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

Justin Thaler (Georgetown)

Joint work with:
Mark Bun (Princeton)
Robin Kothari (MSR Redmond)

Boolean Functions

- Boolean function $f : \{-1, 1\}^n \to \{-1, 1\}$
- $$\mathrm{AND}_n(x) = \begin{cases} -1 \quad \text{(TRUE)} & \text{if } x = (-1)^n \\ 1 \quad \text{(FALSE)} & \text{otherwise} \end{cases}$$

Approximate Degree

- A real polynomial p ϵ-approximates f if

$$|p(x) - f(x)| < \epsilon \quad \forall x \in \{-1,1\}^n$$

- $\widetilde{\deg}_\epsilon(f) =$ minimum degree needed to ϵ-approximate f
- $\widetilde{\deg}(f) := \deg_{1/3}(f)$ is the approximate degree of f

Example 1: The Approximate Degree of AND_n

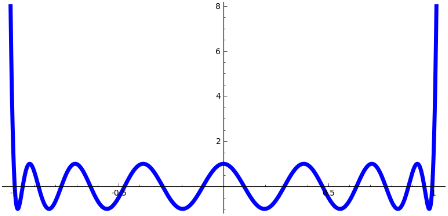
Example: What is the Approximate Degree of AND_n?

$\widetilde{\deg}(\mathrm{AND}_n) = \Theta(\sqrt{n})$.

- Upper bound: Use **Chebyshev Polynomials**.
- Markov's Inequality: Let $G(t)$ be a univariate polynomial s.t. $\deg(G) \leq d$ and $\max_{t \in [-1,1]} |G(t)| \leq 1$. Then

$$\max_{t \in [-1,1]} |G'(t)| \leq d^2.$$

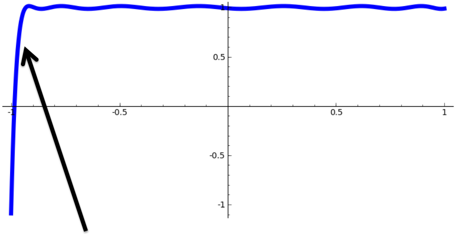
- Chebyshev polynomials are the extremal case.



Example: What is the Approximate Degree of AND_n?

$\widetilde{\deg}(\mathrm{AND}_n) = O(\sqrt{n})$.

- After shifting a scaling, can turn degree $O(\sqrt{n})$ Chebyshev polynomial into a univariate polynomial $Q(t)$ that looks like:

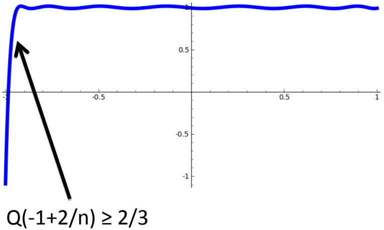


- Define n-variate polynomial p via $p(x) = Q(\sum_{i=1}^{n} x_i/n)$.
- Then $|p(x) - \mathrm{AND}_n(x)| \leq 1/3 \quad \forall x \in \{-1,1\}^n$.

Example: What is the Approximate Degree of AND_n?

[NS92] $\widetilde{\deg}(\mathrm{AND}_n) = \Omega(\sqrt{n})$.

- Lower bound: Use **symmetrization**.
- Suppose $|p(x) - \mathrm{AND}_n(x)| \leq 1/3 \quad \forall x \in \{-1,1\}^n$.
- There is a way to turn p into a univariate polynomial p^{sym} that looks like this:



- Claim 1: $\deg(p^{\mathsf{sym}}) \leq \deg(p)$.
- Claim 2: Markov's inequality $\implies \deg(p^{\mathsf{sym}}) = \Omega(n^{1/2})$.

Why Care about Approximate Degree?

Applications of $\widetilde{\deg}$ Upper Bounds

Upper bounds on $\widetilde{\deg}_\epsilon(f)$ yield efficient learning algorithms.

- $\epsilon \approx 1/3$: Agnostic Learning [KKMS05]
- $\epsilon \approx 1 - 2^{-n^\delta}$: Attribute-Efficient Learning [KS04, STT12]
- $\epsilon \to 1$ (i.e., threshold degree, $\deg_\pm(f)$): PAC learning [KS01]

Applications of $\widetilde{\deg}$ Upper Bounds

Upper bounds on $\widetilde{\deg}_\epsilon(f)$ yield efficient learning algorithms.

- $\epsilon \approx 1/3$: Agnostic Learning [KKMS05]
- $\epsilon \approx 1 - 2^{-n^\delta}$: Attribute-Efficient Learning [KS04, STT12]
- $\epsilon \to 1$ (i.e., threshold degree, $\deg_\pm(f)$): PAC learning [KS01]

- Upper bounds on $\widetilde{\deg}_{1/3}(f)$ also:
  - Imply fast algorithms for differentially private data release [TUV12, CTUW14].

Applications of $\widetilde{\deg}$ Upper Bounds

Upper bounds on $\widetilde{\deg}_\epsilon(f)$ yield efficient learning algorithms.

- $\epsilon \approx 1/3$: Agnostic Learning [KKMS05]
- $\epsilon \approx 1 - 2^{-n^\delta}$: Attribute-Efficient Learning [KS04, STT12]
- $\epsilon \to 1$ (i.e., threshold degree, $\deg_\pm(f)$): PAC learning [KS01]

- Upper bounds on $\widetilde{\deg}_{1/3}(f)$ also:
  - Imply fast algorithms for differentially private data release [TUV12, CTUW14].
  - Underly the best known lower bounds on formula complexity and graph complexity [Tal2014, 2016a, 2016b]

This Talk: Two Focuses Involving $\widetilde{\deg}$ Lower Bounds

- Focus 1: A nearly optimal bound on the approximate degree of AC^0, and its applications [BT17].
- Focus 2: Proving tight quantum query lower bounds for specific functions [BKT17].

First Focus: Approximate Degree of AC^0

- Approximate degree is a key tool for understanding AC^0.
- At the heart of the best known bounds on the complexity of AC^0 under measures such as:
  - Multi-Party (Quantum) Communication Complexity
  - Approximate Rank
  - Sign-rank $\approx$ Unbounded Error Communication (UPP)
  - Discrepancy $\approx$ Margin complexity
  - Majority-of-Threshold circuit size
  - Threshold-of-Majority circuit size
  - and more.

First Focus: Approximate Degree of AC^0

- Approximate degree is a key tool for understanding AC^0.
- At the heart of the best known bounds on the complexity of AC^0 under measures such as:
  - Multi-Party (Quantum) Communication Complexity
  - Approximate Rank
  - Sign-rank $\approx$ Unbounded Error Communication (UPP)
  - Discrepancy $\approx$ Margin complexity
  - Majority-of-Threshold circuit size
  - Threshold-of-Majority circuit size
  - and more.

Problem 1: Is there a function on n variables that is in AC^0, and has approximate degree $\Omega(n)$?

Approximate Degree of AC^0: Details

- Best known result: $\tilde{\Omega}(n^{2/3})$ for the Element Distinctness function (Aaronson and Shi, 2004).

Approximate Degree of AC^0: Details

- Best known result: $\tilde{\Omega}(n^{2/3})$ for the Element Distinctness function (Aaronson and Shi, 2004).
- Our result: For any constant $\delta > 0$, a function in AC^0 with approximate degree $\Omega(n^{1-\delta})$.
  - More precisely, circuit depth is $O(\log(1/\delta))$.

Approximate Degree of AC^0: Details

- Best known result: $\tilde{\Omega}(n^{2/3})$ for the Element Distinctness function (Aaronson and Shi, 2004).
- Our result: For any constant $\delta > 0$, a function in AC^0 with approximate degree $\Omega(n^{1-\delta})$.
  - More precisely, circuit depth is $O(\log(1/\delta))$.
  - Lower bound also applies to DNFs of polylogarithmic width (and quasipolynomial size).

Applications

- Nearly optimal $\Omega(n^{1-\delta})$ lower bounds on quantum communication complexity of AC^0.
- Essentially optimal (quadratic) separation of certificate complexity and approximate degree.
- Better secret sharing schemes with reconstruction in AC^0.

Second Focus: Quantum Query Complexity

- In the quantum query model, a quantum algorithm is given query access to the bits of an input x.
- Goal: compute some function f of x while minimizing the number of queried bits.
- Most quantum algorithms were discovered in or can easily be described in the query setting.

Connecting $\widetilde{\deg}$ and Quantum Query Complexity

- Let $\mathcal{A}$ be a quantum algorithm making at most T queries.
- [BBC$^+$01] there is a polynomial p of degree $2T$ such that

$$p(x) = \Pr[\mathcal{A}(x) = 1].$$

  - So $\mathcal{A}$ computes f to error $\epsilon \Longrightarrow 2p(x) - 1$ approximates f to error 2ϵ.
  - So $\widetilde{\deg}(f)$ is a lower bound on the quantum query complexity of f.
- This is called the **polynomial method** in quantum query complexity.

Our Results

Problem	Prior Upper Bound	Our Lower Bound	Prior Lower Bound
k-distinctness	$O(n^{3/4-1/(2^{k+2}-4)})$	$\tilde{\Omega}(n^{3/4-1/(2k)})$	$\tilde{\Omega}(n^{2/3})$
Image Size Testing	$O(\sqrt{n}\log n)$	$\tilde{\Omega}(\sqrt{n})$	$\tilde{\Omega}(n^{1/3})$
k-junta Testing	$O(\sqrt{k}\log k)$	$\tilde{\Omega}(\sqrt{k})$	$\tilde{\Omega}(k^{1/3})$
SDU	$O(\sqrt{n})$	$\tilde{\Omega}(\sqrt{n})$	$\tilde{\Omega}(n^{1/3})$
Shannon Entropy	$\tilde{O}(\sqrt{n})$	$\tilde{\Omega}(\sqrt{n})$	$\tilde{\Omega}(n^{1/3})$

Our lower bounds on quantum query complexity and $\widetilde{\deg}$ vs. prior work.

Problem	Prior Upper Bound	Our Upper and Lower Bounds	Prior Lower Bound
Surjectivity	$\tilde{O}(n^{3/4})$	$\tilde{O}(n^{3/4})$ and $\tilde{\Omega}(n^{3/4})$	$\tilde{\Omega}(n^{2/3})$

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

Lower Bound Methods in Quantum Query Complexity

- Since 2002, the positive-weights adversary method, and the newer negative-weights adversary method have been tools of choice for proving quantum query lower bounds.
  - Negative-weights method can prove a tight lower bound for any function [Rei11, LMR+11].
  - But is often challenging to apply to specific functions.
- Quantum query bounds proved via approximate degree "lift" to communication lower bounds [She11].
  - Not known to hold for adversary methods.

Ruminations on the Polynomial Method

- Intuitively, how do we resolve questions that have resisted adversary methods?
  - A key fact exploited in our analysis is:

Fact (1)

Any polynomial $p\colon \{-1,1\}^n \to \mathbb{R}$ *satisfying the following conditions requires degree* $\Omega(n^{1/4})$*:*

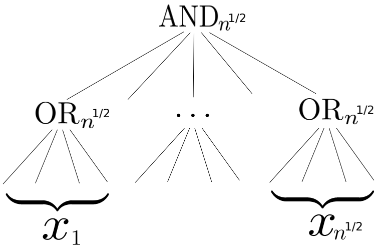
$$\begin{cases} |p(x) - \mathrm{OR}_n(x)| \le 1/3 & \textit{if } |x| \le n^{1/4} \\ |p(x)| \le \exp(|x| \cdot n^{-1/4}) & \textit{if } |x| > n^{1/4}. \end{cases}$$

  - Fact (1) is "non-quantum" because any quantum query algorithm always produces polynomials bounded in $[0,1]$.
- Reasoning about such "non-quantum" polynomials seems difficult to capture by adversary methods.

Prior Work: The Method of Dual Polynomials and the AND-OR Tree

Beyond Symmetrization

- Symmetrization is "lossy": in turning an n-variate poly p into a univariate poly p^{sym}, we throw away information about p.
- **Challenge Problem**: What is $\widetilde{\deg}(\text{AND-OR}_n)$?



History of the AND-OR Tree

Theorem

$\widetilde{\deg}(\text{AND-OR}_n) = \Theta(n^{1/2})$.

History of the AND-OR Tree

Theorem

$\widetilde{\deg}(\text{AND-OR}_n) = \Theta(n^{1/2})$.

Tight Upper Bound of $O(n^{1/2})$

[HMW03] via quantum algorithms
[BNRdW07] different proof of $O(n^{1/2} \cdot \log n)$ (via error reduction+composition)
[She13] different proof of tight upper bound (via robustification)

History of the AND-OR Tree

Theorem

$\widetilde{\deg}(\text{AND-OR}_n) = \Theta(n^{1/2})$.

Tight Upper Bound of $O(n^{1/2})$

[HMW03] via quantum algorithms
[BNRdW07] different proof of $O(n^{1/2} \cdot \log n)$ (via error reduction+composition)
[She13] different proof of tight upper bound (via robustification)

Tight Lower Bound of $\Omega(n^{1/2})$

[BT13] and [She13] via the method of dual polynomials

Linear Programming Formulation of Approximate Degree

What is best error achievable by **any** degree d approximation of f?
Primal LP (Linear in ϵ and coefficients of p):

$$\begin{aligned} \min_{p,\epsilon} \quad & \epsilon \\ \text{s.t.} \quad & |p(x) - f(x)| \leq \epsilon && \text{for all } x \in \{-1,1\}^n \\ & \deg p \leq d \end{aligned}$$

Dual LP:

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

Dual Characterization of Approximate Degree

Theorem: $\deg_\epsilon(f) > d$ iff there exists a "dual polynomial" $\psi \colon \{-1,1\}^n \to \mathbb{R}$ with

(1) $$\sum_{x \in \{-1,1\}^n} \psi(x) f(x) > \epsilon$$ "high correlation with f"

(2) $$\sum_{x \in \{-1,1\}^n} |\psi(x)| = 1$$ "L_1-norm 1"

(3) $$\sum_{x \in \{-1,1\}^n} \psi(x) q(x) = 0, \text{ when } \deg q \leq d$$ "pure high degree d"

A **lossless** technique. Strong duality implies any approximate degree lower bound can be witnessed by dual polynomial.

Dual Characterization of Approximate Degree

Theorem: $\widetilde{\deg}_\epsilon(f) > d$ iff there exists a "dual polynomial" $\psi\colon \{-1,1\}^n \to \mathbb{R}$ with

(1) $\displaystyle\sum_{x\in\{-1,1\}^n} \psi(x)f(x) > \epsilon$ "high correlation with f"

(2) $\displaystyle\sum_{x\in\{-1,1\}^n} |\psi(x)| = 1$ "L_1-norm 1"

(3) $\displaystyle\sum_{x\in\{-1,1\}^n} \psi(x)q(x) = 0,$ when $\deg q \leq d$ "pure high degree d"

Example: $2^{-n} \cdot$ PARITY$_n$ witnesses the fact that $\widetilde{\deg}_\epsilon(\text{PARITY}_n) = n$ for any $\epsilon < 1$.

Goal: Construct an explicit dual polynomial $\psi_{\textsf{AND-OR}}$ for AND-OR

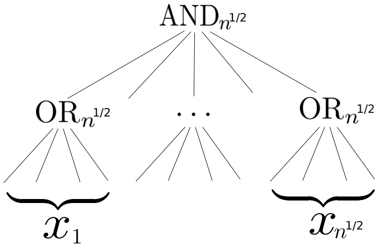
Constructing a Dual Polynomial

- By [NS92], there are dual polynomials
  $\psi_{\textsf{OUT}}$ for $\widetilde{\deg}\,(\mathrm{AND}_{n^{1/2}}) = \Omega(n^{1/4})$ and
  $\psi_{\textsf{IN}}$ for $\widetilde{\deg}\,(\mathrm{OR}_{n^{1/2}}) = \Omega(n^{1/4})$
- Both [She13] and [BT13] combine $\psi_{\textsf{OUT}}$ and $\psi_{\textsf{IN}}$ to obtain a dual polynomial $\psi_{\textsf{AND-OR}}$ for $\mathrm{AND\text{-}OR}$.
- The combining method was proposed in earlier work by [SZ09, Lee09, She09].

The Combining Method [SZ09, She09, Lee09]

$$\psi_{\textsf{AND-OR}}(x_1, \ldots, x_{n^{1/2}}) := C \cdot \psi_{\textsf{OUT}}(\ldots, \mathrm{sgn}(\psi_{\textsf{IN}}(x_i)), \ldots) \prod_{i=1}^{n^{1/2}} |\psi_{\textsf{IN}}(x_i)|$$

(C chosen to ensure $\psi_{\textsf{AND-OR}}$ has L_1-norm 1).



The Combining Method [SZ09, She09, Lee09]

$$\psi_{\textbf{AND-OR}}(x_1, \ldots, x_{n^{1/2}}) := C \cdot \psi_{\textbf{OUT}}(\ldots, \mathrm{sgn}(\psi_{\textbf{IN}}(x_i)), \ldots) \prod_{i=1}^{n^{1/2}} |\psi_{\textbf{IN}}(x_i)|$$

(C chosen to ensure $\psi_{\textbf{AND-OR}}$ has L_1-norm 1).

Must verify:

1. $\psi_{\textbf{AND-OR}}$ has pure high degree $\geq n^{1/4} \cdot n^{1/4} = n^{1/2}$.
2. $\psi_{\textbf{AND-OR}}$ has high correlation with $\mathrm{AND\text{-}OR}$.

The Combining Method [SZ09, She09, Lee09]

$$\psi_{\textsf{AND-OR}}(x_1, \ldots, x_{n^{1/2}}) := C \cdot \psi_{\textsf{OUT}}(\ldots, \operatorname{sgn}(\psi_{\textsf{IN}}(x_i)), \ldots) \prod_{i=1}^{n^{1/2}} |\psi_{\textsf{IN}}(x_i)|$$

(C chosen to ensure $\psi_{\textsf{AND-OR}}$ has L_1-norm 1).

Must verify:

1. $\psi_{\textsf{AND-OR}}$ has pure high degree $\geq n^{1/4} \cdot n^{1/4} = n^{1/2}$.✓[She09]
2. $\psi_{\textsf{AND-OR}}$ has high correlation with AND-OR. [BT13, She13]

Recent Progress on the Complexity of AC^0: Applying the Method of Dual Polynomials to Block-Composed Functions

(Negative) One-Sided Approximate Degree

- Negative one-sided approximate degree is an intermediate notion between approximate degree and threshold degree.
- A real polynomial p is a negative one-sided ϵ-approximation for f if

$$|p(x) - 1| < \epsilon \quad \forall x \in f^{-1}(1)$$

$$p(x) \leq -1 \quad \forall x \in f^{-1}(-1)$$

- $\widetilde{\text{odeg}}_{-,\epsilon}(f) =$ min degree of a negative one-sided ϵ-approximation for f.
- Examples: $\widetilde{\text{odeg}}_{-,1/3}(\text{AND}_n) = \Theta(\sqrt{n})$; $\widetilde{\text{odeg}}_{-,1/3}(\text{OR}_n) = 1$.

Recent Theorems

Theorem (BT13, She13)

Let f be a Boolean function with $\widetilde{odeg}_{-,1/2}(f) \geq d$. Let $F = \mathrm{OR}_t(f, \ldots, f)$. Then $\widetilde{\deg}_{1/2}(F) \geq d \cdot \sqrt{t}$.

Recent Theorems

Theorem (BT13, She13)

Let f be a Boolean function with $\widetilde{odeg}_{-,1/2}(f) \geq d$. Let $F = \mathrm{OR}_t(f, \ldots, f)$. Then $\widetilde{\deg}_{1/2}(F) \geq d \cdot \sqrt{t}$.

Theorem (BT14)

Let f be a Boolean function with $\widetilde{odeg}_{-,1/2}(f) \geq d$. Let $F = \mathrm{OR}_t(f, \ldots, f)$. Then $\widetilde{\deg}_{1-2^{-t}}(F) \geq d$.

Recent Theorems

Theorem (BT13, She13)

Let f be a Boolean function with $\widetilde{odeg}_{-,1/2}(f) \geq d$. *Let* $F = \mathrm{OR}_t(f, \ldots, f)$. *Then* $\widetilde{\deg}_{1/2}(F) \geq d \cdot \sqrt{t}$.

Theorem (BT14)

Let f be a Boolean function with $\widetilde{odeg}_{-,1/2}(f) \geq d$. *Let* $F = \mathrm{OR}_t(f, \ldots, f)$. *Then* $\widetilde{\deg}_{1-2^{-t}}(F) \geq d$.

Theorem (She14)

Let f be a Boolean function with $\widetilde{odeg}_{-,1/2}(f) \geq d$. *Let* $F = \mathrm{OR}_t(f, \ldots, f)$. *Then* $\deg_{\pm}(F) = \Omega(\min\{d, t\})$.

Recent Theorems

Theorem (BT13, She13)

Let f be a Boolean function with $\widetilde{odeg}_{-,1/2}(f) \geq d$. Let $F = \mathrm{OR}_t(f, \ldots, f)$. Then $\widetilde{\deg}_{1/2}(F) \geq d \cdot \sqrt{t}$.

Theorem (BT14)

Let f be a Boolean function with $\widetilde{odeg}_{-,1/2}(f) \geq d$. Let $F = \mathrm{OR}_t(f, \ldots, f)$. Then $\widetilde{\deg}_{1-2^{-t}}(F) \geq d$.

Theorem (She14)

Let f be a Boolean function with $\widetilde{odeg}_{-,1/2}(f) \geq d$. Let $F = \mathrm{OR}_t(f, \ldots, f)$. Then $\deg_{\pm}(F) = \Omega(\min\{d, t\})$.

Theorem (BCHTV16)

Let f be a Boolean function with $\widetilde{\deg}_{1/2}(f) \geq d$. Let $F = GAPMAJ_t(f, \ldots, f)$. Then $\deg_{\pm}(F) \geq \Omega(\min\{d, t\})$.

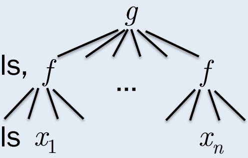
Reminder

Problem 1: Is there a function on n variables that is in AC^0, and has approximate degree $\Omega(n)$?

Our Techniques

Hardness Amplification in AC^0

<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



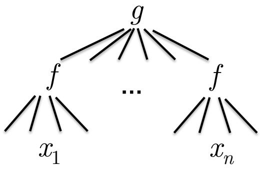
<u>"Block Composition Barrier"</u>

Robust approximations, i.e.,

$$\widetilde{\deg}(g \circ f) \leq O(\widetilde{\deg}(g) \cdot \widetilde{\deg}(f))$$

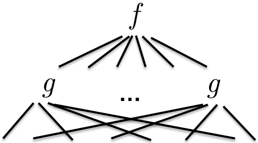
imply that block composition *cannot* increase approximate degree as a function of n

Around the Block-Composition Barrier



Prior work:

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



This work:

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

A General Hardness Amplification Result

Theorem (Strong Hardness Amplification Within AC^0)

Let $f\colon \{-1,1\}^n \to \{-1,1\}$

- *be computed by an AC^0 circuit of depth k, and*
- $\widetilde{\deg}(f) \geq d$.

Then there exists an F on $O(n \log^2 n)$ variables that

- *is computed by an AC^0 circuit of depth $k+3$, and*
- $\widetilde{\deg}(F) \geq n^{1/2} \cdot d^{1/2}$

A General Hardness Amplification Result

Theorem (Strong Hardness Amplification Within AC^0)

Let $f\colon \{-1,1\}^n \to \{-1,1\}$

- *be computed by an AC^0 circuit of depth* k, *and*
- $\widetilde{\deg}(f) \geq d$.

Then there exists an F *on* $O(n \log^2 n)$ *variables that*

- *is computed by an AC^0 circuit of depth* $k+3$, *and*
- $\widetilde{\deg}(F) \geq n^{1/2} \cdot d^{1/2}$

Remarks:

- E.g.: If $f = \mathrm{AND}$, then $\widetilde{\deg}(F) \geq n^{3/4}$.
- Recursive application yields $\Omega(n^{1-\delta})$ bound for AC^0 function.
- Analogous result holds for monotone DNF.

Idea of the Hardness Amplification Construction

Idea of the Hardness-Amplifying Construction

- Consider the function SURJECTIVITY: $\{-1,1\}^n \to \{-1,1\}$.
  - Let $n = N \log R$. SURJ interprets its input x as a list of N numbers $(x_1, \ldots, x_N)$ from a range $[R]$.
  - $\text{SURJ}_{R,N}(x) = -1$ if and only if every element of the range $[R]$ appears at least once in the list.
- When we apply Main Theorem to $f = \text{AND}_R$, the "harder" function F is precisely $\text{SURJ}_{R,N}$.
- We show that $\widetilde{\deg}(\text{SURJ}_{R,N}) = \tilde{\Theta}(R^{1/4} \cdot N^{1/2})$.
  - If $R = \Theta(N)$, this is $\tilde{\Theta}(n^{3/4})$.

Idea of the Hardness-Amplifying Construction

- Consider the function SURJECTIVITY: $\{-1,1\}^n \to \{-1,1\}$.
  - Let $n = N \log R$. SURJ interprets its input x as a list of N numbers $(x_1, \ldots, x_N)$ from a range $[R]$.
  - $\text{SURJ}_{R,N}(x) = -1$ if and only if every element of the range $[R]$ appears at least once in the list.
- When we apply Main Theorem to $f = \text{AND}_R$, the "harder" function F is precisely $\text{SURJ}_{R,N}$.
- We show that $\widetilde{\deg}(\text{SURJ}_{R,N}) = \tilde{\Theta}(R^{1/4} \cdot N^{1/2})$.
  - If $R = \Theta(N)$, this is $\tilde{\Theta}(n^{3/4})$.
- For convenience: let's change the domain and range of all Boolean functions to $\{0,1\}^n$ and $\{0,1\}$.

Resolving the Approximate Degree of SURJ

The $\tilde{O}(R^{1/4} \cdot N^{1/2})$ Upper Bound For SURJ: First Try

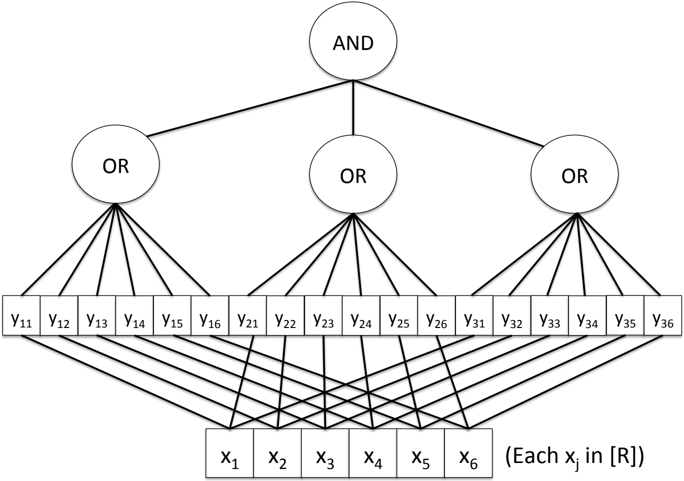
- Let's start with how to achieve a (loose) bound of $\widetilde{\deg}(\mathsf{SURJ}_{R,N}) = \tilde{O}(R^{1/2} \cdot N^{1/2})$.

The $\tilde{O}(R^{1/4} \cdot N^{1/2})$ Upper Bound For SURJ: First Try

- Let's start with how to achieve a (loose) bound of $\widetilde{\deg}(\mathsf{SURJ}_{R,N}) = \tilde{O}(R^{1/2} \cdot N^{1/2})$.
  - Let
$$y_{ij} = \begin{cases} 1 \text{ if } x_j = i \\ 0 \text{ otherwise} \end{cases}$$
  - Then

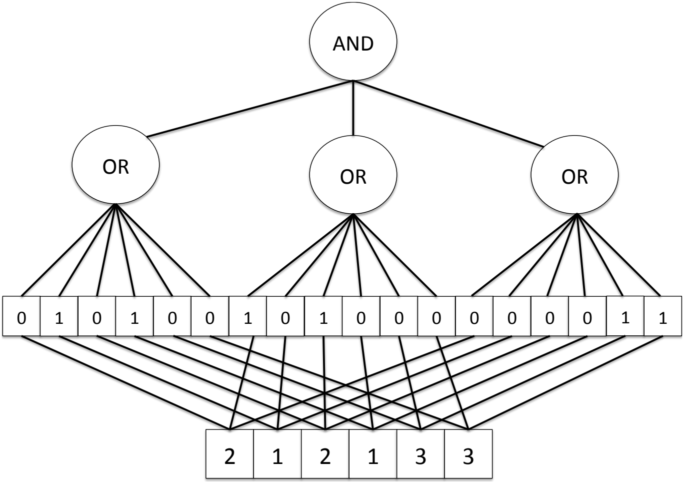
$$\mathsf{SURJ}(x) = \mathrm{AND}_R(\mathrm{OR}_N(y_{1,1}, \ldots, y_{1,N}), \ldots, \mathrm{OR}_N(y_{R,1} \ldots, y_{R,N})).$$

SURJ Illustrated $(R = 3, N = 6)$



(Each x_j in [R])

SURJ Illustrated ($R = 3, N = 6$)



The Upper Bound For SURJ: First Try

- Let's start with how to achieve a (loose) bound of $\widetilde{\deg}(\mathsf{SURJ}_{R,N}) = \tilde{O}(R^{1/2} \cdot N^{1/2})$.
  - Let

$$y_{ij} = \begin{cases} 1 \text{ if } x_j = i \\ 0 \text{ otherwise} \end{cases}$$

  - Then

$$\mathsf{SURJ}(x) = \mathrm{AND}_R(\mathrm{OR}_N(y_{1,1}, \ldots, y_{1,N}), \ldots, \mathrm{OR}_N(y_{R,1} \ldots, y_{R,N})).$$

The Upper Bound For SURJ: First Try

- Let's start with how to achieve a (loose) bound of $\widetilde{\deg}(\mathsf{SURJ}_{R,N}) = \tilde{O}(R^{1/2} \cdot N^{1/2})$.
  - Let
$$y_{ij} = \begin{cases} 1 \text{ if } x_j = i \\ 0 \text{ otherwise} \end{cases}$$
  - Then

$$\mathsf{SURJ}(x) = \mathrm{AND}_R(\mathrm{OR}_N(y_{1,1}, \ldots, y_{1,N}), \ldots, \mathrm{OR}_N(y_{R,1} \ldots, y_{R,N})).$$

  - Let p be a degree $O(R^{1/2} \cdot N^{1/2})$ polynomial approximating $\mathrm{AND}_R(\mathrm{OR}_N, \ldots, \mathrm{OR}_N)$.
  - Then $p(y_{1,1}, \ldots, y_{1,N}, \ldots, y_{R,1}, \ldots, y_{R,N})$ approximates SURJ, with degree $O(\deg(p) \cdot \log R) = O(R^{1/2} \cdot N^{1/2} \cdot \log R)$.

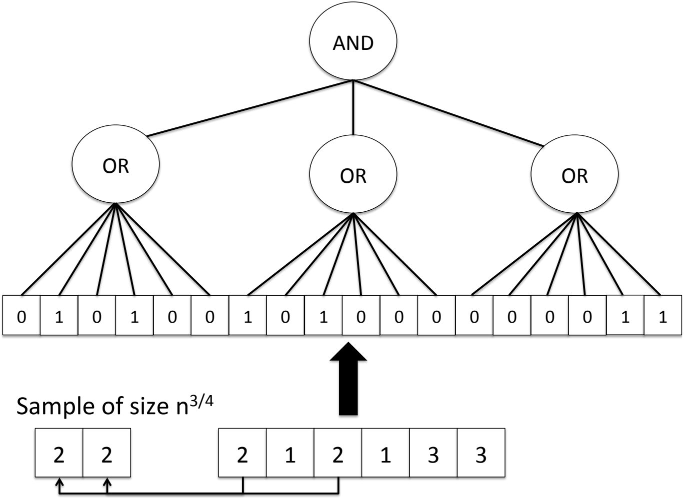
The Upper Bound For SURJ: Second Try

- Fix $R = N/2$. We'll show $\widetilde{\deg}(\mathsf{SURJ}_{R,N}) = \tilde{O}(R^{1/4} \cdot N^{1/2})$.
- We'll want to think of polynomials as computing the probability that a query algorithm outputs 1.
  - E.g., we can think of our "first try" as composing an query algorithm for computing AND_R with R copies of a query algorithm computing OR_N.
- We'll approximate SURJ via a "two-stage" construction.

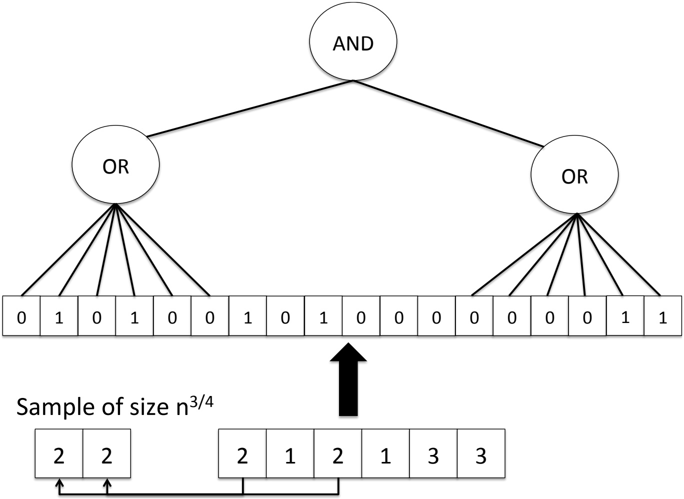
Stage 1

- Consider a query algorithm that samples $O(n^{3/4})$ inputs.
- Any range item appearing in the sample definitely has frequency at least 1, so we can just "remove it from consideration."
- Stage 2 just needs to determine whether all range items not appearing in the sample have frequency at least 1.
- Let $\mathsf{SURJ}_{\mathsf{unsamp}}$ be the function we need to compute in Stage 2.

Stage 1 Illustrated ($R = 3, N = 6$)



Stage 1 Illustrated $(R = 3, N = 6)$



Stage 2

- Key observation: any range item with frequency larger than $T = n^{1/2}$ will appear in the sample at least once, with probability $1 - \exp(-n^{1/4})$.
- i.e., if a range item doesn't appear in the sample, we are <u>really confident</u> that it does not have a very high frequency.

Stage 2

- Key observation: any range item with frequency larger than $T = n^{1/2}$ will appear in the sample at least once, with probability $1 - \exp(-n^{1/4})$.
- i.e., if a range item doesn't appear in the sample, we are really confident that it does not have a very high frequency.
- So Stage 2 only needs an approximation p to $\mathsf{SURJ}_{\mathsf{unsamp}}$ that is accurate under the assumption that no range item has frequency higher than T.

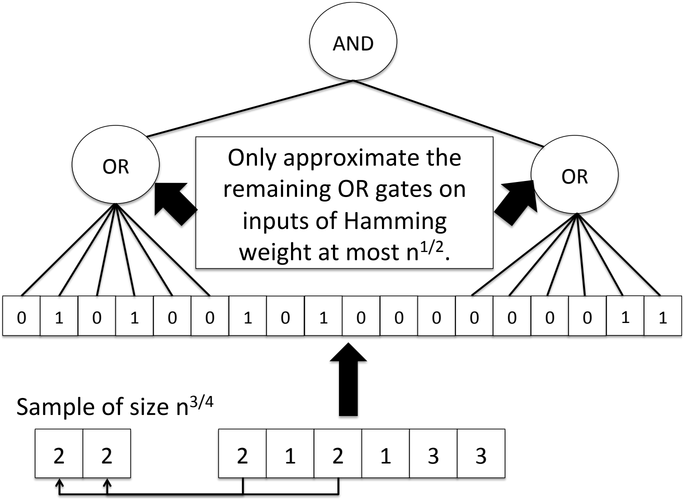
Stage 2

- Key observation: any range item with frequency larger than $T = n^{1/2}$ will appear in the sample at least once, with probability $1 - \exp(-n^{1/4})$.
- i.e., if a range item doesn't appear in the sample, we are really confident that it does not have a very high frequency.
- So Stage 2 only needs an approximation p to $\text{SURJ}_{\text{unsamp}}$ that is accurate under the assumption that no range item has frequency higher than T.
  - If p is fed an input in which some range item has frequency higher than T, then p is allowed to be exponentially large on that input.

Stage 2

- Key observation: any range item with frequency larger than $T = n^{1/2}$ will appear in the sample at least once, with probability $1 - \exp(-n^{1/4})$.
- i.e., if a range item doesn't appear in the sample, we are really confident that it does not have a very high frequency.
- So Stage 2 only needs an approximation p to $\mathsf{SURJ}_{\mathsf{unsamp}}$ that is accurate under the assumption that no range item has frequency higher than T.
  - If p is fed an input in which some range item has frequency higher than T, then p is allowed to be exponentially large on that input.
  - Specifically, if b unsampled range items have frequency larger than T, then it is okay for $|p(x)|$ to be as large as $\exp(n^{1/4} \cdot b)$.

Stage 2 Illustrated $(R = 3, N = 6)$



Stage 2 Details

Lemma (Chebyshev polynomials)

There is a polynomial q of degree $\tilde{O}(n^{1/4})$ such that

- $|q(x) - \mathrm{OR}_n(x)| \ll 1/n$ *for all* $|x| \leq n^{1/2}$.
- $|q(x)| \leq \exp\left(\tilde{O}(n^{1/4})\right)$ *otherwise.*

Theorem

For $x = (x_1, \ldots, x_R)$, *let* $b(x_1, \ldots, x_R) = \#\{i \colon |x_i| > n^{1/2}\}$. *There is a polynomial q of degree* $\tilde{O}(R^{1/2} \cdot N^{1/4})$ *such that:*

- $|q(x) - \mathrm{AND}_R \circ \mathrm{OR}_N(x)| \leq 1/3$ *if* $b(x) = 0$.
- $|p(x)| \leq \exp\left(\tilde{O}(b(x) \cdot n^{1/4})\right)$ *otherwise.*

Proof.

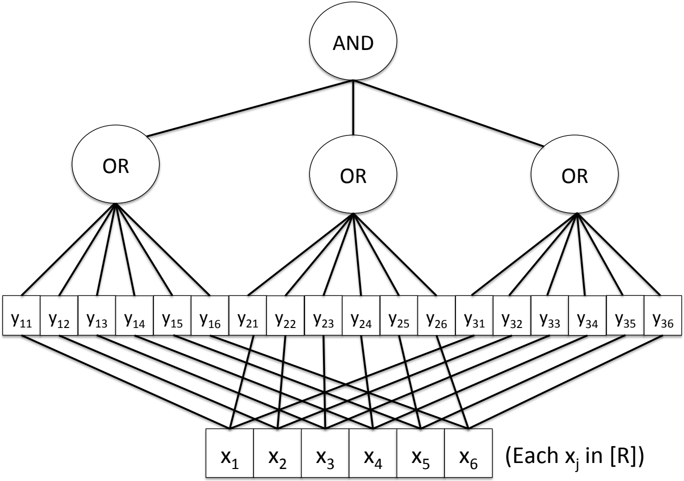
Let h approximate AND_R, and let $p = h \circ q$. □

Lower Bound Analysis for SURJ

Lower Bound Analysis for SURJ

- Recall: to approximate $\mathsf{SURJ}_{R,N}$, it is **sufficient** to approximate the block-composed function $\mathrm{AND}_R(\mathrm{OR}_N, \ldots, \mathrm{OR}_N)$ on $N \cdot R$ bits, on inputs of Hamming weight exactly N.

SURJ Illustrated $(R = 3, N = 6)$



(Each x_j in [R])

Lower Bound Analysis for SURJ

- Recall: to approximate $\mathsf{SURJ}_{R,N}$, it is **sufficient** to approximate the block-composed function $\mathrm{AND}_R(\mathrm{OR}_N, \ldots, \mathrm{OR}_N)$ on $N \cdot R$ bits, on inputs of Hamming weight exactly N.

Lower Bound Analysis for SURJ

- Recall: to approximate $\mathsf{SURJ}_{R,N}$, it is **sufficient** to approximate the block-composed function $\mathrm{AND}_R(\mathrm{OR}_N, \ldots, \mathrm{OR}_N)$ on $N \cdot R$ bits, on inputs of Hamming weight exactly N.
- Step 1: Show the converse.

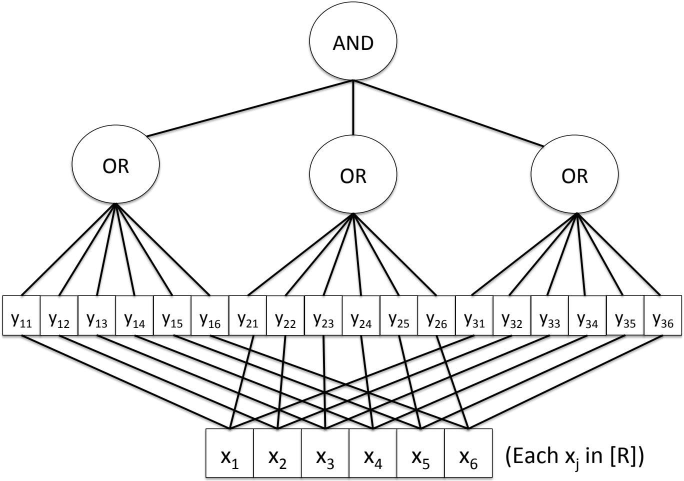
Lower Bound Analysis for SURJ

- Recall: to approximate $\mathsf{SURJ}_{R,N}$, it is **sufficient** to approximate the block-composed function $\mathrm{AND}_R(\mathrm{OR}_N, \ldots, \mathrm{OR}_N)$ on $N \cdot R$ bits, on inputs of Hamming weight exactly N.
- Step 1: Show the converse.
  - i.e., to approximate $\mathsf{SURJ}(x)$, it is **necessary** to approximate $\mathrm{AND}_R(\mathrm{OR}_N, \ldots, \mathrm{OR}_N)$, under the promise that the input has Hamming weight **at most*** N.

Lower Bound Analysis for SURJ

- Recall: to approximate $\mathsf{SURJ}_{R,N}$, it is **sufficient** to approximate the block-composed function $\mathrm{AND}_R(\mathrm{OR}_N, \ldots, \mathrm{OR}_N)$ on $N \cdot R$ bits, on inputs of Hamming weight exactly N.
- Step 1: Show the converse.
  - i.e., to approximate $\mathsf{SURJ}(x)$, it is **necessary** to approximate $\mathrm{AND}_R(\mathrm{OR}_N, \ldots, \mathrm{OR}_N)$, under the promise that the input has Hamming weight **at most*** N.
    - Follows from a symmetrization argument (Ambainis 2003).
    - *To get "at most N" rather than "equal to N", we need to introduce a dummy range item that is ignored by the function.

SURJ Illustrated $(R = 3, N = 6)$



(Each x_j in [R])

Lower Bound Analysis for SURJ

- Let $n = N \log R$.
- Recall: to approximate SURJ: $\{-1, 1\}^n \to \{-1, 1\}$, it is **sufficient** to approximate the block-composed function $\mathrm{AND}_R(\mathrm{OR}_N, \ldots, \mathrm{OR}_N)$ on $N \cdot R$ bits, on inputs of Hamming weight exactly N.
- Step 1: Show the converse.
  - To approximate SURJ(x), it is **necessary** to approximate $\mathrm{AND}_R(\mathrm{OR}_N, \ldots, \mathrm{OR}_N)$, under the promise that the input has Hamming weight **at most*** N.
    - Follows from a symmetrization argument (Ambainis 2003).
    - *To get "at most N" rather than "equal to N", we need to introduce a dummy range item that is ignored by the function.

Lower Bound Analysis for SURJ

- Let $n = N \log R$.
- Recall: to approximate SURJ: $\{-1,1\}^n \to \{-1,1\}$, it is **sufficient** to approximate the block-composed function $\mathrm{AND}_R(\mathrm{OR}_N, \ldots, \mathrm{OR}_N)$ on $N \cdot R$ bits, on inputs of Hamming weight exactly N.
- Step 1: Show the converse.
  - To approximate SURJ(x), it is **necessary** to approximate $\mathrm{AND}_R(\mathrm{OR}_N, \ldots, \mathrm{OR}_N)$, under the promise that the input has Hamming weight **at most*** N.
    - Follows from a symmetrization argument (Ambainis 2003).
    - *To get "at most N" rather than "equal to N", we need to introduce a dummy range item that is ignored by the function.
- Step 2: Prove that for some $N = O(R)$, this promise problem requires degree $\gtrsim \Omega(R^{3/4})$.

Lower Bound Analysis for SURJ

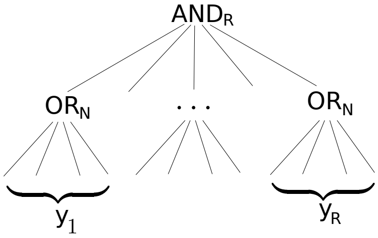
- Let $n = N \log R$.
- Recall: to approximate SURJ: $\{-1,1\}^n \to \{-1,1\}$, it is **sufficient** to approximate the block-composed function $\mathrm{AND}_R(\mathrm{OR}_N, \ldots, \mathrm{OR}_N)$ on $N \cdot R$ bits, on inputs of Hamming weight exactly N.
- Step 1: Show the converse.
  - To approximate SURJ(x), it is **necessary** to approximate $\mathrm{AND}_R(\mathrm{OR}_N, \ldots, \mathrm{OR}_N)$, under the promise that the input has Hamming weight **at most*** N.
    - Follows from a symmetrization argument (Ambainis 2003).
    - *To get "at most N" rather than "equal to N", we need to introduce a dummy range item that is ignored by the function.
- Step 2: Prove that for some $N = O(R)$, this promise problem requires degree $\gtrsim \Omega(R^{3/4})$.
  - Builds on the "dual combining technique" used earlier to analyze $\mathrm{AND\text{-}OR}_n$ (with no promise).

Overview of Step 2

Prove That For Some $N = O(R)$, Approximating $\mathrm{AND}_R \circ \mathrm{OR}_N$ Under the Promise That The Input Has Hamming Weight **At Most** N Requires Degree $\gtrsim R^{3/4}$.

Attempt 1

- For some $N = O(R)$, want a dual witness for $\mathrm{AND}_R(\mathrm{OR}_N, \ldots, \mathrm{OR}_N)$ that **only places mass on inputs of Hamming weight at most** N.



Attempt 1

- For some $N = O(R)$, want a dual witness for $\mathrm{AND}_R(\mathrm{OR}_N, \ldots, \mathrm{OR}_N)$ that **only places mass on inputs of Hamming weight at most** N.
- Attempt 1: Use the dual witness for $\mathrm{AND}_R(\mathrm{OR}_N, \ldots, \mathrm{OR}_N)$ from prior work [She09, Lee09, BT13, She13].

$$\psi_{\textsf{AND-OR}}(y_1, \ldots, y_R) := C \cdot \psi_{\textsf{AND}}(\ldots, \mathrm{sgn}(\psi_{\textsf{OR}}(y_j)), \ldots) \prod_{j=1}^{R} |\psi_{\textsf{OR}}(y_j)|$$

(C chosen to ensure $\psi_{\textsf{AND-OR}}$ has L_1-norm 1).

Attempt 1

- For some $N = O(R)$, want a dual witness for $\mathrm{AND}_R(\mathrm{OR}_N, \ldots, \mathrm{OR}_N)$ that **only places mass on inputs of Hamming weight at most** N.
- Attempt 1: Use the dual witness for $\mathrm{AND}_R(\mathrm{OR}_N, \ldots, \mathrm{OR}_N)$ from prior work [She09, Lee09, BT13, She13].

$$\psi_{\textbf{AND-OR}}(y_1, \ldots, y_R) := C \cdot \psi_{\textbf{AND}}(\ldots, \mathrm{sgn}(\psi_{\textbf{OR}}(y_j)), \ldots) \prod_{j=1}^{R} |\psi_{\textbf{OR}}(y_j)|$$

(C chosen to ensure $\psi_{\textbf{AND-OR}}$ has L_1-norm 1).

Must verify:

1. $\psi_{\textbf{AND-OR}}$ has pure high degree $\geq R^{1/2} \cdot N^{1/2} = \Omega(N)$.
2. $\psi_{\textbf{AND-OR}}$ well-correlated with AND-OR.
3. $\psi_{\textbf{AND-OR}}$ places mass only on inputs of Hamming weight $\leq N$.

Attempt 1

- For some $N = O(R)$, want a dual witness for $\mathrm{AND}_R(\mathrm{OR}_N, \ldots, \mathrm{OR}_N)$ that **only places mass on inputs of Hamming weight at most** N.
- Attempt 1: Use the dual witness for $\mathrm{AND}_R(\mathrm{OR}_N, \ldots, \mathrm{OR}_N)$ from prior work [She09, Lee09, BT13, She13].

$$\psi_{\textbf{AND-OR}}(y_1, \ldots, y_R) := C \cdot \psi_{\textbf{AND}}(\ldots, \mathrm{sgn}(\psi_{\textbf{OR}}(y_j)), \ldots) \prod_{j=1}^{R} |\psi_{\textbf{OR}}(y_j)|$$

(C chosen to ensure $\psi_{\textbf{AND-OR}}$ has L_1-norm 1).

Must verify:

1. $\psi_{\textbf{AND-OR}}$ has pure high degree $\geq R^{1/2} \cdot N^{1/2} = \Omega(N)$. ✓ [She09]
2. $\psi_{\textbf{AND-OR}}$ well-correlated with AND-OR. ✓ [BT13, She13]
3. $\psi_{\textbf{AND-OR}}$ places mass only on inputs of Hamming weight $\leq N$. ✗

Patching Attempt 1

- Goal: Fix Property 3 without destroying Properties 1 or 2.

Patching Attempt 1

- Goal: Fix Property 3 without destroying Properties 1 or 2.
- Fact (cf. Razborov and Sherstov 2008): Suppose

$$\sum_{|y|>N} |\psi_{\textbf{AND-OR}}(y)| \ll R^{-D}.$$

  - Then we can "post-process" $\psi_{\textbf{AND-OR}}$ to "zero out" any mass it places it inputs of Hamming weight larger than N.
  - While ensuring that the resulting dual witness still has pure high degree $\min\{D, \mathsf{PHD}(\psi_{\textbf{AND-OR}})\}$.

Patching Attempt 1

- New Goal: Show that, for $D \approx R^{3/4}$,

$$\sum_{|y|>N} |\psi_{\textsf{AND-OR}}(y)| \ll R^{-D}. \tag{1}$$

- Recall:

$$\psi_{\textsf{AND-OR}}(y_1, \ldots, y_R) := C \cdot \psi_{\textsf{AND}}(\ldots, \mathrm{sgn}(\psi_{\textsf{OR}}(y_j)), \ldots) \prod_{j=1}^{R} |\psi_{\textsf{OR}}(y_j)|$$

Patching Attempt 1

- New Goal: Show that, for $D \approx R^{3/4}$,

$$\sum_{|y|>N} |\psi_{\textsf{AND-OR}}(y)| \ll R^{-D}. \quad (1)$$

- Recall:

$$\psi_{\textsf{AND-OR}}(y_1, \ldots, y_R) := C \cdot \psi_{\textsf{AND}}(\ldots, \mathrm{sgn}(\psi_{\textsf{OR}}(y_j)), \ldots) \prod_{j=1}^{R} |\psi_{\textsf{OR}}(y_j)|$$

- A dual witness $\psi_{\textsf{OR}}$ for OR can be made "weakly" biased toward low Hamming weight inputs.
  - Specifically, can ensure:
    - $\mathrm{PHD}(\psi_{\textsf{OR}}) \geq n^{1/4}$.
    - For all t, $\sum_{|y_i|=t} |\psi_{\textsf{OR}}(y_i)| \leq t^{-2} \cdot \exp(-t/n^{1/4})$. (2)

Patching Attempt 1

- New Goal: Show that, for $D \approx R^{3/4}$,

$$\sum_{|y|>N} |\psi_{\textbf{AND-OR}}(y)| \ll R^{-D}. \qquad (1)$$

- Recall:

$$\psi_{\textbf{AND-OR}}(y_1, \ldots, y_R) := C \cdot \psi_{\textbf{AND}}(\ldots, \mathrm{sgn}(\psi_{\textbf{OR}}(y_j)), \ldots) \prod_{j=1}^{R} |\psi_{\textbf{OR}}(y_j)|$$

- A dual witness $\psi_{\textbf{OR}}$ for OR can be made "weakly" biased toward low Hamming weight inputs.
  - Specifically, can ensure:
    - PHD$(\psi_{\textbf{OR}}) \geq n^{1/4}$.
    - For all t, $\sum_{|y_i|=t} |\psi_{\textbf{OR}}(y_i)| \leq t^{-2} \cdot \exp(-t/n^{1/4})$. (2)
- $|\psi_{\textbf{AND-OR}}(y_1,\ldots,y_R)|$ resembles product distribution: $\prod_{j=1}^{R} |\psi_{\textbf{OR}}(y_j)|$
- So it is exponentially more biased toward low Hamming weight inputs than $\psi_{\textbf{OR}}$ itself.

Patching Attempt 1

- New Goal: Show that, for $D \approx R^{3/4}$,

$$\sum_{|y|>N} |\psi_{\textsf{AND-OR}}(y)| \ll R^{-D}. \qquad (1)$$

- Recall:

$$\psi_{\textsf{AND-OR}}(y_1, \ldots, y_R) := C \cdot \psi_{\textsf{AND}}(\ldots, \mathrm{sgn}(\psi_{\textsf{OR}}(y_j)), \ldots) \prod_{j=1}^{R} |\psi_{\textsf{OR}}(y_j)|$$

- A dual witness $\psi_{\textsf{OR}}$ for OR can be made "weakly" biased toward low Hamming weight inputs.
  - Specifically, can ensure:
    - PHD$(\psi_{\textsf{OR}}) \geq n^{1/4}$.
    - For all t, $\sum_{|y_i|=t} |\psi_{\textsf{OR}}(y_i)| \leq t^{-2} \cdot \exp(-t/n^{1/4})$. (2)
- Intuition: By (2): the mass that $\prod_{j=1}^{R} |\psi_{\textsf{OR}}(y_j)|$ places on inputs of Hamming weight $> N$ is dominated by inputs with $|y_i| = N^{1/4}$ for at least $N^{3/4}$ values of i.
- Also by (2), each $|y_i| = N^{1/4}$ contributes a factor of $1/\mathrm{poly}(N)$.

Patching Attempt 1

- New Goal: Show that, for $D \approx R^{3/4}$,

$$\sum_{|y|>N} |\psi_{\textbf{AND-OR}}(y)| \ll R^{-D}. \qquad (1)$$

- Recall:

$$\psi_{\textbf{AND-OR}}(y_1, \ldots, y_R) := C \cdot \psi_{\textbf{AND}}(\ldots, \mathrm{sgn}(\psi_{\textbf{OR}}(y_j)), \ldots) \prod_{j=1}^{R} |\psi_{\textbf{OR}}(y_j)|$$

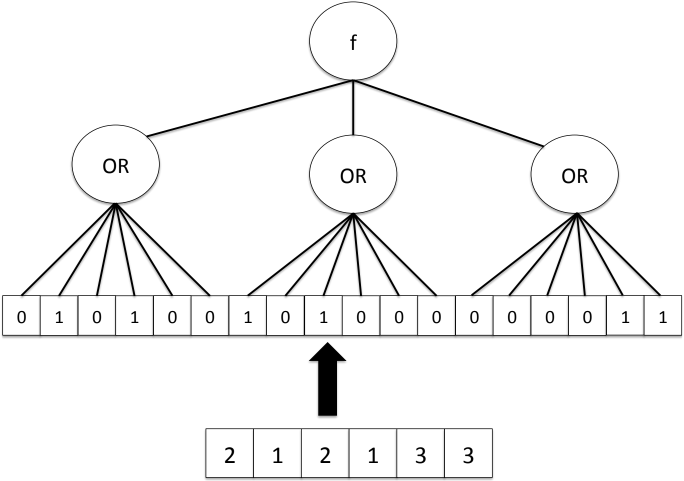
- A dual witness $\psi_{\textbf{OR}}$ for OR can be made "weakly" biased toward low Hamming weight inputs.
  - Specifically, can ensure:
    - $\mathrm{PHD}(\psi_{\textbf{OR}}) \geq n^{1/4}$.
    - For all t, $\sum_{|y_i|=t} |\psi_{\textbf{OR}}(y_i)| \leq t^{-2} \cdot \exp(-t/n^{1/4})$. (2)
- Intuition: By (2): the mass that $\prod_{j=1}^{R} |\psi_{\textbf{OR}}(y_j)|$ places on inputs of Hamming weight $> N$ is dominated by inputs with $|y_i| = N^{1/4}$ for at least $N^{3/4}$ values of i.
- So total mass on these inputs is $\exp(-\Omega(N^{3/4}))$.

General Hardness Amplification Within AC^0

General Hardness Amplification Within AC^0

- Recall: When we apply our hardness amplification to $f = \mathrm{AND}_R$, the "harder" function F is precisely SURJ.
- For a general function f, what is the "harder" function F?

First Attempt: Amplifying Hardness of $f: \{-1,1\}^R \to \{-1,1\}$ ($R=3, N=6$)



Hardness-Amplifying Construction: Second Attempt

- First attempt at handling general f fails when $f = \mathrm{OR}$.
  - $F(x) = \mathrm{OR}_R(\mathrm{OR}_N(y_{1,1}, \ldots, y_{1,N}), \ldots, \mathrm{OR}_N(y_{R,1} \ldots, y_{R,N}))$ has (exact) degree 0.

Hardness-Amplifying Construction: Second Attempt

- First attempt at handling general f fails when $f = \mathrm{OR}$.
  - $F(x) = \mathrm{OR}_R(\mathrm{OR}_N(y_{1,1}, \ldots, y_{1,N}), \ldots, \mathrm{OR}_N(y_{R,1} \ldots, y_{R,N}))$ has (exact) degree 0.
- Let $R' = R \log R$. For $f \colon \{-1,1\}^R \to \{-1,1\}$, the real* definition of F is:

$$F(x) = (f \circ \mathrm{AND}_{\log R})(\mathrm{OR}_N(y_{1,1}, \ldots, y_{1,N}), \ldots, \mathrm{OR}_N(y_{R',1}, \ldots, y_{R',N}))$$

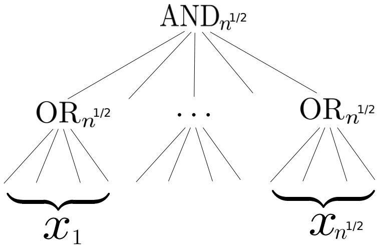
*This is still a slight simplification. Also need to introduce a dummy range item that is ignored by F.

Future Directions

- Resolve the quantum query complexity of k-distinctness, counting triangles, graph collision, etc.
- Prove an $\Omega(n^{k/(k+1)})$ lower bound on approximate degree of the k-sum function?
  - Its quantum query complexity is known to be $\Theta(n^{k/(k+1)})$.
- An $\Omega(n)$ lower bound on the approximate degree of AC^0?
- A sublinear <u>upper bound</u> for DNFs of polynomial size? Or even polynomial size AC^0 circuits?
  - Either result would yield new circuit lower bounds (namely, for $\mathsf{AC}^0 \circ \mathsf{MOD}_2$ circuits).
- Extend our bounds on $\widetilde{\deg}_\epsilon(f)$ from $\epsilon = 1/3$ to ϵ much closer to 1.
  - We believe our techniques can extend to give a $\Omega(n^{1-\delta})$ lower bound on the <u>threshold degree</u> of AC^0.

Thank you!

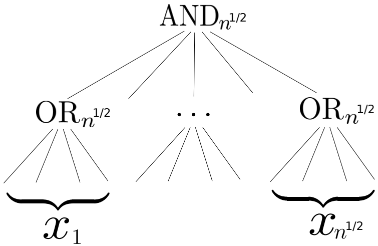
Analysis of the Dual Witness for the AND-OR Tree



The Combining Method [SZ09, She09, Lee09]

$$\psi_{\textbf{AND-OR}}(x_1, \ldots, x_{n^{1/2}}) := C \cdot \psi_{\textbf{OUT}}(\ldots, \mathrm{sgn}(\psi_{\textbf{IN}}(x_i)), \ldots) \prod_{i=1}^{n^{1/2}} |\psi_{\textbf{IN}}(x_i)|$$

(C chosen to ensure $\psi_{\textbf{AND-OR}}$ has L_1-norm 1).



The Combining Method [SZ09, She09, Lee09]

$$\psi_{\textbf{AND-OR}}(x_1, \ldots, x_{n^{1/2}}) := C \cdot \psi_{\textbf{OUT}}(\ldots, \operatorname{sgn}(\psi_{\textbf{IN}}(x_i)), \ldots) \prod_{i=1}^{n^{1/2}} |\psi_{\textbf{IN}}(x_i)|$$

(C chosen to ensure $\psi_{\textbf{AND-OR}}$ has L_1-norm 1).

Must verify:

1. $\psi_{\textbf{AND-OR}}$ has pure high degree $\geq n^{1/4} \cdot n^{1/4} = n^{1/2}$.
2. $\psi_{\textbf{AND-OR}}$ has high correlation with $\mathrm{AND\text{-}OR}$.

The Combining Method [SZ09, She09, Lee09]

$$\psi_{\textbf{AND-OR}}(x_1, \ldots, x_{n^{1/2}}) := C \cdot \psi_{\textbf{OUT}}(\ldots, \operatorname{sgn}(\psi_{\textbf{IN}}(x_i)), \ldots) \prod_{i=1}^{n^{1/2}} |\psi_{\textbf{IN}}(x_i)|$$

(C chosen to ensure $\psi_{\textbf{AND-OR}}$ has L_1-norm 1).

Must verify:

1. $\psi_{\textbf{AND-OR}}$ has pure high degree $\geq n^{1/4} \cdot n^{1/4} = n^{1/2}$.✓[She09]
2. $\psi_{\textbf{AND-OR}}$ has high correlation with AND-OR. [BT13, She13]

Pure High Degree Analysis [She09]

$$\psi_{\textbf{AND-OR}}(x_1, \ldots, x_{n^{1/2}}) := C \cdot \psi_{\textbf{OUT}}(\ldots, \operatorname{sgn}(\psi_{\textbf{IN}}(x_i)), \ldots) \prod_{i=1}^{n^{1/2}} |\psi_{\textbf{IN}}(x_i)|$$

- Intuition: Consider $\psi_{\textbf{OUT}}(y_1, y_2, y_3) = y_1 y_2$. Then $\psi_{\textbf{AND-OR}}(x_1, x_2, x_3)$ equals:

$$C \cdot \operatorname{sgn}(\psi_{\textbf{IN}}(x_1)) \cdot \operatorname{sgn}(\psi_{\textbf{IN}}(x_2)) \cdot \prod_{i=1}^{3} |\psi_{\textbf{IN}}(x_i)|$$
$$= \psi_{\textbf{IN}}(x_1) \cdot \psi_{\textbf{IN}}(x_2) \cdot |\psi_{\textbf{IN}}(x_3)|$$

Pure High Degree Analysis [She09]

$$\psi_{\textbf{AND-OR}}(x_1, \ldots, x_{n^{1/2}}) := C \cdot \psi_{\textbf{OUT}}(\ldots, \text{sgn}(\psi_{\textbf{IN}}(x_i)), \ldots) \prod_{i=1}^{n^{1/2}} |\psi_{\textbf{IN}}(x_i)|$$

- Intuition: Consider $\psi_{\textbf{OUT}}(y_1, y_2, y_3) = y_1 y_2$. Then $\psi_{\textbf{AND-OR}}(x_1, x_2, x_3)$ equals:

$$C \cdot \text{sgn}(\psi_{\textbf{IN}}(x_1)) \cdot \text{sgn}(\psi_{\textbf{IN}}(x_2)) \cdot \prod_{i=1}^{3} |\psi_{\textbf{IN}}(x_i)|$$

$$= \psi_{\textbf{IN}}(x_1) \cdot \psi_{\textbf{IN}}(x_2) \cdot |\psi_{\textbf{IN}}(x_3)|$$

- Each term of $\psi_{\textbf{AND-OR}}$ is the product of PHD($\psi_{\textbf{OUT}}$) polynomials over disjoint variable sets, each of pure high degree at least PHD($\psi_{\textbf{IN}}$).
- So $\psi_{\textbf{AND-OR}}$ has pure high degree at least PHD($\psi_{\textbf{OUT}}$) $\cdot$ PHD($\psi_{\textbf{IN}}$).

(Sub)Goal: Show $\psi_{\textbf{AND-OR}}$ has high correlation with $\mathrm{AND\text{-}OR}$

Correlation Analysis

$$\psi_{\textsf{\textbf{AND-OR}}}(x_1, \ldots, x_{n^{1/2}}) := C \cdot \psi_{\textsf{\textbf{OUT}}}(\ldots, \operatorname{sgn}(\psi_{\textsf{\textbf{IN}}}(x_i)), \ldots) \prod_{i=1}^{n^{1/2}} |\psi_{\textsf{\textbf{IN}}}(x_i)|$$

- Idea: Show

$$\sum_{x \in \{-1,1\}^n} \psi_{\textsf{\textbf{AND-OR}}}(x) \cdot \text{AND-OR}_n(x) \approx \sum_{y \in \{-1,1\}^{n^{1/2}}} \psi_{\textsf{\textbf{OUT}}}(y) \cdot \text{AND}_{n^{1/2}}(y).$$

- Intuition: We are feeding $\operatorname{sgn}(\psi_{\textsf{\textbf{IN}}}(x_i))$ into $\psi_{\textsf{\textbf{OUT}}}$.
- $\psi_{\textsf{\textbf{IN}}}$ is **correlated** with $\text{OR}_{n^{1/2}}$, so $\operatorname{sgn}(\psi_{\textsf{\textbf{IN}}}(x_i))$ is a "decent predictor" of $\text{OR}_{n^{1/2}}$.
- But there are errors. Need to show errors don't "build up".

Correlation Analysis

- Goal: Show

$$\sum_{x \in \{-1,1\}^n} \psi_{\textsf{AND-OR}}(x) \cdot \mathrm{AND\text{-}OR}_n(x) \approx \sum_{y \in \{-1,1\}^{n^{1/2}}} \psi_{\textsf{OUT}}(y) \cdot \mathrm{AND}_{n^{1/2}}(y).$$

- Case 1: Consider any $y = (\operatorname{sgn} \psi_{\textsf{IN}}(x_1), \ldots, \operatorname{sgn} \psi_{\textsf{IN}}(x_{n^{1/2}})) \neq$ **All-True**.
- There is some coordinate of y that equals FALSE. Only need to "trust" this coordinate to force $\mathrm{AND\text{-}OR}_n$ to evaluate to FALSE on $(x_1, \ldots, x_{n^{1/2}})$. So errors do not build up!

Correlation Analysis

- Case 2: Consider $y =$ **All-True**.
- $\mathrm{AND}_{n^{1/2}}(y) = \mathrm{AND\text{-}OR}_n(x_1, \ldots, x_{n^{1/2}})$ only if all coordinates of y are "error-free".
- Fortunately, ψ_{IN} has a special **one-sided error** property:

  If $\mathrm{sgn}(\psi_{\mathsf{IN}}(x_i)) = -1$, then $\mathrm{OR}_{n^{1/2}}(x_i)$ is **guaranteed** to equal -1.

Summary of Correlation Analysis

- Two Cases.
- In first case (feeding at least one FALSE into $\psi_{\mathbf{OUT}}$), errors did not build up, because we only needed to "trust" the FALSE value.
- In second case (all values fed into $\psi_{\mathbf{OUT}}$ are TRUE), we needed to trust all values. But we could do this because $\psi_{\mathbf{IN}}$ had one-sided error.

One-Sided Approximate Degree

- A real polynomial p is a <u>one-sided ϵ-approximation</u> for f if

$$|p(x) - 1| < \epsilon \quad \forall x \in f^{-1}(-1)$$

$$p(x) \geq 1 \quad \forall x \in f^{-1}(1)$$

- $\widetilde{\text{odeg}}_{-,\epsilon}(f) =$ min degree of a one-sided ϵ-approximation for f.
- $\widetilde{\text{odeg}}_{-}(f) := \widetilde{\text{odeg}}_{-,1/3}(f)$ is the one-sided approximate degree of f.

Dual Formulation of $\widetilde{\text{odeg}}_{-}$

Primal LP (Linear in ϵ and coefficients of p):

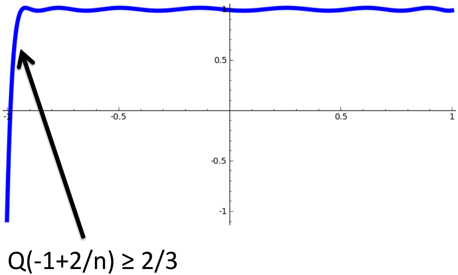
$$\begin{aligned} \min_{p,\epsilon} \quad & \epsilon \\ \text{s.t.} \quad & |p(x) - 1| \leq \epsilon && \text{for all } x \in f^{-1}(-1) \\ & p(x) \geq 1 && \text{for all } x \in f^{-1}(1) \\ & \deg p \leq d \end{aligned}$$

Dual LP:

$$\begin{aligned} \max_{\psi} \quad & \sum_{x \in \{-1,1\}^n} \psi(x) f(x) \\ \text{s.t.} \quad & \sum_{x \in \{-1,1\}^n} |\psi(x)| = 1 \\ & \sum_{x \in \{-1,1\}^n} \psi(x) q(x) = 0 && \text{whenever } \deg q \leq d \\ & \psi(x) \geq 0 \quad \forall x \in f^{-1}(1) \end{aligned}$$

Proof that $\widetilde{\text{odeg}}\_(\text{AND}_n) = \Omega(\sqrt{n})$

We argued that the symmetrization of any $1/3$-approximation to AND_n had to look like this:



Andris Ambainis, Aleksandrs Belovs, Oded Regev, and Ronald de Wolf.
Efficient quantum algorithms for (gapped) group testing and junta testing.
In Proceedings of the Twenty-Seventh Annual ACM-SIAM Symposium on Discrete Algorithms, pages 903–922. Society for Industrial and Applied Mathematics, 2016.

Scott Aaronson and Yaoyun Shi.
Quantum lower bounds for the collision and the element distinctness problems.
J. ACM, 51(4):595–605, 2004.

Robert Beals, Harry Buhrman, Richard Cleve, Michele Mosca, and Ronald de Wolf.
Quantum lower bounds by polynomials.
J. ACM, 48(4):778–797, 2001.

Aleksandrs Belovs.
Learning-graph-based quantum algorithm for k-distinctness.
In Foundations of Computer Science (FOCS), 2012 IEEE 53rd Annual Symposium on, pages 207–216. IEEE, 2012.

Sergey Bravyi, Aram Wettroth Harrow, and Avinatan Hassidim.
Quantum algorithms for testing properties of distributions.
IEEE Trans. Information Theory, 57(6):3971–3981, 2011.

Harry Buhrman, Ilan Newman, Hein Röhrig, and Ronald de Wolf.
Robust polynomials and quantum algorithms.
Theory Comput. Syst., 40(4):379–395, 2007.

Troy Lee, Rajat Mittal, Ben W. Reichardt, Robert Špalek, and Mario Szegedy.

Quantum query complexity of state conversion.
In Proceedings of the 52nd Symposium on Foundations of Computer Science (FOCS 2011), pages 344–353, 2011.

Tongyang Li and Xiaodi Wu.
Quantum query complexity of entropy estimation.
arXiv preprint arXiv:1710.06025, 2017.

Ben W Reichardt.
Reflections for quantum query algorithms.
In Proceedings of the twenty-second annual ACM-SIAM symposium on Discrete Algorithms, pages 560–569. Society for Industrial and Applied Mathematics, 2011.

Alexander A. Sherstov.
The pattern matrix method.
SIAM J. Comput., 40(6):1969–2000, 2011.
Preliminary version in STOC 2008.

Alexander A. Sherstov.
Approximating the AND-OR Tree.
Theory of Computing, 9(20):653–663, 2013.

Alexander A. Sherstov, 2017.
Personal communication.