Characterizing Quantum Supremacy in Near-Term Devices

S. Boixo

S. Isakov, V. Smelyanskiy, R. Babbush, M. Smelyanskiy, N. Ding, Z. Jiang, J.Martinis, H. Neven

ICTP

August 27th

Quantum Supremacy
J. Preskill, 2012

With a quantum device

- perform a well-defined computational task
- beyond the capabilities of state-of-the-art classical supercomputers
- in the near-term
    - without error correction.

Not necessarily solving a practical problem.

Approaches to quantum supremacy

- Optimization of a classical function:
  - Quantum Annealing.
  - Quantum Approximate Optimization Algorithm (E. Farhi et. al.).
- Non-simulable Hamiltonian Evolution.
- Variational Quantum Eigensolver (Ground state energy of a Hamiltonian).
- Approximate sampling from a well defined distribution:
  - Commuting Quantum Circuits (M. Bremner et. al.).
  - BosonSampling. (Aaronson and Arkhipov).
  - Random Universal Circuits. “Randomized benchmarking for complex circuits.”

Requisites for quantum supremacy in the near-term

- Classically, nothing must work for the computational task, except direct simulation of quantum evolution.
  - Cost exponential in size of Hilbert space.
  - Typical of chaotic systems.
- Specific figure of merit for the computational task.
  - We should measure the figure of merit up to quantum supremacy frontier.
  - Naturally related to fidelity.
- Well understood extrapolation of figure of merit beyond the quantum supremacy frontier where it can not be measured. (Unfortunately, we lack witness.)
- Predictions from theory for figure of merit.
- Relation to Computational Complexity is a plus.
  - Formal Computational Complexity is asymptotic, requires error correction (Strong Church-Turing Thesis).

Random Universal Quantum Circuits

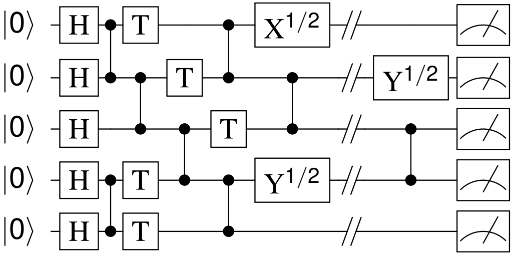


Figure: Vertical lines correspond to controlled-phase gates .

- Random quantum circuits are examples of quantum chaos.
- Classically sampling $p_U(x) = |\langle x| U |0\rangle|^2$ requires direct simulations. Cost in $2D$ exponential in $\propto \min(n, d\sqrt{n})$, depth d, qubits n. (With 7×7 qubits requires $d \simeq 25$.)
- Good benchmark for quantum computers.
- New results in computational complexity.

Porter-Thomas distribution

- (Pseudo-)random circuit U

$$|\Psi\rangle = U\,|0\rangle = \sum_{j=1}^{N} c_i\,|x_i\rangle \ .$$

- Sample the output distribution with probabilities

$$p_i = |c_i|^2 = |\,\langle x_i|\,U\,|\Psi\rangle\,|^2\ .$$

- Real and imaginary parts of c_i are distributed (quasi) uniformly on a $2N$ dimensional sphere (Hilbert space) if the circuit (or Hamiltonian evolution) has sufficient depth (evolution time).
  - The distribution of c_i is, up to finite moments, Gaussian with mean 0 and variance $\propto 1/N$ (random unitary matrices, delocalization, level repulsion...).
- Porter-Thomas distribution: $\Pr(Np) = e^{-Np}$.

Porter-Thomas distribution

Histogram of the output distribution for different values of the two-qubit gate error rate *r*.

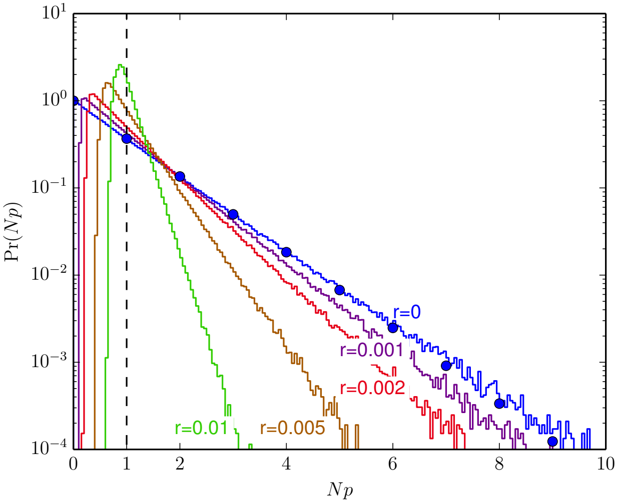


Figure: Circuit with 5 $\times$ 4 qubits (2D lattice) and depth 25.

Verification and uniformity test

- The PT distribution is very flat: $p(x_j) \sim 1/N$.
- The ℓ_1 distance between PT and uniform distribution is

$$\sum_j |p(x_j) - 1/N| = 2/e\ .$$

- If we calculate $p(x_j)$ given circuit U, we can distinguish these distributions with a constant number of measurements.
- If we don't know anything about $p(x_j)$ (black-box setting) we need $\Theta(\sqrt{N})$ measurements.
- There is no polynomial witness for this sampling problem. This problem is much harder than NP.
  - This is required for near-term (few qubits) supremacy.

Sampling from ideal circuit U

Sample $S = \{x_1, \ldots, x_m\}$ of bit-strings x_j from circuit U (measurements in the computational basis).

$$\log \mathrm{Pr}_U(S) = \sum_{x_j \in S} \log p_U(x_j) = -m\,\mathrm{H}(p_U) + O(m^{1/2}) \ ,$$

where $H(p_U)$ is the entropy of PT

$$\mathrm{H}(p_U) = -\int_0^\infty pN^2 e^{-Np} \log p \, dp = \log N - 1 + \gamma \ .$$

and $\gamma \simeq 0.577$.

Sampling with polynomial classical circuit $A_{\mathrm{pcl}}(U)$

A *polynomial* classical algorithm $A_{\mathrm{pcl}}(U)$ produces sample $S_{\mathrm{pcl}} = \{x_1^{\mathrm{pcl}}, \ldots, x_m^{\mathrm{pcl}}\}$. The probability $\mathrm{Pr}_U(S_{\mathrm{pcl}})$ that this sample S_{pcl} is observed from the output $|\psi\rangle$ of the circuit U is

$$\log \mathrm{Pr}_U(S_{\mathrm{pcl}}) = -m\,\mathrm{H}(p_{\mathrm{pcl}}, p_U) + O(m^{1/2}) \ ,$$

where

$$\mathrm{H}(p_{\mathrm{pcl}}, p_U) \equiv -\sum_{j=1}^{N} p_{\mathrm{pcl}}(x_j|U) \log p_U(x_j)$$

is the cross entropy.

Sensitivity to single Pauli error

A single Pauli error (almost) destroys the output distribution p_U.

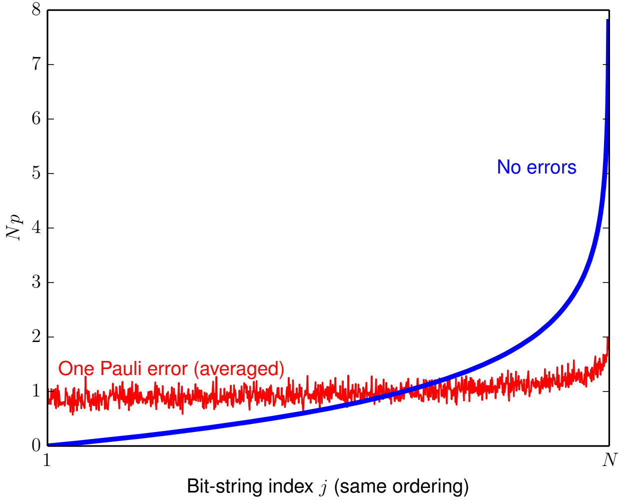


Figure: Blue line shows sorted probabilities $p_U(x_j)$ (universal quantum chaos distribution, Porter-Thomas). Red line average of a single Pauli error in all different locations, same ordering.

Sampling with polynomial classical circuit $A_{\mathrm{pcl}}(U)$ (II)

We are interested in the average over $\{U\}$ of random circuits (or chaotic evolutions)

$$\mathbb{E}_U\left[\mathrm{H}(p_{\mathrm{pcl}}, p_U)\right] = \mathbb{E}_U\left[\sum_{j=1}^{N} p_{\mathrm{pcl}}(x_j|U)\log\frac{1}{p_U(x_j)}\right] .$$

Because U is chaotic, Hilbert space has exponential dimension, and $A_{\mathrm{pcl}}(U)$ is polynomial, we conjecture that p_{pcl} and p_U are (almost) uncorrelated (more reasons later). We can take averages independently.

$$-\mathbb{E}_U\left[\log p_U(x_j)\right] \approx -\int_0^\infty N e^{-Np}\log p\, dp = \log N + \gamma .$$

$$\mathbb{E}_U\left[\mathrm{H}(p_{\mathrm{pcl}}, p_U)\right] = \log N + \gamma \equiv \mathrm{H}_0 .$$

Cross entropy difference

- The average cross entropy of a polynomial classical algorithm is the same as for a uniform distribution $p(x) = 1/N$.
- For algorithm A (quantum or classical of any cost) define the cross entropy difference

$$\alpha \equiv \Delta \mathrm{H}(p_A) \equiv \log N + \gamma - \mathrm{H}(p_A, p_U)$$
$$= \sum_j \left(\frac{1}{N} - p_A(x_j|U) \right) \log \frac{1}{p_U(x_j)} .$$

- The cross entropy goes between $\alpha = 0$ for no correlation, and $\alpha = 1$ for the ideal circuit.

Cross entropy and fidelity

- The output of an evolution with fidelity $\tilde{\alpha}$ is

$$\rho = \tilde{\alpha} U |0\rangle\langle 0| U^\dagger + (1-\tilde{\alpha})\sigma_U ,$$

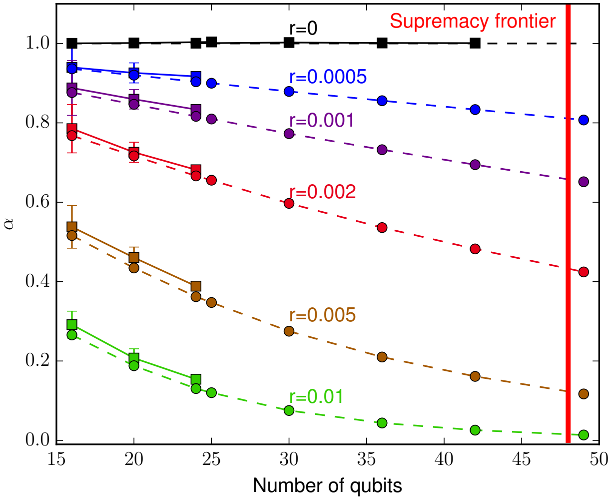
  with $p_{\rm exp}(x) = \langle x|\,\rho\,|x\rangle = \tilde{\alpha} p_U(x) + (1-\tilde{\alpha})\,\langle x|\,\sigma_U\,|x\rangle$.
- We again conjecture that $\langle x|\,\sigma_U\,|x\rangle$ is uncorrelated with $p_U(x)$.

$$\begin{aligned}
\alpha &= \mathbb{E}_U[\Delta \mathrm{H}(p_{\rm exp})] \\
&= \mathrm{H}_0 + \sum_j \left(\tilde{\alpha} p_U(x_j) + (1-\tilde{\alpha})\,\langle x_j|\,\sigma_U\,|x_j\rangle\, x\right) \log p_U(x_j) \\
&= \mathrm{H}_0 - \tilde{\alpha}\mathrm{H}(p_U) - (1-\tilde{\alpha})\mathrm{H}_0 = \tilde{\alpha} \; .
\end{aligned}$$

- The cross entropy α approximates the fidelity $\tilde{\alpha}$.

Numerics and theory for realistic 2D circuits

Cross entropy difference □ and estimated fidelity ∘.



r is two-qubit gate error rate. $\alpha = 1$ for chaotic state. $d = 25$.

Experimental proposal

1. Implement a random universal circuit U (chaotic evolution).
2. Take large sample $S_{\rm exp} = \{x_1^{\rm exp}, \ldots, x_m^{\rm exp}\}$ of bit-strings x in the computational basis ($m \sim 10^3 - 10^6$).
3. Compute quantities $\log p_U(x_j^{\rm exp})$ with supercomputer.

Cross entropy difference (figure of merit)

$$\alpha = \frac{1}{m}\sum_{j=1}^{m} \log p_U(x_j^{\rm exp}) + \log 2^n + \gamma \pm \frac{\kappa}{\sqrt{m}} , \quad \kappa \simeq 1,\ \gamma = 0.577$$

Measure and extrapolate α (size, depth, T gates).
Fit to theory: α approx. circuit fidelity, chaotic state very sensitive to errors.

$$\alpha \approx \exp(-r_1 g_1 - r_2 g_2 - r_{\rm init} n - r_{\rm mes} n) ,$$

$r_1, r_2 \ll 1$ one and two-qubit gates Pauli error rates, $g_1, g_2 \gg 1$ number of one and two-qubit gates, $r_{\rm init}, r_{\rm mes} \ll 1$ initialization and measurement error rates.

Convergence to chaos

Depth required for PT distribution, in 2D is $\propto \sqrt{n}$.

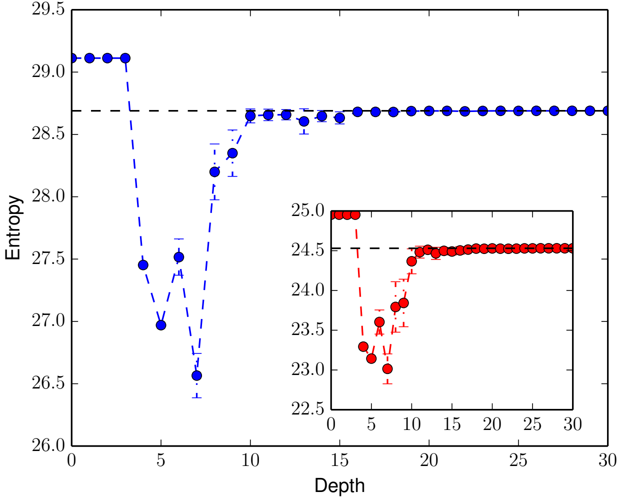


Figure: 2D circuit 7 $\times$ 6 qubits. Inset 6 $\times$ 6 qubits.

Dashed line is known $\mathrm{H}(p_U)$ for PT.

Convergence to chaos (II)

Moments of p_U converge to PT distribution.

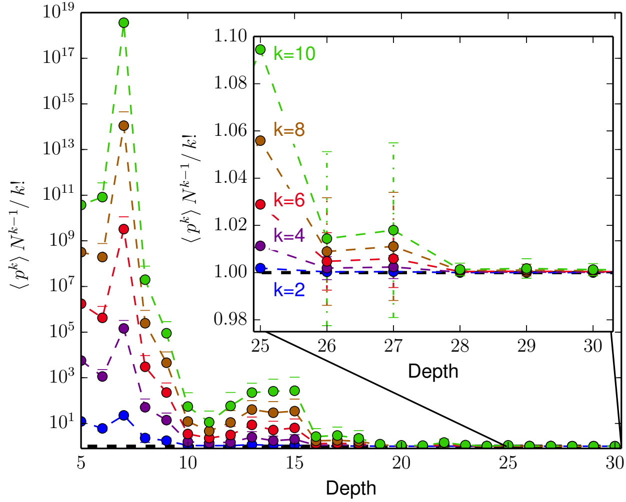


Figure: Moments $\langle p^k \rangle$ with $k = 2, 4, 6, 8, 10$, normalized to 1 for PT distribution. 7×6 circuit.

Convergence to chaos (III)

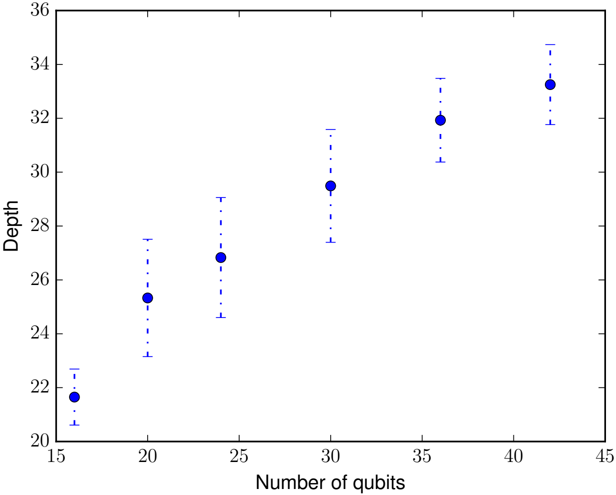


Figure: First cycle such that the entropy remains within $2^{-n/2}$ of PT entropy.

Complex Ising models from universal circuits

- As in a path integral, the output amplitude of U is

$$\langle x|\, U \,|0\rangle = \sum_{\{s^t\}} \prod_{t=0}^{d} \langle s^t|\, U^{(t)} \,|s^{t-1}\rangle \,, \quad |s^d\rangle = |x\rangle \ .$$

  where $|s^t\rangle = \otimes_{j=1}^{n} |s_j^t\rangle$ is the computational basis, $s_j^t = \pm 1$, and $U^{(t)}$ are gates at clock cycle t.
- Gates give Ising couplings between spins s_j^k, like in path integral QMC. For instance, for $\mathrm{X}^{1/2}$ gates

$$\frac{i\pi}{4} H_s^{\mathrm{X}^{1/2}}(x) = \frac{i\pi}{2} \sum_{j=1}^{n} \sum_{k=0}^{d(j)} \alpha_j^k \frac{1 + s_j^{k-1} s_j^k}{2} \ .$$

  where $\alpha_j^k = 1$ denotes that a $\mathrm{X}^{1/2}$ gate was applied at qubit j in (clock cycle) k.

Computational complexity

- For universal circuits, $p_U(x) = \lambda|Z|^2$ is proportional to the partition function $Z = \sum_s e^{i\theta H_x(s)}$ of an Ising model $H_x(s) = h_x \cdot s + s \cdot \hat{J} \cdot s$ with complex temperature $i\theta (= i\pi/8)$ and no structure.
- Z has a strong sign problem: $Z = \sum_j M_j e^{i\theta E_j}$, $|M_j|$ exponentially larger than $|Z|$.
- Worst-case complexity: Z can not be probabilistically approximated asymptotically with an NP-oracle (is #P-hard). (Fujii and Morimae 2013, Goldbert and Guo 2014).
- Computational complexity conjecture: average case = worst case complexity. There is no structure. (Bremner et. al. 2015).
- Theorem: if $p_U(x)$ can be classically sampled, then Z can be approximated with an NP-oracle (Bremner et. al. 2015). Contradiction.
- Classical factoring has no computational complexity implications.

Simulation time

% of comm	# of sockets	# of fused	Avg. time per gate (sec)	Time per Depth-25 (sec)
5×4 *circuit*: 20 qubits, 10.3 gates per level, **17 MB of memory**				
0.0%	1	0.00	0.00015	**0.039**
6×4 *circuit*: 24 qubits, 12.5 gates per level, **268 MB of memory**				
0.0%	1	7.01	0.0041	**1.294**
6×5 *circuit*: 30 qubits, 16.2 gates per level, **17 GB of memory**				
0.0%	1	5.64	0.349	**141.3**
6×6 *circuit*: 36 qubits, 19.5 gates per level, **1 TB of memory**				
6.2%	64	5.40	0.76	**369.0**
7×6 *circuit*: 42 qubits, 23.0 gates per level, **70 TB of memory**				
11.2%	4,096	5.54	1.72	**989.0**

On Edison, a Cray XC30 with 5,576 nodes. Each node is dual-socket Intel®Xeon E5 2695-V2 with 12 cores per socket, 2.4GHz. 64GB per node (32GB per socket). Nodes connected via Cray Aries with Dragonfly topology. (Mikhail Smelyanskiy).

Conclusions

- We expect to be able to approximately sample the output distribution of shallow random circuits of 7×7 qubits with significant fidelity in the near term.
- It is impossible to approximately sample the output distribution of shallow random quantum circuits of ≈ 48 qubits with state-of-the-art supercomputers ($d \sim 25$).
- Quantum supremacy.
- New method to benchmark complex quantum circuits efficiently.
- Relation to quantum chaos.
- Relation to computational complexity.
- The theory applies to other chaotic systems: chaotic Hamiltonians, commuting quantum circuits, BosonSampling.
- The cross entropy method applies to other sampling problems.