Last time: • Fourier analysis over $\{0,1\}^n$, simple applic. of
- fooling size-s DTs (ε/s-biased)
- fooling k-juntas better (k-wise $\varepsilon 2^{-k}$-biased)

• sandwiching approximators & fooling

• Viola's thm: sum of d ε-biased RVs fools DEG_d. (except for L2)

Today: • pf of L2 (finish pf of Viola's thm)

• tools for Braverman's thm: $(\log n)^{O(1)}$-wise indep. fools AC^0
- BRS poly. approx. of AC^0
- LMN " " " "

• (hopefully) start pf of Braverman's thm

Scribe: Shiv

Questions?

Reminder: • 2nd Official HW Problem due Wed March 20
• project progress report due Thurs April 4

Recall lemma we need to complete Viola's Thm:

L2 (bal. case): Sps W γ-fools DEG_{i-1}. Let Y be indep. of W, Y is δ-biased. Then

$W+Y$ $\left(\text{imbal}(f)+\sqrt{\gamma/2}+\frac{\delta}{2}\right)$-fools any $f\in DEG_i$.

Tech. lemma we'll use: Fix even n.

Let $g : \mathbb{F}_2^{n/2} \to \{\pm 1\}$, let $f(x,y) = g(x) \overset{\text{XOR}}{\cdot} g(y) \in \{\pm 1\}$.

For $U \sim \mathcal{U}_{n/2}$, Y any ε-biased RV over $\mathbb{F}_2^{n/2}$, have $(U, U+Y)$ ε-fools f.

i.e. $\left| \mathbb{E}_{U,U'}[f(U,U')] - \mathbb{E}_{U,Y}[f(U,U+Y)] \right| \le \varepsilon$.

Pf: Consider $F : \mathbb{F}_2^{n/2} \to [-1,1]$,

$$F(x) = \mathbb{E}_U\Big[\overbrace{g(U)\cdot g(U+x)}^{=f(U,U+x)}\Big].$$

Note: to prove, suff. to show Y ε-fools F.

We'll show $L_1(F) = 1$: means ε-biased Y (last time) $(\varepsilon \cdot L_1(F))$-fools F, so ☺.

Fix any $S \subseteq [n/2]$.

What's $\hat{F}(S)$? It's (recall $\hat{F}(S) = \mathbb{E}_{U'}[\chi_S(U')F(U')]$)

$$\mathbb{E}_{U,U'}[g(U)\cdot g(U+U')\cdot \chi_S(U')]$$

$$= \mathbb{E}_{U,U'}[g(U)\cdot g(U')\cdot \chi_S(U+U')]$$

$$= \mathbb{E}_{U,U'}[g(U)\cdot g(U')\cdot \chi_S(U)\cdot \chi_S(U')]$$

$$= \left(\mathbb{E}_U[g(U)\chi_S(U)]\right)^2 = \hat{g}(S)^2.$$

So $L_1(F) = \sum |\hat{F}(S)| = \sum \hat{g}(S)^2 \overset{\text{Parseval}}{=} \mathbb{E}_U[g(U)^2] = \mathbb{E}_U[1] = 1.$ ■

Pf of L2:

L2 (bal. case): Sps W γ-fools DEG_{i-1}. Let Y be indep. of W, Y is δ-biased. Then $W+Y$ $\left(\text{imbal}(f)+\sqrt{\gamma/2}+\frac{\delta}{2}\right)$-fools any $f \in DEG_i$.

Have

$$\left|\mathbb{E}[f(W+Y)] - \mathbb{E}[f(U)]\right|$$

$$= \frac{1}{2}\left|\mathbb{E}\left[(-1)^{f(W+Y)}\right] - \overbrace{\mathbb{E}\left[(-1)^{f(U)}\right]}^{=\text{imbal}(f)}\right| \quad \triangle\text{ ineq:}$$

$$\le \frac{1}{2}\underbrace{\left|\mathbb{E}\left[(-1)^{f(W+Y)}\right]\right|}_{} + \frac{1}{2}\cdot\text{imbal}(f).$$

So let's bound $\cdot$. Square + use C-S: $\mathbb{E}[A]^2 \le \mathbb{E}[A^2]$

$$\left(\mathbb{E}_{W,Y}\left[(-1)^{f(W+Y)}\right]\right)^2 \le \mathbb{E}_W\left[\left(\mathbb{E}_Y\left[(-1)^{f(W+Y)}\right]\right)^2\right]$$

$$= \mathbb{E}_{W,Y,Y'}\left[(-1)^{f(W+Y)}\cdot(-1)^{f(W+Y')}\right] \quad (Y' \text{ indep., dist. like } Y)$$

$$= \mathbb{E}_{W,Y,Y'}\left[(-1)^{f(W+Y)+f(W+Y')}\right]$$

For any fixed Y outcome, $f^{+Y}(x) = f(Y+x)$ is a deg-i poly in x.

And for any fixed outcome of $Y+Y'$, have

$f(x+Y)+f(x+Y')$ is $\partial_{Y+Y'} f^{+Y}(x)$, so $\deg(\cdot) \le i-1$.

Since W γ-fools DEG_{i-1}, have

$$\mathbb{E}_{w,Y,Y'}\left[(-1)^{f(w+Y)+f(w+Y')}\right] \le \mathbb{E}_{u,Y,Y'}\left[(-1)^{f(u+Y)+f(u+Y')}\right] + 2\gamma$$

Dist. $(u+Y, u+Y')$ $\equiv$ as dist. $(u, u+Y+Y')$, so

$Y+Y$: δ^2-biased

$$= \mathbb{E}_{u,Y,Y'}\left[(-1)^{f(u)+f(u+Y+Y')}\right]$$

tech lemma

$$\le \mathbb{E}\left[(-1)^{f(u)}\right]^2 + \delta^2 = \mathrm{imbal}(f)^2 + \delta^2.$$

Showed:

$$\left|\mathbb{E}[f(w+Y)] - \mathbb{E}[f(u)]\right|$$

$$\le \frac{1}{2}\left|\mathbb{E}\left[(-1)^{f(w+Y)}\right]\right| + \frac{1}{2}\cdot \mathrm{imbal}(f)$$

$$\le \frac{1}{2}\sqrt{\mathrm{imbal}(f)^2 + \delta^2 + 2\gamma} + \frac{1}{2}\mathrm{imbal}(f)$$

since $\sqrt{a+b+c} \le \sqrt{a} + \sqrt{b} + \sqrt{c}$,

$$\le \frac{1}{2}\left(\mathrm{imbal}(f) + \delta + \sqrt{2\gamma}\right) + \frac{1}{2}\mathrm{imbal}(f)$$

$$= \mathrm{imbal}(f) + \frac{\delta}{2} + \sqrt{\frac{\gamma}{2}}.$$

Bounded Indepence Fools AC^0

Motiv:

- we ♡ AC^0
- we got $\left\{\begin{array}{l}\text{worst-case}\\ \text{avg-case}\end{array}\right\}$ lb's against AC^0 :)

PRGs? Yes.

: not simplest
not first
not best params

most general machinery: k-wise indep.

History: $\boxed{'90}$ Linial & Nisan:

Conj: poly(n)-size, O(1)-depth ckts is fooled by any k-wise indep dist. for $k =$ polylog(n).

2007: Bazzi : $d = 2$ case (50 p)
8: Razborov : $d = 2$ (4 p)
9: Braverman: all d

Thm: Let $k = (\log \frac{s}{\varepsilon})^{O(d)}$.

SOTA: $k = (\log s)^{O(d)} \cdot \log \frac{1}{\varepsilon}$

Let $\mathcal{D}$ be any k-wise indep RV over $\{0,1\}^n$.

Then $\mathcal{D}$ ε-fools $AC^0_{s,d}$.

We'll do 1.

Why Conjecture?

- k-wise indep does fool depth-k DTs, k-juntas...

- k-wise inp. does fool real polys of deg k.

Known since '80s, '90s that AC^0 ckts are well-approx. by low-degree real polys, in 2 ways.

Each of the 2 approx. for AC^0: not good enough.
Braverman: clever combo. of them, to get sandwiching poly approximators. ✓

Rest of today: the 2 kinds of poly approx.

1st poly. approx for AC^0:

"pointwise" [B, R, S].

Lemma: (BRS) Let $f \in AC^0_{s,d}$.
Let $\mathcal{D}$ be any dist. over $\{0,1\}^n$.
There's a real poly p s.t.

$$\Pr_{x \sim \mathcal{D}}[p(x) = f(x)] \geq 1-\varepsilon, \text{ where}$$

(i) $\deg(p) \leq (\log s/\varepsilon)^{O(d)}$, +

(ii) $\forall x \in \{0,1\}^n, \quad |p(x)| \leq \exp\left(\left(\log \frac{s}{\varepsilon}\right)^{O(d)}\right).$

---

this p may be bad on average.....

First step in BRS pf: do it for

$f(x) = x_1 \vee \dots \vee x_t$ ($d=1$ case)

So, let $f = x_1 \vee \dots \vee x_t$. We'll describe a <u>dist over polys</u>:

Let $V_0 = \{x_1, \dots, x_t\}$ set of vars.

$p_0(x) = x_1 + \dots + x_t$.

For $i = 1, \dots, \log_2(t) + 1$, let

- V_i be constr. from V_{i-1} by discarding each var w.p. $1/2$
- $p_i(x) = \sum_{x_j \in V_i} x_j$.

$p_0, p_1, \dots, p_{1+\log t}$ all deg-1 polys, taking values in $[0, t]$ on $\{0,1\}^t$.

Fix any input asst $z \in \{0,1\}^t$ s.t. some $z_i = 1$, i.e. $z \neq 0^t$. (so $z_1 \vee \dots \vee z_t = 1$)

<u>Claim:</u> $\Pr\left[\text{at least one of } p_0(z), \dots, p_{1+\log t}(z) = 1\right] \geq \frac{1}{3}$.

<u>Pf</u>: 3 cases:

(a) $p_i(z) > 1 \;\; \forall\, i = 0, \dots, \log_2(t) + 1$.

Each z_j survives all stages w. p. $\le \frac{1}{2t}$.

So $\Pr[\text{any } z_j \text{ surv. all stages}] \le 1/2$.

So $\Pr[ⓐ] \le \frac{1}{2}$.

ⓑ $p_0(z) = 1$.

ⓒ for some i, $p_i(z) > 1$ but $p_{i+1}(z) \le 1$.

In this case, given value of $p_j(z)$, have

$$\Pr[p_{j+1}(z) = 0] = \frac{1}{2^{p_j(z)}}, \ +$$

$$\Pr[p_{j+1}(z) = 1] = p_j \cdot \frac{1}{2^{p_j(z)}}$$

000
100
010
001

$p_j(z) = 2$:

00
01
10

So if i is s.t. $p_i(z) > 1$ + $p_{i+1}(z) \le 1$,

have $\Pr[p_{i+1}(z) = 1] \ge 2/3$.

$\frac{2}{3}$

So since ⓐ $\le \frac{1}{2}$

one of ⓑ, ⓒ $\ge \frac{1}{2}$

$\frac{1}{2} \cdot \frac{2}{3} = \frac{1}{3}$ ∎

Break

Need to improve success prob... + get a single poly...

$$\text{Define } \cancel{r(x) = \prod_{i=0}^{\log(t)+1} (1 - p_i(x))}$$

rand. poly, deg $\leq \log(t)+1$,
values in $[-t^{O(\log t)}, t^{O(\log t)}]$.

$x = 0^t$: $r(x) = 1$ b/c $p_i(x) = 0$ $\forall i$.
Any other x: some $p_i(x) = 1$ w.p. $\geq \frac{1}{3}$.
So $\Rightarrow$, w.p. $\geq \frac{1}{3}$ $r(x) = 0$.

$\Rightarrow$: Let $r'(x)$ = prod. of $O(\log \frac{1}{\varepsilon})$ indep $r(x)$'s.

$\mathrm{Deg}(r') \leq O(\log \frac{1}{\varepsilon} \cdot \log t)$,

values of $r' \in [-t^{O(\log \frac{1}{\varepsilon}) \cdot \log t}, t^{O(\log \frac{1}{\varepsilon}) \cdot \log t}]$.

$r'(0^t) = 1$

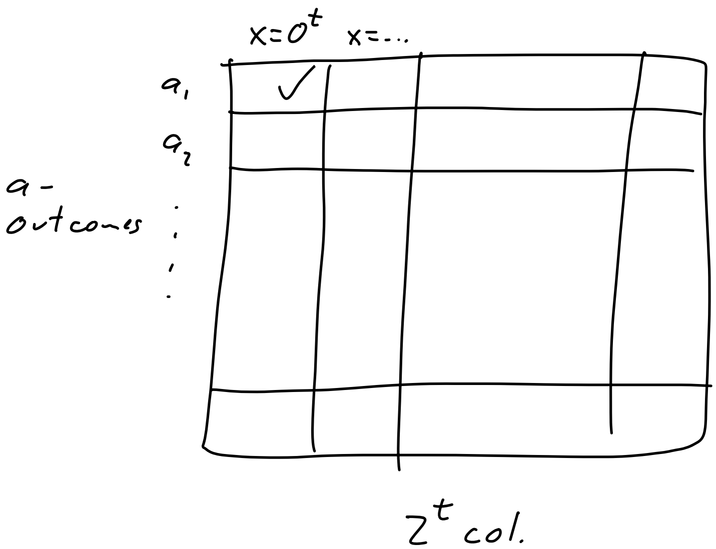
any $x \neq 0^t$: have $r'(x) \neq 0$ w.p. $\leq (2/3)^{O(\log \frac{1}{\varepsilon})} \leq \varepsilon$.

Let $a(x) = 1 - r(x)$: have

$(\star)$ $\forall x \quad \Pr_a[a(x) = \overbrace{x_1 \vee \dots \vee x_t}^{0/1 \text{ value}}] \geq 1 - \varepsilon$.

$\Longrightarrow$ $\forall$ dist $\mathcal{D}$, $\exists$ (poly) outcome of a s.t.

$(\star\star)$ $\Pr_{x \sim \mathcal{D}}[a(x) = x_1 \vee \dots \vee x_t] \geq 1 - \varepsilon$.



✓: $a(x) = x_1 \vee \ldots \vee x_t$.

☆: $\forall$ col., $1-\varepsilon$ frac. entries are ✓.

So $\forall\, \vartheta$ over cols, ✓-density in mtx (a la ϑ) is $\geq 1-\varepsilon$.

So $\exists$ row w/ $\geq 1-\varepsilon$ density of ✓ (a la ϑ)

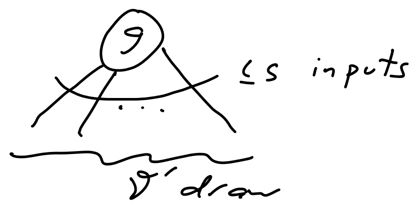
& that's the a of ☆☆

Did $d=1$ case of $f = x_1 \vee \ldots \vee x_t$.

Also true for AND, ¬ $\quad \neg x \quad 1-x$

For each gate g in $f \in AC^0_{s,d}$,

Consider dist. of inputs to that gate when $x \sim \vartheta$.
ϑ'



By prev lemma,

there's a deg-$O(\log\frac{s}{\varepsilon}\cdot\log s)$ poly.

which computes value of g correctly w.p. $\geq 1 - \varepsilon/s$ for (args to g) $\sim \mathcal{D}'$.

Replace each of s gates by its poly:

composition is a final poly $p(x)$,

of deg $O(\log(s/\varepsilon) \cdot \log s)^d$, which computes f

correctly w.p. $\geq 1 - \frac{\varepsilon}{s} \cdot s = 1 - \varepsilon$.

And easy to check (ii) as well.

End of [BRS] approx.

2^{nd} poly. approx for AC^0:

" L_2 approx " [L, M, N].

Thm: Let $f \in AC^0_{s,d}$.
LMN
There's a real poly p_2, of deg $O((\log \frac{s}{\varepsilon})^d)$,

s.t.
$$\mathbb{E}_{\mathcal{U}}\left[(f(\mathcal{U}) - p_2(\mathcal{U}))^2\right] \leq \varepsilon.$$

---

p_2 may not sandwich, $f(x) - p_2(x)$ may (rarely) be big...

Notation: Write
$$\boxed{W^k(f) := \sum_{S \subseteq [n],\, |S| = k} \hat{f}(S)^2}$$

$$\boxed{W^{\geq k}(f) := \sum_{\substack{S \subseteq [n] \\ |S| \geq k}} \hat{f}(S)^2}$$

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

$$W^{\geq 0}(f) = \sum_{S \subseteq [n]} \hat{f}(S)^2 = 1.$$

Thm (LMN): Let $f \in AC^0_{s,d}$. There's a real poly p_2, of deg $O((\log \frac{s}{\varepsilon})^d)$, s.t. $\mathbb{E}_u[(f(u) - p_2(u))^2] \leq \varepsilon$.

follows from:

"Fourier concentration"

Thm': If $f \in AC^0_{s,d}$, then $\forall r$, have

$$W^{>r}(f) \leq 2s \cdot 2^{-\frac{r^{1/d}}{20}}$$

Given , take $r = (20 \log \frac{2s}{\varepsilon})^d$. Makes $2s \cdot 2^{-\frac{r^{1/d}}{20}} \leq \varepsilon$.

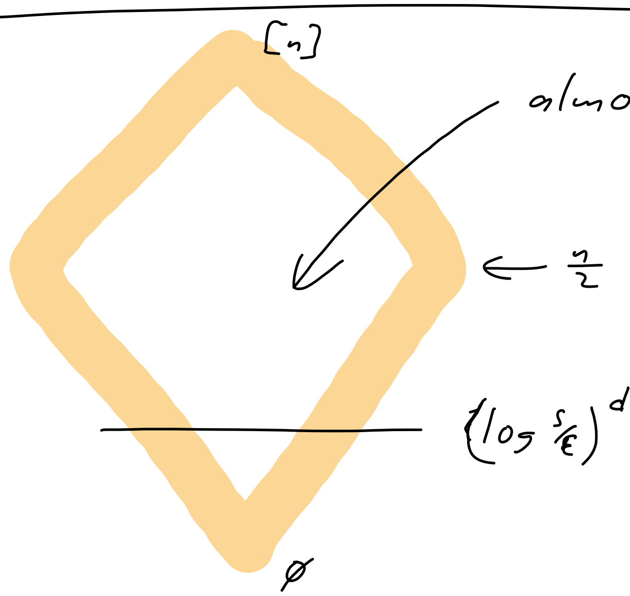
so the poly p_2 is just the truncated (at level r) Fourier part of f:

$$p_2(x) = \sum_{|S| \leq r} \hat{f}(S) \underbrace{\chi_S(x)}_{= \prod_{i \in S} x_i \text{, for } \pm 1 \text{ inputs}}, \quad \text{b/c by}$$

Parseval,

$$\mathbb{E}[(f - p_2)^2] = \sum_{S \subseteq [n]} \widehat{f - p_2}(S)^2$$

$$= \sum_{S \subseteq [n]} \left(\hat{f}(S) - \hat{p_2}(S) \right)^2 \leq \quad \leq \varepsilon \text{ by our choice of } r.$$



almost no Fourier wt up here, for AC^0!

Goal: . 2 main steps.

(1): based on switching lemma:

(L1): Let $f \in AC^0_{s,d}$. Fix t + let

$\rho \sim R_p$, $p \overset{*\text{-prob.}}{\leq} \dfrac{1}{10^d \cdot t^{d-1}}$.

Then $\Pr_{\rho \sim R_p}[\text{DT-depth}(f \restriction \rho) \geq t] \leq s \cdot 2^{-d}$.

☺ we did this!

(L2): generic statement about Fourier + rand. restric.

(L2): For any $f:\{\pm1\}^n \to \{\pm1\}$, any $p \le \frac{1}{10}$, have

$$W^{\ge t/p}(f) \le 2 \cdot \mathop{\mathbb{E}}_{(J,z)\sim R_p}\left[W^{\ge t}[f_{J,z}]\right].$$

"$(J,z)\sim R_p$" same as "$\rho \sim R_p$":

$J \subseteq [n]$, $J = *$-vars,

$z \in \{\pm1\}^{\bar{J}}$ asst to non-$*$ vars.

Pf that (L1) + (L2) $\Longrightarrow$

take $p = \frac{1}{10 \cdot r^{\frac{d-1}{d}}}$, so $pr = \frac{r^{1/d}}{10}$.

Let $t = \frac{pr}{2} = \frac{r^{1/d}}{20}$, get

$$\frac{1}{10^d t^{d-1}} = \frac{20^{d-1}}{10^d r^{\frac{d-1}{d}}} = \frac{2^{d-1}}{10 r^{\frac{d-1}{d}}} = 2^{d-1} \cdot p \ge p,$$

so can use (L1): gives

$$\Pr_\rho[\text{DT-depth}(f_\rho) \ge t] \le s \cdot 2^{-t}.$$

☆ This means $\mathop{\mathbb{E}}_{(J,z)\sim R_p}[W^{\ge t}(f_{J,z})] \le s \cdot 2^{-t}$.

$t = \frac{r^{1/d}}{20}$, $\frac{t}{p} = \frac{r}{2}$;

So by (L2), get

$$W^{\geq \frac{r}{2}}(f) \leq 2 \cdot \mathop{\mathbb{E}}_{(J,z)\sim R_p}\left[W^{\geq t}(f_{J,z})\right] \leq 2s \cdot 2^{-t}$$

$$= 2s \cdot 2^{-\frac{r^{1/d}}{20}}.$$

Left to do: prove (L2).

(L2): For any $f : \{\pm 1\}^n \to \{\pm 1\}$, any $p \leq \frac{1}{10}$, have

$$W^{\geq t/p}(f) \leq 2 \cdot \mathop{\mathbb{E}}_{(J,z)\sim R_p}\left[W^{\geq t}[f_{J,z}]\right].$$

No ckt involved :)

Note: gloss some of steps in L2 pf...