Today:

- Probabilistically Checkable Proof Systems
- Proofs of NP statements can be verified with $O(1)$ queries!

Useful Fact:

Given vectors $\bar{a} \neq \bar{b}$

$$\Pr_{\bar{r} \in \{0,1\}} [\, \bar{a} \cdot \bar{r} \neq \bar{b} \cdot \bar{r} \,] \geq 1/2$$

Given matrices A, B, C

if $A \cdot B \neq C$ then

$$\Pr_{\bar{r}} [\, \underbrace{A \cdot \underbrace{B \cdot \bar{r}}}_{O(n^2)\text{ time}} \neq C \cdot \bar{r} \,] \geq 1/2$$

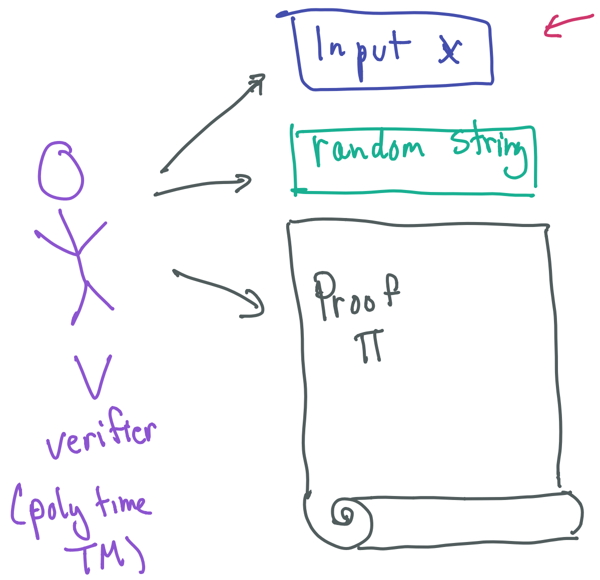
also true for equality mod 2

why?

Homework 1 optional problem 2

also: same argument used to show Fourier basis is orthogonal $\langle \chi_S, \chi_T \rangle = 0$ for $S \neq T$

Probabilistically Checkable Proofs



← theorem to prove } for today x is 3-CNF

Thm x is satisfiable

fixed fctn

differs from prover because can't adapt to V's questions

created by adversary with unlimited computational power

<u>def</u> $L \in PCP(r, q)$ if $\exists V$ s.t.

- poly time TM
- uses $\leq r(n)$ randombits
- uses $\leq q(n)$ queries to π ↑ 1 bit each

1) $\forall x \in L \quad \exists \pi$ s.t.

$$\Pr_{\text{V's random strings}}[\,V, \pi \text{ accepts}\,] = 1$$

2) $\forall x \notin L \quad \forall \pi' \quad \Pr_{\text{V's random strings}}[\,V, \pi' \text{ accepts}\,] < 1/4$

$$\mathrm{SAT} \in \mathrm{PCP}(0, \underset{\substack{\nwarrow \\ \text{look at all settings} \\ \text{of vars}}}{n})$$

<u>Today</u>:

<u>Thm</u> $\mathrm{NP} \subseteq \mathrm{PCP}(\underset{\substack{\uparrow \\ \$}}{O(n^3)}, \underset{\substack{\uparrow \\ \text{queries}}}{O(1)})$

<u>Actually</u>: <u>Thm</u> $\mathrm{NP} \subseteq \mathrm{PCP}(O(\log n), O(1))$

3SAT: $F = \bigwedge C_i$ s.t. $C_i = (y_{i_1} \vee y_{i_2} \vee y_{i_3})$

where $y_{ij} \in \{x_1 .. x_n \; \bar{x}_1 ... \bar{x}_n\}$

I. Encode satisfiability of F as a collection of polys in variables of assignment

- one for each clause
- low degree
- evaluate to 0 if assignment satisfies clause
- V knows coeffs – depend on structure of clause & vars of clause

Arithmetization of 3SAT:

boolean formula F $\Longleftrightarrow$ arithmetic formula $A(F)$ over $\mathbb{Z}_2$

$$\begin{aligned}
T &\longleftrightarrow 1\\
F &\longleftrightarrow 0\\
x_i &\longleftrightarrow x_i\\
\bar{x}_i &\longleftrightarrow 1-x_i\\
\alpha \wedge \beta &\longleftrightarrow \alpha \cdot \beta\\
\alpha \vee \beta &\longleftrightarrow 1-(1-\alpha)(1-\beta)\\
\alpha \vee \beta \vee \gamma &\longleftrightarrow 1-(1-\alpha)(1-\beta)(1-\gamma)
\end{aligned}$$

example:

$$\begin{aligned}
x_1 \vee \bar{x}_2 \vee x_3 \Longleftrightarrow \quad & 1-(1-x_1)(1-(1-x_2))(1-x_3)\\
&= 1-(1-x_1)(x_2)(1-x_3)
\end{aligned}$$

F satisfied by $\bar{a}$ iff $A(F)(\bar{a})=1$

Consider $\mathcal{C}(\bar{x}) = (\hat{C}_1(\bar{x}), \hat{C}_2(\bar{x}), \dots)$

Note: (1) Complements of arithmetization of clause C_i

$\Rightarrow$ evaluate to 0 if x satisfies C_i

(2) each $\hat{C}_i$ is $\deg \leq 3$ poly in x

(3) V knows coeffs of each $\hat{C}_i$

Need to convince V that

$$\mathcal{C}(\bar{a}) = (0, 0, \dots 0)$$

w/o sending $\bar{a}$

"weird idea"

assume $\exists$ "little birdie" who tells V dot products of $\mathcal{C}$ with random vectors mod 2.

(V inputs $\bar{r}$

birdie answers $\mathcal{C}(\bar{a}) \cdot \bar{r}$)

Fix $\bar{a}$

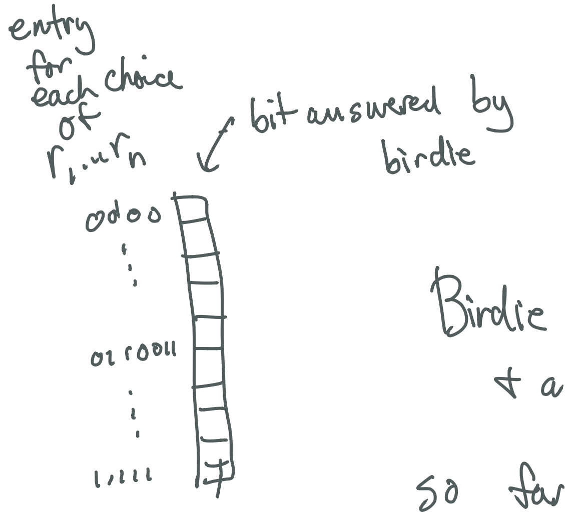
$$\left(\hat{C}_1(\bar{a}), \ldots, \hat{C}_m(\bar{a})\right) \cdot (r_1 \cdots r_m)$$

$$\equiv \sum r_i \hat{C}_i(\bar{a}) \bmod 2$$

$$\Pr\left[\sum r_i \hat{C}_i(\bar{a}) = 0\right] = \begin{cases} 1 & \text{if } \forall_i \ \hat{C}_i(\bar{a}) = 0 \\ 1/2 & \text{o.w.} \end{cases}$$

$$\left(\exists i \text{ s.t. } \hat{C}_i(\bar{a}) \neq 0 \Downarrow C(\bar{a}) \text{ not satisfied}\right)$$

At this point can write a very long proof



Birdie can cheat + always answer 0!!

so far - no check for consistency with $\hat{C}_i(\bar{a})$

So, why believe the birdie?

recall:

we know r_i's

we know coeff of polys of $\hat{C}_i$'s

$\hat{C}_i$'s have deg ≤ 3 in a_i's

we do not know a_i's

V doesnt know these

(mod 2)

$$\sum_i r_i \hat{C}_i(a) = \Gamma + \sum_i a_i \alpha_i + \sum_{ij} a_i a_j \beta_{ij} + \sum_{ijk} a_i a_j a_k \gamma_{ijk}$$

from here on:

$\alpha_i \to x_i$

$\beta_{ij} \to y_{ij}$

$\gamma_{ijk} \to z_{ijk}$

no reln to vars of 3SAT

- V knows these (s, does proof) depend on r_i's, coeffs of polys do not depend on a_i's
- since working mod 2, all values ∈ {0,1}

idea: make birdie write all answers for all choices of r_i's & check consistency (and later check satisfying the assignment)

We will do something stronger & easier to check

better idea

make birdie write out answers to all

3 seperate parts of proof

$$\left\{\begin{array}{llll} \text{linear} & \text{fctns.} & \text{of} & \bar{a} \\ \text{deg } 2 & \text{"} & \text{"} & \text{"} \\ \text{deg } 3 & \text{"} & \text{"} & \text{"} \end{array}\right.$$

- we only care about 1 lin / deg 2 / deg 3 fctn of $\bar{a}$

- will use to check that birdie wrote down a proper encoding of $\bar{a}$

V doesn't know

V knows

def $A : \mathbb{F}_2^n \to \mathbb{F}_2 \quad A(\bar{x}) = \sum a_i x_i = a^T \cdot \bar{x}$

$B : \mathbb{F}_2^{n^2} \to \mathbb{F}_2 \quad B(\bar{y}) = \sum_{i,j} a_i a_j y_{ij} = (\underbrace{a \circ a}_{\text{outer product}})^T \cdot \bar{y}$

if $z = b \circ c$

$z_{ij} := b_i \cdot c_j$

$C : \mathbb{F}_2^{n^3} \to \mathbb{F}_3 \quad C(\bar{z}) = \sum_{ijk} a_i a_j a_k z_{ijk}$

$= (a \circ a \circ a)^T \cdot \bar{z}$

Proof contains:

Complete description of truth tables of $\tilde{A}, \hat{B}, \tilde{C}$ for all inputs $\bar{x}, \bar{y}, \bar{z}$

supposed to be A, B, C

but V needs to check

What to check?

(1) $\tilde{A}, \hat{B}, \tilde{C}$ are of right form

- all are linear fctns ⇒ sc-$\tilde{A}$ will always answer according to closest lin fctn
  - linearity test + self-correct passes if $\tilde{A}$ close to linear
- correspond to same assignment $\bar{a}$
  - test all self-corrections consistent

(2) $\bar{a}$ is a sat assignment

all $\hat{C}_\alpha$'s evaluate to 0 on $\bar{a}$

How to do (1):

- Test $\hat{A}, \hat{B}, \underline{\tilde{C}}$ are all $\frac{1}{8}$ close to linear fctns

#random bits: $O(n^3)$
#queries $O(1)$
runtime $O(n^3)$

- · Pass if linear
- Fail if $\geq \frac{1}{8}$-far from linear

in $O(1)$ queries

per query to self-corr:
#randombits $O(n^3)$
queries $O(1)$
runtime $O(n^3)$

- From now on use self-corrector to get $sc\text{-}\hat{A}$, $sc\text{-}\hat{B}$, $\underline{sc\text{-}\tilde{C}}$ lin fctns can query on all inputs (use really small error bound on S-C s.t. if union bound over all calls to $Sc\hat{A}$ $Sc\hat{B}$ + $Sc\tilde{C}$ will never see error)

Consistency Test:

are sc-$\tilde{A}$, sc-$\tilde{B}$ + sc-$\tilde{C}$ from

<u>same</u> assignment $\bar{a}$?

Tester:

pick random $\bar{x}_1\ \bar{x}_2\ \ \bar{x}\ \bar{y}$

assume $\tilde{A}$ + $\tilde{B}$ + $\tilde{C}$ correspond to <u>same</u> $\bar{a}$

test that sc-$\tilde{A}(\bar{x}_1)\cdot$ sc$\tilde{A}(\bar{x}_2)$

$$= \sum a_i x_{1i} \cdot \sum a_j x_{2j}$$

$$= \sum_{ij} a_i a_j\ x_{1i} x_{2j}$$

$$= sc\text{-}\tilde{B}(\bar{x}_1 \circ \bar{x}_2)$$

randombits $O(n^2)$

queries $O(1)$

runtime $O(n^3)$

test that sc-$\tilde{A}(\bar{x})\cdot$ sc$\tilde{B}(\bar{y})$ =

$$= \sum_i a_i x_i \cdot \sum_{jk} a_j a_k y_{jk}$$

$$= \sum a_i a_j a_k\ x_i y_{jk}$$

$$= sc\text{-}\tilde{C}(\bar{x} \circ \bar{y})$$

<u>note</u>:

not unif dist queries

<u>but</u> s-c helps here

Is it a good test?

given sc-$\tilde{A}$, sc-$\hat{B}$, sc-$\tilde{C}$ $\Big\}$ all lin fctns

$A(x) = a^T x$

$B(y) = b^T y$

$C(y) = c^T z$

hopefully $b^T = (a \circ a)^T$

$c^T = (a \circ b)^T$

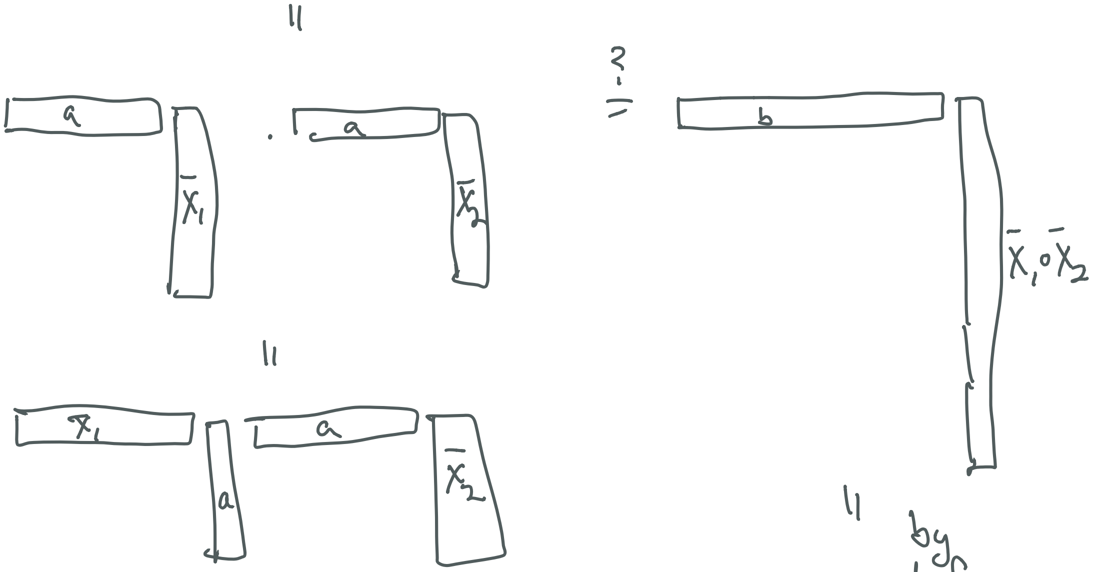
$= (a \circ a \circ a)^T$

If $b = a \circ a$, $c = a \circ a \circ a$ then test pass via green argument ✓

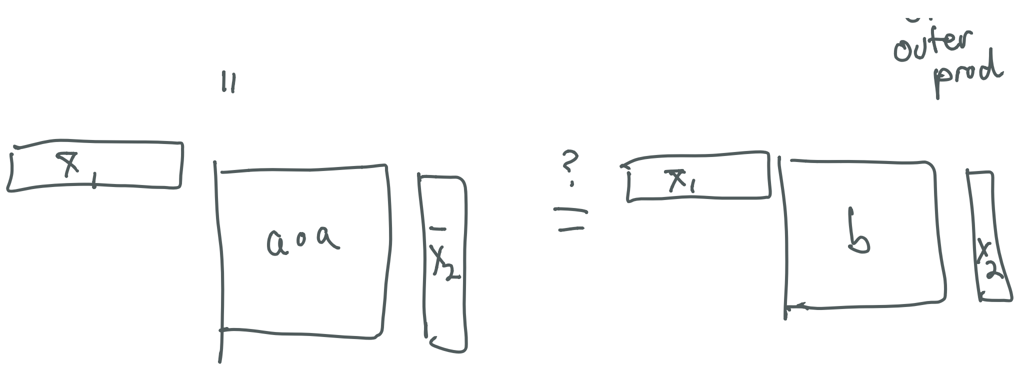
else $\boxed{\text{if } b \neq a \circ a}$

with what prob?

$$A(\bar{x}_1) \cdot A(\bar{x}_2) \overset{?}{=} B(\bar{x}_1 \circ \bar{x}_2)$$



by def of



if $a \circ a \neq b$ then

$$\Rightarrow \Pr_{x_2}\left[(a \circ a)\, x_2 \neq b \cdot x_2\right] \geq 1/2$$

$$\Pr_{x_1 x_2}\left[x_1 \cdot (a \circ a) \cdot x_2 \neq x_1 \cdot b \cdot x_2\right] \geq \frac{1}{2} \cdot \frac{1}{2}$$

$$\geq 1/4$$

so test fails with prob $\geq 1/4$

∎