arXiv:2211.00697v4 [quant-ph] 9 Aug 2023

A Converse for Fault-tolerant Quantum Computation

Uthirakalyani G[†1], Anuj K. Nayak[†2], and Avhishek Chatterjee[1]

[1]Department of Electrical Engineering, Indian Institute of Technology Madras, Chennai, India.
[2]Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, USA.

As techniques for fault-tolerant quantum computation keep improving, it is natural to ask: what is the fundamental lower bound on space overhead? In this paper, we obtain a lower bound on the space overhead required for ϵ-accurate implementation of a large class of operations that includes unitary operators. For the practically relevant case of sub-exponential depth and sub-linear gate size, our bound on space overhead is tighter than the known lower bounds. We obtain this bound by connecting fault-tolerant computation with a set of finite blocklength quantum communication problems whose accuracy requirements satisfy a joint constraint. The lower bound on space overhead obtained here leads to a strictly smaller upper bound on the noise threshold for noise that are not degradable. Our bound directly extends to the case where noise at the outputs of a gate are non-i.i.d. but noise across gates are i.i.d.

1 Introduction

The idea of a quantum computer was proposed in [3]. The computational advantage of a quantum computer over its classical counterpart was first shown mathematically in [5], followed by [6]. However, this initial excitement for quantum computers was confronted with a practical issue, noise in quantum circuits.

To tackle noise in quantum circuits, the exciting area of quantum fault-tolerance emerged, thanks to the work in [18], followed by [19, 1, 14] and many others. The seminal works in [18, 19, 1, 14] essentially showed that if the strength of the noise at the gates is below a threshold, almost accurate quantum computation can be realized. However, this is achieved at the cost of a poly-logarithmic increase in the size of the circuit with respect to the ideal circuit made of noiseless gates. In subsequent works, this poly-logarithmic space overhead has been improved. In [10, 8], it was shown that a constant space overhead can be achieved if the noise strength is below a threshold. This naturally raises the question: what is the minimum space overhead requirement?

Relatively fewer works have addressed the question of minimum space overhead requirement. However, as better and better codes are found, the interest in understanding the fundamental limit on space overhead is growing [17, 12, 9].

In this paper, we obtain a lower bound on the required space overhead for a broad class of noise models. In the practically likely regime of sub-exponential (in the number of input qubits) depth and sub-linear gate size, [10, 8], our bound is strictly better than the

Uthirakalyani G[†]: ee19d404@smail.iitm.ac.in, [†] Equally contributing student authors.
Anuj K. Nayak[†]: anujk4@illinois.edu
Avhishek Chatterjee: avhishek@ee.iitm.ac.in.

existing bounds in [11, 17, 12, 9]. This bound is obtained by connecting the fault-tolerant computation problem to a set of finite blocklength communication problems whose accuracy requirements satisfy a joint constraint.

1.1 Related Work

Influenced by the gradual improvements in the space overhead of the error correcting schemes for quantum circuits [19, 1, 10, 8], improved lower bounds on space overhead were sought.

Harrow and Nielsen obtained a threshold of 0.74 for depolarizing noise [11]. In [17], Razborov obtained an improved gate size dependent threshold $1 - \frac{1}{g}$, where g is the gate size, i.e., the maximum number of inputs to a gate. Kempe et al. [12] improved this bound to $1 - \sqrt{2^{\frac{1}{g}} - 1}$ for mixture of unitary gates of size g.

Buhrman et al. [4] and Virmani et al. [20] showed classical simulability of noisy quantum circuits beyond a threshold under assumptions on special gate operations and noise. These indirectly showed that, under certain assumptions, a quantum computer loses its edge over a classical computer if the noise is more than a threshold.

Recently, Fawzi et al. [9] obtained a threshold for quantum computation in terms of the quantum capacity of a channel with the same noise. Their model includes any arbitrary gate operation and allows free noiseless classical computation. Though their threshold does not depend on the gate size g, it is strictly better than the threshold in [12] for $g \geq 2$. Moreover, they provide a lower bound on the space overhead when the noise is below the threshold.

1.2 Our Contribution

The main result in this paper is a lower bound on the number of physical qubits needed for fault-tolerant quantum computation with a quantum state of d (logical) qubits. The bound is obtained by first establishing a connection between fault-tolerant computing and a set of finite blocklength quantum communication problems [13] whose accuracy requirements share a joint constraint. Then, using results from finite blocklength quantum communication [13] we obtain a converse, which is then optimized to obtain the final lower bound. Our approach is inspired by information theoretic converse for classical noisy circuits in [16, 7], but the techniques used are quite different.

In the likely practical scenario of sub-exponential circuit depth and constant gate size, our bound is tighter than the best known bound from [9]. Specifically, if the input has d qubits and the gate noise is given by the quantum channel $\mathcal{N}$, then [9] shows that the number of physical qubits needed is $\frac{d}{Q(\mathcal{N})}$, where $Q(\mathcal{N})$ is the quantum capacity of the channel $\mathcal{N}$. We show that for sub-linear gate size, i.e., $g = o(d)$, the lower bound is $\frac{d\ g}{I_c(\mathcal{N}^{\otimes g})}$, where $I_c(\cdot)$ denotes the coherent information of the corresponding quantum channel. Since $Q(\mathcal{N}) = \sup_k \frac{I_c(\mathcal{N}^{\otimes k})}{k}$, our bound is tighter.

An implication of our bound is an improved upper bound on noise threshold for a broad class of noise and gate models. For noise that are not degradable, this upper bound is strictly lower than the known ones and depends on gate size g like [17, 12]. Interestingly, in the case where noise at a gate are correlated but noise across gates are i.i.d., bounds on space overhead and noise threshold can be obtained by replacing $I_c(\mathcal{N}^{\otimes g})$ with $I_c(\mathcal{N}^{(g)})$ in our bound. Here, $\mathcal{N}^{(g)}$ is an arbitrary non-i.i.d. noise at the output of a gate of size g.

1.3 Organization

In Sec. 2, the quantum computation model is presented. The main result is presented in Sec. 3 followed by a discussion about its implications. Proof of the main result is presented in Sec. 4, which draws on a few intermediate results, whose proofs are presented in Appendix A and B. We conclude in Sec. 5 with a short discussion on interesting directions for further explorations.

2 Quantum Computation Model

We consider a fault-tolerant quantum circuit whose goal is to compute a function $f(\cdot)$ on a d-qubit input, i.e., a 2^d dimensional input. At the input to the circuit, there are total N physical qubits on which quantum operations are performed sequentially in layers (Fig. 1). The initial input, $\rho^{(d)}$, is placed on d physical qubits and the remaining $N-d$ physical qubits are ancillas, which can be used for error correction. Thus, the initial state of the N physical qubits is given by $\rho^{(N)} = \rho^{(d)} \otimes \left(\otimes_{a=1}^{N-d} \sigma_a\right)$, where σ_a are ancilla qubits.

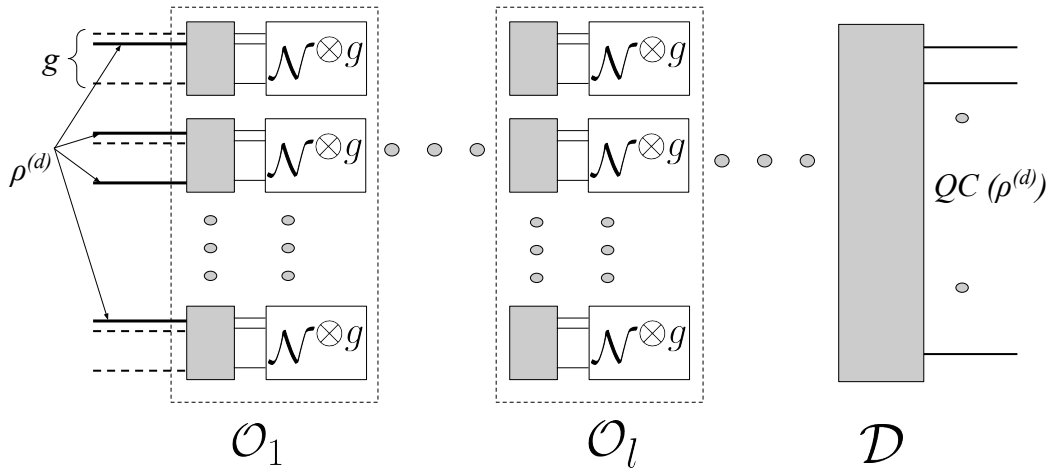

Figure 1: A schematic of the quantum computation model.

A quantum circuit of depth D has D layers and does D operations on $\rho^{(N)}$ in a sequential manner. That is, given the quantum operations $\mathcal{O}_l$ corresponding layers $1 \leq l \leq D$, the output at the end of the final layer D is $\mathcal{O}_D \circ \mathcal{O}_{D-1} \cdots \circ \mathcal{O}_1 \left(\rho^{(N)}\right)$.

The operation $\mathcal{O}_l$ in layer l is realized using noisy quantum gates of size g. If a set of qubits are not processed by any gate at a particular layer, we model them to be processed by a dummy g-input identity gate. Thus, in the first layer, there are at least $\lceil \frac{N}{g} \rceil$ gates, including the dummy identity gates. Like [17, 9], at each layer, introduction of fresh ancilla qubits and noise free classical computations are allowed.

Following prior work [17, 12, 9], a noisy gate is modeled by a perfect gate followed by g i.i.d. noisy quantum channels $\mathcal{N}$, given by $\mathcal{N}^{\otimes g}$. Noise across gates and layers are assumed to be i.i.d.

Finally, an arbitrary noiseless quantum operation $\mathcal{D}$ is allowed after the final layer. It maps $\mathcal{O}_D \circ \mathcal{O}_{D-1} \cdots \circ \mathcal{O}_1 \left(\rho^{(N)}\right)$ to a quantum state of the same dimension as that of

$f(\rho^{(d)})$. This operation is equivalent to the *noiseless* decoder in [9] and the partial trace operation in [17]. We refer to the final output of the noisy circuit after the operation $\mathcal{D}$ as $QC(\rho^{(d)})$.

2.1 Criteria for fault-tolerance

In the above computational model, the criteria for ϵ-accurate (or fault-tolerance) computation of $f(\cdot)$ using $QC(\cdot)$ is

$$\text{C0: } \forall \rho^{(d)} \text{ (d-qubit state)}, F(QC(\rho^{(d)}), f(\rho^{(d)})) > 1 - \epsilon, \tag{1}$$

where, $F(\cdot, \cdot)$ is the standard fidelity [13].

The usual arguments for achieving fault-tolerance against sufficiently small but constant noise using concatenated codes [15, Ch. 10.6.1] imply that criteria C0 can be achieved. In particular, it can be achieved using concatenated codes and poly-logarithmic space overhead for any sufficiently small $\epsilon > 0$, where for any state σ and a completely positive trace preserving map $\mathcal{E}$, $\mathcal{N}(\sigma) = (1 - \epsilon)\sigma + \epsilon\ \mathcal{E}(\sigma)$. Space overhead may be improved using better fault tolerance schemes like [10, 8]. However, the focus of this work is on converse results: a lower bound on required space overhead and an upper bound on noise threshold. Our computational model, fault-tolerance criteria and objective are similar to that in [9, 17].

3 Converse for ϵ-accurate Computing

The following proposition gives a lower bound on the number of physical qubits needed for fault-tolerant computation.

Proposition 1. *There exists a class of functions on d-qubit states such that the number of physical qubits needed for implementing ϵ-accurate circuits for computing these functions is lower bounded by*

$$\frac{d}{\frac{I_c(\mathcal{N}^{\otimes g})}{g} + \frac{1}{2(d-g)\ \ln 2}\left(\ln \frac{4d}{g} + \frac{8}{7}\right)} - 2g,$$

for $\epsilon \in (0, 0.11)$ and $d \geq 2g$. Here, $I_c(\mathcal{N}^{\otimes g})$ is the coherent information of the channel $\mathcal{N}^{\otimes g}$.

Proof. The proof of Prop. 1 is given in Sec. 4. □

First, we would like to mention that the class of functions in Proposition 1 includes all unitary transformations. Hence, this class of functions can simulate evolutions of physical systems and implement important quantum computation modules like quantum Fourier transform.

The bound in Proposition 1 directly extends to circuits where noise at the outputs of a gate are not independent, but noise are i.i.d. across gates. In that case, instead of $I_c(\mathcal{N}^{\otimes g})$, the bound would have $I_c(\mathcal{N}^{(g)})$, where $\mathcal{N}^{(g)}$ represents the potentially correlated noise acting on the g-qubit output of a gate. This would be evident from the proof of Proposition 1.

3.1 Lower bound on space overhead

Next, we consider circuits where D is sub-exponential in d and the gate size $g = o(d)$ since the probable practical implementations lie in this regime. The following corollary, which follows directly from Proposition 1, gives a lower bound on the minimum required space overhead in this regime.

Corollary 1. *For $g = o(d)$ and $\epsilon \in (0, 0.11)$, there exist a class of functions that include unitary transformations, such that the limiting space overhead, $\lim_{d\to\infty} \frac{N}{d}$, required for ϵ-accurate implementation of those functions is lower bounded by $\frac{g}{I_c(\mathcal{N}^{\otimes g})}$.*

When D is sub-exponential and $g = o(d)$, the best known lower bound on limiting space overhead is $\inf_{k\geq 1} \frac{k}{I_c(\mathcal{N}^{\otimes k})}$ [9]. Note that this is equal to the bound in Corollary 1 when $\mathcal{N}$ is degradable [13, §3.25]. However, for depolarizing noise, and in general, for noise that are not degradable, the bound in Corollary 1 is strictly higher.

Another interesting aspect of the bound in Corollary 1 is that it captures the effect of the gate size on minimum space overhead, which is not captured by the existing best bound $\inf_{k\geq 1} \frac{k}{I_c(\mathcal{N}^{\otimes k})}$.

3.2 Upper bound on noise threshold

Noise threshold for a noise model parameterized by a single parameter (e.g., depolarizing) is defined as the strength of noise beyond which quantum computation is not possible. Its dependence on the gate size was captured in [17] and [12]. On the other hand, in [9], a tighter bound involving only quantum capacity was given. Here, we obtain a tighter upper bound on noise threshold in terms of both gate size and a quantity closely related to quantum capacity.

We state results for generic noise that may have multiple parameters and characterize the parameter region where fault-tolerant computation is not possible using reasonable space overhead. The following is a direct corollary of Proposition 1.

Corollary 2. *For $g = o(d)$ and $\epsilon \in (0, 0.11)$, there exist a class of functions that include unitary transformations, such that for any parameter region of the noise $\mathcal{N}$ where $I_c(\mathcal{N}^{\otimes g}) = 0$, ϵ-accurate computation requires $\frac{N}{d} = \Omega(\frac{d}{\ln d})$, i.e, sub-linear space overhead (upto a logarithmic factor) is necessary.*

This corollary of Proposition 1, unlike the threshold results in [9, 17, 12], does not give an impossibility result. Despite that it says something quite useful: constant or poly-logarithmic space overhead, which are the gold standards for fault-tolerant schemes, cannot be achieved when $I_c(\mathcal{N}^{\otimes g}) = 0$. In comparison, the best existing result says that linear or poly-logarithmic space overhead is not possible when $\sup_{k\geq 1} \frac{I_c(\mathcal{N}^{\otimes k})}{k} = 0$ [9].

Thus, for noise that are not degradable, Corollary 2 gives a strictly larger parameter region where no scheme can achieve poly-logarithmic space overhead. An immediate implication of Corollary 2 is an upper bound on noise threshold for depolarizing noise which is strictly lower than the best existing upper bound.

Next, we derive a much stronger phase transition of the required space overhead for the same parameter region as given by Corollary 2. However, this result does not directly follow from Proposition 1 and is derived using intermediate results from the proof of Proposition 1.

Proposition 2. *For $g = o(d)$ and $\epsilon \in (0, 0.11)$, there exist a class of functions that include unitary transformations, such that for any parameter region of the noise $\mathcal{N}$ where $I_c(\mathcal{N}^{\otimes g}) = 0$, ϵ-accurate computation for $\epsilon < 0.11$ is not possible for $N = o(\exp(d))$.*

Proof of this proposition is presented in Appendix C. This result also extends to noise that are correlated at a gate but are independent across gates.

For all practical purposes, Proposition 2 is an impossibility result since it proves that in the parameter region where $I_c(\mathcal{N}^{\otimes g}) = 0$, accurate computation requires at least exponential space overhead. An even stronger impossibility result is discussed in Appendix D.

In summary, Proposition 2 gives an upper bound on the noise threshold and Proposition 1 gives a lower on bound the minimum space overhead when the noise strength is below that threshold. These bounds are applicable to a broad class of noise and gate models, and for circuits with sub-exponential depth, these are tighter than the respective existing bounds.

3.3 Discussions

The gate size dependent bounds on noise threshold and minimum space overhead in Proposition. 2 and 1, respectively, can be useful in choosing the right experimental implementations. Consider the scenario where there are multiple options for experimental implementations using different approaches. These possible implementations have different gate sizes and encounter different types of noise. In this scenario, the bound in Proposition. 1 can be a thumb rule for choosing the best experimental implementation in terms of space overhead requirement. As the bounds in Propositions 1 and 2 extend to any qudit circuit, this can also be used as a thumb rule for comparison across different qudit and qubit technologies.

When the noise at the outputs of a gate are correlated, the minimum required space overhead and the noise threshold are decided by $I_c(\mathcal{N}^{(g)})$, where $\mathcal{N}^{(g)}$ is the generic non-i.i.d. noise at the outputs of a gate. In this context, it is important to understand the kind of correlations that hurt the computations most and plan to avoid such correlations in the experimental realizations.

Unlike [17, 12], the bound in Proposition. 1 and that in [9] require the knowledge of coherent information of a k-fold channel, $I_c(\mathcal{N}^{\otimes k})$. This often does not have a closed form and obtaining $I_c(\mathcal{N}^{\otimes k})$ requires to solve a 2^k-dimensional non-convex optimization problem. With the increasing interest in understanding non-convex optimization in the machine learning community, a search for provably efficient algorithm for computing $I_c(\mathcal{N}^{\otimes k})$ can be of independent interest.

4 Proof of Proposition 1

For proving the lower bound in Proposition 1, we first state a converse for computing a class of functions f on d-qubit states of dimension.

Theorem 1. *Suppose the function f to be computed using the noisy quantum circuit satisfies the following conditions.*

(i) f has an inverse f^{-1} that can be accurately computed if we have access to a noiseless quantum circuit, (ii) f^{-1} exists and for any two d-qubit states $\eta_1^{(d)}$ and $\eta_2^{(d)}$, $1 - F(f^{-1}(\eta_1^{(d)}), f^{-1}(\eta_2^{(d)})) \leq L(1 - F(\eta_1^{(d)}, \eta_2^{(d)}))$ for some $L > 0$ and (iii) $L \leq \frac{1-e^{-\frac{1}{8}}}{\epsilon}$.

Then, for ϵ-accurate computation of f, for $d \geq 2g$, the required number of physical qubits is lower bounded by

$$\frac{d - \frac{1}{1-4\ln\frac{1}{1-\epsilon L}} I_c(\mathcal{N}^{\otimes g})}{\frac{I_c(\mathcal{N}^{\otimes g})}{g} + \frac{G}{g(G-1)} \frac{h_2(\frac{2g}{d}\ln\frac{1}{1-\epsilon L})}{1-4\ln\frac{1}{1-\epsilon L}}},$$

where $G = \lceil \frac{N}{g} \rceil$, and $h_2(\cdot)$ is the binary entropy function. Proof of this theorem is presented in Appendix A. Proof of Proposition 1 follows from this theorem.

Proof of Proposition 1. For proving Proposition 1, we first consider the second term in the bound in Theorem 1:

$$\frac{-\frac{1}{1-4\ln\frac{1}{1-\epsilon L}} I_c(\mathcal{N}^{\otimes g})}{\frac{I_c(\mathcal{N}^{\otimes g})}{g} + \frac{G}{g(G-1)} \frac{h_2(\frac{2g}{d}\ln\frac{1}{1-\epsilon L})}{1-4\ln\frac{1}{1-\epsilon L}}}.$$

Note that as $\frac{G}{g(G-1)} \frac{h_2(\frac{2g}{d}\ln\frac{1}{1-\epsilon L})}{1-\frac{4g}{d}\ln\frac{1}{1-\epsilon L}} > 0$ if $L \leq \frac{1-e^{-\frac{1}{8}}}{\epsilon}$,

$$-\frac{1}{\frac{I_c(\mathcal{N}^{\otimes g})}{g} + \frac{G}{g(G-1)} \frac{h_2(\frac{2g}{d}\ln\frac{1}{1-\epsilon L})}{1-4\ln\frac{1}{1-\epsilon L}}} \geq -\frac{g}{I_c(\mathcal{N}^{\otimes g})}. \tag{2}$$

On the other hand, $\frac{1}{1-4\ln\frac{1}{1-\epsilon L}} \leq 2$ if $L \leq \frac{1-e^{-\frac{1}{8}}}{\epsilon}$

Hence, the second term in the bound in Theorem 1 is $\geq -2\ g$ if $L \leq \frac{1-e^{-\frac{1}{8}}}{\epsilon}$.

Next, we consider the first term in the bound in Theorem 1:

$$\frac{d}{\frac{I_c(\mathcal{N}^{\otimes g})}{g} + \frac{G}{g(G-1)} \frac{h_2(\frac{2g}{d}\ln\frac{1}{1-\epsilon L})}{1-4\ln\frac{1}{1-\epsilon L}}},$$

and obtain a lower bound for this term. Note that an upper bound on $h_2(\frac{2g}{d}\ln\frac{1}{1-\epsilon L})$ will give a lower bound on this term.

Next, we use the fact that

$$\begin{aligned}
h_2(x)\ln 2 &= x\ln\frac{1}{x} + (1-x)\ln\frac{1}{1-x} \\
&\leq x\ln\frac{1}{x} + \ln\frac{1}{1-x} \\
&= x\ln\frac{1}{x} - \ln(1-x) \\
&= x\ln\frac{1}{x} - (-x - \frac{x^2}{2} - \frac{x^3}{3} - \cdots) \\
&\leq x\ln\frac{1}{x} + x(1 + x + x^2 + \cdots) \\
&= x\left(\ln\frac{1}{x} + \frac{1}{1-x}\right),
\end{aligned}$$

and under condition (iii) in Theorem 1, $h_2(\frac{2g}{d}\ln\frac{1}{1-\epsilon L}) \leq h_2(\frac{g}{4d})$.

Thus, we obtain

$$h_2(\frac{2g}{d}\ln\frac{1}{1-\epsilon L}) \leq \frac{g}{4d\ln 2}\left(\ln\frac{4d}{g} + \frac{1}{1-\frac{g}{4d}}\right) \leq \frac{g}{4d\ln 2}\left(\ln\frac{4d}{g} + \frac{1}{1-\frac{1}{8}}\right) \tag{3}$$

as $d \geq 2g$. Under condition (iii) in Theorem 1, $\frac{1}{1-4\ln\frac{1}{1-\epsilon L}} \leq 2$ and hence,

$$
\begin{aligned}
&\frac{d}{\frac{I_c(\mathcal{N}^{\otimes g})}{g} + \frac{G}{g(G-1)} \frac{h_2(\frac{2g}{d} \ln \frac{1}{1-\epsilon L})}{1-4 \ln \frac{1}{1-\epsilon L}}} \\
&\geq \frac{d}{\frac{I_c(\mathcal{N}^{\otimes g})}{g} + \frac{2G}{g(G-1)} h_2(\frac{2g}{d} \ln \frac{1}{1-\epsilon L})} \\
&\geq \frac{d}{\frac{I_c(\mathcal{N}^{\otimes g})}{g} + \frac{2G}{g(G-1)} \frac{g}{4d \ln 2} \left(\ln \frac{4d}{g} + \frac{8}{7} \right)} \\
&\geq \frac{d}{\frac{I_c(\mathcal{N}^{\otimes g})}{g} + \frac{1}{2(d-g) \ln 2} \left(\ln \frac{4d}{g} + \frac{8}{7} \right)}.
\end{aligned}
$$

The last inequality follows from the fact that $\frac{G}{G-1}$ is monotonically decreasing in G and $G \geq \frac{d}{g}$.

If f is unitary, then f^{-1} satisfies all the three conditions in Theorem 1 for $L = 1$. Hence, the condition $0 < L \leq \frac{1-e^{-\frac{1}{8}}}{\epsilon}$, is satisfied by a large class of f, including unitary transformations, if $0 < \epsilon < 1 - e^{-\frac{1}{8}} \leq 0.11$. Thus, the derived bound is applicable to all practically important f, including the well known quantum Fourier transform, which is the fundamental building block of many interesting algorithms.

This completes the proof of Proposition 1. □

5 Conclusion and Future Work

In this paper, inspired by the information theoretic bounds for noisy classical circuits [16, 7], a connection between fault-tolerant quantum computation and finite blocklength communication is cultivated. This leads to a lower bound on the required space overhead for fault-tolerant computation and is given by gate size g divided by the coherent information of a g-fold noisy channel. This bound is tighter than the existing bounds and can be extended to the case where the noise on the outputs of a gate are correlated. It gives a tighter upper bound on the noise threshold.

In future, we would like to combine our approach with techniques developed in [17, 12, 9] for tightening the bound and design fault-tolerant schemes for achieving that bound.

As discussed in Sec. 3, obtaining $I_c(\mathcal{N}^{\otimes k})$ involves a 2^k-dimensional non-convex optimization problem. Since understanding non-convex optimization is of mathematical interest due to its usefulness in modern machine learning, an exploration for efficient computation of $I_c(\mathcal{N}^{\otimes k})$ can be of independent mathematical interest.

Acknowledgment

The authors are thankful to the two anonymous reviewers and the editor, Anurag Anshu, for their many helpful comments. In particular, they thank Reviewer 1 for pointing out an error in the lower order term of the main bound.

AC thanks Lav R. Varshney for nudging him multiple times to work on quantum circuits. AC also thanks Lav R. Varshney and Arul Lakshminarayan for useful feedback on the paper. AC gratefully acknowledges supports through the grants NFIG IIT Madras, DST/INSPIRE/04/2016/001171 and SERB/SRG/2019/001809.

AN expresses gratitude for the support provided in part by National Science Foundation grant PHY-2112890.

A Proof of Theorem 1

For proving 1, we first state an important lemma.

Lemma 1. *Suppose the conditions (i)–(iii) in Theorem 1 hold. Then, for ϵ-accurate computation we need $d \leq \sum_{i=1}^{G} \frac{1}{1-2\ \epsilon_i} \left(I_c(\mathcal{N}^{\otimes g}) + h_2(\epsilon_i)\right)$, where, $\{\epsilon_i\}$ are tunable auxiliary variables taking values in $[0,1]$ and satisfying the condition: $\prod_{i=1}^{G}(1-\frac{\epsilon_i}{2}) \geq 1-\epsilon\ L$.*

Proof of Theorem 1. Proof of this lemma is presented in Appendix B. Here, we prove Theorem 1 using this lemma.

First, by using the fact that $1-x \leq \exp(-x)$, we obtain the following upper bound.

$$\begin{aligned} &\max_{\{\epsilon_i\}} \sum_{i=1}^{G} \frac{1}{1-2\ \epsilon_i} \left(I_c(\mathcal{N}^{\otimes g}) + h_2(\epsilon_i)\right) \\ &\text{s.t.} \ \sum_{i=1}^{G} \epsilon_i \leq 2 \ln \frac{1}{1-\epsilon L}. \end{aligned} \tag{4}$$

Clearly the above upper bound is bounded by the sum of the maximum of the following optimization problems: P1 and P2.

$$\text{P1:} \ \max_{\{\epsilon_i\}} \sum_{i=1}^{G} \frac{1}{1-2\ \epsilon_i} I_c(\mathcal{N}^{\otimes g}) \ \text{s.t.} \ \sum_{i=1}^{G} \epsilon_i \leq 2 \ln \frac{1}{1-\epsilon L}. \tag{5}$$

$$\text{P2:} \ \max_{\{\epsilon_i\}} \sum_{i=1}^{G} \frac{h_2(\epsilon_i)}{1-2\ \epsilon_i} \ \text{s.t.} \ \sum_{i=1}^{G} \epsilon_i \leq 2 \ln \frac{1}{1-\epsilon L}. \tag{6}$$

As P1 is a maximization of sum of convex functions subject to a linear constraint, P1 is maximized at an extreme point, as it is a convex maximization. Under condition (iii) in Theorem 1, $2\ln\frac{1}{1-\epsilon L} < \frac{1}{2}$. Thus, the optimum in P1 is obtained when $\epsilon_1 = 2\ln\frac{1}{1-\epsilon L}$ and $\epsilon_i = 0$ for $i \geq 2$.

Note that the optimum of P2 is upper-bounded by the optimum of the following problem.

$$\text{P3:} \ \max_{\{\epsilon_i\}} \frac{1}{1-4\ln\frac{1}{1-\epsilon L}} \sum_{i=1}^{G} h_2(\epsilon_i) \ \text{s.t.} \ \sum_{i=1}^{G} \epsilon_i \leq 2 \ln \frac{1}{1-\epsilon L}. \tag{7}$$

As P3 is a maximization of sum of concave functions subject to a linear constraint, by the symmetry of the problem, the optimum solution is $\epsilon_i^* = \frac{2}{G}\ln\frac{1}{1-\epsilon L}$.

Summing the optimum of P1 and P3, the resulting upper bound on d becomes

$$\begin{aligned} d &\leq (G-1) I_c(\mathcal{N}^{\otimes g}) + \frac{1}{1-4\ln\frac{1}{1-\epsilon L}} I_c(\mathcal{N}^{\otimes g}) + \frac{G}{1-4\ln\frac{1}{1-\epsilon L}} h_2\left(\frac{2}{G}\ln\frac{1}{1-\epsilon L}\right) \\ &= (G-1)\left(I_c(\mathcal{N}^{\otimes g}) + \frac{G}{G-1}\frac{h_2(\frac{2}{G}\ln\frac{1}{1-\epsilon L})}{1-4\ln\frac{1}{1-\epsilon L}}\right) + \frac{1}{1-4\ln\frac{1}{1-\epsilon L}} I_c(\mathcal{N}^{\otimes g}). \end{aligned} \tag{8}$$

Note that as every physical qubit goes through a gate of size g (including dummy identity gates as discussed in Sec. 2), $N \geq (G-1)g$ and hence,

$$d \leq N\left(\frac{I_c(\mathcal{N}^{\otimes g})}{g} + \frac{G}{g(G-1)}\frac{h_2(\frac{2}{G}\ln\frac{1}{1-\epsilon L})}{1-4\ln\frac{1}{1-\epsilon L}}\right) + \frac{1}{1-4\ln\frac{1}{1-\epsilon L}} I_c(\mathcal{N}^{\otimes g})$$

As all inputs must pass through a gate, $G \geq \frac{d}{g}$. So, under condition (iii), $\frac{2}{G}\ln\frac{1}{1-\epsilon L} \leq \frac{2g}{d}\ln\frac{1}{1-\epsilon L} \leq 0.5$ for any $d \geq \frac{g}{2}$. As $h_2(x)$ is monotonic over $[0, 0.5]$,

$$d \leq N\left(\frac{I_c(\mathcal{N}^{\otimes g})}{g} + \frac{G}{g(G-1)}\frac{h_2(\frac{2g}{d}\ln\frac{1}{1-\epsilon L})}{1-4\ln\frac{1}{1-\epsilon L}}\right) + \frac{1}{1-4\ln\frac{1}{1-\epsilon L}} I_c(\mathcal{N}^{\otimes g}). \tag{9}$$

This implies

$$N \geq \frac{d - \frac{1}{1-4\ln\frac{1}{1-\epsilon L}} I_c(\mathcal{N}^{\otimes g})}{\frac{I_c(\mathcal{N}^{\otimes g})}{g} + \frac{G}{g(G-1)}\frac{h_2(\frac{2g}{d}\ln\frac{1}{1-\epsilon L})}{1-4\ln\frac{1}{1-\epsilon L}}}. \tag{10}$$

This completes the proof of Theorem 1. □

B Proof of Lemma 1

For f^{-1} satisfying conditions (i)-(iii) in Theorem 1, satisfaction of the ϵ-accuracy condition C0 (Eq. 1) implies satisfaction of the following condition.

$$\text{C1: } \forall \rho^{(d)} \text{ (d-qubit state)}, F(f^{-1} \circ QC(\rho^{(d)}), \rho^{(d)}) > 1 - \epsilon\ L. \tag{11}$$

Let us denote $\mathcal{D} \circ \mathcal{O}_D \circ \mathcal{O}_{D-1} \cdots \circ \mathcal{O}_2$ by $\mathcal{O}_{2:D}$, then $QC(\rho^{(d)}) = \mathcal{O}_{2:D} \circ \mathcal{N}^{\otimes N} \circ \bar{\mathcal{O}}_1(\rho^{(N)})$ and hence, $f^{-1} \circ QC(\rho^{(d)})$ can be written as $\mathcal{H} \circ \mathcal{N}^{\otimes N} \circ \bar{\mathcal{O}}_1(\rho^{(N)})$ for some $\mathcal{H}$. Here $\bar{\mathcal{O}}_1$ is the noiseless operation at layer 1, i.e., $\mathcal{O}_1 = \mathcal{N}^{\otimes N} \circ \bar{\mathcal{O}}_1$. Hence, satisfaction of condition C1 implies satisfaction of the following condition.

$$\exists \mathcal{H} \text{ s.t. } \forall \rho^{(d)}, F(\mathcal{H} \circ \mathcal{N}^{\otimes N} \circ \bar{\mathcal{O}}_1(\rho^{(N)}), \rho^{(d)}) > 1 - \epsilon\ L. \tag{12}$$

The input to $\bar{\mathcal{O}}_1$ is $\rho^{(d)} \otimes \left(\otimes_{a=1}^{N-d}\sigma_a\right)$, which is a state of N physical qubits, where d of them are input qubits (from $\rho^{(d)}$) and $N-d$ of them are ancilla qubits used for error correction.

Let in an ϵ-accurate implementation of QC, $\mathcal{O}_1$ is implemented using G gates of size g in parallel. In that implementation, let the ith gate on $\mathcal{O}_1$, for $1 \leq i \leq G$, take d_i qubits from $\rho^{(d)}$ and rest from ancilla qubits as inputs.

Without loss of generality, let us assume that in $\mathcal{O}_1$, the first d_1 qubits are input to gate 1, next d_2 qubits are input to gate 2 and so on. Let the noiseless computation by gate i in layer 1 be given by $\mathcal{G}_i(\cdot)$. Then, for an input of the form $\rho^{(d)} = \otimes_{i=1}^{G}\rho_{d_i}$, where ρ_{d_i} are d_i-qubit states, the output of $\mathcal{O}_1$ is

$\mathcal{N}^{\otimes N} \circ \otimes_{i=1}^{G}\mathcal{G}_i\left(\rho_{d_i} \otimes \left(\otimes_{a=1}^{g-d_i}\sigma_a^{(i)}\right)\right)$, where, $\sigma_a^{(i)}$ are the ancilla inputs to gate i.

So, for an ϵ-accurate implementation, the following special case of Eq. 12 must be satisfied: $\exists \mathcal{H}$ s.t. for all $\rho^{(d)} = \otimes_{i=1}^{G}\rho_{d_i}$,

$$F\left(\mathcal{H} \circ \mathcal{N}^{\otimes N} \circ \otimes_{i=1}^{G}\mathcal{G}_i\left(\rho_{d_i} \otimes \left(\otimes_{a=1}^{g-d_i}\sigma_a^{(i)}\right)\right), \otimes_{i=1}^{G}\rho_{d_i}\right) > 1 - \epsilon\ L, \tag{13}$$

Since $\mathcal{N}^{\otimes N} \circ \otimes_{i=1}^{G} \mathcal{G}_i(\rho_{d_i} \otimes \left(\otimes_{a=1}^{g-d_i} \sigma_a^{(i)}\right) = \otimes_{i=1}^{G} \mathcal{N}^{\otimes g} \circ \mathcal{G}_i(\rho_{d_i} \otimes \left(\otimes_{a=1}^{g-d_i} \sigma_a^{(i)}\right)$, and $\rho^{(d)} = \otimes_{i=1}^{G} \rho_{d_i}$, if there exists an $\mathcal{H}$ satisfying the condition in Eq. 13, then there also exists an $\mathcal{H} = \otimes_{i=1}^{G} \mathcal{H}_i$ satisfying the same condition. This is because $\mathcal{G}_i \left(\rho_{d_i} \otimes \left(\otimes_{a=1}^{g-d_i} \sigma_a^{(i)}\right)\right)$ is independent of $\mathcal{G}_j(\cdot)$ and ρ_{d_j} for $j \neq i$ and noise are independent across gates.

Thus, a necessary condition for satisfaction of C0 is: $\exists\{\mathcal{H}_i\}$ s.t. for all $\rho^{(d)} = \otimes_{i=1}^{G} \rho_{d_i}$,

$$F\left(\otimes_{i=1}^{G} \mathcal{H}_i \circ \mathcal{N}^{\otimes g} \circ \mathcal{G}_i(\rho_{d_i} \otimes \left(\otimes_{a=1}^{g-d_i} \sigma_a^{(i)}\right), \otimes_{i=1}^{G} \rho_{d_i}\right) > 1 - \epsilon\ L. \tag{14}$$

By the fact that $F(\sigma_1 \otimes \sigma_2, \sigma_1' \otimes \sigma_2') = F(\sigma_1, \sigma_1') \cdot F(\sigma_2, \sigma_2')$, condition (14) is equivalent to: for some $\{\epsilon_i \in (0,1) : i\}$, for each i,

$$\text{C2: } \exists \mathcal{H}_i \text{ s.t. } \forall \rho_{d_i}, F(\mathcal{H}_i \circ \mathcal{N}^{\otimes g} \circ \mathcal{G}_i(\rho_{d_i} \otimes \left(\otimes_{a=1}^{g-d_i} \sigma_a^{(i)}\right)), \rho_{d_i}) > 1 - \epsilon_i, \tag{15}$$

where, $\prod_{i=1}^{G}(1-\epsilon_i) \geq 1 - \epsilon\ L$.

For satisfaction of condition C2, a necessary condition is the following: for some $\{\epsilon_i \in (0,1) : i\}$, for each i,

$$\text{C3: } \exists\{\mathcal{H}_i, \bar{\mathcal{G}}_i\} \text{ s.t. } \forall \rho_{d_i}, F(\mathcal{H}_i \circ \mathcal{N}^{\otimes g} \circ \bar{\mathcal{G}}_i(\rho_{d_i} \otimes \left(\otimes_{a=1}^{g-d_i} \sigma_a^{(i)}\right)), \rho_{d_i}) > 1 - \epsilon_i, \tag{16}$$

where, $\prod_{i=1}^{G}(1-\epsilon_i) \geq 1 - \epsilon\ L$. Note that under condition (iii) in Theorem 1, $1 - \epsilon L > \frac{1}{2}$ and hence, a necessary condition is $\epsilon_i < \frac{1}{2}$ for $1 \leq i \leq G$.

Note that condition C3 is equivalent to the ϵ_i-accuracy criteria for one-shot communication of quantum information over the quantum channel $\mathcal{N}^{\otimes g}$ [13]. Hence, by [13, Sec. 9.1.2], a necessary condition for satisfaction of C3 is: for $1 \leq i \leq G$

$$d_i \leq \frac{1}{1 - 2\ \epsilon_i} \left(I_c(\mathcal{N}^{\otimes g}) + h_2(\epsilon_i)\right), \tag{17}$$

where h_2 is the binary entropy function.

From this, the following upper bound on d for ϵ-accurate computation is obtained: for some $\{\epsilon_i\}$,

$d = \sum_{i=1}^{G} d_i \leq \sum_{i=1}^{G} \frac{1}{1-2\ \epsilon_i} \left(I_c(\mathcal{N}^{\otimes g}) + h_2(\epsilon_i)\right)$, where, $\prod_{i=1}^{G}(1-\epsilon_i) \geq 1 - \epsilon\ L$.

By allowing noiseless forward classical communication, for the same rate of quantum communication, fidelity of $1 - \eta$ can at most improve to $1 - \frac{\eta}{2}$ [2, Sec. VIII]. Using this fact, if follows that by allowing noiseless classical computation and classical buffers, the bound can at most be the following.

$d = \sum_{i=1}^{G} d_i \leq \sum_{i=1}^{G} \frac{1}{1-2\ \epsilon_i} \left(I_c(\mathcal{N}^{\otimes g}) + h_2(\epsilon_i)\right)$, where, $\prod_{i=1}^{G}(1 - \frac{\epsilon_i}{2}) \geq 1 - \epsilon\ L$. This completes the proof of Lemma 1.

The same proof goes through after replacing $\mathcal{N}^{\otimes g}$ by a general 2^g-dimensional channel $\mathcal{N}^{(g)}$. Hence, the bound applies if the noise at the outputs of a gate are not independent, but noise are independent across gates.

C Proof of Proposition 2

Consider the bound (8) obtained in the proof of Theorem 1. When $I_c(\mathcal{N}^{\otimes g}) = 0$ and $\epsilon < 0.11$ this bounds become

$$d \leq 2G \cdot h_2(\frac{2}{G} \ln \frac{1}{1 - \epsilon L}).$$

Then, using the facts that $h_2(x)$ is monotonic over $[0, 0.5]$ and $h_2(x)\ln 2 \leq x\left(\ln\frac{1}{x} + \frac{1}{1-x}\right)$, we obtain the following bound.

$$\begin{aligned} d &\leq 2G \cdot \frac{2}{G\ln 2} \ln\frac{1}{1-\epsilon L}\left(\ln\frac{G}{2\ln\frac{1}{1-\epsilon L}} + \frac{G}{G - 2\ln\frac{1}{1-\epsilon L}}\right) \\ &\leq \frac{4}{\ln 2}\ln\frac{1}{1-\epsilon L}(\ln G + \frac{4}{3}) - \frac{4}{\ln 2}\ln\frac{1}{1-\epsilon L}\ln\left(2\ln\frac{1}{1-\epsilon L}\right) \\ &\leq \frac{1}{2\ln 2}(\ln G + \frac{4}{3}) + \frac{1}{2\ln 2}\ln 4. \end{aligned}$$

The first inequality follows because $2\ln\frac{1}{1-\epsilon L} \leq \frac{1}{4}$ under conditions (i)-(iii) in Theorem 1, $G \geq 1$ and $\frac{G}{G-\frac{1}{4}}$ is decreasing with G.

For the last inequality, the second term is maximized when $L = \frac{1-e^{-1/8}}{\epsilon}$, which implies that $-4\ln\frac{1}{1-\epsilon L}\ln\left(2\ln\frac{1}{1-\epsilon L}\right) \leq \frac{1}{2}\ln 4$.

As $G = \lceil\frac{N}{g}\rceil \leq \frac{N}{g} + 1$, the bound implies

$$N \geq \frac{g}{4}\ \exp(2\ \ln 2\ d - \frac{4}{3}) - 1. \tag{18}$$

This completes the proof of Proposition 2.

D Towards a Stronger Impossibility Result

Using the same proof as that of Lemma 1 and the results in [13, Sec. 9.1.2], one can obtain another bound on d, in terms of α-sandwiched Renyi coherent information $I_c(\mathcal{N}^{\otimes g};\alpha)$ of the channel. Here, α is a parameter more than 1.

$d = \sum_{i=1}^{G} d_i \leq \sum_{i=1}^{G}\left(I_c(\mathcal{N}^{\otimes g};\alpha) + \frac{\alpha}{\alpha-1}\log_2\frac{1}{1-\epsilon_i}\right)$, where, $\prod_{i=1}^{G}(1-\frac{\epsilon_i}{2}) \geq 1 - \epsilon\ L$.

Therefore, the bound becomes

$$\begin{aligned} &GI_c(\mathcal{N}^{\otimes g};\alpha) + \frac{\alpha}{\alpha-1}\max_{\{\epsilon_i\}}\sum_{i=1}^{G}\log_2\frac{1}{1-\epsilon_i} \\ &\qquad \text{s.t. } \prod_{i=1}^{G}(1-\frac{\epsilon_i}{2}) \geq 1 - \epsilon\ L. \end{aligned} \tag{19}$$

Since $1 - x \leq \exp(-x)$ and loosening the constraint increases the maximum, the maximum of the above optimization is upper-bounded by

$$\begin{aligned} &GI_c(\mathcal{N}^{\otimes g};\alpha) + \frac{\alpha}{\ln 2(\alpha-1)}\max_{\{\epsilon_i\}}\sum_{i=1}^{G}\ln\frac{1}{1-\epsilon_i} \\ &\qquad \text{s.t. } \sum_{i=1}^{G}\epsilon_i \leq 2\ln\frac{1}{1-\epsilon L}. \end{aligned} \tag{20}$$

Note that $-\ln(1-x) = x + \frac{x^2}{2} + \frac{x^3}{3} + \cdots \leq x + x^2 + x^3 + \cdots = \frac{x}{1-x}$, and hence, for getting an upper-bound, it is enough to maximize $\sum_{i=1}^{G}\frac{\epsilon_i}{1-\epsilon_i}$.

As $\frac{x}{1-x}$ is convex on $[0, 1)$, the maximization is obtained at an extreme point, i.e., $\epsilon_1 = 2\ln\frac{1}{1-\epsilon L}$ and $\epsilon_i = 0$ for $i \geq 2$. Thus, the bound becomes $GI_c(\mathcal{N}^{\otimes g};\alpha) + \frac{\alpha}{\ln 2(\alpha-1)}\frac{2\ln\frac{1}{1-\epsilon L}}{1-2\ln\frac{1}{1-\epsilon L}}$.

Under the conditions in Theorem 1, $\frac{2\ln\frac{1}{1-\epsilon L}}{1-2\ln\frac{1}{1-\epsilon L}} \leq \frac{1}{3}$ and hence, the upper bound on d is $\frac{\alpha}{3\ln 2(\alpha-1)}$, whenever $I_c(\mathcal{N}^{\otimes g};\alpha) = 0$.

Thus, for noise strengths that result into $I_c(\mathcal{N}^{\otimes g};\alpha) = 0$, fault-tolerant computation is impossible for $d \geq \frac{\alpha}{3\ln 2(\alpha-1)}$, i.e., there is an upper bound on the number of (logical) qubits that can be computed with ϵ-accuracy.

It is known that $I_c(\mathcal{N}^{\otimes g};\alpha) \downarrow I_c(\mathcal{N}^{\otimes g})$ as $\alpha \downarrow 1$ [13] and hence, it can be possible to take the impossibility threshold arbitrarily close to $I_c(\mathcal{N}^{\otimes g}) = 0$.

The fact $I_c(\mathcal{N}^{\otimes g};\alpha) \downarrow I_c(\mathcal{N}^{\otimes g})$ as $\alpha \downarrow 1$ implies for $\alpha \leq \alpha_{th}$, for some $\alpha_{th} > 1$, $\frac{I_c(\mathcal{N}^{\otimes g})}{g} \leq \frac{I_c(\mathcal{N}^{\otimes g};\alpha)}{g} \leq \sup_{k\geq 1}\frac{I_c(\mathcal{N}^{\otimes k};\alpha)}{k}$.

This implies that no computation with number of qubits $d > \frac{\alpha_{th}}{3\ln 2(\alpha_{th}-1)}$ is possible for certain noise strengths, which include the noise strengths for which the quantum capacity is zero, i.e., $\sup_{k\geq 1}\frac{I_c(\mathcal{N}^{\otimes k};\alpha)}{k} = 0$.

References

[1] Dorit Aharonov and Michael Ben-Or. "Fault-tolerant quantum computation with constant error". In *Proceedings of the Twenty-Ninth Annual ACM Symposium on Theory of Computing*, pages 176–188, 1997.

[2] Howard Barnum, Emanuel Knill, and Michael A. Nielsen. "On quantum fidelities and channel capacities". *IEEE Transactions on Information Theory*, 46(4):1317–1329, 2000.

[3] Paul Benioff. "The computer as a physical system: A microscopic quantum mechanical Hamiltonian model of computers as represented by Turing machines". *Journal of Statistical Physics*, 22(5):563–591, 1980.

[4] Harry Buhrman, Richard Cleve, Monique Laurent, Noah Linden, Alexander Schrijver, and Falk Unger. "New limits on fault-tolerant quantum computation". In *Proceedings of the 2006 47th Annual IEEE Symposium on Foundations of Computer Science (FOCS'06)*, pages 411–419. IEEE, 2006.

[5] David Deutsch. "Quantum theory, the Church–Turing principle and the universal quantum computer". *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 400(1818):97–117, 1985.

[6] David Deutsch and Richard Jozsa. "Rapid solution of problems by quantum computation". *Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences*, 439(1907):553–558, 1992.

[7] William S Evans and Leonard J Schulman. "Signal propagation and noisy circuits". *IEEE Transactions on Information Theory*, 45(7):2367–2373, 1999.

[8] Omar Fawzi, Antoine Grospellier, and Anthony Leverrier. "Constant overhead quantum fault tolerance with quantum expander codes". *2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)*, 2018.

[9] Omar Fawzi, Alexander Müller-Hermes, and Ala Shayeghi. "A lower bound on the space overhead of fault-tolerant quantum computation". In *Proceedings of the 13th Innovations in Theoretical Computer Science Conference (ITCS 2022)*, 2022.

[10] Daniel Gottesman. "Fault-tolerant quantum computation with constant overhead". *Quantum Information & Computation*, 14(15-16):1338–1372, 2014.

[11] Aram W Harrow and Michael A Nielsen. "Robustness of quantum gates in the presence of noise". *Physical Review A*, 68(1):012308, 2003.

[12] Julia Kempe, Oded Regev, Falk Unger, and Ronald de Wolf. "Upper bounds on the noise threshold for fault-tolerant quantum computing". In *International Colloquium on Automata, Languages, and Programming*, pages 845–856. Springer, 2008.

[13] Sumeet Khatri and Mark M Wilde. "Principles of quantum communication theory: A modern approach". *arXiv preprint arXiv:2011.04672*, 2020.

[14] A Yu Kitaev. "Quantum computations: algorithms and error correction". *Russian Mathematical Surveys*, 52(6):1191, 1997.

[15] Michael A. Nielsen and Isaac L. Chuang. *"Quantum Computation and Quantum Information: 10th Anniversary Edition"*. Cambridge University Press, 2010.

[16] Nicholas Pippenger. "Reliable computation by formulas in the presence of noise". *IEEE Transactions on Information Theory*, 34(2):194–197, 1988.

[17] Alexander A Razborov. "An upper bound on the threshold quantum decoherence rate". *Quantum Information & Computation*, 4(3):222–228, 2004.

[18] Peter W Shor. "Fault-tolerant quantum computation". In *Proceedings of 37th Conference on Foundations of Computer Science*, pages 56–65. IEEE, 1996.

[19] Andrew M Steane. "Error correcting codes in quantum theory". *Physical Review Letters*, 77(5):793, 1996.

[20] Shashank Virmani, Susana F Huelga, and Martin B Plenio. "Classical simulability, entanglement breaking, and quantum computation thresholds". *Physical Review A*, 71(4):042328, 2005.