Lecture 17: Expanders and Random Walks

Leduc

October 22, 2025

1 Expander Graphs

Expander graphs are almost magical graphs that have many applications in mathematics and computer science.

Definition 1. A multigraph $G = (V, E)$ is an α-edge expander if

$$|E(S, \bar{S})| \geq \alpha |S|$$

for all $S \subseteq V$ with $|S| \leq \frac{|V|}{2}$.

The surprising fact is that a random 3-regular graph has expansion ≈ 0.18 with high probability.

Theorem 1 (Bollobás)**.** *Let $d \geq 3$ be an integer and $\epsilon \in (0, 1)$. If $n \geq c(\epsilon, d)n_0$ for some constant c depending on ϵ and d, then a random d-regular graph has expansion $\geq \frac{d-2}{d} - \epsilon$ with high probability.*

Corollary 1. *As $n \to \infty$, a random d-regular graph has expansion $\to \frac{d-2}{d}$ with high probability. In particular, for $d = 3$ we can obtain expansion ≈ 0.18.*

The proof is via the probabilistic method and not very difficult. Nevertheless, it is somewhat technical. Simple proofs give weaker expansion.

1.1 Explicit Constructions

Although random regular graphs are expanders and one can generate them relatively easily, it is hard to compute the expansion. Therefore, there have been several works that construct expanders explicitly. Many of these are based on group theoretic constructions.

One of the early explicit constructions is given by Margulis and analyzed by Gabber and Galil.

Construction: Fix integer m and let $n = 2m^2$. We construct a graph on n vertices. It is a bipartite graph with parts A and B with $|A| = |B| = m^2$.

$$A = \{(u, x, y) : x, y \in \mathbb{Z}_m\}$$
$$B = \{(v, x, y) : x, y \in \mathbb{Z}_m\}$$

For each vertex (u, x, y) in A, we add 5 edges to vertices:

$$(v, x, y), \quad (v, x, x + y), \quad (v, x, x + y + 1), \quad (v, x + y, y), \quad (v, x + y + 1, y)$$

Here addition is done mod m. It can be shown that expansion is $\Omega(1)$.

1.2 Conductance

A notion related to expansion is conductance.

Definition 2. The conductance ϕ of a graph is

$$\phi = \min_{S : \mathrm{Vol}(S) \leq \frac{\mathrm{Vol}(V)}{2}} \frac{|E(S, \bar{S})|}{\mathrm{Vol}(S)}$$

where $\mathrm{Vol}(S) = \sum_{u \in S} \deg(u)$.

Note that if G is d-regular then $\mathrm{Vol}(S) = d|S|$ and hence for those graphs $\phi = \frac{\alpha}{d}$.

2 Cheeger's Inequality

Recall that we did a spectral analysis of the convergence of a random walk in undirected graphs.

Let A be the adjacency matrix of G. Then the random walk matrix is

$$W = AD^{-1}$$

where D is the diagonal matrix with $D_{ii} = d_i$.

The lazy walk matrix is $\frac{1}{2}(I + AD^{-1})$.

If G is d-regular then $W = \frac{1}{d}A$ is symmetric. Otherwise, we considered the normalized adjacency matrix

$$\mathcal{A} = D^{-1/2}AD^{-1/2}$$

and noted that $W = D^{-1}A = D^{-1/2}\mathcal{A}D^{1/2}$, which implies that W is similar to the symmetric matrix $\mathcal{A}$. Note that $\mathcal{A} = \frac{1}{d}A$ if G is d-regular.

Let $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ be the real eigenvalues of $\mathcal{A}$. Then $1 = \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq -1$. If G is connected, we assume $\lambda_2 < 1$.

We saw that the random walk on G converges in $O(\beta^{-1} \log \epsilon^{-1})$ steps to a distribution that has $\leq \epsilon$ total variation distance from the stationary distribution, where

$$\beta = \min\{1 - \lambda_2, 1 + \lambda_n\}$$

is the spectral gap. For the lazy random walk, $\beta = \frac{1-\lambda_2}{2}$.

How do we know when β is not too small?

Cheeger's inequality allows us to bound conductance via another important matrix of graphs called the Laplacian:

$$L = D - A$$

The normalized Laplacian is

$$\mathcal{L} = I - \mathcal{A}$$

where $\mathcal{A}$ is the normalized adjacency matrix.

$\mathcal{L}$ and L are positive semi-definite matrices.

Observation 1. Let $0 = \gamma_1 \leq \gamma_2 \leq \cdots \leq \gamma_n$ be the eigenvalues of $\mathcal{L}$. Then $\gamma_i = 1 - \lambda_i$ where $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ are the eigenvalues of A_n. Thus $\beta = \gamma_2$.

Theorem 2 (Cheeger's Inequality)**.** *For any graph,*

$$\frac{\phi^2}{2} \leq \gamma_2 \leq 2\phi$$

Equivalently: $\phi \geq \sqrt{2\gamma_2}$ and $\gamma_2 \leq 2\phi$.

Corollary 2. *Suppose G is d-regular. Let $1 = \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ be eigenvalues of A/d. Then*

$$\frac{\phi^2}{2d} \leq 1 - \lambda_2 \leq \frac{2\phi}{d}$$

For d-regular graphs, note that $\phi = \frac{\alpha}{d}$ and also we have $\beta = 1 - \lambda_2 = \gamma_2$.

Thus if we have a constant degree expander, $\beta = \Omega(1)$ and the random walk mixes in $O(\log n)$ steps.

3 Randomized Complexity Classes

Recall **P** is the set of decision problems that have a deterministic poly-time algorithm.

3.1 RP (Randomized Polynomial Time)

$L \in \mathbf{RP}$ if there exists a poly-time randomized algorithm A such that for all inputs $x \in \Sigma^*$:

(i) If $x \notin L$, $A(x)$ says No

(ii) If $x \in L$, $A(x)$ says YES with probability $\geq \frac{1}{2}$

(One-sided error)

3.2 co-RP

co-RP is one-sided error for $x \notin L$. That is, $L \in \mathbf{co\text{-}RP}$ if $\bar{L} \in \mathbf{RP}$.

3.3 BPP (Bounded-Error Probabilistic Polynomial Time)

BPP is the set of languages that admit poly-time randomized algorithms that can make 2-sided errors.

$L \in \mathbf{BPP}$ if there exists a randomized polytime algorithm A such that for all $x \in \Sigma^*$:

(i) If $x \in L$, $A(x)$ outputs YES with probability $\geq \frac{2}{3}$

(ii) If $x \notin L$, $A(x)$ outputs YES with probability $\leq \frac{1}{3}$

RP and **BPP** algorithms are called Monte Carlo algorithms. They run in polytime but can make a mistake.

Clear that $\mathbf{P} \subseteq \mathbf{RP} \subseteq \mathbf{BPP}$.

3.4 ZPP (Zero-Error Probabilistic Polynomial Time)

ZPP is the set of problems for which there is a randomized algorithm A such that for all x:

(i) If $x \in L$, $A(x)$ returns YES

(ii) If $x \notin L$, $A(x)$ returns No

(iii) The expected run time of A on x is $p(|x|)$ for some fixed polynomial p

Claim 1. $\boldsymbol{ZPP = RP \cap co\text{-}RP}$

Proof. Exercise. □

(Las Vegas algorithms)

4 Error Reduction in Randomized Algorithms

Easy lemma based on repetition:

Lemma 1. *Suppose $L \in \boldsymbol{RP}$ and A is a randomized poly-time algorithm for L. Suppose on input x, A takes n random bits and is correct with probability $\geq \frac{1}{2}$. Then by running A k times independently, we can reduce error to 2^{-k}.*

Now suppose we have $L \in \mathbf{BPP}$. How do we reduce error?

We run A k times independently and take majority vote of the outputs.

Lemma 2. *Error is $2^{-\Omega(k)}$ for some fixed c.*

(Use Chernoff bounds)

Repeating k times independently requires kn random bits where n is the number of random bits for each run. Is this optimal? Can we do better?

Turns out that one can reduce error to 2^{-k} by using only $O(n + k)$ bits.

How? By random walks on expanders!

4.1 Setup

Let $N = 2^n$. We will assume that we can construct implicitly a constant degree expander on N vertices. Typically we will not be able to construct expanders for all N but we will not worry about that technicality for now.

Let $G = (V, E)$ be the expander with expansion α and degree d. We assume $d = O(1)$ and $\alpha = \Omega(1)$.

Each vertex $v \in V$ corresponds to a n-bit binary string.

We will assume that given v, we can find the d neighbors of v in poly(n) time. For example, in the Margulis (explicit Gabber-Galil) expander, we can do this. Thus we can implement a random walk on G for t steps in poly(t, n) time and the number of random bits required to implement t-step walk is $O(t)$ since each step requires only $O(1)$ bits to pick a random neighbor. We will assume the walk on G is ergodic; otherwise we can do the lazy random walk.

4.2 Algorithm for RP and BPP Amplification

For both **RP** and **BPP** amplification, we do the following:

1. Pick a uniformly random $v_1 \in V$
2. Do a random walk for t steps and let $v_1, v_2, \ldots, v_t$ be the vertices
3. Let $r_1, r_2, \ldots, r_t$ be the n-bit random strings associated with $v_1, v_2, \ldots, v_t$ respectively
4. Let $b_i = A(x, r_i)$ be the output of A on input x with random string r_i

Lemma 3. *Let $G = (V, E)$ be an undirected graph whose random walk matrix has spectral gap β. Consider a t-step random walk $v_1, v_2, \ldots, v_t \in V$ where $v_1 \in V$ is chosen uniformly at random. For any set $B \subseteq V$,*

$$\Pr[\forall i \in [t] : v_i \in B] \leq \mu^t + \beta^t$$

where $\mu = \frac{|B|}{|V|}$.

Assume lemma is true. Now consider the algorithm we had using random walks on expanders with each vertex being a n-bit random string.

4.3 RP Amplification

Suppose we have $L \in \mathbf{RP}$ and A on input x outputs No if $x \notin L$ and outputs YES with probability $\geq 1 - \mu$ for $x \in L$.

Let $b_1, \ldots, b_t$ be the outputs of $A(x, r_1), \ldots, A(x, r_t)$.

The algorithm outputs Yes if any of the outputs is Yes. Otherwise it outputs No.

Suppose $x \notin L$, then it is clear that A will output No.

Suppose $x \in L$. What is the probability it will output No?

$$\leq \mu^t + \beta^t$$

Expander gives us β is a fixed constant. By basic repetition we can ensure that $\mu \leq \frac{1}{4}$.

Hence $\mu^t + \beta^t \leq 2(\frac{1}{4})^t$.

If we choose $t \geq c \log(\frac{1}{\epsilon})$ for suitable c, then $(\frac{1}{4})^t \leq \frac{\epsilon}{2}$, we will have failure probability $\leq \epsilon$.

$t = O(k)$ suffices (the $O(\cdot)$ notation hides a dependence on β which we assume is a fixed constant).

5 Proof of the Lemma

Let W be the random walk matrix $W = AD^{-1}$ and since G is d-regular, W is symmetric.

Let $1 = \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq -1$ be the eigenvalues of W.

$B \subseteq V$ and $\mu = \frac{|B|}{|V|}$.

Let P be a $|V| \times |V|$ diagonal matrix with $P_{vv} = 1$ if $v \in B$, 0 otherwise.

For any vector $\bar{x} \in \mathbb{R}^{|V|}$, $(P\bar{x})_i = \begin{cases} x_i & \text{if } i \in B \\ 0 & \text{otherwise} \end{cases}$.

P is like the identity matrix restricted to B.

What is the probability that $v_1, v_2, \ldots, v_t \in B$?

We claim it is $\|\frac{1}{n}PWP \cdots P\mathbb{1}\|_1$ where $\mathbb{1} = (1, 1, \ldots, 1)^T$.

Here since v_1 is chosen uniformly at random, $p(v_1) = \frac{1}{n}\mathbb{1}$ where $n = |V|$.

PWP is also a symmetric matrix.

Recall W can be written as $W = \sum_{i=1}^{n} \lambda_i \bar{z}_i \bar{z}_i^T$ where $\bar{z}_1, \bar{z}_2, \ldots, \bar{z}_n$ are the orthonormal eigenvectors. $\lambda_1 = 1$ and $\bar{z}_1 = \frac{1}{\sqrt{n}}\mathbb{1}$ since we normalize to unit vectors.

Lemma 4. *For any vector $\bar{y}$,*

$$\|PWP\bar{y}\|_2 \leq \mu + \beta\|\bar{y}\|_2$$

Proof. We can assume that $y_i = 0$ if $i \notin B$ because it can only help the inequality. We can also assume $\bar{y} \geq 0$. We can then assume that $\sum_i y_i = 1$ by scaling $\bar{y}$ since it doesn't change the inequality.

Thus $\bar{y}$ can be written as

$$\bar{y} = \mu\mathbb{1} + \bar{\epsilon}$$

where $\mathbb{1} = \frac{1}{\sqrt{n}}(1, 1, \ldots, 1)^T$ is the uniform distribution and $\bar{\epsilon}$ is orthogonal to it.

$$\begin{aligned} PWP\bar{y} &= PW(\mu\mathbb{1} + \bar{\epsilon}) \\ &= PW\mu\mathbb{1} + PW\bar{\epsilon} \end{aligned}$$

because $\bar{y}$ is in support of B.

$$PW\mu\mathbb{1} = \mu P\mathbb{1}$$
$$\|PWP\bar{y}\|_2 \leq \|P\mu\mathbb{1}\|_2 + \|PW\bar{\epsilon}\|_2$$

First we show $\|P\mu\mathbb{1}\|_2 \leq \mu$:

$$\|P\mu\mathbb{1}\|_2 = \mu\|P\mathbb{1}\|_2 \leq \mu\sqrt{n} \cdot \frac{1}{\sqrt{n}} = \mu$$

since P has at most μn 1's on the diagonal.

By Cauchy-Schwarz and the fact that $\bar{y}$ has support at most μn:

$$\sum_i y_i = \mu n \cdot \frac{1}{\sqrt{\mu n}}\|\bar{y}\|_2 = \sqrt{\mu n}\|\bar{y}\|_2$$

$$\|P\mu\mathbb{1}\|_2 \leq \sqrt{\mu}\|\bar{y}\|_2$$

Now consider $\|PW\bar{\epsilon}\|_2 \leq \|W\bar{\epsilon}\|_2$ since P is a contraction in ℓ_2.

Since $\bar{\epsilon}$ is orthogonal to $\mathbb{1}$, the largest eigenvector of W, and $|\lambda_i| \leq \beta$ for $i \geq 2$:

$$\|W\bar{\epsilon}\|_2 \leq \beta\|\bar{\epsilon}\|_2$$

Thus $\|PWP\bar{y}\|_2 \leq \mu + \beta\|\bar{y}\|_2$.

Max eigenvalue of $PWP \leq \mu + \beta$. □

Now we prove the lemma:

We have $\|PWP\bar{y}\|_2 \leq (\mu + \beta)\|\bar{y}\|_2$, so

$$\|(PWP)^t\mathbb{1}\|_1 \leq (\mu + \beta)^t\|\mathbb{1}\|_1 = \sqrt{n}(\mu + \beta)^t$$

$$\frac{1}{n}\|(PWP)^t\mathbb{1}\|_1 \leq \frac{1}{\sqrt{n}}(\mu + \beta)^t$$

5.1 BPP Derandomization

BPP derandomization is a bit more technical to prove.

We state a Chernoff bound for walks on expanders:

Theorem 3. *Let $G = (V, E)$ be a d-regular graph. Let $v_1, v_2, \ldots, v_t$ be vertices of a random walk on G where v_1 is chosen uniformly at random from V. Let $f : V \to [0, 1]$ be any bounded function. Then*

$$\Pr\left[\left|\frac{1}{t}\sum_{i=1}^{t} f(v_i) - \mathbb{E}[f]\right| \geq \delta\right] \leq 2e^{-\Omega(\beta\delta^2 t)}$$

where $\mathbb{E}[f] = \mathbb{E}_v[f(v)]$ where v is chosen at random from V, and β is the spectral gap. Note $\beta = 1 - \lambda_2$ if G is ergodic.

Using the above powerful theorem, one can generalize the majority vote algorithm for **BPP** to obtain error reduction to $2^{-\Omega(k)}$ using $O(n + k)$ random bits.