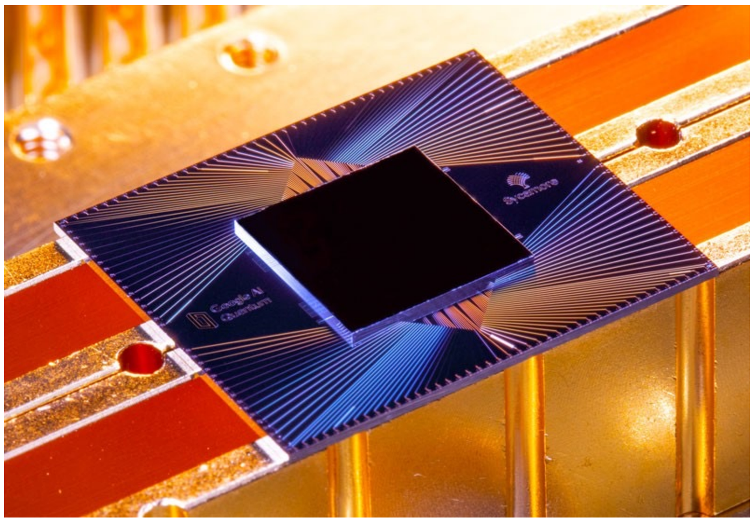
On the power of random quantum circuits

Bill Fefferman

Entangle This: Randomness, Complexity and Quantum Circuits, Benasque, Spain

The first “Quantum advantage” claims have now been made…



Random Circuit Sampling (Google “Sycamore”) in 2019, 2023

Gaussian BosonSampling (USTC, Xanadu) in 2021,2022,2023

This talk: the latest complexity theoretic arguments to understand these “random quantum circuit” experiments

What is the *ideal* goal of quantum advantage?

- Find a problem:

1. Can be solved using a near-term quantum experiment
2. Is classically hard to solve – can't be solved by any classical algorithm in polynomial time
3. Solution can be efficiently verified with a classical computer with minimal trust in the experiment

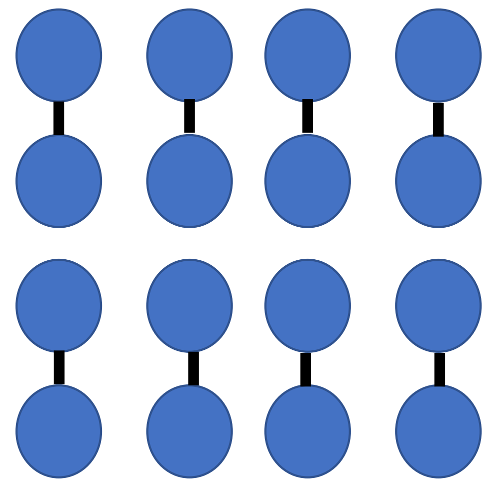
What is the *status quo*?

- Current quantum advantage experiments solve "sampling problems" in which the goal is to sample from a complicated distribution
- We'll discuss *evidence* that these problems cannot be solved classically in polynomial time
- But current experiments *are not scalable*
    1. The verification is inefficient
    2. Noise causes the signal to rapidly decay
- These issues force current quantum advantage candidates to find "Goldilocks" parameter regimes
    - *Is this inevitable?* Can classical hardness in RCS survive the noise asymptotically?

Goldilocks and the three bears

What is Random Circuit Sampling? [e.g., Boixo et. al. 2017]

- Generate a quantum circuit C on n qubits on a 2D lattice, with d layers of (Haar) random nearest-neighbor gates
  - In practice use a discrete approximation to the Haar random distribution
- Start with $|0^n\rangle$ input state, apply random quantum circuit and measure all qubits in computational basis
  - i.e., Sample from a distribution D_C over $\{0,1\}^n$
- Has now been implemented:
  - n = 53 qubits, d = 20 [Google, 2019]
  - n = 60 qubits, d = 24 [USTC, 2021]
  - n = 70 qubits, d = 24 [Google, 2023]



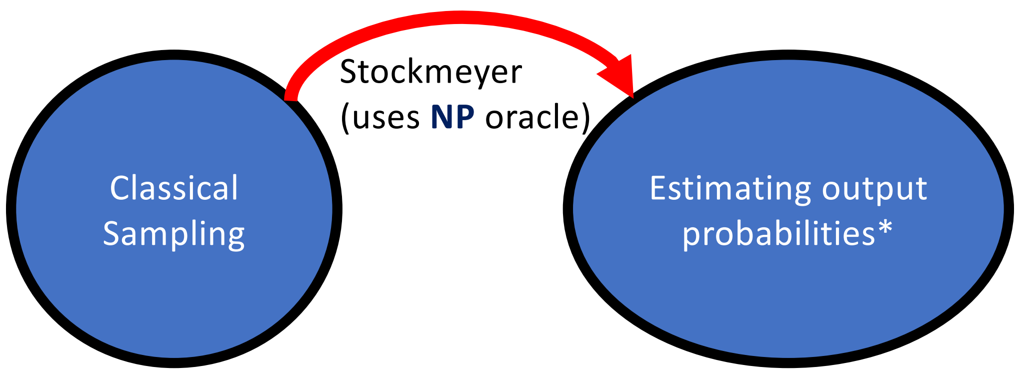
(single layer of Haar random two qubit gates applied on 2D grid of qubits)

Why should RCS be classically hard?

- **First goal:** prove impossibility of an efficient *"classical Sampler"* algorithm that:
  - takes as input a random circuit C
  - outputs a sample from D_C whp over C
- Here we **aren't modelling physical noise**, we're just asking if there's a hard to simulate quantum signal in the *ideal* case

Proof first step: from sampling to computing

- By a well-known reduction [Stockmeyer '85] it suffices to prove that ***estimating*** the output probability of a ***random*** quantum circuit is **#P**-hard
  - i.e., need to prove hardness of estimating $p_{0^n} = |\langle 0^n | C | 0^n \rangle|^2$ within additive error $O(2^{-n})$ wp $2/3$



Inspiration: average-case hardness of Permanent [Lipton '91]

- **Permanent** of $n \times n$ matrix is **#P**-hard in the worst-case [Valiant '79]
  - $Per[X] = \sum_{\sigma \in S_n} \prod_{i=1}^{n} X_{i,\sigma(i)}$
- *Algebraic property*: $Per[X]$ is a degree n polynomial with n^2 variables
- Need to compute $Per[X]$ of worst-case matrix X over $\mathbb{F}_p$
  - But we only have access to algorithm O that correctly computes *most* permanents
  - i.e., $\Pr_{Y \in_R \mathbb{F}_p^{n \times n}}\left[O(Y) = Per[Y]\right] \geq 1 - \frac{1}{poly(n)}$
- Choose $n+1$ fixed non-zero points $t_1, t_2 \ldots, t_{n+1} \in \mathbb{F}_p$ and uniformly random matrix R
- Consider line $A(t) = X + tR$
  - *Observation 1 "scrambling property"*: for each i, $A(t_i)$ is a random matrix over $\mathbb{F}_p^{n \times n}$
  - *Observation 2*: "univariate polynomial": $Per[A(t)]$ is a degree n polynomial in t
- But now these $n+1$ points uniquely determine the polynomial, so use polynomial extrapolation to evaluate $Per[A(0)] = Per[X]$

[BFNV'18]: Hardness for Random Quantum Circuits

- *Algebraic property*: much like $Per[X]$, output probability of random quantum circuits has polynomial structure
  - Consider circuit $C = C_m C_{m-1} \dots C_1$
  - Polynomial structure comes from path integral:
    - $\langle 0^n|C|0^n\rangle = \sum_{y_2,y_3,\dots,y_m\in\{0,1\}^n}\langle 0^n|C_m|y_m\rangle\langle y_m|C_{m-1}|y_{m-1}\rangle \dots \langle y_2|C_1|0^n\rangle$
- This is a polynomial of degree m in the gate entries of the circuit
- So the output probability $p_{0^n}(C)$ is a polynomial of degree $2m$

How to "scramble" worst-case circuit, $\boldsymbol{C}$?

- Fix m Haar random two qubit gates $\{H_i\}_{i\in[m]}$
- ***Main idea***: Implement tiny fraction of H_i^{-1}
  - i.e., each $C_i' = C_i H_i e^{-ih_i\theta}$
  - This scrambles C if $\theta \approx small$, since each gate is close to Haar random
  - However, if $\theta = 1$ the corresponding circuit $C' = C$
- ***Strategy (in style of Lipton):*** take several non-zero but small θ, for each angle we have "random but correlated" circuit $C'_{\theta_1}, C'_{\theta_2} \ldots, C'_{\theta_{2m}}$ then compute output probabilities and apply polynomial extrapolation, evaluate at $\theta = 1$ to retrieve $p_{0^n}(C)$

This is not quite the “right way” to scramble!

- ***Problem***: $e^{-ih_i\theta}$ is not polynomial in θ
- ***Solution:*** take fixed truncation of Taylor series for $e^{-ih_i\theta}$
  - i.e., each gate of C'_θ is $C_i H_i \sum_{k=0}^{K} \frac{(-ih_i\theta)^k}{k!}$
  - So each gate entry is a polynomial in θ and so is $p_{0^n}(C'_\theta)$
  - Now extrapolate and compute $p(1) = p_{0^n}(C)$

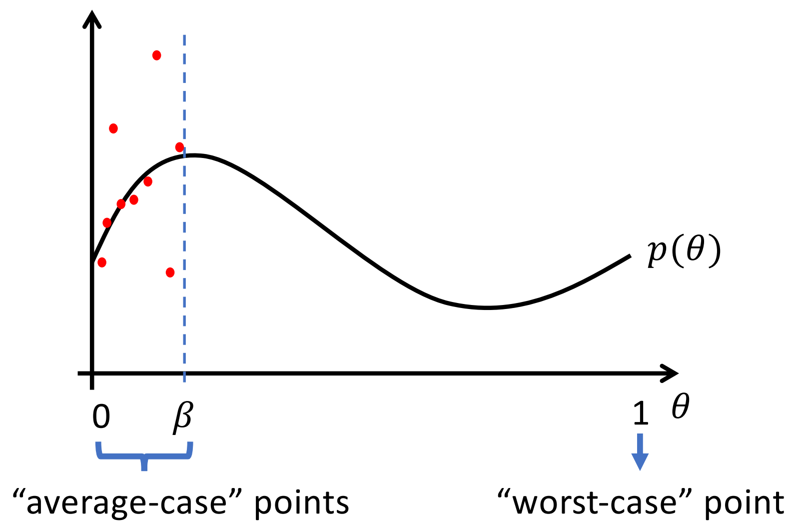
Subtleties in this argument

Truncations make the distribution supported on circuits that are ***slightly non-unitary!***

- [BFNV’18] addressed this by proving that **estimating** the *truncated* random circuit probability is hard iff **estimating** the *unitary* random circuit probability is hard
- See also follow-up work which gets rid of these truncations entirely [Movassagh’19’20]

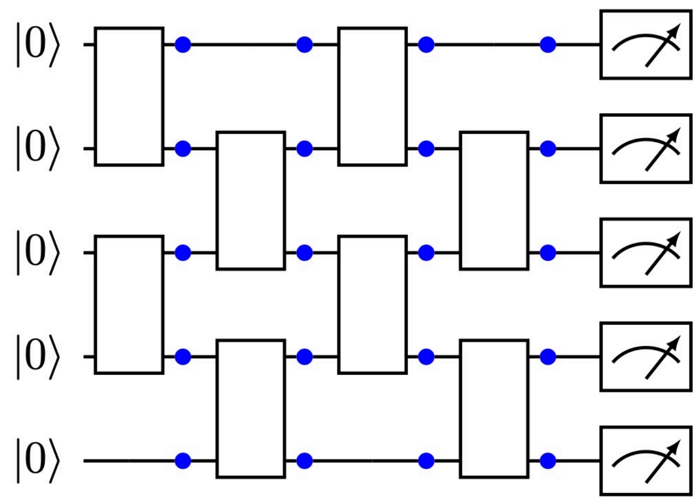
On robustness to *imprecision*

- So far we assumed the ability to compute the output probabilities of random circuits $\{p_{0^n}(C'_{\theta_i})\}$ *exactly*
- **Actual setting:** Given faulty evaluation points $\{(\theta_i, y_i)\}$ so that for *most* i:
  - $|y_i - p_{0^n}(C'_{\theta_i})| \leq \delta$
  - There's "ideal" polynomial $p(\theta_i) = p_{0^n}(C'_{\theta_i})$ of degree m and need an estimate for $p(1)$
- **State of the art [BFLL'21, KMM'21]:** There's an algorithm (uses **NP** oracle) that outputs a polynomial $q(\theta)$ so that:
  - $|q(1) - p(1)| \leq \delta 2^{m \log m}$
- $\Rightarrow$ need $\delta \sim 2^{-O(m \log m)}$
- (for BosonSampling: have hardness at $\frac{1}{e^{6n \log n}}$ but we need $\frac{1}{e^{n \log n}}$ [BFLL'21])



Does the “quantum signal” survive uncorrected noise?

- Noise is overwhelming in near-term experiments
  - e.g., Google RCS: ~0.2% signal, 99.8% noise
- How to theoretically model this? First, consider just single qubit depolarizing – i.e., each layer random gates followed by:
  - $\mathcal{E}(\rho) = (1-\gamma)\rho + \frac{\gamma I}{2} Tr[\rho]$
  - Where the noise strength, γ is positive constant
  - This is a popular model, but oversimplified!



Depolarizing noise and complexity

- Intuitively, uncorrected depolarizing noise increases entropy. As the circuit gets deeper the output distribution converges to uniform
- **First question:** how close are the output distribution of noisy (i.e., depolarizing) random circuit and uniform distribution?
  - $2^{-\Theta(d)}$ close in TVD [Aharonov et. al. '96][Deshpande et. al.'22]
- This rules out scalable noisy quantum advantage at *super-logarithmic depth*

What about noisy *shallow* circuits?

- If depth is at most $\log(n)$ then output distribution is **far from uniform**
- [Aharonov et. al. '22] give a classical algorithm for sampling from the output distribution of noisy, $\log(n)$ depth random quantum circuits
- **Idea:** Write noisy output probabilities as path integral in the Pauli basis
  - i.e., as $\tilde{p}_x = \sum_{s \in P_n^{d+1}} (1-\gamma)^{|s|} f(C, s, x)$
  - Where $|s|$ is the "weight of the path", i.e., the number of non-identity operators
- **Key point:** output probabilities of noisy circuit in Pauli basis are *exponentially suppressed* in weight of path
- **Classical algorithm:** throw away paths with sufficiently high Pauli weight

Analysis of this algorithm uses *anti-concentration*

- Bounding approximation error relies on "anti-concentration" property
  - i.e., Output distribution of random circuit is well-spread over outcomes
  - **Formal:** for any outcome $x \in \{0,1\}^n$ there exists constants $\alpha \in (0,1], c > 0$ so that $\Pr_C \left[p_x(C) \geq \frac{\alpha}{2^n} \right] \geq c$
- Anti-concentration is a property of *sufficiently deep* random quantum circuits
  - For noiseless circuits, or for circuits with depolarizing noise, at least $\log(n)$ depth is known to be necessary and sufficient [Dalzell et. al. '20] [Deshpande et. al. '22]

Adapting [Aharonov et. al. ’22] to other noise

- For BosonSampling with “Gaussian” noise, we show that a similar classical algorithm **works** [Oh et. al., ‘23] building on [Kalai & Kindler ‘14]
  - Gaussian noise means $U \rightarrow \sqrt{\gamma}\, U + \sqrt{1-\gamma}\, G$ where G is Gaussian matrix
  - But we don’t know how to make this work for other noise models e.g., photon loss
- Anticoncentration **fails** for random circuits with depolarizing noise **together with** many non-unital noise channels, at any depth [Ghosh et. al., unpublished]
  - e.g., Amplitude damping channel: $K_0 = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-\gamma} \end{pmatrix}, K_1 = \begin{pmatrix} 0 & \sqrt{\gamma} \\ 0 & 0 \end{pmatrix}$
  - *So in these cases we know neither hardness, nor easiness!*

Open questions

- Can we prove hardness of sampling from random circuits in the *noiseless* case? (i.e., involves improving robustness of hardness results)
- How hard are random quantum circuits with “low noise” i.e., $\gamma = O\left(\frac{1}{n}\right)$
  - Motivated by recent results showing that in this regime cross-entropy approximates fidelity [Google group ‘23]
- Can we find better RCS verification protocols?

Thanks!