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

Lecture 5: February 13, 2024

Lecturer: Rocco Servedio *Scribe: Mark Chen*

Last Time:

- Used Håstad's Switching Lemma (HSL) to get $2^{\Omega\left(n^{\frac{1}{d-1}}\right)}$ size lower bound (l.b.) for depth-d circuits for PAR.
- Proof of "weak switching lemma."
- Proof of HSL (with a "key fact" left to prove today).

Today:

- Finish proof of HSL by proving "key fact."
- Average-case l.b. for AC^0 circuits for PAR (small extension of worst-case l.b. we did).
- Depth-2 average-case l.b. (O'Donnell and Wimmer): Any CNF that agrees with (some explicit function) on 90% of all inputs must have $2^{\Omega\left(\frac{n}{\log(n)}\right)}$ clauses.
- Start $\mathbb{F}_2$-polynomials: Definitions and basic properties that lead up to showing an average-case l.b. for them (next time).

1 Håstad's Switching Lemma (Continued)

Many of these results can be found in chapter 12 of [Juk12].

Proof. [Of HSL] Last time, we had a key fact left unproven for HSL, which was stated as such:

Lemma 1. *Any restriction σ is $\mathsf{Angel}(\rho)$ for $\leq (4w)^t$ many bad ρ's.*

[**Intuition:** Recall from last time that $\mathsf{Angel}(\rho)$ is similar to $\mathsf{Devil}(\rho)$ in that it fixes t additional variables beyond ρ, but it is different from $\mathsf{Devil}(\rho)$ in that they necessarily disagree in each block, such that $\mathsf{Angel}(\rho)$ path is able to reach the 1-leaf of the block V_n with one query instead. This poses us a natural question that we are showing with this lemma here, because the hope is that $\mathsf{Angel}(\rho)$ is much more likely under R_p than ρ, because it has more fixed bits, and fixed bits are more likely than $*$'s.]

Proof of Lemma 1. Suppose we know F to begin with, though it is not explicit.

Let $\sigma = \mathsf{Angel}(\rho)$ for some bad ρ. The idea is to decode ρ from "σ" and "little extra information" (where the bound on the possibilities for the "little extra information" gives a bound on the number of of possible ρ's such that $\sigma = \mathsf{Angel}(\rho)$).

Now, we use the following auxiliary information: 2 rows of t numbers in each row and a little extra information (see Example 2):

- First row: An element of $[w]^t \implies$ gives w^t possibilities.
- Second row: An element of $\{0,1\}^t \implies$ gives 2^t possibilities.
- Extra info: In each of the $t-1$ possibilities between 2 elements of the first row, we can put ";" or not $\implies$ note that this can be represented by "0" and "1" for on and off in between each bit of the first row, and so there are $t-1$ such positions between t bits. That gives us 2^{t-1} possibilities.

This means a total of $\leq (4w)^t$ possible combinations. Since this is the upper bound for the total possible number of auxiliary information. Once we show we can decode ρ from σ and auxiliary information, we have the key lemma. ■

Example 2.

$$\boxed{2\ \_\ 3}\ ;\ \_\ \boxed{5\ \_\ 4\ \_\ 5}\ ;\ \_\ \cdots\ \_\ 3$$
$$0 \quad 0 \quad 1 \quad 0 \quad 1 \quad \cdots \quad 0$$

1.1 How to decode ρ from ["σ" + "auxiliary info"?

Intuition: The problem is we don't know which t fixed bits of σ are from $\mathsf{Angel}(\rho)$ (in contrast, these bits are already fixed in ρ). If we knew, we could replace them by $*$ in σ and get ρ.]

We'll identify those variables by first finding V_1, then, V_2, ... Here's how: Recall $\mathsf{CDT}(F, \rho)$ has V_1 = surviving variables in first of F that's not killed to 0 by ρ.

Imagine restricting F by just σ. The original σ adds to the existence of ρ, so any term killed to 0 by ρ is killed to 0 by σ. But, the first term in F that is not killed by ρ is satisfied by $\sigma = \mathsf{Angel}(\rho)$! So, the first term in F that's satisfiable by ρ is where V_1 came from.

Example 3. *Suppose we start with something like:*

$$F = (x_1 \wedge \overline{x}_2) \vee (\overline{x}_1 \wedge \overline{x}_2 \wedge x_8) \vee (\overline{x}_2 \vee x_4 \vee x_5 \vee \overline{x}_6) \vee \ldots.$$

Also, let's say that σ is fixed to be

$$\begin{array}{cccccccccc} & x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & x_7 & x_8 \\ \sigma = & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1, \end{array}$$

then

$$F = (\underbrace{\cancel{x_1 \wedge \overline{x}_2}}_{x_1=0,\ \textit{killed}}) \vee (\overline{x}_1 \wedge \overline{x}_2 \wedge x_8) \vee (\overline{x}_2 \vee x_4 \vee x_5 \vee \overline{x}_6) \vee \ldots.$$

Notice that $(\overline{x}_1 \wedge \overline{x}_2 \wedge x_8)$ is the first term that evaluates to true with the σ assignment, so it is where V_1 came from.

We use the following steps to inductively find variables in V_i:

Step 1: To decide which variables in the term we found above are V_1 ones: We read the elements of the first row to get info about which of the w variables in that term are "$*$" in ρ. Use ";" to mark the last position of this term. We have, thus, found V_1!

Step 2: To find V_2: Use the second row of the auxiliary info to

- learn how to traverse V_1-block of CDT to follow $\mathsf{Devil}(\rho)$ path (this is because the second row tells us how $\mathsf{Devil}(\rho)$ fixes the V_1 variables).
- Map $\sigma \rightsquigarrow \sigma'$ by replacing V_1 variables with those bits and continue.
- Now, the first term in F that is satisfiable by σ' must be where V_2 came from.

Then, we find variables that are V_2 ones analogous to how Step 1 does for V_1.

Step 3: We can find V_i inductively following the process described for V_2 based on the previously found variables and auxiliary information. So, like this, we can continue and end up recovering ρ, obtained by replacing all $V_1, V_2, \ldots$ variables in σ with $*$'s.

This marks the end of the HSL proof (see example 4 for how this decoding works). ■

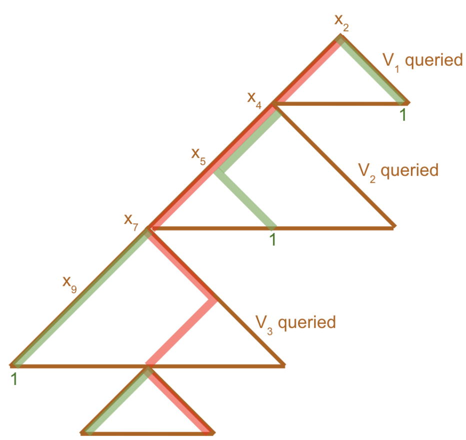
Example 4. *Recall the example from last time about Angel and Devil, for which the set-up is:*

- *Consider a "canonical decision tree" for $F \upharpoonright \rho$, denoted $\mathsf{CDT}(F, \rho)$. Furthermore, $\mathsf{CDT}(F, \rho)$ is to obliviously query all surviving variables (meaning to fix them to 0 or 1) in each block unkilled by ρ, and recurse through all such unkilled blocks. In the example illustrated below, we have $\{x_2\}$ in V_1 block unkilled, $\{x_4, x_5\}$ in V_2 block unkilled, ... Then, $\mathsf{Devil}(\rho)$ and $\mathsf{Angel}(\rho)$ are specific ways to fix surviving variables in unkilled blocks.*

- *The green path segments denote the $\mathsf{Devil}(\rho)$; the red path segments denote the $\mathsf{Angel}(\rho)$. In particular, we may have the following assignments (for x_1 to x_5 only, but it is easy to infer what x_7 and x_9 could be):*

	x_1	x_2	x_3	x_4	x_5
$\rho :$	1	$*$	0	$*$	$*$
$\mathsf{Devil}(\rho) :$	1	0	0	0	0
$\mathsf{Angel}(\rho) :$	1	1	0	0	1

[Notice how $\mathsf{Devil}(\rho)$ and $\mathsf{Angel}(\rho)$ fix additional bits to those that were initially assigned "$*$" by ρ ($\mathsf{Devil}(\rho)$ and $\mathsf{Angel}(\rho)$ always assign different values if it is the bit right before the end of a block).]



1.2 Project Topic

Variants / extensions of HSL:

- Proof complexity.
- Derandomization versions.
- ...

2 Average-Case L.B. for AC^0

Many of these results can be found in chapter 12 of [Juk12].

2.1 Recall Worst-Case Lower Bound

Reusing previous argument, we can see that all our "w.p. $\geq \frac{1}{2}$" are very strong. Let's set M, the size of the circuits against which we'll give an average-case l.b., to be

$$M := 2^{cn^{1/d}},$$

where $c = c_\alpha = \frac{1}{100^d}$. Then, overall, for M, we can verify each failure probability is at most $\frac{1}{M^5}$. So, $O(d)$ many: overall, w.p. $1 - \frac{O(d)}{M}$, the circuit C_{d-2} [the thing after random restrictions] is depth-2 circuit with bottom fanin $\leq 10 \log M$, over $n_{d-2} \geq \frac{M}{200(200 \log M)^{d-2}}$ variables.

Now, do one more round of random restriction. With $p = \frac{1}{100 \log M}$,

$$\Pr\left[C_{d-2} \text{ doesn't collapse to depth-}(10 \log M) \text{ DT}\right] \leq \frac{1}{M^5},$$

so, by Chernoff Bound, just like before,

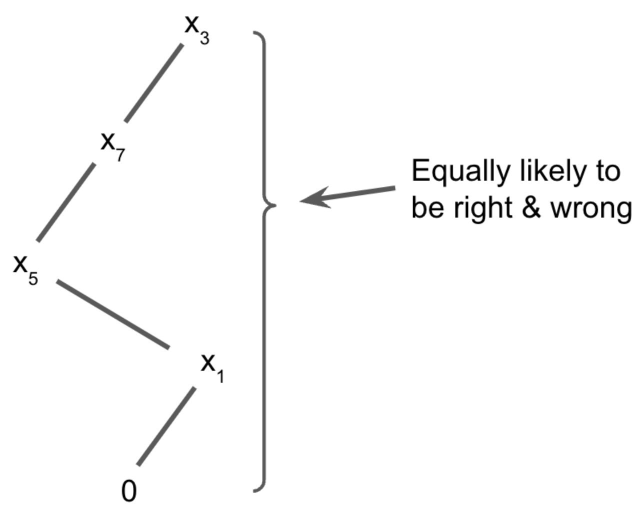
$$\Pr\left[\text{fewer than } \frac{n_{d-2}}{200 \log M} \text{ variables survive}\right] \leq \exp\left(-\frac{n}{c \cdot (\log M)^{d-2}}\right) \leq \frac{1}{M^5}.$$

So, overall, w.p. at least $1 - \frac{O(d)}{M^5}$, we get $(10 \log M)$-depth DT, and at least $\frac{n}{c \cdot (\log M)^{d-1}}$ variables survive.

2.2 Moving on to Average-Case Lower Bound

Fact 5. *Any DT of depth-d has correlation* 0*, under* $\mathcal{U}$*, with any* PAR *on more than d variables. See example 6*

Example 6. *Consider* $\text{PAR}(x_1, x_3, x_5, x_7, x_8)$*. Below is an example where half of all the assignments that reach this leaf satisfy the* PAR *and half don't. For this particular example, we see all variables other than* x_8 *are on the path. Given this path, we can see that if* $x_8 = 0$*, then* PAR *isn't satisfied; if* $x_8 = 1$*, then* PAR *is satisfied.*



Reinterpreting the key fact gives the following theorem:

Theorem 7. *Let C be a circuit of size* $M = 2^{cn^{1/d}}$*, depth d. Then,* PAR_n *is* ϵ*-hard for C, where* $\epsilon \leq 2^{-cn^{1/d}}$ *(which implies the average-case l.b. for* PAR*).*

2.2.1 Project Topic:

Refinements of the average-case l.b.'s for AC^0, by either making the circuit size bigger or making the ϵ bound smaller.

Goal 8. *In fact, it would be awesome to show that, for some explicit f, we can make* M *bigger while making* ϵ *smaller. But we don't know how to achieve both at the same time.*

Remark 9. *Theorem 7 is a compromise to that hope as described in goal 8, as in they made* M *bigger but also made* ϵ *bigger. Furthermore, what they did was restricted to depth-2 circuits, not* PAR_n*. Particularly, they achieved:* $M = 2^{n^{1/d}} = 2^{\sqrt{n}}$ *and* $\epsilon = \frac{1}{2^{\sqrt{n}}}$ *when* $d = 2$*.*

2.3 O'Donnell & Wimmer Statement and Proof [RO07]

Definition 10. *Let $F^* :=$ DNFTRIBES f_n on $n = w2^w$ variables be defined as*

$$F^*(x_1, \ldots, x_n) = \underbrace{(x_1 \wedge \ldots \wedge x_w) \vee (x_{w+1} \wedge \ldots \wedge x_{2w}) \vee \ldots}_{\frac{n}{\log n} = \frac{n}{w} = 2^w \text{ terms, } w \text{ variables per term}},$$

where we let $w \approx \log n - \log\log n, n = w2^w$.

Theorem 11. *[O'D&W] Any CNF g that agrees with F^* on 90% of all 2^n inputs must have $2^{\Omega\left(\frac{n}{\log n}\right)}$ clauses. Then, we can extend to get average-case l.b. for all depth-2 circuits.*

Proof. [$O'D\&W$] The first step is to show a fact:

Fact 12. *If g is s-clause CNF, there's a CNF g' which is ϵ-close to g, i.e.*

$$\Pr_{\mathbf{x} \in \mathcal{U}}[g(\mathbf{x}) = g'(\mathbf{x})] \geq 1 - \epsilon$$

s.t. g' has width of every clause being at most $\log(\frac{s}{\epsilon})$.

Proof. [Of fact 12] Any clause of length t falsifies w.p. $\frac{1}{2^t}$, so, removing a clause of width greater than $\log\left(\frac{s}{\epsilon}\right)$ changes g on $\leq \frac{1}{2^{\log\left(\frac{s}{\epsilon}\right)}} = \frac{\epsilon}{s}$ fraction of inputs. By union bound over all such clauses removed (at most s), we have the desired upper bound. ■

So, in order to show $O'D\&W$, it is sufficient to argue that:

Claim 13. *Any CNF g' that 0.2-approximates F^* must have width $\geq \frac{1}{4} \cdot 2^w = \Omega\left(\frac{n}{\log n}\right)$.*

Proof. [Claim 13 $\implies$ $O'D\&W$] Suppose g 0.1-approximates F^* and has s clauses. Then, fact 12 $\implies$ $\exists$ a width-$\log(10s)$ CNF g' s.t. it 0.1-approximates g. So, g' 0.2-approximates F^*, so claim 13 say $\log(10s) \geq \Omega\left(\frac{n}{\log n}\right)$. ■

Thus, the goal of the proof for $O'D\&W$ (theorem 11) is to show claim 13. To prove that, first observe:

Observation 14. *Random restrictions / switching lemma won't help*

Observation 14 is because F^*, like g', is a depth-2 circuit, so random restrictions will simplify F^* like g'. However, F^* is a width-w DNF and g' is a width-$\left(\frac{1}{4}2^w\right)$ CNF, so simplifications won't work because F^* will simplify at least as much as g' (and F^* is much smaller than g'). Instead, we need a way to "keep F^* complex" while "making g' simple." This is why we introduce the method of random projections.

Definition 15 (Random Projections)**.** *A projection $\boldsymbol{\rho}$ is a mapping $\{x_1, \ldots, x_n\} \to \{0, 1, \mathbf{y_1}, \ldots, \mathbf{y_t}\}$.*

The purpose of a random projection, ρ, is to both fix variables and identify groups of variables.

Example 16. *One example for $f \upharpoonright \boldsymbol{\rho}$ is:*

$$\begin{aligned}
\boldsymbol{\rho}(x_1) &= 1 \\
\boldsymbol{\rho}(x_2) &= 0 \\
\boldsymbol{\rho}(x_3) = \boldsymbol{\rho}(x_4) &= \mathbf{y_1} \\
\boldsymbol{\rho}(x_5) = \boldsymbol{\rho}(x_6) &= \mathbf{y_2} \\
&\vdots
\end{aligned}$$

The point of random projections is that they let us "carefully preserve structures" in target function F^* so it "survives."

Now, the key to prove claim 13, is to draw uniform n-bit string, using random projections:

Lemma 17. *Let $\boldsymbol{\rho} \sim \left\{\mathbf{y}_{\frac{1}{2}}, 1_{\frac{1}{2}}\right\}^w \backslash \{1^w\}$ and $\mathbf{y} \sim \left\{0_{1-\frac{1}{2^w}}, 1_{\frac{1}{2^w}}\right\}$ (the subscript denotes the probability that a specific position of a string is assigned a character). Doing $\boldsymbol{\rho}$ then $\mathbf{y}$ gives uniform string in $\{0,1\}^w$.*

Proof. [Of lemma 17] If $z = 1^w$, then we will need $\boldsymbol{\rho}$ to be all 1 or y to be 1. The former has a probability of $\frac{1}{2^w}$ and the latter has a probability of $\frac{1}{2^w}$, so the total probability is just $\frac{1}{2^w}$.

Otherwise, $z \neq 1^w$, then,

$$\Pr[\text{get } z] = \underbrace{\frac{1}{2^w - 1}}_{\text{get } \boldsymbol{\rho} \text{ compatible w/ } z} \cdot \underbrace{(1 - \frac{1}{2^w})}_{\text{get } \mathbf{y}=0} = \frac{1}{2^w}.$$

■

Finally, we prove the one missing piece, which is claim 13. We can do a global version of the trick described in lemma 17. Firstly, we have $\underbrace{\text{independent copies of } \boldsymbol{\rho}}_{(!)}$ for each of the 2^w terms of F^*. Recall F^* from definition 10. In draw of $\boldsymbol{\rho}_1, \boldsymbol{\rho}_2, \ldots, \boldsymbol{\rho}_{2^w}$, for each $i \in [2^w]$, all surviving variables under $\boldsymbol{\rho}_i \rightsquigarrow \mathbf{y_i}$.

- F^* "stays complex and balanced under $\boldsymbol{\rho}$": w.p. 1 over $\boldsymbol{\rho} = (\boldsymbol{\rho}_1, \boldsymbol{\rho}_2, \ldots, \boldsymbol{\rho}_{2^w})$,

  $$F^* \upharpoonright \boldsymbol{\rho} = \mathbf{y_1} \vee \mathbf{y_2} \vee \cdots \vee \mathbf{y_{2^w}},$$

  where $\mathbf{y_i} \sim \left\{0_{1-\frac{1}{2^w}}, 1_{\frac{1}{2^w}}\right\}$, so

  $$\mathop{\mathbb{E}}_{\mathbf{y_1},\ldots,\mathbf{y_{2^w}}}[F^* \upharpoonright \boldsymbol{\rho}(\mathbf{y})] = 1 - \left(1 - \frac{1}{2^w}\right)^{2^w} \approx 1 - \frac{1}{e} \approx 0.63.$$

- Any non-super-wide CNF is very biased (either towards 0 or towards 1) after $\boldsymbol{\rho}$: Fix any CNF g' of width $\leq \frac{1}{4}2^w$, consider any fixed outcome $g' \upharpoonright \boldsymbol{\rho}$, a CNF over $\mathbf{y_1}, \ldots, \mathbf{y_{2^w}}$: there are two possibilities:

  1) Every clause of $g' \upharpoonright \boldsymbol{\rho}$ has at least 1 negated variable, so

     $$g' \upharpoonright \boldsymbol{\rho} = (\overline{\mathbf{y}}_\mathbf{1} \vee \ldots) \wedge (\overline{\mathbf{y}}_\mathbf{7} \vee \ldots) \wedge \ldots,$$

     which means that
     $$g' \upharpoonright \boldsymbol{\rho}\left(0^{2^w}\right) = 1.$$
     Each $\mathbf{y_i}$ is 0 w.p. $1 - \frac{1}{2^w}$, but $F^* \upharpoonright \boldsymbol{\rho}\left(0^{2^w}\right) = 0$. So,

     $$\Pr[y = 0^{2^w}] = \left(1 - \frac{1}{2^w}\right)^{2^w} \approx 0.37.$$

     So, in this case,

     $$\boxed{g' \upharpoonright \boldsymbol{\rho} \text{ and } F^* \upharpoonright \boldsymbol{\rho} \text{ disagree on 37\% of } \mathbf{y}\text{-outcomes.}} \tag{1}$$

  2) Not every clause of $g' \upharpoonright \boldsymbol{\rho}$ has at least 1 negated variable, i.e. $g' \upharpoonright \boldsymbol{\rho}$ contains a clause
     $$C = \mathbf{y_1} \vee \mathbf{y_2} \vee \cdots \vee \mathbf{y_k} \text{ where } k \leq \frac{1}{4}2^w \text{ variables.}$$

Then,

$$\Pr_{\mathbf{y}}[g' \upharpoonright \boldsymbol{\rho}(\mathbf{y}) = 1] \le \Pr_{\mathbf{y}}[C(\mathbf{y}) = 1] \overset{\text{Union Bound}}{\le} \frac{1}{4} \cdot 2^w \cdot \frac{1}{2^w} = \frac{1}{4}.$$

But, we already concluded above that

$$\Pr_{\mathbf{y}}[F^* \upharpoonright \boldsymbol{\rho}(\mathbf{y}) = 1] = \Pr_{\mathbf{y}}[\mathbf{y_1} \vee \cdots \vee \mathbf{y_{2^w}} = 1] \approx 0.63.$$

So, in this case

$$\boxed{g' \upharpoonright \boldsymbol{\rho} \text{ and } F^* \upharpoonright \boldsymbol{\rho} \text{ disagree on } \approx \left(0.63 - \frac{1}{4}\right) = 0.38 \text{ of } \mathbf{y}\text{-outcomes.}} \quad (2)$$

In either case, we have

$$\Pr_{\mathbf{x} \sim \mathcal{U}(\{0,1\}^n)}[F^*(\mathbf{x}) \neq g'(\mathbf{x})] = \mathop{\mathbb{E}}_{\boldsymbol{\rho} \sim (!)}[\Pr_{\mathbf{y}}[F^* \upharpoonright \boldsymbol{\rho}(\mathbf{y}) \neq g' \upharpoonright \boldsymbol{\rho}(\mathbf{y})]]^{[1]}$$
$$\ge 0.2, \text{ based on (1) and (2) we have shown for two cases of } \boldsymbol{\rho}$$

.

And, thus, we have shown claim 13. Then, finally, as we have pointed out as soon as we stated claim 13, showing claim 13 shows $O'D\&W$, as desired. Thus, we have proved $O'D\&W$ (theorem 11). ■

2.3.1 Beyond $O'D\&W$:

To defeat all depth-2 circuits by considering DNF as well, instead of just CNF as theorem 11 did, we simply follow the same proof by switch between the following pairs: 1. 0/1 and 2. $\wedge/\vee$.

Corollary 18. *Any DNF g' that 0.1-approximates n-variable CNFTRIBES must have a size of at least $2^{\Omega(n/\log n)}$.*

Proof. Also analogous to what we have already shown for the CNF and DNFTRIBES equivalent above. ■

[1] This equality is a trick specified in the paper by O'Donnell and Wimmer in 2007, titled "Approximation by DNF: Examples and Counterexamples", link; similar approaches in chapter 12 of [Juk12]

2.3.2 Official Homework Problem:

Consider

$$f : \{0,1\}^{2n=2w2^w} \to \{0,1\}$$
$$f(a,b) = \mathsf{DNFTRIBES}(a) \vee \mathsf{CNFTRIBES}(b)$$

It is official homework problem to show that any depth-2 circuit that 0.01-approximates $f(a,b)$ must at least have a size of $2^{\Omega(n/\log n)}$.

2.3.3 Project Topic:

Other applications of random projections.

3 New Unit: Lower Bounds for $\mathbb{F}_2$-polynomials.

The following are the set-ups towards the sections (section 1.2 and section 1, respectively) on correlation bounds for polynomials over $\mathbb{F}_2$ in [Vio09, Vio22].

Definition 19 ($\mathbb{F}_2$)**.** *Recall from abstract algebra that, for $\mathbb{F}_p$, when p is prime, $\mathbb{F}_p \cong \mathbb{Z}_p$. So, $\mathbb{F}_2 \cong \mathbb{Z}_2 = \{[0],[1]\}$, and it is a field. For example, $[0] \equiv 4 \ (mod\ 2)$; $[1] \equiv 123 \ (mod\ 2)$.*

Definition 20 (Monomial over $\mathbb{F}_2$)**.** *First of all, monomials are polynomials that only have a single term (e.g., $x_1x_3^2x_4$ is a monomial but $x_1 + x_2$ is not). A monomial over $\mathbb{F}_2$ is a monomial such that the coefficient of the monomial is an element in $\mathbb{F}_2 \cong \mathbb{Z}_2$ (e.g., $x_1x_3^2x_4 \cong x_1x_3x_4 \in \mathbb{F}_2$ [we see here that powers greater than 1 can be erased for $\mathbb{F}_2$ polynomials, for a reason specified in the second bullet point]). There are a few simplifications / definitions we can make about $\mathbb{F}_2$ nomomials to see why they are special:*

- *$x_{i_1}x_{i_2}\dots x_{i_k}$ is a* deg-*k monomial with distinct $i_1,\dots$. Note that* AND $(x_{i_1}, x_{i_2}, \dots, x_{i_k})$ *has no negations.*
- *Never need to consider higher powers of x_i, because $x_i \in \{0,1\}$ and $0^2 = 0$, 1^1. By induction, this generalizes to a higher order finite n.*
- *All monomials over $\mathbb{F}_2$ are multilinear.*

Definition 21 ($\mathbb{F}_2$-Polynomial: Sum of Monomials over $\mathbb{F}_2$)**.** *Sum can be defined as:*

$$a + b \equiv a \oplus b \equiv \mathrm{PAR}(a,b).$$

Degree of a polynomial is the highest degree of any monomial in the sum.

Notation 22. *Note that there are 2^n n-variable multi-linear monomials because each of the n variables can take either an exponent of 0 or 1. Then, there are 2^{2^n} multi-linear polynomials, because the following map exists between polynomials and monomials:*

$$f : \{0,1\}^n \to \{0,1\}$$

(see fact 23 for details why).

Fact 23. *Every $f : \overset{\mathbb{F}_2^n}{\{0,1\}^n} \to \overset{\mathbb{F}_2}{\{0,1\}}$ has a unique expression as an $\mathbb{F}_2$-poly.*

Proof. This is an easy fact that can be proved by induction on n. ■

Remark 24. *Our goal, next, is to find the degree l.b.'s. However, $\text{AND}(x_1, \ldots, x_n) = x_1 \cdots x_n$ which already has degree n, so worst case l.b.'s for these polynomials are easy. Next time, we will show an average-case l.b.*

Notation 25. $\text{DEG}_d = \{$*all functions* $f : \{0,1\}^n \to \{0,1\}$ *that have deg-d $\mathbb{F}_2$-polys*$\}$.

3.1 Correlation Bound

In general, correlation bounds are hard!

3.1.1 Open question:

Prove some $f : \{0,1\}^n \to \{0,1\}, f \in \mathsf{NP}$, is $\frac{1}{n}$-hard for $\text{DEG}_{\log n}$ for some distribution $\mathcal{D}$ over $\{0,1\}^n$.

3.1.2 Next:

We will do 2 correlation bounds:

1) Degree $\theta(\sqrt{n})$, but correlation $= \theta(1)$.

2) Degree $<<\log n$, but tiny correlation.

References

[Juk12] Stasys Jukna. *Boolean Function Complexity.* Springer, New York, NY, 2012. 1, 2, 1

[RO07] Karl Wimmer Ryan O'Donnell. https://www.cs.cmu.edu/ odonnell/papers/approx-by-dnf.pdf. 4596:195–206, 2007. 2.3

[Vio09] Emanuele Viola. *On the Power of Small-Depth Computation*. Now Publishers Inc, New York, 2009. 3

[Vio22] Emanuele Viola. *Correlation bounds against polynomials*. 2022. 3