

# 4

---

## 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 this survey, *expander graphs*. These are graphs that are “sparse” yet very “well-connected.”

### 4.1 Measures of Expansion

We will typically interpret the properties of expander graphs 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 the degree  $D_i$  of  $G_i$  should be very slowly growing as a function of  $N_i$ . The “well-connectedness” 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. As in Section 2.4.2, we will state many of our definitions for *directed multigraphs* (*digraphs* for short), though in the end we will mostly study undirected multigraphs.

#### 4.1.1 Vertex Expansion

The classic measure of well-connectedness in expanders follows:

---

**Definition 4.1.** A digraph  $G$  is a  $(K, A)$  *vertex expander* if for all sets  $S$  of at most  $K$  vertices, the *neighborhood*  $N(S) \stackrel{\text{def}}{=} \{u \mid \exists v \in S \text{ s.t. } (u, v) \in E\}$  is of size at least  $A \cdot |S|$ .

---

Ideally, we would like  $D = O(1)$ ,  $K = \Omega(N)$  where  $N$  is the number of vertices, and  $A$  as close to  $D$  as possible.

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

- 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. the second eigenvalue  $\lambda(G)$  is small.

- “Quasi-randomness” (a.k.a “Mixing”): 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 vertex 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 4.2.** For all constants  $D \geq 3$ , there is a constant  $\alpha > 0$  such that for all sufficiently large  $N$ , a random  $D$ -regular undirected 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 4.3.** A bipartite multigraph  $G$  is a  $(K, A)$  *vertex expander* if for all sets  $S$  of *left*-vertices of size at most  $K$ , the neighborhood  $N(S)$  is of size at least  $A \cdot |S|$ .

Now, let  $\text{Bip}_{N,D}$  be the set of bipartite multigraphs that have  $N$  vertices on each side and are  *$D$ -leftregular*, meaning that every vertex on the left has  $D$  neighbors, numbered from  $1, \dots, D$  (but vertices on the right may have varying degrees).

**Theorem 4.4.** 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  $G \stackrel{\text{R}}{\leftarrow} \text{Bip}_{N,D}$  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  $(i - 1)/N \leq KD/N$ . So,

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

Thus, we find that

$$p_K \leq \binom{N}{K} \binom{KD}{2K} \left(\frac{KD}{N}\right)^{2K} \leq \left(\frac{Ne}{K}\right)^K \left(\frac{KDe}{2K}\right)^{2K} \left(\frac{KD}{N}\right)^{2K} = \left(\frac{e^3 D^4 K}{4N}\right)^K$$

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

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

□

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

- We can obtain a bipartite multigraph that is  $D$ -regular on both sides 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 as 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 bipartite multigraph 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 natural combinatorial objects, 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, and 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.

#### 4.1.2 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. As we have seen in Section 2.4.2, the mixing rate of random walks in turn is captured well by the second largest eigenvalue of the transition matrix, and this turns out to be a very useful measure of expansion for many purposes.

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

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

where  $u = (1/N, \dots, 1/N) \in \mathbb{R}^N$  is the uniform distribution on  $[N]$ , the first maximum is over all probability distributions  $\pi \in [0, 1]^N$ , and the second maximum is over all vectors  $x \in \mathbb{R}^N$  that are orthogonal to  $u$ . We write  $\gamma(G) \stackrel{\text{def}}{=} 1 - \lambda(G)$  to denote the *spectral gap* of  $G$ .

---

**Definition 4.5.** For  $\gamma \in [0, 1]$ , we say that a regular digraph  $G$  has *spectral expansion*  $\gamma$  if  $\gamma(G) \geq \gamma$  (equivalently,  $\lambda(G) \leq 1 - \gamma$ ).<sup>1</sup>

---

Larger values of  $\gamma(G)$  (or smaller values of  $\lambda(G)$ ) correspond to better expansion. Sometimes it is more natural to state results in terms of  $\gamma(G)$  and sometimes in terms of  $\lambda(G)$ . 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 4.6 (spectral expansion  $\Rightarrow$  vertex expansion).** If  $G$  is a regular digraph with spectral expansion  $\gamma = 1 - \lambda$  for some  $\lambda \in [0, 1]$ , then, for every  $\alpha \in [0, 1]$ ,  $G$  is an  $(\alpha N, 1/((1 - \alpha)\lambda^2 + \alpha))$  vertex expander. In particular,  $G$  is an  $(N/2, 1 + \gamma)$  expander.

---

We prove this theorem using the following two useful statistics of probability distributions.

---

**Definition 4.7.** 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 4.8.** For every probability distribution  $\pi \in [0, 1]^N$ , we have:

- (1)  $\text{CP}(\pi) = \|\pi\|^2 = \|\pi - u\|^2 + 1/N$ , where  $u$  is the uniform distribution on  $[N]$ .
  - (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  $\pi - u$  is orthogonal to  $u$ , we have  $\|\pi\|^2 = \|u\|^2 + \|\pi - u\|^2 = 1/N + \|\pi - u\|^2$ .

For Part 2, by Cauchy-Schwarz we have

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

with equality iff  $\pi$  is uniform on  $\text{Supp}(\pi)$ . □

---

<sup>1</sup>In other sources (including the original lecture notes on which this survey was based), the spectral expansion referred to  $\lambda$  rather than  $\gamma$ . Here we use  $\gamma$ , because it has the more natural feature that larger values of  $\gamma$  correspond to the graph being “more expanding”.

*Proof.* [of Theorem 4.6] The condition that  $G$  has spectral expansion  $\gamma = 1 - \lambda$  is equivalent to saying that  $\lambda(G) \leq \lambda$ . By the definition of  $\lambda(G)$  and Part 1 of Lemma 4.8, 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  $\text{CP}(\pi M) \geq 1/|\text{Supp}(\pi M)| = 1/|N(S)|$ . Thus,

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

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

The other direction, i.e. obtaining spectral expansion from vertex expansion, is more difficult (and we will not prove it here).

**Theorem 4.9 (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  $\gamma = \Omega((\delta/D)^2)$ .

Note first the dependence on subset size being  $N/2$ : this is necessary, because a graph can have vertex expansion  $(\alpha N, 1 + \Omega(1))$  for  $\alpha < 1/2$  and be disconnected (eg the disjoint union of two good expanders), thereby having no spectral expansion. Also note 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, indeed equivalent for many interesting settings of parameters:

**Corollary 4.10.** Let  $\mathcal{G}$  be an infinite family of  $D$ -regular multigraphs, for a constant  $D \in \mathbb{N}$ . Then the following two conditions are equivalent:

- There is a constant  $\delta > 0$  such that every  $G \in \mathcal{G}$  is an  $(N/2, 1 + \delta)$  vertex expander.
- There is a constant  $\gamma > 0$  such that every  $G \in \mathcal{G}$  has spectral expansion  $\gamma$ .

When people informally use the term “expander,” they often mean a family of regular graphs of *constant degree*  $D$  satisfying one of the two equivalent conditions above.

However, the 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 4.6, a graph with spectral expansion  $1 - \lambda$  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 4.11.** For every constant  $D \in \mathbb{N}$ , any  $D$ -regular,  $N$ -vertex multigraph  $G$  satisfies  $\lambda(G) \geq 2\sqrt{D-1}/D - o(1)$ , where the  $o(1)$  term vanishes as  $N \rightarrow \infty$  (and  $D$  is held constant).

---

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

---

**Theorem 4.12.** For any constant  $D \in \mathbb{N}$ , a random  $D$ -regular  $N$ -vertex graph  $\lambda(G) \leq 2\sqrt{D-1}/D + o(1)$  with probability  $1 - o(1)$  where both  $o(1)$  terms vanish as  $N \rightarrow \infty$  (and  $D$  is held constant).

---

Now let us see what these results for spectral expansion imply in the world of vertex expansion. With Ramanujan graphs ( $\lambda(G) \leq 2\sqrt{D-1}/D$ ), the bound from Theorem 4.6 gives a vertex expansion factor of  $A \approx D/4$  for small sets. This is not tight, and it is known that Ramanujan graphs actually have vertex expansion  $D/2 - o(1)$  for sets of density  $o(1)$ , which is tight in the sense that there are families of graphs with  $\lambda(G) \rightarrow 2\sqrt{D-1}/D$  but vertex expansion at most  $D/2$ . Still, this vertex expansion is not as good as we obtained via the Probabilistic Method (Theorem 4.2), 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 these are crucial).

### 4.1.3 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  $e(S, T) = \{(u, v) \in S \times T \mid \{u, v\} \in E\}$ . Here  $(u, v)$  refers to an *ordered* pair, in contrast to the definition of  $\text{cut}(S, T)$  in Section 2.3.4. Thus, we count edges entirely within  $S \cap T$  twice, corresponding to both orientations.

---

**Definition 4.13.** A  $D$ -regular digraph  $G$  is a  $(K, \varepsilon)$  *edge expander* if for all sets  $S$  of at most  $K$  vertices, the cut  $e(S, \bar{S})$  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  $D$ , only requiring  $|e(S, \bar{S})| \geq \varepsilon \cdot |S|$ .) When viewed in terms of the random walk on  $G$ , the ratio  $e(S, \bar{S})/(|S| \cdot D)$  is the probability that, if we condition the stationary distribution on being in  $S$ , the random walk leaves  $S$  in one step. It turns out that if we fix  $K = N/2$ , then edge expansion turns out to be even more closely related to spectral expansion than is vertex expansion. Indeed:

---

**Theorem 4.14.**

- (1) If a  $D$ -regular,  $N$ -vertex digraph  $G$  has spectral expansion  $\gamma$ , then  $G$  is an  $(N/2, \gamma/2)$  edge expander.
- (2) If a  $D$ -regular,  $N$ -vertex digraph  $G$  is a  $(N/2, \varepsilon)$  edge expander and at least an  $\alpha$  fraction of edges leaving each vertex are self-loops for some  $\alpha \in [0, 1]$ , then  $G$  has spectral expansion  $\alpha \cdot \varepsilon^2/2$ .

---

The condition about self-loops in Part 2 is to ensure that the graph is far from being bipartite (or more generally “periodic” in the sense that all cycle lengths are divisible by some number larger than 1), because a bipartite graph has spectral expansion 0 but can have positive edge expansion. For graphs with a constant fraction of self-loops at each vertex, the theorem implies that the edge expansion is bounded away from 0 iff the spectral expansion is bounded away from 0. Unlike Corollary 4.10, this equivalence holds even for graphs of unbounded degree. The intuition for the relation 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*. Part 1 of Theorem 4.14 will follow as a special case of the Expander Mixing Lemma below; we omit the proof of Part 2.

Next, we consider a generalization of edge expansion, where we look at 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 approximately equal to the product of the densities of  $S$  and  $T$ . The following result shows that this intuition is correct:

---

**Lemma 4.15 (Expander Mixing Lemma).** Let  $G$  be a  $D$ -regular,  $N$ -vertex digraph with spectral expansion  $1 - \lambda$ . Then for all sets of vertices  $S, T$  of densities  $\alpha = |S|/N$  and  $\beta = |T|/N$ , we have

$$\begin{aligned} \left| \frac{e(S, T)}{N \cdot D} - \alpha\beta \right| &\leq \lambda \sqrt{\alpha \cdot (1 - \alpha) \cdot \beta \cdot (1 - \beta)}. \\ &\leq \lambda \sqrt{\alpha\beta} \leq \lambda. \end{aligned}$$


---

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 we were to choose  $G$  randomly is “small”, roughly  $\lambda$  times the square root of this fraction. Finally, note that Part 1 of Theorem 4.14 follows from the Expander Mixing Lemma by setting  $T = S^c$ , so  $\beta = 1 - \alpha$  and  $e(S, T)/ND \geq (1 - \lambda) \cdot \alpha \cdot (1 - \alpha) \geq \gamma\alpha/2$ .

When a digraph  $G = (V, E)$  has the property that  $|e(S, T)/|E| - \alpha\beta| = o(1)$  for all sets  $S, T$  (with densities  $\alpha, \beta$ ), the graph is called *quasirandom*. Thus, the Expander Mixing Lemma implies that a regular digraph with  $\lambda(G) = o(1)$  is quasirandom. Quasirandomness has been studied extensively for dense graphs, in which case it has numerous equivalent formulations. Here we are most interested in sparse graphs, especially constant-degree graphs (for which  $\lambda(G) = o(1)$  is impossible).

*Proof.* Let  $\chi_S$  be the characteristic (row) 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  $e(S, T) = \chi_S A \chi_T^t = \chi_S (DM) \chi_T^t$ , where the superscript  $t$  denotes the transpose.

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  $\langle \chi_S, u \rangle / \|u\| = \sum_i (\chi_S)_i = |S| = \alpha N$ . Then  $\chi_S = (\alpha N)u + \chi_S^\perp$  and similarly  $\chi_T = (\beta N)u + \chi_T^\perp$ . Intuitively, the components parallel to the uniform distribution “spread” the weight of  $S$  and  $T$  uniformly over the entire graph, and  $\chi_S^\perp$  and  $\chi_T^\perp$  will yield the error term.

Formally, we have

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

Since  $uM = u$  and  $Mu^t = u^t$ , and both  $\chi_S^\perp$  and  $\chi_T^\perp$  are orthogonal to  $u$ , the above expression simplifies to:

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

Thus,

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

To complete the proof, we note that

$$\alpha N = \|\chi_S\|^2 = \|(\alpha N)u\|^2 + \|\chi_S^\perp\|^2 = \alpha^2 N + \|\chi_S^\perp\|^2,$$

so  $\|\chi_S^\perp\| = \sqrt{\alpha - \alpha^2} = \sqrt{\alpha \cdot (1 - \alpha)}$  and similarly  $\|\chi_T^\perp\| = \sqrt{\beta \cdot (1 - \beta)}$ .  $\square$

Similarly to vertex expansion and edge expansion, a natural question is to what extent the converse holds. That is, if  $e(S, T)/ND$  is always “close” to the product of the densities of  $S$  and  $T$ , then is  $\lambda(G)$  necessarily small? This is indeed true:

---

**Theorem 4.16 (Converse to Expander Mixing Lemma).** Let  $G$  be a  $D$ -regular,  $N$ -vertex undirected graph. Suppose that for all pairs of disjoint vertex sets  $S, T$ , we have  $\left| \frac{e(S, T)}{N \cdot D} - \mu(S)\mu(T) \right| \leq \theta \sqrt{\mu(S)\mu(T)}$  for some  $\theta \in [0, 1]$ , where  $\mu(R) = |R|/N$  for any set  $R$  of vertices. Then  $\lambda(G) = O(\theta \log(1/\theta))$ .

---

Putting the two theorems together, we see that  $\lambda$  and  $\theta$  differ by at most a logarithmic factor. Thus, unlike the other connections we have seen, this connection is good for highly expanding graphs (i.e.  $\lambda(G)$  close to zero,  $\gamma(G)$  close to 1).

## 4.2 Random Walks on Expanders

From the previous section, we know that one way of characterizing an expander graph  $G$  is by having a bound on its second eigenvalue  $\lambda(G)$ , and in fact there exist constant-degree expanders where  $\lambda(G)$  is bounded by a constant less than 1. From Section 2.4.3, we know that this implies that the random walk on  $G$  converges quickly to the uniform distribution. Specifically, a walk of length  $t$  started at any vertex ends at  $\ell_2$  distance at most  $\lambda^t$  from the uniform distribution. Thus after  $t = O(\log N)$  steps, the distribution is very close to uniform, e.g. the probability of every vertex is  $(1 \pm .01)/N$ . Note that, if  $G$  has constant degree, the number of random bits invested

here is  $O(t) = O(\log N)$ , which is within a constant factor of optimal; clearly  $\log N - O(1)$  random bits are also necessary to sample an almost uniform vertex. Thus, expander walks give a very good tradeoff between the number of random bits invested and the “randomness” of the final vertex in the walk. Remarkably, expander walks give good randomness properties not only for the final vertex in the walk, but also for the *sequence* of vertices traversed in the walk. Indeed, in several ways to be formalized below, this sequence of vertices “behaves” like uniform independent samples of the vertex set.

A canonical application of expander walks is for randomness-efficient error reduction of randomized algorithms: Suppose we have an algorithm with constant error probability, which uses some  $m$  random bits. Our goal is to reduce the error to  $2^{-k}$ , with a minimal penalty in random bits and time. Independent repetitions of the algorithm suffers just an  $O(k)$  multiplicative penalty in time, but needs  $O(km)$  random bits. We have already seen that with pairwise independence we can use just  $O(m+k)$  random bits, but the time blows up by  $O(2^k)$ . Expander graphs let us have the best of both worlds, using just  $m + O(k)$  random bits, and increasing the time by only an  $O(k)$  factor. Note that for  $k = o(m)$ , the number of random bits is  $(1 + o(1)) \cdot m$ , even better than what pairwise independence gives.

The general approach is to consider an expander graph with vertex set  $\{0, 1\}^m$ , where each vertex is associated with a setting of the random bits. We will choose a uniformly random vertex  $v_1$  and then do a random walk of length  $t - 1$ , visiting additional vertices  $v_2, \dots, v_t$ . (Note that unlike the rapid mixing analysis, here we start at a uniformly random vertex.) This requires  $m$  random bits for the initial choice, and  $\log D$  for each of the  $t - 1$  steps. For every vertex  $v_i$  on the random walk, we will run the algorithm with  $v_i$  as the setting of the random coins.

First, we consider the special case of randomized algorithms with one-sided error (**RP**). For these, we should accept if at least one execution of the algorithm accepts, and reject otherwise. If the input is a NO instance, the algorithm never accepts, so we also reject. If the input is a YES instance, we want our random walk to hit at least one vertex that makes the algorithm accept. Let  $B$  denote the set of “bad” vertices giving coin tosses that make the algorithm reject. By definition, the density of  $B$  is at most  $1/2$ . Thus, our aim is to show that the probability that *all* the vertices in the walk  $v_1, \dots, v_t$  are in  $B$  vanishes exponentially fast in  $t$ , if  $G$  is a good expander.

The case  $t = 2$  follows from the Expander Mixing Lemma given last time. If we choose a random edge in a graph with spectral expansion  $1 - \lambda$ , the probability that both endpoints are in a set  $B$  is at most  $\mu(B)^2 + \lambda \cdot \mu(B)$ . So if  $\lambda \ll \mu(B)$ , then the probability is roughly  $\mu(B)^2$ , just like two independent random samples. The case of larger  $t$  is given by the following theorem.

---

**Theorem 4.17 (Hitting Property of Expander Walks).** If  $G$  is a regular digraph with spectral expansion  $1 - \lambda$ , then for any  $B \subset V(G)$  of density  $\mu$ , the probability that a random walk  $(V_1, \dots, V_t)$  of  $t - 1$  steps in  $G$  starting in a uniformly random vertex  $V_1$  always remains in  $B$  is

$$\Pr \left[ \bigwedge_{i=1}^t V_i \in B \right] \leq (\mu + \lambda \cdot (1 - \mu))^t$$

---

Equivalently, a random walk “hits” the complement of  $B$  with high probability. Note that if  $\mu$  and  $\lambda$  are constants less than 1, then the probability of staying in  $B$  is  $2^{-\Omega(t)}$ , completing the analysis of the efficient error-reduction algorithm for **RP**.

Before proving the theorem, we discuss general approaches to analyzing spectral expanders and random walks on them. Typically, the first step is to express the quantities of interest linear-algebraically, involving applications of the random-walk (or adjacency) matrix  $M$  to some vectors  $v$ . For example, last time when proving the Expander Mixing Lemma, we expressed the fraction of edges between sets  $S$  and  $T$  as  $\chi_S^t M \chi_T$  (up to some normalization factor). Then we can proceed in one of the two following ways:

**Vector Decomposition** Decompose the input vector  $v$  as  $v = v^\parallel + v^\perp$ , where  $v^\parallel = (\langle v, u \rangle / \langle u, u \rangle)u$  is the component of  $v$  in the direction of the uniform distribution  $u$  and  $v^\perp$  is the component of  $v$  orthogonal to  $u$ . Then this induces a similar orthogonal decomposition of the output vector  $vM$  into  $vM = (vM)^\parallel + (vM)^\perp = v^\parallel M + v^\perp M$ , where  $v^\parallel M = v^\parallel$  and  $\|v^\perp M\| \leq \lambda \cdot \|v^\perp\|$ . Thus, from information about how  $v$ 's lengths are divided into the uniform and non-uniform components, we deduce information about how  $vM$  is divided into the uniform and non-uniform components. This is the approach we took in the proof of the Expander Mixing Lemma.

**Matrix Decomposition** This corresponds to a different decomposition of the output vector  $vM$  that can be expressed in a way that is independent of the decomposition of the input vector  $v$ . Specifically, if  $G$  has spectral expansion  $\gamma = 1 - \lambda$ , then

$$vM = v^\parallel + v^\perp M = \gamma \cdot v^\parallel + (\lambda \cdot v^\parallel + v^\perp M) = \gamma \cdot vJ + \lambda \cdot vE = v(\gamma J + \lambda E),$$

where  $J$  is the matrix in which every entry is  $1/N$  and the error matrix  $E$  satisfies  $\|vE\| \leq \|v\|$ . The advantage of this decomposition is that we can apply it even when we have no information about how  $v$  decomposes (only its length). The fact that  $M$  is a convex combination of  $J$  and  $E$  means that we can often treat each of these components separately and then just apply triangle inequality. However, it is less refined than the vector decomposition approach, and sometimes gives weaker bounds. Indeed, if we used it to prove the Expander Mixing Lemma (without decomposing  $\chi_S$  and  $\chi_T$ ), we would get a slightly worse error term of  $\lambda\sqrt{\mu(S)\mu(T)} + \lambda\mu(S)\mu(T)$ .

The Matrix Decomposition Approach can be formalized using the following notion.

---

**Definition 4.18.** The (*spectral*) *norm* of an  $N \times N$  real matrix  $M$  is defined to be

$$\|M\| = \max_{x \in \mathbb{R}^N} \frac{\|xM\|}{\|x\|}$$

(If  $M$  is symmetric, then  $\|M\|$  equals the *largest* absolute value of any eigenvalue of  $M$ .)

---

Some basic properties of the matrix norm are that  $\|cA\| = |c| \cdot \|A\|$ ,  $\|A + B\| \leq \|A\| + \|B\|$ , and  $\|A \cdot B\| \leq \|A\| \cdot \|B\|$  for every two matrices  $A, B$ , and  $c \in \mathbb{R}$ . From the discussion above, we have the following lemma:

---

**Lemma 4.19.** Let  $G$  be a regular digraph on  $N$  vertices with random-walk matrix  $M$ . Then  $G$  has spectral expansion  $\gamma = 1 - \lambda$  iff  $M = \gamma J + \lambda E$ , where  $J$  is the  $N \times N$  matrix where every entry is  $1/N$  (i.e. the random-walk matrix for the complete graph with self-loops) and  $\|E\| \leq 1$ .

---

This lemma has a nice intuition: we can think of a random step on a graph with spectral expansion  $\gamma$  as being a random step on the complete graph with probability  $\gamma$  and “not doing damage” with probability  $1 - \gamma$ . This intuition would be completely accurate if  $E$  were a stochastic matrix, but it is typically not (e.g. it may have negative entries). Still, note that the bound given in Theorem 4.17 exactly matches this intuition: in every step, the probability of remaining in  $B$  is at most  $\gamma \cdot \mu + \lambda = \mu + \lambda \cdot (1 - \mu)$ .

Now we can return to the proof of the theorem.

*Proof.* We need a way to express getting stuck in  $B$  linear-algebraically. For that, we define  $P$  to be the diagonal matrix with  $P_{i,i} = 1$  if  $i \in B$  and  $P_{i,i} = 0$  otherwise. An example of using  $P$  would be to say that the probability a distribution  $\pi$  picks a node in  $B$  is  $|\pi P|_1$ , where  $|\cdot|_1$  is the  $\ell_1$  norm,  $|x|_1 = \sum |x_i|$  (which in our case is equal to the sum of the components of the vector, since all values are nonnegative).

Let  $M$  be the random-walk matrix of  $G$ . The probability distribution for  $V_1$  is given by the vector  $u$ . Now we can state the following crucial fact:

---

**Claim 4.20.** The probability that the random walk stays entirely within  $B$  is precisely  $|uP(MP)^{t-1}|_1$ .

---

**Proof of claim:** By induction on  $\ell$ , we show that  $(uP(MP)^\ell)_i$  is the probability that the random walk is entirely within  $B$  for the first  $\ell + 1$  steps and is at node  $i$  at step  $\ell + 1$ . The base case is  $\ell = 0$ . If  $i \in B$ ,  $(uP)_i = 1/N$ ; if  $i \notin B$ ,  $(uP)_i = 0$ . Now assume the hypothesis holds up to some  $\ell$ . Then  $(uP(MP)^\ell M)_i$  is the probability that the random walk is entirely within  $B$  for the first  $\ell + 1$  steps and is at node  $i$  at step  $\ell + 2$ . Multiplying by  $P$ , we zero out all components for nodes not in  $B$  and leave the others unchanged. Thus, we obtain the probability that the random walk is at  $i$  after  $\ell + 2$  steps, and never leaves  $B$ .  $\square$

To get a bound in terms of the spectral expansion, we will now switch to the  $\ell_2$  norm. The intuition is that multiplying by  $M$  shrinks the component that is perpendicular to  $u$  (by expansion) and multiplying by  $P$  shrinks the component parallel to  $u$  (because it zeroes out some entries). Thus, we should be able to show that the norm  $\|MP\|$  is strictly less than 1. Actually, to get the best bound, we note that  $uP(MP)^{t-1} = uP(PMP)^{t-1}$ , because  $P^2 = P$ , so we instead bound  $\|PMP\|$ . Specifically:

---

**Claim 4.21.**  $\|PMP\| \leq \mu + \lambda \cdot (1 - \mu)$ .

---

**Proof of claim:** Using the Matrix Decomposition Lemma (Lemma 4.19), we have:

$$\begin{aligned} \|PMP\| &= \|P(\gamma J + \lambda E)P\| \\ &\leq \gamma \cdot \|PJP\| + \lambda \cdot \|PEP\| \\ &\leq \gamma \cdot \|PJP\| + \lambda \end{aligned}$$

Thus, we only need to analyze the case of  $J$ , the random walk on the complete graph. Given any vector  $x$ , let  $y = xP$ . Note that  $\|y\| \leq \|x\|$  and  $y$  has at most  $\mu N$

nonzero coordinates. Then

$$xPJP = yJP = \left( \sum_i y_i \right) uP,$$

so

$$\|xPJP\| \leq \left| \sum_i y_i \right| \cdot \|uP\| \leq \left( \sqrt{\mu N} \cdot \|y\| \right) \cdot \sqrt{\frac{\mu}{N}} \leq \mu \cdot \|x\|.$$

Thus,

$$\|PMP\| \leq \gamma \cdot \mu + \lambda = \mu + \lambda \cdot (1 - \mu).$$

□

Using Claim 4.21, the probability of never leaving  $B$  in a  $t$ -step random walk is

$$\begin{aligned} |uP(MP)^{t-1}|_1 &\leq \sqrt{\mu N} \cdot \|uP(MP)^{t-1}\| \\ &\leq \sqrt{\mu N} \cdot \|uP\| \cdot \|PMP\|^{t-1} \\ &\leq \sqrt{\mu N} \cdot \sqrt{\frac{\mu}{N}} \cdot (\mu + \lambda \cdot (1 - \mu))^{t-1} \\ &\leq (\mu + \lambda \cdot (1 - \mu))^t \end{aligned}$$

□

The hitting properties described above suffice for reducing the error of **RP** algorithms. What about **BPP** algorithms, which have two-sided error? They are handled by the following.

---

**Theorem 4.22 (Chernoff Bound for Expander Walks).** Let  $G$  be a regular digraph with on  $N$  vertices with spectral expansion  $1 - \lambda$ , and let  $f : [N] \rightarrow [0, 1]$  be any function. Consider a random walk  $V_1, \dots, V_t$  in  $G$  from a uniform start vertex  $V_1$ . Then for any  $\varepsilon > 0$

$$\Pr \left[ \left| \frac{1}{t} \sum_i f(V_i) - \mu(f) \right| \geq \lambda + \varepsilon \right] \leq 2e^{-\Omega(\varepsilon^2 t)}.$$

---

Note that this is just like the standard Chernoff Bound (Theorem 2.21), except that our additive approximation error increases by  $\lambda = 1 - \gamma$ . Thus, unlike the Hitting Property we proved above, this bound is only useful when  $\lambda$  is sufficiently small (as opposed to bounded away from 1). This can be achieved by taking an appropriate power of the initial expander. However, there is a better Chernoff Bound for Expander Walks, where  $\lambda$  does not appear in the approximation error, but the exponent in the probability of error is  $\Omega(\gamma\varepsilon^2 t)$  instead of  $\Omega(\varepsilon^2 t)$ . The bound above suffices in the common case that a small constant approximation error can be tolerated, as in error reduction for **BPP**.

*Proof.* Let  $X_i$  be the random variable  $f(V_i)$ , and  $X = \sum_i X_i$ . Just like in the standard proof of the Chernoff Bound (Problem 2.7), we show that the expectation of the moment generating function  $e^{rX} = \prod_i e^{rX_i}$  is not much larger than  $e^{rE[X]}$  and apply Markov's Inequality, for a suitable choice

of  $r$ . However, here the factors  $e^{rX_i}$  are not independent, so the expectation does not commute with the product. Instead, we express  $\mathbb{E}[e^{rX}]$  linear-algebraically as follows. Define a diagonal matrix  $P$  whose  $(i, i)$ 'th entry is  $e^{rf(i)}$ . Then, similarly to the hitting proof above, we observe that

$$\mathbb{E}[e^{rX}] = |uP(MP)^{t-1}|_1 = |u(MP)^t|_1 \leq \sqrt{N} \cdot \|u\| \cdot \|MP\|^t = \|MP\|^t.$$

To see this, we simply note that each cross-term in the matrix product  $uP(MP)^{t-1}$  corresponds to exactly one expander walk  $v_1, \dots, v_t$ , with a coefficient equal to the probability of this walk times  $\prod_i e^{f(v_i)}$ . By the Matrix Decomposition Lemma (Lemma 4.19), we can bound

$$\|MP\| \leq (1 - \lambda) \cdot \|JP\| + \lambda \cdot \|EP\|.$$

Since  $J$  simply projects onto the uniform direction, we have

$$\begin{aligned} \|JP\|^2 &= \frac{\|uP\|^2}{\|u\|^2} \\ &= \frac{\sum_v (e^{r \cdot f(v)} / N)^2}{\sum_v (1/N)^2} \\ &= \frac{1}{N} \cdot \sum_v e^{2rf(v)} \\ &= \frac{1}{N} \cdot \sum_v (1 + 2rf(v) + O(r^2)) \\ &= 1 + 2r\mu + O(r^2) \end{aligned}$$

for  $r \leq 1$ , and thus

$$\|JP\| = \sqrt{1 + 2r\mu + O(r^2)} = 1 + r\mu + O(r^2).$$

For the error term, we have

$$\|EP\| \leq \|P\| \leq e^r = 1 + r + O(r^2).$$

Thus,

$$\|MP\| \leq (1 - \lambda) \cdot (1 + r\mu + O(r^2)) + \lambda \cdot (1 + r + O(r^2)) \leq 1 + (\mu + \lambda)r + O(r^2),$$

and we have

$$\mathbb{E}[e^{rX}] \leq (1 + (\mu + \lambda)r + O(r^2))^t \leq e^{(\mu + \lambda)rt + O(r^2t)}.$$

By Markov's Inequality,

$$\Pr[X \geq (\mu + \lambda + \varepsilon) \cdot t] \leq e^{-\varepsilon rt + O(r^2t)} = e^{-\Omega(\varepsilon^2 t)},$$

if we set  $r = \varepsilon/c$  for a large enough constant  $c$ . By applying the same analysis to the function  $1 - f$ , we see that  $\Pr[X \leq (\mu - \lambda - \varepsilon)t] = e^{-\Omega(\varepsilon^2 t)}$ , and this establishes the theorem.  $\square$

We now summarize the properties that expander walks give us for randomness-efficient error reduction and sampling.

For reducing the error of a **BPP** algorithm from  $1/3$  to  $2^{-k}$ , we can apply Theorem 4.22 with  $\lambda = \varepsilon = 1/12$ , so that a walk of length  $t = O(k)$  suffices. If the original **BPP** algorithm used

$m$  random bits and the expander is of constant degree (which is possible with  $\lambda = 1/12$ ), then the number of random bits needed is only  $m + O(k)$ . Comparing with previous methods for error reduction, we have:

	Number of Repetitions	Number of Random Bits
Independent Repetitions	$O(k)$	$O(km)$
Pairwise Independent Repetitions	$O(2^k)$	$O(k + m)$
Expander Walks	$O(k)$	$m + O(k)$

For SAMPLING, where we are given an oracle to a function  $f : \{0, 1\}^m \rightarrow [0, 1]$  and we want to approximate  $\mu(f)$  to within an additive error of  $\varepsilon$ , we can apply Theorem 4.22 with error  $\varepsilon/2$  and  $\lambda = \varepsilon/2$ . The needed expander can be obtained by taking an  $O(\log(1/\varepsilon))$ 'th power of a constant-degree expander, yielding the following bounds:

	Number of Samples	Number of Random Bits
Truly Random Sample	$O(\frac{1}{\varepsilon^2} \log \frac{1}{\delta})$	$O(\frac{m}{\varepsilon^2} \log \frac{1}{\delta})$
Pairwise Independent Samples	$O(\frac{1}{\varepsilon^2 \delta})$	$O(m + \log \frac{1}{\varepsilon} + \log \frac{1}{\delta})$
Expander Walks	$O(\frac{1}{\varepsilon^2} \log \frac{1}{\delta})$	$m + O((\log \frac{1}{\delta}) \cdot (\log \frac{1}{\varepsilon}) / \varepsilon^2)$

The  $\log(1/\varepsilon)$  factor in the number of random bits used by expander walks is actually not necessary and comes from the slightly weaker Chernoff Bound we proved. In any case, note that expander walks have a much better dependence on the error probability  $\delta$  in the number of samples (as compared to pairwise independence), but have a worse dependence on the approximation error  $\varepsilon$  in the number of random bits.

Before we end, we make an important remark: we have not actually given an algorithm for randomness-efficient error reduction! Our algorithm assumes an expander graph of exponential size, namely  $2^m$  where  $m$  is the number of random bits used by the algorithm. Generating such a graph at random would use far too many coins. Even generating it deterministically would not suffice, since we would have to write down an exponential-size object. In the following section, we will see how to implicitly construct an expander (without writing it down), and do random walks in such a graph.

### 4.3 Explicit Constructions

As discussed in previous sections, expander graphs have numerous applications in theoretical computer science. (See also the Chapter Notes and Exercises.) For some of these applications, it may be acceptable to simply choose the graph at random, as we know that a random graph will be a good expander with high probability. For many applications, however, this simple approach does not suffice. Some reasons (in increasing order of significance):

- We may not want to tolerate the error probability introduced by the (unlikely) event that the graph is not an expander. To deal with this, we could try checking that the graph is an expander. Computing most combinatorial measures of expansion (e.g. vertex expansion or edge expansion) of a given graph is **NP**-hard, but the spectral expansion can be computed to high precision in time polynomial in the size of the graph (as it is just an eigenvalue computation). As we saw, spectral expansion does yield estimates on vertex expansion and edge expansion (but cannot give optimal expansion in these

measures).

- Some of the applications of expanders (like the one from the previous section) are for reducing the amount of randomness needed for certain tasks. Thus choosing the graph at random defeats the purpose.
- A number of the applications require *exponentially large* expander graphs, and thus we cannot even write down a randomly chosen expander. For example, for randomness-efficient error reduction of randomized algorithms, we need an expander on  $2^m$  nodes where  $m$  is the number of random bits used by the algorithm.

From a more philosophical perspective, finding explicit constructions is a way of developing and measuring our understanding of these fundamental combinatorial objects.

A couple of alternatives for defining explicit constructions of expanders on  $N$  nodes are:

**Mildly Explicit:** Construct a complete representation of the graph in time  $\text{poly}(N)$ .

**Fully Explicit:** Given a node  $u \in [N]$  and  $i \in [D]$ , where  $D$  is the degree of the expander, compute the  $i^{\text{th}}$  neighbor of  $u$  in time  $\text{poly}(\log N)$ .

Consider the randomness-efficient error reduction application discussed in the previous section, in which we performed a random walk on an expander graph with exponentially many nodes. Mild explicitness is insufficient for this application, as the desired expander graph is of exponential size, and hence cannot be even entirely stored, let alone constructed. But full explicitness is perfectly suited for efficiently conducting a random walk on a huge graph. So now our goal is the following:

**Goal:** Devise a fully explicit construction of an infinite family  $\{G_i\}$  of  $D$ -regular graphs with spectral expansion at least  $\gamma$ , where  $D$  and  $\gamma > 0$  are constants independent of  $i$ .

We remark that we would also like the set  $\{N_i\}$ , where  $N_i$  is the number of vertices in  $G_i$ , to be not too sparse, so that the family of graphs  $\{G_i\}$  has graphs of size close to any desired size.

### 4.3.1 Algebraic Constructions

Here we mention a few known explicit constructions that are of interest because of their simple description, the parameters achieved, and/or the mathematics that goes into their analysis. We will not prove the expansion properties of any of these constructions (but will rather give a different explicit construction in the subsequent sections).

---

**Construction 4.23 (Gabber–Galil expanders).** This is the graph  $G = (V, E)$  with nodes  $V = [M]^2$  (namely, each node is a pair of numbers from  $\{0, \dots, M - 1\}$ ), and edges connecting each node  $(x, y)$  with each of the following nodes:  $(x, y)$ ,  $(x, x + y)$ ,  $(x, x + y + 1)$ ,  $(x + y, y)$ , and  $(x + y + 1, y)$ , where all arithmetic is mod  $M$ .

---

This is a fully explicit 5-regular digraph with  $N = M^2$  nodes and spectral expansion  $\gamma = \Omega(1)$ , proved using Fourier analysis. We note that this graph is directed, but it can be made undirected by adding a reverse copy of each edge.

---

**Construction 4.24 ( $p$ -Cycle with inverse chords).** This is the graph  $G = (V, E)$  with vertex set  $V = \mathbb{Z}_p$ , the integers modulo  $p$ , and edges that connect each node  $x$  with the nodes:  $x + 1$ ,  $x - 1$  and  $x^{-1}$  (where all arithmetic is mod  $p$  and we define  $0^{-1}$  to be 0).

---

This graph is only mildly explicit since we do not know how to construct  $n$ -bit primes deterministically in time  $\text{poly}(n)$  (though it is conjectured that we can do so by simply checking the first  $\text{poly}(n)$   $n$ -bit numbers). The proof of expansion relies on the “Selberg 3/16 Theorem” from number theory.

---

**Construction 4.25 (Ramanujan graphs).** The graph  $G = (V, E)$  with nodes  $V = \mathbb{F}_q \cup \{\infty\}$  elements in the finite field of prime order  $q$  s.t.  $q \equiv 1 \pmod{4}$  plus one extra node representing infinity. The edges in this graph connect each node  $z$  with all  $z'$  of the form

$$z' = \frac{(a_0 + ia_1)z + (a_2 + ia_3)}{(-a_2 + ia_3)z + (a_0 - ia_1)}$$

for  $a_0, a_1, a_2, a_3 \in \mathbb{Z}$  such that  $a_0^2 + a_1^2 + a_2^2 + a_3^2 = p$ ,  $a_0$  is odd and positive, and  $a_1, a_2, a_3$  are even, for some fixed prime  $p \neq q$  satisfying  $p \equiv 1 \pmod{4}$  and where  $i^2 = -1 \pmod{q}$ .

---

The degree of the graph is the number of solutions to the equation  $a_0^2 + a_1^2 + a_2^2 + a_3^2 = p$ , which turns out to be  $D = p + 1$ , and it has  $\lambda(G) \leq 2\sqrt{D-1}/D$ , so it is an *optimal* spectral expander. (See Theorems 4.11 and 4.12, and note that this bound is even better than we know for random graphs, which have an additive  $o(1)$  term in the spectral expansion.) These graphs are also only mildly explicit.

These are called Ramanujan Graphs because the proof of their spectral expansion relies on results in number theory concerning the “Ramanujan Conjectures.” Subsequently, the term *Ramanujan graphs* came to refer to any infinite family of graphs with optimal spectral expansion  $\gamma \geq 2\sqrt{D-1}/D$ .

### 4.3.2 Graph Operations

The explicit construction of expanders given in the next section will be an iterative one, where we start with a “constant size” expander  $H$  and repeatedly apply graph operations to get bigger expanders. The operations that we apply should increase the number of nodes in the graph, while keeping the degree and the second eigenvalue  $\lambda$  bounded. We’ll see three operations, each improving one property while paying a price on the others; however, combined together, they yield the desired expander. It turns out that this approach for constructing expanders will also be useful in derandomizing the logspace algorithm for UNDIRECTED S-T CONNECTIVITY, as we will see in Section 4.4.

The following concise notation will be useful to keep track of each of the parameters:

---

**Definition 4.26.** An  $(N, D, \gamma)$ -graph is a  $D$ -regular digraph on  $N$  vertices with spectral expansion  $\gamma$ .

---

#### 4.3.2.1 Squaring

**Definition 4.27 (Squaring of Graphs).** If  $G = (V, E)$  is a  $D$ -regular digraph, then  $G^2 = (V, E')$  is a  $D^2$ -regular digraph on the same vertex set, where the  $(i, j)$ ’th neighbor of a vertex  $x$  is the  $j$ ’th neighbor of the  $i$ ’th neighbor of  $x$ . In particular, a random step on  $G^2$  consists of two random steps on  $G$ .

---

---

**Lemma 4.28.** If  $G$  is a  $(N, D, 1 - \lambda)$ -graph, then  $G^2$  is a  $(N, D^2, 1 - \lambda^2)$ -graph.

---

Namely, the degree deteriorates by squaring, while the spectral expansion is improved from  $\gamma = 1 - \lambda$  to  $\gamma' = 1 - \lambda^2 = 2\gamma - \gamma^2$ .

*Proof.* The effect of squaring on the number of nodes  $N$  and the degree  $D$  is immediate from the definition. For the spectral expansion, note that if  $M$  is the random-walk matrix for  $G$ , then  $M^2$  is the random-walk matrix for  $G^2$ . So for any vector  $x \perp u$ ,

$$\|xM^2\| \leq \lambda \cdot \|xM\| \leq \lambda^2 \cdot \|x\|.$$

□

### 4.3.2.2 Tensoring

The next operation we consider increases the size of the graph at the price of increasing the degree.

---

**Definition 4.29 (Tensor Product of Graphs).** Let  $G_1 = (V_1, E_1)$  be  $D_1$ -regular and  $G_2 = (V_2, E_2)$  be  $D_2$ -regular. Then their *tensor product* is the  $D_1D_2$ -regular graph  $G_1 \otimes G_2 = (V_1 \times V_2, E)$ , where the  $(i_1, i_2)$ 'th neighbor of a vertex  $(x_1, x_2)$  is  $(y_1, y_2)$ , where  $y_b$  is the  $i_b$ 'th neighbor of  $x_b$  in  $G_b$ . That is, a random step on  $G_1 \otimes G_2$  consists of a random step on  $G_1$  in the first component and a random step on  $G_2$  in the second component.

---

Often this operation is simply called the “product” of  $G_1$  and  $G_2$ , but we use “tensor product” to avoid confusion with squaring and to reflect its connection with the standard tensor products in linear algebra:

---

**Definition 4.30 (Tensor Products of Vectors and Matrices).** Let  $x \in \mathbb{R}^{N_1}, y \in \mathbb{R}^{N_2}$ , then their *tensor product*  $z = x \otimes y \in \mathbb{R}^{N_1N_2}$  is the vector<sup>2</sup>  $z_{ij} = x_i y_j$ .

Similarly, for matrices  $A = (a_{ij}) \in \mathbb{R}_{N_1 \times N_1}, B = (b_{ij}) \in \mathbb{R}_{N_2 \times N_2}$ , their *tensor product*  $C = A \otimes B \in \mathbb{R}_{N_1N_2 \times N_1N_2}$  is the matrix  $C = (c_{ij})$  where  $c_{ij, i'j'} = a_{ii'} b_{jj'}$ .

---

A few comments on the tensor operation:

- A random walk on a tensor graph  $G_1 \otimes G_2$  is equivalent to taking two independent random walks on  $G_1$  and  $G_2$ .
- For vectors  $x \in \mathbb{R}^{N_1}, y \in \mathbb{R}^{N_2}$  that are probability distributions (i.e. nonnegative vectors with  $\ell_1$  norm 1), their tensor product  $x \otimes y$  is a probability distribution on  $[N_1] \times [N_2]$  elements where the two components are independently distributed according to  $x$  and  $y$ , respectively.
- $(x \otimes y)(A \otimes B) = (xA) \otimes (yB)$  for every  $x \in \mathbb{R}^{N_1}, y \in \mathbb{R}^{N_2}$ , and in fact  $A \otimes B$  is the unique matrix with this property.

---

<sup>2</sup>For convenience we index the vector  $z$  by two indices  $i, j$ . To transform to a standard indexing we can map  $ij$  to  $iN_1 + j$ .

- Not all vectors  $z \in \mathbb{R}^{N_1 N_2}$  are decomposable as  $x \otimes y$  for  $x \in \mathbb{R}^{N_1}$  and  $y \in \mathbb{R}^{N_2}$ . Nevertheless, the set of all decomposable tensors  $x \otimes y$  spans  $\mathbb{R}^{N_1 N_2}$ .
- If  $M_1, M_2$  are the random-walk matrices for graphs  $G_1, G_2$  respectively, then the random-walk matrix for the graph  $G_1 \otimes G_2$  is

$$M_1 \otimes M_2 = (I_{N_1} \otimes M_2)(M_1 \otimes I_{N_2}) = (M_1 \otimes I_{N_2})(I_{N_1} \otimes M_2),$$

where  $I_N$  denotes the  $N \times N$  identity matrix. That is, we can view a random step on  $G_1 \otimes G_2$  as being a random step on the  $G_1$  component followed by one on the  $G_2$  component or vice-versa.

---

**Lemma 4.31.** If  $G_1$  is an  $(N_1, D_1, \gamma_1)$ -graph and  $G_2$  is an  $(N_2, D_2, \gamma_2)$ -graph, then  $G_1 \otimes G_2$  is an  $(N_1 N_2, D_1 D_2, \min\{\gamma_1, \gamma_2\})$ -graph.

---

In particular, if  $G_1 = G_2$ , then the number of nodes improves, the degree deteriorates, and the spectral expansion remains unchanged.

*Proof.* As usual, we write  $\gamma_1 = 1 - \lambda_1$ ,  $\gamma_2 = 1 - \lambda_2$ ; then our goal is to show that  $G_1 \otimes G_2$  has spectral expansion  $1 - \max\{\lambda_1, \lambda_2\}$ . The intuition for the construction is as follows. Any probability distribution  $(V_1, V_2)$  on the vertices  $(v_1, v_2)$  of  $G_1 \otimes G_2$  can be thought of as picking a cloud  $v_1$  according to the marginal distribution<sup>3</sup>  $V_1$  and then picking the vertex  $v_2$  within the cloud  $v_1$  according to the conditional distribution  $V_2|_{V_1=v_1}$ . If the overall distribution on pairs is far from uniform, then either

- (1) The marginal distribution on the clouds must be far from uniform, *or*
- (2) the conditional distribution within the clouds must be far from uniform.

When we take a random step, the expansion of  $G_1$  will bring us closer to uniform in Case 1 and the expansion of  $G_2$  will bring us closer to uniform in Case 2.

One way to prove the bound in the case of undirected graphs is to use the fact that the eigenvalues of  $M_1 \otimes M_2$  are all the products of eigenvalues of  $M_1$  and  $M_2$ , so the three largest absolute values are  $\{1 \cdot 1, \lambda_1 \cdot 1, 1 \cdot \lambda_2\}$ . Instead, we instead use the Vector Decomposition Method to give a proof that matches the intuition more closely and is a good warm-up for the analysis of the zig-zag product in the next section. Given any vector  $x \in \mathbb{R}^{N_1 N_2}$  that is orthogonal to  $u_{N_1 N_2}$ , we can decompose  $x$  as  $x = x^\parallel + x^\perp$  where  $x^\parallel$  is a multiple of  $u_{N_2}$  on each cloud of size  $N_2$  and  $x^\perp$  is orthogonal to  $u_{N_2}$  on each cloud. Note that  $x^\parallel = y \otimes u_{N_2}$ , where  $y \in \mathbb{R}^{N_1}$  is orthogonal to  $u_{N_1}$  (because  $x^\parallel = x - x^\perp$  is orthogonal to  $u_{N_1 N_2}$ ). If we think of  $x$  as the nonuniform component of a probability distribution, then  $x^\parallel$  and  $x^\perp$  correspond to the two cases in the intuition above.

For the first case, we have

$$x^\parallel M = (y \otimes u_{N_2})(M_1 \otimes M_2) = (y M_1) \otimes u_{N_2}.$$

---

<sup>3</sup>For two jointly distributed random variables  $(X, Y)$ , the *marginal distribution* of  $X$  is simply the distribution of  $X$  alone, ignoring information about  $Y$ .

The expansion of  $G_1$  tells us that  $M_1$  shrinks  $y$  by a factor of  $\lambda_1$ , and thus  $M$  shrinks  $x^\parallel$  by the same factor. For the second case, we write

$$x^\perp M = x^\perp (I_{N_1} \otimes M_2) (M_1 \otimes I_{N_2}).$$

The expansion of  $G_2$  tells us that  $M_2$  will shrink  $x^\perp$  by a factor of  $\lambda_2$  on each cloud, and thus  $I_{N_1} \otimes M_2$  will shrink  $x^\perp$  by the same factor. The subsequent application of  $M_1 \otimes I_{N_2}$  cannot increase the length (being the random-walk matrix for a regular graph, albeit a disconnected one). Thus,  $\|x^\perp M\| \leq \lambda_2 \|x^\perp\|$ .

Finally, we argue that  $x^\parallel M$  and  $x^\perp M$  are orthogonal. Note that  $x^\parallel M = (yM_1) \otimes u_{N_2}$  is a multiple of  $u_{N_2}$  on every cloud. Thus it suffices to argue that  $x^\perp$  remains orthogonal to  $u_{N_2}$  on every cloud after we apply  $M$ . Applying  $(I_{N_1} \otimes M_2)$  retains this property (because applying  $M_2$  preserves orthogonality to  $u_{N_2}$ , by regularity of  $G_2$ ) and applying  $(M_1 \otimes I_{N_2})$  retains this property because it assigns each cloud a linear combination of several other clouds (and a linear combination of vectors orthogonal to  $u_{N_2}$  is also orthogonal to  $u_{N_2}$ ).

Thus,

$$\begin{aligned} \|xM\|^2 &= \|x^\parallel M\|^2 + \|x^\perp M\|^2 \\ &\leq \lambda_1^2 \cdot \|x^\parallel\|^2 + \lambda_2^2 \cdot \|x^\perp\|^2 \\ &\leq \max\{\lambda_1, \lambda_2\}^2 \cdot (\|x^\parallel\|^2 + \|x^\perp\|^2) \\ &= \max\{\lambda_1, \lambda_2\}^2 \cdot \|x\|^2, \end{aligned}$$

as desired. □

### 4.3.2.3 The Zig-Zag Product

Of the two operations we have seen, one (squaring) improves expansion and one (tensoring) increases size, but both have the deleterious effect of increasing the degree. Now we will see a third operation that decreases the degree, without losing too much in the expansion. By repeatedly applying these three operations, we will be able to construct arbitrarily large expanders while keeping both the degree and expansion constant.

Let  $G$  be an  $(N_1, D_1, \gamma_1)$  expander and  $H$  be a  $(D_1, D_2, \gamma_2)$  expander. The *zig-zag product* of  $G$  and  $H$ , denoted  $G \mathbb{Z} H$ , will be defined as follows. The nodes of  $G \mathbb{Z} H$  are the pairs  $(u, i)$  where  $u \in V(G)$  and  $i \in V(H)$ . The edges of  $G \mathbb{Z} H$  will be defined so that a random step on  $G \mathbb{Z} H$  corresponds to a random step on  $G$ , but using a random steps on  $H$  to choose the edge in  $G$ . (This is the reason why we require the number of vertices in  $H$  to be equal to the degree of  $G$ .) A step in  $G \mathbb{Z} H$  will therefore involve a step to a random neighbor in  $H$  and then a step in  $G$  to a neighbor whose index is equal to the label of the current node in  $H$ . Intuitively, a random walk on an a “good” expander graph  $H$  should generate choices that are sufficiently random to produce a “good” random walk on  $G$ . One problem with this definition is that it is not symmetrical. That is, the fact that you can go from  $(u, i)$  to  $(v, j)$  does not mean that you can go from  $(v, j)$  to  $(u, i)$ . We correct this by adding another step in  $H$  after the step in  $G$ . In addition to allowing us to construct undirected expander graphs, this extra step will also turn out to be important for the expansion of  $G \mathbb{Z} H$ .

More formally,

---

**Definition 4.32 (Zig-zag Product).** Let  $G$  be an  $D_1$ -regular digraph on  $N_1$  vertices, and  $H$  a  $D_2$ -regular digraph on  $D_1$  vertices. Then  $G \otimes H$  is a graph whose vertices are pairs  $(u, i) \in [N_1] \times [D_1]$ . For  $a, b \in D_2$ , the  $(a, b)$ 'th neighbor of a vertex  $(u, i)$  is the vertex  $(v, j)$  computed as follows:

- (1) Let  $i'$  be the  $a$ 'th neighbor of  $i$  in  $H$ .
  - (2) Let  $v$  be the  $i'$ 'th neighbor of  $u$  in  $G$ , so  $e = (u, v)$  is the  $i'$ 'th edge leaving  $u$ . Let  $j'$  be such that  $e$  is the  $j'$ 'th edge entering  $v$  in  $G$ . (In an undirected graph, this means that  $u$  is the  $j'$ 'th neighbor of  $v$ .)
  - (3) Let  $j$  be the  $b$ 'th neighbor of  $j'$  in  $H$ .
- 

**Theorem 4.33.** If  $G$  is a  $(N_1, D_1, \gamma_1)$ -graph, and  $H$  is a  $(D_1, D_2, \gamma_2)$ -graph then  $G \otimes H$  is a  $(N_1 D_1, D_2^2, \gamma = \gamma_1 \cdot \gamma_2^2)$ -graph. In particular, if  $\gamma_1 = 1 - \lambda_1$  and  $\gamma_2 = 1 - \lambda_2$ , then  $\gamma = 1 - \lambda$  for  $\lambda \leq \lambda_1 + 2\lambda_2$ .

---

$G$  should be thought of as a big graph and  $H$  as a small graph, where  $D_1$  is a large constant and  $D_2$  is a small constant. Note that the number of nodes  $D_1$  in  $H$  is required to equal the degree of  $G$ . Observe that when  $D_1 > D_2^2$  the degree is reduced by the zig-zag product.

There are two different intuitions underlying the expansion of the zig-zag product:

- Given an initial distribution on the vertices of  $G_1 \otimes G_2$  that is far from uniform, there are two extreme cases just as in the intuition for the tensor product. Either
  - (1) All the (conditional) distributions within the clouds are far from uniform, *or*
  - (2) All the (conditional) distributions within the clouds of size  $D_1$  are uniform (in which case the marginal distribution on the clouds must be far from uniform).

In Case 1, the first  $H$ -step already brings us closer to the uniform distribution, and the other two steps cannot hurt (as they are steps on regular graphs). In Case 2, the first  $H$ -step has no effect, but the  $G$ -step has the effect of making the marginal distribution on clouds closer to uniform. But note that we haven't actually gotten closer to the uniform distribution on the vertices of  $G_1 \otimes G_2$  because the  $G$ -step is a permutation. Still, if the marginal distribution on clouds has become closer to uniform, then the conditional distributions within the clouds must have become further from uniform, and thus the second  $H$ -step brings us closer to uniform. This leads to a proof by *Vector Decomposition*, where we decompose any vector  $x$  that is orthogonal to uniform into components  $x^{\parallel}$  and  $x^{\perp}$ , where  $x^{\parallel}$  is uniform on each cloud, and  $x^{\perp}$  is orthogonal to uniform on each cloud. This approach gives the best known bounds on the spectral expansion of the zig-zag product, but it can be a bit messy since the two components generally do not remain orthogonal (unlike the case of the tensor product, where we were able to show that  $x^{\parallel} M$  is orthogonal to  $x^{\perp} M$ ).

- The second intuition is to think of the expander  $H$  as behaving “similarly” to the complete graph on  $D_1$  vertices (with self-loops). In the case that  $H$  equals the complete graph, then

it is easy to see that  $G \circledast H = G \otimes H$ . Thus it is natural to apply *Matrix Decomposition*, writing the random-walk matrix for an arbitrary expander  $H$  as a convex combination of the random-walk matrix for the complete graph and an error matrix. This gives a very clean analysis, but slightly worse bounds than the Vector Decomposition Method.

We now proceed with the formal proof, following the Matrix Decomposition approach.

*Proof.* [of Theorem 4.33] Let  $A$ ,  $B$ , and  $M$  be the random-walk matrices for  $G_1$ ,  $G_2$ , and  $G_1 \circledast G_2$ , respectively. We decompose  $M$  into the product of three matrices, corresponding to the three steps in the definition of  $G_1 \circledast G_2$ 's edges. Let  $\tilde{B}$  be the transition matrix for taking a random  $G_2$ -step on the second component of  $[N_1] \times [D_1]$ , i.e.  $\tilde{B} = I_{N_1} \otimes B$ , where  $I_{N_1}$  is the  $N_1 \times N_1$  identity matrix. Let  $\hat{A}$  be the permutation matrix corresponding to the  $G_1$ -step. That is,  $\hat{A}_{(u,i),(v,j)}$  is 1 iff  $(u, v)$  is the  $i$ 'th edge leaving  $u$  and the  $j$ 'th edge entering  $v$ . By the definition of  $G_1 \circledast G_2$ , we have  $M = \tilde{B} \hat{A} \tilde{B}$ .

By the Matrix Decomposition Lemma (Lemma 4.19),  $B = \gamma_2 J + (1 - \gamma_2) E$ , where every entry of  $J$  equals  $1/D_1$  and  $E$  has norm at most 1. Then  $\tilde{B} = \gamma_2 \tilde{J} + (1 - \gamma_2) \tilde{E}$ , where  $\tilde{J} = I_{N_1} \otimes J$  and  $\tilde{E} = I_{N_1} \otimes E$  has norm at most 1.

This gives

$$M = \left( \gamma_2 \tilde{J} + (1 - \gamma_2) \tilde{E} \right) \hat{A} \left( \gamma_2 \tilde{J} + (1 - \gamma_2) \tilde{E} \right) = \gamma_2^2 \tilde{J} \hat{A} \tilde{J} + (1 - \gamma_2^2) F,$$

where  $F$  has norm at most 1. Now, the key observation is that  $\tilde{J} \hat{A} \tilde{J} = A \otimes J$ .

Thus,

$$M = \gamma_2^2 \cdot A \otimes J + (1 - \gamma_2^2) F,$$

and thus

$$\begin{aligned} \lambda(M) &\leq \gamma_2^2 \cdot \lambda(A \otimes J) + (1 - \gamma_2^2) \\ &\leq \gamma_2^2 \cdot (1 - \gamma_1) + (1 - \gamma_2^2) \\ &= 1 - \gamma_1 \gamma_2^2, \end{aligned}$$

as desired. □

### 4.3.3 The Expander Construction

**Construction 4.34 (Mildly Explicit Expanders).** As a first attempt for constructing a family of expanders, we construct an infinite family  $G_1, G_2, \dots$  of graphs utilizing only the squaring and the zig-zag operations: Let  $H$  be a  $(D^4, D, 1 - \lambda_0)$ -graph for  $\lambda_0 = 1/8$  (e.g., as constructed in Problem 4.3).

$$\begin{aligned} G_1 &= H^2 \\ G_{t+1} &= G_t^2 \circledast H \end{aligned}$$

---

**Proposition 4.35.** For all  $t$ ,  $G_t$  is a  $(D^{4t}, D^2, 1/2)$ -graph.

---

*Proof.* By induction on  $t$ .

Base Case: by the definition of  $H$  and Lemma 4.28,  $G_1 = H^2$  is a  $(D^4, D^2, 1 - \lambda_0^2)$ -graph and  $\lambda_0^2 \leq 1/2$ .

Induction Step: First note that  $G_t^2 \otimes H$  is well-defined because  $\deg(G_t^2) = \deg(G_t)^2 = (D^2)^2 = \#\text{nodes}(H)$ . Then,

$$\begin{aligned} \deg(G_{t+1}) &= \deg(H)^2 = D^2 \\ \#\text{nodes}(G_{t+1}) &= \#\text{nodes}(G_t^2) \cdot \#\text{nodes}(H) = N_t \cdot D^4 = D^{4t} D^4 = D^{4(t+1)} \\ \lambda(G_{t+1}) &\leq \lambda(G_t)^2 + 2\lambda_0 \leq (1/2)^2 + 2 \cdot (1/8) = 1/2 \end{aligned}$$

□

Now, we recursively bound the time to compute neighbors in  $G_t$ . Actually, due to the way the  $G$ -step in the zig-zag product is defined, we actually bound the time to compute the *edge-rotation map*  $(u, i) \mapsto (v, j)$ , where the  $i$ 'th edge leaving  $u$  equals the  $j$ 'th edge entering  $v$ . Denote by  $\text{time}(G_t)$  the time required for one evaluation of the edge-rotation map for  $G_t$ . This requires two evaluations of the edge-rotation map for  $G_{t-1}$  (the squaring requires two applications, while the zig-zag part does not increase the number of applications), plus time  $\text{poly}(\log N_t)$  for manipulating strings of length  $O(\log N_t)$ . Therefore,

$$\begin{aligned} \text{time}(G_t) &= 2 \cdot \text{time}(G_{t-1}) + \text{poly}(\log N_t) \\ &= 2^t \cdot \text{poly}(\log N_t) \\ &= N_t^{\Theta(1)}, \end{aligned}$$

where the last equality holds because  $N_t = D^{4t}$  for a constant  $D$ . Thus, this construction is only mildly explicit.

---

**Construction 4.36 (Fully Explicit Expanders).** We remedy the above difficulty by using tensoring to make the sizes of the graphs grow more quickly. Let  $H$  be a  $(D^8, D, 1/8)$ -graph, and define:

$$\begin{aligned} G_1 &= H^2 \\ G_{t+1} &= (G_t \otimes G_t)^2 \otimes H \end{aligned}$$


---

In this family of graphs, the number of nodes grow doubly exponentially  $N_t \approx c^{2^t}$ , while computation time grows only exponentially as before. Namely,

$$\text{time}(G_t) = 4^t \cdot \text{poly}(\log N_t) = \text{poly}(\log N_t).$$

We remark that the above family is rather sparse, i.e. the numbers in  $\{N_t\}$  are far apart. To overcome this shortcoming, we can amend the above definition to have

$$G_t = (G_{\lceil t/2 \rceil} \otimes G_{\lfloor t/2 \rfloor})^2 \otimes H.$$

Now  $N_t = D^{8t}$ , so given a number  $N$ , we can find a graph  $G_t$  in the family whose size is at most  $D^8 \cdot N = O(N)$ . Moreover, the construction remains fully explicit because  $\text{time}(G_t) = O(\text{time}(G_{\lceil t/2 \rceil}) + \text{time}(G_{\lfloor t/2 \rfloor})) = \text{poly}(t)$ . Thus we have established:

---

**Theorem 4.37.** There is a constant  $D$  such that for every  $t \in \mathbb{N}$ , there is a fully explicit expander graph  $G_t$  with degree  $D$ , spectral expansion  $7/8$ , and  $N_t = D^{4t}$  nodes.

---

#### 4.3.4 Open Problems

As we have seen, spectral expanders such as those in Theorem 4.37 are also vertex expanders (Theorem 4.6 and Corollary 4.10) and edge expanders (Theorem 4.14), but these equivalences do not extend to optimizing the various expansion measures.

As mentioned in Section 4.3.1, there are explicit constructions of optimal spectral expanders, namely Ramanujan graphs. However, unlike the expanders of Theorem 4.37, those constructions rely on deep results in number theory. The lack of a more elementary construction seems to signify a limitation in our understanding of expander graphs.

---

**Open Problem 4.38.** Give an explicit “combinatorial” construction of constant-degree expander graphs  $G$  with  $\lambda(G) \leq 2\sqrt{D-1}/D$  (or even  $\lambda(G) = O(\sqrt{D})$ ), where  $D$  is the degree.

---

For vertex expansion, it is known how to construct *bipartite* (or directed) expanders with constant left-degree (or out-degree)  $D$  and expansion  $(1-\varepsilon) \cdot D$  for an arbitrarily small constant  $\varepsilon$  (see Chapter ??), but achieving the optimal expansion of  $D - O(1)$  (cf., Theorem 4.4) or constructing undirected vertex expanders with high expansion remains open.

---

**Open Problem 4.39.** For an arbitrarily large constant  $D$ , give an explicit construction of bipartite  $(\Omega(N), D - c)$  vertex expanders with  $N$  vertices on each side and left-degree  $D$ , where  $c$  is a universal constant independent of  $D$ .

---

---

**Open Problem 4.40.** For an arbitrarily small constant  $\varepsilon > 0$ , give an explicit construction of *undirected*  $(\Omega(N), (1-\varepsilon)D)$  vertex expanders with  $N$  vertices and constant degree  $D$  that depends only on  $\varepsilon$ .

---

We remark that the case of highly imbalanced bipartite expanders is even harder (despite them being useful in a number of applications); see Chapter ??.

## 4.4 UNDIRECTED S-T CONNECTIVITY in Deterministic Logspace

Recall the UNDIRECTED S-T CONNECTIVITY problem: given an undirected graph  $G$  and two vertices  $s, t$ , decide whether there is a path from  $s$  to  $t$ . In Section 2.4, we saw that this problem can be solved in randomized logspace (**RL**). Here we will see how we can use expanders and the operations above to solve this problem in deterministic logspace (**L**).

The algorithm is based on the following two ideas:

- UNDIRECTED S-T CONNECTIVITY can be solved in logspace on constant-degree expander graphs. More precisely, it is easy on constant-degree graphs where every connected component is promised to be an expander (i.e. has spectral expansion bounded away from 1):

we can try all paths of length  $O(\log N)$  from  $s$  in logarithmic space; this works because expanders have logarithmic diameter. (See Problem 4.2.)

- The same operations we used to construct an infinite expander family above can also be used to turn *any* graph into an expander (in logarithmic space). Above, we started with a constant-sized expander and used various operations to build larger and larger expanders. There, the goal was to increase the size of the graph (which was accomplished by tensoring and/or zig-zag), while preserving the degree and the expansion (which was accomplished by zig-zag and squaring, which made up for losses in these parameters). Here, we want to improve the expansion (which will be accomplished by squaring), while preserving the degree (as will be handled by zig-zag) and ensuring the graph remains of polynomial size (so we will not use tensoring).

Specifically, the algorithm is as follows.

---

**Algorithm 4.41** (UNDIRECTED S-T CONNECTIVITY in L).

Input: An undirected graph  $G$  with  $N$  edges and vertices  $s$  and  $t$ .

- (1) Let  $H$  be a fixed  $(D^4, D, 3/4)$  graph for some constant  $D$ .
  - (2) Reduce  $(G, s, t)$  to  $(G_0, s_0, t_0)$ , where  $G_0$  is a  $D^2$ -regular graph in which every connected component is nonbipartite and  $s_0$  and  $t_0$  are connected in  $G_0$  iff  $s$  and  $t$  are connected in  $G$ .
  - (3) For  $k = 1, \dots, \ell = O(\log N)$ , define:
    - (a) Let  $G_k = G_{k-1}^2 \otimes H$
    - (b) Let  $s_k$  and  $t_k$  be any two vertices in the “clouds” of  $G_k$  corresponding to  $s_{k-1}$  and  $t_{k-1}$ , respectively. (Note that if  $s_k$  and  $t_k$  are connected in  $G_k$ , then  $s_{k-1}$  and  $t_{k-1}$  are connected in  $G_{k-1}$ .)
  - (4) Try all paths of length  $O(\log N)$  in  $G_\ell$  from  $s_\ell$  and accept if any of them visit  $t_\ell$ .
- 

We will discuss how to implement this algorithm in logspace later, and first analyze its correctness. Let  $C_k$  be the connected component of  $G_k$  containing  $s_k$ . Observe that  $C_k$  is a connected component of  $C_{k-1}^2 \otimes H$ ; below we will show that  $C_{k-1}^2 \otimes H$  is connected and hence  $C_k = C_{k-1}^2 \otimes H$ . Since  $C_0$  is undirected, connected, and nonbipartite, we have  $\gamma(C_0) \geq 1/\text{poly}(N)$  by Theorem 2.55. We will argue that in each iteration the spectral gap increases by a constant factor, and thus after  $O(\log N)$  iterations we have an expander.

By Lemma 4.28, we have

$$\gamma(C_k^2) \geq 2 \cdot \gamma(C_{k-1}) \cdot (1 - \gamma(C_{k-1})/2) \approx 2\gamma(C_k)$$

for small  $\gamma(C_{k-1})$ . By Theorem 4.33, we have

$$\begin{aligned} \gamma(C_{k-1}^2 \otimes H) &\geq \gamma(H)^2 \cdot \gamma(C_{k-1}^2) \\ &\geq \left(\frac{3}{4}\right)^2 \cdot 2 \cdot \gamma(C_{k-1}) \cdot (1 - \gamma(C_{k-1})/2) \\ &\geq \min \left\{ \frac{17}{16} \cdot \gamma(C_{k-1}), \frac{1}{18} \right\} \end{aligned}$$

where the last inequality is obtained by considering whether  $\gamma(C_{k-1}) \leq 1/18$  or  $\gamma(C_{k-1}) > 1/18$ . In particular,  $C_{k-1}^2 \otimes H$  is connected, so we have  $C_k = C_{k-1}^2 \otimes H$  and

$$\gamma(C_k) \geq \min \left\{ \frac{17}{16} \cdot \gamma(C_{k-1}), \frac{1}{18} \right\}.$$

Thus, after  $\ell = O(\log N)$  iterations, we must have  $\gamma(C_\ell) \geq 1/18$ . Moreover, observe that the number of vertices  $N_\ell$  in  $G_\ell$  is at most  $N_0 \cdot (D^4)^\ell = \text{poly}(N)$ , so considering paths of length  $O(\log N)$  will suffice to decide  $s$ - $t$  connectivity in  $G_\ell$ .

To show that the algorithm can be implemented in logarithmic space, we argue that the edge-rotation map of each  $G_k$  can be computed with only  $O(1)$  more space than the edge-rotation map of  $G_{k-1}$ , so that  $G_\ell$  requires space  $O(\log N) + O(\ell) = O(\log N)$ . Since the inductive claim here refers to sublogarithmic differences of space (indeed  $O(1)$  space) and sublogarithmic space is model-dependent (even keeping a pointer into the input requires logarithmic space), we will refer to a specific model of computation in establishing it. (The final result, that UNDIRECTED S-T CONNECTIVITY is in  $\mathbf{L}$ , is, however, model-independent.) Formally, let  $\text{space}(G_k)$  denote the workspace needed to compute the edge-rotation map of  $G_\ell$  on a multi-tape Turing machine with the following input/output conventions:

- Input Description:
  - Tape 1 (read-only): Contains the initial input graph  $G$ , with the head at the leftmost position of the tape.
  - Tape 2 (read-write): Contains the input pair  $(v, i)$ , where  $v$  is a vertex of  $G_i$  and  $i \in [D^2]$  is an index of the a neighbor on a *read-write* tape, with the head at the *rightmost* position of  $i$ . The rest of the tape may contain additional data.
  - Tapes 3+ (read-write): Blank worktapes with the head at the leftmost position.
- Output Description:
  - Tape 1: The head should be returned to the leftmost position.
  - Tape 2: In place of  $(v, i)$ , it should contain the output  $(w, j)$  where  $w$  is the  $i$ 'th neighbor of  $v$  and  $v$  is the  $j$ 'th neighbor of  $w$ . The head should be at the rightmost position of  $j$  and the rest of the tape should remain unchanged from its state at the beginning of the computation.
  - Tapes 3+ (read-write): Are returned to the blank state with the heads at the leftmost position.

With these conventions, it is not difficult to argue that  $\text{space}(G_0) = O(\log N)$ , and  $\text{space}(G_k) = \text{space}(G_{k-1}) + O(1)$ . For the latter, we first argue that  $\text{space}(G_{k-1}^2) = \text{space}(G_{k-1}) + O(1)$ , and then that  $\text{space}(G_{k-1}^2 \otimes H) = \text{space}(G_{k-1}^2) + O(1)$ . For  $G_{k-1}^2$ , we are given a triple  $(v, (i_1, i_2))$ , with the head on the rightmost position of  $i_2$ , and both  $i_1$  and  $i_2$  are elements of  $[D^2]$  (and thus of constant size). We move the head left to the rightmost position of  $i_1$ , compute the edge-rotation map of  $G_{k-1}$  on  $(v, i_1)$  so that the tape contents are now  $(w, j_1, i_2)$ . Then we swap  $j_1$  and  $i_2$ , and run the edge-rotation map of  $G_{k-1}$  on  $(w, i_2)$  to get  $(w, j_2, j_1)$ , which is the final output. For  $G_{k-1}^2 \otimes H$ , we are given a tuple  $((v, i), (a_1, a_2))$ , where  $v$  is a vertex of  $G_{k-1}^2$ ,  $i$  is a vertex of  $H$  (equivalently,

an edge-label for  $G_{k-1}^2$ ), and  $a_1, a_2$  are edge labels for  $H$ . Evaluating the rotation map requires two evaluations of the rotation map for  $H$  (both of which are “constant-size” operations) and one evaluation of the rotation map of  $G_{k-1}^2$ .

Thus we have proven:

---

**Theorem 4.42.** **UNDIRECTED S-T CONNECTIVITY** is in **L**.

---

We remark that proving  $\mathbf{RL} = \mathbf{L}$  in general remains open. The best deterministic simulation known for  $\mathbf{RL}$  is essentially  $\mathbf{L}^{3/2} = \mathbf{DSPACE}(\log^{3/2} n)$ , which makes beautiful use of known pseudorandom generators for logspace computation. (Unfortunately, we do not have space to cover this line of work in this survey.) Historically, improved derandomizations for **UNDIRECTED S-T CONNECTIVITY** have inspired improved derandomizations of  $\mathbf{RL}$  (and vice-versa). Since Theorem 4.42 is still quite recent (2005), there is a good chance that we have not yet exhausted the ideas in it.

---

**Open Problem 4.43.** Show that  $\mathbf{RL} \subseteq \mathbf{L}^c$  for some constant  $c < 3/2$ .

---

Another open problem is the construction of *universal traversal sequences* — fixed walks of polynomial length that are guaranteed to visit all vertices in any connected undirected regular graph of a given size. (See Example 3.8 and Open Problem 3.9.) Using the ideas from the algorithm above, it is possible to obtain logspace-constructible, polynomial-length universal traversal sequences for all regular graphs that are *consistently labelled* in the sense that no pair of distinct vertices have the same  $i$ 'th neighbor for any  $i \in [D]$ . For general labellings, the best known universal traversal sequences are of length  $N^{O(\log N)}$  (and are constructible in space  $O(\log^2 N)$ ).

---

**Open Problem 4.44 (Open Problem 3.9, restated).** Give an explicit construction of universal traversal sequences of polynomial length for arbitrarily labelled undirected graphs (or even for an arbitrary labelling of the complete graph!).

---

We remark that handling general labellings (for “pseudorandom walk generators” rather than universal traversal sequences) seems to be the main obstacle in extending the techniques of Theorem 4.42 to prove  $\mathbf{RL} = \mathbf{L}$ .

## 4.5 Exercises

**Problem 4.1.** (Bipartite vs. Nonbipartite Expanders) Show that constructing bipartite expanders is equivalent to constructing (standard, nonbipartite) expanders. That is, show how given an explicit construction of one of the following, you can obtain an explicit construction of the other:

- (1)  $D$ -regular  $(\alpha N, A)$  expanders on  $N$  vertices for infinitely many  $N$ , where  $\alpha > 0$ ,  $A > 1$ , and  $D$  are constants independent of  $N$ .
- (2)  $D$ -regular (on both sides)  $(\alpha N, A)$  bipartite expanders with  $N$  vertices on each side for infinitely many  $N$ , where  $\alpha > 0$ ,  $A > 1$ , and  $D$  are constants independent of  $N$ .

(Your transformations need not preserve the constants.)

---

---

**Problem 4.2.** (More Combinatorial Consequences of Spectral Expansion) Let  $G$  be a graph on  $N$  vertices with spectral expansion  $\gamma = 1 - \lambda$ . Prove that:

- (1) The *independence number*  $\alpha(G)$  is at most  $\lambda N$ , where  $\alpha(G)$  is defined to be the size of the largest independent set, i.e. subset  $S$  of vertices s.t. there are no edges with both endpoints in  $S$ .
- (2) The *chromatic number*  $\chi(G)$  is at least  $1/\lambda$ , where  $\chi(G)$  is defined to be the smallest number of colors for which the vertices of  $G$  can be colored s.t. all pairs of adjacent vertices have different colors.
- (3) The *diameter* of  $G$  is  $O(\log_{1/\lambda} N)$ .

Recall that computing  $\alpha(G)$  and  $\chi(G)$  exactly are **NP**-complete problems. However, the above shows that for expanders, nontrivial bounds on these quantities can be computed in polynomial time.

---

**Problem 4.3.** (A “Constant-Sized” Expander)

- (1) Let  $\mathbb{F}$  be a finite field. Consider a graph  $G$  with vertex set  $\mathbb{F}^2$  and edge set  $\{((a, b), (c, d)) : ac = b + d\}$ . That is, we connect vertex  $(a, b)$  to all points on the line  $y = ax - b$ . Prove that  $G$  is  $|\mathbb{F}|$ -regular and  $\lambda(G) \leq 1/\sqrt{|\mathbb{F}|}$ . (Hint: consider  $G^2$ .)
  - (2) Show that if  $|\mathbb{F}|$  is sufficiently large (but still constant), then by applying appropriate operations to  $G$ , we can obtain a base graph for the expander construction given in Section 4.3.3, i.e. a  $(D^8, D, 7/8)$  graph for some constant  $D$ .
- 

**Problem 4.4.** (The Replacement Product) Given a  $D_1$ -regular graph  $G_1$  on  $N_1$  vertices and a  $D_2$ -regular graph  $G_2$  on  $D_1$  vertices, consider the following graph  $G_1 \textcircled{R} G_2$  on vertex set  $[N_1] \times [D_1]$ : vertex  $(u, i)$  is connected to  $(v, j)$  iff (a)  $u = v$  and  $(i, j)$  is an edge in  $G_2$ , or (b)  $v$  is the  $i$ 'th neighbor of  $u$  in  $G_1$  and  $u$  is the  $j$ 'th neighbor of  $v$ . That is, we “replace” each vertex  $v$  in  $G_1$  with a copy of  $G_2$ , associating each edge incident to  $v$  with one vertex of  $G_2$ .

- (1) Prove that there is a function  $g$  such that if  $G_1$  has spectral expansion  $\gamma_1$  and  $G_2$  has spectral expansion  $\gamma_2$ , then  $G_1 \textcircled{R} G_2$  has spectral expansion  $g(\gamma_1, \gamma_2, D_2) > 0$ . (Hint: Note that  $(G_1 \textcircled{R} G_2)^3$  has  $G_1 \textcircled{Z} G_2$  as a subgraph.)
  - (2) Show how to convert an explicit construction of constant-degree (spectral) expanders into an explicit construction of degree 3 (spectral) expanders.
  - (3) Prove that a dependence on  $D_2$  in Part 1 is necessary by showing that  $\gamma(G_1 \textcircled{R} G_2) = O(1/D_2)$  for sufficiently large  $N_1$ .
-

---

**Problem 4.5.** (Bounds on the Zig-Zag Product) In this problem, you will show some limitations on the zig-zag product, namely that its expansion is limited by that of the two composed graphs. Show that:

- (1) The vertex expansion of  $G \otimes H$  is at most the degree of  $H$ . (That is, find a “small” set  $S$  which does not expand by more than the degree of  $H$ . By small, we mean that  $|S|$  should be sublinear in the number of vertices of  $G \otimes H$ .)
  - (2)  $\lambda(G \otimes H) \geq \lambda(G)$ .
  - (3)  $\lambda(G \otimes H) \geq \lambda(H)^2$  if the labelling of clouds is consistent, i.e. if  $e = (u, v)$  is the  $i$ 'th edge leaving  $u$ , then it is also the  $i$ 'th edge leaving  $v$ . (In other words, the same vertex of  $H$  is mapped to the two vertices of  $G \otimes H$  corresponding to  $e$  (one in cloud  $u$  and one in cloud  $v$ .)
- 

**Problem 4.6.** (Unbalanced Vertex Expanders and Data Structures) Consider a  $(K, (1 - \varepsilon)D)$  bipartite vertex expander  $G$  with  $N$  left vertices,  $M$  right vertices, and left degree  $D$ .

- (1) For a set  $S$  of left vertices, a  $y \in N(S)$  is called a *unique neighbor* of  $S$  if  $y$  is incident to exactly one edge from  $S$ . Prove that every left-set  $S$  of size at most  $K$  has at least  $(1 - 2\varepsilon)D|S|$  unique neighbors.
- (2) For a set  $S$  of size at most  $K/2$ , prove that at most  $|S|/2$  vertices outside  $S$  have at least  $\delta D$  neighbors in  $N(S)$ , for  $\delta = O(\varepsilon)$ .

Now we'll see a beautiful application of such expanders to data structures. Suppose we want to store a small subset  $S$  of a large universe  $[N]$  such that we can test membership in  $S$  by probing just 1 bit of our data structure. A trivial way to achieve this is to store the characteristic vector of  $S$ , but this requires  $N$  bits of storage. The hashing-based data structures mentioned in Section 3.5.3 only require storing  $O(|S|)$  words, each of  $O(\log N)$  bits, but testing membership requires reading an entire word (rather than just one bit.)

Our data structure will consist of  $M$  bits, which we think of as a  $\{0, 1\}$ -assignment to the right vertices of our expander. This assignment will have the following property.

**Property II:** For all left vertices  $x$ , all but a  $\delta = O(\varepsilon)$  fraction of the neighbors of  $x$  are assigned the value  $\chi_S(x)$  (where  $\chi_S(x) = 1$  iff  $x \in S$ ).

- (3) Show that if we store an assignment satisfying Property II, then we can probabilistically test membership in  $S$  with error probability  $\delta$  by reading just one bit of the data structure.
- (4) Show that an assignment satisfying Property II exists provided  $|S| \leq K/2$ . (Hint: first assign 1 to all of  $S$ 's neighbors and 0 to all its nonneighbors, then try to correct the errors.)

It turns out that the needed expanders exist with  $M = O(K \log N)$  (for any constant  $\varepsilon$ ), so the size of this data structure matches the hashing-based scheme while admitting 1-bit probes.

However, note that such bipartite vertex expanders do *not* follow from explicit spectral expanders as given in Theorem 4.37, because the latter do not provide vertex expansion beyond  $D/2$  nor do they yield highly imbalanced expanders (with  $M \ll N$ ) as needed here. But later in this survey, we will see how to explicitly construct expanders that are quite good for this application (specifically, with  $M = K^{1.01} \cdot \text{polylog}N$ ).

---

**Problem 4.7.** (Error Reduction For Free\*) Show that if a problem has a **BPP** algorithm with constant error probability, then it has a **BPP** algorithm with error probability  $1/n$  that uses *exactly* the same number of random bits.

---

## 4.6 Chapter Notes and References

A detailed coverage of expander graphs and their applications is given by Hoory, Linial, and Wigderson [HLW].

The first papers on expander graphs appeared in conferences on telephone networks. Specifically, Pinsker [Pin] proved that random graphs are good expanders, and used these to demonstrate the existence of graphs called “concentrators.” Bassalygo [Bas] improved Pinsker’s results, in particular giving the general tradeoff between the degree  $D$ , expansion factor  $A$ , and set density  $\alpha$  mentioned after Theorem 4.4. The first computer science application of expanders (and “superconcentrators”) came in an approach by Valiant [Val] to proving circuit lower bounds. A early and striking algorithmic application was the  $O(\log n)$ -depth sorting network by Ajtai, Komlós, and Szemerédi [AKS2], which also illustrated the usefulness of expanders for derandomization.

The fact that spectral expansion implies vertex expansion and edge expansion was shown by Tanner [Tan] (for vertex expansion) and Alon and Milman [AM] (for edge expansion). The converses are discrete analogues of Cheeger’s Inequality for Riemannian manifolds [Che1], and various forms of these were proven by Alon [Alo1] (for vertex expansion) and Jerrum and Sinclair [JS] (for edge expansion in undirected graphs and, more generally, conductance in reversible Markov chains) and Mihail [Mih] (for edge expansion in regular digraphs and conductance in non-reversible Markov chains).

The “Ramanujan” upper bound on spectral expansion given by Theorem 4.11 was proven by Alon and Boppana (see [Alo1, Nil]). Theorem 4.12, stating that random graphs are asymptotically Ramanujan, was conjectured by Alon [Alo1], but was only proven recently by Friedman [Fri]. Kahale [Kah] proved that Ramanujan graphs have expansion roughly  $D/2$  for small sets.

Forms of the Expander Mixing Lemma date back to Alon and Chung [AC2], who considered the case that  $T = S^c$ . The converse to the Expander Mixing Lemma (Theorem 4.16) is due to Bilu and Linial [BL]. For more on quasirandomness, see [CGW, AS] for the case of dense graphs and [CG2, CG3] for sparse graphs.

The sampling properties of random walks on expanders were analyzed in a series of works starting with Ajtai, Komlós, and Szemerédi [AKS3]. The hitting bound of Theorem 4.17 is due to Kahale [Kah], and the Chernoff Bound for expander walks (cf., Theorem 4.22) is due to Gillman [Gil2]. Our proof of the Chernoff Bound is inspired by that of Healy [Hea], who also provides some other variants and generalizations. Problem 4.7 is due to Karp, Pippenger, and Sipser [KPS], who initiated the study of randomness-efficient error reduction of randomized algorithms.

The first explicit construction of constant-degree expanders was given by Margulis [Mar1], albeit with a nonconstructive proof that the spectral expansion is bounded away from 0. Construction 4.23 is due to Gabber and Galil [GG], who simplified Margulis’ construction and gave an elementary proof of expansion, with an explicit bound on the second eigenvalue. Ramanujan graphs (Construction 4.25) were constructed independently by Lubotzky, Phillips, and Sarnak [LPS] and Margulis [Mar2]. For more on Ramanujan graphs and the mathematical machinery that goes into their analysis, see the books [Lub, DSV].

The zig-zag product and the expander constructions of Section 4.3.3 are due to Reingold, Vadhan, and Wigderson [RVW]. Our analysis of the zig-zag product is from [RTV], which in turn builds on [RV], who used matrix decomposition (Lemma 4.19) for analyzing other graph products. Earlier uses of graph products in constructing expanders include the use of the tensor product in [Tan]. Problem 4.4, on the replacement product, is from [RVW], and can be used in place of the zig-zag product in both the expander constructions and the UNDIRECTED S-T CONNECTIVITY algorithm (Algorithm 4.41). Independently of [RVW], Martin and Randall [MR] proved a “decomposition theorem” for Markov chains that implies a better bound on the spectral expansion of the replacement product.

There has been substantial progress on Open Problem 4.38. Bilu and Linial [BL] give a mildly explicit construction achieving  $\lambda(G) = \tilde{O}(\sqrt{D})$ , Ben-Aroya and Ta-Shma [BT] give a fully explicit construction achieving  $\lambda(G) = D^{1/2+o(1)}$ , and Spielman et al. [BSS] give a mildly explicit construction of a *weighted* graph achieving  $\lambda(G) = O(\sqrt{D})$ .

Constant-degree bipartite expanders with expansion  $(1 - \varepsilon) \cdot D$  have been constructed by Capalbo et al. [CRVW]. Alon and Capalbo [AC1] have made progress on Open Problem 4.40 by giving an explicit construction of nonbipartite constant-degree “unique-neighbor” expanders (see Problem 4.6).

The deterministic logspace algorithm for UNDIRECTED S-T CONNECTIVITY (Algorithm 4.41) is due to Reingold [Rei]. Saks and Zhou [SZ] showed that  $\mathbf{RL} \subseteq \mathbf{L}^{3/2}$ , making nontrivial use of Nisan’s pseudorandom generator for space-bounded computation [Nis]. This was slightly improved by Armoni [Arm], who showed that  $\mathbf{RL} \subseteq \mathbf{DSPACE}((\log^{3/2} n)/\sqrt{\log \log n})$ , which remains the best known derandomization of  $\mathbf{RL}$ .

Based on Algorithm 4.41, explicit polynomial-length universal traversal sequences for “consistently labelled” regular digraphs, as well as “pseudorandom walk generators” for such graphs, were constructed in [Rei, RTV]. (See also [RV].) In [RTV], it is shown that pseudorandom walk generators for arbitrarily labelled regular digraphs would imply  $\mathbf{RL} = \mathbf{L}$ . The best known explicit construction of a full-fledged universal traversal sequence is due to Nisan [Nis] has length  $n^{O(\log n)}$ , and can be constructed in time  $n^{O(\log n)}$  and space  $O(\log^2 n)$ .

Problem 4.3, Part 1 is a variant of a construction of Alon [Alo2]; Part 2 is from [RVW]. The result of Problem 4.6, on bit-probe data structures for set membership, is due to Buhrman, Miltersen, Vinodchandran, and Miltersen [BMRV].