

## Lecture 16

Prof. Nick Harvey

University of British Columbia

Today we will make a brief excursion into the fascinating world of expander graphs. It is hard to overstate the importance of these objects to theoretical computer science. They play an important role in some astonishing algorithms, the PCP theorem, derandomization, circuit complexity, etc. Today we will show how expanders are used in designing special routing networks called **superconcentrators**.

## 1 Expanders

The easiest one-sentence description of an expander graph is: *a graph that looks in many important ways like a random graph*. So, although expander graphs are often viewed as a difficult field of study, in many ways they are very simple. In particular, here's an easy one-sentence construction of an expander graph: *a random graph is very likely to be a good expander graph*. Proving that sentence is not difficult, and we will do so today.

Beyond the basics, the study of expanders becomes quite non-trivial. For example, in many scenarios we need a *deterministic* algorithm to construct an expander; such constructions often involve very sophisticated mathematics. Alternatively, we might like to give very precise guarantees on the quality of randomly generated expanders. That also gets very difficult.

But today we will be content with a simple, crude, randomized construction. This suffices for many amazing applications.

**The Definition.** Let  $G = (L, R, E)$  be a bipartite graph, where  $L$  is the set of left-vertices,  $R$  is the set of right-vertices, and  $E$  is the set of edges (each of which have one endpoint in  $L$  and one in  $R$ ). For any subset  $S \subseteq L$ , let

$$\Gamma(S) = \{ v \in R : \{u, v\} \in E \text{ for some } u \in S \}.$$

There are several different and interrelated definitions of expanders. We will say that  $G$  is  $(n, m, d)$ -**expander** if  $|L| = n$ ,  $|R| = m$ , every vertex in  $L$  has degree  $d$ , and

$$|\Gamma(S)| \geq |S| \quad \forall S \subseteq L \text{ s.t. } |S| \leq \frac{n}{2}. \quad (1)$$

**Theorem 1** *For some sufficiently large constant  $d$ , sufficiently large  $n$  and  $m \geq \frac{7}{8}n$ , there exists an  $(n, m, d)$ -expander.*

**Remark.** The theorem is trivial when  $m \geq n$  but it is interesting even in the case  $m = 0.999n$ .

**PROOF:** We generate the graph as follows. First we generate  $n$  vertices to form the set  $L$ , and  $m$  vertices to form the set  $R$ . Next, for every vertex  $u \in L$ , we randomly choose exactly  $d$  vertices from  $R$  (with repetition), and we add edges from  $u$  to all of those vertices. (So technically the resulting graph will be a multigraph, but this is not a problem.)

Let  $S \subseteq L$  be an arbitrary subset with  $s := |S| \leq \frac{n}{2}$ . We need to show that  $|\Gamma(S)| \geq |S|$ . We do this in a very naive way. We simply consider every set  $T \subseteq R$  with  $|T| < |S|$  and show that it is unlikely that

$\Gamma(S) \subseteq T$ . That probability is easy to analyze:

$$\Pr[\Gamma(S) \subseteq T] = \left(\frac{|T|}{m}\right)^{|S|d}.$$

Observe that we don't need to consider all sets  $T$  with  $|T| < |S|$ , it suffices to consider those with  $|T| = |S| - 1$ . (In fact, we will simplify our notation by considering the negligibly harder case of  $|T| = |S|$ .) As long as  $\Gamma(S)$  is not contained in any such  $T$ , then we have  $|\Gamma(S)| \geq |S|$  as required.

The remainder of the proof simply does a union bound over all such  $S$  and  $T$ .

$$\begin{aligned} \Pr[G \text{ is not a } (n, m, d)\text{-expander}] &\leq \sum_{\substack{S \subseteq L \\ |S| \leq n/2}} \sum_{\substack{T \subseteq R \\ |T|=|S|}} \Pr[\Gamma(S) \subseteq T] \\ &\leq \sum_{s \leq n/2} \binom{n}{s} \binom{m}{s} (s/m)^{sd} \\ &\leq \sum_{s \leq n/2} \left(\frac{ne}{s}\right)^s \left(\frac{me}{s}\right)^s \left(\frac{s}{m}\right)^{sd} \\ &\leq \sum_{s \leq n/2} \left(\frac{ne}{m} \frac{me}{m} \frac{s^{d-2}}{m^{d-2}}\right)^s \\ &\leq \sum_{s \leq n/2} \left(\frac{ne^2}{(7/8)n} \frac{(n/2)^{d-2}}{((7/8)n)^{d-2}}\right)^s \\ &\leq \sum_{s \leq n/2} \left(\frac{8e^2}{7} (4/7)^{d-2}\right)^s \\ &\leq \sum_{s=1}^{\infty} \left(8.5 \cdot 0.58^{d-2}\right)^s. \end{aligned}$$

For  $d \geq 9$ , the base of the exponent is less than 0.19, and so the infinite series adds up to less than  $\frac{0.19}{1-0.19} < 0.25$ .

So the probability that the random graph fails to satisfy the second inequality in (1) is less than 1/4.  $\square$

## 2 Superconcentrators

Let  $G = (V, E)$  be a graph. There are two disjoint subsets  $I \subseteq V$  and  $O \subseteq V$  with  $|I| = |O| = n$ . The vertices in  $I$  are called the **inputs** and the vertices in  $O$  are called the **outputs**. The graph  $G$  is called a **superconcentrator** if, for every  $k \leq n$  and every  $S \subseteq I$  and  $T \subseteq O$  with  $|S| = |T| = k$ , there are  $k$  vertex-disjoint paths from  $S$  to  $T$ .

It is important to note that we do not insist that the path starting at some  $s \in S$  has to arrive at a *particular*  $t \in T$ ; it is acceptable for it to arrive at an arbitrary  $t \in T$ . In other words, the paths from  $S$  to  $T$  induce some bijective mapping  $\pi : S \rightarrow T$  but we do not care what that mapping is.

Why might we be interested in such graphs? Suppose we want to design a telephone network with  $n$  input lines and  $n$  output lines such that any set of  $k$  input lines can be connected to any set of  $k$  output lines. To minimize the cost of this network, we would like the number of wires and routing points inside the network to be as small as possible.

It is easy to design a superconcentrator with  $O(n^2)$  edges: simply add an edge between every vertex in  $I$  and every vertex in  $O$ . Can one do better than that? It was conjectured by Valiant that  $\Omega(n \log n)$  edges are necessary. His conjecture came from an attempt to prove a lower bound in circuit complexity. But a short while later he realized that his own conjecture was false:

**Theorem 2** *For any  $n$ , there are superconcentrators with  $n$  inputs,  $n$  outputs, and  $O(n)$  edges.*

The proof uses expanders, recursion and Hall's theorem in a very clever way.

First we observe the following interesting property of expanders: *they have matchings covering any small set of vertices.*

**Claim 3** *Let  $H = (L, R, E)$  be a  $(n, 7n/8, 9)$ -expander. For every set  $S \subseteq L$  of size  $|S| \leq n/2$ , there is a **matching** in  $H$  covering  $S$ . In other words, there exists an injective map  $\pi_S : S \rightarrow R$  such that  $\{s, \pi_S(s)\} \in E$  for every  $s \in S$ .*

PROOF: [Hall's marriage theorem](#) says that  $H$  has a matching covering  $S$  if and only if every  $S' \subseteq S$  has  $|\Gamma(S')| \geq |S'|$ . That condition holds by (1), since  $H$  is an expander.  $\square$

**The Main Idea.** The construction proceeds as follows.

- First we create the input vertices  $I$  and output vertices  $O$ , where  $|I| = |O| = n$ .
- Next we create some internal vertices  $I'$  and  $O'$  with  $|I'| = |O'| = 7n/8$ .
- We directly connect  $I$  to  $O$  using  $n$  edges that form a matching. (Say, the  $i$ th vertex in  $I$  is connected to the  $i$ th vertex in  $O$ , for every  $i$ .)
- Let  $H$  be a  $(n, 7n/8, 9)$ -expander. We make one copy of  $H$  with left-vertices  $I$  and right-vertices  $I'$ , and an additional copy with left-vertices  $O$  and right-vertices  $O'$ .
- Finally, we create a smaller superconcentrator with inputs  $I'$  and outputs  $O'$ .

**The Size.** First let us analyze the total number of edges in this construction. We use a simple recurrence relation: let  $f(n)$  be the number of edges used in a superconcentrator on  $n$  inputs and outputs. The total number of edges used in the construction is:

- $n$  edges for the matching from  $I$  to  $O$ .
- $9n$  edges for each of the two expanders.
- $f(7n/8)$  edges for the smaller superconcentrator.

Thus  $f(n) = f(7n/8) + 19n$ . For the base case, we can take  $f(n) = n^2$  whenever  $n$  is a sufficiently small constant. This recurrence has the solution  $f(n) = O(n)$ .

**The Superconcentrator Property.** Consider any  $S \subseteq I$  and  $T \subseteq O$  with  $|S| = |T| = k$ . We must find vertex-disjoint paths from  $S$  to  $T$ .

*Case 1:  $k \leq n/2$ .* By Claim 3, the first expander contains a matching covering  $S$ . Let  $S' \subseteq I'$  be the other endpoints of that matching. Similarly, the second expander contains a matching covering  $T$ . Let  $T' \subseteq O'$  be the other endpoints of that matching. Note that  $|S'| = |T'| = k$ . By induction, the smaller

superconcentrator contains  $k$  vertex-disjoint paths between  $S'$  and  $T'$ . Combining those paths with the edges of the matchings gives the desired paths between  $S$  and  $T$ .

*Case 2:*  $k > n/2$ . Claim 3 only provides a matching covering sets of size at most  $n/2$ , so the previous argument cannot handle the case  $k > n/2$ . The solution is to use the edges that directly connect  $I$  and  $O$ .

By the pigeonhole principle, there are at least  $k - n/2$  vertices in  $S$  that are directly connected by the matching to vertices in  $T$ . The remaining vertices (of which there are at most  $n/2$ ) are handled by the argument of Case 1.