CSCI-B609: A Theorist's Toolkit, Fall 2016 Oct 25

Lecture 17: Expander Graphs

Lecturer: Yuan Zhou *Scribe: Xuan Dong*

In this lecture, we will talk about expander graphs. Expander graphs are graphs with the special property: sparse yet very highly-connected, i.e., any set of vertices S (unless very large) has a number of outgoing edges proportional to $|S|$. Expanders are very useful in computer science. We will mention some applications later.

1 Overview of Expander Graphs

Let $G = (V, E)$ be an undirected d-regular graph, here, $|V| = n$, $\deg(u) = d$ for all $u \in V$. We will typically interpret the properties of expander graphs in an asymptotic sense. That is, there will be an infinite family of graphs G, with a growing number of vertices n. By "sparse", we mean that the degree d of G should be very slowly growing as a function of n. When n goes to infinity ($n \to \infty$), d is thought as a constant, so the graph automatically becomes sparse as n grows large, since the number of edges $|E| = \frac{d}{2}n \sim O(n)$. The "highly-connected" property has a variety of different interpretations, like in terms of edge expansion, vertex expansion or spectral expansion. Informally, we define a graph is well-connectness that for every $S \subseteq V$, S is a not-too-large set of vertices (say $0 < |S| \le \frac{n}{2}$) but has lots of (say $\Omega(|S|)$) edges for edge expansion or vertices for vertex expansion on its boundary.

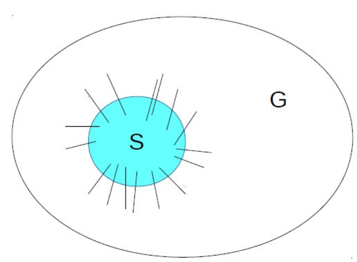

Figure 1: An example of highly-connected expander graph. The lines coming from S represent edges from vertices in S to those in $\bar{S}$.

More formally, let us give a few definitions for an expander family.

Definition 1. *A d-regular graph is an (α, ε)-edge expander if for $\forall S \subseteq V$ and $|S| \le \alpha n$,* $\text{edges}(S, \bar{S}) \ge \varepsilon \cdot |S| \cdot d$.

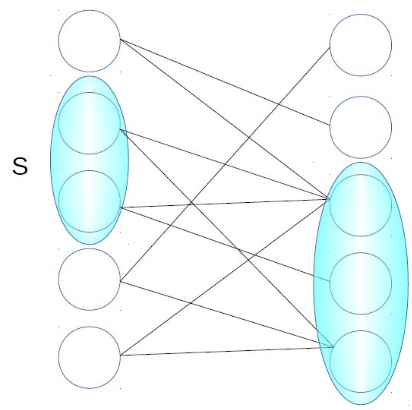

Figure 2: An example of a d-left regular graph where $d = 2$ and $n = 5$.

Definition 2. *A d-regular graph is an (α, A)-vertex expander if for $\forall S \subseteq V$ and $|S| \leq \alpha n$, $|\Gamma(S)| \geq A \cdot |S|$.*

Definition 3. *A d-left regular bipartite graph $G = (u, v, E)$, which is a graph $G = (V, E)$ where V is partitioned into two disjoint sets left-side u and right-side v, is an (α, A)-left vertex expander if for all sets S of left-vertices of size at most $\alpha|u|$, $|\Gamma(S)| \geq A \cdot |S|$.*

Ideally, we would like our A be as close to d as possible. Graphs with A comparable to d are very good expanders.

By Cheeger's inequality, where λ_2 is the second-smallest eigenvalue of the normalized Laplacian matrix of G, L_G, there is a fact relating edge expansion to the second eigenvalue:

$$\min_{0<|S|\leq \frac{n}{2}} \{\frac{\text{edges}(S, \bar{S})}{d|S|}\} \geq \frac{1}{2}\lambda_2(G) \Longleftrightarrow G \text{ is always a } \left(\frac{1}{2}, \frac{1}{2}\lambda_2(L_G)\right) \text{-edge expander.}$$

Now let us explore the connection between expanders and the spectrum of the graph. There is a connection between the expansion of a graph and the eigengap (or spectral gap) of the normalized adjacency matrix, it turns to a theorem on the spectral expansion of random graphs.

Theorem 4 (Alon'1986 [1], Friedman'2003 [6])**.** *Let G be a random n-vertex d-regular graph with normalized adjacency matrix A_G, for any constant $\varepsilon > 0$ and $\forall i \geq 2$, we have*

$$\Pr[|\lambda_i(A_G)| \leq 2\sqrt{d-1} + \varepsilon] = 1 - o(1).$$

Corollary 5. *With probability $1 - o(1)$, a random n-vertex d-regular graph ($d \geq 3$) is a $(\frac{1}{2}, \Omega(1))$-edge expander.*

We will show the existence of bipartite expanders.

Theorem 6. *Let $Bip_{n,d}$ be the uniform distribution on all simple $n \times n$ d-left regular bipartite graphs with partite sets L and R of cardinality n. For all d, there exists $\alpha = \frac{1}{e^3 d^4}$ for*

$$\Pr[G \textit{ is an } (\alpha, d-2)\textit{-left vertex expander }] \geq \frac{1}{2},$$

where G is chosen uniformly at random from $Bip_{n,d}$.

Proof. To choose G in $Bip_{n,d}$ uniformly at random, we choose d (not necessarily distinct) neighbors for each vertex L at random. For $k \leq \alpha n$, let

$$p_k = \Pr[\exists S \subseteq L \text{ such that } |S| = k, |\Gamma(S)| < (D-2)|S|].$$

Thus p_k is the probability that there exists a left-set L of size exactly k that does not expand by at least $d-2$. To prove the theorem, it suffices to show that $\sum_k p_k \leq 1/2$.

If $S \subseteq L$ has cardinality k, then the total number of neighbors of vertices in S, counted with multiplicity, is dk. We can imagine these vertices in $\Gamma(S)$ be $v_1, v_2, \cdots, v_{kd}$ being chosen in sequence. Call v_i a repeat if $v_i \in \{v_1, \cdots, v_{i-1}\}$. Then the probability that v_i is a repeat is

$$\Pr[v_i \text{ is a repeat}] \leq \frac{i-1}{n} \leq \frac{kd}{n}.$$

So if $|\Gamma(S)| < (d-2)k$, then there must be $2k$ repeats among the neighbors of vertices in S. We can compute this probability:

$$\begin{aligned} &\Pr[|\Gamma(S)| < (D-2)|S|] \\ \leq &\Pr[\text{There are at least } 2|S| = 2k \text{ repeats among } v_1, \cdots, v_{kd}] \\ \leq &\binom{kd}{2k}\left(\frac{kd}{n}\right)^{2k}. \end{aligned}$$

Here, the binomial coefficient represents the number of ways to choose $2k$ neighbors to be repeats, and the fraction kd/n represents an upper bound on the probability that any given choice of a neighbor is a repeat. That this is an upper bound follows from the union bound. Since there are $\binom{n}{k}$ possibilities for S, we have

$$\begin{aligned} p_k &\leq \binom{n}{k}\binom{kd}{2k}\left(\frac{kd}{n}\right)^{2k} \\ &\leq \left(\frac{en}{k}\right)^k \left(\frac{ekd}{2k}\right)^{2k} \left(\frac{kd}{n}\right)^{2k} \\ &= \left(\frac{e^3 d^4 k}{4n}\right)^k, \end{aligned}$$

where e is the base of the natural logarithm. Since $k \leq \alpha n$, we can set $k = \frac{1}{e^3 d^4} n$ to obtain $p_k \leq 4^{-k}$. Therefore,

$$\Pr_{G \sim \mathrm{Bip}_{n,d}}[G \text{ is not an } (\alpha, D-2)\text{-left vertex expander}] \leq \sum_{k=1}^{\alpha n} (\frac{1}{4})^k < \frac{1}{2}.$$

This completes the proof. □

Using similar probabilistic method, we can prove:

Theorem 7 (Bassalygo'1981 [2])**.** *Let G be a random $n \times \frac{3}{4} n$ d-regular bipartite graph, for $d \geq 64$, we have*

$$\Pr[G \textit{ is a } (\frac{.02}{d}, .8d)\textit{-left vertex expander}] = \Omega(1).$$

2 Application in Coding Theory

Now we introduce how to obtain an asymptotically good code, the Tanner code [7], which is obtained from the expander graphs, and it will have the following properties:

- positive constant rate,
- positive constant minimum distance,
- efficient to decode and encode.

Let $G = (L, R, E)$ be an instance of Theorem 7 with $d = 64$, $|L| = n$, $|R| = \frac{3}{4} n$. Thus, for all $S \subseteq L$, $|S| \leq \frac{.02}{64} n$, we have $|\Gamma(S)| \geq .8 \times 64|S|$.

Claim 1. *If $S \subseteq L$, and $0 \neq |S| \leq \frac{.02}{64} n$, there exists a $v \in \Gamma(S)$ so that v has exactly one neighbor in S.*

Proof 1. *Suppose for contraction that for all $v \in \Gamma(S)$, v has ≥ 2 neighbors in S. We have*

$$\mathrm{edges}(S, \Gamma(S)) \geq 2|\Gamma(S)| \geq 2 \times .8 \times 64|S| > 64|S|.$$

This is a contradiction, since the left partition is 64-regular, so $\mathrm{edges}(S, \Gamma(S)) \leq 64|S|$*, which means there exists a $v \in \Gamma(S)$ so that v has exactly one neighbor in S.*

Now consider the Tanner code defined using the adjacent matrix of G as the parity check matrix H, i.e., $H \in \mathbb{F}_2^{|R| \times |L|}$ where each node in L corresponds to a bit in the codeword, and each node in R corresponds to a parity check.

The dimension of the code $k = |L| - |R| = n = \frac{3}{4} n = \frac{1}{4} n$. The rate is $\frac{1}{4}$, which is a constant, as promised.

Claim 2. *The minimum distance of the code is greater than* $\frac{.02}{64}n$ *(assume* $\frac{.02}{64}n$ *is an integer), and the relative minimum distance is greater than* $\frac{.02}{64}$*. Note that this is also an constant, as desired.*

Proof 2. *Assume for contradiction that the minimum distance* d *is at most* $\frac{.02}{64}n$*. Thus, there exists a non-zero codeword* $z \in \mathbb{F}_2^{|L|}$ *with Hamming weight* $\mathrm{wt}(z) \leq \frac{.02}{64}n$*, and* $Hz = \vec{0}$*. Let* $S = \{u \in L : z_u = 1\}$*. Since* $z \neq \vec{0}$*, we know that* $S \neq \emptyset$*, and the Hamming weight of* z *implies* $|S| \leq \frac{.02}{64}n$*. Since* $Hz = \vec{0}$*, each* $v \in R$ *which corresponds to a parity check must have even number of neighbors in* S*. This is particular true for each* $v \in \Gamma(S) \leq |R|$*. Therefore, for all* $v \in \Gamma(S)$*,* v *has* ≥ 2 *neighbors in* S*, which contracts to Claim 1.*

Decoding the Tanner codes is efficient. It critically uses the fact $2 \times .8 - 1 > .5$ (i.e., $.8 > .75$). Let $y \in \mathbb{F}_2^{|L|}$ be the received message. Call a parity check node $v \in R$ satisfied if $\sum_{u \in \Gamma(v)} y_u = 0$. The algorithm used here is simple. Where for $\exists u \in L$, there are more unsatisfied check nodes than satisfied nodes in $\Gamma(u)$, then flip y_u.

Claim 3. *If the number of errors in* y *is at least* 1 *and at most* $\frac{.02}{64}n$*, then there exists* $u \in L$*, there are more than* $d/2 = 32$ *unsatisfied nodes in* $\Gamma(u)$*.*

Proof 3. *Let* $T \neq \emptyset$ *be the error locations. Since* $|T| \leq \frac{.02}{64}n$*,* $|\Gamma(T)| \geq .8 \times 64|T|$*. Now let* $u(T) \subseteq \Gamma(T)$ *be set of nodes having exactly* 1 *neighbor in* T*, we have*

$$\begin{aligned} 2 \times |\Gamma(T) - u(T)| + |u(T)| &\leq d|T| \\ |u(T)| &\geq 2\Gamma(T) - d|T| \\ &\geq 2 \times .8 \times 64|T| - 64|T| \\ &= .6 \times 64|T| \\ &> 32|T|. \end{aligned}$$

Clearly there are more than 32 *check nodes in* $u(T)$ *are not satisfied. Proved.*

The correctness and the running time of the algorithm will be introduced in next lecture.

References

[1] Alon, Noga. Eigenvalues and expanders. *Combinatorica* 6.2 (1986): 83-96.

[2] Bassalygo, Leonid Alexandrovich. Asymptotically optimal switching circuits. *Problemy Peredachi Informatsii* 17.3 (1981): 81-88.

[3] Cynthia Dwork. Eigenvalues and Expanders. *Lecture notes.* `http://www.tcs.tifr.res.in/~prahladh/teaching/05spring/lectures/lec2.pdf`

[4] Expander Graphs. Harvard University. *Lecture notes.* `http://people.seas.harvard.edu/~salil/pseudorandomness/expanders.pdf`

[5] Ryan O'Donnell. Expanders. *Lecture notes.* `http://www.cs.cmu.edu/~odonnell/toolkit13/lecture12.pdf`

[6] Friedman, Joel. A proof of alon's second eigenvalue conjecture. *Proceedings of the thirty-fifth annual ACM symposium on Theory of computing.* ACM, 2003.

[7] Tanner, R. A recursive approach to low complexity codes. *IEEE Transactions on information theory* 27.5 (1981): 533-547.