Microsoft

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

Mark Bun
Princeton University

Robin Kothari
Microsoft Research

Justin Thaler
Georgetown University

arXiv:1710.09079

THE POLYNOMIAL METHOD STRIKES BACK

We use the polynomial method to prove (nearly) optimal lower bounds on the quantum query complexity of several problems, resolving several open questions from prior work.

The problems include k-distinctness, image size testing, k-junta testing, approximating statistical

THE POLYNOMIAL METHOD STRIKES BACK

We use the polynomial method to prove (nearly) optimal lower bounds on the quantum query complexity of several problems, resolving several open questions from prior work.

The problems include *k*-distinctness, image size testing, *k*-junta testing, approximating statistical distance, approximating Shannon entropy, and surjectivity.

Introduction

Query complexity

Let $f: \{0,1\}^n \to \{0,1\}$ be a function and $x \in \{0,1\}^n$ be an input to f.

$$x = \begin{array}{|c|c|c|c|c|}\hline x_1 & x_2 & x_3 & \quad\cdots\quad & x_n \\ \hline\end{array}$$

Goal: Compute $f(x)$ by reading as few bits of x as possible.

Equivalently, compute $f(x)$ using a circuit/algorithm with the least number of uses of this oracle:

$$i \longrightarrow \boxed{O_x} \longrightarrow x_i$$

In the quantum setting, we have this oracle:

$$|i\rangle|b\rangle \longrightarrow \boxed{O_x} \longrightarrow |i\rangle|b \oplus x_i\rangle$$

Why query complexity?

Complexity theoretic motivation

- We can prove statements about the power of different computational models! (E.g., exponential separation between classical and quantum algorithms)

Algorithmic motivation

- Algorithms often transfer to the circuit model, while the abstraction of query complexity often gets rid of unnecessary details.
- Most quantum algorithms are naturally phrased as query algorithms. E.g., Shor, Grover, Hidden Subgroup, Linear systems (HHL), etc.

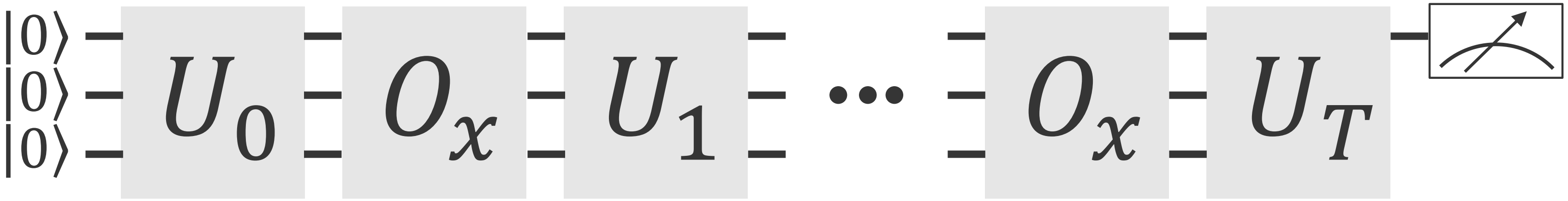
Other applications

- Oracle separations between classes, lower bounds on restricted models, upper and lower bounds in communication complexity, circuit complexity, data structures, etc.

Quantum query complexity

Quantum query complexity: Minimum number of uses of O_x in a quantum circuit that for every input x, outputs $f(x)$ with error $\leq 1/3$.

$Q(f)$



Example: Let $\mathrm{OR}_n(x) = \bigvee_{i=1}^{n} x_i$ and $\mathrm{AND}_n(x) = \bigwedge_{i=1}^{n} x_i$.

Then $Q(\mathrm{OR}_n) = Q(\mathrm{AND}_n) = \Theta\left(\sqrt{n}\right)$ [Grover96, Bennett-Bernstein-Brassard-Vazirani97]

Classically, we need $\Theta(n)$ queries for both problems.

Lower bounds on quantum query complexity

Positive-weights adversary method

Easy to use, but has many limitations.
Cannot show any of the results of this paper.

Negative-weights adversary method

Equals (up to constants) quantum query complexity, but difficult to use.

In recent years, the adversary methods have become the tools of choice for proving lower bounds.

Polynomial method

- Equals (up to constants) quantum query complexity for many natural functions
- Can show lower bounds for algorithms with unbounded error, small error, and no error
- Works when the positive-weights adversary fails (e.g., the collision problem)
- Lower bounds “lift” to lower bounds in communication complexity!
(For an application of this, see the talk by Shalev Ben-David at 10:40 in this room.)

Approximate degree

Approximate degree: Minimum degree of a polynomial $p(x_1, \ldots, x_n)$ with real coefficients such that $\forall x \in \{0,1\}^n$, $|f(x) - p(x)| \leq 1/3$.

$\widetilde{\deg}(f)$

Examples:
- $p(x_1, \ldots, x_n) = x_1 x_2 \cdots x_n$ exactly computes the AND_n function.
- $p(x_1, x_2) = \frac{1}{3} x_1 + \frac{1}{3} x_2$ approximates the AND_2 function.

$$\widetilde{\deg}(\mathrm{OR}_n) = \widetilde{\deg}(\mathrm{AND}_n) = \Theta(\sqrt{n}) \qquad Q(\mathrm{OR}_n) = Q(\mathrm{AND}_n) = \Theta(\sqrt{n})$$

Theorem ([Beals-Buhrman-Cleve-Mosca-de Wolf01]): For any f,

$$Q(f) \geq \frac{1}{2} \widetilde{\deg}(f)$$

The polynomial method

Other applications of approximate degree

Upper bounds

- Learning algorithms [Klivans-Servedio04, Klivans-Servedio06, Kalai-Klivans-Mansour-Servedio08]
- Algorithmic approximations of inclusion-exclusion [Kahn-Linial-Samorodnitsky96, Sherstov09]
- Differentially private data release [Thaler-Ullman-Vadhan12, Chandrasekaran-Thaler-Ullman-Wan14]
- Formula & Graph Complexity *Lower* Bounds [Tal14, Tal17]

Lower bounds

- Communication Complexity [Sherstov07, Shi-Zhu07, Chattopadhyay-Ada08, Lee-Shraibman08,…]
- Circuit Complexity [Minsky-Papert69, Beigel93, Sherstov08]
- Oracle Separations [Beigel94, Bouland-Chen-Holden-Thaler-Vasudevan16]
- Secret Sharing Schemes [Bogdanov-Ishai-Viola-Williamson16]

Results

The $\boldsymbol{k}$-distinctness problem

k-distinctness: Given n numbers in $[R] = \{1, \dots, R\}$, does any number appear $\geq k$ times?

This generalizes element distinctness, which is 2-distinctness.

Upper bounds

- $Q(\mathrm{Dist}_k) = O\left(n^{k/(k+1)}\right)$, using quantum walks [Ambainis07]
- $Q(\mathrm{Dist}_k) = O\left(n^{3/4-1/\exp(k)}\right)$, using learning graphs [Belovs12]

Lower bounds

- $Q(\mathrm{Dist}_k) = \Omega(Q(\mathrm{Dist}_2)) = \Omega\left(n^{2/3}\right)$, using the polynomial method [Aaronson-Shi04]

Our result: $Q(\mathrm{Dist}_k) = \widetilde{\Omega}\left(n^{3/4-1/(2k)}\right)$.

k-junta testing

k-junta testing: Given the truth table of a Boolean function, decide if
(YES) the function depends on at most k variables, or
(NO) the function is far (at least δn in Hamming distance) from having this property.

Upper bounds

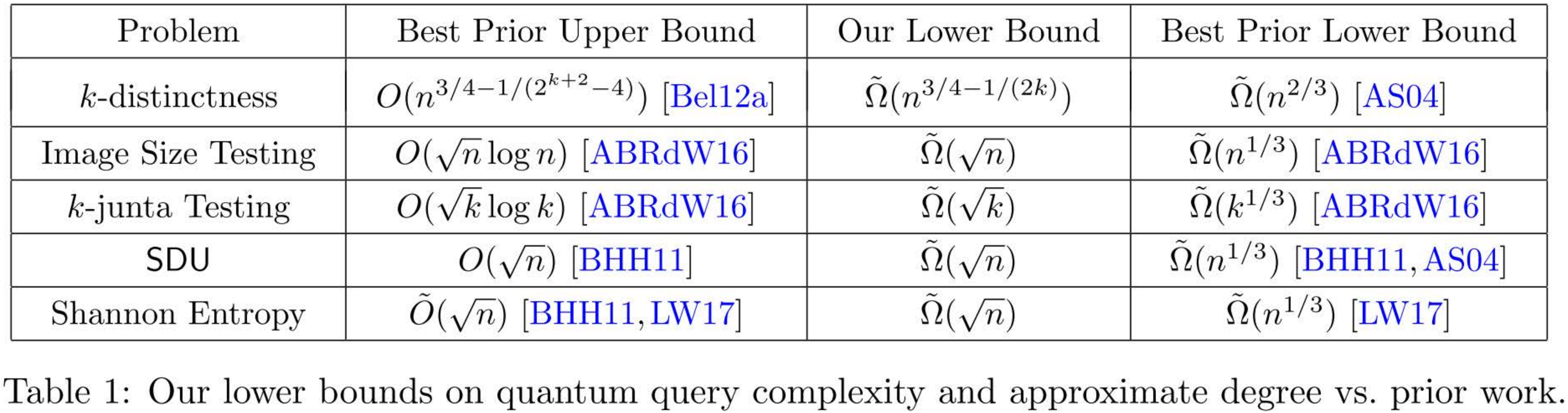
- $Q(\mathrm{Junta}_k) = O(k)$ [Atıcı-Servedio07]
- $Q(\mathrm{Junta}_k) = \tilde{O}(\sqrt{k})$ [Ambainis-Belovs-Regev-deWolf16]

Lower bounds

- $Q_{\mathrm{nonadaptive}}(\mathrm{Junta}_k) = \Omega(\sqrt{k})$ [Atıcı-Servedio07]
- $Q(\mathrm{Junta}_k) = \Omega(k^{1/3})$ [Ambainis-Belovs-Regev-deWolf16]

Our result: $Q(\mathrm{Junta}_k) = \tilde{\Omega}(\sqrt{k})$.

Summary of 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.

Surjectivity

Surjectivity: Given n numbers in $[R]$ ($R = \Theta(n)$), does every $r \in [R]$ appear in the list?

Quantum query complexity

- $Q(\mathrm{SURJ}) = \Theta(n)$ [Beame-Machmouchi12, Sherstov15]

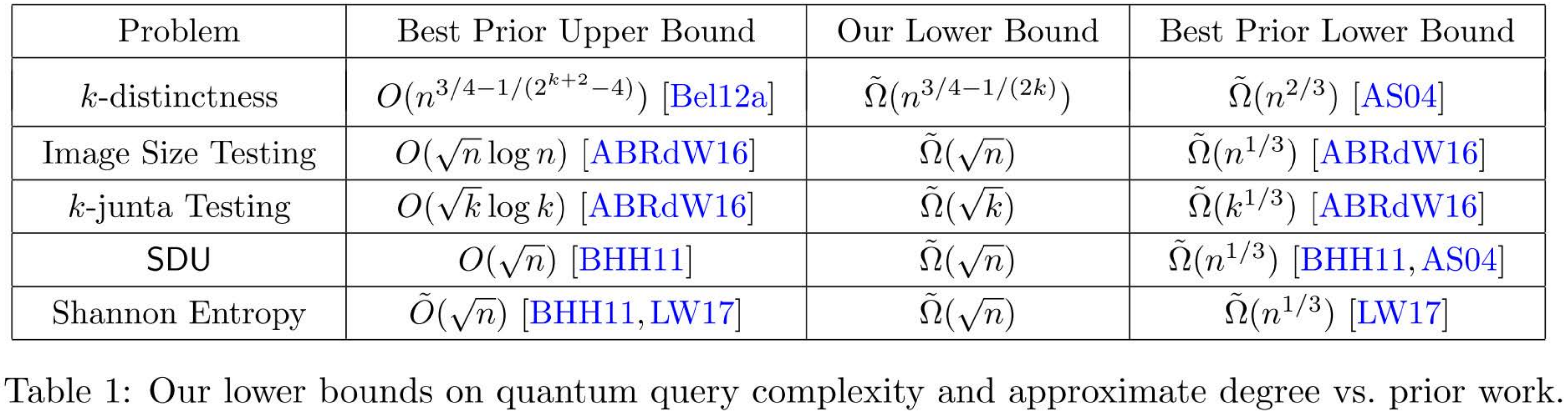
Approximate degree

- Conjecture: $\widetilde{\deg}(\mathrm{SURJ}) = \widetilde{\Omega}(n)$.
- $\widetilde{\deg}(\mathrm{SURJ}) = \widetilde{\Omega}(n^{2/3})$ [Aaronson-Shi04, Ambainis05, Bun-Thaler17]
- $\widetilde{\deg}(\mathrm{SURJ}) = \tilde{O}(n^{3/4})$ [Sherstov18]

Our result: $\widetilde{\deg}(\mathrm{SURJ}) = \widetilde{\Omega}(n^{3/4})$ and a new proof of $\widetilde{\deg}(\mathrm{SURJ}) = \tilde{O}(n^{3/4})$.

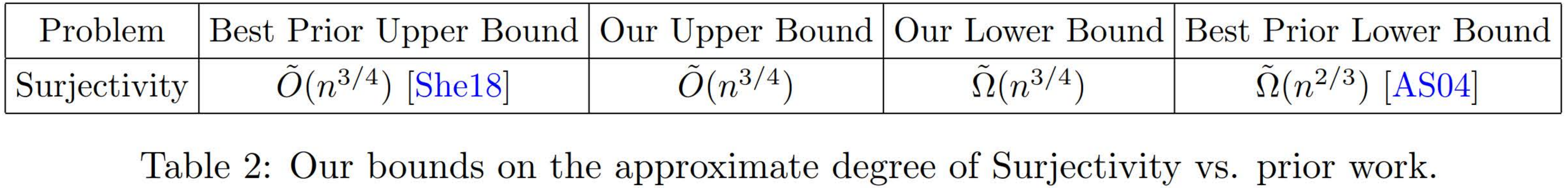
SURJ is the first natural function to have $Q(f) \gg \widetilde{\deg}(f)$!

Summary of results



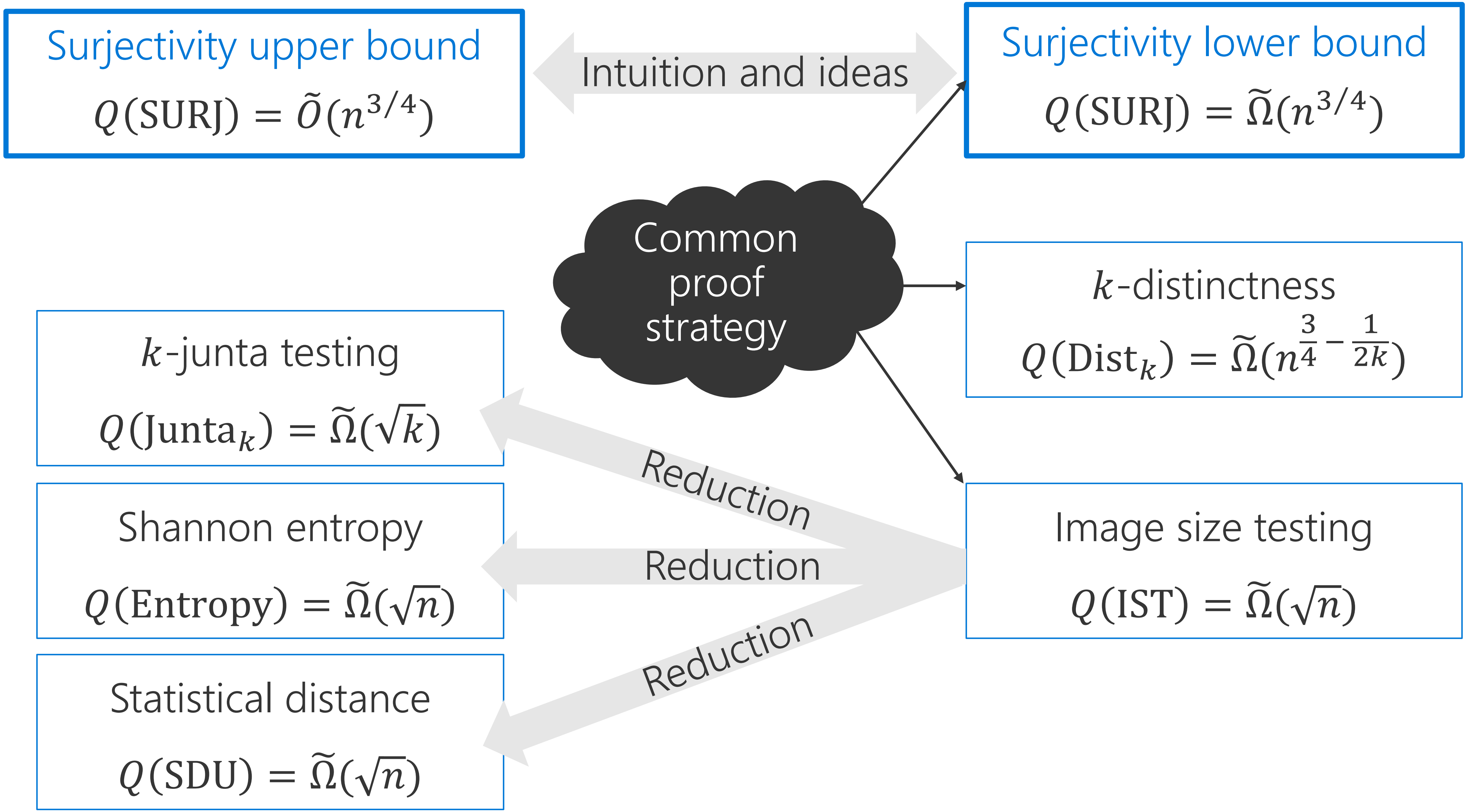
Problem	Best Prior Upper Bound	Our Lower Bound	Best Prior Lower Bound
k-distinctness	$O(n^{3/4-1/(2^{k+2}-4)})$ [Bel12a]	$\tilde{\Omega}(n^{3/4-1/(2k)})$	$\tilde{\Omega}(n^{2/3})$ [AS04]
Image Size Testing	$O(\sqrt{n}\log n)$ [ABRdW16]	$\tilde{\Omega}(\sqrt{n})$	$\tilde{\Omega}(n^{1/3})$ [ABRdW16]
k-junta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



Surjectivity upper bound

$Q(\mathrm{SURJ}) = \tilde{O}(n^{3/4})$

Intuition and ideas

Surjectivity lower bound

$Q(\mathrm{SURJ}) = \widetilde{\Omega}(n^{3/4})$

Common proof strategy

k-distinctness

$Q(\mathrm{Dist}_k) = \widetilde{\Omega}(n^{\frac{3}{4}-\frac{1}{2k}})$

k-junta testing

$Q(\mathrm{Junta}_k) = \widetilde{\Omega}(\sqrt{k})$

Reduction

Shannon entropy

$Q(\mathrm{Entropy}) = \widetilde{\Omega}(\sqrt{n})$

Reduction

Image size testing

$Q(\mathrm{IST}) = \widetilde{\Omega}(\sqrt{n})$

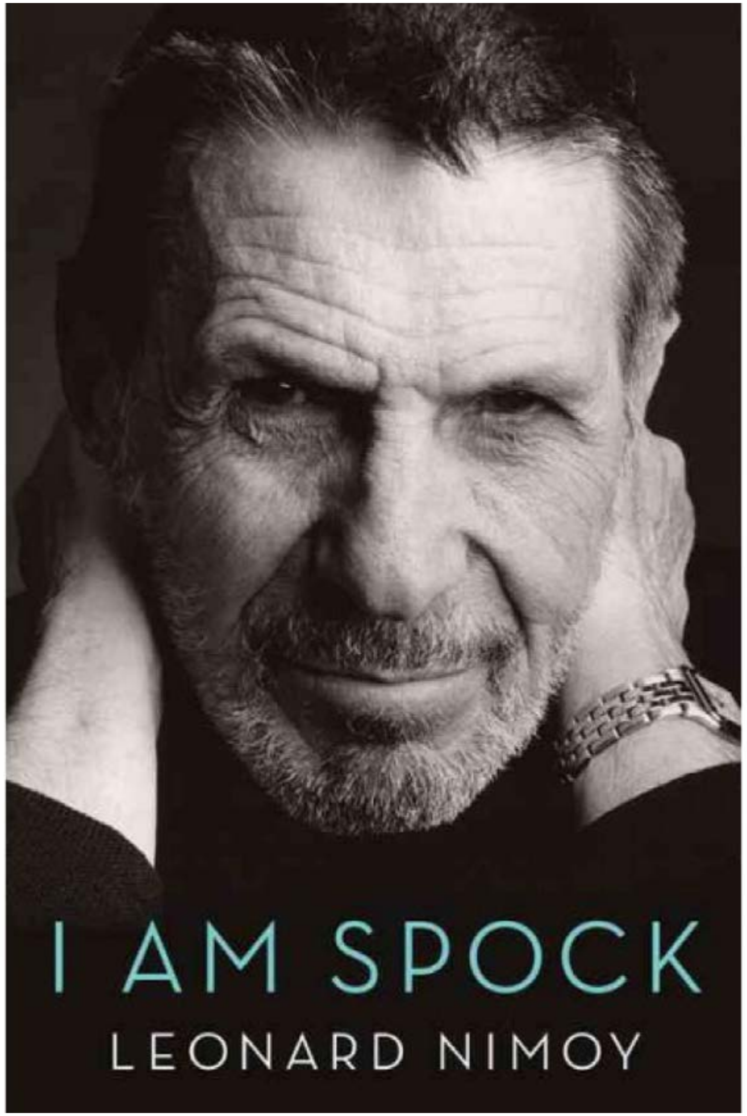
Statistical distance

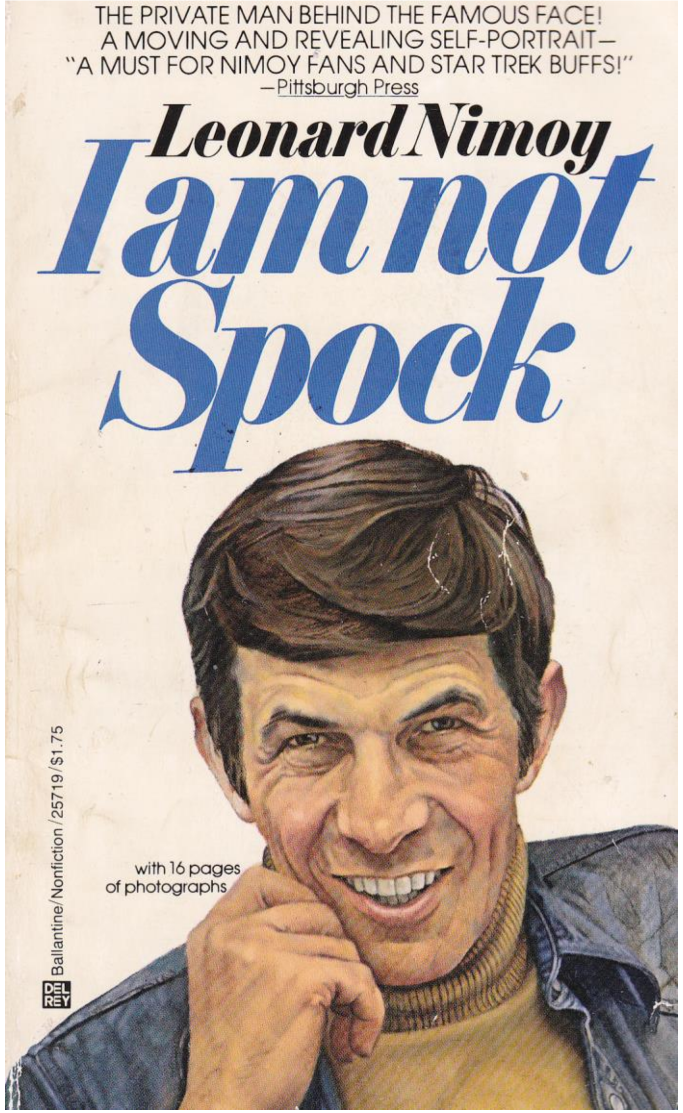
$Q(\mathrm{SDU}) = \widetilde{\Omega}(\sqrt{n})$

Reduction

High level overview of techniques

Overview of the upper bound




THE PRIVATE MAN BEHIND THE FAMOUS FACE!
A MOVING AND REVEALING SELF-PORTRAIT—
"A MUST FOR NIMOY FANS AND STAR TREK BUFFS!"
—Pittsburgh Press
Leonard Nimoy
I am not Spock
with 16 pages of photographs
Ballantine/Nonfiction/25719/$1.75
DEL REY

Polynomials are algorithms

Polynomials are not algorithms

Overview of the upper bound

Idea 1: Polynomials are algorithms

Polynomials can mimic algorithmic primitives like If-then-else, majority voting, reductions, sampling, etc.

Example: Implementing an if-then-else statement

Imagine that polynomials p_1, p_2, and p_3 represent the acceptance probability of algorithms (that output 0 or 1) A_1, A_2, and A_3.

Algorithm: If A_1 outputs 1, then output A_2, else output A_3.

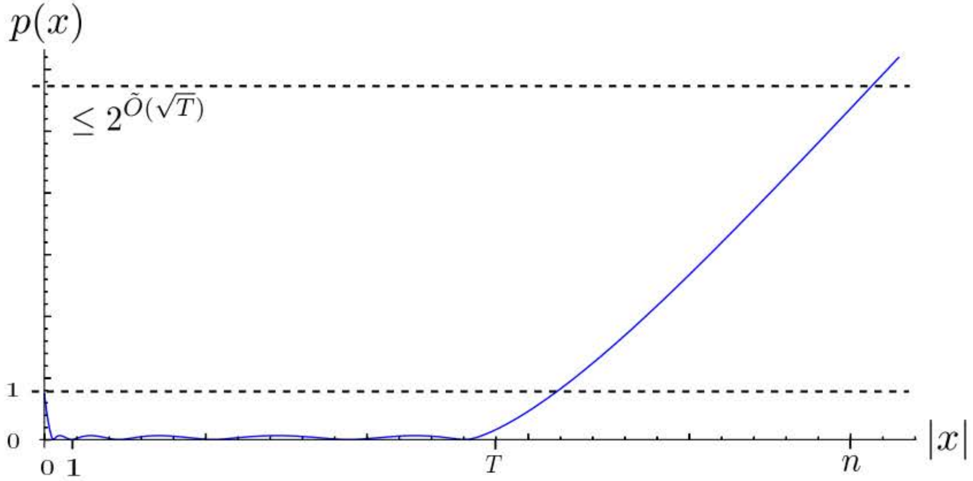
Polynomial: $p_1(x)p_2(x) + \left(1 - p_1(x)\right)p_3(x)$.

Key idea: This is well defined even if $p_i \notin [0,1]$ and do not represent probabilities.

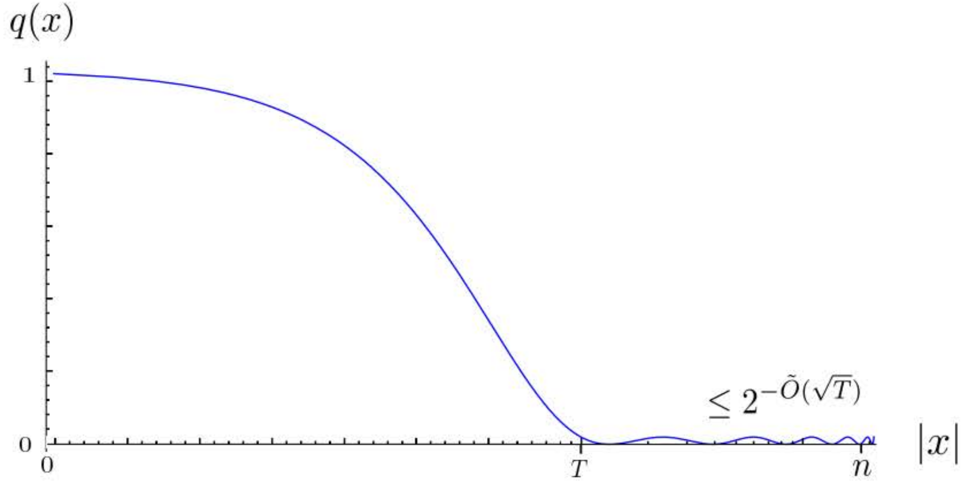
Overview of the upper bound

Idea 2: Polynomials are not algorithms

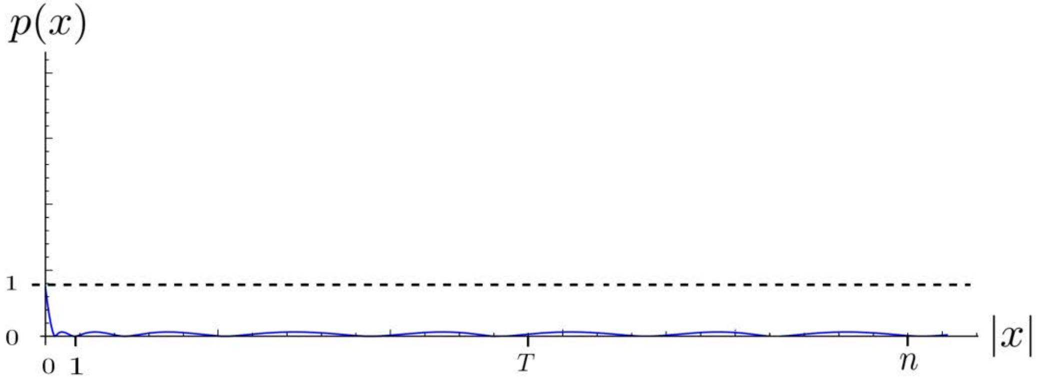
We can use polynomials taking values outside $[0,1]$, even if the final polynomial is bounded in $[0,1]$.



×



=

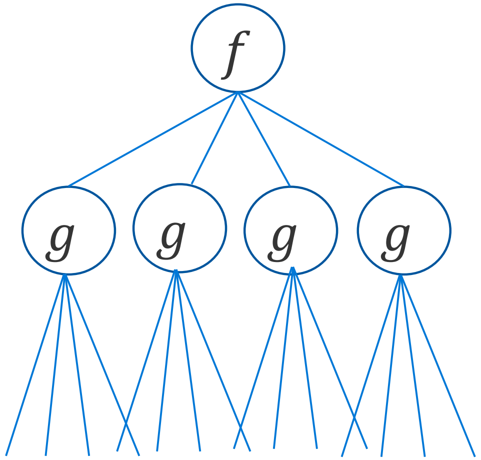


Overview of lower bounds

1. Use linear programming formulation of approximate degree.
   Show degree $\geq d$ by exhibiting a dual witness.
   (Works for simple functions, like AND, OR, etc.)

2. For composed functions, combine dual witnesses of individual functions.
   (Difficult step)

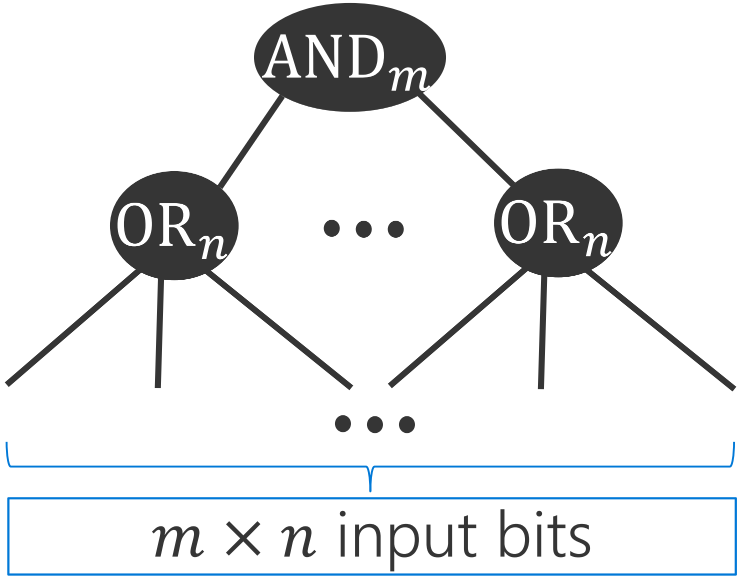
3. Express surjectivity, k-distinctness, and Image size testing as composed functions.



Remainder of the talk: More about lower bounds

Approximate degree of $\mathbf{AND} \circ \mathbf{OR}$

$\mathrm{AND}_m \circ \mathrm{OR}_n$



Upper bound: $\widetilde{\deg}(\mathrm{AND}_m \circ \mathrm{OR}_n) = O(\sqrt{mn})$.

Proof 1: Use robust quantum search [Høyer-Mosca-de Wolf03]

Proof 2: $\forall f, g$, $\widetilde{\deg}(f \circ g) = O\left(\widetilde{\deg}(f)\widetilde{\deg}(g)\right)$ [Sherstov13]

Lower bound [Sherstov13, Bun-Thaler13]: $\widetilde{\deg}(\mathrm{AND}_m \circ \mathrm{OR}_n) = \Omega(\sqrt{mn})$.

Open for 10+ years! Proof uses the method of dual polynomials.

PS: See Adam Bouland's talk at 11:15 for an alternate proof using quantum arguments.

Dual polynomials

Approximate degree can be expressed as a linear program

Primal formulation

$\widetilde{\deg}(f) \le d$ <u>iff</u> there exists a polynomial p of degree d, i.e., $p = \sum_{S:|S|\le d} \alpha_S\, x^S$, s.t.,

$$\forall x \in \{0,1\}^n, \qquad |f(x) - p(x)| \le 1/3$$

Dual formulation

$\widetilde{\deg}(f) > d$ <u>iff</u> there exists $\psi\colon \{0,1\}^n \to \mathbb{R}$,

1. $\sum_x |\psi(x)| = 1$ — (1) ψ is ℓ_1 normalized
2. If $\deg(q) \le d$ then $\sum_x \psi(x) q(x) = 0$ — (2) ψ has pure high degree d
3. $\sum_x \psi(x)(-1)^{f(x)} > 1/3.$ — (3) ψ is well correlated with f

Lower bound for $\mathbf{AND}_m \circ \mathbf{OR}_n$

Dual formulation

$\widetilde{\deg}(f) > d$ iff there exists $\psi: \{0,1\}^n \to \mathbb{R}$,

1. $\sum_x |\psi(x)| = 1$ — (1) ψ is ℓ_1 normalized
2. If $\deg(q) \le d$ then $\sum_x \psi(x) q(x) = 0$ — (2) ψ has pure high degree d
3. $\sum_x \psi(x)(-1)^{f(x)} > 1/3.$ — (3) ψ is well correlated with f

Proof strategy for $\widetilde{\deg}(\mathrm{AND}_m \circ \mathrm{OR}_n) = \Omega(\sqrt{mn})$:

1. Start with ψ_{AND} and ψ_{OR} witnessing $\widetilde{\deg}(\mathrm{AND}_m) = \Omega(\sqrt{m})$ and $\widetilde{\deg}(\mathrm{OR}_n) = \Omega(\sqrt{n})$
2. Combine these into ψ witnessing $\widetilde{\deg}(\mathrm{AND}_m \circ \mathrm{OR}_n) = \Omega(\sqrt{mn})$ using the technique of dual block composition.

Dual block composition for $f \circ g$

Dual formulation

$\widetilde{\deg}(f) > d$ iff there exists $\psi: \{0,1\}^n \to \mathbb{R}$,

1. $\sum_x |\psi(x)| = 1$ — (1) ψ is ℓ_1 normalized
2. If $\deg(q) \leq d$ then $\sum_x \psi(x) q(x) = 0$ — (2) ψ has pure high degree d
3. $\sum_x \psi(x)(-1)^{f(x)} > 1/3.$ — (3) ψ is well correlated with f

Given two dual witnesses ψ_f for f and ψ_g for g, we can define $\psi_{f \circ g}$ for $f \circ g$ as follows:

$$\psi_{f \circ g} = 2^n\, \psi_f\left(\operatorname{sgn}\left(\psi_g(x_1)\right), \ldots, \operatorname{sgn}\left(\psi_g(x_n)\right)\right) \prod_{i=1}^n |\psi_g(x_i)|$$

[Shi-Zhu09, Lee09, Sherstov13]

Composed dual automatically satisfies (1) and (2).

[Sherstov13, Bun-Thaler13] show that property (3) is also satisfied for $\mathrm{AND}_m \circ \mathrm{OR}_n$.

Surjectivity lower bound

$$\widetilde{\deg}(\mathrm{SURJ}) = \widetilde{\Omega}\left(n^{3/4}\right)$$

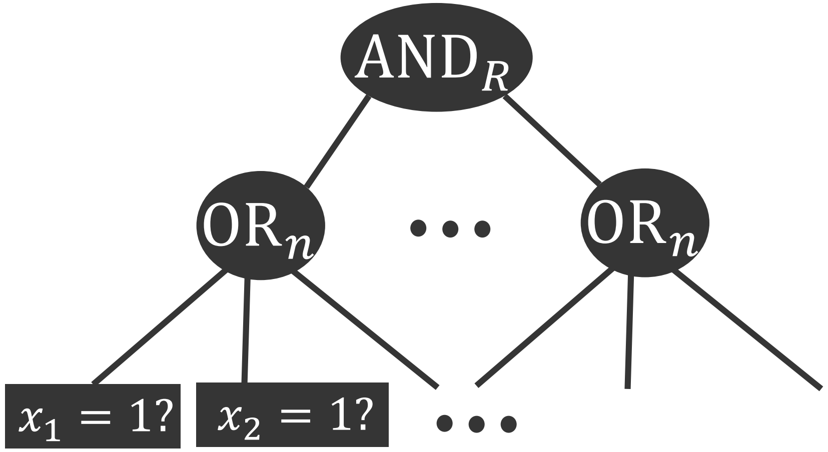
Reduction to a composed function

Surjectivity: Given n numbers in $[R]$ ($R = \Theta(n)$), does every $r \in [R]$ appear in the list?

$$\mathrm{SURJ}(x_1, \ldots, x_n) = \bigwedge_{r\in[R]} \bigvee_{i\in[n]} (x_i = r?)$$

SURJ reduces to $\mathrm{AND}_R \circ \mathrm{OR}_n$ function, restricted to inputs with Hamming weight $\leq n$.

We denote this function $(\mathrm{AND}_R \circ \mathrm{OR}_n)^{\leq n}$.



$$\Rightarrow \widetilde{\deg}(\mathrm{SURJ}) = \tilde{O}\left(\widetilde{\deg}\left((\mathrm{AND}_R \circ \mathrm{OR}_n)^{\leq n}\right)\right)$$

Converse [Ambainis05, Bun-Thaler17]: $\widetilde{\deg}(\mathrm{SURJ}) = \tilde{\Omega}\left(\widetilde{\deg}\left((\mathrm{AND}_R \circ \mathrm{OR}_n)^{\leq n}\right)\right)$

Important: $\mathrm{AND}_R \circ \mathrm{OR}_n \neq (\mathrm{AND}_R \circ \mathrm{OR}_n)^{\leq n}$

$$\widetilde{\deg}(\mathrm{AND}_R \circ \mathrm{OR}_n) = \Theta\left(\sqrt{Rn}\right) = \Theta(n)$$

$$\widetilde{\deg}\left((\mathrm{AND}_R \circ \mathrm{OR}_n)^{\leq n}\right) = \widetilde{\Theta}\left(\widetilde{\deg}(\mathrm{SURJ})\right) = \Theta(n^{3/4})$$

Progress so far towards $\widetilde{\deg}(\mathrm{SURJ}) = \tilde{\Omega}\left(n^{3/4}\right)$

1. We saw that $\widetilde{\deg}(\mathrm{SURJ}) = \tilde{\Theta}\left(\widetilde{\deg}\left((\mathrm{AND}_R \circ \mathrm{OR}_n)^{\leq n}\right)\right)$.
2. We saw using dual block composition that

$$\widetilde{\deg}(\mathrm{AND}_R \circ \mathrm{OR}_n) = \Omega\left(\sqrt{Rn}\right) = \Omega(n), \text{ when } R = \Theta(n).$$

Does the constructed dual also work for $(\mathrm{AND}_R \circ \mathrm{OR}_n)^{\leq n}$? No.

Dual formulation for problems where we only care about Hamming weight $\leq H$

$\widetilde{\deg}\left(f^{\leq H}\right) > d$ iff there exists ψ,

1. $\sum_x \lvert\psi(x)\rvert = 1$	(1) ψ is ℓ_1 normalized
2. If $\deg(q) \leq d$ then $\sum_x \psi(x) q(x) = 0$	(2) ψ has pure high degree d
3. $\sum_x \psi(x)(-1)^{f(x)} > 1/3$.	(3) ψ is well correlated with f
4. $\psi(x) = 0$ if $\lvert x\rvert > H$	(4) ψ is only supported on the promise

I'M ALTERING THE ~~DEAL~~ DUAL
PRAY
I DON'T ALTER IT ANY FURTHER
quickmeme.com

Dual witness for $\widetilde{\deg}((\mathrm{AND}_R \circ \mathrm{OR}_n)^{\le n})$

Dual formulation for problems where we only care about Hamming weight $\le H$

$\widetilde{\deg}(f^{\le H}) > d$ iff there exists ψ,

1. $\sum_x |\psi(x)| = 1$ — (1) ψ is ℓ_1 normalized
2. If $\deg(q) \le d$ then $\sum_x \psi(x) q(x) = 0$ — (2) ψ has pure high degree d
3. $\sum_x \psi(x)(-1)^{f(x)} > 1/3.$ — (3) ψ is well correlated with f
4. $\psi(x) = 0$ if $|x| > H$ — (4) ψ is only supported on the promise

Fix 1: Use a dual witness ψ_{OR} for OR_n that only certifies $\widetilde{\deg}(\mathrm{OR}_n) = \Omega(n^{1/4})$ and satisfies a "dual decay condition", i.e., $|\psi_{\mathrm{OR}}(x)|$ is exponentially small for $|x| \gg n^{1/4}$. Thus the composed dual has degree $\Omega(\sqrt{R} n^{1/4}) = \Omega(n^{3/4})$ and almost satisfies condition (4).

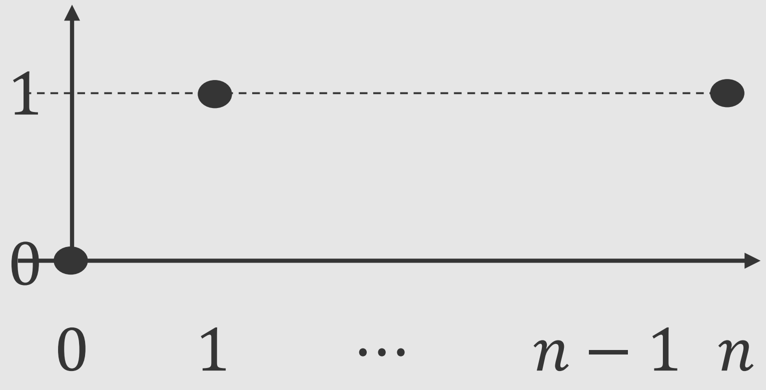
Fix 2: Although condition (4) is only "almost satisfied" in our dual witness, we can postprocess the dual to have it be exactly satisfied [Razborov-Sherstov10].

Looking back at the lower bounds

How did we resolve questions that have resisted attack by the adversary method?

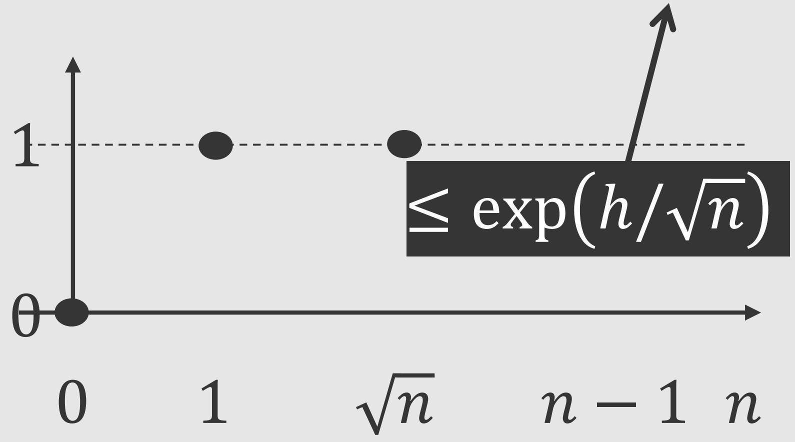
What is the key new ingredient in these lower bounds?

Lower bound for OR:



Any polynomial like this must have degree $\Omega(\sqrt{n})$.

Key property we exploit:



Any polynomial like this must still have degree $\Omega(\sqrt{n})$!

Open problems

Open problems

1. What is the quantum query complexity (or approximate degree) of
   - Triangle finding
   - Graph collision
   - Matrix product verification
   - k-distinctness (pin down the exponent precisely)

2. What is the approximate degree of k-sum? The quantum query complexity is $\Theta\left(n^{k/k+1}\right)$ [Ambainis07, Belovs-Špalek13].

3. Is there a function in AC^0 with approximate degree $\widetilde{\Omega}(n)$? The best known lower bound is $\widetilde{\Omega}(n^{1-2^{-d}})$ for a depth-d AC^0 function (follows from our results).

4. Do all polynomial size DNFs have approximate degree $o(n)$? Best lower bound is from k-distinctness. What about the quantum query complexity?

Microsoft

Thanks!

Microsoft Quantum internship applications:
microsoft.com/en-us/research/opportunity/internship-microsoft-quantum/

Microsoft Quantum Development Kit:
microsoft.com/quantum