Unconditional Pseudorandomness: Expanders and Extractors

October 31, 2023

Contents

1 Random walks and eigenvalues

We consider random walks on undirected graphs. The graphs can have "multi edges" (*i.e.* there may be multiple parallel edges between the same set of vertices), and self-loops.

We first recall some basic facts from linear algebra. We use the linear space $\mathbb{R}^n$. If $u, v \in \mathbb{R}^n$ are two vectors, then their *inner product* is

$$\langle u, v \rangle = \sum_{i=[n]} u_i v_i.$$

The vectors are orthogonal if their inner product is 0. The L_2-norm of any vector $v \in \mathbb{R}^n$, denoted $||v||$, is

$$||v|| = \sqrt{\langle v, v \rangle}.$$

A *unit vector* is one whose L_2-norm is 1.

The *Pythagorean theorem* says that if $u, v \in \mathbb{R}^n$are orthogonal, then $||u+v||^2 = ||u||^2 + ||v||^2$.

The L_1-norm of a vector $v \in \mathbb{R}^n$, denoted $|v|$, is $|v| = \sum_{i \in [n]} |v_i|$.

The relation between these norms is as follows.

Fact 1. *For every $v \in \mathbb{R}^n$, we have*

$$\frac{|v|}{\sqrt{n}} \leq ||v|| \leq |v|$$

Proof. The upper bound on $||v||$ follows from the following observation. We have

$$||v||^2 = \sum_{i \in [n]} v_i^2 = \sum_{i \in [n]} |v_i|^2 \leq \left(\sum_{i \in [n]} |v|_i \right)^2,$$

where the last inequality is due to the fact that the last term contains an additional term $\sum_{i,j \in [n], i<j} 2|v|_i |v|_j$ which is non-negative. Conversely, we have,

$$|v| = \sum_{i \in [n]} |v_i| = \sum_{i \in [n]} |v_i \times 1| \leq \left(\sum_{i \in [n]} |v_i|^2 \right) \left(\sum_{i \in [n]} 1^2 \right) = n||v||,$$

using the Cauchy-Schwartz inequality. □

1.1 Distributions as vectors, and the spectral gap $\lambda(G)$

Let G be a d-regular graph (with self-loops and parallel edges), and $\boldsymbol{p}$be a probability distribution on the vertices of G.

Pick a vertex i from G at random according to the distribution p. Now, let q be the distribution over the vertices defined over the neighbors $N(i)$ of the selected vertex i.

Then for every vertex $1 \leq j \leq n$, $q(j) = \sum_{i \in N(j)} p(i) \times \frac{1}{d}$.

Thus, $q = Ap$, where A is the normalized adjacency matrix of G-$A[i][j]$ is defined to be the number of edges between i and j, divided by $1/d$. We call A the *random-walk matrix* of G. Since G is an undirected graph, A is a symmetric matrix. Since G is d-regular, the sum of entries in each row and each column is exactly one. Such a matrix is called a *doubly stochastic matrix*.

By induction, it is easy to show that the probability distribution over the vertices when we start a random walk at vertex i and take ℓ steps is $A^\ell e_i$.

Definition 2. Let $\mathbf{1}$ be the vector $\left(\frac{1}{n}, \ldots, \frac{1}{n}\right) \in \mathbb{R}^n$. Denote by $1^\perp$the set of vertices orthogonal to 1. The parameter $\lambda(G)$ is defined to be

$$\lambda(G) = \max_{\substack{v \in 1^\perp \\ ||v||=1}} ||Av||.$$

We now show that $\lambda(G)$ is the second eigenvalue of the normalized adjacency matrix of a d-regular graph. This needs some pre-requisite results.

Lemma 3. *All eigenvalues of a Hermetian matrix are real.*

Proof. Let M be a Hermetian matrix, *i.e.* $M = M^*$(entries of $M^*$are complex conjugates of the corresponding entries of M). Let x be an eigenvector corresponding to an eigenvalue λ of M. Then

$$\begin{aligned}\langle \lambda x, x\rangle &= \lambda^* x^* x \\ &= x^* M^* x \\ &= x^* M x \\ &= \lambda x x^*,\end{aligned}$$

since $M = M^*$for a Hermetian matrix. Since $x \neq 0$, we have $\lambda = \lambda^*$, implying that λ is real. □

Hence the eigenvalues of the normalized adjacency matrices of undirected graphs are real. We now have a special property for the eigenvalues of d-*regular* undirected graphs. Here, the real eigenvalues are upper bounded in absolute value by 1.

Lemma 4. *Let A be the normalized adjacency matrix of a d-regular graph, and λ be an arbitrary eigenvalue of A. Then $|\lambda| \leq 1$.*

Proof. We show that the spectral norm of A is at most 1. Recall that the spectral norm is

$$||A|| = \max\{||Av|| : v \in \mathbb{R}^n, ||v|| = 1\} = \max\{|\lambda| : \lambda \text{is an eigenvalue of } A\}.$$

First we show that $||A|| \leq n^2$.Note that all powers of A are also stochastic. Now, since

$$\langle w, Bz\rangle = \langle B^* w, z\rangle$$

and $\langle w, z\rangle \leq ||w||||z||$ by the Cauchy-Schwarz inequality, we have □

The above lemma stated that the largest absolute value of any eigenvalue for the adjacency matrix of a d-regular graph is 1. The following states that 1 is indeed

Lemma 5. *If A is the adjacency matrix of a d-regular graph, then its largest eigenvalue is 1.*

Proof. HW2 $\square$

Lemma 6. *1 is an eigenvalue of the normalized adjacency matrix of a d-regular graph.*

Proof. It suffices to observe that $u = (1/n, \dots, 1/n)$ is an eigenvector corresponding to the eigenvalue 1. Indeed, we have, for every $i \in [n]$, $(Au)_i = \sum_{j\in[n]} a_{ij}u_j = 1/n\left(\sum_{j\in[n]} a_{ij}\right) = 1/n = u_i$. $\square$

Hence, we know that the absolute values of the eigenvalues of the normalized adjacency matrix of a d-regular graph can be sorted as $1 = |\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_n|$,with corresponding eigenvectors $\mathbf{1}, v_2, \dots, v_n$. Also, since $1^{\perp} = \text{Span}\{v_2, \dots, v_n\}$, the value $\lambda(G)$ will be maximized by the vector v_2, hence, $\lambda(G) = |\lambda_2|$.

The quantity $1 - \lambda(G)$ (*i.e.* the difference between the first and the second eigenvalues) of the normalized adjacency matrix of a d-regular graph is called the *spectral gap* of the matrix. Its importance is that it controls the rate at which iterated applications of A to a probability vector p causes it to converge to $\mathbf{1}$in L_2-norm.

Lemma 7. *Let G be a regular graph, and p be a probability distribution over its vertices. Then for every $\ell \in \mathbb{N}$, we have*

$$||A^{\ell}p - 1|| \leq \lambda^{\ell}.$$

Since $\lambda < 1$, we conclude that iterated applications of A to any distribution p will lead to the uniform distribution on the vertices.

Every connected graph has a non-trivial spectral gap.

Lemma 8. *Let G be a d-regular connected graph with n vertices, and self-loops at each vertex. Then $\lambda(G) \geq 1 - \frac{1}{12n^2}$.*

Proof. Let $\epsilon = \frac{1}{6n^2}$. Let $u \perp 1$ be a unit vector (*i.e.* $\sum_{i\in[n]} u_i = 0$ and $||u|| = 1$: note here that u is not a probability vector), and let $v = Au$.

We show that $||v|| \leq 1 - \frac{\epsilon}{2}$. If $||v|| \leq 1 - \frac{\epsilon}{2}$, then $||v||^2 \leq 1 + \epsilon^2/2 - 2\epsilon \leq 1 - \epsilon$, hence, $||v||^2 \leq 1 - \epsilon$. Thus it suffices to show that $1 - ||v||^2 \geq \epsilon$. Since u is a unit vector, we show that $||u||^2 - ||v||^2 \geq \epsilon$.

We first show that

$$||u||^2 - ||v||^2 = \sum_{i\in[n]} \sum_{j\in[n]} A_{i,j}(u_i - v_j)^2. \tag{1}$$

Indeed,

$$\begin{aligned}\sum_{i,j} A_{ij}\left(u_i - v_j\right)^2 &= \sum_{i,j} A_{i,j}u_i^2 - 2\sum_{i,j} A_{ij}u_iv_j + \sum_{i,j} A_{ij}v_j^2 \\ &= ||u||^2 - 2\langle Au, v\rangle + ||v||^2 \\ &= ||u||^2 - 2\langle v, v\rangle + ||v||^2 \\ &= ||u||^2 - ||v||^2\end{aligned}$$

Thus, we have established (1)

Now, we show that $\sum_{i,j} A_{i,j}\left(u_i - v_j\right)^2 \geq \epsilon$. Since u is a unit vector whose coordinates sum to zero, there must be coordinates $u_i > 0$ and $u_j < 0$ such that at least one of these co-ordinates must have absolute value $\geq \frac{1}{\sqrt{n}}$, implying $u_i - u_j \geq \frac{1}{\sqrt{n}}$. Let us rename u_i as u_1 and u_j as u_{D+1}, where D is the diameter of G. Since G is connected, there is a path $u_1, u_2, \ldots, u_{D+1}$. We have

$$\begin{aligned}\frac{1}{\sqrt{n}} &\leq u_1 - u_{D+1} \\ &= (u_1 - v_1) + (v_1 - u_2) + \cdots + (v_D - u_{D+1}) \\ &\leq |u_1 - v_1| + |v_1 - u_2| + \cdots + |v_D - u_{D+1}| \\ &\leq \sqrt{\left(u_1 - v_1\right)^2 + \left(v_1 - u_2\right)^2 + \cdots + \left(v_D - u_{D+1}\right)^2}\sqrt{2D+1}\end{aligned}$$

where the last two inequalities follow from the relation between the $L_2$and L_1norms of the vector $(u_1 - v_1, v_1 - u_2, \ldots, v_D - u_{D+1})$. Thus

$$\sum_{i,j} A_{ij}\left(u_i - v_j\right)^2 \quad \geq \frac{1}{dn(2D+1)} \geq \frac{1}{2dn^2},$$

using the trivial estimate $D \leq n - 1$. Thus, we have $\lambda(G) \geq 1 - \frac{1}{4dn^2}$. Using the tighter estimate $D \leq \frac{3n}{(d+1)}$ for regular graphs yields the result as stated. □

2 Expander Graphs

We now give two approaches to the notion of an expander graph - an algebraic approach, and a combinatorial approach. We show the relationships between the two.

2.1 Algebraic Definition (Spectral expansion)

First, we give the definition of an expander graph based on the parameter $\lambda(G)$.

Definition 9. If G is an n-vertex d-regular graph with $\lambda(G) < 1$, then we say that G is an (n, d, λ)-graph.
A family of graphs $(G_n)_{n\in\mathbb{N}}$ is an *expander graph family* if there are constants d and $\lambda < 1$ such that for every $n \in \mathbb{N}$, $G_n$is an (n, d, λ)-graph.

The smallest value of $\lambda(G)$ is $(1 - o(1))\frac{2\sqrt{d-1}}{d}$, which are attained by *Ramanujan graphs.*

Definition 10. We say that the expander graph family $(G_n)_{n\in\mathbb{N}}$ is *explicit* if there is a polynomial-time aglorithm that, given 1^n, outputs the *adjacency matrix* of G_n. We say that $(G_n)_{n\in\mathbb{N}}$ is *strongly explicit* if there is a polynomial-time algorithm that, given $\langle n, v, i\rangle \in \mathbb{N} \times V \times [d]$, outputs the index of the i^{th}neighbor of the vertex v in G_n. [n and d are as in Definition 9.

2.2 Combinatorial Definition (Edge expansion)

Let $G = (V, E)$ denote a graph with the set of vertices V and the set of edges E.

Definition 11. A constant (d)-degree regular graph $G = (V, E)$ is an *expander* if $\exists c > 0\ \forall S \subseteq V$, $|S| \leq |V|/2$ implies that

$$\frac{|\{(a, b) \in E \mid a \in S, b \in V - S\}|}{|\{(a, b) \in E \mid a \in S\}|} \geq c.$$

That is, a constant fraction of edges incident on vertices of S must be edges from S to its complement.

We now outline a probabilistic argument for the existence of expander graphs. This construction is only one of many possible constructions.

Let $V = \{v_1, \ldots, v_n\}$ be the set of vertices in a graph G we are about to construct. Consider permutations $\pi_1, \ldots, \pi_d : [n] \to [n]$ chosen independently and uniformly at random from the set of all permutations. We add edges to the graph as follows. For each permutation π_i, $1 \leq i \leq d$, if $\pi_i(k) = \ell$, then we add the edge $\{k, \ell\}$ to G. Then G is a $2d$-regular graph. [For each permutation π_i and each k, we have $\pi_i(k) = \ell$, and for some integer j (possibly ℓ), we have $\pi_i(j) = k$. If j, k, ℓ are all distinct, this adds two edges per permutation to the vertex k. If $j = \ell \neq k$, then this adds parallel edges $\{k, \ell\}$. If $j = k = \ell$, then this adds two self-loops to k.]

Theorem 12. *(Existence of expander graphs) The above construction yields a* $\left(n, 2d, \frac{1}{10}\right)$ *expander.*

Proof. (Outline) The proof is a probabilistic argument. We show that the probability that G is not an expander is less than 1. This implies that there is some expander following the above construction.

Consider $S \subset V$, $|S| \leq n/2$, and let $\overline{S} = V \setminus S$. Let $S = \{v_{s_1}, \ldots, v_{s_k}\}$. Select a vertex v_{s_i}. Define the indicator random variables

$$X_j = \begin{cases} 1 & \text{if the jth neighbor of } \ v_{s_i} \text{is in } \overline{S} \\ 0 & \text{otherwise.} \end{cases}$$

Then $\mathbb{E}\left[E(v_{s_i}, \overline{S})\right] = \mathbb{E}\left[\sum_{j \in [2d]} X_{ij}\right] = \frac{d}{n}(n-k)$.

Using Hoeffding bound, it is possible to bound the probability that a random choice of $\pi_1, \ldots, \pi_d$ violates the expander property for the given cut S. Using union bound, we show that the probability that there is *some* cut which violates the expander property, is strictly less than 1. $\square$

2.3 Expander Mixing Lemma

Lemma 13. *Let G be a d-regular undirected graph, and S, T be disjoint subsets of vertices. Then*

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

Proof. Note that

$$|E(S,T)| = 1_S^\top A 1_T$$

and

$$|S||T| = 1_s^\top J 1_T,$$

where J is the all-one matrix. Hence,

$$\begin{aligned} \left| E(S,T) - \frac{d|S||T|}{|V|} \right| &= \left| 1_S^\top A 1_T - \frac{d}{|V|} 1_s^\top J 1_T \right| \\ &= \left| 1_S^\top \left(A - \frac{d}{|V|} J \right) 1_T \right| \\ &\leq \left\| 1_S^\top \right\| \left\| A - \frac{d}{|V|} J \right\| \|1_T\| \\ &= \lambda(G)\sqrt{|S||T|}. \end{aligned}$$

$\square$

This explains why expander graphs are considered "pseudorandom" - the number of edges from any S to any disjoint T is similar to what you would expect in a random graph.

2.4 Algebraic expansion implies combinatorial expansion

Lemma 14. *If $G$is an (n, d, λ)-expander, then it is an $\left(n, d, \frac{(1-\lambda)}{2}\right)$-edge expander.*

Proof. Let A be the normalized adjacency matrix of G. Let S and T be arbitrary disjoint subsets of vertices in G. Using the property that the second largest (in absolute value) eigenvalue of A is λ, we show that at least $\frac{1-\lambda}{2}$fraction of the edges in S cross to T.

Define $x \in \mathbb{R}^n$by

$$x_i = \begin{cases} +|T| & i \in S \\ -|S| & i \in T \\ 0 & \text{otherwise.} \end{cases}$$

Then,

$$\|x\|^2 \,| = |S||T|^2 + |T||S|^2 = |S||T|\left(|S| + |T|\right)$$

and

$$\sum_{i \in [n]} x_i = \sum_{i \in S} |T| - \sum_{i \in T} |S| = |S||T| - |T||S| = 0.$$

Thus $x \in 1^\top$.

Consider the sum of weights obtained by placing the weight $\frac{(x_i - x_j)^2}{2d}$ on every edge $\{i, j\}$ in the graph G. If the edge is inside S or inside T, then this weight is 0, otherwise it is $\left(|S| + |T|\right)^2$. That is, if we define $Z = \sum_{i,j} A_{ij} \left(x_i - x_j\right)^2$, then it is clear that

$$Z = \frac{2}{d}\ E(S, T)\ \left(|S| + |T|\right)^2.$$

On the other hand, we also have

$$Z = \sum_{i,j} A_{ij} x_{i^2}\ - 2 \sum_{i,j} A_{ij} x_i x_j\ + \sum_{i,j} A_{ij} x_j^2 \quad = \quad 2\,\|x\|^2 - 2\langle x, Ax\rangle.$$

Since $||Ax|| \leq \lambda ||x||$, we have $2\langle x, Ax\rangle \leq 2\lambda ||x||^2$. This yields that

$$\frac{1}{d}\ E(S, T)\ \left(|S| + |T|\right)^2 \geq (1 - \lambda)\,\|x\|^2\,,$$

hence,

$$\frac{1}{d}\ E(S,T)\ \ (|S|+|T|) \geq (1-\lambda)|S||T|,$$

thus yielding

$$E(S,T)\ \ \geq (1-\lambda)\ d\ \frac{|S||T|}{|S|+|T|} \geq (1-\lambda)\ d\ \frac{|S||T|}{|V|},$$

as required. □

2.5 Combinatorial expansion implies algebraic expansion

Lemma 15. *If G is an (n,d,ρ) edge expander, then its second largest eigenvalue is at most $\left(1-\frac{\rho^2}{2}\right)$.*

Proof. Let $G=(V,E)$ be an n-vertex d-regular graph such that for every $S \subset V$, $|S|<n/2$, there are $\rho d|S|$ edges between S and its omplement. Let A be G's normalized adjacency matrix.

Let λ be the ssecond largest eigenvalue of A. We need to establish that $\lambda \leq \left(1-\frac{\rho^2}{2}\right)$. Then there is a vector $u \perp \mathbf{1}$ such that $Au=\lambda u$. Since $u \perp 1$, we have $\langle u,1\rangle = \sum_{i\in[n]} u_i = 0$, it is clear that u has positive and negative co-ordinates. Write $u=v+w$ where v is non-zero in co-ordinates where u are positive, and w is non-zero in co-ordinates where u has negative entries. Assume, without loss of generality, that v contains at most $n/2$ non-zero entries (otherwise, we can use $-u$). Define

$$Z=\sum_{i,j} A_{ij}\left|v_i^2-v_j^2\right|.$$

If we show that

$$\begin{aligned} Z &\geq 2\rho\left\|v\right\|^2 \\ Z &\leq \sqrt{8(1-\lambda)}\left\|v\right\|^2, \end{aligned} \tag{2}$$

then it follows that

$$\lambda \geq 1-\frac{\rho^2}{2}.$$

§ We first show that $Z \geq 2\rho \|v\|^2$. Sort the co-ordinates of v so that $v_1 \geq v_2 \geq \cdots \geq v_n$, with $v_i = 0$ for all $i \geq \frac{n}{2}$. Note that

$$v_i^2 - v_j^2 = \left(v_i^2 - v_{i+1}^2\right) + \left(v_{i+1}^2 - v_{i+2}^2\right) + \cdots + \left(v_{j-1}^2 - v_j^2\right).$$

Hence,

$$Z = \sum_{i\in[n]} \sum_{j\in[n]} A_{ij} \left|v_i^2 - v_j^2\right| = \sum_{i\in[n]} \sum_{j\in[i+1,n]} 2A_{ij} \left(v_i^2 - v_j^2\right) = 2 \sum_{i\in[n]} \sum_{j\in[i+1,n]} \sum_{k\in[i+1,j-1]} A_{ij} \left(v_k^2 - v_{k+1}^2\right).$$

We now estimate the sum above. For every edge $\{i, j\}$, for every $k \in [i, j)$, the term $\left(v_k^2 - v_{k+1}^2\right)$ appears once with a weight $2/d$. Since $v_k = 0$ for $k \geq n/2$, this means that the above sum is

$$\frac{2}{d} \sum_{k\in[n/2]} |E([k], [k+1, n]| \ \left(v_k^2 - v_{k+1}^2\right) \geq \frac{2}{d} \sum_{k\in[n/2]} \rho d k \left(v_k^2 - v_{k+1}^2\right).$$

We have, using the fact that $v_i = 0$ for all $i \geq n/2$,

$$\sum_{k\in[n/2]} k \left(v_k^2 - v_{k+1}^2\right) = v_1^2 - v_2^2 + 2v_2^2 - 2v_3^2 + \cdots + (n-1)v_{n-1}^2 - (n-1)v_n^2 = \|v\|^2.$$

Hence

$$Z \geq \frac{2}{d} d\rho \|v\|^2,$$

establishing the lower bound for Z.

Proof. We now show the upper bound $Z \leq \sqrt{8(1-\lambda)}\,\|v\|^2$. Since $u = v+w$ as stated before, we have $\langle v, w\rangle = 0$. Further, $Au = \lambda u$. Thus, we have

$$\langle Av, v\rangle + \langle Aw, v\rangle = \langle A(w+v), v\rangle = \langle Au, v\rangle = \langle \lambda u, v\rangle = \langle \lambda(v+w), w\rangle = \lambda\langle v, v\rangle = \lambda\|v\|^2.$$

Since $\langle Aw, v\rangle$ is not positive, we have that $\frac{\langle Av,v\rangle}{\|v\|^2} \geq \lambda$. Hence,

$$1-\lambda \geq 1 - \frac{\langle Av, v\rangle}{\|v\|^2} = \frac{\|v\|^2 - \langle Av, v\rangle}{\|v\|^2} = \frac{\sum_{i,j} A_{ij}\,(v_i - v_j)^2}{2\|v\|^2}.$$

Multiplying the numerator and the denominator with the term $\sum_{i,j} A_{ij}\,(v_i + v_j)^2$, we get

$$\frac{\left(\sum_{i,j} A_{ij}\,(v_i - v_j)^2\right)\left(\sum_{i,j} A_{ij}\,(v_i + v_j)^2\right)}{2\|v\|^2\left(\sum_{i,j} A_{ij}\,(v_i + v_j)^2\right)} \geq \frac{\left(\sum_{i,j} A_{ij}\,(v_i - v_j)\,(v_i + v_j)\right)^2}{2\|v\|^2\left(\sum_{i,j} A_{ij}\,(v_i + v_j)^2\right)}$$

using the Cauchy Schwarz inequality. Hence,

$$1-\lambda \geq \frac{\left(\sum_{i,j} A_{ij}\left(v_i^2 - v_j^2\right)\right)^2}{2\|v\|^2\left(\sum_{i,j} A_{ij}\,(v_i + v_j)^2\right)} = \frac{Z^2}{2\|v\|^2 2\left(\|v\|^2 + \langle Av, v\rangle\right)} \geq \frac{Z^2}{8\|v\|^2},$$

establishing the upper bound. □

This completes the proof. □

2.6 Error reduction using expanders

One way of reducing the probability of error in a probabilistic algorithm for a decision problem is to execute it k times independently, and taking the majority output. By Chernoff bounds, it is possible to show that if the probability of error of one execution is at most $1/3$, then the probability of the *majority* of k executions being wrong is $2^{-\Omega(k)}$. If one execution of the algorithm uses m random coins, then the multiple executions take mk random coins.

Using expanders, we can reduce the number of random coins to $m+O(k)$ random coins.

The idea is as follows. Let G be a $(2^m, d, 1/10)$-graph from a strongly explicit expander graph family.[1] Note especially that the number of vertices

[1]This means $\lambda \leq 1/10$, using the algebraic definition.

in the graph G is equal to the total number of possible random coins used by the algorithm. Then, we do as follows:

let v_1 at random For i=2 to k do: From the d neighbors of v_{i-1}, choose a vertex v_i at random. [Requires $O(log\ d)$ random bits] Run the algorithm with the random coins being $v_1, \ldots, v_m$ and output the majority

Algorithm 1: Error reduction using a random walk on an expander

We now analyze the probability of error of the above algorithm. Assume that we have an algorithm which makes one-sided errors: for strings not in the language, the original algorithm always says no, and for strings in the language, the algorithm says "yes" with probability 2/3.

In the following theorem, setting $\beta = 1/3$ and $\lambda = 1/10$ implies that the probability that the above algorithm will reject an input in the language is bounded by $2^{-\Omega(k)}$.

Theorem 16. *(Expander walks)*

Let G be an (n, d, λ) expander. Let $B \subseteq [n]$ have at most βn vertices, where $\beta \in (0, 1)$. Let $X_1, \ldots, X_k \in [n]$ be random variables denoting a $k-1$-step random walk in G starting from X_1, with X_1 being uniformly chosen from $[n]$. Then,

$$Pr\left[\wedge_{i-1}^{k}(X_i \in B)\right] \ \le\ \left((1-\lambda)\sqrt{\beta}\ +\ \lambda\right)^{k-1}.$$

To get some intuition about what the theorem says, think of B being a "bad set" which we want to escape. If B is very large, *i.e.* $\beta \approx 1$, then the right-side expression is approximately 1, so the above bound is useless - so the bad set cannot be very large. If, on the other hand, the bad set is, say, $0.1n$, then the probability that all the randomly chosen vertices are in the bad set slowly decays with k - the rate of decay is not as fast as λ^{k-1}, but it is still some γ^{k-1},where $\gamma = (1-\lambda)\sqrt{\beta}\ +\ \lambda < 1$.

Proof. For $1 \le i \le k$, let B_i denote the event $X_i \in B$.

Let $T : \mathbb{R}^n \to \mathbb{R}^n$be defined by $(Tu)_i = u_i$ if $i \in B$ and 0 otherwise. Then this transformation "zeroes out" coordinates not in B. It is easy to verify that for every probability vector p over $[n]$, we have Tp is a vector whose co-ordinates sum to the probability (according to p) of chosing a vertex in B. If we normalize so that the sum of coordinates of Tp is 1, then we obtain the cpnditional probability distribution of p conditioned on the event of chosing a vertex in B.

Hence, let $\mathbf{1} = \left(\frac{1}{n}, \ldots, \frac{1}{n}\right)$ be the uniform distribution, and let p^i be the conditional distribution of X conditioned on the events $B_1, \ldots, B_n$. Then

$$\begin{aligned}
p^1 &= \frac{1}{Pr[B_1]} T\mathbf{1} \\
p^2 &= \frac{1}{Pr[B_2|B_1]Pr[B_1]} TAT\mathbf{1} \\
&\ldots \\
p^i &= \frac{1}{Pr[B_i \mid B_1, \ldots, B_{i-1}]} (TA)^{i-1} T\mathbf{1}.
\end{aligned}$$

Since p is a probability vector, $|p| = 1$. Hence, $Pr\left[\wedge_{i-1}^{k} (X_i \in B)\right] = \left|(TA)^{k-1} T\mathbf{1}\right| \leq \sqrt{n}\|(TA)^{k-1} T\mathbf{1}\|$, where the last inequality is due to the relation between $L_1$and L_2 norms. It suffices now to show that

$$\sqrt{n}\|(TA)^{k-1} T\mathbf{1}\| \quad \leq \quad \left((1-\lambda)\sqrt{\beta} \ + \ \lambda\right).$$

We now assume the following fact, given as Lemma 21.4 in the book.
Let A be a random-walk matrix of an (n, d, λ)-expander graph G. Let J be the matrix with all entries equal to $1/n$. (This corresponds to the adjacency matrix of an n-clique with self-loops. Then

$$A = (1-\lambda)J + \lambda C,$$

where $\|C\| =: \max\{\|Av_2\| \ : \ v_2 \perp \mathbf{1}\} \leq 1$.
The use of this fact is as follows. For any probability vector p, we may view the probability vector Ap a convex combination of Jp (the uniform distribution) and Cp. This can be interpreted as Ap goes to Jp with probability $1-\lambda$ and Cp with probability λ.

Let TA be written as $(1-\lambda)TJ + \lambda TC$. Then

$$\begin{aligned}
\|TA\| &= (1-\lambda)\|TJ\| + \lambda\|TC\| \\
&= (1-\lambda)\sqrt{\frac{\beta n}{n^2}} + \lambda\|TC\|.
\end{aligned}$$

Since $\|T\| \leq 1$, we have $\|TC\| \leq 1$. Hence

$$\|TA\| \leq (1-\lambda)\sqrt{\beta} + \lambda,$$

whence

$$\left\|(TA)^{k-1} T\mathbf{1}\right\| \leq \left((1-\lambda)\sqrt{\beta} + \lambda\right)^{k-1} \frac{\sqrt{\beta}}{\sqrt{n}},$$

as required. $\square$

3 Undirected graph reachability in non-deterministic logspace

This section contains the result by Reingold that the randomized logspace algorithm for s-t connectivity in undirected graphs, which relies on random walks, can be completely derandomized.

Theorem 17. *(Reingold)*
$UPATH \in \boldsymbol{L}$.

making graph regular

We consider undirected graphs that have a lot of parallel edges. We can assume that we have 4-regular graphs, without loss of generality because of the following reason. If a vertex has degree at most 3, then we can add parallel self-loops. If a vertex v has degree greater d' than 3, then we can replace it with a cycle of d' vertices such that each of the edges incident on v now becomes incident on a unique vertex in the cycle. Since this is a local transformation, it can be performed in logspace.

checking connectivity in expanders

It is easy to check connectivity in expander graphs. Indeed, Lemma 7 shows that the probability distribution of the k^{th}vertex in a random walk is approximately $\frac{1}{n} \pm \sqrt{n}\lambda^k$. Then a random walk of length $k - O(\log n)$ from s will reach t with positive probability. Hence, If every connected component in G is an expander, then there is a number $k = O(\log n)$ such that every pair of connected vertices have a path of length $\leq k$ between them.

G, H and G_ks

Proof. Assume that the input graph G has degree d^{50} for some sufficiently large d so that there is an $(d^{50}, d/2, 0.01)$-expander H. We can ensure this by sufficiently many self-loops if necessary. Since $d^{50}$is a constant independent of the number of vertices in G, we can store H in $O(1)$ bits.

Let

$$G_0 = G$$
$$G_k = (G_{k-1} \circledR H)^{50},$$

replacement product

where $\circledR$ denotes the replacement product defined as follows.

Let R, R' be two graphs such that R has n vertices, and degree D and R' has D vertices and degree d'. The *(balanced) replacement product* of R and R', denoted $R \circledR R'$, is the nD-vertex, 2d'-degree graph defined as follows.

1. For every vertex u of R, the graph $R \circledR R'$ has a copy of R'.

2. If $\{u, v\}$ is an edge in R, then we place d' parallel edges between the i^{th}vertex in the copy of R' corresponding to u and the j^{th} vertex in the copy of R' corresponding to v.

degree and connectivity of G_k

Coming back to the construction, if G_{k-1}has N vertices and degree d^{50}, then $G_{k-1} \circledR H$ is a $d^{50}N$-vertex graph with degree d, hence $(G_{k-1} \circledR H)^{50}$ is a $d^{50}N$-vertex graph with degree d. If two vertices were connected in G_{k-1}, then they are connected in G_k, and if they were disconnected in G_{k-1}, they remain disconnected in G_k.

components in G_k are good expanders

We show that $G_{10\log n}$ is an expander, and this is an easy instance of $UPATH$. Specifically, we can show that every connected component in G_k is an $\left(d^{50}n, d^{20}, 1-\epsilon\right)$-expander, where $\epsilon = \min\left(\frac{1}{20}, \frac{1.5^k}{12n^2}\right)$.

Set $k = \log n$. Then $G_{10\log n}$is an expander with expansion parameter $\leq 1 - \frac{1}{20}$. Hence to find whether s and t are connected, we enumerate over all paths in $G_{10\log n}$.

Now we argue that, given the input $G = G_0$, we can perform a single step of a random walk in G_k. That is, given a descriptionof a vertex s in G_k, and an $i \in \left[d^{20}\right]$, we can compute the i^{th}neighbor of G_k.

indexing vertices

A vertex in G_{k-1}®H is represented by a pair $[u,v]$, $u \in V(G_{k-1})$, $v \in V(H)$, and the index of a neighbor is represented by a pair $[b,i]$, $b \in 0,1$, $i \in [d/2]$. If $b = 0$, then the designated neighbor is $[u,v']$ where v' is the i^{th}neigbor of u in H. Otherwise, $[b,i]$ designates $[u',v']$ such that u' is the v^{th} neighbor of u in G_{k-1} and v' is the index of u as a neighbor of u' in G_{k-1}. This can be computed by a recursive algorithm, in logspace. $\square$