

“Quantum Supremacy” and the Complexity of Random Circuit Sampling

Bill Fefferman (UC Berkeley and NIST)

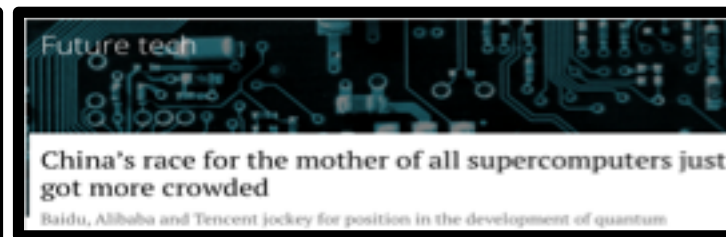
Joint with Adam Bouland, Chinmay Nirkhe, and Umesh Vazirani

arXiv: 1803.04402



Interest in quantum computing is exploding

- Huge media hype:



- Enormous industry interest for quantum computing
 - Google, Microsoft, IBM, Intel, Alibaba, Baidu, ...
- Even new start-ups “bridging” industry and academia
 - IonQ/UMD, Quantum Circuits Inc/Yale, Zapata/Harvard...

Why is there so much excitement?

- Experimental advances are leading to the “Noisy Intermediate Scale Quantum” (NISQ) era
 - Several experimental groups will soon implement noisy 50-70 qubit systems [e.g., Google/UCSB, IBM, U. Maryland]
 - No error-correction – can’t run many quantum algorithms
 - But also too big to be *naively* simulated by classical computers
- This has greatly accelerated the development of full-scale quantum computing
- *Major challenge is algorithmic*

A first step: “Quantum Supremacy”

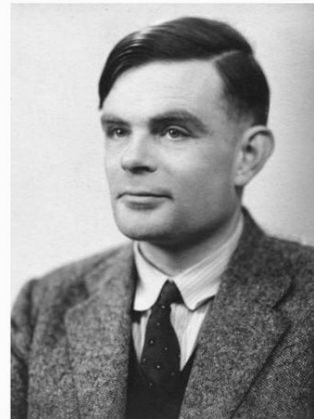
- **Quantum Supremacy:** A practical demonstration of a quantum computation that is prohibitively hard for classical computers
 - Needs to be experimentally feasible
 - Need theoretical evidence for hardness (i.e., problem couldn't be solved efficiently on classical computer)
 - Like early quantum algorithms, no need to be useful!
- Stepping stone to scalable, fault tolerant, universal quantum computers
- But it's much more than that!

Importance of quantum supremacy: foundations of computation

- *Experimental* violation of the Extended Church-Turing thesis
 - i.e., If we want to model computation, we must consider quantum mechanics!
- Complements *theoretical* evidence given by earlier speedups (e.g., [Bernstein-Vazirani '93][Simon'94][Shor '94])



Alonzo Church



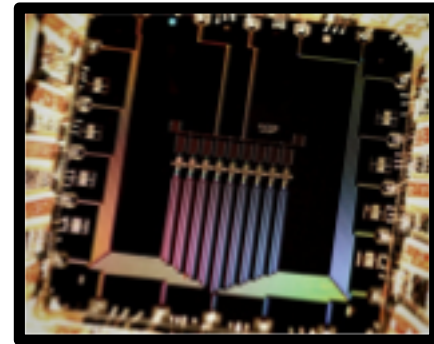
Alan Turing

Importance of quantum supremacy: validating quantum physics

- Exponential growth arguably the most counter-intuitive aspect of quantum mechanics.
 - Is the exponential description of a quantum state really necessary?
- New limit in which to test physics: **high complexity**.
- *Difficulty*: how to verify something that's exponentially complex?

Importance of quantum supremacy: validating near term devices

- Quantum supremacy: necessary to have a large quantity of high quality qubits
 - Achieving both is quite difficult experimentally
- In recent years, tools from quantum supremacy have become more and more central to experimental efforts in validating NISQ devices
 - E.g., to “tuning qubits” and “diagnosing errors”



Martinis group: Google/UCSB

Existing quantum supremacy proposals

Broadly speaking, fall into two categories:

1. *Theoretically driven proposals*

- Special purpose devices with good evidence for hardness
- Are not yet experimentally feasible at sufficiently large scale
- e.g., BosonSampling [Aaronson & Arkhipov '11]

2. Experimentally driven proposals

- Will be realizable in the near term, on the path to scalable quantum computing
- But do not yet have as strong theoretical evidence of hardness
- e.g., Random Circuit Sampling proposal of the Google/UCSB group [Boixo et. al. '16]

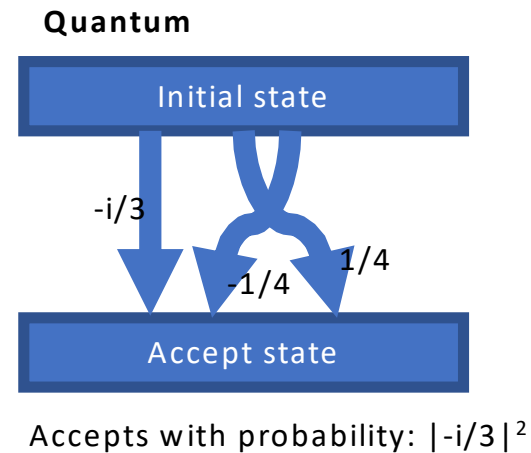
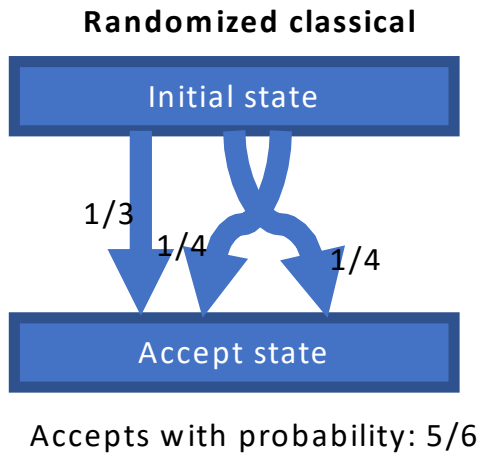
Our results

1. We provide ***strong theoretical backing*** to the leading *experimental* candidate for quantum supremacy: Random Circuit Sampling [Google/UCSB group: Boixo et al '16]
2. We study verification, clarifying when existing proposals work to ***verify these devices***, as well as give new verification measure which is in some sense “optimal”

1. Quantum supremacy from average-case interference patterns

Quantum computing and *interference*

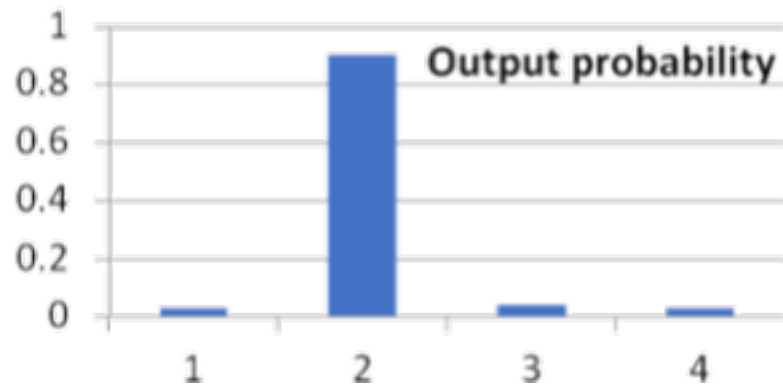
- Fundamental difference between quantum and randomized classical computation: *interference*



- This exponential cancellation is *critical ingredient* for quantum speedups

Quantum algorithms from the 90's are hard to implement!

- The speedups come from carefully engineered interference patterns with large amounts of constructive and destructive interference

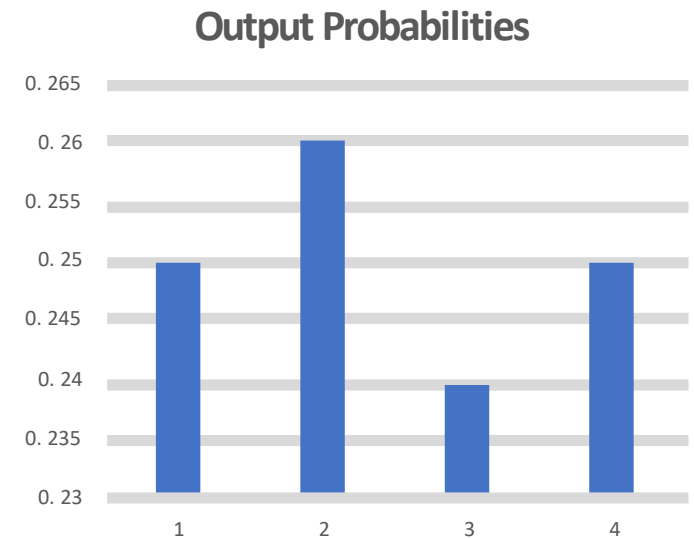


This behavior is far from “typical”.
It’s hard to make this happen in the lab!!

“Proving a quantum system’s computational power by having it factor integers is a bit like proving a dolphin’s intelligence by teaching it to solve arithmetic problems” [Aaronson & Arkhipov ‘11]

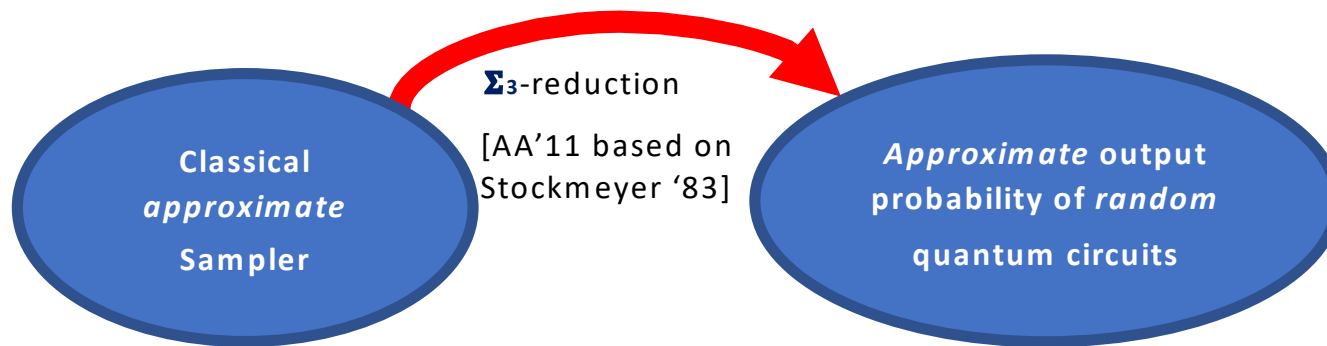
NISQ interference patterns are generic!

- Random, *average-case* interference patterns
- *Supremacy proposal*: Given random quantum circuit, sample from distribution *close* to its ideal output distribution
- ***Our question***: How hard is this problem for classical computers?



How to prove classical hardness of sampling?

- **Key point:**



Our question: How hard is *approximating* output probabilities of *random* quantum circuits?

If it's **#P**-hard then classical sampler implies **PH** collapse (by Toda's theorem)

BosonSampling [Aaronson & Arkhipov'11]

- **Task:** sample from the outcome distribution of a random linear optical quantum circuit
- **Key point:** Output probabilities of random linear optical circuits are *permanents* of random matrices
- Permanent has a **worst-to-average** case reduction, and so is **#P**-hard to compute permanent of **most** random matrices [AA'11, building on Lipton '91]
- **Open question:** extend *exact* average-case hardness to *approximate* average-case hardness: *Is it #P-hard to approximate most random matrices?*

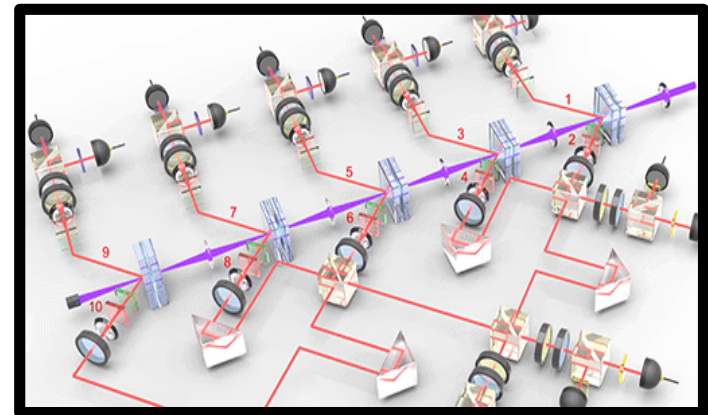


Photo credit: X.-L. Wang *et al.* (2016)

But BosonSampling seems hard to experimentally implement...

- *Yet to see sufficiently large experiments to test Extended Church Turing thesis*
- *Current experiments: 6 photon experiments [O'Brien et. al., '16]*
- Recent classical simulation algorithms indicate need ~50 photons [e.g., Clifford & Clifford '17, Neville et. al.'17]
- Further, it's a special purpose device– not necessarily on path to universal scalable quantum computation

Random Circuit Sampling [e.g., Boixo et. al., '16]

- **Task:** sample from the output distribution of a random quantum circuit
 - Generate a quantum circuit C on n qubits on a 2D lattice, with $d=O(n)$ layers of Haar random nearest-neighbor gates
 - Start with $|0^n\rangle$ input state and measure in computational basis
- *Experimentally compelling:* 72? qubits coming soon [Google/UCSB]
- **RCS Conjecture:** **#P**-hard to *estimate* output probability of *most* random quantum circuit
 - But unlike BosonSampling, no connection to permanents
 - **Missing:** average-case hardness!

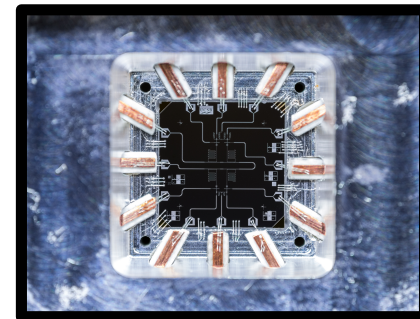
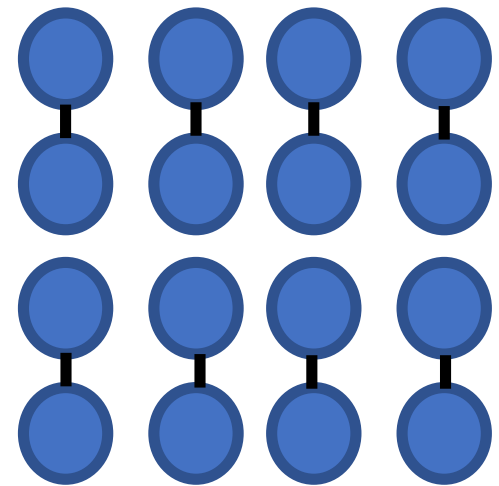


Photo Credit:
Michael Fang

2. *Main result: Average-case hardness for RCS*

Average-case hardness for permanent of matrices over finite fields [Lipton '91]

- **Permanent** of $n \times n$ matrix is (worst-case) **#P**-hard [Valiant '79]

$$\text{per}[X] = \sum_{\sigma \in S_n} \prod_{i=1}^n X_{i,\sigma(i)}$$

- *Algebraic property*: **permanent** is a degree n polynomial on n^2 variables
- Lipton shows “worst-to-average case reduction”
 - Need compute **permanent** of worst-case matrix X
 - But we only have access to algorithm that correctly computes *most* permanents
- Choose $n+1$ fixed non-zero points $t_1, t_2, \dots, t_{n+1} \in \mathbb{F}_q$ and uniformly random matrix R
- Consider line $A(t) = X + tR$
 - *Observation 1 “marginal property”*: for each i , $A(t_i)$ is a random matrix over \mathbb{F}_q
 - *Observation 2: “univariate polynomial”*: $\text{per}[A(t)]$ is a degree n polynomial in t
- But now these $n+1$ evaluation points uniquely define the polynomial, so use error-correction (noisy polynomial interpolation) and evaluate $\text{per}[A(0)] = \text{per}[X]$

RCS also has polynomial structure

- The polynomial structure comes from the Feynman path integral:
 - Consider circuit $C=C_m C_{m-1} \dots C_1$

$$\langle 0^n | C | 0^n \rangle = \sum_{y_2, y_3, \dots, y_m \in \{0,1\}^n} \langle 0^n | C_m | y_m \rangle \langle y_m | C_{m-1} | y_{m-1} \rangle \dots \langle y_2 | C_1 | 0^n \rangle$$

- This is polynomial of degree m in the gate entries of the circuit
- Output probability is polynomial of degree $2m$

Worst-to-Average Reduction-Attempt 1: Copy Lipton's proof

- Our case: want to compute $|\langle 0^n | C | 0^n \rangle|^2$ for worst case C
 - But we only have the ability to compute output probability of *most* circuits
- *Recall*: Lipton wanted to compute $\text{per}[X]$, choose random R , considered line $A(t) = X + tR$
- *Problem*: can't just perturb gates in a random linear direction (quantum gates don't have linear structure... i.e., if A is unitary, B is unitary, $A + tB$ is not necessary unitary)

New approach to *scramble* gates of fixed circuit

- Choose $\{H_i\}_{i \in [m]}$ Haar random gates
- Now consider new circuit $C' = C'_m C'_{m-1} \dots C'_1$ so that for each gate $C'_i = C_i H_i$
 - Notice that each gate in C' is completely random – “marginal property”
- But no univariate polynomial structure connects worst-case circuit C with the new circuit C' !!

Correlating via quantumness

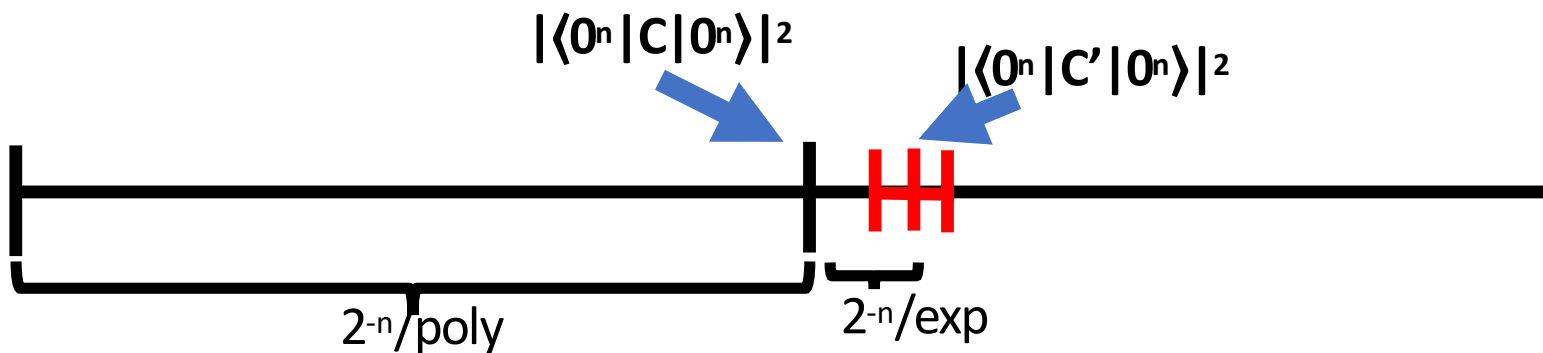
- We need the analogue to Lipton's "*univariate* polynomial structure"
- ***Uniquely quantum capability***: Implementing a fraction of a gate!
 - i.e., if G is a quantum gate then $G=e^{ig}$ where $g=\log G$
 - Can implement $e^{ig\theta}$ where θ is a small angle
- *Main idea*: "Implement tiny fraction of H_i^{-1} " (i.e., $C'_i=C_i H_i e^{-ih_i\theta}$)
 - If $\theta=1$ the corresponding circuit $C'=C$, and if $\theta \approx$ small, each gate is close to Haar random
 - Now take several non-zero but small θ and apply polynomial interpolation...

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

- *Problem:* rotation back by angle θ isn't low-degree polynomial in θ
 - Because $e^{-ih_i\theta}$ is not polynomial in θ
- *Solution:* take fixed truncation of Taylor series for $e^{-ih_i\theta}$
 - i.e., instead of $C'_i = C_i H_i e^{-ih_i\theta}$ each gate is now $C'_i = C_i H_i \sum_{k=1}^K \frac{(-ih_j \theta)^k}{k!}$
 - So each gate entry is a polynomial in θ and so is $|\langle \mathbf{0}^n | \mathbf{C} | \mathbf{0}^n \rangle|^2$ by path integral
 - Now interpolate and compute $p(1) = |\langle \mathbf{0}^n | \mathbf{C} | \mathbf{0}^n \rangle|^2$

This shows average-case exact hardness for a (very slightly) different circuit distribution!

- But recall, the real question is about hardness of *estimating* a typical circuit amplitude
- We show that approximate hardness over this “truncated” circuit distribution is **equivalent** to the approximate RCS hardness conjecture



3. Using statistical tests to verify RCS

Verifying RCS in the NISQ era

- **Challenge:** Need to develop a statistical measure to verify the RCS output distribution from samples of device, but...
 - **Constraint 1:** don't know the output distribution (only given a description of circuit)
 - **Constraint 2:** can only take a small ($\text{poly}(n)$) number of samples from the quantum device
- **Compromise:** OK to use exponential postprocessing time on supercomputer to compute "a few" ideal output probabilities (doable for $n=49$ qubits)
- Complexity arguments require closeness in total variation distance. **But we can't hope to unconditionally verify this with few samples from the device.**

A candidate test for verifying RCS: cross-entropy [Boixo et. al., 16]

- Proposal is to compute:

$$CE(p_{dev}, p_{id}) = \sum_x p_{dev}(x) \log \frac{1}{p_{id}(x)} = \mathbb{E}_{p_{dev}} \log \left(\frac{1}{p_{id}} \right)$$

- This can be well-approximated in few samples using concentration of measure arguments!
- Then accept if score is sufficiently close to the expected ideal cross-entropy, which can be calculated analytically

Why might one believe this verifies RCS?

- This is a “one-dimensional projection” of observed data
- Does not verify closeness in total variation distance directly
- (Theorem: exist distributions score well on CE but are far in total variation)
- [Boixo et al. '16]: Assume that

$$\rho_{\text{dev}} = \alpha \rho_{\text{id}} + (1-\alpha) I_d$$

In this case, achieving near-perfect cross-entropy certifies closeness in total variation distance

Deeper reasons to believe in Cross-Entropy?

Claim: If Cross-Entropy is close to ideal **and** $H(\mathbf{p}_{dev}) \geq H(\mathbf{p}_{id})$, then the output distribution is close to ideal in total variation distance

This assumption would follow from certain noise models (e.g., local depolarizing noise) but not from others (e.g., correlated noise, erasure channel etc...)

Proof:

- Pinsker's inequality: $|p_{dev} - p_{id}|_{TV} \leq \sqrt{\frac{1}{2}|p_{dev} - p_{id}|_{KL}}$
 - Where $|p_{dev} - p_{id}|_{KL} = CE(\mathbf{p}_{dev}, \mathbf{p}_{id}) - H(\mathbf{p}_{dev})$
- So if we find cross-entropy ε -close to ideal, we've certified closeness in total variation distance to error $O(\varepsilon^{1/2})$

More on Cross-Entropy

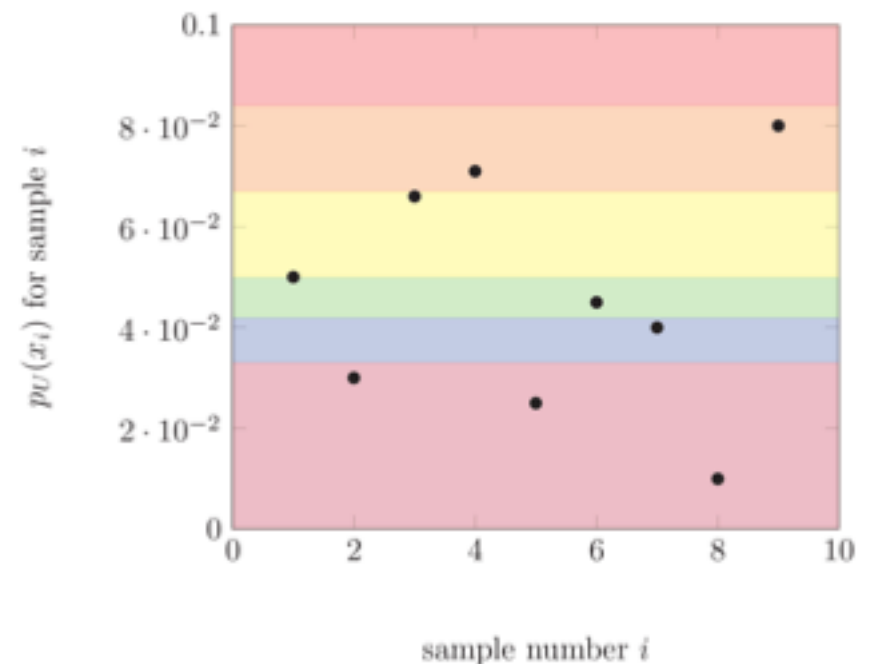
- The output distributions of RCS are “Porter-Thomas”
 - $\Pr[p_x = q/N] = e^{-q}$
- This is **not** a signature of quantum effects
 - Can be reproduced classically (e.g., by Poisson processes)
- However, pairs of distributions scoring highly on CE test share similar “heavy” outcomes
 - This intuition was sharpened by a recent proposal of Aaronson & Chen called “HOG”
$$\mathbb{E}_{p_{dev}} \delta(p_{id} \text{ is “heavier than median”})$$
- Scoring above some threshold conjectured to be intrinsically hard
 - Can be connected to nonstandard complexity assumptions [AC’16]
 - But don’t know how to connect to well-studied complexity assumptions

An alternative verification method:

- To compute cross-entropy or HOG, you take a lot of data, and reduce it to a single number, why not use more of it?
- We show that you can create a verification procedure with the same data, which verifies both cross-entropy and HOG – so inherits advantages of both

Introducing...Binned Output Generation (BOG)

- Consider dividing the $[0,1]$ interval into **poly(n)** bins
- Observe k samples x_1, x_2, \dots, x_k and calculate ideal probabilities for each sample on supercomputer
- Accept if the number of outcome probabilities in each bin are approximately equal to expected frequency in each bin (from P-T)



Conclusions

- Average case hardness gives evidence that circuit sampling hard even for random circuits which exhibit generic interference patterns.
- For sufficiently small supremacy experiments we can verify supremacy if we make strong enough assumptions about the device output distribution: e.g., $H(\text{dev}) \geq H(\text{ideal})$.

Open Questions

Missing piece: extend hardness of *exactly* computing output probability of a random quantum circuit to *approximate* case (this is open for *all* supremacy proposals!)

At what system size should we conclude “quantum supremacy”? What is the importance of implementing a particular system size, like 49 qubits?

Can recent classical heuristics for RCS simulation, such as those of the Alibaba group [Chen et. al., '18] be used to verify RCS experiments for larger system sizes?

Many other proposals may be relevant for sufficiently large system size (see e.g., [Brakerski, Christiano, Mahadev, Vazirani, Vidick][Aaronson]...)

Thanks!