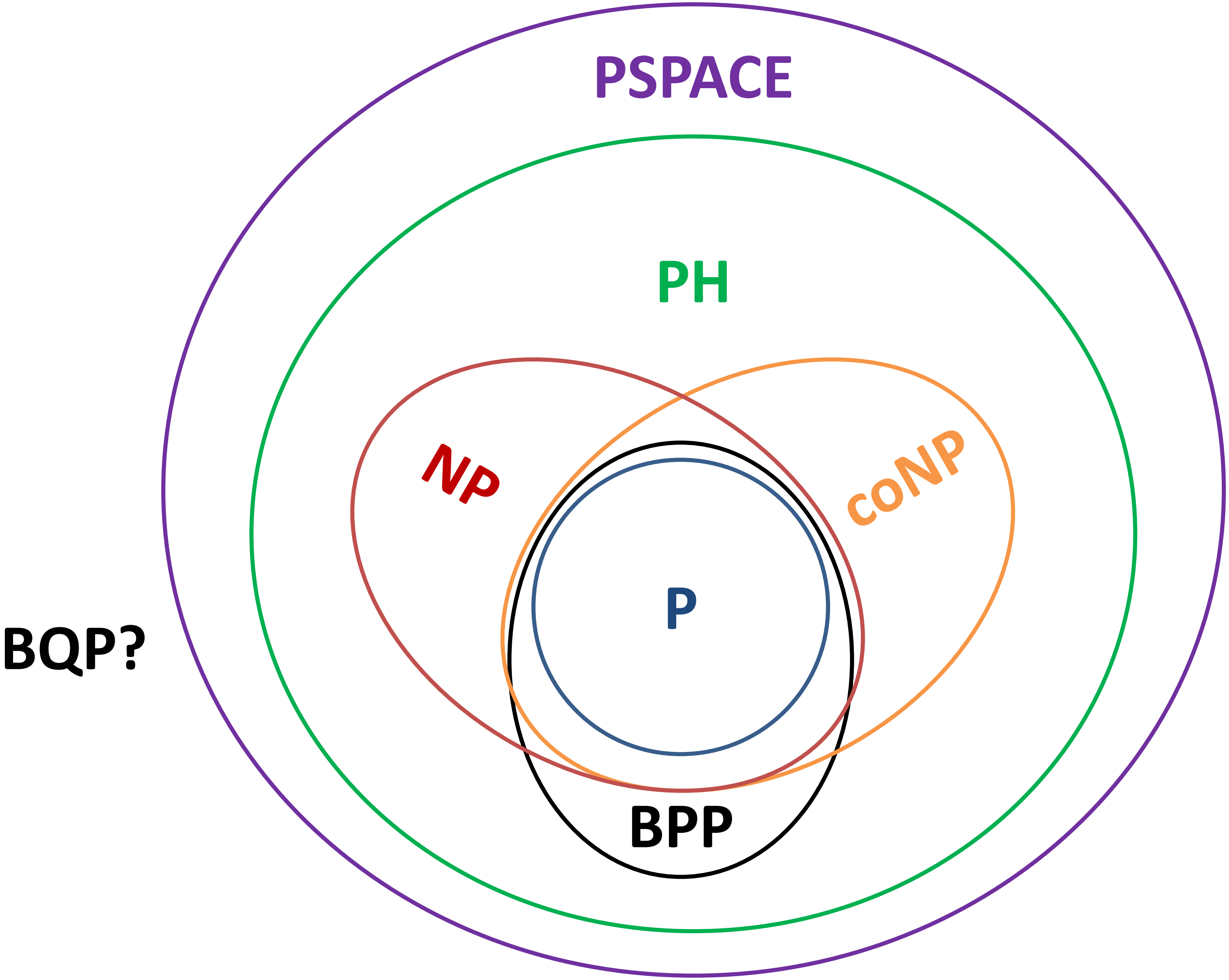
Oracle Separation of BQP and PH

Avishay Tal (Stanford University)
joint with **Ran Raz** (Princeton University)

The Landscape of Complexity Classes



Where does BQP fit in the landscape?

BQP: Bounded Error Quantum Polynomial Time

We know: **BPP ⊆ BQP ⊆ PSPACE**

Oracle Separations:

- $\exists$oracle A: $\mathbf{NP}^{A} \not\subseteq \mathbf{BQP}^{A}$ [BBBV'97]
- $\exists$oracle A: $\mathbf{BQP}^{A} \not\subseteq \mathbf{BPP}^{A}$ [BV'93]
- $\exists$oracle A: $\mathbf{BQP}^{A} \not\subseteq \mathbf{MA}^{A}$ [Watrous'00]

Could it be possible that **BQP ⊆ PH ?**

BQP ⊆ AM ?

Our Main Result: BQP vs. PH

Recall: a language L in **PH** iff there exists a constant k, and a poly-time computable relation R s.t.

$$x \in L \quad \Leftrightarrow \quad \exists y_1 \forall y_2 \exists y_3 \ldots Q_k y_k : R(x, y_1, \ldots, y_k)$$

$$|y_1| + |y_2| + \ldots + |y_k| \le \mathrm{poly}(|x|)$$

Our Main Result:

$$\exists \text{ oracle } A: \quad \mathbf{BQP}^{A} \not\subseteq \mathbf{PH}^{A}$$

The Black-Box/Query Model

$$\sum_{i=1}^{N} \alpha_i |i\rangle \longrightarrow \boxed{x \in \{\pm 1\}^N} \longrightarrow \sum_{i=1}^{N} \alpha_i x_i |i\rangle$$

Complexity measure: number of queries to the black box.

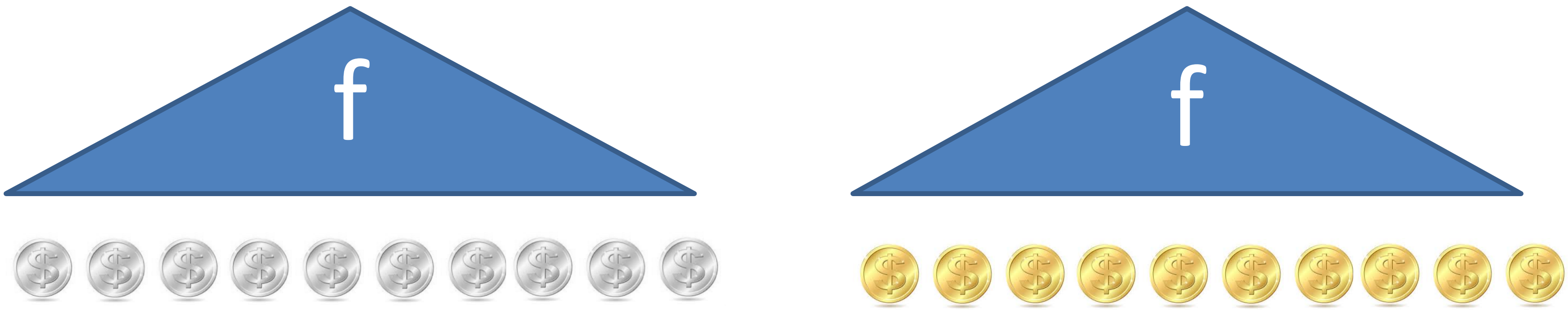
Deterministic Query Complexity = **Decision Tree Complexity**

Quantum Query Complexity = Queries are made in **superposition**

PH analog = **AC^0** circuits

Known reductions: Black-box separations imply oracle separations

The Pseudorandomness Setting



Def'n: a distribution D is **pseudorandom** against a class of functions $\mathcal{C}$ if

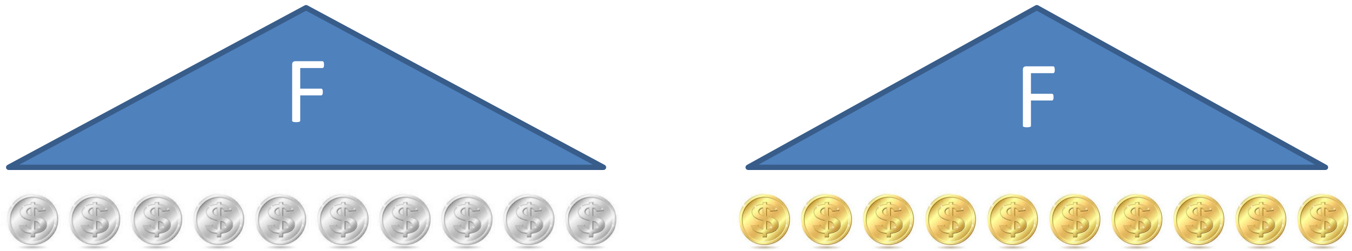
$$\forall f \in \mathcal{C}: \qquad \mathbf{E}_{x\sim D}[f(x)] \approx \mathbf{E}_{x\sim U}[f(x)]$$

The Pseudorandomness Setting

[Aaronson'10, Fefferman-Shaltiel-Umans-Viola'12]

Can you find a distribution which is pseudorandom for **AC^0** but not pseudorandom for **poly-log-time quantum algorithms**?

→ an oracle separation between **BQP** from **PH**



Let D be a distribution over $\{-1,1\}^{2N}$.

We say that an algorithm A distinguishes between D and U with **advantage** α if $\alpha = |\mathbf{E}_{x\sim D}[A(x)] - \mathbf{E}_{x\sim U}[A(x)]|$.

Main Result: We present a distribution D such that:

1. $\exists$a **log(N)** time quantum algorithm distinguishing between D and U with advantage $\Omega\,(1/\log N)$
2. Any **quasipoly(N)**-size constant-depth circuit distinguishes between D and U with advantage $\tilde{O}\left(\frac{1}{\sqrt{N}}\right)$

Standard techniques ➔ amplify advantage of quantum algorithm to be 0.99 or even 1-1/poly(N).

The Separating Distribution D

(Based on Aaronson's Forrelation distribution)

- Let N be a power of 2. Let $\epsilon = 1/O(\log N)$.
- Let G to be a multi-variate gaussian (**MVG**) distribution on $\mathbb{R}^{2N}$ with zero-means and covariance matrix

$$\epsilon \cdot \begin{pmatrix} I_N & H \\ H & I_N \end{pmatrix}$$

where H is the $N \times N$ **Hadamard** matrix with

$$H_{i,j} = \frac{1}{\sqrt{N}} \cdot \ (-1)^{<i,j>}$$

Sampling $z' \sim D$:

1. Sample $z \sim G$, truncate each z_i to be within $[-1,1]$
2. For $i = 1, \ldots, 2N$, sample independently $z'_i \in \{-1,1\}$ with $\mathbf{E}[z'_i] = z_i$.

Quantum Algorithm Distinguishing D

[Aaronson'10, Aaronson-Ambainis'15]:

1-query O(log N)-time quantum algorithm Q s.t.

$$\Pr[Q \text{ accepts input } (x,y)] = \frac{1+\Phi(x,y)}{2}$$

where

$$\Phi(x,y) = \frac{1}{N^{3/2}} \cdot \sum_{i=1}^{N} \sum_{j=1}^{N} (-1)^{<i,j>} \cdot x_i \cdot y_j$$

$$\mathbf{E}_{(x,y)\sim U}[\Phi(x,y)] = 0$$

$$\mathbf{E}_{(x,y)\sim D}[\Phi(x,y)] \approx \epsilon = \Omega\left(\frac{1}{\log N}\right)$$

D is Pseudorandom for $\mathbf{AC}^0$

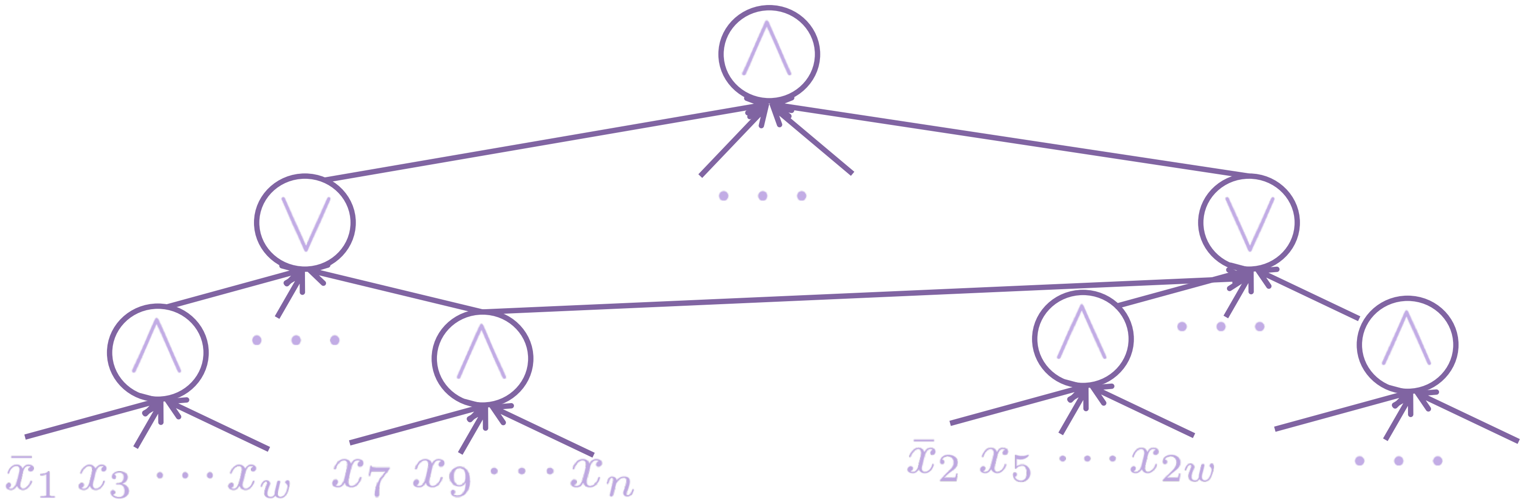
We are left to prove:

D is pseudorandom for $\mathbf{AC}^0$.

Main Ingredients & Techniques:

- Fourier Analysis
- $\mathbf{AC}^0$ circuits are well-approximated by **sparse** low-degree polynomials.
- Fractional **PRG** approach of **[CHHL18]**.
- Sum of independent Gaussians is a Gaussian.

Bounded Depth Circuits



$\mathrm{AC}^0[s, d]$:

- s gates (**size** of the circuit)
- depth d
- alternating gates

We focus on

$$\mathrm{AC}^0\left[N^{\mathrm{polylog}(N)}, O(1)\right]$$

What do we know about AC^0?

[Ajtai'83, Furst-Saxe-Sipser'84, Yao'85,Håstad '86]:

- **Parity** not in $\mathrm{AC}^0\left[N^{\mathrm{polylog}(N)}, O(1)\right]$.
- **Parity** requires $\exp\left(N^{1/(d-1)}\right)$ size for depth d.

➔ $\exists$ oracle A: $\mathbf{PSPACE}^A \not\subseteq \mathbf{PH}^A$

Fourier-analytical proof technique:

- **AC^0** circuits can be well-approximated (in ℓ_2) by **low-degree polynomials** (over $\mathbb{R}$). [**Håstad '86,LMN'93**]
- **Parity** cannot.

Potential problem with the approach:

$O(\log N)$ time quantum algorithms (**BQLogTime**) are also well-approximated by low-degree polys.

The Difference between BQLogTime and AC^0

Both **BQLogtime** & **AC^0** are approximated by low-degree polynomials, but **these polynomials are different**!

BQLogtime can have **dense** low-degree polynomials, e.g.

$$\Phi(x,y) = \frac{1}{N^{3/2}} \cdot \sum_{i=1}^{N} \sum_{j=1}^{N} (-1)^{<i,j>} \cdot x_i \cdot y_j$$

[T'14]: **AC^0** has **sparse** low-degree approximations

$$\forall k: \sum_{S \subseteq [n], |S|=k} \left|\hat{f}(S)\right| \leq (\text{polylog } N)^k$$

Fourier Analytical Approach – First Attempt

The Fourier expansion of $f : \{-1,1\}^{2N} \to \{-1,1\}$:

$$f(x) = \sum_{S \subseteq [2N]} \hat{f}(S) \cdot \prod_{i \in S} x_i$$

Goal: $\left|\mathbf{E}_{z' \sim D}[f(z')] - \mathbf{E}_{x \sim U}[f(x)]\right| = \tilde{O}\left(\frac{1}{\sqrt{N}}\right)$

Recall: Sampling $z' \sim D$:

1. Sample $z \sim G$, truncate each z_i to be within $[-1,1]$
2. For $i = 1, \ldots, 2N$, sample independently $z_i' \in \{-1,1\}$ with $\mathbf{E}[z_i'] = z_i$

Using multilinearity of f and that whp $\mathrm{trunc}(z) = z$:

$$\mathbf{E}_{z' \sim D}[f(z')] = \mathbf{E}_{z \sim G}[f(\mathrm{trunc}(z))] \approx \mathbf{E}_{z \sim G}[f(z)]$$

➔ Suffices to show $\left|\mathbf{E}_{z \sim G}[f(z)] - \mathbf{E}_{x \sim U}[f(x)]\right| = \tilde{O}\left(\frac{1}{\sqrt{N}}\right)$

Fourier Analytical Approach – First Attempt

$$\mathbf{E}_{z\sim G}[f(z)] - \mathbf{E}_{x\sim U}[f(x)]$$

$$= \sum_{S\subseteq[2N]} \hat{f}(S) \cdot \left(\mathbf{E}_{z\sim G}\left[\prod_{i\in S} z_i\right] - \mathbf{E}_{x\sim U}\left[\prod_{i\in S} x_i\right]\right)$$

$$= \sum_{S\subseteq[2N],|S|\geq 1} \hat{f}(S) \cdot \mathbf{E}_{z\sim G}\left[\prod_{i\in S} z_i\right]$$

$$= \sum_{\ell=1}^{N} \sum_{|S|=2\ell} \hat{f}(S) \cdot \mathbf{E}_{z\sim G}\left[\prod_{i\in S} z_i\right]$$

$$\leq \sum_{\ell=1}^{N} \sum_{|S|=2\ell} \left|\hat{f}(S)\right| \cdot \epsilon^{\ell} \cdot \frac{\ell!}{\sqrt{N}^{\ell}}$$

$$\leq \sum_{\ell=1}^{N} \mathrm{polylog}(N)^{2\ell} \cdot \epsilon^{\ell} \cdot \frac{\ell!}{\sqrt{N}^{\ell}}$$

Contribution of first $\widetilde{\mathrm{O}}(\sqrt{N})$ terms:

$\epsilon \cdot \mathrm{polylog}(N)/\sqrt{N}$

Contribution of larger terms?

Main Technical Lemma

Suppose $Z \sim G$ is a zero-mean **MVG** on $\mathbb{R}^{2N}$ with

- $\forall i$: $\mathbf{var}(Z_i) \leq 1/O\left(\log\left(\frac{N}{\delta}\right)\right)$
- $\forall i, j$: $\mathbf{cov}\left(Z_i, Z_j\right) \leq \delta$

➔ whp $Z \in [-1,1]^{2N}$

Then, for any quasi-poly size constant depth **AC⁰** circuit f,

$$|\mathbf{E}_{z \sim G}[f(z)] - \mathbf{E}_{x \sim U}[f(x)]| \leq \delta \cdot \mathrm{polylog}(N)$$

Which properties of **AC⁰** circuits are used in the proof?

- The bound $\sum_{|S|=2} \left|\hat{f}(S)\right| \leq \mathrm{polylog}(N)$
- Closure under restrictions.

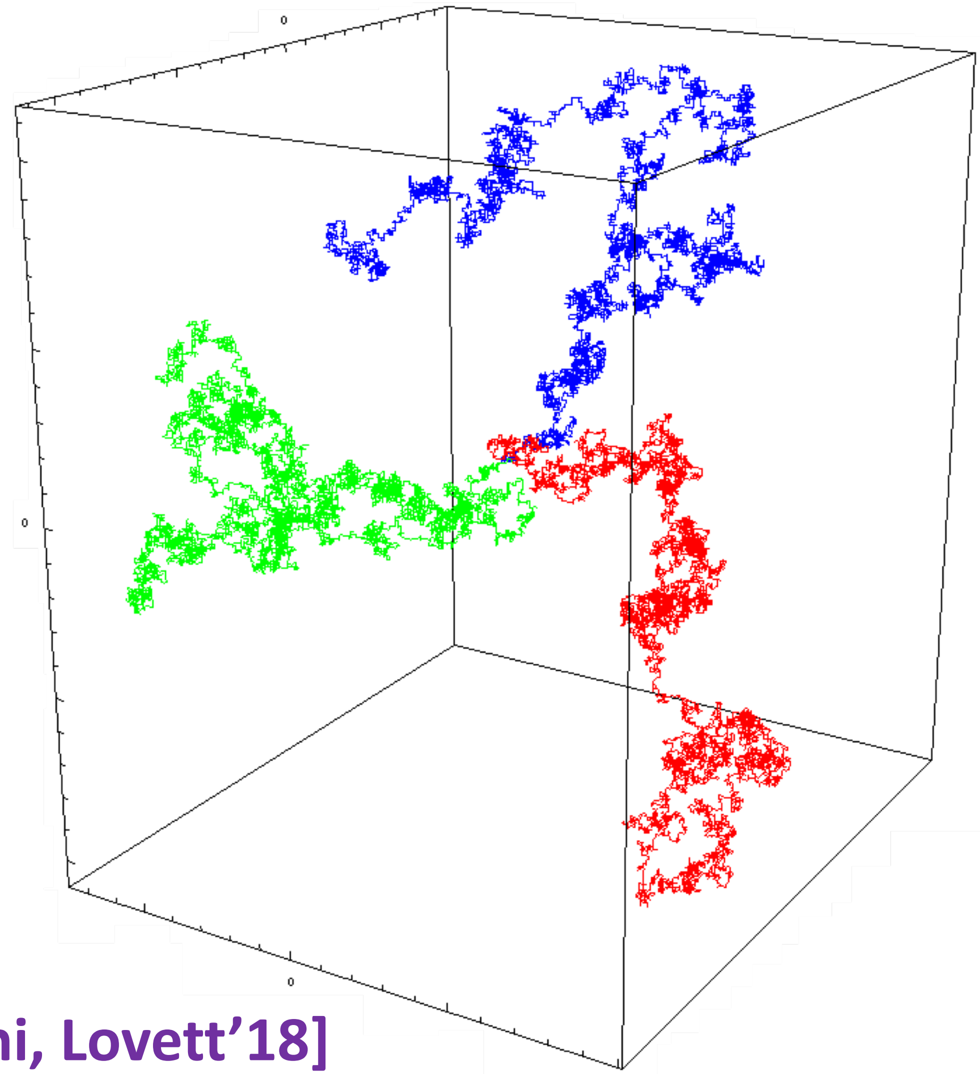
G fools any class of functions with these two properties

Viewing $Z \sim G$ as a result of a random walk

A Thought Experiment:

Instead of sampling $Z \sim G$ at once, we sample t vectors $Z^{(1)}, \ldots, Z^{(t)} \sim G$ independently, and take

$$Z = \frac{1}{\sqrt{t}} \cdot \left(Z^{(1)} + \cdots + Z^{(t)}\right)$$



Based on the work of **[Chattopadhyay, Hatami, Hosseini, Lovett'18]**

Picture from http://en.wikipedia.org/wiki/Random_walk

Viewing $Z \sim G$ as a result of a random walk

Sample t vectors $Z^{(1)}, \dots, Z^{(t)} \sim G$

Define $t+1$ hybrids:

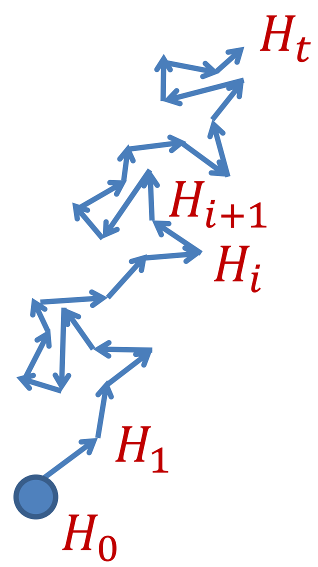
- $H_0 = \vec{0}$
- For $i = 1, \dots, t$

$$H_i = \frac{1}{\sqrt{t}} \cdot \left(Z^{(1)} + \cdots + Z^{(i)} \right)$$

Observe: $H_t \sim G$.

Taking $t \to \infty$ yields a Brownian motion.

We take $t = \mathrm{poly}(N)$.



Claim: for $i = 0, \dots, t-1$,

$$\left| \mathbf{E}\left[f\left(H_{(i+1)} \right) \right] - \mathbf{E}[f(H_i)] \right| \le \frac{\delta}{t} \cdot \mathrm{polylog}(N).$$

Claim - Base Case

Base Case:

$$\mathbf{E}[f(H_1)] - \mathbf{E}[f(H_0)] = \mathbf{E}\left[f\left(\frac{1}{\sqrt{t}} \cdot Z^{(1)}\right)\right] - f(\vec{0})$$

$$= \sum_{\ell=1}^{N} \sum_{|S|=2\ell} \hat{f}(S) \cdot \mathbf{E}_{z\sim G}\left[\left(\frac{1}{\sqrt{t}}\right)^{2\ell} \cdot \prod_{i\in S} z_i\right]$$

$$\leq \sum_{\ell=1}^{N} \sum_{|S|=2\ell} |\hat{f}(S)| \cdot \frac{\delta^{\ell} \cdot O(\ell)^{\ell}}{t^{\ell}}$$

$$\leq \frac{\delta}{t} \cdot \mathrm{polylog}(N) \; + \; o\left(\frac{\delta}{t}\right)$$

Reducing the General Case to the Base Case

Lemma [CHHL'18]: for all $z_0 \in [-1/2,1/2]^{2N}$

$$g(z) = f(z + z_0) - f(z_0)$$

can be written as $\mathbf{E}_\rho\left[f_\rho(2 \cdot z) - f_\rho\left(\vec{0}\right)\right]$ where f_ρ is a random restriction of f (whose marginals depend on z_0).

Conditioned on $H_i \in [-1/2,1/2]^{2N}$ (happens whp):

$$\left|\mathbf{E}\left[f\left(H_{(i+1)}\right)\right] - \mathbf{E}[f(H_i)]\right|$$

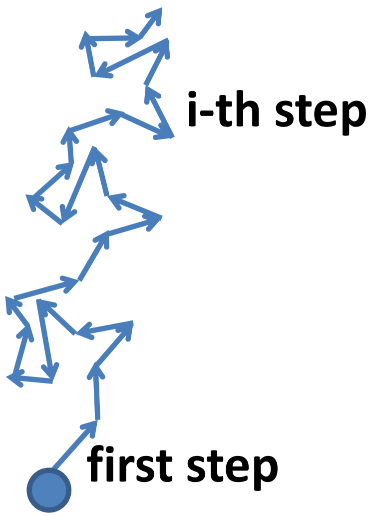
$$\leq \left|\mathbf{E}\left[f\left(H_i + \frac{1}{\sqrt{t}} Z^{(i+1)}\right) - f(H_i)\right]\right|$$

$$\leq \left|\mathbf{E}\left[f_\rho\left(\frac{2}{\sqrt{t}} \cdot Z^{(i+1)}\right) - f_\rho\left(\vec{0}\right)\right]\right| \leq \frac{4\delta}{t} \cdot \mathrm{polylog}(N)$$

Recap: Proof by Picture

[CHHL'18]: i-th step $\approx$ first step, using closure under restrictions.

First Step: Simple Fourier Analysis
Only second level matters.



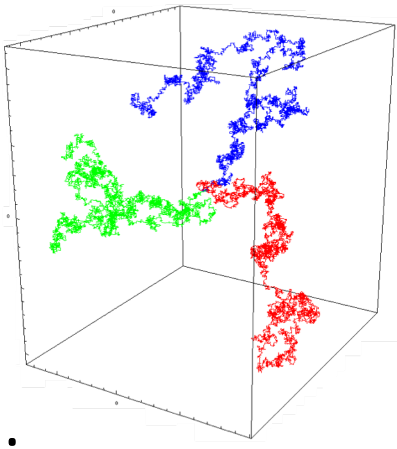
Recap

- Defined a distribution D based on **MVG** G.
- D is **not pseudorandom** for **log**(**N**)-time **quantum** algorithms. **[Aaronson'09, Aaronson-Ambainis'15]**
- D is **pseudorandom** for **AC⁰** **(our contribution)**

$$|\mathbf{E}_{z\sim G}[f(z)] - \mathbf{E}_{x\sim U}[f(x)]| \le \delta \cdot \mathrm{polylog}(N):$$

  - **Thought experiment:** Viewing $Z \sim G$ as a result of a random walk with t tiny steps.
  - **AC⁰** circuits are well-approximated by **sparse** low-degree polynomials **[T'14]**

    ➔ first step has advantage $\left(\frac{\delta}{t}\right) \cdot \mathrm{polylog}(N)$



  - **[Chattopadhyay, Hatami, Hosseini, Lovett '18]**:

    ➔ i-th step has advantage $\left(\frac{\delta}{t}\right) \cdot \mathrm{polylog}(N)$

Open Problems & New results

Follow-ups:

- [Aaronson, Fortnow]: an oracle A s.t.
$$\mathbf{BQP}^{A} \not\subseteq \mathbf{P}^{A} = \mathbf{NP}^{A}$$
- [Fortnow]: under our oracle **PH** is infinite.

Open Problems:

- Does the original suggestion of [Aaronson'09] (without $1/\log(N)$ noise) work?
- [Aaronson]: Find an oracle A s.t.
  - $\mathbf{NP}^{A} \subseteq \mathbf{BQP}^{A}$
  - $\mathbf{PH}^{A} \not\subseteq \mathbf{BQP}^{A}$
- [Fortnow]: Does $\mathbf{NP}^{\mathbf{BQP}} \not\subseteq \mathbf{BQP}^{\mathbf{NP}}$?

Open Problems 2: Pseudorandomness

Separate **BQLogTime** and $\mathbf{AC^0}[\oplus]$.

Suffices to show for all f in $\mathbf{AC^0}[\oplus]$:

$$\sum_{|S|=2} |\hat{f}(S)| \leq \frac{\sqrt{N}}{\mathrm{polylog}(N)}$$

Conjecture [CHLT'18]: for all f in $\mathbf{AC^0}[\oplus]$

$$\sum_{|S|=2} |\hat{f}(S)| \leq \mathrm{polylog}(N)$$

Claim [CHLT'18]: Conjecture implies a **PRG** for $\mathbf{AC^0}[\oplus]$ with $\mathrm{polylog}(N)$ seed length.

Thank You!

© Kevin Hong for Quanta Magazine