CSE431: Complexity Theory February 27, 2012

Lecture 15 Proof of PCP Theorem (Part 2)

Lecturer: Anup Rao Scribe:

References

1 Powering

We start with a constraint graph G that is an expander with eigenvalue bound λ, degree d and alphabet size k. Our goal is to generate G' that is at most a constant factor bigger than G, such that G' is satisfiable if and only if G is satisfiable, and if every assignment to G leaves at least γ fraction of constraints unsatisfied, then every assignment to G' leaves at least $\Omega(t\gamma)$ fraction of the constraints unsatisfied. Here we don't care about the dependence on d, k, λ.

2 Construction

2.1 Before Stopping Random Walk

In order to describe the construction, we need to understand a particular kind of random walk on regular graphs. Consider the following process that samples a walk in the graph:

1. Sample a random vertex u_1. Set $i = 1$.
2. With probability $1/t$ terminate and return the walk sampled so far.
3. Sample a random neighbor u_{i+1} of u_i.
4. Set $i = i + 1$ and go to step 2.

2.2 After Stopping Random Walk

1. Sample a random vertex u_1. Set $i = 1$.
2. Sample a random neighbor u_{i+1} of u_i.
3. With probability $1/t$ terminate and return the walk sampled so far.
4. Set $i = i + 1$ and go to step 2.

The result of the sampling process is a sequence $u_1, \ldots, u_r$.

Lemma 1. *In the after stopping random walk,* $\mathbb{E}[r] = t$.

Proof Since after one step, the distribution of the remaining steps is the same, we have

$$\mathbb{E}[r] = 1/t + (1 - 1/t)(1 + \mathbb{E}[r])$$
$$\Rightarrow \mathbb{E}[r] = t.$$

■

Next we describe the constraint graph G':

Vertices The vertex set is the same as in G.

Edges Every edge in G' with end points u, v corresponds to a unique random walk whose length is at most $2t$. The degree of G' is thus $\sum_{i=1}^{2t} d^i$. We shall add multiple edges so that picking a random edge in G' corresponds to taking a random walk in G, conditioned on the event that the walk has length at most $2t$.

Alphabet Let $B(u, t) = \{v | \mathsf{dist}(u, v) \leq B\}$, namely $B(u, t)$ is the ball of radius t around u in G'. We shall interpret an assignment of ℓ' to the vertices of G' as defining a function on each vertex u,

$$\ell'_u : B(u, t) \to [k],$$

that gives an opinion about what the assignment to all the vertices in the ball should be in the original graph G. The alphabet size is $k^{O(d^B)}$, where $B = O(t)$ is a parameter that we shall set later.

Constraints Given an edge $\{u, v\}$ in G', let $u = u_0, \ldots, u_t = v$ be the corresponding path. The constraint will check that for each u_i, $\ell'_u(u_i) = \ell'_v(u_i)$ and that $\ell'_u(u_i), \ell'_u(u_{i+1})$ satisfy the constraint of the edge $\{u_i, u_{i+1}\}$ in G.

3 Analysis

It is clear that if G is satisfiable, then so is G'. We shall prove:

Theorem 2. *Suppose every assignment to the vertices of G must leave $\gamma < 1/t$ fraction of the edge constraints unsatisfied. Then every assignment to the vertices of G' leaves $\Omega(t\gamma)$ fraction of the constraints in G' unsatisfied.*

We shall first outline the proof at a high level, and then fill in the details. Let ℓ' be an arbitrary assignment to G'. Then we shall first construct a related assignment ℓ to G. For now, think of ℓ_v as being set to the "majority" opinion of $\ell'_u(v)$ where $u \in B(v, t)$. We shall define it more concretely later when we see how it should be defined.

Now consider the set of edges F that are violated by the assignment ℓ. We know that $|F|/(nd/2) = \gamma$. We will show that $\Omega(t\gamma)$ fraction of the edges in G' will correspond to paths that use an edge of F.

To this end, we use the concept of *line-graphs*. The line graph of a graph H, denoted $L(H)$ is the graph whose vertex set is the edge set of H, and two vertices are connected in $L(H)$ if and only if the corresponding edges share a vertex in H. We shall prove:

Lemma 3. *If H is a regular graph, then the eigenvalues of $L(H)$ are the same as the eigenvalues of H except that $L(H)$ may have additional eigenvalues which are 0.*

In particular, Lemma 3 implies that $L(G)$ is an expander.

Then we shall prove:

Theorem 4. *Let S be any subset of the vertices in an expander graph such that the density of S is γ. Then the probability that a random path of length t touches S is at least $\Omega(\gamma t)$.*

Let us prove this theorem now.

3.1 Analyzing ASRW

Lemma 5. *Consider the distribution of an ASRW conditioned on making exactly k (u, v) steps, for a fixed edge $\{u, v\}$. Let a be the initial vertex and b be the final vertex. Then b has the same distribution as the final vertex of a BSRW starting at v, and a is like the final vertex of BSRW starting at u, and the initial and final vertices are independent.*

Proof If we were conditioning on $k \geq 1$ then the case of the final vertex would be clear, since we would just fix the path up to the first k visited vertices.

Let Y denote number of u, v steps, then

$$\begin{aligned}\Pr[b = w|Y \geq k] &= \Pr[b = w|Y \geq 1] \\ &= \frac{\Pr[b = w \wedge Y \geq 1]}{\Pr[Y \geq 1]} \\ &= \frac{\Pr[b = w \wedge Y = 1] + \Pr[b = w \wedge Y \geq 2]}{\Pr[Y = 1] + \Pr[Y \geq 2]},\end{aligned}$$

but $\frac{\Pr[b=w\wedge Y\geq 2]}{\Pr[Y\geq 2]} = \Pr[b = w|Y \geq 1]$. Thus $\Pr[b = w|Y = 1] = \Pr[b = 2|Y \geq 1]$. ■

Consider any random path $X_1, \ldots, X_t$ in the graph. Let N denote the number of vertices of this path that lie in S. Then we have:

Lemma 6. $\Pr[N > 0] \geq \mathbb{E}[N]^2 / \mathbb{E}\left[N^2\right]$.

This is an easy application of the second moment method:

Lemma 7. *For any real valued random variables A, B, $|\mathbb{E}[AB]| \leq \sqrt{\mathbb{E}[A^2] \cdot \mathbb{E}[B^2]}$.*

Proof Let $p_{a,b}$ denote $\Pr[A = a, B = b]$. Then

$$\begin{aligned}|\mathbb{E}[AB]| &\leq \sum_{a,b} p_{a,b}|a||b| \\ &= \sum_{a,b} \sqrt{p_{a,b}}|a| \cdot \sqrt{p_{a,b}}|b| \\ &\leq \sqrt{\sum_{a,b} p_{a,b}a^2} \cdot \sqrt{\sum_{a,b} p_{a,b}b^2} \\ &= \sqrt{\mathbb{E}[A^2] \cdot \mathbb{E}[B^2]}\end{aligned}$$

■

Proof of Lemma 6: Define

$$Y = \begin{cases} 1 & \text{if } N > 0, \\ 0 & \text{else.} \end{cases}$$

Then by Lemma 7, $\mathbb{E}[N]^2 = \mathbb{E}[NY]^2 \le \mathbb{E}[N] \cdot \mathbb{E}[Y] = \mathbb{E}\left[N^2\right] \cdot \Pr[N > 0]$, which proves what we want. ■

To use Lemma 6, first note that each vertex of a random path is uniformly distributed, and so lands inside S with probability γ. Thus,

Claim 8. *$\mathbb{E}[N] = t\gamma$.*

To bound $\mathbb{E}\left[N^2\right]$, we shall need the expander mixing lemma, which you have proved in class:

Lemma 9. *If U, V are subsets of vertices in a d-regular expander graph with eigenvalue bound $\lambda < 1$, and E denotes the number of edges from U to V, then*

$$\left| E - \frac{d|U||V|}{n} \right| \le \lambda d \sqrt{|U||V|}$$

Lemma 10. *If $\gamma < 1/t$, $\mathbb{E}\left[N^2\right] = O(t\gamma)$.*

Proof Define the binary random variables

$$Y_i = \begin{cases} 1 & \text{if the } i\text{'th vertex of the path lies in } S, \\ 0 & \text{else.} \end{cases}$$

Thus $N = \sum_{i=1}^{t} Y_i$, and $N^2 = \sum_{i \le j} Y_i Y_j$. If $i = j$, $\mathbb{E}[Y_i Y_j] = \mathbb{E}[Y_i] = \gamma$. If $i < j$, then

$$\begin{aligned} \mathbb{E}[Y_i Y_j] &= \Pr[Y_j = 1 | Y_i = 1] \Pr[Y_i = 1] \\ &= \gamma \Pr[Y_j = 1 | Y_i = 1] \\ &= \gamma \Pr[X_j \in S | X_i \in S]. \end{aligned}$$

Now X_j, X_i have the same distribution as taking a random step in the expander that is the $j - i + 1$'th power of G. The eigenvalue of this expander is λ^{j-i+1}. Thus, the expander mixing lemma says that of the total $d^{j-i+1}|S|$ edges coming out of S in this power, at most $d^{j-i+1}|S|^2/n + \lambda^{j-i+1} d^{j-i+1}|S| = d^{j-i+1}|S|(\gamma + \lambda^{j-i+1})$ of them go back to S.

Thus,

$$\mathbb{E}\left[N^2\right] = \mathbb{E}\left[\sum_{i\le j} Y_i Y_j\right]$$
$$\le \sum_{i=1}^{t} \gamma \left(1 + \sum_{j=i+1}^{t} (\gamma + \lambda^{j-i+1})\right)$$
$$\le \sum_{i=1}^{t} \gamma \left(1 + t\gamma + O(1)\right)$$
$$\le \sum_{i=1}^{t} \gamma \left(1 + O(1)\right) \le O(t\gamma)$$

■

By Lemma 10, Claim 8 and Lemma 6, we have proved Theorem 4.

Define ℓ in such a way that for a random walk $u_1, \ldots, u_t$ and a random $i \in [t]$, $\Pr[\ell'_{u_1}(u_i) = \ell_{u_i}]$ is maximized. More concretely, we think of each path of length t, $u_1, \ldots, u_t$ as voting the value $\ell'_{u_1}(u_i)$ for the value of ℓ_{u_1}, and then we set ℓ_{u_i} to be the symbol that got the most votes.

Now returning to the PCP. There are two cases. In the first case, the assignments $\ell'_v(v)$ violate half the edges of F (i.e. $\gamma/2$ fraction). In this case, every path that passes through F will detect this violation, so we are done.

If this is not true, then consider the set of vertices $S = \{u : \ell_u \ne \ell'_u(u)\}$. It must be that S constitutes an $\Omega(\gamma)$ fraction of all vertices.

Now let $u_1, \ldots, u_t$ be a random path, and let $N = \sum_{i=1}^{t} N_i$ denote the number of vertices u_i such that $\ell_{u_i} \ne \ell'_{u_1}(u_i)$. Then, since ℓ_u the value that got at least $1/k$ fraction of the votes, we must have that $1/k$ fraction of all paths that