Lecture 13

PCP Continued

May 13, 2004
Lecturer: Paul Beame
Notes: Tian Sang

Last time we began the proof of the theorem that $\mathsf{PCP}(\mathsf{poly}, \mathsf{poly}) = \mathsf{NEXP}$.

We showed that IMPLICIT-3SAT is NEXP-complete where IMPLICIT-3SAT takes as input a Boolean formula B defined on $m' + 3n + 3$ variables where $m' = m + g$ and B is in the language if and only if there is an oracle truth assignment $A : \{0,1\}^n \rightarrow \{0,1\}$ such that

$$\exists A \forall w, z, v_1, v_2, v_3 \; B(w, z, v_1, v_2, v_3, A(v_1), A(v_2), A(v_3)).$$

(We interpreted w as an m-bit long clause index, z as gate variables, each v_i as an n-bit long variable index in the clause indexed by w, and $A(v_i)$ is the truth assignment to variable x_{v_i}.)

Given A, the verification that A satisfies B is in coNP^A. This could be viewed as an oracle special case of the $\mathsf{IP} = \mathsf{PSPACE}$ protocol. An alternative to this is a sum-check protocol of Lund, Fortnow, Karloff, and Nisan protocol for $\#\mathsf{P}$ (given in Sipser's text and discussed in the next lecture) which instead verifies the value of

$$\sum_{w,z,v_1,v_2,v_3} B(w, z, v_1, v_2, v_3, A(v_1), A(v_2), A(v_3)).$$

In either case we need to be able to convert B to a low degree polynomial in w, z, v_1, v_2, v_3 over $\mathbb{F}$ and evaluate that polynomial on random variables in $\mathbb{F}$ instead of $\{0,1\}$. To do this we needed to be able to evaluate A on such assignments so we use the fact shown last time that there is a (unique) multilinear extension of an assignment $A : \{0,1\}^n \rightarrow \mathbb{F}$. Let $\overline{A} : \mathbb{F}^n \rightarrow \mathbb{F}$ be this multilinear extension.

Thus the proof table gives values for $\overline{A}$, as well as a full tree of all possible executions of the prover's strategy for a $\mathsf{coNP}^{\overline{A}}$ ($\mathsf{IP} = \mathsf{PSPACE}$ style) proof that A is a satisfying assignment to $\forall w, z, v_1, v_2, v_3 \; B(w, z, v_1, v_2, v_3, A(v_1), A(v_2), A(v_3))$. Such a tree is given in Figure 13.1.

This would be OK if the proof table correctly gives an $\overline{A}$ that really is multilinear. However, since the verifier only will examine a polynomial number of places out of the exponentially many in the proof the verifier can never be sure that the table truly is multilinear. The verifier will only be able to verify that the table is close (in Hamming distance) to a multilinear function.

A useful property of the $\mathsf{IP} = \mathsf{PSPACE}$ protocol (and the $\#\mathsf{P}$ protocol) is that the final polynomial is evaluated only once on random inputs chosen by the verifier. Thus on any run, the verifier will only examine $\overline{A}$ in 3 places, randomly chosen by the verifier.

If $\overline{A}$ instead is merely close to a multilinear $\widehat{A} : \mathbb{F}^n \rightarrow \mathbb{F}$, such that say $dist(\overline{A}, \widehat{A}) \leq \delta$, then with probability $\geq 1 - 3\delta$ the verifier will only see values on which $\overline{A}$ and $\widehat{A}$ agree and thus the additional error contributed to the acceptance probability of the protocol by the difference between $\overline{A}$ and $\widehat{A}$ is at most 3δ.

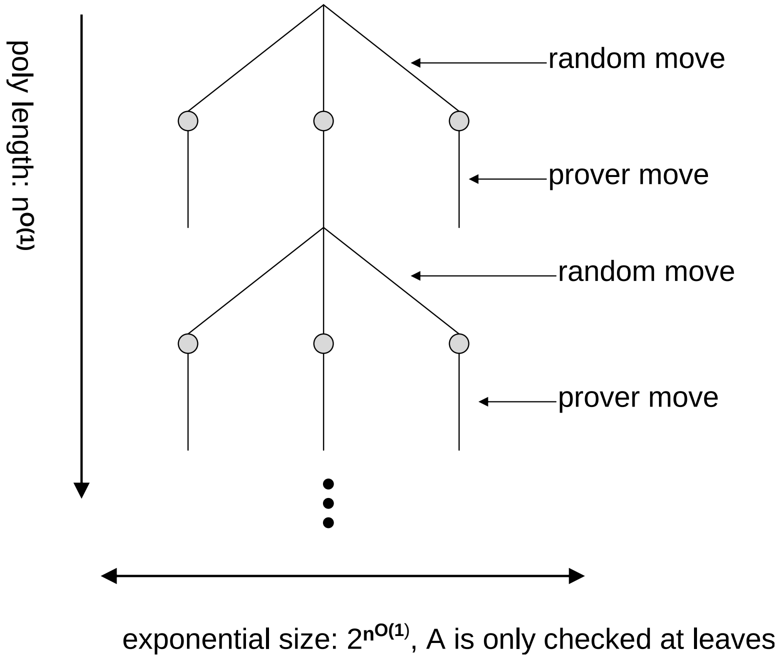

Figure 13.1: Proof Tree Corresponding to an Interactive Protocol

So we need to check that $dist(\overline{A}, \widehat{A}) \leq \delta$ for some multilinear function $\widehat{A}$. This is done using a multilinearity test that is a special case of a test that checks whether or not a polynomial has degree at most k in each variable.

13.1 The Max-degree-k test

Definition 13.1. $f : I^n \to \mathbb{F}$ is called *max-degree-k* iff it can be extended to $f : \mathbb{F}^n \to \mathbb{F}$, that has degree at most k in each variable.

Definition 13.2. For $u_1, \ldots, u_n \in I^n$ an *i-line* is a set $\{(u_1, \ldots, u_{i-1}, z, u_{i+1}, \ldots, u_n) \mid z \in I)\}$.

The following test is a generalization to a max-degree-k test of a multilinearity test due to Feige, Goldwasser, Lovasz, Safra, and Szegedy that improved on the original multilinearity test used by Babai, Fortnow, and Lund. The following analysis is from Friedl, Hatsagi, and Shen.

Aligned Line Test: Choose $k+1$ distinct elements $a_1, \ldots, a_{k+1} \in I$
Repeat t times:

(1) Choose $i \in_R \{1, \ldots, n\}$

(2) Choose a random point $(u_1, \ldots, u_n) \in_R |I|^n$

(3) Check that on the i-line through $(u_1, \ldots, u_n)$ at $u_i, a_1, \ldots, a_{k+1}$, f looks like a degree $\leq$ k polynomial, if not, reject. That is, check that f on the $k+2$ points $(u_1, \ldots, u_n), (u_1, \ldots, u_{i-1}, a_1, u_{i+1}, \ldots, u_n), \ldots, (u_1, \ldots, u_{i-1}, a_{k+1}, u_{i+1}, \ldots, u_n)$ fits a degree k polynomial evaluated at $u_i, a_1, \ldots, a_{k+1}$.

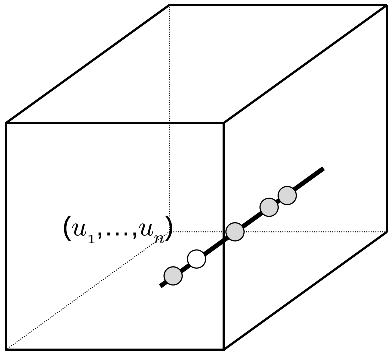

Figure 13.2: Aligned Line Test

In step (3) it would be more natural simply to check f on $k+2$ random points on a randomly chosen i-line but this version is both easier to analyze and uses less randomness.

Definition 13.3. Let $P(n,k) = \{f : I^n \to \mathbb{F}$ that are max-degree-k $\}$ and $P_i(n,k) = \{f : I^n \to \mathbb{F}$ that has degree $\leq k$ in $x_i\}$. Observe that $P(n,k) = \bigcap_{i=1}^{k} P_i(n,k)$.

Definition 13.4. Define $d(f, P(n,k)) = Pr_{x \in_R I^n}[f(x) \neq g(x)]$, and $d(f,S) = \min_{g \in S} d(f,g)$.

Lemma 13.1. *In a single round of the aligned line test,* $\Pr[$*test rejects* $\mid i$ *is chosen*$] \geq d(f, P_i(n,k))$.

Proof. For each choice of $u_1, \ldots, u_{i-1}, u_{i+1}, \ldots, u_n$, there is a unique degree k polynomial $h_{u_1,\ldots,u_{i-1},u_{i+1},\ldots,u_n}(z)$ that equals $f(u_1,\ldots,u_{i-1},z,u_{i+1},\ldots,u_n)$ for $z = a_1,\ldots,a_{k+1}$. The probability that the test does not reject in step (3) is the probability that for $u_i \in_R I$, $f(u_1,\ldots,u_{i-1},u_i,u_{i+1},\ldots,u_n) = h_{u_1,\ldots,u_{i-1},u_{i+1},\ldots,u_n}(u_i)$. Combining all these $h_{u_1,\ldots,u_{i-1},u_{i+1},\ldots,u_n}$ functions for different values of $u_1,\ldots,u_{i-1},u_{i+1},\ldots,u_n$ this yields a function $h \in P_i(n,k)$ and the probability that the test does not reject is the probability that on a random $\overrightarrow{u} = (u_1,\ldots,u_n) \in I^n$, $f(\overrightarrow{u}) = h(\overrightarrow{u})$ and this is precisely $1 - d(f,h) \leq 1 - d(f, P_i(n,k))$. □

The correctness of the test is based on the lemma above and the following analysis.

Lemma 13.2. *For any function $f : I^n \to \mathbb{F}$ and any k,*

$$d(f, P(n,k)) \leq 6(k+1) \sum_{i=1}^{n} d(f, P_i(n,k)) + 2nk/\sqrt{|I|}.$$

We first see how this implies that the aligned line test successfully detects functions that are far from max-degree k.

Corollary 13.3. *If $d(f, P(n,k)) \geq \delta \geq 4nk/\sqrt{|I|}$ then running the aligned test for $t = \Theta(nk/\delta)$ rounds will ensure that the test rejects with arbitrarily high constant probability. The total number of queries is $\Theta(nk^2/\delta)$ and the total number of random bits is $\Theta(\frac{n^2k}{\delta}\log|I|)$.*

Proof. By Lemma 13.1,

$$\begin{aligned}
\Pr[\text{test rejects in a round}] &\geq \frac{1}{n}\sum_{i=1}^{n}\frac{d(f,P_i(n,k))}{n}\\
&\geq \frac{d(f,P(n,k)) - \frac{2nk}{\sqrt{|I|}}}{6(k+1)n}\\
&\geq \frac{\delta - \frac{2nk}{\sqrt{|I|}}}{6(k+1)n}\\
&\geq \frac{\delta}{12(k+1)n},
\end{aligned}$$

by the assumption on δ. Thus the expected number of rounds before the test rejects is at most $12(k+1)n/\delta$. Repeating this $t = \Omega(kn/\delta)$ rounds yields arbitrarily high constant probability of detection. The aligned line test makes $k+1$ queries per round and uses $n \log_2 |I|$ random bits per round. □

Proof of Theorem 12.4. This Corollary is enough to complete the proof that $\mathsf{PCP}(\mathsf{poly}, \mathsf{poly}) = \mathsf{NEXP}$. Using $\delta = 1/10$, say, and using a field $\mathbb{F}$ with $I \subseteq \mathbb{F}$ and $|I| \geq 160n^2$ (since $k = 1$ in the application), running the aligned line test for $O(n)$ rounds (which yields $O(n)$ queries and $O(n^2 \log n)$ random bits is sufficient to ensure that with probability $\geq 9/10$, $\overline{A}$ is within Hamming distance δ of some multilinear $\widehat{A}$. Using the proof table for a $\mathsf{coNP} \subseteq \mathsf{IP}$ (oracle) protocol with error at most $1/10$ to verify that $\overline{A}$ satisfies B, yields total failure probability at most $1/10 + 1/10 + 3\delta = 1/2$. □

We now sketch some of the ideas involved in the proof of Lemma 13.2. The main idea behind all the low degree tests we will use is the following generalization to multivariate polynomials of the fact that a low degree polynomial only has a small number of roots.

Lemma 13.4 (Schwartz, Zippel). *If $p \in \mathbb{F}[x_1, \ldots, x_n]$ is a polynomial of total degree $\leq d$, and $p \neq 0$, then for $a_1, \ldots, a_n \in_R I \subseteq \mathbb{F}$, $\Pr[p(a_1, \ldots, a_n) = 0] \leq \frac{d}{|I|}$*

Proof. By induction on n. The base case $n = 1$ follows because the polynomial p has $\leq d$ roots.

For the induction step, write

$$p(x_1, \ldots, x_n) = \sum_{i=0}^{m} p_i(x_1, \ldots, x_{n-1})x_n^i.$$

Since the total degree of p is at most d, the total degree of p_i is at most $d - i$.

$$\begin{aligned}
\Pr[p(a_1, \ldots, a_n) = 0] &\leq \Pr[p_m(a_1, \ldots, a_{n-1}) = 0] + \Pr[p(a_1, \ldots, a_n) = 0] \mid p_m(a_1, \ldots, a_{n-1}) \neq 0]\\
&\leq \frac{d-m}{|I|} + \frac{m}{|I|} = \frac{d}{|I|}
\end{aligned}$$

where the bound for the first term follows from the inductive hypothesis applied to p_m and the bound for the second term follows from the application of the base case to the polynomial $q(x_n) = \sum_{i=0}^{m} p_i(a_1, \ldots, a_{n-1})x_n^i$. □

Corollary 13.5. *If $|I| > 3nk$ and $d(f, P(n,k)) \leq 1/3$ then there is a unique g such that $d(f,g) = d(f, P(n,k))$*

Proof. Let $g \in P(n,k)$ be a polynomial witnessing the fact that $d(f, P(n,k)) < 1/3$ so $d(f,g) < 1/3$. Suppose $h \in P(n,k)$ and $h \neq g$. The total degree of each of g and h is at most nk. Applying the Schwartz-Zippel Lemma to $g - h$ we see that $d(g,h) \geq 1 - \frac{nk}{|I|} > \frac{2}{3}$. Since by the triangle inequality $d(g,h) \leq d(f,g) + d(f,h) \leq 1/3 + d(f,h)$, we obtain that $d(f,h) > d(g,h) - 1/3 > 2/3 - 1/3 = 1/3$ implying that g is unique. □

Let $f^i \in P_i(n,k)$ be such that $d(f, f^i) = d(f, P_i(n,k))$. Observe that by definition, $d(f^i, g) = \frac{1}{|I|} \sum_{c \in I} d(f^i|_{x_i=c}, g|_{x_i=c})$ and that if $g \in P(n,k)$ then $g|_{x_i=c} \in P(n-1,k)$ for any c. In particular this means that $d(f^i, g) \geq \frac{1}{|I|} \sum_{c \in I} d(f^i|_{x_i=c}, P(n-1,k))$.

The following is an immediate consequence of Corollary 13.5.

Lemma 13.6. *If $|I| > 3nk$ and $f^i \in P_i(n,k)$ and $g \in P(n,k)$ agree on at least 2/3 of all i-lines then $d(f^i, g) = d(f, P(n,k))$ and for all $c \in I$, $d(f^i|_{x_i=c}, g|_{x_i=c}) = d(f^i|_{x_i=c}, P(n-1,k))$ and thus*

$$d(f^i, g) = \frac{1}{|I|} \sum_{c \in I} d(f^i|_{x_i=c}, P(n-1,k)).$$

For any $k+1$ hyperplanes $x_i = c_1, \ldots, x_i = c_{k+1}$, the fraction of all i-lines on which f^i and g disagree is at most $\sum_{j=1}^{k+1} d(f^i_{x_i=c_j}, g)$ since f and g are in complete agreement on any i-line on which f^i and g agree on all of these hyperplanes. (This follows because f^i and g are both degree k polynomials along any i-line; see Figure 13.3.) For the same reason, on any i-line on which f^i and g disagree, they agree on at most k points.

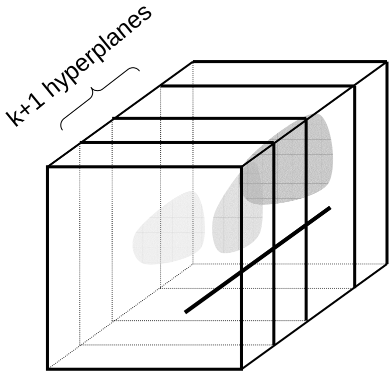

Figure 13.3: Behavior on i-lines related to $k+1$-hyperplanes

Moreover, even if there are $k+1$ hyperplanes $x_i = c_1, \ldots, x_i = c_{k+1}$ on which $\sum_{j=1}^{k+1} d(f^i_{x_i=c_j}, P(n-1,k)) \leq 1/3$ then letting $g_1, \ldots, g_{k+1} \in P(n-1,k)$ be the witnessing polynomials close to $f^i_{x_i=c_j}$ then we can interpolate the g_j to a single polynomial $g' \in P(n,k)$ for which f^i and g' agree on a fraction $1 - \alpha \geq 2/3$ of all i-lines, $d(f^i, g') = d(f^i, P(n,k)) \leq \alpha$, and $d(f^i_{x_i=c}, P(n-1,k)) = d(f^i_{x_i=c}, g'_{x_i=c})$ for all $c \in I$. Again, on the α fraction of i-lines on which f^i and g' do not completely agree, there are at most k points of agreement, so in total there are at most an $\alpha k/|I| \leq k/(3|I|)$ fraction of all points in I^n that could contribute to places where $d(f^i_{x_i=c}, P(n-1,k)) < \alpha$. By an averaging argument we obtain the following lemma.

Lemma 13.7. *For $|I| > 3nk$ for any $\mu > 0$, either*
(a) $\Pr_{c \in_R I}[d(f^i|_{x_i=c}, P(n-1,k)) \leq 1/(3k+3)] \leq \frac{k+1}{|I|}$ *or*
(b) $\Pr_{c \in_R I}[d(f^i|_{x_i=c}, P(n-1,k)) \leq d(f^i, P(n,k)) - \mu] \leq \frac{k}{3\mu|I|}$.

The final argument follows, for suitable μ, using induction by iteratively choosing $c_i \in_R I$ and setting each $x_i = c_i$ for $i = 1, \ldots, n$. We begin by observing that $d(f, P(n,k)) \leq d(f, f^1) + d(f^1, P(n,k))$ and expanding $d(f^1, P(n,k))$. Intuitively, at each step, either (a) holds and except for a $\frac{k+1}{|I|}$ fraction of choices so far, the average distance between f^i and a max-degree k polynomial on the remaining choices is at least $1/(3k+3) \geq d(f, P(n,k))/(3k+3)$, or (b) holds and except for a $\frac{k}{3\mu|I|}$ fraction of choices so far, the distance between f^i and a max-degree k polynomial is well represented by its error on the remaining choices (except for an additive error of at most μ). Furthermore, the distance at each choice c_i is at most the sum of the error based on setting one more variable in f^i, $d(f^i, f^{i+1})$, and the distance between f^{i+1} and a max-degree k polynomial in the remaining variables. Since after all values are set, the function defined on a single input can be exactly represented by a max-degree k polynomial if (b) always holds then

$$d(f^1, P(n,k)) \leq n\mu + \sum_{i=1}^{n-1} d(f^i, f^{i+1}) + n\frac{k}{3\mu|I|} \leq n\mu + \sum_{i=1}^{n-1}(d(f, f^i) + d(f, f^{i+1})) + n\frac{k}{3\mu|I|}$$

and thus

$$\begin{aligned} d(f, P(n,k)) &\leq 2\sum_{i=1}^{n} d(f, f^i) + n\mu + \frac{nk}{3\mu|I|} \\ &= 2\sum_{i=1}^{n} d(f, P_i(n,k)) + n\mu + \frac{nk}{3\mu|I|}. \end{aligned}$$

However, if (a) holds at some stage, consider the last stage j in which (a) holds. Except for a $(k+1)/|I|$ fraction of inputs, $d(f, P(n,k))/(3k+3)$ is a lower bound on the distance between f^j and a max-degree k polynomial on the remaining inputs, and thus

$$\begin{aligned} d(f, P(n,k))/(3k+3) &\leq (n-j)\mu + \sum_{i=j}^{n-1}(d(f, f^i) + d(f, f^{i+1})) + (n-j)\frac{k}{3\mu|I|} + \frac{k+1}{|I|} \\ &= 2\sum_{i=j}^{n} d(f, P_i(n,k)) + (n-j)\mu + \frac{(n-j)k}{3\mu|I|} + \frac{k+1}{|I|}. \end{aligned}$$

Strictly speaking, the stages at which (a) holds depend on the choices of the c_i so this argument is not fully rigorous as stated; however, it can be made rigorous by maintaining the sequence of choices made explicitly and summing appropriately. The bounds in Lemma 13.2 follow.