We are entering a period of time where experimental quantum devices of growing size and fidelity will perform well defined computational tasks, of progressing practical value, beyond the capabilities of state-of-the-art classical supercomputers. Sampling problems, where the goal is to approximately sample from a well defined probability distribution, are prominent examples of such tasks, and also possess interesting relations to computational complexity theory [1•5].

Reference [3] proposed sampling from a chaotic quantum evolution as a particularly promising computational task for a near-term “quantum supremacy” demonstration. More specifically, the task is to approximately sample from the output distribution \( \{p_U(x)\} \) (over bit-strings \( \{x\} \)) of a pseudo-random universal quantum circuit \( U \). We are particularly interested in 2D circuits of depth \( O(\sqrt{n}) \), where \( n \) is the number of qubits. Classically sampling from a distribution correlated with the output distribution of \( U \) likely requires exponential classical resources (when averaged over instances in an ensemble). This can be argued in two different ways. From a physical point of view, pseudo-random universal quantum circuits, as examples of chaotic quantum evolutions, are hypersensitive to perturbations [7•8]. Furthermore, quantum states from instances of an ensemble of chaotic evolutions spread quasi-uniformly in Hilbert space, and each output distribution quickly approaches the entropy and lower moments of the characteristic Porter-Thomas (or exponential) distribution [3•4•9•10]. Therefore, sampling from the output distribution is expected to require a direct simulation of the quantum dynamics, which classically implies exponential cost. The same paper [3] also extends previous results in computational complexity theory [11•21•20•27], and in particular of Ref. [2], to argue the same point.

More specifically, assume a polynomial classical polynomial algorithm which takes a description of circuit \( U \) as input. We write its output probabilities as \( \{p_{\text{pol}}(x|U)\} \). Our assumption is that there is no polynomial time classical algorithm for which the average cross-entropy between \( \{p_{\text{pol}}(x|U)\} \) and the ideal probabilities \( \{p_U(x)\} \) is less than 3.

\[
E_U \left[ \sum_{j=1}^{N} p_{\text{pol}}(x_j|U) \log \frac{1}{p_U(x_j)} \right] \leq \log N + \gamma - \Omega \left( \frac{1}{N} \right),
\]

where \( N = 2^n \) and \( \gamma \) is the Euler’s constant. The value \( \log N + \gamma \) corresponds to the case where \( p_{\text{pol}}(x;U) \) is uncorrelated with \( p_U(x) \), assuming enough depth so that the entropy of \( \{p_U(x)\} \) is well approximated by the entropy from the Porter-Thomas distribution \( E_U[H(p_U)] = \log N + \gamma - 1 \), with \( U \) dependent fluctuations of order \( 2^{-n/2} \). The reason why we focus on the cross entropy is because it is approximately linear in the circuit fidelity. We use this to estimate the fidelity using cross-entropy benchmarking. A related quantum threshold conjecture was presented recently, stating that there is no polynomial classical algorithm which can guess if \( p_U(x) \) is above or below the median with probability at least \( 1/2 + \Omega(1/N) \) [28•].

Nevertheless, under fairly general noise models, the output distribution of chaotic quantum dynamics converges to the classical uniform distribution [3]. Therefore, in the asymptotic limit, it suffices to sample uniformly at random to approximate the experimental output for any fixed distance \( \delta \) in the \( \ell_1 \) norm. This can be done classically in linear cost on the number of qubits \( n \).

---

1 It is formally possible that an algorithm could brake the cross-entropy assumption without braking the quantum threshold conjecture [28].
In other words, “achieving a constant error in the limit of large \( n \) requires a fault tolerant quantum computer, which will not be available in the near term” [29, 31] [3]. It is therefore critical for experimentally accessible near-term quantum supremacy experiments to count with a well-defined metric for the relevant computational task. In this vein, one of the main contributions of Ref. [3] is the introduction of cross-entropy benchmarking as an approximation to the fidelity, and to the distance to the uniform distribution, for complex quantum systems.

Reference [3] also gives an ansatz for the density matrix produced by a noisy chaotic quantum evolution

\[
\rho \simeq \alpha |\psi\rangle \langle \psi| + (1 - \alpha) \frac{1}{N},
\]

where \( |\psi\rangle \) is the ideal output and \( \alpha \) is the fidelity. This ansatz is supported numerically by the observation that the output distribution after adding a single discrete error is (almost) uncorrelated with the ideal output distribution, and noticing that \( \alpha \) is (almost) the probability of no-error in the quantum circuit. It is also consistent with simulations where each ideal gate is followed by a depolarizing channel with error rate \( \epsilon \), a common error model well matched by experimental results [32, 39]. The fidelity \( \alpha \) can be approximated by \( \alpha \simeq e^{-\epsilon m} \), where \( m \) is the number of gates and \( \epsilon \) is the error per gate. Therefore, classically it suffices to choose bit-strings \( x \) uniformly at random to approximate the output distribution of a noisy quantum circuit with statistical distance \( \delta \simeq e^{-\epsilon m} \). We note also that noisy quantum circuits do not violate asymptotically the cross-entropy assumption or the quantum threshold conjecture stated above for polynomial classical algorithms.

Nonetheless, a quantum computer with gate error rates within reach in the near term would be able to approximate the ideal distribution with an statistical distance beyond the capabilities of state-of-the-art classical supercomputers. For instance, two-qubit, initialization and measurement error rates of 0.3\%, with single-qubit error rates of 0.06\%, would result in a final fidelity of approximately 10\% for circuits of 7 \times 7 qubits in a 2D lattice and depth 40. This depth takes into account current constraints in the layout of two qubit gates for superconducting qubits [32, 35]. Sampling a correlated distribution classically would require \( 2^{40} \times 8 \times 2 \) bytes in Rapid Access Memory, which is likely to be unfeasible. The conjectures stated above give support to the estimated exponential cost of classical approximation algorithms.

A recent interesting paper [40] extends results on commuting or IQP quantum circuits [4] to universal quantum circuits, and obtains a polynomial time classical algorithm for approximating the output distribution under certain models of noise. The upper bound on the cost is \( (n + m)^{O(\frac{1}{\epsilon} \log \frac{1}{\delta})} \) where \( n \) is the number qubits, \( m \) the number of gates, \( \epsilon \) the error per gate, and \( \delta \) the desired distance in the \( \ell_1 \) norm from the noisy distribution. A first observation is that for the experimental error rates \( \epsilon \simeq 0.3\% \) within reach of experimental quantum computers, the estimate of this upper bound is much worse than a direct simulation for sizes of practical interest [4]. Furthermore, the upper bound to improve over the statistical distance \( \delta \simeq e^{-\epsilon m} \) achieved by random sampling of bit-strings is \( (n + m)^{O(m)} \), again worse than exact simulation. We next analyze this algorithm in more detail and conclude that indeed, when restricted to polynomial time, it results in an output distribution whose correlation with the ideal distribution is exponentially small in the number of gates \( m \gg n \) for chaotic quantum circuits.

The methodology of Ref. [40] proceeds in two steps. First, a universal random circuit is mapped to an IQP circuit using techniques related to measurement-based quantum computation. The resulting IQP circuit has \( n' = n + m \) qubits. Second, the noisy output distribution is approximated using the Fourier analysis algorithm from Ref. [4]. For our purposes it will suffice to focus in the second step.

Let’s now review the algorithm from Ref. [4]. Consider

---

**FIG. 1.** Histograms of the entropy \( H(p_{\text{IQP}}) \) of the output distribution of sparse IQP circuits [4] with 20 qubits showing the convergence to the entropy of the corresponding Porter-Thomas distribution \( H(p_{\text{PT}}) \) with increasing circuit density, controlled by \( \gamma \). For each possible choice of a pair \((j,k)\) of distinct qubits, we apply a controlled-phase gate across those qubits with probability \( \gamma (\log n)/n \).
Therefore define a Fourier transform as
\[
\hat{p}_{\text{IQP}}(s) = \frac{1}{2^n} \sum_y f^*(y) f(y + s) .
\] (9)

It can be seen that \[1\], from the Chernoff bound, we can approximate \(2^n \hat{p}_{\text{IQP}}(s)\) up to an additive error \(\eta\) in time \(O(1/\eta^2)\) using the same order of evaluations of \(f\).

Consider now an output distribution \(p'_{\text{IQP}}(x)\) where a depolarizing channel with error rate \(\epsilon\) is added to each qubit, exactly before measurement. Note that a limitation of this model is that errors do not spread, because there are no errors prior to any gate. The motivation is that this noise model can be studied using tools from Fourier analysis. Consider the Fourier transform \(\hat{p}'_{\text{IQP}}(s)\) of \(p'_{\text{IQP}}(x)\). It can be shown that this error model diminishes the Fourier coefficients \(\hat{p}'_{\text{IQP}}(s)\) exponentially in the Hamming weight \(l = |s|\) of the argument \(s\) [3, 11].

\[
\hat{p}'_{\text{IQP}}(s) = (1 - \epsilon)^{|s|} \hat{p}_{\text{IQP}}(s) .
\] (10)

Note that the uniform distribution over bit-string has Fourier components \(\hat{p}_{\text{uniform}}(s) = 2^{-n} \delta_{s0}\). Equation (10) then says that the Fourier components of \(p_{\text{IQP}}\) converge to the Fourier components of the uniform distribution exponentially in \(|s|\). There is some discrepancy with the ansatz of Eq. (2), which implies that the Fourier components converge to those of the uniform distribution exponentially in the number of gates \(n\), and independently of \(|s|\). Nevertheless, for this ansatz we are assuming a very different error model: we apply a depolarizing channel after each gate in a circuit with significant depth, so errors can propagate to most qubits. This is in contrast to the noise model that led to Eq. (10), where errors are constrained to the qubit in which they occur.

The algorithm in Ref. [4] consists in approximating a polynomial number \(O(n!)\) of Fourier components with low weight \(|s| \leq l\), up to polynomial error \(O(\delta n^{-l/2})\). According to Eq. (10), Fourier components with higher weight are diminished by a factor of \((1 - \epsilon)^l\). Now assume that \(\sum p^2_{\text{IQP}}(s) \leq 2^{-n}\) for some constant \(\beta\) independent of \(n\). It is then shown that it suffices to choose \(l = O(\log \beta/\delta)/\epsilon\) for a target statistical distance \(\delta\). An upper bound \(O(n^{O(\log \beta)})\) on the cost of this algorithm follows from a bound on the number of Fourier components \(O(n!)\) being approximated, and the previous bound on the cost to approximate each component.

We now estimate the distribution of Fourier components. The output probabilities \(p_{\text{PR}}(x) = |\langle x | \psi \rangle|^2\) from a state \(|\psi\rangle\) chosen uniformly at random in Hilbert space have a characteristic distribution called the Porter-Thomas (or exponential) distribution [9, 42, 43]. They are i.i.d distributed (up to normalization) with \(\text{Pr}(p_{\text{PR}}) = Ne^{-Np_{\text{PR}}}\). Figure 1 shows that the entropy of the output distribution of IQP circuits approximates the Porter-Thomas distribution. This is also true a fortiori for the output probabilities \(\{p_{\text{PR}}(x)\}\) of pseudo-random universal quantum circuits [3]. We therefore will use the ap-

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{fig2.png}
\caption{Numerical Fourier components for the output distribution of pseudo-random universal quantum circuits [3] with 5 \times 4 qubits and depth 40, rescaled by the standard deviation \(2^{-3n/2}\) from Porter-Thomas. The two-qubit gate error rate is, from blue to green, \(\epsilon = [0, 0.0001, 0.0002, 0.0005, 0.001, 0.002, 0.005, 0.01, 0.02, 0.05]\), the single-qubit gate error rate is 10 times less. We plot 10 instances and 10 Fourier components for each error rate and weight \(l = |s|\). Numerics confirm the Porter-Thomas estimates, and all components converge to 0 (uniform random distribution) with increased noise.}
\end{figure}
Therefore, guessing the uniform random distribution is 0 (the random uniform distribution) with increasing \( \epsilon \) by 2 pseudo-random universal quantum circuits \([3]\), rescaled increasing noise \( \epsilon \) are of order \( \hat{\epsilon} \) the Porter-Thomas estimates: all Fourier components standard deviation, also rescaled. The numerics confirm the Porter-Thomas distribution is Gaussian with mean 0 and standard deviation of Fourier components of probabilities with a Porter-Thomas approximation. That is, the expectation of each Fourier component is the same as for the uniform distribution. The variance is

\[
\text{Var} (\hat{\hat{p}}_{\text{PT}} (s)) = \frac{1}{2n} \text{Var} \left( \sum_x p_{\text{PT}}(x)(-1)^{s \cdot x} \right) = 0 .
\]

That is, the expectation of each Fourier component is the same as for the uniform distribution. The variance is

\[
\text{Var} (\hat{p}_{\text{PT}} (s)) = \frac{1}{2n} \text{Var} \left( \sum_x p_{\text{PT}}(x)(-1)^{s \cdot x} \right)
\]

\[
= \frac{1}{2n} \text{Var} (p_{\text{PT}}) = 2^{-3n} .
\]

More specifically, we show in App. \([A]\) that the distribution of Fourier components of probabilities with a Porter-Thomas distribution is Gaussian with mean 0 and standard deviation \( 2^{-3n/2} \).

Figure 2 shows Fourier components obtained numerically from the output distribution \( \{p_{\text{PT}}(x)\} \) of low-depth pseudo-random universal quantum circuits \([3]\), rescaled by \( 2^{-3n/2} \). Colors from blue to green correspond to increasing noise \( \epsilon \). Figure 3 shows the numerical standard deviation, also rescaled. The numerics confirm the Porter-Thomas estimates: all Fourier components are of order \( \hat{p}_{\text{PT}}(s) \in O(2^{-3n/2}), \) and they converge to 0 (the random uniform distribution) with increasing \( \epsilon \). Therefore, guessing the uniform random distribution is already a good approximation to all the Fourier components \( 2^n \hat{p}_{\text{PT}} (s) \) with an additive error \( O(2^{-n/2}) \).

A critical subroutine of the polynomial approximation algorithm to sample from IQP circuits \([4]\) (and extended to pseudo-random circuits in Ref. \([10]\)), is the approximation of a Fourier component \( \hat{p}_{\text{IQP}}(s) \) with additive error \( \eta/2^n \) at cost \( O(1/\eta^2) \). We just showed above, both numerically for pseudo-random circuits, and analytically for the Porter-Thomas distribution, that \( 2^n \hat{p}(s) \in O(2^{-n/2}) \). Therefore, improving this trivial estimate for a single Fourier component requires \( \eta \lesssim 2^{-n/2} \) and, according to the Chernoff bound, cost \( O(2^n) \). Even for IQP circuits, using the Porter-Thomas approximation, this is already of the same order as evaluating the probability exactly. When this technique is used for chaotic random circuits, the cost is exponential in \( n = m + n \).

In conclusion, noisy pseudo-random universal quantum circuits, as well as noisy IQP circuits, converge to the uniform random distribution exponentially in the error rate \( \epsilon \). Therefore, approximating an experimental noisy output distribution for large \( n \) can be done in cost \( O(n) \). The polynomial algorithm based on Fourier analysis \([4,40]\) does not output a distribution correlated with the ideal distribution, because it does not improve over guessing uniformly at random. Therefore, it does not represent an additional challenge over previous results suggesting the failure of classical devices in simulating noisy chaotic circuits with about 48 qubits and depth 40, as long as the error rate per gate is low (\( \sim 0.3\% \)). This extreme sensitivity of chaotic evolutions to noise allows experimentalists to approximate the global fidelity of complex evolutions using cross-entropy benchmarking, and important metric of experimental progress.

ACKNOWLEDGMENTS

We would like to thank Michael J. Brenner, Edward Farhi and Ashley Montanaro for discussions.

Appendix A: Distribution of Fourier Components

We give here a direct derivation of the form of the distribution of the probability of Fourier components \( \hat{p}_T(p_{\text{PT}}) \) assuming that the probabilities \( \{p_{\text{PT}}(x)\} \) have a Porter-Thomas (exponential) distribution \([4,42,43]\).

For a given bit-string \( s \), the value of the Fourier coefficient is

\[
\hat{p}(s) = \frac{1}{N} \sum_x p(x)(-1)^{s \cdot x} = \frac{1}{N} \left( 1 - 2 \sum_{x \in S} p(x) \right) ,
\]

\[2\] Nevertheless, note that there are an exponential number of Fourier components and that the statistical distance between the ideal and the uniform distribution is \( \sum_x |p_{\text{PT}}(x) - 1/N| = 2/\epsilon \).
where $x \in S$ if and only if $s \cdot x$ is odd, and $N = 2^n$. The size of $S$ is $|S| = 2^{n-1}$ (with $s \neq 0$). We first calculate the distribution

$$
\Pr_S(u) = \left\langle \delta \left( \sum_{x \in S} p(x) - u \right) \right\rangle. \tag{A2}
$$

Assume a quantum state $|\psi\rangle$ sampled uniformly at random in Hilbert space. We write a generic state in the basis $\{|x\rangle\}$ as $|\psi\rangle = \sum_x (a_x + i b_x) |x\rangle$. Then

$$
\Pr_S(u) = \frac{\text{num}(u)}{\text{dem}} \tag{A3}
$$

where

$$
\text{num}(u) = \int_{-\infty}^{\infty} \Pi_x da_x db_x \delta \left( \sum_x a_x^2 + b_x^2 - 1 \right) \delta \left( \sum_{j=1}^{N/2} a_j^2 + b_j^2 - u \right), \tag{A4}
$$

and

$$
\text{dem} = \int_{-\infty}^{\infty} \Pi_x da_x db_x \delta \left( \sum_x a_x^2 + b_x^2 - 1 \right). \tag{A5}
$$

We calculate the numerator as

$$
\begin{align*}
\int_{-\infty}^{\infty} & \Pi_x da_x db_x \delta \left( \sum_x a_x^2 + b_x^2 - 1 \right) \delta \left( \sum_{j=1}^{N/2} a_j^2 + b_j^2 - u \right) \\
= & \int_{-\infty}^{\infty} \Pi_x da_x db_x \frac{1}{2\pi} \int_{-\infty}^{\infty} dt e^{it \sum_x (a_x^2 + b_x^2) - it u} \\
= & \frac{1}{2\pi} \int_{-\infty}^{\infty} dw e^{i \sum_{j=1}^{N/2} (a_j^2 + b_j^2) - iw u} \\
= & \frac{1}{2\pi} \int_{-\infty}^{\infty} dw e^{-iwu} \left( \int_{-\infty}^{\infty} da_x db_x e^{i t x (a_x^2 + b_x^2)} \right)^{N/2} \\
= & \frac{1}{2\pi} \int_{-\infty}^{\infty+it} dw e^{-it/u} \left( \frac{\pi}{i(t+u)} \right)^{N/2} \\
= & \pi^{N} \frac{1}{((-i)^N (N/2 - 1)! \lim_{t \to 0} d^{N/2-1} t^{-N/2} e^{-it(1-u)})} \\
= & \frac{1}{(N/2 - 1)! \lim_{u \to 0} d^{N/2} e^{-iu}} (1 - u)^{N/2 - 1} u^{N/2 - 1}. \tag{A6}
\end{align*}
$$

The denominator is calculated in a similar way

$$
\begin{align*}
\int_{-\infty}^{\infty} & \Pi_x da_x db_x \delta \left( \sum_x a_x^2 + b_x^2 - 1 \right) \\
= & \frac{1}{2\pi} \int_{-\infty}^{\infty} dt e^{-it} \left( \int_{-\infty}^{\infty} da_x db_x e^{i t x (a_x^2 + b_x^2)} \right)^N \\
= & \pi^{N} \frac{1}{(N - 1)!} \\
\end{align*} \tag{A7}
$$

Therefore

$$
\Pr_S(u) = \frac{(N - 1)!}{((N/2 - 1)!)^2} (1 - u)^{N/2 - 1} u^{N/2 - 1}. \tag{A8}
$$

For the distribution of a Fourier component

$$
\hat{\Pr}_{\nuT}(s) = \frac{1}{N} \left( 1 - 2 \sum_{x \in S} p(x) \right) = \frac{1}{N} (1 - 2u). \tag{A9}
$$

we have

$$
\Pr_{\nuT}(\hat{\nuT}) = \frac{N}{2} \Pr \left( (1 - \hat{\nuT} N)/2 \right) \tag{A10}
$$

$$
\begin{align*}
= & \frac{N}{2} \frac{(N - 1)!}{((N/2 - 1)!)^2} \left( \frac{1}{2} + \frac{\hat{\nuT} N}{2} \right)^{N/2 - 1} \\
= & \frac{N^2}{2^{N+1}} \frac{N}{2} \left( 1 - \frac{\hat{\nuT}^2}{2} \right)^{N/2 - 1} \\
\approx & \frac{N^{3/2}}{\sqrt{2\pi}} \exp \left( -\frac{1}{2} \hat{\nuT}^2 N^3 \right). \tag{A11}
\end{align*}
$$

In conclusion, the distribution of Fourier components is Gaussian with mean 0 and standard deviation $N^{-3/2}$. 

