Oracle Separation of BQP and PH

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

© Kevin Hong for Quanta Magazine
The Landscape of Complexity Classes

- P
- BPP
- BQP?
- PSPACE
- PH
- NP
- coNP

Diagram:

- P ⊆ BPP
- NP ⊆ coNP
- PH ⊆ PSPACE

Questions:

- BQP?
Where does BQP fit in the landscape?

**BQP:** Bounded Error Quantum Polynomial Time

We know: \( \text{BPP} \subseteq \text{BQP} \subseteq \text{PSPACE} \)

**Oracle Separations:**

- \( \exists \text{oracle } A: \text{NP}^A \not\subseteq \text{BQP}^A \) \text{[BBBV’97]}
- \( \exists \text{oracle } A: \text{BQP}^A \not\subseteq \text{BPP}^A \) \text{[BV’93]}
- \( \exists \text{oracle } A: \text{BQP}^A \not\subseteq \text{MA}^A \) \text{[Watrous’00]}

Could it be possible that \( \text{BQP} \subseteq \text{PH} \)?

\( \text{BQP} \subseteq \text{AM} \)?
Recall: a language $L$ in $\text{PH}$ iff there exists a constant $k$, and a poly-time computable relation $R$ s.t.

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

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

Our Main Result: $\exists$ oracle $A$: $\text{BQP}^A \not\subseteq \text{PH}^A$
The Black-Box/Query Model

\[ \sum_{i=1}^{N} \alpha_i |i\rangle \quad \xrightarrow{x \in \{\pm 1\}^N} \quad \sum_{i=1}^{N} \alpha_i x_i |i\rangle \]

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 $C$ if

$$\forall f \in C: \quad E_{x \sim D}[f(x)] \approx E_{x \sim U}[f(x)]$$
Can you find a distribution which is pseudorandom for $\text{AC}^0$ but not pseudorandom for poly-log-time quantum algorithms?

⇒ an oracle separation between $\text{BQP}$ from $\text{PH}$
Let $D$ be a distribution over $\{-1,1\}^{2N}$.

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

**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 $\Rightarrow$ 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 $\epsilon = 1/\mathcal{O}(\log N)$.

• 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 $\mathbb{E}[z_i'] = z_i$. 

(Based on Aaronson’s Forrelation distribution)
Quantum Algorithm Distinguishing

[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

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

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

$$E_{(x,y) \sim D}[\Phi(x, y)] \approx \epsilon = \Omega\left(\frac{1}{\log N}\right)$$
We are left to prove:

\[ D \text{ is pseudorandom for } \mathbf{AC}^0. \]

**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

$\text{AC}^0[s, d]$:
- $s$ gates (size of the circuit)
- depth $d$
- alternating gates

We focus on $\text{AC}^0[N^{\text{polylog}(N)}, O(1)]$
What do we know about $\text{AC}^0$?

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

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

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

Fourier-analytical proof technique:

- $\text{AC}^0$ circuits can be well-approximated (in $\ell_2$) by low-degree polynomials (over $\mathbb{R}$). [Håstad ’86, LMN’93]
- **Parity** cannot.

Potential problem with the approach:

$O(\log N)$ time quantum algorithms ($\text{BQLogTime}$) 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

$$\forall k:\sum_{S \subseteq [n], |S|=k} |\hat{f}(S)| \leq (\text{polylog } N)^k$$
The Fourier expansion of \( f : \{-1,1\}^{2N} \rightarrow \{-1,1\} \):

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

**Goal:**

\[
|E_{z' \sim D}[f(z')] - E_{x \sim U}[f(x)]| = \tilde{O}\left(\frac{1}{\sqrt{N}}\right)
\]

**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 \( E[z'_i] = z_i \)

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

\[
E_{z' \sim D}[f(z')] = E_{z \sim G}[f(\text{trunc}(z))] \approx E_{z \sim G}[f(z)]
\]

\( \Rightarrow \) Suffices to show \( |E_{z \sim G}[f(z)] - E_{x \sim U}[f(x)]| = \tilde{O}\left(\frac{1}{\sqrt{N}}\right) \)
\[ E_{z \sim G}[f(z)] - E_{x \sim U}[f(x)] = \sum_{S \subseteq [2N]} \hat{f}(S) \cdot \left( E_{z \sim G} \left[ \prod_{i \in S} z_i \right] - E_{x \sim U} \left[ \prod_{i \in S} x_i \right] \right) \]

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

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

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

Contribution of first \( \widetilde{O}(\sqrt{N}) \) terms:
\[ \epsilon \cdot \text{polylog}(N)/\sqrt{N} \]

Contribution of larger terms?

Fourier Analytical Approach – First Attempt
Main Technical Lemma

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

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

Then, for any quasi-poly size constant depth $\text{AC}^0$ circuit $f$,

$$|\mathbb{E}_{Z \sim G}[f(Z)] - \mathbb{E}_{x \sim U}[f(x)]| \leq \delta \cdot \text{polylog}(N)$$

Which properties of $\text{AC}^0$ circuits are used in the proof?

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

$G$ fools any class of functions with these two properties.
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

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

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

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 $i = 1, \ldots, t$

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

Observe: $H_t \sim G$.

Taking $t \to \infty$ yields a Brownian motion.
We take $t = \text{poly}(N)$.

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

$$|\mathbb{E}[f(H_{(i+1)})] - \mathbb{E}[f(H_i)]| \leq \frac{\delta}{t} \cdot \text{polylog}(N).$$
Claim - Base Case

Base Case:

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

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

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

\[\leq \frac{\delta}{t} \cdot \text{polylog}(N) + o\left(\frac{\delta}{t}\right) \]
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 $\mathbb{E}_\rho [f_\rho (2 \cdot z) - f_\rho (\vec{0})]$ where $f_\rho$ is a random restriction of $f$ (whose marginals depend on $z_0$).

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

$$|\mathbb{E} [f(H_{i+1})] - \mathbb{E} [f(H_i)]| \leq |\mathbb{E} [f(H_i + \frac{1}{\sqrt{t}} Z^{i+1}) - f(H_i)]| \leq \frac{4\delta}{t} \cdot \text{polylog}(N)$$
Recap: Proof by Picture

[CHHL’18]: i-th step ≈ 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 $AC^0$ (our contribution)
  
  $$|E_{Z \sim G}[f(Z)] - E_{X \sim U}[f(X)]| \leq \delta \cdot \text{polylog}(N):$$

  - **Thought experiment:** Viewing $Z \sim G$ as a result of a random walk with $t$ tiny steps.

  - $AC^0$ circuits are well-approximated by **sparse** low-degree polynomials [T’14]

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

  - [Chattopadhyay, Hatami, Hosseini, Lovett ’18]:

    $\Rightarrow i$-th step has advantage $\left(\frac{\delta}{t}\right) \cdot \text{polylog}(N)$
Follow-ups:

• [Aaronson, Fortnow]: an oracle \( A \) s.t.
\[
\text{BQP}^A \not\subseteq \text{P}^A = \text{NP}^A
\]

• [Fortnow]: under our oracle \( \text{PH} \) is infinite.

Open Problems:

• Does the original suggestion of [Aaronson’09] (without \( 1/\log(N) \) noise) work?

• [Aaronson]: Find an oracle \( A \) s.t.
  - \( \text{NP}^A \subseteq \text{BQP}^A \)
  - \( \text{PH}^A \not\subseteq \text{BQP}^A \)

• [Fortnow]: Does \( \text{NP}^\text{BQP} \not\subseteq \text{BQP}^\text{NP} \)?
Open Problems 2: Pseudorandomness

Separate \( \text{BQLogTime} \) and \( \text{AC}^0[\oplus] \).
Suffices to show for all \( f \) in \( \text{AC}^0[\oplus] \):

\[
\sum_{|S|=2} |\hat{f}(S)| \leq \frac{\sqrt{N}}{\text{polylog}(N)}
\]

**Conjecture [CHLT’18]:** for all \( f \) in \( \text{AC}^0[\oplus] \)

\[
\sum_{|S|=2} |\hat{f}(S)| \leq \text{polylog}(N)
\]

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