COMS 6998: Unconditional Lower Bounds and Derandomization **Spring 2024**

Lecture 10: March 26, 2024

Lecturer: Rocco Servedio *Scribes: Szymon Snoeck and Patrick Yeh*

Last Time

We previously went over results from the [BRS91] and [LMN93] papers, with the overall goal to combine the best of both results using the [Bra08] construction.

Today

1 $[L, M, N]$ $\mathbf{L}_2$ Approximator for $\mathbf{AC}^0$

Theorem 1. *Let $f \in AC^0_{s,d}$, then there is a real polynomial p_2 of degree $O((\log(\frac{S}{d})^d)$ such that:*

$$\mathbb{E}_{\mathbf{x}\sim\mathcal{U}}[(f(\mathbf{x}) - p_2(\mathbf{x}))^2] \leq \epsilon.$$

Continuing the proof from the previous class, all that is left to show is **lemma 2**. Recall from the previous class, that a random restriction $\rho \sim \mathcal{R}_p$ can be written as $(\mathbf{J}, \mathbf{z})$ where $\mathbf{J}$ are the variables that remain unrestricted (i.e. are $*$'s) and $\mathbf{z}$ is the assignment of the restricted variables. Furthermore, $f \upharpoonright \rho = f_{\mathbf{J},\mathbf{z}}$.

Lemma 2. *For all $f : \{+1, -1\}^n \to \{+1, -1\}$ and any $p \leq \frac{1}{10}$, it holds that:*

$$W^{\geq t/p}(f) = 2 \cdot \mathbb{E}_{(\mathbf{J},\mathbf{z})\sim\mathcal{R}_p}[W^{\geq t}(f_{\mathbf{J},\mathbf{z}})].$$

To prove **lemma 2**, we prove a sequence of compounding claims:

Claim 3. *Fix some $(\mathbf{J}, \mathbf{z}) \sim \mathcal{R}_p$, then for all $f : \{+1, -1\}^n \to \{+1, -1\}$ and any $S \subseteq [n]$:*

$$\widehat{f_{\mathbf{J},\mathbf{z}}}(s) = \begin{cases} 0 & \text{if } S \not\subseteq \mathbf{J} \\ \sum_{T\subseteq \mathbf{J}^c} \widehat{f}(S \cup T)\chi_T(\mathbf{z}) & \text{if } S \subseteq \mathbf{J} \end{cases} \tag{1}$$

Proof. We can view $f_{\mathbf{J},\mathbf{z}}$ as a $|\mathbf{J}|$-junta over n-variables. In other words, $f_{\mathbf{J},\mathbf{z}}(x) = f(x_{\mathbf{J}}, \mathbf{z})$. Suppose $S \not\subseteq \mathbf{J}$, then S contains an irrelevant variable for $f_{\mathbf{J},\mathbf{z}}$ so $\widehat{f_{\mathbf{J},\mathbf{z}}}(S) = \mathbb{E}[f_{\mathbf{J},\mathbf{z}} \cdot \chi_S] = 0$. Alternatively suppose $S \subset \mathbf{J}$, then:

$$f_{\mathbf{J},\mathbf{z}}(x) = f(x_{\mathbf{J}}, \mathbf{z}) = \sum_{R \subseteq [n]} \widehat{f}(R)\chi_R(x_{\mathbf{J}}, \mathbf{z}) = \sum_{S \subseteq \mathbf{J}} \sum_{T \subseteq \mathbf{J}^c} \widehat{f}(S \cup T)\chi_S(x_{\mathbf{J}}) \cdot \chi_T(\mathbf{z})$$
$$= \sum_{S \subseteq \mathbf{J}} \chi_S(x_{\mathbf{J}}) \cdot \sum_{T \subseteq \mathbf{J}^c} \widehat{f}(S \cup T)\chi_T(\mathbf{z}).$$

Hence $\widehat{f_{\mathbf{J},\mathbf{z}}}(s) = \sum_{T \subseteq \mathbf{J}^c} \widehat{f}(S \cup T)\chi_T(\mathbf{z})$. ■

Building on the above claim, the following sequence of claims can be proven:

Claim 4. *For fixed $\mathbf{J}, S \subseteq [n]$ and uniformly random $\mathbf{z} \sim \{+1, -1\}^{\mathbf{J}^c}$, it holds that:*

$$\mathbb{E}_{\mathbf{z} \sim \{+1,-1\}}[\widehat{f_{\mathbf{J},\mathbf{z}}}(S)] = \mathbf{1}[S \subseteq \mathbf{J}]\widehat{f}(S)$$
$$\mathbb{E}_{\mathbf{z} \sim \{+1,-1\}}[\widehat{f_{\mathbf{J},\mathbf{z}}}(S)^2] = \mathbf{1}[S \subseteq \mathbf{J}] \sum_{T \subseteq \mathbf{J}^c} \widehat{f}(S \cup T)^2.$$

Claim 5. *For $(\mathbf{J}, \mathbf{z}) \sim \mathcal{R}_p$, it holds that:*

$$\mathbb{E}_{(\mathbf{J},\mathbf{z}) \sim \mathcal{R}_p}[\widehat{f_{\mathbf{J},\mathbf{z}}}(S)] = p^{|S|}\widehat{f}(S)$$
$$\mathbb{E}_{(\mathbf{J},\mathbf{z}) \sim \mathcal{R}_p}[\widehat{f_{\mathbf{J},\mathbf{z}}}(S)^2] = \sum_{T \subseteq [n]} \widehat{f}(T)^2 \cdot \mathbb{P}_{\mathbf{J}}[T \cap \mathbf{J} = S].$$

Claim 6.

$$\mathbb{E}_{(\mathbf{J},\mathbf{z}) \sim \mathcal{R}_p}[W^{\geq k}(f_{\mathbf{J},\mathbf{z}})] = \sum_{r \geq k} W^r(f) \cdot \mathbb{P}[Bin(r, p) \geq k]$$

where $Bin(r, p)$ is a binomial random variable with r trials and probability p of success.

Now we can finally prove **lemma 2**:

Proof. **Claim 4** gives us that:

$$\mathbb{E}_{(\mathbf{J},\mathbf{z}) \sim \mathcal{R}_p}[W^{\geq k}(f_{\mathbf{J},\mathbf{z}})] \geq \sum_{r \geq k/p} W^r(f) \cdot \mathbb{P}[\text{Bin}(r, p) \geq k].$$

For each $r \geq k/p$, we have $\mathbb{P}[\text{Bin}(r, p) \geq k] \geq \frac{1}{2}$. Thus:

$$\mathbb{E}_{(\mathbf{J},\mathbf{z}) \sim \mathcal{R}_p}[W^{\geq k}(f_{\mathbf{J},\mathbf{z}})] \geq \frac{1}{2} \sum_{r \geq k/p} W^r(f) = \frac{1}{2} W^{\geq k/p}(f).$$

■

This concludes the proof of **theorem 1**.

2 Proof of Braverman's Theorem

In 2010, Braverman proved the following theorem:

Theorem 7. *(Braverman's Theorem) Let* $k = (\log \frac{S}{\epsilon})^{O(d^2)}$ *and* $\mathcal{D}$ *be any* k*-wise independent random variable over* $\{0,1\}^n$*, then* $\mathcal{D}$ ϵ*-fools* $AC^0_{s,d}$

Note that the state of the art result gives $k = \log(S)^{O(d)} \log(\frac{1}{\epsilon})$. Before proving Braverman's Theorem, we must first improve the BRS (i.e. pointwise) approximator from the last lecture:

Theorem 8. *Let* $f \in AC^0_{s,d}$*. Consider any* $\mathcal{D}$ *over* $\{0,1\}^n$*. There exists a real-valued polynomial, p, such that:*

i) $\mathbb{P}_{\mathbf{x}\sim\mathcal{D}}[p(\mathbf{x}) = f(\mathbf{x})] \geq 1 - \epsilon$

ii) $deg(p) \leq (\log \frac{S}{\epsilon})^{O(d)}$

iii) $\forall x \in \{0,1\}^n$, $|p(x)| \leq \exp((\log \frac{S}{\epsilon})^{O(d)})$

iv) There exists a circuit $E \in AC^0_{poly(s),d+O(1)}$ *such that* $E(x) = 0 \implies p(x) = f(x)$ *(i.e.* $p(x) \neq f(x) \implies E(x) = 1$*) and* $\mathbb{P}_{\mathbf{x}\sim\mathcal{D}}[E(\mathbf{x}) = 1] \leq \epsilon$*.*

Proof. Items *i), ii), iii)* were proved last class hence it suffices to show *iv)*. The idea is that E functions as an indicator of when something went wrong during the building of the polynomial p. Consider some fixed OR-gate circuit $g = g_1 \vee \cdots \vee g_t$ where $t \leq s$. Our polynomial approximation is $p(g_1, \cdots, g_t) = 1 - \prod_{i=1}^{\text{polylog}(t/\epsilon)}(1 - \sum_{j\in S_i} g_j)$ where $S_i \subseteq [t]$. Thus, $p(g_1, \cdots, g_t) \neq g(g_1, \cdots, g_t)$ is true only if:

1) At least one $g_1, \cdots, g_t$ is equal to one.

2) Each set $\{g_j \mid j \in S_i\}$ does not contain 1 or contains ≥ 2 1's.

Fortunately, these conditions can be checked with constant depth:

$$E' = \bigvee_{1\leq a<b\leq t} g_a \wedge g_b.$$

E' is satisfied if and only if any two $g_1, \cdots, g_t$ are 1. Thus indeed, if $E' = 0$ then $p(x) = g(x)$. Repeat this process for each gate in f and OR the results together the results to get E. To finish the proof, note that in order to obtain the bound $\mathbb{P}_{\mathbf{x}\sim\mathcal{D}}[p(\mathbf{x}) \neq f(\mathbf{x})] \leq \epsilon$ during the previous class, it was proved that $\mathbb{P}_{\mathbf{x}\sim\mathcal{D}}[E(\mathbf{x}) = 1] \leq \epsilon$. ■

Before we dive into proving Braverman's theorem, recall the sandwiching lemma from 2 classes ago:

Theorem 9. *A function $f : \{0,1\}^n \to \{0,1\}$ is ϵ-fooled by a k-wise independent distribution if there exists an "ϵ-sandwiching" by real polynomials $q_l, q_u : \{0,1\}^n \to \mathbb{R}$ of degree at most k such that:*

i) $q_l \leq f \leq q_u$

ii) $\mathbb{E}_{\mathbf{x}\sim\mathcal{U}}[q_u(\mathbf{x}) - q_l(\mathbf{x})] \leq \epsilon$

Observe that since $\text{AC}^0_{s,d}$ is closed under negation, to show the above, it is sufficient to show there exists q_l such that $q_l \leq f$ and $\mathbb{E}[f - q_l] \leq \frac{\epsilon}{2}$. Indeed this implies $q_u = 1 - q_l$ is a valid upper sandwich real polynomial.

Another key observation is that it is sufficient enough to provide a q_l, such that it can depend on the particular k-wise distribution $\mathcal{D}$ and also is a lower sandwich polynomial for a function f' that is "close" to f under both $\mathcal{D}$ and $\mathcal{U}$.

Keeping this in mind, we will consider the following lemma:

Lemma 10. *Suppose for every k-wise distribution $\mathcal{D}$ there exists a boolean function f' and a degree-k polynomial q_l such that:*

i) $\mathbb{P}_{\mathbf{x}\sim\mathcal{D}}[f(\mathbf{x}) \neq f'(\mathbf{x})] \leq \frac{\epsilon}{3}$ *and* $\mathbb{P}_{\mathbf{x}\sim\mathcal{U}}[f(\mathbf{x}) \neq f'(\mathbf{x})] \leq \frac{\epsilon}{3}$

ii) $q_l \leq f$ *and* $\mathbb{E}_{\mathbf{x}\sim\mathcal{U}}[f(\mathbf{x}) - q_l(\mathbf{x})] \leq \frac{\epsilon}{3}$

Then $\mathbb{E}[f(\mathcal{U})] - \mathbb{E}_{\mathbf{x}\sim\mathcal{D}}[f(\mathbf{x})] \leq \epsilon$. Combining this with the version for q_u, this means that $\mathcal{D}$ ϵ-fools f.

Proof. Condition *i)* and the fact that $q_l \leq f$ gives us the following bound:

$$\mathbb{E}_{\mathbf{x}\sim\mathcal{D}}[f(\mathbf{x})] \geq \mathbb{E}_{\mathbf{x}\sim\mathcal{D}}[f'(\mathbf{x})] - \frac{\epsilon}{3} \geq \mathbb{E}_{\mathbf{x}\sim\mathcal{D}}[q_l(\mathbf{x})] - \frac{\epsilon}{3}$$

Since $\mathcal{D}$ is a k-wise independent distribution and $deg(q_l) \leq k$, we get that the righthand side of the inequality above is equivalent to:

$$\mathbb{E}_{\mathbf{x}\sim\mathcal{U}}[q_l(\mathbf{x})] - \frac{\epsilon}{3}$$

By the second half of condition *ii)*, and then applying condition *i)*, we get:

$$\mathbb{E}_{\mathbf{x}\sim\mathcal{U}}[q_l(\mathbf{x})] - \frac{\epsilon}{3} \geq \mathbb{E}_{\mathbf{x}\sim\mathcal{U}}[f'(\mathbf{x})] - \frac{2\epsilon}{3} \geq \mathbb{E}_{\mathbf{x}\sim\mathcal{U}}[f(\mathbf{x})] - \epsilon$$

Applying the same result to $q_u = 1 - q_l$, we get that $\mathcal{D}$ ϵ-fools f. ■

As a result, to finish our proof of Braverman's theorem, it suffices to show the following lemma:

Lemma 11. *Let* $f \in AC^0_{s,d}$*,* $k = \left(O\left(\log \frac{s}{\epsilon}\right)\right)^{O(d^2)}$*, and* $\mathcal{D}$ *be a* k*-wise independent distribution. Then there exists a boolean function* f' *and a degree-*k *polynomial* q_l *such that the following conditions both hold:*

i) $\mathbb{P}_{\mathbf{x}\sim\mathcal{D}}[f(\mathbf{x}) \neq f'(\mathbf{x})] \leq \frac{\epsilon}{3}$ *and* $\mathbb{P}_{\mathbf{x}\sim\mathcal{U}}[f(\mathbf{x}) \neq f'(\mathbf{x})] \leq \frac{\epsilon}{3}$

ii) $q_l \leq f$ *and* $\mathbb{E}_{\mathbf{x}\sim\mathcal{U}}[f(\mathbf{x}) - q_l(\mathbf{x})] \leq \frac{\epsilon}{3}$

Proof. Apply the BRS approximator to f using the distribution $\frac{1}{2}(\mathcal{D}+\mathcal{U})$ and error parameter $\frac{\epsilon}{8}$. This gives a polynomial p_0 such that

a) $\mathbb{P}_{\mathbf{x}\sim\mathcal{D}}[p_0(\mathbf{x}) \neq f(\mathbf{x})] \leq \frac{\epsilon}{4}$

b) $\mathbb{P}_{\mathbf{x}\sim\mathcal{U}}[p_0(\mathbf{x}) \neq f(\mathbf{x})] \leq \frac{\epsilon}{4}$

We are also given a $poly(s)$-size, $d + O(1)$-depth error-detecting circuit E such that

a) $f(\mathbf{x}) \neq p_0(\mathbf{x}) \implies E(\mathbf{x}) = 1$

b) $\mathbb{P}_{\mathbf{x}\sim\mathcal{D}}[E(\mathbf{x}) = 1] \leq \frac{\epsilon}{4}$ and $\mathbb{P}_{\mathbf{x}\sim\mathcal{U}}[E(\mathbf{x}) = 1] \leq \frac{\epsilon}{4}$

We will now apply the LMN result on E. Let $p_{E,2}$ be the polynomial of degree $\left(\log\left(\frac{s}{\delta}\right)\right)^{O(d)}$ for some value δ that will be fixed later such that

$$\mathbb{E}_{\mathbf{u}\sim\mathcal{U}}\left[(E(\mathbf{u}) - p_{E,2}(\mathbf{u}))^2\right] \leq \delta$$

For the actual construction, set $f' = f \vee E$, $q = p_0(1-p_{E,2})$, and our desired polynomial $p = q_l = 1 - (1-q)^2$. We will show that f' and p satisfy conditions *i)* and *ii)*.

The intuition is that since the error region is small, f' is close to f. Moreover, p_0 may make wild errors on f' when $E(x) = 1$, but we will tame this error by multiplying by $1 - p_{E,2}$. However, $q = p_0(1 - p_{E,2})$ may not be a lower sandwiching polynomial when $f'(x) = 1$. Thus, we let $p = 1 - (1-q)^2 \leq 1$ to force our polynomial to be a lower sandwiching polynomial. (Please refer to Figure 1 for further intuition.)

To show condition *i)* holds, note that $f'(x) \neq f(x)$ only if $E(x) = 1$. Furthermore, under both distributions $\mathcal{D}$ and $\mathcal{U}$, we have $\mathbb{P}_{\mathbf{x}\sim\mathcal{D}}[E(\mathbf{x}) = 1] \leq \frac{\epsilon}{4}$ and $\mathbb{P}_{\mathbf{x}\sim\mathcal{U}}[E(\mathbf{x}) = 1] \leq \frac{\epsilon}{4}$. Thus, condition *i)* is satisfied.

For condition *ii)* to hold, we will prove two claims:

Claim 12. *If* $f'(x) = 0$*, then* $q(x) = 0$*.*

Proof. If $f'(x) = f(x) \vee E(x) = 0$, then $E(x) = 0$. In other words, we are not in the error region, so $p_0(x) = f(x) = 0$ and $q(x) = 0$. ■

Claim 13. *Let* $\delta = \epsilon \cdot \exp\left(-\log\left(\frac{s}{\epsilon}\right)\right)^{O(d)}$ *such that*

$$\|f' - q\|_2 \leq \sqrt{\frac{\epsilon}{4}} + \exp\left(-\log\left(\frac{s}{\epsilon}\right)\right)^{O(d)} \cdot \sqrt{\delta} \leq \sqrt{\frac{\epsilon}{3}}$$

Proof. Recall that for functions $a, b : \{+1, -1\}^n \to \mathbb{R}$, $\|a-b\|_2 = \mathbb{E}_{x\sim\mathcal{U}}\left[(a(x) - b(x))^2\right]^{\frac{1}{2}}$. We will make use of the triangle inequality:

$$\|f' - q\|_2 \leq \|f' - p_0(1 - E)\|_2 + \|p_0(1 - E) - q\|_2$$

For the first term, by considering the two cases $E(x) = 0$ (and recognizing that this case provides zero contribution) and $E(x) = 1$, we get

$$\|f' - p_0(1 - E)\|_2 \leq \sqrt{\mathbb{P}_{\mathbf{x}\sim\mathcal{U}}[E(\mathbf{x}) = 1]} \leq \sqrt{\frac{\epsilon}{4}}$$

For the second term, we can write $p_0(1 - E) - q = p_0(p_{E,2} - E)$. By applying our pointwise bound on p_0, we have (from the previous lecture's results):

$$\max_x |p_0(x)| \leq \exp\left(\log\left(\frac{s}{\epsilon}\right)\right)^{O(d)}$$

As a result, we get

$$\|p_0(1 - E) - q\|_2 \leq \exp\left(\log\left(\frac{s}{\epsilon}\right)\right)^{O(d)} \cdot \|p_{E,2} - E\|_2$$

The LMN result gives the bound $\|p_{E,2} - E\|_2 \leq \sqrt{\delta}$, so we find

$$\|p_0(1 - E) - q\|_2 \leq \exp\left(\log\left(\frac{s}{\epsilon}\right)\right)^{O(d)} \cdot \sqrt{\delta}$$

This gives us our desired bound:

$$\|f' - q\|_2 \leq \sqrt{\frac{\epsilon}{4}} + \exp\left(\log\left(\frac{s}{\epsilon}\right)\right)^{O(d)} \cdot \sqrt{\delta}$$

■

Given the previous two claims, we show that condition *ii)* naturally follows. We will first verify that $p = q_l = 1 - (1-q)^2$ is indeed a lower sandwiching polynomial for f'. We know that $p < f'$ pointwise. Consider the two cases $f'(x) = 0$ and $f'(x) = 1$.

If $f'(x) = 0$, then by Claim 1 $q(x) = 0$, so $p(x) = 0 \leq f'(x)$.

Otherwise, if $f'(x) = 1$, then $f'(x) - p(x) = (1 - q(x))^2 = (f'(x) - q(x))^2$. We now get the following result, showing that p is indeed a lower sandwicher in this case:

$$\mathbb{E}_{\mathbf{x}\sim\mathcal{U}}\left[f'(\mathbf{x}) - p(\mathbf{x})\right] = \mathbb{E}_{\mathbf{x}\sim\mathcal{U}}\left[|f'(\mathbf{x}) - p(\mathbf{x})|\right] \leq \mathbb{E}_{\mathbf{u}\sim\mathcal{U}}\left[\left(f'(\mathbf{u}) - q(\mathbf{u})\right)^2\right] \leq \frac{\epsilon}{3}$$

Finally, we will verify that $deg(p) \leq k = \left(O\left(\log\frac{s}{\epsilon}\right)\right)^{O(d^2)}$. From our polynomials $q = p_0(1 - p_{E,2})$ and $p = q_l = 1 - (1-q)^2$, we get

$$deg(p) \leq 2 \cdot (deg(p_0) + deg(p_{E,2}))$$

We then apply the BRS and LMN results to get

$$deg(p) \leq 2 \cdot \left(\log\left(\frac{s}{\epsilon}\right)^{O(d)} + \log\left(\frac{s}{\delta}\right)^{O(d)}\right) = 2 \cdot \left(\log\left(\frac{s}{\epsilon}\right)^{O(d)} + \log\left(\frac{s}{\epsilon}\right)^{O(d^2)}\right)$$

This finishes the proof of our lemma and consequently Braverman's theorem. ■

3 Introduction to Linear Threshold Functions

Definition 14. *(Linear Threshold Functions) A function $f : \{+1,-1\}^n \to \{+1,-1\}$ is a linear threshold function (LTF) if $f(x) = sign(w \cdot x - \theta)$ for some $w \in \mathbb{R}^n$, $\theta \in \mathbb{R}$.*

Intuitively, an LTF is a hyperplane that divides $\mathbb{R}^n$ into half-spaces that separate the set of vectors for which $f(x) = +1$ from the set of vectors where $f(x) = -1$.

One notable example of an LTF is the majority function:

$$\text{MAJ}(x_1, ..., x_n) = sign\left(\sum_{i=1}^{n} x_i\right)$$

Note that lower bounds for LTFs are trivial. For example, the parity function $\text{PAR}(x_1, x_2)$ is not computable by any LTF since there does not exist a hyperplane that cleanly separates the preimage of $\{+1\}$ from that of $\{-1\}$.

As a result, we will focus our attention of pseudorandom generators (PRGs) and deterministic approximate counting for LTFs. An interesting question is whether we

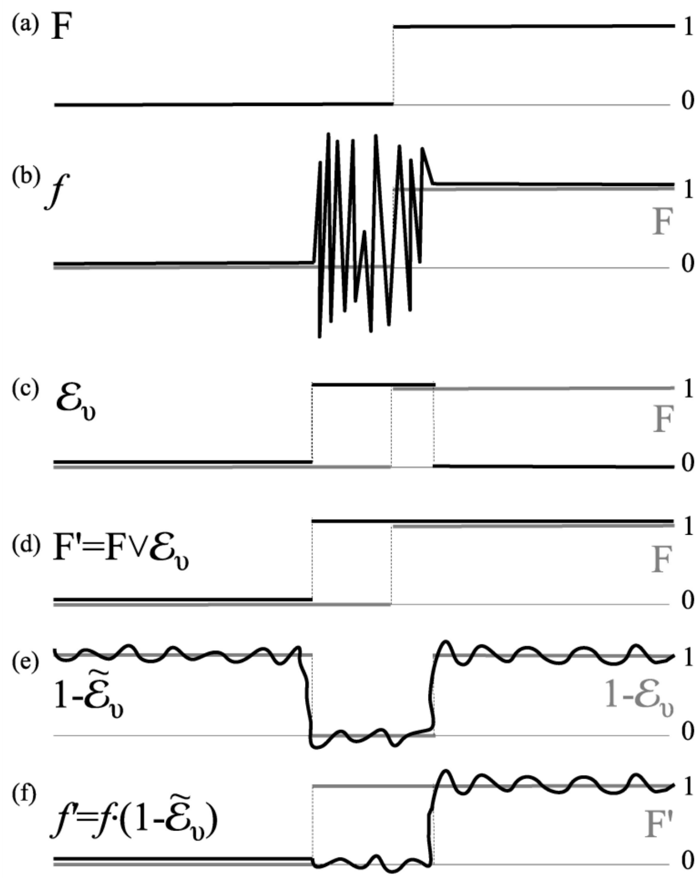

Figure 1: This figure was taken from [Bra08]. In this case, graph a) shows our function f. Graph b) gives the polynomial p_0. Graph c) is the error-detecting circuit E. Graph d) is the function f'. Graph e) shows the polynomial $1 - p_{E,2}$. Graph f) depicts $q = p_0(1 - p_{E,2})$.

can, given an LTF f, determine $|f^{-1}|$ in $poly(n)$ time. In other words, can we perform exact counting efficiently? The answer turns out to be no: exact counting is #P-hard, which is why we settle for approximate counting instead.

When we consider $|f^{-1}(1)|$ for an LTF f, it is helpful to have two perspectives. The first is to picture a hyperplane dividing the set of inputs into half-spaces. The second is to picture a discrete probability distribution over $\mathbb{R}$ for the 2^n values of $w \cdot x - \theta$ as x ranges over $\{-1, +1\}^n$. To visualize this, we will go through several examples. Note that we will be considering the locations of different $w \cdot x$, and θ will be used to determine what fraction of the points are satisfying assignments.

First, consider $w \cdot x = x_1$. The distribution ends up dividing half of the points to -1 and the other half to $+1$.

Second, consider $w \cdot x = \sum_{i=1}^{n} x_i$. This gives us a binomial distribution, and thus a bell-shaped curve that looks roughly like the Gaussian distribution $\mathcal{N}(0, n)$.

Finally, the LTF $w \cdot x = \sum_{i=1}^{n} 2^i x_i$ ends up giving a uniform distribution.

Of course, we could also apply this perspective to other arbitrary LTFs such as $w \cdot x = \sum_{i=1}^{n} i^{\log^2 i} x_i$. However, the distribution in our second example ends up being the "nicest." This will pertain to the "regularity" of an LTF, which is a concept that will be explored further in the next lecture. One can refer to [DGJ+10] to preview the results we will be showing.

References

[Bra08] Mark Braverman. Polylogarithmic independence fools ac0 circuits. *J. ACM*, 57(5), jun 2008. (document), 1

[BRS91] R. Beigel, N. Reingold, and D. Spielman. The perceptron strikes back. In *[1991] Proceedings of the Sixth Annual Structure in Complexity Theory Conference*, pages 286–291, 1991. (document)

[DGJ+10] Ilias Diakonikolas, Parikshit Gopalan, Ragesh Jaiswal, Rocco A. Servedio, and Emanuele Viola. Bounded independence fools halfspaces. *SIAM Journal on Computing*, 39(8):3441–3462, 2010. 3

[LMN93] Nathan Linial, Yishay Mansour, and Noam Nisan. Constant depth circuits, fourier transform, and learnability. *J. ACM*, 40(3):607–620, jul 1993. (document)

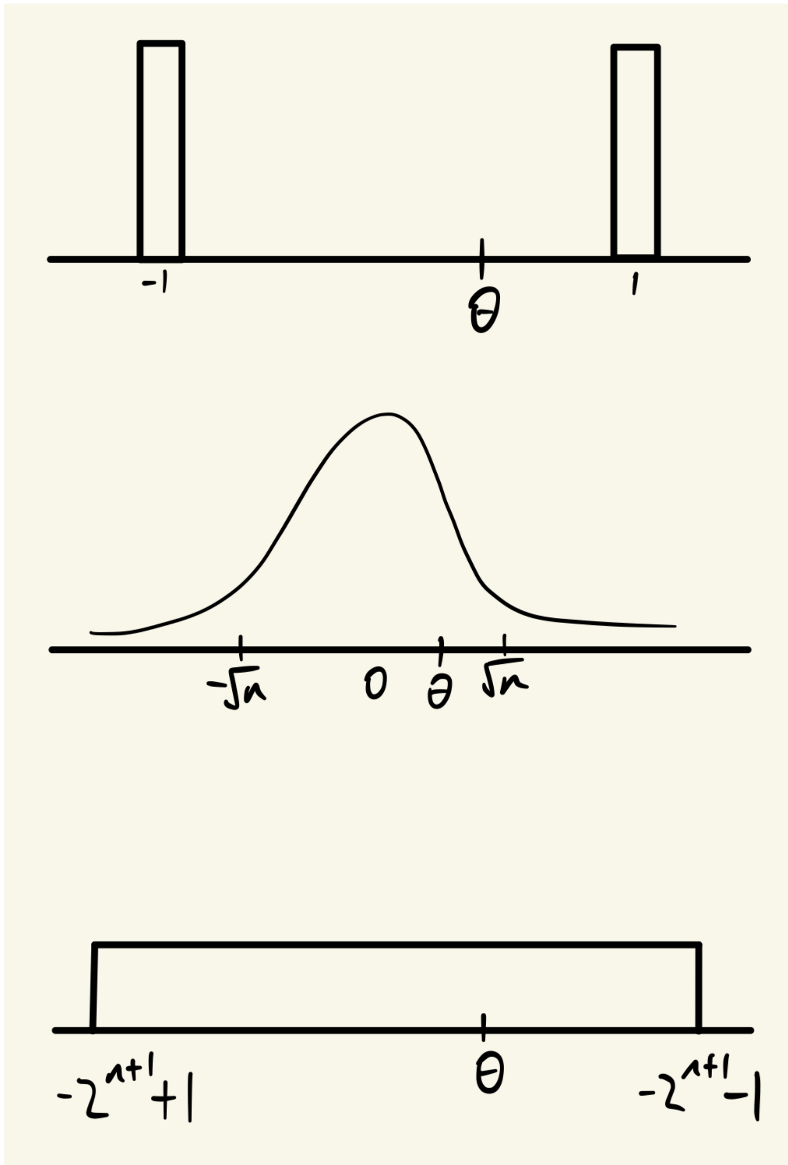

Figure 2: Visualizing LTFs using probability distributions.