Lecture 8: Expanders and Applications

Topics in Complexity Theory and Pseudorandomness (Spring 2013)
Rutgers University
Swastik Kopparty
Scribes: Amey Bhangale, Mrinal Kumar

1 Overview

In this lecture, we will introduce some notions of expanders and then explore some of their useful properties and applications.

2 Expanders

In one of the previous lectures, we have already used one notion of expander graphs while studying data structures for the set membership problem. We will now look at some other definitions typically used to describe Expanders and then explore their properties.

Definition 1. *For a real number $\alpha > 0$ and a natural number k, a graph $G(V, E)$ is said to be an (α, k)-Edge Expander if for every $S \subseteq V$, such that $|S| \leq k$, the number of edges from S to $V \setminus S$, denoted as $e(S, \bar{S})$ is at least $\alpha|S|$.*

Now, it is easy to see from the definition that a complete graph is definitely an edge expander as per the above definition. For our applications, we will be mostly interested in expander graphs which have a much smaller number of edges. For example, we would be interested in d-regular graphs which are expanders with parameters $d = O(1)$ and $\alpha = \theta(1)$. Clearly, α cannot exceed d. The existence of such graphs is guaranteed by the following theorem which we will come back and prove at a later point in time.

Theorem 2. *For any natural number $d \geq 3$, and sufficiently large n, there exist d-regular graphs on n vertices which are $(\frac{d}{10}, \frac{n}{2})$ edge expanders.*

Let us now define another notion of expanders, this time based upon the number of vertices in the neighborhood of small subsets of vertices.

Definition 3. *For a real number α and a natural number k, a graph $G(V, E)$ is said to be an (α, k)-Vertex Expander if for every $S \subseteq V$, such that $|S| \leq k$, $|N(S)| \geq \alpha|S|$.*
Here, $N(S) = \{x \in V : \exists y \in S$ such that $(x, y) \in E\}$.

The following theorem guarantees the existence of vertex expanders for some choice of parameters.

Theorem 4. *For any natural number $d \geq 3$, and sufficiently large n, there exist d-regular graphs on n vertices which are $(\frac{d}{10}, \frac{n}{10d})$ vertex expanders.*

These definitions of expanders were defined in terms of the expansion properties of small sized subsets of vertices. We will now define another notion of expanders in term of the eigenvalues of the adjacency matrix of a graph.

Definition 5. *For any natural number d, a d-regular graph $G(V, E)$ is said to be a λ-absolute eigenvalue expander if $|\lambda_2|, |\lambda_3|, \ldots, |\lambda_n| \leq \lambda$. Here, $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$ are the eigenvalues of the adjacency matrix A of G.*

The following theorem which we will believe without a proof tells us that there exist expanders whose all eigenvalues except the first one are bounded away from d.

Theorem 6 (Broder, Shamir)**.** *For all positive integers $d \geq 3$ there exists a $\lambda < d$ such that for all sufficiently large n, there is a d-regular graph which is a λ-absolute eigenvalue expander.*

In fact, we also know that the there exist λ-absolute expanders with λ around $\sqrt{d}$.

3 Some properties of eigenvectors and eigenvalues

To understand the definitions better and to be able to use them for our ends, let us first look at some basic properties of the eigenvalues and eigenvectors of the adjacency matrix of a d-regular graph. During the course of this entire discussion, we will sometimes refer to the eigenvalues and eigenvectors of the adjacency matrix A of G as eigenvalues and eigenvectors of G.

Lemma 7. *Let $G(V, E)$ be an n vertex undirected d- regular graph for some natural number d. Let $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$ be its n eigenvalues. Then,*

1. *For all $i \in [n]$, -$d \leq \lambda_i \leq d$*
2. *$\lambda_1 = d$*
3. *G is connected $\iff$ $\lambda_2 < d$*
4. *G is nonbipartite $\iff$ $\lambda_n >$ -d*

Proof. Before going into the proof, let us list down some basic properties of the adjacency matrix A of G.

- A is symmetric
- Every row and every column of A have exactly d ones.

Now, let us prove the items in the lemma.

1. Let v be an eigenvector of A with eigenvalue λ. Let $v(x)$ be the component of v with the maximum absolute value. Then,we know that

$$\lambda v(x) = \sum_{j \in [n]} A_{ij} v(j) \tag{1}$$

Now, using -$|v(x)| \leq |v(j)| \leq |v(i)|$ in the equation above, we get

$$|\lambda v(x)| = |\sum_{j \in [n]} A_{ij} v(j)| \leq \sum_{j \in [n]} A_{ij} |v(j)| \leq d|v(x)| \tag{2}$$

This gives us

$$\text{-}d \leq \lambda \leq d \tag{3}$$

2. To show that the maximum eigenvalue is d, it is sufficient to show that there is a vector v_1 so that $Av = dv_1$. From the item above, it follows that $\lambda_1 = d$. Consider the vector $v_1 \in \mathbb{R}^n$ which is 1 in all coordinates. Then, $Av_1 = dv_1$. Thus, d is the maximum eigenvalue.

3. Let us first prove the reverse direction. Let G be disconnected. Let C_1 and C_2 be its two connected components. Let $v_1 = 1_{C_1}$ and $v_2 = 1_{C_2}$. Observe that $Av_1 = dv_1$ and $Av_2 = dv_2$. Since v_1 and v_2 are linearly independent, this gives us that the second largest eigenvalue is also d.

   Let us now argue the converse. Let G have the second eigenvalue d. This means that there is a vector v_2 which is orthogonal to 1_V such that $Av_2 = dv_2$. Let x be the component such that $v_2(x)$ has the maximum value for $x \in V$. Now, $v_2(x) = \frac{1}{d} \sum_{y \in N(x)} v_2(y)$. Since there are precisely d nonzero terms in the sum on the right hand side, the maximality of x implies that the $v_2(y) = v_2(y)$ at every neighbor y of x. This argument can be now extended similarly to imply that every vertex z in the same connected component as x, satisfies $v_2(x) = v_2(z)$. In particular, all the cells of v_2 indexed by the vertices in the same connected component as x have the same sign. But from the fact that v_2 is orthogonal to 1_V, we know that there are cells with entries with different signs in v_2. Hence, not all vertices in the graph lie in the same connected component as x. Therefore, G is disconnected.

4. Let $G(V, E)$ be a d-regular bipartite graph with bipartitions L and R. To show that -d is an eigenvalue, it is sufficient to give a vector v such that $Av = -dv$. Consider a vector v defined as follows:

$$v(x) = 1, \forall x \in L \tag{4}$$

$$v(x) = -1, \forall x \in R \tag{5}$$

   Now it is not difficult to see that $Av =$ -dv and hence there is an eigenvalue which is not greater than -d. Since we know that the all eigenvalues are $\geq -d$, so $\lambda_n = -d$.

   Let us now show that the converse also holds. Let us now consider a graph which has an eigenvalue less than or equal to -d. Item 1 of this lemma tells us that $\lambda_n = -d$. Let us work with a graph which is connected. For disconnected graphs, the argument can be applied by applying it to each of the connected components. Then, let the eigenvector for λ_n be v_n. Let x be the component of v with the largest absolute value. From the eigenvalue relation, we will get

$$-dv(x) = \sum_{y \in N(x)} v(y) \tag{6}$$

   From the choice of x, the only way this equality can hold is when all neighbors of x have the same absolute value as $v(x)$ with a different sign. This argument can be applied again with one of the neighbors of x as the component of interest to conclude that all the components

corresponding to the vertices in the same connected component as x have the same same absolute value and the sign at any vertex differs from all its neighbors. Therefore, there are no edges among the vertices with the same sign and hence the positive and the negatively signed vertices form a bipartition of G.

□

Let us now look at how eigenvalues of a matrix change with respect to some operations on them. We will be using these properties very cruicially in our analysis later in the lecture.

Lemma 8. *If λ is an eigenvalue of a $n \times n$ matrix M with eigenvector v, then*

1. *$\frac{\lambda}{d}$ is an eigenvalue of $\frac{1}{d}M$ with eigenvector v for a nonzero scalar d.*
2. *λ^i is an eigenvalue for M^i with eigenvector v for a natural number i.*

Proof. The proofs basically follow from the basic definition.

1. Since λ is an eigenvalue of M with eigenvector v, we get $Mv = \lambda v$. Multipliying both sides by the scalar $\frac{1}{d}$, we get the desired claim.
2. Since λ is an eigenvalue of M with eigenvector v, we get $Mv = \lambda v$. Now, multiplying both sides of the equality by M, we get $MMv = M\lambda v$, which is the same as $M^2v = \lambda Mv = \lambda^2 v$. Repeating this procedure $i - 2$ more times, we get the required claim.

□

We will also cruicially use the following fact about the spectrum of a real symmetric matrix.

Fact 9. *For a real symmetric square matrix M, the following is true*

- *All its eigenvalues are real.*
- *There is set of eigenvectors of M, $\{v_i : i \in [n]\}$ which form an orthonormal basis for $\mathbb{R}^n$.*

3.0.1 Some Notations

For the rest of the lecture, we will always use $G(V, E)$ to refer to a n vertex d-regular graph which is a λ-absolute eigenvalue expander and A will be its adjacency matrix. We will refer by P, the matrix $\frac{1}{d}A$. We will use $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$ to refer to the eigenvalues of A and $\{v_i : i \in [n]\}$ as the set of orthonormal eigenvectors, where $Av_i = \lambda_i v_i$ for every $i \in [n]$. From Claim 8, we know that $\frac{\lambda_1}{d} \geq \frac{\lambda_2}{d} \geq \ldots \geq \frac{\lambda_n}{d}$ are the eigenvalues of P and the set of eigenvectors remains the same.

4 Random walks on expanders

From the definition of edge and vertex expanders, it is intuitively clear that while doing a random walk on an expander, the chances of being "trapped" inside a small subset of vertices should be small. We will now analyse a random walk on an expander and show that this is indeed correct. Let G be d-regular graph on n vertices, which is a λ-absolute eigenvalue expander. Let us define a random walk on it as follows:

- We start at a fixed vertex x_0 in the first step.
- In the i^th step, we chose a uniformly random neighbor of x_{i-1} as x_i.

This procedure will give us a distribution on the vertex set of the graph. Let us call the distribution obtained at the end of the i^{th} step as f_i. Clearly, f_0 is 1 at x_o and 0 everywhere else. The next claim tells us the relation between f_{i-1} and f_i.

Claim 10. *For all integers $i \geq 1$ $f_i = \frac{1}{d}Af_{i-1}$.*

Proof. The probability of being at a vertex x in the i^{th} step is precisely the probability that we are at a veterx y, which is a neighbor of x in the $i-1$ step and we pick the edge (x, y) in the i^th step. So we get,

$f_i(x) = \sum_{y \in N[x]} \Pr[\text{we are at } y \text{ after } (i-1)^{th} \text{ step}] \Pr[\text{edge } (x,y) \text{ is picked in the next step}]$

This gives us $f_i(x) = \sum_{y \in N(x)} f_{i-1}(x)\frac{1}{d}$. Hence, $f_i = \frac{1}{d}Af_{i-1}$. □

Recall that from the definition of matrix P, we get $f_i = Pf_{i-a}$. Now, in general, applying the Claim 10 multiple times, we will get the following

Claim 11. *For all integers $i \geq 0$, $f_i = P^i f_0$*

Let us now show that after sufficiently large number of steps i, the distribution f_i is "close" to the uniform distribution on the vertex set of G.

Theorem 12. *For an n vertex d-regular λ-absolute eigenvalue expander G, the distribution f_i obtained at the end of i steps of a random walk as defined above satisfies $\|f_i - U\|_2 \leq \left(\frac{\lambda}{d}\right)^i \left(1 - \frac{1}{n}\right)$.*

Proof. Using Fact 9 above, we know that we can express any vector $u \in \mathbb{R}^n$ as a linear combination of $\{v_i : i \in [n]\}$. In particular, $f_0 = \sum_{i \in [n]} \alpha_i v_i$, where $\alpha_i = \langle f_0, v_i \rangle$. Now, using the Claim 11, we obtain

$$f_i = P^i f_0 = P^i \sum_{j \in [n]} \alpha_j v_j \tag{7}$$

Now we can separate this sum into two parts, keeping in mind that $\lambda_1 = d$ and P^i has eigenvalues $\{\lambda_k^i : k \in [n]\}$. We get

$$f_i = \alpha_1 v_1 + \sum_{j \in [n] \setminus \{1\}} \alpha_j \left(\frac{\lambda_j}{d}\right)^i v_j \tag{8}$$

Taking the ℓ_2 norms on both sides,

$$\|f_i - \alpha_1 v_1\|_2 = \left\| \sum_{j \in [n] \setminus \{1\}} \alpha_j \left(\frac{\lambda_j}{d}\right)^i v_j \right\|_2 \tag{9}$$

Since G is a λ-absolute eigenvalue expander, we can know that $\|\lambda_i\| \leq \lambda$ for $i \in \{2, 3, 4 \ldots, n\}$. So, the right hand side gets simplified, and we obtain

$$\|f_i - \alpha_1 v_1\|_2 \leq \left(\frac{\lambda}{d}\right)^i \left\| \sum_{j \in [n] \setminus \{1\}} \alpha_j v_j \right\|_2 \tag{10}$$

Now, from the definition of α_i, we know that $f_0 = \alpha_1 v_1 + \sum_{j \in [n] \setminus \{1\}} \alpha_j v_j$. So taking the ℓ_2 norm both sides, we obtain using orthonormality of v_i's

$$\|f_0\|_2{}^2 = |\alpha_1|^2 + \left\| \sum_{j \in [n] \setminus \{1\}} \alpha_j v_j \right\|_2^2 \tag{11}$$

We also know that $\alpha_1 = \langle f_0, v_1 \rangle$. Recall that f_0 is a vector with all 0's and 1 at one position. v_1 is a normalized all ones vector, and so has each entry $\frac{1}{\sqrt{n}}$. So, $\alpha_1 = \langle f_0, v_1 \rangle = \frac{1}{\sqrt{n}}$.

$$\|f_0\|_2{}^2 - \frac{1}{n} = \left\| \sum_{j \in [n] \setminus \{1\}} \alpha_j v_j \right\|_2^2 \tag{12}$$

Besides, we can also observe that $\alpha_1 v_1$ is a vector whose all components are equal to $\frac{1}{n}$, which is precisely equal to the uniform distribution U on n vertices. Using both these observations, we obtain fron Equation 10, the following

$$\|f_i - U\|_2 \leq \left(\frac{\lambda}{d}\right)^i \left(\|f_0\|_2{}^2 - \frac{1}{n}\right)^{\frac{1}{2}} \tag{13}$$

Now, the ℓ_2 norm of f_0 is 1. So, we obtain

$$\|f_i - U\|_2 \leq \left(\frac{\lambda}{d}\right)^i \left(1 - \frac{1}{n}\right)^{\frac{1}{2}} \tag{14}$$

□

From here, we can also get something about the ℓ_1 norm using the relation between the ℓ_2 and the ℓ_1 norm. We obtain the following corollary.

Corollary 13. $\|f_i - U\|_1 \leq \sqrt{n}\left(\frac{\lambda}{d}\right)^i \left(1 - \frac{1}{n}\right)^{\frac{1}{2}}$

In particular, if λ and d are absolute constants, for $i = \theta(\log(n))$, we can conclude that f_1 is $\frac{1}{n^{100}}$ close to uniform distribution in ℓ_1 norm. It also follows from this statement that the support of f_i has to be n for $i = c \log n$, where c is a sufficiently large constant. This gives us the following corollary,

Corollary 14. *The diameter of a λ-absolute eigenvalue expander is $\theta(\log n)$.*

5 Properties of λ-absolute eigenvalue expander

5.1 Expander Mixing Lemma

It is not hard to prove that a random d-regular graph is a good expander. We can think of *expander mixing lemma* as showing somewhat converse of the previous statement. Informally it says that, any d-regular λ-absolute eigenvalue expander graph is close to a random d-regular graph.

Lemma 15. *(Expander mixing lemma) If $G(V, E)$ be a d-regular λ-absolute eigenvalue expander then for all $S, T \subseteq V$,*

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

where $e(S, T)$ is the number of edges between the vertex set S and T.

In the above expression $\frac{d}{n}|S||T|$ is the expected number of edges between S, T in a random d-regular graph. So the above lemma says that for any d-regular λ-absolute eigenvalue expander graph, the quantity $e(S, T)$ is close to the the expected value in a random d-regular graph.

Proof of expander mixing lemma :

Proof. Let 1_S and 1_T be the indicator $0/1$ vectors for vertex set S and T respectively. The number of edges between S and T is given by following expression.

$$\begin{aligned} e(S,T) &= \sum_{u \in S, v \in T} A_{uv} \\ &= 1_S^T A 1_T \end{aligned}$$

We can write vectors 1_S and 1_T as a linear combinations of eigen vectors. So,

$$\begin{aligned} e(S,T) &= \left(\sum_i \alpha_i v_i \right)^T A \left(\sum_j \beta_j v_j \right) \\ &= \sum_i \sum_j \beta_j \alpha_i v_i^T \lambda_j v_j \\ &= \sum_i \alpha_i \beta_i \lambda_i \\ &= \alpha_1 \beta_1 \lambda_1 + \sum_{i=2}^{n} \alpha_i \beta_i \lambda_i \end{aligned}$$

We know

$$\alpha_1 = \frac{|S|}{\sqrt{n}}, \beta_1 = \frac{|T|}{\sqrt{n}}, \lambda_1 = d$$
$$\alpha_1\beta_1\lambda_1 = \frac{d}{n}|S||T|$$

Therefore,

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

We can bound the right hand side as,

$$\begin{aligned}\left|\sum_{i=2}^{n} \alpha_i\beta_i\lambda_i\right| &\leq \lambda \sum_{i=2}^{n} |\alpha_i\beta_i| \\ &\leq \lambda \left(\sum_{i=2}^{n} |\alpha_i|^2\right)^{\frac{1}{2}} \left(\sum_{i=2}^{n} |\beta_i|^2\right)^{\frac{1}{2}} \\ &\leq \lambda \|1_S\|_2 \|1_T\|_2 \\ &= \lambda\sqrt{|S||T|}\end{aligned}$$

Therefore,

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

□

6 Error Reduction

Consider a randomized circuit C computing some function f on n variables. The property of circuit C is that for every input $x \in \{0,1\}^n$, with probability at least $2/3$, it outputs the correct answer using r *truly* random bits. We have already seen the following methods to bring down the error probability.

1. Repeating the compution of C with fresh random bits every time and taking the *majority* of the outputs. In this case, if we repeat the computation m times independently then we can bring down the error probability to $exp(-m)$. But we pay for it in total number of random bits used by the circuit which is rm.

2. Instead of using truly random bits every time, if we use m pairwise independent strings, we can bring down the probability to $1/m$. Here the total number of *truly* random bits used is $2r$.

In this section we will see an application of expander graphs in bringing down the error probability using few random bits. Consider the universe of r random bits $\{0,1\}^r$. Circuit C has an error probability atmost $1/3$ means for every input $x \in \{0,1\}^n$ there are atmost $2^r/3$ values in $\{0,1\}^r$ which are bad for x. Let B be the bad set. If we can generate d points *efficiently* in the set $\{0,1\}^r$ such that probability that more than half of them lie inside B is small then we can just use these d points as seeds to circuit C and output the $majority$. Hence with very small probability th circuits $errs$.

Take a d-regular λ-eigenvalue expander graph G on the vertex set $\{0,1\}^r$. We want this graph to be *explicit*, that is, given a vertex $x \in \{0,1\}^r$ in G, we should be able to find its d neighbors in $\mathsf{poly}(r)$ time.

6.1 Approach 1

1. Pick a uniformly random $x \in \{0,1\}^r$
2. Let $x_1, x_2, \cdots, x_d$ be its neighbors in G.
3. Run circuit C with seeds $x_1, x_2, \cdots, x_d$
4. Output $majority$ of the answers.

Randomness used : The only randomness used in above procedure is in step 1 which is just r random bits.

We will show that the error probability of above algorithm is very small.

Claim 16. *The error probability of the above algorithm is atmost* $O(\frac{1}{\sqrt{m}})$

Proof. For an input $y \in \{0,1\}^n$, let B be the subset of $\{0,1\}^r$ which is bad for y i.e. $C(y,r) = f(y)$ if and only if $r \notin B$. Define a set D as follows,

$$D = \{x \text{ such that at least d/2 of its neighbors } x_1, x_2 \cdots, x_d \text{ are in } B\}$$

In order to show the error probability is small we want to argue that the size of D is small, since the error probability of the algorithm is $|D|/2^r$,

Consider these two subsets B and D of vertex set of a garph G. Applying the expander mixing lemma (Lemma 15):

$$\left| e(D,B) - \frac{d}{2^r}|B||D| \right| \leq \lambda\sqrt{|B||D|}$$

Number of edges going across B and D is at least $|D|.d/2$,

$$\begin{aligned}
|D|\frac{d}{2} &\leq \frac{d}{2^r}|B||D| + \lambda\sqrt{|B||D|} \\
|D|\frac{d}{2} &\leq \frac{d}{3}|D| + \lambda\sqrt{|B||D|} \\
\frac{d}{6}|D| &\leq \lambda\sqrt{|B||D|} \\
|D| &\leq O(\frac{\lambda^2|B|}{d^2})
\end{aligned}$$

Setting $d = m$ and $\lambda = m^{3/4}$,

$$\text{Error Probability} = \frac{|D|}{2^r} \leq O(\frac{1}{\sqrt{m}})$$

□

So this algorithm uses only r random bits and brings down the error probability from $1/3$ to $O(1/\sqrt{m})$!

6.2 Approach 2

1. Pick $x \in \{0,1\}^r$ uniformly at random.
2. Take a random walk of length m
   $x_0 = x, x_1, x_2, \cdots, x_m$.
3. Output the majority of $(C(x_i) \mid i \in [m])$.

Randomness used : First step requires r random bits, to pick a random vertex in $\{0,1\}^r$. Since G is a d-regular graph, we can think of neighbors of a vertex in a graph are labeled by numbers in $[d]$. So each step of random walk is same as picking a random number between 1 to d and moving to that neighbor of a vertex. Picking a random number between 1 to d needs roughly $\log d$ random bits, so for a random walk of length m we need total $m \log d$ random bits. Since d is a constant in this case, the randomness used by above algorithn is $r + O(m)$.

In order to bound the error probability of above algorithm we will be interested in following quantities.

1. Pr[all x_i lie in B]
2. For a fixed set of indices $I \subseteq [m], |I| = m/2$.
   Pr[$x_i, i \in I$ are all in B]
3. Union bound over all $I's$.

Since we pick the vertex x_0 uniformly at random, we know

$$Pr[x_0 \in B] = \frac{|B|}{2^r}$$

Want to estimate $Pr[x_0, x_1 \in B] =?$. In order to estimate this quantity, we start with a distribution f_0 which corresponds to distribution of vertex x_0, an uniform distribution. Let π be the restriction onto B i.e.

$$\pi : \mathbb{R}^V \to \mathbb{R}^V,$$
$$\pi(f)_i = \begin{cases} f_i & \text{if } i \in B \\ 0 & \text{otherwise} \end{cases}$$

Let P be the normalized adjacency matrix of expander graph G. Then,

$$Pr[x_0 \in B] = \|\pi f_0\|_1 = \beta$$
$$Pr[x_0, x_1 \in B] = \|\pi P \pi f_0\|_1$$
$$Pr[x_0, x_1, \cdots, x_j \in B] = \|(\pi P)^j \pi f_0\|_1$$

Since $\pi^2 = \pi$, $(\pi P)^j \pi = (\pi P \pi)^j$. Hence if we can get an upper bound on $\|(\pi P \pi)^j f_0\|_2$, we get an upper bound on $\|(\pi P \pi)^j f_0\|_1$.

Claim 17. *For a d-regular, λ eigenvalue expander graph,*

$$\forall f, \|\pi P \pi f\|_2 \leq \left(\beta + \frac{\lambda}{d}\right) \|f\|_2$$

where P, β and π as defined above.

Proof. We can write a vector πf as a linear combination of eigenvectors $v_1, v_2, \cdots, v_{2^r}$ of G. Let v^* be the component of vector πf orthogonal to v_1.

$$\pi f = \alpha_1 v_1 + v^*$$

$$\begin{aligned} \pi P \pi f &= \pi P \alpha_1 v_1 + \pi P v^* \\ &= \alpha_1 \pi P v_1 + \pi(P v^*) \end{aligned}$$

By triangle inequality,

$$\|\pi P \pi f\|_2 \leq \|\alpha_1 \pi P v_1\|_2 + \|\pi(P v^*)\|_2 \tag{15}$$

We can bound the first expression in equation 15 as,

$$\begin{aligned}\|\alpha_1 \pi P v_1\|_2 &= \alpha_1 \|\pi P v_1\|_2 \\ &\leq \alpha_1 \|\pi v_1\|_2 \\ &= \alpha_1 \sqrt{\beta}\end{aligned}$$

We know,

$$\alpha_1 = \langle \pi f, v_1 \rangle$$

By Cauchy-Schwarz inequality and the fact that πf has atmost $|B|/2^r$ fraction of non-zero entries,

$$\begin{aligned}\alpha_1 &\leq \|\pi f\|_2 \sqrt{\beta} \\ &\leq \|f\|_2 \sqrt{\beta}\end{aligned}$$

Hence,

$$\begin{aligned}\|\alpha_1 \pi P v_1\|_2 &\leq \|f\|_2 \sqrt{\beta}\sqrt{\beta} \\ &= \beta \|f\|_2\end{aligned}$$

Also for the second expression in equation 15,

$$\begin{aligned}\|\pi(P v^*)\|_2 &\leq \|P v^*\|_2 \\ &\leq \frac{\lambda}{d} \|v^*\|_2 \\ &\leq \frac{\lambda}{d} \|f\|_2\end{aligned}$$

as required. □

Now,

$$\begin{aligned}Pr[x_1, x_2, \cdots, x_m \text{ all in } B] &\leq \|(\pi P \pi)^m f_0\|_1 \\ &\leq \sqrt{2^r} \|(\pi P \pi)^m f_0\|_2 \\ &\leq \sqrt{2^r} \left(\beta + \frac{\lambda}{d}\right)^m \|f_0\|_2 \qquad \text{... by claim 17} \\ &= \sqrt{2^r} \left(\beta + \frac{\lambda}{d}\right)^m \frac{1}{\sqrt{2^r}} \\ &= \left(\beta + \frac{\lambda}{d}\right)^m\end{aligned}$$

Using the claim we proved, we will try to estimate the following probability,

- For a fixed set of indices $I \subseteq [m], |I| = m/2$.
  Pr[$x_i, i \in I$ are all in B].

The above probability is exactly some expression of the form

$$\Pr[\ x_i, i \in I \text{ are all in } B] = \|..(\pi P)PP(\pi P)(\pi P)...\|_1$$

which contains exactly $m/2$ (πP) terms and $m/2$ P terms. By combining the terms, we can rewrite the above expression in the form,

$$\Pr[\ x_i, i \in I \text{ are all in } B] = \|(\pi P^{k_1}\pi)(\pi P^{k_2}\pi)\cdots f_0\|_1$$

For some $k_1, k_2, \cdots k_{m/2-1}$.

Claim 18.

$$\forall f, \|\pi P^k \pi f\|_2 \leq \left(\beta + \left(\frac{\lambda}{d}\right)^k\right)\|f\|_2$$

Proof. Proof of this claim is similar to the claim 17 except in equation 15 we have the second term $\|P^k v^*\|_2$ instead of $\|Pv^*\|_2$ which is atmost $(\lambda/d)^k\|f\|_2$. □

Using above claim, For a fixed set of indices $I \subseteq [m], |I| = m/2$,

$$\begin{aligned}
\Pr[\ x_i, i \in I \text{ are all in } B] &= \|(\pi P^{k_1}\pi)(\pi P^{k_2}\pi)\cdots f_0\|_1 \\
&\leq \sqrt{2^r}\|(\pi P^{k_1}\pi)(\pi P^{k_2}\pi)\cdots f_0\|_2 \\
&\leq \sqrt{2^r}\left(\beta + \left(\frac{\lambda}{d}\right)^{k_1}\right)\left(\beta + \left(\frac{\lambda}{d}\right)^{k_2}\right)\cdots\left(\beta + \left(\frac{\lambda}{d}\right)^{k_{m/2-1}}\right)\|f_0\|_2 \\
&\leq \left(\beta + \frac{\lambda}{d}\right)^{m/2-1}
\end{aligned}$$

If we choose λ and d such that $(\beta + \frac{\lambda}{d})^{1/2} < \frac{1}{5}$, then the error probability of the algorithm is,

$$\begin{aligned}
\text{Error Probability} &= \Pr[\text{ majority of } x_i\text{'s are in } B] \\
&= \bigcup_{I\subseteq[m],|I|\geq m/2} \Pr[x_i, i \in I \text{ are in } B] \\
&\leq O\left(2^m\right) \max_{I, I\subseteq[m],|I|\geq m/2}\{\Pr[\ x_i, i \in I \text{ are all in } B]\} \\
&\leq O\left(\frac{2^m}{5^m}\right) \\
&= exp(-m)
\end{aligned}$$

Hence the algorithm uses $r + O(m)$ random bits and reduces error probability to $exp(-m)$.

7 Connectivity of a d-regular graph

In this section we will be looking at the following problem.

Problem : Given an undirected graph G on n vertices which is a d-regular , determine whether it is *connected*.

The graph is given as input in the form of adjacency matrix in readonly memory. This problem is simple if we have access to $poly(n)$ bits of space for computation.

1. Start with any arbitrary node in a graph G.
2. Preform a DFS/BFS from the starting node and count the number of nodes in the DFS/BFS tree.
3. If the count is equal to n then the graph is connected otherwise it is disconnected.

But the problem is not trivial if we have access to only $O(\log n)$ space. We will discuss a randomized algorithm to solve this problem in $O(\log n)$ space. The algorithm is as follows:

- For every pair of vertices s, t , take n^{10} independent random walks of length n^{10} each starting from s and check if it ends at vertex t.
- If for all pairs of vertices the above condition is satisfied for at least one random walk then G is connected ($true$) otherwise disconnected($false$).

Claim 19. *The above algrithm fails with exponential small probability in n.*

Proof. If the graph is disconnected then the algorithm always return $false$. We will show if G is connected then the algorithm returns $false$ with very small probability.

If G is a *connected* d-regular graph on n vertices, then except λ_1 the absolute value of other eigenvalues is atmost $d - d/n^2$. Let e_s be an indicator vector of a vertex s. Let P be the normalized adjacency matrix of graph G. By theorem 12- the property of random walks on an expander graph,

$$\begin{aligned}
\Pr[\text{ a random walk from } s \text{ lands on } t \text{ at the end of } n^{10^{th}} \text{ step}] &\geq \frac{1}{n} - \|U - P^{n^{10}} e_s\|_1 \\
&\geq \frac{1}{n} - \sqrt{n}\left(\frac{\lambda}{d}\right)^{n^{10}} \\
&\geq \frac{1}{n} - \sqrt{n}\left(1 - \frac{1}{n^2}\right)^{n^{10}} \\
&\geq \frac{1}{n} - e^{-n^8} \\
&\geq \frac{1}{2n}
\end{aligned}$$

Therefore,

$$Pr\left[\begin{array}{r}\text{none of the } n^{10} \text{ random walks from } s \\ \text{lands on } t \text{ at the end of } n^{10^{th}} \text{ step}\end{array}\right] \leq \left(1 - \frac{1}{2n}\right)^{n^{10}}$$
$$\leq exp(-n^9)$$

Hence by union bound,

$$Pr[\text{failing}] \leq n^2 exp(-n^9)$$
$$\leq exp(-n^8)$$

Hence the algorithm fails with exponential small probability. □

In a later lecture, we will see a *deterministic* algorithm for this problem which uses only $O(\log n)$ space.