1.

AGENDA

1. Worst-case $\to$ avg case
2. XOR lemma : avg case $\to$ harder avg case

$$\exists \text{ worst-case hard fns} \overset{\text{today}}{\Longrightarrow} \exists \text{ avg-case hard fns} \overset{\text{last week}}{\Longrightarrow} BPP = P$$

$$\left(H_{wrs}(f) \geq 2^{\varepsilon n}\right) \qquad \left(H_{avg}(f) \geq 2^{\varepsilon n}\right)$$

1. Worst-case $\to$ avg case

Main idea: coding theory

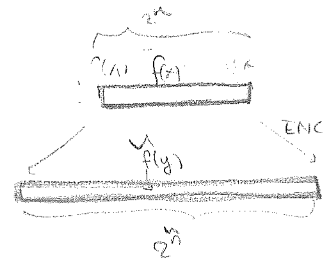
say we had ENC :

$$f : \{0,1\}^n \to \{0,1\} \xrightarrow[\text{ENC}]{} \hat{f} : \{0,1\}^{\hat{n}} \to \{0,1\}$$

so that $\forall\ g : \{0,1\}^{\hat{n}} \to \{0,1\}$, with

$$\Delta(\hat{f}, g) := \left|\{ y \in \{0,1\}^{\hat{n}} : \hat{f}(y) \neq g(y)\}\right| \leq \delta \cdot 2^{\hat{n}},$$

we can (efficiently) recover f from g.



If so, we can amplify hardness:

Say f is worst-case hard. Then $\text{ENC}(f) = \hat{f}$ is avg-case hard.

Otherwise, $\exists$ C that computes $\hat{f}$ on $(1-\delta)2^{\hat{n}}$ inputs (say)

$$\Rightarrow \Delta(C, \hat{f}) \leq \delta \cdot 2^{\hat{n}}$$

$\Rightarrow$ We can use C to find f, using our efficient decoding algorithm.

This should raise some red flags: efficient needs to mean REALLY efficient.

Lif I even look at every entry $C(y)$, that's still $2^{\hat{n}}$ copies of C, so

$$H_{wrs}(f) \geq t \Rightarrow \underbrace{H^{1-\delta}_{avg}(\hat{f}) \geq t / 2^{\hat{n}}}_{\text{this is extremely uninteresting.}}$$

To fix this, LOCAL DECODING.

Def We say (ENC, DEC) is (efficiently) δ-locally-decodable w/ q queries

maps $f:\{0,1\}^n \to \{0,1\}$ to $\hat{f}:\{0,1\}^{\hat{n}} \to \{0,1\}$ (ENC)

randomized alg w/ an oracle (DEC)

if $\forall\ g:\{0,1\}^{\hat{n}} \to \{0,1\}$ w/ $\Delta(\hat{f}, g) \le \delta 2^{\hat{n}}$, $\forall x \in \{0,1\}^{n}$

$$\mathbb{P}_{DEC}\{ DEC^{g}(x) = f(x) \} \ge 2/3$$

and DEC makes q calls to the oracle g (and runs in poly(q) time, say)

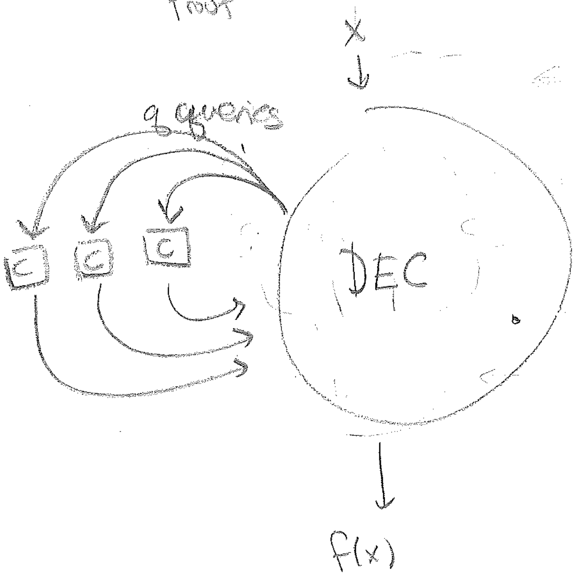
Previous arg shows:

"LEMMA": Say (ENC, DEC) Local-δ-dec. w/ q queries.

not enough quantifiers to make this a real lemma.

$$H_{avg}^{1-\delta}(ENC(f)) \ge \frac{H_{wrs}(f)}{poly(q)}$$

"Proof" ← here's a circuit for computing $f(x)$ given C for $\hat{f}(x)$.



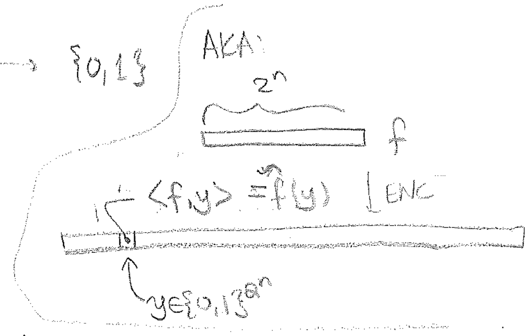
← size poly(q) by cook-levin

NOTE: we can hard-code in the randomness of DEC, a la BPP $\subseteq$ P/poly

size is $\approx |C| \cdot q + \text{poly}(q)$

3 **Example:** Hadamard encoding.

Define $f : \{0,1\}^n \to \{0,1\} \xmapsto{\text{ENC}} \tilde{f} : \{0,1\}^{(2^n)} \to \{0,1\}$



By $\tilde{f}(y) = \sum_{x \in \{0,1\}^n} f(x) \cdot y(x)$

think of $y \in \{0,1\}^{2^n}$ as $y : \{0,1\}^n \to \{0,1\}$

So if $y_0 = \mathbb{1}_{x_0}$ (aka, $y(x_0) = 1$, $y(x) = 0 \ \forall x \neq x_0$) then $\tilde{f}(\mathbb{1}_{x_0}) = f(x_0)$.

(That's a pretty dumb way to get $f(x_0)$ from $\tilde{f}$, though.)

We also have, $\forall y \in \{0,1\}^{2^n}$, $\tilde{f}(\mathbb{1}_{x_0} \oplus y) = \tilde{f}(\mathbb{1}_{x_0}) \oplus \tilde{f}(y)$, so

$$\tilde{f}(\mathbb{1}_{x_0} \oplus y) \oplus \tilde{f}(y) = \tilde{f}(\mathbb{1}_{x_0}) = f(x_0).$$

So, local decoding algorithm is: (supposing $C \equiv \tilde{f}$ on a $1-\delta$ fraction)

1. pick y at random
2. Ask the oracle for $C(\mathbb{1}_{x_0} \oplus y)$, $C(y)$
3. Return $C(\mathbb{1}_{x_0} \oplus y) \oplus C(y)$.

Analysis: correct w/ prob... $(1-\delta)^2 > 2/3$ for δ small enough. Hooray!

What's wrong with that?

We've just shown that

$$H_{avg}^{1-\delta}(\tilde{f}) \geq \frac{H_{wrs}(f)}{poly(n)}$$

So, even if f were REALLY REALLY hard, $H_{wrs}(f) \geq 2^n$, then

$H_{avg}^{1-\delta}(\tilde{f}) \gtrsim 2^n \ \ldots = \tilde{n}$. . . :(D'oh.

$\boxed{4}$ How can we do better? Reed Muller Codes.

$$f: \cancel{\{0,1\}}^n \to \cancel{\{0,1\}} \quad \xrightarrow{\text{ENC}} \quad \hat{f}: \cancel{\{0,1\}}^{\hat{n}} \to \cancel{\{0,1\}}$$

$\mathbb{F} \quad \mathbb{F} \qquad\qquad \mathbb{F} \quad \mathbb{F}$

We could have phrased what we had before as: $\hat{f} \in \{0,1\}^{2^{2^n}}$

$x \in \{0,1\}^n$ — 2^n — $y \in \{0,1\}^{2^n}$

$f \in \{0,1\}^{2^n}$

$\hat{f}(y) = \langle f, y \rangle$ (aka, inner product mod 2)

aka

$1_x \in \{0,1\}^{2^n}$ — $y \in \{0,1\}^{2^n}$

$\langle f, 1_x \rangle$ — $\langle f, y \rangle$

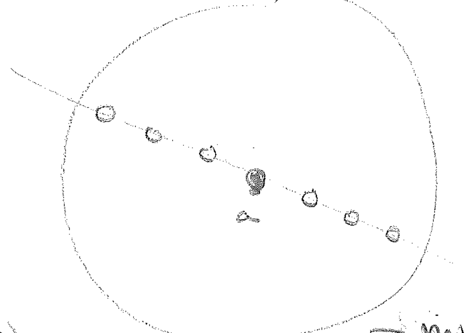
So f is playing the role of some linear fn. Why does it have to be linear?

$A \subseteq \mathbb{F}^m$ — $y \in \mathbb{F}^m$ — $\mathbb{F}^m$

$p(a),\ a \in A$ — $p(y)$

for $p(X_1, \ldots, X_n)$ an ~~m~~-variate deg-d polys say $d < (1 - 10\delta)|\mathbb{F}|$

decoding alg: Say I want $p(a)$.



$\mathbb{F}^m$

- Pick a random line through a,
  $L = \{a + \lambda b : \lambda \in \mathbb{F}\}$
  for $b \in \mathbb{F}^n$ random. (not zero)
- Restricted to L, $p|_L$ is a UNIVARIATE deg d poly.
- whp (at least $2/3$), there are $\leq 5\delta|\mathbb{F}|$ places where C messes up
- There's a unique poly $p|_L$ that agrees w/ these in $\geq (1 - 5\delta)|\mathbb{F}|$ places and we can find it efficiently.

Issue: over $\mathbb{F}$, which is large — can fix that w/ concatenated codes.

5

one can show

Thm There is a δ-LDC (ENC, DEC) for $\delta = 0.01$ s.t.

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

  $ENC(f) : \{0,1\}^{O(n)} \to \{0,1\}$

  and $ENC(f)$ can be computed in time $2^{O(n)}$

- DEC uses poly(n) queries and has poly(n) running time.

This implies:

Thm If $f \in E$ has worst-case hardness $H_{wrs}(f) = H(n)$ then there is some $\hat{f} \in E$ w/

$$H_{avg}^{0.99}(\tilde{f}) \geq H(n/c) / n^c$$

for some constant c.

This is actual progress, unlike our previous tries, but it's still not enough.

From last time,

$$\underbrace{H_{avg}(\hat{f})}_{\max\{s \,:\, \underbrace{H_{avg}^{1/2+1/s}}(\hat{f}) \geq s\}} \geq 2^{n\epsilon} \Rightarrow BPP = P$$

this is a REALLY hard fn.
Can't even guess it w/ prob ½ + negl.

Above, we have hard to guess w/ prob 0.99.

We CAN push this coding technique through w/ list-decoding.

But let's go a different route.

6 ~~X~~OR LEMMA. "Mild" avg-case hardness (0.99-hard) $\longrightarrow$ "Extreme" avg-case hardness ($1/2$ + negligible - hard).

Here's what we'll do:

Given $f : \{0,1\}^n \longrightarrow \{0,1\}$ mildly hard, consider

$f^{\oplus k} : \{0,1\}^{nk} \longrightarrow \{0,1\}$ given by

$$f^{\oplus k}(x_1, \ldots, x_k) = f(x_1) \oplus f(x_2) \oplus \cdots \oplus f(x_k).$$

Intuitively, to compute $f^{\oplus k}$, you need to compute all k copies of f.

$\boxed{\text{Thm}}$ (Yao's XOR-lemma)

If $H_{avg}^{1-\delta}(f) = s$, then

$$H_{avg}^{1/2+(1-\delta)^k+\gamma}(f^{\oplus k}) = \Omega\left(\gamma^2 \log\left(\tfrac{1}{\varepsilon\gamma}\right)\right) \cdot s$$

Before we prove this, notice that this will get us derandomization:

$$\exists f \in E,\ H_{wrs}(f) = 2^{\varepsilon n} \overset{\text{coding}}{\Rightarrow} \overset{\exists \hat{f} \in E}{H_{avg}^{0.99}(\hat{f}) = 2^{\hat{\varepsilon} n}}$$

$$\overset{\text{choose } k \approx \hat{\varepsilon} n,\ \gamma \approx 1/2^{\hat{\varepsilon} n}}{\Rightarrow} H_{avg}(\hat{f}^{\oplus k}) = 2^{\hat{\hat{\varepsilon}} n} \overset{\text{NW}}{\Rightarrow} BPP = ~~Q~~ QUASI\text{-}P$$

and similarly

$$H_{wrs}(f) = 2^{n^\epsilon} \Rightarrow BPP \subseteq QUASI\text{-}P$$

$$H_{wrs}(f) = n^{\omega(1)} \Rightarrow BPP \subseteq SUBEXP$$

7 proof of XOR LEMMA.

First, let's understand what makes functions hard on avg.

0.99^-

Theory 1. Hardness is "spread out" among inputs. ← eg, any small circuit fails on a bunch of inputs, and which inputs vary from circuit to circuit.

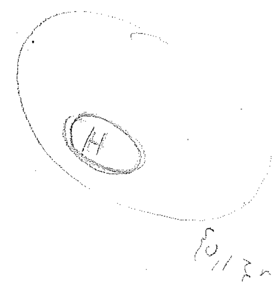
Theory 2. A small fraction of inputs are REALLY HARD. ← eg, all circuits have a hard time w/ a few inputs.

Turns out, it's theory 2.

Def $H \subseteq \{0,1\}^n$ is a γ-HARD CORE against size s circuits for $f:\{0,1\}^n \to \{0,1\}$ if

$$\forall \text{circuits } |C| \le s, \quad \mathbb{P}_{x \in_R H}\{C(x) = f(x)\} \le 1/2 + \gamma$$

← replace uniform on $\{0,1\}^n$ w/ uniform on H.

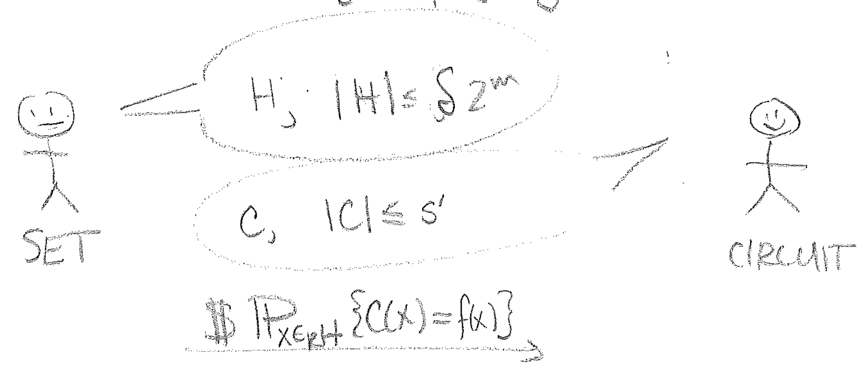


Thm (Impagliazzo '95).

If $H_{avg}^{1-\delta}(f) = s$, then $\exists$ γ-hardcore set H for

$$s' = \Omega\left(\gamma^2 / \log(1/\delta\gamma)\right) \cdot s, \text{ with } |H| \ge \delta \cdot 2^n.$$

Proof Consider the following 2-player game.



It's unclear why SET would ever agree to play this game.

18 Now, by Von Neumann min-max thm,

MIN-MAX THM: In a 0-sumgame, $\min_{p \in \Delta^n} \max_{q \in \Delta^n} q^T A p = \max_{q \in \Delta^n} \min_{p \in \Delta^n} q^T A p$ (Where A is a payoff matrix)

$\exists V$ s.t.

(1) $\exists$ randomized strategy for SET (aka, dist on sets H) s.t. $\forall$ rand. strategies for CIRCUIT,

$$\mathbb{E}\{ \$ \} \leq V$$

— and —

(2) $\exists$ rand. strat. for CIRCUIT (dist on circuits) s.t. $\forall$ rand strategies for SET,

$$\mathbb{E}\{ \$ \} \geq V$$

If $V \leq 1/2 + \gamma$, then we're done! Just use SET's strategy in (1) to pick H.

Not quite — actually you get a dist on H. But you can do it w/ some avging trickery.

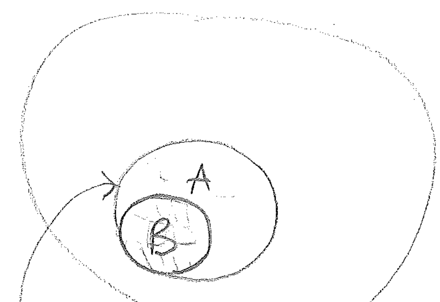
If $V \not\leq 1/2 + \gamma$ ... CLAIM, this never happens.

Suppose this is the case. Then CIRCUIT has a strategy $\mathcal{C}$ (dist on circuits C) s.t. for every set H of size $\geq \delta 2^n$, $\mathbb{E}_{C \sim \mathcal{C}} \mathbb{P}_{x \in_R H}\{C(x) = f(x)\} \geq 1/2 + \gamma$.

Let $B = \{ x : \mathbb{P}_{C \sim \mathcal{C}}\{C(x) = f(x)\} < \frac{1}{2} + \gamma/10 \}$ be the x's on which $\mathcal{C}$ is bad.

CLAIM: $|B| \leq \delta(1 - \gamma/2) \cdot 2^n$

Pf. $\mathcal{C}$ has to do well on $B \cup \overbrace{\{\text{any set of size } \delta 2^n - |B|\}}^{A}$ (disjoint)



extend B to a set of size $\delta 2^n$ by adding A

so if $|B| > \delta(1 - \gamma/2) \cdot 2^n$, then

$$|A| < \frac{\delta \gamma}{2} \cdot 2^n$$

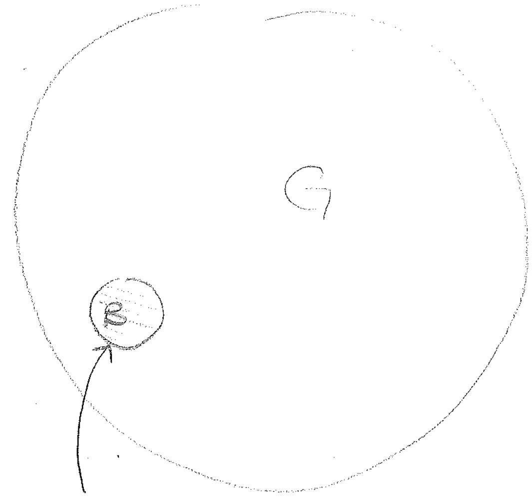
Then

$$\mathbb{E}_{C \sim \mathcal{C}} \mathbb{P}_{x \in_R (B \cup A)}\{C(x) = f(x)\} \leq \left(\frac{1}{2} + \frac{\gamma}{10}\right)(1 - \gamma/2) + \gamma/2 < \frac{1}{2} + \gamma \;\lightning$$

Now use $\mathcal{C}$ to violate hardness of f:

Pick $C_1, \dots, C_t$ indep. from $\mathcal{C}$, let $D = \mathrm{MAJ}(C_1, \dots, C_t)$

(Notice $|D| \le t \cdot s' + O(t)$)



$\mathcal{C}$ gets $\le \frac{1}{2} + \gamma/10$,
has size $< \delta(1-\gamma/2)2^n$

D does well on any x not in B:

$\forall x \notin B$,

$$\mathbb{P}_D\{D(x) \ne f(x)\} \le \exp(-\gamma^2 t)$$

by Chernoff.

Choose $t \approx \frac{1}{\gamma^2} \log(2/\gamma\delta)$ so

$\forall x \notin B$,

$$\mathbb{P}_D\{D(x) \ne f(x)\} \le \frac{\gamma\delta}{2}$$

But then $\mathbb{P}_{x,D}\{D(x) \ne f(x)\} < \frac{\gamma\delta}{2} + \delta(1 - \frac{\gamma}{2}) = \delta$

Hard-coding in D's random choices, we get a circuit of size

$$\Omega\left(\frac{1}{\gamma^2} \log(2/\gamma\delta)\right) \cdot s'$$

that is $(1-\delta)$-accurate on f $\lightning$

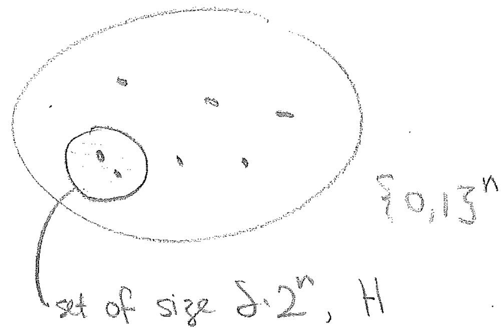
That contradicts the hardness of f, so the $V \ge 1/2 + \gamma$ case never could have happened, and we're done.

$\blacksquare$

Now we use our new understanding of hardness to prove the XOR lemma.

WTS that if f is $(1-\delta)$-hard, then $f^{\oplus k}$ is $\left(\frac{1}{2}+(1-\delta)^k+\gamma\right)$-hard.

Argue by contradiction: Suppose C has size s' that guesses $f^{\oplus k}$ w/ prob $\uparrow$.



$f^{\oplus k}$ takes k random inputs $x_1, \dots, x_k$. Some of them land in H.

Let $\mathcal{E}_\ell$ = event that (# x_i that land in H is ℓ)

(Now, $\mathbb{P}\{\mathcal{E}_0\} = (1-\delta)^k$ is pretty unlikely. So if C really does do well of $f^{\oplus k}$, it must do well on some $\mathcal{E}_\ell$.)

$$\frac{1}{2}+(1-\delta)^k+\gamma \overset{\text{assm TWOC}}{\leq} \mathbb{P}\{C(\vec{x}) = f^{\oplus}(\vec{x})\} \leq \underbrace{(1-\delta)^k}_{\mathbb{P}\{\mathcal{E}_0\}} + \mathbb{P}\{C(\vec{x}) = f^{\oplus k}(\vec{x}) \mid \overline{\mathcal{E}_0}\}$$

so $\exists \ell$ s.t. $\mathbb{P}\{C(x) = f^{\oplus k}(\vec{x}) \mid \mathcal{E}_\ell\} \geq \frac{1}{2}+\gamma$.

Using that ℓ, we'll contradict the hardness of f on H:

A: Given $x \in_R H$

Choose $z_1, \dots, z_{\ell-1} \in_R H$, $y_{\ell+1}, \dots, y_k \in_R \overline{H}$, $\pi \in_R S_n$

Return $C(\pi(z_1, \dots, z_{\ell-1}, x, y_{\ell+1}, \dots, y_k))$

Now, $$\mathbb{P}_{x, \underbrace{y, z, \pi}_{\text{hard-code these in to get a circuit } D \text{ w/}}}\{A(x) = f(z_1)\oplus \cdots \oplus f(z_{\ell-1}) \oplus f(x) \oplus f(y_{\ell+1}) \oplus \cdots \oplus f(y_k)\} \geq \frac{1}{2}+\gamma$$

now these guys are all fixed and can be hard-coded in.

$\mathbb{P}_{x \in_R H}\{D(x) = f(x)\} \geq \frac{1}{2}+\gamma$ $\lightning$