

Hardness Amplification and the Approximate Degree of Constant Depth Circuits

Mark Bun¹ and Justin Thaler²

¹Harvard University

²Simons Institute for the Theory of Computing, UC Berkeley

5 December 2013

Boolean Functions

- Boolean function $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$



$$\text{AND}_n(x) = \begin{cases} -1 & \text{(TRUE)} & \text{if } x = (-1)^n \\ 1 & \text{(FALSE)} & \text{otherwise} \end{cases}$$

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

Why Care About Approximate Degree?

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

- $\epsilon \rightarrow 1$: PAC learning [KS01]
- ϵ “close to” 1: Attribute-Efficient Learning [KS04, STT12]
- $\epsilon < 1$ a constant: Agnostic Learning [KKMS05]

Why Care About Approximate Degree?

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

- Quantum query complexity [BBCMW98] [AS01] [Amb03] [KSW04]
- Communication complexity [BVW07] [She07] [SZ07] [CA08] [LS08] [She12]
- Circuit complexity [MP69] [Bei93] [Bei94] [She08]

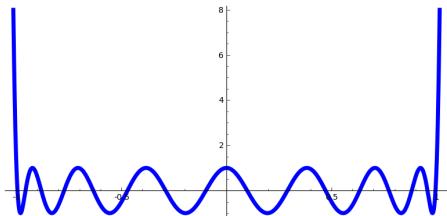
Example: What is the Approximate Degree of AND_n ?

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

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

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

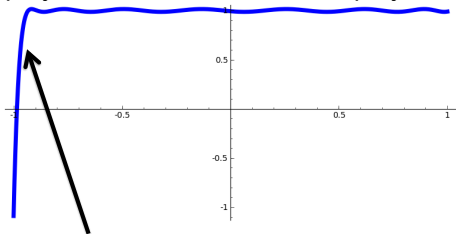
- Chebyshev polynomials are the extremal case.



Example: What is the Approximate Degree of AND_n ?

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

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



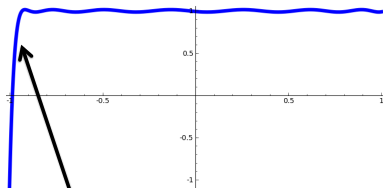
$$Q(-1+2/n) = 2/3$$

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

Example: What is the Approximate Degree of AND_n ?

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

- 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:



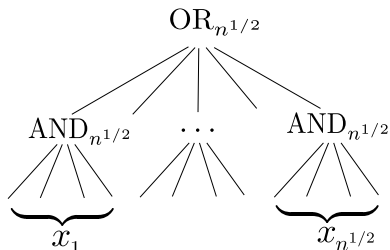
$$Q(-1+2/n) \geq 2/3$$

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

Beyond Symmetrization: Analyzing the OR-AND 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 $\widetilde{\text{deg}}(\text{OR-AND}_n)$?



History of the OR-AND Tree

Upper bounds

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

Lower bounds

$$[\text{NS92}] \quad \Omega(n^{1/4})$$

$$[\text{Shi01}] \quad \Omega(n^{1/4} \sqrt{\log n})$$

$$[\text{Amb03}] \quad \Omega(n^{1/3})$$

[Aar08] Reposed Question

$$[\text{She09}] \quad \Omega(n^{3/8})$$

$$[\text{BT13a}] \quad \Omega(n^{1/2})$$

$$[\text{She13a}] \quad \Omega(n^{1/2}), \text{ independently}$$

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):

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

Dual LP:

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

Dual Characterization of Approximate Degree

Theorem: $\deg_\epsilon(f) > d$ iff there exists a “dual polynomial” ψ 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) $\sum_{x \in \{-1,1\}^n} \psi(x)q(x) = 0, \deg q \leq d$ “pure high degree d ”

Key technique in, e.g., [She07] [Lee09] [She09]

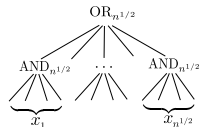
A **lossless** technique. Strong duality implies any approximate degree lower bound can be witnessed by dual polynomial.

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

Constructing a Dual Polynomial

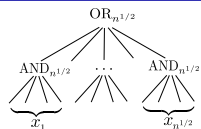
- By [NS92], there are dual polynomials
 ψ_{OUT} for $\widetilde{\text{deg}}(\text{OR}_{n^{1/2}}) = \Omega(n^{1/4})$ and
 ψ_{IN} for $\widetilde{\text{deg}}(\text{AND}_{n^{1/2}}) = \Omega(n^{1/4})$
- Can we combine ψ_{OUT} and ψ_{IN} to obtain a dual polynomial $\psi_{\text{OR-AND}}$ for OR-AND?

A First Attempt



$$\psi_{\mathbf{OR-AND}}(x_1, \dots, x_{n^{1/2}}) := \psi_{\mathbf{OUT}}(\dots, \psi_{\mathbf{IN}}(x_i), \dots)$$

A First Attempt

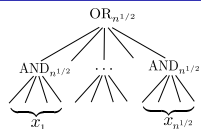


$$\psi_{\text{OR-AND}}(x_1, \dots, x_{n^{1/2}}) := \psi_{\text{OUT}}(\dots, \psi_{\text{IN}}(x_i), \dots)$$

- Easy to check: $\psi_{\text{OR-AND}}$ has pure high degree at least $n^{1/4} \cdot n^{1/4} = n^{1/2}$.
- E.g. If $\psi_{\text{OUT}}(y_1, y_2) = y_1 y_2$ and $\psi_{\text{IN}}(z_1, z_2) = z_1 z_2$, then

$$\psi_{\text{OR-AND}}(x_{11}, x_{12}, x_{21}, x_{22}) = (x_{11}x_{12})(x_{21}x_{22}) = x_{11}x_{12}x_{21}x_{22}.$$

A First Attempt



$$\psi_{\text{OR-AND}}(x_1, \dots, x_{n^{1/2}}) := \psi_{\text{OUT}}(\dots, \psi_{\text{IN}}(x_i), \dots)$$

- Easy to check: $\psi_{\text{OR-AND}}$ has pure high degree at least $n^{1/4} \cdot n^{1/4} = n^{1/2}$.
- E.g. If $\psi_{\text{OUT}}(y_1, y_2) = y_1 y_2$ and $\psi_{\text{IN}}(z_1, z_2) = z_1 z_2$, then
$$\psi_{\text{OR-AND}}(x_{11}, x_{12}, x_{21}, x_{22}) = (x_{11} x_{12})(x_{21} x_{22}) = x_{11} x_{12} x_{21} x_{22}.$$
- Does $\psi_{\text{OR-AND}}$ have high correlation with OR-AND_n ?
- Problem: Proposed definition of $\psi_{\text{OR-AND}}$ may feed non-Boolean values into ψ_{OUT} . But we only have control over ψ_{OUT} on **Boolean** inputs.

A Second (and Final) Attempt [She09, Lee09]

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

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

A Second (and Final) Attempt [She09, Lee09]

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

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

Must verify:

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

A Second (and Final) Attempt [She09, Lee09]

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

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

Must verify:

- 1 $\psi_{\text{OR-AND}}$ has pure high degree $\geq n^{1/4} \cdot n^{1/4} = n^{1/2}$. ✓ [She09]
- 2 $\psi_{\text{OR-AND}}$ has high correlation with OR-AND. [BT13a]

(Sub)Goal: Show $\psi_{\text{OR-AND}}$ has pure high degree at least $n^{1/2}$ [She09]

Pure High Degree Analysis [She09]

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

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

$$\begin{aligned} & C \cdot \text{sgn}(\psi_{\text{IN}}(x_1)) \cdot \text{sgn}(\psi_{\text{IN}}(x_2)) \cdot \prod_{i=1}^3 |\psi_{\text{IN}}(x_i)| \\ & = \psi_{\text{IN}}(x_1) \cdot \psi_{\text{IN}}(x_2) \cdot |\psi_{\text{IN}}(x_3)| \end{aligned}$$

Pure High Degree Analysis [She09]

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

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

$$\begin{aligned} & C \cdot \text{sgn}(\psi_{\text{IN}}(x_1)) \cdot \text{sgn}(\psi_{\text{IN}}(x_2)) \cdot \prod_{i=1}^3 |\psi_{\text{IN}}(x_i)| \\ & = \psi_{\text{IN}}(x_1) \cdot \psi_{\text{IN}}(x_2) \cdot |\psi_{\text{IN}}(x_3)| \end{aligned}$$

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

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

Correlation Analysis

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

- Idea: Show

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

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

Correlation Analysis

- Goal: Show

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

- Case 1: Consider any $y = (\text{sgn } \psi_{\text{IN}}(x_1), \dots, \text{sgn } \psi_{\text{IN}}(x_{n^{1/2}})) \neq \mathbf{All-False}$.
- There is some coordinate of y that equals TRUE. Only need to “trust” this coordinate to force OR-AND_n to evaluate to True on $(x_1, \dots, x_{n^{1/2}})$. So errors do not build up!

Correlation Analysis

- Case 2: Consider $y = \mathbf{All-False}$.
- $\text{OR}_{n^{1/2}}(y) = \text{OR-AND}_n(x_1, \dots, x_{n^{1/2}})$ only if all coordinates of y are “error-free”.
- Fortunately, $\psi_{\mathbf{IN}}$ has a special **one-sided error** property:
If $\text{sgn}(\psi_{\mathbf{IN}}(x_i)) = 1$, then $\text{AND}_{n^{1/2}}(x_i)$ is **guaranteed** to equal 1.

Summary of Correlation Analysis

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

One-Sided Approximate Degree

- A real polynomial p is a 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 one-sided ϵ -approximation for f .
- $\widetilde{\text{odeg}}(f) := \widetilde{\text{odeg}}_{1/3}(f)$ is the **one-sided approximate degree** of f .

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

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

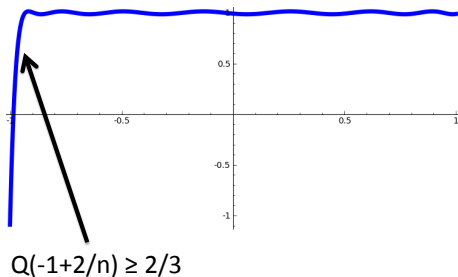
$$\begin{aligned} \min_{p, \epsilon} \quad & \epsilon \\ \text{s.t.} \quad & |p(x) - 1| \leq \epsilon \quad \text{for all } x \in f^{-1}(1) \\ & p(x) \leq -1 \quad \text{for all } x \in f^{-1}(-1) \\ & \deg p \leq d \end{aligned}$$

Dual LP:

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

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

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



Hardness Amplification for Constant-Depth Circuits [BT13b]

Main Theorem

- Given: A “simple” Boolean function f that is “hard to approximate to low error” by degree d polynomials.
- Can we turn f into a “still-simple” F that is hard to approximate even to very high error?

Main Theorem

- Given: A “simple” Boolean function f that is “hard to approximate to low error” by degree d polynomials.
- Can we turn f into a “still-simple” F that is hard to approximate even to very high error?

A: Yes.

Theorem

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

Proof of Main Theorem

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

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

- Combine ψ_{OUT} and ψ_{IN} exactly as before to obtain a dual witness ψ_F for F .

Must verify:

- 1 $\psi_{\text{OR-AND}}$ has pure high degree d .
- 2 $\psi_{\text{OR-AND}}$ has correlation at least $1 - 2^{-t}$ with F .

Proof of Main Theorem: Pure High Degree

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

Proof of Main Theorem: Correlation Analysis

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

- Idea: Show

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

- Case 1: Consider $y = (\text{sgn } \psi_{\text{IN}}(x_1), \dots, \text{sgn } \psi_{\text{IN}}(x_t)) =$
All-True.
- If even a single coordinate y_i of y is “error-free”, then
 $F(x) = \text{OR}_t(f(x_1), \dots, f(x_t)) = -1. \therefore \text{D}$
- Any individual coordinate of y is in error with probability at most $1/2$, since ψ_{IN} is well-correlated with f .
- So **all** coordinates of y are in error with probability only 2^{-t} .

Proof of Main Theorem: Correlation Analysis

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

- Idea: Show

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

- Case 2: Consider $y = (\text{sgn } \psi_{\text{IN}}(x_1), \dots, \text{sgn } \psi_{\text{IN}}(x_t)) =$
All-False. Then $\text{sgn}(\psi_F(x)) = \text{sgn}(\psi_{\text{OUT}}(y)) = 1$.
- Then $F(y) = \text{OR}_t(f(x_1), \dots, f(x_t)) = 1$ only if all coordinates of y are “error-free”.
- Fortunately, ψ_{IN} has one-sided error: If $\text{sgn}(\psi_{\text{IN}}(x_i)) = 1$, then $f(x_i)$ is **guaranteed** to equal 1.

A New $\widetilde{\text{odeg}}$ Bound for AC^0

- We want to apply amplification to functions in AC^0 , getting out very “hard” functions that are still in AC^0 .
- Let $\text{ED} : \{-1, 1\}^n \rightarrow \{-1, 1\}$ denote the ELEMENT DISTINCTNESS function.
- [AS04] showed $\widetilde{\text{deg}}(\text{ED}) = \Omega(N^{2/3})$.
- This is the best known lower bound on the approximate degree of an AC^0 function.
- We show that in fact $\widetilde{\text{odeg}}(\text{ED}) = \Omega(N^{2/3})$.

New Lower Bounds for AC^0

Theorem

Let $F = \text{OR}_{n^{2/5}}(\text{ED}_{n^{3/5}}, \dots, \text{ED}_{n^{3/5}})$ and $\epsilon = 1 - 2^{-n^{2/5}}$. Then $\widetilde{\text{odeg}}_\epsilon(F) = \tilde{\Omega}(n^{2/5})$.

Proof: Combine lower bound on $\widetilde{\text{odeg}}(\text{ED})$ with Main Theorem.

New Lower Bounds for AC^0

Definition

Let $f : X \times Y \rightarrow \{-1, 1\}$ be a function, and μ a probability distribution on $X \times Y$. The discrepancy of f under μ is

$$\text{disc}_\mu(f) := \max_{S \subseteq X, T \subseteq Y} \left| \sum_{x \in S} \sum_{y \in T} \mu(x, y) f(x, y) \right|.$$

The discrepancy of f is: $\text{disc}(f) := \min_\mu \text{disc}_\mu(f)$.

- Low discrepancy implies high communication complexity in nearly every communication model.
- Also a central quantity in learning theory and circuit complexity.

New Lower Bounds for AC^0

Theorem (She08, "Pattern Matrix Method")

Let $F : \{-1, 1\}^n$ be any function satisfying $\widetilde{\text{deg}}_{1-1/W}(F) \geq d$. Let $F' : \{-1, 1\}^{4n} \times \{-1, 1\}^{4n} \rightarrow \{-1, 1\}$ by

$$F'(x, y) = F(\dots, \bigvee_{j=1}^4 (x_{i,j} \wedge y_{i,j}), \dots).$$

Then $\text{disc}(F') \lesssim \max\{1/W, 2^{-d}\}$.

Corollary

There is an AC^0 function f (computed by a depth four circuit) with discrepancy $\exp(-\Omega(n^{2/5}))$.

Proof: Apply Pattern Mat. Meth. to $OR_{n^{2/5}}(ED_{n^{3/5}}, \dots, ED_{n^{3/5}})$.

Previous best bound: $\exp(-\Omega(n^{1/3}))$ [She08, BVW08].

More applications

Corollary

There is an AC^0 function f that cannot be computed by $MAJ \circ THR$ circuits of size $\exp(\Omega(n^{2/5}))$.

Corollary

There is an AC^0 function f with threshold weight $\exp(\Omega(n^{2/5}))$.

Back to OR-AND Trees

- Let $\text{OR-AND}_{d,n}$ denote the balanced OR-AND tree of depth d (with an OR gate at the top).
- Earlier, we proved $\widetilde{\text{deg}}(\text{OR-AND}_{2,n}) = \Theta(n^{1/2})$.
- But proving equivalent lower bound for depth 3 or greater remained open.

Back to OR-AND Trees

- Let $\text{OR-AND}_{d,n}$ denote the balanced OR-AND tree of depth d (with an OR gate at the top).
- Earlier, we proved $\widetilde{\text{deg}}(\text{OR-AND}_{2,n}) = \Theta(n^{1/2})$.
- But proving equivalent lower bound for depth 3 or greater remained open.

Theorem

For any constant $d > 1$, $\widetilde{\text{deg}}(\text{OR-AND}_{d,n}) = \Omega(n^{1/2} / \log^{d-2}(n))$.

(Upper bound of $O(n^{1/2})$ for any constant d follows from [She12]).

First Proof Attempt

- Goal: construct a dual polynomial for $\text{OR-AND}_{3,n}$.
- Let ψ_{IN} denote the dual polynomial for $\text{OR-AND}_{2,n^{2/3}}$ constructed earlier.
- Let ψ_{OUT} denote a dual polynomial witnessing $\widetilde{\text{deg}}(\text{OR}_{n^{1/3}}) = \Omega(n^{1/6})$
- Combine ψ_{IN} and ψ_{OUT} exactly as before:

$$\psi_{\text{COMB}}(x_1, \dots, x_{n^{1/3}}) := C \cdot \psi_{\text{OUT}}(\dots, \text{sgn}(\psi_{\text{IN}}(x_i)), \dots) \prod_{i=1}^{n^{1/3}} |\psi_{\text{IN}}(x_i)|$$

First Proof Attempt

- Goal: construct a dual polynomial for $\text{OR-AND}_{3,n}$.
- Let ψ_{IN} denote the dual polynomial for $\text{OR-AND}_{2,n^{2/3}}$ constructed earlier.
- Let ψ_{OUT} denote a dual polynomial witnessing $\widetilde{\text{deg}}(\text{OR}_{n^{1/3}}) = \Omega(n^{1/6})$
- Combine ψ_{IN} and ψ_{OUT} exactly as before:

$$\psi_{\text{COMB}}(x_1, \dots, x_{n^{1/3}}) := C \cdot \psi_{\text{OUT}}(\dots, \text{sgn}(\psi_{\text{IN}}(x_i)), \dots) \prod_{i=1}^{n^{1/3}} |\psi_{\text{IN}}(x_i)|$$

- ψ_{COMB} has p.h.d. $\Omega(n^{1/6} \cdot n^{1/3}) = \Omega(n^{1/2})$. ✓
- But ψ_{COMB} may have poor correlation with $\text{OR-AND}_{3,n}$.
Problem: ψ_{IN} does not have one-sided error.

Actual Proof

- Instead, use a different dual polynomial ψ_{IN} for $\text{OR-AND}_{2,n^{2/3}}$.
- Construction of ψ_{IN} uses hardness amplification to achieve the following:
 - ψ_{IN} has error “on both sides”, but the error from the “wrong side” will be very small.
 - Hardness amplification step causes ψ_{IN} to have p.h.d. $\Omega(n^{1/3}/\sqrt{\log n})$, rather than $\Omega(n^{1/3})$.

Subsequent Work by Sherstov [She13b]

Threshold Degree

Definition

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

Threshold Degree of AC^0

- Minsky and Papert [MP68] proved an $\Omega(n^{1/3})$ lower bound on the threshold degree of a specific DNF.
- It has been open ever since to prove a lower bound of $\Omega(n^{1/3+\delta})$ for any function in AC^0 .
- Only progress: $\Omega(n^{1/3} \log^k n)$ for any constant k [OS03].
- We conjectured in [BT13b] that $OR_{n^{2/5}}(ED_{n^{3/5}}, \dots, ED_{n^{3/5}})$ has threshold degree $\Omega(n^{2/5})$.

Subsequent Work

- Sherstov [She13b] has recently proved our conjecture.
- More generally, he exhibits a depth k circuit of polynomial size with threshold degree $\Omega(n^{(k-1)/(2k-1)})$.