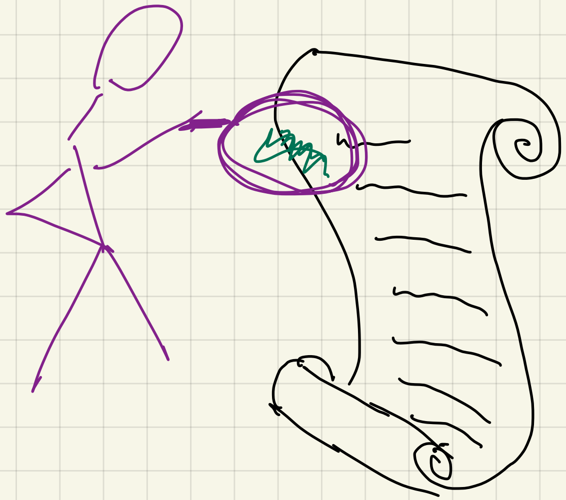
Probabilistically Checkable Proof Systems
(cont.)



linear fctn : $\forall x, y \quad f(x) + f(y) = f(x+y)$

Self-correcting :

if f is $\frac{1}{8}$ - close to linear g

Do $O(\log \frac{1}{\beta})$ times

Pick y randomly

$\text{answer}_i \leftarrow f(y) + f(x-y)$

Output most common answer_i

then

$\forall x, \Pr[\text{output} = g(x)] \geq 1 - \beta$

Self-testing: Given f

Do $O(\frac{1}{\varepsilon})$ times:

Pick x, y randomly

if $f(x) + f(y) \neq f(x+y)$ Fail

Pass

if f linear passes

if f ε-far from linear, fails

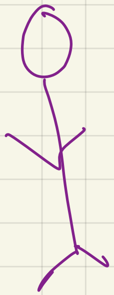
Probabilistically Checkable Proofs

$\boxed{\text{Input } x}$ ⟵ Theorem you want to prove

for today: x is 3CNF

Thm x is satisfiable

$\boxed{\text{random string}}$



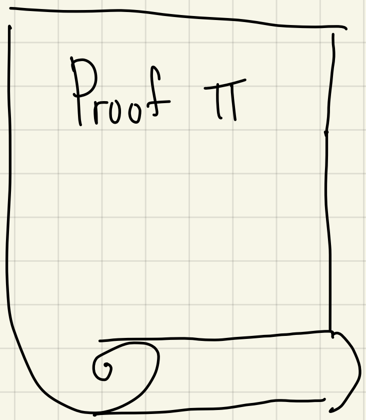
V

verifier

(poly time TM)

using r random bits

+ q queries to proof



fixed fctn

Verifier can query: what is ith bit?

Charged per query

proof doesn't change based on past questions of verifier

Created by adversary who knows verifier's algorithm + has unlimited computational power

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

1) $\forall x \in L \quad \exists\, \pi$ s.t. $\Pr_{v\text{'s random string}}[v, \pi \text{ accepts}] = 1$

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

e.g. $SAT \in PCP(0, n)$ ← proof settings of all n vars
V doesn't need any randomness

Today: $NP \subseteq PCP(O(n^3), O(1))$ ← crazy?

Actually: $NP \subseteq PCP(O(\log n), O(1))$

Let's start with a "warmup":

$x \cdot y = \sum x_i \cdot y_i$ "inner product"

$x \circ y = (x_1 y_1, x_1 y_2, x_1 y_3, \dots, x_i y_j, \dots, x_n y_n)$ "outer product"

$\underbrace{x \circ y}_{\text{n-bit vectors}}$, $\underbrace{(x_1 y_1, \dots, x_n y_n)}_{n^2 \text{ bit vector}}$

Fact: if $\underbrace{\bar{a} \neq \bar{b}}_{\text{n-bit vector}}$ then $\Pr_{\bar{r} \in_R \{0,1\}^n}[\bar{a} \cdot \bar{r} \neq \bar{b} \cdot \bar{r}] \geq \frac{1}{2}$

if $\underbrace{A \cdot B \neq C}_{n \times n \text{ matrices}}$ then $\Pr_{\bar{r}}[A \cdot \underbrace{B \cdot \bar{r}} \neq \underbrace{C \cdot \bar{r}}] \geq 1/2$

also true for "= mod 2"

$A \cdot (B \cdot \bar{r})$ take $O(n^2)$ to compute

Proof of fact: if $a_i \neq b_i$ pair n-bit strings that agree on all but i^{th} location

so pair $\begin{cases} \bar{r} = (r_1 \dots r_i \dots r_n) \\ \bar{s} = (r_1 \dots \bar{r}_i \dots r_n) \end{cases}$ then either $\bar{a} \cdot \bar{r} \neq \bar{b} \cdot \bar{r}$ or $\bar{a} \cdot \bar{s} \neq \bar{b} \cdot \bar{s}$

why?

$\bar{a} \cdot \bar{s} = (\bar{a} \cdot \bar{r}) \pm a_i$

$\bar{b} \cdot \bar{s} = (\bar{b} \cdot \bar{r}) \pm b_i$

different so $\bar{a} \cdot \bar{s} \neq \bar{b} \cdot \bar{s}$

$\Rightarrow$ In each pair at least one is $\neq$

note that this proof works "mod 2"

Fact: if $\bar{a} \neq \bar{b}$ then $\Pr_{\bar{r} \in_R \{0,1\}^n}[\bar{a}\cdot\bar{r} \neq \bar{b}\cdot\bar{r}] \geq \frac{1}{2}$

if $A \cdot B \neq C$ then $\Pr_{\bar{r}}[A \cdot B \cdot \bar{r} \neq C \cdot \bar{r}] \geq 1/2$

Example "application": setting: given vector $\bar{a} = (a_1, a_2, \ldots a_n)$

in <u>one</u> step:
- can query a_i
- can specify $\bar{y}$ & query $\underbrace{\bar{a}\cdot\bar{y}}$

what if these are all written for you? why believe correct?

to test if $\bar{a} \equiv (0, 0, \ldots, 0)$:

Do several times:

pick $\bar{r} \in_R \{0,1\}^n$

if $\bar{a}\cdot\bar{r} \neq 0$ output "Fail"

Output PASS

<u>behavior</u>: if $\bar{a} = (0, \ldots 0)$ will always PASS

if $\bar{a} \neq (0, \ldots 0)$ then FACT $\Rightarrow \Pr_{\bar{r}}[\bar{a}\cdot\bar{r} \neq 0] \geq 1/2$

$\Rightarrow O(1)$ query 0-testing algorithm for n-bit vector in strange model setting: given vector $\bar{a} = (a_1, a_2, \dots a_n)$

in one step: • can query a_i

• can specify $\bar{y}$ & query $\bar{a}\cdot\bar{y}$

Making the model "less strange":

first idea:

"Proof" = write out all answers to $\bar{a}\cdot\bar{y}$

$\bar{a}\cdot\bar{r}$	answer vector
$\bar{a}\cdot(0,0,\dots 0)$	0
$\bar{a}\cdot(0,0,\dots 1)$	0
$\bar{a}\cdot(0,0,\dots 1,0)$	0
$\bar{a}\cdot(0,0,\dots 1,1)$	⋮

to test if $\bar{a} \equiv (0,0,\dots,0)$:

Do several times:

pick $\bar{r} \in_R \{0,1\}^n$

ask proof for value of $\bar{a}\cdot\bar{r}$

if $\bar{a}\cdot\bar{r} \neq 0$ output "Fail"

Output PASS

Problem: proof can cheat

write all 0's in answer vector

How can we check that proof doesn't cheat?

test on $\bar{r}$'s that we know answer to?

is this easier than just looking at every entry of $\bar{a}$?

WILL COME BACK TO THIS

3SAT:

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

where $y_{i_j} \in \{ x_1 \cdots x_n, \bar{x}_1 \cdots \bar{x}_n \}$

here use $\bar{x}$ notation for complement

First crack:

π = setting of sat assignment a

$a_1 = T \quad a_2 = F \quad a_3 = T \ldots$ $\boxed{1 \ 0 \ 1 \ \ldots}$

V's protocol given formula + a:

Pick random clause C_i + check if $\bar{a}$ satisfies

good? $\bar{a}$ satisfies $\bar{c}$ ✓

$\bar{a}$ doesn't satisfy $\bar{c} \Rightarrow \exists i$ s.t. $C_i(\bar{a}) \neq T$

Pick i with prob $= \frac{1}{\# \text{clauses}}$ ☹

$F = (x_1 \vee \bar{x}_2 \vee x_3)(x_2 \vee \bar{x}_3 \vee x_4)$

$\bar{a} = (x_1 = T, x_2 = F, x_3 = F, x_4 = F, \ldots)$

random clause $(x_2 \vee \bar{x}_3 \vee x_4)$ ✓

$\quad F \quad T \quad F$

Arithmetization of 3SAT:

Boolean formula F $\iff$ arithmetic formula $A(F)$ over $\underbrace{\mathbb{Z}_2}_{\text{mod } 2}$

$$T \iff 1$$

$$F \iff 0$$

$$x_i \iff x_i$$

$$\bar{x}_i \iff 1 - x_i$$

$$\alpha \wedge \beta \iff \alpha \cdot \beta$$

$$\alpha \vee \beta \iff 1 - (1-\alpha)(1-\beta)$$

$$\alpha \vee \beta \vee \gamma \iff 1 - (1-\alpha)(1-\beta)(1-\gamma)$$

<u>example</u>: $x_1 \vee \bar{x}_2 \vee x_3 \iff 1 - (1-x_1)\overbrace{(x_2)}^{1-(1-x_2)}(1-x_3)$

<u>Key point</u> F satisfied by assignment a iff $[A(F)](a) = 1$

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

where $y_{i_j} \in \{ x_1 \dots x_n, \bar{x}_1 \dots \bar{x}_n \}$

$T \Leftrightarrow 1$

$F \Leftrightarrow 0$

$x_i \Leftrightarrow x_i$

$\bar{x}_i \Leftrightarrow 1 - x_i$

$\alpha \wedge \beta \Leftrightarrow \alpha \cdot \beta$

$\alpha \vee \beta \Leftrightarrow 1 - (1-\alpha)(1-\beta)$

$\alpha \vee \beta \vee \gamma \Leftrightarrow 1 - (1-\alpha)(1-\beta)(1-\gamma)$

Consider $C^\sigma(x) = (\hat{C}_1(x), \hat{C}_2(x), \dots)$

s.t. $\hat{C}_i(x)$ = complement of arithmetization of clause C_i

$\Rightarrow$ evaluates to 0 if x satisfies C_i

$\Rightarrow C^\sigma(x) = (0, \dots 0)$ if x satifies F

Observe (1) each $\hat{C}_i$ is deg ≤ 3 poly in x

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

Need to convince V that $C^\sigma(a) = (\hat{C}_1(a), \hat{C}_2(a) \dots) = (0, \dots 0)$ WITHOUT SENDING assignment a

High level idea: special encoding of assignment

Encode satisfiability of F as a collection of polys in vars of assignment

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

Note: We are only concerned that V is poly time, here will not be sublinear

note that solving SAT in poly time would be impressive ☺

However, want # queries to proof to be <u>constant</u>

Idea for proof:

- proof contains $\mathcal{C}(a)\cdot r \quad \forall \; r \in \{0,1\}^n$
- if $\forall\, i,\ \hat{C}_i(a)=0$, $\Pr_r[\mathcal{C}(a)\cdot r=0]=1$

  if $\exists\, i$ s.t. $\hat{C}_i(a)\neq 0$, $\Pr[\mathcal{C}(a)\cdot r=0]=\frac{1}{2}$

  $\qquad = \Pr[\mathcal{C}(a)\cdot r=1]$

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

where $y_{i_j} \in \{x_1 \dots x_n, \bar{x}_1 \dots \bar{x}_n\}$

$\mathcal{C}(a) = (\underbrace{\hat{C}_1(a)}_{\text{complement}}, \hat{C}_2(a), \dots) \overset{?}{=} (0,0,\dots 0)$

$T \Leftrightarrow 1$

$F \Leftrightarrow 0$

$x_i \Leftrightarrow x_i$

$\bar{x}_i \Leftrightarrow 1-x_i$

$\alpha \wedge \beta \Leftrightarrow \alpha\cdot\beta$

$\alpha \vee \beta \Leftrightarrow 1-(1-\alpha)(1-\beta)$

$\alpha \vee \beta \vee \gamma \Leftrightarrow 1-(1-\alpha)(1-\beta)(1-\gamma)$

mod 2 arithmetic

challenge: proof could write 0's even if $\mathcal{C}(a)\cdot r \neq 0$

so need to do more

What does $\mathcal{C}(a)\cdot r$ look like?

V doesn't know

$$\sum_i r_i \hat{\mathcal{C}}_i(a) = \Gamma + \sum_i a_i \alpha_i + \sum_{i,j} a_i a_j \beta_{ij} + \sum_{i,j,k} a_i a_j a_k \gamma_{ijk} \pmod 2$$

V does know: depend on r_i's + coefts in $\hat{C}_i$

from here on:

$\alpha_i \to x_i$

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

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

no relation to vars of 3SAT!!!

example

$$G = (x_1 \vee x_2) \wedge (\overline{x}_1 \vee x_2)$$

evaluates to 1 if C_1 satisfied

$$A(C_1) = 1 - (1-x_1)(1-x_2) = x_1 + x_2 - x_1 x_2$$

$$\Rightarrow \quad C_1(a) = 1 - a_1 - a_2 + a_1 a_2$$

evals to 0 if C_i satisfied

$$A(C_2) = 1 - (x_1)(1-x_2) = 1 - x_1 + x_1 x_2$$

$$\Rightarrow \quad C_2(a) = a_1 - a_1 a_2$$

$$\sum r_i \cdot C_i(a) = r_1 (1 - a_1 - a_2 + a_1 a_2) + r_2 (a_1 - a_1 a_2)$$

$$= \underbrace{r_1 \cdot 1 + r_2 \cdot 0}_{\text{deg } 0} + \underbrace{(-r_1 + r_2) \cdot a_1 + (-r_1) a_2}_{\text{deg } 1} + \underbrace{(r_1 - r_2)}_{\text{deg } 2} \cdot a_1 a_2$$

Freivald's trick over all clauses (should be 0 for all choices of r_i's)

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

where $y_{i_j} \in \{ x_1 \cdots x_n, \overline{x}_1 \cdots \overline{x}_n \}$

$$\hat{C}(a) = (\underbrace{\hat{C}_1}_{\text{complement}}(a), \hat{C}_2(a), \ldots) \overset{?}{=} (0, 0, \ldots 0)$$

$T \Leftrightarrow 1$

$F \Leftrightarrow 0$

$x_i \Leftrightarrow x_i$

$\overline{x}_i \Leftrightarrow 1 - x_i$

$\alpha \wedge \beta \Leftrightarrow \alpha \cdot \beta$

$\alpha \vee \beta \Leftrightarrow 1 - (1-\alpha)(1-\beta)$

$\alpha \vee \beta \vee \gamma \Leftrightarrow 1 - (1-\alpha)(1-\beta)(1-\gamma)$

$$\sum_i r_i \hat{C}_i(a) = r + \sum_i a_i \alpha_i + \sum_{i,j} a_i a_j \beta_{ij} + \sum_{i,j,k} a_i a_j a_k \gamma_{ijk} \pmod 2$$

r_1	r_2	$\sum r_i C_i(a)$	sat case $a^+ = (0,1)$	unsat case $a^- = (0,0)$
0	0	0	0	0
0	1	$a_1 - a_1 a_2$	0	0
1	0	$1 - a_1 - a_2 + a_1 a_2$	$1-0-1+0=0$	$1-0-0+0=1$
1	1	$1 - a_2$	$1-1=0$	$1-0=1$

High level idea: Special encoding of assignment

- proof writes out all linear fctns of assignment
  deg 2
  deg 3

- possible "confusion": "symmetric" for linear case

$$f_x(a) = x \cdot a = A_a(x)$$

inner product

- for deg 2,3: $B_a(y) = (a \circ a)^T \cdot y$

$$C_a(y) = (a \circ a \circ a)^T \cdot z$$

A_a, B_a, C_a are all linear fctns ⟹ can test linearity & self-correct

Proof can cheat!
- what if A_a, B_a, C_a come from different assignments
- is a satisfying?

These are fctns (hopefully all of same a) + we only care about their values at one input corresponding to what V computes from coefficients of deg 3 poly's + r_i's

def

A = all linear fctns evaluated at assignment a

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

V knows x, y, z but not a

B = all deg 2 fctns evaluated at a

$B : \mathbb{F}_2^{n^2} \to \mathbb{F}_2$ $\qquad B(y) = \sum_{i,j} a_i a_j y_{ij} = (a \circ a)^T \cdot y$

C = all deg 3 fctns evaluated at a

$C : \mathbb{F}_2^{n^3} \to \mathbb{F}_2$ $\qquad C(y) = \sum_{ijk} a_i a_j a_k z_{ijk} = (a \circ a \circ a)^T \cdot z$

recall: $x \circ y = (x_1 y_1, x_1 y_2, x_1 y_3, \dots, x_i y_j, \dots, x_n y_n)$

Proof contains:

hopefully A, B, C but we need to check

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

only need value at $x = \alpha$, $y = \beta$, $z = \gamma$ but extra info helps us check consistency

def

$x \circ y = (x_1y_1, x_1y_2, x_1y_3, \ldots, x_iy_j, \ldots, x_ny_n)$

A = all linear fctns evaluated at assignment a

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

B = all deg 2 fctns evaluated at a

$B: \mathbb{F}_2^{n^2} \to \mathbb{F}_2 \qquad B(y) = \sum_{i,j} a_i a_j y_{ij} = (a \circ a)^T \cdot y$

C = all deg 3 fctns evaluated at a

$C: \mathbb{F}_2^{n^3} \to \mathbb{F}_2 \qquad C(y) = \sum_{ijk} a_i a_j a_k z_{ijk} = (a \circ a \circ a)^T \cdot z$

Proof contains: **HUGE**

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

only need value at $x=\alpha,\ y=\beta,\ z=\gamma$ but extra info helps us check consistency

---

What does verifier need to check in proof?

(1) $\tilde{A}, \tilde{B}, \tilde{C}$ in right form (good encoding)

- all are linear fctns

Can only test ε-close to linear but can self-correct to access the linear fctns.

- correspond to same assignment a

i.e. $\tilde{A}(x) = a^T \cdot x \Rightarrow \tilde{B}(y) = (a \circ a)^T \cdot y \Rightarrow \tilde{C}(z) = (a \circ a \circ a)^T \cdot z$

Test consistency of self corrections

(2) a is satisfying assignment (check the entries in encoding which give value of $\mathcal{C}(a)$)

- all $\hat{C}_i$'s evaluate to 0 on a

(recall $\mathcal{C}(a) = (\underbrace{\hat{C}_1}_{\text{complement}}(a), \hat{C}_2(a), \ldots) \stackrel{?}{=} (0, 0, \ldots 0)$)

$x \circ y = (x_1 y_1, x_1 y_2, x_4 y_3, \ldots, x_i y_j, \ldots, x_n y_n)$

def

A = all linear fctns evaluated at assignment a

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

B = all deg 2 fctns evaluated at a

$B: \mathbb{F}_2^{n^2} \to \mathbb{F}_2 \qquad B(y) = \sum_{i,j} a_i a_j y_{ij} = (a \circ a)^T \cdot y$

C = all deg 3 fctns evaluated at a

$C: \mathbb{F}_2^{n^3} \to \mathbb{F}_2 \qquad C(y) = \sum_{ijk} a_i a_j a_k z_{ijk} = (a \circ a \circ a)^T \cdot z$

Proof contains:

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

only need value at $x = \alpha$, $y = \beta$, $z = \gamma$ but extra info helps us check consistency

---

Test (1) $\tilde{A}, \tilde{B}, \tilde{C}$ in right form: all are linear fctns

Can only test ε-close to linear but can self-correct to access the linear fctns.

- Test $\tilde{A}, \tilde{B}, \tilde{C}$ are all $1/8$-close to linear (i.e. if all linear, PASS; if any one is $1/8$-far FAIL) in $O(1)$ queries
- From now on, use self corrector to get

$sc\text{-}\tilde{A}, \quad sc\text{-}\tilde{B}, \quad sc\text{-}\tilde{C}$ for all inputs

$\Updownarrow \qquad \Updownarrow \qquad \Updownarrow$

$a \qquad b = a \circ a? \qquad c = a \circ a \circ a?$

use β = prob of getting wrong answer in SC that is so small ($\leq \frac{1}{\text{big enough constant}}$) that union bnd over all queries to $sc\text{-}\tilde{A}, sc\text{-}\tilde{B}, sc\text{-}\tilde{C}$

$\Rightarrow$ unlikely to see error

def $\quad x \circ y = (x_1 y_1, x_1 y_2, x_4 y_3, \dots, x_i y_j, \dots, x_n y_n)$

A = all linear fctns evaluated at assignment a $\quad A: \mathbb{F}_2^n \to \mathbb{F}_2 \qquad A(x) = \sum a_i x_i = a^T \cdot x$

B = all deg 2 fctns evaluated at a $\quad B: \mathbb{F}_2^{n^2} \to \mathbb{F}_2 \qquad B(y) = \sum_{i,j} a_i a_j y_{ij} = (a \circ a)^T \cdot y$

C = all deg 3 fctns evaluated at a $\quad C: \mathbb{F}_2^{n^3} \to \mathbb{F}_2 \qquad C(y) = \sum_{ijk} a_i a_j a_k z_{ijk} = (a \circ a \circ a)^T \cdot z$

Proof contains:

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

only need value at $x = \alpha$, $y = \beta$, $z = \gamma$ but extra info helps us check consistency

---

Test (1) $\tilde{A}, \tilde{B}, \tilde{C}$ in right form:

- all are linear fctns
- correspond to same assignment a

ie. $\tilde{A}(x) = a^T \cdot x \Rightarrow \tilde{B}(y) = (a \circ a)^T \cdot y \Rightarrow \tilde{C}(z) = (a \circ a \circ a)^T \cdot z$

Test consistency of self corrections

Goal: Pass if $sc\text{-}\tilde{B} = sc\text{-}\tilde{A} \circ sc\text{-}\tilde{A}$

$sc\text{-}\tilde{C} = sc\text{-}\tilde{A} \circ sc\text{-}\tilde{B}$

test $sc\text{-}\tilde{A}$ & $sc\text{-}\tilde{B}$ correspond to same a_i's

Outer Product Tester: Pick random x_1, x_2, x, y

Test $\quad sc\text{-}\tilde{A}(x_1) \cdot sc\text{-}\tilde{A}(x_2) = \left[\sum a_i x_{1i} \circ \sum a_j x_{2j} = \sum_{i,j} a_i a_j x_{1i} x_{2j} = \sum_{i,j} b_{ij} x_{1i} x_{2j}\right]$

$= sc\text{-}\tilde{B}(x_1 \circ x_2)$ (*)

$sc\text{-}\tilde{A}(x) \cdot sc\text{-}\tilde{B}(y) = \left[\sum a_i x_i \circ \sum_{jk} b_{jk} y_{jk} = \sum_{i,j,k} a_i b_{jk} x_i y_{jk} = \sum a_i a_j a_k x_i y_{jk}\right]$

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

(*) = not uniformly distributed

def

A = all linear fctns evaluated at assignment a — $A : \mathbb{F}_2^n \to \mathbb{F}_2$ — $A(x) = \sum a_i x_i = a^T \cdot x$

B = all deg 2 fctns evaluated at a — $B : \mathbb{F}_2^{n^2} \to \mathbb{F}_2$ — $B(y) = \sum_{i,j} a_i a_j y_{ij} = (a \circ a)^T \cdot y$

C = all deg 3 fctns evaluated at a — $C : \mathbb{F}_2^{n^3} \to \mathbb{F}_2$ — $C(y) = \sum_{ijk} a_i a_j a_k z_{ijk} = (a \circ a \circ a)^T \cdot z$

Proof contains:

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

only need value at $x = \alpha$, $y = \beta$, $z = \gamma$ but extra info helps us check consistency

Test $\quad$ sc-$\tilde{A}(x_1) \cdot$ sc-$\tilde{A}(x_2) = \left[\sum a_i x_{1i} \circ \sum a_j x_{2j} = \sum_{i,j} a_i a_j x_{1i} x_{2j} = \sum_{ij} b_{ij} x_{1i} x_{2j}\right]$

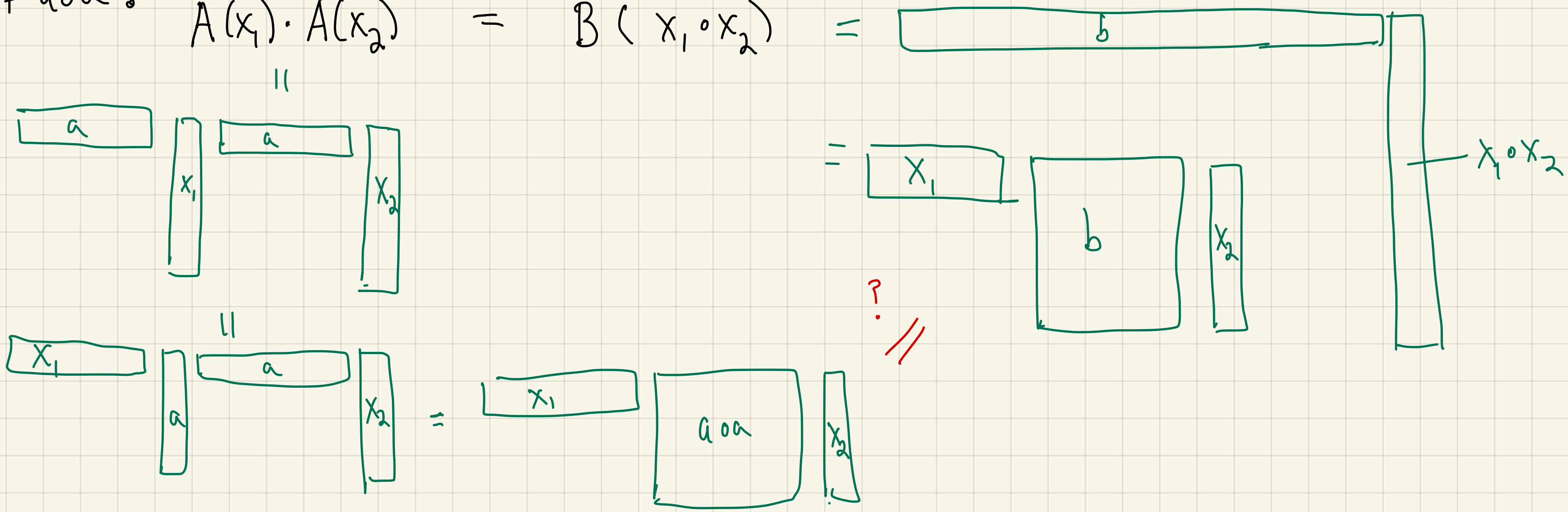
picked randomly

$= \text{sc-}\tilde{B}(x_1 \circ x_2)$

if $b = a \circ a$ test passes $\leftarrow$ since "blue" equalities hold

if $b \neq a \circ a$:

$$A(x_1) \cdot A(x_2) = B(x_1 \circ x_2) =$$



Fact: if $\bar{a} \neq \bar{b}$ then $\Pr_{\bar{r} \in_R \{0,1\}^n}[\bar{a} \cdot \bar{r} \neq \bar{b} \cdot \bar{r}] \geq \frac{1}{2}$

if $A \cdot B \neq C$ then $\Pr_{\bar{r}}[A \cdot B \cdot \bar{r} \neq C \cdot \bar{r}] \geq 1/2$

if $b \neq a \circ a$:

What is prob

$$x_1 \cdot (a \circ a) \cdot x_2 = x_1 \cdot b \cdot x_2 \quad ?$$

$(a \circ a) \cdot x_2 \stackrel{?}{=} b \cdot x_2$

Yes: Pass with prob 1

No: Pass with prob ?

Fact $\Rightarrow$ $\Pr_{x_2}[(a \circ a) \cdot x_2 \neq b \cdot x_2] = 1/2$

if $(a \circ a) \cdot x_2 \neq b \cdot x_2$

then Fact $\Rightarrow$ $\Pr_{x_1}[x_1 \cdot (a \circ a) \cdot x_2 \neq x_1 \cdot b \cdot x_2] = 1/2$

$\Rightarrow \Pr[\text{fail test}] \geq 1/4$

Test: $SC\text{-}\tilde{A}(x_1) \cdot SC\text{-}\tilde{A}(x_2) = \left[\sum a_i x_{1i} \circ \sum a_j x_{2j} = \sum_{i,j} a_i a_j x_{1i} x_{2j} = \sum_{i,j} b_{ij} x_{1i} x_{2j}\right]$

$= SC\text{-}\tilde{B}(x_1 \circ x_2)$

picked randomly

So passing test $\Rightarrow$ safe to assume $b = a \circ a$!

Similarly passing other test $\Rightarrow$ safe to assume $c = a \circ a \circ a$

def

$$x \circ y = (x_1y_1, x_1y_2, x_1y_3, \dots, x_iy_j, \dots, x_ny_n)$$

A = all linear fctns evaluated at assignment a

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

B = all deg 2 fctns evaluated at a

$$B: \mathbb{F}_2^{n^2} \to \mathbb{F}_2 \qquad B(y) = \sum_{i,j} a_i a_j y_{ij} = (a \circ a)^T \cdot y$$

C = all deg 3 fctns evaluated at a

$$C: \mathbb{F}_2^{n^3} \to \mathbb{F}_2 \qquad C(y) = \sum_{ijk} a_i a_j a_k z_{ijk} = (a \circ a \circ a)^T \cdot z$$

Proof contains:

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

only need value at $x = \alpha$, $y = \beta$, $z = \gamma$ but extra info helps us check consistency

---

Test (1) $\tilde{A}, \tilde{B}, \tilde{C}$ in right form:

- all are linear fctns
- correspond to same assignment a

i.e. $\tilde{A}(x) = a^T \cdot x \Rightarrow \tilde{B}(y) = (a \circ a)^T \cdot y \Rightarrow \tilde{C}(z) = (a \circ a \circ a)^T \cdot z$

Test consistency of self corrections

Goal: Pass if $sc\text{-}\tilde{B} = sc\text{-}\tilde{A} \circ sc\text{-}\tilde{A}$

$sc\text{-}\tilde{C} = sc\text{-}\tilde{A} \circ sc\text{-}\tilde{B}$

test $sc\text{-}\tilde{A}$ & $sc\text{-}\tilde{B}$ correspond to same a_i's

Outer Product Tester: Pick random x_1, x_2, x, y

Test $sc\text{-}\tilde{A}(x_1) \cdot sc\text{-}\tilde{A}(x_2) = \left[\sum a_i x_{1i} \circ \sum a_j x_{2j} = \sum_{i,j} a_i a_j x_{1i} x_{2j} = \sum_{i,j} b_{ij} x_{1i} x_{2j}\right]$

$= sc\text{-}\tilde{B}(x_1 \circ x_2)$ (*)

$sc\text{-}\tilde{A}(x) \cdot sc\text{-}\tilde{B}(y) = \left[\sum a_i x_i \circ \sum_{jk} b_{jk} y_{jk} = \sum_{i,j,k} a_i b_{jk} x_i y_{jk} = \sum a_i a_j a_k x_i y_{jk}\right]$

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

(*) = not uniformly distributed

$x \circ y = (x_1 y_1, x_1 y_2, x_4 y_3, \dots, x_i y_j, \dots, x_n y_n)$

def

A = all linear fctns evaluated at assignment a

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

B = all deg 2 fctns evaluated at a

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

C = all deg 3 fctns evaluated at a

$C : \mathbb{F}_2^{n^3} \to \mathbb{F}_2$ $\quad C(y) = \sum_{ijk} a_i a_j a_k z_{ijk} = (a \circ a \circ a)^T \cdot z$

Proof contains:

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

only need value at $x = \alpha$, $y = \beta$, $z = \gamma$ but extra info helps us check consistency

---

Test (1) $\tilde{A}, \tilde{B}, \tilde{C}$ in right form: all are linear fctns

Can only test ε-close to linear
but can self-correct to access the linear fctns.

- Test $\tilde{A}, \tilde{B}, \tilde{C}$ are all $1/8$-close to linear (i.e. if all linear, PASS; if any one is $1/8$-far FAIL) in $O(1)$ queries
- From now on, use self corrector to get

$sc\text{-}\tilde{A}$, $sc\text{-}\tilde{B}$, $sc\text{-}\tilde{C}$ for all inputs

$\Updownarrow$ a $\quad \Updownarrow$ $b = a \circ a$? $\quad \Updownarrow$ $c = a \circ a \circ a$?

use β = prob of getting wrong answer in SC that is so small ($\le \frac{1}{\text{big enough constant}}$) that union bnd over all queries to $sc\text{-}\tilde{A}$, $sc\text{-}\tilde{B}$, $sc\text{-}\tilde{C}$
$\Rightarrow$ unlikely to see error

$x \circ y = (x_1y_1, x_1y_2, x_1y_3, \dots, x_iy_j, \dots, x_ny_n)$

def

A = all linear fctns evaluated at assignment a

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

B = all deg 2 fctns evaluated at a

$B : \mathbb{F}_2^{n^2} \to \mathbb{F}_2 \qquad B(y) = \sum_{i,j} a_i a_j y_{ij} = (a \circ a)^T \cdot y$

C = all deg 3 fctns evaluated at a

$C : \mathbb{F}_2^{n^3} \to \mathbb{F}_2 \qquad C(y) = \sum_{ijk} a_i a_j a_k z_{ijk} = (a \circ a \circ a)^T \cdot z$

Proof contains: HUGE

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

only need value at $x = \alpha$, $y = \beta$, $z = \gamma$ but extra info helps us check consistency

---

What does verifier need to check in proof?

(1) $\tilde{A}, \tilde{B}, \tilde{C}$ in right form

- all are linear fctns

Can only test ε-close to linear but can self-correct to access the linear fctns.

- correspond to same assignment a

i.e. $\tilde{A}(x) = a^T \cdot x \Rightarrow \tilde{B}(y) = (a \circ a)^T \cdot y \Rightarrow \tilde{C}(z) = (a \circ a \circ a)^T \cdot z$

Test consistency of self corrections

random bits $= O(n^3)$

queries $= O(1)$

(2) a is satisfying assignment

- all $\hat{C}_i$'s evaluate to 0 on a

(recall $C(a) = (\hat{C}_1(a), \hat{C}_2(a), \dots) \overset{?}{=} (0, 0, \dots 0)$)

complement

$$\sum_i r_i \hat{C}_i(a) = \Gamma + \sum_i a_i \alpha_i + \sum_{i,j} a_i a_j \beta_{ij} + \sum_{i,j,k} a_i a_j a_k \gamma_{ijk} \pmod 2$$

How to do (2):

- Call self-correctors $\Rightarrow$ recover linear fctns a, $a \circ a$, $a \circ a \circ a$
- a represents assignment, but we don't know it
- a satisfying $\Leftrightarrow$ $\hat{C}(a) = (\hat{C}_1(a), \hat{C}_2(a), \dots) = (0, 0, \dots, 0)$

Satisfiability Test:

\# random bits $= O(n)$

Pick $r \in \mathbb{F}_2^n$

Compute Γ, $\underset{\downarrow\, x_i\text{'s}}{\alpha_i\text{'s}}$, $\underset{\downarrow\, y_{ij}\text{'s}}{\beta_{ij}\text{'s}}$, $\underset{\downarrow\, z_{ijk}\text{'s}}{\gamma_{ijk}\text{'s}}$ $\leftarrow$ fctns of r & of coeffs of deg 3 polys

\# queries $= O(1)$

query proof to get

$$SC\text{-}\tilde{A}(\alpha_1 \dots \alpha_n) = w_0$$
$$SC\text{-}\tilde{B}(\beta_{11} \dots \beta_{nn}) = w_1$$
$$SC\text{-}\tilde{C}(\gamma_{111} \dots \gamma_{nnn}) = w_2$$

Verify $0 = \Gamma + w_0 + w_1 + w_2 \pmod 2$

$\uparrow$ hopefully means $\sum r_i \hat{C}_i(a) = 0$

Why do this?

if $\forall i$ $\hat{C}_i(a) = 0$

$\Pr[\text{pass}] = 1$ ✓

if $\exists i$ s.t. $\hat{C}_i(a) \neq 0$

Fact $\Rightarrow$ $\Pr[\sum_j r_j \hat{C}_j(a) = 0] = \frac{1}{2}$

so after k times

$\Pr[\text{pass}] = \frac{1}{2^k}$ ✓