$= \max \{0, r - I_{\max}(\mathcal{N}) - I_{\text{doe}}^\infty(\mathcal{N})\}, \quad (\text{S194})$$

which completes the proof of the desired statement.  $\square$

## VII. EXAMPLES

In this section, we apply our results to some special classes of channels, and we numerically compare the retrocausal capacity of these channels to their entanglement-assisted and unassisted capacities.

### A. Depolarizing channels

A depolarizing channel  $\mathcal{D}_{A \rightarrow A}^p \in \mathcal{C}_{A \rightarrow A}$  with parameter  $p \in [0, 1]$  is defined as

$$\mathcal{D}_{A \rightarrow A}^p \equiv (1 - p) \text{id}_{A \rightarrow A} + p \mathcal{R}_{A \rightarrow A}^\pi. \quad (\text{S195})$$

Since  $\mathcal{D}_{A \rightarrow A}^p$  is a mictodiactive covariant channel, its regularized Doeblin information is equal to its Doeblin information [61, Theorem 4], both of which are given by

$$I_{\text{doe}}^\infty(\mathcal{D}^p) = I_{\text{doe}}(\mathcal{D}^p) \quad (\text{S196})$$

$$= \inf_{\lambda \in \mathbb{R}} \left\{ \log_2 \lambda : \pi_{A'} \otimes \pi_A \leq \lambda ((1 - p) \Phi_{A'A} + p \pi_{A'} \otimes \pi_A) \right\} \quad (\text{S197})$$

$$= -\log_2 p, \quad (\text{S198})$$

where Eq. (S197) follows from Ref. [61, Lemma 3] and Eq. (S195). The max-information of  $\mathcal{D}_{A \rightarrow A}^p$  is equal to [71, Eq. (66)]

$$I_{\max}(\mathcal{D}^p) = \log_2 \left( (1 - p) d_A^2 + p \right). \quad (\text{S199})$$

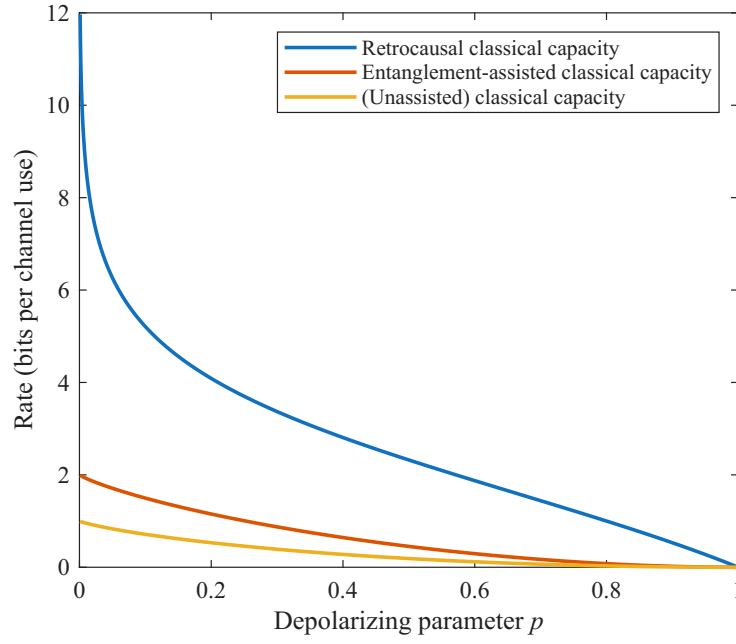


FIG. S8. For qubit depolarizing channels, the asymptotic retrocausal classical capacity, the asymptotic entanglement-assisted classical capacity, and the asymptotic (unassisted) classical capacity are compared numerically as functions of the depolarizing parameter  $p \in [0, 1]$ . The asymptotic retrocausal classical capacity approaches  $\infty$  as  $p \searrow 0$ .

By Theorem 5, the asymptotic retrocausal classical capacity of  $\mathcal{D}_{A \rightarrow A}^p$  is thus given by

$$C_{\text{retro}}(\mathcal{D}^p) = I_{\text{max}}(\mathcal{D}^p) + I_{\text{doe}}^{\infty}(\mathcal{D}^p) \quad (\text{S200})$$

$$= \log_2 \left( (1-p) d_A^2 + p \right) - \log_2 p. \quad (\text{S201})$$

The asymptotic entanglement-assisted classical capacity of  $\mathcal{D}_{A \rightarrow A}^p$  is equal to [37, Eq. (3)]

$$C_{\text{EA}}(\mathcal{D}^p) = 2 \log_2 d_A - h_2 \left( \frac{p(d_A^2 - 1)}{d_A^2} \right) - \frac{p(d_A^2 - 1)}{d_A^2} \log_2 (d_A^2 - 1), \quad (\text{S202})$$

where  $h_2(x) \equiv -x \log_2 x - (1-x) \log_2 (1-x)$  is the binary entropy. The asymptotic (unassisted) classical capacity of  $\mathcal{D}_{A \rightarrow A}^p$  is equal to [94, Theorem 1]

$$C(\mathcal{D}^p) = \log_2 d_A - h_2 \left( \frac{p d_A}{d_A + 1} \right) - \frac{p d_A}{d_A + 1} \log_2 (d_A - 1). \quad (\text{S203})$$

See Fig. S8 for a numerical comparison between the asymptotic retrocausal classical capacity, the asymptotic entanglement-assisted classical capacity, and the asymptotic (unassisted) classical capacity of depolarizing channels for when  $d_A = 2$ .

## B. Erasure channels

An erasure channel  $\mathcal{E}_{A \rightarrow B}^p \in \mathcal{C}_{A \rightarrow B}$  with  $\mathbb{H}_B = \mathbb{H}_A \oplus \text{span}(|e\rangle)$  and parameter  $p \in [0, 1]$  is defined as

$$\mathcal{E}_{A \rightarrow B}^p \equiv (1-p) \text{id}_{A \rightarrow B} + p \mathcal{R}_{A \rightarrow B}^{|e\rangle\langle e|}, \quad (\text{S204})$$

where  $\overline{\text{id}}_{A \rightarrow B}[\cdot] \equiv \sum_{i,j=0}^{d_A-1} \text{Tr}[|j\rangle\langle i|_A (\cdot)_A] |i\rangle\langle j|_B$  with  $\{|i\rangle\}_{i=0}^{d_A-1}$  an orthonormal basis of  $\mathbb{H}_A$ . It follows from Eq. (S11) that

$$I_{\pm}(\mathcal{E}^p) = \inf_{\substack{\lambda \in \mathbb{R}_+ \\ \tau_B \in \text{aff}(\mathcal{D}_B)}} \left\{ \log_2 \lambda : -\lambda \Phi_{A'B}^{\mathcal{E}^p} \leq \pi_{A'} \otimes \tau_B \leq \lambda \Phi_{A'B}^{\mathcal{E}^p} \right\} \quad (\text{S205})$$

$$\leq \inf_{\lambda \in \mathbb{R}_+} \left\{ \log_2 \lambda : -\lambda \Phi_{A'B}^{\mathcal{E}^p} \leq \pi_{A'} \otimes |e\rangle\langle e|_B \leq \lambda \Phi_{A'B}^{\mathcal{E}^p} \right\} \quad (\text{S206})$$

$$= \inf_{\lambda \in \mathbb{R}_+} \left\{ \log_2 \lambda : \pi_{A'} \otimes |e\rangle\langle e|_B \leq \lambda \left( (1-p) \overline{\Phi}_{A'B} + p \pi_{A'} \otimes |e\rangle\langle e|_B \right) \right\} \quad (\text{S207})$$

$$= -\log_2 p, \quad (\text{S208})$$

where  $\overline{\Phi}_{A'B} \equiv (1/d_A) \sum_{i,j=0}^{d_A-1} |i\rangle\langle j|_{A'} \otimes |i\rangle\langle j|_B$ . Now define the following channel:

$$\mathcal{P}_{B \rightarrow A}[\cdot] := \sum_{i,j=0}^{d_A-1} \text{Tr}[|j\rangle\langle i|_B (\cdot)_B] |(i+1) \bmod d_A\rangle\langle (j+1) \bmod d_A|_A + \text{Tr}[|e\rangle\langle e|_B (\cdot)_B] \pi_A. \quad (\text{S209})$$

It follows from the interpretation of the Doeblin information in terms of the minimum achievable singlet fraction (Lemma 1) that

$$I_{\text{doe}}(\mathcal{E}^p) = -2 \log_2 d_A - \log_2 \min_{\mathcal{K}_{B \rightarrow A} \in \mathcal{C}_{B \rightarrow A}} \text{Tr} \left[ \Phi_{AA'} \left( \mathcal{K}_{B \rightarrow A} \circ \mathcal{E}_{A \rightarrow B}^p \right) [\Phi_{AA'}] \right] \quad (\text{S210})$$

$$\geq -2 \log_2 d_A - \log_2 \text{Tr} \left[ \Phi_{AA'} \left( \mathcal{P}_{B \rightarrow A} \circ \mathcal{E}_{A \rightarrow B}^p \right) [\Phi_{AA'}] \right] \quad (\text{S211})$$

$$= -2 \log_2 d_A - \log_2 (p \text{Tr}[\Phi_{AA'} (\pi_A \otimes \pi_{A'})]) \quad (\text{S212})$$

$$= -\log_2 p. \quad (\text{S213})$$

Combining Eqs. (S208) and (S213), it follows from Eq. (S12) that

$$-\log_2 p \leq I_{\text{doe}}(\mathcal{E}^p) \leq I_{\text{doe}}^{\infty}(\mathcal{E}^p) \leq I_{\pm}(\mathcal{E}^p) \leq -\log_2 p, \quad (\text{S214})$$

which implies that

$$I_{\text{doe}}(\mathcal{E}^p) = I_{\text{doe}}^{\infty}(\mathcal{E}^p) = I_{\pm}(\mathcal{E}^p) = -\log_2 p. \quad (\text{S215})$$

The max-information of  $\mathcal{E}_{A \rightarrow B}^p$  is equal to [71, Eq. (75)]

$$I_{\text{max}}(\mathcal{E}^p) = \log_2 \left( (1-p) d_A^2 + p \right). \quad (\text{S216})$$

By Theorem 5, the asymptotic retrocausal classical capacity of  $\mathcal{E}_{A \rightarrow B}^p$  is thus given by

$$C_{\text{retro}}(\mathcal{E}^p) = I_{\text{max}}(\mathcal{E}^p) + I_{\text{doe}}^{\infty}(\mathcal{E}^p) \quad (\text{S217})$$

$$= \log_2 \left( (1-p) d_A^2 + p \right) - \log_2 p. \quad (\text{S218})$$

The asymptotic entanglement-assisted classical capacity of  $\mathcal{E}_{A \rightarrow B}^p$  is equal to [37, p. 3083]

$$C_{\text{EA}}(\mathcal{E}^p) = 2(1-p) \log_2 d_A. \quad (\text{S219})$$

The asymptotic (unassisted) classical capacity of  $\mathcal{E}_{A \rightarrow B}^p$  is equal to [95]

$$C(\mathcal{E}^p) = (1-p) \log_2 d_A. \quad (\text{S220})$$

See Fig. S9 for a numerical comparison between the asymptotic retrocausal classical capacity, the asymptotic entanglement-assisted classical capacity, and the asymptotic (unassisted) classical capacity of erasure channels for when  $d_A = 2$ .

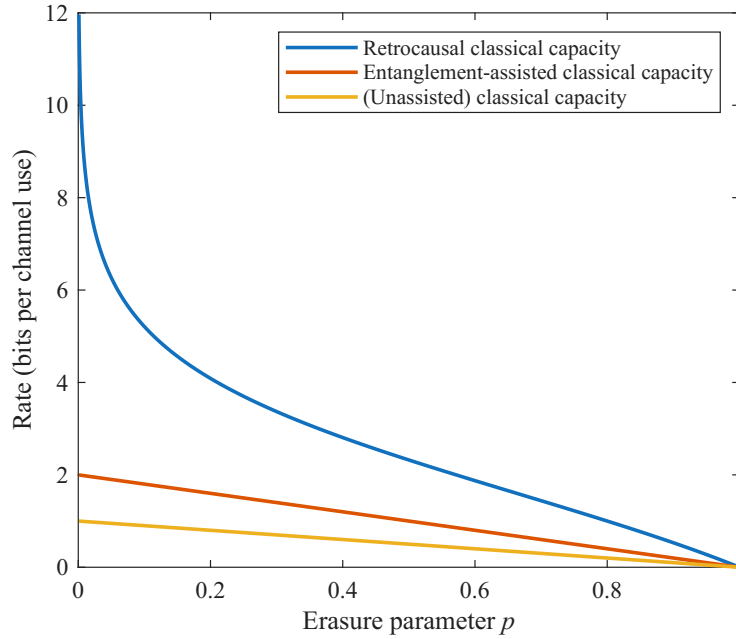


FIG. S9. For qubit erasure channels, the asymptotic retrocausal classical capacity, the asymptotic entanglement-assisted classical capacity, and the asymptotic (unassisted) classical capacity are compared numerically as functions of the erasure parameter  $p \in [0, 1]$ . The asymptotic retrocausal classical capacity approaches  $\infty$  as  $p \searrow 0$ .

### C. Amplitude-damping channels

An amplitude-damping channel  $\mathcal{A}_{A \rightarrow A}^\gamma \in \mathcal{C}_{A \rightarrow A}$  with  $d_A = 2$  and parameter  $\gamma \in [0, 1]$  is defined as

$$\mathcal{A}_{A \rightarrow A}^\gamma[\cdot] \equiv \left( |0\rangle\langle 0|_A + \sqrt{1-\gamma}|1\rangle\langle 1|_A \right) (\cdot)_A \left( |0\rangle\langle 0|_A + \sqrt{1-\gamma}|1\rangle\langle 1|_A \right) + \gamma|0\rangle\langle 1|_A (\cdot)_A |1\rangle\langle 0|_A. \quad (\text{S221})$$

The Doeblin information of  $\mathcal{A}_{A \rightarrow A}^\gamma$  is equal to [61, Lemma 9]

$$I_{\text{doe}}(\mathcal{A}^\gamma) = -2 \log_2 \left( 1 - \sqrt{1-\gamma} \right). \quad (\text{S222})$$

The max-information of  $\mathcal{A}_{A \rightarrow A}^\gamma$  is given by [71, Eq. (69)]

$$I_{\text{max}}(\mathcal{A}^\gamma) = \log_2 \left( 2 \left( 1 + \sqrt{1-\gamma} \right) - \gamma \right). \quad (\text{S223})$$

By Theorem 2, the asymptotic retrocausal quantum capacity of  $\mathcal{A}_{A \rightarrow A}^\gamma$  has the following lower bound:

$$Q_{\text{retro}}(\mathcal{A}^\gamma) = \frac{1}{2} \left( I_{\text{max}}(\mathcal{A}^\gamma) + I_{\text{doe}}^\infty(\mathcal{A}^\gamma) \right) \quad (\text{S224})$$

$$\geq \frac{1}{2} \left( I_{\text{max}}(\mathcal{A}^\gamma) + I_{\text{doe}}(\mathcal{A}^\gamma) \right) \quad (\text{S225})$$

$$= \frac{1}{2} \left( \log_2 \left( 2 \left( 1 + \sqrt{1-\gamma} \right) - \gamma \right) - 2 \log_2 \left( 1 - \sqrt{1-\gamma} \right) \right). \quad (\text{S226})$$

The asymptotic entanglement-assisted quantum capacity of  $\mathcal{A}_{A \rightarrow A}^\gamma$  is equal to [96, Eq. (38)]

$$Q_{\text{EA}}(\mathcal{A}^\gamma) = \max_{\lambda \in [0,1]} \{ h_2(\lambda) + h_2((1-\gamma)\lambda) - h_2(\gamma\lambda) \}. \quad (\text{S227})$$

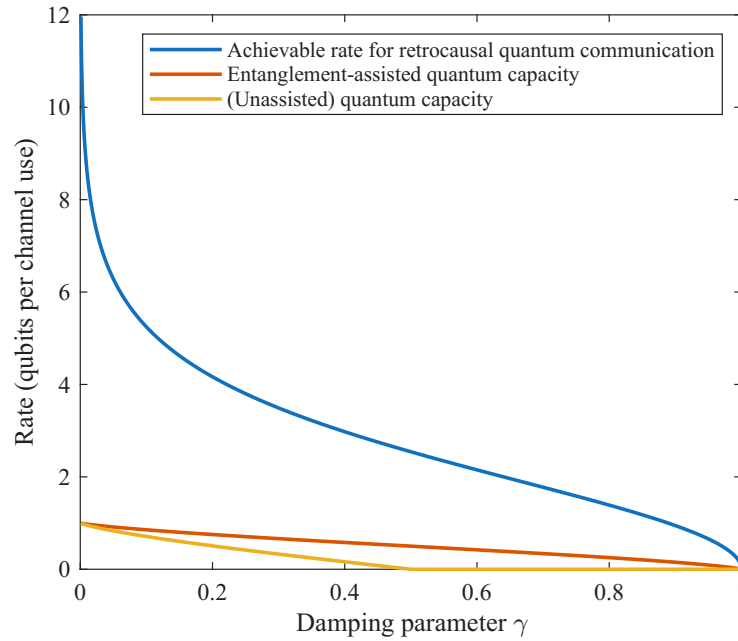


FIG. S10. For amplitude-damping channels, an achievable asymptotic rate for retrocausal quantum communication, the asymptotic entanglement-assisted quantum capacity, and the asymptotic (unassisted) quantum capacity are compared numerically as functions of the damping parameter  $\gamma \in [0, 1]$ . The asymptotic retrocausal quantum capacity approaches  $\infty$  as  $\gamma \searrow 0$ .

The asymptotic (unassisted) quantum capacity of  $\mathcal{A}_{A \rightarrow A}^\gamma$  is equal to [96, Eq. (36)]

$$Q(\mathcal{A}^\gamma) = \max_{\lambda \in [0,1]} \{h_2((1-\gamma)\lambda) - h_2(\gamma\lambda)\}. \quad (\text{S228})$$

See Fig. S10 for a numerical comparison between the achievable asymptotic rate for retrocausal quantum communication in Eq. (S226), the asymptotic entanglement-assisted quantum capacity, and the asymptotic (unassisted) quantum capacity of amplitude-damping channels.