

## Lecture 7: Expander Graphs

February 22, 2007

Based on scribe notes by Kartik Venkatram and Mihai Pătraşcu.

## 1 Expander Graphs

Now that we have seen a variety of basic derandomization techniques, we will move on to study the first major “pseudorandom object” in the course, *expander graphs*. These are graphs that are “sparse” yet very “well-connected.” We will typically interpret these properties in an asymptotic sense. That is, there will be an infinite family of graphs  $G_i$ , with a growing number of vertices  $N_i$ . By *sparse*, we mean that degree  $D_i$  of  $G_i$  should be very slowly growing as a function of  $N_i$ . The “well-connected” property has a variety of different interpretations, which we will discuss below. Typically, we will drop the subscripts of  $i$  and the fact that we are talking about an infinite family of graphs will be implicit in our theorems.

The classic measure of well-connectedness in expanders follows:

**Definition 1** *A graph  $G$  is a  $(K, A)$  vertex expander if for all sets  $S$  of at most  $K$  vertices, the neighborhood  $N(S) = \{u \mid \exists v \in S \text{ s.t. } (u, v) \in E\}$  is of size at least  $A \cdot |S|$ .*

Ideally, we would like  $D = O(1)$ ,  $K = \Omega(N)$ , and  $A$  as close to  $D$  as possible.

There are several other measures of expansion, some of which we will examine in forthcoming lectures:

- Edge expansion (cuts): instead of  $N(S)$ , use the number of edges leaving  $S$ .
- Random walks: random walks converge quickly to uniform distribution, i.e. if the second eigenvalue  $\lambda_2(G)$  is small.
- “Well-mixed edges”: for every two sets  $S$  and  $T$  (say of constant density), the fraction of edges between  $S$  and  $T$  is roughly the product of their densities.

All of these measures are very closely related to each other, and are even equivalent for certain settings of parameters.

It is not obvious from the definition that good expanders (say, with  $D = O(1)$ ,  $K = \Omega(N)$ , and  $A = 1 + \Omega(1)$ ) even exist. We will show this using the *probabilistic method*.

**Theorem 2** *For all constants  $D \geq 3$ , there is a constant  $\alpha > 0$  such that for all sufficiently large  $N$ , a random  $D$ -regular graph on  $N$  vertices is an  $(\alpha N, D - 1.01)$  vertex expander with probability at least  $1/2$ .*

Note that the degree bound of 3 is the smallest possible, as every graph of degree 2 is a poor expander (being a union of cycles and chains).

We prove a slightly simpler theorem for *bipartite expanders*.

**Definition 3** A bipartite graph  $G$  is a  $(K, A)$  vertex expander if for all sets  $S$  of size at most  $K$  left-vertices, the neighborhood  $N(S)$  is of size at least  $A \cdot |S|$ .

Now, let  $\text{Bip}_{N,D}$  be the set of bipartite multi-graphs that have  $N$  vertices on each side and  $D$ -regular on the left.

**Theorem 4** For every constant  $D$ , there exists a constant  $\alpha > 0$ , such that for all sufficiently large  $N$ , a uniformly random graph from  $\text{Bip}_{N,D}$  is an  $(\alpha N, D - 2)$  vertex expander with probability at least  $1/2$ .

**Proof:** First, note that choosing  $G \stackrel{R}{\leftarrow} \text{Bip}_{N,D}$  is equivalent to uniformly and independently choosing  $D$  neighbors on the right for each left vertex  $v$ . Now, for  $K \leq \alpha N$ , let  $p_K$  be the probability that there exists a left-set  $S$  of size exactly  $K$  that does not expand by  $D - 2$ . Fixing a subset  $S$  of size  $K$ ,  $N(S)$  is a set of  $KD$  random vertices in  $R$  (chosen with replacement). We can imagine these vertices  $V_1, V_2, \dots, V_{KD}$  being chosen in sequence. Call  $V_i$  a *repeat* if  $V_i \in \{V_1, \dots, V_{i-1}\}$ . Then the probability that  $V_i$  is a repeat, even conditioned on  $V_1, \dots, V_{i-1}$ , is at most  $(i - 1)/N \leq KD/N$ .

Observe that

$$\Pr[|N(S)| \leq (D-2) \cdot K] \leq \Pr[\text{there are at least } 2K \text{ repeats among } V_1, \dots, V_{KD}] \leq \binom{KD}{2K} \left(\frac{KD}{N}\right)^{2K}.$$

Thus, we find that

$$p_K \leq \binom{N}{K} \binom{KD}{2K} \left(\frac{KD}{N}\right)^{2K} \leq \left(\frac{Ne}{K}\right)^K \left(\frac{KDe}{2K}\right)^K \left(\frac{KD}{N}\right)^{2K} = \left(\frac{cD^3K}{N}\right)^K \quad (1)$$

where  $e$  is the base of the natural logarithm and  $c = e^2/2$ . Since  $K \leq \alpha N$ , we can set  $\alpha = \frac{1}{4cD^3}$  to obtain  $p_K \leq 4^{-K}$ . Thus

$$\Pr_{G \in \mathcal{G}_D} [G \text{ is not an } (\alpha N, D - 2) \text{ expander}] \leq \sum_{K=1}^{\lfloor \alpha N \rfloor} 4^{-K} < \frac{1}{2} \quad (2)$$

■

There are a number of variants to the above probabilistic construction of expanders.

- We can obtain a  $D$ -regular multigraph by taking the union of  $D$  random perfect matchings. This can be analyzed using a small modification of the analysis above; even though  $V_1, \dots, V_{KD}$  are not independent, the probability of a  $V_i$  being a repeat conditioned on  $V_1, \dots, V_{i-1}$  can still be bounded by  $KD/(N - K)$ . Also, the multiple edges in the resulting graph can be eliminated or redistributed to obtain a simple graph that is at least a good an expander.

- One can optimize  $\alpha$  rather than the expansion factor  $A$ , showing that for all constants  $\alpha < 1$  and  $D > 2$ , there exists a constant  $A > 1$  such that for all sufficiently large  $N$ , a random graph in  $\text{Bip}_{N,D}$  is an  $(\alpha N, A)$  vertex expander with high probability.
- In fact, a very general tradeoff between  $D$ ,  $\alpha$ , and  $A$  is known: a random  $D$ -regular  $N$ -vertex graph is an  $(\alpha N, A)$  vertex expander with high probability if  $D > \frac{H(\alpha) + H(\alpha A)}{H(\alpha) - \alpha A H(1/A)}$ , where  $H(p) = p \log(1/p) + (1 - p) \log(1/(1 - p))$  is the binary entropy function.
- The results can also be extended to unbalanced bipartite graphs (where the right side is smaller than the left), and non-bipartite graphs as well, and both of these cases are important in some applications.

In addition to being a natural combinatorial object, expander graphs have numerous applications in theoretical computer science, including the construction of fault-tolerant networks (indeed, the first papers on expanders were in conferences on telephone networks), sorting in  $O(\log n)$  time in parallel, derandomization (as we will see), lower bounds in circuit complexity and proof complexity, error-correcting codes, negative results regarding integrality gaps for linear programming relaxations and metric embeddings, distributed routing, data structures. For many of these applications, it is not enough to know that a random graph is a good expander — we need *explicit* constructions. That is, constructions that are deterministic and efficient. We view explicit expanders as ‘pseudorandom objects’ because they are fixed graphs that possess many of the properties of random graphs.

## 2 Spectral Expansion

Intuitively, another way of saying that a graph is well-connected is to require that random walks on the graph converge quickly to the stationary distribution. The mixing rate of random walks in turn is captured well by the second largest eigenvalue of the transition matrix (as we have seen in Lecture 4), and this turns out to be a very useful measure of expansion.

Recall that for a regular directed graph  $G$  with random-walk matrix  $M$ , we define

$$\lambda(G) \stackrel{\text{def}}{=} \max_{\pi} \frac{\|\pi M - u\|}{\|\pi - u\|} = \max_{x \perp u} \frac{\|xM\|}{\|x\|},$$

where the first maximization is over all probability distributions  $\pi \in [0, 1]^n$  and the second is over all vectors  $x \in \mathbb{R}^n$  such that  $x \perp u$ . We write  $\gamma(G) \stackrel{\text{def}}{=} 1 - \lambda(G)$ .

**Definition 5** For  $\lambda \in [0, 1]$ , we say that  $G$  is a  $\lambda$  spectral expander if  $\lambda(G) \leq \lambda$ .

Smaller values of  $\lambda$  correspond to better expansion. Sometimes we will state results in terms of the spectral gap  $\gamma = 1 - \lambda$ .

Surprisingly, this linear-algebraic measure of expansion turns out to be equivalent to the combinatorial measure of vertex expansion for common parameters of interest.

One direction is given by the following:

**Theorem 6 (spectral expansion  $\Rightarrow$  vertex expansion)** *If  $G$  is a  $\lambda$  spectral expander for some  $\lambda \in [0, 1]$ , then, for every  $\alpha \in [0, 1]$ ,  $G$  is an  $\left(\alpha N, \frac{1}{(1-\alpha)\lambda^2 + \alpha}\right)$  vertex expander. In particular, for  $\alpha = 1/2$  and  $\lambda = 1 - \gamma$ , we conclude that  $G$  is a  $(N/2, 1 + \gamma)$  expander.*

Before proving this theorem, we introduce some properties of probability distributions.

**Definition 7** *For a probability distribution  $\pi$ , the collision probability of  $\pi$  is defined to be the probability that two independent samples from  $\pi$  are equal, namely  $\text{CP}(\pi) = \sum_x \pi_x^2$ .*

*The support of  $\pi$  is  $\text{Supp}(\pi) = \{x : \pi_x > 0\}$ .*

**Lemma 8** *For every probability distribution  $\pi \in [0, 1]^N$ , we have:*

1.  $\text{CP}(\pi) = \|\pi\|^2 = \|\pi - u\|^2 + 1/N$ .
2.  $\text{CP}(\pi) \geq 1/|\text{Supp}(\pi)|$ , with equality iff  $\pi$  is uniform on  $\text{Supp}(\pi)$ .

**Proof:** For Part 1, the fact that  $\text{CP}(\pi) = \|\pi\|^2$  follows immediately from the definition. Then, writing  $\pi = u + (\pi - u)$ , and noting that  $(\pi - u) \perp u$ , we have  $\|\pi\|^2 = \|u\|^2 + \|\pi - u\|^2 = 1/N + \|\pi - u\|^2$ .

For Part 2, by Cauchy-Schwarz we have

$$1 = \sum_{x \in \text{Supp}(\pi)} \pi_x \leq \sqrt{|\text{Supp}(\pi)|} \cdot \sqrt{\sum_x \pi_x^2}.$$

■

**Proof (of Theorem 6):** By the definition of spectral expansion and Part 1 of Lemma 8, we have

$$\text{CP}(\pi M) - \frac{1}{N} \leq \lambda^2 \cdot \left( \text{CP}(\pi) - \frac{1}{N} \right)$$

for every probability distribution  $\pi$ . Letting  $S$  be any subset of the vertices of size at most  $\alpha N$  and  $\pi$  the uniform distribution on  $S$ , we have  $\text{CP}(\pi) = 1/|S|$  and  $\text{CP}(\pi M) \geq 1/|\text{Supp}(\pi M)| = 1/|N(S)|$ , implying that

$$\left( \frac{1}{|N(S)|} - \frac{1}{N} \right) \leq \lambda^2 \cdot \left( \frac{1}{|S|} - \frac{1}{N} \right)$$

Solving for  $|N(S)|$  and using  $N \geq |S|/\alpha$ , we obtain  $|N(S)| \geq |S|/(\lambda^2(1 - \alpha) + \alpha)$ , as desired. ■

The other direction, i.e. obtaining spectral expansion from vertex expansion, is somewhat more difficult.

**Theorem 9 (vertex expansion  $\Rightarrow$  spectral expansion)** *For every  $\delta > 0$  and  $D > 0$ , there exists  $\gamma > 0$  such that if  $G$  is a  $D$ -regular  $(N/2, 1 + \delta)$  vertex expander, then it is also  $(1 - \gamma)$  spectral expander. Specifically, we can take  $\gamma = \Omega(\delta^2/D)$ .*

Note first the dependence on subset size being  $\frac{N}{2}$ : this is necessary, because a graph can have vertex expansion  $(\alpha N, \Omega(1))$  for  $\alpha < 1/2$  and be disconnected (eg the disjoint union of two good expanders), thereby having no spectral expansion. The other problem is that the bound on  $\gamma$  depends on  $D$ . This is also necessary, because adding edges to a good expander cannot hurt its vertex expansion, but can hurt its spectral expansion.

Still, roughly speaking, these two results show that vertex expansion and spectral expansion are closely related, and equivalent for many interesting settings of parameters. Indeed, when people informally use the term “expander,” they often mean a family of  $D$ -regular graphs for *constant* degree  $D = O(1)$  satisfying one of the following two equivalent conditions:

- Every graph in the family is a  $\lambda$  spectral expander for some constant  $\lambda < 1$ .
- Every graph in the family is an  $(N/2, 1 + \delta)$  expander for some constant  $\delta > 0$ .

However, these two measures are no longer equivalent if one wants to optimize the expansion constants. For vertex expansion, we have already seen that if we allow  $\alpha$  to be a small constant (depending on  $D$ ), then there exist  $(\alpha N, A)$  vertex expanders with  $A$  very close to  $D$ , e.g.  $A = D - 1.01$ , and clearly one cannot have  $A$  to be any larger than  $D$ . The optimal value for the spectral expansion is also well-understood. First note that, by taking  $\alpha \rightarrow 0$  in Theorem 6, a  $\lambda$  spectral expander has vertex expansion  $A \approx 1/\lambda^2$  for small sets. Thus, a lower bound on  $\lambda$  is  $1/\sqrt{D} - o(1)$ . In fact, this lower bound can be improved:

**Theorem 10** *Any infinite family of  $D$ -regular graphs with spectral expansion  $\lambda$  has  $\lambda \geq \frac{2\sqrt{D-1}}{D} - o(1)$ , where the additive term vanishes as  $|V(G)| \rightarrow \infty$ .*

Surprisingly, there exist explicit constructions giving  $\lambda < \frac{2\sqrt{D-1}}{D}$ . Graphs meeting this bound are called *Ramanujan graphs*. Random graphs almost match this bound, as well:

**Theorem 11** *For any constant  $D$  and any constant  $\epsilon > 0$ , a random  $D$ -regular graph has spectral expansion at most  $\frac{2\sqrt{D-1}}{D} + \epsilon$  with probability  $1 - \frac{1}{N^{\Omega(\sqrt{D})}}$ .*

Now let us see what these results for spectral expansion imply in the world of vertex expansion. With Ramanujan graphs ( $\lambda = \frac{2\sqrt{D-1}}{D}$ ), our bound from last time gives a vertex expansion of  $A \approx \frac{D}{4}$  (for small sets). This is not tight, and it is known that Ramanujan graphs actually have vertex expansion  $D/2$ , which is tight in the sense that there are families of graphs with  $\lambda \rightarrow \frac{2\sqrt{D-1}}{D}$  with vertex expansion at most  $D/2$ . Still, this is not as good as we showed by the probabilistic method, where we achieved vertex expansion  $D - O(1)$ . This means that we cannot obtain optimal vertex expansion by going through spectral expansion. Similarly, we cannot obtain optimal spectral expansion by going through vertex expansion. The conclusion is that vertex and spectral expansion are loosely equivalent, but only if we are not interested in optimizing the constants in the tradeoffs between various parameters (and for some applications this is crucial).

### 3 Other Measures of Expansion

In this section, we mention two other useful measures of expansion involving edges crossing cuts in the graph. For two sets  $S, T \subset V(G)$ , let  $e(S, T) = \{(u, v) \in S \times T \mid \{u, v\} \in E\}$ , where  $(u, v)$  is

interpreted as an ordered pair. Here  $S$  and  $T$  may not be disjoint, in which case some edges may be counted twice, corresponding to both orientations.

**Definition 12** A  $D$ -regular graph  $G$  is a  $(K, \varepsilon)$  edge expander if for all sets  $S$  of at most  $K$  vertices, the cut  $e(S, \overline{S})$  is of size at least  $\varepsilon \cdot |S| \cdot D$ .

That is, at least an  $\varepsilon$  fraction of the edges from  $S$  lead outside  $S$ . (Sometimes edge expansion is defined without the normalization factor of  $S$ , only requiring  $|e(S, \overline{S})| \geq \varepsilon \cdot |S|$ .) When viewed in terms of stationary distribution of the random walk on  $G$ , the ratio  $e(S, \overline{S})/(|S| \cdot D)$  is the probability that, conditioned on being in  $S$ , the random walk leaves  $S$  in one step. It turns out that, if we let  $\varepsilon$  be the minimum of this quantity over all sets  $S$  of density at most  $1/2$ , i.e.  $\varepsilon$  is the largest value such that  $G$  is an  $(N/2, \varepsilon)$  edge expander, then this turns out to be even more closely related to spectral expansion than vertex expansion. Indeed, it is known that the spectral gap  $\gamma$  satisfies  $2\varepsilon \geq \gamma \geq \varepsilon^2/2$ . The intuition is that a large edge expansion  $\varepsilon$  implies that the random walk on the graph has no “bottlenecks” and thus should mix rapidly. This connection also holds for Markov chains in general (when the definitions are appropriately generalized), where the edge expansion is known as the *conductance*.

There is yet another measure of expansion in terms of edges not just from a set  $S$  to its complement but between any two sets  $S$  and  $T$ . If we think of an expander as being like a random graph, we would expect the fraction of graph edges that are in  $e(S, T)$  to be proportional to the fraction of nodes that are in  $S$  and  $T$ . The following result shows that this intuition is correct:

**Lemma 13 (Expander Mixing Lemma)** Let  $G$  be a  $D$ -regular  $\lambda$  spectral expander on  $N$  vertices. Then for all sets of vertices  $S, T$ , we have

$$\left| \frac{e(S, T)}{N \cdot D} - \mu(S)\mu(T) \right| \leq \lambda \sqrt{\mu(S)\mu(T)},$$

where  $\mu(R) = |R|/N$  for any set  $R$ .

Observe that the denominator  $N \cdot D$  counts all edges of the graph (as ordered pairs). The lemma states that the difference between the fraction of edges in  $e(S, T)$  and the expected value if  $G, S$  and  $T$  would be chosen randomly is “small”, roughly the square root of this fraction. The error term should not be surprising, since the theorem holds for *all* sets  $S$  and  $T$ , so we would expect a standard deviation. Finally, observe that the error term is at most  $\lambda$ , which is an interesting result if  $\lambda$  is small.

**Proof:** Let  $\chi_S$  be the characteristic vector of  $S$  and  $\chi_T$  the characteristic vector of  $T$ . Let  $A$  be the adjacency matrix of  $G$ , and  $M = A/D$  be the random-walk matrix for  $G$ . Note that  $e(S, T) = \chi_S^t A \chi_T = \chi_S^t (DM) \chi_T$ .

Let  $\alpha = \mu(S)$  and  $\beta = \mu(T)$ . As usual, we can express  $\chi_S$  as the sum of two components, one parallel to the uniform distribution  $u$ , and the other a vector  $\chi_S^\perp$ , where  $\chi_S^\perp \perp u$ . The coefficient of  $u$  should be  $\frac{\chi_S \cdot u}{\|u\|^2} = \sum_i (\chi_S)_i = |S| = \alpha N$ . Then  $\chi_S = (\alpha N)u + \chi_S^\perp$  and similarly  $\chi_T = (\beta N)u + \chi_T^\perp$ . Intuitively,  $\chi_S^\perp$  and  $\chi_T^\perp$  give the error term, while the component parallel to the uniform distribution “spreads” the weight of  $S$  and  $T$  uniformly over the entire graph.

Formally, we have

$$\begin{aligned}\frac{e(S, T)}{D \cdot N} &= \frac{1}{N}((\alpha N)u + \chi_S^\perp)^t M((\beta N)u + \chi_T^\perp) \\ &= \frac{1}{N}(\alpha\beta N^2)u^t M u + \frac{1}{N}(\alpha N)u^t M \chi_T^\perp + \frac{1}{N}(\beta N)(\chi_S^\perp)^t M u + (\chi_S^\perp)^t M \chi_T^\perp.\end{aligned}$$

Since the component of  $\chi_T^\perp$  along  $u$  is zero, and  $u$  is an eigenvector of  $A$ ,  $M\chi_T^\perp$  also has a component of 0 along  $u$ , so  $u^t M \chi_T^\perp = 0$ . In addition,  $(\chi_S^\perp)^t M u = \chi_S^\perp \cdot u = 0$ . Our expression thus simplifies to:

$$(\alpha\beta N)u \cdot u + \frac{(\chi_S^\perp)^t M \chi_T^\perp}{N} = \alpha\beta + \frac{\chi_S^\perp \cdot (M\chi_T^\perp)}{N}.$$

Thus,

$$\begin{aligned}\left| \frac{e(S, T)}{N \cdot D} - \mu(S)\mu(T) \right| &= \frac{\chi_S^\perp \cdot (M\chi_T^\perp)}{N} \\ &\leq \frac{1}{N} \|\chi_S^\perp\| \cdot \|M\chi_T^\perp\| \\ &\leq \frac{1}{N} \|\chi_S^\perp\| \cdot \lambda \|\chi_T^\perp\| \\ &\leq \frac{\lambda}{N} \|\chi_S\| \cdot \|\chi_T\| \\ &= \frac{\lambda\sqrt{\alpha N}\sqrt{\beta N}}{N} = \lambda\sqrt{\alpha\beta},\end{aligned}$$

completing the proof. ■

A natural question is whether having such “well-mixed” edges characterizes a good expander. This is indeed true:

**Theorem 14 (Converse to Expander Mixing Lemma)** *Let  $G$  be a  $D$ -regular graph. If for all  $S, T \subset V(G)$ ,  $\left| \frac{e(S, T)}{N \cdot D} - \mu(S)\mu(T) \right| \leq \theta \sqrt{\mu(S)\mu(T)}$ , for some fixed  $\theta$ , then  $G$  is a  $\lambda$  spectral expander for  $\lambda = O\left(\theta \log \frac{D}{\theta}\right)$ .*

Putting the two theorems together, we see that  $\lambda$  and  $\theta$  differ by at most an  $O(\log D)$  factor, and in particular  $\lambda = 1/D^{\Omega(1)}$  iff  $\theta = 1/D^{\Omega(1)}$  (for constant). Thus, unlike the other connections we have seen, this connection is good for highly expanding graphs (as opposed to the case  $\lambda$  slightly less than 1). Unlike the case of vertex expansion vs. spectral expansion, here it is not known whether the dependence of the relationship on the degree  $D$  is necessary.

We will sketch the idea of the proof. It is known that  $\lambda_2(A) = \max_{x, y \perp u} \frac{x^t A y}{\|x\| \cdot \|y\|}$ . The proof relates this to the maximum taken when  $x, y \in \{\pm 1\}^N$ , and bounds this based on the relation from the hypothesis, for appropriate  $S$  and  $T$ .