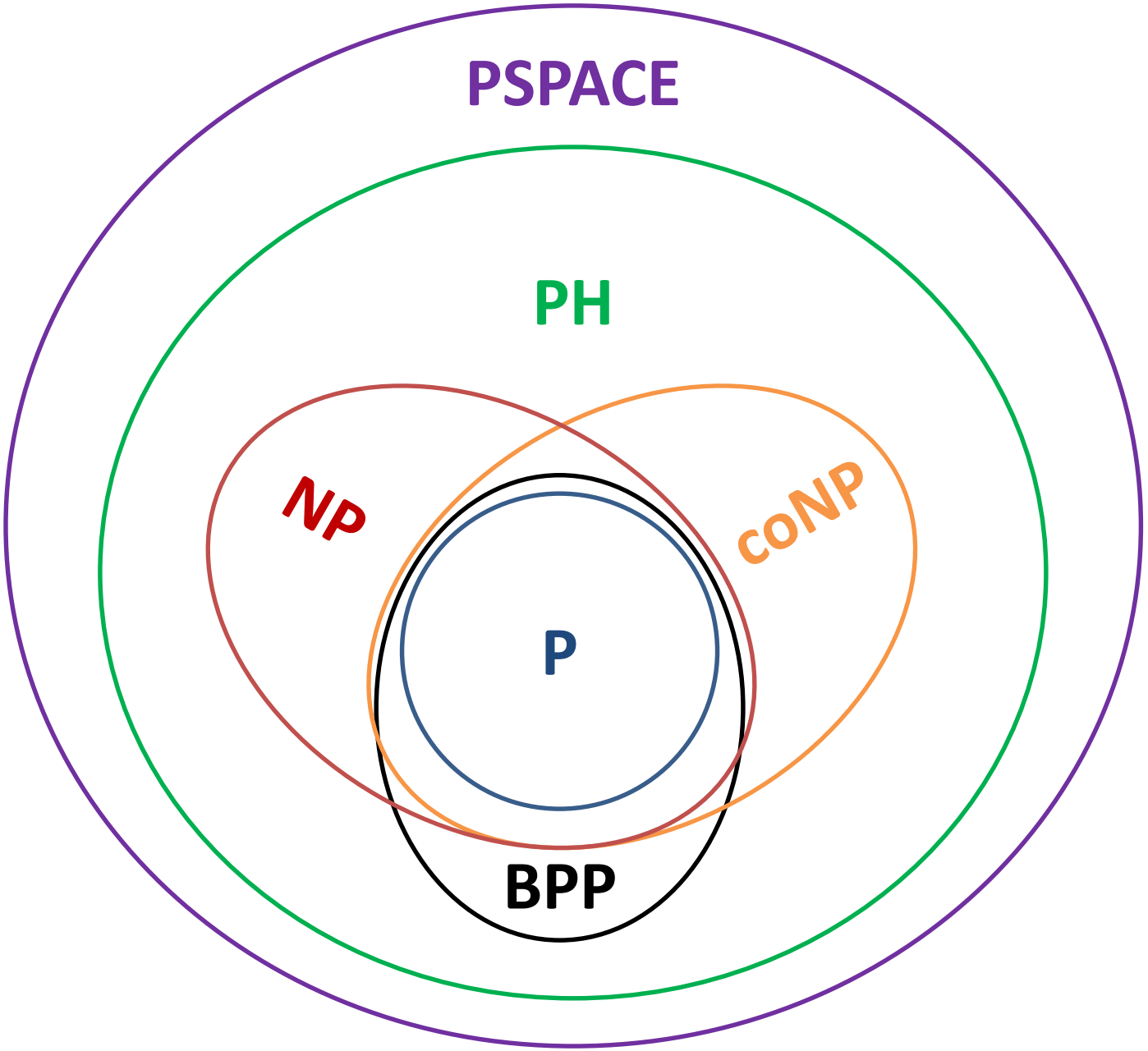
Oracle Separation of BQP and PH

Avishay Tal (**Stanford, Simons Institute**)

joint with **Ran Raz (Princeton University)**

The Landscape of Complexity Classes



BQP?

Where does BQP fit in the landscape?

BQP: Bounded-error Quantum Polynomial time

We know: **BPP ⊆ BQP ⊆ PSPACE**

Open:

BQP ⊆ NP? **NP ⊆ BQP?**

BQP ⊆ PH? **PH ⊆ BQP?**

Common Belief: all are false.

Any separation would imply **PSPACE ≠ P,**

and thus seems out of reach.

Oracles

Oracles

An oracle is a language $A: \{0,1\}^* \to \{0,1\}$.

$\mathbf{P}^A$, $\mathbf{NP}^A$, … are the classes of decision problems solvable using devices in **P, NP**, … that could additionally ask queries to A at a unit cost.

For example:

L in $\mathbf{P}$ $\Leftrightarrow$ ∃ a poly-time Turing-machine M, that decides whether $x \in L$.

L in $\mathbf{P}^A$ $\Leftrightarrow$ ∃ a poly-time Turing-machine M, making queries to the oracle A, that decides whether $x \in L$.

Oracle Separations

Separating Classical Classes:

- $\exists$oracle A: $\mathbf{P}^{A} \neq \mathbf{NP}^{A}$ **[BGS'75]**
- $\exists$oracle A: $\mathbf{NP}^{A} \neq \mathbf{PH}^{A}$ **[BGS'75]**
- $\exists$oracle A: $\mathbf{PH}^{A} \neq \mathbf{PSPACE}^{A}$ **[FSS'81, A'83, Y'85]**

Quantum vs. Classical 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: $\mathbf{BQP}^{A} \not\subseteq \mathbf{NP}^{A}$ **[Watrous'00]**

Could it be possible that **BQP** $\subseteq$ **PH ?**

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.

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

$$|y_1| + |y_2| + \ \ldots \ + \ |y_k| \leq \text{poly}(|x|)$$

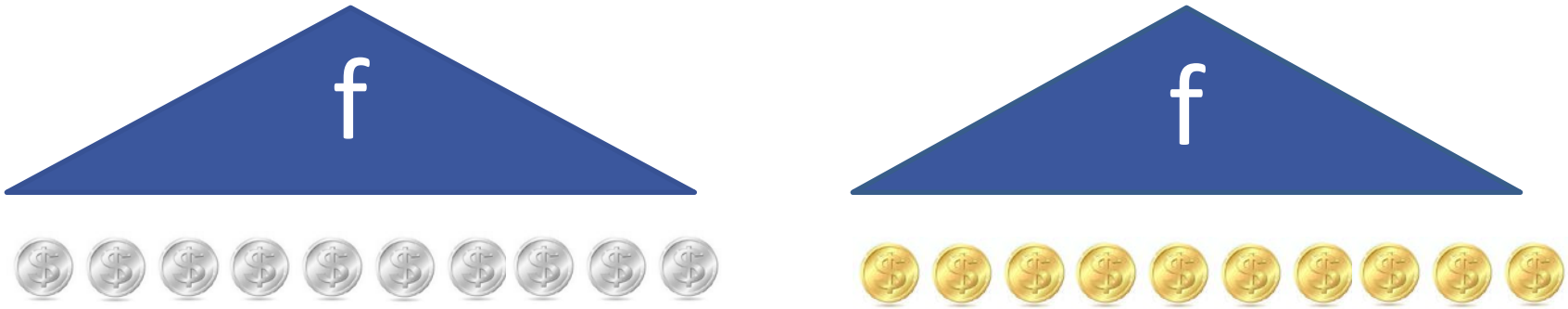
Our Main Result:

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

"Even if **P** were equal to **NP**, even making that strong assumption, that's not going to be enough to capture quantum computing."

(Lance Fortnow)

The Pseudorandomness Setting



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

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

The Pseudorandomness Setting

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

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 $\{-1,1\}^N$.

Def'n: f has **advantage** α distinguishing between D and U with if $\alpha = |\mathbf{E}_{x\sim D}[f(x)] - \mathbf{E}_{x\sim U}[f(x)]|$.

Main Result: We present a distribution D such that:

1. ∃a **log(N)** time quantum algorithm distinguishing between D and U with advantage $\Omega\left(\frac{1}{\log N}\right)$.
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 alg to **1-1/poly(N)**.

Plan:

1. Definition of D
2. Quantum algorithm distinguishing D from U
3. D is pseudorandom for AC^0

The Separating Distribution D

(Based on Aaronson's Forrelation distribution)

- Let N be a power of 2. Let $\epsilon = 1/O(\log N)$.
- **The Distribution G:** draw $x_1, \ldots, x_{N/2}$ i.i.d. from $\mathcal{N}(0, \epsilon)$

$$\begin{pmatrix} y_1 \\ y_2 \\ \ldots \\ y_{N/2} \end{pmatrix} = H \cdot \begin{pmatrix} x_1 \\ x_2 \\ \ldots \\ x_{N/2} \end{pmatrix}$$

where H is the $(N/2) \times (N/2)$ **Hadamard** matrix.
Output $z = (x_1, \ldots., x_{N/2}, y_1, \ldots, y_{N/2})$.

The Separating Distribution D

(Based on Aaronson's Forrelation distribution)

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

$$\epsilon \cdot \begin{pmatrix} I_{N/2} & H \\ H & I_{N/2} \end{pmatrix}$$

where H is the $(N/2) \times (N/2)$ **Hadamard** matrix with

$$H_{i,j} = \frac{1}{\sqrt{N/2}} \cdot (-1)^{<i,j>}$$

Quantum Algorithm Distinguishing D from U

Quantum Algorithm Distinguishing D

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

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

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

$$\mathbf{Pr}_{(x,y)\sim U}[Q \text{ accepts input } (x,y)] = 1/2$$

$$\mathbf{Pr}_{(x,y)\sim D}[Q \text{ accepts input } (x,y)] \geq \frac{1}{2} + \Omega(\epsilon)$$

The Quantum Algorithm

1. **Prepare the state**:

$$\sum_{i\in[N/2]} \frac{1}{\sqrt{N}} \cdot |0,i\rangle + \sum_{i\in[N/2]} \frac{1}{\sqrt{N}} \cdot |1,i\rangle$$

2. **Query**:

$$\sum_{i\in[N/2]} \frac{x_i}{\sqrt{N}} \cdot |0,i\rangle + \sum_{i\in[N/2]} \frac{y_i}{\sqrt{N}} \cdot |1,i\rangle$$

3. Apply the **Hadamard** transform to first half

$$\sum_{i\in[N/2]} \frac{(Hx)_i}{\sqrt{N}} \cdot |0,i\rangle + \sum_{i\in[N/2]} \frac{y_i}{\sqrt{N}} \cdot |1,i\rangle$$

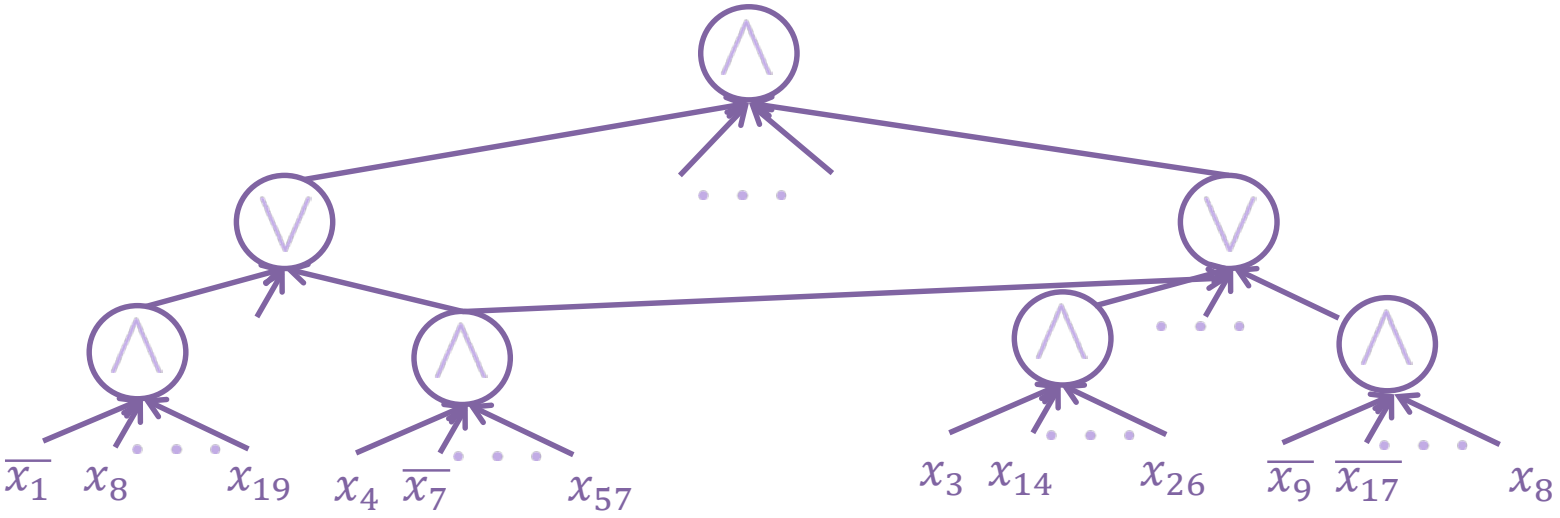
4. **Measure** the first qbit in $\{|+\rangle, |-\rangle\}$ basis

$$\Pr[Q \text{ accepts } (x,y)] = \frac{1}{2} + \frac{1}{N} \cdot \sum_{i=1}^{N/2} (Hx)_i y_i$$

The Main Result:

D is Pseudorandom for AC^0

Bounded Depth Circuits



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

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

We focus on

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

What do we know about AC^0?

[Furst-Saxe-Sipser'81, Ajtai'83, 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.

Fourier-analytical proof technique:

- **AC^0** circuits can be well-approximated (in ℓ_2) by **low-degree polynomials** (over $\mathbb{R}$). **[Håstad'86, LMN'89]**
- **Parity** cannot.

Potential problem with the approach:

$O(\log N)$ time quantum algorithms (**BQLogTime**) are also well-approximated by low-degree polys. **[BBCMW'98]**

The Difference between BQLogTime and AC^0

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

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

$$\frac{1}{2}+\frac{\langle Hx,y\rangle}{N} = \frac{1}{2}+\frac{1}{N}\cdot\sum_{i=1}^{N/2}\sum_{j=1}^{N/2} y_i\cdot H_{i,j}\cdot x_j$$

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

$$\forall k\colon \sum_{S\subseteq\{1,\ldots,N\}:|S|=k} \left|\hat{f}(S)\right| \le (\text{polylog } N)^k$$

First Attempt:

Fourier Analytical Approach

Fourier Analytical Approach – First Attempt

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

$$f(x) = \sum_{S \subseteq \{1,\dots,N\}} \hat{f}(S) \cdot \prod_{i \in S} x_i$$

The **Fourier expansion** extends f to **non-Boolean** inputs!

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

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

$\mathbf{E}_{u \sim U}[f(u)] = f(\vec{0})$

New Goal: $\left|\mathbf{E}_{z \sim G}[f(z)] - f(\vec{0})\right| = \tilde{O}\left(\frac{1}{\sqrt{N}}\right)$

Fourier Analytical Approach – First Attempt

$$\mathbf{E}_{z\sim G}[f(z)] - f(\vec{0}) =$$

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

$$= \sum_{\ell=1}^{N/2} \sum_{|S|=2\ell} \hat{f}(S) \cdot \mathbf{E}_{z\sim G}\left[\prod_{i\in S} z_i\right]$$ **(odd moments = 0)**

$$\leq \sum_{\ell=1}^{N/2} \sum_{|S|=2\ell} \left|\hat{f}(S)\right| \cdot O\left(\frac{\epsilon\ell}{\sqrt{N}}\right)^{\ell}$$ **(Isserlis' Theorem)**

$$\leq \sum_{\ell=1}^{N/2} \mathrm{polylog}(N)^{2\ell} \cdot O\left(\frac{\epsilon\ell}{\sqrt{N}}\right)^{\ell}$$ **[T'14]**

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

What about the larger terms?

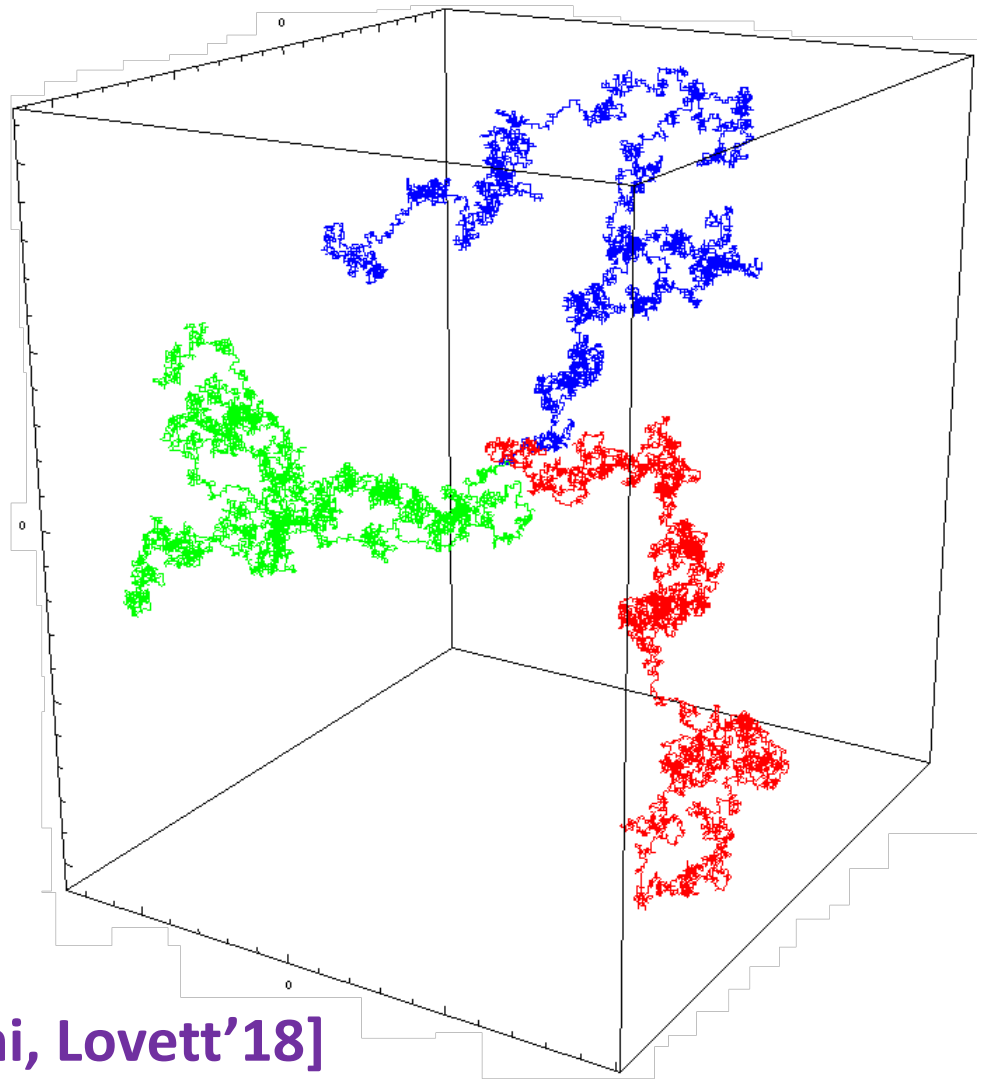
Second Attempt:
The Random Walk Approach

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

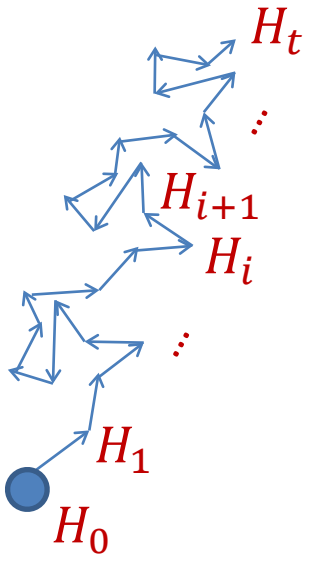
Viewing $z \sim G$ as a result of a random walk

Sample t vectors $z^{(1)}, \ldots, z^{(t)} \sim G$

Define $t+1$ hybrids:

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

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



Observe: $H_t \sim G$.

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

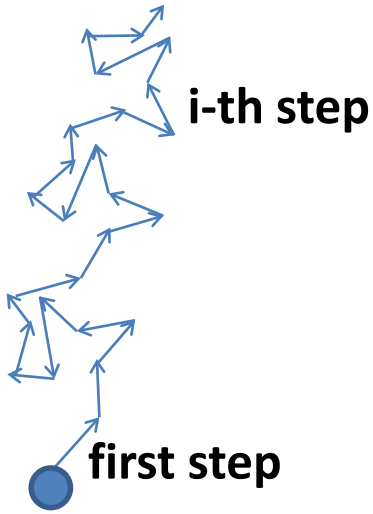
We take $t = \text{poly}(N)$.

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

$$\left| \mathbf{E}[f(H_{i+1})] - \mathbf{E}[f(H_i)] \right| \leq \frac{\text{polylog}(N)}{t\sqrt{N}}.$$

Proof by Picture

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



First Step: Simple Fourier Analysis
Only second level matters.

Base Case

$$\mathbf{E}[f(H_1)] - \mathbf{E}[f(H_0)]$$

$$= \mathbf{E}_{z\sim G}\left[f\left(\frac{1}{\sqrt{t}}z\right)\right] - f(\vec{0})$$

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

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

$$\leq \sum_{\ell=1}^{N/2} \mathrm{polylog}(N)^{2\ell} \cdot O\left(\frac{\epsilon\ell}{t\sqrt{N}}\right)^{\ell}$$

$$\leq \frac{\mathrm{polylog}(N)}{t\sqrt{N}} + o\left(\frac{1}{t\sqrt{N}}\right)$$

(for t large enough)

General Case: Reduction to Base Case

Lemma [CHHL'18]: for any fixed $v \in [-0.5, 0.5]^N$ the fnc

$$g(z) \stackrel{\text{def}}{=} f(v+z) - f(v)$$

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

Analysis of step i+1:

Conditioned on $H_i \in [-0.5, 0.5]^N$ (happens whp):

$$\begin{aligned}
&|\mathbf{E}[f(H_{i+1})] - \mathbf{E}[f(H_i)]| \\
&\leq \left|\mathbf{E}\left[f\left(H_i + \frac{1}{\sqrt{t}} \cdot z^{(i+1)}\right) - f(H_i)\right]\right| \\
&\leq \left|\mathbf{E}\left[f_\rho\left(\frac{2}{\sqrt{t}} \cdot z^{(i+1)}\right) - f_\rho(\vec{0})\right]\right| \leq \frac{\text{polylog}(N)}{t\sqrt{N}}
\end{aligned}$$

Recap

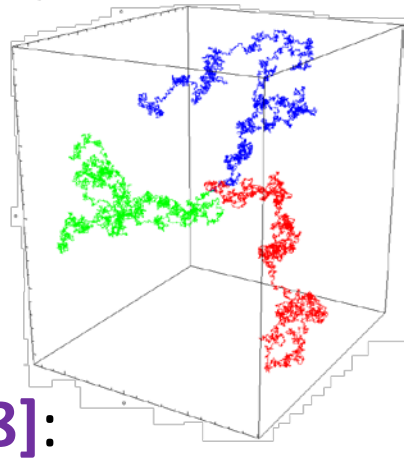
1. Defined a distribution D based on a **MVG** G.
2. D is **not pseudorandom** for **log**(**N**)-time **quantum** algorithms. **[Aaronson'10, Aaronson-Ambainis'15]**
3. D is **pseudorandom** for **AC⁰** **(our contribution)**

$$|\mathbf{E}_{z\sim G}[f(z)] - \mathbf{E}_{u\sim U}[f(u)]| \leq \tilde{O}\left(\frac{1}{\sqrt{N}}\right).$$

- **Thought Experiment:** View $z\sim G$ as a result of a random walk making t tiny steps.
- **AC⁰** circuits are approximated by **sparse** low-degree polynomials **[T'14]**

  ➔ first step has advantage $\tilde{O}\left(\frac{1}{t\sqrt{N}}\right)$
- **[Chattopadhyay, Hatami, Hosseini, Lovett '18]**:

  ➔ i-th step has advantage $\tilde{O}\left(\frac{1}{t\sqrt{N}}\right)$



Thank You!