CS 252, Lecture 11: Expander Graphs

1 Introduction

Expander graphs are sparse yet highly connected graphs. That is, for every subset S of vertices of the graph, there are a lot of edges leaving S. In other words, we cannot disconnect the graph into near-equal pieces by cutting a few edges. This notion is useful in various contexts ranging from design of computer networks to proving the existence of probabilistically checkable proofs.

There are many ways to define an expander graph formally: in terms of spectral expansion, vertex expansion or edge expansion. We can also define them separately for bipartite and general graphs. However, in this lecture, we will stick with the spectral notion of expanders. We will also consider general graphs. This connects well with random walks and pseudorandom properties of the expanders.

Recall from the spectral graph theory lecture that the adjacency matrix A of a d-regular graph $G = (V, E)$ with n vertices has n eigenvalues $d = \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$. We have also proved that the graph is connected if and only if $\lambda_2 < d$. Let $\lambda(G) = \max\{|\lambda_2|, |\lambda_n|\}$. Note that[1]

$$\lambda(G) = \max_{x:\langle x,u\rangle=0} \frac{\|Ax\|}{\|x\|}$$

where $u = \left(\frac{1}{n}, \frac{1}{n}, \ldots, \frac{1}{n}\right)$. Recall that we have proved earlier that $\lambda(G) \leq d$.

Definition 1. (Expander graphs) A graph $G = (V, E)$ is said to be an (n, d, λ)-expander if $|V| = n$, each vertex of the graph has degree d, and $\lambda(G) = \lambda$.

The parameter $d - \lambda(G)$ is known as the *spectral expansion* of the graph G. As we alluded to earlier, a different way to define expander graphs is in terms of *edge expansion*:

$$\min_{S\subseteq V, |S|\leq \frac{n}{2}} \frac{E(S, \overline{S})}{|S|}$$

Both the notions of expansion are actually closely related. We have

$$\frac{E(S, \overline{S})}{|S|} \geq \frac{d - \lambda_2}{2} \, \forall S : |S| \leq \frac{n}{2}.$$

[1] Unless stated otherwise, $\|.\|$ denotes the ℓ_2 norm.

The converse of this is also true. Graphs with good edge expansion also have good spectral expansion. This is given by the Cheeger inequality:

$$\exists S : \frac{E(S, \overline{S})}{|S|} \leq O\left(\sqrt{d(d - \lambda_2)}\right)$$

2 Construction of expander graphs

Of course expander graphs have a lot of useful properties, but we really need to construct them so that we can make use of them. By constructing expander graphs, we mean constructing a family of d-regular graphs $G_1, G_2, \ldots$, such that G_i has i vertices. The goal is to find a polynomial time algorithm that takes input n, d and outputs the graph G_n in the family.

Random graphs are good expanders. In fact, one can view expander graphs as *pseudorandom* in the sense that they have a lot of properties of random graphs. Random d-regular graphs actually achieve $\lambda_2 \leq O(\sqrt{d})$. This is roughly, the best expansion that we can get. Such graphs, whose $\lambda_2 = 2\sqrt{d-1}$ are known as Ramanujan graphs.

Even though random graphs are good expanders with high probability, finding explicit construction of expander graphs is a challenging question. There are various known such explicit constructions:

1. (Discrete torus expanders) The first known explicit construction of expanders is due to Margulis, Gabber and Galil, where $G = (V, E)$ with the vertex set $V = \mathbb{Z}_m^2$, and a vertex (x, y) is adjacent to the following vertices: $(x \pm y, y), (x \pm (y + 1), y), (x, y \pm x), (x, y \pm (x + 1))$. It is a 8-regular graph. We can prove that λ_2 of this graph is $8c$ for $c < 1$.

2. (p-cycle with inverse chords) The graph $G = (V, E)$ with vertex set $V = \mathbb{Z}_p$, and edges connect each node x with $x + 1, x - 1, x^{-1}$. The arithmetic is $\bmod p$, and set $0^{-1} = 0$.

3 Pseudorandom properties of expander graphs

A key property of expander graphs is the so called expander mixing lemma. In a random graph with constant degree d, the number of edges between two sets S and T is roughly equal to $\frac{d|S||T|}{n}$. Expander graphs also mimic this property:

Theorem 2. *(Expander Mixing Lemma) Let G be an (n, d, λ)-expander. Then, for any two subsets S, T of vertices of the graph,*

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

Proof. We will estimate the value of $(\mathbf{1}_S)^T A \mathbf{1}_T$, where $\mathbf{1}_S$ is the indicator vector of the set S:

$$(\mathbf{1}_S)_j = \begin{cases} 0 & \text{if } j \notin S \\ 1 & \text{if } j \in S. \end{cases}$$

$\mathbf{1}_T$ is also defined analogously. As before, let $u = \left(\frac{1}{n}, \frac{1}{n}, \ldots, \frac{1}{n}\right)$. Let $\mathbf{1}_S = |S|u + v_1, \mathbf{1}_T = |T|u + v_2$. Note that $\langle v_1, u\rangle = \langle v_2, u\rangle = 0$. Furthermore, we get

$$\|v_1\| = \sqrt{(n - |S|)\frac{1}{n^2} + |S|\left(1 + n^2 - \frac{2}{n}\right)} = \sqrt{\frac{1}{n} + |S| - \frac{2|S|}{n}} \leq \sqrt{|S|}$$

Similarly, $\|v_2\| \leq \sqrt{|T|}$.

The number of edges between S and T is equal to $(\mathbf{1}_S)^T A \mathbf{1}_T$. We can write it as

$$\begin{aligned}
&= \left(|S|u^T + v_1^T\right) A \left(|T|u + v_2\right) \\
&= \left(|S|u^T + v_1^T\right) \left(d|T|u + Av_2\right) \\
&= \frac{d|S||T|}{n} + d|T|v_1^T u + |S|u^T A v_2 + v_1^T A v_2 \\
&= \frac{d|S||T|}{n} + v_1^T A v_2 \quad (\text{since } \langle v_1, u\rangle = 0, u^T A v_2 = (Au)^T v_2 = d\langle u, v_2\rangle = 0)
\end{aligned}$$

Thus, we get

$$\begin{aligned}
\left|E(S,T) - \frac{d|S||T|}{n}\right| &= v_1^T A v_2 \\
&\leq \|v_1\| \, \|Av_2\| \quad (\text{By Cauchy Schwartz}) \\
&\leq \|v_1\| \, \lambda \, \|v_2\| \quad (\text{since } \langle v_2, u\rangle = 0) \\
&\leq \lambda \sqrt{|S|}\sqrt{|T|}
\end{aligned}$$

□

A couple remarks:

1. Note that the above theorem is applicable even if S and T intersect. In fact, both S and T can be the same set as well: in this case, we get a good estimate on the number of edges present inside a given set.
2. If we set $S = T$ to be an ϵ sized set, we can infer that most of the edges adjacent to S cross the set S. This is a very useful expansion property.

4 Application of expanders to derandomization

Suppose that there is a randomized algorithm for a language L using n bits such that

1. If a string $x \in L$, then the algorithm accepts with probability 1.
2. If a string $x \notin L$, then the algorithm rejects with probability at least $\frac{1}{2}$.

Our goal is to reduce the error probability $\frac{1}{2}$ of the algorithm. A natural idea is to repeat the algorithm multiple times. Suppose that we repeat the algorithm t times. Then, the error probability goes down to $\frac{1}{2^t}$, which is great, but the number of random bits used by the algorithm is equal to nt. Can we get the same exponential reduction in error probability, but with smaller amount of randomness?

The idea is to cleverly reuse the randomness by making "correlated" choices, rather than picking each set of bits independently. One way to achieve this is by performing a random walk on an expander graph. Let G be an expander on 2^n vertices, and with degree d. We first start at a random vertex v in G, using n random bits. Next, we pick a random neighbor of v, say v_1. We need $\log d = O(1)$ bits for this. Next, we continue the random walk from v_1. We repeat this process till we pick t vertices overall. The overall random bits that we needed is equal to $n + O(t)$.

Note that each vertex in the graph corresponds to a set of n bits. We run the algorithm using all the t sets of random bits obtained by the random walk. If the algorithm outputs No for any of these choices of random bits, we output No, else we output Yes. We still accept all the strings $x \in L$ with probability 1. How do we argue about the probability that we accept incorrect strings? In order to do so, we need an estimate on the probability that a t-step random walk on expander graph stays inside a set of size $\frac{n}{2}$ through out the walk. Turns out this probability is also exponentially small in t! Using this, we obtained a "randomness efficient" algorithm that uses $n+O(t)$ random bits that still gets error probability exponentially small in t.

Now we prove the above mentioned bound:

Theorem 3. *(Random walk on expander graph) Let $G = (V, E)$ be an (n, d, λ)-expander. Let $B \subseteq V$ be a subset of vertices such that $|B| = (1 - \delta)n$. Then, the probability that a random walk $V_1, V_2, \ldots, V_t$ of $t - 1$ steps starting at a uniformly random vertex of G completely stays inside B is*

$$Pr(V_i \in B \ \forall i \in [t]) \leq \left(1 - \delta\left(1 - \frac{\lambda}{d}\right)\right)^{t-1}$$

Proof. Let P be the projection matrix that zeroes out all the indices not in B. It is a diagonal matrix such that

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

Let $A' = \frac{A}{d}$ be the random walk matrix of the graph G i.e. for a probability distribution μ on the vertices, $A'\mu$ gives the probability distribution after taking a single random walk step. Thus, the probability that all the t vertices of the random walk are in B is equal to

$$\Pr(V_i \in B \ \forall i \in [t]) = \left\|PA'PA' \ldots PA'Pu\right\|_1$$

Since $P^2 = P$, we can write the above as

$$\left\|(PA'P)^{t-1}Pu\right\|_1 \leq \sqrt{n}\left\|(PA'P)^{t-1}Pu\right\|$$

where we have used Cauchy-Schwartz inequality to relate ℓ_1 and ℓ_2 norms of a vector in $\mathbb{R}^n$. We now bound the largest absolute value of eigenvalues of $PA'P$: Let $\nu = \max_{\|x\|=1} x^T PA'Px$. Let $y = Px$.

We have $\nu = y^T A' y$. As before, we write $y = y^{\|} + y^{\perp}$, where $y^{\|} = \alpha u$, and $\langle y^{\perp}, y^{\|} \rangle = 0$. Note that $\left\|y^{\|}\right\|^2 + \left\|y^{\perp}\right\|^2 = \|y\|^2 \leq \|x\|^2 = 1$. Let $\mathbf{1} = (1, 1, \ldots, 1)$. We have $y^{\|} = \alpha u$, where $\alpha = \langle y, \mathbf{1} \rangle$. Thus, $\left\|y^{\|}\right\|^2 = \frac{1}{n}\langle y, \mathbf{1}\rangle^2 = \frac{1}{n}\langle y, P\mathbf{1}\rangle^2 \leq \frac{1}{n}\|y\|^2 \|P\mathbf{1}\|^2 = (1-\delta)\|y\|^2$.

We now bound ν:

$$\begin{aligned}
\nu &= y^T A' y \\
&= \left((y^{\|})^T + (y^{\perp})^T\right) A' (y^{\|} + y^{\perp}) \\
&\leq \left\|y^{\|}\right\|^2 + \frac{\lambda}{d}\left\|y^{\perp}\right\|^2 \\
&= \left\|y^{\|}\right\|^2 + \frac{\lambda}{d}\left(\|y\|^2 - \left\|y^{\|}\right\|^2\right) \\
&= \left(1 - \frac{\lambda}{d}\right)\left\|y^{\|}\right\|^2 + \frac{\lambda}{d}\|y\|^2 \\
&\leq \|y\|^2 \left(\left(1 - \frac{\lambda}{d}\right)(1-\delta) + \frac{\lambda}{d}\right) \\
&= \|y\|^2 \left(1 - \delta\left(1 - \frac{\lambda}{d}\right)\right) \\
&\leq 1 - \delta\left(1 - \frac{\lambda}{d}\right)
\end{aligned}$$

Substituting it in the above probability, we get

$$\begin{aligned}
\Pr(V_i \in B \ \forall i \in [t]) &\leq \sqrt{n}\left\|(PA'P)^{t-1}Pu\right\| \\
&\leq \sqrt{n}\left(1 - \delta\left(1 - \frac{\lambda}{d}\right)\right)^{t-1}\|Pu\| \\
&= \sqrt{n}\sqrt{\frac{1-\delta}{n}}\left(1 - \delta\left(1 - \frac{\lambda}{d}\right)\right)^{t-1} \\
&\leq \left(1 - \delta\left(1 - \frac{\lambda}{d}\right)\right)^{t-1}
\end{aligned}$$

□