CSE 533: The PCP Theorem and Hardness of Approximation **(Autumn 2005)**

Lecture 3: Expander Graphs and PCP Theorem Proof Overview

Oct. 5, 2005

Lecturer: Venkatesan Guruswami *Scribe: Matt Cary*

1 Key Expander Graph Lemmas

Recall in last lecture that we defined a (n, d, λ)-expander to be a d-regular n-vertex undirected graph with second eigenvalue λ. We also defined the edge expansion of a graph G with vertex set V to be

$$\phi(G) = \min_{\substack{S \subset V \\ |S| \leq n/2}} \frac{|E(S, \overline{S})|}{|S|},$$

where $E(S, \overline{S})$ is the set of edges between a vertex set S and its complement.

The following lemma shows that the eigenvalue formulation of an expander is essentially equivalent to edge expansion.

Lemma 1.1. *Let G be a (n, d, λ) expander. Then*

$$\phi(G) \geq (d - \lambda)/2.$$

Remark 1.2. *In the other, harder, direction, it is possible to show that $\phi(G) \leq \sqrt{2d(d-\lambda)}$. For our purposes of constructing and using expanders, the easier direction shown in this lemma is enough.*

Proof. Let V and E be the vertex and edge sets of G. Let $S \subset V$ with $|S| \leq n/2$. We will set up a vector x that is similar to the characteristic vector of S, but perpendicular to $\overrightarrow{1}$. Then by the Rayleigh coefficient formulation of λ, we have that $\|Ax\| \leq \lambda\|x\|$, where $A = A(G)$ is the adjacency matrix of G.

Accordingly, we define x by

$$x_v = \begin{cases} -|\overline{S}| & \text{if } v \in S \\ |S| & \text{if } v \in \overline{S} \end{cases},$$

and you can confirm that $\sum x_v = 0$ so that $x \perp \overrightarrow{1}$. Now combining the Raleigh coefficient with the fact that $\langle Ax, x\rangle \leq \|Ax\| \cdot \|x\|$ we get

$$\langle Ax, x\rangle \leq \lambda\|x\|^2.$$

Note that $(Ax)_u = \sum_{(u,v)\in E} x_v$, so that as A is symmetric

$$\begin{aligned}\langle Ax, x\rangle &= \sum_u x_u \sum_{(u,v)\in E} x_v \\ &= 2\sum_{(u,v)\in E} x_u x_v \\ &= 2|E(S,\overline{S})|\cdot\big(-|S|\cdot|\overline{S}|\big) + \big(d|S| - E(S,\overline{S})\big)|\overline{S}|^2 + \big(d|\overline{S}| - E(S,\overline{S})\big)|S|^2\end{aligned}$$

where in the last two terms we count the number of edges wholly in S and $\overline{S}$, respectively: there are $d|S|$ edge originating in S, minus those that cross over to $\overline{S}$, cut in half as we have double counted. We then simplify by noting that $|S| + |\overline{S}| = n$ and $(|S|^2 + 2|S||\overline{S}| + |\overline{S}|^2 = (|S| + |\overline{S}|)^2 = n^2$ to get

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

Hence as $\langle Ax, x\rangle \le \lambda\|x\|^2$ and $\|x\|^2 = |S||\overline{S}|^2 + |\overline{S}||S|^2 = |S||\overline{S}|n$, we can say

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

implying

$$|E(S,\overline{S})| \ge \frac{d-\lambda}{n}|S||\overline{S}|$$

which shows

$$\frac{|E(S,\overline{S})|}{|S|} \ge \frac{d-\lambda}{2}$$

as $|\overline{S}|/n \ge 1/2$ by our assumption on S. □

Suppose we take a walk of length t from a random vertex v in G, as shown in Figure 1, and are interested in the probability that the entire walk takes place within the set B of size βn. If each vertex in the walk were picked independently, the probability would be β^t. The next lemma shows that even when the vertices are taken from a walk, in a good expander we are not that far off from independent vertex choices. This illustrates why expanders are so useful: structures that should be very dependent, such as walks, look very close to independent. We will not give the proof of this lemma. We will prove a different (easier) result concerning random walks in expanders in Lemma 3.1, and later use it when proving the PCP theorem.

Lemma 1.3. *Let G be a (n, d, λ)-expander, and $B \subset V(G)$ a set of size βn. Then the probability that a t step walk starting from a vertex v never leaves B is at most*

$$\left(\sqrt{\beta^2 + \left(\frac{\lambda}{d}\right)^2(1-\beta^2)}\right)^t,$$

where the probability is taken over the uniform choice of v as well as the steps in the walk.

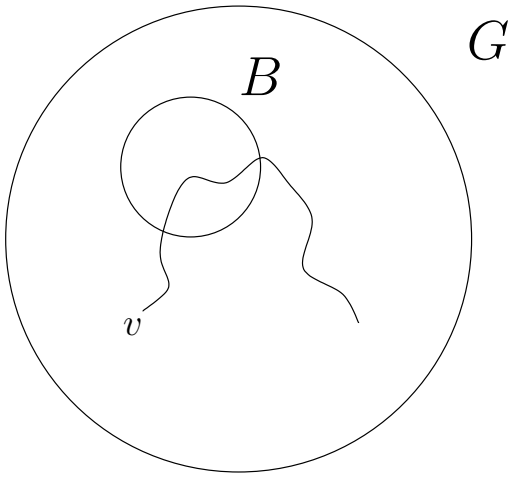

Figure 1: Hitting a set B during a walk in a graph G. Here the walk only intersects B; Lemma 1.3 bounds the probability that the entire walk starts and remains inside B.

Remark 1.4. *Note that as $\lambda \to 0$, the probability approaches β^t as our intuition about independent vertex choices suggests. Also, if $\beta = 1 - \epsilon$ for $\epsilon \to 0$, then the probability is at most $(1 - (1 - \lambda^2/d^2)(1 - \beta^2))^{t/2} \leq 1 - O(t\epsilon)$, or in other words, a t-step random walk has probability $\Omega(t\epsilon)$ of hitting a set of density ϵ, for small ϵ. This fact will later be appealed to sketch the basic intuition behind Dinur's proof.*

2 Constructions of Explicit Expanders

In case you are starting to wonder if such marvelous objects as expanders can exist at all, let alone with interesting parameters, we survey couple of constructions that show that some constant-degree regular graphs exist with good expansion and they can be described very explicitly.

2.1 The Margulis/Gaber-Galil Expander

We construct an n^2 vertex graph G whose vertex set is $\mathbb{Z}_n \times \mathbb{Z}_n$, where $\mathbb{Z}_n$ is the ring of integers modulo n. Given a vertex $v = (x, y)$, we connect it to the following vertices:

$$\begin{array}{cc} (x + 2y, y) & (x, 2x + y) \\ (x + 2y + 1) & (x, 2x + y + 1) \end{array},$$

where all operations are done modulo n. We also add the edges corresponding to the inverse transformations. This is then an 8-regular undirected graph, where there may be self loops or multiedges, depending on n. One can prove that for this graph $\lambda \leq 5\sqrt{2} < 8$.

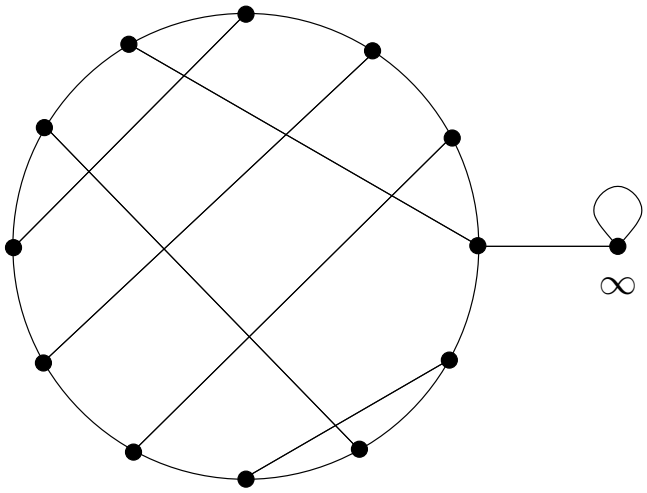

Figure 2: The LPS Expander

2.2 The Lubotzky-Phillips-Sarnak Expander

The construction presented here is much simplified from the original construction. Let $V = \mathbb{Z}_p \cup \{\infty\}$, where p is prime. We view V as a p-element field defined by addition and multiplication modulo p, where we extend the multiplicative inverse by defining 0^{-1} to be the special point ∞, and $\infty + x = \infty$ for all $x \in V$. Given any vertex x, connect it to $x+1$, $x-1$ and x^{-1}. That's it! This gives a 3-regular graph with second largest eigenvalue $\lambda \leq \lambda_0 < 3$ for some absolute constant λ_0. The structure of the graph as shown in Figure 2. The graph is a cycle with a matching between the edges. Note that the extra point with a self-loop that is introduced by the point ∞ can be removed along with 0, and 1 connected to $p-1$ without much disturbing the expansion. Actually since 1 and $p-1$ also have self loops for their inverses (instead of matching edges, though this isn't reflected in the figure!), we can remove them, and simply have a simple graph that is a cycle on $p-3$ nodes with a matching.

3 A Final Expander Lemma

In this section we prove a version of Lemma 1.3 that will be used in the proof of the PCP theorem. Here the set of interest will be an *edge* set $F \subset E$, we consider a walk that starts along a particular edge in F, and consider the chance that the edge used in the t^{th} step is also in F.

Lemma 3.1. *Let G be an (n, d, λ)-expander and $F \subset E(G) = E$. Then the probability that a random walk, starting in the zero-th step from a random edge in F, passes through F on its t^{th} step is bounded by*

$$\frac{|F|}{|E|} + \left(\frac{\lambda}{d}\right)^{t-1}.$$

Proof. To prove this lemma we will use a very useful technique essentially that used for random walks on Markov chains. Let x be the distribution on the vertices of G for the start of the walk. That is, x_v is the probability that our walk begins at vertex v. Consider the first step in the random walk. The probability that this ends at a vertex u is the sum, over all edges (v, u), of the probability that we were on v before this step, times the probability we chose the edge (v, u) out of all the other edges leaving v. As G is d-regular, v has exactly d edges leaving it, the chance we take the one heading to u is just $1/d$ (we will ignore the possibilities of multi-edges—as you will see, the final expression we actually use takes this into account). Hence if x' is the distribution on the vertices of G after the first step,

$$x'_u = \sum_{(v,u)\in E} x_v/d.$$

Now let A be the adjacency matrix of G. Then the row A_u has ones in exactly the columns (v, u) where there is an edge (v, u) in G. Hence we can write the above expression compactly as $x' = Ax/d$. Note that in this case, multi-edges are handled correctly, for if there is a multi-edge of multiplicity m between v and u, the corresponding entry in A will be m, giving the probability we take that edge from v as m/d as desired. This notation is so convenient we will normalize by defining $\tilde{A} = A/d$ so that simply

$$x' = \tilde{A}x.$$

If we take i steps, the distribution we reach is given by $\tilde{A}^i x$. Let P be the probability we're interested in, which is that of traversing an edge of F in the t^{th} step. Suppose w is the vertex we arrive at after the $(t-1)^{\text{th}}$ step. Let y_w be the number of edges of F incident on w, divided by d. Then $P = \sum_{w\in V} (\tilde{A}^{i-1}x)_w y_w$, where x is the initial distribution.

To calculate x, we pick an edge in F, then pick one of the endpoints of that edge to start on. If v has k edges of F incident on it, we have a $k/|F|$ chance to pick that edge, then a further $1/2$ chance to pick v. Now, y_w is the same quantity k, but divided by d instead of $2|F|$. Hence, we can write that $y_w = x_w \cdot 2|F|/d$. Without calculating x further, we now write

$$\begin{aligned} P &= \sum_{w\in V} (\tilde{A}^{i-1}x)_w y_w \\ &= \sum (\tilde{A}^{i-1}x)_w x_w \cdot \frac{2|F|}{d} \\ &= \frac{2|F|}{d} \langle \tilde{A}^{i-1}x, x\rangle. \end{aligned}$$

To finish our calculation, we rely on an as-yet unused property of G: its regularity. As each vertex in G has exactly d neighbors, each row in $\tilde{A}$ sums to one. Hence if $x^{\|}$ is the uniform distribution on G—$x^{\|}_v = 1/n$—then $\tilde{A}x^{\|} = x^{\|}$. As x is a probability distribution, we can decompose it as $x = x^{\|} + x^{\perp}$ with $\langle x^{\|}, x^{\perp}\rangle = 0$. Then by linearity and the fact just mentioned,

$$\begin{aligned} \tilde{A}^{i-1}x &= \tilde{A}^{i-1}x^{\|} + \tilde{A}^{i-1}x^{\perp} \\ &= x^{\|} + \tilde{A}^{i-1}x^{\perp}. \end{aligned}$$

Hence,

$$\begin{aligned}
\langle \tilde{A}^{i-1}x, x\rangle &= \langle \tilde{A}^{i-1}x^{\|}, x\rangle + \langle \tilde{A}^{i-1}x^{\perp}, x\rangle \\
&= \langle x^{\|}, x\rangle + \langle \tilde{A}^{i-1}x^{\perp}, x\rangle \\
&= \|x^{\|}\|^2 + \langle \tilde{A}^{i-1}x^{\perp}, x\rangle \\
&= \frac{1}{n} + \langle \tilde{A}^{i-1}x^{\perp}, x\rangle \\
&\leq \frac{1}{n} + \|\tilde{A}^{i-1}x^{\perp}\| \cdot \|x\| \\
&\leq \frac{1}{n} + \left(\frac{\lambda}{d}\right)^{i-1} \|x^{\perp}\| \cdot \|x\| \\
&\leq \frac{1}{n} + \left(\frac{\lambda}{d}\right)^{i-1} \|x\|^2,
\end{aligned}$$

as $\|x^{\perp}\| \leq \|x\|$. Now notice as the entries of x are positive that $\|x\|^2 = \sum x_v^2 \leq \max x_v \sum x_v = \max x_v$, as $\sum x_v = 1$, x being a probability distribution. The maximum x_v is achieved when all edges incident to v are in F, and in that case $x_v = d/(2|F|)$, so the calculation above continues

$$\leq \frac{1}{n} + \left(\frac{\lambda}{d}\right)^{i-1} \frac{d}{2|F|}.$$

Hence

$$P \leq \frac{2|F|}{dn} + \left(\frac{\lambda}{d}\right)^{i-1}$$

which finishes the proof as $|E| = nd/2$.

□

4 Overview of the GAP-3SAT Hardness Proof

In this section we give an overview of the proof of hardness for GAP-3SAT that will occupy us over the next several lectures. We will actually prove the hardness of a problem that can be seen as a generalization of graph optimization problems, which has an easy reduction to GAP-3SAT.

Definition 4.1. *A* constraint graph *is given by an alphabet Σ, a graph $G = (V, E)$ and a set of constraints $C = \{c_e \subseteq \Sigma \times \Sigma \mid e \in E\}$. A* labeling *on G is an assignment $\sigma : V \to \Sigma$ of elements from Σ to each vertex of G. A labeling σ* satisfies *an edge $(u, v) \in E$ if $\big(\sigma(u), \sigma(v)\big) \in c_e$ (for each edge a canonical orientation $u \to v$ is assumed).*

The optimization problem *for a constraint graph is to find a labeling that maximizes the number of satisfied edges. The* gap problem *for constraint graphs with* gap parameter ϵ, $0 < \epsilon \leq 1$, *is the following: given a constraint graph, such that either (i) there is a labeling that satisfies all the edges, or (ii) every labeling fails to satisfy at least a fraction ϵ of edges, determine which of the cases (i) or (ii) holds.*

Remark 4.2. *Many graph problems can be expressed as a constraint graph problem. For example, given a graph G to k-color, let Σ be a k-element alphabet, and define the set of constraints C as $\{(a, b)\}_{a \neq b}$. The optimization problem is to find a coloring for G that maximizes the number of valid edges, those whose endpoints are different colors.*

Note that the hardness of the decision problem for constraint graphs implies the hardness of the gap problem with gap parameter $1/|E|$. We will amplify this hardness in a series of stages, where at each stage we take a constraint graph G_i over alphabet Σ, and produce G_{i+1} also over alphabet Σ, where the number of unsatisfied edges in G_{i+1} is at least twice the number in G_i. If the size of G_{i+1} increases polynomially over the size of G_i, this will not be enough, as we will apply this $\log |E|$ times. Hence we must also insure that the size of G_{i+1} is at most a constant factor larger than that of G_i, so that the size of $G_{\log |E|}$ is a polynomial in the size of G_1.

Each stage will be split into 4 steps. Let $\mathrm{gap}(G)$ denote the minimum fraction of unsatisfied edges over all labelings of G.

Sparsification (degree-reduce): $G_i \rightarrow G^{(1)}$, a d-regular graph where $\mathrm{gap}(G_i^{(1)}) \geq \beta_1 \mathrm{gap}(G_i)$, with integer d, $\beta_1 > 0$ being absolute constants. This step is achieved by placing a $(d-1)$-regular expander at each vertex of G_i with number of vertices of the j'th expander being equal to the degree of vertex j.

Expanderize: $G^{(1)} \rightarrow G^{(2)}$, that is a good expander. We can achieve this by unioning $G^{(1)}$ with an expander, and apply Lemma 3.6 of the previous lecture. This step will also reduce the gap by an absolute constant factor.

Amplify the Gap: $G^{(2)} \rightarrow G^{(3)}$ by powering, which will increase $\mathrm{gap}(G^{(2)})$ a lot, more than making up for the loss of the other steps. This will also increase Σ, where we want the final graph to be over the same alphabet as the original graph. This step was the main innovation of Dinur.

Composition (alphabet-reduce): $G^{(3)} \rightarrow G_{i+1}$ by reducing the alphabet back to the original Σ.

In addition, each step will only increase the size of the graph by a constant factor.
It is worth pointing out that expanders are used in each of the first three steps above!