Lecture 17 Circuit lower bounds

$\exists$ Oracle A, $P^A \neq NP^A$
$\exists$ Oracle B, $P^B = NP^B$ ($B = TQBF$)
} Relativization barrier

$\Downarrow$

Limitation to diagonalization

<u>Thm</u>: PARITY $\in AC^0$ [Furst – Saxe – Sipser; Yao; Håstad 86]

$\Downarrow$ depth d requires size $2^{\Omega(n^{\frac{1}{d-1}})}$

(AC^0: poly size constant depth circuits, unbounded fan-in AND/OR gates, NOT gates at the leaves)

Remark: Recently [Rossman'15] depth-d Formulae need size $2^{\Omega(d n^{\frac{1}{d-1}})}$

Today we will prove in the way as [Razborov – Smolensky 87']

~~a~~ algebraic ~~we~~ approach, "method of approximation", which gives a weaker bound $2^{\Omega(n^{\frac{1}{2d}})}$

<u>High level idea</u>:

- Prove small AC^0 circuits are "simple" (in some sense)
- Prove PARITY is not simple

<u>Def</u>: algebraic degree

Every function $f: \{0,1\}^n \to \mathbb{R}$ can be ~~uniquely~~ uniquely expressed as a polynomial of degree $\leq n$ in $x_1, x_2, \cdots, x_n$ with real coefficients.

$$f(x_1, \cdots, x_n) = a_0 + \sum a_i x_i + \sum a_{ij} x_i x_j + \cdots$$

$$f(x_1, \cdots, x_n) = f(1,1,\cdots,1)\, x_1 x_2 \cdots x_n + f(0,1,\cdots,1)(1-x_1)\, x_2 \cdots x_n + \cdots f(0,0,\cdots,0)(1-x_1)(1-x_2)\cdots(1-x_n)$$

$\deg(f) :=$ the degree of the multilinear polynomial that expresses f.

Example:

$AND(x_1, \dots, x_n) = x_1 x_2 \cdots x_n$ $\qquad deg(AND) = n$

$OR(x_1, \dots, x_n) = 1 - \prod_{i=1}^{n} (1 - x_i)$ $\qquad deg(OR) = n$

$\oplus(x_1, \dots, x_n) = \frac{1}{2}(1 - \prod_{i=1}^{n} (1 - 2x_i))$ $\qquad deg(\oplus) = n$

Notice that linear transform does not change the degree,

If we define $\oplus$ on $\{-1, 1\}^n \to \{-1, 1\}$, $\oplus(y_1, \dots, y_n) = y_1 \cdots y_n$, where $y_i = 1 - 2x_i$

Approximate degree

For boolean function $f: \{0,1\}^n \to \{0,1\}$, we say polynomial g $\frac{3}{4}$-approx f if polynomial $g: \mathbb{R}^n \to \mathbb{R}$ satisfies $g(x) = f(x)$ for at least $\frac{3}{4}$ inputs on $\{0,1\}^n$

define: $deg_{\frac{3}{4}}(f) := \min_{g \ \frac{3}{4}\text{-approx } f} deg(g)$

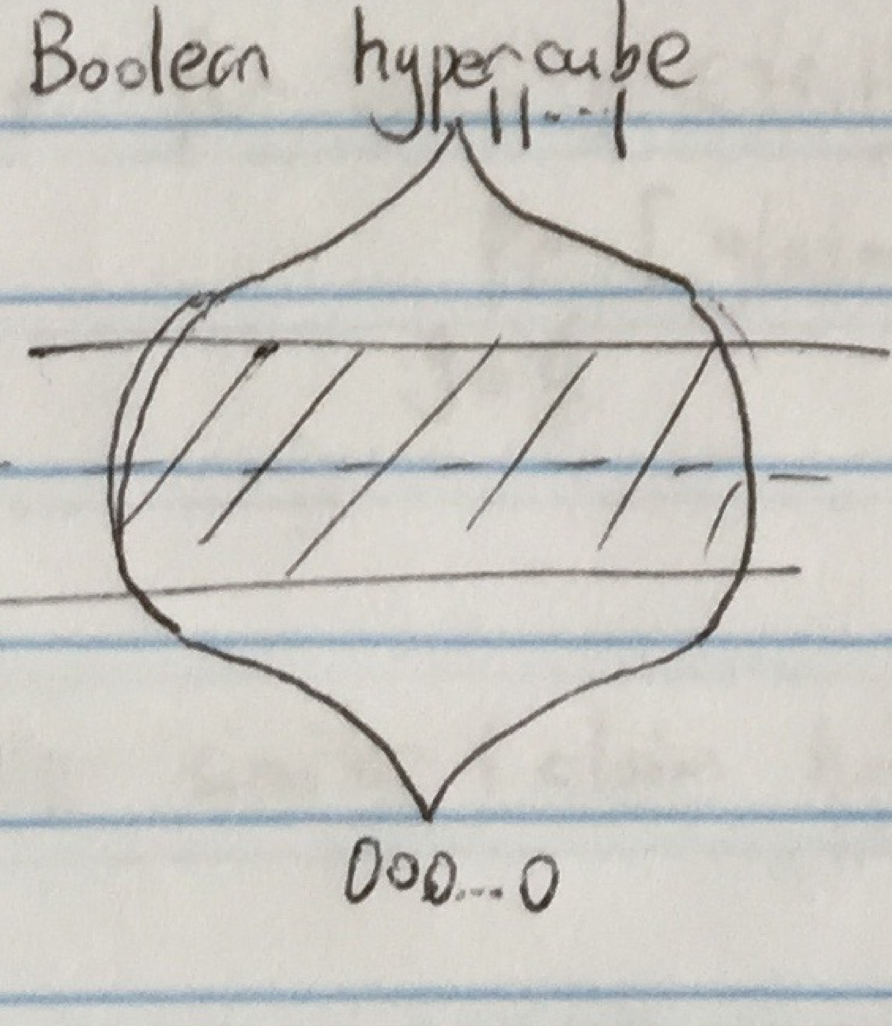
Proof of $S \geq 2^{\Omega(n^{\frac{1}{4d}})}$ will follow from:

Lemma 1: $deg_{\frac{3}{4}}(\oplus_n) \geq \Omega(\sqrt{n})$

Lemma 2: If f is computed by size S, depth d, AC^0 circuit,

$$deg_{\frac{3}{4}}(f) \leq O((\log S)^{2d})$$

In the Boolean hypercube



$\sum x_i \in [\frac{n}{2} - c\sqrt{n}, \frac{n}{2} + c\sqrt{n}]$ w.h.p.

This shows the tightness of Lemma 1.

Pf of Lemma 1: (In $\{-1,1\}$ notation)

Let $g: \mathbb{R}^n \to \mathbb{R}$ agrees with $\oplus_n$ on $\geq \frac{3}{4} \cdot 2^n$ inputs,

where $\oplus_n : \{-1,1\}^n \to \{-1,1\}$, and $\deg(g) = t$

Define $A = \{y \in \{-1,1\}^n \mid g(y) = \oplus_n(y)\}$, $|A| \geq \frac{3}{4} \cdot 2^n$

Define $H = \{h : A \to \mathbb{R}\}$, $\dim(H) = |A| \geq \frac{3}{4} \cdot 2^n$.

We can write $h(y_1, \dots, y_n) = \sum_{T \subseteq [n]} \hat{h}(T) \cdot \prod_{i \in T} y_i$

Notice on set A

$$\prod_{i \in T} y_i = \left(\prod_{i=1}^{n} y_i\right)\left(\prod_{i \notin T} y_i\right) = g(y) \prod_{i \notin T} y_i$$

Either $\prod_{i \in T} y_i \leq \frac{n}{2} + t$

or $g(y) \prod_{i \notin T} y_i \leq t + (n - \frac{n}{2} - t) = \frac{n}{2} \leq \frac{n}{2} + t$

Therefore each $h \in H$ has $\deg(h) \leq \frac{n}{2} + t$

There fore

$$\frac{3}{4} \cdot 2^n \leq \dim(H) \leq \sum_{i=0}^{\frac{n}{2}+t} \binom{n}{i}$$

$$\Leftrightarrow \frac{1}{4} \cdot 2^n \leq \sum_{i=\frac{n}{2}}^{\frac{n}{2}+t} \binom{n}{i} \leq (t+1)\binom{n}{n/2} \leq O\left(\frac{t \cdot 2^n}{\sqrt{n}}\right)$$

$$\Leftrightarrow t \geq \Omega(\sqrt{n})$$

Claim: $\forall\, m, \varepsilon$, $\exists$ a distribution $\mathcal{G}$ on $\{0,1\}^m \to \mathbb{R}$ of degree $O(\log m \log \frac{1}{\varepsilon})$

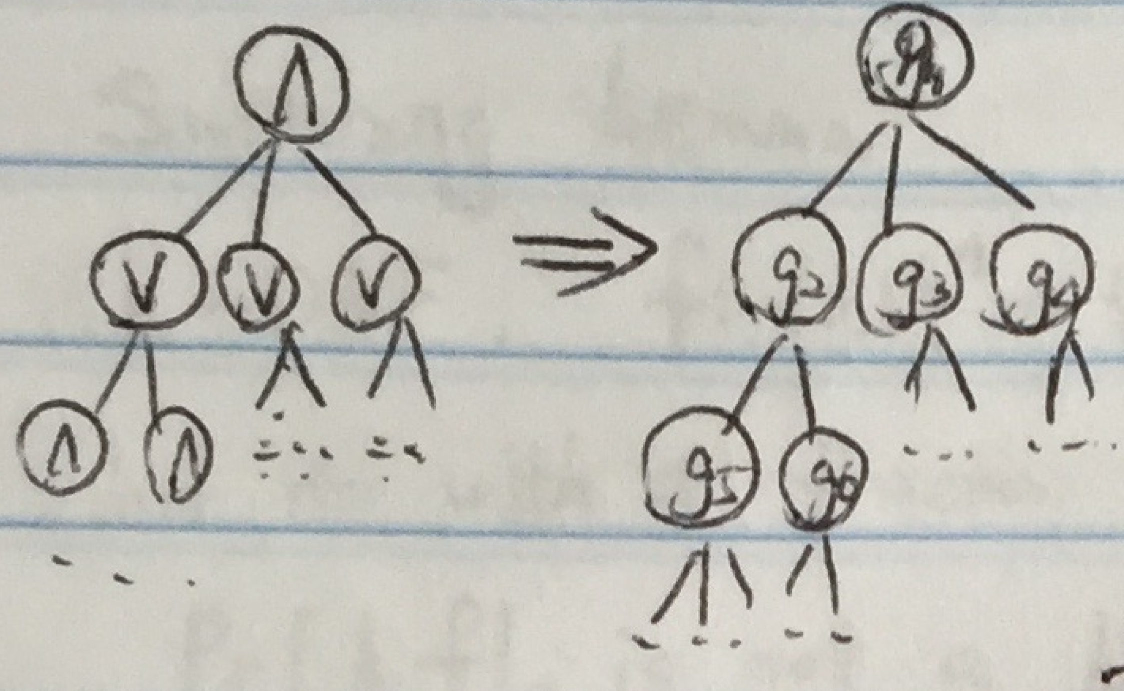
s.t. for each $a \in \{0,1\}^m$

$$\Pr_{g \in \mathcal{G}} [g(a) = OR(a)] \geq 1 - \varepsilon$$

Note: similar claim holds for AND by de Morgan's law

Pf of Lemma 2:

Sample a random polynomial for each gate from the distribution in the Claim.
If each polynomial has degree at most L, the whole polynomial G has degree at most L^d.



$$G = g_1(g_2(g_5(\dots), g_6(\dots)), g_3(\dots), g_4(\dots))$$

By union bound, if $\varepsilon = \frac{1}{4S}$, $\overbrace{\Pr_{G\sim\mathcal{G}^S}}^{\text{for any } x}[f(x) \neq G(x)] \le \varepsilon \cdot S = \frac{1}{4}$

$$\Downarrow$$

$$\exists G, \ \Pr_x[f(x) \neq G(x)] \le \frac{1}{4}$$

And also $m \le S$.

Therefore $L \le O(\log m \log \frac{1}{\varepsilon}) = O(\log^2 S)$

$$\deg_{\frac{3}{4}}(f) \le L^d = O((\log S)^{2d})$$

Pf of Claim:

$OR(x_1, \dots, x_n) = 1 - \prod_{i=1}^{n} (1 - x_i)$

We generate a random family $\mathcal{S}$, $g_{\mathcal{S}} = 1 - \prod_{S \in \mathcal{S}} (1 - \sum_{i \in S} x_i)$ $\deg(g_{\mathcal{S}}) = |\mathcal{S}|$

If $x_1 = x_2 = \dots = x_n = 0$, $g_{\mathcal{S}} = 0$ always

Fix any a s.t. $OR(a_1, \dots, a_n) = 1$

Let $T = \{i \mid a_i = 1\}$, we know $T \neq \emptyset$, suppose $2^k \le |T| \le 2^{k+1}$,

Repeat $O(\log \frac{1}{\varepsilon})$ times, each time pick each $i \in [n]$ w.p. $\frac{1}{2^k}$.

Then $\Pr[g(a) = OR(a)] \ge 1 - \varepsilon$.

~~For diffide~~ There are $\log m$ different choices of k among all a.

Therefore we will ~~regpeat~~ repeat $\log m$ times for different size of k

$$\deg(g_{\mathcal{S}}) = |\mathcal{S}| = O(\log m \cdot \log \frac{1}{\varepsilon})$$

Original Proof: [Håstad Switching Lemma]

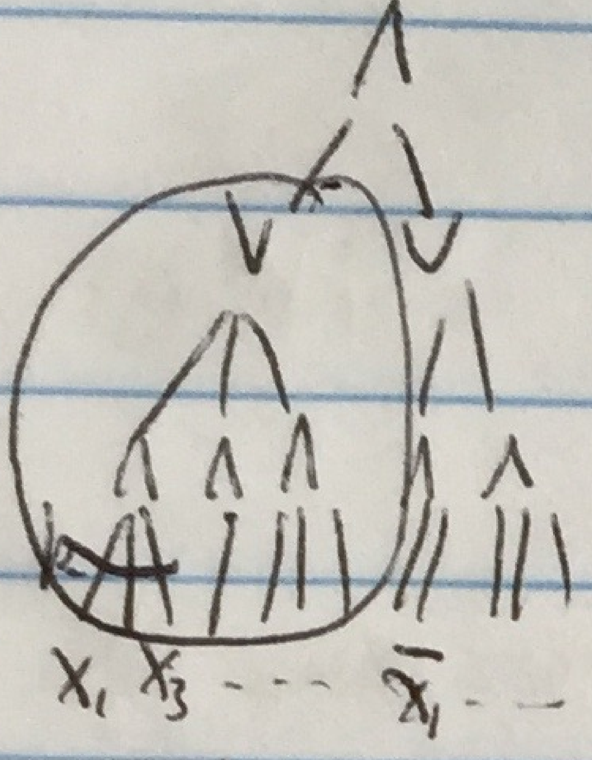
Hard property for $\oplus^n$:

Even if one restricts $n-n^{\varepsilon}$ variables arbitrarily, $\oplus$ is still undetermined.

It is still a parity on fewer variables (n^{ε}).

AC^0:

Assume bottom fan-in is k.



Random restriction

$$x_i = \begin{cases} 0 & \text{w.p. } \frac{1-p}{2} \\ 1 & \text{w.p. } \frac{1-p}{2} \\ * & \text{w.p. } p \quad \text{(remain them)} \end{cases}$$

Assign $(1-p)m$ variables with random 0/1,

remain pm variables

→ After restriction, ~~th~~ DNF can also be expressed to a small CNF.

Then we shrink the depth by 1.

Switching Lemma: f be a k-DNF formula, let ρ be a random p-restriction.

$$\Pr[f|_{\rho} \text{ is a } t\text{-CNF}] > 1-(7pk)^t$$

choose $t = O(\log S)$, $k = \Theta(\log S)$, $p = \Theta\left(\frac{1}{\log S}\right)$