CSC2429 / MAT1304: Circuit Complexity February 28, 2019

Lecture 7: $\mathrm{AC}^0[p]$ lower bounds (continued) and approximation by real polynomials

Instructor: Benjamin Rossman

Picking up from last lecture, we consider (random) polynomials over the field $\mathbb{F}_p$ for an arbitrary prime p.

Definition 1. $\deg_\varepsilon(f)$ is the minimum degree of an ε-approximating random polynomial for f, that is, a random polynomial $A \in \mathbb{F}_p[x_1, \ldots, x_n]$ such that $\mathbb{P}_A[\ A(x) \neq f(x)\ ] \leq \varepsilon$ for every $x \in \{0,1\}^n$.

Last time we showed:

- $\deg_\varepsilon(\mathrm{MOD}_p) \leq p - 1$,
- $\deg_\varepsilon(\mathrm{OR}_p), \deg_\varepsilon(\mathrm{AND}_p) \leq (p-1)\lceil \log_p(1/\varepsilon) \rceil$

As a consequence:

Proposition 2. *If f is computed by an $\mathrm{AC}^0[p]$ circuit of depth d and size S, then $\deg_\varepsilon(f) = O(\log(S/\varepsilon))^d$. In particular, $\deg_{1/4}(f) = O(\log S)^d$.*

Today we will show:

Theorem 3 (Smolensky 1987)**.** *Depth-d $\mathrm{AC}^0[3]$ circuits for XOR_n require size $2^{\Omega(n^{1/2d})}$.*

In fact, we will prove a stronger lower bound for *formulas*:

Theorem 4. *Depth-d $\mathrm{AC}^0[3]$ formulas for XOR_n require size $2^{\Omega(dn^{1/2d})}$ (really: $2^{\Omega(d(n^{1/2d}-1))}$).*

(Note that Theorem 12 implies Theorem 3, since every depth-d circuit of size S is equivalent to a depth-d formula of size S^{d-1} at most.) As a corollary, Theorem 12 gives super-polynomial formula size lower bounds up to any depth $o(\log n)$.

1 Better approximating random polynomials for $\mathrm{AC}^0[p]$ formulas

We first prove a lemma that improves the dependence on ε in our random approximating polynomials.

Lemma 5 (Kopparty-Srinivasan)**.** *For any boolean function f, $\deg_\varepsilon(f) \leq \deg_{1/4}(f) \cdot O(\log(1/\varepsilon))$*

Proof. Let $A(x)$ be a 1/4-approximating random polynomial for $f(x)$ of degree $\deg_{1/4}(f)$. Let $A^{(1)}, \ldots, A^{(t)}$ be independent random copies of A. Let $M \in \mathbb{F}_p[y_1, \ldots, y_t]$ be a degree-t polynomial

such that $M(y) = \mathrm{MAJ}_t(y)$ for all $y \in \{0,1\}^n$. Let $B(x) = M(A^{(1)}(x), \ldots, A^{(t)}(x))$. Then $\deg(B) = \deg(M)\deg(A) = t \cdot \deg_{1/4}(f)$. For any $x \in \{0,1\}^n$,

$$\mathop{\mathbb{P}}_{B}[\ B(x) \neq f(x)\ ] \leq \mathbb{P}[\ \mathrm{Bin}(t, 1/4) \geq 1/2\ ] \leq \exp(-\Omega(t))$$

So it suffices to choose $t = O(\log(1/\varepsilon))$. □

We restate a lemma from the previous lecture:

Lemma 6. *Suppose $f(x) = g(h_1(x), \ldots, h_m(x))$. Then for all $\delta, \varepsilon_1, \ldots, \varepsilon_m$,*

$$\deg_{\delta+\varepsilon_1+\cdots+\varepsilon_m}(f) \leq \deg_\delta(g) \cdot \max_i \deg_{\varepsilon_i}(h_i).$$

In the application of Lemma 6 to prove Proposition 2, we set $\delta = \varepsilon_1 = \cdots = \varepsilon_m = \varepsilon/S$ for each gate g in a circuit of size S. We optimize the setting of these parameters in our improved approximating polynomials for formulas. We will make use of the following inequality.

Lemma 7. *For all $d \geq 2$ and $a, b \geq 0$, we have*

$$(\tfrac{1}{d-1}a + 1)^{d-1}(b+1) \leq (\tfrac{1}{d}(a+b) + 1)^d.$$

Proof. Let $\varphi(a,b) := (\mathrm{RHS}) - (\mathrm{LHS})$. Then

$$\frac{\partial}{\partial b}\varphi(a,b) = \left(\frac{a+b}{d} + 1\right)^{d-1} - \left(\frac{a}{d-1} + 1\right)^{d-1}.$$

This is zero iff $b = a/(d-1)$. This is a minimum of the function with $\phi(a, a/(d-1)) = 0$. □

We are ready to give the construction of approximating polynomials for formulas.

Proposition 8. *If F is an $\mathrm{AC}^0[p]$ formula of depth d and size S, then $\deg_{1/4}(F) \leq O(\frac{1}{d}p\log_p(S))^d$.*

Proof. By induction on d. The base case $d = 0$ is trivial. Suppose $F = g(F_1, \ldots, F_m)$ where $g \in \{\mathrm{MOD}_p, \mathrm{OR}, \mathrm{AND}\}$ and $F_1, \ldots, F_m$ have depth $d-1$ and sizes $S_1, \ldots, S_m$.

By Lemma 5,

$$\deg_{1/4}(F) \leq \deg_{1/8}(g) \cdot \max_i \deg_{S_i/8S}(F_i).$$

From last lecture, we know that $\deg_{1/8}(g) = O(p)$ for each $g \in \{\mathrm{MOD}_p, \mathrm{OR}, \mathrm{AND}\}$.

For each i, we have

$$\begin{aligned}
\deg_{S_i/8S}(F_i) &\leq \deg_{1/4}(F_i) \cdot O(\log(S/S_i)) && \text{Lemma 6} \\
&\leq p^{d-1} \cdot O(\tfrac{1}{d-1}\log_p(S_i))^{d-1} \cdot O(\log(S/S_i)) && \text{induction hypothesis} \\
&\leq p^{d-1} \cdot O(\tfrac{1}{d}\log_p(S))^d && \text{Lemma 7.}
\end{aligned}$$

Choose an appropriately large constant in the big-O, we conclude that

$$\deg_{1/4}(F) \leq O(\tfrac{1}{d}p\log_p(S))^d.$$ □

2 Lower bound for XOR_n

We now use Proposition 8 to prove a lower bound on the $\mathrm{AC}^0[3]$ formula size of XOR_n.

Lemma 9 (Main Lemma)**.** *Let p be a prime ≥ 3. If $A \in \mathbb{F}_p[x_1, \ldots, x_n]$ is a degree-Δ polynomial, then*

$$\mathop{\mathbb{P}}_{x \in \{0,1\}^n}[\ A(x) = \mathrm{XOR}_n(x)\ ] \leq \frac{1}{2} + O\left(\frac{\Delta}{\sqrt{n}}\right).$$

Proof. Let $\lambda : \mathbb{F}_p \to \mathbb{F}_p$ be the map $\lambda(x) = 1 - 2x$. We have $\lambda(0) = 1$ and $\lambda(1) = -1$; so $\lambda(b) = (-1)^b$ for $b \in \{0, 1\}$.

For $x \in \{0,1\}^n$, we have

$$\lambda(x_1 \oplus \cdots \oplus x_n) = (-1)^{x_1 \oplus \cdots \oplus x_n} = \prod_{i \in [n]} (-1)^{x_i} = \prod_{i \in [n]} \lambda(x_i).$$

So for $y \in \{1, -1\}^n$, if we apply the inverse $\lambda^{-1}(y_i) = (1 - y_i)/2 \in \{0, 1\}$ to each coordinate, we get

$$\lambda(\mathrm{XOR}_n(\lambda^{-1}(y_1), \ldots, \lambda^{-1}(y_n))) = \prod_{i \in [n]} y_i.$$

Therefore,

$$\mathop{\mathbb{P}}_{x \in \{0,1\}^n}\Big[\ A(x) = \mathrm{XOR}_n(x)\ \Big] = \mathop{\mathbb{P}}_{y \in \{-1,1\}^n}\Big[\ \lambda(A(\lambda^{-1}(y_1), \ldots, \lambda^{-1}(y_n))) = \prod_{i \in [n]} y_i\ \Big].$$

Define $\widetilde{A} \in \mathbb{F}_p[y_1, \ldots, y_n]$ by

$$\widetilde{A}(y) = \lambda(A(\lambda^{-1}(y_1), \ldots, \lambda^{-1}(y_n))).$$

Clearly, $\deg(\widetilde{A}) = \deg(A) = \Delta$. It now suffices to show

$$\mathop{\mathbb{P}}_{y \in \{1,-1\}^n}\Big[\ \widetilde{A}(y) = \prod_{i \in [n]} y_i\ \Big] \leq \frac{1}{2} + O\left(\frac{\Delta}{\sqrt{n}}\right).$$

Let $S \subseteq \{1, -1\}^n$ be the set

$$S \stackrel{\text{def}}{=} \Big\{y \in \{1, -1\}^n : \widetilde{A}(y) = \prod_{i \in [n]} y_i\Big\}.$$

Consider any function $f : S \to \mathbb{F}_p$. We claim that f is equivalent (over S) to a <u>multilinear</u> polynomial $M \in \mathbb{F}_p[y_1, \ldots, y_n]$ of degree at most $\frac{n+\Delta}{2}$ (that is, each monomial of M has the form $\prod_{i \in I} y_i$ where $|I| \leq \frac{n+\Delta}{2}$).

To see why, let $B \in \mathbb{F}_p[y_1, \ldots, y_n]$ be an *arbitrary* polynomial that is equivalent to f over S. First, we multi-linearize B by repeatedly substituting 1 for y_i^2 whenever possible until B. That

is, for all even (odd) exponents c, replace $(y_i)^c$ with 1 (respectively y_i. Let M' be the resulting multilinear polynomial, which computes the same function as B over $\{1,-1\}^n$.

We next transform M' to M by substituting each monomial $\prod_{i\in I} y_i$ of degree $|I| > \frac{n+\Delta}{2}$ with the polynomial

$$\widetilde{A}(y) \cdot \prod_{j\in[n]\setminus I} y_j$$

of degree

$$\deg(A) + (n - |I|) = \Delta + n - \frac{n+\Delta}{2} < \frac{n+\Delta}{2}.$$

Note that M computes the same function as M' over S, since for $y \in S$ we have

$$\widetilde{A}(y) \cdot \prod_{j\in[n]\setminus I} y_j = \prod_{i\in[n]} y_i \cdot \prod_{j\in[n]\setminus I} y_j = \prod_{i\in I} y_i \cdot \prod_{j\in[n]\setminus I} y_j^2 = \prod_{i\in I} y_i.$$

We now have

$$\begin{aligned}
|S| &= \log_p(\#\{\text{functions from } S \text{ to } \mathbb{F}_p\}) \\
&\le \log_p(\#\{\text{multilinear polynomials in } \mathbb{F}_p[y_1,\ldots,y_n] \text{ of degree at most } \tfrac{n+\Delta}{2}\}) \\
&= \#\{\text{multilinear monomials of degree at most } \tfrac{n+\Delta}{2}\} \\
&= \underbrace{\binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{\frac{n}{2}}}_{=\,2^{n-1}} + \underbrace{\binom{n}{\frac{n}{2}+1} + \cdots + \binom{n}{\frac{n}{2}+\frac{\Delta}{2}}}_{\text{each term has size } < 2^n/\sqrt{n}}.
\end{aligned}$$

We conclude that

$$\mathop{\mathbb{P}}_{x\in\{0,1\}^n}\Big[\ A(x) = \mathrm{XOR}_n(x)\ \Big] = \mathop{\mathbb{P}}_{y\in\{1,-1\}^n}\Big[\ \widetilde{A}(y) = \prod_{i\in[n]} y_i\ \Big] = \frac{|S|}{2^n} \le \frac{1}{2} + O\Big(\frac{\Delta}{\sqrt{n}}\Big).$$

□

Theorem 10. *Depth-d $\mathrm{AC}^0[3]$ formulas for XOR_n require size $2^{\Omega(d(n^{1/2d}-1))}$.*

Proof. Suppose XOR_n is computed by an $\mathrm{AC}^0[3]$ formula of depth d and size S. By Proposition 8, there is a polynomial $A \in \mathbb{F}_3[x_1,\ldots,x_n]$ of degree $O(\frac{1}{d}\log S + 1)^{2d}$ such that

$$\mathop{\mathbb{P}}_{x\in\{0,1\}^n}[\ A(x) \ge \mathrm{XOR}_n(x)\ ] \le \frac{3}{4}.$$

By Lemma 9,

$$\mathop{\mathbb{P}}_{x\in\{0,1\}^n}[\ A(x) = \mathrm{XOR}_n(x)\ ] \le \frac{1}{2} + \frac{O(\frac{1}{d}\log S + 1)^d}{\sqrt{n}}.$$

It follows that $S \ge 2^{\Omega(d(n^{1/2d}-1))}$. □

As mentioned earlier, this lower bound for formulas implies a $2^{\Omega(n^{1/2d})}$ lower bound for circuits (quantitatively the strongest known lower bound for $\mathrm{AC}^0[p]$). By essentially the same argument, we can show:

Theorem 11. *For all distinct primes p and q, depth-d $\mathrm{AC}^0[p]$ circuits for $\mathrm{MOD}_{q,n}$ require size $2^{\Omega_{p,q}(n^{1/2d})}$.*

By an analogue of Lemma 9 for the majority function, Razborov (1987) proved a similar lower bound for MAJ_n.

Theorem 12. *Depth-d $\mathrm{AC}^0[p]$ circuits for MAJ_n require size $2^{\Omega_p(n^{1/2d})}$.*

These lower bounds may also be stated as *correlation bounds* (a.k.a. average-case lower bounds). For example, the argument presented here shows that depth-d $\mathrm{AC}^0[3]$ circuit of size $2^{o(n^{1/2d})}$ fail to approximate XOR_n on more than $\frac{1}{2} + O(\frac{1}{\sqrt{n}})$ fraction of inputs in $\{0,1\}^n$. It is an open problem to prove a quantitatively stronger correlation bound for $\mathrm{AC}^0[p]$.

2.1 Prime powers

So far we have considered $\mathrm{AC}^0[p]$ for a fixed prime p. What about prime powers p^k? It turns out that $\mathrm{AC}^0[p^k]$ and $\mathrm{AC}^0[p]$ have the same power (up to a polynomial blow-up in size and a constant blow-up in depth). In the homework exercises, you are asked to show that the function $\mathrm{MOD}_4(x_1, \ldots, x_n)$ is computable by polynomial-size constant-depth $\mathrm{AC}^0[2]$ circuits; the similar result for arbitrary p^k and p is a straightforward generalization.

2.2 Mod-6 gates

The lower bound technique using approximating polynomials breaks down in the presence of MOD_m gates where m is not a prime power (such as $m = 6$). Though it is widely believe that simple functions like MAJ_n do not have polynomial-size constant-depth $\mathrm{AC}^0[6]$ circuits, proving this is a major open problem.

2.3 The class TC^0

TC^0 is the class of functions computable by poly-size constant-depth circuits in the basis of unbounded fan-in AND, OR, NOT and MAJ gates. Note that every threshold function $\mathrm{THR}_{k,n}$ (including $\mathrm{AND}_n = \mathrm{THR}_{n,n}$ and $\mathrm{OR}_n = \mathrm{THR}_{n,1}$) is a subfunction of MAJ_{2n+1} (by appropriately fixing $n+1$ variables to 0's or 1's). Therefore, it suffices to consider circuits with only MAJ gates and inputs labeled by literals or constants.

Recall that a *symmetric function* is a boolean function $f : \{0,1\}^n \to \{0,1\}$ such that $f(x)$ is determined by $|x|$; examples are XOR_n, MAJ_n and $\mathrm{THR}_{k,n}$. It's not difficult to show that every symmetric function is computable by a polynomial-size depth-2 circuit consisting of two layers of MAJ gates. It follows that TC^0 contains $\mathsf{AC}[m]$ for every constant m.

Recall that NC^1 is the class of boolean functions computable by polynomial size DeMorgan formulas. Since MAJ is computable by polynomial-size DeMorgan formulas (Valiant's construction), we see that $\mathrm{TC}^0 \subseteq \mathrm{NC}^1$. Therefore, we have the following picture of circuit classes within P/poly:

$$\mathrm{AC}^0 \subsetneq \begin{matrix}\mathrm{AC}^0[2] \\ \mathrm{AC}^0[3]\end{matrix} \subsetneq \mathrm{AC}^0[6] \subseteq \mathrm{ACC}^0 \subseteq \mathrm{TC}^0 \subseteq \mathrm{NC}^1 \subseteq \mathrm{P/poly}.$$

Here ACC^0 denote the union of classes $\bigcup_m \mathrm{AC}^0[m]$.

It is open whether $\mathrm{AC}^0[6] = \mathrm{P/poly}$. Super-polynomial lower bounds for $\mathrm{AC}^0[6]$ (moreover ACC^0) are known for explicit functions of high complexity. Ryan Williams in 2011 showed $\mathrm{ACC}^0 \nsubseteq \mathrm{NEXP}$ (nondeterministic exponential time) via a lower bound for the NEXP-complete problem SUCCINCT-3SAT. This result was improved to $\mathrm{ACC}^0 \nsubseteq \mathrm{NQP}$ (nondeterministic quasi-polynomial time) by Murray and Williams in 2018.

3 Real approximating polynomials (in the ℓ_0, ℓ_∞ and ℓ_2 norms)

We have discussed approximations of boolean functions $f : \{0,1\}^n \to \{0,1\}$ by polynomials over finite fields. What about real polynomials $A \in \mathbb{R}[x_1, \ldots, x_n]$? Here are there are different norms we may consider. Interesting things can be said about low-degree approximating polynomials in the ℓ_0, ℓ_∞ and ℓ_2 norms:

$$\begin{aligned}
\|f - A\|_0 &:= |\{x \in \{0,1\}^n : f(x) \neq A(x)\}|, \\
\|f - A\|_\infty &:= \max_{x \in \{0,1\}^n} |f(x) - A(x)|, \\
\|f - A\|_2^2 &:= \sum_{x \in \{0,1\}^n} (f(x) - A(x))^2.
\end{aligned}$$

(For polynomials over finite fields, only the ℓ_0-norm approximation makes sense.)

3.1 ℓ_0-approximation

Similar to approximation of OR_n over $\mathbb{F}_p[x_1, \ldots, x_n]$, there exist low-degree ε-approximating random polynomials over the reals (in the ℓ_0-norm). The construction uses the following special case of the Valiant-Vazirani Isolation Lemma.

Lemma 13. *Let $S_0 = [n]$ and for $j = 1, \ldots, \log n + 1$, let S_j be a uniform random subset of S_{j-1}. Then for every nonempty subset $X \subseteq [n]$,*

$$\mathbb{P}[\ (\exists j)\ (|X \cap S_j| = 1)\ ] \geq 1/6.$$

Proof.

$$\mathbb{P}[\ (\forall j)\ |X \cap S_j| \neq 1\ ] \leq \mathbb{P}[\ |X \cap S_{\log n}| \geq 2\ ] + \mathbb{P}[\ (\exists j)\ |X \cap S_j| \geq 2 \text{ and } |X \cap S_{j-1}| = 0\ ].$$

We have

$$\mathbb{P}[\ |X \cap S_{\log n+1}| \geq 2\ ] \leq \mathbb{P}[\ S_{\log n+1} \neq \emptyset\ ] \leq n(1/2)^{\log n+1} \leq 1/2.$$

And we have

$$\mathbb{P}[\ (\exists j)\ |X \cap S_j| \geq 2 \text{ and } |X \cap S_{j-1}| = 0\ ] \leq \max_{j,\ k\geq 2} \mathbb{P}[\ |X \cap S_{j-1}| = 0 \mid |X \cap S_j| = k \text{ and } |X \cap S_{j-1}| \leq 1\ ]$$
$$\leq \max_{k\geq 2} \frac{2^{-k}}{2^{-k} + k2^{-k}} \quad = \quad \frac{1/4}{1/4 + 1/2} \quad = \quad 1/3.$$

It follows that $\mathbb{P}[\ (\forall j)\ |X \cap S_j| \neq 1\ ] \geq 1 - (1/2) - (1/3) = 1/6$. □

Lemma 14 (Aspnes et al '93)**.** *There exists a random polynomial $A \in \mathbb{R}[x_1, \ldots, x_n]$ of degree $O(\log(1/\varepsilon) \cdot \log(n))$ such that, for every $x \in \{0,1\}^n$,*

$$\mathbb{P}_A[\ A(x) \neq \mathrm{OR}_n(x)\ ] \leq \varepsilon.$$

Proof. Let random sets $S_0, \ldots, S_{\log n+1}$ be as in Lemma 13. For each $0 \leq j \leq \log n + 1$, let $B_j(x)$ be the random degree-1 polynomial $B_j(x) := \sum_{i\in S_j} x_i$. Let

$$B(x) := 1 - \prod_{j=0}^{\log n+1} (1 - B_j(x)).$$

We claim that $B(x)$ is a $(1/6)$-approximating random polynomial for $\mathrm{OR}_n(x)$. To see why: if $\mathrm{OR}_n(x) = 0$, then

$$\mathbb{P}[\ B(x) = \mathrm{OR}_n(x)\ ] = \mathbb{P}[\ B(x) = 0\ ] = 1.$$

If $\mathrm{OR}_n(x) = 1$, then

$$\begin{aligned}\mathbb{P}[\ B(x) = \mathrm{OR}_n(x)\ ] &= \mathbb{P}[\ B(x) = 0\ ] \\ &= \mathbb{P}[\ (\exists j)\ B_j(x) = 1\ ] \\ &= \mathbb{P}[\ (\exists j)\ |S_j \cap \{i \in [n] : x_i = 1\}| = 1\ ] \\ &\geq 1/6.\end{aligned}$$

Let $B^{(1)}, \ldots, B^{(t)}$ be independent copies of the random polynomial B where $t = \log_{5/6}(1/\varepsilon)$, and let A be the random polynomial

$$A := 1 - \prod_{r=1}^{t} (1 - B^{(r)}).$$

We have error probability

$$\mathbb{P}[\ A(x) = \mathrm{OR}_n(x)\ ] \leq \varepsilon.$$

Finally, note that A has degree $\log_{5/6}(1/\varepsilon) \cdot O(\log n) = O(\log(1/\varepsilon) \cdot \log n)$. □

We have the following corollary.

Corollary 15. *If f is computed by AC^0 circuit of depth d and size S, there is a real polynomial of degree $O(\log(S))^{2d}$ that agrees with f on all but $\frac{1}{4}$-fraction of inputs in $\{0,1\}^n$.*

This result can be used to show:

Theorem 16. *$2^{\Omega(n^{1/4d})}$ lower bound for $\mathrm{MAJ} \circ \mathrm{AC}^0$ circuits computing XOR_n.*

3.2 ℓ_∞-approximation

The ℓ_∞-*approximate degree* of a boolean function f, denoted $\widetilde{\deg}(f)$, is the minimum degree of a real polynomial $\widetilde{f} \in \mathbb{R}[x_1, \ldots, x_n]$ such that $|f(x) - \widetilde{f}(x)| \leq 1/3$ for all $x \in \{0,1\}^n$. In Lecture 2, we mentioned the result $O(\widetilde{\deg}(f)) \leq O(\sqrt{\mathcal{L}(f)})$, which was proved via quantum query complexity. Using this bound, Tal gives an elegant proof of the shrinkage property for DeMorgan formulas: $\mathbb{E}[\ \mathcal{L}(f\restriction \mathbf{R}_p)\ ] = O(p^2\mathcal{L}(f) + 1)$.

ℓ_∞-approximate degree has also been studied for functions in AC^0. It is an open question there is any AC^0 function $f : \{0,1\}^n \to \{0,1\}$ with $\widetilde{\deg}(f) = \Omega(n)$. For depth-$d$ functions, the best construction (of Bun and Thaler) has degree $n^{1-1/2^{O(d)}}$; see also work of Sherstov on this question.

3.3 ℓ_2-approximation

Perhaps the most natural and well-studied polynomial approximation is under the ℓ_2-norm. Here is again natural to consider boolean function with range $\{1,-1\}$ instead of $\{0,1\}$. For every boolean function $f : \{0,1\}^n \to \{1,-1\}$, there is a unique multilinear real polynomial $\widetilde{f}(x) = \sum_{I \subseteq [n]} c_I \prod_{i \in I} x_i$ that agrees with f over $\{-1,1\}^n$. (Coefficients $c_I \in \mathbb{R}$ are called the *Fourier coefficients* of f.)

For any $0 \leq k \leq n$, we may truncate this polynomial to its degree-k part:

$$\widetilde{f}_k(x) := \sum_{I \subseteq [n] \,:\, |I| \leq k} c_I \prod_{i \in I} x_i.$$

Among all degree-k polynomials, this polynomial $\widetilde{f}_k$ minimizes the ℓ_2-distance from f over $\{0,1\}^n$.

One notable result about AC^0 functions is the following:

Theorem 17 (Linial-Mansour-Nisan '93, improvement by Tal '17)**.** *If f is computable by an AC^0 circuit of depth $d+1$ and size S, then for all k,*

$$\mathop{\mathbb{E}}_{x \in \{1,-1\}^n}[\ (f(x) \neq \widetilde{f}_k(x))^2\ ] \leq \exp\left(1 - \frac{k}{O(\log S)^d}\right).$$

The proof involves the random restriction $\mathbf{R}_p$, which interacts nicely with the Fourier coefficients of f. A fairly straightforward proof of Theorem 17 can be derived from the switching-lemma like bound:

Theorem 18. *Suppose f is computable by an AC^0 circuit of depth $d+1$ and size S, and let $p = 1/O(\log S)^d$. Then for all $\ell > 0$,*

$$\mathbb{P}[\ \deg(f\restriction R_p) \geq \ell\ ] \leq \mathbb{P}[\ \mathrm{DT}_{\mathrm{depth}}(f\restriction R_p) \geq \ell\ ] \leq \frac{1}{2^\ell}.$$

(The second inequality is Theorem 3 from Lecture 5 without additional the error term $\frac{1}{S}$.)