(80630) Analytical Methods in Combinatorics and CS	January 17, 2007
Lecture 12	
Lecturer: Irit Dinur	*Scribe: Limor Ben Efraim*

1 Reminder from last lecture

In the last lecture we started to prove the following theorem related to the graphs $G = (V, E)$ and $G^t = (V, E')$:

Theorem 1 *There exist constants c, γ dependent only on d and $|\Sigma|$ such that*

$$unsat(G^t) \geq \min\{\gamma, c \cdot t \cdot unsat(G)\}$$

We recall the main idea of the proof: Let $t \in \mathbb{N}$ such that t is divisible by 8. We defined

$$T_8 = [-\frac{t}{8}, \frac{t}{8}], \quad T_4 = [-\frac{t}{4}, \frac{t}{4}], \quad T_2 = [-\frac{t}{2}, \frac{t}{2}]$$

The distribution $\mathcal{D}$ Let $\mathcal{D}$ be the following distribution on the set T_2:

- choose independently $j_1, j_2 \in T_4$.
- answer $\ell = j_1 + j_2$. (Note that $\ell \in T_2$)

Let P be the set of paths in the graph G.

The distribution $\mathcal{A}$ Let $\mathcal{A}$ be the following distribution on the set $\mathrm{P} \times \mathrm{T}_4$:

- choose $\ell \in_{\mathcal{D}} T_2$.
- choose random walk $\bar{p}$ in G at length $t + \ell$.
- choose uniformly $s \in T_4 \cap (\ell + T_4)$.
- answer $(\bar{p}, s)$.

Let Σ be finite set (The alphabet). Let $\Sigma' = \Sigma^{d^0+d^1+\ldots+d^t}$.
Let $H : V \to \Sigma'$ be the best assignment to V. That is, $unsat_H(G^t) = unsat(G^t)$.

We define $h : V \to \Sigma$ as follows:

- We choose uniformly $k \sim \frac{t}{2} + T_8 = [\frac{t}{2} - \frac{t}{8}, \frac{t}{2} - \frac{t}{8}]$.
- For $v \in V$, we choose random walk of length k which starts at v. We denote by w the end vertex of that walk.

- let

$$h(v) = \arg\max_{a\in\Sigma} \Pr_{w}(H(w)_{(v)} = a)$$

I.e., $h(v)$ is the most common value from Σ under that distribution, the most common opinion at the end of the random walks.

Let $F \subseteq E$ be the edges that h violates.
Let $F' \subseteq E'$ be the edges that H violates.
Actually, our target is to see that

$$unsat(G^t) = \frac{|F'|}{|E'|} \geq c \cdot t \cdot \frac{|F|}{|E|} = c \cdot t \cdot unsat(G).$$

Let $i \in T_8$. Let $\bar{p} = (v_1, ..., v_{t+\ell}) \in \mathrm{P}$, $v = v_{i+s+\frac{t}{2}+1}$, $s \in T_4$, $u = v_{i+s+\frac{t}{2}}$. Let $A_i(\bar{p}, s)$ be an event on $\mathrm{P} \times T_4$ which occurs iff:

- $H(v_1)_{(u)} = h(u)$
- $H(v_{t+\ell})_{(v)} = h(v)$
- $(u, v) \in F$

Notice that if there exists i such that $A_i(\bar{p}, s)$ occurs then $(v_1, v_{t+\ell}) \in F'$.
Let $x_i(\bar{p}, s) = 1_{A_i(\bar{p},s)}$ and $N = \sum_i x_i(\bar{p}, s)$.

2 Finishing the proof of Theorem 1

Claim 2

$$unsat_H(G^t) \geq Prob_{(\bar{p},s)\sim\mathcal{A}}(N(\bar{p}, s) > 0)$$

Proof If $N(\bar{p}, s) > 0$, then at least there exist i such that $A_i(\bar{p}, s)$ occurs, and there are edges in G^t that are violated.
Let us notice that by looking at the events A_i we did not count all the possibilities to violations of edges in G^t. Therefore $Prob_{(\bar{p},s)\sim\mathcal{A}}(N(\bar{p}, s) > 0)$ is actually a lower bound to $unsat_H(G^t)$. ■

Therefore, it is left to show:

Lemma 3

$$Prob_{(\bar{p},s)\sim\mathcal{A}}\Big[N(\bar{p}, s) > 0\Big] > c\,t\,\frac{|F|}{|E|}.$$

Before we prove the Lemma we need few help Lemmas.

Claim 4 *Let Y be non negative random variable. Then*

$$Prob\Big[Y > 0\Big] \geq \frac{(\mathbb{E}[Y])^2}{\mathbb{E}(Y^2)}$$

Proof

$$\mathbb{E}[Y] = \mathbb{E}\Big[Y \cdot 1_{\{Y>0\}}\Big] \leq \sqrt{\mathbb{E}(Y^2)}\sqrt{\mathbb{E}(1_{\{Y>0\}^2})} = \sqrt{\mathbb{E}(Y^2)} \cdot \sqrt{Prob(Y>0)}$$

The inequality is by Cauchy-Schwartz inequality. ■

Lemma 5

$$\mathbb{E}_{\mathcal{A}}[N] \geq \frac{1}{16|\Sigma|^2} \cdot t \cdot \frac{|F|}{|E|}.$$

Proof We will prove that for each $i \in T_8$ we have $\mathbb{E}[x_i] \geq \frac{1}{4|\Sigma|^2} \cdot \frac{|F|}{|E|}$. Since there are $\frac{t}{4}$ elements in T_8, we will get the inequality.

The distribution $\mathcal{B}_i$ We define a distribution $\mathcal{B}_i$ on the probability space $P \times T_4$ as follows:

- We choose $(u, v) \in E$.
- We choose $j_1, j_2 \in T_4$ independently.
- We choose path of length $\frac{t}{2} + i + j_1$ starts from u and denote it by $(v_1, ..., v_{\frac{t}{2}+i+j_1} = u)$.
- We choose path of length $\frac{t}{2} - i + j_2$ starts from v and denote it by $(v = v_{\frac{t}{2}+i+j_1+1}, ..., v_{t+j_1+j_2})$.
- The answer is $s = j_1$ and $\bar{p} = (v_1, ..., v_{t+j_1+j_2})$.

We claim that $\mathcal{B}_i = \mathcal{A}$. Indeed, let $\ell = j_1 + j_2$. Then $\ell \in T_2$ distributes according to $\mathcal{D}$. In both distributions we pick random path at the same length, (the length of the random path, $t + \ell$ is fixed by the same distribution $\mathcal{D}$ at both distributions $\mathcal{A}$ and $\mathcal{B}_i$). The graph is regular, so it does not matter which start point we start.

Calculate $\mathbb{E}_{\mathcal{B}_i}[x_i] = \mathbb{E}_{\mathcal{A}}[x_i]$: By the definition of the event A_i we have

$$Prob[x_i > 0] = Prob[(u, v) \in F] \cdot Prob\Big[H(v_1)_{(u)} = h(u) \mid (u, v)\Big] \cdot Prob\Big[H(v_{t+j_1+j_2})_{(v)} = h(v) \mid (u, v)\Big] \tag{1}$$

We notice that since the graph is regular $Prob[(u, v) \in F] = \frac{|F|}{|E|}$, and the above events are independent, so we can calculate them separately.

We now calculate $Prob[H(v_1)_{(u)} = h(u) \mid (u, v)]$:

$$Prob\Big[H(v_1)_{(u)} = h(u) \mid (u, v)\Big] = Prob\Big[i + \frac{t}{2} + j_1 \in \frac{t}{2} + T_8\Big] \cdot Prob\Big[H(v_1)_{(u)} = h(u) \mid u = v_{\frac{t}{2}+i+j_1},\ j_1 \in T_4\Big] \tag{2}$$

Lower bound to first term in the above equation: from the definition of $\mathcal{B}_i$, the length of the path $(v_1, ..., v_{\frac{t}{2}+i+j_1} = u)$ is distributed uniformly in the set $S_{i,j_1} = \{\ \frac{t}{2} + i + j_1 :\ \ j_1 \in T_4\}$. Noticing that $\frac{t}{2} + T_8 \subset S_{i,j+1},\ |T_4| = \frac{t}{2} + 1,\ |T_8| = \frac{t}{4} + 1$ and

$$|S_{i,j+1}| = |T_4| < 2\,|\frac{t}{2} + T_8| = 2|T_8|,$$

we get

$$Prob\Big[i + \frac{t}{2} + j_1 \in \frac{t}{2} + T_8\Big] \geq \frac{1}{2}$$

Since there are at most $|\Sigma|$ possible 'opinions' on u, we can bound the second term in equation (2):

$$Prob\Big[H(v_1)_{(u)} = h(u)\ |\ u = v_{\frac{t}{2}+i+j_1},\ j_1 \in T_4\ \Big] \geq \frac{1}{|\Sigma|}$$

We get that

$$Prob\Big[H(v_1)_{(u)} = h(u)\ |\ (u,v)\Big] > \frac{1}{2\,|\Sigma|}$$

Similarly we have,

$$Prob\Big[H(v_{t+j_1+j_2})_{(v)} = h(v)\ |\ (u,v)\Big] > \frac{1}{2\,|\Sigma|}$$

Summing up all the terms in equation (1), we get

$$\mathbb{E}[x_i] = Prob[x_i > 0] > \frac{|F|}{|E|} \cdot \frac{1}{4\,|\Sigma|^2}$$

And

$$\mathbb{E}[N] = \sum_{i \in T_8} \mathbb{E}[x_i] \geq \frac{|F|}{|E|} \cdot t \cdot \frac{1}{16\,|\Sigma|^2}$$

■

Lemma 6

$$\mathbb{E}_{\mathcal{A}}[N^2] \leq c'\,t\,\frac{|F|}{|E|}$$

For the proof of Lemma 6 we need the following claim:

Claim 7 *For $i < j$,*

$$\mathbb{E}_{(\bar{p},s)\sim\mathcal{A}}\Big[Y_i(\bar{p},s) \cdot Y_j(\bar{p},s)\Big] \leq \frac{|F|}{|E|} \cdot \left(\frac{|F|}{|E|} + (\frac{\lambda}{d})^{j-i}\right)$$

Notice that

$$\mathbb{E}_{(\bar{p},s)\sim\mathcal{A}}\Big[Y_i(\bar{p},s) \cdot Y_j(\bar{p},s)\Big] = \sum_{s \in T_4, l \in T_2} Prob(Y_i \cdot Y_j > 0|\ s,\ell) \cdot Prob_{s\in T_4,\ell\in T_2}\Big[(s,\ell)\Big]$$

thus we can prove the inequality to each couple (s,ℓ) separately. Actually this reduces to the following question: given random walk in the graph and the first step is in F, what is the probability that the $j-i$'s step is also in F?

We sketch a proof for similar easier claim on the vertices, which is enough, since we can take the set of vertices as

$$S = \{v \in V : \ \exists u, \ (v,u) \in F \ \}.$$

Claim 8 *Let $G=(V,E)$ be a (n,d,λ) expander, and let $S \subset V$. We start random walk in G, $(v_1, v_2, ...)$. Then*

$$Prob\Big[v_i \in S\Big] \leq \frac{|S|}{|V|} + (\frac{\lambda}{d})^i$$

Sketch of Proof The idea of the proof is very similar to the idea of the proof of Lemma 7 in Lecture 6 and as we did in exercise 2.
Let 1_S be the indicator vector of S. Then

$$Prob\Big[v_i \in S\Big] =< 1_S, \frac{1}{|S|} \cdot A^i 1_S > .$$

And we can continue in a similar way to claim 7 in lecture 6. I.e., we write $1_S = x_0 + x_1$ where x_0 is the orthogonal projection of 1_S on the vector $\frac{s}{n} \cdot \mathbf{1}$ (vector of ones in all the coordinates) and let x_1 be the projection on the subspace orthogonal to $\mathbf{1}$. Then,

$$Prob\Big[v_i \in S\Big] = \frac{1}{|S|} < 1_S, A^i 1_S >= \frac{1}{|S|} < x_o, A^i x_0 > + \frac{1}{|S|} < x_1, A^i x_1 >$$

$$\leq \frac{1}{|S|}\|x_0\|^2 + \frac{1}{|S|}(\frac{\lambda}{d})^i \|x_1\|^2 \leq \frac{|S|}{n} + (\frac{\lambda}{d})^i$$

■

Proof of Lemma 6: Given $\bar{p} = (v_1, ..., v_{t+l})$, $u = v_{i+s+\frac{t}{2}}, v = v_{i+s+\frac{t}{2}+1}$, let $A'_i(\bar{p}, s)$ be the event that $(u,v) \in F$. Let $Y_i(\bar{p}, s) = 1_{A'_i(\bar{p},s)}$. Note that if $A_i(\bar{p}, s)$ occurs then $A'_i(\bar{p}, s)$ also occurs, therefore for all $(\bar{p}, s)$:

$$Y_i(\bar{p}, s) \geq X_i(\bar{p}, s).$$

Hence, and by the fact that $X_i^2 = X_i$,

$$\mathbb{E}[N^2] = \mathbb{E}\Big[(\sum_{i \in T_8} x_i)^2\Big] = \mathbb{E}_{(\bar{p},s)\sim\mathcal{A}}\Big[\sum_{i,j \in T_8} X_i(\bar{p},s) X_j(\bar{p},s)\Big]$$

$$= \sum_i \mathbb{E}_{(\bar{p},s)\sim\mathcal{A}}\Big[X_i(\bar{p},s)\Big] + \sum_{i \neq j} \mathbb{E}_{(\bar{p},s)\sim\mathcal{A}}\Big[X_i(\bar{p},s) X_j(\bar{p},s)\Big]$$

$$\leq \sum_{i \in T_8} \mathbb{E}_{(\bar{p},s)\sim\mathcal{A}}\Big[Y_i(\bar{p},s)\Big] + 2\sum_{i<j} \mathbb{E}_{(\bar{p},s)\sim\mathcal{A}}\Big[Y_i(\bar{p},s) Y_j(\bar{p},s)\Big]$$

By the regularity of the graph G we have,

$$\mathbb{E}[Y_i] = Prob\Big[Y_i > 0\Big] = Prob\Big[(u,v) \in F\Big] = \frac{|F|}{|E|}$$

Hence,

$$\sum_{i\in T_8} \mathbb{E}Y_i = |T_8| \cdot \frac{|F|}{|E|}$$

As a corollary from claim 7 we get

$$2\sum_{i<j} \mathbb{E}Y_i Y_j = 2\sum_{i<j} \frac{|F|}{|E|} \cdot \left(\frac{|F|}{|E|} + (\frac{\lambda}{d})^{j-i}\right) = \sum_{i} \frac{|F|}{|E|} \cdot \left(\frac{|F|}{|E|} + (\frac{\lambda}{d})^{i}\right)$$

$$\leq c'(\lambda, d) \sum_{i} \frac{|F|}{|E|} = c' \cdot t \cdot \frac{|F|}{|E|}$$

and we calculate c as follows:

$$c' = \sum_i (\frac{\lambda}{d})^i = \frac{1}{1-(\frac{\lambda}{d})}.$$

■

Proof of Lemma 3: By claim 4 and Lemmas 5,6 we have

$$Prob_{(\bar{p},s)\sim\mathcal{A}}\Big[N(\bar{p},s) > 0\Big] \geq \frac{(\mathbb{E}N)^2}{\mathbb{E}(N^2)} > c \cdot t \cdot \frac{|F|}{|E|}$$

■