Approximate Degree: A Survey

Justin Thaler[1]

Georgetown University

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
- $\widetilde{\deg}(f) := \widetilde{\deg}_{1/3}(f)$ is the approximate degree of f

Threshold Degree

Definition

Let $f : \{-1,1\}^n \to \{-1,1\}$ be a Boolean function. A polynomial p sign-represents f if $\operatorname{sgn}(p(x)) = f(x)$ for all $x \in \{-1,1\}^n$.

Definition

The threshold degree of f is $\min \deg(p)$, where the minimum is over all sign-representations of f.

- An equivalent definition of threshold degree is $\lim_{\epsilon \nearrow 1} \widetilde{\deg}_\epsilon(f)$.

Why Care About Approximate and Threshold Degree?

Upper bounds on $\widetilde{\deg}_\epsilon(f)$ and $\deg_\pm(f)$ 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., $\deg_\pm(f)$ upper bounds): PAC learning [KS01]

Why Care About Approximate and Threshold Degree?

Upper bounds on $\widetilde{\deg}_\epsilon(f)$ and $\deg_\pm(f)$ 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., $\deg_\pm(f)$ upper bounds): PAC learning [KS01]

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

Why Care About Approximate and Threshold Degree?

Upper bounds on $\widetilde{\deg}_\epsilon(f)$ and $\deg_\pm(f)$ 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., $\deg_\pm(f)$ upper bounds): PAC learning [KS01]

- Upper bounds on $\widetilde{\deg}_{1/3}(f)$ also imply fast algorithms for **differentially private data release** [TUV12, CTUW14].
- Upper bounds on $\widetilde{\deg}_\epsilon(f)$ and $\deg_\pm(f)$ for small formulas and threshold circuits f yield state of the art **formula size and threshold circuit lower bounds** [Tal17, Forster02].

Why Care About Approximate and Threshold Degree?

Lower bounds on $\widetilde{\deg}_\epsilon(f)$ and $\deg_\pm(f)$ yield lower bounds on:

- **Oracle Separations** [Bei94, BCHTV16]
- **Quantum query complexity** [BBCMW98]
- **Communication complexity** [She08, SZ08, CA08, LS08, She12]
  - Lower bounds hold for a communication problem **related** to f.
  - Via, e.g., a technique called the Pattern Matrix Method [She08].

Why Care About Approximate and Threshold Degree?

Lower bounds on $\widetilde{\deg}_\epsilon(f)$ and $\deg_\pm(f)$ yield lower bounds on:

- **Oracle Separations** [Bei94, BCHTV16]
- **Quantum query complexity** [BBCMW98]
- **Communication complexity** [She08, SZ08, CA08, LS08, She12]
  - Lower bounds hold for a communication problem **related** to f.
  - Via, e.g., a technique called the Pattern Matrix Method [She08].
    - $\epsilon \approx 1/3 \implies \mathbf{BQP}^{\mathrm{cc}}$ lower bounds.
    - $\epsilon \approx 1 - 2^{-n^{\delta}} \implies$: $\mathbf{PP}^{\mathrm{cc}}$ lower bounds
    - $\epsilon \to 1$ (i.e., $\deg_\pm(f)$ lower bounds) $\implies \mathbf{UPP}^{\mathrm{cc}}$ lower bounds.

Why Care About Approximate and Threshold Degree?

Lower bounds on $\widetilde{\deg}_\epsilon(f)$ and $\deg_\pm(f)$ yield lower bounds on:

- **Oracle Separations** [Bei94, BCHTV16]
- **Quantum query complexity** [BBCMW98]
- **Communication complexity** [She08, SZ08, CA08, LS08, She12]
  - Lower bounds hold for a communication problem **related** to f.
  - Via, e.g., a technique called the Pattern Matrix Method [She08].
    - $\epsilon \approx 1/3 \implies \mathbf{BQP}^{\mathsf{cc}}$ lower bounds.
    - $\epsilon \approx 1 - 2^{-n^\delta} \implies$: $\mathbf{PP}^{\mathsf{cc}}$ lower bounds
    - $\epsilon \to 1$ (i.e., $\deg_\pm(f)$ lower bounds) $\implies \mathbf{UPP}^{\mathsf{cc}}$ lower bounds.
- Lower bounds on $\widetilde{\deg}_\epsilon(f)$ and $\deg_\pm(f)$ also yield efficient **secret-sharing schemes** [BIVW16]

Example 1: The Approximate Degree of AND_n

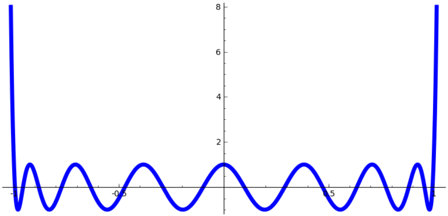
Example: What is the Approximate Degree of AND_n?

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

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

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

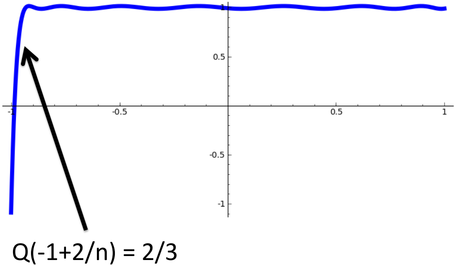
- Chebyshev polynomials are the extremal case.



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:

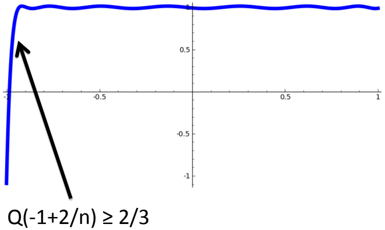


- Define n-variate polynomial p via $p(x) = Q(\sum_{i=1}^{n} x_i/n)$.
- Then $|p(x) - \mathrm{AND}_n(x)| \leq 1/3 \quad \forall x \in \{-1,1\}^n$.

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

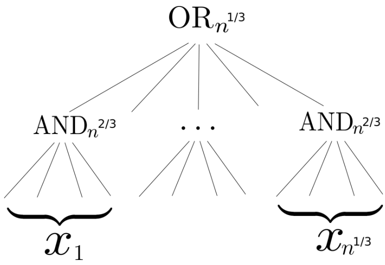
AND Has Low Threshold Degree

- Fact: $\deg_{\pm}(\mathrm{AND}_n) = 1$.
- Proof: $\mathrm{AND}_n(x) = \mathrm{sgn}(p(x))$ for $p(x) = n - 1 + \sum_{i=1}^n x_i$.
- In fact, $p(x)/n$ approximates AND_n to error $1 - 1/n$.

Example 2: The Threshold Degree of the Minsky-Papert DNF

The Minsky-Papert DNF

- The Minsky-Papert DNF is $\mathsf{MP}(x) := \mathrm{OR}_{n^{1/3}} \circ \mathrm{AND}_{n^{2/3}}$.



The Minsky-Papert DNF

- Claim: $\deg_{\pm}(\mathsf{MP}) = \tilde{\Omega}(n^{1/3})$.
- More generally, $\deg_{\pm}(\mathrm{OR}_t \circ \mathrm{AND}_b) \geq \Omega(\min(b^{1/2}, t))$.
- Proved by Minsky and Papert in 1969 via an ad hoc symmetrization argument.

The Minsky-Papert DNF

- Claim: $\deg_{\pm}(\mathsf{MP}) = \tilde{\Omega}(n^{1/3})$.
- More generally, $\deg_{\pm}(\mathrm{OR}_t \circ \mathrm{AND}_b) \geq \Omega(\min(b^{1/2}, t))$.
- Proved by Minsky and Papert in 1969 via an ad hoc symmetrization argument.
- (Klivans-Servedio 2004): **All** polysize DNFs have threshold degree $\tilde{O}(n^{1/3})$.
  - Yields fastest known algorithm for PAC learning DNFs.

The Minsky-Papert DNF

- Claim: $\deg_{\pm}(\mathsf{MP}) = \tilde{\Omega}(n^{1/3})$.
- More generally, $\deg_{\pm}(\mathrm{OR}_t \circ \mathrm{AND}_b) \geq \Omega(\min(b^{1/2}, t))$.
- Proved by Minsky and Papert in 1969 via an ad hoc symmetrization argument.
- (Klivans-Servedio 2004): **All** polysize DNFs have threshold degree $\tilde{O}(n^{1/3})$.
  - Yields fastest known algorithm for PAC learning DNFs.
- We will prove the matching upper bound:

$$\deg_{\pm}(\mathrm{OR}_t \circ \mathrm{AND}_b) \leq \tilde{O}(\min(b^{1/2}, t)).$$

  - First, we'll construct a sign-representation of degree $\tilde{O}(b^{1/2})$ using Chebyshev approximations to AND_b.
  - Then we'll construct a sign-representation of degree $\tilde{O}(t)$ using rational approximations to AND_b.

A Sign-Representation for $\mathrm{OR}_t \circ \mathrm{AND}_b$ of degree $\tilde{O}(b^{1/2})$

- Let p_1 be a (Chebyshev-derived) polynomial of degree $O\left(\sqrt{b \cdot \log t}\right)$ approximating AND_b to error $\frac{1}{8t}$.
- Let $p = \frac{1}{2} \cdot (1 - p_1)$.
- $p(x_i)$ is "close to 0" if $\mathrm{AND}_b(x_i)$ is FALSE, and "close to 1" otherwise.

A Sign-Representation for $\mathrm{OR}_t \circ \mathrm{AND}_b$ of degree $\tilde{O}(b^{1/2})$

- Let p_1 be a (Chebyshev-derived) polynomial of degree $O\left(\sqrt{b \cdot \log t}\right)$ approximating AND_b to error $\frac{1}{8t}$.
- Let $p = \frac{1}{2} \cdot (1 - p_1)$.
- $p(x_i)$ is "close to 0" if $\mathrm{AND}_b(x_i)$ is FALSE, and "close to 1" otherwise.
- Then $\frac{1}{2} - \sum_{i=1}^{t} p(x_i)$ sign-represents $\mathrm{OR}_t \circ \mathrm{AND}_b$.

A Sign-Representation for $\mathrm{OR}_t \circ \mathrm{AND}_b$ of degree $\tilde{O}(t)$

- Fact: there exist p_1, q_1 of degree $O(\log b \cdot \log t)$ such that
$$\left| \mathrm{AND}_b(x) - \frac{p_1(x)}{q_1(x)} \right| \leq \frac{1}{8t} \text{ for all } x \in \{-1, 1\}^b.$$
- Let $\frac{p(x)}{q(x)} = \frac{1}{2} \cdot \left(1 - \frac{p_1(x)}{q_1(x)}\right)$.

A Sign-Representation for $\mathrm{OR}_t \circ \mathrm{AND}_b$ of degree $\tilde{O}(t)$

- Fact: there exist p_1, q_1 of degree $O(\log b \cdot \log t)$ such that

$$\left| \mathrm{AND}_b(x) - \frac{p_1(x)}{q_1(x)} \right| \leq \frac{1}{8t} \text{ for all } x \in \{-1,1\}^b.$$

- Let $\frac{p(x)}{q(x)} = \frac{1}{2} \cdot \left(1 - \frac{p_1(x)}{q_1(x)}\right)$.
- Then $\mathrm{sgn}(\mathrm{OR}_t \circ \mathrm{AND}_b(x)) = \frac{1}{2} - \sum_{i=1}^{t} \frac{p(x_i)}{q(x_i)}$

A Sign-Representation for $\mathrm{OR}_t \circ \mathrm{AND}_b$ of degree $\tilde{O}(t)$

- Fact: there exist p_1, q_1 of degree $O(\log b \cdot \log t)$ such that

$$\left| \mathrm{AND}_b(x) - \frac{p_1(x)}{q_1(x)} \right| \leq \frac{1}{8t} \text{ for all } x \in \{-1,1\}^b.$$

- Let $\frac{p(x)}{q(x)} = \frac{1}{2} \cdot \left(1 - \frac{p_1(x)}{q_1(x)}\right)$.
- Then $\mathrm{sgn}(\mathrm{OR}_t \circ \mathrm{AND}_b(x)) = \frac{1}{2} - \sum_{i=1}^{t} \frac{p(x_i)}{q(x_i)}$

$$= \frac{1}{2} - \sum_{i=1}^{t} \frac{p(x_i) \cdot q(x_i)}{q^2(x_i)}.$$

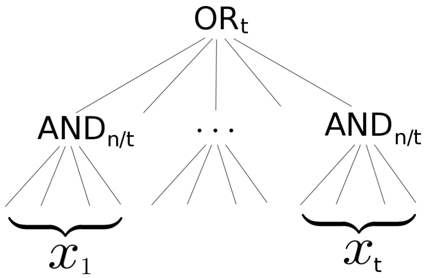
A Sign-Representation for $\mathrm{OR}_t \circ \mathrm{AND}_b$ of degree $\tilde{O}(t)$

- Fact: there exist p_1, q_1 of degree $O(\log b \cdot \log t)$ such that
$$\left| \mathrm{AND}_b(x) - \frac{p_1(x)}{q_1(x)} \right| \leq \frac{1}{8t} \text{ for all } x \in \{-1,1\}^b.$$
- Let $\frac{p(x)}{q(x)} = \frac{1}{2} \cdot \left(1 - \frac{p_1(x)}{q_1(x)}\right)$.
- Then $\mathrm{sgn}(\mathrm{OR}_t \circ \mathrm{AND}_b(x)) = \frac{1}{2} - \sum_{i=1}^{t} \frac{p(x_i)}{q(x_i)}$
$$= \frac{1}{2} - \sum_{i=1}^{t} \frac{p(x_i) \cdot q(x_i)}{q^2(x_i)}.$$
- Put the sum over common denominator $\prod_{i=1}^{t} q^2(x_i)$ to obtain:
$$\mathrm{sgn}(\mathrm{OR}_t \circ \mathrm{AND}_b(x)) = r(x) / \prod_{i=1}^{t} q^2(x_i)$$
for $r(x) := \left(\frac{1}{2} \cdot \prod_{1 \leq i \leq t} q^2(x_i)\right) - \sum_{i=1}^{t} \left(p(x_i) \cdot q(x_i) \cdot \prod_{1 \leq i \leq t, i' \neq i} q^2(x_{i'})\right).$

Recent Progress on Lower Bounds: Beyond Symmetrization

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 $\widetilde{\deg}(\text{OR}_t \circ \text{AND}_{n/t})$?



History of the OR-AND Tree

Theorem

$\widetilde{\deg}(\mathsf{OR}_t \circ \mathsf{AND}_{n/t}) = \Theta(n^{1/2})$.

History of the OR-AND Tree

Theorem

$\widetilde{\deg}(\mathsf{OR}_t \circ \mathsf{AND}_{n/t}) = \Theta(n^{1/2})$.

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

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

History of the OR-AND Tree

Theorem

$\widetilde{\deg}(\mathsf{OR}_t \circ \mathsf{AND}_{n/t}) = \Theta(n^{1/2})$.

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

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

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):

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

Dual LP:

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

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

Example: $2^{-n} \cdot \mathsf{PARITY}_n$ witnesses the fact that $\lim_{\epsilon \nearrow 1} \widetilde{\deg}_\epsilon(\mathsf{PARITY}_n) = n$.

Goal: Construct an explicit dual polynomial $\psi_{\textbf{OR-AND}}$ for $\text{OR}_t \circ \text{AND}_{n/t}$

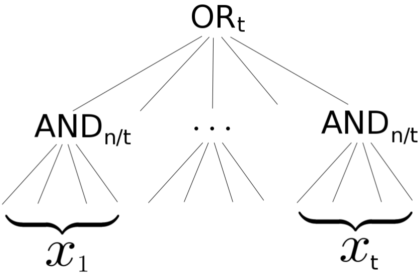
Constructing a Dual Polynomial

- By [NS92], there are dual polynomials
  $\psi_{\textsf{OUT}}$ for $\widetilde{\deg}\left(\mathrm{OR}_t\right) = \Omega(t^{1/2})$ and
  $\psi_{\textsf{IN}}$ for $\widetilde{\deg}\left(\mathrm{AND}_{n/t}\right) = \Omega\left((n/t)^{1/2}\right)$
- Both [She13] and [BT13] combine $\psi_{\textsf{OUT}}$ and $\psi_{\textsf{IN}}$ to obtain a dual polynomial $\psi_{\textsf{OR-AND}}$ for $\mathrm{OR}_t \circ \mathrm{AND}_{n/t}$.
- The combining method was proposed in earlier works [SZ09, She09, Lee09]. We call it **dual block composition**.

Dual Block Composition [SZ09, She09, Lee09]

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

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



Dual Block Composition [SZ09, She09, Lee09]

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

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

Must verify:

1. $\psi_{\textbf{OR-AND}}$ has pure high degree $\geq t^{1/2} \cdot (n/t)^{1/2} = n^{1/2}$.
2. $\psi_{\textbf{OR-AND}}$ has high correlation with $\mathsf{OR}_t \circ \mathsf{AND}_{n/t}$.

Dual Block Composition [SZ09, She09, Lee09]

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

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

Must verify:

1. $\psi_{\textbf{OR-AND}}$ has pure high degree $\geq t^{1/2} \cdot (n/t)^{1/2} = n^{1/2}$. ✓[She09]
2. $\psi_{\textbf{OR-AND}}$ has high correlation with $\text{OR}_t \circ \text{AND}_{n/t}$. [BT13, She13]

Proving Hardness Amplification Theorems Via Dual Block Composition

Proving Hardness Amplification Theorems Via Dual Block Composition

These theorems show that $g \circ f$ is "harder to approximate" by low-degree polynomials than is f alone.

(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.

(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

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

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

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

Hardness-Amplification Theorems: Part 1

Theorem (BT13, She13)

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

Hardness-Amplification Theorems: Part 1

Theorem (BT13, She13)

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

Hardness-Amplification Theorems: Part 1

Theorem (BT13, She13)

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

Recent Theorems: Part 2

- For some applications in complexity theory, one needs an even simpler "hardness-amplifying function" than OR_t.

Recent Theorems: Part 2

- For some applications in complexity theory, one needs an even simpler "hardness-amplifying function" than OR_t.
- Define $\mathsf{GAPMAJ}_t \colon \{-1,1\}^t \to \{-1,1\}$ to be the partial function that equals:
  - -1 if at least $2/3$ of its inputs are -1
  - $+1$ if at least $2/3$ of its inputs are $+1$
  - undefined otherwise.

Theorem (BCHTV16)

Let f be a Boolean function with $\widetilde{\deg}_{1/2}(f) \geq d$. Let $F = \mathsf{GAPMAJ}_t \circ f$. Then $\widetilde{\deg}_{1-2^{-\Omega(t)}}(F) \geq d$ and $\deg_{\pm}(F) \geq \Omega(\min\{d, t\})$.

Proving the Theorem

Theorem (BCHTV16, BT14, BIVW16)

Let f *be a Boolean function with* $\widetilde{\deg}_{1/2}(f) \geq d$. *Let* $F = \text{GAPMAJ}_t \circ f$. *Then* $\deg_{1-2^{-\Omega(t)}}(F) \geq d$.

Proving the Theorem

Theorem (BCHTV16, BT14, BIVW16)

Let f be a Boolean function with $\widetilde{\deg}_{1/2}(f) \geq d$. Let $F = \text{GAPMAJ}_t \circ f$. Then $\deg_{1-2^{-\Omega(t)}}(F) \geq d$.

- Let ψ_{IN} be any dual witness to the fact that $\widetilde{\deg}_{1/2}(f) \geq d$.
- Define $\psi_{\text{OUT}} : \{-1,1\}^t \to \mathbb{R}$ via:

$$\psi_{\text{OUT}}(y) = \begin{cases} 1/2 & \text{if } y = \textbf{ ALL-FALSE} \\ -1/2 & \text{if } y = \textbf{ ALL-TRUE} \\ 0 & \text{otherwise} \end{cases}$$

- Combine ψ_{OUT} and ψ_{IN} via **dual block composition** to obtain a dual witness ψ_F for F.

Proving the Theorem

Theorem (BCHTV16, BT14, BIVW16)

Let f be a Boolean function with $\widetilde{\deg}_{1/2}(f) \geq d$. Let $F = \text{GAPMAJ}_t \circ f$. Then $\deg_{1-2^{-\Omega(t)}}(F) \geq d$.

- Let $\psi_{\textbf{IN}}$ be any dual witness to the fact that $\widetilde{\deg}_{1/2}(f) \geq d$.
- Define $\psi_{\textbf{OUT}} : \{-1, 1\}^t \to \mathbb{R}$ via:

$$\psi_{\textbf{OUT}}(y) = \begin{cases} 1/2 & \text{if } y = \textbf{ ALL-FALSE} \\ -1/2 & \text{if } y = \textbf{ ALL-TRUE} \\ 0 & \text{otherwise} \end{cases}$$

- Combine $\psi_{\textbf{OUT}}$ and $\psi_{\textbf{IN}}$ via **dual block composition** to obtain a dual witness ψ_F for F.

Must verify:

1. ψ_F has pure high degree d.
2. ψ_F has correlation at least $1 - 2^{-\Omega(t)}$ with F.

Proving the Theorem: Pure High Degree

- Notice $\psi_{\textbf{OUT}}$ is balanced (i.e., it has pure high degree 1).
- So previous analysis shows ψ_F has pure high degree at least $1 \cdot d = d$.

Proving the Theorem: Correlation Analysis

Recall: $F = \mathsf{GAPMAJ}_t \circ f$

$$\psi_F(x_1, \ldots, x_t) := C \cdot \psi_{\mathsf{OUT}}(\ldots, \mathrm{sgn}(\psi_{\mathsf{IN}}(x_i)), \ldots) \prod_{i=1}^{t} |\psi_{\mathsf{IN}}(x_i)|$$

Proving the Theorem: Correlation Analysis

Recall: $F = \mathsf{GAPMAJ}_t \circ f$

$$\psi_F(x_1, \ldots, x_t) := C \cdot \psi_{\mathbf{OUT}}(\ldots, \mathrm{sgn}(\psi_{\mathbf{IN}}(x_i)), \ldots) \prod_{i=1}^{t} |\psi_{\mathbf{IN}}(x_i)|$$

- Goal: Show $\sum_{x \in \{-1,1\}^n} \psi_F(x) \cdot F(x) \geq 1 - 2^{-\Omega(t)}$.

Proving the Theorem: Correlation Analysis

Recall: $F = \mathsf{GAPMAJ}_t \circ f$

$$\psi_F(x_1, \ldots, x_t) := C \cdot \psi_{\mathsf{OUT}}(\ldots, \mathrm{sgn}(\psi_{\mathsf{IN}}(x_i)), \ldots) \prod_{i=1}^{t} |\psi_{\mathsf{IN}}(x_i)|$$

- Goal: Show $\sum_{x \in \{-1,1\}^n} \psi_F(x) \cdot F(x) \geq 1 - 2^{-\Omega(t)}$.
- Idea: Show

$$\sum_{x \in \{-1,1\}^n} \psi_F(x) \cdot F(x) = \sum_{y \in \{-1,1\}^t} \psi_{\mathsf{OUT}}(y) \cdot \mathsf{GAPMAJ}_t(y) - 2^{-\Omega(t)} = 1 - 2^{-\Omega(t)}.$$

Proving the Theorem: Correlation Analysis

Recall: $F = \mathsf{GAPMAJ}_t \circ f$

$$\psi_F(x_1, \ldots, x_t) := C \cdot \psi_{\mathbf{OUT}}(\ldots, \operatorname{sgn}(\psi_{\mathbf{IN}}(x_i)), \ldots) \prod_{i=1}^{t} |\psi_{\mathbf{IN}}(x_i)|$$

- Goal: Show $\sum_{x \in \{-1,1\}^n} \psi_F(x) \cdot F(x) \geq 1 - 2^{-\Omega(t)}$.
- Idea: Show

$$\sum_{x \in \{-1,1\}^n} \psi_F(x) \cdot F(x) = \sum_{y \in \{-1,1\}^t} \psi_{\mathbf{OUT}}(y) \cdot \mathsf{GAPMAJ}_t(y) - 2^{-\Omega(t)} = 1 - 2^{-\Omega(t)}.$$

- Consider $y = (\operatorname{sgn}(\psi_{\mathbf{IN}}(x_1)), \ldots, \operatorname{sgn}(\psi_{\mathbf{IN}}(x_t))) =$ **ALL-TRUE**.

Proving the Theorem: Correlation Analysis

Recall: $F = \mathsf{GAPMAJ}_t \circ f$

$$\psi_F(x_1, \ldots, x_t) := C \cdot \psi_{\mathbf{OUT}}(\ldots, \mathrm{sgn}(\psi_{\mathbf{IN}}(x_i)), \ldots) \prod_{i=1}^{t} |\psi_{\mathbf{IN}}(x_i)|$$

- Goal: Show $\sum_{x \in \{-1,1\}^n} \psi_F(x) \cdot F(x) \geq 1 - 2^{-\Omega(t)}$.
- Idea: Show

$$\sum_{x \in \{-1,1\}^n} \psi_F(x) \cdot F(x) = \sum_{y \in \{-1,1\}^t} \psi_{\mathbf{OUT}}(y) \cdot \mathsf{GAPMAJ}_t(y) - 2^{-\Omega(t)} = 1 - 2^{-\Omega(t)}.$$

- Consider $y = (\mathrm{sgn}(\psi_{\mathbf{IN}}(x_1)), \ldots, \mathrm{sgn}(\psi_{\mathbf{IN}}(x_t))) =$ **ALL-TRUE**.
- If a $\leq 1/3$ fraction of the coordinates y_i of y are "in error", then $F(x) = \psi_{\mathbf{OUT}}(y) = -1$. ☺

Proving the Theorem: Correlation Analysis

Recall: $F = \mathsf{GAPMAJ}_t \circ f$

$$\psi_F(x_1, \ldots, x_t) := C \cdot \psi_{\mathbf{OUT}}(\ldots, \mathrm{sgn}(\psi_{\mathbf{IN}}(x_i)), \ldots) \prod_{i=1}^{t} |\psi_{\mathbf{IN}}(x_i)|$$

- Goal: Show $\sum_{x \in \{-1,1\}^n} \psi_F(x) \cdot F(x) \geq 1 - 2^{-\Omega(t)}$.
- Idea: Show

$$\sum_{x \in \{-1,1\}^n} \psi_F(x) \cdot F(x) = \sum_{y \in \{-1,1\}^t} \psi_{\mathbf{OUT}}(y) \cdot \mathsf{GAPMAJ}_t(y) - 2^{-\Omega(t)} = 1 - 2^{-\Omega(t)}.$$

- Consider $y = (\mathrm{sgn}(\psi_{\mathbf{IN}}(x_1)), \ldots, \mathrm{sgn}(\psi_{\mathbf{IN}}(x_t))) = \mathbf{ALL\text{-}TRUE}$.
- If a $\leq 1/3$ fraction of the coordinates y_i of y are "in error", then $F(x) = \psi_{\mathbf{OUT}}(y) = -1$. ☺
- Any coordinate y_i is "in error" with probability $\leq 1/4$ under distribution $|\psi_{\mathbf{IN}}(x_i)|$, since $\psi_{\mathbf{IN}}$ is well-correlated with f.

Proving the Theorem: Correlation Analysis

Recall: $F = \text{GAPMAJ}_t \circ f$

$$\psi_F(x_1, \ldots, x_t) := C \cdot \psi_{\textbf{OUT}}(\ldots, \text{sgn}(\psi_{\textbf{IN}}(x_i)), \ldots) \prod_{i=1}^{t} |\psi_{\textbf{IN}}(x_i)|$$

- Goal: Show $\sum_{x \in \{-1,1\}^n} \psi_F(x) \cdot F(x) \geq 1 - 2^{-\Omega(t)}$.
- Idea: Show

$$\sum_{x \in \{-1,1\}^n} \psi_F(x) \cdot F(x) = \sum_{y \in \{-1,1\}^t} \psi_{\textbf{OUT}}(y) \cdot \text{GAPMAJ}_t(y) - 2^{-\Omega(t)} = 1 - 2^{-\Omega(t)}.$$

- Consider $y = (\text{sgn}(\psi_{\textbf{IN}}(x_1)), \ldots, \text{sgn}(\psi_{\textbf{IN}}(x_t))) =$ **ALL-TRUE**.
- If a $\leq 1/3$ fraction of the coordinates y_i of y are "in error", then $F(x) = \psi_{\textbf{OUT}}(y) = -1$. ☺
- Any coordinate y_i is "in error" with probability $\leq 1/4$ under distribution $|\psi_{\textbf{IN}}(x_i)|$, since $\psi_{\textbf{IN}}$ is well-correlated with f.
- Under product distribution $\prod_{i=1}^{t} |\psi_{\textbf{IN}}(x_i)|$, a $\geq 1/3$ fraction of the coordinates of y are in error with probability only $2^{-\Omega(t)}$.

Proving the Theorem: Correlation Analysis

Recall: $F = \mathsf{GAPMAJ}_t \circ f$

$$\psi_F(x_1, \ldots, x_t) := C \cdot \psi_{\mathsf{OUT}}(\ldots, \operatorname{sgn}(\psi_{\mathsf{IN}}(x_i)), \ldots) \prod_{i=1}^{t} |\psi_{\mathsf{IN}}(x_i)|$$

- Goal: Show $\sum_{x \in \{-1,1\}^n} \psi_F(x) \cdot F(x) \geq 1 - 2^{-\Omega(t)}$.
- Idea: Show

$$\sum_{x \in \{-1,1\}^n} \psi_F(x) \cdot F(x) = \sum_{y \in \{-1,1\}^t} \psi_{\mathsf{OUT}}(y) \cdot \mathsf{GAPMAJ}_t(y) - 2^{-\Omega(t)} = 1 - 2^{-\Omega(t)}.$$

- Consider $y = (\operatorname{sgn}(\psi_{\mathsf{IN}}(x_1)), \ldots, \operatorname{sgn}(\psi_{\mathsf{IN}}(x_t))) =$ **ALL-TRUE**.
- If a $\leq 1/3$ fraction of the coordinates y_i of y are "in error", then $F(x) = \psi_{\mathsf{OUT}}(y) = -1$. ☺
- Any coordinate y_i is "in error" with probability $\leq 1/4$ under distribution $|\psi_{\mathsf{IN}}(x_i)|$, since ψ_{IN} is well-correlated with f.
- Under product distribution $\prod_{i=1}^{t} |\psi_{\mathsf{IN}}(x_i)|$, a $\geq 1/3$ fraction of the coordinates of y are in error with probability only $2^{-\Omega(t)}$.
- Identical analysis applies for $y =$ **ALL-FALSE**.

Applying the Theorem: Oracle Separations for Statistical Zero Knowledge

Some Delicious Alphabet Soup

- **PP** is the class of all languages solvable by polynomial time randomized algorithms that output the correct answer with probability strictly better than $1/2$.
- **SZK** is the class of all languages with efficient interactive proofs, in which convincing proofs reveal no information other than their own validity.

Some Delicious Alphabet Soup

- $\mathbf{PP}$ is the class of all languages solvable by polynomial time randomized algorithms that output the correct answer with probability strictly better than $1/2$.
- $\mathbf{SZK}$ is the class of all languages with efficient interactive proofs, in which convincing proofs reveal no information other than their own validity.
- Open Problem (Watrous, 2002): Give an oracle $\mathcal{O}$ such that $\mathbf{PP}^{\mathcal{O}} \not\subset \mathbf{SZK}^{\mathcal{O}}$.

Some Delicious Alphabet Soup

- $\mathbf{PP}$ is the class of all languages solvable by polynomial time randomized algorithms that output the correct answer with probability strictly better than $1/2$.
- $\mathbf{SZK}$ is the class of all languages with efficient interactive proofs, in which convincing proofs reveal no information other than their own validity.
- Open Problem (Watrous, 2002): Give an oracle $\mathcal{O}$ such that $\mathbf{PP}^{\mathcal{O}} \not\subset \mathbf{SZK}^{\mathcal{O}}$.
- Remainder of the talk: Solving this problem using the Theorem just proved.
- This is the strongest relativized evidence that $\mathbf{SZK}$ contains intractable problems.

Some Delicious Alphabet Soup

- $\mathbf{PP}$ is the class of all languages solvable by polynomial time randomized algorithms that output the correct answer with probability strictly better than $1/2$.
- $\mathbf{SZK}$ is the class of all languages with efficient interactive proofs, in which convincing proofs reveal no information other than their own validity.
- Open Problem (Watrous, 2002): Give an oracle $\mathcal{O}$ such that $\mathbf{PP}^{\mathcal{O}} \not\subset \mathbf{SZK}^{\mathcal{O}}$.
- Remainder of the talk: Solving this problem using the Theorem just proved.
- This is the strongest relativized evidence that $\mathbf{SZK}$ contains intractable problems.
- Other consequences of the Theorem: $\mathbf{SZK}^{\mathcal{O}} \not\subset \mathbf{PZK}^{\mathcal{O}}$, $\mathbf{NISZK}^{\mathcal{O}} \not\subset \mathbf{NIPZK}^{\mathcal{O}}$, $\mathbf{PZK}^{\mathcal{O}} \not\subset \mathbf{coPZK}^{\mathcal{O}}$, and more.

Query (Decision Tree) Complexity

- Let $f\colon \{-1,1\}^n \to \{-1,1\}$ be a function and $x \in \{-1,1\}^n$ be an input to f.
- Goal: Compute $f(x)$ by reading as few bits of x as possible.

Query (Decision Tree) Complexity

- Let $f\colon \{-1,1\}^n \to \{-1,1\}$ be a function and $x \in \{-1,1\}^n$ be an input to f.
- Goal: Compute $f(x)$ by reading as few bits of x as possible.
- The $\mathbf{PP}^{\textsf{dt}}$ cost of f is the least d such that there is some randomized algorithm making at most d queries that outputs $f(x)$ with probability at least $1/2 + 2^{-d}$.

Query (Decision Tree) Complexity

- Let $f\colon \{-1,1\}^n \to \{-1,1\}$ be a function and $x \in \{-1,1\}^n$ be an input to f.
- Goal: Compute $f(x)$ by reading as few bits of x as possible.
- The $\mathbf{PP}^{\mathbf{dt}}$ cost of f is the least d such that there is some randomized algorithm making at most d queries that outputs $f(x)$ with probability at least $1/2 + 2^{-d}$.
- The $\mathbf{SZK}^{\mathbf{dt}}$ cost of f is the least d such that there is an interactive proof for the claim that $f(x) = -1$, where:
  - The total communication between prover and verifier is $\leq d$.
  - The verifier makes $\leq d$ queries to x.
  - A convincing proof reveals nothing to the verifier (other than $f(x) = -1$) that the verifier could not have learned by making d queries to f without ever talking to the prover.

Query (Decision Tree) Complexity

- Let $f\colon \{-1,1\}^n \to \{-1,1\}$ be a function and $x \in \{-1,1\}^n$ be an input to f.
- Goal: Compute $f(x)$ by reading as few bits of x as possible.
- The $\mathbf{PP}^{\textsf{dt}}$ cost of f is the least d such that there is some randomized algorithm making at most d queries that outputs $f(x)$ with probability at least $1/2 + 2^{-d}$.
- The $\mathbf{SZK}^{\textsf{dt}}$ cost of f is the least d such that there is an interactive proof for the claim that $f(x) = -1$, where:
  - The total communication between prover and verifier is $\leq d$.
  - The verifier makes $\leq d$ queries to x.
  - A convincing proof reveals nothing to the verifier (other than $f(x) = -1$) that the verifier could not have learned by making d queries to f without ever talking to the prover.
- Fact: To give an oracle $\mathcal{O}$ s.t. $\mathbf{SZK}^{\mathcal{O}} \not\subset \mathbf{PP}^{\mathcal{O}}$, it's enough to give an f s.t. $\mathbf{SZK}^{\textsf{dt}}(f) = O(\log n)$ and $\mathbf{PP}^{\textsf{dt}}(f) = n^{\Omega(1)}$.

Connecting $\mathrm{PP}^{\mathsf{dt}}$ and Approximate Degree

Connecting $\mathbf{PP}^{\textsf{dt}}$ and Approximate Degree

- Fact: $\mathbf{PP}^{\textsf{dt}}(f) \leq d \iff \widetilde{\deg}_\epsilon(f) \leq d$ for $\epsilon = 1 - 2^{-d}$.

Connecting $\mathbf{PP}^{\mathsf{dt}}$ and Approximate Degree

- Fact: $\mathbf{PP}^{\mathsf{dt}}(f) \leq d \iff \widetilde{\deg}_\epsilon(f) \leq d$ for $\epsilon = 1 - 2^{-d}$.
- Idea for $\Longrightarrow$: For any randomized algorithm $\mathcal{A}$ making at most T queries to x, there is a degree T polynomial p such that $p(x) = \Pr[\mathcal{A}(x) = -1]$.

Connecting $\mathrm{PP}^{\textsf{dt}}$ and Approximate Degree

- Fact: $\mathbf{PP}^{\textsf{dt}}(f) \leq d \iff \widetilde{\deg}_\epsilon(f) \leq d$ for $\epsilon = 1 - 2^{-d}$.
- Idea for $\implies$: For any randomized algorithm $\mathcal{A}$ making at most T queries to x, there is a degree T polynomial p such that $p(x) = \Pr[\mathcal{A}(x) = -1]$.
  - If $\mathbf{PP}^{\textsf{dt}}(f) = d$, then there is a d-query algorithm $\mathcal{A}$ such that

$$\begin{cases} f(x) = 1 \implies \Pr[\mathcal{A}(x) = 1] \geq 1/2 + 2^{-d} \\ f(x) = -1 \implies Pr[\mathcal{A}(x) = 1] \leq 1/2 - 2^{-d}. \end{cases}$$

Connecting $\mathbf{PP}^{\mathsf{dt}}$ and Approximate Degree

- Fact: $\mathbf{PP}^{\mathsf{dt}}(f) \leq d \iff \widetilde{\deg}_\epsilon(f) \leq d$ for $\epsilon = 1 - 2^{-d}$.
- Idea for $\implies$: For any randomized algorithm $\mathcal{A}$ making at most T queries to x, there is a degree T polynomial p such that $p(x) = \Pr[\mathcal{A}(x) = -1]$.
  - If $\mathbf{PP}^{\mathsf{dt}}(f) = d$, then there is a d-query algorithm $\mathcal{A}$ such that

$$\begin{cases} f(x) = 1 \implies \Pr[\mathcal{A}(x) = 1] \geq 1/2 + 2^{-d} \\ f(x) = -1 \implies Pr[\mathcal{A}(x) = 1] \leq 1/2 - 2^{-d}. \end{cases}$$

  - So there is a degree d polynomial p such that:

$$\begin{cases} f(x) = 1 \implies p(x) - 1/2 \in [2^{-d}, 1] \\ f(x) = -1 \implies p(x) - 1/2 \in [-1, 2^{-d}] \end{cases} .$$

Summary So Far

Giving an oracle O such that $\boldsymbol{SZK^O} \not\subseteq \boldsymbol{PP^O}$

Reduces To

Giving a function f such that $\boldsymbol{SZK}^{\text{dt}}(f) = O(\log n)$ and $\boldsymbol{PP}^{\text{dt}}(f) = n^{\Omega(1)}$

Reduces To

Giving a function f such that $\boldsymbol{SZK}^{\text{dt}}(f) = O(\log n)$ and $\widetilde{\deg}_{1-2^{-d}}(f) \geq d$ for some $d = n^{\Omega(1)}$

A Problem in $\mathsf{SZK}^{\mathsf{dt}}$ With Large $(1/3)$-Approximate Degree

- The Permutation Testing Problem (PTP) interprets its input as a list of N numbers $(x_1, \ldots, x_N)$ from range $\{1, \ldots, N\}$.
  - $\mathsf{PTP}(x) = -1$ if every number between 1 and N appears exactly once in the list.
  - $\mathsf{PTP}(x) = 1$ if at least $N/2$ range items do not appear in the list.
  - $\mathsf{PTP}(x)$ is undefined otherwise.

A Problem in $\mathsf{SZK}^{\mathsf{dt}}$ With Large $(1/3)$-Approximate Degree

- The Permutation Testing Problem (PTP) interprets its input as a list of N numbers $(x_1, \ldots, x_N)$ from range $\{1, \ldots, N\}$.
  - PTP$(x) = -1$ if every number between 1 and N appears exactly once in the list.
  - PTP$(x) = 1$ if at least $N/2$ range items do not appear in the list.
  - PTP(x) is undefined otherwise.
- Fact: $\mathbf{SZK^{dt}}(\text{PTP}) = O(\log n)$.
  - Idea: Verifier picks a range item j at random, and demands that prover provide an i such that $x_i = j$.

A Problem in SZK$^{\text{dt}}$ With Large $(1/3)$-Approximate Degree

- The Permutation Testing Problem (PTP) interprets its input as a list of N numbers $(x_1, \ldots, x_N)$ from range $\{1, \ldots, N\}$.
  - PTP$(x) = -1$ if every number between 1 and N appears exactly once in the list.
  - PTP$(x) = 1$ if at least $N/2$ range items do not appear in the list.
  - PTP(x) is undefined otherwise.
- Fact: $\mathbf{SZK^{dt}}$(PTP) $= O(\log n)$.
  - Idea: Verifier picks a range item j at random, and demands that prover provide an i such that $x_i = j$.
- Fact: $\widetilde{\deg}$(PTP) $= \tilde{\Theta}(n^{1/3})$ [Aaronson 2012, AS 2004].

The Punchline: A Problem Separating $\mathbf{SZK}^{\mathsf{dt}}$ And $\mathbf{PP}^{\mathsf{dt}}$

- Recall: we seek a function f such that: $\mathbf{SZK}^{\mathsf{dt}}(f) = O(\log n)$ and $\widetilde{\deg}_{1-2^{-d}}(f) = \Omega(d)$, for some $d = n^{\Omega(1)}$.
- Recall: $\mathbf{SZK}^{\mathsf{dt}}(\mathsf{PTP}) = O(\log n)$, and $\widetilde{\deg}(\mathsf{PTP}) = \tilde{\Theta}(n^{1/3})$.

The Punchline: A Problem Separating $\mathbf{SZK}^{\mathsf{dt}}$ And $\mathbf{PP}^{\mathsf{dt}}$

- Recall: we seek a function f such that: $\mathbf{SZK}^{\mathsf{dt}}(f) = O(\log n)$ and $\widetilde{\deg}_{1-2^{-d}}(f) = \Omega(d)$, for some $d = n^{\Omega(1)}$.
- Recall: $\mathbf{SZK}^{\mathsf{dt}}(\mathsf{PTP}) = O(\log n)$, and $\widetilde{\deg}(\mathsf{PTP}) = \tilde{\Theta}(n^{1/3})$.
- But $\widetilde{\deg}_{1-1/n}(\mathsf{PTP}) = O(\log n)$. ☹
- Can we turn PTP into a function F such that $\mathbf{SZK}^{\mathsf{dt}}(F) = O(\log n)$, yet $\widetilde{\deg}_{1-2^{-d}}(F) \geq d$ for $d = n^{\Omega(1)}$?

The Punchline: A Problem Separating $\mathbf{SZK}^{\mathsf{dt}}$ And $\mathbf{PP}^{\mathsf{dt}}$

- Recall: we seek a function f such that: $\mathbf{SZK}^{\mathsf{dt}}(f) = O(\log n)$ and $\widetilde{\deg}_{1-2^{-d}}(f) = \Omega(d)$, for some $d = n^{\Omega(1)}$.
- Recall: $\mathbf{SZK}^{\mathsf{dt}}(\mathsf{PTP}) = O(\log n)$, and $\widetilde{\deg}(\mathsf{PTP}) = \tilde{\Theta}(n^{1/3})$.
- But $\widetilde{\deg}_{1-1/n}(\mathsf{PTP}) = O(\log n)$. ☹
- Can we turn PTP into a function F such that $\mathbf{SZK}^{\mathsf{dt}}(F) = O(\log n)$, yet $\widetilde{\deg}_{1-2^{-d}}(F) \geq d$ for $d = n^{\Omega(1)}$?
- Yes! Let $F = \mathsf{GAPMAJ}_{n^{1/4}} \circ \mathsf{PTP}_{n^{3/4}}$.

The Punchline: A Problem Separating $\mathbf{SZK}^{\mathsf{dt}}$ And $\mathbf{PP}^{\mathsf{dt}}$

- Recall: we seek a function f such that: $\mathbf{SZK}^{\mathsf{dt}}(f) = O(\log n)$ and $\widetilde{\deg}_{1-2^{-d}}(f) = \Omega(d)$, for some $d = n^{\Omega(1)}$.
- Recall: $\mathbf{SZK}^{\mathsf{dt}}(\mathsf{PTP}) = O(\log n)$, and $\widetilde{\deg}(\mathsf{PTP}) = \tilde{\Theta}(n^{1/3})$.
- But $\widetilde{\deg}_{1-1/n}(\mathsf{PTP}) = O(\log n)$. ☹
- Can we turn PTP into a function F such that $\mathbf{SZK}^{\mathsf{dt}}(F) = O(\log n)$, yet $\widetilde{\deg}_{1-2^{-d}}(F) \geq d$ for $d = n^{\Omega(1)}$?
- Yes! Let $F = \mathsf{GAPMAJ}_{n^{1/4}} \circ \mathsf{PTP}_{n^{3/4}}$.
- Claim 1: $\widetilde{\deg}_{1-2^{-n^{1/4}}}(F) = \Omega(n^{1/4})$.
- Proof: By Theorem from earlier.

The Punchline: A Problem Separating $\mathbf{SZK}^{\mathsf{dt}}$ And $\mathbf{PP}^{\mathsf{dt}}$

- Recall: we seek a function f such that: $\mathbf{SZK}^{\mathsf{dt}}(f) = O(\log n)$ and $\widetilde{\deg}_{1-2^{-d}}(f) = \Omega(d)$, for some $d = n^{\Omega(1)}$.
- Recall: $\mathbf{SZK}^{\mathsf{dt}}(\mathsf{PTP}) = O(\log n)$, and $\widetilde{\deg}(\mathsf{PTP}) = \tilde{\Theta}(n^{1/3})$.
- But $\widetilde{\deg}_{1-1/n}(\mathsf{PTP}) = O(\log n)$. ☹
- Can we turn PTP into a function F such that $\mathbf{SZK}^{\mathsf{dt}}(F) = O(\log n)$, yet $\widetilde{\deg}_{1-2^{-d}}(F) \geq d$ for $d = n^{\Omega(1)}$?
- Yes! Let $F = \mathsf{GAPMAJ}_{n^{1/4}} \circ \mathsf{PTP}_{n^{3/4}}$.
- Claim 1: $\widetilde{\deg}_{1-2^{-n^{1/4}}}(F) = \Omega(n^{1/4})$.
- Proof: By Theorem from earlier.
- Claim 2: $\mathbf{SZK}^{\mathsf{dt}}(F) = O(\log n)$.
  - **Rough Intuition**: On input $x = (x_1, \ldots, x_{n^{1/4}})$ to F, Verifier picks a random $i \in \{1, \ldots, n^{1/4}\}$, and prover proves that $\mathsf{PTP}(x_i) = -1$ in zero-knowledge.
  - i.e., $\mathbf{SZK}^{\mathsf{dt}}$ is closed under composition with GAPMAJ.

Summary

- Many important hardness amplifications for approximate degree have been proven in recent years using the method of dual polynomials.
- These theorems show that the block-composed function $g \circ f$ is harder to approximate than f alone, even for very simple "hardness amplifiers" g.
- Most of the proofs use dual block composition and its variants.

Summary

- Many important hardness amplifications for approximate degree have been proven in recent years using the method of dual polynomials.
- These theorems show that the block-composed function $g \circ f$ is harder to approximate than f alone, even for very simple "hardness amplifiers" g.
- Most of the proofs use dual block composition and its variants.
- These results led to:
  - Improved understanding of how subclasses of the polynomial hierarchy (e.g. $\mathbf{SZK}$), and AC^0 circuits, can compute hard functions in query, communication, and relativized settings.
  - Secret-sharing schemes with reconstruction procedures in AC^0.
  - and more.

Summary

- Many important hardness amplifications for approximate degree have been proven in recent years using the method of dual polynomials.
- These theorems show that the block-composed function $g \circ f$ is harder to approximate than f alone, even for very simple "hardness amplifiers" g.
- Most of the proofs use dual block composition and its variants.
- These results led to:
  - Improved understanding of how subclasses of the polynomial hierarchy (e.g. $\mathbf{SZK}$), and AC^0 circuits, can compute hard functions in query, communication, and relativized settings.
  - Secret-sharing schemes with reconstruction procedures in AC^0.
  - and more.
- Next talk by Mark Bun: beyond block-composed functions.

Thank you!