Lecture 7: Expander Graphs in Computer Science

Lecturer: Kurt Mehlhorn & He Sun

Over the past decades expanders have become one basic tool in various areas of computer science, including algorithm design, complexity theory, coding theory, and etc. Informally, expanders are regular graphs with low degree and high connectivity. We can use different ways to define expander graphs.

1. **Combinatorically**, expanders are highly connected graphs, and to disconnect a large part of the graph, one has to sever many edges.
2. **Geometrically**, every vertex set has a relatively large boundary.
3. From the **Probabilistic** view, expanders are graphs whose behavior is "like" random graphs.
4. **Algebraically**, expanders are the real-symmetric matrix whose first positive eigenvalue of the Laplace operator is bounded away from zero.

1 Spectra of Graphs

Definition 1. *Let $G = (V, E)$ be an undirected graph with vertex set $[n] \triangleq \{1, \ldots, n\}$. The adjacency matrix of G is an n by n matrix $\mathbf{A}$ given by*

$$\mathbf{A}_{i,j} = \begin{cases} 1 & \text{if } i \text{ and } j \text{ are adjacent} \\ 0 & \text{otherwise} \end{cases}$$

If G is a multi-graph, then $\mathbf{A}_{i,j}$ is the number of edges between vertex i and vertex j.

By definition, the adjacency matrix of graphs has the following properties: (1) The sum of elements in every row/column equals the degree of the corresponding vertex. (2) If G is undirected, then $\mathbf{A}$ is symmetric.

Example. *The adjacency matrix of a triangle is*

$$\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

Given a matrix $\mathbf{A}$, a vector $\mathbf{x} \neq 0$ is defined to be an eigenvector of $\mathbf{A}$ if and only if there is a $\lambda \in \mathbb{C}$ such that $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$. In this case, λ is called an eigenvalue of $\mathbf{A}$.

Definition 2 (graph spectrum)**.** *Let* $\mathbf{A}$ *be the adjacency matrix of an undirected graph* G *with* n *vertices. Then* $\mathbf{A}$ *has* n *real eigenvalues, denoted by* $\lambda_1 \geq \cdots \geq \lambda_n$*. These eigenvalues associated with their multiplicities compose the spectrum of* G*.*

Here are some basic facts about the graph spectrum.

Lemma 3. *Let* G *be any undirected simple graph with* n *vertices. Then*

1. $\sum_{i=1}^{n} \lambda_i = 0$.
2. $\sum_{i=1}^{n} \lambda_i^2 = \sum_{i=1}^{n} \deg(i)$.
3. *If* $\lambda_1 = \cdots = \lambda_n$*, then* $E[G] = \emptyset$.
4. $\deg_{\text{avg}} \leq \lambda_1 \leq \deg_{\max}$.
5. $\sqrt{\deg_{\max}} \leq \lambda_1 \leq \deg_{\max}$.

Proof. We only prove the first three items.

(1) Since G does not have self-loops, all the diagonal elements of $\mathbf{A}$ are zero. By the definition of trace, we have $\sum_{1=1}^{n} \lambda_i = \operatorname{tr}(\mathbf{A}) = \sum_{i=1}^{n} \mathbf{A}_{i,i} = 0$.

(2) By the properties of matrix trace, we have $\sum_{1=1}^{n} \lambda_i^2 = \operatorname{tr}\left(\mathbf{A}^2\right) = \sum_{i=1}^{n} \mathbf{A}_{i,i}^2$. Since $\mathbf{A}_{i,i}^2$ is the degree of vertex i, $\operatorname{tr}(\mathbf{A}^2)$ equals the sum of all vertices' degrees in G.

(3) Combing $\sum_{i=1}^{n} \lambda_i = 0$ with $\lambda_1 = \cdots = \lambda_n$, we have $\lambda_i = 0$ for every vertex i. By item (2), we have $\deg(i) = 0$ for any vertex i. Therefore $E[G] = \emptyset$. □

For a graph G with adjacency matrix $\mathbf{A}$ and integer $k \geq 1$, $\mathbf{A}_{u,v}^k$ is the number of walks of length k from u to v .

> Let $\|\mathbf{x}\|_p \triangleq \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}$. Then for any $1 \leq p \leq q < \infty$, it holds that
>
> $$\|\mathbf{x}\|_q \leq \|\mathbf{x}\|_p \leq n^{1/p-1/q} \cdot \|\mathbf{x}\|_q.$$

Lemma 4. *For any graph* G *with* m *edges, the number of cycles of length* k *in* G *is bounded by* $\mathcal{O}\left(m^{k/2}\right)$*.*

Proof. Let $\mathbf{A}$ be the adjacency matrix of G with eigenvalues $\lambda_1, \ldots, \lambda_n$. Thus the number of C_k (cycles of length k) in G is bounded by $\operatorname{tr}\left(\mathbf{A}^k\right)/(2k) = \left(\sum_{i=1}^{n} \lambda_i^k\right)/(2k)$. For any $k \geq 3$, it holds that

$$\left(\sum_{i=1}^{n} \lambda_i^k\right)^{1/k} \leq \left(\sum_{i=1}^{n} |\lambda_i|^k\right)^{1/k} \leq \left(\sum_{i=1}^{n} |\lambda_i|^2\right)^{1/2} = (2 \cdot m)^{1/2}.$$

Hence $\operatorname{tr}\left(\mathbf{A}^k\right) \leq (2m)^{k/2}$ and the number of C_k is at most $\mathcal{O}\left(m^{k/2}\right)$. □

Example. *Some examples for different spectra of graphs:*

- *For the complete graph* K_n*, the eigenvalues are* $n-1$ *with multiplicity 1 and* -1 *with multiplicity* $n-1$*.*
- *For the complete bipartite graph* $K_{m,n}$*, the eigenvalues are* $+\sqrt{mn}$*,* $-\sqrt{mn}$ *and* 0 *with multiplicity* $m+n-2$*.*

- *For the cycle C_n, the spectrum is* $2\cos(2\pi j/n)$ $(j = 0, 1, \ldots, n-1)$.

Two assumptions that we make throughout the lecture are as follows:

1. We only consider *undirected graphs*. Note that if G is not undirected, then $\mathbf{A}$ is not symmetric any more and the eigenvalues of $\mathbf{A}$ could be complex numbers.

> A matrix $\mathbf{A} = (a_{i,j})_{n\times n}$ is called a Hermitian matrix if $a_{i,j} = \overline{a_{j,i}}$ for any element $a_{i,j}$. Hermitian matrices always have real eigenvalues.

2. Unless mentioned otherwise, we consider *regular graphs*.

Lemma 5. *Consider any undirected graph G with adjacency matrix $\mathbf{A}$.*

1. *If G is d-regular, then $\lambda_1 = d$ and $|\lambda_i| \le d$ for $i = 2, \ldots, n$.*
2. *G is connected iff $\lambda_2 < d$, i.e., the eigenvalue d has multiplicity 1. Moreover, the number of connected components of G equals the multiplicity of eigenvalue d.*
3. *If G is connected, then G is bipartite iff $\lambda_n = -d$.*

For studying regular graphs, it is convenient to work with *the normalized adjacency matrix* $\mathbf{M}$ of graph G. For any d-regular graph with adjacency matrix $\mathbf{A}$, define

$$\mathbf{M} \triangleq \frac{1}{d} \cdot \mathbf{A}.$$

We use $\lambda_1 \ge \cdots \ge \lambda_n$ to denote the eigenvalues of matrix $\mathbf{M}$ of graph G. For regular graphs, $\lambda_1 = 1$ and we mainly consider the second largest eigenvalue in absolute value. The formal definition is as follows.

Definition 6 (spectral expansion)**.** *The spectral expansion of graph G is defined by $\lambda \triangleq \max\{|\lambda_2|, |\lambda_n|\}$, i.e.*

$$\lambda = \max_{\|\mathbf{x}\|=1, \mathbf{x}\perp\mathbf{u}} \|\mathbf{A}\mathbf{x}\|,$$

where $\mathbf{u} = (1/\sqrt{n}, \ldots, 1/\sqrt{n}) \in \mathbb{R}^n$.

> **Courant-Fischer Formula**. Let $\mathbf{B}$ be an n by n symmetric matrix with eigenvalues $\lambda_1 \le \cdots \le \lambda_n$ and corresponding eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$. Then
>
> $$\lambda_1 = \min_{\|\mathbf{x}\|=1} \mathbf{x}^{\mathrm{T}}\mathbf{B}\mathbf{x} = \min_{\mathbf{x}\neq\mathbf{0}} \frac{\mathbf{x}^{\mathrm{T}}\mathbf{B}\mathbf{x}}{\mathbf{x}^{\mathrm{T}}\cdot\mathbf{x}},$$
>
> $$\lambda_2 = \min_{\substack{\|\mathbf{x}\|=1 \\ \mathbf{x}\perp v_1}} \mathbf{x}^{\mathrm{T}}\mathbf{B}\mathbf{x} = \min_{\substack{\mathbf{x}\neq\mathbf{0} \\ \mathbf{x}\perp v_1}} \frac{\mathbf{x}^{\mathrm{T}}\mathbf{B}\mathbf{x}}{\mathbf{x}^{\mathrm{T}}\cdot\mathbf{x}},$$
>
> $$\lambda_n = \max_{\|\mathbf{x}\|=1} \mathbf{x}^{\mathrm{T}}\mathbf{B}\mathbf{x} = \max_{\mathbf{x}\neq\mathbf{0}} \frac{\mathbf{x}^{\mathrm{T}}\mathbf{B}\mathbf{x}}{\mathbf{x}^{\mathrm{T}}\mathbf{x}}.$$

It is well known that λ relates to various graph properties. In particular, we will see that there is a close connection between λ and the expansion of the graph.

Lemma 7.

$$\lambda \geq \sqrt{\frac{n-d}{d(n-1)}} \quad .$$

Proof. Follow from $\mathrm{tr}\,(\mathbf{M}^2) = n/d = \sum_{i=1}^{n} \lambda_i^2 \leq 1 + (n-1)\lambda^2$. □

Theorem 8. *[2] Any infinite family of d-regular graphs $\{G_n\}_{n\in\mathbb{N}}$ has spectral expansion (as $n \to \infty$) at least $2\sqrt{d-1}/d - o(1)$.*

Definition 9 (Ramanujan graphs)**.** *A family of d-regular graphs with spectral expansion at most $2\sqrt{d-1}/d$ is called Ramanujan graphs.*

Although Friedman [6] showed that random d-regular graphs are close to being Ramanujan in the sense that λ satisfies

$$\lambda \leq 2\sqrt{d-1}/d + 2\log(d)/d + o(1),$$

constructing families of Ramanujan graphs with arbitrary degrees is one of the biggest open problems in this area. So far, we only know the construction of Ramanujan graphs with certain degrees and these constructions are based on deep algebraic knowledge. See [8] for example. Another quite important problem is to find a combinatorial construction of Ramanujan graphs.

We list some more interesting facts on eigenvalues of graphs:

1. If graphs G and H are isomorphic, then there is a permutation matrix P such that $\mathbf{P} \cdot \mathbf{A}(G) \cdot \mathbf{P}^{\mathrm{T}} = \mathbf{A}(H)$ and hence the matrices $\mathbf{A}(G)$ and $\mathbf{A}(H)$ are similar.
2. There are nonisomorphic graphs with the same spectrum. See Figure 1.

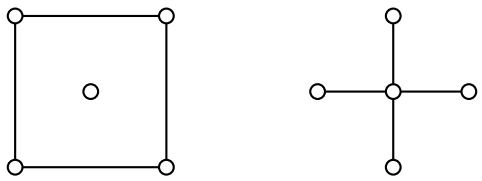

Figure 1: An example for two graphs which are not isomorphic but have the same spectrum. Their common graph spectrum is $2, 0, 0, 0, -2$.

2 Combinatorial Expansion of Graphs

For any d-regular graph $G = (V, E)$, let $\Gamma(v)$ be the set of neighbors of v, i.e.,

$$\Gamma(v) \triangleq \{u \mid (u, v) \in E\}.$$

For any subset $S \subseteq V$, let $\Gamma(S) \triangleq \cup_{v\in S}\Gamma(v)$ and $\Gamma'(S) \triangleq \Gamma(S) \cup S$. Moreoever, for any set $S \subseteq V$ we define $\partial S \triangleq E(S, \overline{S})$.

Definition 10 (vertex expansion)**.** *A graph G with n vertices is said to have vertex expansion (K, A) if*

$$\min_{S:\,|S|\leq K} \frac{|\Gamma(S)|}{|S|} \geq A.$$

If $K = n/2$, then for simplicity we call G an A-expander.

Definition 11 (edge expansion)**.** *The edge expansion of a graph $G = (V, E)$ is defined by*

$$h(G) \triangleq \min_{S:\, |S| \le |V|/2} \frac{|\partial S|}{|S|}.$$

To explain edge expansion, let us see two examples. (1) If G is not connected, we choose one connected component as S so that $|E(S, \overline{S})| = 0$. Therefore $h(G) = 0$. (2) If G is a complete graph K_n, then $|E(S, \overline{S})| = |S| \cdot (n - |S|)$ and $h(G) = \lceil n/2 \rceil$.

Definition 12 (expanders)**.** *Let $d \in \mathbb{N}$. A sequence of d-regular graphs $\{G_i\}_{i\in\mathbb{N}}$ of size increasing with i is a family of expanders if there is a constant $\varepsilon > 0$ such that $h(G_i) \ge \varepsilon$ for all i.*

Usually, when speaking of an expander G_i, we actually mean a family of graphs $\{G_i\}_{i\in\mathbb{N}}$, where each graph in $\{G_i\}_{i\in\mathbb{N}}$ is d-regular and its expansion is lower bounded by a constant $\varepsilon > 0$.

Observation 13. *Any expander graph is a connected graph.*

3 Spectral Expansion vs. Combinatorial Expansion

The next result shows that small spectral expansion implies large vertex expansion.

Theorem 14 (spectral expansion $\Rightarrow$ vertex expansion)**.** *If G has spectral expansion λ, then for all $0 < \alpha < 1$, G has vertex expansion $\left(\alpha n, \frac{1}{(1-\alpha)\lambda^2+\alpha}\right)$.*

Before showing the proof, we introduce some notations at first. For any probability distribution π, the support of π is defined by $\mathrm{support}(\boldsymbol{\pi}) = \{x : \boldsymbol{\pi}_x > 0\}$.

Definition 15. *Given a probability distribution $\boldsymbol{\pi}$, the collision probability of $\boldsymbol{\pi}$ is defined to be the probability that two independent samples from $\boldsymbol{\pi}$ are equal, i.e. $\mathrm{CP}(\boldsymbol{\pi}) = \sum_x \boldsymbol{\pi}_x^2$.*

Lemma 16. *Let $\mathbf{u} = (1/n, \ldots, 1/n)$ be the uniform distribution. Then for every probability distribution $\boldsymbol{\pi} \in [0, 1]^n$, we have*

1. $\mathrm{CP}(\boldsymbol{\pi}) = ||\boldsymbol{\pi}||^2 = ||\boldsymbol{\pi} - \mathbf{u}||^2 + 1/n$.
2. $\mathrm{CP}(\boldsymbol{\pi}) \ge 1/|\mathrm{support}(\boldsymbol{\pi})|$.

Proof. (1) We write $\boldsymbol{\pi}$ as $\boldsymbol{\pi} = \mathbf{u} + (\boldsymbol{\pi} - \mathbf{u})$ where $\mathbf{u} \perp (\boldsymbol{\pi} - \mathbf{u})$. By Pythagorean theorem

$$\mathrm{CP}(\boldsymbol{\pi}) = ||\boldsymbol{\pi}||^2 = ||\boldsymbol{\pi} - \mathbf{u}||^2 + ||\mathbf{u}||^2 = ||\boldsymbol{\pi} - \mathbf{u}||^2 + 1/n.$$

(2) By Cauchy-Schwarz inequality, we get

$$1 = \left(\sum_{x \in \mathrm{support}(\boldsymbol{\pi})} \boldsymbol{\pi}_x \right)^2 \le |\mathrm{support}(\boldsymbol{\pi})| \cdot \sum_x \boldsymbol{\pi}_x^2$$

and hence

$$\mathrm{CP}(\boldsymbol{\pi}) = \sum_x \boldsymbol{\pi}_x^2 \ge \frac{1}{|\mathrm{support}(\boldsymbol{\pi})|}.$$

□

Let us turn to the proof of Theorem 14.

Proof. Let $|S| \leq \alpha n$. Choose a probability distribution π that is uniform on S and 0 on the $\overline{S}$, i.e.

$$\boldsymbol{\pi} = \left(\frac{1}{|S|}, \frac{1}{|S|}, \ldots, \frac{1}{|S|}, 0, \ldots, 0 \right).$$

Note that $\mathbf{M}$ is a real symmetric matrix, then $\mathbf{M}$ has n orthonormal eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$. Hence we can decompose $\boldsymbol{\pi}$ as $\sum_{i=1}^{n} \boldsymbol{\pi}_i$ where $\boldsymbol{\pi}_i$ is a constant multiplicity of $\mathbf{v}_i$. Then $\mathrm{CP}(\boldsymbol{\pi}) = 1/|S|$ and by Lemma 16 (2) it holds that

$$\mathrm{CP}(\mathbf{M}\boldsymbol{\pi}) \geq \frac{1}{|\mathrm{support}(\mathbf{M}\boldsymbol{\pi})|} = \frac{1}{|\Gamma(S)|}.$$

On the other hand, by item (1) of Lemma 16 we have

$$\begin{aligned}
\mathrm{CP}(\mathbf{M}\boldsymbol{\pi}) - \frac{1}{n} &= \|\mathbf{M}\boldsymbol{\pi} - \mathbf{u}\|^2 \\
&= \|\mathbf{M}\mathbf{u} + \mathbf{M}\boldsymbol{\pi}_2 + \cdots + \mathbf{M}\boldsymbol{\pi}_n - \mathbf{u}\|^2 \\
&= \|\lambda_2 \boldsymbol{\pi}_2 + \cdots + \lambda_n \boldsymbol{\pi}_n\|^2 \\
&\leq \lambda^2 \|\boldsymbol{\pi} - \mathbf{u}\|^2 = \lambda^2 \left(\mathrm{CP}(\boldsymbol{\pi}) - \frac{1}{n} \right) = \lambda^2 \left(\frac{1}{|S|} - \frac{1}{n} \right).
\end{aligned}$$

Hence

$$\frac{1}{|\Gamma(S)|} - \frac{1}{n} \leq \mathrm{CP}(\mathbf{M}\pi) - \frac{1}{n} \leq \lambda^2 \left(\frac{1}{|S|} - \frac{1}{n} \right),$$

and

$$\begin{aligned}
|\Gamma(S)| &\geq \frac{1}{\lambda^2 \left(\frac{1}{|S|} - \frac{1}{n} \right) + \frac{1}{n}} = \frac{|S|}{\lambda^2 \left(1 - \frac{|S|}{n} \right) + \frac{|S|}{n}} = \frac{|S|}{\lambda^2 + (1 - \lambda^2) \cdot |S|/n} \\
&\geq \frac{|S|}{\lambda^2 + (1 - \lambda^2)\alpha} = \frac{|S|}{\alpha + (1 - \alpha)\lambda^2}.
\end{aligned}$$

□

Theorem 17 (vertex expansion ⇒ spectral expansion)**.** *Let G be a d-regular graph. For every $\delta > 0$ and $d > 0$, there exists $\gamma > 0$ such that if G is a d-regular $(1+\delta)$-expander according to Definition 10, then G has spectral expansion $(1 - \gamma)$. Specifically, we can take $\gamma = \Omega(\delta^2/d)$.*

When talking about expanders, we often mean a family of d-regular graphs satisfying one of the following two equivalent properties:

- Every graph in the family has spectral expansion λ.
- Every graph in the family is a $(1+\delta)$-expander for some constant δ.

4 Existence and Constructibility of Expander Graphs

Expander graphs have two seemingly contradictory properties: low degree and high connectivity. Two general problems are *existence* and *constructibility* of expander graphs. Among these two problems, existential proofs of expanders are easier, as one can resort to probabilistic techniques. Further, the existence of expanders can be often used as a black-box to show the existence of other interesting combinatorial objects. On the other hand, many applications of expanders need explicit constructions. We will mention some explicit constructions in this lecture, but they do not always match the bounds from the existential proofs.

Let $\mathcal{G}_{d,N}$ be the set of bipartite graphs with bipartite sets L, R of size N and left degree d. The following lemma shows the existence of expanders.

Theorem 18. *For any d, there exists an $\alpha(d) > 0$, such that for all N*

$$\mathbf{Pr}\left[\, G \text{ is an } (\alpha N, d-2)\text{-expander } \right] \geq 1/2,$$

where G is chosen uniformly from $\mathcal{G}_{d,N}$.

Proof. Define

$$p_k \triangleq \mathbf{Pr}\left[\, \exists S \subseteq L : |S| = k, |\Gamma(S)| < (d-2)|S| \,\right].$$

So G is not an $(\alpha N, d-2)$-expander iff $\sum_k p_k > 0$.

Assume that there is a set S of size K and $|\Gamma(S)| < (d-2)|S|$. Then there are at least $2k$ repeats among all the neighbors of vertices in S. We calculate the probability

$$\mathbf{Pr}\left[\, \text{at least } 2k \text{ repeats among all the neighbors of vertices in } S \,\right] \leq \binom{dk}{2k}\left(\frac{dk}{N}\right)^{2k}.$$

Therefore

$$\begin{aligned} p_k &\leq \binom{N}{k}\binom{dk}{2k}\left(\frac{dk}{N}\right)^{2k} \\ &\leq \left(\frac{N\varepsilon}{k}\right)^{k} \cdot \left(\frac{dk\varepsilon}{2k}\right)^{2k} \cdot \left(\frac{dk}{N}\right)^{2k} \\ &= \left(\frac{cd^4k}{N}\right)^{k} \end{aligned}$$

where $c = \varepsilon^3$. By setting $\alpha = 1/(cd^4)$ and $k \leq \alpha N$, we know that $p_k \leq 4^{-k}$ and

$$\mathbf{Pr}\left[\, G \text{ is not an } (\alpha N, d-2)\text{-expander} \,\right] \leq \sum_{k=1}^{\alpha N} p_k \leq \sum_{k=1}^{\alpha N} 4^{-k} \leq 1/2.$$

□

Let us now turn to the constructibility of expanders.

Definition 19. *Let $\{G_i\}_{i\in\mathbb{N}}$ be a family of expander graphs where G_i is a d-regular graph on n_i vertices and the integers $\{n_i\}$ are increasing, but not too fast.(e.g. $n_{i+1} \leq n_i^2$ will do)*

1. *The family is called* ***Mildly Explicit*** *if there is an algorithm that generates the j-th graph in the family $\{G_i\}_{i\in\mathbb{N}}$ in time polynomial in j.*

2. *The family is called* ***Very Explicit*** *if there is an algorithm that on input of an integer i, a vertex $v \in V(G_i)$ and $k \in \{1, \cdots, d\}$ computes the kth neighbor of the vertex v in the graph G_i. The algorithm's running time should be polynomial in its input length.*

Example. *A family of 3-regular p vertex graph for a prime number p. Let $G = (\mathbb{Z}_p, E)$. For any vertex $x \in \mathbb{Z}_p$, vertex x is connected to $x+1, x-1$ and x^{-1}. (The inverse of 0 is defined to be 0.)*

The family of expanders above is just mildly explicit, since at present we are unable to generate large prime numbers deterministically.

Example (Margulis, 1973)**.** *Fix a positive integer M and let $[M] = \{1, 2, \cdots, M\}$. Define the bipartite graph $G = (V, E)$ as follows. Let $V = [M]^2 \cup [M]^2$, where vertices in the first partite set as denoted $(x, y)_1$ and vertices in the second partite set are denoted $(x, y)_2$. From each vertex $(x, y)_1$, put in edges*

$$(x, y)_2, (x, x + y)_2, (x, x + y + 1)_2, (x + y, y)_2, (x + y + 1, y)_2,$$

where all arithmetic is done modulo M. Then G is an expander.

Example (Jimbo and Maruoka, 1987)**.** *Let $G = (L \cup R, E)$ be the graph described above, then $\forall X \subset L$, $|\Gamma(X)| \geq |X|(1 + d_0|\overline{X}|/n)$, where $d_0 = (2 - \sqrt{3})/4$ is the optimal constant.*

5 Expander Mixing Lemma

Consider two experiments on a d-regular graph G. (1) Pick a random vertex $u \in V$ and then pick one of its neighbors v. (2) Pick two random vertices $u, v \in V$ randomly and independently from $V \times V$. What is the probability of the event $(u, v) \in S \times T$, $S, T \subseteq V$ for these two experiments? For the first experiment, the probability is $|E(S, T)|/(nd)$. The probability for the second probability is $\mu(S) \cdot \mu(T)$, where $\mu(S) \triangleq |S|/n$ is the density of set S.

For the random bits used in these two experiments, it is easy to show that the first experiment uses $\log n + \log d$ random bits and the second one uses $2 \log n$ random bits. However, we will show that for graphs with good expansion these two probabilities are quite close to each other.

Lemma 20 (Expander Mixing Lemma)**.** *[3] Let $G = (V, E)$ be a d-regular n-vertex graph with spectral expansion λ. Then for any subset $S, T \subseteq V$, we have*

$$\left| |E(S, T)| - \frac{d|S| \cdot |T|}{n} \right| \leq \lambda d \sqrt{|S||T|}.$$

Let us look at the two terms in the left side: the size of $E(S, T)$ is the number of edges between two sets, and $d|S| \cdot |T|/n$ is the expected number of edges between S and T in a random graph with edge density d/n. So small λ implies that G is "more" random.

Proof. Let $\mathbf{1}_S, \mathbf{1}_T$ be the characteristic vectors of S and T. Expand these vectors in the orthonormal basis of eigenvectors $\mathbf{v}_1, \cdots, \mathbf{v}_n$, i. e. $\mathbf{1}_S = \sum_i \alpha_i \mathbf{v}_i$, and $\mathbf{1}_T = \sum_i \beta_i \mathbf{v}_i$. Then

$$|E(S,T)| = \mathbf{1}_S \cdot A \cdot \mathbf{1}_T = \left(\sum_i \alpha_i \mathbf{v}_i\right) A \left(\sum_i \beta_i \mathbf{v}_i\right) = \sum_i \lambda_i \alpha_i \beta_i,$$

where λ_is are eigenvalues of A. Since $\alpha_1 = \langle \mathbf{1}_S, \frac{1}{\sqrt{n}} \rangle = \frac{|S|}{\sqrt{n}}, \beta_1 = \frac{|T|}{\sqrt{n}}$ and $\lambda_1 = d$, then

$$|E(S,T)| = d \cdot \frac{|S| \cdot |T|}{n} + \sum_{i=2}^{n} \lambda_i \alpha_i \beta_i.$$

Thus

$$\left| |E(S,T)| - \frac{d|S||T|}{n} \right| \leq \sum_{i=2}^{n} \lambda_i \alpha_i \beta_i \leq \lambda \cdot d \cdot \sum_{i=2}^{n} |\alpha_i \beta_i|$$

By Cauchy-Schwartz inequality, we have

$$\left| |E(S,T)| - \frac{d|S||T|}{n} \right| \leq \lambda ||\mathbf{1}_S|| \cdot ||\mathbf{1}_T|| = \lambda \cdot d \cdot \sqrt{|S| \cdot |T|}.$$

□

Some remarks about the expander mixing lemma.

Lemma 21 (Converse of the Expander Mixing Lemma)**.** *[5] Let G be a d-regular graph and suppose that*

$$\left| |E(S,T)| - \frac{d|S||T|}{n} \right| \leq \theta d \sqrt{|S||T|}$$

holds for every two disjoint sets S, T and for some positive θ. Then $\lambda = \mathcal{O}(\theta(1 + \log(d/\theta)))$.

In the following, we use a three-tuple (n, d, λ) to represent an n-vertex d-regular graph with spectral expansion λ.

Corollary 22. *The size of the independent set for any (n, d, λ)-graph is at most λn.*

Proof. Let $T = S$. By Expander Mixing Lemma, we get $|S| \leq \lambda n$. □

Corollary 23. *For any (n, d, λ)-graph G, the chromatic number $\chi(G) \geq 1/\lambda$.*

Proof. Let $c : V \to \{1, \ldots, k\}$ be a coloring of G. Then for every $1 \leq i \leq k$, $c^{-1}(i)$ is an independent set. Since the size of every independent set is at most λn, so the chromatic number is at least $1/\lambda$. □

6 Cheeger's Inequality

The expander Mixing Lemma states that on graphs with good expansion, the graph's edges are *well distributed* and the spectral expansion of graphs are closely related to behavior of *any cut* in a graph. Cheeger's inequality provides another result of that flavour.

Theorem 24 (Cheeger's Inequality)**.** *Let G be a d-regular graph and let the eigenvalues of $\mathbf{A}(G)$ be $\lambda_1 \geq \ldots \geq \lambda_n$. Then*

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

7 Random Walks on Graphs

We now focus on random walks on graphs. We assume $G = (V, E)$ is an undirected, unweighted and connected graph which is also d-regular. Recall that $\mathbf{M} = \frac{1}{d} \cdot \mathbf{A}$ is the normalized adjacency matrix which will be transition matrix of the random walk on G.

Lemma 25. *Let* $\mathbf{M}$ *be any symmetric transition matrix. Then for any probability vector* $\mathbf{x}$*:*

$$\left\|\mathbf{M}^t\mathbf{x} - \boldsymbol{\pi}\right\|_2 \leq \lambda^t,$$

where $\boldsymbol{\pi} = (1/n, \ldots, 1/n)$ *is the uniform vector and* $\lambda = \max\{|\lambda_2|, |\lambda_n|\}$*. In particular, for* $t = \mathcal{O}(\log n/\log(1/\lambda)) = \mathcal{O}(\log n/(1-\lambda))$*, it holds that*

$$\left\|\mathbf{M}^t\mathbf{x} - \boldsymbol{\pi}\right\|_\infty \leq 1/(2n),$$

i.e., for all $u, v \in V$*,* $\mathbf{M}^t_{u,v} \in [\frac{1}{2} \cdot \frac{1}{n}, \frac{3}{2} \cdot \frac{1}{n}]$*.*

Proof. Since $\mathbf{M}$ is a real, symmetric matrix, $\mathbf{M}$ has n orthogonal eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$. Let $\mathbf{x}$ be any starting probability distribution on the vertices of G. Then we can decompose $\mathbf{x}$ uniquely as $\sum_{i=1}^n$ where $\mathbf{x}_i = \alpha_i \mathbf{v}_i$. Since $\mathbf{x}$ is a probability vector and all $\mathbf{v}_i, i \geq 2$ are orthogonal to $\boldsymbol{\pi} = (1/n, \ldots, 1/n)$, it follows that $\alpha_i = 1$. Then,

$$\begin{aligned}
\|\mathbf{M}\mathbf{x} - \boldsymbol{\pi}\|^2 &= \left\|\mathbf{M}\left(\sum_{i=1}^n \alpha_i \mathbf{v}_i\right) - \pi\right\|^2 \\
&= \left\|\pi + \sum_{i=2}^n \alpha_i \lambda_i \mathbf{v}_i - \pi\right\|^2 = \left\|\sum_{i=2}^n \alpha_i \lambda_i \mathbf{v}_i\right\|^2 \\
&= \sum_{i=2}^n \|\alpha_i \lambda_i \mathbf{v}_i\|^2 \\
&\leq \lambda^2 \sum_{i=2}^n \|\alpha_i \mathbf{v}_i\|^2 \\
&= \lambda^2 \left\|\sum_{i=2}^n \alpha_i \mathbf{v}_i\right\|^2 \\
&= \lambda^2 \|\mathbf{x} - \boldsymbol{\pi}\|^2 .
\end{aligned}$$

Taking square roots yields

$$\|\mathbf{M}\mathbf{x} - \boldsymbol{\pi}\| \leq \lambda \|\mathbf{x} - \boldsymbol{\pi}\| ,$$

so that for any $t \geq 1$,

$$\left\|\mathbf{M}^t\mathbf{x} - \boldsymbol{\pi}\right\| = \left\|\mathbf{M}(\mathbf{M}^{t-1}\mathbf{x}) - \boldsymbol{\pi}\right\| \leq \lambda \left\|\mathbf{M}^{t-1}\mathbf{x} - \boldsymbol{\pi}\right\| \leq \cdots \leq \lambda^t \|\mathbf{x} - \boldsymbol{\pi}\| .$$

Finally,

$$\begin{aligned}
\left\|\mathbf{M}^t\mathbf{x} - \boldsymbol{\pi}\right\| &\leq \lambda^t \cdot \|\mathbf{x} - \boldsymbol{\pi}\| \\
&\leq \lambda^t \cdot \|\mathbf{x}\| \\
&\leq \lambda^t \cdot \|\mathbf{x}\|_1 = \lambda^t,
\end{aligned}$$

where in the first inequality we have used that

$$\|\mathbf{x} - \boldsymbol{\pi}\|_2^2 + \|\boldsymbol{\pi}\|_2^2 = \|\mathbf{x}\|_2^2 ,$$

since $\mathbf{x} - \boldsymbol{\pi}$ and $\boldsymbol{\pi}$ are orthogonal, which immediately implies $\|\mathbf{x} - \boldsymbol{\pi}\| \leq \|\mathbf{x}\|$.

Hence $\|\mathbf{M}^t\boldsymbol{\pi} - \mathbf{u}\|_\infty < \frac{1}{2n}$, when $t = O\left(\frac{\log n}{\log(1/\lambda)}\right) = O\left(\frac{\log n}{1-\lambda}\right)$. To see the last step, we note that

$$\log(1+x) = 1 - \frac{1}{1+x} + O\left(\frac{1}{(1+x)^2}\right)$$

by taking the Taylor expansion of both sides. □

8 Probability Amplification by Random Walks on Expanders

Suppose we run a randomized sampling algorithm for which there is an unknown (bad) set $B \subseteq V$ for which the algorithm cannot solve the problem. For instance, B could describe all possible choices for the (random) bits the algorithm could use. Then after t repetitions of the algorithm we find the correct answer if and only if at least one sample lies outside B. While the obvious way to amplify the success probability is to generate t independent samples, there is a more clever and somewhat surprising solution. One performs a t-step random walk with a random starting vertex on an expander graph with vertex set V. Despite the large dependencies among two consecutive vertices, it turns out the probability for the random walk to hit at least one vertex outside B is very close to the probability that we have when we sample all t vertices independently from V.

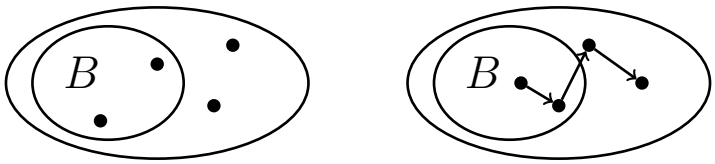

Figure 2: Illustration of Independent Samples and the corresponding random walk. B represents the set of "bad" inputs that we seek to avoid.

Theorem 26 (Ajtai-Komlos-Szemerdi (1987), Alon-Feige-Wigderson-Zuckerman (1995), [1, 4])**.** *Let G be a d-regular graph with n vertices and spectral expansion λ. Let $B \subseteq V$ with $|B| = \beta n$. Then,*

$$\mathbf{Pr}\left[\,\mathcal{B}\,\right] \triangleq \mathbf{Pr}\left[\,\forall t\colon X_t \in B\,\right] \leq (\beta + \lambda)^t .$$

Proof. Define a new matrix $\widetilde{\mathbf{P}}$ as follows:

$$\widetilde{\mathbf{P}}_{u,v} = \begin{cases} 1 & \text{if } u = v \in B \\ 0 & \text{otherwise.} \end{cases}$$

Lemma 27. *Let $\boldsymbol{\pi} = (1/n, \ldots, 1/n)$ be the uniform vector. Then,*

$$\mathbf{Pr}\left[\,\mathcal{B}\,\right] = \left\|(\widetilde{\mathbf{P}}\mathbf{M})^t\widetilde{\mathbf{P}}\boldsymbol{\pi}\right\|_1$$

Proof. Note that $\widetilde{\mathbf{P}}\boldsymbol{\pi}$ gives a vector which is $1/n$ at components corresponding to vertices in B and 0 otherwise. Moreover, $\mathbf{M}^t_{u,v}$ is the probability for a random walk on G starting from u to be located at v at step t. Since $\widetilde{\mathbf{P}}$ is a projection matrix, it follows that

$$(\widetilde{\mathbf{P}}\mathbf{M})^t_{u,v}$$

is the same probability as before, but now the random walk is additionally required to use only vertices in B. We now argue (slightly) more formally:

$$\begin{aligned}\left[(\widetilde{\mathbf{P}}\mathbf{M})^t\widetilde{\mathbf{P}}\boldsymbol{\pi}\right]_u &= \sum_{w\in V}\left[(\widetilde{\mathbf{P}}\mathbf{M})^t\right]_{u,w}\cdot\left[\widetilde{\mathbf{P}}\boldsymbol{\pi}\right]_w\\ &= \sum_{w\in V}\left[(\widetilde{\mathbf{P}}\mathbf{M})^t\right]_{u,w}\cdot\mathbf{1}_{w\in B}\cdot\frac{1}{n}\\ &= \sum_{w\in B}\frac{1}{n}\left[(\widetilde{\mathbf{P}}\mathbf{M})^t\right]_{w,u}\end{aligned}$$

Hence the claim of the lemma follows.

□

We continue with another lemma.

Lemma 28. *For any probability vector $\mathbf{v}\in\mathbb{R}^n$, it holds that*

$$\|\widetilde{\mathbf{P}}\mathbf{M}\widetilde{\mathbf{P}}\mathbf{v}\|_2 \le (\beta+\lambda)\cdot\|\mathbf{v}\|_2.$$

Proof. We first note that we can assume that $\widetilde{\mathbf{P}}\mathbf{v}=\mathbf{v}$. Otherwise, we replace v by $\widetilde{\mathbf{P}}v$ which doesn't change the left-hand-side, and can only make the right-hand-side smaller. For the same reason, we can also assume that $\mathbf{v}$ is non-negative and by scaling, we may also assume that $\sum_i \mathbf{v}_i = 1$.

Hence, we can express v as follows:

$$\widetilde{\mathbf{P}}\mathbf{v} = \mathbf{v} = \boldsymbol{\pi}+\mathbf{z},$$

where $\mathbf{z}=\mathbf{v}-\boldsymbol{\pi}$ is orthogonal to $\boldsymbol{\pi}$. With this, we obtain

$$\begin{aligned}\widetilde{\mathbf{P}}\mathbf{M}\widetilde{\mathbf{P}}\mathbf{v} &= \widetilde{\mathbf{P}}\mathbf{M}\mathbf{v}\\ &= \widetilde{\mathbf{P}}\mathbf{M}\boldsymbol{\pi}+\widetilde{\mathbf{P}}\mathbf{M}\mathbf{z}\\ &= \widetilde{\mathbf{P}}\boldsymbol{\pi}+\widetilde{\mathbf{P}}\mathbf{M}\mathbf{z},\end{aligned}$$

and thus

$$\left\|\widetilde{\mathbf{P}}\mathbf{M}\widetilde{\mathbf{P}}\mathbf{v}\right\|_2 \le \left\|\widetilde{\mathbf{P}}\boldsymbol{\pi}\right\|_2 + \left\|\widetilde{\mathbf{P}}\mathbf{M}\mathbf{z}\right\|_2. \tag{1}$$

We now bound the two summands on the right-hand side separately. By Cauchy-Schwartz inequality we have

$$1 = \sum_i \mathbf{v}_i = \sum_i \mathbf{1}_{i\in B}\cdot\mathbf{v}_i \le \sqrt{\beta n}\cdot\|\mathbf{v}\|_2.$$

Since $\|\widetilde{\mathbf{P}}u\|_2 = \sqrt{\beta/n}$, it follows that

$$\|\widetilde{\mathbf{P}}\mathbf{u}\|_2 \leq \beta \cdot \|\mathbf{v}\|_2.$$

For the other summand, recall that $\mathbf{z}$ is orthogonal to $\boldsymbol{\pi}$. Hence, $\mathbf{z} = \sum_{i=2}^{n} c_i \mathbf{v}_i$ where v_i is the ith eigenvector of $\mathbf{M}$. Moreover, since $\widetilde{\mathbf{P}}$ is a projection,

$$\left\|\widetilde{\mathbf{P}}\mathbf{M}\mathbf{z}\right\|_2 \leq \|\mathbf{M}\mathbf{z}\|_2 \leq \left\|\sum_{i=2}^{n} c_i \lambda_i \mathbf{v}_i\right\|_2 \leq \lambda \cdot \left\|\sum_{i=2}^{n} c_i \mathbf{v}_i\right\|_2 = \lambda \cdot \|\mathbf{z}\|_2 \leq \lambda \cdot \|\mathbf{v}\|_2,$$

where the last inequality holds since $\mathbf{v} = \boldsymbol{\pi} + \mathbf{z}$ and $\mathbf{z}$ is orthogonal to $\boldsymbol{\pi}$. Hence,

$$\|\mathbf{z}\|^2 + \|\boldsymbol{\pi}\|^2 = \|\mathbf{v} - \boldsymbol{\pi}\|^2 + \|\boldsymbol{\pi}\|^2 = \|\mathbf{v}\|^2$$

Plugging in the two inequalities in (1) yields the lemma. □

By combining the two lemmas, we obtain the theorem as follows.

$$\begin{aligned}\left\|(\widetilde{\mathbf{P}}\mathbf{M})^t\widetilde{\mathbf{P}}\boldsymbol{\pi}\right\|_1 &\leq \sqrt{n} \cdot \left\|(\widetilde{\mathbf{P}}\mathbf{M})^t\widetilde{\mathbf{P}}\boldsymbol{\pi}\right\|_2 \\ &= \sqrt{n} \cdot \left\|(\widetilde{\mathbf{P}}\mathbf{M}\widetilde{\mathbf{P}})^t\boldsymbol{\pi}\right\|_2 \\ &\leq \sqrt{n} \cdot (\beta + \lambda)^t \cdot \|\boldsymbol{\pi}\|_2 = (\beta + \lambda)^t .\end{aligned}$$

□

There exist various extensions of Theorem 26; for instance, where we only consider a subset of the time-steps in $\{1, \ldots, t\}$ or the set to be avoided changes over time. We refer to [7] for further details and the references therein.

Let us now apply Theorem 26 for a probabilistic algorithm A for the language $L \in \mathbf{RP}$ (the class of problems with one-sided error).

Complexity Class RP. The complexity class **RP** consists of all languages L for which there exists a polynomial-time randomized algorithm A such that

$$\begin{aligned}x \in L &\Longrightarrow \mathbf{Pr}\left[\, A(x) = 1 \,\right] \geq \frac{3}{4}, \\ x \notin L &\Longrightarrow \mathbf{Pr}\left[\, A(x) = 1 \,\right] = 0.\end{aligned}$$

To decide whether a given input x is in L, the algorithm A samples a random string $r \in \{0,1\}^\ell$ of length ℓ and computes in polynomial time a boolean function $A(x, r)$. If $x \notin L$, then $A(x, r) = 0$ for all r. If $x \in L$, then the probability (over r) that $A(x, r) = 0$ is at most β.

Now take a graph $G = (V, E)$ with $V = \{0,1\}^k$ which has spectral expansion at least λ and suppose that λ is sufficiently smaller than β which is the error of the given algorithm. Then the new algorithm $\widetilde{A}$ based on random walks is defined as follows:

(1) Pick a vertex $u_0 \in V$ uniformly at random.
(2) Perform a random walk of length t $(X_0, X_1, \ldots, X_t)$
(3) Return $\bigvee_{i=0}^{t} A(x, v_i)$.

Algorithm	Error Probability	Random Bits
Rand. Algorithm	$1/4$	k
t Repetitions	$(1/4)^t$	$t \cdot k$
t-step Random Walk	$(1/4)^t$	$k + \mathcal{O}(t) \cdot \log(d)$

Figure 3: Comparison of the methods for probability amplification. If k is a sufficiently large constant, then for any value of t, the t-step Random Walk algorithm requires less random bits for achieving the same error probability as the t-repetitions.

By definition of **RP**, we only need to consider the case where $x \in L$. By Theorem 26,

$$\mathbf{Pr}\left[\ \widetilde{A} \text{ fails } \right] \leq \mathbf{Pr}\left[\, \forall i \colon X_i \in B \,\right] \leq (\beta + \lambda)^t \,,$$

which is at most $2^{-\Omega(t)}$ assuming that λ is a constant smaller than $< 3/4$. Adjusting the constants, we conclude that the error probability is at most 4^{-t} if we use $k + \mathcal{O}(t) \cdot \log(d)$ random bits.

Complexity Class BPP. The complexity class **BPP** consists of all languages L for which there exists a polynomial-time randomized algorithm A such that

$$x \in L \Longrightarrow \mathbf{Pr}\left[\, A(x) = 1 \,\right] \geq \frac{3}{4},$$
$$x \notin L \Longrightarrow \mathbf{Pr}\left[\, A(x) = 0 \,\right] \geq \frac{3}{4}.$$

For this class, we can achieve a similar probability amplification by performing a random walk of length t and returning the majority vote. For more details, see [7].

9 Application: Superconcentrators

References

[1] Miklós Ajtai, János Komlós, and Endre Szemerédi. Deterministic simulation in logspace. In *Proceedings of the 19th Annual ACM Symposium on Theory of Computing*, pages 132–140, 1987.

[2] N. Alon. Eigenvalues and expanders. *Combinatorica*, 6(2):83–96, 1986.

[3] N. Alon and F. R. K. Chung. Explicit construction of linear sized tolerant networks. *Discrete Mathematics*, 72(1-3):15–19, 1988.

[4] Noga Alon, Uriel Feige, Avi Wigderson, and David Zuckerman. Derandomized graph products. *Computational Complexity*, 5(1):60–75, 1995.

[5] Y. Bilu and N. Linial. Lifts, discrepancy and nearly optimal spectral gap. *Combinatorica*, 26(5):495–519, 2006.

[6] Joel Friedman. On the second eigenvalue and random walks in random d-regular graphs. *Combinatorica*, 11:331–362, 1991.

[7] Shlomo Hoory, Nathan Linial, and Avi Widgerson. Expander graphs and their applications. *Bulletin (New series) of the American Mathematical Society*, 43(4):439?561, 2006.

[8] A. Lubotzky, R. Phillips, and P. Sarnak. Ramanujan graphs. *Combinatorica*, 8:261–277, 1988.