Asymptotic Limits of Entanglement Transmission

Piotr Masajada, Aby Philip, and Alexander Streltsov*
Institute of Fundamental Technological Research, Polish Academy of Sciences, Pawińskiego 5B, 02-106 Warsaw, Poland

Reliable distribution of quantum entanglement over long distances is a central challenge in quantum information science, fundamentally limited by decoherence in noisy communication channels. In this work, we investigate the asymptotic limits of entanglement distribution across homogeneous linear repeater chains with arbitrary intermediate LOCC processing. We establish a strict dichotomy: the asymptotic preservation of entanglement over arbitrarily long distances is possible if and only if the underlying quantum channel admits a correctable subspace. For channels lacking such a subspace, we prove that the transmitted state converges exponentially fast to the set of separable states, rendering standard LOCC filtering insufficient. To counteract this exponential degradation, we analyze "networks" employing parallel channel uses per link. We derive a fundamental lower bound on the required number of parallel channel uses per elementary link, proving that for broad classes of channels without a correctable subspace, the number of parallel channels per link must scale at least logarithmically with the number of intermediate stations to sustain a non-zero amount of entanglement. This provides a code-independent bound for the number of physical links for the considered homogeneous one-way architecture. For depolarizing noise below the corresponding coding threshold, the logarithmic lower bound is achievable.

Introduction. Quantum entanglement [1] is a fundamental resource for a wide array of quantum information processing tasks, including quantum key distribution [2], quantum teleportation [3], and distributed quantum computing [4, 5]. A central challenge in realizing these technologies is the reliable distribution of entanglement between distant parties. In realistic scenarios, quantum states must be transmitted through noisy quantum channels, which inevitably interact with the environment and induce decoherence. This continuous degradation strictly limits the amount of entanglement that can be shared and maintained between remote parties.

In the standard entanglement distribution scenario [6–9] two distant parties, Alice and Bob, have access to a noisy quantum channel Λ. Alice first prepares a bipartite quantum state ρ in her laboratory, keeping one subsystem and using the channel to transmit the other subsystem to Bob. Because the transmission is imperfect, the channel acts nontrivially on the particle sent to Bob, while the particle retained by Alice remains unchanged. As a result, the initially prepared state ρ is transformed into the final bipartite state $\Lambda \otimes \mathbb{1}[\rho]$. Any entanglement initially present in ρ is generally degraded by the action of the channel, and the properties of Λ determine how much entanglement can ultimately be shared between the two parties.

If the noisy channel Λ is used sequentially, the resulting state shared by Alice and Bob is given by

$$(\Lambda \circ \cdots \circ \Lambda) \otimes \mathbb{1}[\rho]. \tag{1}$$

This setting is relevant, for example, in long-distance quantum communication through a chain of noisy transmission segments, where the quantum system effectively undergoes several consecutive applications of the same channel before reaching Bob. In general, the detrimental effects of noise accumulate with the number of channel uses, leading to a progressive degradation of the shared entanglement [10]. In many physically relevant situations, this entanglement eventually disappears altogether in the limit of infinitely many channel applications [11].

To counteract the accumulation of noise, one can introduce intermediate stations to process the transmitted systems between successive channel segments. This general principle underlies quantum repeaters, although different repeater architectures employ different physical resources and classical communication methods [12–18]. Here, we consider a homogeneous chain in which every elementary link is described by the same finite-dimensional channel Λ. Each transmission round consists of a trace-preserving LOCC operation $\mathcal{F}_i$ followed by an application of $\Lambda \otimes \mathbb{1}$. The operations may depend on all previous measurement outcomes; equivalently, the classical transcript required for such adaptivity may be included in the state, see Fig. 1. Denoting the overall transformation after n rounds by Λ_n, we have

$$\Lambda_n[\rho] = (\Lambda \otimes \mathbb{1}) \circ \mathcal{F}_n \circ \cdots \circ (\Lambda \otimes \mathbb{1}) \circ \mathcal{F}_1[\rho]. \tag{2}$$

For deterministic protocols, our entanglement measure of choice is the entanglement of formation [19, 20],

$$E_{\mathrm{f}}(\rho^{AB}) = \min_{\{p_i, \psi_i^{AB}\}} \sum_i p_i S\left(\mathrm{Tr}_B\, \psi_i^{AB}\right), \tag{3}$$

where the minimum is taken over all pure-state decompositions $\rho^{AB} = \sum_i p_i \psi_i^{AB}$ and $S(\rho^{AB})$ denotes von Neumann entropy. For two-qubit states, E_{f} is a strictly increasing function of the concurrence C,

$$E_{\mathrm{f}}(\rho) = h_2\left(\frac{1 + \sqrt{1 - C(\rho)^2}}{2}\right), \tag{4}$$

where h_2 is the binary entropy. For the definitions of concurrence see Supplementary Material. Thus, for two qubits, convergence of concurrence to zero is equivalent to convergence of the entanglement of formation to zero.

We say that a deterministic protocol asymptotically preserves entanglement if, for some initial state ρ^{AB} and some admissible sequence of intermediate operations,

$$\lim_{n\to\infty} E_{\mathrm{f}}\left(\Lambda_n[\rho^{AB}]\right) > 0. \tag{5}$$

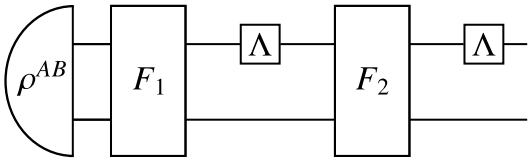

Figure 1. **Entanglement distribution through a noisy channel with intermediate repeater stations.** An initial bipartite state ρ^{AB} is shared between parties A and B. Subsystem A is transmitted through a sequence of n noisy channel uses Λ, which may be understood as propagation through n successive transmission segments with intermediate stations located between Alice and Bob. After each channel use, the parties, possibly assisted by these intermediate stations, may apply an LOCC filter F_i. The figure illustrates the resulting state in Eq. (2) for $n = 2$.

Without intermediate processing, the related problem was studied in the theory of asymptotically entanglement-saving channels [21]. Our setting additionally permits arbitrary adaptive operations between successive channel uses.

We subsequently consider another transmission architecture in which each occurrence of Λ is replaced by m parallel uses $\Lambda^{\otimes m}$. The intermediate operations may include decoding, quantum error correction, recovery, and re-encoding. Consequently, deterministic error-correction-based third-generation repeaters form a special case of our model [13]. Concrete realizations of this architecture include direct one-way protocols based on teleportation-based error correction, encoded Bell measurements, block quantum codes, and code concatenation [15, 22–27]. Our analysis does not assume any particular code, decoding procedure, or recovery scheme.

This problem is complementary to the literature on quantum network capacities, routing, and cut-based bounds [28–32]. These frameworks typically fix a network and optimize an asymptotic communication rate, allowing general topologies, adaptive protocols, and different routing strategies. Instead, we let the number of links n in a homogeneous chain grow and ask how the number of physical channels $m = m(n)$ assigned to every link must scale for the end-to-end entanglement to remain nonzero. Within this setting, we derive code-independent constraints on asymptotic entanglement distribution and determine when the resulting resource scaling is tight. Techniques like network capacities and min-cut bounds do not directly determine this finite-block scaling with the chain length.

Single-qubit channels. To begin, we investigate entanglement distribution through a single-qubit quantum channel Λ. We assume Alice holds a single-qubit memory that she aims to entangle with Bob's system across a sequence of n transmission segments, each linked via the quantum channel Λ. We recall that the Choi state of Λ is defined as [33]

$$\Gamma = (\Lambda \otimes \mathbb{1})(\phi^+), \tag{6}$$

where $|\phi^+\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$ is a Bell state.

We consider the scenario in which the intermediate filtering operations F_i are allowed to be stochastic local operations and classical communication (SLOCC) [34] rather than deterministic LOCC. This broader framework is important because probabilistic filters can, upon success, increase the entanglement of the post-selected state which is not possible with deterministic LOCC. The key question is thus whether this additional freedom provided by SLOCC can preserve a nonzero amount of entanglement after many sequential applications of the channel by conditioning on successful filtering events. In the following, we show that although SLOCC filters may temporarily improve the quality of the surviving states, this does not overcome the fundamental limitations of the channel: any such apparent entanglement preservation comes at the price of a success probability that necessarily decreases with the number of channel uses. To show this, we use concurrence whose definition we recall in the Supplementary Material. Note that going forward, $P(C \geq c, n)$ denotes the probability that the final state $\Lambda_n[\rho^{AB}]$ has concurrence at least c.

Theorem 1. *Let Λ be a single-qubit channel and consider an adaptive n-round protocol in which an SLOCC instrument is followed by $\Lambda \otimes \mathbb{1}$ in each round. The instrument in round i may depend on the previous outcomes $h_{i-1} = (x_1, \ldots, x_{i-1})$. If ρ_{H_n} denotes the random final state, then, for every $c > 0$,*

$$\Pr[C(\rho_{H_n}) \geq c] \leq \frac{q^n C(\rho_0)}{c} \leq \frac{q^n}{c}, \tag{7}$$

where $q = C[(\Lambda \otimes \mathbb{1})(\phi^+)]$ is the concurrence of Choi state of Λ.

The proof of the theorem can be found in the supplemental materials. If we demand that $P(C \geq c, n) = 1$ in Theorem 1, we recover the following statement.

Corollary 2. *For any single-qubit channel Λ the amount of concurrence decreases as*

$$C(\Lambda_n[\rho^{AB}]) \leq q^n. \tag{8}$$

We also note that this corollary has been proven using methods in [35]. We provide a proof in the supplemental materials. Concurrence is related with entanglement of formation by the following inequality:

$$E_f(\rho) \leq C(\rho). \tag{9}$$

Thus, if concurrence converges to 0, so does entanglement of formation. Since a two-qubit state has concurrence one only if it is a pure maximally entangled state [20], the corresponding Choi state has rank one, implying that Λ is unitary. Thus, only qubit unitary channels lead to preservation of entanglement in the limit of infinite filter–channel sequence.

This kind of problem was previously studied in two important aspects. First, in Ref. [10], authors studied under which conditions an infinite sequence of qubit filter–channel becomes an entanglement breaking channel. We demand stronger property, we demand that entanglement of formation after infinite filter–channel sequence does not converge to 0.

In Ref. [21], it was shown that, without intermediate operations, an entanglement of a state does not converge to 0 if and only if the channel is unitary. Corollary 2 strengthens these qualitative results by showing that, even when intermediate LOCC operations are allowed, every nonunitary qubit channel imposes an exponentially decaying upper bound on the surviving concurrence. Consequently, unitary channels are the only qubit channels for which such exponential decay can be avoided.

General channels. We now consider the more general setting in which subsystem A has an arbitrary finite dimension $d_A \geq 2$, and the channel Λ acts on $\mathcal{H}_A$. This extension goes beyond the single-qubit case and requires more refined tools, as two-qubit entanglement monotones such as concurrence are no longer directly applicable in higher dimensions. In the following, we say that a quantum channel Λ admits a correctable subspace $\mathcal{H}_c \subseteq \mathcal{H}$ if there exists a recovery channel $\mathcal{R}$ such that

$$\mathcal{R} \circ \Lambda[\psi] = \psi \tag{10}$$

for all $|\psi\rangle \in \mathcal{H}_c$ [36].

Having established the necessary groundwork, we now prove the following theorem.

Theorem 3. *For any quantum channel Λ, there exists a state ρ^{AB} such that*

$$\lim_{n\to\infty} E_{\mathrm{f}}(\Lambda_n[\rho^{AB}]) > 0, \tag{11}$$

if and only if Λ contains a correctable subspace.

We refer to the supplemental material for the proof.

A related problem was studied in Ref. [21], where the authors consider asymptotically entanglement-saving (AES) channels, that preserve a nonzero amount of entanglement after inifinitely many repeated applications of the channel without intermediate operations. Our setting is more general, since we allow intermediate LOCC recovery operations. To see that this leads to a strictly larger class of channels, consider a channel Λ_{ex} acting on a four-dimensional system, with Kraus operators

$$K_1 = |2\rangle\langle 0| + |3\rangle\langle 1|, \qquad K_2 = |0\rangle\langle 2|, \qquad K_3 = |0\rangle\langle 3|. \tag{12}$$

The subspace $C = \mathrm{span}\{|0\rangle, |1\rangle\}$ is correctable, since it is mapped isometrically onto $\mathcal{V} = \mathrm{span}\{|2\rangle, |3\rangle\}$. However,

$$\Lambda^2_{\mathrm{ex}}(\rho) = \mathrm{Tr}(P_C\rho)\,|0\rangle\langle 0| + \mathrm{Tr}(P_{\mathcal{V}}\rho)\,|2\rangle\langle 2|, \tag{13}$$

which is a measure-and-prepare channel and is therefore entanglement breaking [37]. Hence, Λ_{ex} is not AES, although entanglement can be preserved indefinitely by applying a recovery operation between successive channel uses. Therefore, the class characterized in Theorem 3 strictly contains the AES class.

Theorem 3 establishes a strict qualitative dichotomy for asymptotic entanglement based entirely on the existence of a correctable subspace. A natural question that follows is how the system behaves when such a subspace is absent. In this regime, the entanglement must vanish, but understanding the rate at which it disappears is crucial for characterizing the channel's noise properties. The following theorem provides a quantitative bound, showing that without a correctable subspace, the output state converges exponentially fast to the set of separable states, denoted by $\mathcal{S}$. In the following, $\|X\|_1 = \mathrm{Tr}\sqrt{X^\dagger X}$ denotes the trace norm of a matrix X.

Theorem 4. *For any quantum channel Λ acting on a subsystem A of dimension d_A that possesses no correctable subspace, there exists a constant $\kappa \in (0,1)$ such that for any initial bipartite state ρ^{AB}, the trace distance to the set of separable states $\mathcal{S}$ satisfies*

$$\min_{\sigma^{AB}\in\mathcal{S}} \|\Lambda_n[\rho^{AB}] - \sigma^{AB}\|_1 \leq 4\kappa^{n/[2(d_A-1)]}(\sqrt{d_A} - 1). \tag{14}$$

We refer to the supplemental material for the proof.

This exponential decay has severe implications for long-distance quantum communication and entanglement distribution. If we restrict each connection to a single use of a noisy channel lacking a correctable subspace, the state inevitably becomes separable over large distances. To combat this exponential degradation, we will extend the scenario to parallel uses of channels between the links. By utilizing multiple channels simultaneously, we can attempt to offset the noise accumulation. The next theorem illustrates the physical resource cost of this approach, quantifying exactly how many parallel channel uses are required to successfully establish entanglement across a "network" of a given length.

First, we assume that Alice and Bob are connected via linear "network" consisting of n sequential links. Second, the connection between any two adjacent nodes is modeled by m independent identical channels. Third, at each node we can apply arbitrary LOCC filter, which aims to improve the entanglement between the parties. Specifically, we can apply decoding, error correction and encoding operations. Consequently, direct one-way idealized quantum repeaters of the 3rd generation (in classification from [13]) are included as a special case of our model. We assume that all LOCC operations at the repeater stations are noiseless. Since noisy local operations cannot improve the performance of the protocol, a lower bound on the number of channel uses per link obtained in this idealized setting remains valid for realistic noisy repeaters.

Finally, the operational goal is to distribute a bipartite entangled state across the entire "network" such that the endpoint state retains a strictly positive amount of entanglement. Under these conditions, Theorem 5 provides a lower bound on the resources required to counter the exponentially fast decay of entanglement: maintaining a nonvanishing amount of end-to-end entanglement over n links requires $m = \Omega(\log n)$ channel uses per elementary link. In the following theorem, by establishing entanglement between A and B, we mean that:

$$\liminf_{n\to\infty} E_f\left(\rho^{AB}_{n,m(n)}\right) > 0, \tag{15}$$

where $\rho^{AB}_{n,m(n)}$ is a state after n segments of a network.

Theorem 5. *Assume that we have a quantum "network" consisting of n segments. Assume that we send a bipartite quantum state $\rho_0^{A_1\cdots A_m|B} = \rho_0^{A|B}$ through this "network". Each segment consists of LOCC filter which can perform arbitrary $A \times B \to A \times B$ LOCC operation in bipartition $A|B$. After LOCC filter we have m parallel uses of a channel Λ. If*

1. *Λ can be decomposed as*

$$\Lambda = p\mathcal{E} + (1-p)\mathcal{M}, \tag{16}$$

where $\mathcal{E}$ is some EB channel and $\mathcal{M}$ is an arbitrary quantum channel, we can establish entanglement at the output of a "network" only if m scales as:

$$m = \Omega(\ln n). \tag{17}$$

2. *Λ is a qudit amplitude damping channel with $\eta_j > 0$, we can establish entanglement at the output of the network only if m scales as:*

$$m = \Omega(\ln n). \tag{18}$$

The Kraus operators of a qudit amplitude damping channel are:

$$\begin{aligned} K_0 &= \sum_{j=0}^{d-1} \sqrt{q_j}\,|j\rangle\langle j| \\ K_{i\leftarrow j} &= \sqrt{\gamma_{i\leftarrow j}}\,|i\rangle\langle j|\,, \end{aligned} \tag{19}$$

where $\eta_j = \sum_{i<j}\gamma_{i\leftarrow j}$ is total decay probability of the j-th level and $q_j = 1-\eta_j$ is no decay probability. If for every $j > 0$ $\eta_j > 0$ the channel does not have a correctable subspace.

The complete proof can be found in the supplemental material. We provide a brief proof outline here. We first bound the contraction parameter κ for $\Lambda^{\otimes m}$. For channels admitting the decomposition (16), we exploit the existence of a rank-one Kraus representation of the entanglement-breaking component. For the amplitude-damping channel, we instead use a suitable explicit Kraus representation. We then apply Theorem 4 and a uniform continuity bound for the entanglement of formation to relate the distance from the separable set to the number of physical channels per link m and the number of links n. Solving the resulting inequality yields the logarithmic lower bound.

A related logarithmic lower bound was obtained for fault-tolerant qubit circuits under local i.i.d. noise [38]. They showed that preserving an arbitrary unknown logical qubit over n sequential noisy stages requires $m = \Omega(\log n)$ noisy physical qubits per stage. Reliable qubit transmission entails the preservation of entanglement with a reference system and is therefore closely related to the task considered here. This result differs from those presented in this paper in both operational criteria and dimension of the channels involved. Ref. [38] requires the effective channel to approximate the identity on every input qubit state, whereas we only require that some input state retain a nonvanishing amount of entanglement at the very end. Also, their result is restricted to single-qubit noise, while Theorem 5 applies, under its stated assumptions, to channels of arbitrary fixed finite dimension.

Every full-Kraus-rank d-dimensional channel, i.e., a channel of Choi rank d^2, can be written in the form (16) [39]. However, not every channel without a correctable subspace admits such a decomposition. An important example is the amplitude-damping channel, which is covered separately by Theorem 5. The theorem therefore applies to broad classes of physically relevant finite-dimensional channels.

Finally, we show that the logarithmic lower bound is tight for depolarizing noise below the correctable-error threshold of the chosen code family. For a family of quantum codes with linear minimum distance, the failure probability per elementary link decreases exponentially with the number of physical channels and takes the form $e^{-\alpha m}$. Hence, over n links, the total failure probability is bounded by $ne^{-\alpha m}$, so that $m = O(\log n)$ is sufficient. Together with Theorem 5, this establishes the optimal scaling $m = \Theta(\log n)$ within our model. The full derivation of our achievability result is provided in the supplemental material. A related measurement-based qLDPC repeater protocol was recently developed by Shi *et al.* [40], who demonstrated constant-yield Bell-state distribution over arbitrary distances and exponential suppression of the logical error with increasing code size. Their protocol employs entanglement distillation and swapping and uses a different resource accounting; therefore, it does not establish the same minimal number of physical links statement as Theorem 5, but provides a complementary qLDPC-based approach to long-distance entanglement distribution.

Discussion. Our results form a sequence of structural, quantitative, and operational limitations on long-distance entanglement distribution. Theorem 3 first identifies the exact structural condition under which intermediate deterministic LOCC can preserve some entanglement indefinitely: the channel must possess a nontrivial correctable subspace. Theorem 4 provides a quantitative statement regarding entanglement preservation in absence of a correctable subspace: the output approaches the set of separable states exponentially fast in the number of channel segments, uniformly over the initial state and regardless of the adaptive intermediate LOCC operations. By applying theorem 4 to tensor powers of the channel, theorem 5 provides a physical resource constraint to counter the exponential decay. For channels with a nonzero entanglement-breaking component—including all full-Kraus-rank channels—and for qudit amplitude-damping channels, theorem 5 proves that maintaining

$$\liminf_{n\to\infty} E_{\mathrm{f}}\left(\rho^{AB}_{n,m(n)}\right) > 0 \tag{20}$$

requires

$$m = \Omega(\log n). \tag{21}$$

Thus, the logarithmic bound is not an isolated network estimate, but an operational consequence of the underlying entanglement-contraction theorem.

A logarithmic lower bound with an analogous scaling was previously obtained for preserving a single unknown logical qubit in fault-tolerant circuits subject to local i.i.d. qubit noise [38]. Theorem 5 instead provides an analogous lower bound directly in terms of nonvanishing endpoint entanglement and extends it to the specified classes of channels of arbitrary fixed finite dimension. For depolarizing noise in the corresponding correctable regime, quantum-code families with linear minimum distance achieve $m = O(\log n)$. Together with theorem 5, this establishes the tight scaling

$$m = \Theta(\log n) \tag{22}$$

within the considered model. The converse is independent of the particular encoding, decoding, recovery operation, or quantum code employed at the intermediate stations.

The operational meaning of m should be distinguished from the asymptotic rates studied in general quantum-network capacity frameworks. Here, m is the number of physical channels allocated to every elementary link. It need not only represent m distinct transmission lines operating simultaneously, it could represent the same physical channel being used sequentially m times, provided the transmitted systems are stored until all relevant physical systems are received. Such serialization does not improve the resource requirement, but transfers the bulk of the cost to long-lived quantum memory.

The scope of the result is the homogeneous sequential architecture defined above, with noiseless intermediate processing. It captures the idealized logical layer of a broad family of direct one-way error-correction-based repeaters, as discussed in the introduction, but is not intended to describe all repeater architectures. In particular, protocols based on separately generated elementary-link entanglement, two-way purification, and entanglement swapping employ a different resource model. This includes recent measurement-based distillation architectures based on locally prepared resource states and elementary-link Bell pairs [40]. Likewise, genuinely bosonic encoding schemes [41] are not directly covered by our finite-dimensional theorems without an effective finite-dimensional reduction.

Acknowledgments. This work was supported by the National Science Centre Poland (Grant No. 2022/46/E/ST2/00115 and 2024/55/B/ST2/01590) and within the QuantERA II Programme (Grant No. 2021/03/Y/ST2/00178, acronym ExTRaQT) that has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No. 101017733.

* streltsov.physics@gmail.com

[1] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Quantum entanglement, Reviews of Modern Physics **81**, 865 (2009).
[2] C. H. Bennett and G. Brassard, Quantum cryptography: Public key distribution and coin tossing, in *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing* (Bangalore, India, 1984) pp. 175–179.
[3] 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).
[4] H. J. Kimble, The quantum internet, Nature **453**, 1023 (2008).
[5] R. V. Meter and S. J. Devitt, The Path to Scalable Distributed Quantum Computing, Computer **49**, 31 (2016).
[6] R. Pal, S. Bandyopadhyay, and S. Ghosh, Entanglement sharing through noisy qubit channels: One-shot optimal singlet fraction, Physical Review A **90**, 052304 (2014).
[7] A. Streltsov, R. Augusiak, M. Demianowicz, and M. Lewenstein, Progress towards a unified approach to entanglement distribution, Physical Review A **92**, 012335 (2015).
[8] T. Krisnanda, Distribution of quantum entanglement: Principles and applications (2020), arXiv:2003.08657 [quant-ph].
[9] M. Zuppardo, T. Krisnanda, T. Paterek, S. Bandyopadhyay, A. Banerjee, P. Deb, S. Halder, K. Modi, and M. Paternostro, Excessive distribution of quantum entanglement, Physical Review A **93**, 012305 (2016).
[10] L. Lami and V. Giovannetti, Entanglement-breaking indices, Journal of Mathematical Physics **56**, 092201 (2015).
[11] M. Rahaman, S. Jaques, and V. I. Paulsen, Eventually entanglement breaking maps, Journal of Mathematical Physics **59**, 062201 (2018).
[12] H.-J. Briegel, W. Dür, J. I. Cirac, and P. Zoller, Quantum repeaters: The role of imperfect local operations in quantum communication, Phys. Rev. Lett. **81**, 5932 (1998).
[13] S. Muralidharan, L. Li, J. Kim, N. Lütkenhaus, M. D. Lukin, and L. Jiang, Optimal architectures for long distance quantum communication, Scientific Reports **6**, 20463 (2016).
[14] L.-M. Duan, M. D. Lukin, J. I. Cirac, and P. Zoller, Long-distance quantum communication with atomic ensembles and linear optics, Nature **414**, 413 (2001).
[15] S. Muralidharan, J. Kim, N. Lütkenhaus, M. D. Lukin, and L. Jiang, Ultrafast and fault-tolerant quantum communication across long distances, Physical Review Letters **112**, 250501 (2014).
[16] W. J. Munro, A. M. Stephens, S. J. Devitt, K. A. Harrison, and K. Nemoto, Quantum communication without the necessity of quantum memories, Nature Photonics **6**, 777 (2012).
[17] K. Azuma, S. E. Economou, D. Elkouss, P. Hilaire, L. Jiang, H.-K. Lo, and I. Tzitrin, Quantum repeaters: From quantum networks to the quantum internet, Review of Modern Physics **95**, 045006 (2023).
[18] K. Azuma, K. Tamaki, and H.-K. Lo, All-photonic quantum repeaters, Nature Communications **6**, 6787 (2015).
[19] S. A. Hill and W. K. Wootters, Entanglement of a pair of quantum bits, Physical Review Letters **78**, 5022 (1997).
[20] W. K. Wootters, Entanglement of formation of an arbitrary state of two qubits, Physical Review Letters **80**, 2245 (1998).
[21] L. Lami and V. Giovannetti, Entanglement-saving channels, Journal of Mathematical Physics **57**, 032201 (2016).
[22] R. Namiki, L. Jiang, J. Kim, and N. Lütkenhaus, Role of syndrome information on a one-way quantum repeater using teleportation-based error correction, Phys. Rev. A **94**, 052304 (2016).
[23] F. Ewert, M. Bergmann, and P. van Loock, Ultrafast long-

distance quantum communication with static linear optics, Phys. Rev. Lett. **117**, 210501 (2016).
[24] S. Muralidharan, C.-L. Zou, L. Li, and L. Jiang, One-way quantum repeaters with quantum reed-solomon codes, Phys. Rev. A **97**, 052316 (2018).
[25] K. J. Wo, G. Avis, F. Rozpędek, M. F. Mor-Ruiz, G. Pieplow, T. Schröder, L. Jiang, A. S. Sørensen, and J. Borregaard, Resource-efficient fault-tolerant one-way quantum repeater with code concatenation, npj Quantum Information **9**, 123 (2023).
[26] D. Niu, Y. Zhang, A. Shabani, and H. Shapourian, All-photonic one-way quantum repeaters with measurement-based error correction, npj Quantum Information **9**, 106 (2023).
[27] A. N. Glaudell, E. Waks, and J. M. Taylor, Serialized quantum error correction protocol for high-bandwidth quantum repeaters, New Journal of Physics **18**, 093008 (2016).
[28] K. Azuma, A. Mizutani, and H.-K. Lo, Fundamental rate-loss trade-off for the quantum internet, Nature Communications **7**, 13523 (2016).
[29] L. Rigovacca, G. Kato, S. Bäuml, M. S. Kim, W. J. Munro, and K. Azuma, Versatile relative entropy bounds for quantum networks, New Journal of Physics **20**, 013033 (2018).
[30] S. Pirandola, End-to-end capacities of a quantum communication network, Communications Physics **2**, 51 (2019).
[31] S. Bäuml, K. Azuma, G. Kato, and D. Elkouss, Linear programs for entanglement and key distribution in the quantum internet, Communications Physics **3**, 55 (2020).
[32] M. Pant, H. Krovi, D. Towsley, L. Tassiulas, L. Jiang, P. Basu, D. Englund, and S. Guha, Routing entanglement in the quantum internet, npj Quantum Information **5**, 25 (2019).
[33] M.-D. Choi, Completely positive linear maps on complex matrices, Linear Algebra and its Applications **10**, 285 (1975).
[34] W. Dür, G. Vidal, and J. I. Cirac, Three qubits can be entangled in two inequivalent ways, Physical Review A **62**, 062314 (2000).
[35] T. Konrad, F. de Melo, M. Tiersch, C. Kasztelan, A. Aragão, and A. Buchleitner, Evolution equation for quantum entanglement, Nature Physics **4**, 99 (2008).
[36] R. Blume-Kohout, H. K. Ng, D. Poulin, and L. Viola, Characterizing the structure of preserved information in quantum processes, Physical Review Letters **100**, 030501 (2008).
[37] M. Horodecki, P. W. Shor, and M. B. Ruskai, Entanglement breaking channels, Reviews in Mathematical Physics **15**, 629 (2003), https://doi.org/10.1142/S0129055X03001709.
[38] O. Fawzi, A. Müller-Hermes, and A. Shayeghi, A lower bound on the space overhead of fault-tolerant quantum computation, in *13th Innovations in Theoretical Computer Science Conference (ITCS 2022)*, Leibniz International Proceedings in Informatics (LIPIcs), Vol. 215 (Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2022) pp. 68:1–68:20.
[39] I. George, C. Hirche, T. Nuradha, and M. M. Wilde, Quantum Doeblin Coefficients: Interpretations and Applications, Quantum **10**, 2115 (2026).
[40] Y. Shi, A. Patil, and S. Guha, Measurement-based entanglement distillation and constant-rate quantum repeaters over arbitrary distances, Phys. Rev. Lett. **135**, 130803 (2025).
[41] S. N. Swain and T. C. Ralph, Loss-tolerant quantum communication via bosonic-gkp-parity-encoding (2026), arXiv:2604.09002 [quant-ph].
[42] G. Vidal, Entanglement monotones, Journal of Modern Optics **47**, 355 (2000).
[43] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press, 2010).
[44] Y. Guo, Strict entanglement monotonicity under local operations and classical communication, Physical Review A **99**, 022338 (2019).
[45] S. Khatri and M. M. Wilde, Principles of Quantum Communication Theory: A Modern Approach (2024), arXiv:2011.04672 [quant-ph].
[46] E. Schmidt, Zur theorie der linearen und nichtlinearen integralgleichungen, Mathematische Annalen **63**, 433 (1907).
[47] G. Gour, Family of concurrence monotones and its applications, Physical Review A **71**, 012318 (2005).
[48] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis* (Cambridge University Press, 1991).
[49] A. Uhlmann, The "transition probability" in the state space of a *-algebra, Reports on Mathematical Physics **9**, 273 (1976).
[50] A. Philip, S. Rethinasamy, V. Russo, and M. M. Wilde, Schrödinger as a Quantum Programmer: Estimating Entanglement via Steering, Quantum **8**, 1366 (2024).
[51] C. A. Fuchs and J. van de Graaf, Cryptographic distinguishability measures for quantum-mechanical states, IEEE Transactions on Information Theory **45**, 1216 (1999).
[52] R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth, On the Lambert W function, Advances in Computational Mathematics **5**, 329 (1996).
[53] S. Chessa and V. Giovannetti, Quantum capacity analysis of multi-level amplitude damping channels, Communications Physics **4**, 22 (2021).
[54] D. I. Cartwright and M. J. Field, A refinement of the arithmetic mean-geometric mean inequality, Proceedings of the American Mathematical Society **71**, 36 (1978).
[55] A. Winter, Tight uniform continuity bounds for quantum entropies: Conditional entropy, relative entropy distance and energy constraints, Communications in Mathematical Physics **347**, 291 (2016), arXiv:1507.07775 [quant-ph].
[56] W. Hoeffding, Probability inequalities for sums of bounded random variables, Journal of the American Statistical Association **58**, 13 (1963), https://doi.org/10.1080/01621459.1963.10500830.

SUPPLEMENTAL MATERIAL

Concurrence

For a two-qubit state ρ the concurrence is defined as [19]

$$C(\rho) = \max\{\xi_1 - \xi_2 - \xi_3 - \xi_4, 0\}, \tag{23}$$

where ξ_i are the square roots of the eigenvalues of $\rho\tilde{\rho}$, arranged in decreasing order, and

$$\tilde{\rho} = (\sigma_y \otimes \sigma_y)\rho^*(\sigma_y \otimes \sigma_y), \tag{24}$$

where σ_y is the Pauli y-matrix and ρ^* denotes complex conjugation in the computational basis. For two-qubit states, the concurrence is directly related to the entanglement of formation via

$$E_f(\rho) = h\left(\frac{1}{2} + \frac{1}{2}\sqrt{1 - C(\rho)^2}\right), \tag{25}$$

where $h(x) = -x\log_2 x - (1-x)\log_2(1-x)$ is the binary entropy function.

Proof of Theorem 1

In the proof, we employ the concurrence factorization law [35], which states that for any completely positive trace-preserving map Λ one has

$$C((\Lambda \otimes \mathbb{1})[\rho^{AB}]) \leq C((\Lambda \otimes \mathbb{1})[\phi^+])\, C(\rho^{AB}) = C(\Gamma)\, C(\rho^{AB}). \tag{26}$$

Our protocol consists of n rounds. In each round, an SLOCC instrument is applied first, followed by the channel Λ. The SLOCC instrument used in the i-th round may depend on the complete history of all outcomes obtained in the previous rounds. Let $h_{i-1} = (x_1, \ldots, x_{i-1})$ denote a particular history before the i-th round. At the beginning of this round, the parties share the normalized state $\rho_{h_{i-1}}$. Let $\left\{\mathcal{F}_{i,x_i}^{h_{i-1}}\right\}_{x_i}$ be the SLOCC instrument chosen after observing the history h_{i-1}. The outcome x_i occurs with conditional probability

$$p(x_i \mid h_{i-1}) = \mathrm{Tr}\left[\mathcal{F}_{i,x_i}^{h_{i-1}}(\rho_{h_{i-1}})\right], \tag{27}$$

and the corresponding normalized post-filter state is

$$\sigma_{h_i} = \frac{\mathcal{F}_{i,x_i}^{h_{i-1}}(\rho_{h_{i-1}})}{p(x_i \mid h_{i-1})}, \qquad h_i = (h_{i-1}, x_i). \tag{28}$$

Since concurrence is an entanglement monotone, it cannot increase on average under SLOCC. Therefore,

$$\sum_{x_i} p(x_i|h_{i-1})\, C(\sigma_{h_i}) \leq C(\rho_{h_{i-1}}). \tag{29}$$

After the filter, the channel Λ acts on one subsystem. The state at the end of the i-th round is therefore

$$\rho_{h_i} = (\Lambda \otimes \mathbb{1})(\sigma_{h_i}). \tag{30}$$

We define

$$q = C\left[(\Lambda \otimes \mathbb{1})(\phi^+)\right]. \tag{31}$$

The factorization law for concurrence implies, for every outcome x_i,

$$C(\rho_{h_i}) = C\left[(\Lambda \otimes \mathbb{1})(\sigma_{h_i})\right] \leq q\, C(\sigma_{h_i}). \tag{32}$$

Multiplying Eq. (32) by $p(x_i \mid h_{i-1})$ and summing over all possible outcomes gives

$$\sum_{x_i} p(x_i \mid h_{i-1}) C(\rho_{h_i}) \leq q \sum_{x_i} p(x_i \mid h_{i-1}) C(\sigma_{h_i}). \tag{33}$$

Using Eq. (29), we obtain

$$\sum_{x_i} p(x_i \mid h_{i-1}) C(\rho_{h_i}) \leq q\, C(\rho_{h_{i-1}}). \tag{34}$$

Let $H_i = (X_1, \ldots, X_i)$ denote the random history after the i-th round, and define the random variable

$$C_i = C(\rho_{H_i}). \tag{35}$$

For every fixed history h_{i-1}, the left-hand side of Eq. (34) is precisely the conditional expectation

$$\mathbb{E}\left[C_i \mid H_{i-1} = h_{i-1}\right] = \sum_{x_i} p(x_i \mid h_{i-1}) C(\rho_{h_i}). \tag{36}$$

Consequently,

$$\mathbb{E}\left[C_i \mid H_{i-1} = h_{i-1}\right] \leq q\, C(\rho_{h_{i-1}}) \tag{37}$$

for every possible history h_{i-1}. Equivalently,

$$\mathbb{E}\left[C_i \mid H_{i-1}\right] \leq q C_{i-1}. \tag{38}$$

Taking the expectation of both sides with respect to the random history H_{i-1} and using the law of total expectations, we find

$$\begin{aligned}\mathbb{E}[C_i] &= \mathbb{E}\left[\mathbb{E}\left[C_i \mid H_{i-1}\right]\right] \\ &\leq q\, \mathbb{E}[C_{i-1}].\end{aligned} \tag{39}$$

Iterating this inequality over n rounds yields

$$\mathbb{E}[C_n] \leq q^n \mathbb{E}[C_0] = q^n C(\rho_0), \tag{40}$$

where in the last step we used a fact that we start with a fixed state so there is only one possibility, and therefore expectation value is equal to the concurrence value itself.

$$\mathbb{E}[C_n] \leq q^n C(\rho_0). \tag{41}$$

Finally, since C_n is a non-negative random variable, Markov's inequality gives, for every $c > 0$,

$$\begin{aligned}\Pr\left[C_n \geq c\right] &\leq \frac{\mathbb{E}[C_n]}{c} \\ &\leq \frac{q^n C(\rho_0)}{c} \\ &\leq \frac{q^n}{c}.\end{aligned} \tag{42}$$

Proof of Corollary 2

In this section we present a proof of Corollary 2 using factorization law from [35]. We prove the claim by induction. Specifically, we establish that

$$C((\mathbb{1}\otimes\Lambda_n)[\rho^{AB}]) \leq C(\Gamma)^n\, C(\rho^{AB}). \tag{43}$$

To do so, we employ the factorization law for concurrence, which extends to mixed two-qubit states in the form [35]

$$C((\Lambda\otimes\mathbb{1})[\rho^{AB}]) \leq C((\Lambda\otimes\mathbb{1})[\phi^{+}])\, C(\rho^{AB}) = C(\Gamma)\, C(\rho^{AB}). \tag{44}$$

For the base case $n = 1$, we obtain

$$C((\Lambda\otimes\mathbb{1})\circ F_1[\rho^{AB}]) \leq C(\Gamma)\, C(F_1[\rho^{AB}]) \leq C(\Gamma)\, C(\rho^{AB}), \tag{45}$$

where F_1 is an LOCC filter. In the second inequality, we used the fact that LOCC operations cannot increase concurrence. We have thus established that Eq. (43) holds for $n = 1$.

For the induction step, note that

$$\Lambda_{n+1}[\rho^{AB}] = (\Lambda\otimes\mathbb{1})\circ F_{n+1}\circ\Lambda_n[\rho^{AB}] = (\Lambda\otimes\mathbb{1})\circ F_{n+1}[\sigma^{AB}], \tag{46}$$

where we defined the state $\sigma^{AB} = \Lambda_n[\rho^{AB}]$. Now assume that Eq. (43) holds for some n. For $n+1$, we have

$$\begin{aligned} C(\Lambda_{n+1}[\rho^{AB}]) &= C((\Lambda\otimes\mathbb{1})\circ F_{n+1}[\sigma^{AB}]) \\ &\leq C(\Gamma)\, C(F_{n+1}[\sigma^{AB}]) \\ &\leq C(\Gamma)\, C(\sigma^{AB}). \end{aligned} \tag{47}$$

By the induction hypothesis,

$$C(\sigma^{AB}) \leq C(\Gamma)^n\, C(\rho^{AB}). \tag{48}$$

Combining the above inequalities, we obtain

$$C(\Lambda_{n+1}[\rho^{AB}]) \leq C(\Gamma)^{n+1}\, C(\rho^{AB}), \tag{49}$$

which completes the induction.

Proof of Theorem 3

Proof. It is clear to see that if Λ contains a correctable subspace then Eq. (11) holds. We will now deal with the reverse implication.

If Eq. (11) holds, then there must exist an entangled state ρ^{AB} such that

$$E_{\mathrm{f}}(\Lambda\otimes\mathbb{1}[\rho^{AB}]) = E_{\mathrm{f}}(\rho^{AB}). \tag{50}$$

See Lemma 6 in the supplemental material for proof. We now show that this condition implies the existence of a pure entangled state $|\psi^{AB}\rangle$ satisfying

$$E_{\mathrm{f}}(\Lambda\otimes\mathbb{1}[\psi^{AB}]) = E_{\mathrm{f}}(\psi^{AB}). \tag{51}$$

For this, let $\{p_i, \psi_i^{AB}\}$ be an optimal pure state decomposition of ρ^{AB}, i.e.,

$$E_{\mathrm{f}}(\rho^{AB}) = \sum_i p_i E_{\mathrm{f}}(\psi_i^{AB}). \tag{52}$$

Since E_{f} does not increase under local operations, for each of the states $|\psi_i^{AB}\rangle$ we have

$$E_{\mathrm{f}}(\Lambda\otimes\mathbb{1}[\psi_i^{AB}]) \leq E_{\mathrm{f}}(\psi_i^{AB}). \tag{53}$$

By contradiction, assume that for all entangled states $|\psi_j^{AB}\rangle$ this inequality is strict, i.e.,

$$E_{\mathrm{f}}(\Lambda\otimes\mathbb{1}[\psi_j^{AB}]) < E_{\mathrm{f}}(\psi_j^{AB}). \tag{54}$$

Using these assumptions and convexity [42] of E_{f} we obtain

$$\begin{aligned} E_{\mathrm{f}}(\Lambda\otimes\mathbb{1}[\rho^{AB}]) &= E_{\mathrm{f}}\left(\sum_i p_i \Lambda\otimes\mathbb{1}[\psi_i^{AB}]\right) \\ &\leq \sum_i p_i E_{\mathrm{f}}(\Lambda\otimes\mathbb{1}[\psi_i^{AB}]) \\ &< \sum_i p_i E_{\mathrm{f}}(\psi_i^{AB}) = E_{\mathrm{f}}(\rho^{AB}). \end{aligned} \tag{55}$$

This contradicts Eq. (50), which implies that Eq. (51) is true for some entangled state $|\psi\rangle^{AB}$.

Let now $|\psi\rangle^{AB}$ be an entangled state such that Eq. (51) holds. In the following, let $\{K_i\}$ be a set of Kraus operators corresponding to the quantum channel Λ. We further define the probabilities q_i and pure quantum states ϕ_i^{AB} as follows:

$$q_i = \mathrm{Tr}\left[K_i\otimes\mathbb{1}\psi^{AB}K_i^{\dagger}\otimes\mathbb{1}\right], \tag{56}$$

$$\phi_i^{AB} = \frac{1}{q_i} K_i\otimes\mathbb{1}\psi^{AB}K_i^{\dagger}\otimes\mathbb{1}. \tag{57}$$

Without loss of generality we assume that $q_i > 0$ for all i.

Consider now the average entanglement of the ensemble $\{q_i, \phi_i^{AB}\}$:

$$\sum_i q_i E_{\mathrm{f}}(\phi_i^{AB}) = \sum_i q_i S(\phi_i^B). \tag{58}$$

By the concavity of the von Neumann entropy it holds that

$$\sum_i q_i S(\phi_i^B) \leq S\left(\sum_i q_i \phi_i^B\right) = S(\psi^B). \tag{59}$$

Since the von Neumann entropy is strictly concave [43], Eq. (59) holds with equality if and only if

$$\phi_i^B = \psi^B \quad \text{for all } i. \tag{60}$$

By convexity of the entanglement of formation we further have

$$E_{\mathrm{f}}(\Lambda\otimes\mathbb{1}[\psi^{AB}]) \leq \sum_i q_i E_{\mathrm{f}}(\phi_i^{AB}) = \sum_i q_i S(\phi_i^B). \tag{61}$$

Summarizing these arguments, the condition (51) implies that $\phi_i^B = \psi^B$ holds for all i. A similar argument has also been used in [44].

In the next step we introduce a reference system R. By the Stinespring dilation theorem [43], for any quantum channel Λ there exists a unitary U_{RA} on RA such that

$$\Lambda \otimes \mathbb{1}[\psi^{AB}] = \mathrm{Tr}_R\left[U_{RA}\left(|0\rangle\langle 0|^R \otimes \psi^{AB}\right) U_{RA}^\dagger\right] \tag{62}$$

with a pure state $|0\rangle^R$. We further define the pure state

$$|\mu\rangle^{RAB} = U_{RA}\left(|0\rangle^R \otimes |\psi\rangle^{AB}\right). \tag{63}$$

Note that any rank-1 POVM $\{M_j^R\}$ on the reference system R gives rise to a set of Kraus operators $\{L_j\}$ for the channel Λ, such that

$$\mathrm{Tr}_R\left[M_j^R \mu^{RAB}\right] = \left(L_j \otimes \mathbb{1}\right)\psi^{AB}\left(L_j^\dagger \otimes \mathbb{1}\right). \tag{64}$$

From the arguments given above we see that Eq. (51) implies that the following equality holds for all rank-1 POVMs $\{M_j^R\}$ and all j:

$$\mathrm{Tr}_{RA}\left[M_j^R \mu^{RAB}\right] = \mathrm{Tr}_{RAB}\left[M_j^R \mu^{RAB}\right] \times \psi^B. \tag{65}$$

We will now take a closer look at the state μ^{RB}. Due to Eq. (65), it holds that

$$\mathrm{Tr}_R\left[M_j^R \mu^{RB}\right] = \mathrm{Tr}_{RB}\left[M_j^R \mu^{RB}\right] \times \psi^B \tag{66}$$

for all POVMs $\{M_j^R\}$ and all j. As we show in Lemma 7 in the supplemental material, this is only possible if μ^{RB} is a product state, i.e.,

$$\mu^{RB} = \mu^R \otimes \psi^B. \tag{67}$$

Recalling that the state μ^{RAB} is pure and using Eq. (67), we see that it is always possible to attach an ancillary system A' on Alice's side in a pure state $|0\rangle^{A'}$ and perform a unitary $U_{AA'}$ on Alice's part of the total state $|\mu\rangle^{RAB} \otimes |0\rangle^{A'}$ such that

$$U_{AA'}\left(|\mu\rangle^{RAB} \otimes |0\rangle^{A'}\right) = |\nu\rangle^{RA'} \otimes |\psi\rangle^{AB}, \tag{68}$$

where $|\nu\rangle^{RA'}$ is a purification of μ^R. This implies that the state $\mu^{AB} \otimes |0\rangle\langle 0|^{A'}$ is transformed as follows:

$$U_{AA'}\left(\mu^{AB} \otimes |0\rangle\langle 0|^{A'}\right) U_{AA'}^\dagger = |\psi\rangle\langle\psi|^{AB} \otimes \nu^{A'}. \tag{69}$$

Recalling that $\mu^{AB} = \Lambda \otimes \mathbb{1}[\psi^{AB}]$, these arguments imply that there exists a quantum operation Λ' on Alice's subsystem such that

$$(\Lambda' \circ \Lambda) \otimes \mathbb{1}[\psi^{AB}] = \psi^{AB}. \tag{70}$$

Since $|\psi\rangle^{AB}$ is entangled, it has a Schmidt decomposition of the form $|\psi\rangle^{AB} = \sum_i \sqrt{\lambda_i}\,|ii\rangle$ with at least two nonzero Schmidt coefficients λ_i. Consider now a pure state $|\eta\rangle$ of the form

$$|\eta\rangle = \sum_i c_i\,|i\rangle^A, \tag{71}$$

where c_i are complex coefficients and $|i\rangle^A$ are local basis states on Alice's subsystem for which the state $|\psi\rangle^{AB}$ has a nonzero Schmidt coefficient. Using Lemma 8 in the supplemental material, it follows that

$$\Lambda' \circ \Lambda[\eta] = \eta \tag{72}$$

and the proof of the theorem is complete. □

Lemma 6. *Let Λ be a quantum channel and ρ^{AB} an initial state. If*

$$\lim_{n\to\infty} E_{\mathrm{f}}(\Lambda_n[\rho^{AB}]) > 0, \tag{73}$$

then there exists an entangled state μ^{AB} such that

$$E_{\mathrm{f}}(\Lambda \otimes \mathbb{1}[\mu^{AB}]) = E_{\mathrm{f}}(\mu^{AB}). \tag{74}$$

Proof. Let

$$c = \lim_{n\to\infty} E_{\mathrm{f}}(\Lambda_n[\rho^{AB}]), \tag{75}$$

so that $c > 0$ by assumption. Then, for every $\varepsilon > 0$, there exists N such that

$$\left|E_{\mathrm{f}}(\Lambda_k[\rho^{AB}]) - c\right| < \varepsilon \tag{76}$$

for all $k > N$. In particular, for all $k > N$ we also have

$$\left|E_{\mathrm{f}}(\Lambda_{k+1}[\rho^{AB}]) - c\right| < \varepsilon. \tag{77}$$

Now observe that

$$\Lambda_{k+1}[\rho^{AB}] = (\Lambda \otimes \mathbb{1}) \circ F_{k+1} \circ \Lambda_k[\rho^{AB}] = (\Lambda \otimes \mathbb{1})[\sigma^{AB}], \tag{78}$$

where we define

$$\sigma^{AB} = F_{k+1} \circ \Lambda_k[\rho^{AB}]. \tag{79}$$

Since the entanglement of formation is non-increasing under LOCC, it follows that

$$E_{\mathrm{f}}(\Lambda_{k+1}[\rho^{AB}]) \leq E_{\mathrm{f}}(\sigma^{AB}) \leq E_{\mathrm{f}}(\Lambda_k[\rho^{AB}]). \tag{80}$$

Combining this with Eqs. (76) and (77), we obtain

$$c - \varepsilon < E_{\mathrm{f}}(\sigma^{AB}) < c + \varepsilon, \tag{81}$$

or equivalently,

$$\left|E_{\mathrm{f}}(\sigma^{AB}) - c\right| < \varepsilon. \tag{82}$$

Using the triangle inequality together with Eq. (78), we therefore obtain

$$\begin{aligned}\left|E_{\mathrm{f}}(\Lambda \otimes \mathbb{1}[\sigma^{AB}]) - E_{\mathrm{f}}(\sigma^{AB})\right| &\leq \left|E_{\mathrm{f}}(\Lambda \otimes \mathbb{1}[\sigma^{AB}]) - c\right| \\ &\quad + \left|E_{\mathrm{f}}(\sigma^{AB}) - c\right| < 2\varepsilon.\end{aligned} \tag{83}$$

Since $\varepsilon > 0$ can be chosen arbitrarily small, the proof is complete. □

Lemma 7. *A quantum state ρ^{AB} fulfills*

$$\mathrm{Tr}_A\left[M_i^A\rho^{AB}\right] = \mathrm{Tr}_{AB}\left[M_i^A\rho^{AB}\right]\times\rho^B \tag{84}$$

for all POVMs $\{M_i^A\}$ on the subsystem A and all i if and only if ρ^{AB} is of the form

$$\rho^{AB} = \rho^A\otimes\rho^B. \tag{85}$$

Proof. The implication $\Leftarrow$ is immediate. We now prove the implication $\Rightarrow$. Let us write a state ρ^{AB} in its Operator-Schmidt decomposition [45] such that $\rho^{AB} = \sum_k \sqrt{\lambda_k}E_k\otimes F_k$, where $\{E_k\}$ and $\{F_k\}$ form orthonormal sets of Hermitian operators acting on systems A and B, respectively. Without loss of generality, we restrict ourselves to $\lambda_k \neq 0$. We also expand the operator M_i^A in the basis $\{E_\alpha\}$ as $M_i^A = \sum_\alpha m_\alpha E_\alpha$, where m_α are real coefficients. We have:

$$\begin{aligned}\mathrm{Tr}_A\left[M_i^A\rho^{AB}\right] &= \sum_{k,\alpha}\sqrt{\lambda_k}m_\alpha\mathrm{Tr}\left(E_kE_\alpha\right)F_k\\ &= \sum_k\sqrt{\lambda_k}m_kF_k = \mathrm{Tr}\left[M_i^A\rho^A\right]\sum_k b_kF_k,\end{aligned} \tag{86}$$

where in the last equality we used Eq. (84) and $\mathrm{Tr}\left[M_i^A\rho^{AB}\right] = \mathrm{Tr}\left[M_i^A\rho^A\right]$. We also decomposed ρ^B in a basis F_k obtaining $\rho^B = \sum_k b_kF_k$ with real coefficients b_k. Coefficient m_k satisfies $m_k = \mathrm{Tr}\left(M_i^AE_k\right)$. Using this, and comparing coefficients in Eq. (86) we obtain:

$$\mathrm{Tr}\left(M_i^AE_k\right) = \frac{b_k}{\sqrt{\lambda_k}}\mathrm{Tr}\left(M_i^A\rho^A\right). \tag{87}$$

It must hold for all M_i^A, thus $E_k = \frac{b_k}{\sqrt{\lambda_k}}\rho^A$. Plugging it into ρ^{AB} we obtain:

$$\rho^{AB} = \sum_k\sqrt{\lambda_k}E_k\otimes F_k = \sum_k\rho^A\otimes b_kF_k = \rho^A\otimes\rho^B. \tag{88}$$

□

Lemma 8. *Let $|\psi\rangle^{AB}$ be a pure entangled state satisfying*

$$(N\otimes\mathbb{1})[\psi^{AB}] = \psi^{AB}, \tag{89}$$

where N is a quantum channel and

$$|\psi\rangle^{AB} = \sum_i\sqrt{\lambda_i}\,|i\rangle^A\,|i\rangle^B, \tag{90}$$

such that $|\psi\rangle^{AB}$ has at least two non-zero Schmidt coefficients. Here, $\{|i\rangle^A\}$ and $\{|i\rangle^B\}$ denote the Schmidt bases for subsystems A and B, respectively.

Then, for any pure state $|\eta\rangle$ belonging to the subspace spanned by $\{|i\rangle^A\}$, i.e.

$$|\eta\rangle = \sum_i c_i\,|i\rangle^A, \tag{91}$$

with $c_i\in\mathbb{C}$, it holds that

$$N[\eta] = \eta. \tag{92}$$

Proof. From Eq. (89) and the assumed Schmidt decomposition of $|\psi\rangle^{AB}$, we obtain

$$\sum_{i,j}\sqrt{\lambda_i\lambda_j}\,N\left[|i\rangle\langle j|^A\right]\otimes|i\rangle\langle j|^B = \sum_{i,j}\sqrt{\lambda_i\lambda_j}\,|i\rangle\langle j|^A\otimes|i\rangle\langle j|^B. \tag{93}$$

Left multiplying by $\mathbb{1}\otimes|k\rangle\langle l|$ and tracing over subsystem B, we obtain

$$N\left[|l\rangle\langle k|^A\right] = |l\rangle\langle k|^A. \tag{94}$$

Since N is a quantum channel, it is linear on operators. This completes the proof. □

Proof of Theorem 4

In the following we consider a pure bipartite state $|\psi\rangle^{AB} = \sum_{i=0}^{k-1}\sqrt{\lambda_i}\,|ii\rangle$ with Schmidt rank at most k [43, 46], such that the Schmidt coefficients λ_i are sorted in non-increasing order. We define the G_k-concurrence as, (similar quantity was also defined in [47])

$$G_k(\psi^{AB}) = k\left(\prod_{i=0}^{k-1}\lambda_i\right)^{1/k}. \tag{95}$$

This definition can be extended in a straightforward way also to non-normalized vectors in $\mathcal{H}_{AB}$. We further define the state

$$|\phi\rangle_k^+ = \frac{1}{\sqrt{k}}\sum_{i=0}^{k-1}|ii\rangle \tag{96}$$

to be maximally entangled state of rank k.

Let us consider a quantum channel Λ. We define:

$$\kappa = \min_{K_i}\sum_i G_k[(K_i\otimes\mathbb{1})\,\phi_k^+(K_i^\dagger\otimes\mathbb{1})], \tag{97}$$

where the minimization is taken over all Kraus decompositions of Λ. The κ parameter quantifies how good is channel in entanglement preservation.

Lemma 9. *For any pure state $|\psi\rangle^{AB}$ with Schmidt rank at most k and any matrix M on $\mathcal{H}_A$, the following equality holds:*

$$\begin{aligned}&G_k[(M\otimes\mathbb{1})\,\psi^{AB}(M^\dagger\otimes\mathbb{1})]\\ &\qquad = G_k[(M\otimes\mathbb{1})\,\phi_k^+(M^\dagger\otimes\mathbb{1})]\times G_k(\psi^{AB}).\end{aligned} \tag{98}$$

Proof. If k is greater than Schmidt rank of $|\psi^{AB}\rangle$, both sides of the Eq. (98) are 0. Assume that Schmidt rank of $|\psi^{AB}\rangle$ is k. We can write our state as [45]:

$$|\psi^{AB}\rangle = \sum_{i,j}X_{i,j}\,|i\rangle_A\,|j\rangle_B = \sum_{i=1}^k\sqrt{\lambda_i}\,|e_i\rangle_A\,|f_i\rangle_B, \tag{99}$$

where $|i\rangle_A, |e_i\rangle_A$ are orthonormal bases in space $\mathcal{H}_A$, $|j\rangle_B, |f_i\rangle_B$ are orthonormal bases in space $\mathcal{H}_B$ and $\sqrt{\lambda_i}$ are singular values of X. We further have $M\otimes\mathbb{1}\,|\psi^{AB}\rangle = \sum_{i,j}(MX)_{i,j}\,|i\rangle_A\,|j\rangle_B$.

Thus, the Schmidt coefficients of $M\otimes\mathbb{1}\,|\psi^{AB}\rangle$ are singular values of MX.

Let

$$\mathcal{S}_A = \mathrm{span}\{|e_1\rangle,\ldots,|e_k\rangle\},\qquad \mathcal{S}_B = \mathrm{span}\{|f_1\rangle,\ldots,|f_k\rangle\} \tag{100}$$

be the Schmidt supports on subsystems A and B, respectively. Let

$$X_k = \mathrm{diag}\left(\sqrt{\lambda_1},\ldots,\sqrt{\lambda_k}\right) \tag{101}$$

be the coefficient matrix represented in the corresponding Schmidt bases, and let

$$M_k := M|_{\mathcal{S}_A} : \mathcal{S}_A \to \mathcal{H}_A \tag{102}$$

be the restriction of M to the Schmidt support. The nonzero singular values of MX coincide with the singular values of M_kX_k. Singular values of M_kX_k satisfy [48]:

$$\begin{aligned}\prod_{i=1}^{k}\sigma_i(M_kX_k) &= \sqrt{\det\left(X_k^\dagger M_k^\dagger M_kX_k\right)}\\ &= \sqrt{\det(M_k^\dagger M_k)}\sqrt{\det(X_k^\dagger X_k)} = \prod_{i=1}^{k}s_i\prod_{i=1}^{k}\sqrt{\lambda_i},\end{aligned} \tag{103}$$

where s_i are singular values for a matrix M_k. Using above relation we can obtain formula for G_k-concurrence.

$$\begin{aligned}G_k[(M\otimes\mathbb{1})\,\psi^{AB}(M^\dagger\otimes\mathbb{1})] &= k\left(\prod_{i=1}^{k}s_i^2\prod_{i=1}^{k}\lambda_i\right)^{1/k}\\ &= \left(\prod_{i=1}^{k}s_i^2\right)^{1/k}G_k(\psi^{AB}).\end{aligned} \tag{104}$$

Substituting $\psi^{AB}=\phi_k^+$ in Eq. (104), we see that

$$G_k[(M\otimes\mathbb{1})\,\phi_k^+(M^\dagger\otimes\mathbb{1})] = \left(\prod_{i=1}^{k}s_i^2\right)^{1/k}. \tag{105}$$

Plugging it into Eq. (104) finishes the proof. □

Proposition 10. *For a pure state $|\psi\rangle^{AB}$ with Schmidt rank at most k and a local quantum operation Λ which has no correctable subspace the state $\Lambda\otimes\mathbb{1}[\psi^{AB}]$ can be written as*

$$\Lambda\otimes\mathbb{1}[\psi^{AB}] = \sum_i p_i\,|\phi_i\rangle\langle\phi_i|^{AB} \tag{106}$$

where each state $|\phi_i\rangle^{AB}$ has Schmidt rank at most k and

$$\sum_i p_iG_k(\phi_i^{AB}) = \kappa_sG_k(\psi^{AB}) \tag{107}$$

with $\kappa_s<1$.

Proof. Let $\{K_i\}$ be a set of Kraus operators for Λ. Using Lemma 9, we have

$$\begin{aligned}&\sum_i G_k[(K_i\otimes\mathbb{1})\psi^{AB}(K_i^\dagger\otimes\mathbb{1})]\\ &=\sum_i G_k[(K_i\otimes\mathbb{1})\phi_k^+(K_i^\dagger\otimes\mathbb{1})]\times G_k(\psi^{AB}).\end{aligned} \tag{108}$$

As we prove in Lemma 11, for any quantum channel Λ which has no correctable subspace, the set of Kraus operators $\{K_i\}$ can always be chosen such that

$$\sum_i G_k[(K_i\otimes\mathbb{1})\phi_k^+(K_i^\dagger\otimes\mathbb{1})]<1. \tag{109}$$

Defining

$$\kappa_s = \sum_i G_k[(K_i\otimes\mathbb{1})\phi_k^+(K_i^\dagger\otimes\mathbb{1})] \tag{110}$$

we obtain

$$\sum_i G_k[(K_i\otimes\mathbb{1})\psi^{AB}(K_i^\dagger\otimes\mathbb{1})] = \kappa_sG_k(\psi^{AB}). \tag{111}$$

The proof of the proposition is then completed by defining probabilities p_i and pure states $|\phi_i\rangle^{AB}$ as follows:

$$p_i = \mathrm{Tr}\left[K_i\otimes\mathbb{1}\psi^{AB}K_i^\dagger\otimes\mathbb{1}\right], \tag{112}$$

$$|\phi_i\rangle^{AB} = \frac{1}{\sqrt{p_i}}K_i\otimes\mathbb{1}\,|\psi\rangle^{AB}, \tag{113}$$

and noting that

$$\sum_i G_k[(K_i\otimes\mathbb{1})\psi^{AB}(K_i^\dagger\otimes\mathbb{1})] = \sum_i p_iG_k(\phi_i^{AB}). \tag{114}$$

□

Lemma 11. *Any quantum channel Λ which has no correctable subspace has a set of Kraus operators $\{K_i\}$ such that that*

$$\kappa_s = \sum_i G_k[(K_i\otimes\mathbb{1})\phi_k^+(K_i^\dagger\otimes\mathbb{1})]<1 \tag{115}$$

for any $k\geq 2$.

Proof. We begin by showing that if Λ does not have a correctable subspace, then $E_f(\Lambda\otimes\mathbb{1}\,|\phi_k^+\rangle) < E_f(|\phi_k^+\rangle)$. We proceed by contradiction. Assume that there is no correctable subspace and that $E_f(\Lambda\otimes\mathbb{1}\,|\phi_k^+\rangle) = E_f(|\phi_k^+\rangle)$ for some k. Repeating the steps of the proof of Theorem 3 with ϕ_k^+ in place of ψ^{AB}, we obtain that a correctable subspace must exist. This contradicts the assumption that no correctable subspace exists. Therefore, $E_f(\Lambda\otimes\mathbb{1}\,|\phi_k^+\rangle) < E_f(|\phi_k^+\rangle)$.

Let $\rho=\Lambda\otimes\mathbb{1}(|\phi_k^+\rangle)$. Let us denote by $\{q_i,|\psi_i\rangle\}$ the ensemble decomposition of ρ that minimizes $\sum_i q_iE_f(|\psi_i\rangle)$. There exist Kraus operators $\{L_i\}$ such that $L_i\otimes\mathbb{1}\,|\phi_k^+\rangle\langle\phi_k^+|\,L_i^\dagger\otimes\mathbb{1} = q_i\,|\psi_i\rangle\langle\psi_i|$. Additionally, we have $\sum_i q_iE_f(|\psi_i\rangle) < E_f(|\phi_k^+\rangle)$. Consequently, for at least one i there is inequality $E_f(|\psi_i\rangle) <$

$E_f(|\phi_k^+\rangle)$. Both the entanglement of formation and the G_k-concurrence reach their global maximum for pure states if and only if all their Schmidt coefficients are equal. Therefore, any pure state $|\psi_i\rangle$ that is not maximally entangled (i.e. satisfying $E_f(|\psi_i\rangle) < E_f(|\phi_k^+\rangle)$) must necessarily satisfy $G_k(|\psi_i\rangle) < G_k(|\phi_k^+\rangle)$. Consequently, since for at least one i there is relation $G_k(|\psi_i\rangle) < G_k(|\phi_k^+\rangle)$, the following inequality must hold

$$\sum_i q_i G_k(|\psi_i\rangle) < G_k(|\phi_k^+\rangle) = 1 \tag{116}$$

This concludes the proof. □

Consider now a mixed state ρ^{AB} which can be written as $\rho^{AB} = \sum_j p_j \psi_j^{AB}$ with each of the pure states $|\psi_j\rangle^{AB}$ having Schmidt rank at most k. Using Proposition 10, each of the states $\Lambda \otimes \mathbb{1}[\psi_j^{AB}]$ can be expressed as

$$\Lambda \otimes \mathbb{1}[\psi_j^{AB}] = \sum_i q_i \psi_{ij}^{AB}, \tag{117}$$

where the pure states $|\psi_{ij}\rangle^{AB}$ have Schmidt rank at most k and

$$\sum_i q_i G_k(\psi_{ij}^{AB}) = \kappa_s G_k(\psi_j^{AB}). \tag{118}$$

This means that state $\Lambda \otimes \mathbb{1}[\rho^{AB}]$ can be written as a convex combination of states $|\psi_{ij}\rangle^{AB}$ with probabilities $q_i p_j$ such that

$$\sum_{ij} q_i p_j G_k(\psi_{ij}^{AB}) = \kappa_s \left(\sum_j p_j G_k(\psi_j^{AB}) \right). \tag{119}$$

With these results, we can now prove the following proposition.

Proposition 12. *Let Λ_n be defined as in Eq. (2) and let ρ^{AB} be a mixed state with Schmidt rank at most k. Then, $\Lambda_n[\rho^{AB}] = \sum_i s_i \phi_i^{AB}$, where each ϕ_i^{AB} has Schmidt rank at most k and $\sum_i s_i G_k(\phi_i^{AB}) \leq \kappa_s^n$, where κ_s is defined in Eq. (110).*

Proof. We will prove this proposition by induction. First, we show that the above conjecture holds for $n = 1$. Let $\{K_i\}$ be a set of Kraus operators for Λ and $\{M_j^i \otimes N_k^i\}$ be a set of Kraus operators for F_i. Let $\rho^{AB} = \sum_j q_j |\phi_j^{AB}\rangle\langle\phi_j^{AB}|$, where each $|\phi_j^{AB}\rangle$ has Schmidt rank at most k. For $n = 1$ we have:

$$\Lambda_1[\rho^{AB}] = (\Lambda \otimes \mathbb{I})[F_1 \circ \rho^{AB}] = \sum_{ijlg} r_{lgi} q_j \phi_{ijgl}^{AB}, \tag{120}$$

such that

$$|\phi_{lgij}\rangle^{AB} = \frac{1}{\sqrt{r_{lgi}}} (K_i \otimes \mathbb{I})(M_l \otimes N_g) \, |\phi_j\rangle^{AB}, \tag{121}$$

$$r_{lgi} = \mathrm{Tr}[(K_i \otimes \mathbb{I})(M_l \otimes N_g) \, \phi_j^{AB} (M_l \otimes N_g)^\dagger (K_i \otimes \mathbb{I})^\dagger]. \tag{122}$$

Note that each $|\phi_{lgij}\rangle^{AB}$ cannot have higher Schmidt rank than $|\phi_j^{AB}\rangle$. Defining new indices h as $h = ijgl$ we obtain:

$$\Lambda_1(\rho^{AB}) = \sum_h p_h \phi_h^{AB}, \tag{123}$$

where each ϕ_h^{AB} has Schmidt rank at most k.

We have shown the first part of the Proposition for $n = 1$. We proceed to the second part for $n = 1$.

$$\sum_{ijgl} q_i \, r_{lgi} \, G_k(\phi_{lgij}^{AB}) = \sum_{ijgl} q_i \, G_k((K_i \otimes \mathbb{I})(M_l \otimes N_g) \, \phi_{ij}^{AB} (M_l \otimes N_g)^\dagger (K_i \otimes \mathbb{I})^\dagger). \tag{124}$$

Now, we recall that, see Lemma 9

$$G_k((K_i \otimes \mathbb{I})(M_l \otimes N_g) \, \phi_{ij}^{AB} (M_l \otimes N_g)^\dagger (K_i \otimes \mathbb{I})^\dagger) = G_k((K_i \otimes \mathbb{I})\phi_k^+ (K_i \otimes \mathbb{I})^\dagger) \, G_k((M_l \otimes N_g)\phi_k^+ (M_l \otimes N_g)^\dagger) \times G_k(\phi_{ij}^{AB}). \tag{125}$$

Using Eqs. (110), (124) and (125) we get

$$\sum_{ijgl} q_j \, r_{lgi} \, G_k(\phi_{ijgl}^{AB}) \leq \sum_{ijh} q_j G_k((K_i \otimes \mathbb{I})\phi_k^+ (K_i \otimes \mathbb{I})^\dagger) \, G_k(\phi_{ij}^{AB}) \leq \kappa_s \left(\sum_{ij} q_i G_k(\phi_{ij}^{AB}) \right) \leq \kappa_s. \tag{126}$$

In the first inequality we used that $G_k((M_l \otimes N_g)\phi_k^+ (M_l \otimes N_g)^\dagger) \leq 1$.

We have showed that our conjecture holds for $n = 1$. Let us assume it holds for some n. Let $\Lambda_n(\rho^{AB}) = \sum_i s_i \Xi_i^{AB}$, where each Ξ_i^{AB} has Schmidt rank at most k because of the induction assumption. We can write:

$$\Lambda_{n+1}(\rho^{AB}) = \Lambda \circ F_{n+1} \circ \Lambda_n(\rho^{AB}) = \Lambda \circ F_{n+1}(\sum_i s_i \Xi_i^{AB}) = \sum_{ighl} s_i r_{ghl} \phi_{ihgl}^{AB}, \tag{127}$$

where

$$|\phi_{hikl}\rangle^{AB} = \frac{1}{\sqrt{r_h r'_{lg}}} (K_h \otimes \mathbb{I})(M_l \otimes N_g) \, |\Xi_i\rangle^{AB}, \tag{128}$$

$$r_{hgl} = \mathrm{Tr}[(K_h \otimes \mathbb{I})(M_l \otimes N_g) \, \Xi_i^{AB} (M_l \otimes N_k)^\dagger (K_h \otimes \mathbb{I})^\dagger], \tag{129}$$

where each $|\phi_{hikl}\rangle^{AB}$ has Schmidt rank not greater than k. This completes the first part of the Proposition. Using the induction assumption and Eq. (110) , we obtain:

$$\sum_{ighl} s_i r_{ghl} G_k(\phi_{ihgl}^{AB}) = \sum_{ighl} s_i G_k((K_h \otimes \mathbb{I})\phi_k^+ (K_h \otimes \mathbb{I})^\dagger) \, G_k((M_l \otimes N_g)\phi_k^+ (M_l \otimes N_g)^\dagger) \times G_k(\Xi_i^{AB}) \leq \sum_h G_k((K_h \otimes \mathbb{I})\phi_k^+ (K_h \otimes \mathbb{I})^\dagger)(\sum s_i G_k(\Xi_i^{AB})) \leq \kappa_s^{n+1}. \tag{130}$$

In the last step, we used the induction assumption and Eq. (110). It completes the proof. □

Now, let us define the k-Schmidt fidelity of a bipartite pure state $|\psi\rangle^{AB}$ as:

$$F_k(\psi^{AB}) = \max_{\phi \in \mathcal{S}_k} F(\psi^{AB}, \phi^{AB}). \tag{131}$$

where $\mathcal{S}_k$ denotes the set of states that can be expressed as convex combinations of bipartite pure states with Schmidt rank at most k. Since $|\psi^{AB}\rangle$ is pure, the maximum in (131) is attained by a pure state.

Proposition 13. *For a bipartite pure state $|\psi\rangle^{AB}$, the following holds:*

$$F_k(\psi^{AB}) = \sum_{i=0}^{k-1} \lambda_i, \tag{132}$$

where λ_i are Schmidt coefficients of $|\psi\rangle$ in decreasing order. Additionally, the state $|\phi\rangle^{AB} \in \mathcal{S}_k$ that achieves the maximum in Eq. (131) is given by

$$|\phi\rangle^{AB} = \frac{1}{\sqrt{\sum_{j=0}^{k-1} \lambda_j}} \sum_{i=0}^{k-1} \sqrt{\lambda_i}\, |ii\rangle^{AB} \tag{133}$$

where $|i\rangle$ are the Schmidt vectors of $|\psi\rangle$.

Proof. Let $|\psi\rangle = \sum_{i,j} M_{ij} |i\rangle |j\rangle$ and $|\phi\rangle = \sum_{i,j} N_{ij} |i\rangle |j\rangle$. Their inner product can be written as

$$|\langle\phi|\psi\rangle| = \left| \sum_{i,j} N^*_{ij} M_{ij} \right| = \left| \mathrm{Tr}\left(N^\dagger M\right) \right|. \tag{134}$$

Using the von Neumann trace inequality [48], we obtain:

$$\left| \mathrm{Tr}\left(N^\dagger M\right) \right| \leq \sum_{i=0}^{d-1} s_i(N) s_i(M) = \sum_{i=0}^{k-1} s_i(N) \sqrt{\lambda_i}, \tag{135}$$

where $s_i(X)$ denote the singular values of X in non-increasing order. In the last step, we used that the singular values of M are the Schmidt coefficients of $|\psi\rangle$.

Since $|\phi\rangle$ has Schmidt rank k, the matrix N has rank at most k, and hence only the first k singular values are non-zero. Applying the Cauchy–Schwarz inequality, we obtain

$$\sum_{i=0}^{k-1} s_i(N) \sqrt{\lambda_i} \leq \sqrt{\sum_{i=0}^{k-1} s_i(N)^2} \sqrt{\sum_{i=0}^{k-1} \lambda_i} = \sqrt{\sum_{i=0}^{k-1} \lambda_i}. \tag{136}$$

In the last step we used that squares of singular values of N must add up to 1. Thus, we have the inequality $|\langle\phi|\psi\rangle| \leq \sqrt{\sum_{i=0}^{k-1} \lambda_i}$. Taking squares we obtain:

$$|\langle\phi|\psi\rangle|^2 \leq \sum_{i=0}^{k-1} \lambda_i. \tag{137}$$

We have thus established an upper bound on the fidelity. One can verify that the bound is saturated by the state given in Eq. (133), which completes the proof. □

Analogous to the definition of k-Schmidt fidelity for pure states, we define k-Schmidt fidelity for a bipartite mixed state ρ^{AB} as follows:

$$F_k(\rho^{AB}) = \max_{\widetilde{\sigma}^{AB} \in \mathcal{S}_k} F(\rho^{AB}, \widetilde{\sigma}^{AB}). \tag{138}$$

Proposition 14. *For any pure-state decomposition $\{p_i, \psi_i^{AB}\}$ of ρ^{AB} it holds that*

$$F_k(\rho^{AB}) \geq \sum_i p_i F_k(\psi_i^{AB}). \tag{139}$$

Proof. For every pure state $|\psi_i^{AB}\rangle$ appearing in a decomposition of ρ, there exists a Schmidt-rank-at-most-k state $|\phi_i^{AB}\rangle$ that optimally achieves the fidelity, i.e.,

$$F_k(\psi_i^{AB}) = |\langle\psi_i^{AB}|\phi_i^{AB}\rangle|^2. \tag{140}$$

We can always set phases such that inner product $\langle\psi_i^{AB}|\phi_i^{AB}\rangle$ is non-negative. We obtain:

$$\sqrt{F_k(\psi_i^{AB})} = \langle\psi_i^{AB}|\phi_i^{AB}\rangle\,. \tag{141}$$

Let us define new probabilities q_i such that:

$$q_i = \frac{p_i F_k(\psi_i^{AB})}{T}, \tag{142}$$

where $T = \sum_j p_j F_k(\psi_j^{AB})$ is a normalization constant. Let us define:

$$\sigma^{AB} = \sum_i q_i\, |\phi_i^{AB}\rangle\langle\phi_i^{AB}|\,. \tag{143}$$

Since every state $|\phi_i^{AB}\rangle$ has Schmidt rank at most k, $\sigma^{AB} \in S_k$.

Let us purify σ^{AB} and ρ^{AB}. We obtain:

$$\begin{aligned} |\Psi^\rho\rangle &= \sum_i \sqrt{p_i}\, |i\rangle^R |\psi_i\rangle^{AB} \\ |\Phi^\sigma\rangle &= \sum_i \sqrt{q_i}\, |i\rangle^R |\phi_i\rangle^{AB}\,. \end{aligned} \tag{144}$$

By Uhlmann's theorem we obtain [49]

$$F(\rho^{AB}, \sigma^{AB}) = \max_{U^R} |\langle\Phi^\sigma|\, U^R \otimes \mathbb{1}^{AB}\, |\Psi^\rho\rangle|^2, \tag{145}$$

where maximization is taken over all unitaries U^R on purifying system. When we take specific $U^R = \mathbb{1}^R$ we obtain:

$$F(\rho^{AB}, \sigma^{AB}) \geq |\langle\Phi^\sigma|\Psi^\rho\rangle|^2. \tag{146}$$

We can further write:

$$\langle\Phi^\sigma|\Psi^\rho\rangle = \sum_i \sqrt{p_i q_i}\, \langle\psi_i^{AB}|\phi_i^{AB}\rangle \tag{147}$$

Using Eq. (141) and (142) we obtain:

$$\begin{aligned} \langle\Phi^\sigma|\Psi^\rho\rangle = \sum_i \sqrt{p_i \frac{p_i F_k(\psi_i^{AB})}{T}} \sqrt{F_k(\psi_i^{AB})} = \\ \frac{\sum_i p_i F_k(\psi_i^{AB})}{\sqrt{\sum_i p_i F_k(\psi_i^{AB})}} = \sqrt{\sum_i p_i F_k(\psi_i^{AB})}. \end{aligned} \tag{148}$$

Using relations (146) and (148) we obtain:

$$F(\rho^{AB},\sigma^{AB}) \geq \sum_i p_i F_k(\psi_i^{AB}) \tag{149}$$

Since in Eq. (143) we constructed specific σ^{AB} there is a relation:

$$\max_{\widetilde{\sigma}^{AB}\in\mathcal{S}_k} F(\rho^{AB},\widetilde{\sigma}^{AB}) \geq F(\rho^{AB},\sigma^{AB}) \geq \sum_i p_i F_k(\psi_i^{AB}), \tag{150}$$

which completes the proof. Similar technique was used in [50]. □

For a pure state $|\psi\rangle$ with Schmidt rank at most k, we further have

$$\lambda_{k-1} \leq \left(\prod_{i=0}^{k-1}\lambda_i\right)^{1/k} = \frac{G_k(\psi)}{k}. \tag{151}$$

Using Proposition 13 we arrive at

$$F_{k-1}(\psi) = 1-\lambda_{k-1} \geq 1-\frac{G_k(\psi)}{k}. \tag{152}$$

Using now Propositions 12 and 14, it follows that the state $\Lambda_n[\rho^{AB}]$ can be written as $\Lambda_n[\rho^{AB}] = \sum_i s_i\phi_i^{AB}$ with probabilities s_i and pure states ϕ_i^{AB} having Schmidt rank at most k such that

$$\begin{aligned} F_{k-1}\left(\Lambda_n[\rho^{AB}]\right) &\geq \sum_i s_i F_{k-1}\left(\phi_i^{AB}\right) \\ &\geq 1-\frac{1}{k}\sum_i s_i G_k\left(\phi_i^{AB}\right) \\ &\geq 1-\frac{\kappa_s^n}{k}. \end{aligned} \tag{153}$$

These arguments together with the Fuchs–van de Graaf inequality [51] imply that we can bound the trace-distance to the set $\mathcal{S}_{k-1}$ as follows:

$$\begin{aligned} \min_{\sigma^{AB}\in\mathcal{S}_{k-1}} \|\Lambda_n[\rho^{AB}]-\sigma^{AB}\|_1 &\leq 2\sqrt{1-F_{k-1}\left(\Lambda_n[\rho^{AB}]\right)} \\ &\leq 2\frac{\kappa_s^{n/2}}{\sqrt{k}}. \end{aligned} \tag{154}$$

Note now that the overall protocol Λ_n can always be decomposed into two protocols $\Lambda_{n_2}\circ\Lambda_{n_1}$ with $n = n_1+n_2$. We now define the state $\nu_{k-1}^{AB}\in\mathcal{S}_{k-1}$ to be the closest state to $\Lambda_{n_1}[\rho^{AB}]$. We further define $\nu_{k-2}\in\mathcal{S}_{k-2}$ to be the closest state to $\Lambda_{n_2}[\nu_{k-1}^{AB}]$. Using triangle inequality and data processing inequality we find

$$\begin{aligned} \|\Lambda_{n_2}\circ\Lambda_{n_1}[\rho^{AB}]-\nu_{k-2}^{AB}\|_1 &\leq \|\Lambda_{n_2}\circ\Lambda_{n_1}[\rho^{AB}]-\Lambda_{n_2}[\nu_{k-1}^{AB}]\|_1 \\ &\quad+\|\Lambda_{n_2}[\nu_{k-1}^{AB}]-\nu_{k-2}^{AB}\|_1 \\ &\leq \|\Lambda_{n_1}[\rho^{AB}]-\nu_{k-1}^{AB}\|_1 \\ &\quad+\|\Lambda_{n_2}[\nu_{k-1}^{AB}]-\nu_{k-2}^{AB}\|_1 \\ &\leq 2\left(\frac{\kappa_s^{n_1/2}}{\sqrt{k}}+\frac{\kappa_s^{n_2/2}}{\sqrt{k-1}}\right). \end{aligned} \tag{155}$$

Note that we can decompose the overall protocol Λ_n into $k-1$ smaller protocols: $\Lambda_{n_{k-1}}\circ\Lambda_{n_{k-2}}\circ\cdots\circ\Lambda_{n_1}$ with $n=\sum_{i=1}^{k-1}n_i$. As before, we define the state $\nu_{k-1}^{AB}\in\mathcal{S}_{k-1}$ to be the closest state to $\Lambda_{n_1}[\rho^{AB}]$. We then define an analogous state for each n_i in the following fashion: let $\nu_{k-i}^{AB}\in\mathcal{S}_{k-i}$ be the state closest to $\Lambda_{n_i}[\nu_{k-i+1}^{AB}]$. We obtain

$$\begin{aligned} &\|\Lambda_{n_{k-1}}\circ\cdots\circ\Lambda_{n_2}\circ\Lambda_{n_1}[\rho^{AB}]-\nu_1^{AB}\|_1 \\ &\leq \|\Lambda_{n_{k-1}}\circ\cdots\circ\Lambda_{n_2}\circ\Lambda_{n_1}[\rho^{AB}]-\Lambda_{n_{k-1}}\circ\cdots\circ\Lambda_{n_2}[\nu_{k-1}^{AB}]\|_1 \\ &\quad+\|\Lambda_{n_{k-1}}\circ\cdots\circ\Lambda_{n_2}[\nu_{k-1}^{AB}]-\nu_1^{AB}\|_1. \end{aligned} \tag{156}$$

Then using data processing [43], we get

$$\begin{aligned} &\|\Lambda_{n_{k-1}}\circ\cdots\circ\Lambda_{n_2}\circ\Lambda_{n_1}[\rho^{AB}]-\nu_1^{AB}\|_1 \\ &\leq \|\Lambda_{n_1}[\rho^{AB}]-[\nu_{k-1}^{AB}]\|_1 \\ &\quad+\|\Lambda_{n_{k-1}}\circ\cdots\circ\Lambda_{n_2}[\nu_{k-1}^{AB}]-\nu_1^{AB}\|_1. \end{aligned} \tag{157}$$

Using the triangle inequality once more, we get

$$\begin{aligned} &\|\Lambda_{n_{k-1}}\circ\cdots\circ\Lambda_{n_2}\circ\Lambda_{n_1}[\rho^{AB}]-\nu_1^{AB}\|_1 \\ &\leq \|\Lambda_{n_1}[\rho^{AB}]-[\nu_{k-1}^{AB}]\|_1 \\ &+\|\Lambda_{n_{k-1}}\circ\cdots\circ\Lambda_{n_3}\circ\Lambda_{n_2}[\rho^{AB}]-\Lambda_{n_{k-1}}\circ\cdots\circ\Lambda_{n_3}[\nu_{k-2}^{AB}]\|_1 \\ &\quad+\|\Lambda_{n_{k-1}}\circ\cdots\circ\Lambda_{n_3}[\nu_{k-2}^{AB}]-\nu_1^{AB}\|_1. \end{aligned} \tag{158}$$

Then using data processing, we arrive at

$$\begin{aligned} &\|\Lambda_{n_{k-1}}\circ\cdots\circ\Lambda_{n_2}\circ\Lambda_{n_1}[\rho^{AB}]-\nu_1^{AB}\|_1 \\ &\leq \|\Lambda_{n_1}[\rho^{AB}]-[\nu_{k-1}^{AB}]\|_1+\|\Lambda_{n_2}[\nu_{k-1}^{AB}]-[\nu_{k-2}^{AB}]\|_1 \\ &\quad+\|\Lambda_{n_{k-1}}\circ\cdots\circ\Lambda_{n_3}[\nu_{k-2}^{AB}]-\nu_1^{AB}\|_1. \end{aligned} \tag{159}$$

Iterating this procedure, we obtain the inequality

$$\begin{aligned} &\|\Lambda_{n_{k-1}}\circ\cdots\circ\Lambda_{n_2}\circ\Lambda_{n_1}[\rho^{AB}]-\nu_1^{AB}\|_1 \\ &\leq \|\Lambda_{n_1}[\rho^{AB}]-\nu_{k-1}^{AB}\|_1+\sum_{i=2}^{k-1}\|\Lambda_{n_i}[\nu_{k-i+1}^{AB}]-\nu_{k-i}^{AB}\|_1. \end{aligned} \tag{160}$$

Using (154), we get

$$\min_{\sigma^{AB}\in\mathcal{S}_1}\|\Lambda_{n_{k-1}}\circ\cdots\circ\Lambda_{n_2}\circ\Lambda_{n_1}[\rho^{AB}]-\sigma^{AB}\|_1 \leq 2\sum_{i=1}^{k-1}\frac{\kappa_s^{n_i/2}}{\sqrt{k-i+1}}. \tag{161}$$

Choosing $n_i = N$ and $n = N(k-1)$ we obtain

$$\begin{aligned} \sum_{i=1}^{k-1}\frac{\kappa_s^{n_i/2}}{\sqrt{k-i+1}} &= \kappa_s^{N/2}\sum_{i=1}^{k-1}\frac{1}{\sqrt{k-i+1}} \leq 2\kappa_s^{N/2}(\sqrt{k}-1) \\ &= 2\kappa_s^{n/[2(k-1)]}(\sqrt{k}-1) \leq 2\kappa_s^{n/[2(d_A-1)]}(\sqrt{d_A}-1). \end{aligned} \tag{162}$$

Collecting the above arguments, we thus have

$$\min_{\sigma^{AB}\in\mathcal{S}_1}\|\Lambda_n[\rho^{AB}]-\sigma^{AB}\|_1 \leq 4\kappa_s^{n/[2(k-1)]}(\sqrt{k}-1). \tag{163}$$

Noting that $k \le d_A$, we obtain

$$\min_{\sigma^{AB}\in\mathcal{S}_1} \|\Lambda_n[\rho^{AB}] - \sigma^{AB}\|_1 \le 4\kappa_s^{n/[2(d_A-1)]}(\sqrt{d_A}-1). \quad (164)$$

Taking the optimal Kraus representation of Λ we obtain:

$$\min_{\sigma^{AB}\in\mathcal{S}_1} \|\Lambda_n[\rho^{AB}] - \sigma^{AB}\|_1 \le 4\kappa^{n/[2(d_A-1)]}(\sqrt{d_A}-1). \quad (165)$$

This completes the proof of the theorem.

Proof of Theorem 5.

First part: Assume that the channel Λ admits the decomposition

$$\Lambda = p\mathcal{E} + (1-p)\mathcal{M}, \qquad 0 < p < 1, \quad (166)$$

where $\mathcal{E}$ is an entanglement-breaking channel and $\mathcal{M}$ is an arbitrary quantum channel.

We will show that, for m parallel uses of Λ, the quantity κ introduced in Eq. (163) satisfies

$$\kappa \le 1 - p^m \quad (167)$$

for every $k \ge 2$. The bound is uniform with respect to the choice of the k-dimensional Schmidt support.

Every entanglement-breaking channel admits a Kraus representation consisting of rank-one operators[37]. We may therefore write

$$\mathcal{E}(\rho) = \sum_a A_a \rho A_a^\dagger, \qquad \mathrm{rank}(A_a) = 1, \qquad \sum_a A_a^\dagger A_a = \mathbb{1}. \quad (168)$$

Let

$$\mathcal{M}(\rho) = \sum_b B_b \rho B_b^\dagger, \qquad \sum_b B_b^\dagger B_b = \mathbb{1} \quad (169)$$

be an arbitrary Kraus representation of $\mathcal{M}$. Consequently, a Kraus representation of Λ is given by

$$\{\sqrt{p}\,A_a\}_a \cup \left\{\sqrt{1-p}\,B_b\right\}_b. \quad (170)$$

Consider now the channel $\Lambda^{\otimes m}$. Its Kraus operators are tensor products of the single-system Kraus operators. In particular, the subset of Kraus operators for which all m factors originate from the entanglement-breaking part is given by

$$R_{\boldsymbol{a}} = p^{m/2} A_{a_1} \otimes \cdots \otimes A_{a_m}, \qquad \boldsymbol{a} = (a_1, \ldots, a_m). \quad (171)$$

Since every A_{a_i} has rank one, we have

$$\mathrm{rank}(R_{\boldsymbol{a}}) = 1. \quad (172)$$

Let $|\phi_k^+\rangle$ be a maximally entangled state of Schmidt rank k, whose Schmidt support on the transmitted subsystem is an arbitrary k-dimensional subspace. For every $\boldsymbol{a}$, the state

$$(R_{\boldsymbol{a}} \otimes \mathbb{1})\,|\phi_k^+\rangle \quad (173)$$

has Schmidt rank at most one. Hence, for every $k \ge 2$,

$$G_k\Big[(R_{\boldsymbol{a}} \otimes \mathbb{1})\phi_k^+(R_{\boldsymbol{a}}^\dagger \otimes \mathbb{1})\Big] = 0. \quad (174)$$

Let

$$q_{\boldsymbol{a}} = \mathrm{Tr}\Big[(R_{\boldsymbol{a}} \otimes \mathbb{1})\phi_k^+(R_{\boldsymbol{a}}^\dagger \otimes \mathbb{1})\Big] \quad (175)$$

denote the probability of the corresponding Kraus outcome. The total probability of all outcomes containing only Kraus operators from the entanglement-breaking part is

$$\sum_{\boldsymbol{a}} q_{\boldsymbol{a}} = \mathrm{Tr}\left[\left(\sum_{\boldsymbol{a}} R_{\boldsymbol{a}}^\dagger R_{\boldsymbol{a}} \otimes \mathbb{1}\right)\phi_k^+\right] \quad (176)$$

$$= \mathrm{Tr}\left[\left(p^m\left(\sum_a A_a^\dagger A_a\right)^{\otimes m} \otimes \mathbb{1}\right)\phi_k^+\right] \quad (177)$$

$$= p^m. \quad (178)$$

Let L_a be a fixed Kraus operators of $\Lambda^{\otimes m}$. Since L_a is a fixed Kraus representation, we obtain:

$$\kappa \le \sum_a G_k(L_a \otimes \mathbb{1})\phi_k^+(L_a^\dagger \otimes \mathbb{1})) \quad (179)$$

We now evaluate the average G_k-concurrence associated with this particular Kraus decomposition. Let $\mathcal{I}_{\mathrm{EB}}$ denote the set of outcomes in which all factors originate from $\mathcal{E}$, and let $\mathcal{I}_{\mathrm{rest}}$ denote the remaining outcomes. For each outcome α, let q_α be its probability and let θ_α be the corresponding normalized state. Since

$$0 \le G_k(\theta_\alpha) \le 1, \quad (180)$$

we obtain

$$\kappa \le \sum_\alpha q_\alpha G_k(\theta_\alpha) = \sum_{\alpha\in\mathcal{I}_{\mathrm{EB}}} q_\alpha G_k(\theta_\alpha) + \sum_{\alpha\in\mathcal{I}_{\mathrm{rest}}} q_\alpha G_k(\theta_\alpha)$$

$$= \sum_{\alpha\in\mathcal{I}_{\mathrm{rest}}} q_\alpha G_k(\theta_\alpha) \le \sum_{\alpha\in\mathcal{I}_{\mathrm{rest}}} q_\alpha = 1 - p^m. \quad (181)$$

Introducing

$$\gamma = -\ln p > 0, \quad (182)$$

we may equivalently write

$$\kappa \le 1 - e^{-\gamma m} \quad (183)$$

On the other hand, Eq. (163), applied to the channel $\Lambda^{\otimes m}$, gives

$$\min_{\sigma^{AB}\in\mathcal{S}} \left\|\rho_{n,m(n)}^{AB} - \sigma^{AB}\right\|_1 \le 4\kappa^{n/(2(k_m-1))}\left(\sqrt{k_m}-1\right), \quad (184)$$

where k_m is the relevant Schmidt-rank parameter.

It is possible to establish entanglement using a channel if

$$\liminf_{n\to\infty} E_f\left(\rho_{n,m(n)}^{AB}\right) > 0. \quad (185)$$

Using lower bound on distance to separable states from Lemma 15 we obtain:

$$\frac{E_0^2}{8m^2(\log_2 d)^2} \leq 2\left(1-e^{-\gamma m}\right)^{\frac{n}{2(k_m-1)}}\left(\sqrt{k_m}-1\right). \tag{186}$$

Equivalently,

$$\left(1-e^{-\gamma m}\right)^{\frac{n}{2(k_m-1)}} \geq \frac{E_0^2}{16m^2(\log_2 d)^2\left(\sqrt{k_m}-1\right)}. \tag{187}$$

Taking the natural logarithm gives

$$\frac{n}{2(k_m-1)}\ln(1-e^{-\gamma m}) \geq -\ln\left[\frac{16m^2(\log_2 d)^2\left(\sqrt{k_m}-1\right)}{E_0^2}\right]. \tag{188}$$

Using

$$\ln(1-x) \leq -x, \qquad 0<x<1, \tag{189}$$

we arrive at

$$\frac{ne^{-\gamma m}}{2(k_m-1)} \leq \ln\left[\frac{16m^2(\log_2 d)^2\left(\sqrt{k_m}-1\right)}{E_0^2}\right]. \tag{190}$$

Consequently,

$$n \leq 2(k_m-1)e^{\gamma m}\ln\left[\frac{16m^2(\log_2 d)^2\left(\sqrt{k_m}-1\right)}{E_0^2}\right]. \tag{191}$$

For a block of m qudits,

$$k_m \leq d^m. \tag{192}$$

It follows that

$$\begin{aligned} n &\leq 2d^m e^{\gamma m}\left[\frac{m}{2}\ln d + 2\ln m + C_0\right] \\ &\leq C_1 m e^{(\gamma+\ln d)m} \end{aligned} \tag{193}$$

for all sufficiently large m, where C_0 and C_1 are positive constants independent of n and m. We used the fact that $m+\ln m < Cm$ for sufficiently large m.

Therefore,

$$m \geq \frac{1}{\gamma+\ln d}W_0\left(\frac{\gamma+\ln d}{C_1}n\right), \tag{194}$$

where W_0 is the principal branch of the Lambert function. Using[52]

$$W_0(x) = \ln x - \ln\ln x + O\left(\frac{\ln\ln x}{\ln x}\right), \tag{195}$$

we obtain

$$m \geq \frac{\ln n - \ln\ln n}{\gamma+\ln d} - O(1) \tag{196}$$

Thus, preserving a nonvanishing amount of entanglement of formation requires at least a logarithmic number of parallel channel uses per link.

Second part:

In the beginning we define a notion that will be used in the proof. Let $\mathcal{P}$ be a projector into a Schmidt support of $|\phi_k^+\rangle^{AB} = \sum_i \sqrt{\frac{1}{k}}|e_i\rangle^A|f_i\rangle^B$. That is $\mathcal{P}$ projects onto span$(|e_1\rangle, \cdots, |e_k\rangle)$. Let $\mathcal{V}$ be isometry $\mathbb{C}^k \to \mathcal{H}_{\mathcal{A}}$ defined by $\mathcal{V}|j\rangle^K = |e_j\rangle^A$, where superscripts highlight that $|j\rangle^K \in \mathbb{C}^k$ and $|e\rangle_j^A \in \mathcal{H}_A$. Operator $\mathcal{V}^\dagger X\mathcal{V}$ is a $k\times k$ matrix on $\mathcal{C}^k$ with matrix elements $(\mathcal{V}^\dagger X\mathcal{V})_{ij} = \langle e_i|X|e_j\rangle$. We define $\det_{\mathcal{P}} X = det(V^\dagger XV)$. Consequently, $\det_{\mathcal{P}} X$ is the product of the eigenvalues of the compression $\mathcal{P}X\mathcal{P}$, regarded as an operator on the Schmidt support $\mathcal{P}\mathcal{H}_A$.

Let us consider a qudit amplitude damping channel $\Lambda_{AD}^{\otimes m}$. Kraus operators of a single qudit amplitude damping channel Λ_{AD} are:

$$\begin{aligned} K_0 &= \sum_{j=0}^{d-1}\sqrt{q_j}\,|j\rangle\langle j| \\ K_{i\leftarrow j} &= \sqrt{\gamma_{i\leftarrow j}}\,|i\rangle\langle j|, \end{aligned} \tag{197}$$

where $\eta_j = \sum_{i<j}\gamma_{i\leftarrow j}$ is total decay probability of the j-th level and $q_j = 1-\eta_j$ is no decay probability. This kind of amplitude damping channel was studied in [53].

Assume that we have a Hilbert space with bipartition $A|B = A_1\cdots A_m|B$. Let us consider states with Schmidt rank at most k. Let us define a maximally entangled state on this Hilbert space $|\phi_k^+\rangle = \frac{1}{\sqrt{k}}\sum_{i=0}^{k-1}|e\rangle_i|f_i\rangle$. Let $\mathcal{P}$ be the projector onto the Schmidt support of $|\phi_k^+\rangle$, i.e. space spanned by span$(|e\rangle_1, \cdots |e\rangle_k)$.

To upper bound κ we need to find:

$$\kappa \leq \sum_i G_k\left(L_i\otimes\mathbb{1}\phi_k^+L_i^\dagger\otimes\mathbb{1}\right), \tag{198}$$

where $L_i = K_{\alpha_1}\otimes\cdots\otimes K_{\alpha_m}$ is the standard Kraus representation of $\Lambda_{AD}^{\otimes m}$. Let $E_l = L_l^\dagger L_l$. We are interested only by the action of L_l on the Schmidt support. We can compute it by $L_l\mathcal{P}$. We have $(L_l\mathcal{P})^\dagger L_l\mathcal{P} = \mathcal{P}E_l\mathcal{P}$ As discussed in Lemma 9, we can express G_k-concurrence of a single Kraus branch as:

$$G_k(L_l\otimes\mathbb{1}\phi_k^+L_l^\dagger\otimes\mathbb{1}) = \left(\prod_{i=1}^k s_i^2(L_l\mathcal{P})\right)^{1/k} = \det_{\mathcal{P}}(\mathcal{P}E_l\mathcal{P})^{1/k}, \tag{199}$$

where $s_i(L_lP)$ are singular values of $L_l\mathcal{P}$. For fixed $\mathcal{P}$ we have:

$$\kappa \leq \sum_l \det_{\mathcal{P}}(\mathcal{P}E_l\mathcal{P})^{1/k}. \tag{200}$$

We define:

$$X_l = \mathcal{P}E_l\mathcal{P}. \tag{201}$$

Since $\sum_l E_l = \mathbb{1}$, we obtain

$$\sum_l X_l = \mathcal{P} \tag{202}$$

Now we proceed to find a bound on Eq. (200). We do it using the eigenvalues $\{x_{l,j}\}_{j=1}^{k}$ of the operator X_l. The arithmetic mean of $x_{l,j}$ is $a_l = \frac{1}{k}\mathrm{Tr}(X_l)$,, while their geometric mean is $\det_{\mathcal{P}}(X_l)^{1/k}$. Using Cartwright-Field inequality [54], we know that:

$$a_l - \det_{\mathcal{P}}(X_l)^{1/k} \geq \frac{1}{2kR_l}\sum_{j=1}^{k}(x_{l,j}-a_l)^2, \tag{203}$$

where $R_l = \max\{x_{l,j}\}_{j=1}^{k}$. Eigenvalues of X_l are bounded by 1. Therefore, we obtain a bound $R_l \leq 1$.

On the subspace $\mathcal{P}\mathcal{H}_A$, the projector $\mathcal{P}$ acts as the identity. Therefore,

$$a_l - [\det_{\mathcal{P}}(X_l)]^{1/k} \geq \frac{1}{2k}\left\|X_l - a_l\mathcal{P}\right\|_2^2. \tag{204}$$

Summing over all Kraus outcomes l, we obtain

$$\sum_l \left(a_l - [\det_{\mathcal{P}}(X_l)]^{1/k}\right) \geq \frac{1}{2k}\sum_l \left\|X_l - a_l\mathcal{P}\right\|_2^2. \tag{205}$$

Recalling that

$$X_l = \mathcal{P}E_l\mathcal{P}, \qquad a_l = \frac{\mathrm{Tr}(\mathcal{P}E_l\mathcal{P})}{k}, \tag{206}$$

we arrive at

$$1 - \sum_l det_{\mathcal{P}}(\mathcal{P}E_l\mathcal{P})^{1/k} \geq \frac{1}{2k}\sum_l \left\|\mathcal{P}E_l\mathcal{P} - \frac{\mathrm{Tr}(\mathcal{P}E_l\mathcal{P})}{k}\mathcal{P}\right\|_2^2. \tag{207}$$

We need to derive a lower bound on the right-hand side of (207). Let us consider the single channel Λ_{AD}. Let $D_j = |j\rangle\langle j|$. We obtain $E_0^s = K_0^\dagger K_0 = \sum_{j=0}^{d-1} q_j D_j$. The supersript s denotes that the operator acts on single party Hilbert space. For $j \to i$ jump, we have $E_{i\leftarrow j}^s = K_{i\leftarrow j}^\dagger K_{i\leftarrow j} = \gamma_{i\leftarrow j}D_j$. Let us define a set:

$$\mathcal{I} = \{0\} \cup \{(i \leftarrow j) : 0 \leq i < j \leq d-1\} \tag{208}$$

For every $\alpha \in \mathcal{I}$ we have Kraus operator K_α. We define $E_\alpha^s = K_\alpha^\dagger K_\alpha$. Let T be defined as:

$$E_\alpha^s = \sum_{i=0}^{d-1} T_{\alpha i}D_i. \tag{209}$$

Let us assume that every excited level has non-zero probability of decay ($\eta_j > 0$ for every $j > 0$). Then, T has full column rank, and hence its smallest singular value is strictly positive,

$$\sigma_{\min}(T) > 0. \tag{210}$$

For convenience, we define

$$\mu := \sigma_{\min}(T)^2 > 0. \tag{211}$$

Let $\beta = (\alpha_1, \cdots \alpha_m)$. We have:

$$E_\beta = E_{\alpha_1}^s \otimes \cdots \otimes E_{\alpha_m}^s. \tag{212}$$

Let $z = (z_1, \cdots, z_m)$, where each $z_i \in \{0, 1, \cdots, d-1\}$ and :

$$D_z = D_{z_1} \otimes \cdots \otimes D_{z_m} = |z\rangle\langle z|. \tag{213}$$

Similarly to the single operator case:

$$\begin{aligned} E_\beta &= E_{\alpha_1}^s \otimes \cdots \otimes E_{\alpha_m}^s \\ &= \left(\sum_{z_1=0}^{d-1} T_{\alpha_1 z_1}D_{z_1}\right) \otimes \cdots \otimes \left(\sum_{z_m=0}^{d-1} T_{\alpha_m z_m}D_{z_m}\right) = \sum_z M_{\beta z}D_z. \end{aligned} \tag{214}$$

In Eq. (214) we have $M_{\beta z} = \prod_{i=1}^{m} T_{\alpha_i z_i}$. Consequently, $M = T^{\otimes m}$. Since the singular values of a tensor product are all pairwise products of the singular values of its factors, we have

$$\begin{aligned} \sigma_{\min}(M)^2 &= \sigma_{\min}\left(T^{\otimes m}\right)^2 \\ &= \sigma_{\min}(T)^{2m} = \mu^m. \end{aligned} \tag{215}$$

Let us define $t_z = \mathrm{Tr}(\mathcal{P}D_z\mathcal{P})$ and

$$Y_z = \mathcal{P}D_z\mathcal{P} - \frac{t_z}{k}\mathcal{P}. \tag{216}$$

We have:

$$a_\beta = \frac{1}{k}\mathrm{Tr}X_\beta = \sum_z M_{\beta z}\frac{t_z}{k}. \tag{217}$$

We obtain

$$X_\beta - a_\beta\mathcal{P} = \sum_z M_{\beta z}Y_z. \tag{218}$$

Inequality by the smallest singular value of M gives:

$$\sum_\beta \left\|X_\beta - a_\beta\mathcal{P}\right\|_2^2 \geq \mu^m \sum_z \left\|Y_z\right\|_2^2. \tag{219}$$

Let us estimate $\|Y_z\|_2^2$. Expanding Eq. (216) and noting that Y_z is hermitian we obtain:

$$\left\|Y_z\right\|_2^2 = \mathrm{Tr}[(\mathcal{P}D_z\mathcal{P} - \frac{t_z}{k}\mathcal{P})^2] = \mathrm{Tr}[A_z^2 - \frac{2t_z}{k}A_z + \frac{t_z^2}{k^2}\mathcal{P}], \tag{220}$$

where $A_z = \mathcal{P}D_z\mathcal{P}$. Using definition of t_z and $\mathrm{Tr}\mathcal{P} = k$ and the fact that A_z is at most rank one, we obtain:

$$\left\|Y_z\right\|_2^2 = (1 - \frac{1}{k})t_z^2. \tag{221}$$

Summing over all z we obtain:

$$\sum_z \left\|Y_z\right\|_2^2 = (1 - \frac{1}{k})\sum_z t_z^2. \tag{222}$$

We further obtain:

$$\sum_z t_z = \mathrm{Tr}(\mathcal{P}\sum_z D_z\mathcal{P}) = \mathrm{Tr}\mathcal{P} = k, \tag{223}$$

where we used the fact that $\sum_z D_z = \mathbb{1}$. Recall that $z \in (z_1, \cdots, z_m)$. Each z_i can take d possible values. Therefore, there is d^m possible choices of z. Using Cauchy-Schwarz:

$$(\sum_z t_z)^2 \leq (\sum_z t_z^2)(\sum_z 1) = d^m \sum_z t_z^2. \tag{224}$$

Joining Eqs. (223) and (224) we obtain:

$$\sum_z t_z^2 \geq \frac{k^2}{d^m}. \tag{225}$$

Substituting it into Eq. (222) we obtain:

$$\sum_z \|Y_z\|_2^2 \geq \frac{k(k-1)}{d^m} \tag{226}$$

Combining the previous bounds, we obtain

$$\begin{aligned} 1 - \sum_\beta \left[\det_{\mathcal{P}}(\mathcal{P} E_\beta \mathcal{P})\right]^{1/k} &\geq \frac{1}{2k} \sum_\beta \left\|X_\beta - a_\beta \mathcal{P}\right\|_2^2 \\ &\geq \frac{\mu^m}{2k} \sum_{\mathbf{z}} \|Y_{\mathbf{z}}\|_2^2 \\ &\geq \frac{\mu^m}{2k} \frac{k(k-1)}{d^m} \\ &= \frac{k-1}{2} \left(\frac{\mu}{d}\right)^m. \end{aligned} \tag{227}$$

Finally, for $k = k_m$, we have

$$\kappa \leq 1 - \frac{k_m - 1}{2} \left(\frac{\mu}{d}\right)^m. \tag{228}$$

As before, we assume that the entanglement of formation does not vanish asymptotically. Using Lemma 15, we obtain the following lower bound on the distance to the set of separable states:

$$\varepsilon_{n,m} \geq \frac{E_0^2}{8m^2(\log_2 d)^2}. \tag{229}$$

On the other hand, Eq. (163), applied to the block channel $\Lambda_{\mathrm{AD}}^{\otimes m}$, gives

$$\varepsilon_{n,m} \leq 2\kappa^{n/[2(k_m-1)]} \left(\sqrt{k_m} - 1\right), \tag{230}$$

where k_m is the Schmidt-rank parameter associated with the input state and satisfies

$$k_m \leq d^m. \tag{231}$$

Consequently,

$$\begin{aligned} \kappa^{\frac{n}{2(k_m-1)}} &\leq \left[1 - \frac{k_m - 1}{2} \left(\frac{\mu}{d}\right)^m\right]^{\frac{n}{2(k_m-1)}} && (232) \\ &\leq \exp\left[-\frac{n}{2(k_m-1)} \frac{k_m - 1}{2} \left(\frac{\mu}{d}\right)^m\right] && (233) \\ &= \exp\left[-\frac{n}{4} \left(\frac{\mu}{d}\right)^m\right] && (234) \\ &= \exp\left[-\frac{n\mu^m}{4d^m}\right], && (235) \end{aligned}$$

where we used the inequality $1 - x \leq e^{-x}$.

It follows that

$$\varepsilon_{n,m} \leq 2\left(\sqrt{k_m} - 1\right) \exp\left[-\frac{n\mu^m}{4d^m}\right]. \tag{236}$$

Using

$$\sqrt{k_m} - 1 \leq \sqrt{k_m} \leq d^{m/2}, \tag{237}$$

we obtain

$$\varepsilon_{n,m} \leq 2d^{m/2} \exp\left[-\frac{n\mu^m}{4d^m}\right]. \tag{238}$$

Combining the lower and upper bounds on $\varepsilon_{n,m}$ gives

$$\frac{E_0^2}{8m^2(\log_2 d)^2} \leq 2d^{m/2} \exp\left[-\frac{n\mu^m}{4d^m}\right]. \tag{239}$$

Equivalently,

$$\exp\left[-\frac{n\mu^m}{4d^m}\right] \geq \frac{E_0^2}{16m^2(\log_2 d)^2 d^{m/2}}. \tag{240}$$

Taking the natural logarithm yields

$$\frac{n\mu^m}{4d^m} \leq \frac{m}{2} \ln d + 2 \ln m + C_0, \tag{241}$$

where

$$C_0 := \ln\left[\frac{16(\log_2 d)^2}{E_0^2}\right] \tag{242}$$

is independent of n and m. Therefore,

$$n \leq 4\left(\frac{d}{\mu}\right)^m \left[\frac{m}{2} \ln d + 2 \ln m + C_0\right]. \tag{243}$$

For all sufficiently large m, the expression in square brackets is upper bounded by a constant times m. Hence, there exists a constant $C_1 > 0$, independent of n and m, such that

$$n \leq C_1 m \left(\frac{d}{\mu}\right)^m. \tag{244}$$

Defining

$$a := \ln\left(\frac{d}{\mu}\right) > 0, \tag{245}$$

we can write

$$n \leq C_1 m e^{am}. \tag{246}$$

Multiplying both sides by a/C_1, we obtain

$$\frac{a}{C_1} n \leq (am) e^{am}. \tag{247}$$

Since the function $x \mapsto xe^x$ is increasing for $x > 0$, the principal branch of the Lambert function gives

$$m \geq \frac{1}{a} W_0\left(\frac{a}{C_1} n\right). \tag{248}$$

Using the asymptotic expansion[52]

$$W_0(x) = \ln x - \ln\ln x + O\left(\frac{\ln\ln x}{\ln x}\right), \tag{249}$$

we finally arrive at

$$m \geq \frac{\ln n - \ln\ln n}{\ln(d/\mu)} - O(1). \tag{250}$$

Therefore, preserving a nonvanishing amount of end-to-end entanglement requires

$$m = \Omega(\ln n). \tag{251}$$

Lemma 15. *Let $\rho^{AB}_{n,m(n)}$ denote the final state produced by a protocol using m parallel channel uses per link and n sequential links. We define its trace distance from the set of separable states as*

$$\varepsilon_{n,m} := \frac{1}{2} \min_{\sigma^{AB} \in \mathcal{S}} \left\| \rho^{AB}_{n,m(n)} - \sigma^{AB} \right\|_1 . \tag{252}$$

The factor $1/2$ is introduced so that $\varepsilon_{n,m} \in [0,1]$. There is the following implication:

$$\liminf_{n\to\infty} E_f\left(\rho^{AB}_{n,m(n)}\right) > 0 \Rightarrow \varepsilon_{n,m} \geq \frac{E_0^2}{8m^2(\log_2 d)^2}, \tag{253}$$

where E_0 is defined by:

$$E_f\left(\rho^{AB}_{n,m(n)}\right) \geq E_0 \tag{254}$$

for all $n \geq n_0$.

We first convert this assumption into a lower bound on $\varepsilon_{n,m}$. Let $\sigma_{n,m} \in \mathcal{S}$ be a separable state attaining the minimum in the definition of $\varepsilon_{n,m}$, and define

$$\delta_{n,m} := \sqrt{\varepsilon_{n,m}\,(2 - \varepsilon_{n,m})}. \tag{255}$$

By the uniform continuity bound for the entanglement of formation [55],

$$\left|E_f(\rho_{n,m}) - E_f(\sigma_{n,m})\right| \leq \delta_{n,m} \log_2 D_m + (1+\delta_{n,m})\, h_2\left(\frac{\delta_{n,m}}{1+\delta_{n,m}}\right),$$

where D_m is the smaller local dimension and h_2 denotes the binary entropy. Since $\sigma_{n,m}$ is separable, $E_f(\sigma_{n,m}) = 0$. Moreover, $\dim \mathcal{H}_{\mathcal{A}} = d^m$ and consequently $D_m \leq d^m$. Hence

$$E_f(\rho_{n,m}) \leq \delta_{n,m} m \log_2 d + (1+\delta_{n,m})\, h_2\left(\frac{\delta_{n,m}}{1+\delta_{n,m}}\right). \tag{256}$$

Using

$$h_2(x) \leq 2\sqrt{x(1-x)}, \tag{257}$$

we further obtain

$$E_f(\rho_{n,m}) \leq \delta_{n,m} m \log_2 d + 2\sqrt{\delta_{n,m}}. \tag{258}$$

If $m(n)$ remained bounded along an infinite subsequence, the upper bound from Theorem 4 would imply $\varepsilon_{n,m(n)} \to 0$ and consequently $E_f(\rho_{n,m(n)}) \to 0$. Thus, the assumption of nonvanishing end-to-end entanglement implies that $m(n) \to \infty$.

Since $m \to \infty$, for all E_0 we can choose m_0 such that for all $m > m_0$, the following inequality holds:

$$m \log_2 d \geq \frac{8}{E_0}. \tag{259}$$

Since E_0 can be arbitrary, we can choose it such that:

$$\liminf_{n\to\infty} E_f\left(\rho^{AB}_{n,m(n)}\right) > 0 \Rightarrow E_f(\rho^{AB}_{n,m(n)}) > E_0 \tag{260}$$

for all sufficiently large n.

Assume that the following inequality holds: $\delta_{n,m} < \frac{E_0}{2m\log_2 d}$. We obtain:

$$\delta_{n,m} m \log_2 d + 2\sqrt{\delta_{n,m}} < \frac{E_0}{2} + 2\sqrt{\frac{E_0}{2m\log_2 d}} \leq \frac{E_0}{2} + \frac{E_0}{2}, \tag{261}$$

where in the last step we used Eq. (259). But combination of Eqs. (258) and (260) states that $E_0 < E_f(\rho_{n,m}) \leq \delta_{n,m} m \log_2 d + 2\sqrt{\delta_{n,m}}$. This contradiction implies that:

$$\delta_{n,m} \geq \frac{E_0}{2m\log_2 d}. \tag{262}$$

Since

$$\delta^2_{n,m} = \varepsilon_{n,m}\,(2 - \varepsilon_{n,m}) \leq 2\varepsilon_{n,m}, \tag{263}$$

we obtain the dimension-dependent lower bound

$$\varepsilon_{n,m} \geq \frac{E_0^2}{8m^2(\log_2 d)^2} \tag{264}$$

for all sufficiently large n.

Achievability of the bound for the depolarizing channel

Consider the d-dimensional depolarizing channel

$$\Lambda_p(\rho) = (1-p)\rho + p\frac{\mathbb{1}}{d}. \tag{265}$$

Using the Weyl twirling identity[45], the channel can equivalently be written as

$$\Lambda_p(\rho) = \left(1 - p + \frac{p}{d^2}\right)\rho + \frac{p}{d^2} \sum_{(a,b)\neq(0,0)} W_{ab}\rho W^\dagger_{ab}. \tag{266}$$

The probability of each non-identity Weyl error is $\frac{p}{d^2}$, since there is d^2-1 possible errors, probability that there will be an error on some qudit is

$$q = p\left(1 - \frac{1}{d^2}\right). \tag{267}$$

Suppose that one elementary link consists of m parallel and independent uses of Λ_p. Let X_m denote the number of physical qudits affected by nonidentity Weyl errors in this block. It follows from Eq. (267) that

$$X_m \sim \mathrm{Bin}(m, q), \tag{268}$$

and hence

$$\Pr(X_m = r) = \binom{m}{r} q^r (1-q)^{m-r}. \tag{269}$$

Assume that there exists a family of quantum codes $\{C_m\}$, encoding at least one logical qudit, whose minimum distance Δ_m satisfies

$$\Delta_m \geq \delta m \tag{270}$$

for some constant $\delta > 0$ and all sufficiently large m. A code with distance Δ_m corrects every error supported on at most

$$t_m = \left\lfloor \frac{\Delta_m - 1}{2} \right\rfloor \tag{271}$$

physical qudits[43].

Choose a constant τ such that

$$q < \tau < \frac{\delta}{2}. \tag{272}$$

The crucial point is that the code distance grows linearly with the number of physical qudits. $\Delta_m \geq \delta m$. Therefore, for every fixed $\tau < \delta/2$ and all sufficiently large m,

$$\lfloor \tau m \rfloor \leq \left\lfloor \frac{\Delta_m - 1}{2} \right\rfloor. \tag{273}$$

Hence, every Weyl-error pattern of weight at most $\lfloor \tau m \rfloor$ is correctable. Therefore, the failure probability of the recovery operation on one elementary link satisfies

$$\varepsilon_m \leq \Pr(X_m > \tau m). \tag{274}$$

Since $\tau > q$, the Chernoff-Hoeffding bound gives [56]

$$\varepsilon_m \leq \exp\left[-mD(\tau\|q)\right], \tag{275}$$

where

$$D(\tau\|q) = \tau \ln \frac{\tau}{q} + (1-\tau)\ln\frac{1-\tau}{1-q} \tag{276}$$

is the binary relative entropy. We denote

$$\alpha := D(\tau\|q) > 0. \tag{277}$$

After each elementary link, the corresponding repeater station applies the recovery operation and re-encodes the logical system before transmission through the next link. The probability $\varepsilon_{n,m}$ that at least one of the n recovery operations fails is bounded, by the union bound, as

$$\varepsilon_{n,m} \leq n\varepsilon_m \leq ne^{-\alpha m}. \tag{278}$$

For any constant $\eta > 0$, choose

$$m(n) = \left\lceil \frac{(1+\eta)\ln n}{\alpha} \right\rceil. \tag{279}$$

It then follows that

$$\varepsilon_{n,m(n)} \leq n^{-\eta}, \tag{280}$$

and consequently

$$\lim_{n\to\infty} \varepsilon_{n,m(n)} = 0. \tag{281}$$

To relate this result to entanglement distribution, let the logical input be one half of a maximally entangled logical state Φ_L. If all error patterns are correctable, the state transmitted through the entire network is exactly Φ_L. The final state can therefore be written as

$$\rho_{n,m} = (1-\varepsilon'_{n,m})\Phi_L + \varepsilon'_{n,m}\sigma_{n,m}, \tag{282}$$

where $\varepsilon'_{n,m} \leq \varepsilon_{n,m}$ and $\sigma_{n,m}$ is some quantum state. Consequently,

$$\frac{1}{2}\left\|\rho_{n,m(n)} - \Phi_L\right\|_1 \leq \varepsilon_{n,m(n)} \longrightarrow 0. \tag{283}$$

By the continuity of the entanglement of formation[55],

$$\lim_{n\to\infty} E_f\left(\rho_{n,m(n)}\right) = E_f(\Phi_L) = \log_2 d_L > 0, \tag{284}$$

where $d_L \geq 2$ is the logical dimension.

We have thus shown that

$$m = O(\ln n) \tag{285}$$

parallel channel uses per elementary link are sufficient. Together with the lower bound of Theorem 5, this establishes the optimal scaling

$$m = \Theta(\ln n) \tag{286}$$

for depolarizing noise satisfying the condition $q < \delta/2$ for the chosen family of codes.