Editors' Suggestion

Tight Bounds on the Convergence of Noisy Random Circuits to the Uniform Distribution

Abhinav Deshpande,[1,2,*] Pradeep Niroula,[1] Oles Shtanko,[1,†] Alexey V. Gorshkov,[1] Bill Fefferman,[3] and Michael J. Gullans[1]

[1] *Joint Center for Quantum Information and Computer Science and Joint Quantum Institute, NIST–University of Maryland, College Park, Maryland 20742, USA*

[2] *Institute for Quantum Information and Matter, Caltech, Pasadena, California 91125, USA*

[3] *Department of Computer Science, University of Chicago, Chicago, Illinois 60637, USA*

(Received 31 January 2022; revised 9 September 2022; accepted 31 October 2022; published 16 December 2022)

We study the properties of output distributions of noisy random circuits. We obtain upper and lower bounds on the expected distance of the output distribution from the "useless" uniform distribution. These bounds are tight with respect to the dependence on circuit depth. Our proof techniques also allow us to make statements about the presence or absence of anticoncentration for both noisy and noiseless circuits. We uncover a number of interesting consequences for hardness proofs of sampling schemes that aim to show a quantum computational advantage over classical computation. Specifically, we discuss recent barrier results for depth-agnostic and/or noise-agnostic proof techniques. We show that in certain depth regimes, noise-agnostic proof techniques might still work in order to prove an often-conjectured claim in the literature on quantum computational advantage, contrary to what has been thought prior to this work.

DOI: 10.1103/PRXQuantum.3.040329

I. INTRODUCTION

Noise is an unavoidable part of any quantum computing experiment today. The importance of considering the limitations of noisy quantum computers is most palpable in the coinage and the popularity of the term *noisy intermediate-scale quantum (NISQ)* computers [1]. Because of these limitations, the study of quantum algorithms and of their robustness to noise is a problem at the forefront of quantum information science today. With regard to the aim of outperforming classical computers at solving computational problems, two obstacles are noise and limited system size. There is a trade-off between the two, since it is generally challenging to realize a quantum computation with both a large enough number of qubits and low enough error rates. This is the reason why recent demonstrations of "quantum computational advantage" [2–5] have been rightly hailed as exciting developments. It is of prime importance in the field of quantum information today to study the trade-off between system size and noise from the viewpoint of whether or not a given experiment is indeed efficiently simulable on a classical computer.

Separately, in recent years, the study of random quantum circuits has seen renewed vigor, because of their ability to model chaotic [6,7] and complex [8,9] quantum dynamics and the ability to study certain universal properties of subclasses of random circuits using methods from statistical physics [10]. Indeed, random circuits and random ans'atze sometimes inform the modeling of near-term variational quantum algorithms, an example being the barren-plateau problem [11,12].

In this work, we study various properties of noisy random circuits relating to their rate of convergence to the uniform distribution. Specifically, we study circuits of depth d on n qubits with Haar-random two-qubit gates and local Pauli noise, with measurements in the computational basis at the output. The precise rate of convergence of the resulting output distribution to the uniform distribution is a question of much significance in the complexity of random circuit sampling [13,14], the theory of benchmarking noisy circuits [15–20], and the investigation of near-term algorithms [12,21,22]. We prove upper and lower bounds on the expected total variation distance δ of the output distribution (when measuring in the computational basis) to the uniform distribution, which take the form $\delta \sim \exp\left[-\tilde{\Theta}(d)\right]$ [23]. These bounds are tight with

*abhinavd@caltech.edu

†Currently at IBM Quantum, MIT-IBM Watson AI Lab, Cambridge, Massachusetts 02139, USA.

respect to the scaling with d, in the sense that the exponent scales linearly with d. We also study a property known as anticoncentration in noisy and noiseless random circuits. Anticoncentration is a measure of the "flatness" of the output distribution, which is why our results on the closeness to the uniform distribution inform anticoncentration properties. In the literature on quantum advantage via sampling tasks, anticoncentration has been cited as being crucial for output distributions to be classically hard to sample from [24–26].

We first briefly describe our main results and their consequences:

(a) We prove a lower bound on the expected total variation distance of the output distribution from the uniform distribution for local Pauli noise, denoted $\mathbb{E}_{\mathcal{B}}[\delta]$ (Theorem 1) [27]. This takes the form $\mathbb{E}_{\mathcal{B}}[\delta] \geq \exp[-O(d)]$.

(b) We prove an upper bound on the above quantity for the case of a stochastic Pauli-noise channel that we call *heralded dephasing* (Theorem 2). The upper bound takes the form $\mathbb{E}_{\mathcal{B}}[\delta] \leq \mathsf{poly}(n)\exp[-\Omega(d)]$.

(c) We also study anticoncentration properties of noisy and noiseless random circuits. We show that at sublogarithmic depth, there is a severe lack of anticoncentration, strengthening the results of Dalzell *et al.* [28] (Theorem 3).

(d) We complement the above result by showing that noisy random circuits with local Pauli noise do anticoncentrate at higher depth, since they anticoncentrate at least as fast as noiseless random circuits (Theorem 4).

(e) As a result of independent interest, we develop a mapping between noisy random circuits and a model in statistical mechanics in the context of proving our results in Appendix A. This model builds upon tools invented in the study of random circuits [28–32].

Our results have important consequences for the tightness of proof techniques in the field of quantum computational advantage, which we discuss below. Finally, we comment on the relation of our work to recently obtained results by Dalzell *et al.* [33], where bounds on the rate of convergence to the uniform distribution for noisy random circuits have been obtained. That work considers a low-noise limit where the local probability of error in each circuit location (denoted as ϵ) satisfies the scaling $\epsilon(n)n \to 0$ as the number of qubits n tends to infinity. Under these conditions, the authors are able to recover the scaling $\mathbb{E}_{\mathcal{B}}[\delta] \sim \exp\left[-\tilde{\Theta}(nd)\right]$ observed in small-scale numerics [18,19]. In contrast, we consider a more physically natural scaling limit where the noise rate stays constant in the large-n limit. In this case, the analysis of Ref. [33] breaks down. On the other hand, our bounds on the convergence rate to uniformity apply in the limit of vanishing noise rates but the upper bound becomes uninformative. Thus, the two sets of results provide complementary and, apparently, largely nonoverlapping insights into the behavior of noisy quantum circuits on finite systems.

II. CONSEQUENCES

A. Barriers to proof techniques in the complexity theory of random circuit sampling

We now elaborate more on the connection between the hardness of sampling and the hardness of computing output probabilities for random quantum circuits. There has been an effort in the literature to prove, under a reasonable complexity assumption, that approximately sampling from the output distribution of random quantum circuits is classically hard. In order to prove this statement, it suffices to prove that approximating an output probability $p_{00\ldots0} := |\langle 00\ldots0|U|00\ldots0\rangle|^2$ of an n-qubit random quantum circuit U is hard on average [24,34]. More specifically, proving that $p_{00\ldots0}$ is #P-hard to compute to within imprecision (measured in terms of the additive error) $2^{-n}/\mathsf{poly}(n)$ for some polynomial $\mathsf{poly}(n)$ would give the desired claim that approximately sampling from the target distribution to within a small imprecision is classically hard. The target imprecision of 2^{-n} arises from the fact that the Hilbert-space dimension is 2^n and a typical output probability of a given bit string is approximately 2^{-n}.

The state-of-the-art results [14,35,36] on the average-case hardness of computing output probabilities of quantum circuits come close to proving the desired result in a certain sense. The "closeness" is measured in terms of the largest imprecision to which computing the output probability $p_{00\ldots0}$ of a random circuit is still hard on average. The state-of-the-art results prove that computing $p_{00\ldots0}$ is hard to within a smaller imprecision of $2^{-\Theta(nd)}$, matching the required imprecision of $2^{-n}/\mathsf{poly}(n)$ when d is a constant. These results improve upon prior results [37,38] that proved hardness with imprecision $2^{-\mathsf{poly}(n)}$. Such results are often viewed as evidence for the conjecture that $p_{00\ldots0}$ is hard to compute on average to a much larger imprecision of $2^{-n}/\mathsf{poly}(n)$.

1. Shallow-depth random circuits

In Ref. [39], the authors have given important no-go results for proving the desired result, i.e., the average-case hardness of computing $p_{00\ldots0}$ to an imprecision $2^{-n}/\mathsf{poly}(n)$ for a specific class of constant-depth random circuits. Specifically, they have shown that it is in fact classically easy to compute $p_{00\ldots0}$ to within this much imprecision, even when previous techniques have implied that computing $p_{00\ldots0}$ to within a much smaller imprecision of $2^{-\mathsf{poly}(n)}$ is average-case hard. Therefore, these results mean that one cannot, in general, view the hardness of computing $p_{00\ldots0}$ to a smaller imprecision as evidence

for the hardness of computing $p_{00\ldots0}$ to a larger imprecision. In other words, these results constitute a *barrier* for any technique purporting to prove the desired average-case hardness result for general quantum circuits. Any such technique must necessarily be sensitive to the depth of the circuit; otherwise, it would work for constant-depth circuits of the sort studied in Ref. [39] and contradict their easiness results. Barrier results such as this are useful because they rule out certain proof techniques and guide the search for a proof technique that is resistant to these barriers. In this case, the barrier result informs us about depth-sensitivity of a proof technique.

2. Noisy random circuits

A second barrier has been identified by Bouland *et al.* [14], concerning the issue of noise. The authors have shown that existing hardness-proof techniques are applicable to noisy random circuits as well, to yield hardness of computing a noisy output probability $p_{00\ldots0}$ to small imprecision. Contrastingly, for the slightly larger imprecision of $2^{-n}/\mathsf{poly}(n)$, it is known to be easy to compute output probabilities, since the output distribution in the presence of noise is believed to converge rapidly to the uniform distribution [15,18,19]. An algorithm that always outputs "$1/2^n$" successfully computes $p_{00\ldots0}$ to within the required imprecision. Therefore, the results of Ref. [14] exhibit another barrier for hardness-proof techniques purporting to work with higher imprecision—these techniques must distinguish between noiseless and noisy random circuits.

3. Consequences of our results on tightness of proof techniques

Bouland *et al.* [14] have also shown that existing noise-agnostic techniques for proving average-case hardness are almost tight. The logic is that current techniques prove average-case hardness for imprecision $2^{-\Omega(nd\log nd)}$ or smaller. On the other hand, assuming that noisy random circuits become $2^{-\Omega(nd)}$ close in total variation distance to the uniform distribution, as suggested by small-scale numerics [18,19], would mean that it is average-case easy to approximate output probabilities to within imprecision $2^{-O(nd)}$ or larger. Thus, knowing the convergence properties of noisy random circuits sheds light on the tightness of noise-agnostic techniques.

We now discuss the implications of our results on these aforementioned barrier results. First, Theorem 3 casts more light on the depth barrier result in Ref. [39]. Napp *et al.* [39] have proved that for certain short depths $d \leq 3$ on certain architectures, approximating output probabilities to additive error 2^{-n} is easy on average. This constitutes a barrier for improving robustness of proof techniques, since the same output probabilities are formally average-case hard to approximate to within error $2^{-O(n\log n)}$. Our Theorem 3 states that for any sublogarithmic depth $d = o(\log n)$, most output probabilities are at most $2^{-n} \times 2^{-\Theta(n)} = o(2^{-n})$. This indicates a severe lack of anticoncentration, a property that has been linked to the classical hardness of sampling from output distributions of random quantum circuits [24–26]. Because of Theorem 3, a trivial algorithm that always guesses "0" as the output probability turns out to work to within imprecision $\leq 2^{-n}$ with high probability. Therefore, we conclude that any technique to show the conjecture on average-case hardness to within imprecision $2^{-\Theta(n)}$ must not work at depth $d = o(\log n)$. This extends the previously proven regime of $d \leq 3$ [39,40].

Regarding the noise barrier, our result, namely Theorem 1, shows that current noise-agnostic techniques may yet be improved. This is because this theorem disproves the hypothesis that the total variation distance to uniformity follows $\mathbb{E}_{\mathcal{B}}[\delta] \leq \exp[-\Theta(nd)]$ in the asymptotic limit. As mentioned earlier, this latter scaling has been numerically observed at small system sizes [18,19] and hypothesized to hold asymptotically. Since we show that the distance to uniform distribution behaves as $2^{-\tilde{\Theta}(d)}$, there is some scope for improving current noise-agnostic techniques. Moreover, we also show in Corollary 1 that as long as the depth satisfies $d \leq c\log n$ for some constant c, the trivial algorithm "output $1/2^n$" fails with high probability. This is achieved by a strengthening of Theorem 1, where, in addition to a lower bound on the expectation value of the total variation distance $\mathbb{E}_{\mathcal{B}}[\delta]$, we show that this lower bound is *typical*.

To summarize, noise-agnostic techniques can only work in the regime $d = \Omega(\log n)$, $d = O(n)$. In fact, in this regime, the possibility of being able to prove hardness results for both noisy and noiseless circuits to imprecision $2^{-O(n)}$ is not ruled out (see Fig. 1).

4. Surmounting barriers

Owing to this work, we now identify a map of parameter regimes showing where a quantum advantage may be obtained over classical computers and where depth- and noise-agnostic techniques can be used to prove the conjecture on average-case hardness of computing output probabilities to within imprecision at most $O(2^{-n})$. The limitations are explored via two trivial algorithms for computing output probabilities—one always outputting $1/2^n$ and the other always outputting 0. We now speculate on how one might avoid, or surmount, these barrier results.

One possible lesson to glean from these results is that we need fundamentally new techniques to prove average-case hardness of computing output probabilities. Indeed, all known average-case hardness results rely on polynomial interpolation, which has been pioneered by Lipton [42]. In the context of quantum advantage, all known proofs of average-case hardness [14,35–38] construct a univariate

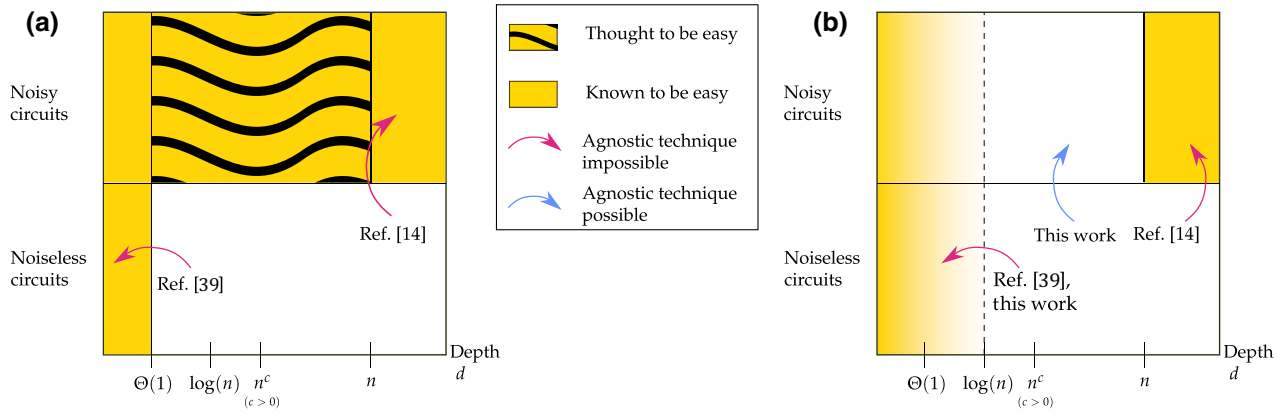


FIG. 1. The status of hardness of approximating random output probabilities to within 2^{-an} imprecision for any constant $a > 0$, (a) before and (b) after our work, for noisy and noiseless circuits at various depths. The results in both (a) and (b) are shown for two dimensions [41]. The solid yellow regions correspond to the regions where the problem is known to be easy (for all a), while the dashed yellow regions are those where the problem has been thought to be easy. The red arrows imply that a technique to show hardness that is either noise-agnostic or depth-agnostic will fail because of the easy region, while the blue arrow indicates that such a technique is not ruled out. The references next to the arrow refer to the works that discuss the presence or absence of a barrier. In (a), the entire shaded region for noisy circuits follows from the assumption that noisy circuits converge to uniformity at a rate $2^{-\Theta(nd)}$. Since we disprove this possibility in Theorem 1, this region is smaller in (b) and allows for a noise-agnostic technique in the regime of $\Theta(\log n) \leq d \leq \Theta(n)$. At large depths, noisy circuits continue to be easy due to their convergence to uniform, as strengthened by Theorem 2. We also extend the easy region at shallow depths from constant to $o(\log n)$ by virtue of Theorem 3. Finally, Theorem 4 implies that the trivial algorithm of outputting 0 for noisy circuits stops working after depth $\Omega(\log n)$.

polynomial in a specific way by perturbing each of the $\Theta(nd)$ gates of the random circuit and using the Feynman path integral. This unavoidably leads to the imprecision depending on the combination nd rather than on just n. An alternative way of constructing polynomials that also explicitly fails for noisy circuits could potentially avoid this drawback and lead to a better dependence of the imprecision.

A second interpretation of our results relies on the following observation. In both known barrier results, the algorithm for computing output probabilities is just a trivial algorithm that outputs a fixed constant independent of the input, i.e., either 0 (by virtue of Theorem 3) or $1/2^n$ (by virtue of Corollary 1 and other results [15]). Suppose that we want to prove a hardness result saying that no PH algorithm that depends nontrivially on the input successfully approximates $p_{00\ldots0}$ on average. Let us call these algorithms "instance-dependent" [43]. Note that instance dependence is an extremely weak requirement; most nontrivial algorithms, e.g., tensor networks and cluster expansions, are instance dependent. Results proving instance-dependent hardness are almost as strong as general hardness results that rule out all algorithms. By our observation from Corollary 1 and Theorem 3 that trivial algorithms generally suffice in the easy cases, we raise the possibility of obtaining these seemingly weaker hardness results that only rule out instance-dependent algorithms, without encountering any barrier. Thus, by finding a suitable enough technique, it may be possible to surmount both the noise and depth barriers. Note that the algorithm of Napp *et al.* [39] does not output a fixed constant. Nevertheless, our result showing that there is an alternative instance-independent algorithm in this regime raises the interesting open question of whether their algorithm fits into the framework of instance-independent algorithms that we have identified here.

B. Benchmarking noise using random circuits

Sampling from random quantum circuits is a leading proposal for demonstrating a quantum computational advantage over classical computers [2,18,37,44,45]. Part of the reason behind the strength of this proposal is the success of the linear cross-entropy measure as a predictor of fidelity [18,20,46] (although there can be exceptions [47]). This is believed to be a major advantage of schemes based on random circuit sampling over schemes based on other quantum sampling problems, such as boson sampling and Gaussian boson sampling. In fact, the resources needed for experimental implementation of benchmarking based on random circuit sampling can scale better than the implementation of randomized benchmarking [20]. The effectiveness of measures such as cross entropy at reflecting the fidelity is crucially related to how the noisy distribution behaves and how close it is to the identity. For example, a crucial fact used in Ref. [20] is that the average fidelity is upper and lower bounded by an exponential function of the

noise strength ($e^{-c\epsilon nd} \leq \mathbb{E}[F] \leq Ke^{-c\epsilon nd}$ for known constants c, K) [48]; therefore, estimating the decay rate of the fidelity serves to estimate the noise strength ϵ. However, like Ref. [33], this result is only applicable to the regime of asymptotically small noise strength $\epsilon \ll 1$.

In the more prevalent and natural regime of $\epsilon = \Theta(1)$, it would be worthwhile to obtain even tighter results on the scaling of the total variation distance between the experimental and uniform distributions with respect to the circuit depth. Our work is the first step in obtaining asymptotically correct decay rates. We prove tight bounds (with respect to scaling with depth) of the form $e^{-c_1\epsilon d} \leq \mathbb{E}_{\mathcal{B}}[\delta] \leq Ke^{-c_2\epsilon d}$ for different c_1, c_2. These results might be useful to show a similar decay behavior for other proxies for the fidelity, such as the linear cross entropy, which is an interesting question for future work. Improving these bounds to obtain the same constants $c_1 = c_2$ for sample-efficient estimators of the fidelity would lead to applications in benchmarking noise based on random circuits.

C. Near-term algorithms

Our results also have important consequences for near-term variational quantum algorithms that can be modeled by random circuits. Variational algorithms with parametrized circuits suffer from the problem of "barren plateaus," which affects the trainability of the circuit due to the vanishing of gradients in the cost-function landscape [11]. Previous work on noise-induced barren plateaus [12] has shown that the gradient of local cost functions vanishes as $\mathsf{poly}(n)\exp[-\Omega(d)]$ in the presence of noise. One possible workaround to avoid the barren-plateau problem would be to use cost functions that are not (sums of) local observables and instead rely on postprocessing the entire data at the output distribution. However, since we give an information-theoretic proof that the output distribution is close to the uniform one, our results (specifically, Theorem 2) imply that even these strategies cannot ameliorate the problem.

More optimistically, at short depths $d = O(\log n)$, Theorem 1 implies that there is enough information content in the output distribution for circuit trainability. We thus avoid a pessimistic conclusion that even constant-depth noisy circuits are untrainable, a conclusion that would have followed from the incorrect hypothesis that noisy circuits converge to uniformity at a rate $2^{-\Theta(nd)}$.

D. Monitored random circuits and entanglement phase transitions

As mentioned above, we study the upper bounds on the expected total variation distance to the uniform distribution for a noise channel called Heralded dephasing corresponds exactly to the situation where, after each layer of the circuit, a random fraction p of qubits is measured. This is the model of monitored random circuits, or random circuits with intermediate measurements, that have been studied in the many-body-physics literature [10,16,49–53].

These models feature entanglement and purity phase transitions as a function of measurement strength (the probability p of measuring a qubit at a given time), exhibiting a volume-law behavior of entanglement entropy of the final pure state for $p < p_c$ and an area-law behavior above the critical point ($p > p_c$). These transitions have also been linked to computational-complexity phase transitions [39,54,55]. One can study the computational complexity of sampling from the output distribution of the random circuit conditioned on knowledge of the measurement locations. This problem can be studied both in the setting where the results of the intermediate measurements are known or unknown. The case where the results of the intermediate measurements are unknown corresponds to heralded dephasing. As a consequence of our work, we conclude that for the problem of approximate sampling from the output distribution of monitored random circuits with unknown intermediate measurement results, the problem becomes easy for classical computers for any depth $d = \omega(\log n)$ and any $p > 0$. Therefore, there is no complexity transition in this setting. It remains open whether there is a complexity transition in the setting where the intermediate measurement results are also known.

III. PRIOR WORK

In this section, we summarize prior work. The foundational work by Aharonov *et al.* [15] dealt with the convergence of states evolving under arbitrary quantum circuits with depolarizing noise to the maximally mixed state. They showed that the entropy of the system reaches its maximal value exponentially fast with d. They concluded that noisy circuits (without error correction) are essentially "worthless" after logarithmic depth $d = \Omega(\log n)$. Translated to our setting, their proof techniques imply that the total variation distance δ to the uniform distribution satisfies $\delta \leq 2^{-\Omega(d)}$. Depolarizing noise is also studied in more detail in Refs. [56,57]. Ben-Or *et al.* [58] have generalized the work of Aharonov *et al.* [15] to other forms of noise and observed that the time after which noisy circuits are worthless depends on the class of channels.

More recently, Gao and Duan [13] have studied the case of circuits with a more generalized form of Pauli noise and have shown an upper bound on the distance to the uniform distribution of the form $\mathbb{E}[\delta] = \exp[-\Omega(d)]$. This result, however, does not work for the case of dephasing noise. In fact, the upper bound for dephasing noise is a constant independent of n and d, which is not informative. Moreover, the result explicitly assumes a property of random quantum circuits known as anticoncentration.

Note that for dephasing noise, it is not possible to prove a general upper bound on δ that works for arbitrary circuits by using the techniques of Ref. [13]. This is because a

state in the computational basis is unaffected by dephasing and hence there are instances (e.g., circuits with diagonal gates in the computational basis) where the state is always in the computational basis throughout the evolution and the output distribution is unaffected by the noise channels. Techniques such as the Pauli twirl do not work either, since we are interested in a quantity that is not a linear observable of ρ.

Prior work on anticoncentration has mostly been via second-moment bounds on the output probabilities [26,28,32,34,59], which can be proved via the design property of random circuits [26,60,61]. An exception is the case of the one-clean-qubit model (DQC1) [62]. A tool to analyze random circuits that has proven particularly fruitful is that of mapping to models in statistical mechanics. This method has been successfully used in prior work (see, e.g., [28–32], among others).

IV. DEFINITIONS

We define a depth-d noisy circuit on n qubits as a sequence of quantum channels

$$\mathcal{N}_d(\rho) = (\mathcal{E}_d \circ \mathcal{C}_d \cdots \mathcal{E}_1 \circ \mathcal{C}_1)(\rho), \tag{1}$$

where $\mathcal{C}_m(\rho) = U_m \rho U_m^\dagger$ is a unitary operation on n qubits and $\mathcal{E}_m(\rho)$ is a noise channel. A noisy Haar-random circuit is a noisy circuit for which each U_m is decomposable into tensor products of two-qubit Haar-random gates.

The results in our work are subject to different assumptions on the noise model. We always work with *local* noise, where the channels $\mathcal{E}_m$ act on at most $k = \Theta(1)$ many qubits at a time. We call a noise channel *stochastic* if it can be written in Kraus form, where at least one Kraus operator is a positive multiple of the identity. *Pauli* noise is described in general by $\mathcal{E}_m(\rho) = \sum_{E\in\mathcal{P}} p_m(E) E\rho E^\dagger$, where $\mathcal{P}$ is the Pauli group on n qubits and $p_m(E)$ is a probability distribution over the Pauli-group elements. The local noise rate in Pauli sector $\mu \in \{I, X, Y, Z\}$ at depth m on site i is the marginal distribution $q_{\mu mi} = \sum_{E\in\mathcal{P}, E_i=\mu} p_m(E)$. Most results in this work are applicable to Pauli-noise channels, though we sometimes discuss what other noise channels the proof techniques could be applicable to. It is not too restrictive to consider Pauli-noise channels, since in practice one may use strategies such as randomized compilation [63] to implement random circuit sampling or benchmarking based on random circuit sampling. It is proved in Ref. [63] that this technique tailors correlated and even coherent noise into stochastic Pauli noise.

We consider circuit architectures where, for simplicity, the gates are applied in parallel. More formally, we define a *parallel architecture* as one for which n is an even number and every qubit is involved in a two-qubit gate at every unit of depth [64]. Our results can be easily extended to more general gate-layering strategies with a suitable redefinition of the depth. For a given site i, we define $n_m(i)$ as the set of *neighbors* of i, which are the sites involved in a two-qubit gate at circuit layer m with i, including i itself. We extend n_m to all subsets of sites $A \subset \{1, \ldots, n\}$ through the composition rule $n_m(A \cup B) = n_m(A) \cup n_m(B)$. We define the *forward* and *backward light cones* $L_d(i)$ and $L_d^\dagger(i)$, respectively, of site i at depth d as the sets

$$L_d(i) = n_d \circ \cdots \circ n_1(i), \tag{2}$$

$$L_d^\dagger(i) = n_1 \circ \cdots \circ n_d(i). \tag{3}$$

In general, we have the bounds $|L_d(i)|, |L_d^\dagger(i)| \leq 2^d$.

We denote the above ensemble of noisy parallel circuits as $\mathcal{B}$ and our results are generally stated in terms of expectation values over this ensemble, denoted $\mathbb{E}_\mathcal{B}$. The ensemble $\mathcal{B}$ describes the distribution over unitaries only and not over the possible noise operations. We also denote $\delta = \|\mathcal{D} - \mathcal{U}\|_{\text{TVD}}$, where $\mathcal{D}$ denotes the (noisy) output distribution when measuring in the computational basis, $\mathcal{U}$ is the uniform distribution over n-bit strings, and "TVD" denotes the total variation distance. A central consequence of our results relates to anticoncentration properties of low-depth noisy random circuits. Here, we provide a formal definition. In the following, we denote by $p_{00\ldots0}$ the output probability of the string $00\ldots0$. For noiseless circuits, we have $p_{00\ldots0} = |\langle 00\ldots0|U|00\ldots0\rangle|^2$, while for noisy circuits it is given by $p_{00\ldots0} = \text{Tr}[|00\ldots0\rangle\langle 00\ldots0| \mathcal{N}_d(|00\ldots0\rangle\langle 00\ldots0|)]$.

Definition 1: A family of random circuit ensembles is anticoncentrated if there exist constants $\alpha \in (0, 1]$ and $c > 0$ such that

$$\Pr_\mathcal{B}\left[p_{00\ldots0} \geq \frac{\alpha}{2^n}\right] \geq c. \tag{4}$$

A stronger definition of anticoncentration is often given in terms of the *collision probability*, which we define for noisy circuits as

$$Z(U, \mathcal{E}) = \sum_{x\in\{0,1\}^n} p_x^2. \tag{5}$$

We denote the collision probability for noiseless circuits as $Z(U, \mathbb{I})$. The strong definition of anticoncentration requires that $2^n \mathbb{E}_\mathcal{B}[Z(U, \mathcal{E})] \leq c$ for some constant $c > 0$ [28,32,65]. It follows from a standard argument provided in the proof of Theorem 4 that, under this stronger definition of anticoncentration, the random circuit ensemble also satisfies Definition 1; however, the converse does not hold.

V. RESULTS

A. Lower bound on the distance to the uniform distribution

Theorem 1: *For a Haar-random circuit on any parallel-circuit architecture subject to local Pauli noise with a uniform upper bound on the local noise rate $q_{\mu mi} \leq q_\mu$ for all Pauli-noise sectors μ, m, and i, we have the lower bound*

$$\mathbb{E}_{\mathcal{B}}[\delta] \geq \frac{(1-2b)^{2d}}{4 \times 30^d}, \tag{6}$$

where $b = \min[q_x + q_y, q_y + q_z, q_z + q_x] \leq 1/2$.

Proof. Let p_x denote the probability of observing the bit string x at the output for a fixed circuit chosen from $\mathcal{B}$ and with a local noise channel applied after every gate. For a region A, let x_A be the value of x restricted to the sites in A. By definition, the total variation distance is the maximum difference in probabilities ascribed to any event E by the two distributions, i.e., $\delta(\mathcal{D},\mathcal{U}) = \max_E[|\Pr_{\mathcal{D}}[E] - \Pr_{\mathcal{U}}[E]|]$. Considering the event that the first qubit is measured to be in the state $|0\rangle$, the total variation distance satisfies

$$\delta \geq |p_0 - \frac{1}{2}|, \tag{7}$$

where p_0 and p_1 are the probabilities of observing a 0 and a 1 on the first qubit at the (noisy) output distribution, respectively. Now since the above is a quantity in $[0,1]$, it satisfies $|p_0 - \frac{1}{2}| \geq |p_0 - \frac{1}{2}|^2$. We now lower bound the expectation value of the latter quantity over parallel circuits, giving us $\mathbb{E}_{\mathcal{B}}[\delta] \geq \mathbb{E}_{\mathcal{B}}[p_0^2 - p_0 + 1/4]$. Observe that for a given circuit,

$$\begin{aligned} p_0 &= \mathrm{Tr}[|0\rangle\langle 0|_1 \mathcal{N}_d(|00\ldots0\rangle\langle 00\ldots0|)] \qquad (8)\\ &= \sum_{x,x_1=0} \sum_{y,z,\ldots} \sum_{y',z',\ldots} \langle x|V_d \ldots U_2|z\rangle\langle z|V_1|y\rangle \\ &\quad \times \langle y|U_1|00\ldots0\rangle \\ &\quad \times \langle 00\ldots0|U_1^\dagger|y'\rangle\langle y'|V_1^\dagger|z'\rangle\langle z'|U_2^\dagger \ldots V_d|x\rangle. \qquad (9) \end{aligned}$$

In the above, the gates V_m are purifications of the noise channels $\mathcal{E}_m$ over ancillary qubits that are traced out. The above equation is simply a Feynman-path-integral representation of a noisy probability, from which it is evident that p_0 is a degree-two polynomial in the gate entries of the Haar-random gates (or, more accurately, a degree-one polynomial in the entries of $U \otimes U^\dagger$).

This fact means that the expectation value $\mathbb{E}_{\mathcal{B}}[p_0^2 - p_0 + 1/4]$, which is over parallel circuits with local Haar-random gates, can be replaced by an expectation value over parallel circuits with local random Clifford gates $\mathbb{E}_C$ because of the 2-design property of the Clifford group. Consider the quantity $\mathbb{E}_C[p_0^2 - p_0 + 1/4]$, which we reexpress as follows:

$$\mathbb{E}_C[p_0^2 - p_0 + 1/4] = \frac{1}{4} - \mathbb{E}_C[p_0(1-p_0)]. \tag{10}$$

We lower bound this quantity by upper bounding $\mathbb{E}_C[p_0(1-p_0)]$. We define the probability density $f(p)$ through $f(p)dp = \Pr_C(p_0 \in [p, p+dp])$ and write

$$\mathbb{E}_C[p_0(1-p_0)] = \int_0^1 dp f(p) p \times (1-p). \tag{11}$$

Our overall strategy is as follows. First, observe that since the quantity $p(1-p)$ takes the maximum value $1/4$ in the interval $p \in [0,1]$, a crude upper bound for Eq. (11) is simply $1/4 \times \int_0^1 dp f(p) = 1/4$. This is a useless bound, since it results in the conclusion $\mathbb{E}_{\mathcal{B}}[\delta] \geq 0$. We refine this useless bound slightly by observing that at least some instances of the Clifford ensemble lead to a value of p bounded away from $1/2$, implying that $p(1-p)$ is bounded away from $1/4$. This will result in a better upper bound on $\mathbb{E}_C[p_0(1-p_0)]$, which will translate into a better lower bound on $\mathbb{E}_{\mathcal{B}}[\delta]$.

We split the integral in Eq. (11) into two parts—those with $p < \frac{1}{2} + \epsilon$ and those with $p \geq \frac{1}{2} + \epsilon$, for some $\epsilon \in (0, 1/2)$:

$$\begin{aligned} \mathbb{E}_C[p_0(1-p_0)] &= \int_0^{(1/2)+\epsilon} dp f(p) p \times (1-p) \\ &\quad + \int_{(1/2)+\epsilon}^1 dp f(p) p \times (1-p) \\ &\leq \frac{1}{4} \int_0^{(1/2)+\epsilon} dp f(p) + \left(\frac{1}{4} - \epsilon^2\right) \\ &\quad \times \int_{(1/2)+\epsilon}^1 dp f(p). \qquad (12) \end{aligned}$$

In the above, we use the fact that if $p \geq 1/2 + \epsilon$, then $p(1-p) \leq 1/4 - \epsilon^2$. Continuing, we obtain

$$\begin{aligned} \mathbb{E}_C[p_0(1-p_0)] &\leq \frac{1}{4}\left(1 - \Pr_C\left(p_0 \geq \frac{1}{2} + \epsilon\right)\right) \\ &\quad + \left(\frac{1}{4} - \epsilon^2\right) \Pr_C\left(p_0 \geq \frac{1}{2} + \epsilon\right) \qquad (13)\\ &= \frac{1}{4} - \epsilon^2 \Pr_C\left(p_0 \geq \frac{1}{2} + \epsilon\right). \qquad (14) \end{aligned}$$

It only remains to lower bound $\Pr_C(p_0 \geq 1/2 + \epsilon)$. For this, we observe that we can take an extreme case over circuits satisfying a certain property. As long as these circuits result in a final state with $p_0 \geq \frac{1}{2} + \epsilon$ and the likelihood

of applying these (Clifford) circuits is large enough, we are done. The extreme case that we choose is simple: it consists of Clifford circuits where all of the d two-qubit Clifford gates that touch the first qubit map the Pauli operator Z_1 to Z_1. In particular, for Pauli-noise channels, this means that the first qubit is never entangled with any other qubit in the system and any mixture of the $|0\rangle\langle 0|$ and $|1\rangle\langle 1|$ states is unchanged by the unitary dynamics. Effectively, the only evolution acting upon it is the noise channel after every layer. Although these are rare events, we prove that they still lead to a nontrivial lower bound $\Pr_C(p_0 \geq 1/2 + \epsilon)$.

We use Lemma 1, which analyzes this case and shows that for a single-qubit evolution under only Pauli noise, $|p_0 - 1/2| = \left(1 - 2(q_x + q_y)\right)^d/2$. This means that we can take $\epsilon = \left(1 - 2(q_x + q_y)\right)^d/2$. Furthermore, the likelihood of observing this extreme event is lower bounded away from 0. In each layer, the probability of applying a Clifford circuit with the property above is at least $1/30$ (this follows, for example, from a brute-force numerical evaluation for each of the 11 520 two-qubit Clifford gates). As a result, we obtain $\Pr_C(p_0 \geq 1/2 + \epsilon) \geq 1/30^d$ for $\epsilon = \left(1 - 2(q_x + q_y)\right)^d/2$.

Wrapping everything up, this results in

$$\mathbb{E}_C[p_0(1-p_0)] \leq \frac{1}{4} - \frac{\left(1 - 2(q_x + q_y)\right)^{2d}/4}{30^d} \tag{15}$$

$$\implies \mathbb{E}_C[p_0^2 - p_0 + 1/4] \geq \frac{\left(1 - 2(q_x + q_y)\right)^{2d}}{4 \times 30^d} \tag{16}$$

$$\implies \mathbb{E}_{\mathcal{B}}[p_0^2 - p_0 + 1/4] \geq \frac{\left(1 - 2(q_x + q_y)\right)^{2d}}{4 \times 30^d} \tag{17}$$

$$\implies \mathbb{E}_{\mathcal{B}}[\delta] \geq \frac{\left(1 - 2(q_x + q_y)\right)^{2d}}{4 \times 30^d}. \tag{18}$$

A comment is in order. First, note that the dependence on $q_x + q_y$ can also be written as $q - q_z$, where $q = q_x + q_y + q_z$. For the extreme event we have considered, it is understandable that dephasing noise (where $q = q_z$) does not affect the quantity p_0. For these cases, we can strengthen the lower bound by considering events where the first gates act as Hadamards or Hadamards plus a phase gate S, the intermediate gates map X_1 to X_1 or Y_1 to Y_1, respectively, and the last gate inverts the first. By symmetry, the previous analysis holds for these events as well. We can combine everything to give the slightly better bound

$$\mathbb{E}_{\mathcal{B}}[\delta] \geq \frac{(1-2b)^{2d}}{4 \times 30^d}, \tag{19}$$

where $b = \min[q_x + q_y, q_y + q_z, q_z + q_x] = q - \max[q_x, q_y, q_z]$. ■

We also note that for the case of perfect depolarizing noise on every qubit, we have $q_x = q_y = q_z = 1/4$. This gives the trivial bound $\mathbb{E}_{\mathcal{B}}[\delta] \geq 0$, as it should, because perfect depolarizing noise immediately gives the identity operator on every qubit and the distance to the uniform distribution is exactly 0. Similarly, for a Haar-random unitary, we have an estimate from the Porter-Thomas distribution $\Pr(p_{00\ldots0} = p) = (2^n - 1)(1-p)^{2^n - 2}$ [18], leading to

$$\mathbb{E}_U[\delta] = 2^{n-1} \int_0^1 dp\, |p - \frac{1}{2^n}| \Pr(p_{00\ldots0} = p) \geq e^{-1}. \tag{20}$$

Thus, in the case of noiseless Haar-random circuits on architectures where the output probability approaches the Porter-Thomas distribution, our lower bound is expected to be a significant underestimate of the true value at large depths. Barak *et al.* [32] have also obtained a similar lower bound for single-qubit marginal probabilities in the noiseless case.

Lemma 1: *Consider a single qubit starting in the state* $\rho = |0\rangle\langle 0|$. *After* d *applications of the channel* $\mathcal{E}(\rho) = (1-q)\rho + q_x X\rho X + q_y Y\rho Y + q_z Z\rho Z$ *(where* $q = q_x + q_y + q_z$*), the resulting state* $\mathcal{E}^d(\rho)$ *obeys*

$$p_0 := \mathrm{Tr}\left[\frac{\mathbb{1} + Z}{2}\mathcal{E}^d(\rho)\right] = \frac{1}{2} + \frac{1}{2}\left(1 - 2(q_x + q_y)\right)^d. \tag{21}$$

Proof. It suffices to show that $\mathcal{E}^d(\rho) = p_0(d)\,|0\rangle\langle 0| + (1 - p_0(d))\,|1\rangle\langle 1|$, where $p_0(d) = \frac{1}{2} + \frac{1}{2}\left(1 - 2(q_x + q_y)\right)^d$. We prove this by induction. The $d = 0$ base case follows from the given initial state. The inductive step is

$$\begin{aligned}
\mathcal{E}^{d+1}(\rho) &= \mathcal{E}(\mathcal{E}^d(\rho)) \\
&= (1 - q_x - q_y)[p_0(d)\,|0\rangle\langle 0| \\
&\quad + (1 - p_0(d))\,|1\rangle\langle 1|] + (q_x + q_y)[p_0(d)\,|1\rangle\langle 1| \\
&\quad + (1 - p_0(d))\,|0\rangle\langle 0|] \\
&= p_0(d+1)\,|0\rangle\langle 0| + (1 - p_0(d+1))\,|1\rangle\langle 1|,
\end{aligned} \tag{22}$$

which completes the proof. ■

We also remark here that the lower bound (with potentially different constants) is applicable to all noise channels for which an analog of Lemma 1 holds. The only condition we need is that after d applications of the single-qubit noise channel to the initial state $|0\rangle\langle 0|$, the resulting state has $|p_0(d) - 1/2| \geq \exp[-ad]$ for some $a > 0$. If this is satisfied, then one can take as an extreme case in the proof of Theorem 1 the Clifford circuits where all of the d two-qubit Cliffords encountered by the first qubit act as identity on it. This would change the denominator 30^d in Eq. (19)

to $11\,520^d$, although similar symmetry arguments could improve this constant. The same considerations apply to Theorem 3, the proof of which makes use of the analysis in Theorem 1.

As we mentioned in Sec. I, Theorem 1 has wide-ranging consequences for the properties of noisy random circuits. To strengthen the result, we now show that a similar lower bound applies to *typical* noisy random circuits (i.e., individual elements of the ensemble that occur with high probability) at sufficiently low depths.

Corollary 1: *For a noisy Haar-random circuit on any parallel-circuit architecture with a uniform upper bound on the local noise rate $q_{\mu mi} \leq q_\mu$ for all Pauli-noise sectors μ, m, and i, we have the bound*

$$\Pr_{\mathcal{B}}(\delta < e^{-2ad}) \leq 8e^{-ad} + 16e^{(2a+\log 4)d}/n, \tag{23}$$

where $a = -2\log(1-2b) + \log(30)$ *and* $b = \min[q_x + q_y, q_y + q_z, q_z + q_x]$.

Proof. Following the arguments from Theorem 1, we have the bound

$$\delta \geq \Delta_i^2 := \frac{1}{4}\langle 0|\mathcal{N}_d^\dagger(Z_i)|0\rangle^2 \tag{24}$$

for every site i, where $\mathcal{N}_d^\dagger$ is the adjoint map of the depth-d noisy circuit $\mathcal{N}_d$. We can, thus, obtain a lower bound by the site-averaged quantity

$$\delta \geq \Delta^2 := \frac{1}{n}\sum_i \Delta_i^2. \tag{25}$$

In Theorem 1, we show that $\mathbb{E}_{\mathcal{B}}[\Delta_i^2] \geq e^{-ad}/4$, which implies $\mathbb{E}_{\mathcal{B}}[\Delta^2] \geq e^{-ad}/4$. When the noise is given by a Pauli-noise channel, as we consider, then the operator $\mathcal{N}_d^\dagger(Z_i)$ has support on at most $L_d^\dagger(i)$ sites. We can use this fact to bound the second moment of the site-averaged quantity:

$$\mathbb{E}_{\mathcal{B}}[\Delta^4] = \frac{1}{n^2}\sum_{i,j}\mathbb{E}_{\mathcal{B}}[\Delta_i^2\Delta_j^2] \tag{26}$$

$$= \frac{1}{n^2}\sum_{i;j\in L_d\circ L_d^\dagger(i)}\mathbb{E}_{\mathcal{B}}[\Delta_i^2\Delta_j^2] + \frac{1}{n^2}\sum_{i;j\notin L_d\circ L_d^\dagger(i)}\mathbb{E}_{\mathcal{B}}[\Delta_i^2]\mathbb{E}_{\mathcal{B}}[\Delta_j^2] \tag{27}$$

$$= \frac{1}{n^2}\sum_{i;j\in L_d\circ L_d^\dagger(i)}(\mathbb{E}_{\mathcal{B}}[\Delta_i^2\Delta_j^2] - \mathbb{E}_{\mathcal{B}}[\Delta_i^2]\mathbb{E}_{\mathcal{B}}[\Delta_j^2]) + \mathbb{E}_{\mathcal{B}}[\Delta^2]^2 \tag{28}$$

$$\leq \mathbb{E}_{\mathcal{B}}[\Delta^2]^2 + \frac{1}{n}\max_{i;j\in L_d\circ L_d^\dagger(i)}|L_d\circ L_d^\dagger(i)|\sigma_i\sigma_j, \tag{29}$$

where we define $\sigma_i^2 = \mathbb{E}_{\mathcal{B}}[\Delta_i^4] - \mathbb{E}_{\mathcal{B}}[\Delta_i^2]^2 \leq 1$. To see this, we use the Cauchy-Schwartz inequality:

$$\mathbb{E}_{\mathcal{B}}\left[(\Delta_i^2 - \mathbb{E}_{\mathcal{B}}[\Delta_i^2])(\Delta_j^2 - \mathbb{E}_{\mathcal{B}}[\Delta_j^2])\right] \leq \sigma_i\sigma_j, \tag{30}$$

whenever $j \in L_d \circ L_d^\dagger(i)$. For j outside this set, the cross-correlation is zero. Applying the lower bound from Theorem 1, we now have the sequence of inequalities

$$\Pr_{\mathcal{B}}[\delta \leq e^{-2ad}] \leq \Pr_{\mathcal{B}}\left[\delta \leq 4e^{-ad}\mathbb{E}_{\mathcal{B}}[\Delta^2]\right] \tag{31}$$

$$\leq \Pr_{\mathcal{B}}\left[\Delta^2 \leq 4e^{-ad}\mathbb{E}_{\mathcal{B}}[\Delta^2]\right] \tag{32}$$

$$= 1 - \Pr_{\mathcal{B}}\left[\Delta^2 > 4e^{-ad}\mathbb{E}_{\mathcal{B}}[\Delta^2]\right] \tag{33}$$

$$\leq 1 - (1-4e^{-ad})^2\frac{\mathbb{E}_{\mathcal{B}}[\Delta^2]^2}{\mathbb{E}_{\mathcal{B}}[\Delta^4]} \tag{34}$$

$$\leq 1 - \frac{(1-4e^{-ad})^2\mathbb{E}_{\mathcal{B}}[\Delta^2]^2}{\mathbb{E}_{\mathcal{B}}[\Delta^2]^2 + (1/n)\max_{i;j\in L_d\circ L_d^\dagger(i)}[|L_d\circ L_d^\dagger(i)|\sigma_i\sigma_j]}$$

$$\leq \frac{8e^{-ad}\mathbb{E}_{\mathcal{B}}[\Delta^2]^2 + (1/n)\max_{i;j\in L_d\circ L_d^\dagger(i)}[|L_d\circ L_d^\dagger(i)|\sigma_i\sigma_j]}{\mathbb{E}_{\mathcal{B}}[\Delta^2]^2 + \frac{1}{n}\max_{i;j\in L_d\circ L_d^\dagger(i)}[|L_d\circ L_d^\dagger(i)|\sigma_i\sigma_j]} \tag{35}$$

$$\leq \frac{8e^{-ad}\mathbb{E}_{\mathcal{B}}[\Delta^2]^2 + (1/n)\max_{i;j\in L_d\circ L_d^\dagger(i)}[|L_d\circ L_d^\dagger(i)|\sigma_i\sigma_j]}{\mathbb{E}_{\mathcal{B}}[\Delta^2]^2} \tag{36}$$

$$\leq 8e^{-ad} + 16\times 4^d e^{2ad}/n, \tag{37}$$

where we apply the Paley-Zygmund inequality in Eq. (34), use Eq. (29) in Eq. (35), and use the fact that $|L_d \circ L_d^\dagger(i)| \leq 4^d$ for every site i to obtain Eq. (36). ■

Corollary 1 shows that for any depth that grows slower than

$$d < \left(\frac{1}{2a + \log 4}\right) \log n, \tag{38}$$

for most circuits the total variation distance δ is lower bounded by e^{-cd} for some fixed constant c. This strengthens Theorem 1 by showing that the lower bound on $\mathbb{E}_{\mathcal{B}}[\delta]$ in Theorem 1 is actually characteristic of typical circuits at these depths. The typicality result rules out the possibility that $\mathbb{E}_{\mathcal{B}}[\delta]$ is dominated by rare circuits with unusually large deviation δ. It is an interesting subject for future work to study similar typicality results for higher depths that scale polynomially with n.

B. Upper bound on the distance to uniform

In this section, we study a partially heralded noise model where first a random set of sites are selected after each layer of the circuit independently with probability p. At each site i in this random subset, a local dephasing channel $\mathcal{E}_i$ is applied with dephasing parameter q, where

$$\mathcal{E}_i(\rho) = (1-q)\rho + qZ_i\rho Z_i. \tag{39}$$

In the limit $p \to 1$, this becomes a standard local dephasing model with parameter q, while $q \to 1/2$ is equivalent to a model where a random set of sites are measured at rate p in the Z basis but without keeping track of the measurement outcomes. For $p < 1$, we absorb the random locations of the dephasing events into the ensemble $\mathcal{B}$.

The "heralding" refers to the fact that the set of sites where the measurements occurred is known but not the measurement outcomes. Note that this is different from a dephasing model, where each site is uniformly dephased with dephasing parameter pq. In particular, the noise locations act as an additional source of randomness in the model.

We focus on this noise model for two reasons. First, we would like to verify the intuition that noise acting during a random unitary circuit renders the output distribution "worthless" (close to the uniform distribution) after logarithmic depth [15], even though there are atypical circuits that can avoid the effects of the noise in a heralded-dephasing model. The second motivation arises from the observation that when the measurement outcomes are also known, such models exhibit an entanglement transition in the conditional evolution of the quantum states as mentioned in Sec. II D [10]. Our analysis of the heralded-dephasing model proves that discarding the measurement outcomes but maintaining knowledge of the noise locations is enough to remove any signature of such entanglement transitions.

For a noisy Haar-random circuit with this noise model, we prove an upper bound on the circuit-averaged total variation distance δ that is independent of the circuit architecture.

Theorem 2: *For a noisy Haar-random circuit on any parallel-circuit architecture with heralded-dephasing noise at rate p with the dephasing parameter q, we have the upper bound*

$$\mathbb{E}_{\mathcal{B}}[\delta] < \frac{3^{2/3}}{2} n^{1/3} e^{-\gamma p d/3}, \tag{40}$$

where $\gamma = 8q(1-q)/3$.

Proof. We start from Pinsker's inequality, which states that the total variation distance between distributions P and Q is related to the corresponding Kullback-Leibler (KL) divergence (the classical relative entropy) by the relation $\delta(P,Q) \leq \sqrt{D_{\mathrm{KL}}(P||Q)/2}$. The KL divergence with respect to the uniform distribution $\mathcal{U}$ is given by $n\log 2 - H(P)$, where $H(P)$ is the Shannon entropy. Let $\mathcal{D}$ denote the distribution of measurements for a circuit. This gives us the following chain of inequality for variation distance to the uniform distribution:

$$2\delta(\mathcal{D},\mathcal{U})^2 \leq D_{KL}(\mathcal{D}||\mathcal{U}) \tag{41}$$

$$= n\log 2 - H(\mathcal{D}) \leq n\log 2 - H_2(\mathcal{D}), \tag{42}$$

where the last inequality follows from the fact that second Rényi entropy $H_2(\mathcal{D}) = -\log\left(\sum_x p_x^2\right)$ is less than or equal to the Shannon entropy $H(\mathcal{D}) = H_{\alpha\to 1}(\mathcal{D}) = -\sum_x p_x \log p_x$. We can now use Markov's inequality to bound the average TVD. For any $\epsilon \in [0,1]$, letting $\Pr_{\mathcal{B}}(\delta = \sigma)$ denote the probability *density* of the continuous variable δ, we have

$$\mathbb{E}_{\mathcal{B}}(\delta) = \int_0^\epsilon d\sigma\, \sigma \Pr_{\mathcal{B}}(\delta = \sigma) + \int_\epsilon^1 d\sigma\, \sigma \Pr_{\mathcal{B}}(\delta = \sigma) \tag{43}$$

$$\leq \epsilon + \Pr_{\mathcal{B}}(\delta \geq \epsilon) \leq \epsilon + \frac{\mathbb{E}_{\mathcal{B}}(\delta^2)}{\epsilon^2}. \tag{44}$$

If $\mathbb{E}_{\mathcal{B}}(\delta^2)$ decays exponentially or faster with depth, i.e., $e^{-\gamma d}$ for some $\gamma > 0$, we can take $\epsilon = e^{-\gamma d/3}$ to ensure that $\mathbb{E}_{\mathcal{B}}(\delta)$ decays exponentially as $e^{-\gamma d/3}$. To show that

the second moment of the TVD must indeed decay exponentially with d, we calculate the expectation of Sec. V B:

$$\begin{aligned}\mathbb{E}_{\mathcal{B}}[2\delta(\mathcal{D},\mathcal{U})^2] &\le n\log 2 - \mathbb{E}_{\mathcal{B}}(H_2(\mathcal{D})) \\ &= n\log 2 + \mathbb{E}_{\mathcal{B}}\left[\log\left(\sum_x p_x^2\right)\right] \\ &\le n\log 2 + \log \mathbb{E}_{\mathcal{B}}\left[\sum_x p_x^2\right]. \end{aligned} \tag{45}$$

In the last inequality, we use Jensen's inequality for concave functions, $\mathbb{E}_{\mathcal{B}}[f(X)] \le f(\mathbb{E}_{\mathcal{B}}[X])$. The term inside the expectation function, $\sum_x p_x^2$, is the collision probability. From Lemma 2, we have that the expectation of the collision probability is upper bounded by $2^{-n}\exp\left[(n/3)e^{-\gamma p d}\right]$, where $\gamma = 8q(1-q)/3$. With this, we have

$$\begin{aligned}\mathbb{E}_{\mathcal{B}}[2\delta(\mathcal{D},\mathcal{U})^2] &\le n\log 2 + \log\left[2^{-n}\exp\left[\frac{n}{3}e^{-\gamma p d}\right]\right] \\ &= \frac{n}{3}e^{-\gamma p d}. \end{aligned} \tag{46}$$

Thus, we have that the second moment of the TVD decays exponentially in circuit depth. The right-hand side of Eq. 44 is minimized at $\epsilon = (n/3)^{1/3}\, e^{-\gamma p d/3}$. This yields the desired bound

$$\mathbb{E}_{\mathcal{B}}(\delta) \le \frac{3^{2/3}}{2} n^{1/3} e^{-\gamma p d/3}. \tag{47}$$

■

To complete the proof, it remains to prove the following lemma.

Lemma 2: *For a noisy Haar-random circuit on any parallel-circuit architecture with heralded-dephasing noise at rate p with the dephasing parameter q, we have the upper bound on the collision probability Z*

$$\mathbb{E}_{\mathcal{B}}[Z] = \mathbb{E}_{\mathcal{B}}\left[\sum_x p_x^2\right] \le 2^{-n}\exp\left[\frac{n}{3}e^{-\gamma p d}\right], \tag{48}$$

where $\gamma = 8q(1-q)/3$.

To prove this bound, we make use of the statistical-mechanics mapping method developed by Dalzell *et al.* [28]. The proof of Lemma 2 can be found in Appendix A.

We compare our proof technique with that of Aharonov *et al.* [15]. In that work, the authors prove that under depolarizing noise, the information [as measured by $I = n\log 2 - S(\rho)$ in terms of the von Neumann entropy $S(\rho)$] decays exponentially in the depth for any arbitrary circuit. As mentioned earlier, since we cannot hope to have a result for arbitrary circuits for dephasing noise, we consider the expectation value of the *classical* information measured through the KL divergence of the classical distribution $\mathcal{D}$—i.e., $D_{\mathrm{KL}}(\mathcal{D},\mathcal{U}) = n\log 2 - H(\mathcal{D})$—and prove that this expectation value decays exponentially in the depth.

C. No-go result for anticoncentration at low depth

In this section, we study the properties of quantum circuit dynamics at sublogarithmic depth, which is defined as a limit where we fix $0 \le c < 1$ and scale depth as $d = O[(\log n)^c]$ while taking $n \to \infty$. At this depth, there is still a notion of locality in the circuit because the light cone $L_d(i)$ of each site cannot extend across the whole system in the large-n limit. We prove that sampling from Haar-random circuits on any parallel-circuit architecture at sublogarithmic depth leads to a poorly anticoncentrated output distribution. We consider the case of both noisy and noiseless circuits.

Theorem 3: *Consider a Haar-random circuit ensemble on any parallel architecture subject to Pauli noise with a uniform upper bound on the local noise rate* $q_{\mu mi} \le q_\mu$ *for all Pauli-noise sectors* μ, *m, and i. Also, let* $b = \min[q_x+q_y, q_y+q_z, q_z+q_x] < 1/2$ *and* $a' = \log 120 - 2\log(1-2b)$. *If the depth of the circuit ensemble is sublogarithmic (i.e., satisfies* $1 \le d = o(\log n)$*), then*

$$\lim_{n\to\infty} \Pr_{\mathcal{B}}\left[p_{00\ldots0} < \frac{1}{2^n e^{ne^{-a'd}/4}}\right] = 1. \tag{49}$$

Proof. The strategy of the proof is to show that a bound on the logarithm,

$$-\frac{1}{n}\log p_{00\ldots0}, \tag{50}$$

is sufficiently concentrated about its mean. This can be thought of as the first step in proving a behavior akin to the central limit theorem for the log-output probability similar to the behavior of the free energy in classical statistical mechanics.

To see this, we express the output probability $p_{00\ldots0}$ in terms of conditional probabilities:

$$\begin{aligned} &p(x_1 = 0,\ldots,x_n = 0) \\ &\quad = p(x_1 = 0)p(x_2 = 0|x_1 = 0)\ldots \\ &\qquad \times p(x_n = 0|x_1 = 0,\ldots,x_{n-1} = 0), \end{aligned} \tag{51}$$

which also holds when the x_i are reordered by any permutation. Taking the logarithm of both sides and summing over all permutation representations of the formula, we arrive at

the expression

$$-\log p_{00\ldots0} = -\frac{1}{n!}\sum_{\sigma\in S_n}\sum_i \log \times p(x_{\sigma(i)}=0|x_{\sigma(1)}=0,\ldots,x_{\sigma(i-1)}=0), \tag{52}$$

where S_n is the permutation group on $(1,\ldots,n)$. We rewrite $2p(x_i=0|\{x_j=0\}_{j\in J_i}) = 1+\langle Z_i\rangle_{J_i}$ to express the sum as

$$\begin{aligned} &-\log p_{00\ldots0} - n\log 2 \\ &= -\frac{1}{n!}\sum_{\sigma\in S_n}\sum_i \log\left(1+\langle Z_{\sigma(i)}\rangle_{\sigma(\{1,\ldots,i-1\})}\right) &(53)\\ &\geq -\frac{1}{n!}\sum_{\sigma\in S_n}\sum_i \langle Z_{\sigma(i)}\rangle_{\sigma(\{1,\ldots,i-1\})} \\ &\quad + \frac{1}{4n!}\sum_{\sigma\in S_n}\sum_i \langle Z_{\sigma(i)}\rangle^2_{\sigma(\{1,\ldots,i-1\})} &(54)\\ &= \frac{1}{n!}\sum_{\sigma\in S_n} A_\sigma + \frac{1}{4n}\sum_{i,j}\frac{1}{(n-1)!}\sum_{\sigma_{ji}}\langle Z_j\rangle^2_{\sigma_{ji}(\{1,\ldots,i-1\})}, &(55)\end{aligned}$$

where we define $A_\sigma := -\sum_i \langle Z_{\sigma(i)}\rangle_{\sigma(\{1,\ldots,i-1\})}$, use the uniform lower bound $-\log(1+x)\geq -x+x^2/4$ for $x\in[-1,1]$, and denote by σ_{ji} the subset of permutations with $\sigma(i)=j$. Now let $J_j^{(i)}$ run over all the subsets of $\{1,\ldots,n\}$ not containing j of length $i-1$. This set has size $\binom{n-1}{i-1}$. For each $J_j^{(i)}$, the term $\langle Z_j\rangle^2_{J_j^{(i)}}$ appears multiple times in the above sum, since there are multiple permutations τ satisfying $\tau(i)=j$ and $\{\tau(1),\tau(2)\ldots\tau(i-1)\} = \{\sigma(1),\sigma(2)\ldots\sigma(i-1)\}$. The number of such permutations is $(i-1)!\times(n-i)!$. Therefore, the above is equal to

$$\begin{aligned} &\frac{1}{n!}\sum_{\sigma\in S_n} A_\sigma + \sum_{i,j}\frac{1}{4n\binom{n-1}{i-1}}\sum_{J_j^{(i)}}\langle Z_j\rangle^2_{J_j^{(i)}} &(56)\\ &\geq \frac{1}{n!}\sum_{\sigma\in S_n} A_\sigma + \sum_j\sum_{i=1}^{n-L_j(d)+1}\frac{1}{4n}\frac{\binom{n-L_j(d)}{i-1}}{\binom{n-1}{i-1}}\langle Z_j\rangle^2, &(57)\end{aligned}$$

where we denote $L_i(d) = |L_d\circ L_d^\dagger(i)|$. In the above, we restrict the sum to only those subsets $J_j^{(i)}$ with $L_d\circ L_d^\dagger(j)\cap J_j^{(i)} = \emptyset$. This ensures conditional independence $p(x_j=0|\{x_k=0\}_{k\in J_j^{(i)}}) = p(x_j=0)$, which removes the conditional dependence on $\langle Z_j\rangle_{J_j^{(i)}}$. The number of subsets $J_j^{(i)}$ such that there is no qubit in $L_d\circ L_d^\dagger(j)\cap J_j^{(i)}$ is $\binom{n-L_j(d)}{i-1}$. The sum over i is truncated to $n-L_j(d)+1$ since for larger i, we get back terms with conditional dependence. Continuing, we use the hockey-stick relation $\sum_{i=1}^{n-L_j(d)+1}\binom{n-1}{n-L_j(d)-i+1} = \binom{n}{n-L_j(d)}$ to obtain

$$\begin{aligned} -\log p_{00\ldots0} &\geq n\log 2 + \frac{1}{n!}\sum_{\sigma\in S_n} A_\sigma + \sum_j \frac{1}{4L_j(d)}\langle Z_j\rangle^2 &(58)\\ &\geq n\log 2 + \frac{1}{n!}\sum_{\sigma\in S_n} A_\sigma \\ &\quad + \frac{1}{4L_{\max}(d)}\sum_j\langle Z_j\rangle^2 =: x, &(59)\end{aligned}$$

where $L_{\max}(d) = \max_i L_i(d)$.

The proof of Theorem 1 provides the lower bound $\mathbb{E}_{\mathcal{B}}\left[(p_0-1/2)^2\right]\geq e^{-ad}/4$ in Eq. (17). This bound is valid for all qubits i, not just the first qubit as in Theorem 1. This bound translates to the lower bound $\mathbb{E}_{\mathcal{B}}[\langle Z_i\rangle^2]\geq e^{-ad}$ because $\langle Z\rangle = p_0-(1-p_0) = 2p_0-1$. We also have $\mathbb{E}_{\mathcal{B}}[A_\sigma]=0$ because the conditional probability $\langle Z_{\sigma(i)}\rangle_{\sigma(\{1,\ldots,i-1\})}$ flips sign upon the action of the single-qubit unitary X_i at the output. By Haar invariance of the layer of single-qubit unitaries, $\mathbb{E}_{\mathcal{B}}\left[\langle Z_{\sigma(i)}\rangle_{\sigma(\{1,\ldots,i-1\})}\right]=0$. As a result, we have a lower bound

$$\mathbb{E}_{\mathcal{B}}[-\log p_{00\ldots0}]\geq n\log 2 + ne^{-ad}/(4L_{\max}(d)). \tag{60}$$

In Lemma 3, we prove that the variance of the lower bound on $-\log p_{00\ldots0}$, denoted by x in Eq. (59), is $\sigma^2\leq\sigma'^2 = 2n = o(n^2)$. This implies that, by Chebyshev's inequality,

$$\Pr_{\mathcal{B}}[|x-\mathbb{E}_{\mathcal{B}}[x]|>k\sigma'] \leq \Pr[|x-\mathbb{E}_{\mathcal{B}}[x]|>k\sigma] \leq \frac{1}{k^2} \tag{61}$$

$$\implies \Pr_{\mathcal{B}}[-x>-\mathbb{E}[x]+k\sigma'] \leq \frac{1}{k^2} \tag{62}$$

$$\implies \Pr_{\mathcal{B}}[\log p_{00\ldots0}>-\mathbb{E}_{\mathcal{B}}[x]+k\sqrt{2n}] \leq \frac{1}{k^2}. \tag{63}$$

Since $L_{\max}(d)\leq 4^d$, we have $-\mathbb{E}_{\mathcal{B}}[x]\leq -n\log 2 - ne^{-a'd}/4$ with $a' = a+\log 4$, which means that

$$\Pr_{\mathcal{B}}\left[p_{00\ldots0}<2^{-n}\exp\left[-\frac{ne^{-a'd}}{4}+k\sqrt{2n}\right]\right]\geq 1-\frac{1}{k^2}. \tag{64}$$

We choose $k=\Theta(n^{0.01})$ so that

$$\lim_{n\to\infty}\Pr_{\mathcal{B}}\left[p_{00\ldots0}<2^{-n}\exp\left(-\frac{ne^{-a'd}}{4}+O(n^{0.51})\right)\right]=1. \tag{65}$$

This means a lack of anticoncentration through $\lim_{n\to\infty} \Pr_{\mathcal{B}}[p_{00\ldots0} = o(2^{-n})] = 1$ whenever $ne^{-a'd} = \omega(n^{0.51})$, which is satisfied for any $d < 0.49 \log n/a'$, with $a' = \log 120 - 2\log(1-2b)$, where $b = \min[q_x + q_y, q_y + q_z, q_z + q_x]$. Thus we establish a lack of anticoncentration for any sublogarithmic depth d subject to the proof of Lemma 3. ■

Before stating Lemma 3, we state the following corollary for noiseless circuits.

Corollary 2: *An ensemble of noiseless Haar-random circuits on any parallel architecture at sublogarithmic depth $1 \le d \le o(\log n)$ is poorly anticoncentrated, satisfying*

$$\lim_{n\to\infty} \Pr_{\mathcal{B}}\left[p_{00\ldots0} < \frac{1}{2^n e^{n/(4\times 120^d)}}\right] = 1. \tag{66}$$

Proof. The proof for the noiseless case follows directly by taking $b = 0$ in Theorem 3. ■

We comment on the relation between this no-go result for anticoncentration at sublogarithmic depth and a related result of Dalzell *et al.* [28]. Dalzell *et al.* have adopted a different definition of anticoncentration in terms of the expected collision probability $\mathbb{E}_{\mathcal{B}}[Z] \le a/2^n$. As mentioned earlier and justified in the proof of Theorem 4, a bound on the collision probability implies anticoncentration as defined in Definition 1. In the original context of quantum computational-advantage proofs [24], Definition 1 of anticoncentration is more relevant. Moreover, there are simple cases where this definition of anticoncentration is satisfied but there is not a good bound on the collision probability. For example, consider a distribution where $1/2$ of the probability mass is concentrated on a single outcome and the remaining $1/2$ of the mass is equally distributed over all the other $2^n - 1$ outcomes. The distribution satisfies Definition 1 by taking $\alpha = 1/2$ and any $c < 1$. The collision probability for this distribution is $Z = 1/4 + O(1/2^n) \gg a/2^n$.

Barak *et al.* [32] and Dalzell *et al.* [28] have shown that random circuits in one-dimensional and other parallel-circuit architectures anticoncentrate (according to both Definition 1 and their definition in terms of the collision probability). Dalzell *et al.* have shown that logarithmic depth is also *necessary* for the collision probability bound $\mathbb{E}_{\mathcal{B}}[Z] \le a/2^n$ but this has left open the possibility that sublogarithmic-depth random circuits anticoncentrate without a corresponding bound on the collision probability. Our Theorem 3 disproves this possibility and thus strengthens the authors' no-go result at sublogarithmic depths.

Lemma 3: *The variance $\sigma^2 := \mathbb{E}_{\mathcal{B}}[x^2] - \mathbb{E}_{\mathcal{B}}[x]^2$ of the lower bound x in Eq. (59) satisfies $\sigma^2 \le 2n$.*

D. Anticoncentration at large enough depth

We now study anticoncentration properties of both noisy and noiseless circuits at depths logarithmic in n or larger. These results are established in terms of the collision probability defined in Eq. (5), which is the second moment of the output probability.

Theorem 4: *Haar-random circuits with Pauli noise anticoncentrate at least as fast as those without noise. More specifically, we show that the output probability distribution for a Haar-random circuit with Pauli noise is anticoncentrated if the noiseless output satisfies $2^n \mathbb{E}_{\mathcal{B}} Z(U, \mathbb{I}) \le c$ for some constant c.*

Proof. To prove the result, we first need Lemma 4, which shows that the collision probability for a noisy Clifford circuit increases upon removing the noise, i.e., $Z(U, \mathcal{E}) \le Z(U, \mathbb{I})$. The proof of the theorem then follows from the 2-design property of the Clifford group through the inequalities

$$2^{-n} \le \mathbb{E}_{\mathcal{B}} Z(U, \mathcal{E}) = \mathbb{E}_C Z(U, \mathcal{E}) \tag{67}$$

$$\le \mathbb{E}_C Z(U, \mathbb{I}) = \mathbb{E}_{\mathcal{B}} Z(U, \mathbb{I}). \tag{68}$$

The first inequality in Eq. (67) follows from a generic lower bound on $Z(U, \mathcal{E})$, while the second inequality in Eq. (68) follows from Lemma 4.

To prove that this bound on the collision probability implies anticoncentration, we can apply the Paley-Zygmund inequality for $0 < \alpha < 1$,

$$\Pr\left[p_{00\ldots0} \ge \frac{\alpha}{2^n}\right] = \Pr\left[p_{00\ldots0} \ge \alpha \mathbb{E}_{\mathcal{B}}[p_{00\ldots0}]\right] \tag{69}$$

$$\ge \frac{(1-\alpha)^2}{4^n \mathbb{E}_{\mathcal{B}}\left[p_{00\ldots0}^2\right]} \ge \frac{(1-\alpha)^2}{c}, \tag{70}$$

which converges to a number greater than 0 in the limit $n \to \infty$. Here, we use the fact that

$$4^n \mathbb{E}_{\mathcal{B}}\left[p_{00\ldots0}^2\right] = 2^n \mathbb{E}_{\mathcal{B}}\left[\sum_x p_x^2\right] = 2^n \mathbb{E}_{\mathcal{B}}[Z(U, \mathcal{E})], \tag{71}$$

and then apply Eq. (68). ■

Theorem 4 implies that noisy Haar-random circuits on architectures that are reasonably well connected anticoncentrate after $\Theta(\log n)$ depth, since $2^n \mathbb{E}_{\mathcal{B}} Z(U,\mathbb{I}) = 2 + o(1)$ after depth $\Theta(\log n)$ for such circuits [28,66]. Theorem 3 also rules out anticoncentration at any sublogarithmic depth for a Haar-random circuit with Pauli noise.

We now present the lemma for noisy Clifford circuits.

Lemma 4: *The collision probability decreases when adding Pauli noise to a noiseless Clifford circuit, i.e.,*

$$Z(U,\mathcal{E}) \leq Z(U,\mathbb{I}). \tag{72}$$

Proof. The defining property of Clifford circuits is that they send Pauli-group elements to Pauli-group elements. Another result we need is that the composition of two Pauli-noise channels is also a Pauli-noise channel. We can use these two facts to move all the noise channels past the Clifford gates to arrive at the expression

$$\mathcal{N}_d(|0\rangle\langle 0|) = (\mathcal{E}_U \circ \mathcal{C})(|00\ldots0\rangle\langle 00\ldots0|), \tag{73}$$

where $\mathcal{E}_U$ is a new Pauli-noise channel with a modified distribution $q(E)$ and $\mathcal{C} = \mathcal{C}_d \circ \cdots \circ \mathcal{C}_1$ is the noiseless Clifford circuit. Now we are left to bound the expression

$$\begin{aligned} Z(U,\mathcal{E}) &= \sum_x \left| \sum_{E\in\mathcal{P}} q(E)\langle x|E\left[\prod_{i=1}^n \frac{\mathbb{I}+g_i}{2}\right]E^\dagger|x\rangle \right|^2 \\ &= \sum_x \left| \sum_{\vec{s}} q_{\vec{s}}\langle x| \prod_{i=1}^n \frac{\mathbb{I}+(-1)^{s_i}g_i}{2}|x\rangle \right|^2, \end{aligned} \tag{74}$$

where $g_i = U_d \cdots U_1 Z_i U_1^\dagger \cdots U_d^\dagger$ are stabilizer generators for the evolved initial state under the noiseless circuit and Z_i is the Pauli Z operator on site i. In the second equality, we organize the sum into syndrome classes defined by the anticommutation pattern $\vec{s}$ of the Pauli-group elements E with the stabilizer generating set $\{g_i\}$ using the definition

$$q_{\vec{s}} := \sum_{E\in\mathcal{P}} \sum_{\text{s.t. } (-1)^{s_i} = [[E,g_i]]} q(E). \tag{75}$$

Here, $[[A,B]] = \mathrm{Tr}[ABA^{-1}B^{-1}]/2^n$ is the scalar commutator. To bound Eq. (74), we first use properties of stabilizer states to evaluate the measurement probabilities as

$$\langle x| \prod_{i=1}^n \frac{\mathbb{I}+(-1)^{s_i}g_i}{2}|x\rangle = p_{\max}(U,\mathbb{I}) f(x,\vec{s},U), \tag{76}$$

where $p_{\max}(U,\mathbb{I}) = \max_x p_x(U,\mathbb{I})$, $p_x(U,\mathbb{I}) := |\langle x|U|00\ldots0\rangle|^2$, and $f(x,\vec{s},U)$ is either 0 or 1 and satisfies the sum rule $\sum_x f(x,\vec{s},U) = p_{\max}^{-1}(U,\mathbb{I})$ for every $\vec{s}$. As a result,

$$Z(U,\mathcal{E}) = p_{\max}^2(U,\mathbb{I}) \sum_{\vec{s},\vec{\ell}} q_{\vec{s}} q_{\vec{\ell}} \sum_x f(x,\vec{s},U) f(x,\vec{\ell},U) \tag{77}$$

$$\leq p_{\max}^2(U,\mathbb{I}) \sum_{\vec{s},\vec{\ell}} q_{\vec{s}} q_{\vec{\ell}} \sum_x f(x,\vec{s},U) \tag{78}$$

$$= p_{\max}(U,\mathbb{I}) = Z(U,\mathbb{I}), \tag{79}$$

where we use the fact that $f(x,\vec{s},U) f(x,\vec{\ell},U) \leq f(x,\vec{s},U)$, $\sum_{\vec{s}} q_{\vec{s}} = 1$ and the property of Clifford circuits that all nonzero probabilities are equal. ■

ACKNOWLEDGMENTS

We thank Alex Dalzell for helpful and inspiring discussions. We thank Igor Boettcher, Dominik Hangleiter, and Grace Sommers for helpful comments on the manuscript. M.J.G., A.V.G., and P.N. acknowledge support from the National Science Foundation (NSF) Quantum Leap Challenge Institutes (QLCI) (Grant No. OMA-2120757). A.D., A.V.G., P.N., and O.S. were supported in part by the U.S. Department of Energy (DOE) Quantum Systems Accelerator (QSA), the DOE Advanced Scientific Computing Research (ASCR) Accelerated Research in Quantum Computing program (Award No. DE-SC0020312), the DOE ASCR Quantum Testbed Pathfinder program (Award No. DE-SC0019040), the NSF Practical Fully-Connected Quantum Computer (PFCQC) program, the Air Force Office of Scientific Research (AFOSR), the DOE (Award No. DE-SC0019449), the Army Research Office (ARO) Multidisciplinary University Research Initiative (MURI), the AFOSR MURI, and the Defense Advanced Research Projects Agency (DARPA) Science of Atomic Vapors for New Technologies (SAVaNT) ADVENT. A.D. also acknowledges support from the NSF under Award No. 1839204 and the National Science Foundation under Grant No. NSF PHY-1748958. B.F. acknowledges support from the AFOSR (Grants No. YIP FA9550-18-1-0148 and FA9550-21-1-0008). This material is based upon work partially support from the NSF under Award CAREER and by the DOE, Office of Science, National Quantum Information Science Research Centers. The Institute for Quantum Information and Matter is an NSF Physics Frontiers Center (PHY-1733907).

APPENDIX A: PROOF OF LEMMA 2

Here, we introduce the concepts that go into proving Lemma 2.

If X is a random variable in $\{0,1\}^n$ denoting the measurement outcome of n qubits after a unitary $U \sim \mathcal{B}$, the collision probability of the random circuit architecture is defined as

$$Z = \mathbb{E}_{\mathcal{B}}\left[\sum_{x\in\{0,1\}^n} \Pr(X=x)^2\right] = \mathbb{E}_{\mathcal{B}}\left[\sum_{x\in\{0,1\}^n} p_U(x)^2\right], \tag{A1}$$

where $p_U(x)$ is the probability that the measurement result is x. If there is at least one gate for each qubit in a parallel-circuit architecture with Haar-random gates, all measurement outcomes are equally random and, thus, there is a symmetry over them:

$$Z = \mathbb{E}_{\mathcal{B}}\left[\sum_{x\in\{0,1\}^n} p_U(x)^2\right] = 2^n \mathbb{E}_{\mathcal{B}}\left[p_U(0^n)^2\right], \tag{A2}$$

where 0^n denotes the state $|00\ldots0\rangle$. Assuming that the input state is also 0^n, the probability of measuring 0^n after the circuit is given by $\mathrm{Tr}(|0^n\rangle\langle 0^n|U|0^n\rangle\langle 0^n|U^\dagger)$. To obtain the second moment of the probability distribution, we consider two copies of the circuit acting on two copies of the input state. Since the trace obeys $\mathrm{Tr}(A\otimes B) = \mathrm{Tr}(A)\mathrm{Tr}(B)$, we obtain

$$\begin{aligned} Z &= 2^n \mathbb{E}_{\mathcal{B}}\left[p_U(0)^2\right] = 2^n \mathbb{E}_{\mathcal{B}} \mathrm{Tr}[\left(|0^n\rangle\langle 0^n|\right)^{\otimes 2} U^{\otimes 2} \left(|0^n\rangle\langle 0^n|\right)^{\otimes 2} (U^\dagger)^{\otimes 2}] \\ &= 2^n \mathrm{Tr}\left[\left(|0^n\rangle\langle 0^n|\right)^{\otimes 2} \mathbb{E}_{\mathcal{B}}\left[U^{\otimes 2} \left(|0^n\rangle\langle 0^n|\right)^{\otimes 2} (U^\dagger)^{\otimes 2}\right]\right]. \end{aligned} \tag{A3}$$

For convenience, we denote the two-copy Haar-averaged channel over k qubits as M_{U_k}:

$$M_{U_k}[\rho] = \mathbb{E}_{\mathcal{B}}\left[{U_k}^{\otimes 2} \rho ({U_k}^\dagger)^{\otimes 2}\right], \tag{A4}$$

where U_k is a unitary acting on k qubits. To study noisy evolution, we define a dephasing noise of strength q given by the noise channel

$$\mathcal{E}[\rho] = (1-q)\rho + qZ\rho Z. \tag{A5}$$

Lemma 2: *For a noisy Haar-random circuit of depth d on any parallel-circuit architecture with heralded-dephasing noise at rate p with the dephasing parameter q, we have the upper bound on the expected collision probability*

$$\mathbb{E}_{\mathcal{B}}[Z] = \mathbb{E}_{\mathcal{B}}\Big[\sum_x p_x^2\Big] \le 2^{-n} \exp\left[\frac{n}{3} e^{-\gamma p d}\right], \tag{A6}$$

where $\gamma = 8q(1-q)/3$.

Proof. It is convenient to separate out the average over random locations of the dephasing events in the heralded-dephasing model from the ensemble $\mathcal{B}$ (an ensemble of both gates and noise locations). We denote the ensemble of gates for a fixed set of locations L by $\mathcal{B}_L$ and averages over the noise locations by $\mathbb{E}_{\mathcal{L}}$. We show in Lemma 6 that given a random circuit ensemble $\mathcal{B}_L$ on a parallel-circuit architecture with heralded-dephasing noise and a fixed set of locations L, there exists another circuit ensemble $\mathcal{B}'_L$, with the gates drawn at random independently of L, composed solely of single-qubit gates and SWAP gates with an average collision probability $\mathbb{E}_{\mathcal{B}'_L}[Z]$ greater than or equal to the average collision probability of the original circuit $\mathbb{E}_{\mathcal{B}_L}[Z]$. Note that for every circuit in the new ensemble $C \in \mathcal{B}'_L$, we can append a network of SWAP gates to C to return all qubits to their original positions. Adding these SWAP gates does not change the collision probability, since these gates only permute the support of the final probability distribution. We further break up

the ensemble $\mathcal{B}'_L$ into an ensemble of single-qubit Haar-random gates $\mathcal{B}'_1$ and the random SWAP network $\mathcal{B}'_{\text{swap}}$. The distribution of gates in the ensemble $\mathcal{B}'_1 \cup \mathcal{B}'_{\text{swap}}$ is defined by taking every two-qubit gate for the circuits in $\mathcal{B}_L$ and replacing it with a SWAP gate or identity gate with equal probability on those sites, followed by Haar-random single-site gates on the two qubits. The joint distribution over SWAP networks and single-site gates is also conditionally independent, allowing us to commute averages over single-site gates, SWAP networks, and noise locations with each other.

Fixing a realization of SWAP gates and noise locations, we can then follow the path of a qubit, count the total number of dephasing events in that path, and merge consecutive single-qubit gates without intervening noise locations. Since we are working with a parallel-circuit architecture, there is never a case of consecutive dephasing events (there is always a single-qubit gate after each dephasing event). Let t_i be the number of dephasing events on the path of qubit i. For the heralded-dephasing model, we can write the random variable t_i as a sum

$$t_i = \sum_{j=1}^{d} x_{ij}, \tag{A7}$$

where the $x_{ij} \in \{0, 1\}$ are independent identically distributed Bernoulli random variables for each i and j with parameter p. We have $\mathbb{E}_{\mathcal{L}}[x_{ij}] = p$ and $\Pr_{\mathcal{L}}[x_{ij} = a, x_{kl} = b] = \Pr_{\mathcal{L}}[x_{ij} = a] \Pr_{\mathcal{L}}[x_{kl} = b]$ for every $ij \neq kl$.

After averaging over $\mathcal{B}'_1$ using Lemma 5, the final two-copy circuit-averaged state is given by

$$\bigotimes_{i=1}^{n} \left(M_{U_1} \circ \underbrace{(\mathcal{E} \circ M_{U_1}) \cdots \circ (\mathcal{E} \circ M_{U_1})}_{t_i} \right) [|0^n\rangle\langle 0^n|^{\otimes 2}] = \bigotimes_{i=1}^{n} \left[\left(\frac{1}{12}(3 - \beta^{t_i}) \right) I + \frac{1}{6}\beta^{t_i} S \right], \tag{A8}$$

where $\beta = 1 - 8q(1-q)/3$, and I and S are the 4×4 identity and SWAP operators, respectively. Using Eq. (A3) and noting that $\text{Tr}(I\,|0\rangle\,\langle 0|^{\otimes 2}) = \text{Tr}(S\,|0\rangle\,\langle 0|^{\otimes 2}) = 1$, the average collision probability for a fixed SWAP network and set of noise locations equals

$$\mathbb{E}_{\mathcal{B}'_1}[Z] = 2^n \prod_{i=1}^{n} \left[\frac{1}{12}(3 - \beta^{t_i}) + \frac{1}{6}\beta^{t_i} \right] = 2^n \prod_{i=1}^{n} \frac{1}{2^2} \left[1 + \frac{1}{3}\beta^{t_i} \right] = \frac{1}{2^n} \prod_{i=1}^{n} \left[1 + \frac{1}{3}\beta^{t_i} \right]. \tag{A9}$$

We now average over the noise locations using our assumption that the noise locations are uncorrelated with each other and the realization of the SWAP network:

$$\begin{aligned} \mathbb{E}_{\mathcal{L}} \mathbb{E}_{\mathcal{B}'_1}[Z] &= \mathbb{E}_{\substack{t_i \\ 1 \leq i \leq n}} \left[\mathbb{E}_{\mathcal{B}'_1}[Z] \right] \\ &= \frac{1}{2^n} \prod_{i=1}^{n} \left[1 + \mathbb{E}_{t_i} \left[\frac{1}{3}\beta^{t_i} \right] \right] \\ &= \frac{1}{2^n} \prod_{i=1}^{n} \left[1 + \frac{1}{3} \prod_{j=1}^{d} \mathbb{E}_{x_{ij}} [\beta^{x_{ij}}] \right]. \end{aligned} \tag{A10}$$

Using the fact that x_{ij} are Bernoulli random variables, we can compute the expectation $\mathbb{E}_{x_{ij}}[\beta^{x_{ij}}]$ as

$$\mathbb{E}_{x_{ij}}[\beta^{x_{ij}}] = p\beta^1 + (1-p)\beta^0 = p\beta + (1-p) = 1 - p(1-\beta) = 1 - p\gamma, \tag{A11}$$

where we define $\gamma = 1 - \beta$. Inserting this expectation in Eq. (A10) and using Lemma 6, we obtain a bound on the average collision probability:

$$\mathbb{E}_{\mathcal{B}}[Z] \leq \mathbb{E}_{\mathcal{B}'}[Z] = \mathbb{E}_{\mathcal{B}'_{\text{swap}}} \mathbb{E}_{\mathcal{L}} \mathbb{E}_{\mathcal{B}_{1'}}[Z] = \frac{1}{2^n} \left[1 + \frac{1}{3}(1 - p\gamma)^d \right]^n \leq \frac{1}{2^n} \left[1 + \frac{1}{3} e^{-\gamma p d} \right]^n \leq 2^{-n} \exp\left[\frac{n}{3} e^{-\gamma p d} \right]. \tag{A12}$$

Here, we use $\mathbb{E}_{\mathcal{B}}[Z] = \mathbb{E}_{\mathcal{L}} \mathbb{E}_{\mathcal{B}_L}[Z] \leq \mathbb{E}_{\mathcal{B}'}[Z] := \mathbb{E}_{\mathcal{L}} \mathbb{E}_{\mathcal{B}'_L}[Z]$ in applying Lemma 6 in the first inequality. We also use the fact that for any $x > 0$, $1 - x \leq e^{-x}$ in the second inequality and $1 + x \leq e^x$ in the third inequality. ■

Lemma 5: *Consider a random circuit consisting of k dephasing error channels of strength q sandwiched between $k+1$ single-qubit Haar-random gates (denoted by U_1). When this circuit acts on two copies of a single-qubit, the circuit-averaged state is given by*

$$M_{U_1} \circ \underbrace{(\mathcal{E} \circ M_{U_1}) \cdots \circ (\mathcal{E} \circ M_{U_1})}_{k}[|0\rangle\langle 0|^{\otimes 2}] = \frac{1}{12}(3-\beta^k)I + \frac{1}{6}\beta^k S, \tag{A13}$$

where $\beta = 1 - \frac{8}{3}q(1-q)$, and I and S are the 4×4 identity and SWAP *operators, respectively.*

Proof. We first make an observation that $M_{U_1}[\rho] = M_{U_1} \circ M_{U_1}[\rho]$, that is, one can split a Haar-random gate into two Haar-random gates without changing the statistics. In the circuit described above, leaving the two terminal unitary gates intact, we split all the inner gates into two. This lets us treat the circuit as a repeating sequence of n units of the composite channel $\tilde{M}_{U_1,\mathcal{E}} = M_{U_1} \circ \mathcal{E} \circ M_{U_1}$.

From Ref. [28], we know the following:

$$M_{U_1}[\sigma] = \frac{1}{3}\left(\mathrm{Tr}(\sigma) - 2^{-1}\mathrm{Tr}(\sigma S)\right) I + \frac{1}{3}\left(\mathrm{Tr}(\sigma S) - 2^{-1}\mathrm{Tr}(\sigma)\right) S. \tag{A14}$$

If we follow this gate by a dephasing error channel, we obtain

$$\begin{aligned}
\mathcal{E} \circ M_{U_1}[\sigma] &= \frac{1}{3}\left(\mathrm{Tr}(\sigma) - 2^{-1}\mathrm{Tr}(\sigma S)\right)\mathcal{E}[I] + \frac{1}{3}\left(\mathrm{Tr}(\sigma S) - 2^{-1}\mathrm{Tr}(\sigma)\right)\mathcal{E}[S] \\
&= \frac{1}{3}\left(\mathrm{Tr}(\sigma) - 2^{-1}\mathrm{Tr}(\sigma S)\right) I + \frac{1}{3}\left(\mathrm{Tr}(\sigma S) - 2^{-1}\mathrm{Tr}(\sigma)\right)\left((1-q)^2 S + 2q(1-q)(ZI)S(ZI) + q^2(ZZ)S(ZZ)\right) \\
&= \frac{1}{3}\left(\mathrm{Tr}(\sigma) - 2^{-1}\mathrm{Tr}(\sigma S)\right) I + \frac{1}{3}\left(\mathrm{Tr}(\sigma S) - 2^{-1}\mathrm{Tr}(\sigma)\right)\left(((1-q)^2 + q^2)S + 2q(1-q)(ZI)S(ZI)\right).
\end{aligned}$$

We follow this channel by another single-qubit random gate to finish the composite block. First, we observe that $M_{U_1}[I] = I$ and $M_{U_1}[S] = S$. Similarly, using Eq. (A14) together with the fact that $\mathrm{Tr}[S] = 2, \mathrm{Tr}[(IZ)S(IZ)S] = 0$,

$$M_{U_1}[(IZ)S(IZ)] = \frac{2}{3}I - \frac{1}{3}S.$$

The composite channel thus gives

$$\begin{aligned}
\tilde{M}_{U_1,\mathcal{E}}[\sigma] &= \frac{1}{3}\left(\mathrm{Tr}(\sigma) - 2^{-1}\mathrm{Tr}(\sigma S)\right) M_{U_1}[I] + \frac{1}{3}\left(\mathrm{Tr}(\sigma S) - 2^{-1}\mathrm{Tr}(\sigma)\right) \\
&\quad \times \left(((1-q)^2 + q^2)M_{U_1}[S] + 2q(1-q)M_{U_1}[(ZI)S(ZI)]\right) \\
&= \frac{1}{3}\left(\mathrm{Tr}(\sigma) - 2^{-1}\mathrm{Tr}(\sigma S)\right) I + \frac{1}{3}\left(\mathrm{Tr}(\sigma S) - 2^{-1}\mathrm{Tr}(\sigma)\right)\left(((1-q)^2 + q^2)S + 2q(1-q)\left(\frac{2}{3}I - \frac{1}{3}S\right)\right) \\
&= \frac{1}{3}\left(\mathrm{Tr}(\sigma) - 2^{-1}\mathrm{Tr}(\sigma S) + \underbrace{\frac{4}{3}q(1-q)}_{\alpha}\left(\mathrm{Tr}(\sigma S) - 2^{-1}\mathrm{Tr}(\sigma)\right)\right) I + \frac{1}{3}\left(\mathrm{Tr}(\sigma S) - 2^{-1}\mathrm{Tr}(\sigma)\right)\underbrace{\left(1 - \frac{8}{3}q(1-q)\right)}_{\beta} S \\
&= \frac{1}{3}\left(\mathrm{Tr}(\sigma) - 2^{-1}\mathrm{Tr}(\sigma S) + \alpha\left(\mathrm{Tr}(\sigma S) - 2^{-1}\mathrm{Tr}(\sigma)\right)\right) I + \frac{1}{3}\left(\mathrm{Tr}(\sigma S) - 2^{-1}\mathrm{Tr}(\sigma)\right)\beta S.
\end{aligned} \tag{A15}$$

The composite sum returns a state of the form $aI + bS$. Acting on this state with other composite blocks only changes the coefficients a and b. In fact, we can work out exactly how a and b change after each block. Knowing that $\mathrm{Tr}(aI + bS) =$

$4a + 2b$ and $\mathrm{Tr}(S(aI + bS)) = 2a + 4b$, we obtain

$$\tilde{M}_{U_1,\mathcal{E}}[aI + bS] = \frac{1}{3}\left(3a + \alpha(3b)\right) I + \frac{1}{3}(3b)\beta S = (a + \alpha b)I + (\beta b)S. \tag{A16}$$

The first composite acts on the state $|0\rangle\langle 0|^{\otimes 2}$. Knowing that $\mathrm{Tr}[|0\rangle\langle 0|^{\otimes 2}] = \mathrm{Tr}(S\,|0\rangle\langle 0|^{\otimes 2}) = 1$ and using Eq. (A15),

$$\tilde{M}_{U_1,\mathcal{E}}[|0\rangle\langle 0|^{\otimes 2}] = \frac{1}{3}\left(1 - 2^{-1} + \alpha\left(1 - 2^{-1}\right)\right) I + \frac{1}{3}\left(1 - 2^{-1}\right)\beta S = \frac{1}{6}\left(1 + \alpha\right) I + \frac{1}{6}\beta S. \tag{A17}$$

We take this state and apply another $k - 1$ composite channels (since there are k in total). We can calculate the final state recursively using Eq. (A16):

$$\tilde{M}_{U_1,\mathcal{E}} \circ \cdots \circ \tilde{M}_{U_1,\mathcal{E}}[|0\rangle\langle 0|^{\otimes 2}] = \frac{1}{6}\left[1 + \alpha + \alpha\sum_{i=1}^{k-1}\beta^i\right] I + \frac{1}{6}\beta^k S = \frac{1}{12}(3 - \beta^k)I + \frac{1}{6}\beta^k S. \tag{A18}$$

This proves the lemma. ■

It now remains to prove the following lemma, which lets us put a bound on the average collision probability in Lemma 2.

Lemma 6: *Consider a random quantum circuit ensemble on a parallel architecture, $\mathcal{B}_L$, with Haar-random two-qubit gates and heralded-dephasing noise with a fixed set of noise locations L. There is a procedure to obtain another circuit ensemble $\mathcal{B}'_L$ with gates drawn randomly independently of L, composed solely of noisy single-qubit channels and* SWAP *gates, with an equal or higher average collision probability, i.e., $\mathbb{E}_{\mathcal{B}_L}[Z] \leq \mathbb{E}_{\mathcal{B}'_L}[Z]$.*

Proof. Note that without loss of generality, we can assume that every two-qubit Haar-random gate is preceded by Haar-random single-qubit gates on both incoming lines, since the two situations correspond to the same ensemble. When the noise is heralded, the circuit consists of three kinds of two-qubit gates:

(1) Type A, where the two-qubit gate is noiseless
(2) Type B, where one of the two outgoing legs of the gate undergoes dephasing, followed by a single-qubit random gate
(3) Type C, where both outgoing legs undergo dephasing followed by single-qubit gates on both legs

We analyze each of these types separately and start with a brief review of the methods of Ref. [28] for noiseless random circuits. In Ref. [28], it has been shown that taking a two copies of an n-qubit state $|0^n\rangle\langle 0^n|^{\otimes 2}$, acted on by a Haar-random circuit $U \otimes U$, and averaged over the unitaries, leads to a density matrix that can be represented by a linear combination of length-n configurations in $\{I, S\}^n$ where I and S are the 4×4 identity and SWAP matrices. Any two-qubit unitary gate takes a linear combination to another linear combination. More precisely, for $\vec{\gamma} \in \{I, S\}^n$, a two-qubit unitary channel M_{U_2}, acting on qubits i and j, transforms it to

$$M_{U_2}[\vec{\gamma}] = M_{U_2}\left[\bigotimes_{a=1}^{n}\gamma_a\right] = \sum_{\vec{\nu}\in\{I,S\}^n} M_{U_2}^{\vec{\gamma},\vec{\nu}}\bigotimes_{b=1}^{n}\nu_b =: \sum_{\vec{\nu}\in\{I,S\}} M_{U_2}^{\vec{\gamma},\vec{\nu}}\,\vec{\nu}, \tag{A19}$$

where $M_{U_2}^{\vec{\gamma},\vec{\nu}}$ are matrix elements determined by qubit locations i, j:

$$M_{U_2}^{\vec{\gamma},\vec{\nu}} = \begin{cases} 1, & \text{if } \gamma_i = \gamma_j \text{ and } \vec{\gamma} = \vec{\nu}, \\ 2/5, & \text{if } \gamma_i \neq \gamma_j \text{ and } \nu_i = \nu_j \text{ and } \gamma_k = \nu_k\ \forall k \in [n]/[i,j], \\ 0, & \text{otherwise.} \end{cases} \tag{A20}$$

Therefore, a state can be represented as a linear combination of trajectories of the configuration strings, with each trajectory weighted according to Eq. (A20). Furthermore, since $\mathrm{Tr}(\vec{\gamma}\,|0^n\rangle\langle 0^n|^{\otimes 2}) = 1$ for each $\vec{\gamma} \in \{I, S\}^n$, the collision probability

can be, similarly, written as a sum over weighted trajectories. More precisely, the average collision probability of a circuit with s gates,

$$\mathbb{E}_{\mathcal{B}_L}[Z_s] = \frac{1}{3^n} \sum_{\gamma \in \{I,S\}^{n\times s}} \prod_{t=1}^{s-1} M_{U_2}^{\vec{\gamma}_t,\vec{\gamma}_{t+1}} = \frac{1}{3^n} \sum_{\gamma \in \{I,S\}^{n\times s}} \prod_{t=1}^{s-1} \mathrm{wt}(\gamma). \tag{A21}$$

In the above, the factor $1/3^n$ comes from the fact that after the first layer of Haar-random single-qubit gates, the Haar-averaged two-copy state is given by $(1/6^n)\sum_{\vec{\gamma}\in\{I,S\}^n}\vec{\gamma}$, the uniform mixture of all configurations in $\{I,S\}^n$. Also, the weight $\mathrm{wt}(\gamma)$ of a configuration is defined as the product of the matrix elements $M_{U_2}^{\vec{\gamma}_t,\vec{\gamma}_{t+1}}$.

Now, we modify this construction to account for noise. We add one more gate of type A, B, or C to this circuit. Since all three types are two-qubit gates, we let $[i,j]$ denote the qubits on which the gate acts. We can isolate the qubits $[i,j]$ from the decomposition in Eq. (A21) as follows:

$$\mathbb{E}_{\mathcal{B}_L}[Z_s] = \frac{1}{3^n}\left[\sum_{\substack{\gamma\in\{I,S\}^{ns}\\ \vec{\gamma}_{ij}^{\,s}=II}} \mathrm{wt}(\gamma) + \sum_{\substack{\gamma\in\{I,S\}^{ns}\\ \vec{\gamma}_{ij}^{\,s}=IS}} \mathrm{wt}(\gamma) + \sum_{\substack{\gamma\in\{I,S\}^{ns)}\\ \vec{\gamma}_{ij}^{\,s}=SI}} \mathrm{wt}(\gamma) + \sum_{\substack{\gamma\in\{I,S\}^{ns}\\ \vec{\gamma}_{ij}^{\,s}=SS}} \mathrm{wt}(\gamma)\right]. \tag{A22}$$

a. Type A When we add a noiseless two-qubit gate, the bit strings transform according to Eq. (A20). Focusing on qubits i and j, the trajectories evolve as follows:

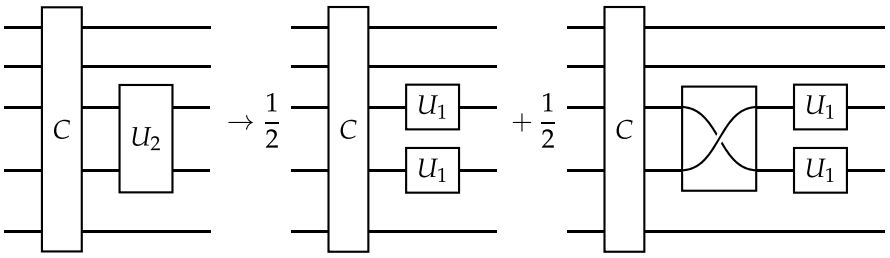
$$M_{U_2}[II] = II, \qquad M_{U_2}[SS] = SS, \qquad M_{U_2}[IS,SI] = \frac{2}{5}(II+SS). \tag{A23}$$

The trajectories for which $\vec{\gamma}_{ij}^{\,s} \in \{II, SS\}$ have their weights unchanged. The trajectories for which $\vec{\gamma}_{ij}^{\,s} \in \{IS, SI\}$ have their weights changed by $4/5$ (the trajectory splits two ways, each weighted by $2/5$):

$$\mathbb{E}_{\mathcal{B}_L}[Z_{s+1}] = \frac{1}{3^n}\left[\sum_{\substack{\gamma\in\{I,S\}^{n(s+1)}\\ \vec{\gamma}_{ij}^{\,s}=II}} \mathrm{wt}(\gamma) + \frac{4}{5}\sum_{\substack{\gamma\in\{I,S\}^{n(s+1)}\\ \vec{\gamma}_{ij}^{\,s}=IS}} \mathrm{wt}(\gamma) + \frac{4}{5}\sum_{\substack{\gamma\in\{I,S\}^{n(s+1)}\\ \vec{\gamma}_{ij}^{\,s}=SI}} \mathrm{wt}(\gamma) + \sum_{\substack{\gamma\in\{I,S\}^{n(s+1)}\\ \vec{\gamma}_{ij}^{\,s}=SS}} \mathrm{wt}(\gamma)\right]. \tag{A24}$$

If, instead, we consider a modified random circuit where the two-qubit gate consists of a SWAP gate or identity with probability $1/2$ followed by Haar-random single-qubit gates, all the trajectories retain their original weights, since the collision probability is invariant under a SWAP gate and $M_{U_1}[I] = I$ and $M_{U_1}[S] = S$. Denoting the locally modified circuit ensemble with the same set of noise locations by $\mathcal{B}'_L$, we have, in both cases,

$$\mathbb{E}_{\mathcal{B}'_L}[Z_{s+1}] = \mathbb{E}_{\mathcal{B}_L}[Z_s] \implies \mathbb{E}_{\mathcal{B}'_L}[Z_{s+1}] > \mathbb{E}_{\mathcal{B}_L}[Z_{s+1}]. \tag{A25}$$



(A26)

b. Type B The gate of type B has a noiseless two-qubit gate followed by dephasing on one of the outgoing legs. The dephasing is also followed by a single-qubit random gate. To simplify things, we first understand the effect of the channel

$M_{U_1} \otimes \mathcal{E}$ on I and S. Of course, $M_{U_1} \circ \mathcal{E}[I] = I$, since neither the error nor the random gate has any effect on the identity matrix. However, for S, we obtain

$$M_{U_1}^{(i)} \circ \mathcal{E}^{(i)}[S] = M_{U_1}[(1-q)^2 S + q(1-q)(IZ)S(IZ) + q(1-q)(ZI)S(ZI) + q^2 S] \tag{A27}$$

$$= \alpha I + \beta S, \tag{A28}$$

with $\alpha = 4q(1-q)/3, \beta = 1 - 8q(1-q)/3$. Without loss of generality, we assume that the dephasing happens on gate i, and the dephasing channel is denoted by $\mathcal{E}$. We now tabulate the effect of this composite channel:

$$M_{U_1}^{(i)} \circ \mathcal{E}^{(i)} \circ M_{U_2}[II] = II, \tag{A29}$$

$$M_{U_1}^{(i)} \circ \mathcal{E}^{(i)} \circ M_{U_2}[SS] = \alpha IS + \beta SS, \tag{A30}$$

$$M_{U_1}^{(i)} \circ \mathcal{E}^{(i)} \circ M_{U_2}[IS, SI] = M_{U_1}^{(i)}\left[\frac{2}{5}(II + SS)\right] = \frac{2}{5}(II + \alpha IS + \beta SS). \tag{A31}$$

The average collision probability of the new circuit is given by

$$\mathbb{E}_{\mathcal{B}_L}[Z_{s+1}] = \frac{1}{3^n}\left[\sum_{\vec{\gamma}_{ij}^s = II} \mathrm{wt}(\gamma) + \frac{2}{5}(1+\alpha+\beta)\left(\sum_{\vec{\gamma}_{ij}^s = IS} \mathrm{wt}(\gamma) + \sum_{\vec{\gamma}_{ij}^s = SI} \mathrm{wt}(\gamma)\right) + (\alpha+\beta)\sum_{\vec{\gamma}_{ij}^s = SS} \mathrm{wt}(\gamma)\right]. \tag{A32}$$

Using the same locally modified circuit ensemble as above, we obtain

$$M_{U_1}^{(i)} \circ \mathcal{E}^{(i)} \circ M_{U_2} \to \frac{1}{2} M_{U_1}^{(i)} \circ \mathcal{E}^{(i)} \circ M_{U_1}^{(i)} \circ M_{U_1}^{(j)} + \frac{1}{2} M_{U_1}^{(i)} \circ \mathcal{E}^{(i)} \circ M_{U_1}^{(i)} \circ M_{U_1}^{(j)} \circ \mathrm{swap} \tag{A33}$$

$$= \frac{1}{2} M_{U_1}^{(i)} \circ \mathcal{E}^{(i)} \circ M_{U_1}^{(i)} \circ M_{U_1}^{(j)} + \frac{1}{2} M_{U_1}^{(j)} \circ \mathcal{E}^{(j)} \circ M_{U_1}^{(j)} \circ M_{U_1}^{(i)}. \tag{A34}$$

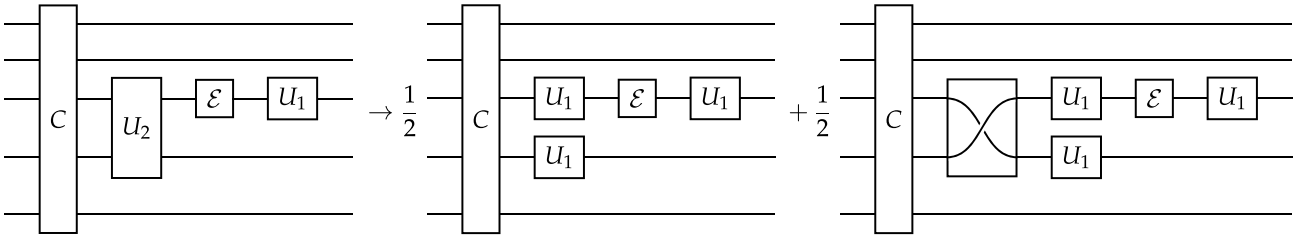
Under this new composite channel, the bit strings evolve as follows:

$$II \to II, \quad SS \to \beta SS + \frac{\alpha}{2}(IS + SI), \quad IS \to \frac{1}{2}(IS + \alpha II + \beta IS), \quad SI \to \frac{1}{2}(\alpha II + \beta SI + SI). \tag{A35}$$

The collision probability of the modified circuit is given by

$$\mathbb{E}_{\mathcal{B}'_L}[Z_s] = \frac{1}{3^n}\left[\sum_{\vec{\gamma}_{ij}^s = II} \mathrm{wt}(\gamma) + \frac{1}{2}(1+\alpha+\beta)\left(\sum_{\vec{\gamma}_{ij}^s = IS} \mathrm{wt}(\gamma) + \sum_{\vec{\gamma}_{ij}^s = SI} \mathrm{wt}(\gamma)\right) + (\alpha+\beta)\sum_{\vec{\gamma}_{ij}^s = SS} \mathrm{wt}(\gamma)\right]. \tag{A36}$$

Since $2/5 < 1/2$, we have $Z'_{s+1} > Z_{s+1}$.



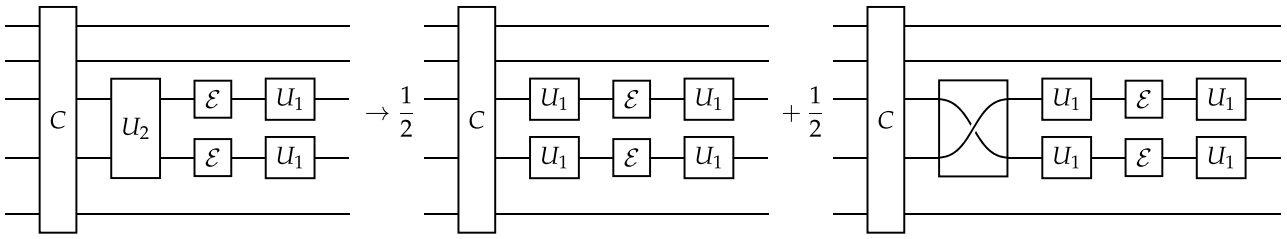
(A37)

c. Type C These gates have dephasing noise on both legs. The noise is followed by single-qubit gates for both legs. The combined channel has the form $M_{U_1}^{(j)} \circ \mathcal{E}^{(j)} \circ M_{U_1}^{(i)} \circ \mathcal{E}^{(i)} \circ M_{U_2}$. The bit strings evolve as follows:

$$II \to II, \quad SS \to \alpha^2 II + \alpha\beta IS + \alpha\beta SI + \beta^2 SS, \quad IS \to \frac{2}{5}((1+\alpha^2)II + \alpha\beta IS + \alpha\beta SI + \beta^2 SS). \tag{A38}$$

The average collision probability is thus given by

$$\mathbb{E}_{\mathcal{B}_L}[Z_{s+1}] = \frac{1}{3^n}\left[\sum_{\vec{\gamma}_{ij}^s=II}\mathrm{wt}(\gamma) + \frac{2}{5}\left(1+(\alpha+\beta)^2\right)\left(\sum_{\vec{\gamma}_{ij}^s=IS}\mathrm{wt}(\gamma)+\sum_{\vec{\gamma}_{ij}^s=SI}\mathrm{wt}(\gamma)\right) + (\alpha+\beta)^2\sum_{\vec{\gamma}_{ij}^s=SS}\mathrm{wt}(\gamma)\right]. \quad \text{(A39)}$$



(A40)

If, instead, we replace the two-qubit gate with two single-qubit Haar-random gates preceded by a SWAP gate with probability 1/2 (as shown in the diagram above), we obtain channels of the form $M_{U_1}^{(j)}\circ\mathcal{E}^{(j)}\circ\mathcal{E}^{(i)}\circ M_{U_1}^{(j)}\circ\mathcal{E}^{(i)}\circ M_{U_1}^{(i)}\circ\frac{1}{2}(\mathrm{swap}+\mathbb{I})$, which, up to the SWAP gate, is same as the composite channel in Lemma 5 applied to both qubits. The states evolve as

$$\begin{aligned} &II\to II, \quad SS\to\alpha^2 II+\alpha\beta IS+\alpha\beta SI+\beta^2 SS,\\ &IS\to\alpha II+\frac{\beta}{2}(IS+SI), \quad SI\to\alpha II+\frac{\beta}{2}(SI+IS). \end{aligned} \quad \text{(A41)}$$

The collision probability of the modified circuit is thus

$$\mathbb{E}_{\mathcal{B}'_L}[Z_{s+1}] = \frac{1}{3^n}\left[\sum_{\vec{\gamma}_{ij}^s=II}\mathrm{wt}(\gamma) + (\alpha+\beta)\left(\sum_{\vec{\gamma}_{ij}^s=IS}\mathrm{wt}(\gamma)+\sum_{\vec{\gamma}_{ij}^s=SI}\mathrm{wt}(\gamma)\right) + (\alpha+\beta)^2\sum_{\vec{\gamma}_{ij}^s=SS}\mathrm{wt}(\gamma)\right]. \quad \text{(A42)}$$

Since $(\alpha+\beta)=1+4q(1-q)/3$, $\alpha+\beta\in[1,4/3]\subset(\frac{1}{2},2)$, we have that $(2/5)\left(1+(\alpha+\beta)^2\right)<(\alpha+\beta)$ and therefore $Z'_{s+1}>Z_s$.

Starting from the input state, we can use the replacement procedure discussed above to iteratively define a new circuit ensemble $\mathcal{B}'_L$ with the gates drawn at random independently of L, composed solely of single-qubit gates and SWAP gates, that has an equal or higher gate-averaged collision probability. This concludes the proof of the lemma. ■

APPENDIX B: PROOF OF LEMMA 3

We restate the lemma here for reference.

Lemma 3: *The variance* $\sigma^2 := \mathbb{E}_{\mathcal{B}}[x^2]-\mathbb{E}_{\mathcal{B}}[x]^2$ *of the random variable* $x := n\log 2+(1/n!)\sum_{\sigma\in S_n}A_\sigma+\frac{1}{4L_{\max}(d)}\sum_j\langle Z_j\rangle^2$ *in Eq. (59) satisfies* $\sigma^2\le 2n$.

Proof of Lemma 3. Recalling that $\mathbb{E}_{\mathcal{B}}[A_\sigma]=0$, the variance of x is

$$\begin{aligned}\mathbb{E}_{\mathcal{B}}[x^2]-\mathbb{E}_{\mathcal{B}}[x]^2 = &\frac{1}{n!^2}\mathbb{E}_{\mathcal{B}}\left[\left(\sum_\sigma A_\sigma\right)^2\right]+\frac{1}{2L_{\max}(d)n!}\mathbb{E}_{\mathcal{B}}\left[\sum_\sigma A_\sigma\sum_j\langle Z_j\rangle^2\right]+\frac{1}{16L_{\max}(d)^2}\mathbb{E}_{\mathcal{B}}\left[\sum_{i,j}\langle Z_i\rangle^2\langle Z_j\rangle^2\right]\\ &-\frac{1}{16L_{\max}(d)^2}\mathbb{E}_{\mathcal{B}}\left[\sum_i\langle Z_i\rangle^2\right]\mathbb{E}_{\mathcal{B}}\left[\sum_j\langle Z_j\rangle^2\right]. \end{aligned} \quad \text{(B1)}$$

We use the covariance bound

$$\mathbb{E}[XY] = \sum_i \sqrt{p_i} x_i \sqrt{p_i} y_i \tag{B2}$$

$$\leq \sqrt{\left(\sum_i p_i x_i^2\right)\left(\sum_j p_j y_j^2\right)} \tag{B3}$$

$$= \sqrt{\mathbb{E}\left[X^2\right]\mathbb{E}\left[Y^2\right]}. \tag{B4}$$

Applying this to the first term gives

$$\frac{1}{n!^2}\mathbb{E}_{\mathcal{B}}\left[\left(\sum_\sigma A_\sigma\right)^2\right] = \frac{1}{n!^2}\sum_{\sigma,\tau}\mathbb{E}_{\mathcal{B}}\left[A_\sigma A_\tau\right] \tag{B5}$$

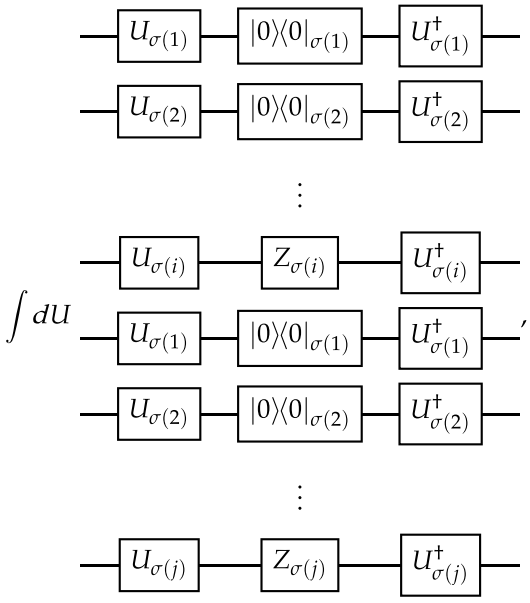
$$\leq \frac{1}{n!^2}\sum_{\sigma,\tau}\sqrt{\mathbb{E}_{\mathcal{B}}[A_\sigma^2]\mathbb{E}_{\mathcal{B}}[A_\tau^2]}. \tag{B6}$$

Now consider

$$\mathbb{E}_{\mathcal{B}}\left[A_\sigma^2\right] = \mathbb{E}_{\mathcal{B}}\left[\sum_{i,j}\langle Z_{\sigma(i)}\rangle_{\sigma(1)\ldots\sigma(i-1)}\langle Z_{\sigma(j)}\rangle_{\sigma(1)\ldots\sigma(j-1)}\right] \tag{B7}$$

$$= \mathbb{E}_{\mathcal{B}}\left[\sum_i \langle Z_{\sigma(i)}\rangle^2_{\sigma(1)\ldots\sigma(i-1)}\right] + 2\mathbb{E}_{\mathcal{B}}\left[\sum_{i<j}\langle Z_{\sigma(i)}\rangle_{\sigma(1)\ldots\sigma(i-1)}\langle Z_{\sigma(j)}\rangle_{\sigma(1)\ldots\sigma(j-1)}\right]. \tag{B8}$$

Consider the expression $\mathbb{E}_{\mathcal{B}}\left[\langle Z_{\sigma(i)}\rangle_{\sigma(1)\ldots\sigma(i-1)}\langle Z_{\sigma(j)}\rangle_{\sigma(1)\ldots\sigma(j-1)}\right]$. The ensemble of circuits that we consider is invariant under the action of single-qubit Haar-random unitaries at each site at the output. Therefore, we can imagine that we have a layer of single-qubit Haar-random unitaries at every site and average over this layer separately. Averaging over this layer, we obtain an expression that is linear in the operator,



, (B9)

where $\int dU$ is a shorthand for $\int dU_1 dU_2 \ldots dU_n$ and we omit the unitaries that have averages of the form $\int dU_k U_k U_k^\dagger = \mathbb{1}$. Now, we observe that the average over $U_{\sigma(j)}$ gives $\int dU_{\sigma(j)} U_{\sigma(j)} Z_{\sigma(j)} U^\dagger_{\sigma(j)} = 0$ (there is no other instance of $U_{\sigma(j)}$ in this

expression, since $\sigma(j) \notin \{\sigma(1), \ldots \sigma(i)\}$ because $j > i$). Therefore, the average over all the gates is 0. This means that $\mathbb{E}_{\mathcal{B}}\left[\langle Z_{\sigma(i)}\rangle_{\sigma(1)\ldots\sigma(i-1)}\langle Z_{\sigma(j)}\rangle_{\sigma(1)\ldots\sigma(j-1)}\right] = 0$, giving

$$\mathbb{E}_{\mathcal{B}}\left[A_\sigma^2\right] \leq n. \tag{B10}$$

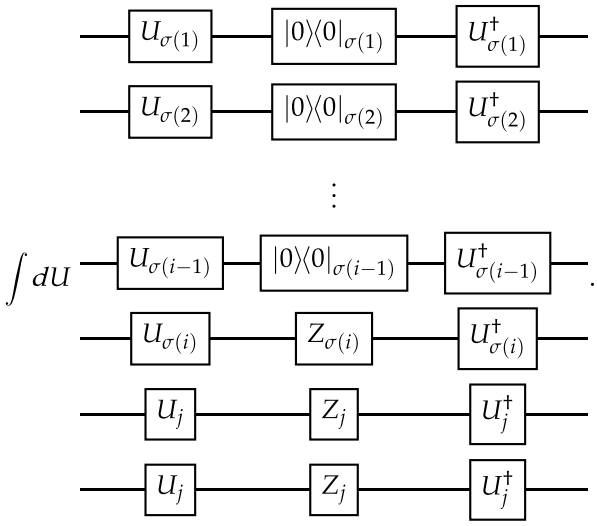
We obtain

$$\frac{1}{n!^2}\mathbb{E}_{\mathcal{B}}\left[\left(\sum_\sigma A_\sigma\right)^2\right] \leq \frac{1}{n!^2}\sum_{\sigma,\tau}\sqrt{n\times n} = n. \tag{B11}$$

The next term is

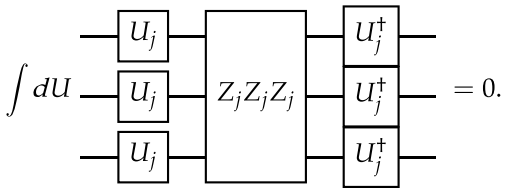
$$\frac{1}{2L_{\max}(d)n!}\mathbb{E}_{\mathcal{B}}\left[\sum_\sigma A_\sigma \sum_j \langle Z_j\rangle^2\right] = \frac{-1}{2L_{\max}(d)n!}\mathbb{E}_{\mathcal{B}}\left[\sum_\sigma\sum_i \langle Z_{\sigma(i)}\rangle_{\sigma(1)\ldots\sigma(i-1)}\sum_j\langle Z_j\rangle^2\right]. \tag{B12}$$

Consider a term in the sum above, which is proportional to $\mathbb{E}_{\mathcal{B}}\left[\langle Z_{\sigma(i)}\rangle_{\sigma(1)\sigma(2)\ldots\sigma(i-1)}\langle Z_j\rangle^2\right]$. By a similar argument as before, the average over the last layer of single-qubit Haar-random gates will be proportional to a three-copy Haar average of the following form:



(B13)

The expectation over $U_{\sigma(i)}$ is zero whenever $\sigma(i) \neq j$. The only (potentially) nonzero term is when $\sigma(i) = j$. But this term is zero as well, since the relevant three-copy Haar average over U_j is



(B14)

This brings us to the final term,

$$\frac{1}{16L_{\max}(d)^2}\mathbb{E}_{\mathcal{B}}\left[\sum_{i,j}\langle Z_i\rangle^2\langle Z_j\rangle^2\right] - \frac{1}{16L_{\max}(d)^2}\mathbb{E}_{\mathcal{B}}\left[\sum_i\langle Z_i\rangle^2\right]\mathbb{E}\left[\sum_j\langle Z_j\rangle^2\right] \tag{B15}$$

$$= \frac{1}{16L_{\max}(d)^2}\sum_{i,j}\left(\mathbb{E}_{\mathcal{B}}\left[\langle Z_i\rangle^2\langle Z_j\rangle^2\right] - \mathbb{E}_{\mathcal{B}}\left[\langle Z_i\rangle^2\right]\mathbb{E}_{\mathcal{B}}\left[\langle Z_j\rangle^2\right]\right). \tag{B16}$$

For a fixed i and j, the term $\mathbb{E}_{\mathcal{B}}\left[\langle Z_i\rangle^2\langle Z_j\rangle^2\right] - \mathbb{E}_{\mathcal{B}}\left[\langle Z_i\rangle^2\right]\mathbb{E}_{\mathcal{B}}\left[\langle Z_j\rangle^2\right]$ is zero unless j is in $L_d \circ L_d^\dagger(i)$. Otherwise, it is at most 1. This means that the sum is at most

$$\frac{1}{16L_{\max}(d)^2}\sum_i |L_d \circ L_d^\dagger(i)| \tag{B17}$$

$$\leq \frac{n}{16L_{\max}(d)} \leq n. \tag{B18}$$

The variance in Eq. (B1) is therefore at most $2n$. ■

[1] J. Preskill, Quantum computing in the NISQ era and beyond, Quantum **2**, 79 (2018).

[2] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. Brandao, D. A. Buell, B. Burkett, Quantum supremacy using a programmable superconducting processor, Nature **574**, 505 (2019).

[3] H.-S. Zhong, *et al.*, Quantum computational advantage using photons, Science **370**, 1460 (2020).

[4] H.-S. Zhong, *et al.*, Phase-Programmable Gaussian Boson Sampling Using Stimulated Squeezed Light, Phys. Rev. Lett. **127**, 180502 (2021).

[5] Y. Wu, *et al.*, Strong Quantum Computational Advantage Using a Superconducting Quantum Processor, Phys. Rev. Lett. **127**, 180501 (2021).

[6] W. Brown and O. Fawzi, Scrambling speed of random quantum circuits, ArXiv:1210.6644 (2012).

[7] W. Brown and O. Fawzi, Decoupling with random quantum circuits, Commun. Math. Phys. **340**, 867 (2015).

[8] F. G. S. L. Brandão, W. Chemissany, N. Hunter-Jones, R. Kueng, and J. Preskill, Models of Quantum Complexity Growth, PRX Quantum **2**, 030316 (2021).

[9] J. Haferkamp, P. Faist, N. B. T. Kothakonda, J. Eisert, and N. Y. Halpern, Linear growth of quantum circuit complexity, Nat. Phys. **18**, 528 (2022).

[10] A. C. Potter and R. Vasseur, in *Entanglement dynamics in hybrid quantum circuits*, edited by A. Bayat, S. Bose, and H. Johannesson (Springer, Cham, 2022).

[11] J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush, and H. Neven, Barren plateaus in quantum neural network training landscapes, Nat. Commun. **9**, 4812 (2018).

[12] S. Wang, E. Fontana, M. Cerezo, K. Sharma, A. Sone, L. Cincio, and P. J. Coles, Noise-induced barren plateaus in variational quantum algorithms, Nat. Commun. **12**, 6961 (2021).

[13] X. Gao and L. Duan, Efficient classical simulation of noisy quantum computation, ArXiv:1810.03176 (2018).

[14] A. Bouland, B. Fefferman, Z. Landau, and Y. Liu, in *2021 IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS)* (IEEE, Denver, CO, USA, 2022).

[15] D. Aharonov, M. Ben-Or, R. Impagliazzo, and N. Nisan, Limitations of noisy reversible computation, ArXiv:quant-ph/9611028 (1996).

[16] D. Aharonov, Quantum to classical phase transition in noisy quantum computers, Phys. Rev. A **62**, 062311 (2000).

[17] J. Emerson, R. Alicki, and K. Zyczkowski, Scalable noise estimation with random unitary operators, J. Opt. B: Quantum Semiclass. Opt. **7**, S347 (2005).

[18] S. Boixo, S. V. Isakov, V. N. Smelyanskiy, R. Babbush, N. Ding, Z. Jiang, M. J. Bremner, J. M. Martinis, and H. Neven, Characterizing quantum supremacy in near-term devices, Nat. Phys. **14**, 595 (2018).

[19] S. Boixo, V. N. Smelyanskiy, and H. Neven, Fourier analysis of sampling from noisy chaotic quantum circuits, ArXiv:1708.01875 (2017).

[20] Y. Liu, M. Otten, R. Bassirianjahromi, L. Jiang, and B. Fefferman, Benchmarking near-term quantum computers via random circuit sampling, ArXiv:2105.05232 (2021).

[21] D. Stilck França and R. García-Patrón, Limitations of optimization algorithms on noisy quantum devices, Nat. Phys. **17**, 1221 (2021).

[22] S. Wang, P. Czarnik, A. Arrasmith, M. Cerezo, L. Cincio, and P. J. Coles, Can error mitigation improve trainability of noisy variational quantum algorithms?, ArXiv:2109.01051 (2021).

[23] In this paper, we use the symbols O, Ω, Θ, o, and ω to denote various relations between the asymptotic scaling of two nonnegative functions $f(n)$ and $g(n)$, studied in the limit $n \to \infty$. We denote $f = O(g)$ if $\lim_{n\to\infty} f(n)/g(n) < \infty$, which is also equivalent to the relation $g = \Omega(f)$. Also, if $\lim_{n\to\infty} f(n)/g(n) = 0$, we have $f = o(g) \Leftrightarrow g = \omega(f)$. Lastly, we say that $f = \Theta(g)$ if both $f = O(g)$ and $f = \Omega(g)$ hold. Symbols with a tilde, such as $\tilde{\Theta}$, suppress potential polylogarithmic factors of n. We also use the asymptotic notation in conjunction with the negative sign in the following sense: $-\Omega(f)$ denotes the set of all functions g such that $-g(n) = \Omega(f(n))$, or $\lim_{n\to\infty}(f/-g) < \infty$. Therefore, $g(n) = -n^2$ satisfies $g(n) = -\Omega(n)$.

[24] S. Aaronson and A. Arkhipov, The computational complexity of linear optics, Theory Comput. **9**, 143 (2013).

[25] A. W. Harrow and A. Montanaro, Quantum computational supremacy, Nature **549**, 203 (2017).

[26] A. Harrow and S. Mehraban, Approximate unitary t-designs by short random quantum circuits using nearest-neighbor and long-range gates, ArXiv:1809.06957 (2018).

[27] The expectation value $\mathbb{E}_{\mathcal{B}}$ here is over the choice of the random unitary gates.

[28] A. M. Dalzell, N. Hunter-Jones, and F. G. S. L. Brandão, Random Quantum Circuits Anticoncentrate in Log Depth, PRX Quantum **3**, 010333 (2022).

[29] A. Nahum, J. Ruhman, S. Vijay, and J. Haah, Quantum Entanglement Growth under Random Unitary Dynamics, Phys. Rev. X **7**, 031016 (2017).

[30] A. Nahum, S. Vijay, and J. Haah, Operator Spreading in Random Unitary Circuits, Phys. Rev. X **8**, 021014 (2018).

[31] N. Hunter-Jones, Unitary designs from statistical mechanics in random quantum circuits, ArXiv:1905.12053 (2019).

[32] B. Barak, C.-N. Chou, and X. Gao, Spoofing linear cross-entropy benchmarking in shallow quantum circuits, ArXiv:2005.02421 (2020).

[33] A. M. Dalzell, N. Hunter-Jones, and F. G. S. L. Brandão, Random quantum circuits transform local noise into global white noise, ArXiv:2111.14907 (2021).

[34] M. J. Bremner, A. Montanaro, and D. J. Shepherd, Average-Case Complexity versus Approximate Simulation of Commuting Quantum Computations, Phys. Rev. Lett. **117**, 080501 (2016).

[35] H. Krovi, Average-case hardness of estimating probabilities of random quantum circuits with a linear scaling in the error exponent, ArXiv:2206.05642 (2022).

[36] Y. Kondo, R. Mori, and R. Movassagh, in *2021 IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS)* (IEEE, Denver, CO, USA, 2022).

[37] A. Bouland, B. Fefferman, C. Nirkhe, and U. Vazirani, On the complexity and verification of quantum random circuit sampling, Nat. Phys. **15**, 159 (2019).

[38] R. Movassagh, Quantum supremacy and random circuits, ArXiv:1909.06210 (2019).

[39] J. Napp, R. L. La Placa, A. M. Dalzell, F. G. S. L. Brandao, and A. W. Harrow, Efficient Classical Simulation of Random Shallow 2D Quantum Circuits, Phys. Rev. X **12**, 021021 (2022).

[40] Napp *et al.* [39] has also had nonrigorous arguments that appear to give convincing results for depths larger than 3 but smaller than some architecture-dependent constant.

[41] The results in Fig. 1(b) remain the same in other dimensions.

[42] R. J. Lipton, in *Distributed Computing and Cryptography, Proceedings of a DIMACS Workshop, Princeton, New Jersey, USA, October 4–6, 1989*, DIMACS Series in Discrete Mathematics and Theoretical Computer Science, Vol. 2, edited by J. Feigenbaum and M. Merritt (DIMACS/AMS, Princeton, New Jersey, 1989), p. 191.

[43] There are simple ways to make the trivial algorithms depend on the instance; we are interested in an algorithm that depends *nontrivially* on the instance.

[44] S. Aaronson and L. Chen, in *Proceedings of the 32Nd Computational Complexity Conference*, CCC '17 (Schloss Dagstuhl–Leibniz-Zentrum für Informatik, Germany, 2017), p. 22:1.

[45] S. Aaronson and S. Gunn, On the classical hardness of spoofing linear cross-entropy benchmarking, Theory Comput. **16**, 1 (2020).

[46] J. Choi, A. L. Shaw, I. S. Madjarov, X. Xie, J. P. Covey, J. S. Cotler, D. K. Mark, H.-Y. Huang, A. Kale, H. Pichler, F. G. S. L. Brandão, S. Choi, and M. Endres, Emergent randomness and benchmarking from many-body quantum chaos, ArXiv:2103.03535 (2021).

[47] X. Gao, M. Kalinowski, C.-N. Chou, M. D. Lukin, B. Barak, and S. Choi, Limitations of linear cross-entropy as a measure for quantum advantage, ArXiv:2112.01657 (2021).

[48] Note that Liu *et al.* [20] write this in terms of $\lambda = n\epsilon$, the total noise acting on a layer of gates.

[49] N. Yunger Halpern and E. Crosson, Quantum information in the Posner model of quantum cognition, Ann. Phys. **407**, 92 (2019).

[50] A. Chan, R. M. Nandkishore, M. Pretko, and G. Smith, Weak measurements limit entanglement to area law (with possible log corrections), Phys. Rev. B **99**, 224307 (2019).

[51] Y. Li, X. Chen, and M. P. A. Fisher, Quantum Zeno effect and the many-body entanglement transition, Phys. Rev. B **98**, 205136 (2018).

[52] B. Skinner, J. Ruhman, and A. Nahum, Measurement-Induced Phase Transitions in the Dynamics of Entanglement, Phys. Rev. X **9**, 031009 (2019).

[53] M. J. Gullans and D. A. Huse, Dynamical Purification Phase Transition Induced by Quantum Measurements, Phys. Rev. X **10**, 041020 (2020).

[54] S. Vijay, Measurement-driven phase transition within a volume-law entangled phase, ArXiv:2005.03052 (2020).

[55] Y. Bao, M. Block, and E. Altman, Finite time teleportation phase transition in random quantum circuits, ArXiv:2110.06963 (2022).

[56] A. Müuller-Hermes, D. Stilck França, and M. M. Wolf, Relative entropy convergence for depolarizing channels, J. Math. Phys. **57**, 022202 (2016).

[57] C. Hirche, C. Rouzé, and D. S. França, On contraction coefficients, partial orders and approximation of capacities for quantum channels, ArXiv:2011.05949 (2021).

[58] M. Ben-Or, D. Gottesman, and A. Hassidim, Quantum refrigerator, ArXiv:1301.1995 (2013).

[59] D. Hangleiter, J. Bermejo-Vega, M. Schwarz, and J. Eisert, Anticoncentration theorems for schemes showing a quantum speedup, Quantum **2**, 65 (2018).

[60] A. W. Harrow and R. A. Low, Random quantum circuits are approximate 2-designs, Commun. Math. Phys. **291**, 257 (2009).

[61] F. G. S. L. Brandao, A. W. Harrow, and M. Horodecki, Local random quantum circuits are approximate polynomial-designs, Commun. Math. Phys. **346**, 397 (2016).

[62] T. Morimae, Hardness of classically sampling the one-clean-qubit model with constant total variation distance error, Phys. Rev. A **96**, 040302(R) (2017).

[63] J. J. Wallman and J. Emerson, Noise tailoring for scalable quantum computation via randomized compiling, Phys. Rev. A **94**, 052325 (2016).

[64] We do not constrain the "grouping" of these qubits. In fact, this grouping can change from layer to layer.

[65] Note that for a Haar-random unitary, $\mathbb{E}_{\mathcal{U}} Z(U, \mathbb{I}) = 2/(2^n + 1)$.

[66] For a sufficient set of connectivity conditions, see Dalzell *et al.* [28].