Chebyshev Polynomials, Approximate Degree, and Their Applications

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
- $\widetilde{\deg}(f) := \deg_{1/3}(f)$ 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 $\lim_{\epsilon \to 1} \widetilde{\deg}_\epsilon(f)$.

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?

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

- Quantum query complexity [BBCMW98, AS01, Amb03, KSW04]
- Communication complexity [She08, SZ08, CA08, LS08, She12]
  - Lower bounds hold for a communication problem **related** to f.
  - Technique is called the **Pattern Matrix Method** [She08].
- Circuit complexity [MP69, Bei93, Bei94, She08]
- Oracle Separations [Bei94, BCHTV16]

Why Care About Approximate and Threshold Degree?

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

- Quantum query complexity [BBCMW98, AS01, Amb03, KSW04]
- Communication complexity [She08, SZ08, CA08, LS08, She12]
  - Lower bounds hold for a communication problem **related** to f.
  - Technique is called the **Pattern Matrix Method** [She08].
- Circuit complexity [MP69, Bei93, Bei94, She08]
- Oracle Separations [Bei94, BCHTV16]
- Lower bounds on $\widetilde{\deg}(f)$ also yield efficient secret-sharing schemes [BIVW16]

Details of Communication Applications

Lower bounds on $\widetilde{\deg}_\epsilon(f)$ and $\deg_\pm(f)$ yield communication lower bounds (often in a black-box manner) [Sherstov 2008]

- $\epsilon \approx 1/3$: BQP^{cc} lower bounds.
- $\epsilon \approx 1 - 2^{-n^\delta}$: PP^{cc} lower bounds
- $\epsilon \to 1$ (i.e., $\deg_\pm(f)$ lower bounds): UPP^{cc} lower bounds.

Example 1: The Approximate Degree of AND_n

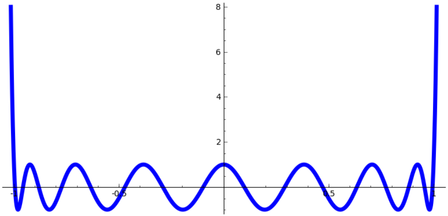
Example: What is the Approximate Degree of AND_n?

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

- 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

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

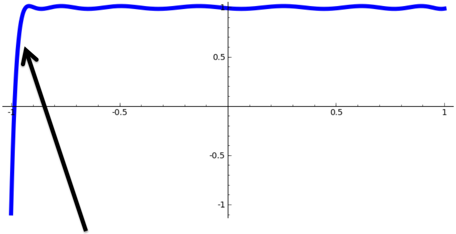
- Chebyshev polynomials are the extremal case.



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:



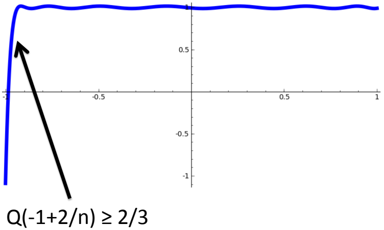
Q(-1+2/n) = 2/3

- 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 <u>univariate</u> polynomial p^{sym} that looks like this:

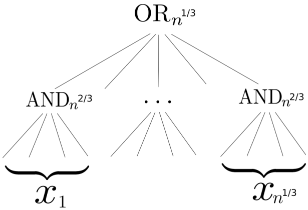


- Claim 1: $\deg(p^{\mathsf{sym}}) \leq \deg(p)$.
- Claim 2: Markov's inequality $\Longrightarrow \deg(p^{\mathsf{sym}}) = \Omega(n^{1/2})$.

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

The Minsky-Papert DNF

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



The Minsky-Papert DNF

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

The Minsky-Papert DNF

- Claim: $\deg_{\pm}(\mathsf{MP}) = \tilde{\Theta}(n^{1/3})$.
- The $\Omega(n^{1/3})$ lower bound was proved by Minsky and Papert in 1969 via a symmetrization argument.
  - More generally, $\deg_{\pm}(\mathrm{OR}_t \circ \mathrm{AND}_b) \geq \Omega(\min(t, b^{1/2}))$.
- We will prove the matching upper bound:

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

  - First, we'll construct a sign-representation of degree $O((b \log t)^{1/2})$ 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 $p = \frac{1}{2} \cdot (1 - p_1)$.
- 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 $\tilde{O}(b^{1/2})$

- 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 $p = \frac{1}{2} \cdot (1 - p_1)$.
- Then $\frac{1}{2} - \sum_{i=1}^{t} p(x_i)$ sign-represents $\mathrm{OR}_t \circ \mathrm{AND}_b$.
  - If $\mathrm{AND}_b(x_i) =$ FALSE for all i, then

$$\frac{1}{2} - \sum_{i=1}^{t} p(x_i) \geq \frac{1}{2} - t \cdot \frac{1}{8t} \geq 3/8.$$

  - If $\mathrm{AND}_b(x_i) =$ TRUE for even one i, then

$$\frac{1}{2} - \sum_{i=1}^{t} p(x_i) \leq \frac{1}{2} - 7/8 + (t-1) \cdot \frac{1}{8t} \leq -1/4.$$

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 $\frac{p(x)}{q(x)} = \frac{1}{2} \cdot \left(1 - \frac{p_1(x)}{q_1(x)}\right)$.
- Claim: The following polynomial sign-represents $\mathrm{OR}_t \circ \mathrm{AND}_b$.

$$r(x) := \left(\frac{1}{2} \cdot \prod_{1 \leq i \leq t} q^2(x_i)\right) - \sum_{i=1}^{t} \left(p(x_i) \cdot q(x_i) \cdot \prod_{1 \leq i \leq t, i' \neq i} q^2(x_{i'})\right).$$

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 $\frac{p(x)}{q(x)} = \frac{1}{2} \cdot \left(1 - \frac{p_1(x)}{q_1(x)}\right)$.
- Claim: The following polynomial sign-represents $\mathrm{OR}_t \circ \mathrm{AND}_b$.

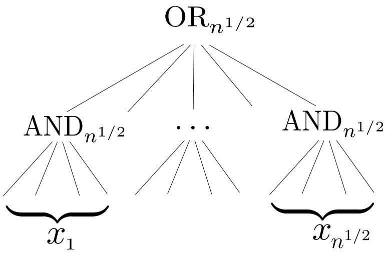
$$r(x) := \left(\frac{1}{2} \cdot \prod_{1 \leq i \leq t} q^2(x_i)\right) - \sum_{i=1}^{t} \left(p(x_i) \cdot q(x_i) \cdot \prod_{1 \leq i \leq t, i' \neq i} q^2(x_{i'})\right).$$

- Proof: $\mathrm{sgn}(\mathrm{OR}_t \circ \mathrm{AND}_b(x)) = \frac{1}{2} - \sum_{i=1}^{t} \frac{p(x_i)}{q(x_i)} = \frac{1}{2} - \sum_{i=1}^{t} \frac{p(x_i) \cdot q(x_i)}{q^2(x_i)} = \frac{r(x)}{\prod_{i=1}^{t} q^2(x_i)}$. The denominator of the RHS is non-negative, so throw it away w/o changing the sign.

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



History of the OR-AND Tree

Upper bounds

[HMW03] $\widetilde{\deg}(\text{OR-AND}_n) = O(n^{1/2})$

Lower bounds

[NS92]	$\Omega(n^{1/4})$
[Shi01]	$\Omega(n^{1/4}\sqrt{\log n})$
[Amb03]	$\Omega(n^{1/3})$
[Aar08]	Reposed Question
[She09]	$\Omega(n^{3/8})$
[BT13]	$\Omega(n^{1/2})$
[She13]	$\Omega(n^{1/2})$, independently

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):

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

Dual LP:

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

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) $\sum_{x \in \{-1,1\}^n} \psi(x) q(x) = 0$, 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.

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

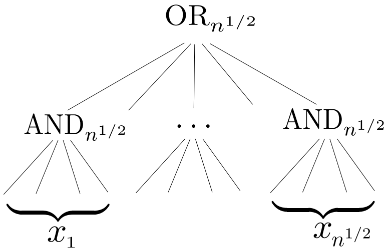
Constructing a Dual Polynomial

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

The Combining Method [She09, Lee09]

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

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



The Combining Method [She09, Lee09]

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

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

Must verify:

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

The Combining Method [She09, Lee09]

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

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

Must verify:

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

Additional Recent Progress on Approximate and Threshold Degree Lower Bounds

(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

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

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

- $\widetilde{\text{odeg}}_{-,\epsilon}(f) =$ 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: Part 1

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: Part 1

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: Part 1

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: Part 2

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

Recent Theorems: Part 2

- For other 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 $F = \mathsf{GAPMAJ}_t(f, \ldots, f)$. Then $\deg_{\pm}(F) \geq \Omega(\min\{d, t\})$.

Recent Theorems: Part 2

- For other 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 $F = \mathsf{GAPMAJ}_t(f, \ldots, f)$. Then $\deg_{\pm}(F) \geq \Omega(\min\{d, t\})$.

Compare to:

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: Part 2

- For other 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 $F = \mathsf{GAPMAJ}_t(f, \ldots, f)$. Then $\deg_{\pm}(F) \geq \Omega(\min\{d, t\})$.

- Implies a number of new oracle separations:

$$\mathsf{SZK}^A \not\subseteq \mathsf{PP}^A,\ \mathsf{SZK}^A \not\subseteq \mathsf{PZK}^A,\ \text{and } \mathsf{NIPZK}^A \not\subseteq \mathsf{coNIPZK}^A.$$

Applications to Communication Complexity

Definition of the UPP^{cc} Communication Model

Alice

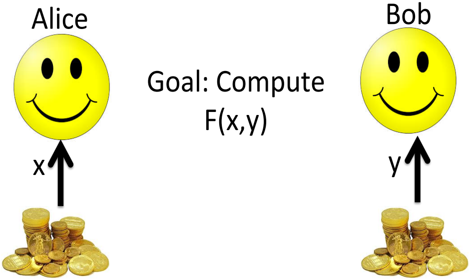
x

Goal: Compute F(x,y)

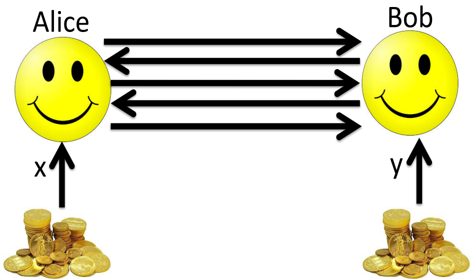
Bob

y

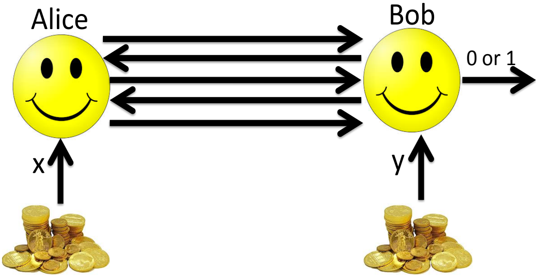
Definition of the UPP^{cc} Communication Model



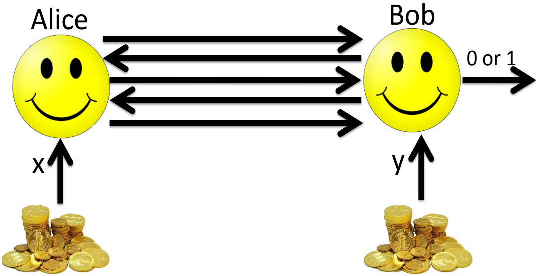
Definition of the UPP^{cc} Communication Model



Definition of the UPP^{cc} Communication Model

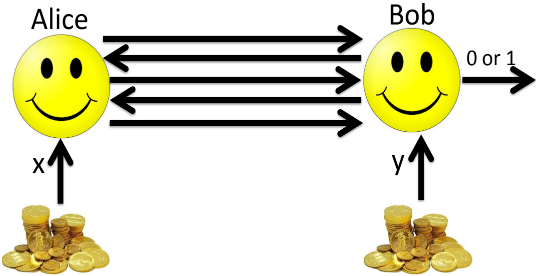


Definition of the UPP^{cc} Communication Model



- Protocol computes F if on every input (x, y), the output is correct with probability greater than $1/2$.
- The cost of a protocol is the worst-case number of bits exchanged on any input (x, y).

Definition of the UPP^{cc} Communication Model



- Protocol computes F if on every input (x, y), the output is correct with probability greater than $1/2$.
- The cost of a protocol is the worst-case number of bits exchanged on any input (x, y).
- UPP$^{cc}(F)$ is the least cost of a protocol that computes F.
- UPP^{cc} is the class of all F computed by UPP^{cc} protocols of polylogarithmic cost.

Importance of UPP^{cc}

- UPP^{cc} is <u>the</u> strongest two-party communication model against which we can prove lower bounds.
- Progress on UPP^{cc} has been slow.

Importance of $\mathsf{UPP}^{\mathsf{cc}}$

- $\mathsf{UPP}^{\mathsf{cc}}$ is the strongest two-party communication model against which we can prove lower bounds.
- Progress on $\mathsf{UPP}^{\mathsf{cc}}$ has been slow.
    - Paturi and Simon (1984) showed that

$$\mathsf{UPP}^{\mathsf{cc}}(F) \approx \log\left(\text{sign-rank}([F(x,y)]_{x,y})\right).$$

    - Forster (2001) nearly-optimal lower bounds on the $\mathsf{UPP}^{\mathsf{cc}}$ complexity of Hadamard matrices.
    - Razborov and Sherstov (2008) proved polynomial $\mathsf{UPP}^{\mathsf{cc}}$ lower bounds for a function in $\mathsf{PH}^{\mathsf{cc}}$ (more context to follow).

Rest of the Talk: How Much of $\mathsf{PH}^{\mathsf{cc}}$ is Contained In $\mathsf{UPP}^{\mathsf{cc}}$?

Background

- An important question in complexity theory is to determine the relative power of alternation (as captured by the polynomial-hierarchy PH), and counting (as captured by #P and its decisional variant PP).
- Both PH and PP generalize NP in natural ways.
- Toda famously showed that their power is related: $\mathsf{PH} \subseteq \mathsf{P}^{\mathsf{PP}}$.
- But it is open how much of PH is contained in PP itself.

Background

- An important question in complexity theory is to determine the relative power of alternation (as captured by the polynomial-hierarchy PH), and counting (as captured by #P and its decisional variant PP).
- Both PH and PP generalize NP in natural ways.
- Toda famously showed that their power is related: $\mathsf{PH} \subseteq \mathsf{P}^{\mathsf{PP}}$.
- But it is open how much of PH is contained in PP itself.
- Babai, Frankl, and Simon (1986) introduced communication analogues of Turing Machine complexity classes.
- Main question they left open was the relationship between PH^{cc} and UPP^{cc}.
  - Is $\mathsf{PH}^{cc} \subseteq \mathsf{UPP}^{cc}$?
  - Is $\mathsf{UPP}^{cc} \subseteq \mathsf{PH}^{cc}$?

Prior Work By Razborov and Sherstov (2008)

- Razborov and Sherstov (2008) resolved the first question left open by Babai, Frankl, and Simon!
- They gave a function F in $\mathsf{PH}^{\mathsf{cc}}$ (actually, in Σ_2^{cc}) such that $\mathsf{UPP}^{\mathsf{cc}}(F) = \Omega(n^{1/3})$.

Remainder of the Talk

- Goal: show that even lower levels of PH^{cc} are not in UPP^{cc}.
- Outline:
  - Proof sketch for Razborov and Sherstov (2008).
    - Threshold degree and its relation to UPP^{cc}.
    - The Pattern Matrix Method (PMM).
    - Combining PMM with "smooth dual witnesses" to prove UPP^{cc} lower bounds.
  - Improving on Razborov and Sherstov.

Communication Upper Bounds from Threshold Degree Upper Bounds

- Let $F\colon \{-1,1\}^n \times \{-1,1\}^n \to \{-1,1\}$.
- Claim: Let $d = \deg_{\pm}(F)$. There is a $\mathsf{UPP}^{\mathsf{cc}}$ protocol of cost $O(d \log n)$ computing $F(x,y)$.

Communication Upper Bounds from Threshold Degree Upper Bounds

- Let $F\colon \{-1,1\}^n \times \{-1,1\}^n \to \{-1,1\}$.
- Claim: Let $d = \deg_{\pm}(F)$. There is a $\mathsf{UPP}^{\mathsf{cc}}$ protocol of cost $O(d \log n)$ computing $F(x,y)$.
- Proof: Let $p(x,y) = \sum_{|T| \leq d} c_T \cdot \chi_T(x,y)$ sign-represent F.
- Alice chooses a parity T with probability proportional to $|c_T|$, and sends to Bob T and $\chi_{T \cap [n]}(y)$.
- From this, Bob can compute and output $\mathrm{sgn}(c_T) \cdot \chi_T(x,y)$.

Communication Upper Bounds from Threshold Degree Upper Bounds

- Let $F\colon \{-1,1\}^n \times \{-1,1\}^n \to \{-1,1\}$.
- Claim: Let $d = \deg_{\pm}(F)$. There is a $\mathsf{UPP}^{\mathsf{cc}}$ protocol of cost $O(d \log n)$ computing $F(x,y)$.
- Proof: Let $p(x,y) = \sum_{|T| \leq d} c_T \cdot \chi_T(x,y)$ sign-represent F.
- Alice chooses a parity T with probability proportional to $|c_T|$, and sends to Bob T and $\chi_{T \cap [n]}(y)$.
- From this, Bob can compute and output $\operatorname{sgn}(c_T) \cdot \chi_T(x,y)$.
- Since p sign-represents F, the output is correct with probability strictly greater than $1/2$.
- Communication cost is $O(d \log n)$.

Communication Lower Bounds from Threshold Degree Lower Bounds

- The previous slide showed that threshold degree upper bounds for $F(x, y)$ imply communication upper bounds for $F(x, y)$.
- Can we use threshold degree lower bounds for $F(x, y)$ to establish communication lower bounds for $F(x, y)$?

Communication Lower Bounds from Threshold Degree Lower Bounds

- The previous slide showed that threshold degree upper bounds for $F(x, y)$ imply communication upper bounds for $F(x, y)$.
- Can we use threshold degree lower bounds for $F(x, y)$ to establish communication lower bounds for $F(x, y)$?
- Answer: No. Bad Example: The parity function has linear threshold degree, but constant communication complexity.

Communication Lower Bounds from Threshold Degree Lower Bounds

- The previous slide showed that threshold degree upper bounds for $F(x, y)$ imply communication upper bounds for $F(x, y)$.
- Can we use threshold degree lower bounds for $F(x, y)$ to establish communication lower bounds for $F(x, y)$?
- Answer: No. Bad Example: The parity function has linear threshold degree, but constant communication complexity.
- Next Slide: Something almost as good.
  - A way to turn threshold degree lower bounds for f into communication lower bounds for a related function $F(x, y)$.

The Pattern Matrix Method (Sherstov, 2008)

- Let $f\colon \{-1,1\}^n \to \{-1,1\}$ satisfy $\deg_{\pm}(f) \geq d$.
- Turn f into a $2^{2n} \times 2^{2n}$ matrix F with $\mathsf{UPP}^{\mathsf{cc}}(F) \geq d$.

The Pattern Matrix Method (Sherstov, 2008)

- Let $f\colon \{-1,1\}^n \to \{-1,1\}$ satisfy $\deg_{\pm}(f) \geq d$.
- Turn f into a $2^{2n} \times 2^{2n}$ matrix F with $\mathsf{UPP^{cc}}(F) \geq d$.
- (Sherstov, 2008) **almost** achieves this.
  - Sherstov turns f into a matrix F, called the "pattern matrix" of f, such that:
    - Any randomized communication protocol that computes F correctly with probability $p = 1/2 + 2^{-d}$ has cost at least d.

The Pattern Matrix Method (Sherstov, 2008)

- Let $f\colon \{-1,1\}^n \to \{-1,1\}$ satisfy $\deg_{\pm}(f) \geq d$.
- Turn f into a $2^{2n} \times 2^{2n}$ matrix F with $\mathrm{UPP}^{\mathrm{cc}}(F) \geq d$.
- (Sherstov, 2008) **almost** achieves this.
  - Sherstov turns f into a matrix F, called the "pattern matrix" of f, such that:
    - Any randomized communication protocol that computes F correctly with probability $p = 1/2 + 2^{-d}$ has cost at least d.
    - Note: to get a $\mathrm{UPP}^{\mathrm{cc}}$ lower bound, we would need the above to hold for any $p > 1/2$.

The Pattern Matrix Method (Sherstov, 2008)

- Let $f\colon \{-1,1\}^n \to \{-1,1\}$ satisfy $\deg_{\pm}(f) \geq d$.
- Turn f into a $2^{2n} \times 2^{2n}$ matrix F with $\mathsf{UPP}^{\mathsf{cc}}(F) \geq d$.
- (Sherstov, 2008) **almost** achieves this.
    - Sherstov turns f into a matrix F, called the "pattern matrix" of f, such that:
        - Any randomized communication protocol that computes F correctly with probability $p = 1/2 + 2^{-d}$ has cost at least d.
        - Note: to get a $\mathsf{UPP}^{\mathsf{cc}}$ lower bound, we would need the above to hold for any $p > 1/2$.
    - Specifically, $F(x, y)$ is set to $f(u)$, where $u(x, y)$ is **derived** from (x, y) in a simple way.

The Pattern Matrix Method (Sherstov, 2008)

- Let $f\colon \{-1,1\}^n \to \{-1,1\}$ satisfy $\deg_{\pm}(f) \geq d$.
- Turn f into a $2^{2n} \times 2^{2n}$ matrix F with $\mathrm{UPP}^{\mathrm{cc}}(F) \geq d$.
- (Sherstov, 2008) **almost** achieves this.
  - Sherstov turns f into a matrix F, called the "pattern matrix" of f, such that:
    - Any randomized communication protocol that computes F correctly with probability $p = 1/2 + 2^{-d}$ has cost at least d.
    - Note: to get a $\mathrm{UPP}^{\mathrm{cc}}$ lower bound, we would need the above to hold for any $p > 1/2$.
  - Specifically, $F(x, y)$ is set to $f(u)$, where $u(x, y)$ is **derived** from (x, y) in a simple way.
    - y "selects" n bits of x and flips some of them to obtain u.

Proof Sketch for the Pattern Matrix Method: Dual Witnesses

- Let μ be a dual "witness" to the fact that the threshold degree of f is large.

Proof Sketch for the Pattern Matrix Method: Dual Witnesses

- Let μ be a dual "witness" to the fact that the threshold degree of f is large.
- Sherstov shows that μ can be "lifted" into a distribution over $\{-1,1\}^{2n} \times \{-1,1\}^{2n}$ under which $F(x,y)$ cannot be computed with probability $1/2 + 2^{-d}$, unless the communication cost is at least d.

Smooth Dual Witnesses Imply UPP^{cc} Lower Bounds

- Let $f: \{-1,1\}^n \to \{-1,1\}$ satisfy $\deg_{\pm}(f) \geq d$.
- Razborov and Sherstov showed that if there is a dual witness μ for f that additionally satisfies a **smoothness** condition, then the pattern matrix F of f actually has $\text{UPP}^{\text{cc}}(F) \geq d$.

Smooth Dual Witnesses Imply UPP$^{\text{cc}}$ Lower Bounds

- Let $f\colon \{-1,1\}^n \to \{-1,1\}$ satisfy $\deg_{\pm}(f) \geq d$.
- Razborov and Sherstov showed that if there is a dual witness μ for f that additionally satisfies a **smoothness** condition, then the pattern matrix F of f actually has UPP$^{\text{cc}}(F) \geq d$.
- The bulk of Razborov-Sherstov is showing that the Minsky-Papert DNF has a smooth dual witness to the fact that its threshold degree is $\Omega(n^{1/3})$.
- Since f is computed by a DNF formula, its pattern matrix is in Σ_2^{cc}.

Improving on Razborov-Sherstov (Part 1)

- Recall:

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 $\deg_{\pm}(F) = \Omega(\min\{d, t\})$.

- The dual witness constructed in (Sherstov 2014) isn't smooth.
- [BT16] showed how to smooth-ify the dual witness of (Sherstov 2014) (under a mild additional restriction on f).
  - Implied more general and quantitatively stronger $\mathsf{UPP}^{\mathsf{cc}}$ lower bounds for Σ_2^{cc} compared to [RS08].

Improving on Razborov-Sherstov (Part 2)

- Recall:

Theorem (BCHTV16)

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

Improving on Razborov-Sherstov (Part 2)

- Recall:

Theorem (BCHTV16)

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

- Moreover, can use the methods of [BT16] to smooth-ify the dual witness!
- Corollary: a function in NISZK^{cc} that is not in UPP^{cc}.
  - Improves on Razborov-Sherstov because:

$$\mathsf{NISZK}^{cc} \subseteq \mathsf{SZK}^{cc} \subseteq \mathsf{AM}^{cc} \cap \mathsf{coAM}^{cc} \subseteq \mathsf{AM}^{cc} \subseteq \Sigma_2^{cc}.$$

Open Questions and Directions

- Beyond Block-Composed Functions.
  - Challenge problem: obtain quantitatively optimal lower bounds on the approximate degree and threshold degree of AC^0.
  - Best lower bound for approximate degree is $\Omega(n^{2/3})$ [AS04].
  - Best lower bound for threshold degree is $\Omega(n^{1/2})$ [She15].
  - Best upper bound for both is the trivial $O(n)$.
- Break the "UPP^{cc} barrier" in communication complexity.
  - i.e., Identify any communication class that is not contained in UPP^{cc} (such as NISZK^{cc}), and then prove a superlogarithmic lower bound on that class for an explicit function.
- Strengthen UPP^{cc} lower bounds into lower bounds on distribution-free Statistical Query learning algorithms.

Thank you!