CS 860 Spectral graph theory, Spring 2019, Waterloo.

Lecture 8: Expander graphs

First, we study a deterministic constructions of expander graphs, called the zig-zag product.

Then, we discuss some properties and various applications of expander graphs.

Expander graphs

There are different possible ways to define expander graphs.

- Combinatorically, expander graphs are graphs with very good connectivity, i.e. $|\delta(S)|/|S|$ is large for all $S \subseteq V$ with $|S| \leq |V|/2$ (edge expansion) or $|N(S)|/|S|$ is large for all $S \subseteq V$ (vertex expansion).
- Probabilistically, expander graphs are graphs in which random walks mix very rapidly.
- Spectrally, expander graphs are graphs with a large spectral gap, i.e. $\alpha_1 - \alpha_2$ is large.

As we have seen, the definitions are closely related.

Cheeger's inequality states that a graph has a large spectral gap if and only if its edge expansion is large.

And we have seen in L06 that lazy random walks mix quickly if and only if the spectral gap is large.

Actually, complete graphs are the best expander graphs in each of the above definitions, but we are interested in sparse expander graphs with linear number of edges, e.g. d-regular graphs for constant d.

It can be shown that a random d-regular graph is an expander graph with high probability using the combinatorial definition, by standard techniques (Chernoff bound and union bound).

Perhaps surprisingly, while almost every graph is an expander graph, it is very difficult to come up with a deterministic construction of expander graphs.

In constructing expander graphs, it turns out that the spectral definition is easier to work with.

We will use the following stronger spectral definition:

<u>Definition</u> Let G be a d-regular graph. Let $d = \alpha_1 \geq \alpha_2 \geq \dots \geq \alpha_n \geq -d$ be the eigenvalues of its adjacency matrix. We say G is an (n,d,ε)-graph if it has n vertices, d-regular and $\max\{\alpha_2, |\alpha_n|\} \leq \varepsilon d$.

The quantity $\alpha := \max\{\alpha_2, |\alpha_n|\}$ is called the spectral radius of the graph.

The smaller is the spectral radius, the stronger the graph is as an expander.

Combinatorically, $|\alpha_n|$ is small if and only if there is no induced subgraph close to a bipartite component.

Probabilistically, α is small if and only if the non-lazy random walks mix quickly.

It can be shown that the spectral radius of a random d-regular graph is $O(\sqrt{d})$ whp (see [HLW]) and we are interested in deterministic constructions of d-regular graphs with small spectral radius.

Deterministic constructions

While almost all d-regular graphs are expander graphs, it is counterintuitive that it is hard to construct expander graphs deterministically.

One possible explanation is that random graphs have high description complexity, while in deterministic constructions the graphs can be described in a succinct way.

There are some explicit constructions of d-regular expander graphs. most of them are algebraic.

- A family of 8-regular graphs G_m for every integer m. The vertex set is $V = \mathbb{Z}_m \times \mathbb{Z}_m$.

  The neighbors of vertex (x,y) is $(x+y, y)$, $(x-y, y)$, $(x, y+x)$, $(x, y-x)$, $(x+y+1, y)$, $(x-y+1, y)$, $(x, y+x+1)$, $(x, y-x+1)$, where all additions are modulo m.

  Note that this family is very explicit, meaning that the neighbors of a vertex is very easy to compute, and this is very useful in applications such as probability amplification.

  Gabber and Galil proved that $\alpha \leq 5\sqrt{2} < 8$. The proof uses Fourier analysis; see chapter 8 of [HLW].

- A family of 3-regular p-vertex graph for every prime p. The vertex set is $V = \mathbb{Z}_p$, and a vertex x is connected to $x+1$, $x-1$ and its multiplicative inverse x^{-1} (for vertex 0 its inverse is 0), where the additions are modulo p. The proof uses some deep results in number theory.

- The main source of explicit deterministic is from Cayley graphs, which are graphs defined by groups. Some of the strongest expanders, called Ramanujan graphs with spectral radius $\leq 2\sqrt{d-1}$, are from Cayley graphs and the proofs require sophisticated mathematical tools.

- In the last part of the course, we will study a new way to show the existence of Ramanujan graphs using combinatorial and probabilistic methods, through interlacing family of polynomials.

We will present a combinatorial construction of expander graphs, known as the zig-zag product, whose proof is simpler and more intuitive, but it is less explicit as the above constructions.

Combinatorial constructions

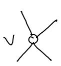
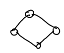
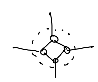
The general idea of the combinatorial constructions is to construct a bigger expander graph from smaller expander graphs.

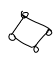
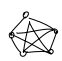
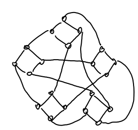
The base case is usually just a constant size complete graph.

Let G be an (n,k,ε_1)-graph and H be an (k,d,ε_2)-graph.

A natural product $G \circledR H$ of the two graphs G and H is called the replacement product.

The vertex set of the product is $V(G) \times V(H)$, the Cartesian product.

Each vertex v in G (of degree k) is replaced by a copy of H, and each edge incident on v is incident on a different vertex of H, e.g. v  H  $\Rightarrow$  . $\leftarrow$ a "cloud"

For example, H is  and G is  , then $G \circledR H$ is  .

Intuitively, $G \circledR H$ is an expander if G and H are.

Consider a set $S \subseteq V(G \circledR H)$.

If S has either large or small intersection with each "cloud", then S should have large expansion because of the large expansion of G (i.e. a S is like a set in G).

If S has medium intersections with many "clouds", then S should have large expansion because of the large expansion of H (i.e. many crossing edges within each cloud).

However, it is not clear how to make this intuition precise, as there seems to be no clean way to decompose a set's contribution into its contribution from G and its contribution from H.

In a way, the spectral proof that we are going to see soon can be thought of as a linear algebraic approach to carry out this idea in a more general setting.

Zig-zag product

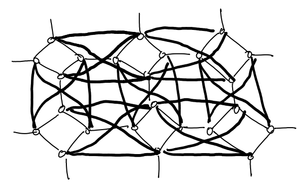
The actual construction that we will analyze is slightly more complicated.

The zig-zag product of G and H, denoted by $G \textcircled{z} H$, has the same vertex set $V(G) \times V(H)$ as the replacement product.

The edges of the zig-zag product is obtained from a length three walk in the replacement product:

there is an edge $(u_1,v_1)-(u_2,v_2)$ in $G ⓩ H$ if and only if there exists w such that $(u_1,v_1)-(u_1,w)$, $(u_1,w)-(u_2,y)$, $(u_2,y)-(u_2,v_2)$ are edges in the replacement product $G ⓡ H$, where $(u_1,w)-(u_2,y)$ is the unique edge incident on (u_1,w) going out the cloud of u_1 in the replacement product.

In other words, each edge in $G ⓩ H$ corresponds to a length three walk in $G ⓡ H$, where the first step is within a cloud, the second step is (the unique way) across two clouds, and the third step is within (the second) cloud.

For example, if G is a grid and H is the 4-cycle, then $G ⓡ H =$ 

where the edges in $G ⓩ H$ are the thick edges.

The intuition that the zig-zag product is an expander graph is from random walks.

Each edge in $G ⓩ H$ corresponds to a random walk in H, a deterministic step in G, and a random step in H.

We can think of the first two steps as going to a random neighbor cloud, and the third step corresponds to moving to a random neighbor within the (second) cloud.

Since both G and H are expanders and so fast mixing, after a few steps of random walks in $G ⓩ H$, we won't know which cloud we are in and also won't know the location within the cloud, and so $G ⓩ H$ is also fast mixing and hence an expander graph.

<u>**Theorem**</u> (zig-zag product) Let G be an (n,k,ε_1)-graph and H be an (k,d,ε_2)-graph. Then $G ⓩ H$ is an $(nk, d^2, \varepsilon_1+\varepsilon_2+\varepsilon_2^2)$-graph.

Before we see the proof of the theorem, let's first see how it can be used to construct bigger and bigger constant degree expander graphs.

A natural graph operation to improve the spectral radius is the graph powering operation.

The k-th power $G^k=(V,E')$ is a graph on the same vertex as G and we add an edge (u,v) in E' for every path (not necessarily simple) of length exactly k in G from u to v.

Note that the adjacency matrix of G^k is A^k where A is the adjacency matrix of G (perhaps this should be used as the definition of G^k).

So, if G is an (n,d,ε)-graph, then G^k is an (n,d^k,ε^k)-graph.

While the spectral radius has improved significantly, the degree also increases significantly.

The idea is to use the zig-zag product to decrease the degree while not losing expansion much.

Let H be a $(d^4, d, 1/16)$-graph. We can prove its existence by probabilistic method. Since d is a constant, we can find it by exhaustive search in constant time.

Using the building block H, we inductively define G_n by $G_1 = H^2$ and $G_{n+1} = G_n^2 ⓩ H$.

We claim that G_n is a $(d^{4n}, d^2, 1/4)$-graph for all $n \geq 1$.

The base case is true, clearly.

Assume G_n is a $(d^{4n}, d^2, 1/4)$-graph, then G_n^2 is a $(d^{4n}, d^4, 1/16)$-graph, and $G_n^2 ⓩ H$ is a $(d^{4(n+1)}, d^2, 1/4)$-graph by the zig-zag theorem.

Proof of zig-zag theorem

It is not difficult to check that $G ⓩ H$ has nk vertices and is d^2-regular.

First, we write down the walk matrix Z of the zig-zag product.

Let B be the walk matrix of H and $\tilde{B}$ be n copies of B on the diagonal.

The matrix $\tilde{B}$ correspond to the movements within the clouds.

The steps between clouds are deterministic: we move from a vertex (v,i) to the unique vertex (u,j) with $v \neq u$.

The walk matrix for this step is a permutation matrix P with $P_{(v,i),(u,j)} = 1$ for each "external" edge and zero otherwise.

By the definition of the zig-zag product, we have $Z = \tilde{B} P \tilde{B}$.

The graph $G ⓩ H$ is regular, and so $\vec{1}_{nk}$ is an eigenvector of Z with eigenvalue one.

To prove the zig-zag product theorem, we will prove that $|f^T Z f| / \|f\|^2 \leq \varepsilon_1 + 2\varepsilon_2 + \varepsilon_2^2$ for all $f \perp \vec{1}_{nk}$, that is, the Rayleigh quotient (with respect to Z) is small for all vectors orthogonal to the first eigenvector.

We decompose f to two vectors to apply the results in G and H.

This is where the power of linear algebra comes from, as in the larger domain of $\mathbb{R}^{nk}$ there is a natural operation for the decomposition, while in the combinatorial domain it is not clear how to decompose.

Define f_G as the average of f on clouds, i.e. $f_G(x,i) = \frac{1}{k} \sum_{j=1}^{k} f(x,j)$, so that two vertices in the same cloud have the same value in f_G.

Define $f_H = f - f_G$. Note that f_H sums to zero in each cloud, i.e. $\sum_{j=1}^{k} f_H(x,j) = 0$ for each $x \in G$.

Roughly speaking, we will use the assumption in G to show that $|f_G^T Z f_G| \le \varepsilon_1 \|f_G\|_2^2$ and use the assumption in H to argue that $|f_H^T Z f_H| \le \varepsilon_2 \|f_H\|_2^2$.

Note that $|f^T Z f| = |f^T \tilde{B} P \tilde{B} f| = |(f_G + f_H)^T \tilde{B} P \tilde{B} (f_G + f_H)| \le |f_G^T \tilde{B} P B f_G| + 2|f_G^T \tilde{B} P \tilde{B} f_H| + |f_H^T \tilde{B} P \tilde{B} f_H|$.

Since $B\vec{1}_k = \vec{1}_k$, it follows that $\tilde{B} f_G = f_G$ as vertices in the same cloud have the same value.

So, $|f^T Z f| \le |f_G^T P f_G| + 2|f_G^T P \tilde{B} f_H| + |f_H^T \tilde{B} P \tilde{B} f_H|$.

We will use the assumption on G to bound the first term, the assumption on H to bound the third term, and a simple bound for the second term.

First, we bound the third term.

Since the spectral radius of B is ε_2, we have $\|Bx\| \le \varepsilon_2 \|x\|$ for any $x \perp \vec{1}_k$. To see this, write $x = \sum_{i=2}^{k} c_i v_i$ where v_i are the orthonormal eigenvectors of B with eigenvalues λ_i. Note that $c_1 = 0$ as $v_1 = \vec{1}/\sqrt{k}$ and $x \perp \vec{1}$. Then, $\|Bx\|_2^2 = \|B(\sum_{i=2}^{k} c_i v_i)\|_2^2 = \|\sum_{i=2}^{k} c_i \lambda_i v_i\|_2^2 = \sum_{i=2}^{k} c_i^2 \lambda_i^2 \le \varepsilon_2^2 \sum_{i=2}^{k} c_i^2 = \varepsilon_2^2 \|x\|_2^2$.

This implies that $\|\tilde{B} f_H\| \le \varepsilon_2 \|f_H\|$ as the sum of each cloud in f_H is zero.

Therefore, the third term is $|f_H^T \tilde{B} P \tilde{B} f_H| \le \|\tilde{B} f_H^T\|_2 \|P \tilde{B} f_H\|_2$ by Cauchy-Schwarz

$= \|\tilde{B} f_H^T\|_2^2$ as $\|Px\| = \|x\|$ since P is a permutation matrix

$\le \varepsilon^2 \|f_H^T\|_2^2$.

Similarly, the second term is $2|f_G^T P \tilde{B} f_H| \le 2\|f_G\| \|P\tilde{B} f_H\| = 2\|f_G\| \|\tilde{B} f_H\| \le 2\varepsilon_2 \|f_G\| \|f_H\|$.

We will prove in the following claim that $\|f_G^T P f_G\| \le \varepsilon_1 \|f_G\|^2$.

This would imply that $|f^T Z f| \le \varepsilon_1 \|f_G\|^2 + 2\varepsilon_2 \|f_G\| \|f_H\| + \varepsilon_2^2 \|f_H\|^2$

$\le \varepsilon_1 \|f_G\|^2 + \varepsilon_2 (\|f_G\|^2 + \|f_H\|^2) + \varepsilon_2^2 \|f_H\|^2 \le (\varepsilon_1 + \varepsilon_2 + \varepsilon_2^2) \|f\|_2^2$, where the last inequality holds because $f_G \perp f_H$ and so $\|f\|^2 = \|f_G + f_H\|^2 = \|f_G\|^2 + \|f_H\|^2$ and so $\|f_G\| \le \|f\|$ and $\|f_H\| \le \|f\|$.

This will complete the proof. It remains to prove the following claim.

Claim $|f_G^T P f_G| \le \varepsilon_1 \|f_G\|^2$.

proof In short, the LHS is equal to the quadratic form of the walk matrix of G.

To work with $|f_G^T P f_G|$, we "contract" each cloud to a single vertex.

Define $g : V(G) \to \mathbb{R}$ as $g(v) = \sqrt{k}\, f_G(v,i)$. Note that $\|g\|_2^2 = \|f_G\|_2^2$.

Note also that $f_G^T P f_G = g^T W g$ where W is the walk matrix of G, as each edge $(u,i)-(v,j)$ contributes

$f_G(u,i) \cdot f_G(v,j)$ to $f_G^T P f_G$ while the corresponding edge $u-v$ in G contributes

$(\sqrt{k}\, f_G(u,i))(\frac{1}{k})(\sqrt{k}\, f_G(v,j)) = f_G(u,i) \cdot f_G(v,j)$ to $g^T W g$.

Therefore, $f_G^T P f_G / \|f_G\|^2 = g^T W g / \|g\|^2$.

Since $f \perp \vec{1}$, we have $f_G \perp \vec{1}$ and thus $g \perp \vec{1}$.

As G is an (n,k,ε_1)-graph, we have $f_G^T P f_G / \|f_G\|^2 = g^T W g / \|g\|^2 \le \varepsilon_1$. □

Properties of expander graphs

A useful property of expander graphs is that it behaves like random graphs.

Let $E(S,T) = \{ (u,v) : u \in S, v \in T, uv \in E \}$ where an edge with $u \in S \cap T$ and $v \in S \cap T$ is counted twice.

In a random graph where each edge presents with probability d/n, we expect that $|E(S,T)|$ is close to $\frac{d}{n}|S||T|$.

The expander mixing lemma says that in an expander graph $|E(S,T)|$ is close to this number.

Theorem (Expander mixing lemma) Let $G=(V,E)$ be a d-regular graph with spectral radius α.

Then for every $S \subseteq V$ and $T \subseteq V$, $\left| |E(S,T)| - \frac{d|S||T|}{n} \right| \le \alpha \sqrt{|S||T|}$.

Proof Let χ_S and χ_T be the characteristic vectors of S and T, i.e. $\chi_S(i) = 1$ if $i \in S$ and zero otherwise.

Let $v_1, v_2, \dots, v_n$ be an orthonormal basis of eigenvectors. Recall that $v_1 = \vec{1}/\sqrt{n}$ and $\alpha_1 = d$.

Write $\chi_S = \sum_{i=1}^n a_i v_i$ and $\chi_T = \sum_{i=1}^n b_i v_i$. So $a_1 = \langle \chi_S, v_1 \rangle = \frac{|S|}{\sqrt{n}}$ and $b_1 = \langle \chi_T, v_1 \rangle = \frac{|T|}{\sqrt{n}}$.

Note that $|E(S,T)| = \chi_S^T A \chi_T = \sum_{i \ge 1} \alpha_i a_i b_i = d|S||T|/n + \sum_{i \ge 2} \alpha_i a_i b_i$.

So,
$$\begin{aligned}
\left| |E(S,T)| - \frac{d|S||T|}{n} \right| &\le \left| \sum_{i \ge 2} \alpha_i a_i b_i \right| \le \sum_{i \ge 2} |\alpha_i||a_i||b_i| \le \alpha \sum_{i \ge 2} |a_i||b_i| && \text{by spectral radius} \\
&\le \alpha \|a\| \|b\| && \text{by Cauchy-Schwarz} \\
&= \alpha \|\chi_S\| \|\chi_T\| = \alpha \sqrt{|S||T|}. \quad \square
\end{aligned}$$

Independent sets and chromatic number

Let $X \subseteq V$ be an independent set, i.e. there is no edge between any pair of vertices of X.

By definition, $|E(X,X)| = 0$.

By the expander mixing lemma with $S=T=X$, we have $d|X|^2/n \le \alpha |X|$, and thus $|X| \le \alpha n / d$.

For graphs with $\alpha \le \varepsilon d$, this implies that the maximum size of an independent set is $\le \varepsilon n$, and it follows that the minimum chromatic number of such graphs is $\geq \frac{1}{\varepsilon}$.

Diameter

We claim that the diameter of an expander graph is $O(\frac{\log n}{1-\varepsilon})$, for graphs with spectral radius $\leq \varepsilon d$.

By Cheeger's inequality, $\phi(S) \geq \frac{\lambda_2}{2} = \frac{d-\alpha_2}{2d}$ where λ_2 is the second eigenvalue of the normalized Laplacian matrix and α_2 is the second largest eigenvalue of A.

This implies that $\frac{|\delta(S)|}{d|S|} \geq \frac{d-\alpha_2}{2d} \geq \frac{d-\varepsilon d}{2d}$, and thus $|\delta(S)| \geq (\frac{1-\varepsilon}{2}) d|S|$ for all $|S| \leq n/2$.

Let $N(S) := \{ i \in V-S \mid i \sim j \text{ for some } j \in S \}$ be the neighbor set of S.

Then $|N(S)| \geq |\delta(S)|/d \geq (\frac{1-\varepsilon}{2})|S|$ for all S with $|S| \leq n/2$.

Let v be a vertex and $B(v,r) := \{ i \in V \mid \text{dist}(i,v) \leq r \}$ be the ball of radius r around v.

Then the above statement implies that $B(v,r) \geq (1+\frac{1-\varepsilon}{2}) B(v,r-1) \geq \ldots \geq (1+\frac{1-\varepsilon}{2})^r B(v,1) = (1+\frac{1-\varepsilon}{2})^r$

If we let $\lambda = 1-\varepsilon$ be the spectral gap, then $B(v,r) > \frac{n}{2}$ for $r = O(\frac{\log n}{\log(1+\lambda)}) = O(\frac{\log n}{\lambda})$.

This implies that the distance between any pair of vertices is at most $2r$, hence the diameter is $O(\frac{\log n}{\lambda})$.

This is not surprising at all, since we know that the mixing time is $O(\frac{\log n}{\lambda})$ from L06.

Vertex expansion

One can give a bound on the vertex expansion through edge expansion as we did above.

Here we prove a stronger bound as the Tanner's theorem.

Theorem (Tanner) Let G be a d-regular graph with spectral radius $\leq \varepsilon d$.

Then $|N(S)|/|S| \geq 1/(c(1-\varepsilon^2)+\varepsilon^2)$ for $|S| = cn$ where $0 < c \leq \frac{1}{2}$ is a constant.

Proof Let $S \subseteq V$ with $|S| = cn$, χ_S be the characteristic vector of S, and A be the adjacency matrix of G with eigenvalues $\alpha_1 \geq \alpha_2 \geq \ldots \geq \alpha_n$.

Write $\chi_S = \sum_{i \geq 1} a_i v_i$, where $v_i = \vec{1}/\sqrt{n}$ and $a_1 = \langle \chi_S, v_1 \rangle = |S|/\sqrt{n}$.

Consider $\|A\chi_S\|_2^2$. Let $N[S] = N(S) \cup S$.

On one hand, $\|A\chi_S\|_2^2 = \sum_{v \in N[S]} |S \cap N(v)|^2 \geq \left(\sum_{v \in N[S]} |S \cap N(v)| \right)^2 / \left(\sum_{v \in N[S]} 1^2 \right)$ by Cauchy-Schwartz

$$= (d|S|)^2 / |N[S]|.$$

On the other hand, $\|A\chi_S\|_2^2 = \| \sum_{i \geq 1} a_i \alpha_i v_i \|_2^2 = \sum_{i \geq 1} a_i^2 \alpha_i^2 \leq \frac{d^2|S|^2}{n} + \alpha^2 \sum_{i \geq 2} a_i^2 = \frac{d^2|S|^2}{n} + \alpha^2(\|\chi_S\|_2^2 - a_1^2)$

$$= \frac{d^2|S|^2}{n} + (\varepsilon d)^2 \left(|S| - \frac{|S|^2}{n} \right) = d^2 cn(c + \varepsilon^2(1-c)).$$

Combining, we get $d^2 cn|S|/N[S] \le d^2 cn(c+\varepsilon^2(1-c)) \Rightarrow |N[S]|/|S| \ge \frac{1}{c+\varepsilon^2(1-c)}$. $\square$

It shows that when $c \ll \varepsilon^2$, then $|N[S]|/|S| \gtrsim \frac{1}{\varepsilon^2}$ which implies that $|N[S]|$ can be much larger than $|S|$ when $|S|$ is small enough.

Converse of expander mixing lemma

Interestingly, Bilu and Linial proved a converse of the expander mixing lemma, showing that it comes close in characterizing graphs with small spectral radius.

Theorem Let G be a d-regular graph on n vertices. Suppose that for any $S, T \subset V(G)$ with $S \cap T = \emptyset$ it holds that $\left| |E(S,T)| - \frac{d|S||T|}{n} \right| \le \alpha \sqrt{|S||T|}$.

Then all but the largest eigenvalue of G are bounded, in absolute value by $O(\alpha(1+\log\frac{d}{\alpha}))$.

The proof is based on the following lemma.

Lemma Let A be an $n \times n$ real symmetric matrix such that the l_1-norm of each row in A is at most d, and all diagonal entries of A are in absolute value $O(\alpha \log(d/\alpha)+1)$.

Assume that for any two vectors, $u, v \in \{0,1\}^n$ with $\mathrm{supp}(u) \cap \mathrm{supp}(v) = \emptyset$ it holds that $\frac{|u^T A v|}{\|u\|\,\|v\|} \le \alpha$.

Then the spectral radius of A is $O(\alpha \log(d/\alpha)+1)$.

The theorem follows from the lemma by letting $A = A(G) - \frac{d}{n}J$, where $A(G)$ is the adjacency matrix of G.

Spectral Gap

How large can the spectral gap be?

Let $\alpha = \max\{|\alpha_2|, |\alpha_n|\}$ where α_i is the i-th largest eigenvalue of the adjacency matrix.

There are graphs, called Ramanujan graphs, that $\alpha \le 2\sqrt{d-1}$.

This is essentially tight, as the following theorem by Alon and Boppana showed.

Theorem $\alpha_2 \ge 2\sqrt{d-1} - O\left(\frac{\sqrt{d-1}}{\mathrm{diam}(G)-4}\right)$.

The theorem implies that if we have a family of constant degree graphs each has $\alpha_2 \le \alpha$, then $\alpha \ge 2\sqrt{d-1}$.

We give an easy proof that $\alpha \ge \sqrt{d}\cdot(1-o_n(1))$.

We know that $nd \le \text{trace}(A^2)$, as each edge contributes one to $\text{trace}(A^2)$.

On the other hand, we know that $\text{trace}(A^2) = \sum_{i=1}^{n} \alpha_i^2 \le d^2 + (n-1)\alpha^2$.

Therefore, $d^2 + (n-1)\alpha^2 \ge nd$ and thus $\alpha^2 \ge d(n-d)/(n-1)$ and hence $\alpha \ge \sqrt{d}\sqrt{\frac{n-d}{n-1}}$.

Random walks in expander graphs

Let $G=(V,E)$ be a d-regular graph. Let's assume $\alpha = \varepsilon d$ and $\varepsilon \le 1/10$. Let $X \subseteq V$ with $|X| \le |V|/100$.

Let v_0 be the initial random vertex, and $v_1, v_2, \dots, v_t$ be the vertices produced by the t steps of random walk.

Let $S = \{i : v_i \in X\}$. We choose v_0 as a uniformly random vertex, each vertex of probability $1/n$.

Theorem $\Pr(|S| > t/2) \le \left(\frac{2}{\sqrt{5}}\right)^{t+1}$.

First, we set up the matrix formulation of the problem.

The initial distribution is $\pi = \vec{1}/n$.

Let χ_X and $\chi_{\bar{X}}$ be the characteristic vectors of X and $\bar{X}$, where $\bar{X} = V - X$.

Let I_X be the diagonal matrix with a 1 in the i-th diagonal entry if $i \in X$, and similarly $I_{\bar{X}}$.

Let p be a probability distribution. Then $I_X p$ is the probability vector on X.

Then $q = W I_X \pi$ is the probability vector where the initial random vertex is in S, where $W = A/d$.

Then, the probability that the walk is in X at precisely the time steps in S is

$$\vec{1}^T I_{Z_t} W I_{Z_{t-1}} W I_{Z_{t-2}} W \dots I_{Z_2} W I_{Z_1} W \pi, \quad \text{where } Z_i = X \text{ if } i \in S \text{ and } Z_i = \bar{X} \text{ if } i \notin S.$$

We will prove that this probability is at most $(1/5)^{|S|}$.

This will imply that $\Pr(|S| > t/2) \le \sum_{|S| > t/2} \Pr(\text{the walk is in } X \text{ at precisely the times in } S)$ by union bound

$$\le 2^{t+1} \left(\tfrac{1}{5}\right)^{\frac{t+1}{2}} = \left(2/\sqrt{5}\right)^{t+1}.$$

Recall that $\|M\| = \max_y \|My\|/\|y\| = \max_y y^T M y / y^T y$ for symmetric M.

We can check that $\|I_X\| = \|I_{\bar{X}}\| = \|W\| = 1$.

We will prove that $\|I_X W\| \le 1/5$, and this would imply that the above probability is at most $(1/5)^{|S|}$.

To see this, $\vec{1}^T I_{Z_t} W I_{Z_{t-1}} W \dots I_{Z_2} W I_{Z_1} W \pi = \vec{1}^T (I_{Z_t} W)(I_{Z_{t-1}} W) \dots (I_{Z_2} W)(I_{Z_1} W)\pi$

$$\le \|\vec{1}^T\| \, \|(I_{Z_t} W)(I_{Z_{t-1}} W) \dots (I_{Z_2} W)(I_{Z_1} W)\pi\| \quad \text{Cauchy-Schwarz}$$

$$\le \|\vec{1}^T\| \left(\prod_{i=1}^{t} \|I_{Z_i} W\|\right) \|\pi\|$$

$$\le \|\vec{1}^T\| (1/5)^{|S|} \|\pi\| \quad \text{as } \|I_{Z_i} W\| \le 1/5 \text{ if } Z_i = X \text{ and}$$

$$\|I_{z_i} W\| \le \|I_{z_i}\|\|W\| = 1 \text{ if } z_i = X$$

$$\le (1/5)^{|S|} \quad \text{as } \|\vec{1}^T\| = \sqrt{n} \text{ and } \|\pi\| = 1/\sqrt{n}.$$

It remains to prove that $\|I_X W\| \le 1/5$.

Let y be any nonzero vector and write $y = c_1 v_1 + \dots + c_n v_n$ where $v_1 = \vec{1}/\sqrt{n}$ and $c_1 = \langle y, v_1 \rangle = \sum y_i / \sqrt{n}$, where $v_1, \dots, v_n$ are the orthonormal vectors of W with eigenvalues $1 = \lambda_1 \ge \lambda_2 \ge \dots \ge \lambda_n \ge -1$.

$$\|I_X W y\|^2 = \|I_X W (c_1 v_1 + \dots + c_n v_n)\|^2 \le \|I_X \sum_{i \ge 1} c_i \lambda_i v_i\|^2 \le 2\|I_X c_1 \lambda_1 v_1\|^2 + 2\|I_X \sum_{i \ge 2} c_i \lambda_i v_i\|^2$$

$$\le 2\left\|I_X \left(\frac{\sum y_i}{n} \vec{1}\right)\right\|^2 + 2\|I_X\|^2 \left\|\sum_{i \ge 2} c_i \lambda_i v_i\right\|^2$$

$$= 2|X| \left(\frac{\sum y_i}{n}\right)^2 + 2\sum_{i \ge 2} \|c_i \lambda_i v_i\|^2 \quad \text{by orthogonality}$$

$$\le 2|X| \cdot \frac{\|y\|^2}{n} + 2\sum_{i \ge 2} \|c_i \lambda_i v_i\|^2 \quad \text{by Cauchy Schwartz as } (\sum y_i)^2 \le (\sum y_i^2)(\sum 1^2)$$

$$\le \frac{1}{50}\|y\|^2 + 2\sum_{i \ge 2} c_i^2 \lambda_i^2$$

$$\le \frac{1}{50}\|y\|^2 + 2\varepsilon^2 \sum_{i \ge 2} c_i^2 \le \frac{1}{50}\|y\|^2 + \frac{1}{50}\|y\|^2$$

Thus $\|I_X W y\| \le \frac{1}{5}\|y\|^2$, finishing the proof.

Application: Probability amplification

Suppose we have a randomized algorithm with error probability $1/100$ by reading n random bits.

This means that among the 2^n n-bit strings, only $2^n/100$ are "bad" strings.

To amplify the success probability, one can pick k random n-bit strings, then the error probability is at most $(1/100)^k$ using kn random bits.

We show how to exponentially decrease the error probability while using only $n + ck$ bits for a constant c.

Construct a d-regular expander graph with 2^n vertices and $\varepsilon \le 1/10$.

In the first step, we use an n-bit random string, with error probability $\le 1/100$.

In the subsequent steps, instead of picking independent n-bit strings, we do a $(k-1)$-step random walk and use the strings corresponding to the vertices in the random walk as "random" n-bit strings.

After we try the k "random" strings, we use the majority answer as our answer.

What is the error probability of this algorithm?

We output the wrong answer if the wrong answer is the majority.

Since the number of bad strings is at most $2^n/100$, the error probability is at most $(2/\sqrt{5})^k$, by letting X to be the set of bad strings.

The number of random bits used is $n+(k-1)\log_2 d$ since each random neighbor can be chosen with $\log_2 d$ bits.

Note that it works for two-sided error randomized algorithms as well.

This is just one example, expander graphs have many applications in derandomization ; see [HLW].

Application : Constructing Efficient Objects

One can think of a d-regular expander graph as a very efficient object, as it only has very few edges but it achieves very high connectivity.

It is not surprising that expander graphs are useful in constructing efficient communication networks.

One example is the construction of superconcentrators, which are directed graphs with n input nodes I and n output nodes, and satisfying the strong connectivity property that there are k vertex disjoint paths for any k input nodes and any k output nodes and for any $k \le n$.

For instance, a complete bipartite graph satisfies this property, but it has $\Theta(n^2)$ edges.

Using expander graphs one can construct a superconcentrator with the following properties :

- it has $O(n)$ nodes, every node is of constant degree, and thus it has $O(n)$ edges.

A superconcentrator can be used as an efficient switching network. Besides, it can be used to design faster algorithms for computing network coding and computing matrix rank.

Another famous example of using expander graphs is to construct an optimal sorting network, with only $O(n\log n)$ edges and with depth $O(\log n)$.

Application: Undirected connectivity in log space

One important application of the zig-zag product is to solve the s-t connectivity problem in log space, i.e. to determine if there is a path from s to t in an undirected graph using minimal space.

As we will see laterr, it is easy to solve the problem if we are allowed to use randomness : just doing random walks for $O(n^3)$ steps would do.

There is a simple deterministic algorithm using $O(\log^2 n)$ space.

Reingold discovered a deterministic algorithm using $O(\log n)$ space using zip-zag product.

We briefly discuss the high level ideas here.

Suppose the graph is a d-regular expander graph for a constant d. Then G has diameter $O(\log n)$

Then we can enumerate all paths of length $O(\log n)$ in $O(\log n)$ space, since the graph is of constant degree.

The idea of Reingold's algorithm is to transform any graph G into a constant degree expander graph H such that s,t are connected in G iff s,t are connected in H.

We can reduce G into a constant degree graph by replacing each vertex of high degree by some low degree graph, say a cycle or an expander, just like what we did in the replacement product.

To improve the expansion, we construct the graph $G^8 \textcircled{z} H$, and it can be shown that the spectral gap doubles in the resulting graph.

So, we just need to repeat this construction $O(\log n)$ times to get constant spectral gap, as initially the spectral gap is at least $\frac{1}{n^2}$, which holds for any connected graph.

Now, we run the exhaustive search algorithm on this resulting constant degree expander graph.

Note that we don't construct this graph completely, as it requires too much space.

We just compute each edge on demand. The observation is that there are only $O(\log n)$ recursion levels for the construction, and in each level we just need constant space, since there are only three steps and the degree is constant.

So, to compute each edge, we only need $O(\log n)$ space in total, and we don't need to store the actual edges for the exhaustive search (we just need to store the current vertex), and the "index" of the edges in the path so far, where each index can be stored in constant space, as the graph is of constant degree).

The total space required is thus $O(\log n)$.

It is open whether there is a $O(\log n)$-space algorithm for directed s-t connectivity.

<u>**Reference**</u> Everything can be found in the excellent survey "expander graphs and their applications" by Hoory, Linial and Wigderson [HLW].