“Quantum Supremacy” and the Complexity of Random Circuit Sampling

Bill Fefferman

UC Berkeley/NIST

Joint with Adam Bouland, Chinmay Nirkhe, and Umesh Vazirani

Quantum mechanics challenges the *foundations* of computation

- *Extended Church-Turing thesis*: everything feasibly computable in the physical world is efficiently computable by a classical Turing machine
- Early 1990's – first evidence that ideal quantum computers violate thesis
  - Initial results didn't solve "natural" problems
    - Bernstein-Vazirani [BV'93]
    - Simon's algorithm [Simon'94]
  - Closely followed by Shor's algorithm [Shor'94], solving the factoring problem
- In all cases, these speedups come from carefully engineered algorithms that exploit "particular interference patterns"
  - "*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]

“Quantum Supremacy”: A demonstration of a quantum computation that is prohibitively hard for classical computers

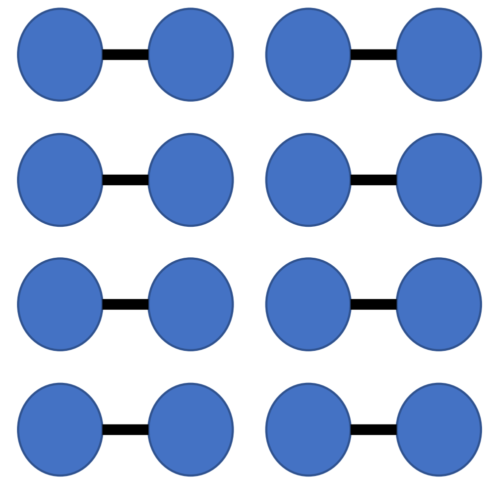
- In the last decade, two concurrent developments:
  1. Theoretical advances in our understanding of restrictive, special-purpose quantum devices like BosonSampling [Aaronson & Arkhipov ‘11]
  2. Experimental advances have led to the “Noisy Intermediate Scale Quantum” era
     - Several experimental groups will soon implement noisy 50-70 qubit systems without error-correction [e.g., Google/UCSB, IBM, U. Maryland]
     - Will not be able demonstrate idealistic speedups from the 90’s
     - At the same time, these systems are too large to be simulated by brute-force classically
- **Major goal**: Combine 1 and 2 to prove that a NISQ era experiment cannot be simulated by *any* classical means – disprove the ECT thesis and validate quantum mechanics in a realm of “high complexity”

BosonSampling

- *Supremacy proposal*: sample from the output distribution of a linear optics experiment
  - Sampling problems are natural since “raw output” of a quantum computer is a sample from an outcome distribution generated by quantum measurement
  - “*If we just watched the dolphin in its natural habitat…it displays equal intelligence with no special training at all*” [Aaronson & Arkhipov ‘11]
- *Theoretically compelling*
  - Linear optical output probabilities are proportional to the “permanent” of random matrices
  - Permanents have “average-case hardness”
    - Allows us to base hardness on a *typical* rather than worst-case experiment
- Yet to see sufficiently large experiments to test ECT
  - Recent classical simulation algorithms indicate need ~50 photons, 50^2 modes [e.g., Clifford & Clifford ‘17, Neville et. al.’17]
  - Recent experiments ~6 photons and 13 modes

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

- *Supremacy proposal*: 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 $|0^n\rangle$ basis state and measure in computational basis
- E*xperimentally compelling*
  - ~49 qubits in the next few months with controllable couplings [Google/UCSB]
- Google/UCSB conjecture this sampling task is classically hard
  - *Challenge*: RCS is different from prior proposals
    - Unlike the carefully engineered speedups of the ‘90s, will have to argue hardness of typical quantum systems with “generic” interference patterns
    - Unlike BosonSampling, no known average-case hardness for RCS
  - Without average-case hardness, there’s no evidence that generic interference patterns cannot be reproduced classically
- ***Main result***: Provide a theoretical foundation backing RCS
  - Prove average-case hardness for RCS: computing output probabilities for most random circuits is as hard as computing them in the worst-case
  - Proof crucially uses error-correcting codes to infer worst-case probabilities from typical output probabilities



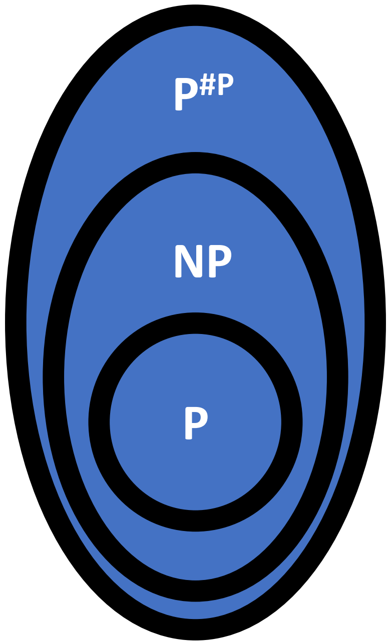
Overview

- Our focus is on two perspectives
    1. Establishing hardness using complexity theory
    2. Verification of supremacy via statistical tests

1. Establishing hardness using complexity theory

Complexity theory basics

- **P**: Class of problems feasible for classical computer
- **NP**: Characterized by **SAT** problem: given boolean formula is it satisfiable?
- **#P**: Generalization of **NP** to *counting* the number of satisfying assignments to boolean formula



The classical hardness of quantum sampling

- ***Premise***: given quantum circuit as input, exactly computing any particular outcome probability is **#P**-hard
  - But these probabilities are exponentially small and cannot be directly estimated
- However, this fact can be leveraged to prove that such quantum outcome distributions cannot be classically sampled *exactly* [e.g., Terhal & DiVincenzo'04, Bremner, Jozsa & Shepherd'10…]
- ***Key challenge*****:** extend hardness of *exact* sampling results to hold in the presence of experimental noise, modelled by closeness in total variation distance
  - Suffices to prove that it is **#P**-hard to *additively estimate most* quantum outcome probabilities

BosonSampling hardness

BosonSampling Conjecture: Efficiently *approximating* 1-δ fraction of the outcome probabilities of *typical* linear optical networks to within additive error ±ε/M in is **#P**-hard

- Evidence
  1. *Average-case exact hardness* (i.e., ε=0 case)
  2. *Anti-concentration conjecture* (unproven for BosonSampling)
     - In a *typical* network, most outcome probabilities are *reasonably large*
     - “Sanity check”
       - “Signal is larger than the noise”: ±ε/M additive estimate can be used to recover (1 ± ε’) multiplicative estimates
     - Computing such multiplicative estimates on *all* outcomes (i.e., δ=0 case) is **#P**-hard
- But general case, where both ε,δ>0 is open, and defies all proof techniques

Proposal	Average-case exact (ε=0)	Worst-case mult. Approximate (δ=0)	Anti-concentration	General (ε,δ>0)
BosonSampling	**#P**-hard	**#P**-hard	?	?

Random Circuit Sampling hardness

RCS Conjecture: Given as input random n qubit quantum circuit **C**, outputting an efficient additive estimate $\alpha \in |\langle y|C|0^n\rangle|^2 \pm \epsilon/2^n$ with probability $1-\delta$ (over choice of C and outcome y) is **#P**-hard

Proposal	Average-case exact (ϵ=0)	Worst-case Mult. Approximate (δ=0)	Anti-concentration	General (ϵ,δ>0)
BosonSampling	**#P**-hard	**#P**-hard	?	?
RCS wrt 2D grid, depth O(n)	**?** ✔ **(Today!)**	**#P**-hard	Yes (e.g., [BH '13])	?

- But what good is anti-concentration without average-case hardness?
  - Anti-concentration tells us most output probabilities are somewhat large
    - For these large probabilities, an additive estimate suffices to prove a multiplicative estimate
  - So this only allows us to compute multiplicative estimates on *average*

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

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

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

- *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 O that correctly computes *most* permanents
    - i.e., $\Pr_{Y \sim U_{\mathbb{F}_q}^{n \times n}}[O(Y) = \mathbf{per}[Y]] \geq 1 - \frac{1}{3(n+1)}$
- Choose n+1 fixed non-zero points $t_1, t_2, \ldots, t_{n+1} \in \mathbb{F}_q$ 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 $\mathbb{F}_q^{n \times n}$
  - *Observation 2*: "univariate polynomial": **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 **per**[A(0)]=**per**[X]

Main result: Worst-to-average-case RCS reduction

- *Algebraic property*: much like **permanent**, fixed amplitudes of random quantum circuits have low-degree polynomial structure
  - Consider circuit $C=C_m C_{m-1} \ldots C_1$
  - Structure comes from Feynman path integral:

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

  - This is a polynomial of degree m in the gate entries of the circuit
  - So the output probability $|\langle 0^n|C|0^n\rangle|^2$ is a 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 probabilities for *most* circuits
- *Recall*: Lipton wanted to compute **per**[X], choose random R, considered line A(t)=X+tR
- *Problem*: can't just perturb gates in a random linear direction (quantum circuits aren't linear… i.e., if A is unitary, B is unitary, A+tB is not necessary unitary)

New approach to *scramble* gates of fixed circuit

- Choose and fix $\{H_i\}_{i\in[m]}$ Haar random gates
- Now consider new circuit $C'=C'_m C'_{m-1} \ldots C'_1$ so that for each gate $C'_i=C_i H_i$
  - Notice that each gate in C' is completely random – "marginal property"
- But recall, Lipton also made use of "*univariate* polynomial structure"
- *Main idea*: "Rotate back towards C by small angle θ" (i.e., $C'_i=C_i H_i e^{-ih_i\theta}$)
  - If θ=1 the corresponding circuit C'=C, and if θ ≈ 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*: $e^{-ih_i\theta}$ is not polynomial in θ
- *Solution:* take fixed truncation of Taylor series for $e^{-ih_i\theta}$
  - So each gate entry is a polynomial in θ and so is $|\langle 0^n|C|0^n\rangle|^2$
  - Now interpolate and compute p(1)= $|\langle 0^n|C|0^n\rangle|^2$
- *This shows average-case exact hardness for a different circuit distribution!*
  - But we show that approximate hardness over this "truncated" circuit distribution is equivalent to the original RCS hardness conjecture (i.e., *approximate average-case* hardness over the gatewise Haar distribution)

2. Using statistical tests to verify RCS

Verifying RCS in the NISQ era

- *Constraint*: can only take a small (poly(n)) number of samples from the quantum device
- *Unique tool in NISQ Era*: It's feasible to take "modestly exponential" classical computation time per sample
- *Challenge*: Complexity arguments require closeness in total variation distance. **But we can't hope to unconditionally verify this with few samples from the device.**

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

- We want to compute:

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

- Note this can be well-approximated– take samples $x_1, x_2, \ldots, x_k$:
  - For each, use exp(n) classical processing time to compute log of ideal probabilities!
  - Mean converges to expectation with k=poly(n) samples from the device by Chernoff
- Then accept if score is sufficiently close to the expected ideal cross-entropy, which can be calculated

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 far in total variation which score well on CE)
- [Boixo et al. ’16]: Assume that

$$\rho_{dev} = \alpha\rho_{id} + (1-\alpha)\, Id$$

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

Deeper reasons to believe in Cross-Entropy

- This assumption can be weakened, if we "merely believe":
  - $H(p_{dev}) \geq H(p_{id})$
- Pinsker's inequality:

$$|p_{dev} - p_{id}|_{TV} \leq \sqrt{\frac{1}{2}|p_{dev} - p_{id}|_{KL}}$$

  - Where $|p_{dev}-p_{id}|_{KL}=CE(p_{dev},p_{id})-H(p_{dev})$
- So if we find cross-entropy ε-close to ideal, we've certified closeness in total variation distance to error $O(\varepsilon^{1/2})$
- This assumption makes sense if you think your device is corrupted by random errors

Removing assumption: Is scoring high on CE "intrinsically" hard?

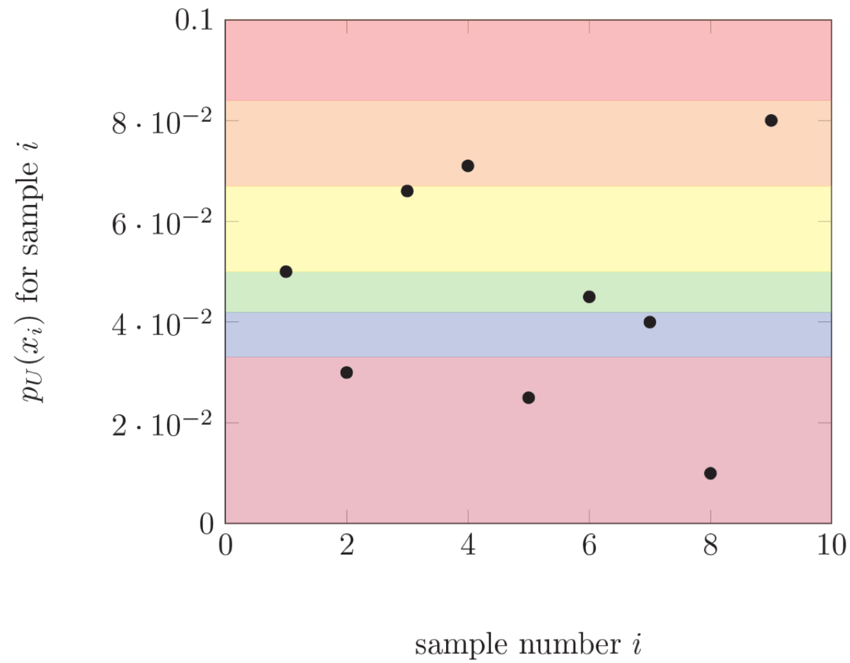
- The output distributions of RCS are "Porter-Thomas"
  - $\Pr[p_x = q/N] = e^{-q}$
- This "shape" of the distribution is **not** a signature of quantum effects!
  - We show the "shape" can be reproduced classically (e.g., by Poisson processes)
- However, pairs of distributions scoring highly on CE test 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
  - But don't know how to give complexity theoretic evidence

Introducing… Binned Output Generation (BOG)

- Why not use the same number of samples and take a multidimensional projection?
- Consider dividing the [0,1] interval into poly(n) bins
- Observe k samples $x_1,x_2,\ldots,x_k$ and calculate ideal probabilities for each sample on supercomputer
- Accept if the number of outcome probabilites in each bin are approximately equal to expected frequency in each bin
- Verifies cross-entropy and HOG – inherits the advantages of both (if you believe in either…)



Thanks!