COMS 6998: Algebraic Techniques in TCS (Fall'21) Oct 5, 2021

# Lecture 4: Probabilistic Polynomials and The Polynomial Method

Instructor: *Josh Alman* Scribe notes by: *Ahmed Shaaban, Yunfeng Guan*

Disclaimer: This draft may be incomplete or have errors. Consult the course webpage for the most up-to-date version.

## 1 Overview

In this lecture, we will introduce the Polynomial Method, and see how it can be applied to solve the following two problems.

1. $\mathsf{AC}^0[\oplus]$ lower bounds.
2. Algorithm design: Orthogonal Vectors, Nearest Neighbor. (next lecture)

## 2 $\mathsf{AC}^0[\oplus]$ lower bounds

### 2.1 Circuit class $\mathsf{AC}^0$

**Definition 1.** *The* $\mathsf{AC}^0$ *circuit class is the set of all functions that can be computed by circuits*

- *with* AND, OR, NOT *gates,*
- *of unbounded fan-in,*
- *with polynomial size and constant depth.*

The size of a circuit is the number of gates in the circuit.
The depth of a circuit is the length of the longest path from the input to the output.
The fan-in of a circuit is the largest number of inputs of a gate.

A natural and important question to ask for a specific circuit class is: What kind of functions can be computed?
Positive results: many important functions are in $\mathsf{AC}^0$, e.g., Add (arithmetic addition for binary representations), polynomial-size CNF/DNF formulas (computable with depth-2 $\mathsf{AC}^0$ circuits).
However a negative result is given by the next theorem:

**Theorem 2** ([FSS81])**.** $\mathsf{PARITY} \notin \mathsf{AC}^0$.

To prove this theorem the switching lemma was used. See [AB09, Ch.14] for a detailed proof.

### 2.2 $\mathsf{AC}^0[\oplus]$: Extension of $\mathsf{AC}^0$

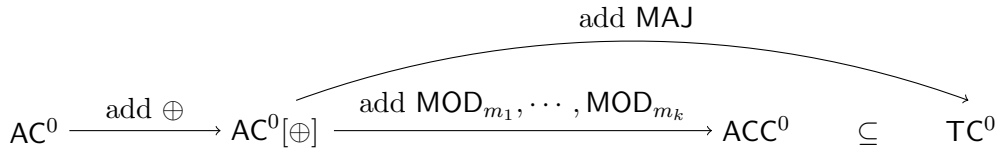
Since $\mathsf{AC}^0$ cannot compute $\mathsf{PARITY}$, one can try to empower it by allowing new gates.

**Definition 3.** $\mathsf{AC}^0[\oplus]$ *is the circuit class in which a circuit is under the restriction of* $\mathsf{AC}^0$*, but allowed to use* $\mathsf{PARITY}$ *gates.*

However, the following theorem, also the main theorem to prove in this lecture, shows that this circuit class is still not powerful enough to compute all functions.

**Theorem 4** (Main theorem, [Raz87, Smo87])**.** $\mathsf{MAJ} \notin \mathsf{AC}^0[\oplus]$.

### 2.3 Relations between circuit classes




Known hardness results:

- $\mathsf{PARITY} \notin \mathsf{AC}^0$.
- $\mathsf{MAJ} \notin \mathsf{AC}^0[\oplus]$.
- $\mathsf{E}^{\mathsf{NP}} \not\subseteq \mathsf{ACC}^0$ ([Wil11]). $\mathsf{E}^{\mathsf{NP}}$ is the class of problems solvable in exponential time with an $\mathsf{NP}$ oracle.
- There is no known hardness result for $\mathsf{TC}^0$.

## 3 Probabilistic Polynomials and The Polynomial Method

**Definition 5.** *Given a Boolean function* $f : \{0,1\}^n \to \{0,1\}$*, a probabilistic polynomial over* $\mathbb{F}$ *with error* $\varepsilon$ *for* $f$ *is a distribution* $\mathcal{P}$ *on polynomials* $p : \mathbb{F}^n \to \mathbb{F}$ *such that*

$$\forall x \in \{0,1\}^n : \Pr_{p \sim \mathcal{P}}[p(x) = f(x)] \geq 1 - \varepsilon.$$

In addition we have two concepts concerning probabilistic polynomials.

- The *degree* of a probabilistic polynomial $\deg(\mathcal{P})$ is the maximum degree of all polynomials in the support of $\mathcal{P}$.
- The *$\varepsilon$-probabilistic degree* of a Boolean function $f$ is the minimum degree of all probabilistic polynomials for $f$ with error $\varepsilon$.

Probabilistic polynomials in fact have proved to be a powerful tool in complexity theory and algorithmic design. In this lecture we will prove our main theorem (Theorem 4) using probabilistic polynomials (often referred to as the Polynomial Method).

The entire proof consists of two parts:

- Part A: Any $\mathsf{AC}^0[\oplus]$ circuit has low probabilistic degree over $\mathbb{F}_2$.
- Part B: $\mathsf{MAJ}$ does not have low probabilistic degree.

# 4 Part A: Low-degree probabilistic polynomial for $\mathsf{AC}^0[\oplus]$ circuits

## 4.1 Low-degree probabilistic polynomial for OR

**Lemma 6.** *For any $\varepsilon \in [0,1]$, the probabilistic degree of* OR *on $n$ inputs over $\mathbb{F}_2$ with error $\varepsilon$ is at most $\lceil \log_2(1/\varepsilon) \rceil$.*

It is worth noting that this bound is independent of $n$.

To prove this lemma, one needs to find a probabilistic polynomial for OR. Some intuition for constructing this probabilistic polynomial can be found from the example presented in Lecture 1 (See Lecture 1 Section 2.2.1 Part 4). We therein showcased a low-degree polynomial which is able to match the Boolean function on a majority of inputs.

To make this example more illuminating, we introduce the following definition.

**Definition 7.** *For any $\varepsilon \in [0,1]$, a polynomial $p : \mathbb{F}^n \to \mathbb{F}$ is a $(1-\varepsilon)$-correct polynomial for a Boolean function $f : \{0,1\}^n \to \{0,1\}$ if $|\{x \in \{0,1\}^n : p(x) = f(x)| \geq (1-\varepsilon) \cdot 2^n$.*

The $(1-\varepsilon)$-correct degree of $f$ over $\mathbb{F}$ if the minimum degree of such a polynomial $p$.

An important observation is that a probabilistic polynomial with error $\varepsilon$ must contain a $(1-\varepsilon)$-correct polynomial in its support. Hence the $(1-\varepsilon)$-correct degree over field $\mathbb{F}$ for a Boolean function $f$ is always a lower bound of $\varepsilon$-probabilistic degree.

Though this observation can't help us directly prove the lemma for OR, it will be of significant use in Part B.

## 4.2 Proof of Lemma 6

We try to design a randomized procedure such that the probabilistic polynomial is defined as the functions generated from this procedure.

Let $k = \lceil \log_2(1/\varepsilon) \rceil$. Pick $k$ subsets $S_1, \cdots, S_k \subseteq [n]$ independently uniformly at random. The procedure outputs the polynomial

$$p(x_1, \cdots, x_n) = 1 - \prod_{l=1}^{k} (1 - \sum_{i \in S_l} x_i) \quad \text{(over } \mathbb{F}_2\text{)}$$

It is easy to see that $p$ has degree $k$.

**Correctness:** Now we try to compute the probability of $p(x) = f(x)$.
Case 1: $x = (0, 0, \cdots, 0)$. $p(x) = 1$ with probability 1 (regardless of $S_1, \cdots, S_k$).
Case 2: $x \neq (0, 0, \cdots, 0)$. $p(x) = 1$ if $\exists l \in [k]$ s.t. $\sum_{i \in S_k} x_i = 1$.

**Claim 8.**

$$\Pr_{S \subseteq [n]}[\sum_{i \in S} x_i = 1] = \frac{1}{2}.$$

*Proof.* Use "Principle of deferred decisions" and pick $S$ in a 2-step fashion:

Fix a $j \in [n]$ such that $x_j = 1$. Step 1: $\forall i \in [n] \backslash \{j\}$, pick $i$ into $S$ with probability 1/2. Step 2: pick $j$ into $S$ with probability 1/2. As the two outcomes of step 2 correspond to different parity, we have that the overall probability is 1/2. □

Therefore for any $x \neq (0, 0, \cdots, 0)$,

$$\Pr[p(x) \neq 1] = \prod_{l=1}^{k} \Pr[\sum_{i \in S_k} x_i \neq 1] = 2^{-k}.$$

Since in the beginning we set $k = \lceil \log_2(1/\epsilon) \rceil$, we have $\Pr[p(x) \neq 1] \leq \varepsilon$.

**Remark:** We can similarly prove AND has low probabilistic degree.

### 4.3 Main result for Part A

**Theorem 9.** *Any circuit in* $\mathsf{AC}^0[\oplus]$ *has low probabilistic degree over* $\mathbb{F}_2$.

*More specifically, any function* $f$ *computable by an* $\mathsf{AC}^0[\oplus]$ *circuit with size* $s$ *and depth* $d$ *has a probabilistic polynomial over* $\mathbb{F}_2$ *with error* $1/3$ *and degree* $O(\log^d n)$.

*Proof.* To prove that any $\mathsf{AC}^0[\oplus]$ has low degree we will use the previous lemma. Let $A$ be the $\mathsf{AC}^0[\oplus]$ circuit that computes the function $f$, and has size $s$ and depth $d$. Replace each OR and AND gate of $A$ with a probabilistic polynomial of error $\dfrac{1}{3s}$.

By a simple union bound the total probability of error is $\leq s \cdot \dfrac{1}{3s} = \dfrac{1}{3}$. Note that $s$ is a polynomial in $n$ so it follows by the previous lemma that the degree of each polynomial is $O(\log n)$.

Next we compose all these polynomials to compute the circuit. For a depth-$d$ circuit, we eventually get a polynomial of degree $O(\log^d n)$. This bound on the degree follows by simply noting that the composition of polynomials multiplies their degrees.

Finally note that parity gates do not pose a problem because the parity of a string can be computed using the degree 1 polynomial $p(x) = x_1 + \cdots + x_n$ (over the field $\mathbb{F}_2$).

In conclusion, we proved that any function computable by $\mathsf{AC}^0[\oplus]$ circuits has poly $\log(n)$ probabilistic degree over $\mathbb{F}_2$ with error $1/3$. □

## 5 Part B: No low-degree probabilistic polynomial for MAJ

### 5.1 Low $(1-\varepsilon)$-correct degree for MAJ ⇒ for all Boolean functions

**Lemma 10.** *For any* $S \subseteq \{0,1\}^n$, *if there is a polynomial* $p : \mathbb{F}_2^n \to \mathbb{F}_2$ *that satisfies* $p(x) = \mathsf{MAJ}(x)$ *for all* $x \in S$, *and* $\deg(p) = d$, *then for all* $f : \{0,1\}^n \to \{0,1\}$, *there is a polynomial* $q : \mathbb{F}_2^n \to \mathbb{F}_2$ *with* $\deg(q) \leq d + n/2$ *s.t.* $q(x) = f(x)$ *for all* $x \in S$.

*Proof.* Write $f$ in 2 equivalent ways.

$$f(x_1, \cdots, x_n) = \sum_{T \subseteq [n]} a_T \prod_{i \in T} x_i = \sum_{T \subseteq [n]} b_T \prod_{i \in T} (1 - x_i).$$

If $\mathsf{MAJ}(x) = 0$, then all $\prod_{i \in T} x_i$ terms involving more than half of the variables must evaluate to 0. Then

$$f(x) = \sum_{T \subseteq [n], |T| \leq n/2} a_T \prod_{i \in T} x_i. \qquad \text{(Denote as } g_1(x)\text{)}$$

Similarly, if $\mathsf{MAJ}(x) = 1$,

$$f(x) = \sum_{T\subseteq[n],|T|\leq n/2} b_T \prod_{i\in T}(1-x_i). \qquad \text{(Denote as } g_2(x)\text{)}$$

Combine the two cases above, we know $f(x)$ can also be written as

$$f(x) = (1 - \mathsf{MAJ}(x)) \cdot g_1(x) + \mathsf{MAJ}(x) \cdot g_2(x).$$

Therefore, if $p(x)$ satisfies $p(x) = \mathsf{MAJ}(x)$ for all $x \in S$, and $\deg(p) = d$, then $q(x) = (1 - p(x)) \cdot g_1(x) + p(x) \cdot g_2(x)$ is a polynomial of degree $d + n/2$ satisfying $q(x) = f(x)$ for all $x \in S$. □

### 5.2 $(1-\varepsilon)$-correct degree lower bound for MAJ

Here we prove the $(1-\varepsilon)$-correct degree lower bound for $\mathsf{MAJ}$. And the main result of Part B is in fact a direct corollary to this theorem. For proving we will use Lemma 10 and the technique of counting argument (dimension argument).

**Theorem 11.** *There is a constant $c > 0$ s.t. for every polynomial $p : \mathbb{F}_2^n \to \mathbb{F}_2$ of degree $< c \cdot \sqrt{n}$ we have*

$$S = \{x \in \{0,1\}^n : p(x) = \mathsf{MAJ}(x)\}, \quad |S| < \frac{2}{3} \cdot 2^n.$$

*Proof.* Assume to the contrary that there exists a polynomial $p$ of degree $< c \cdot \sqrt{n}$ that satisfies $|S| \geq \frac{2}{3} \cdot 2^n$ where $S = \{x \in \{0,1\}^n : p(x) = \mathsf{MAJ}(x)\}$.

First we construct the following function: For any $x \in \mathbb{F}_2^n$, $f_x : \mathbb{F}_2^n \to \mathbb{F}_2$ is the function satisfying $f_x(x) = 1$ and $f_x(y) = 0, \forall y \in \mathbb{F}_2^n, y \neq x$.

By Lemma 10, there exists a polynomial $p_x$ of degree $\leq n/2 + c\sqrt{n}$ for each $f_x$ such that $p_x(y) = f_x(y), \forall y \in S$.

Let $V = \operatorname{span}_{\mathbb{F}_2}\{p_x : x \in S\}$. Since all $f_x$ are linearly independent, we have

$$\dim(V) = |S| \geq \frac{2}{3} \cdot 2^n.$$

On the other hand, since we consider the field $\mathbb{F}_2$, and $x \cdot x = x$ in $\mathbb{F}_2$, all polynomials are mutilinear, so

$$\begin{aligned}
\dim(V) \leq\ & \dim\left(\text{multilinear polynomials } \mathbb{F}_2^n \to \mathbb{F}_2 \text{ of degree} \leq \frac{n}{2} + c\sqrt{n}\right) \\
=\ & \#\text{ of monomials over } n \text{ variables of degree} \leq \frac{n}{2} + c\sqrt{n} \\
=\ & \sum_{i=0}^{\frac{n}{2}+c\sqrt{n}} \binom{n}{i} = 2^n - \sum_{i=0}^{\frac{n}{2}-c\sqrt{n}} \binom{n}{i}
\end{aligned}$$

By using an upper bound for binomial coefficients we will prove later (Claim 12), we know

$$\dim(V) \leq 2^n - \sum_{i=0}^{\frac{n}{2}-c\sqrt{n}} \binom{n}{i} < \frac{2}{3} \cdot 2^n.$$

which leads to contradiction. □

### 5.3 Main result for Part B

After proving the theorem, we can derive the main result easily. Recall the definition of $(1-\varepsilon)$-correct polynomial given in Part A, we can see Theorem 11 is just an equivalent statement of $\mathsf{MAJ}$ does not have 2/3-correct polynomial of degree $< c\sqrt{n}$. Then the 2/3-correct degree of $\mathsf{MAJ}$ is at least $c\sqrt{n}$. Further, from the observation given along with the definition, the 1/3-probabilistic degree is at least $c\sqrt{n}$. This implies the final result we want.

### 5.4 Combining Part A and Part B

In Part A we proved that the 1/3-probabilistic degree of any function computable by $\mathsf{AC}^0[\oplus]$ circuits is $\leq O(\text{poly}\log(n))$, and in Part B we proved that the 1/3-probabilistic degree of $\mathsf{MAJ}$ is $\geq \Omega(\sqrt{n})$. This finishes the proof of $\mathsf{MAJ} \notin \mathsf{AC}^0[\oplus]$.

### 5.5 Binomial coefficient bound

Now it only remains to prove the last step in Theorem 11.

**Claim 12.** $\sum_{i=1}^{\frac{n}{2}-c\sqrt{n}} \binom{n}{i} > \frac{1}{3} \cdot 2^n$.

*Proof.* First observe that the quantity

$$\frac{\sum_{i=0}^{\frac{n}{2}-c\sqrt{n}} \binom{n}{i}}{2^n}$$

is the probability that if we flip $n$ coins with $\Pr[H] = \Pr[T] = \frac{1}{2}$ such that the number of heads is less than or equal to $\frac{n}{2} - c\sqrt{n}$.

Define $n$ random variables $x_1, x_2, \cdots, x_n$ such that $x_i = 1$ iff the outcome of the $i$-th coin is head.

We then employ the following "reverse" Chernoff bound which lower bounds the tail probability:

$$\Pr\left(\sum_{i=1}^{n} x_i \leq (1-\varepsilon)pn\right) \geq \exp(-9pn\varepsilon^2).$$

Where $p$ is the probability that $x_i = 1$ (in this case $\frac{1}{2}$). By setting $\varepsilon = \frac{2c}{\sqrt{n}}$ we get that:

$$\Pr\left(\sum_{i=1}^{n} x_i \leq \frac{n}{2} - c\sqrt{n}\right) \geq \exp(-18c^2).$$

By choosing the right constant $c$ we are able to conclude:

$$\sum_{i=0}^{\frac{n}{2}-c\sqrt{n}} \binom{n}{i} = 2^n \cdot \Pr\left(\sum_{i=1}^{n} x_i \leq \frac{n}{2} - c\sqrt{n}\right) > \frac{1}{3}2^n.$$

□

## References


[AB09] Sanjeev Arora and Boaz Barak. *Computational Complexity: A Modern Approach*. Cambridge University Press, USA, 1st edition, 2009.

[FSS81] Merrick Furst, James B. Saxe, and Michael Sipser. Parity, circuits, and the polynomial-time hierarchy. In *FOCS*, pages 260–270, 1981.

[Raz87] Alexander A. Razborov. Lower bounds on the size of bounded depth circuits over a complete basis with logical addition. *Mathematical notes of the Academy of Sciences of the USSR*, 41:333–338, 1987.

[Smo87] Roman Smolensky. Algebraic methods in the theory of lower bounds for boolean circuit complexity. In *STOC*, page 77–82, 1987.

[Wil11] Ryan Williams. Non-uniform ACC circuit lower bounds. In *CCC*, pages 115–125, 2011.