Expander Graphs and Their Applications (III)

Yijia Chen
Shanghai Jiaotong University

Last Lecture

Expander Graphs are Almost Good

Recall: A graph $G = (V, E)$ is (c, d)-good if each $v \in V$ has degree at most d and for all $u, v \in V$, if there is a path from u to v, then

$$\operatorname{dist}(u, v) \leq c \cdot \log |V|.$$

Definition

A d-regular graph G on n vertices is called an (n, d)-graph. Moreover, it is an (n, d, α)-graph if $\lambda(G) \leq \alpha d$.

Theorem

Let $\alpha < 1/2$. Then all (n, d, α) graphs are (c, d)-good for some appropriate $c \in \mathbb{N}$.

Random Walks

A vector $p = (p_1, \ldots, p_n) \in \mathbb{R}^n$ is a probability (distribution) vector if all $p_i \geq 0$ and $\sum_{i \in [n]} p_i = 1$. The probability vector for the *uniform distribution* is

$$\underline{u} := \frac{\mathbf{1}}{n} = \left\langle \frac{1}{n}, \ldots, \frac{1}{n} \right\rangle.$$

Definition

A random walk on a graph $G = (V, E)$ is a *discrete-time stochastic process* $(X_0, X_1, \ldots)$ taking values in V. The vertex X_0 is sampled from some *initial distribution* on V, and X_{i+1} is chosen *uniformly at random from the neighbors of X_i*.

Normalized Adjacency Matrices

Let G be a d-regular graph with adjacency matrix $A = A(G)$. Then its normalized adjacency matrix is

$$\hat{A} := \frac{1}{d} A.$$

- The matrix $\hat{A}^t$ is the transition matrix of the Markov Chain defined by *random walks of length t*, i.e., $(\hat{A}^t)_{i,j}$ is the probability *a random walk starting at i is at j after t steps*.
- The *stationary distribution* of the random walk on G is the uniform distribution, namely, $u\hat{A} = \hat{A}u = u$.
- If $d = \lambda_1 \geq \ldots \geq \lambda_n$ are the eigenvalues of A, then $1 = \hat{\lambda}_1 \geq \ldots \geq \hat{\lambda}_n$ with $\hat{\lambda}_i = \lambda_i / d$ for $i \in [n]$ are the eigenvalues of $\hat{A}$ with the same eigenvectors as of A.

Random Walks on Expander Graphs

Rapid Mixing of Walks

Lemma

For every probability vector p, $\|\hat{A}p - u\|_2 \leq \|p - u\|_2 \alpha \leq \alpha$.

Theorem

Let G be an (n, d, α)-graph with normalized adjacency matrix $\hat{A}$. Then for any distribution vector p and any positive integer t:

$$\|\hat{A}^t p - u\|_2 \leq \|p - u\|_2 \alpha^t \leq \alpha^t.$$

Proof of the Lemma

Obviously, $\langle p-u, u\rangle = 0$, that is, $p-u$ and u are orthogonal. Consequently, for some appropriate $(\beta_i)_{2\le i\le n}$ we have

$$p-u=\sum_{2\le i\le n}\beta_i v_i.$$

It is also easy to show $\|p-u\|_2 \le 1$.

$$\begin{aligned}
\left\|\hat{A}p-u\right\|_2 &= \left\|\hat{A}p-\hat{A}u\right\|_2 = \left\|\hat{A}(p-u)\right\|_2 \\
&= \left\|\sum_{2\le i\le n}\hat{\lambda}_i\beta_i v_i\right\|_2 = \sqrt{\left\langle \sum_{2\le i\le n}\hat{\lambda}_i\beta_i v_i, \sum_{2\le i\le n}\hat{\lambda}_i\beta_i v_i\right\rangle} \\
&= \sqrt{\sum_{2\le i\le n}\hat{\lambda}_i^2\beta_i^2\left\langle v_i, v_i\right\rangle} \le \alpha\sqrt{\sum_{2\le i\le n}\beta_i^2\left\langle v_i, v_i\right\rangle} \\
&= \alpha\left\|\sum_{2\le i\le n}\beta_i v_i\right\|_2 = \alpha\|p-u\|_2 \le \alpha.
\end{aligned}$$

□

Rapid Mixing of Walks (cont'd)

Lemma

For every $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$

$$\sum_{i \in [n]} |x_i| = \|x\|_1 \leq \sqrt{n}\|x\|_2 = \sqrt{n \sum_{i \in [n]} x_i^2}.$$

Theorem

Let G *be an* (n, d, α)*-graph with normalized adjacency matrix* $\hat{A}$*. Then for any distribution vector* p *and any positive integer* t*:*

$$\|\hat{A}^t p - u\|_1 \leq \sqrt{n}\, \alpha^t$$

Corollary

For every $\varepsilon > 0$*,* (n, d, α)*-graphs are* (c, d)*-good with*

$$c = -\frac{3}{2 \log \alpha} + \varepsilon.$$

Probabilistic Algorithms

Let D be a finite subset of $\mathbb{N}$.

Input:	$f, g : D \to \mathbb{N}$.
Problem:	Decide whether $f \neq g$, i.e., $f(a) \neq g(a)$ for some $a \in D$.

Clearly we have to do $|D|$ many comparisons in any *deterministic* algorithm.

Assume the functions f, g are low degree polynomials, i.e., let $d \in \mathbb{N}$.

Input:	Two polynomials $f, g : D \to \mathbb{N}$ of degree $\leq d$.
Problem:	Decide whether $f \neq g$.

We can do much better by using *randomness*.

RandomPolyTest(f, g)

1. Choose $a \in D$ *uniformly at random*.
2. **if** $f(a) \neq g(a)$ **then** accept.
3. **else** reject.

Probabilistic Algorithms (cont'd)

Lemma

- *If* $f \neq g$, *then* $\Pr_{a \in D}[f(a) = g(a)] \leq d/|D|$. *Then,* RandomPolyTest *will accept* (f, g) *with high probability*

$$1 - \frac{d}{|D|}.$$

- *If* $f = g$, *then* $\Pr_{a \in D}[f(a) = g(a)] = 1$. *Then,* RandomPolyTest *will always reject* (f, g) *correctly.*

Definition

A problem Q is in RP (Randomized Polynomial Time), if there is a *probabilistic polynomial time* algorithm $\mathbb{A}$ such that

$$x \in Q \implies \Pr[\mathbb{A} \text{ accepts } x] \geq \frac{1}{2};$$
$$x \notin Q \implies \Pr[\mathbb{A} \text{ accepts } x] = 0.$$

RP Error Reduction

If we *repeat* $\mathbb{A}$ *twice*, then

$$x \in Q \implies \Pr[\mathbb{A}^2 \text{ accepts } x] \geq \frac{3}{4};$$
$$x \notin Q \implies \Pr[\mathbb{A}^2 \text{ accepts } x] = 0.$$

More generally, for every $k \in \mathbb{N}$, if we repeat $\mathbb{A}$ on x *independently* for k times, then the error probability is *below* 2^{-k}.

Let r_x be *the number of the random bits* that $\mathbb{A}$ uses on x. If we repeat $\mathbb{A}$ on x independently for k times, then we use $k \cdot r_x$ *many random bits*.

Since randomness is *expensive*, thus the question:

Can we use less than $k \cdot r_x$ many random bits?

(The *ultimate goal* is to use no random bits at all.)

RP Error Reduction (cont'd)

Theorem (**Ajtai**, **Komlós**, and **Szemerédi**, 1987)

There is a probabilistic polynomial time that achieves the error probability below 2^k *using random bits*

$$r_x + O(k).$$

Sampling by Random Walks on Expander Graphs

Let $G = (V, E)$ be an (n, d, α)-graph, and $B \subseteq V$ with $|B| = \beta n$.

Intuitively: Fix some input $x \in Q$. Then V consists of all possible random bits that the algorithm $\mathbb{A}$ can use on input x. And B is the set of random bits that make the algorithm falsely reject x.

We carry out the following experiment: We pick $X_0 \in V$ uniformly at random and start from it a random walk $X_0, \ldots, X_t$ on G.

(B, t): this random walk is confined to B, i.e. that $X_i \in B$ for all $0 \leq i \leq t$.

Intuitively: (B, t) corresponds to the following event:

- We first choose a random string X_0 to run the algorithm $\mathbb{A}$ on x.
- Then for each $i \in [t]$, we choose X_i as a *random neighbor of X_{i-1} in G* for which we run the algorithm $\mathbb{A}$ on x.
- *All $t + 1$ runs of $\mathbb{A}$ falsely reject x.*

Sampling by Random Walks on Expander Graphs

Theorem (**Ajtai**, **Komlós**, and **Szemerédi**, 1987)

Let $G = (V, E)$ be an (n, d, α)-graph and $B \subseteq V$ with $B = \beta n$. Then

$$\Pr\left[(B, t)\right] \leq (\beta + \alpha)^t.$$

Compared to:

$$\Pr[\text{Independently choose } t \text{ many } X\text{s from } V \text{ and all in } B] = \beta^t.$$

Proofs

$P = \underline{P_B} = (p_{i,j})_{i,j\in[n]}$ with

$$p_{i,j} = \begin{cases} 1, & i = j \in B; \\ 0, & \text{otherwise.} \end{cases}$$

Lemma

$\Pr\left[(B,t)\right] = \left\|(P\hat{A})^t P u\right\|_1.$

Proofs (cont'd)

Lemma

For any vector $w \in \mathbb{R}^n$

$$\left\| P\hat{A}Pw \right\|_2 \le (\beta + \alpha)\|w\|_2.$$

Proof.

We can assume:

- $w = Pw$, i.e., $w_i = 0$ if $i \notin B$;
- $w = (w_1, \ldots, w_n)$ with $w_i \ge 0$ for all $i \in [n]$;
- $\sum_{i \in [n]} w_i = 1$.

Thus, $Pw = w = u + z$ with $z := w - u$, and $\langle u, z \rangle = 0$. Then

$$P\hat{A}Pw = P\hat{A}u + P\hat{A}z = Pu + P\hat{A}z \quad \text{and} \quad \left\| P\hat{A}Pw \right\|_2 \le \|Pu\|_2 + \|P\hat{A}z\|_2.$$

The result follows if we can show

$$\|Pu\|_2 \le \beta\|w\|_2 \quad \text{and} \quad \|P\hat{A}z\|_2 \le \alpha\|w\|_2.$$

Proof of the lemma (cont'd)

$\|Pu\|_2 \leq \beta\|w\|_2$:

By $Pw = w$, the support of w, i.e., $\{i \mid w_i \neq 0\}$, has at most βn coordinates. Then, by **Cauchy-Schwartz**,

$$1 = \sum_{i \in [n]} w_i \leq \sqrt{\beta n}\|w\|_2.$$

Furthermore, $\|Pu\|_2 = \sqrt{\beta/n}$. ⊣

$\left\|P\hat{A}z\right\|_2 \leq \alpha\|w\|_2$:

Recall $\langle u, z\rangle = 0$, thus we can write $z = \sum_{2 \leq i \leq n} \gamma_i v_i$. Then

$$\begin{aligned}\left\|P\hat{A}z\right\|_2 \leq \|\hat{A}z\|_2 &= \left\|\sum_{2 \leq i \leq n} \hat{\lambda}_i \gamma_i v_i\right\|_2 \leq \alpha \left\|\sum_{2 \leq i \leq n} \gamma_i v_i\right\|_2 \\ &= \alpha\|z\|_2 \leq \alpha\|w\|_2.\end{aligned}$$

⊣

Expansion and Spectral Gap

Recall:

Theorem (**Dodziuk**, 84; **Alon** and **Milman**, 85; **Alon**, 86)
Let G be a d-regular graph with spectrum $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$. Then

$$\frac{d-\lambda_2}{2} \leq h(G) \leq \sqrt{2d(d-\lambda_2)}.$$

$d - \lambda_2 = \lambda_1 - \lambda_2$ *is the spectral gap of* G.

Note, we cannot replace λ_2 by λ: Consider the complete bipartite graph $K_{2,2}$.

Proof of $(d - \lambda_2)/2 \leq h(G)$

Choose $S \subseteq V$ with $|S| \leq n/2$ and

$$h(G) = \frac{|\partial S|}{|S|}.$$

Then let $f := |\bar{S}| 1_S - |S| 1_{\bar{S}}$. So

$$\|f\|_2^2 = |\bar{S}|^2 |S| + |S|^2 |\bar{S}| = n|S||\bar{S}|,$$
$$fAf = E(S,S)|\bar{S}|^2 + |E(\bar{S},\bar{S})||S|^2 - 2|S||\bar{S}||E(S,\bar{S})|.$$

Moreover,

$$E(S,S) = d|S| - |E(S,\bar{S})|, \quad \text{and} \quad E(\bar{S},\bar{S}) = d|\bar{S}| - |E(S,\bar{S})|.$$

As $\langle f, \mathbf{1} \rangle = 0$,

$$\lambda_2 \geq \frac{fAf}{\|f\|_2^2} = \frac{nd|S||\bar{S}| - n^2|E(S,\bar{S})|}{n|S||\bar{S}|} = d - \frac{n|E(S,\bar{S})|}{|S||\bar{S}|} \geq d - 2h(G).$$

$\square$