APPROXIMATE DEGREE AND QUANTUM QUERY LOWER BOUNDS VIA DUAL POLYNOMIALS

14 Agosto 2018

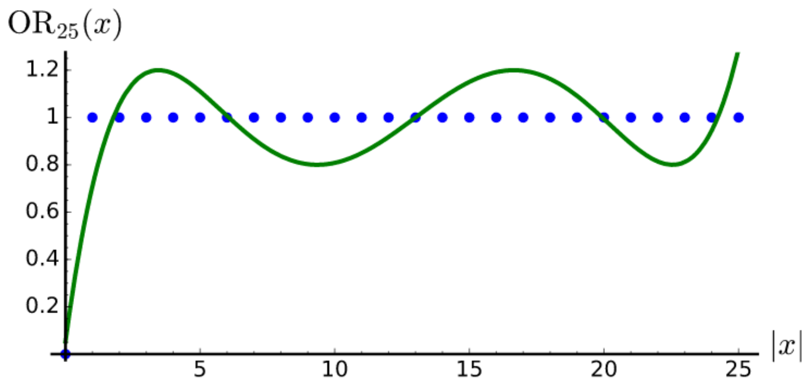
Mark Bun	Princeton → Simons & Boston U.
Robin Kothari	Microsoft Research
Justin Thaler	Georgetown

Approximate Degree [Nisan-Szegedy92]

For a Boolean function $f : \{0,1\}^n \rightarrow \{0,1\}$

Approximate Degree: Minimum degree of a real polynomial $p : \{0,1\}^n \rightarrow \mathbb{R}$ such that

$$|p(x) - f(x)| \leq 1/3 \quad \text{for all } x \in \{0,\ 1\}^n$$



Denoted by $\mathrm{adeg}(f)$

Ex. $\mathrm{adeg}(\mathrm{OR}_n) = \Theta(\sqrt{n})$

Research Directions for the Polynomial Method

(1) Advance our understanding of adeg

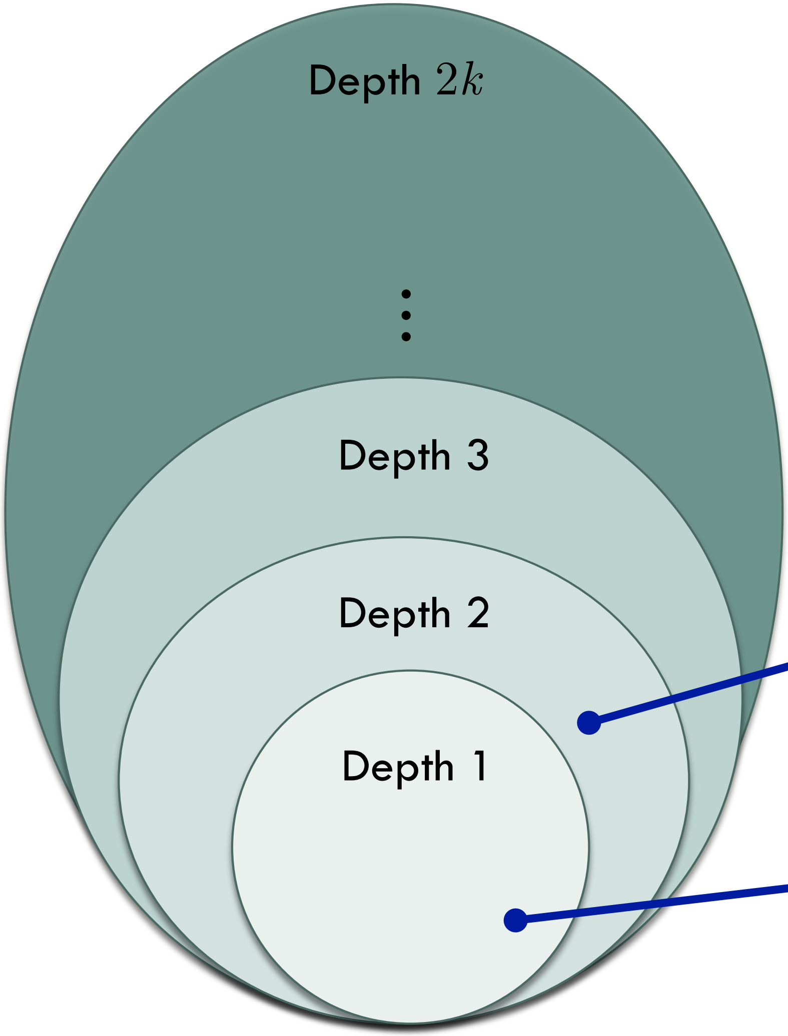
A Nearly Optimal Lower Bound on the Approximate Degree of AC^0

Hardness amplification within AC^0

(2) Use adeg to advance application domains

The Polynomial Method Strikes Back:
Tight Quantum Query Bounds via Dual Polynomials

Approximate Degree of AC^0



Q: Is there an AC^0 function with approximate degree $\Omega(n)$?

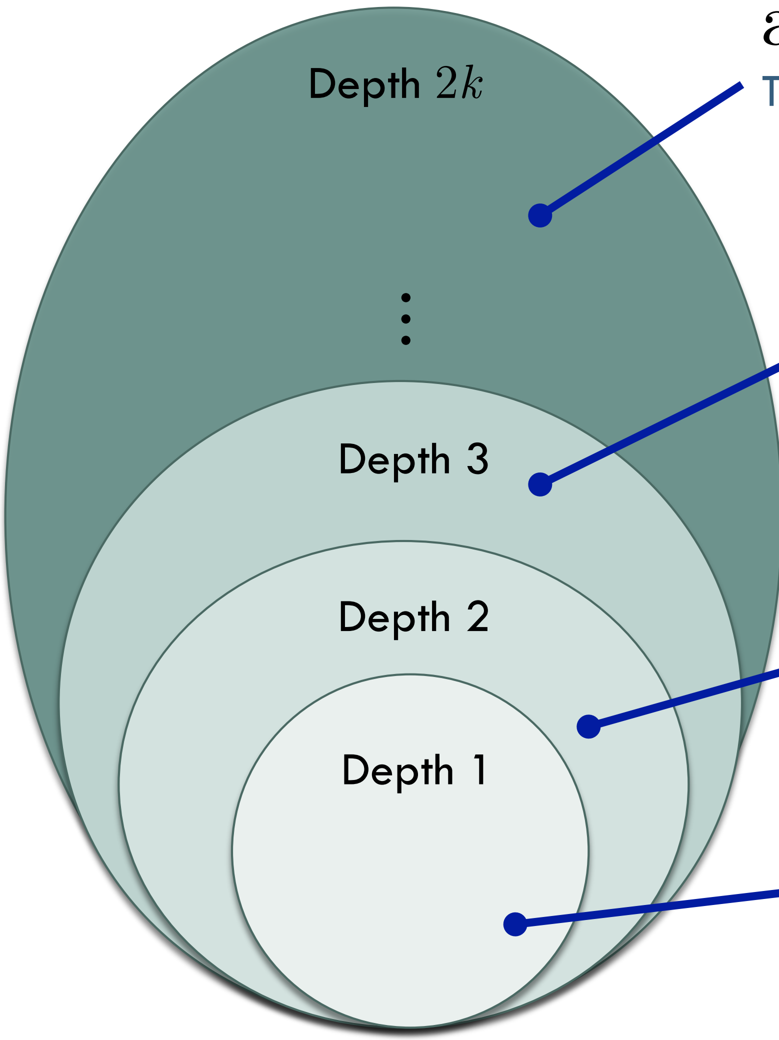
$\mathrm{adeg}(\text{Element-Dist}_n) = \Omega(n^{2/3})$

[Aaronson-Shi01]

$\mathrm{adeg}(\mathrm{OR}_n) = \Omega(n^{1/2})$

[Nisan-Szegedy92]

Approximate Degree of AC^0



$\mathrm{adeg}(F_{k,n}) = \Omega(n^{1-2^{-k}})$ Nearly linear

This work

$\mathrm{adeg}(\mathrm{Surjectivity}_n) = \Omega(n^{3/4})$

This work

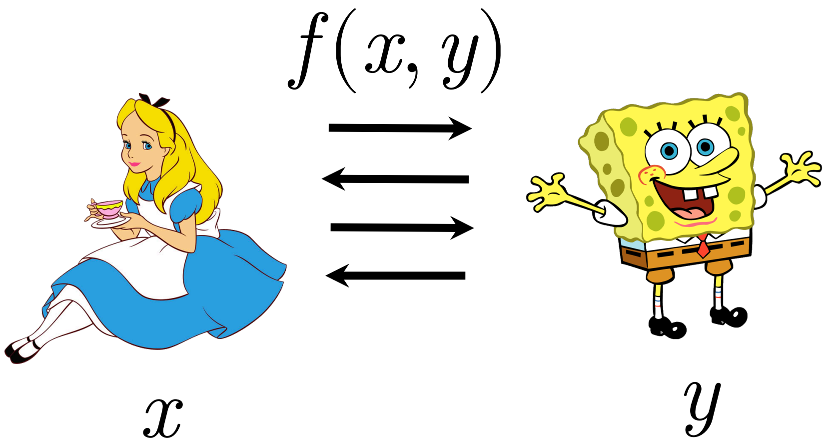
$\mathrm{adeg}(\mathrm{Element\text{-}Dist}_n) = \Omega(n^{2/3})$

[Aaronson-Shi01]

$\mathrm{adeg}(\mathrm{OR}_n) = \Omega(n^{1/2})$

[Nisan-Szegedy92]

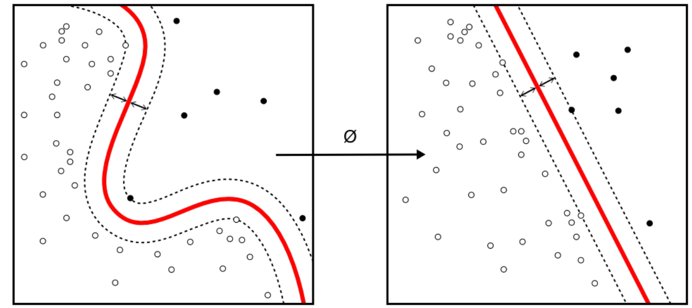
Applications of AC^0 Lower Bound



Nearly optimal $\Omega(n^{1-\delta})$

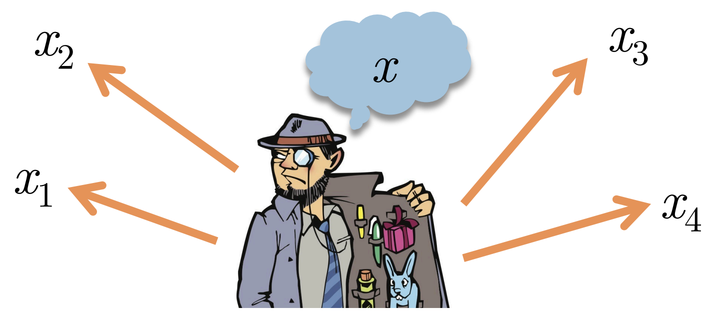
quantum and multiparty

communication lower bounds for AC^0



Learning via regression requires

$\exp(\Omega(n^{1-\delta}))$ features



Improved secret sharing

with reconstruction in AC^0

Research Directions for the Polynomial Method

(1) Advance our understanding of adeg

A Nearly Optimal Lower Bound on the Approximate Degree of AC^0

Hardness amplification within AC^0

(2) Use adeg to advance application domains

The Polynomial Method Strikes Back:
Tight Quantum Query Bounds via Dual Polynomials

(Deterministic) Query Complexity

Let $f : \{0, 1\}^n \rightarrow \{0, 1\}$ be a boolean function

Deterministic Query Complexity:

Minimum number of bits of x that must be read to compute $f(x)$

Ex. Computing OR_n requires n queries

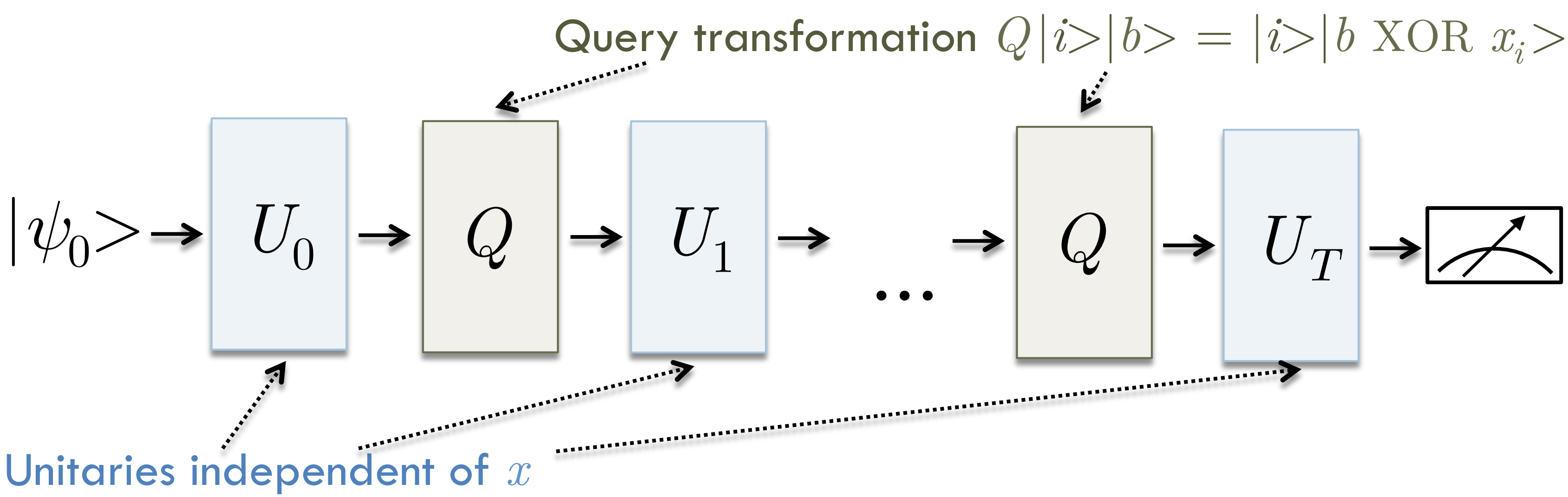
x_1	x_2	x_3	x_4	...	$x_{n\text{-}1}$	x_n

Quantum Query Complexity

Let $f : \{0, 1\}^n \rightarrow \{0, 1\}$ be a boolean function

Quantum Query Complexity:

Minimum number of bits of x that must be read in superposition to compute $f(x)$ with probability $\geq 2/3$



Quantum Query Complexity

Let $f : \{0, 1\}^n$ → $\{0, 1\}$ be a boolean function

Quantum Query Complexity:

Minimum number of bits of x that must be read in superposition to compute $f(x)$ with probability $\geq 2/3$

Ex. Computing OR_n

only needs $\sqrt{n}$ quantum queries [Grover96]

Quantum Query Lower Bounds

"The Polynomial Method" [Beals-Buhrman-Cleve-Mosca-deWolf98]:

Accept prob. of a T query algorithm = Degree $2T$ polynomial in x

$\Rightarrow$ Quantum-query-complexity$(f) \geq$ ½ $\mathrm{adeg}(f)$

Newer "adversary" methods:

- Positive-weights method [Ambainis02]

  Easy to apply, but limited in power

- Negative-weights method [Høyer-Lee-Špalek07, …, Reichardt11]

  Tight characterization, but difficult to apply

This work: New and nearly tight quantum query lower bounds via the polynomial method

Our Results

Problem	Best Prior Upper Bound	Our Lower Bound	Best Prior Lower Bound
k-distinctness	$O(n^{3/4-1/(2^{k+2}-4)})$ [Bel12a]	$\tilde{\Omega}(n^{3/4-1/(2k)})$	$\tilde{\Omega}(n^{2/3})$ [AS04]
Image Size Testing	$O(\sqrt{n}\log n)$ [ABRdW16]	$\tilde{\Omega}(\sqrt{n})$	$\tilde{\Omega}(n^{1/3})$ [ABRdW16]
k-junta Testing	$O(\sqrt{k}\log k)$ [ABRdW16]	$\tilde{\Omega}(\sqrt{k})$	$\tilde{\Omega}(k^{1/3})$ [ABRdW16]
SDU	$O(\sqrt{n})$ [BHH11]	$\tilde{\Omega}(\sqrt{n})$	$\tilde{\Omega}(n^{1/3})$ [BHH11, AS04]
Shannon Entropy	$\tilde{O}(\sqrt{n})$ [BHH11, LW17]	$\tilde{\Omega}(\sqrt{n})$	$\tilde{\Omega}(n^{1/3})$ [LW17]

Table 1: Our lower bounds on quantum query complexity and approximate degree vs. prior work.

Problem	Best Prior Upper Bound	Our Upper Bound	Our Lower Bound	Best Prior Lower Bound
Surjectivity	$\tilde{O}(n^{3/4})$ [She18]	$\tilde{O}(n^{3/4})$	$\tilde{\Omega}(n^{3/4})$	$\tilde{\Omega}(n^{2/3})$ [AS04]

Table 2: Our bounds on the approximate degree of Surjectivity vs. prior work.

Lower Bound for k-distinctness

Define $k\text{-DIST}_{N,R}$: $\{1 ,\ldots, R\}^N$ → $\{0, 1\}$ by

$k\text{-DIST}_{N,R}\,(s_1, \ldots, s_N) = 1$ iff

Some $r \in [R]$ appears $\geq k$ times in the input list

Corresponds to a Boolean function on $\mathrm{O}(N \log_2 R)$ bits

Upper Bounds: $\mathrm{O}(N^{k/(k+1)})$ [Ambainis03] via quantum walks

$\mathrm{O}(N^{3/4-1/\exp(k)})$ [Belovs12] via learning graphs

Lower Bounds: $\Omega(N^{2/3})$ [Aaronson-Shi01] via polynomial method

This work: $\Omega(N^{3/4-1/(2k)})$ via polynomial method

Our Results

$\Omega(n^{1-\delta})$ lower bound for AC^0

Recursive Application

$O(n^{3/4})$ upper bound for Surjectivity

Intuition & Ideas

$\Omega(n^{3/4})$ lower bound for Surjectivity

New proof of [Sherstov17]

Common Proof Strategy

Qualitatively matches [Belovs12]	k-Distinctness:	$\Omega(n^{3/4-1/(2k)})$
Matches [Ambainis-Belovs-Regev-deWolf16]	Image Size Testing:	$\Omega(n^{1/2})$

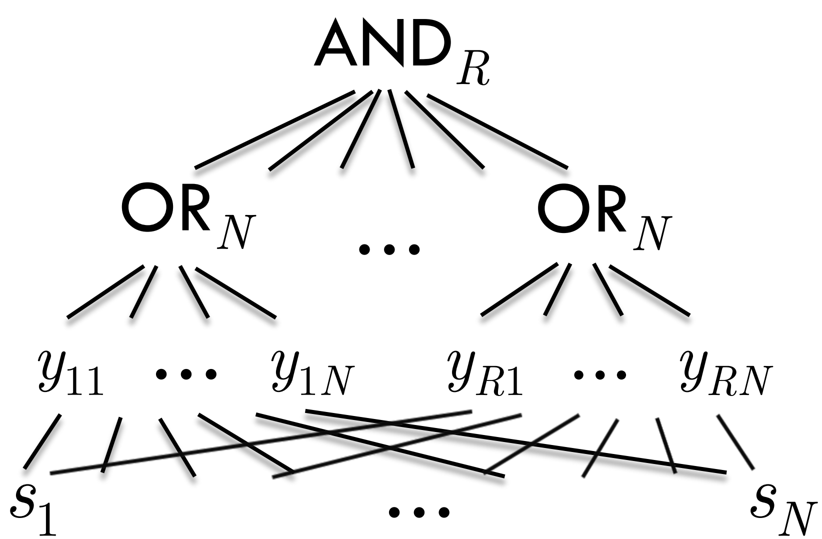
Reductions

Matches [Ambainis-Belovs-Regev-deWolf16]	Statistical Dist. from Uniform:	$\Omega(n^{1/2})$
Matches [Bravyi-Harrow-Hassidim11]	k-Junta Testing:	$\Omega(k^{1/2})$
Matches [Li-Wu17]	Shannon Entropy:	$\Omega(n^{1/2})$

Lower Bound Roadmap

1. Prove a hardness amplification theorem for functions in AC^0

2. Express Surjectivity, k-Distinctness, etc. as amplified versions of functions we understand



Hardness Amplification in AC^0

<u>Theorem 1:</u> If $\mathrm{adeg}(f) > d$, then $\mathrm{adeg}(F) > t^{1/2} d$
for $F = \mathrm{OR}_t \circ f$ [B.-Thaler13, Sherstov13, BenDavid-Bouland-Garg-Kothari17]

<u>Theorem 2:</u> If $\mathrm{adeg}_{-}(f) > d$, then $\mathrm{adeg}_{1-2^{-t}}(F) > d$
for $F = \mathrm{OR}_t \circ f$ [B.-Thaler14]

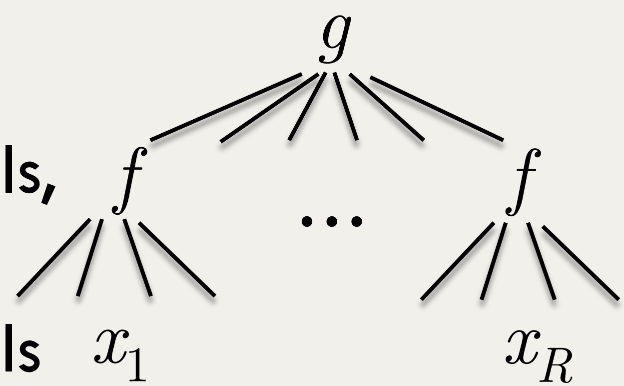
<u>Theorem 3:</u> If $\mathrm{adeg}_{-}(f) > d$, then $\deg_{\pm}(F) > \min\{t, d\}$
for $F = \mathrm{OR}_t \circ f$ [Sherstov14]

<u>Theorem 4:</u> If $\mathrm{adeg}_{+}(f) > d$, then $\mathrm{adeg}_{1-2^{-t}}(F) > d$
for $F = \text{ODD-MAX-BIT}_t \circ f$ [Thaler14]

<u>Theorem 5:</u> If $\mathrm{adeg}(f) > d$, then $\deg_{\pm}(F) > \min\{t, d\}$
for $F = \text{APPROX-MAJ}_t \circ f$ [Bouland-Chen-Holden-Thaler-Vasudevan16]

Hardness Amplification

<u>Theorem Template:</u> If f is "hard" to approximate by low-degree polynomials, then $F = g \circ f$ is "even harder" to approximate by low-degree polynomials



<u>Block Composition Barrier</u>

Robust approximations, i.e.,

$$\mathrm{adeg}(g \circ f) \leq \mathrm{O}(\mathrm{adeg}(g) \cdot \mathrm{adeg}(f))$$

imply that block composition *cannot* give better lower bounds than $\sqrt{n}$

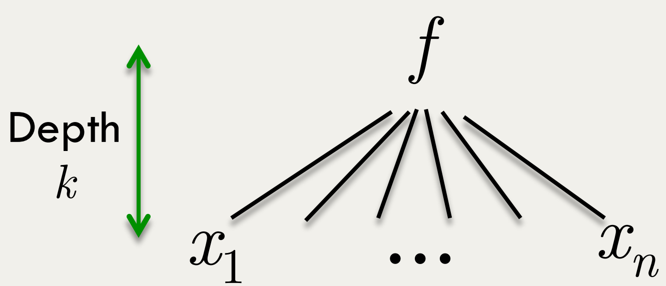
Our Work: A New Hardness Amplification Theorem for Degree

(1) An $\Omega(n^{1-\delta})$ approximate degree lower bound for AC^0

Recursive application

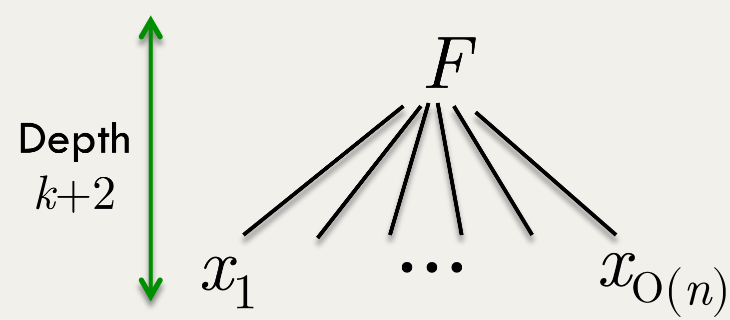
Start with:

$$\text{adeg}(f) \geq d$$



Construct:

$$\text{adeg}(F) \geq \Omega(d^{1/2} \cdot n^{1/2})$$



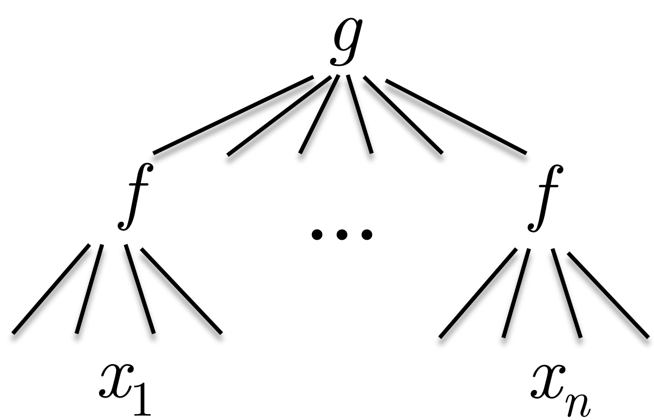
Refined & generalized application

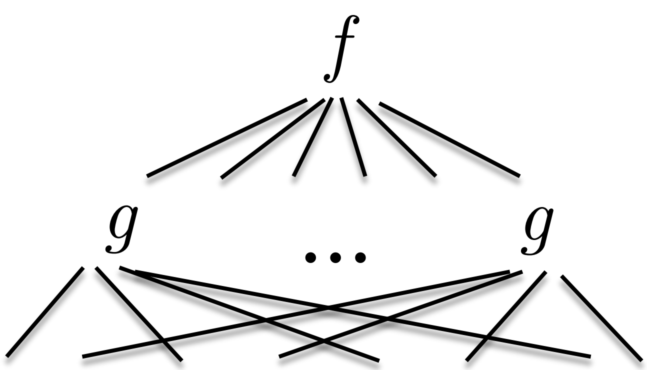
(2) New quantum query lower bounds

Breaking the Block Composition Barrier

<u>Prior work:</u>

- Hardness amplification "from the top"
- Block composed functions

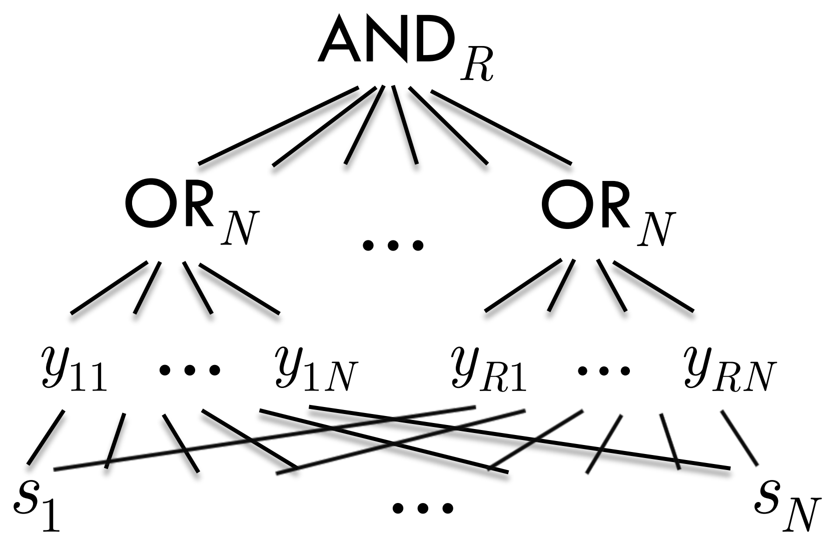




<u>Our new work:</u>

- Hardness amplification "from the bottom"
  - Non-block-composed functions

Remainder of This Talk:
Lower Bound for SURJECTIVITY



Getting to Know Surjectivity

Define $\mathrm{SURJ}_{N,R} : \{1 ,\ldots, R\}^N \to \{0, 1\}$ by

$\mathrm{SURJ}_{N,R}(s_1, \ldots, s_N) = 1$ iff

Every $r \in [R]$ appears in the input list

Corresponds to a Boolean function on $\mathrm{O}(N \log_2 R)$ bits

Has quantum query complexity $\Omega(R)$ [Beame-Machmouchi10] but approximate degree $\mathrm{O}(R^{3/4})$ [Sherstov17]

(For $N = \mathrm{O}(R)$)

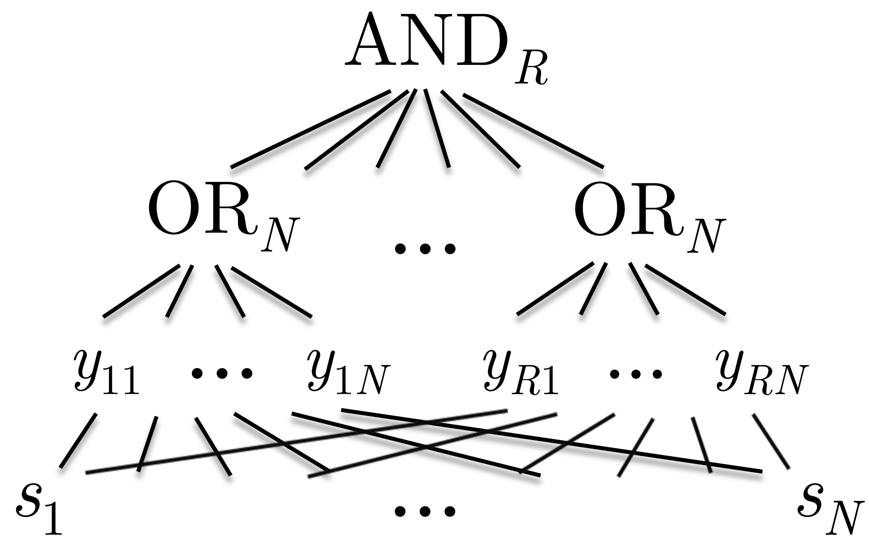
Getting to Know Surjectivity

$\mathrm{SURJ}_{N,R}(s_1, \ldots, s_N) = 1$ iff

Every $r \in [R]$ appears in the input list

Define auxiliary variables

$$y_{r,i}(s) = \begin{cases} 1 & \text{if } s_i = r \\ 0 & \text{otherwise} \end{cases}$$



Then $\mathrm{SURJ}_{N,R}(s_1, \ldots, s_N) =$

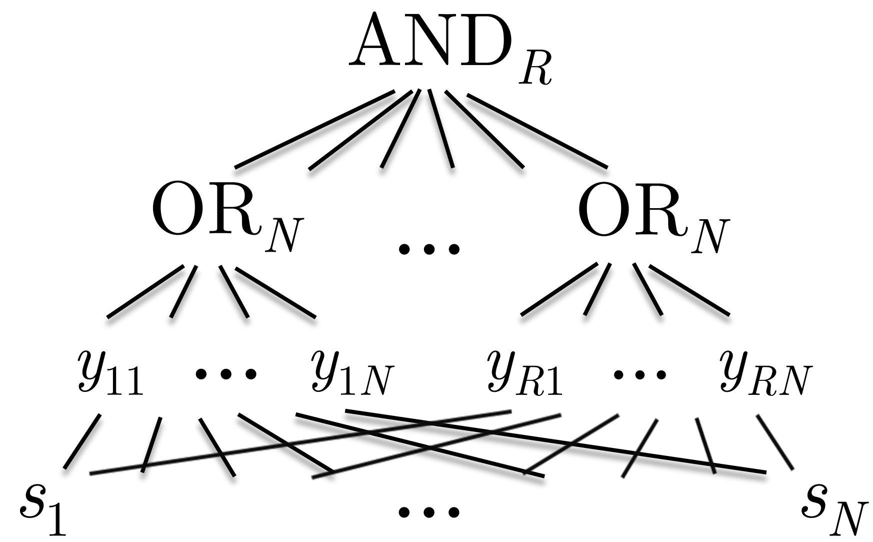
$$\mathrm{AND}_R(\,\mathrm{OR}_N(y_{11}, \ldots, y_{1N}), \ldots, \mathrm{OR}_N(y_{R1}, \ldots, y_{RN})\,)$$

Getting to Know Surjectivity

Observation: To approximate $\mathrm{SURJ}_{N,R}$, suffices to approx. $\mathrm{AND}_R \circ \mathrm{OR}_N$ on inputs of Hamming weight N

Define auxiliary variables

$$y_{r,i}(s) = \begin{cases} 1 & \text{if } s_i = r \\ 0 & \text{otherwise} \end{cases}$$



Then $\mathrm{SURJ}_{N,R}(s_1, \ldots, s_N) =$

$$\mathrm{AND}_R(\ \mathrm{OR}_N(y_{11}, \ldots, y_{1N}), \ldots, \mathrm{OR}_N(y_{R1}, \ldots, y_{RN})\ )$$

Surjectivity Lower Bound

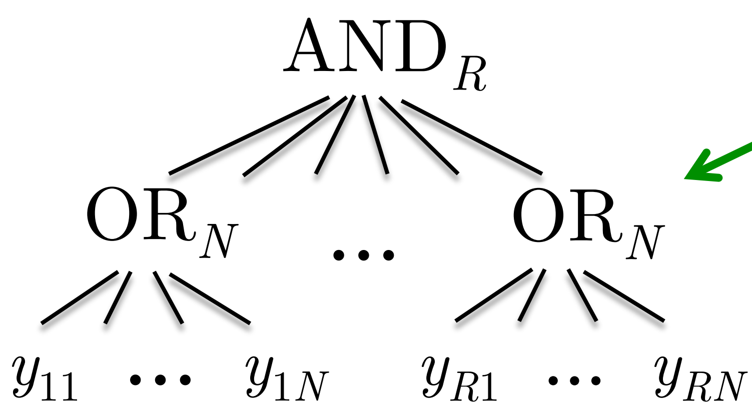
This work:

For some $N = \mathrm{O}(R)$, $\mathrm{adeg}(\mathrm{SURJ}_{N,R}) = \Omega(R^{3/4})$

Stage 1: Reduce to a claim about block composed functions

Lemma: Builds on symmetrization argument of [Ambainis03]

$$\mathrm{adeg}(\mathrm{SURJ}_{N,R}) = \Theta(\mathrm{adeg}(\ (\mathrm{AND}_R \circ \mathrm{OR}_N)^{\leq N}\ )\ )$$



Promised that $|y| \leq N$

Surjectivity Lower Bound

This work:

For some $N = \mathrm{O}(R)$, $\quad \mathrm{adeg}(\mathrm{SURJ}_{N,R}) = \Omega(R^{3/4})$

Stage 2: Prove $\mathrm{adeg}(\ (\mathrm{AND}_R \circ \mathrm{OR}_N)^{\leq N}\ ) = \Omega(R^{3/4})$

Uses method of dual polynomials [Ioffe-Tikhomirov68, Sherstov07, Shi-Zhu07]

Primal

$$\begin{aligned} &\min_{p,\varepsilon} \quad \varepsilon \\ &\text{s.t.} \quad |p(x) - f(x)| \leq \varepsilon \quad \forall x \in \{0,1\}^n \end{aligned}$$

Dual

$$\begin{aligned} &\max_{\Psi} \quad \sum_{x \in \{0,1\}^n} (-1)^{f(x)} \Psi(x) \\ &\text{s.t.} \quad \sum_{x \in \{0,1\}^n} |\Psi(x)| = 1 \\ &\qquad \deg(p) \leq d \implies \sum_{x \in \{0,1\}^n} p(x)\Psi(x) = 0 \end{aligned}$$

Surjectivity Lower Bound

This work:
For some $N = \mathrm{O}(R)$, $\quad \mathrm{adeg}(\mathrm{SURJ}_{N,R}) = \Omega(R^{3/4})$

Stage 2: Prove $\mathrm{adeg}(\ (\mathrm{AND}_R \circ \mathrm{OR}_N)^{\leq N}\ ) = \Omega(R^{3/4})$

Uses method of dual polynomials [Ioffe-Tikhomirov68, Sherstov07, Shi-Zhu07]

From Justin's talk:

Can prove $\mathrm{adeg}(\mathrm{AND}_R \circ \mathrm{OR}_N) = \Omega(R)$ by combining dual polynomials Ψ_{AND} and Ψ_{OR} to construct a dual polynomial $\Psi_{\mathrm{AND\text{-}OR}}$ [B.-Thaler13, Sherstov13]

Details of Stage 2

Claim: $\text{adeg}(\text{AND}_R \circ \text{OR}_N) = \Omega(R^{3/4})$ even under the promise that $|x| \leq N$

is equivalent to

There exists a dual polynomial witnessing $\text{adeg}(\text{AND}_R \circ \text{OR}_N) = \Omega(R^{3/4})$ which is supported on inputs with $|x| \leq N$

Does the dual polynomial we already have for $\text{AND}_R \circ \text{OR}_N$ satisfy this property?

NO

Fixing the AND-OR Dual Polynomial

$$\Psi_{\text{AND-OR}}(x) = 2^R \Psi_{\text{AND}}(\operatorname{sgn} \Psi_{\text{OR}}(x_1), \ldots, \operatorname{sgn} \Psi_{\text{OR}}(x_R)) \prod_{i=1}^{R} |\Psi_{\text{OR}}(x_i)|$$

Ψ_{OR} *must* be nonzero for inputs with Hamming weight up to $\Omega(N)$

$\Rightarrow \Psi_{\text{AND-OR}}$ nonzero up to Hamming weight $\Omega(RN)$

1. $\Psi_{\text{AND-OR}}$ has L_1-norm 1 ✓
2. $\Psi_{\text{AND-OR}}$ has pure high degree $\Omega(R^{1/2}N^{1/2}) = \Omega(R)$ ✓
3. $\Psi_{\text{AND-OR}}$ has high correlation with $\text{AND}_R \circ \text{OR}_N$ ✓
4. $\Psi_{\text{AND-OR}}$ is supported on inputs with $|x| \leq N$ ✗

Fixing the AND-OR Dual Polynomial

$$\Psi_{\text{AND-OR}}(x) = 2^R \Psi_{\text{AND}}(\operatorname{sgn} \Psi_{\text{OR}}(x_1), \ldots, \operatorname{sgn} \Psi_{\text{OR}}(x_R)) \prod_{i=1}^{R} |\Psi_{\text{OR}}(x_i)|$$

Ψ_{OR} *must* be nonzero for inputs with Hamming weight up to $\Omega(N)$

$\Rightarrow \Psi_{\text{AND-OR}}$ nonzero up to Hamming weight $\Omega(RN)$

Fix 1: Trade pure high degree of Ψ_{OR} for "support" size

Fix 2: Zero out high Hamming weight inputs to $\Psi_{\text{AND-OR}}$

Fix 1: Trading PHD for Support Size

For every integer $1 \leq m \leq N$, there is a dual polynomial Ψ^m_{OR} for OR_N which

- has pure high degree $\Omega(m^{1/2})$
- is supported on inputs of Hamming weight $\leq m$

$$\Psi^m_{\text{AND-OR}}(x) = 2^R \Psi_{\mathrm{AND}}(\operatorname{sgn} \Psi^m_{\mathrm{OR}}(x_1), \ldots, \operatorname{sgn} \Psi^m_{\mathrm{OR}}(x_R)) \prod_{i=1}^{R} |\Psi^m_{\mathrm{OR}}(x_i)|$$

Dual polynomial $\Psi^m_{\text{AND-OR}}$

- has pure high degree $\Omega(R^{1/2}\, m^{1/2})$
- is supported on inputs of Hamming weight $\leq mR$

Fix 2: Zeroing Out High Hamming Weight Inputs

Dual polynomial $\Psi^{m}_{\text{AND-OR}}$

- has pure high degree $\Omega(R^{1/2}\, m^{1/2})$
- is supported on inputs of Hamming weight $\leq mR$

Suppose further that $$\sum_{|x|>N} |\Psi^{m}_{\text{AND-OR}}(x)| \ll \text{negl}(R)$$

Can we post-process $\Psi^{m}_{\text{AND-OR}}$ to zero out inputs with Hamming weight $N < |x| \leq mR$…

…without ruining

- pure high degree of $\Psi^{m}_{\text{AND-OR}}$
- correlation between $\Psi^{m}_{\text{AND-OR}}$ and $\text{AND}_R \circ \text{OR}_N$?

YES (Follows from [Razborov-Sherstov-08])

Fix 2: Zeroing Out High Hamming Weight Inputs

Technical Lemma (follows from [Razborov-Sherstov08])

If $0 < D < N$ and

$$\sum_{|x|>N} |\Psi^m_{\text{AND-OR}}(x)| \ll 2^{-D},$$

then there exists a "correction term" Ψ^m_{corr} that

1. Agrees with $\Psi^m_{\text{AND-OR}}$ inputs of Hamming weight $> N$
2. Has L_1-norm 0.01
3. Has pure high degree D

Fix 2: Zeroing Out High Hamming Weight Inputs

Claim: For $1 \leq m \leq N$, $$\sum_{|x|>N} |\Psi^m_{\text{AND-OR}}(x)| \ll 2^{-R/m^{1/2}}$$

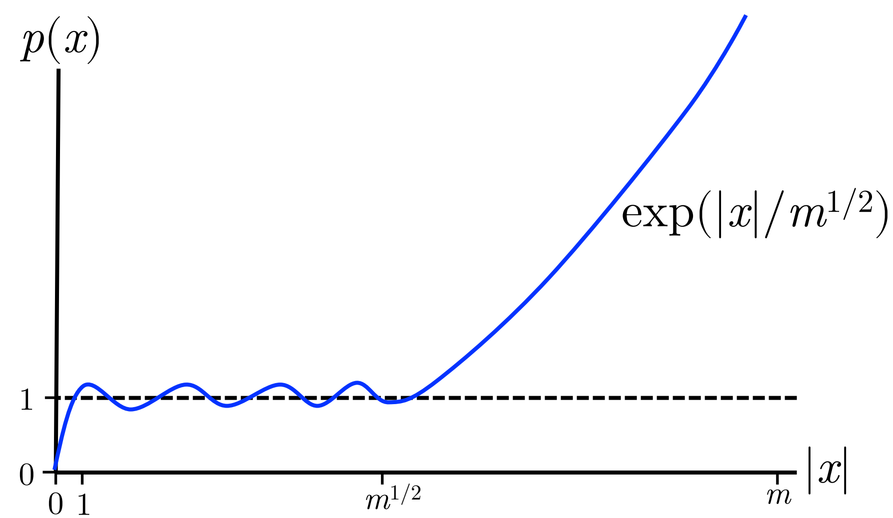
Proof idea:

Ψ^m_{OR} can be made biased toward low Hamming weight inputs:

For all $t > 0$, $$\sum_{|x|=t} |\Psi^m_{\text{OR}}(x)| \lesssim \exp(-t/m^{1/2})$$

Primal interpretation:

Any polynomial that looks like this still has degree $\Omega(m^{1/2})$



Fix 2: Zeroing Out High Hamming Weight Inputs

<u>Claim:</u> For $1 \leq m \leq N$, $$\sum_{|x|>N} |\Psi^m_{\text{AND-OR}}(x)| \ll 2^{-R/m^{1/2}}$$

<u>Proof idea:</u>

Ψ^m_{OR} can be made biased toward low Hamming weight inputs:

For all $t > 0$, $$\sum_{|x|=t} |\Psi^m_{\text{OR}}(x)| \lesssim \exp(-t/m^{1/2})$$

$\Rightarrow$ "Worst" high Hamming weight inputs look like

$|x_1| = m^{1/2}$, …, $|x_{N/m^{1/2}}| = m^{1/2}$, $|x_{(N/m^{1/2})+1}| = 0$, …, $|x_R| = 0$

$$\Psi^m_{\text{AND-OR}}(x) = 2^R \Psi_{\text{AND}}(\text{sgn}\, \Psi^m_{\text{OR}}(x_1), \ldots, \text{sgn}\, \Psi^m_{\text{OR}}(x_R)) \boxed{\prod_{i=1}^{R} |\Psi^m_{\text{OR}}(x_i)|}$$

Weight on such inputs looks like $2^{-N/m^{1/2}}$

Putting the Pieces Together

Dual polynomial $\Psi^{m}_{\text{AND-OR}}$ Fix 1

- has pure high degree $\Omega(R^{1/2}\, m^{1/2})$
- satisfies $\sum_{|x|>N} |\Psi^{m}_{\text{AND-OR}}(x)| \ll 2^{-R/m^{1/2}}$

Balanced at $m = R^{1/2}$
$\Rightarrow$ PHD $\Omega(R^{3/4})$

Correction term Ψ^{m}_{corr}

- has pure high degree $\Omega(R/m^{1/2})$
- agrees with $\Psi^{m}_{\text{AND-OR}}$ inputs of Hammi eight $>N$

$\Rightarrow \Psi_{\text{AND-OR}} = \boxed{\Psi^{m}_{\text{AND-OR}} - \Psi^{m}_{\text{corr}}}$ has

1. L_1-norm $\approx$ 1
2. high correlation with $\text{AND}_R \circ \text{OR}_N$
3. pure high degree $\Omega(\min\{R^{1/2}m^{1/2},\ R/m^{1/2}\})$
4. support on inputs with $|x| \le N$

Recap of SURJECTIVITY Lower Bound

This work:

For some $N = \mathrm{O}(R)$, $\quad \mathrm{adeg}(\mathrm{SURJ}_{N,R}) = \Omega(R^{3/4})$

Stage 1: Apply *symmetrization* to reduce to

Builds on [Ambainis03]

Claim: $\mathrm{adeg}(\mathsf{AND}_R \circ \mathsf{OR}_N) = \Omega(R^{3/4})$ *even under the promise* that $|x| \leq N$

Stage 2: Prove Claim via *method of dual polynomials*

Refines AND-OR dual polynomial w/ techniques of [Razborov-Sherstov08]

Conclusions

Hardness amplification beyond block composition $\Rightarrow$

Nearly optimal lower bounds for AC^0

New quantum query lower bounds

Imminently forthcoming work: $\mathrm{adeg}_{\varepsilon}(F) \geq \Omega(n^{1-\delta})$ for some $\varepsilon \geq 1 - \exp(\Omega(n^{1-\delta}))$ and $F \in AC^0$

Thank you!

Open Problems:

- Approximate degree / quantum query complexity of poly-size DNF? Best lower bound: $\Omega(n^{3/4 - \delta})$
- Lower bounds for quantum problems with different structure (e.g. triangle finding, graph collision, verifying matrix products)