In this lecture we prove the only remaining missing step in the proof of the first half of the PCP Theorem, and we begin a description of the second half of the proof.

1 A Lemma on Random Walks on Expanders

In the last lecture, we stated the following result without proof.

**Lemma 1** Let $G = (V, E)$ be a $d$-regular graph with $\lambda_2(G) \leq \lambda < 1$, let $F \subseteq E$, and define $\delta = |F|/|E|$. Pick a random walk of length $\ell$ in $G$, and let $M$ be the number of edges of $F$ traversed in the walk.

Then

$$E[M^2] = O(\delta \ell + \delta^2 \ell^2)$$

We denote the random walk as the sequence $v_0, v_1, \ldots, v_\ell$ (where each $v_i$ is a random variable), and we introduce the 0/1 random variables $X_1, \ldots, X_\ell$, defined so that $X_i = 1$ if $(v_{i-1}, v_i) \in F$, and $X_i = 0$ otherwise. Hence

$$M = X_1 + X_2 + \cdots + X_\ell$$

and, using linearity of expectation and the fact that $X_i^2 = X_i$,

$$E[M^2] = \sum_{i,j} E[X_i X_j] = \sum_i E[X_i] + 2 \sum_i \sum_{j>i} E[X_i X_j]$$

Every edge in a random walk is uniformly distributed, and so

$$\sum_i E[X_i] = \delta \ell$$

It remains to bound the cross products. Our strategy will be to show that, for every $i$ and $j > 1$, we have

$$E[X_i X_j] \leq \delta^2 + \delta \lambda^{j-i-1} \tag{1}$$

Assuming that we have Equation (1), then for every $i$ we have

$$\sum_{j>i} E[X_i X_j] \leq \delta^2 \ell + \delta \sum_{k=0}^{j-i-1} \lambda^k \leq \delta^2 \ell + \delta \frac{1}{1 - \lambda}$$

and so

$$2 \sum_i \sum_{j>i} E[X_i X_j] = O(\delta^2 \ell^2 + \delta \ell)$$
as required.

We now turn to the proof of Equation (1). First, note that

\[
E[X_iX_j] = \Pr[(v_{i-1}, v_i) \in F \land (v_{j-1}, v_j) \in F] \\
= \Pr[(v_{i-1}, v_i) \in F] \cdot \Pr[(v_{j-1}, v_j) \in F \mid (v_{i-1}, v_i) \in F] \\
= \delta \cdot \Pr[(v_{j-1}, v_j) \in F \mid (v_{i-1}, v_i) \in F]
\]

Which means that proving Equation (1) reduces to proving the bound

\[
\Pr[(v_{j-1}, v_j) \in F \mid (v_{i-1}, v_i) \in F] \leq \delta + \lambda^{j-i-1} \tag{2}
\]

Now, the distribution of the edge \((v_{j-1}, v_j)\) on a random walk conditioned on \((v_{i-1}, v_i)\) is the same as the distribution of the edge \((u_{j-i-1}, u_{j-i})\) in a random walk \(u_0, \ldots, u_{j-i}\) where \(u_0\) is chosen to be a random endpoint of a random edge of \(F\), and the subsequent steps are a length-\((j-i)\) random walk in \(G\).

Equation (2) can be abstracted as the following claim.

**Lemma 2** Let \(G = (V, E)\) be a \(d\)-regular graph with \(\lambda_2(G) \leq \lambda < 1\) and let \(F \subseteq E\). Let \(u_0, \ldots, u_k\) be a random walk in \(G\) where the starting point \(u_0\) is chosen by picking a random edge in \(F\) and then a random endpoint of the edge.

Then the probability that \((u_{k-1}, u_k)\) is in \(F\) is at most

\[
\frac{|F|}{|E|} + \lambda^{k-1}
\]

**Proof:** Let \(M\) be the transition matrix of a random walk on the graph and let \(d_F(v)\) denote the number of edges incident on \(v\) that belong to \(F\). Then the initial distribution vector \(x\) (describing the distribution of \(u_0\)) is of the form \(x(v) = \frac{d_F(v)}{2|F|}\). The distribution \(z\) after \(k-1\) steps (the distribution of \(u_{k-1}\)) is given by \(z = xP^{k-1}\). If the walk is at vertex \(v\) after \(k-1\) steps, then the probability that the last step will be along an edge in \(F\) is

\[
\Pr[(u_{k-1}, u_k) \in F] = \sum_v z(v) \frac{d_F(v)}{d} = \frac{2|F|}{d} x^T = \frac{2|F|}{d} xM^{k-1}x^T
\]

To obtain a bound on \(xM^{k-1}x^T\), we split \(x\) as \(x = x_\parallel + x_\perp\) where \(x_\parallel\) and \(x_\perp\) are respectively parallel and perpendicular to the uniform distribution. Specifically, \(x_\parallel(v) = \frac{1}{n}\) and \(x_\perp(v) = x(v) - \frac{1}{n}\). Then

\[
xM^{k-1}x^T = x_\parallel M^{k-1}x^T + x_\perp M^{k-1}x^T \\
\leq \langle x_\parallel, x \rangle + \left\| x_\perp M^{k-1} \right\| \left\| x_\parallel \right\| \ (\text{because } x_\parallel M = x_\parallel) \\
\leq \frac{1}{n} + \lambda^{k-1} \left\| x_\parallel \right\|^2 \ (\text{since } \left\| x_\perp M \right\| \leq \lambda \left\| x_\parallel \right\| \text{ and } \left\| x_\parallel \right\| \leq \left\| x \right\|)
\]
Also
\[ ||x||^2 = \sum_v \frac{(d_F(v))^2}{(2|F|)^2} \leq \max_v \frac{d_F(v)}{2|F|} \sum_v \frac{d_F(v)}{2|F|} \]
\[ \leq \frac{d}{2|F|} \quad \text{(since } \sum_v d_F(v) = 2|F| \text{ and } d_F(v) \leq d) \]

Using these, we obtain the required result as
\[ \Pr [(u_{k-1}, u_k) \in F] = \frac{2|F|}{d} xM^{k-1}x^T \]
\[ \leq \frac{2|F|}{d} \left[ \frac{1}{n} + \lambda^{k-1} \frac{d}{2|F|} \right] \]
\[ = \frac{2|F|}{dn} + \lambda^{k-1} \]
\[ = \frac{|F|}{|E|} + \lambda^{k-1}. \]

\[ \square \]

2 An Overview of the Rest of the Proof of the PCP Theorem

To complete the proof of the PCP Theorem we will need to establish the following result.

Lemma 3 (Range Reduction) \( \exists \Sigma_0, \exists c_0 > 0, \) such that for all \( \Sigma, \) there exists a poly-time \( R_2, \) mapping Max-2-CSP-\( \Sigma \) to Max-2-CSP-\( \Sigma_0 \) such that:

- For every \( C, \) size\((R_2(C)) = O(\text{size}(C)); \)
- If \( C \) is satisfiable, then \( R_2(C) \) is satisfiable;
- If \( \text{opt}(C) \leq 1 - \delta, \) then \( \text{opt}(R_2(C)) \leq 1 - c_0 \delta. \)

We say that an instance of Max-2-CSP-\( \Sigma \) is in “projection form” if every constraint is of the form \( x = f(y), \) where \( f : \Sigma \to \Sigma \) can be arbitrary function, possibly dependent on the constraint.

The main result in the proof of Lemma 3 will be the following.

Lemma 4 (Reduction to Boolean CSP) There is a \( q \) and a \( c_1 \) such that for all \( \Sigma, \) there exists a poly-time reduction \( R_B, \) mapping Max-2-CSP-\( \Sigma \) to Max \( q \)-CSP-\{0,1\}. such that:

- For every \( C, \) size\((R_B(C)) = O(\text{size}(C)); \)
- If \( C \) is satisfiable, then \( R_B(C) \) is satisfiable;
- If \( \text{opt}(C) \leq 1 - \delta, \) then \( \text{opt}(R_B(C)) \leq 1 - c_1 \delta. \)
The proof of Lemma 3 is completed by the following easier reduction.

**Lemma 5 (Reduction to Projection Form)** For all $\Sigma$ and $q$, there exists a poly-time reduction $R_P$, mapping Max $q$-CSP-$\Sigma$ to Max-2-CSP-$\Sigma^q$ in projection form such that:

- For every $C$, $\text{size}(R_P(C)) = O(\text{size}(C))$;
- If $C$ is satisfiable, then $R_P(C)$ is satisfiable;
- If $\text{opt}(C) \leq 1 - \delta$, then $\text{opt}(R_P(C)) \leq 1 - \delta/q$.

To prove Lemma 3, we start from an instance $C$ of Max-2-CSP-$\Sigma$ and we use Lemma 5 to reduce it to an instance $C_1$ of Max-2-CSP-$\Sigma^2$ in projection form. Then we use Lemma 4 to reduce $C_1$ to an instance $C_2$ of Max $q$-CSP-$\{0,1\}$. Finally, we use Lemma 5 to reduce $C_2$ to an instance $C_3$ of Max 2-CSP-$\{0,1\}^q$. This proves Lemma 3 with $\Sigma_0 = \{0,1\}^q$, where $q$ is the constant of Lemma 4, and with $c_0 = c_1/2q$, where $c_1$ is the constant of Lemma 4.

We conclude this lecture with a proof of Lemma 5.

**PROOF:** [Of Lemma 5] Let $C$ be an instance of Max $q$-CSP-$\Sigma$ with variables $x_1, \ldots, x_n$ and constraints $C_1, \ldots, C_m$. The reduction produces a new instance that has variables $x_1, \ldots, x_n, y_1, \ldots, y_m$, that is, the same set of original variables, plus an extra variable per constraint of $C$. We also fix a surjective mapping of $\Sigma^q \to \Sigma$ so that we may think of an assignment to the $x_i$ in the new instance as an assignment to the $x_i$ in the original instance.

Each constraint $C_j$, over variables $x_{j_1}, \ldots, x_{j_q}$, is mapped into $q$ new constraints. The $i$-th such constraint, over variables $x_{j_i}$ and $y_j$, requires that, if we think of the value of $y_j$ as specifying an assignment to $x_{j_1}, \ldots, x_{j_q}$, then such assignment must satisfy $C_j$ and must be consistent with the assigned value of $x_{j_i}$.

Now we see that if $\text{opt}(C) \leq 1 - \delta$, then any assignment to the $x_j$ contradicts at least $\delta m$ constraints of $C$, and that, for each such constraint, no matter what is the assignment to the $y_j$, at least one of the $q$ constraints derived from it will be contradicted in the new instance. Hence, every assignment to the new instance contradicts at least $\delta m$ of the $qm$ constraints, and the optimum is at most $1 - \delta/q$. □