5

Expander Graphs in Computer Science

Great Ideas in Theoretical Computer Science
Saarland University, Summer 2014

The protagonists of today's lecture are expander graphs, a family of graphs that have found numerous applications in algorithm design, complexity theory, coding theory, and etc. Informally, expanders are graphs with low degree and high connectivity. We can use different ways to define expander graphs. Combinatorically, expanders are highly connected graphs, and to disconnect a large part of the graph, one has to sever many edges. Geometrically, every vertex set has a relatively large boundary. From the Probabilistic view, expanders are graphs whose behavior is "like" random graphs. Algebraically, expanders correspond to a family of real-symmetric matrix whose first positive eigenvalues of the Laplace operators are bounded away from zero. Due to these equivalent formulations using different languages, studies on expander graphs had an impact on many areas in mathematics.

5.1 Preliminaries

We first list a few results from linear algebra.

Definition 5.1 (Eigenvalues, and Eigenvectors)**.** *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 $\mathbf{Ax} = \lambda\mathbf{x}$. In this case, λ is called an eigenvalue of $\mathbf{A}$.*

Theorem 5.2 (Courant-Fischer Formula)**.** *Let $\mathbf{B}$ be an n by n symmetric matrix with eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n$ and corresponding eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$. Then*

$$\lambda_1 = \max_{\|\mathbf{x}\|=1} \mathbf{x}^{\mathrm{T}}\mathbf{B}\mathbf{x} = \max_{\mathbf{x}\neq\mathbf{0}} \frac{\mathbf{x}^{\mathrm{T}}\mathbf{B}\mathbf{x}}{\mathbf{x}^{\mathrm{T}}\mathbf{x}},$$

$$\lambda_2 = \max_{\substack{\|\mathbf{x}\|=1\\ \mathbf{x}\perp\mathbf{v}_1}} \mathbf{x}^{\mathrm{T}}\mathbf{B}\mathbf{x} = \max_{\substack{\mathbf{x}\neq\mathbf{0}\\ \mathbf{x}\perp\mathbf{v}_1}} \frac{\mathbf{x}^{\mathrm{T}}\mathbf{B}\mathbf{x}}{\mathbf{x}^{\mathrm{T}}\mathbf{x}},$$

$$\lambda_n = \min_{\|\mathbf{x}\|=1} \mathbf{x}^{\mathrm{T}}\mathbf{B}\mathbf{x} = \min_{\mathbf{x}\neq\mathbf{0}} \frac{\mathbf{x}^{\mathrm{T}}\mathbf{B}\mathbf{x}}{\mathbf{x}^{\mathrm{T}}\mathbf{x}}.$$

Lemma 5.3. *Let* $\|\mathbf{x}\|_p \triangleq (\sum_{i=1}^n |x_i|^p)^{1/p}$. *Then for any* $1 \leq p \leq q < \infty$, *it holds that*

$$\|\mathbf{x}\|_q \leq \|\mathbf{x}\|_p \leq n^{1/p-1/q} \cdot \|\mathbf{x}\|_q.$$

5.2 Spectra of Graphs

Spectral graph theory views graphs as matrices, and studies the eigenvalues of these matrices. These eigenvalues are called the spectra of graphs. These eigenvalues reveal basic properties of graphs (bipartiteness, connectivity, and etc.), and provide bounds of combinatorial quantities of graphs. Through spectral graph theory, we link the study of discrete universe to continuous ones which allow us to use geometric, analytic and algebraic techniques. More and more research has shown that eigenvalues play a central role in our fundamental understanding of graphs.

Definition 5.4 (adjacency matrix)**.** *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*

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

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

Adjacency matrices of graphs have 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 5.5. *The adjacency matrix of a triangle is*

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

Definition 5.6 (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.

Lemma 5.7. *Let* G *be any undirected and 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. $\deg_{\text{avg}} \leq \lambda_1 \leq \deg_{\text{max}}$.
5. $\sqrt{\deg_{\text{max}}} \leq \lambda_1 \leq \deg_{\text{max}}$.

Proof. We only prove the first three items.

(1) Since G has no self-loops, all the diagonal elements of $\mathbf{A}$ are zero. Hence $\sum_{i=1}^{n} \lambda_i = \operatorname{tr}(\mathbf{A}) = \sum_{i=1}^{n} \mathbf{A}_{i,i} = 0$.

(2) Note that $\sum_{i=1}^{n} \lambda_i^2 = \operatorname{tr}\left(\mathbf{A}^2\right) = \sum_{i=1}^{n} \mathbf{A}_{i,i}^2$. 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 every vertex i. Therefore $E[G] = \emptyset$. □

Example 5.8. *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 $2\cos(2\pi j/n)$ $(j = 0, 1, \ldots, n-1)$.*

Sometimes spectra of graphs give simpler proofs of combinatorial facts of graphs. We look at the following example.

Lemma 5.9. *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 $\lambda_1, \ldots, \lambda_n$. Then 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 $k \geq 3$, by Lemma 5.3 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 $\mathcal{O}\left(m^{k/2}\right)$. □

In this lecture we only study d-regular and *undirected* graphs. Note that if G is not undirected, then $\mathbf{A}$ is not symmetric anymore and the eigenvalues of $\mathbf{A}$ could be complex numbers. In addition, we only study d-regular graphs.

Lemma 5.10. *Consider any undirected graph G with adjacency matrix $\mathbf{A}$.*

1. *If G is d-regular, then $\lambda_1 = d$ and $|\lambda_i| \leq d$ for $i = 2, \ldots, n$.*
2. *Graph 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 $\lambda_n = -d$.*

Graph spectra is also closely related to graph isomorphism. It is easy to see that if graphs G and H are isomorphic, then there is a permutation matrix $\mathbf{P}$ 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. However, there are nonisomorphic graphs with the same spectrum. See Figure 5.1 for one example.

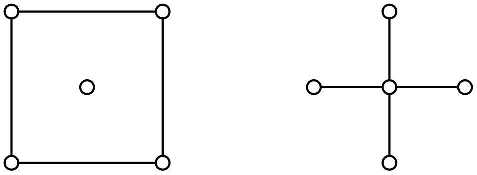

Figure 5.1: An example for two graphs which are not isomorphic but have the same spectrum. Their common graph spectrum is $2, 0, 0, 0, -2$.

5.3 Expansion of Graphs

Combinatorial Expansion. 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 5.11 (vertex expansion). *A graph G with n vertices is said to have vertex expansion (K, A) if*

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

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

Definition 5.12 (edge expansion). *The edge expansion of a d-regular graph $G = (V, E)$ is defined by*

$$h(G) \triangleq \min_{S \colon |S| \leq |V|/2} \frac{|\partial S|}{d \cdot |S|}.$$

To explain edge expansion, let us see two examples. (1) If G is not connected, then there is one connected component S with $|S| \leq |V|/2$. Since $|E(S, \overline{S})| = 0$, we have $h(G) = 0$. (2) If G is a complete graph K_n, then $|E(S, \overline{S})| = |S| \cdot (n - |S|)$ and $h(G) \approx 1/2$.

Spectral Expansioin. One can also define graph expansion by looking at the associated adjacency matrices. For any d-regular graph with adjacency matrix $\mathbf{A}$, define

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

to be the *normalized adjacency matrix* of graph G. We use $\lambda_1 \geq \cdots \geq \lambda_n$ to denote the eigenvalues of matrix $\mathbf{M}$ of graph G, and let $\lambda \triangleq \max \{|\lambda_2|, |\lambda_n|\}$. For regular graphs, $\lambda_1 = 1$ and we mainly consider the second largest eigenvalue in absolute value. The formal definition is as follows.

Definition 5.13 (spectral expansion). *The spectral expansion of graph G is defined by $\lambda \triangleq \max \{|\lambda_2|, |\lambda_n|\}$.*

Relations between Combinatorial and Spectral Expansion. Determining the value of $h(G)$ is co-NP-hard [6], however spectral expansion is easy to compute. So it is desired to approximate combinatorial expansion by using spectral expansion. Here we show several results that build connections between combinatorial and spectral expansion.

Theorem 5.14 (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. For any probability distribution π, the support of π is defined by $\mathrm{support}(\boldsymbol{\pi}) = \{x : \boldsymbol{\pi}_x > 0\}$.

Definition 5.15. *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 5.16. *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. $\mathrm{CP}(\boldsymbol{\pi}) \geq 1/|\mathrm{support}(\boldsymbol{\pi})|$.

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

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

(2) By Cauchy-Schwarz inequality, we have that

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

and hence $\mathrm{CP}(\boldsymbol{\pi}) = \sum_x \boldsymbol{\pi}_x^2 \geq 1/|\mathrm{support}(\boldsymbol{\pi})|$. □

Proof of Theorem 5.14. 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 matrix $\mathbf{M}$ is real and symmetric, then $\mathbf{M}$ has n orthonormal eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$. We 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 the second statement of Lemma 5.16 it holds that

$$\mathrm{CP}(\mathbf{M}\boldsymbol{\pi}) \geq \frac{1}{|\mathrm{support}(\mathbf{M}\boldsymbol{\pi})|} = \frac{1}{|\Gamma(S)|}. \tag{5.1}$$

On the other hand, by the first statement of Lemma 5.16 we have

$$\begin{aligned} \mathrm{CP}(\mathbf{M}\boldsymbol{\pi}) - \frac{1}{n} &= ||\mathbf{M}\boldsymbol{\pi} - \mathbf{u}||^2 \\ &= ||\mathbf{M}\mathbf{u} + \mathbf{M}\boldsymbol{\pi}_2 + \cdots + \mathbf{M}\boldsymbol{\pi}_n - \mathbf{u}||^2 \\ &= ||\lambda_2\boldsymbol{\pi}_2 + \cdots + \lambda_n\boldsymbol{\pi}_n||^2 \\ &\leq \lambda^2 ||\boldsymbol{\pi} - \mathbf{u}||^2 = \lambda^2 \left(\mathrm{CP}(\boldsymbol{\pi}) - \frac{1}{n} \right) = \lambda^2 \left(\frac{1}{|S|} - \frac{1}{n} \right). \end{aligned} \tag{5.2}$$

Combining (5.1) and (5.2) gives us

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

and

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

□

The following theorem shows that vertex expanders, i.e. graphs with high vertex expansion, are also good spectral expanders. We omit the proof due to space limitation.

Theorem 5.17 (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 5.11, then G has spectral expansion $(1-\gamma)$. Specifically, we can take $\gamma=\Omega(\delta^2/d)$.*

The following inequality relates edge expansion of graphs to spectral expansion.

Theorem 5.18 (Cheeger's Inequality)**.** *Let G be a d-regular graph and let the eigenvalues of $\mathbf{M}(G)$ be $\lambda_1\geq\ldots\geq\lambda_n$. Then*

$$\frac{1-\lambda_2}{2}\leq h(G)\leq\sqrt{2\cdot(1-\lambda_2)}.$$

Proof of the easy direction. By the Courant-Fischer Formula (Theorem 5.2), we have that

$$1-\lambda_2=\min_{\substack{\mathbf{x}\in\mathbb{R}^n\\ \mathbf{x}\neq\mathbf{0},\mathbf{x}\perp\mathbf{1}}}\frac{\mathbf{x}^{\mathrm{T}}\mathbf{x}-\mathbf{x}^{\mathrm{T}}\mathbf{M}\mathbf{x}}{\mathbf{x}^{\mathrm{T}}\mathbf{x}}=\frac{1}{d}\cdot\min_{\substack{\mathbf{x}\in\mathbb{R}^n\\ \mathbf{x}\neq\mathbf{0},\mathbf{x}\perp\mathbf{1}}}\frac{\sum_{u\sim v}(\mathbf{x}_u-\mathbf{x}_v)^2}{\sum_u\mathbf{x}_u^2},$$

where $u\sim v$ stands for $\{u,v\}\in E(G)$. Let $S\subseteq V$ with $|S|\leq|V|/2$ be the subset for which $h(G)$ is achieved. Let $\mathbf{y}\in\mathbb{R}^n$ such that $\mathbf{y}_u=1/|S|$ if $u\in S$, and $-1/|V\setminus S|$ if $u\in V\setminus S$. Since $\mathbf{y}\perp\mathbf{1}$, it holds that

$$\begin{aligned}1-\lambda_2 &\leq\frac{1}{d}\cdot\frac{\sum_{u\sim v}(\mathbf{y}_u-\mathbf{y}_v)^2}{\sum_u\mathbf{y}_u^2}=\frac{1}{d}\cdot\frac{|E(S,V\setminus S)|\cdot(1/|S|+1/|V\setminus S|)^2}{1/|S|+1/|V\setminus S|}\\ &\leq\frac{1}{d}\cdot|E(S,V\setminus S)|\cdot(1/|S|+1/|V\setminus S|)\\ &\leq\frac{1}{d}\cdot\frac{2\cdot|E(S,V\setminus S)|}{|S|}=2\cdot h(G).\end{aligned}$$

□

Ramanujan Graphs. All these results between vertex/edge expansion and spectral expansion show that smaller λ implies better combinatorial expansion. The theorem below gives a lower bound of the spectral expansion for any d-regular graph.

Theorem 5.19 ([2]). *Any infinite family of d-regular graphs $\{G_n\}_{n\in\mathbb{N}}$ has spectral expansion (as $n \to \infty$) at least $2\sqrt{d-1}/d - o(1)$.*

We call any d-regular graph with spectral expansion at most $2\sqrt{d-1}/d$ a *Ramanujan* graph. Friedman [7] proved that random d-regular graphs are close to being Ramanujan in the sense that λ satisfies

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

and it is widely believed that for any fixed d there are infinitely many d-regular Ramanujan graphs. However, constructing families of Ramanujan graphs with arbitrary degrees is one of the major open problems in spectral graph theory. So far, we only know constructions of Ramanujan graphs with certain degrees and these constructions are based on deep algebraic knowledge. See [10] for example. Another important problem is to find a combinatorial construction of Ramanujan graphs.

5.4 Expander Graphs

Definition 5.20 (expander graphs). *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 expander graphs if there is a constant $\varepsilon > 0$ such that $h(G_i) \geq \varepsilon$ for all i.*

Two general problems are *existence* and *constructibility* of expander graphs. Between these two problems, existential proofs of expanders are easier, and one can resort to probabilistic techniques. Moreover, 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.

Proof of Existence. The following lemma shows the existence of expanders.

Theorem 5.21. *Let $\mathcal{G}_{d,N}$ be the family of bipartite graphs with bipartite sets L, R of size N and left degree d. For any d, there exists an $\alpha(d) > 0$, such that for all N it holds that*

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

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

$$\begin{aligned} p_k &\leq \binom{N}{k}\binom{dk}{2k}\left(\frac{dk}{N}\right)^{2k} \\ &\leq \left(\frac{N\varepsilon}{k}\right)^{k}\cdot\left(\frac{dk\varepsilon}{2k}\right)^{2k}\cdot\left(\frac{dk}{N}\right)^{2k} \\ &= \left(\frac{cd^4k}{N}\right)^{k} \end{aligned} \tag{5.3}$$

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

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

Constructions of Expanders. While it is easy to prove the existence of expander graphs, explicit constructions are much harder. We say that a family of expander graphs $\{G_n\}_{n\in\mathbb{N}}$ is *explicit* if there is a polynomial-time algorithm that on input 1^n outputs the adjacency matrix of G_n. We say that the family is *strongly explicit* if there is a polynomial-time algorithm that on input n, u, i, outputs the index of the ith neighbor of vertex u. Note that in the strongly explicit case, the lengths of the algorithm's inputs and outputs are $O(\log n)$ and hence its runtime is $O(\text{poly}\log n)$.

Example 5.22. *Let p be a prime number. A 3-regular expander graph with p vertices can be generated as follows: 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}, where the inverse of 0 is defined to be 0.*

The proof that this 3-regular graph is an expander is based on deep results from Number Theory: Selberg 3/16 theorem. This graph is not strongly explicit, since there is no efficient method to generate large prime numbers deterministically.

Example 5.23 ([11])**.** *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 are denoted by $(x, y)_1$ and vertices in the second partite set are denoted by $(x, y)_2$. 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 graph.

Example 5.24 ([9])**.** *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.*

5.5 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 event, the probability is $|E(S,T)|/(nd)$. The probability for the second event 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, 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 5.25 (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*

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

We 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 smaller 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 $\mathbf{v}_1, \cdots, \mathbf{v}_n$ of A, i. e. $\mathbf{1}_S = \sum_i \alpha_i \mathbf{v}_i$, and $\mathbf{1}_T = \sum_i \beta_i \mathbf{v}_i$. Then we can write $|E(S,T)|$ as

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

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 $\lambda_1 = d$, we have

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

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 that

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

Lemma 5.26 (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 $\lambda = \mathcal{O}(\theta(1 + \log(d/\theta)))$.

Corollary 5.27. *Let G be a d-regular graph with n vertices and spectral expansion λ. Then the size of the independent set of G is at most λn.*

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

Corollary 5.28. *Let G be a d-regular graph with n vertices and spectral expansion λ. Then 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 $1/\lambda$. □

5.6 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 5.29. *Let $\mathbf{M}$ be any symmetric transition matrix. Then for any probability vector $\mathbf{x}$, it holds that*

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

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 and 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 $\sum_{i=1}^n \alpha_i \mathbf{v}_i$. 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

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

By induction for any $t \geq 1$ it holds that

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

Finally,

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

where in the first inequality we used the fact 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. □

5.7 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, V 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 an 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 that the probability for a 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. See Figure 5.2 for an illustration.

Theorem 5.30 ([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,*

$$\mathbf{Pr}\left[\,\mathcal{B}\,\right] \triangleq \mathbf{Pr}\left[\,\textit{all } t \textit{ steps of a random walk stays entirely in } B\,\right] \leq (\beta+\lambda)^t.$$

We define a matrix $\mathbf{P}$ by $\mathbf{P}_{u,v} = 1$ if $u = v \in B$, and $\mathbf{P}_{u,v} = 0$ otherwise.

Lemma 5.31. *Let $\boldsymbol{\pi} = (1/n, \ldots, 1/n)$ be the uniform vector. Then,*

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

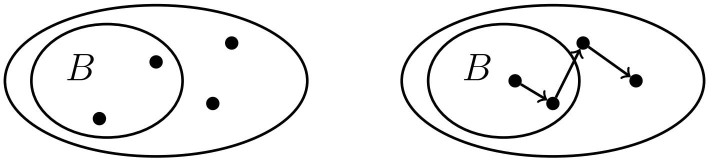

Figure 5.2: Illustration of independent samples and the corresponding random walk. Set B represents the set of "bad" inputs that we want to avoid.

Proof. Note that $\mathbf{P}\boldsymbol{\pi}$ 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 in G starting from u to be located at v at step t. Since $\mathbf{P}$ is a projection matrix, it follows that

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

Hence the claim of the lemma follows. □

> **Lemma 5.32.** *For any probability vector* $\mathbf{v}\in\mathbb{R}^n$*, it holds that*
>
> $$\|\mathbf{PMPv}\|_2 \le (\beta+\lambda)\cdot\|\mathbf{v}\|_2.$$

Proof. We first note that we can assume that $\mathbf{Pv}=\mathbf{v}$. Otherwise, we replace $\mathbf{v}$ by $\mathbf{Pv}$ which does not 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:

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

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

$$\mathbf{PMPv}=\mathbf{PMv}=\mathbf{PM}\boldsymbol{\pi}+\mathbf{PMz}=\mathbf{P}\boldsymbol{\pi}+\mathbf{PMz},$$

and thus

$$\|\mathbf{PMPv}\|_2 \le \|\mathbf{P}\boldsymbol{\pi}\|_2+\|\mathbf{PMz}\|_2\,. \tag{5.4}$$

We now bound these two terms 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 $\|\mathbf{P}u\|_2 = \sqrt{\beta/n}$, it follows that

$$\|\mathbf{Pu}\|_2 \leq \beta \cdot \|\mathbf{v}\|_2.$$

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 $\mathbf{v}_i$ is the ith eigenvector of $\mathbf{M}$. Since $\mathbf{P}$ is a projection, we have that

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

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 (5.4) yields the lemma. □

Proof of Theorem 5.30. By combining Lemma 5.31 and Lemma 5.32, we have that

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

□

There are various extensions of Theorem 5.30. For instance, we may only consider a subset of the time-steps in $\{1, \ldots, t\}$, or the case that the set to be avoided changes over time. We refer to [8] for further details and the references therein.

Let us now apply Theorem 5.30 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 $\mathbf{RP}$ 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 $X_0, X_1, \ldots, X_t$ of length t.

(3) Return $\bigvee_{i=0}^{t} A(x, v_i)$.

By definition of **RP**, we only need to consider the case where $x \in L$. By Theorem 5.30,

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

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.

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

Figure 5.3: 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.

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] Manuel Blum, Richard M. Karp, Oliver Vornberger, Christos H. Papadimitriou, and Mihalis Yannakakis. The complexity of testing whether a graph is a superconcentrator. *Inf. Process. Lett.*, 13(4/5):164–167, 1981.

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

[8] 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.

[9] Shuji Jimbo and Akira Maruoka. Expanders obtained from affine transformations. *Combinatorica*, 7(4):343–355, 1987.

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

[11] G. A. Margulis. Explicit constructions of expanders. *Problemy Peredači Informacii*, 9(4):71–80, 1973.