A Lower Bound on the Space Overhead of Fault-Tolerant Quantum Computation

Omar Fawzi ✉
Univ Lyon, ENS Lyon, UCBL, CNRS, Inria, LIP, F-69342, Lyon Cedex 07, France

Alexander Müller-Hermes ✉
Institut Camille Jordan, Université Claude Bernard Lyon 1, 69622 Villeurbanne cedex, France
Department of Mathematics, University of Oslo, Norway

Ala Shayeghi ✉
Univ Lyon, ENS Lyon, UCBL, CNRS, Inria, LIP, F-69342, Lyon Cedex 07, France

──── **Abstract** ────

The threshold theorem is a fundamental result in the theory of fault-tolerant quantum computation stating that arbitrarily long quantum computations can be performed with a polylogarithmic overhead provided the noise level is below a constant level. A recent work by Fawzi, Grospellier and Leverrier (FOCS 2018) building on a result by Gottesman (QIC 2013) has shown that the space overhead can be asymptotically reduced to a constant independent of the circuit provided we only consider circuits with a length bounded by a polynomial in the width. In this work, using a minimal model for quantum fault tolerance, we establish a general lower bound on the space overhead required to achieve fault tolerance.

For any non-unitary qubit channel $\mathcal{N}$ and any quantum fault tolerance schemes against i.i.d. noise modeled by $\mathcal{N}$, we prove a lower bound of $\max\left\{ Q(\mathcal{N})^{-1}n, \alpha_{\mathcal{N}} \log T \right\}$ on the number of physical qubits, for circuits of length T and width n. Here, $Q(\mathcal{N})$ denotes the quantum capacity of $\mathcal{N}$ and $\alpha_{\mathcal{N}} > 0$ is a constant only depending on the channel $\mathcal{N}$. In our model, we allow for qubits to be replaced by fresh ones during the execution of the circuit and in the case of unital noise, we allow classical computation to be free and perfect. This improves upon results that assumed classical computations to be also affected by noise, and that sometimes did not allow for fresh qubits to be added. Along the way, we prove an exponential upper bound on the maximal length of fault-tolerant quantum computation with amplitude damping noise resolving a conjecture by Ben-Or, Gottesman and Hassidim (2013).

1 Introduction

Quantum computing is capable of solving problems for which no efficient classical algorithms are known. However, the decoherence of quantum systems due to inevitable interactions with the environment makes it challenging to build reliable large-scale quantum computers. To circumvent this obstacle the theory of quantum error correction and fault tolerance was developed [46, 47, 48].

A quantum fault tolerance scheme receives as input any ideal circuit designed for noiseless computation and outputs an encoding of the circuit which is robust against noise. The fault-tolerant threshold theorem was a major achievement in quantum computing: As long as the noise level in a quantum computer is below a certain threshold, any arbitrarily long quantum computation can be reliably performed with a reasonably low overhead, namely, a polylogarithmic factor in space and time [1, 34, 31, 3]. Since then a lot of effort has been focused on introducing fault tolerance schemes with lower resource overheads, improving the lower bounds on the fault-tolerant threshold for different noise models (see, e.g., [33, 4, 6, 53, 56]). Gottesman [27] showed that if quantum error correcting codes with certain properties exist, then the space overhead can be brought down to a constant arbitrarily close to 1, for computations in which the number of time steps is at most polynomial in the number of qubits. Recently, Fawzi *et al.* [26] showed that these properties are indeed satisfied by quantum expander codes and the statement is true even for sub-exponential computations. But what is the fundamental limit on the space overhead for quantum fault-tolerance as a function of the noise model? This question still remains open. More broadly, the study of general properties and limitations of quantum fault tolerance schemes has received far less attention. This is in part due to the difficulty of providing a definition encompassing the different aspects of quantum fault tolerance. In this paper, we take a step towards proving general statements about quantum fault tolerance schemes. In particular, given an *arbitrary* single qubit noise channel $\mathcal{N}$, our main result is a lower bound on the achievable space overhead of quantum fault tolerance schemes against i.i.d. errors modeled by $\mathcal{N}$.

Previous works on the limitations of noisy quantum circuits typically proved statements of the following form (see, e.g., [42, 9, 16, 30]: If the noise level is too high, then any noisy circuit is "useless" after a limited number of time steps. These results do not apply to the low-noise regime, i.e., when the noise channel $\mathcal{N}$ is not far from the identity channel, and in particular, noise levels below the fault-tolerant threshold. Therefore, they cannot be used to derive lower bounds on the overhead of fault tolerance schemes. Moreover, these statements crucially depend on several parameters, including the circuit model, the noise model, and the measure of uselessness they consider. Tweaking any of these variables can make a substantial difference. The circuit model specifies the basic components of the quantum circuits including the set of possible gates, state preparations, and measurements, as well as a model of classical computation for the processing of the measurement outcomes by the circuits. The noise model specifies how each component of the circuit deviates from its ideal action due the noise. Different noise models such as depolarizing, dephasing, or amplitude damping noise are not equivalent in the context of fault tolerance (see, e.g., [9]). Earlier fault tolerance schemes were designed for generic noise models. In practice, however, any physical implementation of a quantum computer has some additional noise structure. With the knowledge of the dominant noise in the system, one can potentially utilize schemes that are tailored for a specific noise model for cheaper suppression of the dominant noise. Indeed, biased-noise fault tolerance beyond the usual Pauli noise models has gained growing interest in recent years (see, e.g., [5, 44, 12, 50, 28]). Earlier relevant results on limitations of noisy quantum computation, however, are mostly focused on a particular noise channel, typically, depolarizing noise (with the exception of Ref. [51]), and it is not clear if the introduced techniques can be extended to other noise models.

With current technology, classical computation can be performed with practically perfect accuracy. Therefore, it is reasonable and conventional to assume classical computation to be error-free when evaluating the performance of quantum fault tolerance schemes. For instance, in this context, given measured syndromes of a quantum stabilizer code, computing

(a) Plain circuit model: all subsystems are subject to the noise

(b) Circuit model with free and noiseless classical computation: classical systems are noise-free
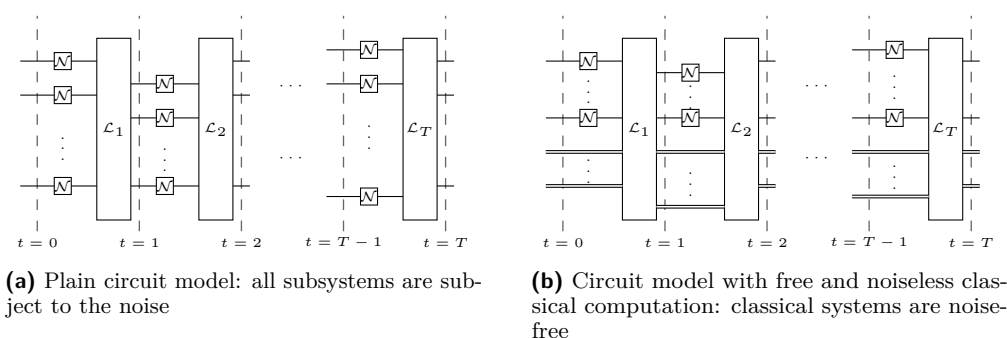
Figure 1 Illustration of the circuit model.

a description of a corresponding error is assumed to be noiseless. While some earlier impossibility results establish very strong statements about noisy circuits (see the discussion in Section 1.3), they do not incorporate the aid of noiseless classical computation and the adaptivity achieved through classical control. In these works, the classical systems are not distinguished from the quantum systems and they are assumed to be subject to the same noise. In fact, without this assumption, the "uselessness" statements established in these works fail to hold even for arbitrarily long circuits. Therefore, these results are not directly applicable to the fault tolerance schemes which rely on perfect classical computation and control. In the case of unital noise, by allowing this additional resource in our model, we address a limitation of the previous results on the power of noisy quantum circuits. Our results in this case apply to adaptive protocols composed of a hybrid of classical and quantum computation.

Before stating our results more formally, we first need to describe our quantum circuit model.

1.1 Our model

A (noiseless) quantum circuit C in our model is defined by a sequence $\{\mathcal{L}_i\}_{i \in [T]}$ of arbitrary completely positive and trace-preserving (CPTP) linear maps, acting on a potentially non-trivial input system. For any density operator ρ on the input Hilbert space, the output of the circuit is given by

$$C(\rho) \coloneqq (\mathcal{L}_T \circ \cdots \circ \mathcal{L}_2 \circ \mathcal{L}_1)(\rho) \ ,$$

where we use the same notation C to denote the overall channel corresponding to the action of the circuit C. We use the term *time* to refer to the vertical cuts between the layers of computation, as shown in Fig. 1. The i-th *time step* of the circuit is defined by $\mathcal{L}_i$ acting between $t = i - 1$ and $t = i$.

Given a qubit noise channel $\mathcal{N}$, for any quantum circuit C, the noisy implementation of C with i.i.d. noise modeled by $\mathcal{N}$ is obtained by interlacing the computational steps $\mathcal{L}_i$ with layer of the noise channel $\mathcal{N}$ acting independently on every qubit. Let n_i denote the number of qubits at $t = i$. Then for every input state ρ, the output of the noisy circuit in this model is given by

$$(\text{i.i.d.}_{\mathcal{N}}(C))(\rho) = \left(\mathcal{L}_T \circ \mathcal{N}^{\otimes n_{T-1}} \circ \cdots \circ \mathcal{L}_2 \circ \mathcal{N}^{\otimes n_1} \circ \mathcal{L}_1 \circ \mathcal{N}^{\otimes n_0}\right)(\rho) \ ,$$

where i.i.d.$_{\mathcal{N}}(C)$ denotes the overall channel corresponding to the noisy implementation of C. We also consider a circuit model with free and noiseless classical computation. A quantum

circuit in this model is defined as above, except that now we distinguish between classical and quantum subsystems and only the quantum subsystems are subject to the noise. In both models, for a circuit C, the *width $Width(C)$* of the circuit is the maximum number of *qubits* at any time step and its *length $Length(C)$* is the number of time steps of the circuit. Note that in the model with free and noiseless classical computation, we do not count the size of the classical subsystems in $Width(C)$ – in other words, the width of the circuit is determined by the number of qubit systems that are subject to the noise. Moreover, classical computations do not contribute to $Length(C)$ in this model.

1.2   Discussion of our main results, proof techniques, and related work

Limitations on quantum memory times. We consider noisy quantum circuits in the i.i.d.$_\mathcal{N}$ error model, for an arbitrary non-unitary qubit channel $\mathcal{N}$. Our main technical contribution is the following:

▶ **Theorem 1.** *Let $\mathcal{I}$ denote the qubit identity channel and $\mathcal{N}$ be a non-unitary qubit channel. Then there exists a constant $p \in (0,1]$ only depending on $\mathcal{N}$ such that the following holds: Let C be an arbitrary circuit of length T and width n such that $T \geq (1/p)^{2n}$. Then for any pair of ideal encoding and decoding maps $\mathcal{E}$ and $\mathcal{D}$ (of appropriate input and output dimensions), we have*

$$\|\mathcal{D} \circ \mathrm{i.i.d.}_\mathcal{N}(C) \circ \mathcal{E} - \mathcal{I}\|_1 \geq \epsilon_0 \ ,$$

for some universal constant $\epsilon_0 \geq 1/128$. For any unital (non-unitary) noise channel $\mathcal{N}$, this statement holds even in the circuit model with free and noiseless classical computation.

The constant $p = p(\mathcal{N})$ in Theorem 1 is given by $p(\mathcal{N}) = \max(p_1(\mathcal{N}), p_2(\mathcal{N}))$ where

$$p_1(\mathcal{N}) = \max\left\{ q \in [0,1] \ : \ \mathcal{N} \geq q\mathcal{M}, \text{ for an entanglement breaking quantum channel } \mathcal{M} \right\}.$$

and

$$p_2(\mathcal{N}) = \max\left\{ q^2 \lambda_{\min}\left(C_{\mathcal{M}^\dagger \circ \mathcal{M}}\right)/2 \ : \ q \in [0,1], \ \mathcal{N} \geq q\mathcal{M}, \text{ for quantum channel } \mathcal{M} \right\}.$$

Here, we write $\mathcal{N} \geq \mathcal{M}$ for linear maps $\mathcal{N}$ and $\mathcal{M}$ if the difference $\mathcal{N} - \mathcal{M}$ is completely positive. We will show that $p_1(\mathcal{N}) > 0$ for any unital qubit channel $\mathcal{N}$ that is not unitary, and that $p_2(\mathcal{N}) > 0$ for every non-unital qubit channel $\mathcal{N}$.

We point out that the statement of Theorem 1 does not hold as is for quantum circuits operating on qudits rather than qubits and affected by any non-unitary noise channel. By taking a direct sum of the ideal quantum channel and a completely depolarizing channel it is easy to construct non-unitary noise channels for which quantum information can be stored indefinitely in a decoherence-free subspace. To obtain an upper bound on the storage time in these cases, further assumptions are needed.

Theorem 1 states that for any noise channel $\mathcal{N}$, no noisy quantum circuit in the i.i.d.$_\mathcal{N}$ error model is capable of preserving even a single (unknown) qubit for more than an exponential number of time steps in its width. We would like to emphasize two features of our result:

1. The quantum channels making up the circuit do not need to be unital. Therefore, they can remove all entropy produced by the noise channel along the computation, for example by introducing ancilla qubits.

2. For unital noise channels, our results also hold for circuits with free and unlimited classical computation. Even after long computation times, it is then possible to prepare orthogonal states conditioned on earlier measurement outcomes that have been kept in the perfect classical memory.

In order to prove the result, we consider two parallel instances of the circuit C and prove that for any input state the distance from the set of separable states with respect to this bipartition is contracted by a factor of $1 - p^{2n}$ (with $p > 0$ as above) in each time step. This implies that the circuit C itself cannot preserve every quantum state for more than $(1/p)^{2n}$ time steps.

Ben-Or *et al.* [9] studied the limitations of fault-tolerant quantum computation under the independent noise model i.i.d.$_{\mathcal{N}}$, for different classes of non-unitary qubit channels. Applying such a channel many times either converges to a limit channel mapping every quantum state to a point in the Bloch sphere (e.g., for the depolarizing channel), or it converges to a limit quantum channel mapping every quantum state to a diameter of the Bloch sphere (e.g., for the dephasing channel). By slightly generalizing an earlier entropic argument[1], Ben-Or *et al.* prove that if the limit channel maps every quantum state to the maximally mixed state, then it is impossible to compute for more than $\mathrm{O}(\log n)$ steps and they show how to achieve this bound up to a polylogarithmic factor in $\log n$. Entropic arguments of this kind are crucially based on the assumption that entropy cannot be removed during the execution of the circuit. In particular, these arguments do not apply for our circuit model (including common settings of fault-tolerance [1, 3]) where the entropy produced by the noise channel can be removed along the computation (cf., feature 1 from above). When the limiting quantum channel maps every quantum state to a diameter of the Bloch sphere (e.g., the dephasing channel), Ben-Or *et al.* prove that it is impossible to store quantum information for more than a polynomial number of time steps, however again assuming no fresh ancilla qubits and no noiseless classical computation. Using standard threshold theorems, they also show that for this class of channels, computations involving $\mathrm{O}(n^a)$ qubits for $\mathrm{O}(n^b)$ time steps are possible, provided $a + b < 1$. Finally, if the limit channel has a fixed point other than the center of the Bloch sphere (e.g., the amplitude damping channel) then exponentially long quantum computations are possible. This is achieved by using the noise itself as a refrigerator to cool down the qubits and produce fresh ancillas. They also conjecture that for this class of channels it is impossible to compute for more than exponential time with no fresh ancilla qubits. To our knowledge, prior to this work, no proof of this statement existed. Theorem 1 proves this conjecture to be true. In addition, our result applies even if fresh ancillas are used and the upper bound on the length becomes $e^{\mathrm{O}(n+m)}$ with m ancilla qubits.

In a related work, Raginsky [40] considers the noise model where a fixed noise channel $\mathcal{T}$ acts collectively on all the qubits in the quantum circuit after each computational step. He shows that, for any quantum circuit involving only unitary operations, if the channel $\mathcal{T}$ modeling the noise is strictly contractive then the distance of possible output states of the circuit is exponentially small in the length of the circuit. Moreover, he shows that the set of strictly contractive channels is dense in the set of all quantum channels. Using this fact and under the assumption that quantum operations can be performed only with a finite precision, he argues that strictly contractive channels serve as a natural abstract model for noise in any physically realizable quantum computer. Our work can be seen as an extension of this argument. In particular, we show how to apply a similar reasoning when the channel $\mathcal{T}$ modeling the noise is given by tensor powers of an arbitrary non-unitary qubit channel, and for unital noise we extend the argument by taking into account free and perfect classical computation (obtaining feature 2 from above). Note that in general strict contractivity of a quantum channel does not necessarily imply the same property for its tensor powers (see the discussion at the end of Section 3 for more details).

---

[1] This argument was stated in [2], but the proof of a key lemma is not correct. This was corrected in [37].

Limitations on quantum fault-tolerance.   Inspired by the observation of Theorem 1, in Section 4, we provide a natural high-level description of quantum fault tolerance schemes with minimal assumptions. A fault tolerance scheme is a circuit encoding associated with a family of quantum error correcting codes. Intuitively, we only require a fault tolerance scheme to be capable of implementing any arbitrary logical circuit on *encoded data* in a code from the family, up to arbitrary accuracy in the presence of noise. Here, we assume that the encoding and decoding maps for the quantum error correcting code can be implemented perfectly. We quantify the accuracy in terms of the induced trace distance between the channels corresponding to the ideal circuit and its fault-tolerant implementation with noise. To our knowledge, this requirement is satisfied by most, if not all, known fault tolerance schemes.

In this work we do not focus on investigating the limitations of *computation* using noisy quantum circuits. Indeed, our results are shown assuming perfect classical computation with *unlimited* power is possible. We are rather interested in proving limitations of fault tolerance schemes as procedures which, *oblivious* to the overall computation being performed, allow the implementation of any ideal input circuit using noisy components.

Our high-level definition of fault tolerance enables us to recast the problem of proving a lower bound on the space overhead for fault-tolerant quantum computation into a more tractable problem in the better understood framework of quantum Shannon theory. Note that a fault tolerance scheme should naturally be able to implement a quantum memory. From this alternative perspective, a quantum memory for T time steps in the i.i.d.$_\mathcal{N}$ noise model corresponds to one-way point-to-point quantum communication with $T-1$ intermediate relay points on a line, linked by T instances of the noisy channel $\mathcal{N}$, where the relay points are allowed to perform arbitrary (noiseless) quantum operations and communicate classically with the following relay points on the line. Similar communication models have been studied earlier in the quantum Shannon theory literature (see, e.g., [39]). Here, given a qubit channel $\mathcal{N}$, we are interested in the optimal rate of communication over the links for reliably sending n qubits as a function of T and n. The case $T=1$ is well-studied in the quantum Shannon theory literature and the optimal rate is given by the quantum capacity of $\mathcal{N}$. Theorem 1 implies an upper bound on the achievable rates (or, equivalently, a lower bound on the necessary overhead) as a function of T. The following is our main result:

▶ **Theorem 2.** *Let $\mathcal{N}$ be a non-unitary qubit channel. For any fault tolerance scheme against the* i.i.d.$_\mathcal{N}$ *noise model, the number of physical qubits is at least*

$$\max \left\{ \mathrm{Q}(\mathcal{N})^{-1} n, \alpha_\mathcal{N} \log T \right\} \ ,$$

for circuits of length T and width n, and any constant accuracy $\epsilon \leq 1/128$. Here, $\mathrm{Q}(\mathcal{N})$ denotes the quantum capacity of $\mathcal{N}$ and $\alpha_\mathcal{N} = \frac{1}{2 \log 1/p}$, where $p \in (0,1]$ is a constant only depending on the channel $\mathcal{N}$. In particular, fault tolerance is not possible even for a single time step if $\mathcal{N}$ has zero quantum capacity or the space overhead is strictly less than $\mathrm{Q}(\mathcal{N})^{-1}$. When $\mathcal{N}$ is a unital (non-unitary) qubit channel, the statement holds even in the circuit model with free and noiseless classical computation.

Theorem 2 in particular implies that fault-tolerant quantum computation with constant space overhead is not achievable beyond exponential computations for any noise channel $\mathcal{N}$.

For the p-depolarizing noise given by $\mathcal{N}(\rho) = (1-p)\rho + p\frac{\mathbb{1}}{2}$, the quantum capacity is known to be zero for $p > \frac{1}{3}$ [15]. Thus, for $p > \frac{1}{3}$, the lower bound becomes $+\infty$, i.e., fault tolerance is not possible in our model. This gives stronger limitations compared to previous results [42, 30, 16] (except for a result in [30] where making an assumption on the

gate set, namely CNOT is the only allowed two-qubit gate, they show fault-tolerance is not possible when $p \geq 29.3\%$). We stress that all these results are incomparable as the models are different. For $p \leq 1/3$, we cannot rule out fault tolerance, but we can show a lower bound on the overhead as a function of the circuit depth.

1.3 Related work on limitations of fault-tolerance under strong noise

The results discussed below are not directly relevant to our work in the sense that they do not apply to noise levels below the fault-tolerant threshold. However, they establish impossibility results for noisy circuits and prove upper bounds on the threshold for fault-tolerant computation.

Razborov [42] considers p-depolarizing i.i.d. noise model and circuits with quantum gates which are allowed to be arbitrary quantum channels. He proves that if $p > 1 - 1/k$ then for any circuit in this model of width n and length T with gates of fan-in at most k, the distance between the states obtained by applying the circuit on any pair of input states decreases as $n2^{-\Omega(T)}$. In particular, any computation of length $T = \Omega(\log n)$ is essentially useless since the output is independent of the initial state. Such a strong notion of uselessness and upper bound on T as a function of n is possible in this case because of the strong assumptions: the noise is sufficiently high and classical information is subject to noise. If classical information can be stored perfectly, one can imagine a circuit in which part of the input is measured in the first step, stored, and then output at the end (as in the second strategy given above): for arbitrarily large values of T this circuit is not useless. Moreover, in our circuit model, there is no restriction on the gate set of the circuit, in particular the fan-in of the gates can be arbitrary. Our bound applies to the i.i.d. noise model for any non-unitary qubit noise channel. Finally, we point out that for the model considered by Razborov, our techniques can be adapted to recover similar bounds for a large class of quantum channels.

Using a different approach, Kempe *et al.* [30] improve the threshold upper bound to $1 - \Theta(1/\sqrt{k})$. However, their result is not quite comparable with Razborov's earlier result since they are obtained in different models. Kempe *et al.* consider circuits composed of arbitrary one qubit CPTP gates which are assumed to be essentially noiseless and k-qubit gates that are probabilistic mixtures of unitary operations preceded by independent p-depolarizing noise on their input qubits. They also assume that the output is the outcome of a measurement of a designated qubit (or a constant number of qubits) in the computational basis. They prove that if $p > 1 - \Theta(1/\sqrt{k})$ then for any circuit of length T in this model, the total variation distance between the output distributions for any pair of input states decreases as $2^{-\Omega(T)}$. Similar to Razborov's, their model does not allow noiseless classical computation and classically controlled operations. Note that a similar bound is known to hold for noisy classical circuits as well [24, 25].

When classical computation is assumed to be free and perfect, a measure of "uselessness" for noisy quantum circuits should naturally capture quantum phenomena that are not observed in classical circuits. One such measure that has been studied in earlier works [51] is being efficiently simulable classically. Along this direction, Buhrman *et al.* [16] show that quantum circuits consisting of noiseless stabilizer operations and arbitrary single qubit unitary gates followed by 45.3% depolarizing noise can be efficiently simulated by classical circuits. Their result is incomparable to ours for the same reasons above.

2   Preliminaries

Let $\mathcal{M}_d$ denote the set of $d \times d$ matrices with complex entries. Consider a quantum system described by the d-dimensional Hilbert space $\mathbb{C}^d$. Quantum states of the system are described by the set of *density operators* given by $\{\rho \in \mathcal{M}_d : \rho \geq 0, \mathrm{Tr}(\rho) = 1\}$. Pure state density operators are rank-one projectors $\rho = |\psi\rangle\langle\psi|$. A density operator ζ on $\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$ is called *separable* if it can be written as $\zeta = \sum_{i \in I} p_i \, \rho_i \otimes \sigma_i$ for some finite set I, a probability distribution $p : I \to [0,1]$, and density operators $\{\rho_i : i \in I\}$ on $\mathbb{C}^{d_A}$ and $\{\sigma_i : i \in I\}$ on $\mathbb{C}^{d_B}$. We denote by $\mathrm{Sep}(\mathrm{A} : \mathrm{B})$ the set of all separable state on $\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$ with respect to the bi-partition $A : B$, where the two subsystems are labeled A and B. A quantum state is called *entangled* if it is not separable.

Representation of Quantum channels.   A *quantum channel* is a completely positive and trace-preserving linear map $\mathcal{T} : \mathcal{M}_{d_A} \to \mathcal{M}_{d_B}$. We denote by $\mathcal{I}_d : \mathcal{M}_d \to \mathcal{M}_d$ the identity channel, i.e., the identity map on $\mathcal{M}_d$. Next, we review some basics about different representations of quantum channels (for more details see [55]). Let $\mathcal{T} : \mathcal{M}_{d_A} \to \mathcal{M}_{d_B}$ be a quantum channel. Let

$$C_\mathcal{T} = \sum_{i,j \in [d]} \mathcal{T}(|i\rangle\langle j|) \otimes |i\rangle\langle j| \quad , \tag{1}$$

where $\{|i\rangle : i \in [d_A]\}$ is the standard basis of $\mathbb{C}^{d_A}$. The operator $C_\mathcal{T}$ is called the *Choi matrix* of $\mathcal{T}$. The rank of $C_\mathcal{T}$ is called the Choi rank of $\mathcal{T}$. For any quantum channel $\mathcal{T}$, there exists a finite set I and linear operators $K_i : \mathbb{C}^{d_A} \to \mathbb{C}^{d_B}$ for any $i \in I$ such that $\sum_{i \in I} K_i^\dagger K_i = \mathbb{1}_{d_A}$ and $\mathcal{T}(X) = \sum_{i \in I} K_i X K_i^\dagger$, for all $X \in \mathcal{M}_{d_A}$. Any representation of this form is called a *Kraus representation* of the channel $\mathcal{T}$ and the operators $\{K_i\}_{i \in I}$ are called the corresponding *Kraus operators*. Note that Kraus representation in general is not unique, and the minimum number of Kraus operators in any Kraus representation of $\mathcal{T}$ is equal to the Choi rank of $\mathcal{T}$ [55, Corollary 2.23]. Important classes of quantum channels can be defined via their Kraus representation: A quantum channel is called *unitary* if it admits a Kraus decomposition with a single unitary Kraus operator, and it is called *entanglement breaking* if it admits a Kraus representation with Kraus operators K_i of rank 1. Consider quantum systems labelled A_1, A_2, B_1 and B_2. A bipartite quantum channel $\mathcal{T} : \mathcal{M}_{d_{A_1}} \otimes \mathcal{M}_{d_{B_1}} \to \mathcal{M}_{d_{A_2}} \otimes \mathcal{M}_{d_{B_2}}$ is called a *separable quantum channel* with respect to the bipartition $(A_1, A_2) : (B_1, B_2)$ if it admits a Kraus representation with Kraus operators of the form $K_i = K_i^A \otimes K_i^B$ with operators $K_i^A : \mathbb{C}^{d_{A_1}} \to \mathbb{C}^{d_{A_2}}$ and $K_i^B : \mathbb{C}^{d_{B_1}} \to \mathbb{C}^{d_{B_2}}$. We denote the set of these quantum channels by $\mathrm{SepC}((A_1, A_2) : (B_1, B_2))$. The set of quantum channels from $\mathcal{M}_{d_A}$ to $\mathcal{M}_{d_B}$ is compact and convex, and its extreme points are given by the quantum channels $\mathcal{T} : \mathcal{M}_{d_A} \to \mathcal{M}_{d_B}$ having a Kraus representation $\mathcal{T}(X) = \sum_i K_i X K_i^\dagger$ such that $\left\{K_i^\dagger K_j\right\}_{i,j}$ is a linearly independent set [18]. For any quantum channel $\mathcal{T} : \mathcal{M}_{d_A} \to \mathcal{M}_{d_B}$, there exists an isometry $V : \mathbb{C}^{d_A} \to \mathbb{C}^{d_B} \otimes \mathbb{C}^{d_E}$, for some $d_E \in \mathbb{N}$, such that $\mathcal{T}(X) = \mathrm{Tr}_E\left(V X V^\dagger\right)$, for every $X \in \mathbb{C}^{d_A}$. Any such representation of $\mathcal{T}$ is called a *Stinespring representation* and the isometry V is referred to as a *Stinespring dilation* of $\mathcal{T}$. The dimension d_E can be chosen to be equal to the Choi rank of $\mathcal{T}$ [55, Corollary 2.27].

Trace-norm and contraction coefficient.   The *induced trace norm* of a linear map $\mathcal{T} : \mathcal{M}_d \to \mathcal{M}_{d'}$ is defined as

$$\|\mathcal{T}\|_1 := \max\left\{\|\mathcal{T}(X)\|_1 : X \in \mathcal{M}_d, \|X\|_1 \leq 1\right\} \quad .$$

We will need the *trace-norm contraction coefficient* of a quantum channel $\mathcal{T} : \mathcal{M}_d \to \mathcal{M}_d$ given by

$$\eta_{\mathrm{tr}}\left(\mathcal{T}\right) := \sup_{\rho,\sigma} \frac{\|\mathcal{T}(\rho) - \mathcal{T}(\sigma)\|_1}{\|\rho - \sigma\|_1}, \tag{2}$$

where the supremum is over all pairs of quantum states. This quantity can be regarded as a quantum analogue of the Dobrushin ergodicity coefficient [21], which is an important tool in the study of Markov processes. We will often use that

$$\eta_{\mathrm{tr}}\left(\mathcal{T}\right) = \frac{1}{2} \max_{\psi \perp \phi} \|\mathcal{T}\left(|\psi\rangle\langle\psi| - |\phi\rangle\langle\phi|\right)\|_1 \;\;, \tag{3}$$

where the maximum is over orthogonal pairs of pure states (see [43]). In particular, $\eta_{\mathrm{tr}}(\mathcal{T}) = 1$ if and only if, there exists a pure state $\psi \in \mathbb{C}^d$ for which $(\mathcal{T}^\dagger \circ \mathcal{T})(|\psi\rangle\langle\psi|)$ is not full-rank.

Quantum capacity. The quantum capacity of a quantum channel is the maximum rate at which it can reliably transmit quantum information over asymptotically many uses of the channel.

▶ **Definition 3.** *For a quantum channel $\mathcal{T} : \mathcal{M}_{d_A} \to \mathcal{M}_{d_B}$, a communication rate R is called ϵ-achievable if there exists n_ϵ such that for all $n \geq n_\epsilon$, there is an encoding quantum channel $\mathcal{E}_n : \mathcal{M}_2^{\otimes Rn} \to \mathcal{M}_{d_A}^{\otimes n}$ and a decoding channel $\mathcal{D}_n : \mathcal{M}_{d_B}^{\otimes n} \to \mathcal{M}_2^{\otimes Rn}$ such that $\|(\mathcal{D}_n \circ \mathcal{T}^{\otimes n} \circ \mathcal{E}_n)(\rho) - \rho\|_1 \leq \epsilon$, for all $\rho \in \mathcal{M}_2^{\otimes Rn}$. A rate R is an achievable communication rate if it is ϵ-achievable for all $\epsilon \in (0, 2]$. The quantum capacity of $\mathcal{T}$ is the supremum over all achievable rates.*

Representation of qubit channels. Consider the set of single-qubit Pauli operators $\mathcal{P} = \{\mathbb{1}, \mathrm{X}, \mathrm{Y}, \mathrm{Z}\}$ defined as

$$\mathbb{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \; \mathrm{X} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \; \mathrm{Y} = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \; \mathrm{Z} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{4}$$

The set of 2×2 Hermitian matrices is a real vector space. One can easily verify that $\mathcal{P}$ forms an orthogonal basis for this set, with respect to the Hilbert-Schmidt inner product defined as $\langle A, B \rangle = \mathrm{Tr}\left(A^\dagger B\right)$. In particular, any single-qubit quantum state ρ can be written in the *Bloch sphere representation* as

$$\rho = \frac{1}{2}\left(\mathbb{1} + r \cdot \sigma\right) \;\;, \tag{5}$$

where $r = (r_x, r_y, r_z) \in \mathbb{R}^3$ with $\|r\|_2 \leq 1$ and σ is the vector of Pauli matrices $(\mathrm{X}, \mathrm{Y}, \mathrm{Z})$. It has been shown in [11], that the action of any qubit channel $\mathcal{N} : \mathcal{M}_2 \to \mathcal{M}_2$ can be written as $\mathcal{N}(\rho) = U\mathcal{C}_{[t,\lambda]}\left(V\rho V^\dagger\right)U^\dagger$, with

$$\mathcal{C}_{[t,\lambda]}\left(\frac{1}{2}\left[\mathbb{1} + r \cdot \sigma\right]\right) = \frac{1}{2}\left(\mathbb{1} + (t + \Lambda r) \cdot \sigma\right) \;\;, \tag{6}$$

where $t = (t_x, t_y, t_z) \in \mathbb{R}^3$ with $\|t\|_2 \leq 1$ and $\Lambda = \mathrm{diag}\left(\lambda\right)$ with $\lambda = (\lambda_x, \lambda_y, \lambda_z) \in \mathbb{R}^3$ and $\|\Lambda\| \leq 1$. The qubit channel $\mathcal{N}$ is *unital*, i.e., $\mathcal{N}(\mathbb{1}) = \mathbb{1}$, if and only if $t = 0$ in this representation. Imposing complete positivity of the unital map $\mathcal{C}_{[0,\lambda]}$ is equivalent to restricting the parameters $\lambda \in \mathbb{R}^3$ to a regular tetrahedron [11] with corners corresponding to

the conjugation by the Pauli operators $\mathbb{1}, X, Y, Z$, respectively. Specifically, the set of unital qubit channels is given by

$$\left\{ U \mathcal{C}_{[0,\lambda]} \left(V \rho V^{\dagger} \right) U^{\dagger} : U, V \, \text{unitary}, \lambda \in \text{conv}\{(1,1,1), (1,-1,-1), (-1,1,-1), (-1,-1,1)\} \right\}. \tag{7}$$

The middle points on the edges of the tetrahedron of Eq. (7) are the permutations of $(\pm 1, 0, 0)$, and they are the vertices of a regular octahedron corresponding to unital entanglement breaking qubit channels. The following lemma is straightforward from these geometric considerations and the fact that entanglement breaking channels stay entanglement breaking under composition by unitary quantum channels:

▶ **Lemma 4.** *Any unital qubit channel $\mathcal{N} : \mathcal{M}_2 \to \mathcal{M}_2$ can be written as a convex combination*

$$\mathcal{N} = (1-p)\mathcal{T} + p\mathcal{B},$$

where $\mathcal{T} : \mathcal{M}_2 \to \mathcal{M}_2$ is a unitary quantum channel and $\mathcal{B} : \mathcal{M}_2 \to \mathcal{M}_2$ is entanglement breaking.

3 Maximal length of noisy quantum computation

Let C be a quantum circuit of width n and length T defined by a sequence of quantum channels $\{\mathcal{L}_i\}_{i \in [T]}$. For $i \in \{0, \ldots, T\}$, let $(A_i : B_i)$ be a bipartition of the registers of the circuit after i time steps into disjoint subsets such that $\mathcal{L}_i \in \text{SepC}\,(A_{i-1}, A_i : B_{i-1}, B_i)$, i.e., $\mathcal{L}_i$ in each time step is a separable channel with respect to the bipartition of the subsystems into the A and B parts. Let $\mathcal{N}$ be an arbitrary non-unitary qubit channel and consider the noisy implementation of C in the i.i.d.$_{\mathcal{N}}$ model. We prove that i.i.d.$_{\mathcal{N}}$(C) is not capable of preserving the entanglement between the A and B subsystems for more than exponential time in the width of C.

We denote by ρ_{AB}^i the state of the noisy circuit after i time steps. Let d_{Sep}^1 denote the 1-norm distance from the set of separable states $\text{Sep}\,(A : B)$ defined for a quantum state τ_{AB} as

$$d_{\text{Sep}}^1 (\tau_{AB}) \coloneqq \min_{\sigma \in \text{Sep}(A:B)} \|\tau - \sigma\|_1 \ . \tag{8}$$

Note that d_{Sep}^1 is a convex and faithful measure of entanglement, i.e., $d_{\text{Sep}}^1 (\tau_{AB}) > 0$ if and only if τ_{AB} is entangled, and it is monotone under separable quantum channels. The following lemma is our main technical contribution which is of independent interest.

▶ **Lemma 5.** *For every non-unitary qubit channel $\mathcal{N}$, there exists a constant $p \in (0, 1]$ such that $d_{\text{Sep}}^1 \left(\rho_{AB}^{i+1} \right) \leq (1 - p^n) \, d_{\text{Sep}}^1 \left(\rho_{AB}^i \right)$, for every $i \in \{0, \ldots, T-1\}$. If $\mathcal{N}$ is unital and non-unitary, then the statement is valid even in the circuit model with free and noiseless classical computation.*

Proof. Let n_i denote the number qubits of the circuit at time i. In time step $i + 1$, each of the n_i qubit in the noisy circuit goes through the noise channel $\mathcal{N} : \mathcal{M}_2 \to \mathcal{M}_2$ followed by a separable quantum channel $\mathcal{L}_{i+1} \in \text{SepC}\,(A_i, A_{i+1} : B_i, B_{i+1})$. We use two different approaches to prove Lemma 5 for unital and non-unital quantum channels.
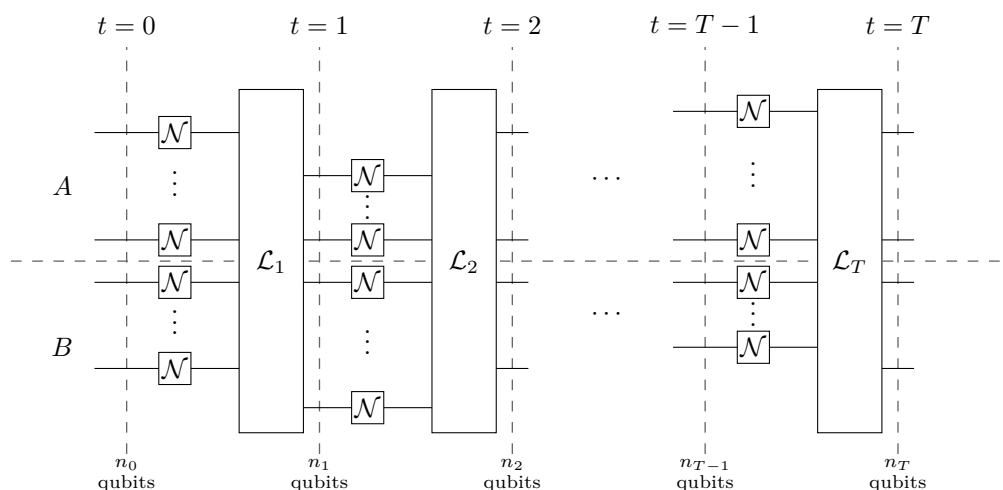
Figure 2 An illustration of i.i.d.$_{\mathcal{N}}(C)$. In each time step, every qubit is subject to the noise channel $\mathcal{N}$. The qubits are partitioned into disjoint subsets A and B such that the quantum channels $\mathcal{L}_i$ are separable with respect to this bipartition. In the model with free and noiseless classical computation, the classical subsystems are assumed to be noise-free.

Case I: $\mathcal{N}$ is non-unital. In this case, we prove the contractivity of d^1_{Sep} via the following simple lemma in terms of the trace-norm contraction coefficient of the channel $\mathcal{N}$.

▶ **Lemma 6.** *For any separable quantum channel $\mathcal{T} : \mathcal{M}_{d_{A_1}} \otimes \mathcal{M}_{d_{B_1}} \to \mathcal{M}_{d_{A_2}} \otimes \mathcal{M}_{d_{B_2}}$ with respect to systems A and B, we have*

$$d^1_{\mathrm{Sep}} \left(\mathcal{T} \left(\rho_{AB} \right) \right) \leq \eta_{tr} \left(\mathcal{T} \right) d^1_{\mathrm{Sep}} \left(\rho_{AB} \right) \tag{9}$$

Proof. Consider

$$\sigma'_{AB} = \operatorname*{argmin}_{\sigma_{AB} \in \mathrm{Sep}(A:B)} \| \rho_{AB} - \sigma_{AB} \|_1 \ . \tag{10}$$

Then, we have

$$d^1_{\mathrm{Sep}} \left(\mathcal{T} \left(\rho_{AB} \right) \right) \leq \| \mathcal{T}(\rho_{AB}) - \mathcal{T}(\sigma'_{AB}) \|_1 \leq \eta_{\mathrm{tr}} \left(\mathcal{T} \right) d^1_{\mathrm{Sep}} \left(\rho_{AB} \right) \ . \tag{11}$$

The definition (2) of the trace-norm contraction coefficient might not be easily computable for all channels, and it is unclear how it behaves under tensor powers. In order to use Lemma 6 for the channel $\mathcal{T} = \mathcal{N}^{\otimes 2n}$, we prove the following upper bound:

▶ **Lemma 7.** *For any quantum channel $\mathcal{T} : \mathcal{M}_d \to \mathcal{M}_d$, we have*

$$\eta_{tr}(\mathcal{T}) \leq \sqrt{1 - \frac{\lambda^{\mathrm{out}}_{\min} \left(\mathcal{T}^\dagger \circ \mathcal{T} \right)}{d^2}}. \tag{12}$$

Here, we used the minimal output eigenvalue $\lambda^{\mathrm{out}}_{\min} \left(\mathcal{T}^\dagger \circ \mathcal{T} \right) = \min_{\psi, \phi} \langle \psi | \left(\mathcal{T}^\dagger \circ \mathcal{T} \right)(|\phi\rangle\langle\phi|) |\psi\rangle$, where the minimum goes over pure quantum states.

Proof. In [54] the fidelity $\mathrm{F}(\rho, \sigma) = \| \sqrt{\rho}\sqrt{\sigma} \|_1^2$ of quantum states ρ and σ on $\mathbb{C}^d$ is expressed as a semidefinite program

$$\mathrm{F}(\rho, \sigma) = \max \left\{ \langle v | W | v \rangle \ : \ W_{AE} \geq 0, \mathrm{Tr}_E \left(W_{AE} \right) \leq \rho \right\}, \tag{13}$$

where $v \in \mathbb{C}^d \otimes \mathbb{C}^{d_E}$ is a purification of σ. Consider the case $\rho = \mathcal{T}(|\psi\rangle\langle\psi|)$ and $\sigma = \mathcal{T}(|\phi\rangle\langle\phi|)$ for pure states $|\psi\rangle, |\phi\rangle \in \mathbb{C}^d$. Let $V : \mathbb{C}^d \to \mathbb{C}^d \otimes \mathbb{C}^{d_E}$ denote a Stinespring dilation of $\mathcal{T}$ such that

$$\mathcal{T}(X) = \mathrm{Tr}_E \left(V X V^\dagger \right) . \tag{14}$$

Note that $V|\phi\rangle \in \mathbb{C}^d \otimes \mathbb{C}^{d_E}$ is a purification of σ. Choosing $W = \mathcal{T}(|\psi\rangle\langle\psi|) \otimes \frac{\mathbb{1}_E}{d_E}$ we find that

$$\mathrm{F}(\rho, \sigma) \geq \langle\phi| V^\dagger \left(\mathcal{T}(|\psi\rangle\langle\psi|) \otimes \frac{\mathbb{1}_E}{d_E} \right) V |\phi\rangle = \frac{1}{d_E} \langle\phi| \mathcal{T}^\dagger \circ \mathcal{T} (|\psi\rangle\langle\psi|) |\phi\rangle , \tag{15}$$

where we used that the adjoint of $\mathcal{T}$ has the Stinespring dilation

$$\mathcal{T}^\dagger(X) = V^\dagger (X \otimes \mathbb{1}_E) V . \tag{16}$$

Finally, we can apply the well-known Fuchs-van-de-Graaf inequality relating the trace norm and the fidelity ($\|\rho - \sigma\|_1 \leq \sqrt{1 - \mathrm{F}(\rho, \sigma)}$ for any quantum states ρ, σ) to conclude

$$
\begin{aligned}
\eta_{\mathrm{tr}}(\mathcal{T}) &= \frac{1}{2} \max_{\psi \perp \phi} \|\mathcal{T}(|\psi\rangle\langle\psi| - |\phi\rangle\langle\phi|)\|_1 \\
&\leq \max_{\psi \perp \phi} \sqrt{1 - \mathrm{F}(\mathcal{T}(|\psi\rangle\langle\psi|), \mathcal{T}(|\phi\rangle\langle\phi|))} \\
&\leq \max_{\psi \perp \phi} \sqrt{1 - \frac{1}{d^2} \langle\phi| (\mathcal{T}^\dagger \circ \mathcal{T})(|\psi\rangle\langle\psi|) |\phi\rangle} \\
&\leq \sqrt{1 - \frac{\lambda_{\min}^{\mathrm{out}} (\mathcal{T}^\dagger \circ \mathcal{T})}{d^2}} ,
\end{aligned}
$$

where we used that $d_E \leq d^2$ for any quantum channel $\mathcal{T} : \mathcal{M}_d \to \mathcal{M}_d$. ◀

Computing $\lambda_{\min}^{\mathrm{out}} \left(\mathcal{T}^\dagger \circ \mathcal{T} \right)$ might be difficult in general, but for our purposes an easy lower bound suffices. This bound also behaves nicely under tensor product of channels; an important property which we use below. For $|\phi\rangle = \sum_{i \in [d]} \alpha_i |i\rangle$, let $|\overline{\phi}\rangle = \sum_{i \in [d]} \overline{\alpha_i} |i\rangle$, where $\overline{\alpha_i}$ denotes the complex conjugate of α_i. Then we have

$$\lambda_{\min}^{\mathrm{out}} \left(\mathcal{T}^\dagger \circ \mathcal{T} \right) = \min_{\psi, \phi} \langle\psi| (\mathcal{T}^\dagger \circ \mathcal{T})(|\phi\rangle\langle\phi|) |\psi\rangle = \min_{\psi, \phi} \langle\psi \otimes \overline{\phi}| C_{\mathcal{T}^\dagger \circ \mathcal{T}} |\psi \otimes \overline{\phi}\rangle \geq \lambda_{\min} (C_{\mathcal{T}^\dagger \circ \mathcal{T}}). \tag{17}$$

Finally, we are in position to present the proof of the case where $\mathcal{N}$ is a non-unital qubit channel. We may write

$$\mathcal{N} = (1 - q) \mathcal{M}' + q \mathcal{M}, \tag{18}$$

where $q \in (0, 1]$ and $\mathcal{M}$ is a non-unital extreme point of the set of qubit channels. By the characterization of extreme points of the set of quantum channels (see preliminaries or [18]), the non-unital extremal qubit channel $\mathcal{M}$ admits a Kraus decomposition with two Kraus operators $K_1, K_2 : \mathbb{C}^2 \to \mathbb{C}^2$ such that the operators $K_1^\dagger K_1, K_1^\dagger K_2, K_2^\dagger K_1, K_2^\dagger K_2$ are linearly independent. Since these operators are exactly the Kraus operators of the completely positive map $\mathcal{M}^\dagger \circ \mathcal{M}$, we conclude that the Choi matrix $C_{\mathcal{M}^\dagger \circ \mathcal{M}}$ is full-rank and that

$$\lambda_{\min} (C_{\mathcal{N}^\dagger \circ \mathcal{N}}) \geq q^2 \lambda_{\min} (C_{\mathcal{M}^\dagger \circ \mathcal{M}}) > 0.$$

Setting $\mathcal{T} = \mathcal{N}^{\otimes n_i}$ and using (17) and Lemma 7, we have

$$\eta_{\mathrm{tr}}(\mathcal{T}) \leq \sqrt{1 - \frac{\lambda_{\min}\left(C_{\mathcal{T}^\dagger \circ \mathcal{T}}\right)}{2^{2n_i}}} = \sqrt{1 - \frac{\lambda_{\min}\left((C_{\mathcal{N}^\dagger \circ \mathcal{N}})^{\otimes n_i}\right)}{2^{2n_i}}} = \sqrt{1 - \left(\frac{\lambda_{\min}\left(C_{\mathcal{N}^\dagger \circ \mathcal{N}}\right)}{4}\right)^{n_i}}$$

$$\leq \sqrt{1 - \left(\frac{q^2 \lambda_{\min}\left(C_{\mathcal{M}^\dagger \circ \mathcal{M}}\right)}{4}\right)^{n_i}} \leq 1 - \left(\frac{q^2 \lambda_{\min}\left(C_{\mathcal{M}^\dagger \circ \mathcal{M}}\right)}{2}\right)^{n_i} \ .$$

By (9), for $p = \left(q^2/2\right) \lambda_{\min}\left(C_{\mathcal{M}^\dagger \otimes \mathcal{M}}\right) > 0$, we have

$$d_{\mathrm{Sep}}^1\left(\rho^{i+1}\right) = d_{\mathrm{Sep}}^1\left(\left(\mathcal{L}_{i+1} \circ \mathcal{N}^{\otimes n_i}\right)\left(\rho^i\right)\right)$$
$$\leq d_{\mathrm{Sep}}^1\left(\left(\mathcal{N}^{\otimes n_i}\right)\left(\rho^i\right)\right) \leq \left(1 - p^{n_i}\right) d_{\mathrm{Sep}}^1\left(\rho^i\right) \leq \left(1 - p^n\right) d_{\mathrm{Sep}}^1\left(\rho^i\right) \ ,$$

where the first inequality holds by monotonicity of d_{Sep}^1 under separable quantum channels.

Case II: $\mathcal{N}$ is unital. We prove the statement of the lemma in the circuit model with free and noiseless classical computation. For every $i \in \{0, \dots, T\}$, let $A_i = A_i^c A_i^q$ and $B_i = B_i^c B_i^q$, where A_i^c and B_i^c are classical subsystems of A_i and B_i, respectively, of arbitrary finite dimensions and A_i^q and B_i^q are the quantum subsystems which contain n_i qubits. Note that only the quantum subsystems are subject to the noise channel $\mathcal{N}$. By Lemma 4, we may write

$$\mathcal{N} = (1 - p)\,\mathcal{T} + p\,\mathcal{B}, \tag{19}$$

where $\mathcal{T} : \mathcal{M}_2 \to \mathcal{M}_2$ is a unitary quantum channel and $\mathcal{B} : \mathcal{M}_2 \to \mathcal{M}_2$ is entanglement breaking. Since $\mathcal{N}$ itself is not unitary, we have $p > 0$. There exists a separable quantum channel $\mathcal{S} : \mathcal{M}_{2^{n_i}} \to \mathcal{M}_{2^{n_i}}$ with respect to the bipartition into A and B, such that

$$\mathcal{N}^{\otimes n_i} = (1 - p^{n_i})\,\mathcal{S} + p^{n_i}\,\mathcal{B}^{\otimes n_i} \ , \tag{20}$$

and we can compute

$$d_{\mathrm{Sep}}^1\left(\rho^{i+1}\right) = d_{\mathrm{Sep}}^1\left(\left(\mathcal{L}_{i+1} \circ \left(\mathcal{N}^{\otimes n_i} \otimes \mathcal{I}^{A_i^c B_i^c}\right)\right)\left(\rho^i\right)\right) \leq d_{\mathrm{Sep}}^1\left(\left(\mathcal{N}^{\otimes n_i} \otimes \mathcal{I}^{A_i^c B_i^c}\right)\left(\rho^i\right)\right) \tag{21}$$

$$\leq (1 - p^{n_i}) d_{\mathrm{Sep}}^1\left(\left(\mathcal{S} \otimes \mathcal{I}^{A_i^c B_i^c}\right)\left(\rho^i\right)\right) + p^{n_i} d_{\mathrm{Sep}}^1\left(\left(\mathcal{B}^{\otimes n_i} \otimes \mathcal{I}^{A_i^c B_i^c}\right)\left(\rho^i\right)\right) \tag{22}$$

$$= (1 - p^{n_i}) d_{\mathrm{Sep}}^1\left(\left(\mathcal{S} \otimes \mathcal{I}^{A_i^c B_i^c}\right)\left(\rho^i\right)\right) \leq \left(1 - p^n\right) d_{\mathrm{Sep}}^1\left(\rho^i\right) \ . \tag{23}$$

Here, the first and the last inequalities follow from monotonicity of d_{Sep}^1 under separable quantum channels, and the second inequality follows from convexity of d_{Sep}^1. Moreover, the second equality holds since $\mathcal{B}$ is entanglement breaking and ρ^i is classical on $A_i^c B_i^c$. This concludes the proof. ◀

Recall the setup of Theorem 1: Let C be an arbitrary quantum circuit in which the gates are allowed to be arbitrary quantum channels acting on several qubits. A supply of fresh ancilla qubits is available and the number of ancillas in C in each time step is restricted by the circuit width. We will now consider two parallel instances of the circuit and apply them to an entangled state. If the parallel instances of the circuits are affected by the i.i.d.$_{\mathcal{N}}$ error model for a non-unitary qubit channel $\mathcal{N}$, then they cannot keep entanglement for arbitrary long time. Theorem 1 will then follow from the following lemma:

▶ **Lemma 8** ([55, Theorem 3.56]). *Let $\mathcal{I}_d : \mathcal{M}_d \to \mathcal{M}_d$ be the identity channel. Let $\mathcal{T} : \mathcal{M}_d \to \mathcal{M}_d$ be a quantum channel satisfying $\|\mathcal{T} - \mathcal{I}_d\|_1 \leq \epsilon$, for some $\epsilon \in [0, 2]$. Then for any $d' \geq 1$, we have*

$$\|\mathcal{T} \otimes \mathcal{I}_{d'} - \mathcal{I}_d \otimes \mathcal{I}_{d'}\|_1 \leq \sqrt{2\epsilon}.$$

This in particular implies that $\left\|\mathcal{T}^{\otimes 2} - \mathcal{I}_d^{\otimes 2}\right\|_1 \leq 2\left\|\mathcal{T} \otimes \mathcal{I}_d - \mathcal{I}_d \otimes \mathcal{I}_d\right\|_1 \leq 2\sqrt{2\epsilon}$.

Combining Lemmas 5 and 8, we have:

▶ **Theorem 1** (**Restated**). *Let $\mathcal{I}$ denote the qubit identity channel and $\mathcal{N}$ be a non-unitary qubit channel. Then there exists a constant $p \in (0, 1]$ only depending on $\mathcal{N}$ such that the following holds: Let C be an arbitrary circuit of length T and width n such that $T \geq (1/p)^{2n}$. Then for any pair of ideal encoding and decoding maps $\mathcal{E}$ and $\mathcal{D}$ (of appropriate input and output dimensions), we have*

$$\|\mathcal{D} \circ \text{i.i.d.}_{\mathcal{N}}(C) \circ \mathcal{E} - \mathcal{I}\|_1 \geq \epsilon_0 \ ,$$

for some universal constant $\epsilon_0 \geq 1/128$. For any unital (non-unitary) noise channel $\mathcal{N}$, this statement holds even in the circuit model with free and noiseless classical computation.

Proof. Consider two parallel instances of the circuit C. By definition of the i.i.d. error model, we have $\text{i.i.d.}_{\mathcal{N}}\left(C^{\otimes 2}\right) = (\text{i.i.d.}_{\mathcal{N}}(C))^{\otimes 2}$. Note that $C^{\otimes 2}$ is a circuit of width $2n$ and length T satisfying the requirements of Lemma 5, with A being the qubits of the first copy of C and B being the qubits of the second copy. For $T \geq (1/p)^{2n}$, by Lemma 5, for any input state ρ^0, we have $d_{\text{Sep}}^1\left((\text{i.i.d.}_{\mathcal{N}}(C))^{\otimes 2}(\rho^0)\right) \leq 1/4$. In particular, for any encoding map $\mathcal{E}$ and any 2-qubit state σ, we have $d_{\text{Sep}}^1\left(\left((\text{i.i.d.}_{\mathcal{N}}(C))^{\otimes 2} \circ \mathcal{E}^{\otimes 2}\right)(\sigma)\right) \leq 1/4$. For any decoding map $\mathcal{D}$, since $\mathcal{D}^{\otimes 2}$ is a separable channel, we have

$$d_{\text{Sep}}^1\left(\left(\mathcal{D}^{\otimes 2} \circ (\text{i.i.d.}_{\mathcal{N}}(C))^{\otimes 2} \circ \mathcal{E}^{\otimes 2}\right)(\sigma)\right) \leq d_{\text{Sep}}^1\left((\text{i.i.d.}_{\mathcal{N}}(C))^{\otimes 2} \circ \mathcal{E}^{\otimes 2}\right)(\sigma) \leq 1/4.$$

Let $\mathcal{M} := \mathcal{D} \circ \text{i.i.d.}_{\mathcal{N}}(C) \circ \mathcal{E}$. Then we have

$$\left\|\mathcal{M}^{\otimes 2} - \mathcal{I}^{\otimes 2}\right\|_1 = \sup_{\sigma} \left\|\mathcal{M}^{\otimes 2}(\sigma) - \sigma\right\|_1 \geq \sup_{\sigma} \left(d_{\text{Sep}}^1(\sigma) - d_{\text{Sep}}^1(\mathcal{M}^{\otimes 2}(\sigma))\right) \ .$$

For σ equal to the maximally entangled state, it is straightforward to see that $d_{\text{Sep}}^1(\rho) \geq 1/2$. Hence, we have $\left\|\mathcal{M}^{\otimes 2} - \mathcal{I}^{\otimes 2}\right\|_1 \geq 1/4$. The statement of the theorem follows from Lemma 8 for $\epsilon_0 \geq 1/128$. Finally, note that by Lemma 5, for unital $\mathcal{N}$, using a similar argument the same conclusion can be made even if classical computation is assumed to be free and noiseless. ◀

We should point out a connection between our proof of Theorem 1 (including Lemma 7) and the zero-error classical capacity of a quantum channel [36, 8]. The zero-error classical capacity $C_0(\mathcal{N})$ of a quantum channel $\mathcal{N}$ arises as the regularization of the logarithm of the (quantum) *independence number* $\alpha(\mathcal{N})$ which equals the maximal number of orthogonal pure states that are mapped to orthogonal outputs by the channel $\mathcal{N}$ (see [23] for details). It is clear from this definition that $\alpha(\mathcal{N}) = 1$ if and only if $\eta_{\text{tr}}(\mathcal{N}) < 1$. Qualitatively, Theorem 1 says that a finite-size quantum memory affected by a non-unitary qubit noise channel $\mathcal{N}$ cannot store information for arbitrary long times. Using our proof technique, this statement can be generalized beyond qubit noise channels as follows: If $\mathcal{N}$ arises as a non-trivial convex combination with an entanglement breaking quantum channel or satisfies

$\alpha(\mathcal{N}^{\otimes k}) = 1$, for some $k \in \mathbb{N}$, then no quantum memory of width k affected by noise $\mathcal{N}$ can store information for arbitrary long times. In particular, if $C_0(\mathcal{N}) = 0$, then no finite width will allow for storing information for arbitrary long times. The zero-error classical capacity can be superactivated in the sense that there are quantum channels $\mathcal{N}$ for which $\alpha(\mathcal{N}) = 1$ but $\alpha(\mathcal{N}^{\otimes k}) > 1$ for some $k \in \mathbb{N}$ (see [22, 19]), and it would be interesting to study the maximal storage times of quantum memories for such examples. Our techniques in Lemma 7 can be used to show more generally that such a superactivation does not happen for channels of maximal Kraus rank. We remark that the impossibility of superactivation of the zero-error classical capacity has been shown previously in [38, 45], for qubit channels. However, unlike Lemma 7 these results did not give any quantitative estimate on the trace-norm contraction coefficient and therefore they do not directly imply our main quantitative bound.

4 Implications for fault-tolerant quantum computation

4.1 A high-level definition of fault tolerance

A *quantum fault tolerance scheme* is a circuit encoding FT which maps any ideal quantum circuit C to a family of its *fault-tolerant simulation* circuits $FT^\ell(C) \coloneqq FT(\ell, C)$. As defined formally below, the circuit encoding is designed to allow the implementation of the circuit C even with faulty components.

Let $T = Length(C)$ denote the length of the circuit C. We denote by $n_t(C)$ the number of (logical) qubits in the circuit C after $t \in \{0, 1, \ldots, T\}$ time steps and use $n = Width(C)$ to denote the width of C, i.e., $n = \max_t n_t(C)$. Similarly, let $n' = Width(FT^\ell(C))$ be the width, i.e, the maximum number of (physical) qubits and $T' = Length(FT^\ell(C))$ be the length of $FT^\ell(C)$. Note that it is important to distinguish between the classical and quantum subsystems, as in the circuit model with free and noiseless classical computation only quantum subsystems are subject to the noise. For the input circuit C, we may assume without loss of generality that there are no classical inputs, as we can consider the circuit obtained by conditioning on the state of the classical input systems.

▶ **Definition 2.** *A circuit encoding FT is a fault tolerance scheme against an error model $\mathcal{E}$ if there exist a family of ideal encoding maps $\{\mathcal{E}_n^\ell\}$ and decoding maps $\{\mathcal{D}_n^\ell\}$ only depending on FT such that for every $\epsilon \in (0, 2]$ and every quantum circuit C, there exists ℓ_0 such that*

$$\left\| \mathcal{D}_{n_T(C)}^\ell \circ \mathcal{E}\left(FT^\ell(C)\right) \circ \mathcal{E}_{n_0(C)}^\ell - C \right\|_1 \le \epsilon, \quad \forall \ell \ge \ell_0 \ . \tag{24}$$

In the above definition, if the input Hilbert space of C is trivial (e.g., if C is a circuit for classical computation), then for every ℓ, the circuit $FT^\ell(C)$ is also naturally defined to have no input, i.e., $\mathcal{E}_0^\ell$ is fixed to be the identity map on $\mathbb{C}$.

Formal definitions of fault tolerance are indeed invaluable in studying general properties and limitations of fault tolerance schemes. We emphasize that Definition 2 is far from capturing all the different aspects of quantum fault tolerance. Indeed, there are several different ways of defining fault tolerance, each with their own advantages and limitations. Here, we aim at providing a high-level definition with minimal assumptions which is easy to work with and applies to most known and potential schemes. Below we discuss some examples of known fault tolerance protocols and explain how they fit into our definition.

Perhaps the most well-known family of fault tolerance schemes are the schemes based on concatenation. These schemes typically involve a quantum error correcting code Q together with fault-tolerant gadgets for the elementary locations, i.e., state preparation gadgets, gate gadgets for a universal set of gates, storage gadgets, and measurement gadgets, in addition to

error correction gadgets. For a quantum circuit C, the circuit $FT(C)$ is obtained by replacing the locations in C by the corresponding gadgets, interlaced with error correction gadgets on the quantum registers, which now correspond to qubits of C encoded in Q. The gadgets are designed to satisfy some fault tolerance properties, such that they not only act correctly in the absence of faults, but also control the propagation of errors if the number of errors is not too high. For $\ell \geq 1$, let $FT^\ell(C) = FT\left(FT^{\ell-1}(C)\right)$, where we define $FT^0(C) = C$. For any such scheme, assuming the distance between the noise channel $\mathcal{N}$ and the identity channel is below a threshold, the output state of i.i.d.$_{\mathcal{N}}\left(FT^\ell(C)\right)$ is arbitrarily close to an encoded version of the output state of C for a sufficiently large ℓ [3].

Fault tolerance schemes based on topological codes are another family of schemes which have a threshold noise level below which arbitrarily long quantum computations are possible [20, 52, 41]. For the surface code [32] (and more generally, low-density parity check codes), a single syndrome bit measurement only involves a few neighboring qubits. This property allows error correction to be performed more efficiently for these codes [49, 14]. Moreover, topological properties of surface codes allow for simpler and more efficient implementation of fault-tolerant gadgets for elementary locations [20, 17, 29, 13]. This in turn results in a desirable scaling of the overhead and reduction of the logical error rate when the size of the underlying surface code is increased. In this case, the index ℓ in definition 2 corresponds to the size of the surface code.

4.2   A lower bound on the space overhead for fault tolerance

For a fault-tolerant simulation circuit $FT^\ell(C)$, the ratios n'/n and T'/T are respectively called the space overhead and the time overhead of $FT^\ell(C)$. We view a fault tolerance scheme as a compiler which reads the instructions given by the input circuit C and encodes them into the fault-tolerant simulation circuits $FT^\ell(C)$, without optimizing the circuit C. This should be true even if the input circuit involves unnecessary steps that are completely independent of the output of the circuit. While in Eq. (24), we only require the circuit $FT^\ell(C)$ to (approximately) implement the *overall* logical operation of the circuit C on a code space, this "faithfulness" property is in fact necessary to have a meaningful notion of overhead for a fault tolerance scheme. As a result of this property, we may assume $n' \geq n$ and $T' \geq T$, for any input circuit C and any fault-tolerant simulation circuit $FT^\ell(C)$. The faithfulness property also rules out circuit encodings which map quantum circuits for classical computation to fully classical (hence, noiseless) circuits.

Our result of the previous section can be used to prove a lower bound on the space overhead of any universal fault tolerance scheme satisfying Definition 2 and the natural requirement that the overhead is at least 1 for any fault-tolerant simulation circuit.

▶ **Theorem 2** (**Restated**). *Let $\mathcal{N}$ be a non-unitary qubit channel. For any fault tolerance scheme against the* i.i.d.$_{\mathcal{N}}$ *noise model, the number of physical qubits is at least*

$$\max\left\{Q(\mathcal{N})^{-1}n, \alpha_{\mathcal{N}} \log T\right\} \ ,$$

for circuits of length T and width n, and any constant accuracy $\epsilon \leq 1/128$. Here, $Q(\mathcal{N})$ denotes the quantum capacity of $\mathcal{N}$ and $\alpha_{\mathcal{N}} = \frac{1}{2\log 1/p}$, where $p \in (0,1]$ is a constant only depending on the channel $\mathcal{N}$. In particular, fault tolerance is not possible even for a single time step if $\mathcal{N}$ has zero quantum capacity or the space overhead is strictly less than $Q(\mathcal{N})^{-1}$. When $\mathcal{N}$ is a unital (non-unitary) qubit channel, the statement holds even in the circuit model with free and noiseless classical computation.

Proof. Let C be a quantum circuit of length T on n qubits in which in every time step an identity gate is applied on every qubit. Consider a fault tolerance scheme FT satisfying Definition 2. For any constant $\epsilon \in (0, 1/128)$, let $C'(\epsilon)$ be a fault-tolerant encoding of C given by FT satisfying Eq. (24), for a pair of encoding and decoding maps $(\mathcal{E}_\epsilon, \mathcal{D}_\epsilon)$. Then by Theorem 1, the width of $C'(\epsilon)$ is bounded in terms of its depth as $n' \geq \frac{\log T'}{2 \log 1/p}$. Since $T' \geq T$ the lower bound of $\alpha_\mathcal{N} \log T$ follows for $\alpha_\mathcal{N} = \frac{1}{2 \log 1/p}$.

For the second lower bound, consider the quantum channel obtained by removing the last layer of noise from i.i.d.$_\mathcal{N}$ $(C'(\epsilon))$ and let $\mathcal{E}'_\epsilon$ denote the composition of $\mathcal{E}_\epsilon$ with this channel. Then for the family of encoding and decoding maps $(\mathcal{E}'_\epsilon, \mathcal{D}_\epsilon)$, the overall channel $\mathcal{T}_\epsilon = \mathcal{D}_\epsilon \circ \mathcal{N}^{\otimes n_\epsilon} \circ \mathcal{E}'_\epsilon$ approximates the qubit identity channel up to arbitrarily high accuracy. But continuity of the quantum capacity [35] implies that, for sufficiently small ϵ, the channel $\mathcal{T}_\epsilon$ (and hence $\mathcal{N}$) has a non-zero quantum capacity. Moreover, the existence of the family of encoding and decoding maps $(\mathcal{E}'_\epsilon, \mathcal{D}_\epsilon)$ implies that, for the channel $\mathcal{N}$, the inverse of the space overhead is an ϵ-achievable communication rate for any arbitrarily small constant ϵ (See Definition 3). Therefore, by the definition of quantum capacity, we have $n' \geq \frac{1}{Q(\mathcal{N})} n$. Since assistance by noiseless forward classical communication does not increase the quantum capacity of a channels [7, 10], the lower bound of $Q(\mathcal{N})^{-1} n$ is still valid in the circuit model with free and noiseless classical computation. Finally, note that the above argument is valid even if there is a single time step of noisy operations. Therefore, fault tolerance is not achievable even for a single time step if $\mathcal{N}$ has zero quantum capacity, or with an space overhead strictly less than $1/Q(\mathcal{N})$. ◀

---- **References** ----

1   D. Aharonov and M. Ben-Or. Fault-tolerant quantum computation with constant error. In *Proceedings of the Twenty-Ninth Annual ACM Symposium on Theory of Computing*, STOC '97, pages 176–188, New York, NY, USA, 1997. Association for Computing Machinery. `doi:10.1145/258533.258579`.

2   D. Aharonov, M. Ben-Or, R. Impagliazzo, and N. Nisan. Limitations of noisy reversible computation. Technical Report arXiv:9611028 [quant-ph], arXiv.org, `https://arxiv.org/abs/quant-ph/9611028`, 1996.

3   Panos Aliferis, Daniel Gottesman, and John Preskill. Quantum accuracy threshold for concatenated distance-3 codes. *Quantum Info. Comput.*, 6(2):97–165, March 2006.

4   Panos Aliferis, Daniel Gottesman, and John Preskill. Accuracy threshold for postselected quantum computation. *Quantum Info. Comput.*, 8(3):181–244, March 2008.

5   Panos Aliferis and John Preskill. Fault-tolerant quantum computation against biased noise. *Phys. Rev. A*, 78:052331, November 2008. `doi:10.1103/PhysRevA.78.052331`.

6   Panos Aliferis and John Preskill. Fibonacci scheme for fault-tolerant quantum computation. *Phys. Rev. A*, 79:012332, January 2009. `doi:10.1103/PhysRevA.79.012332`.

7   H. Barnum, E. Knill, and M.A. Nielsen. On quantum fidelities and channel capacities. *IEEE Transactions on Information Theory*, 46(4):1317–1329, July 2000. `doi:10.1109/18.850671`.

8   Salman Beigi and Peter W Shor. On the complexity of computing zero-error and holevo capacity of quantum channels. *arXiv preprint arXiv:0709.2090*, 2007.

9   Michael Ben-Or, Daniel Gottesman, and Avinatan Hassidim. Quantum refrigerator. *arXiv preprint arXiv:1301.1995*, 2013.

10 Charles H. Bennett, David P. DiVincenzo, John A. Smolin, and William K. Wootters. Mixed-state entanglement and quantum error correction. *Physical Review A*, 54(5):3824–3851, November 1996. `doi:10.1103/physreva.54.3824`.

11 Mary Beth Ruskai, Stanislaw Szarek, and Elisabeth Werner. An analysis of completely-positive trace-preserving maps on m2. *Linear Algebra and its Applications*, 347(1):159–187, 2002. `doi:10.1016/S0024-3795(01)00547-X`.

12   Peter Brooks and John Preskill. Fault-tolerant quantum computation with asymmetric bacon-shor codes. *Phys. Rev. A*, 87:032310, March 2013. `doi:10.1103/PhysRevA.87.032310`.

13   Benjamin J. Brown, Katharina Laubscher, Markus S. Kesselring, and James R. Wootton. Poking holes and cutting corners to achieve clifford gates with the surface code. *Phys. Rev. X*, 7:021029, May 2017. `doi:10.1103/PhysRevX.7.021029`.

14   Benjamin J. Brown, Daniel Loss, Jiannis K. Pachos, Chris N. Self, and James R. Wootton. Quantum memories at finite temperature. *Rev. Mod. Phys.*, 88:045005, November 2016. `doi:10.1103/RevModPhys.88.045005`.

15   Dagmar Bruss, David P. DiVincenzo, Artur Ekert, Christopher A. Fuchs, Chiara Macchiavello, and John A. Smolin. Optimal universal and state-dependent quantum cloning. *Phys. Rev. A*, 57:2368–2378, April 1998. `doi:10.1103/PhysRevA.57.2368`.

16   Harry Buhrman, Richard Cleve, Monique Laurent, Noah Linden, Alexander Schrijver, and Falk Unger. New limits on fault-tolerant quantum computation. In *2006 47th Annual IEEE Symposium on Foundations of Computer Science (FOCS'06)*, pages 411–419, 2006. `doi:10.1109/FOCS.2006.50`.

17   Earl T. Campbell, Barbara M. Terhal, and Christophe Vuillot. Roads towards fault-tolerant universal quantum computation. *Nature*, 549(7671):172–179, September 2017. `doi:10.1038/nature23460`.

18   Man-Duen Choi. Completely positive linear maps on complex matrices. *Linear Algebra and its Applications*, 10(3):285–290, 1975. `doi:10.1016/0024-3795(75)90075-0`.

19   Toby S Cubitt, Jianxin Chen, and Aram W Harrow. Superactivation of the asymptotic zero-error classical capacity of a quantum channel. *IEEE transactions on information theory*, 57(12):8114–8126, 2011. `doi:10.1109/TIT.2011.2169109`.

20   Eric Dennis, Alexei Kitaev, Andrew Landahl, and John Preskill. Topological quantum memory. *Journal of Mathematical Physics*, 43(9):4452–4505, 2002. `doi:10.1063/1.1499754`.

21   R. L. Dobrushin. Central limit theorem for nonstationary markov chains. i. *Theory of Probability & Its Applications*, 1(1):65–80, 1956. `doi:10.1137/1101006`.

22   Runyao Duan. Super-activation of zero-error capacity of noisy quantum channels. *arXiv preprint arXiv:0906.2527*, 2009.

23   Runyao Duan, Simone Severini, and Andreas Winter. Zero-error communication via quantum channels, noncommutative graphs, and a quantum lovász number. *IEEE Transactions on Information Theory*, 59(2):1164–1174, 2012. `doi:10.1109/TIT.2012.2221677`.

24   W.S. Evans and L.J. Schulman. Signal propagation and noisy circuits. *IEEE Transactions on Information Theory*, 45(7):2367–2373, 1999. `doi:10.1109/18.796377`.

25   W.S. Evans and L.J. Schulman. On the maximum tolerable noise of k-input gates for reliable computation by formulas. *IEEE Transactions on Information Theory*, 49(11):3094–3098, 2003. `doi:10.1109/TIT.2003.818405`.

26   Omar Fawzi, Antoine Grospellier, and Anthony Leverrier. Constant overhead quantum fault-tolerance with quantum expander codes. In *2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 743–754. IEEE, 2018.

27   Daniel Gottesman. Fault-tolerant quantum computation with constant overhead. *Quantum Info. Comput.*, 14(15-16):1338–1372, November 2014.

28   Jérémie Guillaud and Mazyar Mirrahimi. Repetition cat qubits for fault-tolerant quantum computation. *Phys. Rev. X*, 9:041053, December 2019. `doi:10.1103/PhysRevX.9.041053`.

29   Clare Horsman, Austin G Fowler, Simon Devitt, and Rodney Van Meter. Surface code quantum computing by lattice surgery. *New Journal of Physics*, 14(12):123011, December 2012. `doi:10.1088/1367-2630/14/12/123011`.

30   Julia Kempe, Oded Regev, Falk Uunger, and Ronald de Wolf. Upper bounds on the noise threshold for fault-tolerant quantum computing. *Quantum Info. Comput.*, 10(5):361–376, May 2010.

31   A Yu Kitaev. Quantum computations: algorithms and error correction. *Russian Mathematical Surveys*, 52(6):1191–1249, December 1997. `doi:10.1070/rm1997v052n06abeh002155`.

32   A.Yu. Kitaev. Fault-tolerant quantum computation by anyons. *Annals of Physics*, 303(1):2–30, 2003. `doi:10.1016/S0003-4916(02)00018-0`.

33   E. Knill. Quantum computing with realistically noisy devices. *Nature*, 434(7029):39–44, March 2005. `doi:10.1038/nature03350`.

34   Emanuel Knill, Raymond Laflamme, and Wojciech H. Zurek. Resilient quantum computation: error models and thresholds. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 454(1969):365–384, January 1998. `doi:10.1098/rspa.1998.0166`.

35   Debbie Leung and Graeme Smith. Continuity of quantum channel capacities. *Communications in Mathematical Physics*, 292(1):201–215, May 2009. `doi:10.1007/s00220-009-0833-1`.

36   Rex AC Medeiros and Francisco M De Assis. Quantum zero-error capacity. *International Journal of Quantum Information*, 3(01):135–139, 2005. `doi:10.1142/S0219749905000682`.

37   Alexander Müller-Hermes, Daniel Stilck França, and Michael M Wolf. Relative entropy convergence for depolarizing channels. *Journal of Mathematical Physics*, 57(2):022202, 2016. `doi:10.1063/1.4939560`.

38   Jeonghoon Park and Soojoon Lee. Zero-error classical capacity of qubit channels cannot be superactivated. *Physical Review A*, 85(5):052321, 2012. `doi:10.1103/PhysRevA.85.052321`.

39   Stefano Pirandola. End-to-end capacities of a quantum communication network. *Communications Physics*, 2(1):51, May 2019. `doi:10.1038/s42005-019-0147-3`.

40   Maxim Raginsky. Strictly contractive quantum channels and physically realizable quantum computers. *Phys. Rev. A*, 65:032306, February 2002. `doi:10.1103/PhysRevA.65.032306`.

41   Robert Raussendorf and Jim Harrington. Fault-tolerant quantum computation with high threshold in two dimensions. *Phys. Rev. Lett.*, 98:190504, May 2007. `doi:10.1103/PhysRevLett.98.190504`.

42   Alexander A. Razborov. An upper bound on the threshold quantum decoherence rate. *Quantum Info. Comput.*, 4(3):222–228, May 2004.

43   Mary Beth Ruskai. Beyond strong subadditivity? improved bounds on the contraction of generalized relative entropy. *Reviews in Mathematical Physics*, 6(05a):1147–1161, 1994. `doi:10.1142/s0129055x94000407`.

44   Pradeep Kiran Sarvepalli, Andreas Klappenecker, and Martin Rötteler. Asymmetric quantum codes: constructions, bounds and performance. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 465(2105):1645–1672, 2009. `doi:10.1098/rspa.2008.0439`.

45   ME Shirokov and Tatiana Shulman. On superactivation of zero-error capacities and reversibility of a quantum channel. *Communications in Mathematical Physics*, 335(3):1159–1179, 2015. `doi:10.1007/s00220-015-2345-5`.

46   Peter W. Shor. Scheme for reducing decoherence in quantum computer memory. *Phys. Rev. A*, 52:R2493–R2496, October 1995. `doi:10.1103/PhysRevA.52.R2493`.

47   Peter W. Shor. Fault-tolerant quantum computation. In *Proceedings of 37th Conference on Foundations of Computer Science*, pages 56–65, 1996. `doi:10.1109/SFCS.1996.548464`.

48   A. M. Steane. Error correcting codes in quantum theory. *Phys. Rev. Lett.*, 77:793–797, July 1996. `doi:10.1103/PhysRevLett.77.793`.

49   Barbara M. Terhal. Quantum error correction for quantum memories. *Rev. Mod. Phys.*, 87:307–346, April 2015. `doi:10.1103/RevModPhys.87.307`.

50   David K. Tuckett, Stephen D. Bartlett, and Steven T. Flammia. Ultrahigh error threshold for surface codes with biased noise. *Phys. Rev. Lett.*, 120:050505, January 2018. `doi:10.1103/PhysRevLett.120.050505`.

51   S. Virmani, Susana F. Huelga, and Martin B. Plenio. Classical simulability, entanglement breaking, and quantum computation thresholds. *Phys. Rev. A*, 71:042328, April 2005. `doi:10.1103/PhysRevA.71.042328`.

52 Chenyang Wang, Jim Harrington, and John Preskill. Confinement-higgs transition in a disordered gauge theory and the accuracy threshold for quantum memory. *Annals of Physics*, 303(1):31–58, 2003. `doi:10.1016/S0003-4916(02)00019-2`.

53 David S. Wang, Austin G. Fowler, and Lloyd C. L. Hollenberg. Surface code quantum computing with error rates over 1%. *Physical Review A*, 83(2), February 2011. `doi:10.1103/physreva.83.020302`.

54 John Watrous. Semidefinite programs for completely bounded norms. *arXiv preprint arXiv:0901.4709*, 2009.

55 John Watrous. *The Theory of Quantum Information*. Cambridge University Press, 2018. `doi:10.1017/9781316848142`.

56 Xing-Can Yao, Tian-Xiong Wang, Hao-Ze Chen, Wei-Bo Gao, Austin G. Fowler, Robert Raussendorf, Zeng-Bing Chen, Nai-Le Liu, Chao-Yang Lu, You-Jin Deng, and et al. Experimental demonstration of topological error correction. *Nature*, 482(7386):489–494, February 2012. `doi:10.1038/nature10770`.