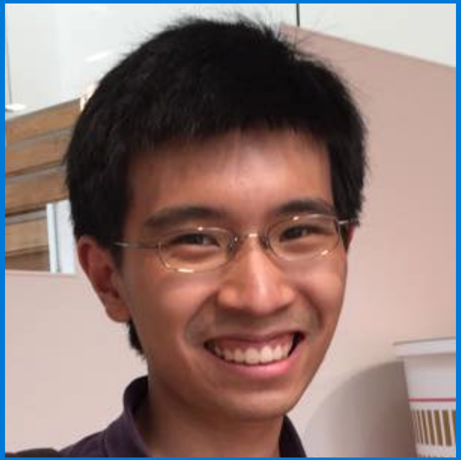
The Polynomial Method Strikes Back:
Tight Quantum Query Bounds via Dual Polynomials



Mark Bun
Princeton University

Robin Kothari
Microsoft Research

Justin Thaler
Georgetown University

arXiv:1710.09079

Query complexity

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

$$x = \boxed{x_1 \,|\, x_2 \,|\, x_3 \,|\, \quad \cdots \quad \,|\, x_n}$$

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:

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

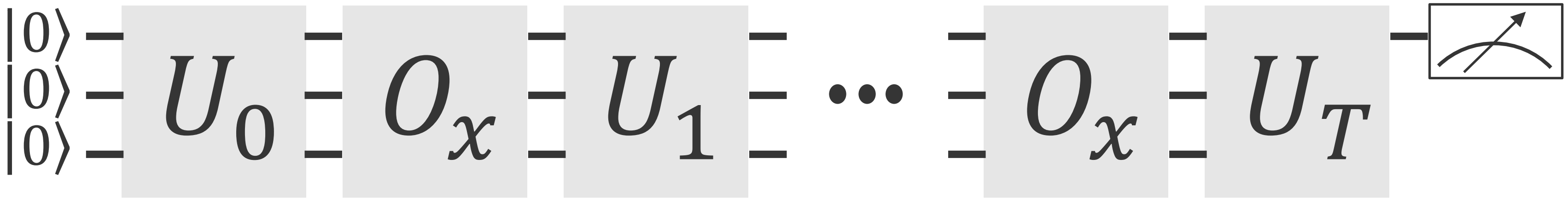
In the quantum setting, we have this oracle:

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

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

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

Why query complexity?

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.

Complexity theoretic motivation

- We can prove statements about the power of different computational models! (E.g., exponential separation between classical and quantum algorithms)
- Oracle separations between classes, lower bounds on restricted models, upper and lower bounds in communication complexity, circuit complexity, data structures, etc.

Lower bounds on quantum query complexity

Positive-weights adversary method

Easy to use, but has many limitations. Cannot show any of the results of our work.

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!

Approximate degree

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

$\widetilde{\deg}(f)$

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

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

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

The polynomial method

- For any T-query quantum algorithm A, there is a polynomial p of degree 2T such that:
  - For all $x \in \{-1,1\}^n, p(x)$ equals the probability that A outputs 1 on input x.

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 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

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

Approximate degree

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

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

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

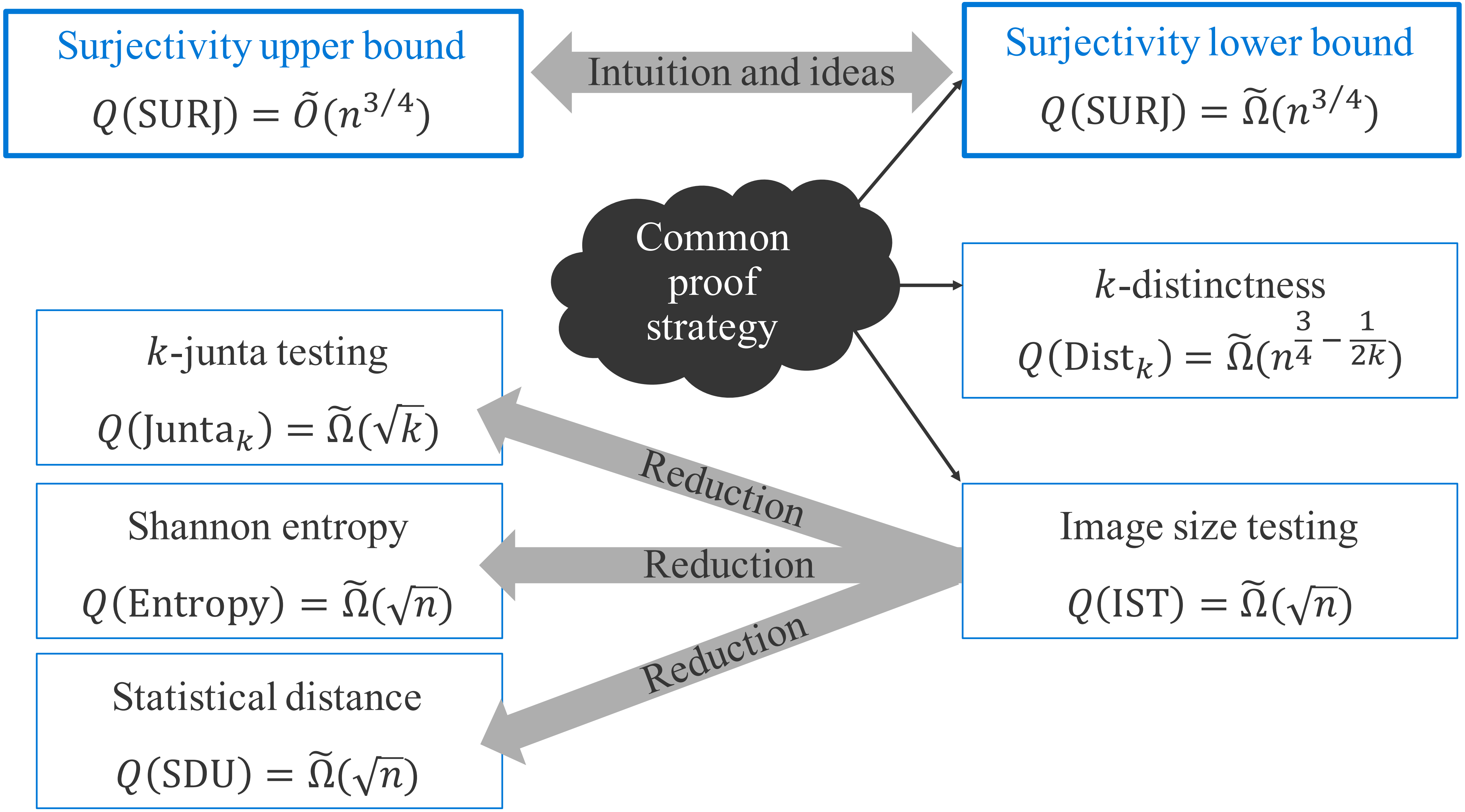
Summary of results

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
$Q(\mathrm{SURJ}) = \tilde{\Omega}(n^{3/4})$
Common proof strategy
k-distinctness
$Q(\mathrm{Dist}_k) = \tilde{\Omega}(n^{\frac{3}{4}-\frac{1}{2k}})$
k-junta testing
$Q(\mathrm{Junta}_k) = \tilde{\Omega}(\sqrt{k})$
Reduction
Shannon entropy
$Q(\mathrm{Entropy}) = \tilde{\Omega}(\sqrt{n})$
Reduction
Image size testing
$Q(\mathrm{IST}) = \tilde{\Omega}(\sqrt{n})$
Statistical distance
$Q(\mathrm{SDU}) = \tilde{\Omega}(\sqrt{n})$
Reduction

Getting To Know Approximate Degree

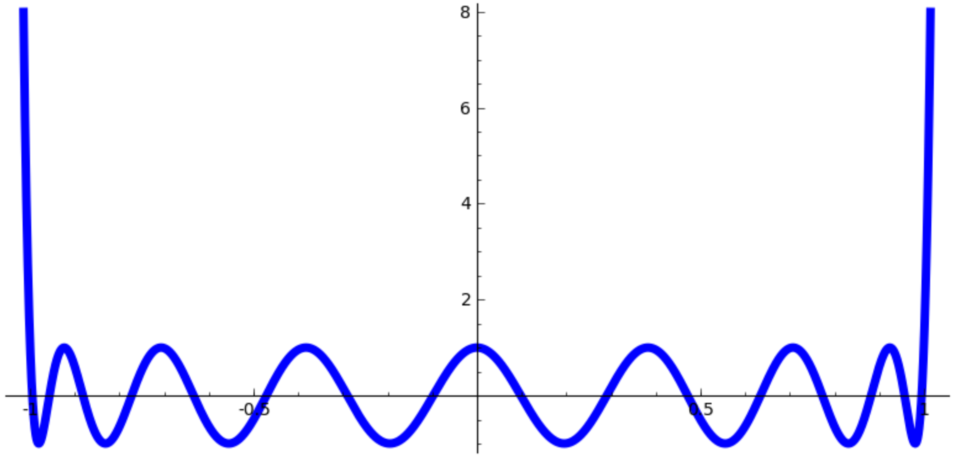
The Approximate Degree of AND_n

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

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

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

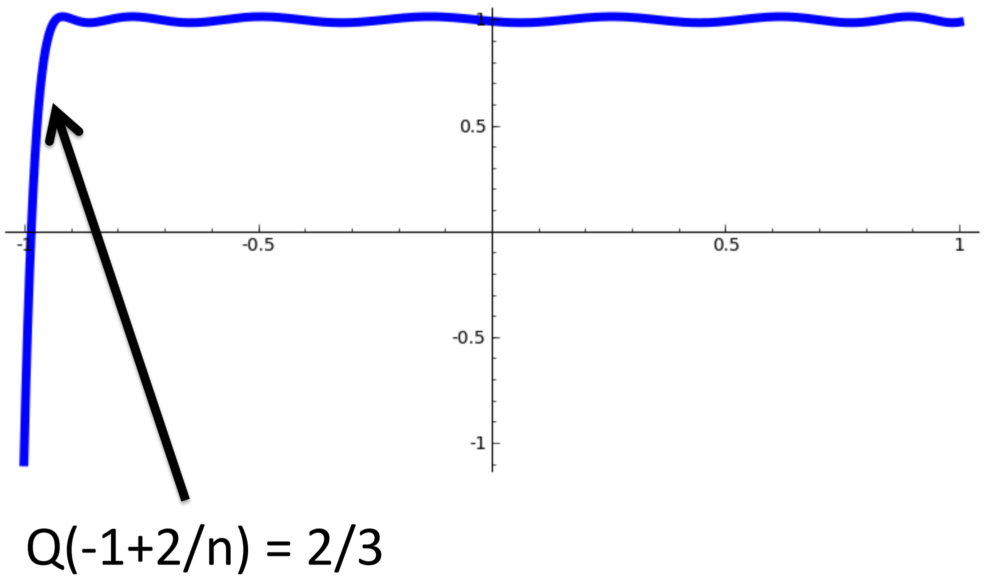
- Chebyshev polynomials are the extremal case.



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:

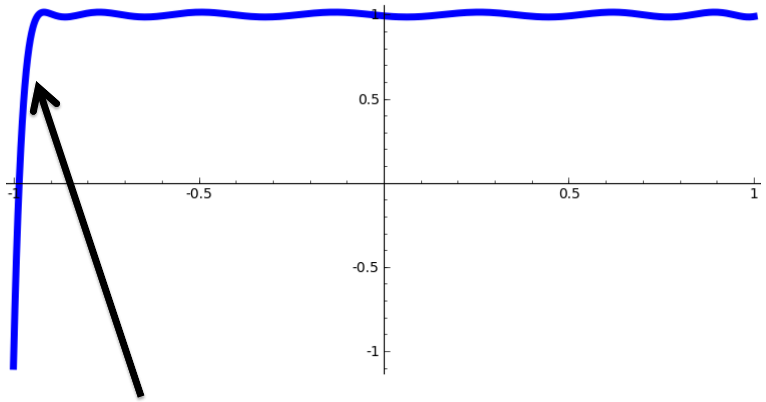


- Define n-variate polynomial p via $p(x) = Q(\sum_{i=1}^{n} x_i/n)$.
- Then $|p(x) - \mathrm{AND}_n(x)| \leq 1/3 \quad \forall x \in \{-1,1\}^n$.

The Approximate Degree of AND_n

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

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



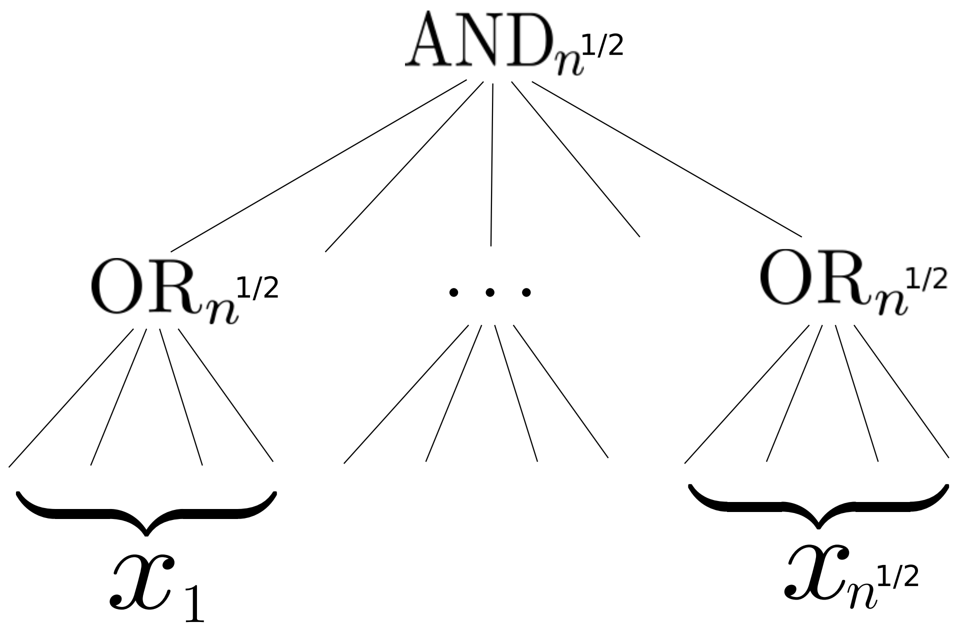
Q(-1+2/n) ≥ 2/3

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

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 $\widetilde{\deg}(\mathrm{AND\text{-}OR}_n)$?



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 <u>method of dual polynomials</u>

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):

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

Dual LP:

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

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 $\deg q \le d$ "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: $\deg_\epsilon(f) > d$ 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"

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

Goal: Construct a Dual Polynomial for the AND-OR Tree

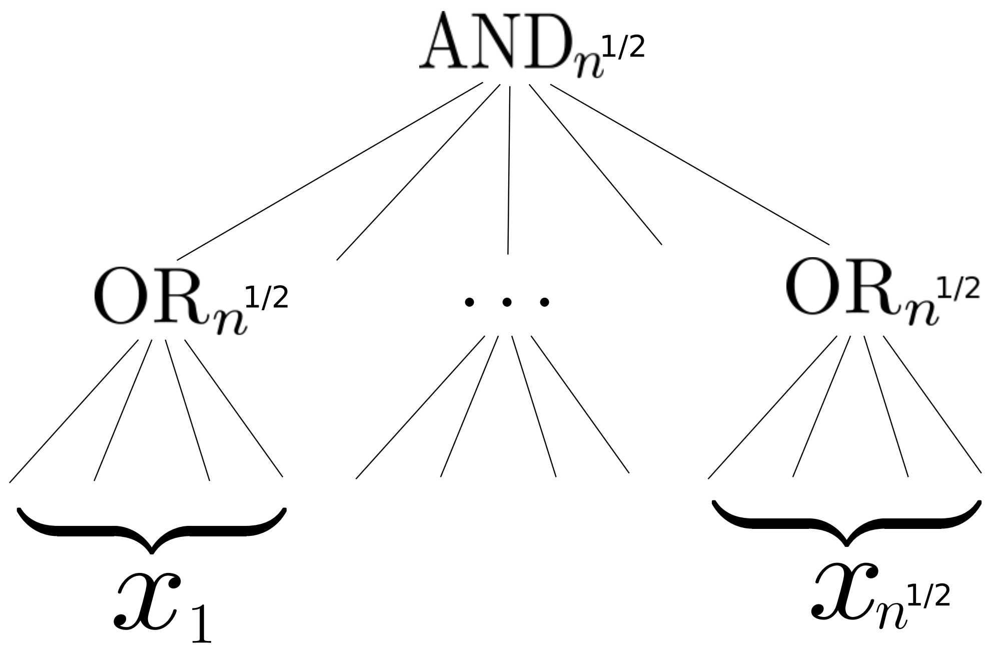
Constructing a Dual Polynomial

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

The Combining Technique

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

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



The Combining Technique

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

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

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 $\mathrm{AND\text{-}OR}$. [BT13, She13]

Our Work: Resolving the Approximate Degree of Surjectivity

Surjectivity

$\mathrm{Surj}_{\mathrm{R,N}}$: Input consists of $n = \mathrm{N} \cdot \log_2(\mathrm{R})$ bits, interpreted as a list of N numbers in $[R]$. Does every $r \in [R]$ appear at least once in the list?

Our result: $\widetilde{\deg}(\mathrm{SURJ}_{R,\,N}) = \widetilde{\Theta}\left(R^{\frac{1}{4}} \cdot N^{\frac{1}{2}}\right)$.

- Let's start with the upper bound.
- For the upper bound, let's change the domain and range of all functions to $\{0,1\}^n$ and $\{0,1\}$.

The SURJ Upper Bound: First Try

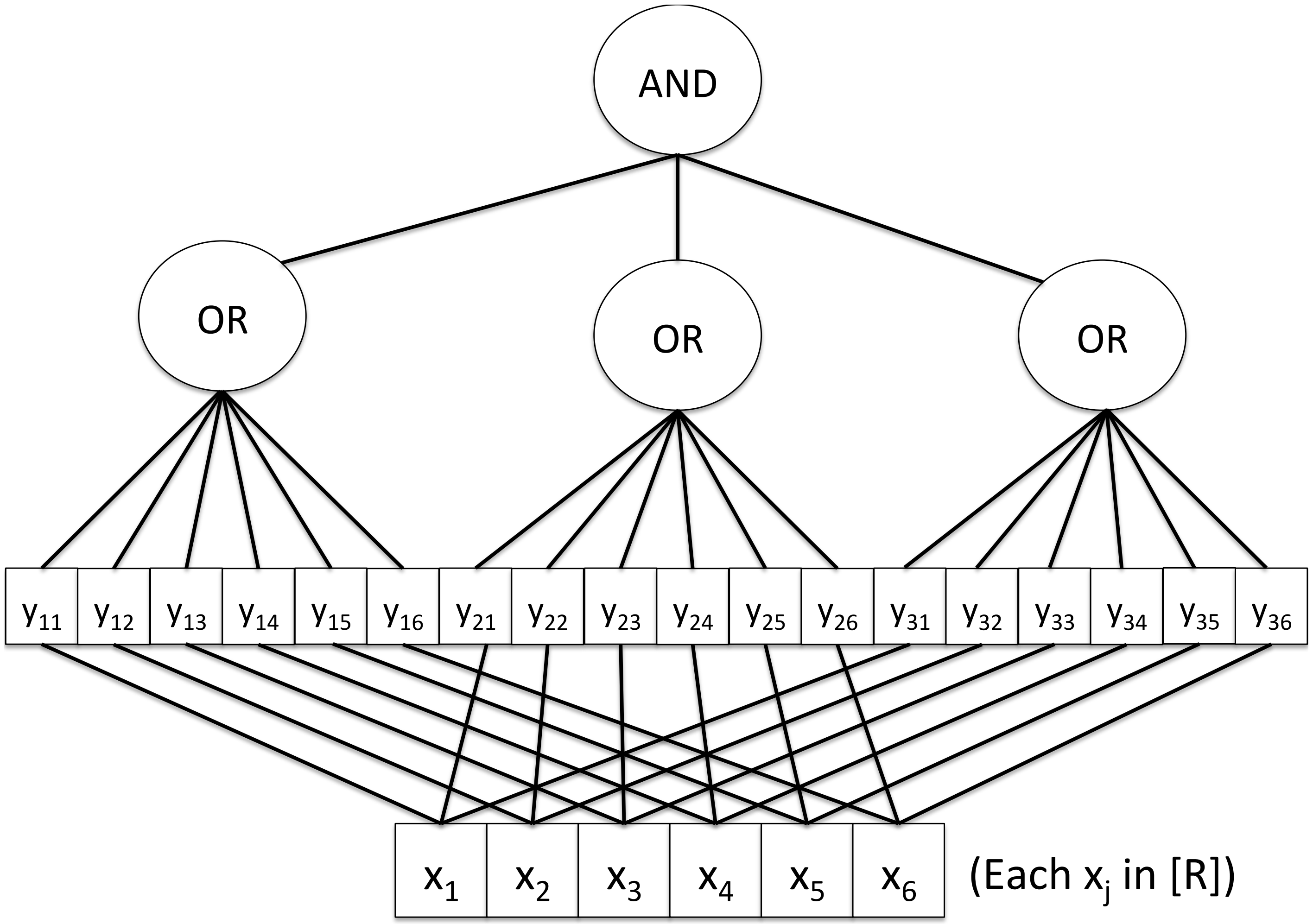
- Let's start with how to achieve a (loose) bound of $\widetilde{\deg}(\mathsf{SURJ}_{R,N}) = \tilde{O}(R^{1/2} \cdot N^{1/2})$.
  - Let

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

  - Then

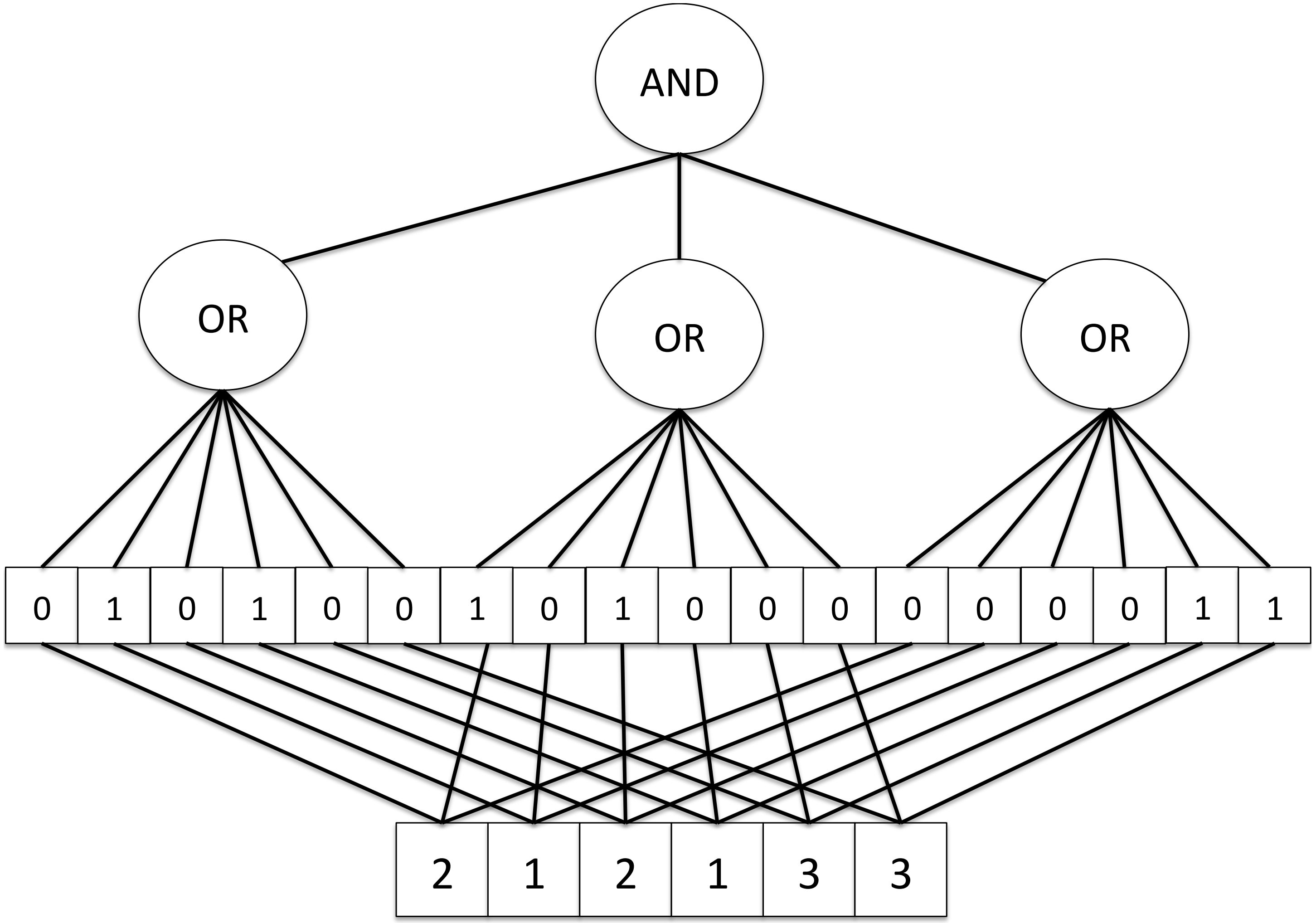
$$\mathsf{SURJ}(x) = \mathrm{AND}_R(\mathrm{OR}_N(y_{1,1}, \dots, y_{1,N}), \dots, \mathrm{OR}_N(y_{R,1} \dots, y_{R,N})).$$

SURJ Illustrated (R=3, N=6)



(Each x_j in [R])

SURJ Illustrated (R=3, N=6)



The SURJ Upper Bound: 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 $O(\deg(p) \cdot \log R) = O(R^{1/2} \cdot N^{1/2} \cdot \log R)$.

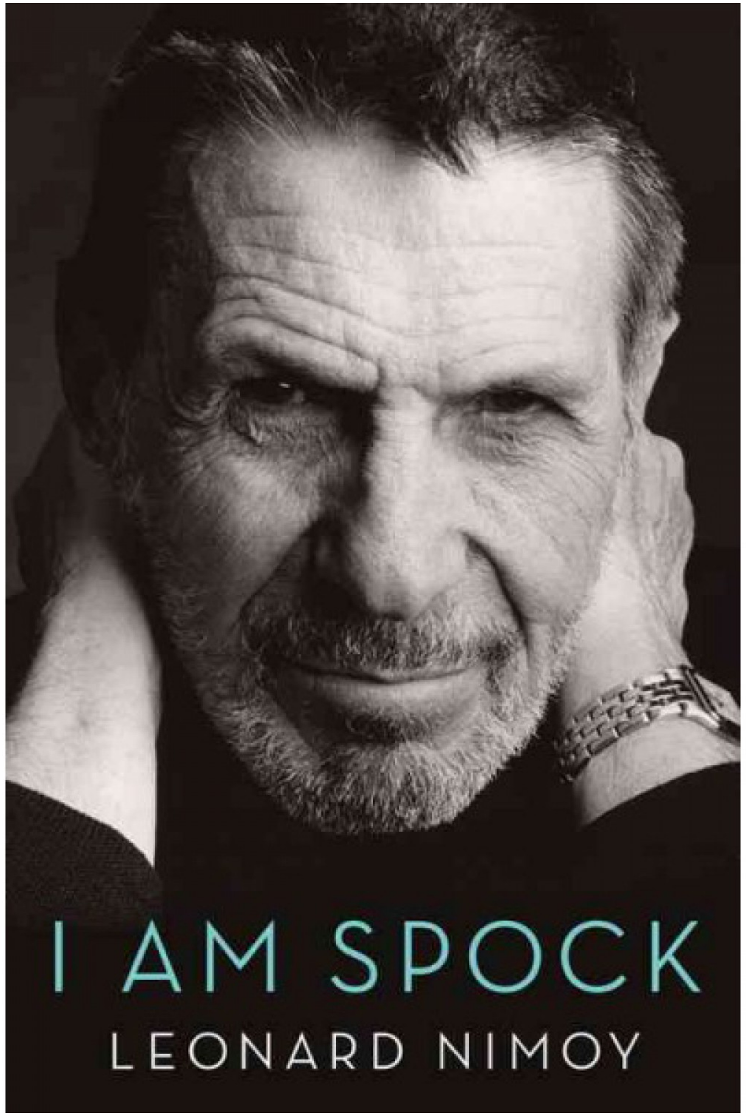
Tight Upper Bound For SURJ

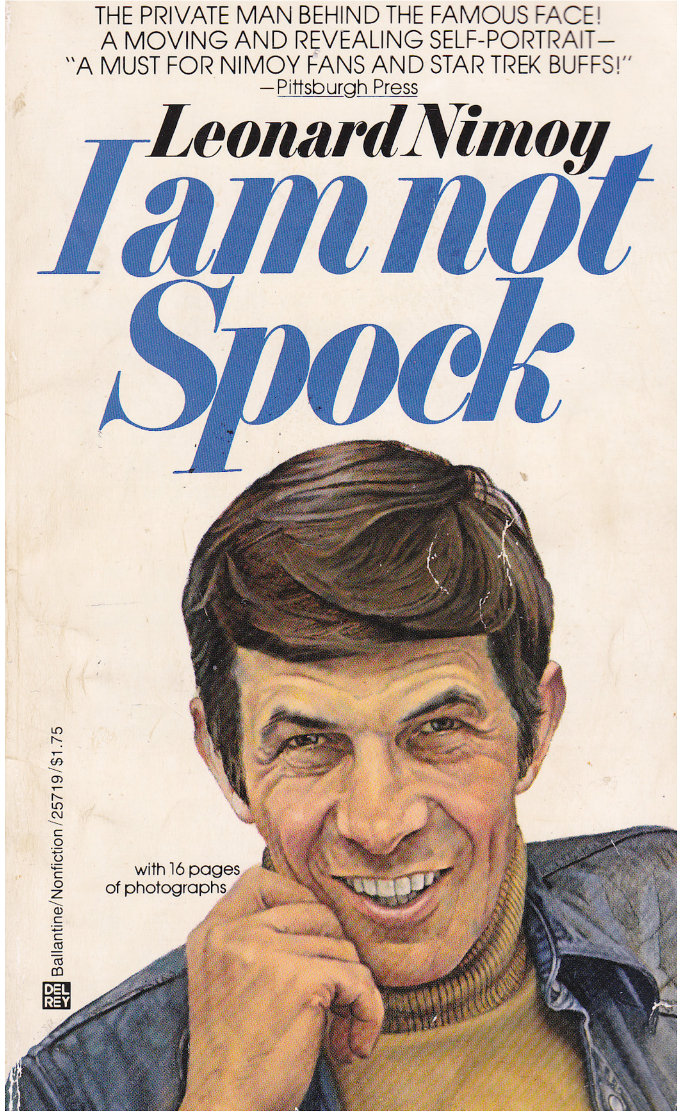
Overview of the tight upper bound

- Previous slide showed that to approximate $\mathrm{SURJ}_{\mathrm{R,N}}$, suffices to approximate the block-composed function $\mathrm{AND}_R \circ \mathrm{OR}_N$ on inputs of Hamming weight exactly N.
- The approximation is allowed to take arbitrary values on all other inputs!
- Denote this function $(\mathrm{AND}_R \circ \mathrm{OR}_N)^N$.
- Important: $\mathrm{AND}_R \circ \mathrm{OR}_N \neq (\mathrm{AND}_R \circ \mathrm{OR}_N)^N$
- $\widetilde{\deg}(\mathrm{AND}_R \circ \mathrm{OR}_N) = \Theta\left(\sqrt{RN}\right)$.
- We'll show that $\widetilde{\deg}\left((\mathrm{AND}_R \circ \mathrm{OR}_N)^N\right) = \widetilde{\Theta}\left(\widetilde{\deg}(\mathrm{SURJ}_{R,\,N})\right) = \Theta(R^{1/4}N^{1/2})$.

Main Idea for approximating $(\mathrm{AND}_R \circ \mathrm{OR}_N)^N$

Main Idea for approximating $(\mathrm{AND}_R \circ \mathrm{OR}_N)^N$





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.

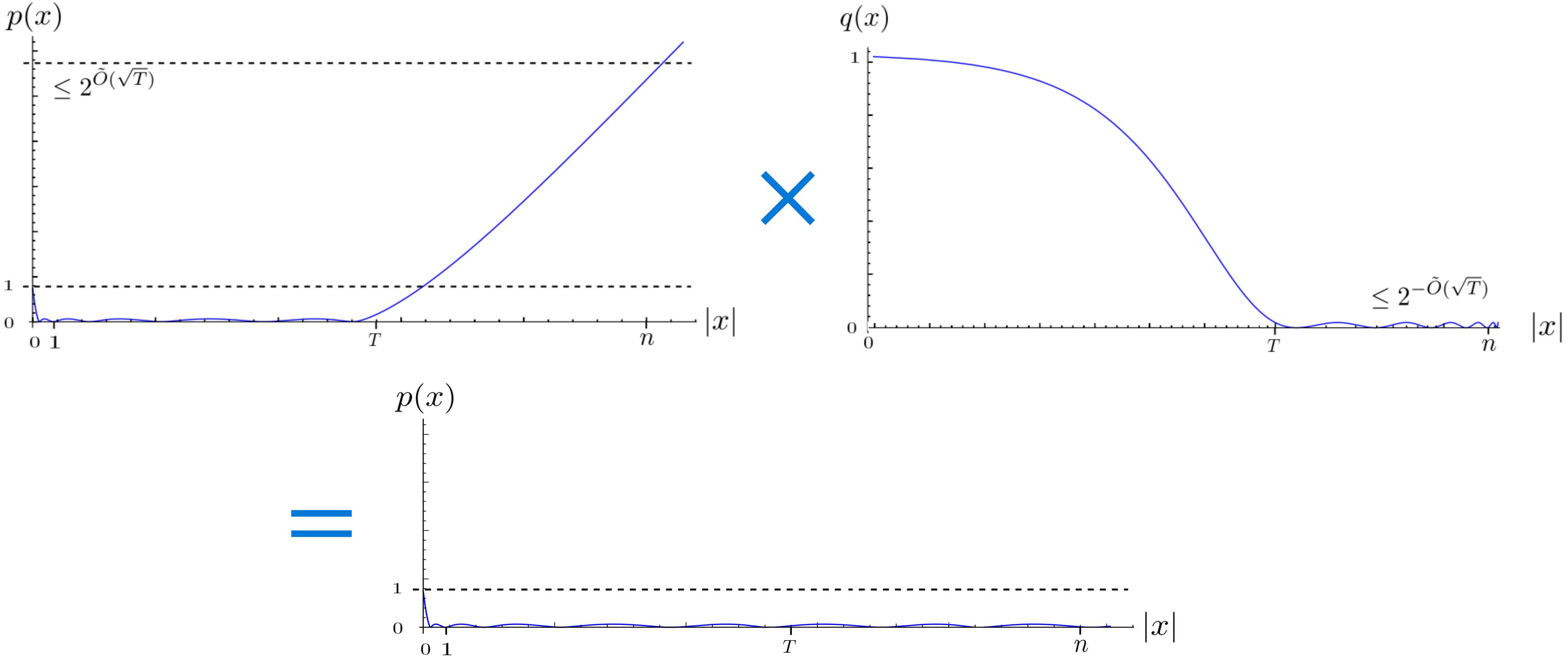
Polynomial: $p_1(x)p_2(x) + (1 - p_1(x))p_3(x)$.

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 must be bounded in [0,1].



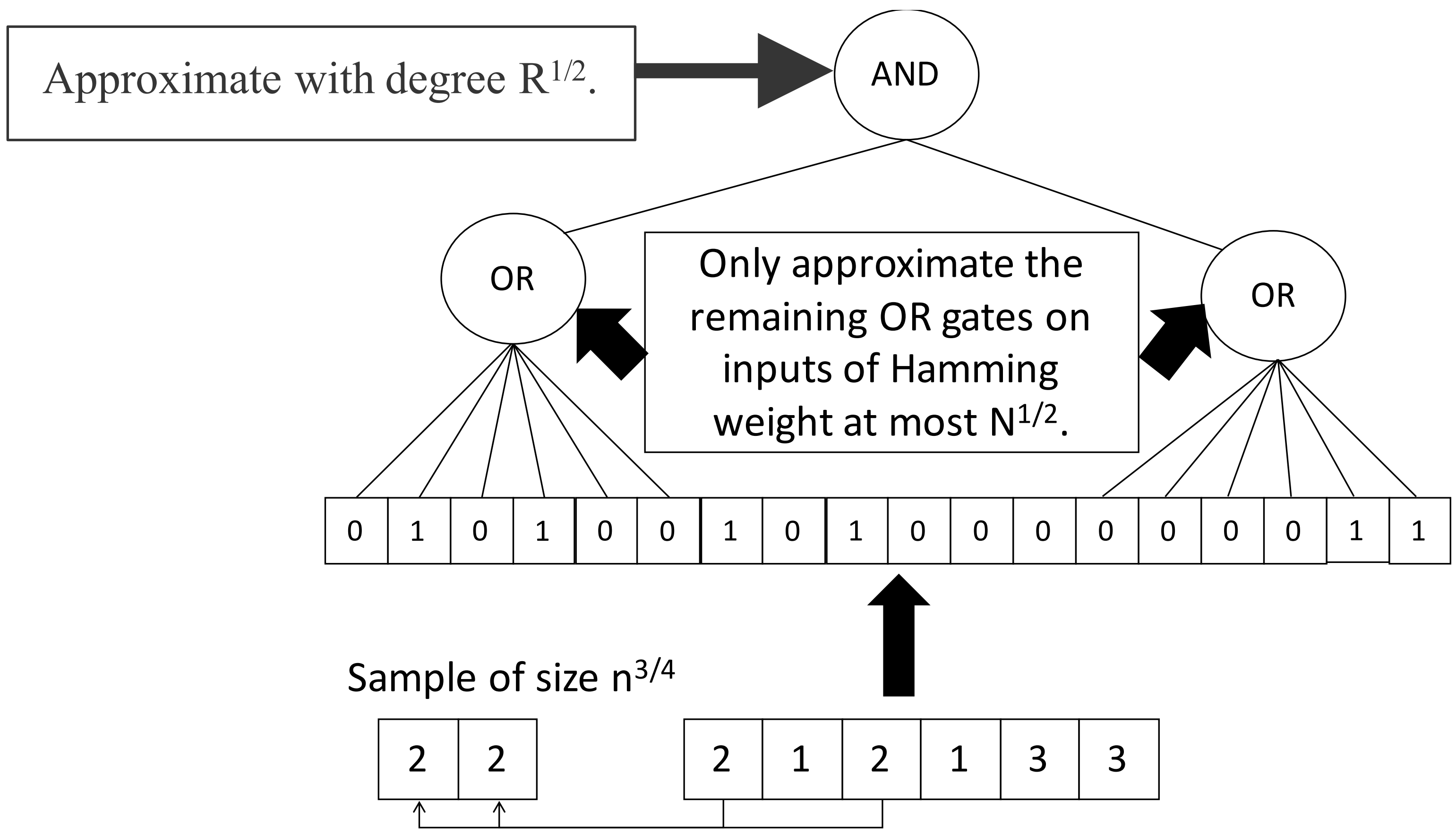
Tight Upper Bound Details

- For simplicity, **fix R=N/2 for duration of talk**. Need to show $\widetilde{\deg}(\mathrm{SURJ}_{R,\,N}) = \widetilde{\Theta}\left(N^{\frac{3}{4}}\right)$.
- We'll approximate SURJ via a "two-stage" construction.
- Think of our construction as a two-stage query algorithm, even though it is not.
- Stage 1: The query algorithm randomly samples $N^{\frac{3}{4}}$ inputs.
- Any range item appearing in the sample definitely appears at least once in the input list, so we can "remove it from consideration".
- Stage 2 just needs to determine whether all range items **not appearing in the sample** appear at least once in the input list.

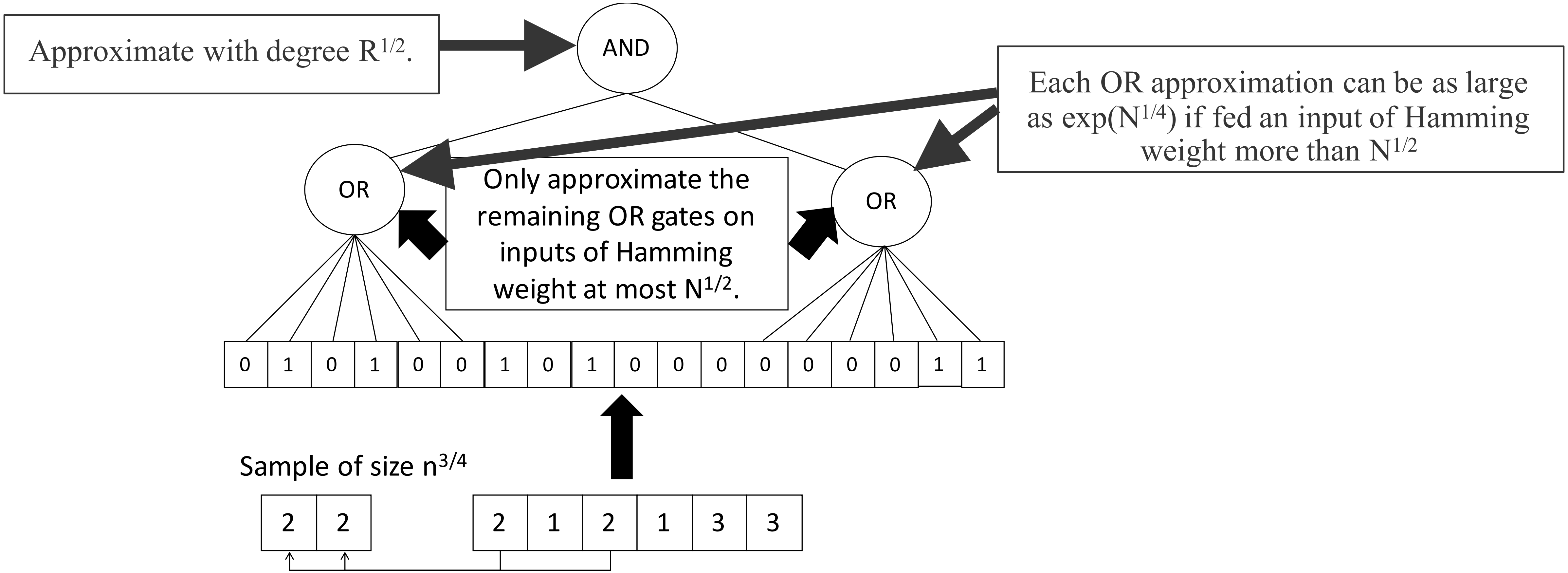
Stage 2

- Key observation: any range item with frequency larger than $T=N^{1/2}$ will appear in the sample at least once with probability at least $1-\exp(-N^{1/4})$.
- i.e., if a range item doesn't appear in the sample, then we are really confident that it does not have high frequency.
- So Stage 2 only needs an approximation p to SURJ that is accurate when no range items have frequency larger than T.
  - When b range items have frequency more than T, p can be as large as $\exp(b \cdot N^{1/4})$.

The Construction in a Picture



The Construction in a Picture



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* $|x| \le n^{1/2}$.
- $|q(x)| \le \exp\left(\tilde{O}(n^{1/4})\right)$ *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:

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

Proof.

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

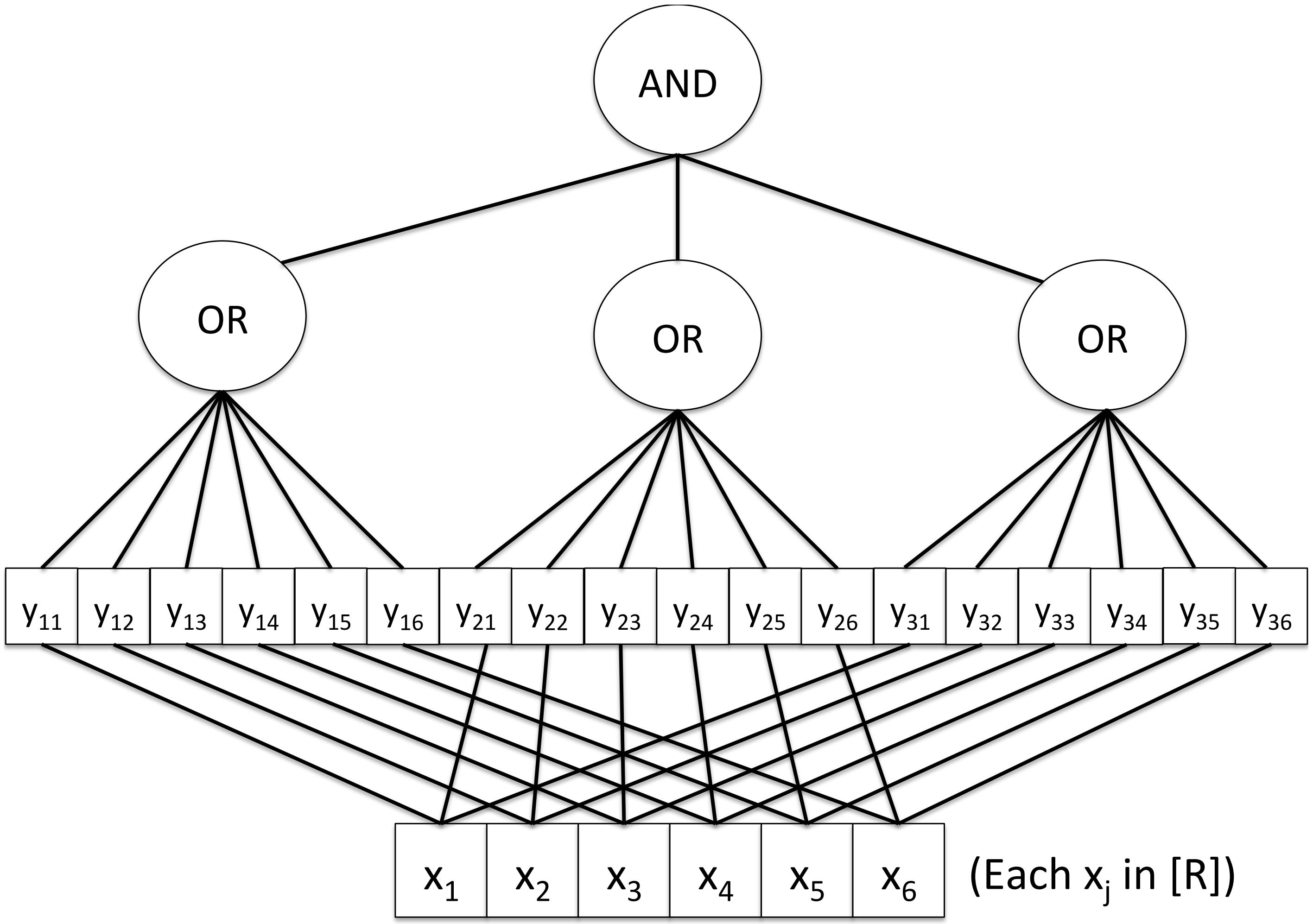
Surjectivity lower bound:

$$\widetilde{\deg}(\mathrm{SURJ}_{R,\,N}) = \Omega(R^{1/4} N^{1/2}).$$

Reduction to a composed function

- 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. [Ambainis05, BunThaler17]
  - 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])

Progress so far towards $\widetilde{\deg}(\mathrm{SURJ}_{R,N}) = \boldsymbol{\Omega}(R^{1/4}N^{1/2})$

1. We saw that $\widetilde{\deg}(\mathrm{SURJ}) = \boldsymbol{\Omega}\left(\widetilde{\deg}\left((\mathrm{AND}_R \circ \mathrm{OR}_N)^{\leq N}\right)\right)$.
2. New goal: show that $\widetilde{\deg}\left((\mathrm{AND}_R \circ \mathrm{OR}_N)^{\leq N}\right) = \boldsymbol{\Omega}\left(R^{\frac{1}{4}} N^{\frac{1}{2}}\right)$.
3. 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 $\leq H$

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

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. $\sum_x \psi(x) f(x) > 1/3$. — (3) ψ is well correlated with f
4. $\psi(x) = 0$ if $|x| > H$ — (4) ψ is only supported on the promise


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

Dual witness for $\widetilde{\deg}\left((\mathrm{AND}_R \circ \mathrm{OR}_n)^{\leq N}\right)$

Dual formulation for problems where we only care about Hamming weight $\leq H$	
$\widetilde{\deg}(f^{\leq H}) > d$ 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. $\sum_x \psi(x) f(x) > 1/3$.	(3) ψ is well correlated with f
4. $\psi(x) = 0$ if $\lvert x\rvert > H$	(4) ψ is only supported on the promise

Fix 1: Use a dual witness ψ_{OR} for OR_N that only certifies $\widetilde{\deg}(\mathrm{OR}_N) = \Omega\left(N^{1/4}\right)$ and satisfies a "dual decay condition", i.e., $\lvert\psi_{\mathrm{OR}}(x)\rvert$ is exponentially small for $\lvert x\rvert \gg N^{1/4}$. Then the composed dual has pure high degree $\Omega\left(\sqrt{R} \cdot N^{1/4}\right) = \Omega(N^{3/4})$ 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-Sherstov08].

Details of Fixes 1 and 2

- Fact (cf. Razborov and Sherstov 2008): Suppose

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

  - Then we can "post-process" $\psi_{\textsf{AND-OR}}$ 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 $\min\{D, \textsf{PHD}(\psi_{\textsf{AND-OR}})\}$.

Details of Fixes 1 and 2

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

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

- Recall:

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

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

- $|\psi_{\text{AND-OR}}(y_1, \ldots, y_R)|$ resembles product distribution: $\prod_{j=1}^{R} |\psi_{\text{OR}}(y_j)|$
- So it is exponentially more biased toward low Hamming weight inputs than ψ_{OR} itself.

Details of Fixes 1 and 2

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

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

- Recall:

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

- A dual witness $\psi_{\textsf{OR}}$ for OR can be made "weakly" biased toward low Hamming weight inputs.
  - Specifically, can ensure:
    - $\mathsf{PHD}(\psi_{\textsf{OR}}) \geq N^{1/4}$.
    - 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 $1/\mathrm{poly}(N)$.

Details of Fixes 1 and 2

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

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

- 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.

- So total mass on these inputs is $\exp(-\Omega(N^{3/4}))$.

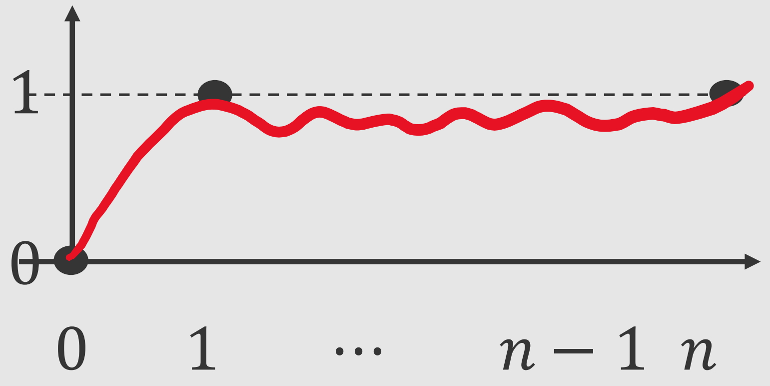
Closing Thoughts

Looking back at the lower bounds

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

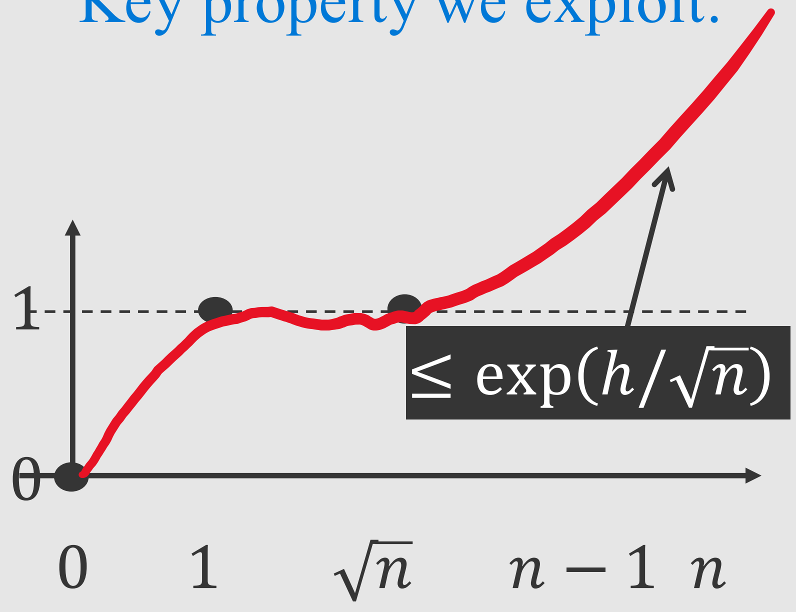
What is the key new ingredient in these lower bounds?

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-$(2d)$ 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?

Thanks!

Approximating distance and entropy

Given n numbers in $[R]$, where $R = \Theta(n)$, interpret them as a probability distribution:

$$p_r = \text{the fraction of times } r \in [R] \text{ appears in the list}$$

Statistical distance from uniform: Compute $\left\| p - \frac{1}{n}\vec{1} \right\|_1$ to additive error ϵ.

Shannon entropy: Compute the Shannon entropy of p to additive error ϵ.

Upper bounds: $\tilde{O}(\sqrt{n})$ for both problems [Bravyi-Harrow-Hassidim09, Li-Wu17]

Lower bounds: $\tilde{\Omega}\left(n^{1/3}\right)$ for both problems [Bravyi-Harrow-Hassidim09, Li-Wu17]

Our result: Optimal lower bound of $\tilde{\Omega}(\sqrt{n})$ for both problems.

Image size testing

Image size testing: Given n numbers in $[R]$, where $R = \Theta(n)$, decide if
(YES) there are at most ℓ distinct range items $r \in [R]$ in the list, or
(NO) the input string is far (at least δn in Hamming distance) from having this property.

Upper bounds

- $Q(\mathrm{IST}) = \tilde{O}(\sqrt{n})$, using the adversary bound dual SDP [Ambainis-Belovs-Regev-deWolf16]

Lower bounds

- $Q(\mathrm{IST}) = \tilde{\Omega}(n^{1/3})$, by a reduction to $\mathrm{Collision}_n$ [Ambainis-Belovs-Regev-deWolf16]

Our result: $Q(\mathrm{IST}) = \tilde{\Omega}(\sqrt{n})$. Lower bound holds for the task of distinguishing between
(YES) every range item $r \in [R]$ appears at least once, or
(NO) at most γn range items appear at least once.

Upper bound: $\widetilde{\deg}(\mathrm{AND}_n) = O(\sqrt{n})$

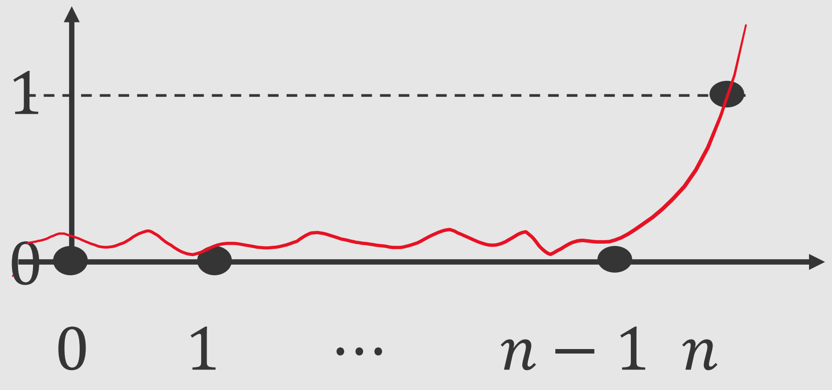
$$\mathrm{AND}_n(x_1, \dots, x_n) = \begin{cases} 0, & \text{if } 0 \le |x| \le n-1 \\ 1, & \text{if } |x| = n \end{cases}, \quad \text{where } |x| \text{ is the Hamming weight of } x.$$

Proof 1. $\widetilde{\deg}(\mathrm{AND}_n) \le 2Q(\mathrm{AND}_n) = O(\sqrt{n})$ by Grover's algorithm.

Proof 2. Say we had a univariate polynomial $q(h) = \sum_{k=0}^{d} \alpha_k h^k$, such that

$q(h) \le 1/3$ for $0 \le h \le n-1$

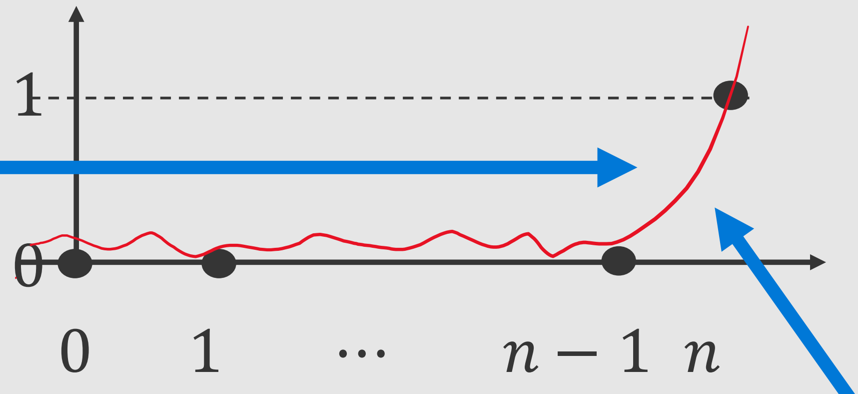
$q(h) \ge 2/3$ for $h = n$



Then the polynomial $p(x_1, \dots, x_n) = q(\sum_i x_i)$ approximates AND_n.

Lower bound: $\widetilde{\deg}(\mathrm{AND}_n) = \Omega(\sqrt{n})$

Summary of upper bound:
Univariate polynomial like this
⇒ polynomial for AND



Symmetrization [Misky-Papert69]: Polynomial for AND ⇒ univariate polynomial like this

Theorem (using Markov's inequality): Any univariate polynomial like this must have degree $\Omega(\sqrt{n})$.

Advantages of the polynomial method

For all symmetric f, $Q(f) = \Theta(\widetilde{\deg}(f))$.	[BBCMdW01]	For most natural functions, $Q(f) = \Theta(\widetilde{\deg}(f))$
For $\epsilon < 1/2$, $Q_\epsilon(\mathrm{XOR}_n) = n/2$.	[BBCMdW01]	Works for unbounded error
For $\epsilon > 1/2^n$, $Q_\epsilon(\mathrm{OR}_n) = \Theta\left(\sqrt{n \log(1/\epsilon)}\right)$.	[BCdWZ99]	Works for small error
For $\epsilon = 0$, $Q_0(\mathrm{OR}_n) = n$.	[BCdWZ99]	Works for zero error
$Q(\mathrm{Collision}_n) = \Theta(n^{1/3})$.	[AS04]	Works when the positive-weights adversary fails

Bonus: Polynomial method lower bounds “lift” to lower bounds in communication complexity! (For more, see the next two talks by Shalev Ben-David and Adam Bouland)

[BCdWZ99] = Buhrman, Cleve, de Wolf, and Zalka (1999)

[AS04] Aaronson and Shi (2004)

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

Dual formulation for problems where we only care about Hamming weight $\leq H$	
4. $\psi(x) = 0$ if $\lvert x \rvert > H$	(4) ψ is only supported on the promise

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].

Dual	Primal
$\sum_{x:\lvert x\rvert>H} \lvert\psi(x)\rvert = 0$	$p(x)$ can be unbounded when $\lvert x\rvert > H$
$\sum_{x:\lvert x\rvert>H} \lvert\psi(x)\rvert = 0.01$	$p(x)$ must be $O(1)$ when $\lvert x\rvert > H$
$\sum_{x:\lvert x\rvert>H} \lvert\psi(x)\rvert$ is small	$p(x)$ can be large, but not too large

Intuition: "Large, but not too large" is sufficient for our bounds, because $p(x)$ is already bounded in $[0,1]$ for Hamming weight $\leq H$. So it cannot grow too large for $\lvert x\rvert > H$.