12/1/2003

Notes for Lecture 26

We continue towards a proof of the PCP theorem. Previously, we constructed a proof system of $m' + 2 = 3m + 5$ polynomials which was complete, and sound given the promise that the polynomials were of low degree. We need to add two refinements. First, we would like to reduce the query complexity from $O(\log^2 n)$ to constant. Second, we would like to remove the promise that the polynomials be of low degree. In this lecture, we address this second requirement.

26.1 Reduction to a single polynomial

As a trivial simplification, notice that the polynomials $p_0 : \mathbf{F}^m \to \mathbf{F}$ and $p_j : \mathbf{F}^{m'} \to \mathbf{F}$, $j = 1, \ldots, m' + 1$, can be stored in just a single polynomial $p : \mathbf{F}^{m'+1} \to \mathbf{F}$, defined by $p(j, x) = p_j(x)$. Since the original polynomials were all of degree at most hm', p has degree at most $d = hm' + m' + 1$.

26.2 Point-line low-degree test

Recall the point-line low-degree test: Given oracles $f : \mathbf{F}^m \to \mathbf{F}$, $L : (F^m)^2 \to \mathbf{F}^{d+1}$, we pick $x, y \in_R \mathbf{F}^m$ and $t \in_R \mathbf{F}$, and accept iff $L(x,y)(t) = f(x + ty)$ (where $L(x, y)$ is regarded as a degree-d polynomial).

Theorem 1 *For $\delta < \frac{1}{4}$, $|\mathbf{F}| = \Omega(d)$, if test accepts w/ prob. $\geq 1 - \frac{\delta}{2}$, then f is δ-close to a polynomial of degree $\leq d$.*

To use this, we fix $\delta = \frac{1}{100}$, say, and run the test say 200 times to verify that with good probability the test accepts with probability at least $1 - \frac{\delta}{2}$.

Thus we can assume we are promised that the function f passed to the verifier is δ-close to a polynomial p of degree $\leq d$. However, the verifier requires the stronger promise, that f itself have low degree. We will show now how to use f to reconstruct p on the verifier's query points.

26.3 Polynomial reconstruction

We give a method which takes the verifier's q query points to f and using an extra oracle, either fails (if it detects an inconsistency) or with good probability returns p correctly evaluated on *all* q queries. (A simple special case is $q = 1$, but in our application actually $q = O(\log^2 n / \log\log n)$.)

For $q = 1$ we can use a line oracle similarly to the point-line low-degree test. For higher q, we might try using an oracle for polynomials on q-dimensional subspaces. However, a degree-d polynomial on a q-dimension subspace is specified by $(q^{d+1} - 1)/(q - 1)$ coefficients – too many for us to query. Instead of generalizing to q-dimensional subspaces, we'll generalize to degree q *curves*.

Definition 1 *A degree q curve in $\mathbf{F}^m$ is $c : \mathbf{F} \to \mathbf{F}^m$*

$$c(t) = (c_1(t), \ldots, c_m(t)) \ , \tag{1}$$

each component c_i being a degree q polynomial.

Remark 1

- *For $|\mathbf{F}| > q+1$, given $t_1, \ldots, t_{q+1} \in \mathbf{F}$ distinct and $x^1, \ldots, x^{q+1} \in \mathbf{F}^m$ distinct, there is a degree q curve c with $c(t_i) = x^i$, $i = 1, \ldots, q+1$. Indeed, we can interpolate to determine each coordinate of c.*
- *Given $t_1, \ldots, t_{q+1}$ and $x^1, \ldots, x^q$, pick x^{q+1} at random, and interpolate c. For $t \notin \{t_1, \ldots, t_q\}$, $c(t)$ is random.*
- *For $p : \mathbf{F}^m \to \mathbf{F}$ a degree d polynomial, $c : \mathbf{F} \to \mathbf{F}^m$ a degree q curve, $p \circ c : \mathbf{F} \to \mathbf{F}$ is a univariate polynomial of degree qd.*

Now add the oracle $C : R \times \mathbf{F}^m \to \mathbf{F}^{qd+1}$, where $R \subset (\mathbf{F}^m)^q$ is the polynomial-size set of possible verifier queries. Here is our reconstruction algorithm:

1. Pick $x^{q+1} \in_R \mathbf{F}^m$ and let c be the degree q curve with $c(t_i) = x^i$.
2. Set $u = C(x^1, \ldots, x^q, x^{q+1})$ a degree qd polynomial.
3. Pick $t \in_R \mathbf{F} \setminus \{t_1, \ldots, t_q\}$, and reject if $u(t) \neq f(c(t))$.
4. Output $u(t_1), \ldots, u(t_q)$.

Theorem 2 *If f is δ-close to a degree d polynomial p, then $\forall\ (x^1, \ldots, x^q) \in R$, the probability of a wrong answer is*

$$\leq 2\sqrt{\delta} + \frac{qd}{|\mathbf{F}| - q} \ . \tag{2}$$

Proof: Let S be the set of points where f disagrees with p, $|S| \leq \delta |\mathbf{F}|^m$. The expected number of random points of c in S is $\leq \delta(|F| - q)$. By Markov's inequality, with probability at most $\sqrt{\delta}$ are there more than $\sqrt{\delta}(|\mathbf{F}| - q)$ points of c in S.

So assume $\leq \sqrt{\delta}$ fraction of the random points of c are in S. Assume the algorithm returns an incorrect answer, so in particular $u \neq p \circ c$. Then the probability the check $u(t) \stackrel{?}{=} f(c(t))$ succeeds is

$$\leq \sqrt{\delta} + \frac{qd}{|\mathbf{F}| - q} \ . \tag{3}$$

Indeed, with probability at most $\sqrt{\delta}$ is t chosen so $c(t) \in S$. Otherwise, $f(c(t)) = p(c(t))$ is the same as $u(t)$ with probability at most $\frac{qd}{|\mathbf{F}| - q}$, since u and $p \circ c$ are distinct degree qd polynomials. $\square$