PCP Lecture 9

Note Title 4/16/2008

Theorem: $NP \subseteq PCPP[poly, 1]$

Recall that in the previous lecture we defined:

Def: The quadratic functions encoding $Q: \{0,1\}^{n-1} \to \{0,1\}^{2^{n^2}}$
maps $(a_2 \dots a_n) \in \{0,1\}^{n-1}$ into $H(a' \otimes a')$
where $a' \otimes a' \in \{0,1\}^{n^2}$ is defined by $(a' \otimes a')_{ij} = a'_i \cdot a'_j$
and $a' \in \{0,1\}^n$ is the vector $(1, a_2 \dots a_n)$.

* This encoding is self correctable (2 queries)
* It is locally testable (exercise). Namely:

There is a randomized procedure that, given oracle access to w, makes $O(1)$ queries and ① if $w \in QF$ accepts w. prob 1
② if $\delta = dist(w, QF) > 0$ rejects w. prob $\geq c \cdot \delta$ (for an abs. const. c)

Proof of Thm:

Step 1: Recall that zero testing is NP-complete.
Moreover,

Lemma 1: Let φ be a 3-SAT formula over vars $X = \{x_1, \dots, x_n\}$. There are m degree-2 polynomials $P_1 \dots P_m$ over variables $X \cup Y$, $(n' = |Y| = n^{O(1)})$, such that
(i) If $(a_1 \dots a_n) \in \{0,1\}^n$ satisfies φ then $\exists (b_1, \dots b_{n'}) \in \{0,1\}^{n'}$ such that $P_i(\bar{a}\bar{b}) = 0$ for all i.
(ii) If φ is unsatisfiable then for any a there is no b s.t. $P_i(\bar{a}\bar{b}) = 0$ for all i.

Proof: convert each $x_i \vee x_j \vee \bar{x}_k$ to $(1-x_i)(1-x_j)x_k$ etc.
whenever a poly contains $x_i x_j$ replace by
a new var y_{ij} and add a poly $x_i x_j - y_{ij}$.
.... (i) + (ii) are immediate.

$\square$

Step 2:

Note that if $a_1 \ldots a_{n+n'}$ does not zero at least one of $P_1 \ldots P_m$ then for a random $\bar{r} = (r_1 \ldots r_m) \in \{0,1\}^m$

$$\Pr_r \left[Q_r(\bar{a}) \triangleq \sum r_i P_i(\bar{a}) = 0 \right] \leq \frac{1}{2}.$$

Step 3:

We describe a verifier of proximity for 3SAT.
Ver has explicit input φ and implicit input $a_1 \ldots a_n$.
Ver expects as proof, the quad. function encoding of $(\bar{a}, \bar{b})$
where b is the assignment to aux. vars that together zero all $\{P_i\}$.
resulting from the reduction in the Lemma.

1) Test that Π is a legal QF encoding, if not reject.

2) Compute $P_1 \ldots P_m$. Choose $r \in_R \{0,1\}^m$ and compute $Q_r = \sum r_i P_i$. (i.e. compute the coefs of the quad. func.)
Use self-correction to verify that $Q_r(\bar{a}, \bar{b}) = 0$

3) Consistency: Select $i \in_R [n]$. test that a_i equals the appropriate bit encoded by Π, using self correction.
i.e. test that $S^{\Pi}(e_i) = a_i$

Proof of completeness: immed.

Proof of Soundness: If $\bar{a}$ is δ-far from satisfying φ, we prove that Ver rejects w. prob $\geq \Omega(\delta)$.

- if Π is δ-far from a legal QF, step 1 rejects w.prob $\Omega(\delta)$. otherwise, Π is δ-close to $QF(a_1 b_1)$.

- if $a_1 b_1$ are such that $\exists i: P_i(a_1 b_1) \neq 0$, then w. prob $\geq \frac{1}{2}$ $Q_r(a_1 b_1) \neq 0$, and with $\Omega(1)$ probability, step 2 rejects. (depending on the self-correction). Otherwise,

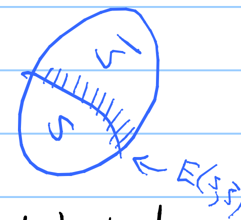
- $a_1 b_1$ is s.t. $\forall i \; P_i(a_1 b_1) = 0$, so $dist(a, a_1) \geq \delta$. (By Lemma 1) So w. prob $\Omega(\delta)$ step 3 rejects.

params: randomness – poly.
queries – $O(1)$
completeness – 1
soundness – constant.

Expander Graphs

Expanders are graphs without "bottlenecks".

Given $G=(V,E)$, and $S \subseteq V$, we consider



$$E(S,\bar{S}) = \{\{u,v\}\in E \mid u \in S \ \ v \notin S\}, \text{ relative to } |S|, |\bar{S}|.$$

Define the edge expansion of G as

$$\Phi(G) = \min_{S,\ |S| \le \frac{|V|}{2}} \frac{|E(S,\bar{S})|}{|S|} \quad \leftarrow \text{size of smallest bottleneck.}$$

A clique is a graph without small bottlenecks, but that's easy since it is dense.
A path has a tiny bottleneck.
A sparse graph without (small) bottlenecks is called an "expander".

Definition: A family $\{G_n\}_{n=n_0}^{\infty}$ of graphs over n vertices is ε-expanding if $\Phi(G_i) \ge \varepsilon$ for all $n \ge n_0$

Thm: There exist $d \in \mathbb{N}$ and $\varepsilon > 0$ and a family $\{G_n\}$ of d-regular graphs that is ε-expanding.
(in fact, a random d-regular graph is quite expanding)

The 2nd Largest Eigenvalue

An alternate way to measure expansion, is via the eigenvalue gap.

Let $G=(V,E)$ be d-regular on n vertices.

Let $A=(a_{ij})$ be its adjacency matrix.

Then A is symmetric and has n real eigenvalues

$$d=\lambda_0 \geq \lambda_1 \geq \dots \geq \lambda_{n-1} \geq -d$$

with eigenvectors $v_0 \dots v_{n-1}$. $(v_0 = \vec{1})$.

Def: A d-regular graph on n vertices is an (n,d,λ)-expander if $\lambda = \max(|\lambda_1|, |\lambda_n|) < d$.

$d-\lambda$ is called the spectral gap of G.

Thm: $$\frac{\phi^2(G)}{2d} \leq d-\lambda \leq 2\phi(G)$$

we will only prove (which is easier) since it is suff. for our application.

Thm: There are explicit $d \geq 3$ and $\lambda < d$ and an explicit (infinite) family of (n,d,λ)-expanders.

Lubotsky - Philips - Sarnak (LPS) construction:

$V = \mathbb{Z}_p \cup \{\infty\}$ x connected with $x+1$, $x-1$, x^{-1}.

Also — A random d-regular graph, has $\lambda \lessapprox 2\sqrt{d-1}$ whp.

The Rayleigh Quotient:

Claim: For a real symmetric matrix A, the following is true

$$\lambda_i = \max_{\substack{x : \|x\|=1 \\ x \perp v_0 \dots v_{i-1}}} \langle x, Ax \rangle$$

where $v_0 \dots v_{n-1}$ are orthonormal eigenvectors of eigenvalues $\lambda_0 \dots \lambda_{n-1}$.

Proof: Clearly taking $x = v_i / \|v_i\|$ we obtain $\leq$.
Write $x = \sum_{i \leq j} \alpha_j v_j$. Then, assuming $\|x\|=1$,

$$x^T A x = \sum \alpha_j v_j^T A (\sum \alpha_j v_j)$$
$$= \sum \alpha_j^2 \lambda_j \leq \lambda_i \quad (\text{since } \sum \alpha_i^2 = \|x\|^2 = 1) \quad \blacksquare$$

Similarly, $\lambda = \max_{x : x \perp \vec{1}} \frac{|x^T A x|}{\|x\|^2}$

Lemma: Let $G_i = ([n], E_i)$ be an (n, d_i, λ_i)-expander for $i = 1, 2$.
Define $H = ([n], E_1 + E_2)$ by adding the adj matrices
$A_H = A_1 + A_2$. (so H is possibly a multigraph).
Then H is an $(n, d_1 + d_2, \lambda_1 + \lambda_2)$ expander.

Proof: Choose x s.t. $\lambda = x^T A_H x$ (and $\|x\|^2 = 1$, and $x \perp 1$)

then clearly $x^T A_H x = x^T(A_1+A_2)x = x^TA_1x + x^TA_2x$

$$\leq \lambda(A_1) + \lambda(A_2) \quad \square$$

This makes sense: if we "union" the edges of two graphs and one of them had no small bottlenecks then the result must also have this property.

We now prove the connection between "bottlenecks" and λ:

Lemma: Let G be an (n,d,λ)-expander, then $d-\lambda \leq 2\phi(G)$.

Proof: suppose $S \subseteq V$ is s.t. $|S| \leq |V|/2$ and $\phi(G) = E(S,\bar{S})/|S|$.

The idea is to consider the following vector x

$$x_v = \begin{cases} |S| & v \notin S \\ -|\bar{S}| & v \in S \end{cases}$$

(for simplicity think that $|S| = \frac{n}{2}$ and then x is a normalized ± 1 vector def. by S)

* $x \perp \vec{1}$: easy to see that $\sum_v x_v = |S| \cdot (-|\bar{S}|) + |\bar{S}| \cdot |S| = 0$.
* $\|x\|^2 = |\bar{S}| \cdot |S|^2 + |S| \cdot |\bar{S}|^2 = n|S||\bar{S}|$.
* $x^TAx = \langle x, Ax \rangle = \sum_v x_v \sum_u A_{vu} x_u = 2\sum_{(u,v)\in E} x_u x_v$.

$$= \underbrace{2E(S,\bar{S}) \cdot (-|\bar{S}||S|)}_{S \leftrightarrow \bar{S}} + \underbrace{(d|S| - E(S,\bar{S}))}_{S \leftrightarrow S}|\bar{S}|^2 + \underbrace{(d|\bar{S}| - E(S,\bar{S}))}_{\bar{S} \leftrightarrow \bar{S}}|S|^2$$

$$= d|S||\bar{S}|n - n^2 E(S,\bar{S}).$$

Since $x^TAx \leq \lambda\|x\|^2$ we get

$$d|S||\bar{S}|n - n^2E(S,\bar{S}) \leq \lambda n|S||\bar{S}|$$

$$d - \lambda \leq n\frac{E(S,\bar{S})}{|S||\bar{S}|} \leq 2 \cdot \phi(G). \quad \square$$

Expander Mixing Lemma

We said that expanders have no "bottlenecks". and that random graphs are expanders. It turns out that expanders "emulate" random graphs in certain precise ways.

For an expander $G=(V,E)$, and for any subsets $S,T \subseteq V$ the # of edges between S and T is "close" to what it would be in a random graph of the same density:

Lemma (EML): Let G be an (n,d,λ)-expander, and let $S,T \subseteq V(G)$.

$$\left| E(S,T) - \underbrace{d|S||T|\tfrac{1}{n}} \right| \leq \lambda\sqrt{|S|\cdot|T|}$$

in a random G: $d|S|$ edges from S, each hits T w. prob $|T|/n$.

Random Walks on Expanders

Any graph naturally gives rise to a Markov chain $X_0, X_1, \dots, X_t, \dots$ (where X_i is a random variable that takes values in V, and

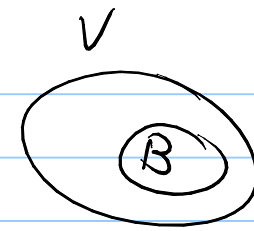
$$\text{Prob}(X_i \mid X_{i-1} \dots X_0) = \text{Prob}(X_i \mid X_{i-1}) \quad)$$

where X_i is the vertex reached in the i-th step of a random walk

Expanders are graphs on which this chain is "rapidly mixing"

i.e.: even if the distribution of X_0 is not very random,

X_t will be "close" to uniformly distributed on V for rather small t's. (i.e the walk quickly "forgets" its starting point)



In particular, consider the following scenario. Let $B \subseteq V$ be a set of density $\frac{|B|}{|V|} = \beta$.

Choosing t vertices independently at random, the probability of them all being in B is β^t.

What if we choose them by taking v_1 at random and then $v_1, v_2, \dots, v_t$ a random walk from v_1?

The probability is $\sim$ similar to β^t, with an error that goes to zero when $\lambda \longrightarrow 0$.

We prove the following edge variant (which will be needed for proving the PCP thm)

Lemma: Let G be an (n,d,λ)-expander, and let $F \subset E$, $\frac{|F|}{|E|} = \rho$. and let $(v_0, v_1, \dots, v_t, v_{t+1})$ be a random walk. (i-th step $= (v_i, v_{i+1})$)
The probability that $(v_t, v_{t+1}) \in F$, conditioned on $(v_0, v_1) \in F$ is at most

$$\underbrace{\frac{|F|}{|E|}}_{\text{the prob if the t-th step were independent of 1st step}} + \underbrace{\left(\frac{\lambda}{d}\right)^{t-1}}_{\text{error term decaying exp. with } t.}$$

Proof: Let A be the normalized adjacency matrix of G.
Suppose $x \in \mathbb{R}^n$ is a probability vector $(x_v \geq 0, \ \sum x_v = 1)$
Let $U_1, \dots, U_t$ be a Markov chain defined by the random walk on G, and assume U_1 is distributed according to x. Then the distribution of U_2 is given by $y = Ax$ and of U_t by $A^{t-1}x$:

$$y_u = \sum_v A_{uv} x_v = \frac{1}{d}\sum_{u \sim v} x_v = \sum_v \text{Prob}[U_1 = v]\cdot \text{Prob}[U_2 = u \,/\, U_1 = v]$$
$$= \text{Prob}[U_2 = u].$$

In the lemma, (v_0, v_1) is a random edge in F, so v_1 is distributed according to $\bar{x}$: $x_v = \frac{k}{2F}$ where k = # of F edges touching v. (check: $\sum x_v = \sum k \cdot \frac{1}{2F} = 1$.)

The probability that the t-th step is in F is $\frac{k}{d}$ where k is the # of F-edges incident on V_t. This is described by $\bar{y}$: $y_w = x_w \cdot \frac{2F}{d}$.

Altogether, we are interested in $\langle A^{t-1}x, y\rangle =$

vector describing prob distribution of V_t

prob of taking last step in F, conditioned on location.

$$\langle A^{t-1}x, y\rangle = \sum_v \underbrace{\text{Prob}[V_t = v]}_{(A^{t-1}x)_v} \cdot \underbrace{\text{Prob}[(V_t, V_{t+1}) \in F \,/\, V_t = v]}_{y_v} = \text{Prob}[t\text{-th step of RW} \in F].$$

Now, lets compute. $\langle A^{t-1}x, y\rangle = \langle A^{t-1}x, x\rangle \cdot \frac{2F}{d}$. $u = \frac{1}{n}\cdot\vec{1}$

We write $x = x^{\parallel} + x^{\perp}$ where $x^{\parallel} = \langle x, u\rangle \cdot u$ and $x^{\perp} = x - x^{\parallel}$.

So (i) $\langle x^{\perp}, \vec{1}\rangle = 0$ and (ii) $(x^{\parallel})_v = \sum x_v \frac{1}{n} = \frac{1}{n}$ so $\|x^{\parallel}\|^2 = \frac{1}{n}$.

$$A^{t-1}x = A^{t-1}x^{\parallel} + A^{t-1}x^{\perp} = x^{\parallel} + A^{t-1}x^{\perp}$$

$$\begin{aligned}
\langle A^{t-1}x, x\rangle &= \langle x^{\parallel}, x\rangle + \langle A^{t-1}x^{\perp}, x^{\parallel} + x^{\perp}\rangle \\
&= \langle x^{\parallel}, x^{\parallel}\rangle + \langle A^{t-1}x^{\perp}, x\rangle \\
&\le \frac{1}{n} + \|A^{t-1}x^{\perp}\| \cdot \|x\| \qquad \text{Cauchy-Schwartz} \\
&\le \frac{1}{n} + \left(\frac{\lambda}{d}\right)^{t-1} \|x^{\perp}\| \cdot \|x\| \\
&\le \frac{1}{n} + \left(\frac{\lambda}{d}\right)^{n-1} \cdot \|x\|^2 \le \frac{1}{n} + \left(\frac{\lambda}{d}\right)^{t-1} \frac{d}{2F}
\end{aligned}$$

$(\|x\|^2 = \sum_v (x_v)^2 \leq \max_v x_v \cdot \sum_v x_v = \max x_v = \frac{d}{2F}.)$

Multiplying by $\frac{2F}{d}$ we get (note that $|E| = \frac{nd}{2}$ so $\frac{2F}{d} \cdot \frac{1}{n} = \frac{|F|}{|E|}$)

$$\text{Prob} = \langle A^{t-1} x, y \rangle = \frac{|F|}{|E|} + \left(\frac{\lambda}{d}\right)^{t-1}$$

$\square$