Quantum–Classical Equivalence for AND Functions

Yogesh Dahiya
UCSD

Joint work with

Sreejata K. Bhattacharya
Arkadev Chattopadhyay
(TIFR)

Farzan Byramji
Shachar Lovett
(UCSD)

When can quantum algorithms achieve
super-polynomial advantage over classical algorithms?

What structure enables a quantum advantage?

What structure rules it out?

Factoring (Shor)

Exponential speedup

Exploits structure (period finding)

Unstructured Search (Grover)

Quadratic speedup only

No exploitable structure

Factoring (Shor)

Exponential speedup

Exploits structure (period finding)

Unstructured Search (Grover)

Quadratic speedup only

No exploitable structure

Is quantum advantage tied to special structure?

Hypothesis

In absence of special structure,
quantum and classical complexities
are polynomially related.

Hypothesis

In absence of special structure,
quantum and classical complexities
are polynomially related.

How do we test this?

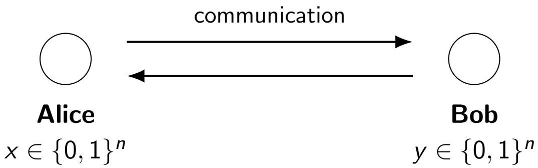
Study concrete computational models.

Lack of structure = f is total

Lack of structure = f is total

Classical and quantum query complexity
are polynomially equivalent for total functions.

Communication Model



$$F : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$$

Compute $F(x, y)$ with minimal communication.

$$D^{cc}(F), \qquad R^{cc}(F), \qquad Q^{cc}(F)$$

For total functions F,

$$R^{cc}(F) \stackrel{?}{=} (Q^{cc}(F))^{O(1)}$$

Are classical and quantum communication polynomially related?

Wide open.

A Natural Approach

The general communication problem seems out of reach.

Restrict to composed functions

$$F = f \circ g$$

Can we understand quantum vs classical power
in such structured subclasses?

AND-Functions

Let $f : \{0,1\}^n \to \{0,1\}$.

$$F(x,y) = f\big(\mathrm{AND}_2(x_1,y_1),\ldots,\mathrm{AND}_2(x_n,y_n)\big)$$

They capture central problems:

Set-Disjointness

$$\mathrm{DISJ}_n = \mathrm{OR}_n \circ \mathrm{AND}_2$$

Inner Product

$$\mathrm{IP}_n = \mathrm{PARITY}_n \circ \mathrm{AND}_2$$

Prior Progress

Bhurman, de Wolf (2001) Initiated study of AND-functions.

Razborov (2003) For symmetric f:

$$R^{cc}(f \circ \mathrm{AND}_2) = (Q^{cc}(f \circ \mathrm{AND}_2))^2.$$

Open: extend to *all* Boolean f.

Our Result

Main Result. For every total Boolean function f,

$$D^{cc}(f \circ \mathrm{AND}_2) = \left(Q^{cc}(f \circ \mathrm{AND}_2)\right)^7 (\log n)^2.$$

$D^{cc},\ R^{cc},\ Q^{cc}$ are polynomially equivalent for AND-functions.

Our Result

Main Result. For every total Boolean function f,

$$D^{cc}(f \circ \mathrm{AND}_2) \leq \left(Q^{cc}(f \circ \mathrm{AND}_2)\right)^7 (\log n)^2.$$

$D^{cc},\ R^{cc},\ Q^{cc}$ are polynomially equivalent for AND-functions.

Corollary. $Q^{cc}(\mathrm{DISJ}_n) = \widetilde{\Omega}(n^{1/7})$

No super-polynomial quantum advantage within this natural class.

More

For $F = f \circ \mathrm{AND}_2$:

Communication measures: $D^{cc}(F),\ R^{cc}(F),\ Q^{cc}(F)$

Matrix measures: $\log \mathrm{rank}(F),\ \log \widetilde{\mathrm{rank}}(F),\ \log \gamma_2(F),\ \log \widetilde{\gamma}_2(F)$

Query / algebraic: $D^{\wedge\text{-}dt}(f),\ \log \mathrm{spar}(f)$

All are polynomially related.

More

For $F = f \circ \mathrm{AND}_2$:

Communication measures: $D^{cc}(F),\ R^{cc}(F),\ Q^{cc}(F)$

Matrix measures: $\log \mathrm{rank}(F),\ \log \widetilde{\mathrm{rank}}(F),\ \log \gamma_2(F),\ \log \widetilde{\gamma}_2(F)$

Query / algebraic: $D^{\wedge\text{-}dt}(f),\ \log \mathrm{spar}(f)$

All are polynomially related.

In particular, the **log approximate-rank conjecture** holds for AND-functions.

Key Tool: Polynomials

Every $f : \{0,1\}^n \to \{0,1\}$ has a unique multilinear polynomial P_f

$$P_f(x) = \sum_{S \subseteq [n]} a_S \prod_{i \in S} x_i.$$

$\deg(f) := \deg(P_f)$ $\qquad\qquad$ $\mathrm{spar}(f) := \#\{\text{monomials in } P_f\}$

$$\widetilde{\deg}(f) = \min_{\substack{p \colon |p(x) - f(x)| \leq 1/3 \\ \forall x}} \deg(p).$$

High Level Proof Idea

Goal:

$$D^{cc}(f \circ \mathrm{AND}_2) \leq \left(Q^{cc}(f \circ \mathrm{AND}_2)\right)^{O(1)}.$$

Characterization of D^{cc} Knop, Lovett, McGuire, Yuan (2021)

$$D^{cc}(f \circ \mathrm{AND}_2) = O\big((\log \mathrm{spar}(f))^5 \log n\big).$$

High Level Proof Idea

Goal:

$$D^{cc}(f \circ \mathrm{AND}_2) \leq \left(Q^{cc}(f \circ \mathrm{AND}_2)\right)^{O(1)}.$$

Characterization of D^{cc} Knop, Lovett, McGuire, Yuan (2021)

$$D^{cc}(f \circ \mathrm{AND}_2) = O\big((\log \mathrm{spar}(f))^5 \log n\big).$$

Therefore, it suffices to prove

$$(\log \mathrm{spar}(f))^{\Omega(1)} \leq Q^{cc}(f \circ \mathrm{AND}_2).$$

From Quantum Protocols to Rectangles

A quantum protocol using c qubits implies:

$$M_F \approx_{1/3} \sum_i \alpha_i R_i = \sum_i \alpha_i\, g_i(x) h_i(y)$$

where $\sum_i |\alpha_i| \;\leq\; 2^{O(c)}$.

The minimum total weight is, up to constants, is same as **approximate** γ_2 **norm**.

Our main technical result:

$$\log \widetilde{wt}(f \circ \mathrm{AND}_2) \;\geq\; (\log \mathrm{spar}(f))^{\Omega(1)}.$$

Why should large sparsity force large $\widetilde{wt}$?

Rough idea: Restriction Argument (Not Exactly).

From large $\mathrm{spar}(f)$, construct a distribution on restrictions $\rho \sim \mathcal{D}$ such that:

1. With high probability, $\deg\big((f \circ \mathrm{AND}_2)|_\rho\big)$ is large.

Why should large sparsity force large $\widetilde{wt}$?

Rough idea: Restriction Argument (Not Exactly).

From large $\mathrm{spar}(f)$, construct a distribution on restrictions $\rho \sim \mathcal{D}$ such that:

1. With high probability, $\deg\big((f \circ \mathrm{AND}_2)|_\rho\big)$ is large.
2. Any small-weight rectangle decomposition $\Pi = \sum \alpha_i\, g_i(x) h_i(y)$ for $f \circ \mathrm{AND}_2$ simplifies under ρ, $\deg(\Pi|_\rho)$ is small w.h.p.

Why should large sparsity force large $\widetilde{wt}$?

Rough idea: Restriction Argument (Not Exactly).

From large $\mathrm{spar}(f)$, construct a distribution on restrictions $\rho \sim \mathcal{D}$ such that:

1. With high probability, $\deg\big((f \circ \mathrm{AND}_2)|_\rho\big)$ is large.
2. Any small-weight rectangle decomposition $\Pi = \sum \alpha_i\, g_i(x)h_i(y)$ for $f \circ \mathrm{AND}_2$ simplifies under ρ, $\deg(\Pi|_\rho)$ is small w.h.p.
3. Therefore, for some ρ, $(f \circ \mathrm{AND}_2)|_\rho$ has large degree but admits a low-degree approximator — contradicting the fact that for every Boolean q, $\deg(q) \leq O(\widetilde{\deg}(q)^2)$.

What actually works

Restriction + Averaging Argument

From large $\mathrm{spar}(f)$, construct a distribution $\rho \sim \mathcal{D}$ such that:

1. With high probability, $\deg\left(\mathbb{E}_{x_{M(\rho)}}\left[(f \circ \mathrm{AND}_2)|_\rho\right]\right)$ is large.
2. Any small-weight rectangle decomposition $\Pi = \sum_i \alpha_i g_i(x) h_i(y)$ simplifies under restriction and averaging, $\deg\left(\mathbb{E}_{x_{M(\rho)}}[\Pi|_\rho]\right)$ is small w.h.p.
3. Contradicting the degree vs. approximate-degree relation.

What actually works

Restriction + Averaging Argument

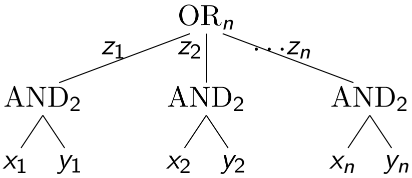
From large $\mathrm{spar}(f)$, construct a distribution $\rho \sim \mathcal{D}$ such that:

1. With high probability, $\deg\left(\mathbb{E}_{x_{M(\rho)}}\left[(f \circ \mathrm{AND}_2)|_\rho\ \right]\right)$ is large.
2. Any small-weight rectangle decomposition $\Pi = \sum_i \alpha_i g_i(x) h_i(y)$ simplifies under restriction and averaging, $\deg\left(\mathbb{E}_{x_{M(\rho)}}[\Pi|_\rho]\right)$ is small w.h.p.
3. Contradicting the degree vs. approximate-degree relation.

Conclusion: Large sparsity $\Rightarrow$ large $\widetilde{wt}$.

Concrete Example: Set-Disjointness

$$\neg\mathrm{DISJ}_n = \mathrm{OR}_n \circ \mathrm{AND}_2$$



High-Level Idea: Restriction + Averaging

1. For a suitable random restriction $\rho \sim \mathcal{D}$, the function $\mathbb{E}_{x_{M(\rho)}}[(\mathrm{OR}_n \circ \mathrm{AND}_2) \mid_\rho]$ has large degree with high probability.
2. Any decomposition $\Pi = \sum_i b_i\, g_i(x) h_i(y) \approx_\varepsilon M_F$ with small $\sum_i |b_i|$ simplifies under restriction and averaging.

Properties of OR_n

- For every restriction $\rho_z \in \{0, \star\}^n$,

$$\deg(\mathrm{OR}_n \mid_{\rho_z}) = |\rho_z^{-1}(\star)|.$$

  $\mathrm{OR}_n \mid_{\rho_z}$ computes OR on the free variables.
- This is exactly what gives OR_n large sparsity.
- We lift these restrictions naturally to restrictions for $\mathrm{OR}_n \circ \mathrm{AND}_2$.

Lifting Restrictions from OR_n to $\mathrm{OR}_n \circ \mathrm{AND}_2$

Given a restriction $\rho_z \in \{0, \star\}^n$, define a lifted restriction

$$\rho := \mathrm{Lift}(\rho_z).$$

For every coordinate $i \in [n]$,

$$(\rho(x_i), \rho(y_i)) = \begin{cases} (\Delta, 0), & \rho_z(z_i) = 0, \\ (\star, 1), & \rho_z(z_i) = \star. \end{cases}$$

Lifting Restrictions from OR_n to $\mathrm{OR}_n \circ \mathrm{AND}_2$

Given a restriction $\rho_z \in \{0, \star\}^n$, define a lifted restriction

$$\rho := \mathrm{Lift}(\rho_z).$$

For every coordinate $i \in [n]$,

$$(\rho(x_i), \rho(y_i)) = \begin{cases} (\Delta, 0), & \rho_z(z_i) = 0, \\ (\star, 1), & \rho_z(z_i) = \star. \end{cases}$$

- ρ induces ρ_z on the z-variables: $(x_i \wedge y_i)\,|_\rho = \rho_z(z_i)$.
- Define the masked variables $M(\rho) := \{x_i : \rho(x_i) = \Delta\}$.
  The masked variables remain free, but $(\mathrm{OR}_n \circ \mathrm{AND}_2)\,|_\rho$ is independent of them.
- Hence $(\mathrm{OR}_n \circ \mathrm{AND}_2)\,|_\rho$ computes OR on the $\star$-variables, and therefore

$$\deg\left(\mathbb{E}_{x_{M(\rho)}}\left[(\mathrm{OR}_n \circ \mathrm{AND}_2)\,|_\rho\right]\right) = |\rho^{-1}(\star)|.$$

Restriction-and-Averaging Process

- Sample a random restriction
$$\rho_z \in \{0, \star\}^n$$
having exactly pn stars uniformly at random, where $p = \Theta(n^{-1/2})$.
- Lift ρ_z to a restriction
$$\rho := \mathrm{Lift}(\rho_z).$$
Let $\mathcal{D}$ denote the resulting distribution over lifted restrictions.
- Apply the restriction-and-averaging operator:
$$\mathbb{E}_{x_{M(\rho)}}[\,\cdot\,], \qquad \rho \sim \mathcal{D}.$$

$\mathrm{OR}_n \circ \mathrm{AND}_2$ Retains Hardness

For every $\rho \sim \mathcal{D}$,

$$\deg\left(\mathbb{E}_{x_{M(\rho)}}\left[(\mathrm{OR}_n \circ \mathrm{AND}_2)\mid_\rho\ \right]\right) = |\rho^{-1}(\star)| = pn = \Theta(n^{1/2}).$$

Simplification of Small-Weight Approximator for $\mathrm{OR}_n \circ \mathrm{AND}_2$

Let $\Pi = \sum_i b_i g_i(x) h_i(y)$ approximate $\mathrm{OR}_n \circ \mathrm{AND}_2$ and suppose $\sum_i |b_i| \le 2^{O(n^{1/4})}$.

Observations / Claims

1. All y-variables are fixed under every $\rho \sim \mathcal{D}$.
2. Under ρ, each rectangle collapses to a function only of the x-variables.
3. (Fourier tail decay property) For every Boolean function $g : \{0,1\}^n \to \{0,1\}$, the expected Fourier mass above level k decays as $2^{-\Omega(k)}$ under the restriction-and-averaging procedure. For $G_\rho := \mathbb{E}_{x_{M(\rho)}}[g|_\rho]$,

$$\mathbb{E}_{\rho\sim\mathcal{D}}\big[\|\widehat{G}_\rho\|_1^{\ge k}\big] \le 2^{-\Omega(k)}.$$

4. Using the tail bound with $k = n^{1/4}$, there is a restriction ρ such that

$$\left\|\widehat{\mathbb{E}_{x_{M(\rho)}}[\Pi|_\rho]}\right\|_1^{\ge k} \le \frac{1}{10}.$$

Putting Things Together

- For every $\rho \in \mathcal{D}$,

$$\deg\Big(\mathbb{E}_{x_{M(\rho)}}\left[(\mathrm{OR}_n \circ \mathrm{AND}_2)\mid_\rho\right]\Big) = \Theta(\sqrt{n}).$$

- Let Π be a $1/3$-approximator for $\mathrm{OR}_n \circ \mathrm{AND}_2$ with $\sum_i |b_i| \leq 2^{O(n^{1/4})}$.
- We saw that there exists $\rho \in \mathcal{D}$ such that, after restriction-and-averaging, the Fourier tail above level $k = n^{1/4}$ has ℓ_1-mass at most $1/10$.
- Truncate above degree k. This gives a degree $n^{1/4}$ approximator to a Boolean function of degree $\Theta(\sqrt{n})$, contradicting

$$\widetilde{\deg}(f) \geq c\sqrt{\deg(f)}.$$

Thus $\sum_i |b_i| \geq 2^{\Omega(n^{1/4})}$.

Step Back: What Structure Did We Use?

For every restriction $\rho_z \in \{0, \star\}^n$,

$$\deg(\mathrm{OR}_n|_{\rho_z}) = |\rho_z^{-1}(\star)|.$$

Step Back: What Structure Did We Use?

For every restriction $\rho_z \in \{0, \star\}^n$,

$$\deg(\mathrm{OR}_n|_{\rho_z}) = |\rho_z^{-1}(\star)|.$$

The same is true for parity:

$$\deg(\mathrm{PARITY}_n|_{\rho_z}) = |\rho_z^{-1}(\star)|.$$

Therefore,

$$\widetilde{wt}(\mathrm{PARITY}_n \circ \mathrm{AND}_2) = 2^{\Omega(n^{1/4})}.$$

Towards General Functions with Large Sparsity

Can we always find such a restriction structure whenever sparsity is large?

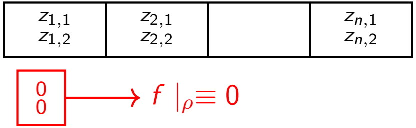
Perhaps after fixing some variables?

Towards General Functions with Large Sparsity

Can we always find such a restriction structure whenever sparsity is large?

Perhaps after fixing some variables?

$$f := \mathrm{AND}_n \circ \mathrm{OR}_2, \qquad \mathrm{spar}(f) = 3^n.$$



Structure of Non-Sparse Functions

Suppose f has large sparsity.

Then we can find a set of *core variables* $V \subseteq \{x_1, \ldots, x_n\}$ such that:

$$\forall \alpha \in \{0, \star\}^V, \quad \exists\, \beta(\alpha) \in \{0,1\}^{[n] \setminus V}$$

for which the restriction $\rho_z := \alpha \cup \beta(\alpha)$ satisfies $\deg(f\,|_{\rho_z}) = |\rho_z^{-1}(\star)|$.

Structure of Non-Sparse Functions

Suppose f has large sparsity.

Then we can find a set of *core variables* $V \subseteq \{x_1, \ldots, x_n\}$ such that:

$$\forall \alpha \in \{0, \star\}^V, \quad \exists\, \beta(\alpha) \in \{0,1\}^{[n] \setminus V}$$

for which the restriction $\rho_z := \alpha \cup \beta(\alpha)$ satisfies $\deg(f\mid_{\rho_z}) = |\rho_z^{-1}(\star)|$.

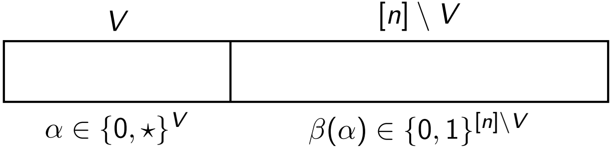
Define the family of restrictions

$$\mathcal{R} := \left\{ \alpha \cup \beta(\alpha) : \alpha \in \{0, \star\}^V \right\}.$$

We will call them *semi-adaptive max-degree restriction trees*.

Structure of Semi-Adaptive Restrictions

A restriction in $\mathcal{R}$ has the following form:



$$\rho_z = \alpha \cup \beta(\alpha), \qquad \deg(f\mid_{\rho_z}) = |\rho_z^{-1}(\star)|.$$

The restriction process is adaptive only on the non-core variables.

Lifting Restrictions to $f \circ \mathrm{AND}_2$

Given $\rho_z \in \mathcal{R}$, define a lifted restriction $\rho = \mathrm{Lift}(\rho_z)$ on the variables (x_i, y_i).

For each coordinate i,

$$(\rho(x_i), \rho(y_i)) = \begin{cases} (\Delta, 0), & \rho_z(z_i) = 0 \text{ and } i \in V, \\ (\star, 1), & \rho_z(z_i) = \star \text{ and } i \in V, \\ (1, 1), & \rho_z(z_i) = 1 \text{ and } i \notin V, \\ (1, 0), & \rho_z(z_i) = 0 \text{ and } i \notin V. \end{cases}$$

Then

$$(x_i \wedge y_i) \mid_\rho = \rho_z(z_i).$$

Properties of Lifted Restrictions

Let $\rho = \mathrm{Lift}(\rho_z)$.

1. The lifted restriction induces ρ_z on the z-variables. Hence
$$\deg((f \circ \mathrm{AND}_2)\restriction_{\rho}) = \deg(f\restriction_{\rho_z}) = |\rho_z^{-1}(\star)| = |\rho^{-1}(\star)|.$$

2. $(f \circ \mathrm{AND}_2)\restriction_{\rho}$ is independent of the masked variables. Therefore,
$$\deg\Big(\mathbb{E}_{x_{M(\rho)}}\big[(f \circ \mathrm{AND}_2)\restriction_{\rho}\big]\Big) = |\rho^{-1}(\star)|.$$

3. All y-variables are fixed, and every non-core x-variable is fixed to 1.

General Functions with Large Sparsity

The lifted restrictions preserve hardness under restriction-and-averaging:

$$\deg\Big(\mathbb{E}_{x_{M(\rho)}}\big[(f \circ \mathrm{AND}_2)\mid_\rho\big]\Big) = |\rho^{-1}(\star)|.$$

Therefore, the previous lower-bound framework applies to $f \circ \mathrm{AND}_2$.

The simplification of small-weight decompositions still works.

- Any function on the x-variables has exponential Fourier tail decay under restriction-and-averaging.
- The only change in the analysis is that the ambient dimension n is replaced by the number of core variables $|V|$.

Semi-Adaptive Max-Degree Restriction Trees

Theorem

$f : \{0,1\}^n \to \{0,1\}$ *admits a semi-adaptive max-degree restriction tree of depth*

$$\Omega\left(\frac{\log \mathrm{spar}(f)}{\log n}\right).$$

Theorem

For every $f : \{0,1\}^n \to \{0,1\}$,

$$\log \widetilde{wt}(f \circ \mathrm{AND}_2) = \Omega\left(\left(\frac{\log \mathrm{spar}(f)}{\log n}\right)^{1/4}\right).$$

Semi-Adaptive Max-Degree Restriction Trees

Theorem

$f : \{0,1\}^n \to \{0,1\}$ admits a semi-adaptive max-degree restriction tree of depth

$$\Omega\left(\frac{\log \mathrm{spar}(f)}{\log n}\right).$$

Theorem

For every $f : \{0,1\}^n \to \{0,1\}$,

$$\log \widetilde{wt}(f \circ \mathrm{AND}_2) = \Omega\left(\left(\frac{\log \mathrm{spar}(f)}{\log n}\right)^{1/4}\right).$$

$$D^{cc}(f \circ \mathrm{AND}_2) = O(D^{\wedge\text{-dt}}(f)) = O\big((\log \mathrm{spar}(f))^{O(1)} \log n\big)$$

$$\log \widetilde{wt}(f \circ \mathrm{AND}_2) = O(Q^{cc}(f \circ \mathrm{AND}_2).$$

Constructing Semi-Adaptive Restriction Trees

Theorem

Let f have sparsity s. Then f admits a semi-adaptive max-degree restriction tree of depth $\Omega\left(\frac{\log s}{\log n}\right)$.

Proof Idea. Write

$$f(x) = \sum_{S \subseteq [n]} a_S \prod_{i \in S} x_i, \qquad \mathcal{M}_f := \{S : a_S \neq 0\}.$$

Since $|\mathcal{M}_f| = s$, Sauer–Shelah gives a shattered set

$$V \subseteq [n], \qquad |V| = \Omega\left(\frac{\log s}{\log n}\right).$$

Shattering means: for every $A \subseteq V$, some monomial has intersection with V exactly A.

Thus, after choosing the free core variables, we can fix non-core variables so that a full-degree monomial on the free variables survives.

Our Results

For $F = f \circ \mathrm{AND}_2$:

$$D^{cc}(F),\ R^{cc}(F),\ Q^{cc}(F)$$

$$\log \mathrm{rank}(F),\ \log \widetilde{\mathrm{rank}}(F),\ \log \gamma_2(F),\ \log \widetilde{\gamma}_2(F)$$

$$D^{\wedge\text{-}dt}(f),\ \log \mathrm{spar}(f)$$

All are polynomially related.

Thank You