Today:

- Probabilistically Checkable Proof Systems
- Proofs of NP statements can be verified with QCDA queries!
Useful Fact:

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

$\Pr [r \cdot \overline{a} \cdot \overline{r} + \overline{b} \cdot \overline{r}] \approx \frac{1}{2}$

Given matrices $A, B, C$

if $A \cdot B \neq C$ then

$\Pr [r \cdot A \cdot \overline{B} \cdot \overline{r} + \overline{C} \cdot \overline{r}] \approx \frac{1}{2}$

$O(n^2)$ time

Why?

Homework 1 optional problem 2
also: same argument used to
show Fourier basis is orthogonal

$\langle X_s, X_T \rangle = 0$ for $s \neq T$
Probabilistically Checkable Proofs

- Input $x$
- Random String
- Proof $\Pi$
- Fixed function differs from prover because can't adapt to verifier's questions created by adversary with unlimited computational power

**Definition:** $L \in \text{PCP}(r, q)$ if $\exists V$ s.t.

1. $\forall x \in L \exists \Pi$ s.t.,
   \[
   \Pr_{\Pi}[V, \Pi \text{ accepts}] = 1
   \]
   $V$'s random strings

2. $\forall x \notin L \forall \Pi' \Pr_{\Pi}[V, \Pi' \text{ accepts}] < \frac{1}{4}$

- Theorem 2: For today, $x$ is 3-CNF
- The $x$ is satisfiable
\[ \text{SAT } \in \text{PCP}(O(1^n)) \]

\text{look at all settings of vars}

\text{Today:}

\[ \text{Thm} \quad \text{NP } \leq \text{PCP}(O(n^3), O(1)) \]

\text{queries}

\text{Actually:}

\[ \text{Thm} \quad \text{NP } \leq \text{PCP}(O(\log n), O(1)) \]

3SAT: \[ F = \bigwedge C_i \text{ s.t. } C_i = (y_{i_1} V y_{i_2} V y_{i_3}) \]

where \( y_{i_j} \in \{1, 0\} \) \( x, \overline{x}, x^n \)

I. Encode satisfiability of \( F \) as a collection of polynomials in variables of assignment

- one for each clause
- low degree
- evaluate to 0 if assignment satisfies clause
- \( V \) knows coefficients depend on structure of clause \& vars of clause
Arithmetization of ZSAT:

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

\[
\begin{align*}
T & \leftrightarrow 1 \\
F & \leftrightarrow 0 \\
X_i & \leftrightarrow x_i \\
\overline{X_i} & \leftrightarrow 1 - x_i \\
\overline{\alpha \land \beta} & \leftrightarrow \alpha \cdot \beta \\
\overline{\alpha \lor \beta} & \leftrightarrow 1 - (1 - \alpha)(1 - \beta) \\
\overline{\alpha \lor \beta \lor \gamma} & \leftrightarrow 1 - (1 - \alpha)(1 - \beta)(1 - \gamma) \\
\overline{(\alpha \lor \beta \lor \gamma)} & \leftrightarrow (1 - \alpha)(1 - \beta)(1 - \gamma)
\end{align*}
\]

Example:

\[
\begin{align*}
X_1 \lor \overline{X_2} \lor X_3 & \leftrightarrow 1 - (1 - x_1)(1 - (1 - x_2))(1 - x_3) \\
& = 1 - (1 - x_1)(x_2)(1 - x_3)
\end{align*}
\]

$F$ satisfied by $\overline{a}$ iff $A(F)(\overline{a}) = 1$
Consider $C^0(\bar{x}) = (\hat{C}_1(\bar{x}), \hat{C}_2(\bar{x}), \ldots)$

Note:
1. Complements of arithmetization of clause $C_i$ 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 $C^0(\bar{a}) = (0, 0, \ldots, 0)$
who sending $\bar{a}$

"weird idea"
assume $F$ "little birdie" who tells $V$
dot products of $C$ with random vectors mod 2
($V$ inputs $\bar{r}$
birdie answers $C(\bar{a}) \cdot \bar{r}$)
Fix $\bar{a}$

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

$$= \sum r_i \hat{C}_i(\bar{a}) \mod 2$$

$$\Pr \left[ \sum r_i \hat{C}_i(\bar{a}) = 0 \right] = \begin{cases} \frac{1}{2} & \text{if } \forall i \; \hat{C}_i(\bar{a}) = 0 \\ \frac{3}{2} & \text{otherwise} \\ \forall i \; \text{s.t. } \hat{C}_i(\bar{a}) \neq 0 \\ \text{closed } (\bar{a}) \text{ not satisfied} \end{cases}$$

At this point can write a very long proof.

entry for each choice of $r_1 \ldots r_m$ bit answered by birdie

Birdie can cheat and always answer 0!!!
so far - no check for consistency with $\hat{C}_m(\bar{a})$

So, why believe the birdie?
Recall:
we know $r_i$'s
we know coeff of polys of $c_i$'s
$c_i$'s have $\deg \leq 3$ in $a_i$'s
we do not know $a_i$'s

$$\sum_i r_i c_i(a) = 1 + \sum_i a_i \alpha_i + \sum_{ij} a_i a_j \beta_{ij} + \sum_{ijk} a_i a_j a_k \gamma_{ijk}$$

From here on:
$$a_i \to x_i$$
$$\beta_{ij} \to y_{ij}$$
$$\gamma_{ijk} \to z_{ijk}$$

- $V$ knows these (so does proof)
- depend on $r_i$'s coeff of polys
- do not depend on $a_i$'s
- since worked mod 2, all values in $\{0, 1\}$

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 separate parts of proof

3 linear ftns. of \( \tilde{a} \)

deg 2 " " "
deg 3 " " "

we only care about 1 lin ftn of \( \tilde{a} \)
deg 2
deg 3

will use to check that birdie wrote down a proper encoding of \( \tilde{a} \)

def \( A : \mathbb{F}_2^n \rightarrow \mathbb{F}_2 \)

\[ A(x) = \sum a_i x_i = a^T \cdot \tilde{x} \]

def \( B : \mathbb{F}_2^n \rightarrow \mathbb{F}_2 \)

\[ B(y) = \sum a_i a_j y_{i,j} = (a \cdot a)^T \cdot \tilde{y} \]

outer product

if \( z = b \cdot c \)

\[ z_{ij} = b_i \cdot c_j \]

def \( C : \mathbb{F}_2^3 \rightarrow \mathbb{F}_3 \)

\[ C(z) = \sum a_{ij} z_{ijk} = (a \cdot a \cdot a)^T \cdot \tilde{z} \]
Proof contains:

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

Supposed to be $A, B, C$

but $V$ needs to check

What to check?

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

- all are linear fits $\Rightarrow$ sc-$\tilde{A}$ will always answer according to closest Lin fit

- linearity test + self-correct passes it if $\tilde{A}$ close to linear

- correspond to same assignment $\tilde{a}$

- test all self-corrections consistent

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

all $\tilde{c}_n$'s evaluate to 0 on $\tilde{a}$
How to do (N):

1. Test \( \hat{A}, \hat{B}, \hat{C} \) are all \( \frac{1}{q} \) close to linear functns
   - random bits: \( O(n^3) \)
   - queries: \( O(1) \)
   - runtime: \( O(n^3) \)
   - Pass if linear
   - Fail if \( \geq \frac{1}{q} \) far from linear in \( O(1) \) queries

2. From now on use self-correcter
   - per query to self-corr:
     - random bits: \( O(n^3) \)
     - queries: \( O(1) \)
     - runtime: \( O(n^3) \)
   - sc-\( \hat{A} \), sc-\( \hat{B} \), sc-\( \hat{C} \) lin functns
   - can query on all inputs
   - use really small error bound on sc-\( C \)
   - if union bound over all calls to sc-\( \hat{A} \) sc-\( \hat{B} \) + sc-\( \hat{C} \)
   - will never see error
Consistency Test:

are $sc-\tilde{A}$, $sc-\tilde{B}$ and $sc-\tilde{C}$ from same assignment $\alpha$?

Tester:

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

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

$$= \sum_{\alpha} x_{1\alpha} \cdot \sum_{\beta} x_{2\beta}$$

$$= \sum_{\alpha} a_{\bar{x}_{1\bar{x}} \bar{x}} x_{1\alpha} x_{2\beta}$$

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

Assume $\bar{x}$ and $\bar{y}$ correspond to same $\alpha$

random $O(n^3)$

generate $O(1)$

random $O(n^3)$

test that $sc-\tilde{A}(\bar{x}) \cdot sc-\tilde{B}(\bar{y}) = $ $sc-\tilde{C}(\bar{x}_0 \bar{y})$

$$= \sum_{\alpha} a_{\bar{x}_x \bar{x}_{1\alpha}} \cdot \sum_{\beta} a_{\bar{y}_y \bar{y}_{1\beta}}$$

$$= \sum_{\alpha} a_{\bar{x}_x \bar{x}_{1\alpha}} \bar{x}_{1\alpha} \bar{y}_{1\beta}$$

$$= sc-\tilde{C}(\bar{x}_0 \bar{y})$$

Note: not unit dist queries

but $s-c$ helps here
Is it a good test?

given $sc-A$
$sc-B$
$sc-C$

\[ A(x) = a^T x \]
\[ B(y) = b^T y \]
\[ C(z) = c^T z \]

hopefully
\[ b^T = (aoa)^T \]
\[ c^T = (aob)^T \]
\[ = (aoa^T a)^T \]

If $b = aoa$ then test pass via green argument $\checkmark$

else

\[ \begin{cases} 
  \text{if } b \neq aoa \\
  A(x_1) \cdot A(x_2) 
\end{cases} \]

\[ \Rightarrow 
  = B(x_1 \circ x_2) 
\]

by def of $x_1 \circ x_2$
\[
\text{if } a \cdot a \neq b \text{ then } \frac{1}{2} \leq \Pr_{x_2} [(a \cdot a) \cdot x_2 \neq b \cdot x_2] \leq \frac{1}{4}
\]

\[
\frac{1}{a} \cdot \frac{1}{2} \geq \frac{1}{4}
\]

so test fails with prob \(\geq \frac{1}{4}\)

Similar argument for if \(c \neq a \cdot b\) then test fails with const prob

\[
\text{if test passes} \quad \text{all three proofs encode same assignment } a
\]
How do we know that \( a \) is a \( \text{sat} \) assignment?

Satisfiability Test:

pick \( r \in \mathbb{Z}_2^n \)

Compute \( \alpha, \beta, \gamma, \delta \)

\[ \begin{align*}
\alpha & : x_1, x_2, \ldots, x_n \\
\beta & : \bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n \\
\gamma & : y_1, y_2, \ldots, y_n \\
\delta & : \bar{y}_1, \bar{y}_2, \ldots, \bar{y}_n
\end{align*} \]

query proof to get

\[ \begin{align*}
SC-A (\alpha_1, \ldots, \alpha_n) & \quad \text{to get} \quad w_0 \\
SC-B (\beta_1, \ldots, \beta_n) & \quad \text{"} \quad w_1 \\
SC-C (\gamma_1, \ldots, \gamma_n) & \quad \text{"} \quad w_2
\end{align*} \]

Verify

\[ 0 = r^T w_0 + w_1 + w_2 \pmod{2} \]

\[ \uparrow \]

hopefully means \( \sum_{r_i} \hat{c}_i (\alpha) = 0 \)
PCP theorems \implies \text{ hardness of approximation theorems}

\[ \text{def } L \in \text{PCP} \text{, s.t. } \]

\[ \exists \text{ prob poly time } V \text{ tosses } r \text{ coins queries } q \text{ bits} \]

\[ \text{st. } x \in L \implies \exists \top \text{ st. } V \text{ accepts with prob} \]

\[ \forall \top \text{, } V \text{ accepts with prob } \leq s \]

(assume \( V \) is "non-adaptive" picks all queries before looking at answers)

FGLSS graph:

for a given input \( x \)

have a large clique iff \( x \in L \)
Construct $G'$:

node $\cap$ entry in table of values
edge $\cap$ entries that are consistent
$\ell_1, \ldots, \ell_q \cap \ell'_1, \ldots, \ell'_q$
if they look at same locus get same answer

for each random string
$R = \ell$
$\exists$ for all possible query answers
$Q = \ell^f$

$\exists$ to $i$
(look at locs $\ell, \ldots, \ell_q$)
if $j$ sees answers $j = (j_1, j_2)$
then outputs pass
1) no edge bet $M_{ij}$ $M_{i'j'}$ for $j \neq j'$

$\Rightarrow$ any clique in $G$ has $\leq 1$ node per row

2) if 2 or more devices v's per row $\Rightarrow$ have complete bipartite graph between their nodes

3) clique corresponds to partial proofs

$\text{size clique} \geq \# \text{ random choices for verifier to accept}$

$\geq (\text{prob of accept}) \cdot 2^c$

if $3\text{SAT} \in \text{PCP}(1/3)$

then $\phi \in 3\text{SAT} \Rightarrow$ pick cols consistent with $11$ (good proof) cause every row to pass
if \( \Phi \notin S \setminus \pi \) then

\[
\forall \varepsilon > 0 \quad \exists i < |\pi| \quad \exists r \in R
\]

else, \( \exists \text{ a good consistent with } \gamma \) is

Consistent with \( \gamma \) and \( \gamma \) is

size

else

\( \Phi \) is consistent with \( \gamma \)

Consistent with \( \gamma \)

\( \forall \varepsilon > 0 \quad \exists i < |\pi| \quad \exists r \in R \)