A Mathematical Introduction to Spectral Expansion

Peter M. R. Rasmussen

October 27, 2019

Abstract

In this introduction to spectral expander graphs, we broadly cover different aspects of the theory, taking care to formally define and explain the various concepts involved. First, we present classic results such as the Cheeger inequalities and the mixing lemma. Second, applications of random walks on spectral expander graphs are touched upon. Third, we prove a generalisation of the Alon-Boppana bound on spectral expansion. Fourth, the expected second largest eigenvalue in absolute value of the adjacency matrix of a random graph is bounded from above, and we show concentration bounds for the distribution of this quantity around its mean. Fifth, we present a construction of spectral expander graphs of non-constant degree from Abelian Cayley graphs. Sixth, we briefly discuss the construction of constant degree spectral expanders from the zig-zag product. Whereas many introductions to expander graphs are written for an audience with a background in computer science, this material is presented so as to be more broadly of interest to mathematicians.

Contents

1 Introduction

Expander graphs are well-connected and often sparse graphs. In applications of graphs or in the study of graph theory, one often requires or intuitively senses that the vertices of a graph are so *entangled* with edges that one cannot easily and nicely partition them, without severing many edges. This phenomenon, among other things, is captured by the notion of an expander graph. For intuition, consider the following concrete example. Picture a road network with cities as vertices and highways between pairs of cities as edges. Given that highways are expensive to build, suppose we still wish to maintain the ability to drive from any city to any other city (possibly passing through other cities on the way there), even while some number of highways are closed. Of course, closing all highways out of any one city disconnects the city from the remaining network. Thus, our requirement will rather be that the closing of any small number of highways leaves all but a few cities connected to each other. A road network satisfying these criteria would be an expander.

Let $G = (V, E)$ be a d-regular graph with vertices V and edges $E \subset V^2$. There are various mathematical characterisations of the above phenomenon. Most notably, G can be a vertex expander, an edge expander, or a spectral expander. Edge expanders are most similar in spirit to the example above. They require that for every subset, $S \subset V$, containing at most half the vertices of G, the number of edges between S and $V \setminus S$ is at least proportional to $|S|$. Stated in terms of the above analogy, this would imply that to disconnect m cities from the remaining cities, one would have to close a number of highways proportional to m. The definition of a vertex expander is similar in nature. A graph is a vertex expander if for every subset, $S \subset V$, containing at most half of the vertices of G, the number of neighbours of S in $V \setminus S$ is proportional to $|S|$. Both these definitions are straight-forward and intuitively corresponds well with the idea of an expanding graph. This brings us to spectral expander graphs. Spectral expanders satisfy that the second largest eigenvalue (in absolute value) of the adjacency matrix, A, of G is small relative to d. That the latter definition should imply expansion is nowhere close to intuitive. However, it turns out to be one of the most flexible ways to work with the notion of expansion. Among other things, spectral expansion implies both vertex and edge expansion. Our focus will be on spectral expanders, and on how their algebraic definition allows for a fine analysis of their properties.

Expander graphs have found wide usage in all corners of computer science, appearing in subfields as diverse as complexity theory, algorithms design, pseudorandomness, cryptography, and more. Within mathematics, expander graphs have mainly been studied by graph theorists, but applications of expander graphs also extend to other areas such as number theory, geometric group theory, and operator algebras.

The goal of this exposition is to present some of the gems of spectral expander graph theory to a mathematical audience. This entails both results, applications, and techniques. The paper is organised as follows. Section 2 contains preliminary definitions and results. After this, we get properly started in Section 3 where we touch upon some classic examples of the usefulness of spectral expansion. We cover the mixing lemma, the Cheeger inequalities, and include a couple of applications from computer science that are based on random walks on spectral expanders. Section 4 turns to bounding spectral expansion from below. In other words, it deals with how small the eigenvalues of the adjacency matrices associated with a family of graphs can be. We first present an elementary argument obtaining a bound on the second largest eigenvalue (in absolute value) of the adjacency matrix. We then proceed with a more advanced approach due to Davidoff, Sarnak, and Valette to establish deeper results regarding the distribution of the eigenvalues of the adjacency matrices of large regular graphs. The fact that random graphs are good expanders, is the topic

of Section 5. Contrary to how difficult it is to construct spectral expanders explicitly, a random, regular graph will be a nearly optimal spectral expander with high probability. We cover the first substantial result to this effect due to Broder and Shamir. Section 6 discusses Cayley graphs in the context of expander graphs and constructs a simple example of a good expander that the author found a direct application for in a recent paper. The final construction of an expander graph is lacking in a vital sense. The degree of the expanders constructed in these notes is not constant. Section 7 demonstrates how constant degree expanders can be constructed combinatorially through the zig-zag product of Reingold, Vadhan, and Wigderson – a classical result in spectral expander theory.

2 Preliminaries

We shall start by introducing basic notation, definitions, and results. The material on generating functions will only be required for Section 4, and the reader may therefore defer reading it until then.

2.1 Mathematical Notation

We start by introducing general mathematical notation.

Integers. We let $\mathbb{N} = \{1, 2, \dots\}$ denote the set of positive integers and $\mathbb{N}_0 = \{0, 1, \dots\}$ the set of non-negative integers. Furthermore, for $n \in \mathbb{N}_0$, we denote by $[n]$ the set $\{k \in \mathbb{N} \mid 1 \leq k \leq n\}$.

Indicator Functions. Let V be a set. For any subset $A \subset V$, the indicator function on A is denoted $1_A \colon V \to \{0, 1\}$.

Iverson Bracket. For some statement P, we shall use the notation $[P]$ as the indicator on whether P is true or not, e.g., $[2 + 2 = 5] = 0$ and $[\pi \in \mathbb{R}] = 1$. This is known as the *Iverson bracket*.

Indices. Let $n, m \in \mathbb{N}$ be positive integers, $v \in \mathbb{R}^n$ be an n-dimensional vector, and $A \in \mathbb{R}^{n \times m}$ be an n by m matrix. For $i \in [n]$ and $j \in [m]$ denote by v_i the ith entry of v and by A_{ij} the entry of A at the ith row and jth column. Furthermore, if V is a set and $x \in \ell^2(V)$, then we denote by x_v the evaluation of x at $v \in V$.

2.2 Graphs

We start by defining a graph and introducing some notation.

Definition 2.1. [Undirected Graph] A *graph* is a tuple, $G = (V, E)$, where V is a set of *vertices*, and E is a multiset $E \subset \{\{u, v\} \mid u, v \in V\}$ of *edges*, pairs, $\{u, v\}$, of vertices $u, v \in V$. We let E be a multi-set so as to allow multiple edges between any two vertices.

- The graph G is d-regular, if every vertex, $v \in V$, has exactly d edges adjacent to it, in other words, if $|E \cap \{\{v, u\} \mid u \in V\}| = d$ for every $v \in V$.

- An edge, $e \in E$, is a *self-loop*, if $e = \{v\}$ for some $v \in V$.
- For a graph G, we denote by $V(G)$ the set of vertices of G.
- For a vertex $v \in V$, we denote by $N(v)$ the set of neighbours of v in G, i.e.,

$$N(v) = \{u \in V \mid \{v, u\} \in E\}.$$

- For a subset of vertices $C \subset V$, we say that C is an *independent set* if no two vertices of C are neighbours.

Remark 2.2. We shall almost exclusively work with undirected graphs. However, we shall occasionally mention directed edges. A directed edge is a tuple (u, v) of vertices $u, v \in V$. In such cases, an edge $\{u, v\}$ represents two directed edges simultaneously, (u, v) and (v, u).

Vital to spectral graph theory is the concept of the adjacency matrix of a graph. Note that the following definition allows for self-loops, multiple edges, and even works for directed graphs.

Definition 2.3. [Adjacency Matrix] Let $G = (V, E)$ be a graph on $n = |V|$ vertices. The adjacency matrix of G, $A \in \mathbb{N}_0^{V \times V}$, is the matrix satisfying that for $u, v \in V$, A_{uv} is the number of edges from u to v in G.

Defining the set of edges of a cut, we find our first use of the adjacency matrix as a simple linear algebraic short form for the number of edges between two subsets of vertices.

Definition 2.4. For a graph $G = (V, E)$ and subsets $A, B \subset V$ of vertices, we denote by $E(A, B)$ the set of directed edges from A to B in G. This means that if $\{u, v\} \in E$ and $u, v \in A \cap B$, then both (u, v) and (v, u) belong to $E(A, B)$.

We have the following algebraic characterisation the size of a cut.

Lemma 2.5. *Let $G = (V, E)$ be a graph with adjacency matrix A. For any two subsets $S, T \subset V$,*

$$|E(S, T)| = \langle A\, 1_S, 1_T \rangle .$$

Proof. Let $v \in V$. The vth entry of $A\, 1_S$ can be expressed as $(A\, 1_S)_v = \sum_{u \in S} A_{vu}$. Thus,

$$\langle A\, 1_S, 1_T \rangle = \sum_{v \in V} \left([v \in T] \cdot \sum_{u \in S} A_{vu} \right) = \sum_{v \in T} \sum_{u \in S} A_{vu} = |E(S, T)| ,$$

since $(1_T)_v = [v \in T]$ for every $v \in V$. □

Finally, we can introduce a metric on any connected graph by the following natural definition.

Definition 2.6. [Distance in Graphs] Let $G = (V, E)$ be a graph and $v, u \in V$ be vertices. We denote by $\mathrm{dist}(v, u)$ the length of the shortest path between v and u in G and set $\mathrm{dist}(v, u) = \infty$ if no such path exists.

Remark 2.7. Let $G = (V, E)$ be a connected graph. Then $\mathrm{dist}(\cdot, \cdot) \colon V^2 \to \mathbb{N}_0$ is a metric on V.

This naturally leads to the notion of a ball of radius $r \in \mathbb{R}^+$ with centre $v \in V$ in G.

Definition 2.8. Let $G = (V, E)$ be a graph and $v \in V$ a vertex. For any non-negative integer $r \in \mathbb{N}_0$ we denote by $B_r(v)$ the set of vertices that have distance at most r to v. i.e.,

$$B_r(v) = \{u \in V \mid \mathrm{dist}(v, u) \leq r\}.$$

2.3 Linear Algebra

A wealth of vectors and matrices will appear in this exposition. Mostly, they will be related to a graph $G = (V, E)$ and be indexed over V. Generally, we will thus be thinking of vectors as functions, $v\colon V \to \mathbb{C}$, or more precisely as members of $\ell^2(V)$, and square matrices as members of $\mathrm{End}(\ell^2(V))$. However, from time to time it will also be convenient to think of vectors as members of $\mathbb{C}^n$ and matrices as members of $\mathbb{C}^{n\times n}$ for some $n \in \mathbb{N}$. For finite n, these two notions are equivalent and will be used interchangeably.

We recall a couple of lemmas from linear algebra, skipping the proof of the first one.

Lemma 2.9 (Min-max theorem)**.** *Let $A \in \mathbb{R}^{n\times n}$ be a symmetric matrix and let $\lambda_1 \geq \cdots \geq \lambda_n$ be the eigenvalues of A counted with multiplicity with associated orthogonal eigenvectors $v_1, \ldots v_n$. Then for every $i \in [n]$,*

$$\lambda_i = \sup_{\substack{x\in\{v_1,\ldots,v_{i-1}\}^\perp \\ ||x||_2=1}} \langle Ax, x\rangle .$$

Furthermore, $\lambda_n = \inf_{||x||_2=1} \langle Ax, x\rangle$.

Lemma 2.10. *Let $A \in \mathbb{C}^{n\times n}$ be a matrix with eigenvalues $\lambda_1, \ldots, \lambda_n$, counted with multiplicity. Then for every polynomial $p \in \mathbb{C}[x]$, $\mathrm{Tr}(p(A)) = \sum_{i=1}^n p(\lambda_i)$.*

Proof. Let $P \in \mathbb{C}^{n\times n}$ be an invertible matrix such that $J = PAP^{-1}$ is the Jordan normal form of A. We apply the cyclic property of the trace to conclude that for every $k \in \mathbb{N}$,

$$\mathrm{Tr}(A^k) = \mathrm{Tr}(P^{-1}A^kP) = \mathrm{Tr}(J^k) = \sum_{i=1}^n \lambda_i^k.$$

For a polynomial, $p(x) = \sum_{i=0}^t a_i x^i \in \mathbb{C}[x]$, it now follows that

$$\mathrm{Tr}(p(A)) = \sum_{i=1}^t a_i \,\mathrm{Tr}(A^i) = \sum_{i=1}^t a_i \sum_{k=1}^n \lambda_k^i = \sum_{k=1}^n p(\lambda_k),$$

since the trace is linear. □

2.4 Generating Functions

We briefly introduce the notion of a generating function. For readers unfamiliar with combinatorial counting using generating functions, a more thorough treatment of the subject may be found in most introductory textbooks on combinatorics.[1]

Definition 2.11. [Generating Function] Let R be a unital ring and $(a_n)_{n\in\mathbb{N}_0}$ be a sequence of elements of R. The *generating function* associated with $(a_n)_{n\in\mathbb{N}_0}$ is the formal power series $f \in R[[t]]$, given by $f(t) = \sum_{n=0}^\infty a_n t^n$.

While addition and multiplication of generating functions are readily inherited from the definition of a formal power series, the concept of division is not always well-defined. The cases in which a formal power series has an inverse are exactly described by the following proposition, the proof of which is an easy exercise.

[1] For example, *Introduction to Enumerative and Analytical Combinatorics* by Miklós Bónas.

Proposition 2.12. *Let R be a unital ring and let $f \in R[[t]]$ be an element of the ring of formal power series over R. Then f is invertible, i.e., there exists $g \in R[[t]]$ satisfying $fg = 1$, if and only if the constant term of f is a unit of R, i.e., $f(0) \in R^\times$. In this case, we shall denote the inverse element by $g(t) = 1/f(t)$.*

3 Spectral Expansion – Definition and Initial Results

We measure the spectral expansion of a graph by the magnitude of the second largest eigenvalue, in absolute value, of its adjacency matrix. It is by no means intuitive why this would be interesting, useful, or related to expansion in a graph. Hence, this section shall, beyond formally defining spectral expansion, provide motivating examples from the theory of spectral expander graphs. We prove basic results regarding the eigenvalues of regular graphs and then showcase three different classic aspects of spectral expansion. These should give the reader some confidence that spectral expander graphs indeed have properties that one would intuitively associate with expander graphs.

Contrary to other forms of expander graphs, most literature on spectral expander graphs require that they be regular graphs. This is because the largest eigenvalue of the adjacency matrix of any d-regular graph is d as shall see in the lemma following the formal definition. The value of the second largest absolute eigenvalue would not carry the same importance without this guarantee.

Definition 3.1. [Spectral Expander Graph] Let $G = (V, E)$ be a d-regular graph on $|V| = n$ vertices, let A be the adjacency matrix of G, and let $\lambda_1 \geq \cdots \geq \lambda_n$ be the eigenvalues of A (counted with multiplicity). We denote by $\lambda(G)$ the quantity $\max\{|\lambda_2|, |\lambda_n|\}$ and say that G is a *$\lambda(G)$-spectral expander.*

For most applications, it is vital to have good expander graphs with arbitrarily many vertices. Hence, finding a single "good" expander graph is of little use. This motivates the definition of a spectral expander family.

Definition 3.2. [Spectral Expander Family] Fix $d > 1$ and let $(G_n)_{n\in\mathbb{N}}$ be a family of d-regular graphs. We say that the family is a *λ-spectral expander family* if for every $n \in \mathbb{N}$, $\lambda(G_n) \leq \lambda$.

We proceed to prove our first results regarding the eigenvalues of the adjacency matrix of a regular graph.

Proposition 3.3. *Let $G = (V, E)$ be a d-regular graph on $n = |V|$ vertices and adjacency matrix A. Let furthermore, $\lambda_1 \geq \cdots \geq \lambda_n$ be the eigenvalues of A counted with multiplicity. Then*

1. *$\lambda_1 = d$.*
2. *For every $i \in [n]$, $|\lambda_i| \leq d$.*
3. *For every $k \in [n]$, $\lambda_k = d$ if and only if G has at least k distinct connected components.*
4. *$\lambda_n = -d$ if and only if G has a bipartite connected component.*

Proof.

1. We observe that $A\,1_V = d \cdot 1_V$ confirming that d is indeed an eigenvalue of A. That $\lambda_1 = d$ now follows from 2.

2. Recall that every row and column of A sums to d. Hence, every $x \in \ell^2(V)$ satisfies

$$|\langle Ax, x\rangle| \leq \sum_{u,v\in V} A_{uv} |x_u x_v| \leq \sum_{u,v\in V} A_{uv} \left(\frac{x_u^2 + x_v^2}{2}\right) = d \cdot \sum_{v\in V} x_v^2 = d\,||x||_2^2\,. \tag{1}$$

The conclusion now follows from Lemma 2.9, the min-max theorem.

3. Suppose that G has k connected components corresponding to the non-empty vertex sets $V_1, \ldots, V_k \subset V$. It is clear that for every $i \in [k]$, the vector 1_{V_i} is an eigenvector of A with eigenvalue d. Thus, there are k linearly independent eigenvectors with eigenvalue d meaning that the eigenspace of d has dimension at least k. It follows that $\lambda_k = d$.

   Conversely, suppose that $\lambda_k = d$ and let $V_1, \ldots, V_t$ be the connected components of G. For every $x \in \mathbb{R}^n$,

$$\langle Ax, x\rangle = \sum_{u,v\in V} A_{uv} x_u x_v = d\,||x||_2^2 - \sum_{u,v\in V} A_{uv}(x_u - x_v)^2,$$

   by a rearrangement of terms. Let $x \in \mathbb{R}^n$ be an eigenvector of A with eigenvalue d. Then $\sum_{u,v\in V} A_{uv}(x_u - x_v)^2 = 0$. It follows that if there is an edge from u to v then $x_u = x_v$. Thus, $x_u = x_v$ whenever u and v belong to the same connected component of G. So x is a linear combination of the vectors $1_{V_1}, \ldots, 1_{V_t}$. Since the eigenspace of d has dimension at least k when $\lambda_k = d$ and the eigenspace is contained in the span of $1_{V_1}, \ldots, 1_{V_t}$, $t \geq k$ as desired.

4. Suppose that G has a bipartite connected component with vertex set $C \subset V$ and let $L, R \subset C$ partition the vertices of C into two independent sets. If we consider the vector $x = 1_L - 1_R$, then, since $A_{uv} = 0$ whenever $(u, v) \notin (L \times R) \cup (R \times L)$,

$$\frac{\langle Ax, x\rangle}{||x||_2^2} = \frac{\sum_{u,v\in V} A_{uv} x_u x_v}{|C|} = \frac{-\sum_{u,v\in C} A_{uv}}{|C|} = -d.$$

   Thus, $\lambda_n = -d$, since $\lambda_n = \inf_{||x||_2=1} \langle Ax, x\rangle \geq -d$ by 2.

   Conversely, suppose that A has $-d$ as an eigenvalue and let $x \in \ell^2(V)$ be a corresponding non-zero eigenvector orthogonal to 1_V. In (1), we must have equality in both inequalities. The first inequality is an equality if and only if every term of the sum is either ≥ 0 or ≤ 0. Since x is an eigenvector of negative eigenvalue, for every u, v with $A_{uv} \neq 0$, $x_u x_v \leq 0$. The second inequality is an equality if and only if $|x_u| = |x_v|$ for every $u, v \in V$ satisfying $A_{uv} \neq 0$.
   Now, let $v \in V$ be a vertex satisfying $x_v \neq 0$, and let $C \subset V$ be the set of vertices of the connected component containing v. By the second observation above, we must have $|x_v| = |x_u| \neq 0$ for every $u \in C$. Partitioning $C = C_1 \uplus C_2$ as

$$C_1 = \{u \in C \mid x_u > 0\}, \qquad C_2 = \{u \in C \mid x_u < 0\},$$

   we see that $A_{uw} = 0$ for every $u, w \in C_1$ and for every $u, w \in C_2$ since otherwise $x_u x_w > 0$, contradicting the first observation above. Hence, C_1 and C_2 are both independent sets of vertices of G, so C is a bipartite connected component of G, which was what we needed to show. $\square$

3.1 The Mixing Lemma and Its Converse

One of the classic results relating to spectral expander graphs is the mixing lemma. Suppose that $S, T \subset V$ are subsets of vertices and let the d-regular graph $G = (V, E)$ be chosen uniformly at random from the set of d-regular graphs on V. Then the expected number of edges between S and T in G is $d\,|S|\,|T|\,/n$. The mixing lemma states in a spectral expander graph, the number of edges between any pair of vertex subsets, S and T, deviates only a little from this expectation. Despite its simple proof, the mixing lemma turns out to be very powerful in combinatorial applications.

Theorem 3.4 (Mixing Lemma)**.** *For every d-regular graph $G = (V, E)$ and every pair of subsets $S, T \subset V$,*

$$\left| |E(S,T)| - \frac{d\,|S|\,|T|}{n} \right| \leq \lambda(G)\sqrt{|S|\,(1 - |S|\,/n)\,|T|\,(1 - |T|\,/n)}.$$

Proof. Let A be the adjacency matrix of G. For any eigenvector $f \in \ell^2(V)$ of A with eigenvalue θ, the orthogonal complement of f, $W = \{f\}^\perp$ is invariant under A, i.e., $AW \subset W$. To see this, observe that for any $g \in W$, $\langle Ag, f\rangle = \langle g, Af\rangle = \theta\,\langle g, f\rangle = 0$, since A is real and symmetric, hence self-adjoint.

Now, decompose $1_S, 1_T \in \ell^2(V)$ into orthogonal components as

$$1_S = \frac{|S|}{n}1_V + \left(1_S - \frac{|S|}{n}1_V\right), \qquad 1_T = \frac{|T|}{n}1_V + \left(1_T - \frac{|T|}{n}1_V\right).$$

Since $\langle A1_V, 1_V\rangle = dn$ and 1_V is an eigenvector of A, we may apply Lemma 2.5 together with the first observation to obtain

$$\begin{aligned}
\left| |E(S,T)| - \frac{d\,|S|\,|T|}{n} \right| &= \left| \langle A1_S, 1_T\rangle - \frac{d\,|S|\,|T|}{n} \right| \\
&= \left| \frac{|S|\,|T|^2}{n}\,\langle A1_V, 1_V\rangle - \frac{d\,|S|\,|T|}{n} + \left\langle A\left(1_S - \frac{|S|}{n}1_V\right), 1_T - \frac{|T|}{n}1_V \right\rangle \right| \\
&\leq \left\| A\left(1_S - \frac{|S|}{n}1_V\right) \right\|_2 \cdot \left\| 1_T - \frac{|T|}{n}1_V \right\|_2 \\
&\leq \lambda(A) \left\| 1_S - \frac{|S|}{n}1_V \right\|_2 \cdot \left\| 1_T - \frac{|T|}{n}1_V \right\|_2 \\
&= \lambda(A) \cdot \sqrt{|S|\left(\frac{n - |S|}{n}\right)|T|\left(\frac{n - |T|}{n}\right)},
\end{aligned}$$

where the first inequality follows from the Cauchy-Schwartz inequality, and the second follows since the vector $1_S - \frac{|S|}{n}1_V$ is orthogonal to 1_V. □

Far from just being a curiosity of spectral expander graphs, "mixing", as described by the mixing lemma, is in fact almost synonymous with spectral expansion. The two properties are equivalent up to logarithmic factors, as the following theorem shows. We follow the proof given in the paper by Bilu and Linial [1].

Theorem 3.5 (Converse Mixing Lemma [1])**.** *Let $G = (V, E)$ be a d-regular graph. If for every pair of vertex subsets $S, T \subset V$, $\left| |E(S,T)| - \frac{d|S||T|}{n} \right| \leq \alpha\sqrt{|S|\,|T|}$, then $\lambda(G) = O\left(\alpha(\log(d/\alpha) + 1)\right)$.*

The proof proceeds through the following proposition, which can very reasonably be interpreted as a generalisation of the theorem.

Proposition 3.6. *Let $\alpha > 0$ be given. Let A be an $n \times n$ real symmetric matrix satisfying that every row of A has ℓ_1 norm $\leq d$, every diagonal element of A has absolute value $O(\alpha(\log(d/\alpha)+1))$, and for every pair of subsets, $S, T \subset [n]$, with $S \cap T = \emptyset$, $|\langle A1_T, 1_S\rangle| \leq \alpha\sqrt{|S| \cdot |T|}$. Then the spectral radius of A is $O(\alpha(\log(d/\alpha)+1))$.*

Proof. First, suppose that the lemma holds for every matrix, B, satisfying the assumptions with the additional constraint that B has only zeroes on its diagonal. Let $A \in \mathbb{R}^{n\times n}$ be arbitrary satisfying the assumptions of the lemma and write $A = B + D$ where B has only zeroes on its diagonal and D is a diagonal matrix with every entry $O(\alpha(\log(d/\alpha)+1)$. Then for every $S, T \subset [n]$ with $S \cap T = \emptyset$, $|\langle B1_T, 1_S\rangle| = |\langle A1_T, 1_S\rangle| \leq \alpha\sqrt{|S| \cdot |T|}$. Hence, since by assumption this implies that B has spectral radius $O(\alpha(\log(d/\alpha)+1))$ and since the spectral radius of D is bounded by the magnitude of its entries, $A = B + D$ must also have spectral radius $O(\alpha(\log(d/\alpha)+1))$. Thus, we shall assume A to have only zeroes on its diagonal.

Suppose that A satisfies the conditions of the lemma. We first wish to show that for any $S \subset V$, $|\langle A1_S, 1_S\rangle| \leq 2\alpha\,|S|$. To this end, let $k = |S|$. We shall proceed by applying our assumption to all partitions of S into two disjoint sets of sizes $\lfloor \frac{k}{2} \rfloor$ and $\lceil \frac{k}{2} \rceil$. Summing, we get the inequality

$$\sum_{\substack{S_1 \subset S \\ |S_1| = \lfloor \frac{k}{2} \rfloor}} \left|\left\langle A1_{S_1}, 1_{S\setminus S_1}\right\rangle\right| \leq \sum_{\substack{S_1 \subset S \\ |S_1| = \lfloor \frac{k}{2} \rfloor}} \alpha\sqrt{|S_1|\,|S \setminus S_1|} \leq \frac{\alpha k}{2}\binom{k}{\lfloor \frac{k}{2} \rfloor}.$$

Since for every $i, j \in S$ with $i \neq j$, the entry $A_{i,j}$ appears exactly $\binom{k-2}{\lfloor \frac{k}{2} \rfloor - 1}$ times in the sum on the right-hand side of the inequality, and since $A_{i,i} = 0$ for every $i \in S$,

$$\binom{k-2}{\lfloor \frac{k}{2} \rfloor - 1} \cdot |\langle A1_S, 1_S\rangle| = \sum_{\substack{S_1 \subset S \\ |S_1| = \lfloor \frac{k}{2} \rfloor}} \left|\left\langle A1_{S_1}, 1_{S\setminus S_1}\right\rangle\right| \leq \frac{\alpha k}{2}\binom{k}{\lfloor \frac{k}{2} \rfloor}.$$

It follows that indeed $|\langle A1_S, 1_S\rangle| \leq 2\alpha k$.

Next, suppose that $f = 1_{S_1} - 1_{S_2} \in \{-1, 0, 1\}^n$ for disjoint sets $S_1, S_2 \in [n]$. Then, by the above,

$$\begin{aligned}
|\langle Af, f\rangle| &\leq \sum_{i,j\in\{1,2\}} \left|\left\langle A1_{S_i}, 1_{S_j}\right\rangle\right| && (2)\\
&\leq 2\alpha \sum_{i,j\in\{1,2\}} \sqrt{|S_i| \cdot |S_j|} && (3)\\
&\leq 4\alpha \sqrt{\sum_{i,j\in\{1,2\}} |S_i| \cdot |S_j|} && (4)\\
&= 4\alpha\,||f||_2^2\,. && (5)
\end{aligned}$$

If instead $f = 1_{S_1} - 1_{S_2} \in \{-1,0,1\}^n$ and $g = 1_{T_1} - 1_{T_2} \in \{-1,0,1\}^n$, for pairwise disjoint sets $S_1, S_2, T_1, T_2 \subset [n]$, then, similarly,

$$
\begin{aligned}
|\langle Af, g\rangle| &\leq \sum_{i,j\in\{1,2\}} \left|\left\langle A1_{S_i}, 1_{T_j}\right\rangle\right| && (6)\\
&\leq \alpha \sum_{i,j\in\{1,2\}} \sqrt{|S_i|\cdot|T_j|} && (7)\\
&\leq 2\alpha \sqrt{\sum_{i,j\in\{1,2\}} |S_i|\cdot|T_j|} && (8)\\
&= 2\alpha \left|\left|f\right|\right|_2 \cdot \left|\left|g\right|\right|_2 . && (9)
\end{aligned}
$$

We are now ready to find a bound on the spectral radius of A. Fix $x \in \mathbb{R}^n$. We shall show that

$$\frac{|\langle Ax, x\rangle|}{||x||_2^2} = O(\alpha(\log(d/\alpha)+1)),$$

which will conclude the proof. We first find a vector $x' \in \mathbb{R}^n$ satisfying that every entry, $i \in [n]$, of x' is of the form $x'_i = \pm 2^{-k_i}$ for some $k_i \in \mathbb{N}$. Furthermore, we require that

$$\frac{|\langle Ax, x\rangle|}{||x||_2^2} \leq 4\frac{|\langle Ax', x'\rangle|}{||x'||_2^2}. \tag{10}$$

We proceed by the probabilistic method. By scaling, we can assume that $||x||_\infty \leq \frac{1}{2}$. For every $i \in [n]$, there exists $k_i \in \mathbb{N}$ and $\delta_i \in [0,1)$, such that we can write the ith coordinate of x as $x_i = \operatorname{sgn}(x_i)(1+\delta_i)\cdot 2^{-k_i}$, where $\operatorname{sgn}\colon \mathbb{R} \to \{-1,0,1\}$ is the sign function. We now stochastically sample a vector $X \in \mathbb{R}^n$. For each $i \in [n]$ make the following independent choice: let $X_i = \operatorname{sgn}(x_i)2^{-k_i}$ with probability $1-\delta_i$ and $X_i = \operatorname{sgn}(x_i)2^{-k_i+1}$ with probability δ_i. Note that $\mathbb{E}[X_i] = x_i$. Since the diagonal of A is zero and for $i \neq j$, X_i is independent of X_j, $\mathbb{E}[|\langle AX, X\rangle|] = |\langle Ax, x\rangle|$. Thus, we can choose x' in the support of X such that $|\langle Ax', x'\rangle| \geq |\langle Ax, x\rangle|$. Furthermore, $||x'||_2^2 \leq 4\,||x||_2^2$ meaning that (10) is satisfied. Fix such a choice of x'.

For each $i \in \mathbb{N}$, we now define the set $S_i = \{j \in [n] \mid x'_j = \pm 2^{-j}\}$. For each set S_i, we define the corresponding signed indicator vector, $x^i \in \{-1,0,1\}^n$ given by $x^i_j = [j \in S_i]\cdot \operatorname{sgn}(x'_j), j \in [n]$, which is 0 at every index except at the indices of S_i, where it takes the value of the sign of x' at that index. Note that by (4) and (8),

$$\left|\left\langle Ax^i, x^j\right\rangle\right| \leq 4\alpha\sqrt{|S_i|\cdot|S_j|} \tag{11}$$

for every $i, j \in [n]$, since $S_i \cap S_j = \emptyset$ for $i \neq j$. Now, writing $q_i = 2^{-2i}\,|S_i|$ for $i \in \mathbb{N}$,

$$
\begin{aligned}
\frac{|\langle Ax, x\rangle|}{4\,||x||_2^2} &\leq \frac{|\langle Ax', x'\rangle|}{||x'||_2^2}\\
&\leq \frac{\sum_{i,j\in\mathbb{N}}\left(2^{-(i+j)}\cdot\left|\left\langle Ax^i, x^j\right\rangle\right|\right)}{\sum_{i\in\mathbb{N}} 2^{-2i}\,|S_i|}\\
&= \underbrace{\frac{\sum_{i\in\mathbb{N}}\left(2^{-2i}\cdot\left|\left\langle Ax^i, x^i\right\rangle\right|\right)}{\sum_{i\in\mathbb{N}} q_i}}_{Q_1} + \underbrace{\frac{\sum_{i<j}\left(2^{-(i+j)+1}\cdot\left|\left\langle Ax^i, x^j\right\rangle\right|\right)}{\sum_{i\in\mathbb{N}} q_i}}_{Q_2}.
\end{aligned}
$$

Applying (11) we bound the first term by

$$Q_1 \leq \frac{\sum_{i\in\mathbb{N}} \left(2^{-2i} \cdot 4\alpha \left|S\right|_i\right)}{\sum_{i\in\mathbb{N}} q_i} = 4\alpha = O(\alpha(\log(d/\alpha)+1))$$

For the second term, set $\gamma = \lceil \log(d/\alpha) \rceil$ and write

$$Q_2 \cdot \sum_{i\in\mathbb{N}} q_i \leq \underbrace{\sum_{i<j\leq i+\gamma} \left(2^{-(i+j)+1} \cdot \left|\left\langle Ax^i, x^j \right\rangle\right|\right)}_{R_1} + \underbrace{\sum_{i+\gamma<j} \left(2^{-(2i+\gamma)+1} \cdot \left|\left\langle Ax^i, x^j \right\rangle\right|\right)}_{R_2}.$$

Applying first (11) and then the inequality of the arithmetic and geometric means,

$$\begin{aligned} R_1 &\leq \sum_{i<j\leq i+\gamma} \left(2^{-(i+j)} \cdot 8\alpha \cdot \sqrt{|S_i| \cdot |S_j|}\right) \\ &= 8\alpha \cdot \sum_{i<j\leq i+\gamma} \sqrt{q_i \cdot q_j} \\ &\leq 4\alpha \sum_{i<j\leq i+\gamma} (q_i + q_j) \\ &\leq 4\alpha(\gamma+1) \sum_{i\in\mathbb{N}} q_i. \end{aligned}$$

Note that for every $i \in \mathbb{N}$, $\left|\left|Ax^i\right|\right|_1 \leq d\left|S_i\right|$, since the ℓ^1-norm of any row or column of A is $\leq d$. Thus, $\sum_{j\in[n]} \left|\left\langle Ax^i, x^j \right\rangle\right| \leq d\left|S_i\right|$, because the non-zero entries of x^{j_1} and x^{j_2} are disjoint for $j_1 \neq j_2$, such that each entry of Ax^i contributes at most once to the sum. It follows that

$$R_2 \leq 2^{-\gamma} \cdot \sum_{i\in\mathbb{N}} \left(2^{-2i} \cdot \sum_{j\in\mathbb{N}} \left|\left\langle Ax^i, x^j \right\rangle\right|\right) \leq 2^{-\gamma} \cdot \sum_{i\in\mathbb{N}} 2^{-2i} d\left|S_i\right| \leq 2^{-\gamma} d \sum_{i\in\mathbb{N}} q_i.$$

Thus,

$$Q_2 = \frac{R_1 + R_2}{\sum_{i\in\mathbb{N}} q_i} \leq 4\alpha\left(\lceil \log(d/\alpha) \rceil + 1\right) + 2^{-\lceil \log(d/\alpha) \rceil} d = O(\alpha(\log(d/\alpha)+1)),$$

which concludes the proof. □

We are now ready to prove Theorem 3.5.

Proof of Theorem 3.5. The largest eigenvalue of the adjacency matrix A of G is d with corresponding eigenvector 1_V. Hence, the matrix $D = A - \frac{d}{n}J$, where $J = 1_V 1_V^T$ is the all-1s matrix, has greatest absolute eigenvalue $\lambda(G)$. Furthermore, the ℓ_1-norm of any row of D is at most $2d$ and every diagonal entry of D is $-\frac{d}{n} = O(1)$. For any vertex subsets $S, T \subset V$ with $S \cap T = \emptyset$, $\left\langle \frac{d}{n} J 1_S, 1_T \right\rangle = \frac{d|S||T|}{n}$, such that

$$\alpha\sqrt{|S|\,|T|} \geq \left| |E(S,T)| - \frac{d\,|S|\,|T|}{n} \right| = \left| \langle A1_S, 1_T \rangle - \frac{d\,|S|\,|T|}{n} \right| = \left| \langle D1_S, 1_T \rangle \right|.$$

It follows that the requirements of Proposition 3.6 are satisfied. Thus, the spectral radius of D is $\lambda(G) = O(\alpha(\log(d/\alpha)+1))$, which was what we needed to prove. □

3.2 The Cheeger Inequalities

Another feature of spectral expansion is its close connection to its more combinatorial cousin, edge expansion. Intuitively, edge expander graphs are graphs with no sparse cuts as the following definition makes precise. For a graph, $G = (V, E)$, on n vertices, and a vertex subset, $S \subset V$, we denote by $\overline{S}$ the set of vertices $\overline{S} = V \setminus S$, and by ∂S the set of edges $\partial S = E(S, \overline{S})$.

Definition 3.7. [Cheeger Constant, Edge Expander] Let $G = (V, E)$ be a graph. The *Cheeger constant* of G is the quantity

$$h(G) = \min_{\emptyset \subsetneq S \subset V} \frac{|\partial S|}{\min\{|S|, |\overline{S}|\}}.$$

We say that the graph G is an *α-edge expander* for every α satisfying $h(G) \geq \alpha \geq 0$.

Suppose now that G is d-regular and let $\lambda_k(G)$ denote the kth largest eigenvalue of the adjacency matrix of G. The Cheeger inequalities, essentially state that up to constant factors, $h(G)$ and $\lambda_2(G)$ are equivalent quantities. Given the mixing lemma (Theorem 3.4), it is by no means surprising that good spectral expansion, i.e., $\lambda(G) = \max\{\lambda_2, |\lambda_n|\}$ being small, implies edge expansion. That the converse is not the case, is clear from considering a bipartite graph. If G is bipartite, $\lambda_n(G) = -d$, such that G has the worst possible spectral expansion, $\lambda(G) = d$. However, it is easy to see that G may still have a strictly positive Cheeger constant. Thus, we disregard $\lambda_n(G)$ and simply consider $\lambda_2(G)$. More specifically, we consider the *spectral gap*, $d - \lambda_2(G)$, which is the smallest non-trivial eigenvalue of the *Laplacian* of G, $L = dI - A$, where A is the adjacency matrix of G. Both the Laplacian and the spectral gap hold significant importance within the theory.

To prove that a large spectral gap implies edge expansion is an easy exercise. The other direction – showing how edge expansion implies a large spectral gap – is a trickier endeavour. For this part, we follow the survey of Hoory, Linial, and Wigderson [6]. Intuitively, the trick is to find a sparse cut using an eigenvector of the Laplacian with eigenvalue $d - \lambda_2(G)$.

Theorem 3.8 (Cheeger Inequalities)**.** *Let $G = (V, E)$ be a d-regular graph. Then*

$$\frac{d - \lambda_2(G)}{2} \leq h(G) \leq \sqrt{2d(d - \lambda_2(G))}.$$

Note that since $d - \lambda(G) \leq d - \lambda_2(G)$, it follows directly that spectral expansion implies edge expansion.

Proof. Denote by $n = |V|$ the number of vertices of G, let A be the adjacency matrix of G, and denote by $L = dI - A$ the Laplacian of G.

Towards proving the left-most inequality, let $S \subset V$ be such that $|S| \leq \frac{n}{2}$ and $h(G) = \frac{|\partial S|}{|S|}$. Observe that the vector $x = |S| \, 1_{\overline{S}} - |\overline{S}| \, 1_S$ is orthogonal to 1_V. We have $||x||_2^2 = n \, |S| \, |\overline{S}|$ and

$$\langle Ax, x \rangle = |S|^2 \, |E(\overline{S}, \overline{S})| + |\overline{S}|^2 \, |E(S, S)| - 2 \, |S| \, |\overline{S}| \, |E(S, \overline{S})|,$$

and since $|E(T, T)| = d \, |T| - |E(T, \overline{T})|$ for any $T \subset V$,

$$\langle Ax, x \rangle = d \, |S|^2 \, |\overline{S}| + d \, |\overline{S}|^2 \, |S| - (|S| + |\overline{S}|)^2 \, |E(S, \overline{S})| = dn \, |S| \, |\overline{S}| - n^2 \, |E(S, \overline{S})|.$$

It follows that $\lambda_2(G) \geq \langle Ax, x\rangle / ||x||_2^2 = d - n \cdot h(G)/\left|\overline{S}\right| \geq d - 2h(G)$, and we are done.

For the right-most inequality, let $f \in \ell^2(V)$ an eigenvector of A with eigenvalue $\lambda_2(G)$, and note that f is also an eigenvector of L with eigenvalue $d - \lambda_2(G)$. Now, let $g = f^+$ to be the positive part of f, i.e., $g(v) = f(v) \cdot [f(v) \geq 0]$ for $v \in V$. Intuitively, we shall use that g can be viewed as a cut of G that approximates the Cheeger constant. Let $V^+ = \sup(g)$ denote the support of g, and assume without loss of generality that $|V^+| \leq \frac{n}{2}$. If this were not the case, we could just consider $g = f^-$ instead. We shall now prove the inequalities

$$d - \lambda_2(G) \geq \frac{\langle Lg, g\rangle}{||g||_2^2} \tag{12}$$

and

$$\frac{\langle Lg, g\rangle}{||g||_2^2} \geq \frac{h(G)^2}{2d}, \tag{13}$$

which will conclude the proof.

First, (12) follows by invoking that $f(v) \leq g(v)$ for every $v \in V$, and that every entry of A is non-negative, yielding

$$\begin{aligned}
\langle Lg, g\rangle &= d\,\langle Ig, g\rangle - \langle Ag, g\rangle \\
&\leq d\,||g||_2^2 - \sum_{v \in V^+} (Af)(v)f(v) \\
&= d\,||g||_2^2 - \lambda_2(G) \sum_{v \in V^+} f(v)^2 \\
&= (d - \lambda_2(G))\,||g||_2^2\,.
\end{aligned}$$

Second, we prove (13) by comparing each side of the inequality to the quantity

$$B_g = \sum_{\{u,v\} \in E} \left|g(u)^2 - g(v)^2\right|.$$

We start with a simple comparison by Cauchy-Schwartz

$$B_g = \sum_{\{u,v\} \in E} |g(u) - g(v)|\,|g(u) + g(v)| \leq \sqrt{\sum_{\{u,v\} \in E} (g(u) - g(v))^2} \cdot \sqrt{\sum_{\{u,v\} \in E} (g(u) + g(v))^2}.$$

Noting that

$$\sum_{\{u,v\} \in E} (g(u) + g(v))^2 \leq 2 \sum_{\{u,v\} \in E} g(u)^2 + g(v)^2 = 2d\,||g||_2^2\,,$$

and that

$$\sum_{\{u,v\} \in E} (g(u) - g(v))^2 = d\,||g||_2^2 - \sum_{\{u,v\} \in E} g(u)g(v) = \langle Lg, g\rangle\,,$$

we conclude that $B_g \leq ||g||_2 \cdot \sqrt{2d \langle Lg, g \rangle}$. On the other hand, we can identify V with the set of integers $[n]$ in such a way that $g\colon [n] \to \mathbb{R}_{\geq 0}$ is weakly decreasing and write

$$B_g = \sum_{\substack{\{i,j\} \in E \\ i<j}} g(i)^2 - g(j)^2 = \sum_{i=1}^{n} g(i)^2 \cdot \left(\left| N(i) \cap \overline{[i]} \right| - |N(i) \cap [i-1]| \right),$$

where we recall that $\overline{[i]} = V \setminus [i] = \{i+1, \ldots, n\}$. Applying the identity

$$\left| N(i) \cap \overline{[i]} \right| - |N(i) \cap [i-1]| = \left| E([i], \overline{[i]}) \right| - \left| E([i-1], \overline{[i-1]}) \right|,$$

and telescoping the sum, we arrive at

$$\begin{aligned} B_g &= \sum_{i=1}^{n} g(i)^2 \cdot \left(\left| E([i], \overline{[i]}) \right| - \left| E([i-1], \overline{[i-1]}) \right| \right) \\ &= \sum_{i=1}^{n-1} (g(i)^2 - g(i+1)^2) \cdot \left| E([i], \overline{[i]}) \right| \\ &\geq h(G) \sum_{i=1}^{n-1} (g(i)^2 - g(i+1)^2) \cdot i \\ &= h(G) \, ||g||_2^2 , \end{aligned}$$

where the inequality follows by definition of the Cheeger constant, since $g(i) \geq g(i+1)$, for every $i \in [n]$, and since $g(i) = 0$ for $i > \frac{n}{2}$. Combining the two bounds on B_g yields (13). □

3.3 Random Walks on Expander Graphs

The following section considers applications of expander graphs to the area of randomised algorithms and pseudorandomness. We begin with a few words on randomised algorithms. The field is huge and complex, so the definitions and explanations found here are simplified ones, but they will suffice for our purposes.

The goal of a traditional deterministic algorithm, A, is to accomplish some task given an input x. On x, the algorithm A will always provide the same output and will always go through the same computational steps. The efficiency of a deterministic algorithm is usually measured by the worst-case running time. The worst-case running time of A is the maximal number of computational steps that A may take to process an input x. In other words, there exists a worst-case input, x, on which A always expends this amount of computational power. The efficiency of a deterministic algorithm is usually measured by its worst-case running time.

A randomised algorithm, B, is a deterministic algorithm that has access to a source of random bits. One can think of this, as the algorithm being able to "flip coins" or repeatedly sample from the uniform distribution on $\{0, 1\}$. The way we measure the efficiency of a randomised algorithm is significantly different from the deterministic setting. Randomised algorithms may answer incorrectly with some probability, and the running time of the algorithm may also be stochastic, since it depends on the random bits. We measure the efficiency of a randomised algorithm in terms of three parameters.

Expected Running Time. If E_x is the expected number of computational steps for B to evaluate on input x, then the *expected running time* of B is the worst-case, $\max_x E_x$.

Success Probability. If P_x is the probability that B outputs the correct answer on input x, then the *success probability* of B is the worst-case over the inputs, $\min_x P_x$.

Expected Number of Random Bits. If R_x is the expected number of times a random bit is sampled on input x, then the *expected number of random bits* of B is the worst-case, $\max_x R_x$.

When designing randomised algorithms, one wishes to optimise for these parameters. Optimally, an algorithms should have a low expected running time while maintaining a high success probability and using few random bits. In this section, we shall only deal with *Monte Carlo algorithms* which have a fixed running time, but may have erroneous output. Thus, we shall only care about optimising success probability and the expected number of random bits.

Reducing the number of bits necessary to perform a task is the subject of pseudorandomness. We shall illustrate how spectral expander graphs have pseudorandom properties by two examples. In both cases, we consider classical problems in computation, where the introduction of spectral expander graphs allows us to use fewer random bits than with the canonical randomised approach. The presentation follows the survey on pseudorandomness by Salil Vadhan [13].

Member Search (or Error Reduction for RP Algorithms). Consider a set U. Suppose that we have an algorithm A that given an input, $x \in U$, outputs 1 on some subset $B \subset U$ and 0 on $U \setminus B$. Given oracle access to A, i.e., we can query A on any $x \in U$, but know nothing of its inner workings, we wish to find a member of B. The input to our algorithm is the algorithm A. Deterministically this task is hard and has a worst-case running time of $|U \setminus B|$. However, consider a Monte Carlo algorithm which simply samples t elements $v_1, \ldots, v_t \in U$ from U uniformly at random and checks, if one of them belongs to B. The algorithm succeeds in finding an element of B with probability $1 - (1 - |B| / |U|)^t$ using $\Theta(t \cdot \log |U|)$ random bits (each sample $v \in U$ takes $\Theta(\log |U|)$ random bits). The success probability increases dramatically with t, for instance picking $t = |U| / |B|$ yields a success probability of around $1 - \frac{1}{e}$.

Average Estimation (or Error Reduction for BPP Algorithms). Consider a set U and a function $f \colon U \to [0, 1]$. We wish to find the average, $\frac{1}{|U|} \sum_{i \in U} f(i)$, up to an additive error of $\pm\varepsilon$ given oracle access to f as an input. Again, this is hard to do with a deterministic algorithm, requiring $\Omega(|U|)$ computational steps. However, simply sampling t uniformly random elements, $v_1, \ldots, v_t \in U$, from U and outputting their average, can be shown, as we will see shortly, to succeed with probability $1 - 2e^{-\Omega(\varepsilon^2 t)}$. Again this uses $\Theta(t \log |U|)$ random bits.

In both cases, there exist simple, efficient randomised algorithms that solve the problems. We shall show that we can use random walks on expander graphs to solve both problems using much fewer random bits than the naive Monte Carlo algorithmic solutions suggest. To this end, we define the random walk matrix of a graph.

Definition 3.9. [Random Walk Matrix] Let G be a d-regular graph with adjacency matrix A. The *random walk matrix* of G is the normalised adjacency matrix $M = \frac{1}{d} A$.

Remark 3.10. The name, *random walk matrix*, derives from the following observation. Let $G = (V, E)$ be a d-regular graph $G = (V, E)$ with random walk matrix M. The probability that a random walk on G, starting at vertex $v \in V$ and taking t steps, ends at vertex $u \in V$ is exactly $(M^t)_{vu}$. This is also essentially the contents of Lemma 4.3.

Both the applications, we shall showcase, rely on the following two lemmas.

Lemma 3.11. *Let $G = (V, E)$ be a d-regular graph on $n = |V|$ vertices with adjacency matrix A and spectral expansion λ. Then there exists a symmetric matrix $E \in \mathbb{R}^{n \times n}$ with $||E||_2 \leq 1$ and*

$$A = \frac{d-\lambda}{n} J + \lambda E,$$

where J is the all-ones matrix.

Proof. We prove that the matrix $E = \frac{1}{\lambda}\left(A - \frac{d-\lambda}{n} J\right)$ has spectral norm at most 1. From this, the conclusion follows. Let $x \in \ell^2(V)$ be any vector of length 1. We can decompose x as $x = c_1 1_V + c_2 v$, where $v \in \ell^2(V)$ is orthogonal to 1_V. Then

$$||Ex||_2^2 = \left|\left|\frac{1}{\lambda}\left(A - \frac{d-\lambda}{n} J\right)(c_1 1_V + c_2 v)\right|\right|_2^2 = \left|\left|\frac{1}{\lambda}\lambda c_1 1_V + c_2 A v\right|\right|_2^2 \leq (c_1^2 \, ||1_V||_2^2 + c_2^2 \, ||v||_2^2) = 1.$$

The conclusion follows. □

Lemma 3.12. *Let $G = (V, E)$ be a graph with random walk matrix M, and let $f \colon V \to \mathbb{R}$ be a value function on the vertices of G. Identify V with $[n]$ for some $n \in \mathbb{N}$ and let T be the matrix*

$$T = \begin{pmatrix} f(1) & & 0 \\ & \ddots & \\ 0 & & f(n) \end{pmatrix}.$$

Then for a random walk of length $t-1$, $(v_1, \ldots, v_t)$, starting at a uniformly random vertex v_1,

$$\mathbb{E}\left[\prod_{i=1}^{t} f(v_i)\right] = \left|\left|(TM)^{t-1} T u\right|\right|_1,$$

where u is the uniform distribution vector on V, i.e., the vector where every entry is $1/n$.

Proof. We shall prove by induction on t that for any vertex $v \in V$,

$$n \cdot \left((TM)^{t-1} T u\right)_v = \mathbb{E}\left[\prod_{i=1}^{t} f(v_i) \,\middle|\, v_t = v\right],$$

from which the conclusion follows. For $t = 1$ the statement is trivial. Suppose that the lemma holds for some $t \in \mathbb{N}$ and let v be arbitrary. Then for $(v_1, \ldots, v_{t+1})$ a random walk of length t starting at

a uniformly random vertex v_1,

$$\begin{aligned}
\mathbb{E}\left[\prod_{i=1}^{t+1} f(v_i) \,\middle|\, v_{t+1} = v\right] &= \frac{1}{d} \sum_{w \in N(v)} \mathbb{E}\left[\prod_{i=1}^{t+1} f(v_i) \,\middle|\, v_{t+1} = v, v_t = w\right] \\
&= \frac{1}{d} \sum_{w \in N(v)} f(v) \cdot \mathbb{E}\left[\prod_{i=1}^{t} f(v_i) \,\middle|\, v_t = w\right] \\
&= f(v) \cdot \sum_{w \in N(v)} \frac{1}{d} \left((TM)^{t-1} Tu\right)_w \\
&= f(v) \left(M(TM)^{t-1} Tu\right)_v \\
&= \left((TM)^t Tu\right)_v .
\end{aligned}$$

The conclusion follows. □

3.3.1 Member Search

Let U be a set and $B \subset U$ a subset with $\mu = |B| / |U|$. Picking elements $v_1, \ldots, v_t \in U$ uniformly at random as described earlier, we find an element of B with probability.

$$P\left(\bigcup_{i=1}^{t} (v_i \in B)\right) = 1 - (1 - \mu)^t .$$

This procedure uses $\Theta(t \cdot \log |U|)$ random bits. Alternatively, we can proceed as follows. Consider U to be the vertex set of a d-regular spectral expander graph G. Sample the random elements according to a random walk on G instead of uniformly at random. This requires only $\Theta((t-1)\log d + \log |U|)$ random bits, $\Theta(\log |U|)$ bits to pick a starting vertex and $\Theta(\log d)$ bits for each step in the random walk. Meanwhile, the procedure matches the error probability up to an additive factor in the base of the exponentiation as the following theorem shows. Using an expander graph of low degree hence yields accurate sampling using few random bits.

Theorem 3.13. *Let U be a set, $B \subset U$ a subset, and write $n = |U|$ and $\mu = |B| / |U|$. Let G be a d-regular graph with vertices $V(G) = U$ and spectral expansion λ. The probability that a random walk $(v_1, \ldots, v_t)$ of length $t-1$ on G starting at a uniformly random vertex $v_1 \in U$ and proceeding to a uniformly random neighbour at each step finds a member of B is*

$$P\left(\bigcup_{i=1}^{t} (v_i \in B)\right) \geq 1 - (1 - \mu + \lambda\mu/d)^t .$$

Proof. We shall show that the probability that the random walk stays entirely inside $U \setminus B$ is $\leq (1 - \mu + \lambda\mu/d)^t$. Let M be the random walk matrix of G. Furthermore, let T be the diagonal matrix satisfying $T_{ii} = [i \in U \setminus B]$ for each $i \in [n]$, and let $u \in \ell^2(V)$ be the uniform distribution vector on the vertices of G, i.e., every entry of u is $\frac{1}{n}$. We can now apply Lemma 3.12 to obtain

$$P\left(\bigcap_{i=1}^{t} (v_i \notin B)\right) = \mathbb{E}\left[\prod_{i=1}^{t} [v_i \in U \setminus B]\right] = \left|\left|(TM)^{t-1} Tu\right|\right|_1 .$$

Using that T is idempotent we can estimate

$$\begin{aligned}
P\left(\bigcap_{i=1}^{t}(v_i \notin B)\right) &= \left|\left|(TMT)^{t-1}Tu\right|\right|_1 \\
&\leq \sqrt{|U|-|B|}\cdot\left|\left|(TMT)^{t-1}Tu\right|\right|_2 \\
&\leq \sqrt{|U|-|B|}\cdot\left|\left|Tu\right|\right|_2\cdot\left|\left|(TMT)^{t-1}\right|\right|_2 \\
&= \frac{|U|-|B|}{|U|}\left|\left|(TMT)^{t-1}\right|\right|_2 \\
&\leq (1-\mu)\left|\left|TMT\right|\right|_2^{t-1},
\end{aligned}$$

where the last inequality uses the fact that TMT is a symmetric matrix. Thus, all that remains, is to show that $||TMT||_2 \leq 1-\mu+\frac{\lambda\mu}{d}$.

To this end, observe that by Lemma 3.11 we can write $M = \frac{d-\lambda}{dn}J + \frac{\lambda}{d}E$, where $||E||_2 \leq 1$ and J is the all-ones matrix. Now,

$$||TMT||_2 \leq \frac{d-\lambda}{nd}\,||TJT||_2 + \frac{\lambda}{d}\,||TET||_2\,,$$

Since $||T||_2 \leq 1$, $||TET||_2 \leq ||T||_2^2\,||E||_2 \leq 1$. Furthermore, for every $x \in \ell^2(U)$,

$$||TJTx||_2 = \sqrt{(|U|-|B|)\left(\sum_{i\in U\setminus B} x_i\right)^2} \leq (|U|-|B|)\,||x||_2\,,$$

implying $\frac{1}{n}\,||TJT||_2 \leq 1-|\mu|$. Thus, $||TMT||_2 \leq 1-\mu+\lambda\mu/d$, as desired. □

3.3.2 Average Estimation

Recall that the goal of average estimation is to estimate the average, $\frac{1}{|U|}\sum_{i\in U} f(i)$, up to an additive factor of ε, given a set U and a function $f\colon U \to [0,1]$. As mentioned above, this can be accomplished by simply sampling t uniformly random elements and outputting their average. By Theorem 3.14 below this will be correct up to an additive error of $\pm\varepsilon$ with probability at least $1-2\exp(-n\varepsilon^2/4)$. This require $\Theta(t\log|U|)$ random bits to be successful.

A recurrent theme in the analysis of randomised algorithms is the application of Chernoff bounds. These are strong concentration bounds for the distribution of sums of independent random variables. The bound presented here addresses additive error as required for our application.

Theorem 3.14. *Let $X_1,\ldots,X_n$ be independent random variables taking values in $[0,1]$ and let $\mu = \frac{1}{n}\mathbb{E}\left[\sum_{i=1}^n X_i\right]$. Then for every $\varepsilon > 0$,*

$$P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right| > \varepsilon\right) \leq 2\exp\left(-\frac{n\varepsilon^2}{4}\right).$$

Proof. We start by observing that by Markov's inequality,

$$P\left(\frac{1}{n}\sum_{i=1}^{n} X_i > \varepsilon + \mu\right) = P\left(\exp\left(r\sum_{i=1}^{n} X_i\right) > \exp(rn(\varepsilon+\mu))\right)$$
$$\leq \frac{\mathbb{E}\left[\exp\left(r\sum_{i=1}^{n} X_i\right)\right]}{\exp(rn(\varepsilon+\mu)},$$

for some $r \in (0,1)$ to be specified later. Recall that $e^x \geq 1 + x$ for every $x \in \mathbb{R}$ and observe that $e^x \leq 1 + x + x^2$ for every $x \in [0,1]$. Now, by independence of $X_1, \ldots, X_n$,

$$\mathbb{E}\left[\exp\left(r\sum_{i=1}^{n} X_i\right)\right] = \prod_{i=1}^{n} \mathbb{E}\left[e^{rX_i}\right] \leq \prod_{i=1}^{n} \mathbb{E}\left[1 + rX_i + (rX_i)^2\right] \leq \prod_{i=1}^{n} e^{r\mathbb{E}[X_i]+r^2} = \exp\left(rn(\mu+r)\right).$$

We may assume that $\varepsilon \in (0,1)$, since the theorem is trivially true for $\varepsilon \geq 1$. Setting $r = \varepsilon/2$, it then follows that

$$P\left(\frac{1}{n}\sum_{i=1}^{n} X_i > \varepsilon + \mu\right) \leq \frac{\exp\left(rn(\mu+r)\right)}{\exp(rn(\varepsilon+\mu)} = \exp(rn(r-\varepsilon) = \exp\left(-\frac{n\varepsilon^2}{4}\right). \tag{14}$$

Now, applying (14) to the random variables $1 - X_1, \ldots, 1 - X_n$,

$$P\left(\frac{1}{n}\sum_{i=1}^{n} X_i < -\varepsilon + \mu\right) \leq \exp\left(-\frac{n\varepsilon^2}{4}\right). \tag{15}$$

Combining (14) and (15) yields our conclusion. □

To avoid the use of $\Theta(t \log |U|)$ random bits in our sampling, we can instead sample elements of U according to a random walk on a d-regular λ-spectral expander graph with vertex set U. This allows us to find the average to within an additive error of $\pm(\varepsilon + \lambda/d)$ with a similar error probability. While being less precise, this estimate only requires $\Theta(t \log d + \log |U|)$ random bits, improving greatly on the more naive approach described above. The proof proceeds very similarly to the proof of the classical Chernoff bound. The two sides of the inequality are exponentiated, which allows for an application of Markov's inequality. The conclusion then follows by suitably applying Lemma 3.11 and Lemma 3.12.

Theorem 3.15. *Let U be a set, $f\colon U \to [0,1]$ be a function, and $\mu = \frac{1}{|U|}\sum_{i\in U} f(i)$ denote the average value of f on U. Furthermore, let G be a d-regular graph with vertex set $V(G) = U$ and spectral expansion λ. For every $\varepsilon > 0$, a random walk, $(v_1, \ldots, v_t)$, of length $t-1$ on G starting at a uniformly random vertex $v_1 \in U$ and proceeding to a uniformly random neighbour at each step, satisfies*

$$P\left(\left|\frac{1}{t}\sum_{i=1}^{t} f(v_i) - \mu\right| \geq \lambda/d + \varepsilon\right) \leq 2\exp\left(-\frac{t\varepsilon^2}{16}\right).$$

Proof. Identify U with $[n]$ for some $n \in \mathbb{N}$, and as in the proof of Theorem 3.14, assume that $\varepsilon \in (0,1)$. For some $r \in (0,1)$ to be specified later, it follows by Markov's inequality that

$$P\left(\sum_{i=1}^{t} f(v_i) \geq (\mu + \varepsilon + \lambda/d)t\right) = P\left(\exp\left(r\sum_{i=1}^{t} f(v_i)\right) \geq \exp\left(rt(\mu + \varepsilon + \lambda/d)\right)\right) \tag{16}$$

$$\leq \frac{\mathbb{E}\left[\exp\left(\sum_{i=1}^{t}(rf(v_i))\right)\right]}{\exp\left(rt(\mu + \varepsilon + \lambda/d)\right)}. \tag{17}$$

Now, define the diagonal matrix

$$T_r = \begin{pmatrix} e^{rf(1)} & & 0 \\ & \ddots & \\ 0 & & e^{rf(n)} \end{pmatrix}. \tag{18}$$

By Lemma 3.12, it follows that $\mathbb{E}\left[\exp\left(\sum_{i=1}^{t}(rf(v_i))\right)\right] = \left|\left|(T_rM)^{t-1}T_ru\right|\right|_1$, where M is the random walk matrix of G and u is the uniform distribution vector on U, i.e., the vector where every entry is $1/n$. Noting that $Mu = u$,

$$\mathbb{E}\left[\exp\left(\sum_{i=1}^{t}(rf(v_i))\right)\right] = \left|\left|(T_rM)^t u\right|\right|_1 \leq \sqrt{n}\left|\left|(T_rM)^t\right|\right|_2 \cdot ||u||_2 \leq ||T_rM||_2^t. \tag{19}$$

As in the proof of Theorem 3.13, we apply Lemma 3.11 to write $M = \frac{d-\lambda}{dn}J + \frac{\lambda}{d}E$, where $||E||_2 \leq 1$ and J is the all-ones matrix. Then

$$||T_rM||_2 \leq \frac{d-\lambda}{dn}||T_rJ||_2 + \frac{\lambda}{d}||T_rE||_2. \tag{20}$$

Solving the equation $e^x = 1 + x + \frac{5}{4}x^2$, one observes that for every $x \in [0,2]$, $e^x \leq 1 + x + \frac{5}{4}x^2$. We apply this to bound the two terms of the right hand side of (20). First,

$$||T_rE||_2 \leq ||T_r||_2 \leq e^r \leq 1 + r + \frac{5r^2}{4},$$

and second,

$$\begin{aligned} ||T_rJ||_2 &= \sup_{x\neq 0}\frac{||T_rJx||_2}{||x||_2} \\ &= \sup_{x\neq 0}\frac{||x||_1 \cdot \sqrt{\sum_{i=1}^{n} e^{2rf(i)}}}{||x||_2} \\ &\leq \sqrt{n}\cdot\sqrt{\sum_{i=1}^{n} 1 + 2rf(i) + 5r^2f(i)^2} \\ &\leq n\cdot\sqrt{1 + 2\mu r + 5r^2} \\ &\leq n\left(1 + \mu r + \frac{5}{2}r^2\right). \end{aligned}$$

Now, we may reexamine (20), to find that

$$\begin{aligned}||T_r M||_2 &\leq \left(1-\frac{\lambda}{d}\right)\left(1+\mu r+\frac{5\mu}{2}r^2\right)+\frac{\lambda}{d}\left(1+r+\frac{5r^2}{4}\right)\\ &\leq 1+(\mu+\lambda/d)r+4r^2\\ &\leq \exp((\mu+\lambda/d)r+4r^2).\end{aligned}$$

Combining this with (17) and (19) yields

$$\begin{aligned}P\left(\sum_{i=1}^{t} f(v_i) \geq (\mu+\varepsilon+\lambda/d)t\right) &\leq \exp\left(rt\left((\mu+\lambda/d+4r)-(\mu+\varepsilon+\lambda/d)\right)\right)\\ &= \exp\left(rt\left(4r-\varepsilon\right)\right).\end{aligned}$$

Setting $r=\frac{\varepsilon}{8}<1$, we obtain

$$P\left(\sum_{i=1}^{t} f(v_i) \geq (\mu+\varepsilon+\lambda/d)t\right) \leq \exp\left(-\frac{t\varepsilon^2}{16}.\right)$$

As in the proof of Theorem 3.14, a symmetry argument now yields the conclusion. □

4 Lower Bounds on Spectral Expansion

Let $d>1$ be fixed and $(G_n)_{n\in\mathbb{N}}$ be a family of d-regular λ-spectral expanders. Alon and Boppana famously showed that λ may be no smaller than $2\sqrt{d-1}$ or to put it differently, they showed that for any d-regular graph, G, on n vertices,

$$\lambda(G) \geq 2\sqrt{d-1}-o_n(1),$$

where $o_n(1)\to 0$ as $n\to\infty$. This bound is in fact tight since there exist graphs, G, on arbitrarily many vertices with $\lambda(G)\leq 2\sqrt{d-1}$. Such graphs are called *Ramanujan graphs* and they are one of the main subjects of study within the field.

Definition 4.1. [Ramanujan Graph] Let G be a d-regular graph. We say that G is a *Ramanujan graph* if $\lambda(G)\leq 2\sqrt{d-1}$.

In this section we present two proofs of the Alon-Boppana lower bound. One directly yields the above statement, while the other is a more general statement regarding the spectrum of the adjacency matrix of any sufficiently large d-regular graph. The reason for presenting two proofs of the same fact is to gain a feel for the method of proof. The first result we prove is a great warm-up towards understanding the methods employed to prove the second more general result.

Denote by $T_d=(V,E)$ the infinite d-regular tree, i.e., an infinite graph where each vertex has degree d and there are no cycles. One may construct the adjacency operator of T_d as an infinite-dimensional linear operator, $A\colon \ell^2(V)\to\ell^2(V)$. The spectral norm of A is $2\sqrt{d-1}$, a fact that we shall not prove in this exposition. We may think of this quantity as the spectral expansion of T_d. The vector 1_V is not contained in $\ell^2(V)$ meaning that d is not a trivial eigenvalue of A as it is for finite d-regular graphs. In many ways, T_d is the optimal d-regular expander. The graph

infinitely expands and never turns back in on itself. This intuition turns out to be quite fruitful. At a high level, the proofs of this section compare a family of d-regular graphs, $(G_n)_{n\in\mathbb{N}}$, with an increasing number of vertices to T_d. The conclusion is that the best expansion achievable arises if for larger and larger neighbourhoods of vertices, G_n resembles T_d as n approaches infinity. Thus, in a sense, T_d has the best possible spectral expansion, yielding the bound mentioned at the start of the section.

The *girth* of a graph is the length of the shortest cycle. If a d-regular graph G has girth $g(G) = 2i$ for some $i \in \mathbb{N}$ then any ball of radius $\leq i$ in G would be indistinguishable from a similar ball in T_d. Thus, if G has large girth, it locally resembles T_d. In fact, this translates to the spectral properties of G. The last theorem of the section concerns the spectrum of a family of graphs $(G_n)_{n\in\mathbb{N}}$ satisfying $|V(G_n)| \to \infty$ and $g(G_n) \to \infty$ as $n \to \infty$. The result states that the spectrum of the adjacency matrix of G_n will approach the spectrum of the adjacency operator of T_d as n tends to infinity in a way that will be made precise later.

In this exposition, we provide three related and increasingly advanced bounds on spectral expansion using the same overall proof idea. Before going into the mathematical details, we shall first provide some intuition. The goal is to compare finite graphs to the infinite d-regular tree, but how to go about this is by no means obvious. The trick that is classically applied, and which will be our focus as well, is to count closed walks. Let G be a d-regular graph, v be a vertex of G and v' a vertex of T_d. Denote by $W_\ell(G, v)$ the number of closed walks on G of length $\ell \in \mathbb{N}$ starting at v. The for any $\ell \in \mathbb{N}$, $W_\ell(G, v) \geq W_\ell(T_d, v')$. To see this, note that a closed walk on T_d starting at v' only can be closed if it backtracks to its origin. Every such backtracking walk can be mapped to a unique backtracking walk on G starting at v. Walks traversing cycles will also be counted by $W_\ell(G, v)$ while such walks are impossible on T_d by definition. Furthermore, closed walks on G relate to the spectrum of the adjacency matrix A of G. One can show that the number of closed walks on G of length ℓ is given by

$$\sum_{v\in V(G)} W_\ell(G, v) = \mathrm{Tr}(A^\ell) = \sum_{i=1}^{n} \lambda_i^\ell$$

where n is the number of vertices of G and $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A counted with multiplicity. Thus, we arrive at an inequality relating T_d to G and a way to translate this into a statement about the spectrum of G.

4.1 A First Bound on Spectral Expansion

Let us begin with a first bound on spectral expansion. It will be non-optimal, but we present it here as a warm-up to show how the aforementioned proof strategy will work. To this end, we first recall Lemma 2.10 from which the following lemma is a corollary.

Lemma 4.2. *Let $M \in \mathbb{R}^{n\times n}$ and $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of M counted with multiplicity. Then for every $k \in \mathbb{N}$,*

$$\mathrm{Tr}\left(M^k\right) = \sum_{i=1}^{n} \lambda_i^k.$$

Second, we note that walks on a graph are counted by powers of its adjacency matrix.

Lemma 4.3. *Let $G = (V, E)$ be a graph and A its adjacency matrix. For every $u, v \in V$ and $k \in \mathbb{N}$, the number of walks on G of length k, starting at u and ending at v, is given by $\left(A^k\right)_{uv}$. In particular, the number of closed walks on G of length $k \in \mathbb{N}$ is given by $\operatorname{Tr}\left(A^k\right)$.*

Proof. We proceed by induction on k. For $k = 1$ the statement is trivially true. Suppose that the statement is true for some $k \in \mathbb{N}$. Let $u, v \in V$ be given and let $w_1, \ldots, w_d$ be the neighbours of v counted with multiplicity. We find that

$$(A^{k+1})_{uv} = \sum_{x \in V} (A^k)_{ux} A_{xv} = \sum_{i=1}^{d} (A^k)_{uw_i}.$$

By the induction hypothesis, $(A^k)_{uw_i}$ is the number of walks from u to w_i of length k. The conclusion follows. □

We are now ready to use the connection between closed walks, the trace of the adjacency matrix, and the spectrum of the adjacency matrix, to get the following naive bound.

Proposition 4.4. *Let $d > 1$ be fixed and $G = (V, E)$ be a d-regular graph on $n = |V|$ vertices. Then $\lambda(G) \geq \sqrt{d} \cdot (1 - o_n(1))$, where $o_n(1) \to 0$ as $n \to \infty$*

Proof. Let A be the adjacency matrix of G and denote by $d = \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ the eigenvalues of A counted with multiplicity. From every vertex v of G there are at least d closed walks of length 2 on G, one for each neighbour of v. Thus, the number of closed walks on G of length 2 is at least dn. Combining this with Lemma 4.2 and Lemma 4.3 yields $\sum_{i=1}^{n} \lambda_i^2 = \operatorname{Tr}(A^2) \geq dn$. It follows that $d^2 + (n-1)\lambda(G)^2 \geq dn$, such that $\lambda(G) \geq \sqrt{d(n-d)/(n-1)}$. The conclusion follows. □

To get a better bound, we generalise and consider the number of walks of length 2ℓ for any $\ell \in \mathbb{N}$. This, however, is not as simple as it sounds. We shall require a much more sophisticated counting argument to arrive at our conclusion.

We shall need two facts regarding the Catalan numbers. The first is a well-known, albeit non-trivial, property of the Catalan numbers and we omit its proof.

Definition 4.5. [Catalan Numbers] The *Catalan numbers*, $(C_n)_{n \geq 0}$, are given by $C_0 = 1$ and $C_n = \frac{1}{n+1}\binom{2n}{n}, n \in \mathbb{N}$.

Lemma 4.6. *The Catalan numbers form the unique sequence recursively given by $C_0 = 1$ and $C_n = \sum_{i=1}^{n} C_{i-1} C_{n-i}, n \in \mathbb{N}$.*

Lemma 4.7. *For every $n \geq 1$, $C_n \geq \frac{4^n}{4n^2}$.*

Proof. The lemma holds for $n = 1$ as $C_1 = 1$. For $n > 1$, we first observe that by the binomial identity, $4^n = (1+1)^{2n} = \sum_{i=0}^{2n} \binom{2n}{i}$, which yields $\binom{2n}{n} \geq \frac{4^n}{2n+1}$, since $\binom{2n}{k} \leq \binom{2n}{n}$ for every $0 \leq k \leq 2n$. With this at hand,

$$C_n = \frac{1}{n+1}\binom{2n}{n} \geq \frac{4^n}{(n+1)(2n+1)} \geq \frac{4^n}{4n^2},$$

which is the desired conclusion. □

With this at hand, we proceed with the main theorem of the subsection.

Theorem 4.8. *Fix $d > 1$. Let $G = (V, E)$ be a d-regular graph on $n = |V|$ vertices. Then $\lambda(G) \geq 2\sqrt{d-1} - o_n(1)$, where $o_n(1) \to 0$ as $n \to \infty$.*

Proof. Denote by T_d the infinite d-regular tree, i.e., an infinite d-regular graph with no cycles. For any graph H and vertex $v \in V(H)$ denote by $W_\ell(H, v)$ the number of closed walks on H of length ℓ starting at v.

Let $e_1, \ldots, e_d$ be the edges adjacent to a vertex $v \in V(T_d)$ of the infinite d-regular tree. Denote by D_ℓ the set of closed walks of length 2ℓ starting at v and never traversing e_d. Clearly, $W_{2\ell}(T_d, v) \geq |D_\ell|$, so we shall derive a closed-form expression for $|D_\ell|$. Let $i \in [\ell]$ be given. We start by counting the number of walks $\omega \in D_\ell$ satisfying that the first time ω returns to v is after $2i$ steps. For every such ω, we choose a starting edge $e \in \{e_1, \ldots, e_{d-1}\}$ traversing which we reach vertex v'. We then choose a closed walk of length $2(i-1)$ starting at v' and not traversing e. After traversing e in the opposite direction to arrive at v as our $2i$th step, we choose any closed walk of $D_{\ell-i}$. Thus, the total number of such walks is $(d-1)D_{i-1}D_{\ell-i}$. Summing over $i \in [\ell]$, we get

$$D_\ell = \sum_{i=1}^{\ell} (d-1) D_{i-1} D_{\ell-i}.$$

Noting that $D_0 = 1$ and rewriting the recursing as

$$\frac{D_\ell}{(d-1)^\ell} = \sum_{i=1}^{\ell} \frac{D_{i-1}}{(d-1)^{i-1}} \cdot \frac{D_{\ell-i}}{(d-1)^{\ell-i}},$$

we observe that in fact $\frac{D_\ell}{(d-1)^\ell} = C_\ell$. In conclusion, $W_{2\ell}(T_d, v) \geq |D_\ell| = (d-1)^\ell C_\ell$.

Returning to the graph G, for every vertex v of G and v' of T_d, enumerate every edge e and e' leaving v and v', respectively, with a unique index $i \in [d]$ (self-loops are considered to correspond to two distinct directed edges). Since now any walk on G or T_d can be uniquely described as a sequence of indices corresponding to the edge chosen at each step in the walk, we obtain a bijection from the set of walks $(e_1, e_2, \ldots, e_\ell) \in E^\ell$ on G or T_d to the set of ℓ-tuples $(i_1, \ldots, i_\ell) \in [d]^\ell$. This immediately gives rise to a natural bijection from the set of walks on G starting at vertex v and the set of walks on T_d starting at vertex v'. Under this bijection, any closed walk on T_d starting at v' is also a closed walk on G starting at v since any closed walk on T_d backtracks to its origin. We conclude that for every $v \in V(G), v' \in V(T_d)$, and $\ell \in \mathbb{N}$, $W_\ell(G, v) \geq W_\ell(T_d, v')$.

We are now ready to bound the spectral expansion of G. Let A be the adjacency matrix of G, and let $\lambda_1 \geq \cdots \geq \lambda_n$ be the eigenvalues of A counted with multiplicity. By Lemma 4.3 and Lemma 4.2, it follows that

$$\sum_{i=1}^{n} \lambda_i^{2\ell} = \operatorname{Tr}\left(A^{2\ell}\right) = \sum_{v \in V} W_{2\ell}(G, v) \geq n W_{2\ell}(T_d, v') \geq n(d-1)^\ell C_\ell.$$

Thus, $d^{2\ell} + (n-1)\lambda(G)^{2\ell} \geq n(d-1)^\ell C_\ell$, since $\lambda_1 = d$. Now, set $n = \ell^{2\ell} + 1$, recall that $C_\ell \geq \frac{4^\ell}{4\ell^2}$

for $\ell \geq 1$, and note that $\frac{1}{\ell^{1/\ell}} \to 1$ as $\ell \to \infty$. Then we obtain

$$\begin{aligned}
\lambda(G) &\geq \sqrt[2\ell]{\frac{n(d-1)^\ell C_\ell}{n-1} - \frac{d^{2\ell}}{n-1}} \\
&\geq 2\sqrt{d-1} \cdot \sqrt[2\ell]{\frac{C_\ell}{2^{2\ell}}} - \frac{d}{(n-1)^{1/(2\ell)}} \\
&\geq 2\sqrt{d-1} - \left(2\sqrt{d-1}\left(1 - \sqrt[2\ell]{\frac{1}{4\ell^2}}\right) + \frac{d}{\ell}\right) \\
&= 2\sqrt{d-1} - o_n(1)
\end{aligned}$$

where $o_n(1) \to 0$ as $n \to \infty$. □

4.2 Distributional Results on Eigenvalues

In the following section, we shall establish a more general machinery for analysing the eigenvalues of the adjacency matrices of d-regular graphs. Let $(G_n)_{n\in\mathbb{N}}$ be a family of d-regular graphs with $|V(G_n)| \to \infty$ as $n \to \infty$, and let A_n be the adjacency matrix of G_n for every $n \in \mathbb{N}$. We prove that for any $\varepsilon > 0$, the number of eigenvalues of A_n in the interval $[(2-\varepsilon)\sqrt{d-1}, d]$ increases linearly with $|V(G_n)|$. Furthermore, we show that if $g(G_n) \to \infty$ as $n \to \infty$, i.e., the girth of $(G_n)_{n\in\mathbb{N}}$ tends to infinity with the number of vertices, then $(A_n)_{n\in\mathbb{N}}$ satisfies Wigner's semicircular law up to a normalisation constant. It is worth noting that the spectrum of the adjacency operator of the infinite d-regular tree is the Wigner semicircle distribution scaled by a constant factor. Theorem 4.8 presented in the previous subsection follows as an easy corollary.

More formally, we prove the following two theorems.

Definition 4.9. [Girth of Graph] Let G be a graph. The *girth* of G, denoted by $g(G)$, is the length of the shortest cycle on G.

Definition 4.10. [Wigner Semicircle Distribution] The *Wigner semicircle distribution* is the probability measure ν supported on $[-1, 1]$ with density function

$$f(x) = \frac{2\sqrt{1-x^2}}{\pi}.$$

Theorem 4.11. *Let $d \in \mathbb{N}$ be given. For every $\varepsilon > 0$, exists a constant $C(\varepsilon, d) > 0$, such that for any d-regular graph G on n vertices with adjacency matrix A, the number of eigenvalues of A (counted with multiplicity) in the interval $[(2-\varepsilon)\sqrt{d-1}, d]$ is at least Cn.*

Theorem 4.12. *Let $d \in \mathbb{N}$ be given. Let $(G_n)_{n\in\mathbb{N}}$ be a family of d-regular graphs satisfying that $|V(G_n)| \to \infty$ and $g(G_n) \to \infty$ as $n \to \infty$. Define the sequence of measures*

$$\nu_n = \frac{1}{|V(G_n)|} \sum_{i=1}^{|V(G_n)|} \delta_{\frac{\lambda_i^{(n)}}{2\sqrt{d-1}}},$$

where $\lambda_1^{(n)}, \ldots, \lambda_{|V(G_n)|}^{(n)}$ are the eigenvalues of the adjacency matrix of G_n counted with multiplicity, and δ_x is the Dirac measure at $x \in \mathbb{R}$. Then $(\nu_n)_{n\in\mathbb{N}}$ converges weakly to the Wigner semicircle distribution.

We follow the presentation of the results as found in [3]. However, the order of the presentation has been changed and care has been taken to highlight the main ideas and insights of the proofs.

4.2.1 The Wigner Semicircle Distribution and Chebyshev polynomials

The proofs of the theorems mentioned above crucially rely on connections between walks on graphs and the Chebyshev polynomials of the second kind. We shall start by introducing the Chebyshev polynomials of the second kind, prove some of their basic properties, and discuss their relation to the Wigner semicircle distribution.

Definition 4.13. The *Chebyshev polynomials of the second kind*, $(U_m)_{m\geq 0}$, are given recursively by $U_0(x) = 1$, $U_1(x) = 2x$, and for $m \in \mathbb{N}$, $U_{m+1}(x) = 2xU_m(x) - U_{m-1}(x)$.

The Chebyshev polynomials of the second kind are also uniquely determined by the following trigonometric identity.

Proposition 4.14. *For every $m \in \mathbb{N}_0$ and $\theta \in \mathbb{R}$, $U_m(\cos\theta) = \sin((m+1)\theta)/\sin\theta$.*

Proof. We prove that the expression satisfies the recurrence relation of the definition of U_m. For $m = 0$, there is nothing to show. For $m = 1$,

$$U_1(\cos\theta) = 2\cos\theta = \frac{2\cos\theta\sin\theta}{\sin\theta} = \frac{\sin(2\theta)}{\sin\theta}.$$

Lastly, for $m \in \mathbb{N}$,

$$\begin{aligned}\frac{\sin((m+2)\theta)}{\sin\theta} &= \frac{\cos\theta\sin((m+1)\theta) + \cos((m+1)\theta)\sin\theta}{\sin\theta} \\ &= 2\cos\theta \cdot \frac{\sin((m+1)\theta)}{\sin\theta} - \frac{\sin(m\theta)}{\sin\theta},\end{aligned}$$

where the last equality uses the fact that

$$\sin(m\theta) = \sin((m+1)\theta + (-\theta)) = \cos(\theta)\sin((m+1)\theta) - \cos((m+1)\theta)\sin(\theta).$$

This indeed matches $U_{m+1}(\cos\theta) = 2\cos\theta \cdot U_m(\cos\theta) - U_{m-1}(\cos\theta)$, and the conclusion follows. □

The recursion formula in the definition of the Chebyshev polynomials of the second kind allows us to find a closed expression for their generating functions.

Proposition 4.15. *The generating function of $(U_m(x))_{m\in\mathbb{N}_0}$ is given, as a formal power series over the ring of polynomials over $\mathbb{Q}[x]$, by*

$$\sum_{m=0}^{\infty} U_m(x)t^m = \frac{1}{1 - 2xt + t^2} \in (\mathbb{Q}[x])[[t]].$$

Proof. First, note that for every $m \in \mathbb{N}_0$, $U_m(x)$ has coefficients in $\mathbb{Q}$, so the generating function is indeed a formal power series over the polynomial ring of $\mathbb{Q}$.

Second, observe that by the recursion formula for U_m,

$$(1 - 2xt + t^2)\sum_{m=0}^{\infty} U_m(x)t^m = U_0(x) + (U_1(x) - 2x)t + \sum_{m=2}^{\infty}\left(U_m(x) - 2xU_{m-1}(x) + U_{m-2}\right)t^m = 1,$$

as we were required to show. □

The Weierstrass approximation theorem states that the set of polynomials over the reals are dense in the space of continuous functions on $[-1,1]$ equipped with the uniform norm, which is denoted by $C([-1,1])$. As the following proposition shows, we may use this fact to prove results regarding weak convergence of measures.

Proposition 4.16. *Let $L \geq 1$ be given, and let $(\nu_n)_{n\in\mathbb{N}}$ be a sequence of measures satisfying that for every $m \in \mathbb{N}_0$,*

$$\lim_{n\to\infty} \int_{-L}^{L} U_m(x)\, d\nu_n(x) = \begin{cases} 1, & m = 0 \\ 0, & m > 0. \end{cases}$$

Then $(\nu_n)_{n\in\mathbb{N}}$ converges weakly to the Wigner semicircle distribution.

Proof. Denote by ν the Wigner semicircle distribution. A change of variables yields

$$\begin{aligned} \int_{-L}^{L} U_m(x)\, d\nu(x) &= \int_{-1}^{1} U_m(x) \cdot \sqrt{1-x^2}\, dx \\ &= \int_0^{\pi} U(\cos(\theta)) \sin^2(\theta)\, d\theta \\ &= \int_0^{\pi} \sin((m+1)\theta)\sin(\theta)\, d\theta \\ &= \begin{cases} 1, & m = 0 \\ 0, & m > 0. \end{cases} \end{aligned}$$

Thus, for every $m \in \mathbb{N}_0$, $\lim_{n\to\infty} \int_{-L}^{L} U_m(x)\, d\nu_n(x) = \int_{-L}^{L} U_m(x)\, d\nu(x)$. To prove weak convergence of $(\nu_n)_{n\in\mathbb{N}}$ to ν, it suffices to extend this statement from the Chebyshev polynomials of the second kind to all functions of $C([-L,L])$.

To this end, note that every polynomial $P(x)$ of degree n is a linear combination of the polynomials $U_0, \ldots, U_n$. Thus, for every $P \in \mathbb{R}[x]$,

$$\lim_{n\to\infty} \int_{-L}^{L} P(x)\, d\nu_n(x) = \int_{-L}^{L} P(x)\, d\nu(x).$$

Let $f \in C([-L,L])$ be given, and let $\varepsilon > 0$. By the Weierstrass approximation theorem, there exists a polynomial $P(x) \in \mathbb{R}[x]$, such that for every $x \in [-L,L]$, $|f(x) - P(x)| < \varepsilon/3$. Furthermore, there exists an $N \in \mathbb{N}$, such that for every $n \geq N$, $\left|\int_{-L}^{L} P(x)\, d\nu_n(x) - \int_{-L}^{L} P(x)\, d\nu(x)\right| < \varepsilon/3$. Thus, for every $n \geq N$,

$$\begin{aligned} \left|\int_{-L}^{L} f(x)\, d\nu_n(x) - \int_{-L}^{L} f(x)\, d\nu(x)\right| &\leq \left|\int_{-L}^{L} P(x)\, d\nu(x) - \int_{-L}^{L} f(x)\, d\nu(x)\right| \\ &+ \left|\int_{-L}^{L} P(x)\, d\nu(x) - \int_{-L}^{L} P(x)\, d\nu_n(x)\right| \\ &+ \left|\int_{-L}^{L} P(x)\, d\nu_n(x) - \int_{-L}^{L} f(x)\, d\nu_n(x)\right| \\ &< \varepsilon. \end{aligned}$$

It follow that, indeed, $\lim_{n\in\mathbb{N}} \int_{-L}^{L} f(x)\, d\nu_n(x) = \int_{-L}^{L} f(x)\, d\nu(x)$, for every $f \in C([-L, L])$. $\square$

The attentive reader may notice that the above proposition reduces the proof of Theorem 4.12 to a statement regarding the evaluation of Chebyshev polynomials of the second kind at certain points relating to the eigenvalues of the adjacency matrices of graphs.

Remark 4.17. Invoking the Chebyshev polynomials of the second kind in the context of the Wigner semicircle distribution, is in fact much more natural than may be immediately clear. Let W be the real Hilbert space $L^2([-1, 1], \nu)$, where ν is the Wigner semicircle distribution on $[-1, 1]$, with associated inner product $\langle f, g\rangle = \int f\overline{g}\, d\nu$, for $f, g \in W$. Gram-Schmidt orthonormalisation in W of the polynomials $1, x, x^2, \cdots \in W$ yields exactly the Chebyshev polynomials of the second kind, $U_0, U_1, \cdots \in W$. To see this, it suffices to observe that $U_0 = 1$; for any $n \in \mathbb{N}_0$,

$$||U_n||^2 = \int_{-1}^{1} U_n^2(x) \cdot \frac{2\sqrt{1-x^2}}{\pi}\, dx = \frac{2}{\pi}\int_0^{\pi} \sin((n+1)\theta)^2\, d\theta = 1;$$

and for any $n \neq m$,

$$\langle U_n, U_m\rangle = \frac{2}{\pi}\int_0^{\pi} \sin((n+1)\theta)\sin((m+1)\theta)\, d\theta = \frac{2}{\pi}\int_0^{\pi} \cos((n-m)\theta) - \cos((n+m+2)\theta)\, d\theta = 0.$$

To prove Theorem 4.11, we shall need another property of the Chebyshev polynomials of the second kind.

Proposition 4.18. *Le $L \geq 2$ and $\varepsilon > 0$ be given. There exists a constant $C = C(L, \varepsilon)$, such that whenever μ is a probability measure on $[-L, L]$ satisfying that for all $m \in \mathbb{N}_0$,*

$$\int_{-L}^{L} U_m\left(\frac{x}{2}\right)\, d\mu(x) \geq 0, \tag{21}$$

then $\mu([2-\varepsilon, L]) \geq C$.

Proof. For simplicity, define $X_m(x) = U_m\left(\frac{x}{2}\right)$, for every $m \in \mathbb{N}_0$. We note that the polynomials $(X_m)_{m\in\mathbb{N}_0}$ satisfy the recursion $X_0(x) = 1$, $X_1(x) = x$, and $X_{m+1}(x) = xX_m(x) - X_{m-1}(x)$, for $m \in \mathbb{N}$. Furthermore, $X_m(2\cos\theta) = \frac{\sin((m+1)\theta)}{\sin\theta}$. It follows that for $m > 0$, U_m has m roots, $\{2\cos(2\pi k/(m+1))\}_{k\in[m]}$. We will denote by $\alpha_m = 2\cos(2\pi/(m+1))$ the largest root of X_m. Before proceeding with the proof, we shall first state some observations regarding the polynomials, $(X_m)_{m\in\mathbb{N}}$, and their roots.

First, we show by induction on k that whenever $k, \ell \in \mathbb{N}_0$, with $k \leq \ell$, $X_kX_\ell = \sum_{i=0}^{k} X_{k+\ell-2i}$. For $k = 0$ this is trivially true. For $k = 1$, $X_1X_\ell = xX_\ell = \sum_{i=0}^{1} X_{\ell-2i+1}$, by the recursion formula. Now, let $k \geq 2$, and assume the statement to be true for every $k' < k$. Then, by the induction

hypothesis and the recursion formula,

$$\begin{aligned}
X_k X_\ell &= x X_{k-1} X_\ell - X_{k-2} X_\ell \\
&= x \sum_{i=0}^{k-1} X_{k+\ell-1-2i} - \sum_{i=0}^{k-2} X_{k+\ell-2-2i} \\
&= \sum_{i=0}^{k-1} (X_{k+\ell-2i} + X_{k+\ell-2-2i}) - \sum_{i=0}^{k-2} X_{k+\ell-2-2i} \\
&= \sum_{i=0}^{k} X_{k+\ell-2i},
\end{aligned}$$

for every $\ell \geq k$.

Second, we prove that for every $m \in \mathbb{N}$,

$$\frac{X_m(x)}{x - \alpha_m} = \sum_{i=0}^{m-1} X_{m-1-i}(\alpha_m) X_i(x).$$

The recursion formula and rearrangement of terms yields

$$\begin{aligned}
x \sum_{i=0}^{m-1} X_{m-1-i}(\alpha_m) X_i(x) &= X_1(x) X_{m-1}(\alpha_m) + \sum_{i=1}^{m-1} X_{m-1-i}(\alpha_m)(X_{i+1}(x) + X_{i-1}(x)) \\
&= \sum_{i=1}^{m} X_{m-i}(\alpha_m) X_i(x) + \sum_{i=0}^{m-2} X_{m-i-2}(\alpha_m) X_i(x) \\
&= X_0(\alpha_m) X_m(x) + X_1(\alpha_m) X_{m-1}(x) + X_{m-2}(\alpha_m) X_0(x) \\
&\quad + \sum_{i=1}^{m-2} (X_{m-i}(\alpha_m) + X_{m-i-2}(\alpha_m)) X_i(x) \\
&= X_m(x) + \alpha_m X_{m-1}(x) + X_{m-2}(\alpha_m) + \alpha \sum_{i=1}^{m-2} X_{m-i-1}(\alpha_m) X_i(x) \\
&= X_m(x) + X_{m-2}(\alpha_m) + \alpha \sum_{i=1}^{m-1} X_{m-i-1}(\alpha_m) X_i(x).
\end{aligned}$$

Now, cancelling terms, applying the recursion formula yet again, and recalling that α_m is a root of X_m,

$$\begin{aligned}
(x - \alpha_m) \cdot \sum_{i=0}^{m-1} X_{m-1-i}(\alpha_m) X_i(x) &= X_m(x) + X_{m-2}(\alpha_m) - \alpha X_{m-1}(\alpha_m) X_0(x) \\
&= X_m(x) - X_m(\alpha_m) \\
&= X_m(x),
\end{aligned}$$

which proves the claim.

Third, for $m \in \mathbb{N}$ define the polynomium $Y_m = X_m(x)^2/(x - \alpha_m)$. We prove that there exist non-negative reals, $y_0^{(m)}, \ldots, y_{2m-1}^{(m)} \geq 0$, such that $Y_m(x) = \sum_{i=0}^{2m-1} y_i^{(m)} X_i(x)$. To this end, note that by the previous observation, $Y_m(x) = \sum_{i=1}^{m-1} X_{m-i-1}(\alpha_m) X_i(x) X_m(x)$. Since $\alpha_m > \alpha_j$ for every $j < m$, and α_j is the largest root of X_j, $X_{m-i-1}(\alpha_m) > 0$, for $1 \leq i \leq m-1$. Hence, the terms $X_i(x)X_m(x)$ all have non-negative coefficients in the above sum. Since $X_i(x)X_m(x)$ can be written as a linear combination of $X_0(x), \ldots, X_{2m-1}$ with non-negative coefficients, as per our first observation, the conclusion follows.

We are now ready to proceed with the proof of the proposition. Let $\varepsilon > 0$ be given. Suppose for contradiction that $\mu([2-\varepsilon, L]) = 0$, and let $m \in \mathbb{N}$ be such that $\alpha_m > 2 - \varepsilon$. Since for every $k \in \mathbb{N}_0$, $\int_{-L}^{L} X_k(x)\, d\mu(x) \geq 0$, it follows by the third observation above that $\int_{-L}^{L} Y_m(x)\, d\mu(x) \geq 0$. As $\mu([2-\varepsilon, L]) = 0$, this implies $\int_{-L}^{-\varepsilon} Y_m(x)\, d\mu(x) \geq 0$. However, $Y_m(x) \leq 0$ for every $x \leq \alpha_m$, so the support of μ must be contained in the roots of Y_m, $\{2\cos(2\pi k/(m+1))\}_{k\in[m]}$. The same argument applies to Y_{m+1}, since $\alpha_{m+1} > \alpha_m$, so the support of μ must also be contained in the roots of Y_{m+1}. The roots of Y_m and of Y_{m+1} are disjoint, however, since $m+1$ and $m+2$ are coprime. This implies that μ is the zero measure, a contradiction.

Let $C([-L, L])$ denote the space of continuous function from $[-L, L]$ to $\mathbb{R}$, endowed with the uniform norm. Then, since $[-L, L]$ is compact, it follows by the Riesz-Markov representation theorem that the dual of $C([-L, L])$ is the space $R([-L, L])$ of signed Radon measures on $[-L, L]$. The closed unit ball of the dual of a normed space is compact in the weak* topology, by the Banach–Alaoglu theorem, so the set of probability measures on $[-L, L]$, $\mathcal{P}([-L, L])$, which is a closed subset of the closed unit ball of $R([-L, L])$, is compact.

Now, let $\mathcal{W} \subset \mathcal{P}([-L, L])$ denote the set of probability measures satisfying (21), for all $m \in \mathbb{N}_0$. Since $\mathcal{W}$ is a countable intersection of closed sets,

$$\mathcal{W} = \bigcap_{m=0}^{\infty} \left\{ \mu \in \mathcal{P}([-L, L]) \mid \int U_m \, d\mu \geq 0 \right\},$$

$\mathcal{W}$ is closed, and hence compact. We may then consider the continuous function $f\colon [-L, L] \to \mathbb{R}$, given by

$$f(x) = \begin{cases} 0, & x \leq 2 - \varepsilon \\ \frac{2}{\varepsilon}(x - (2-\varepsilon)), & 2 - \varepsilon < x < 2 - \frac{\varepsilon}{2} \\ 1, & 2 - \frac{\varepsilon}{2} \leq x, \end{cases}$$

and observe that for every $\mu \in \mathcal{W}$,

$$\mu([2-\varepsilon, L]) \geq \int_{-L}^{L} f(x)\, d\mu(x) \geq \mu\left(\left[2 - \frac{\varepsilon}{2}, L\right]\right) > 0.$$

Since the map, $T\colon \mathcal{P}([-L, L]) \to \mathbb{R}$, given by $\mu \mapsto \int_{-L}^{L} f(x)\, d\mu(x)$, is weak* continuous, the image, $T(\mathcal{W}) > 0$, is compact in $\mathbb{R}$. Thus, there exists $C > 0$, such that for every $\mu \in \mathcal{W}$,

$$\mu([2-\varepsilon, L]) \geq \int_{-L}^{L} f(x)\, d\mu(x) \geq C,$$

which is what we set out to prove. □

4.2.2 Recursions on Non-backtracking Walks

In our proof of Theorem 4.8, we computed the number of closed walks on a graph and compared it to the number of closed walks on the infinite d-regular tree, our "ideal" expander. In this section, we shall instead consider the number of non-backtracking closed walks. The main goal of this subsection is to prove a recurrence relation that will allow us to later relate non-backtracking walks to the Chebyshev polynomials of the second kind. In the following, $G = (V, E)$ is a d-regular graph with adjacency matrix A.

Definition 4.19. [Non-backtracking Walks] A walk, $\omega = (v_0, v_1, \ldots, v_n), v_i \in V$, on G is a *non-backtracking walk*, if for every $i \in [n-1]$, $v_{i+1} \neq v_{i-1}$, i.e., the walk never crosses from vertex v_{i-1} to vertex v_i, and then crosses back to v_{i-1}. For $r \in \mathbb{N}_0$, we shall denote by A_r the matrix counting the number of non-backtracking walks of length r between two vertices. In other words, $(A_r)_{u,v}$ is the number of non-backtracking walks on G from u to v.

Lemma 4.20. *The matrices $(A_r)_{r\in\mathbb{N}_0}$ satisfy the following relations.*

a) $A_0 = I$ and $A_1 = A$.

b) $A_1^2 = A_2 + d \cdot I$.

c) For $r \geq 2$, $A_1 A_r = A_r A_1 = A_{r+1} + (d-1)A_{r-1}$.

Proof.

a) This follows by definition.

b) The walks of length two, where backtracking occurs, are exactly the walks of the form (v, u, v), where $(v, u) \in E$. These walks are thus captured by the matrix $d \cdot I$, since each vertex v has degree d. Since $A_1^2 = A^2$ counts the walks of length two, $A_1^2 - d \cdot I$ exactly counts the number of non-backtracking walks of length two.

c) Let $u_0, v \in V$. The entry $(A_r A_1)_{u_0 v}$ is exactly the number of walks of the form $(u_0, u_1, \ldots, u_r, v)$ where for each $i \in [r-1]$, $u_{i-1} \neq u_{i+1}$. In other words, the number of walks from u_0 to v of length $r+1$ where only the last edge traversal may be back-tracking. On the other hand, $(A_{r+1})_{u_0 v}$ is the number of walks of the form $(u_0, u_1, \ldots, u_r, v)$ where for each $i \in [r-1]$, $u_{i-1} \neq u_{i+1}$, and furthermore, $v \neq u_{r-1}$. The paths of the form $(u_0, u_1, \ldots, u_{r-2}, u_{r-1} = v, u_r, v)$ where for each $i \in [r-1]$, $u_{i-1} \neq u_{i+1}$, are exactly counted by $(d-1) \cdot (A_{r-1})_{u_0 v}$, since given a non-backtracking walk $(u_0, u_1, \ldots, v)$ of length $r-1$, there are exactly $(d-1)$ ways of choosing u_r, such that $(u, u_1, \ldots, u_{r-2}, v, u_r)$ is non-backtracking. It follows that $A_1 A_r = A_{r+1} + (d-1)A_{r-1}$ since both sides of the equation count the same thing. A symmetric argument establishes $A_1 A_r = A_{r+1} + (d-1)A_{r-1}$. □

Generating functions will be the key tool in connecting non-backtracking walks to the Chebyshev polynomials of the second kind. We start by finding an expression for the generating function of $(A_r)_{r\in\mathbb{N}_0}$. Note that this generating function is a formal power series over the ring $\mathrm{End}(\ell^2(V))$.

Lemma 4.21. *The generating function of $(A_r)_{r\in\mathbb{N}_0}$ is given by*

$$\sum_{r=0}^{\infty} A_r t^r = \frac{I - It^2}{I - At + ((d-1) \cdot I)t^2}.$$

Proof. Repeated application of Lemma 4.20 yields

$$\begin{aligned}(I - At + ((d-1)\cdot I)t^2)\cdot\sum_{r=0}^{\infty} A_r t^r &= A_0 + (A_1 - AA_0)t + \sum_{r=2}^{\infty}(A_r - A_1A_{r-1} + (d-1)A_{r-2})\,t^r \\ &= I + (A_2 - A_1A_1 + (d-1)A_0)\,t^2 \\ &= I - It^2,\end{aligned}$$

from which the conclusion follows. □

For simplicity, we shall henceforth write the identity matrix as 1, or omit it if it is the coefficient of some term t^n in a formal power series. We now introduce the quantity $T_m = \sum_{r=0}^{\lfloor \frac{m}{2} \rfloor} A_{m-2r}$, since its generating function resembles that of the Chebyshev polynomials of the second kind.

Lemma 4.22. *The generating function of* $(T_m)_{m\in\mathbb{N}_0}$ *is*

$$\sum_{m=0}^{\infty} T_m t^m = \frac{1}{1 - At + (d-1)t^2}.$$

Proof. We observe that

$$\begin{aligned}\sum_{m=0}^{\infty} T_m t^m &= \sum_{r=0}^{\infty}\sum_{m=2r}^{\infty} A_{m-2r}t^m \\ &= \left(\sum_{r=0}^{\infty} t^{2r}\right)\left(\sum_{m=0}^{\infty} A_m t^m\right) \\ &= \frac{1}{1-t^2}\cdot\frac{1-t^2}{1-At+(d-1)t^2} \\ &= \frac{1}{1-At+(d-1)t^2},\end{aligned}$$

from which the conclusion follows. □

4.2.3 Walks and Eigenvalues of Regular Graphs

With the above at hand, we shall draw a connection between the Chebyshev polynomials of the second kind and closed non-backtracking walks on graphs. One may notice that the generating functions of T_m and U_m, as presented in Lemma 4.22 and Proposition 4.15, are quite alike. A simple change of variables allows us to equate the two in a striking manner.

Lemma 4.23. *For every* $m \in \mathbb{N}_0$, $T_m = (d-1)^{\frac{m}{2}} U_m\left(\frac{A}{2\sqrt{d-1}}\right)$.

Proof. A change of variables in Proposition 4.15, to $x = \frac{A}{2\sqrt{d-1}}$ and $t = s\cdot\sqrt{d-1}$, yields

$$\sum_{m=0}^{\infty}(d-1)^{\frac{m}{2}} U_m\left(\frac{A}{2\sqrt{d-1}}\right)s^m = \frac{1}{1-As+(d-1)s^2}.$$

Noting that this is exactly the generating function of T_m completes the proof. □

We turn our attention to closed non-backtracking walks on G. Let $K_\ell(v)$ denote the number of closed non-backtracking walks on G of length ℓ starting and ending at $v \in V$. Then $K_\ell(v) = (A_\ell)_{v,v}$, such that

$$\sum_{v\in V} K_\ell(v) = \sum_{v\in V} (A_\ell)_{v,v} = \mathrm{Tr}(A_\ell).$$

The following key theorem is immediate.

Theorem 4.24. *Let G be a d-regular graph with adjacency matrix A. Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of A counted with multiplicity. For every $m \in \mathbb{N}_0$,*

$$\sum_{v\in V} \sum_{0\le r\le m/2} K_{m-2r}(v) = (d-1)^{m/2} \sum_{i=1}^{n} U_m\left(\frac{\lambda_i}{2\sqrt{d-1}}\right).$$

Proof. By Lemma 2.10 and Lemma 4.23,

$$\begin{aligned}\sum_{v\in V} \sum_{0\le r\le m/2} K_{m-2r}(v) &= \sum_{0\le r\le m/2} \mathrm{Tr}\left(A_{m-2r}\right)\\ &= \mathrm{Tr}\left(T_m\right)\\ &= \mathrm{Tr}\left((d-1)^{\frac{m}{2}} U_m\left(\frac{A}{2\sqrt{d-1}}\right)\right)\\ &= (d-1)^{\frac{m}{2}} \sum_{i=1}^{n} U_m\left(\frac{\lambda_i}{2\sqrt{d-1}}\right).\end{aligned}$$

The theorem follows. □

We are now ready to prove the two main theorems of the section.

Theorem 4.11. *Let $d \in \mathbb{N}$ be given. For every $\varepsilon > 0$, exists a constant $C(\varepsilon, d) > 0$, such that for any d-regular graph G on n vertices with adjacency matrix A, the number of eigenvalues of A (counted with multiplicity) in the interval $[(2-\varepsilon)\sqrt{d-1}, d]$ is at least Cn.*

Proof. Let $\varepsilon > 0$ be given, and let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of A counted with multiplicity. Define the probability measure μ on $\left[-\frac{d}{\sqrt{d-1}}, \frac{d}{\sqrt{d-1}}\right]$ by

$$\mu = \frac{1}{n}\sum_{i=1}^{n} \delta_{\frac{\lambda_i}{\sqrt{d-1}}},$$

where δ_x is the Dirac measure at $x \in \mathbb{R}$. By Theorem 4.24, it follows that for every $m \in \mathbb{N}_0$,

$$\int_{-d}^{d} U_m\left(\frac{x}{2}\right)\, d\mu(x) = \frac{1}{n}\sum_{i=1}^{n} U_m\left(\frac{\lambda_i}{2\sqrt{d-1}}\right) = \frac{1}{n(d-1)^{m/2}} \sum_{v\in V} \sum_{0\le r\le m/2} K_{m-2r}(v) \ge 0.$$

According to Proposition 4.18 this implies that there exists a constant, C, depending only on d and ε, such that $\mu([2-\varepsilon, d]) \ge C$. Observing that

$$\mu([2-\varepsilon, d]) = \frac{1}{n}\left|\{i \in [n] \mid \lambda_i \ge (2-\varepsilon)\sqrt{d-1}\}\right|,$$

we find that there are at least Cn eigenvalues of A in the interval $[(2-\varepsilon)\sqrt{d-1}, d]$. □

Note that the above theorem has Theorem 4.8 as a corollary. Finally, we arrive at the second major theorem of the section.

Theorem 4.12. *Let $d \in \mathbb{N}$ be given. Let $(G_n)_{n\in\mathbb{N}}$ be a family of d-regular graphs satisfying that $|V(G_n)| \to \infty$ and $g(G_n) \to \infty$ as $n \to \infty$. Define the sequence of measures*

$$\nu_n = \frac{1}{|V(G_n)|} \sum_{i=1}^{|V(G_n)|} \delta_{\frac{\lambda_i^{(n)}}{2\sqrt{d-1}}},$$

where $\lambda_1^{(n)}, \ldots, \lambda_{|V(G_n)|}^{(n)}$ are the eigenvalues of the adjacency matrix of G_n counted with multiplicity, and δ_x is the Dirac measure at $x \in \mathbb{R}$. Then $(\nu_n)_{n\in\mathbb{N}}$ converges weakly to the Wigner semicircle distribution.

Proof. By Proposition 4.16, it suffices to show that as n tends to infinity, $\int_{-d}^{d} U_m(x)\, d\nu_n$ converges to 0 when $m \in \mathbb{N}$, and converges to 1 when $m = 0$. To this end, first note that when $m = 0$, $U_m(x) = 1$ for every x, so the statement is trivially true. Second, observe that for $m \in \mathbb{N}$, there exists $N \in \mathbb{N}$, such that for every $n \geq N$, $g(G_n) > m$. Denote by $K_\ell^{(n)}(v)$ the number of closed non-backtracking walks of length ℓ on G_n starting at $v \in V(G_n)$. When $g(G_n) > m$, there are no closed non-backtracking walks of length $\leq m$. Thus, whenever $n \geq N$, $K_\ell^{(n)}(v) = 0$ for every $v \in V(G_n)$ and $\ell \leq m$. It now follows by Theorem 4.24 that for every $n \geq N$,

$$\int_{-d}^{d} U_m(x)\, d\nu_n = \frac{1}{n} \sum_{i=1} U_m\left(\frac{\lambda_i^{(n)}}{2\sqrt{d-1}}\right) = \frac{1}{n(d-1)^{m/2}} \sum_{v\in V} \sum_{0\leq r\leq m/2} K_{m-2r}^{(n)}(v) = 0.$$

The was what we needed to show. □

5 Spectral Expansion of Random Graphs

A body of research has been dedicated to the study of the spectral expansion of random regular graphs. We shall consider two such lines of research. First, applying the probabilistic method, one may establish the existence of good spectral expanders. Notable examples are the work by Bilu and Linial [1] proving the existence of near-Ramanujan spectral expanders, and the recent establishment of the existence of bipartite Ramanujan graphs of every degree and size by Marcus, Spielman, and Srivastava [8, 9]. Second, let G be a random graph. Another body focuses on the expected value of $\lambda(G)$. As we show below, the distribution of $\lambda(G)$ is strongly concentrated around its mean, so bounding the mean value equates to proving a bound on the spectral expansion of "most" regular graphs. Friedman [4, 5] managed to find the optimal bound in this regard, showing that for every $\varepsilon > 0$ and $d \in \mathbb{N}$, a random d-regular graph G satisfies $\lambda(G) < 2\sqrt{d-1} + \varepsilon$ with probability $1 - O_\varepsilon(|V(G)|^{-C})$, for a constant, $C > 0$, depending on d. This resolved a long-standing conjecture by Noga Alon. Thus, a random graph is guaranteed to almost be Ramanujan with high probability. A consequence of this, is that one can find a good spectral expander by simply picking random graphs until one is found. In expectation, only a constant number of samples would be necessary, implying an efficient procedure for producing good spectral expanders.[2]

[2]This sweeps under the rug, the issue of determining the eigenvalues of the adjacency matrix of a sampled graph, something that takes time polynomial in the number of vertices to even approximate.

In the following, we shall start by describing a framework for modelling random graphs. We then proceed to prove that for a random graph G, $\lambda(G)$ is concentrated around its mean. Lastly, we present an upper bound on $\lambda(G)$. In our treatment of the subject, we follow Broder and Shamir [2] with inspiration from the presentation of the proof found in the survey by Hoory, Linial, and Wigderson [6].

5.1 Modelling a Random Regular Graph

There are multiple ways to define random regular graphs. We shall consider the *permutation model* for a random $2d$-regular graph. It is defined as a union of d permutations on the vertices of the graph.

Definition 5.1. [Permutation Model] The distribution $\mathcal{G}_{n,2d}$ is given as follows. For each $i \in [d]$, let $\pi_i \colon [n] \to [n]$ be a permutation. The graph $G(\pi_1, \ldots, \pi_d) = (V, E)$ with vertices $V = [n]$ is generated by adding an edge between j and $\pi_i(j)$ for each $i \in [d]$ and $j \in [n]$. One samples from $\mathcal{G}_{n,d}$ by sampling eachpermutation $\pi_i \in S_n$ independently and uniformly at random and then constructing $G(\pi_1, \ldots, \pi_d)$.

We observe that every G sampled from $\mathcal{G}_{n,2d}$ is indeed $2d$-regular, but may contain self-loops or multiple edges. We shall allow this for two reasons. First, the expansion properties of the graph are not influenced by multiple edges and self-loops, so restricting to simple graphs is not really necessary for the discussion of expander graphs. Second, the distribution $\mathcal{G}_{n,2d}$ is easier to work with than other models of random graphs, such as, for instance, the uniform distribution on the simple $2d$-regular graphs. The results presented carry over to most other models of random graphs by general arguments from the theory of random graphs. We shall, however, not cover this phenomenon in detail here.

5.2 Concentration Around the Mean

We start our treatment of the spectral expansion of random graphs with a proof that the spectral expansion is concentrated around its mean. Towards such a result, we state the Azuma-Hoeffding inequality, a Chernoff-style bound for martingales.

Proposition 5.2 (Azuma-Hoeffding)**.** *Let the adapted sequence $(X_i, \mathcal{F}_i)_{0 \le i \le n}$ be a martingale with X_0 constant. If $|X_{i+1} - X_i| \le c$, for every $0 \le i < n$, then $P\left(|X_n - X_0| > \gamma c \sqrt{n}\right) \le 2e^{-\gamma^2/2}$.*

With this at hand, we are ready to prove the following concentration bound.

Theorem 5.3. *Let $d \ge 2$ be given, and let G be a graph drawn from the distribution $\mathcal{G}_{n,2d}$. Then*

$$P\left(|\lambda(G) - \mathbb{E}\left[\lambda(G)\right]| > \gamma\sqrt{d}\right) \le 2e^{-\frac{\gamma^2}{32}}.$$

Proof. The $2d$-regular graph, $G = G(\pi_1, \ldots, \pi_d)$, is generated by uniformly random permutations $\pi_1, \ldots, \pi_d \in S_n$, as discussed above. Denote by Ω the associated probability space, and write $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and $\mathcal{F}_i = \sigma(\pi_1, \ldots, \pi_i), i \in [d]$. For $0 \le i \le n$, define $X_i = \mathbb{E}\left[\lambda(G) \mid \mathcal{F}_i\right]$, and note that the adapted sequence $(X_i, \mathcal{F}_i)$ is a martingale. Clearly, $X_0 = \mathbb{E}\left[\lambda(G)\right]$ and $X_d = \lambda(G)$, such that

$$P\left(|\lambda(G) - \mathbb{E}\left[\lambda(G)\right]| > \gamma\sqrt{d}\right) = P\left(|X_d - X_0| > \gamma\sqrt{d}\right) \le 2e^{-\frac{\gamma^2}{32}},$$

will follow directly from Proposition 5.2, if we can show that $|X_i - X_{i-1}| \leq 4$ almost surely, for every $i \in [d]$. Since $X_{i-1} = \mathbb{E}\left[\mathbb{E}\left[\lambda(G) \mid \mathcal{F}_{i-1}, \pi_i\right] \mid \mathcal{F}_{i-1}\right]$, it follows that almost surely,

$$|X_i - X_{i-1}| \leq \max_{\pi,\pi' \in S_n} \left|\mathbb{E}\left[\lambda(G) \mid \mathcal{F}_{i-1}, \pi_i = \pi\right] - \mathbb{E}\left[\lambda(G) \mid \mathcal{F}_{i-1}, \pi_i = \pi'\right]\right|. \tag{22}$$

Let $\lambda_1 \geq \cdots \geq \lambda_n$ be the eigenvalues of the adjacency matrix, A, of G, counted with multiplicity. Note that $A_{ij} = \sum_{k=1}^{d}([\pi_k(i) = j] + [\pi_k(j) = i])$, such that by the min-max theorem,

$$\lambda_2 = \sup_{\substack{x \in \{1_V\}^\perp \\ ||x||_2 = 1}} \langle Ax, x\rangle = \sup_{\substack{x \in \{1_V\}^\perp \\ ||x||_2 = 1}} \sum_{1 \leq i,j \leq n} A_{ij} x_i x_j = \sup_{\substack{x \in \{1_V\}^\perp \\ ||x||_2 = 1}} \sum_{k=1}^{d} \sum_{i=1}^{n} 2 x_i x_{\pi_k(i)}.$$

In the same fashion, $\lambda_n = \inf_{||x||_2=1} \sum_{k=1}^{d} \sum_{i=1}^{n} 2x_i x_{\pi_k(i)}$. Let $i \in [d]$ be given, and fix $\pi_1, \ldots, \pi_{d-1} \in S_n$. For $\pi \in S_n$, denote by $\lambda_2(\pi)$ and $\lambda_n(\pi)$ the corresponding eigenvalues of the adjacency matrix of the graph generated by $\pi_1, \ldots, \pi_{d-1}$, and π. Then for any pair of permutations, $\pi, \pi' \in S_n$,

$$\begin{aligned}
|\lambda_2(\pi) - \lambda_2(\pi')| &\leq \sup_{\substack{x \in \{1_V\}^\perp \\ ||x||_2 = 1}} \left(\left|\sum_{i=1}^{n} 2x_i x_{\pi(i)} - 2x_i x_{\pi'(i)}\right|\right) \\
&\leq \sup_{\substack{x \in \{1_V\}^\perp \\ ||x||_2 = 1}} 2\left(\left|\sum_{i=1}^{n} x_i x_{\pi(i)}\right| + \left|\sum_{i=1}^{n} x_i x_{\pi'(i)}\right|\right) \\
&\leq 4,
\end{aligned}$$

where the last inequality follows by the Cauchy-Schwartz inequality. A very similar computation yields $|\lambda_n(\pi) - \lambda_n(\pi')| \leq 4$. It follows that any two graphs, G and G', generated by permutations $\pi_1, \ldots, \pi_d$ and $\tau_1, \ldots, \tau_d$, respectively, such that $\tau_i = \pi_i$ for all but one $i \in [d]$, satisfy $|\lambda(G) - \lambda(G')| \leq 4$. Applying this insight to (22) yields $|X_i - X_{i-1}| \leq 4$ almost surely, which completes the proof. □

5.3 A Bound on the Mean Spectral Expansion

We have established that the spectral expansion of a random graph in the permutation model is concentrated around its mean. We now turn to bounding this mean from above. The best known such bound is due to Friedman [5] and states that $\mathbb{E}\left[\lambda(G)\right] \leq 2\sqrt{d-1} + o(1)$. However, that result is far beyond the scope of this exposition. Instead, we shall prove a bound by Broder and Shamir [2], the first bound achieving even $\mathbb{E}\left[\lambda(G)\right] = o(d)$. With its elegance and simplicity, the proof and the framework of Broder and Shamir influenced the later series of papers culminating in the result by Friedman.

Theorem 5.4. *Let $d \geq 2$ be given and let G be a graph drawn from the distribution $\mathcal{G}_{n,2d}$. Then $\mathbb{E}\left[\lambda(G)\right] \leq \sqrt[4]{2} \cdot d^{3/4} \cdot (1 + o(1))$ as n tends to infinity.*

Before proceeding to the proof of the above theorem, let us first discuss the strategy and framework of the proof. As in Section 4, we shall discuss walks on graphs to estimate $\lambda(G)$. Recall that for a graph G with adjacency matrix A and any $k \in \mathbb{N}$, $\lambda(G)^{2k} \leq \sum_{i=2}^{n} \lambda_i^{2k} = \mathrm{Tr}(A^{2k}) - d^{2k}$, where

$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ are the eigenvalues of A counted with multiplicity. Expressed using the random walk matrix, M, of the random graph G, this becomes $\left(\frac{\lambda(G)}{d}\right)^{2k} \leq \mathrm{Tr}(M^{2k}) - 1$. By Jensen's inequality,

$$\mathbb{E}\left[\lambda(G)\right] \leq d \cdot \left(\mathbb{E}\left[\mathrm{Tr}(M^{2k})\right] - 1\right)^{\frac{1}{2k}}. \tag{23}$$

Thus, we shall aim to bound $\mathbb{E}\left[\mathrm{Tr}(M^{2k})\right]$, which is equal to n times the probability that a random walk on the random graph G of length $2k$, starting at a given vertex $v \in V(G)$, is closed. For the remainder of the section, we identify V with the set $[n]$, and let the starting vertex be $1 \in V = [n]$.

Now, consider $G = (V, E)$ to be drawn from the distribution $\mathcal{G}_{n,2d}$. The graph G is then defined by random permutations $\pi_1, \ldots, \pi_d \in S_n$ and any edge of G is of the form $\{u, \pi_i(u)\}$ or $\{u, \pi_i^{-1}(u)\}$, $u \in V, i \in [d]$. Thus, there is a bijective correspondence between words ω of length t over the alphabet $\Sigma = \{\pi_1, \pi_1^{-1}, \ldots, \pi_d, \pi_d^{-1}\}$ and the walks of length t on G starting at vertex 1. As an example, the word $\omega = \pi_1\pi_2^{-1}\pi_1^{-1}$ corresponds to the walk

$$1 \to v_1 = \pi_1(1) \to v_2 = \pi_2^{-1}(v_1) \to v_3 = \pi_1^{-1}(v_2).$$

A word of length t over the alphabet Σ can be regarded as a member of the free group with generating set $\{\pi_1, \ldots, \pi_d\}$. It will be natural to consider the reduced word $\omega' = \mathrm{red}(\omega)$, where each substring of ω of the form $\alpha\alpha^{-1}, \alpha \in \Sigma$, has been repeatedly removed until no such string occurs.

To perform a random walk of length $2k$ on the random graph G, we pick a random word, ω, of length $2k$ over the alphabet Σ, and then choose a random assignment of permutations to $\pi_1, \ldots, \pi_d \in S_n$. The word ω then describes a walk on $G(\pi_1, \ldots, \pi_d)$. We shall be interested in whether or not the walk described by ω is closed. We shall analyse this in two steps. We call a reduced word ω' *bad* if it is of the form $\omega' = \omega_1\omega_2^j\omega_1^{-1}$ for some $j \geq 2$ and words ω_1, ω_2, where taking inverses should be interpreted as in the free group over $\{\pi_1, \ldots, \pi_d\}$. It is worth noting that the empty word is bad. First, we show that the reduced word, $\omega' = \mathrm{red}(\omega)$, of a random word is unlikely to be bad. Second, we show that any word that is not bad is unlikely to yield a closed walk on G. We prove this in two lemmas and then finally conclude with a proof of Theorem 5.4.

Lemma 5.5. *Let ω be a uniformly drawn word of length $2k$ from Σ and let ω' be the reduced word. The probability that ω' is bad is less than $4k^2 \cdot \left(\frac{2}{d}\right)^k$.*

Proof. We bound the number of words of length $2k$ that reduce to a bad word of length 2ℓ. Note that a word of even length can only reduce to another word of even length. Let $\omega = a_1a_2\ldots a_{2k}$ be a word, let $a_{i_1}a_{i_2}\ldots a_{i_{2\ell}} = \omega' = \mathrm{red}(\omega)$ be the reduced word of ω, and suppose that ω' is bad. We call an index j *staying* if $j \in \{i_1, \ldots, i_{2\ell}\}$. If the index $j \in [k]$ satisfies that a_j is eliminated when reducing from ω to ω', we call it *opening* if a_j gets cancelled out with some $a_k, k > j$ and *closing* otherwise. We make this notion well-defined by requiring that the reduction occurs by repeatedly passing over ω from left to right and cancelling each subword of the form $\alpha\alpha^{-1}, \alpha \in \Sigma$ in that order until no such subwords occur. Denote by $A, B, C \subset [2k]$, respectively, the set of opening, closing, and staying indices. We claim that A uniquely defines B and C. This follows as for every $j \in [2k] \setminus A$, $j \in B$ if and only if $|A \cap [j-1]| > |B \cap [j-1]|$, since if some a_j and a_k, $j < k$, cancel out in the reduction, then there can be no staying index $m \in C$ with $j < m < k$. Thus, if $|A \cap [j-1]| > |B \cap [j-1]|$ there is an index of $A \cap [j-1]$ the letter of which is not cancelled out by any letter with index $m < j$. Now, as $|A| = k - \ell$ it follows that there are at most $\binom{2k}{k-\ell}$ configurations of A, B, C that lead to a valid cancelling of letters. Fix such a configuration A, B, C.

The letters at the indices of A and B can be chosen in at most $(2d)^{k-\ell}$ ways since the choice for the indices of A determine the choice of indices for B. By assumption the reduced word ω' can be written in the form $\omega' = \omega_1\omega_2^j\omega_1^{-1}$ with $j \geq 2$. Then $|\omega_1| + |\omega_2| \leq \ell$ since $|\omega'| = 2\ell$. Thus, for every possible choice of lengths $t_1, t_2 \in \{0, 1, \ldots, \ell\}$ for ω_1 and ω_2, i.e.,$|\omega_1| = t_1$ and $|\omega_2| = t_2$, there are at most $(2d)^\ell$ ways of choosing ω' since it is fixed after choosing ω_1 and ω_2. It follows that we can choose ω' in at most $(\ell+1)^2(2d)^\ell$ ways.

In conclusion there are at most

$$\sum_{\ell=0}^{k} \binom{2k}{k-\ell}(2d)^{k-\ell} \cdot (\ell+1)^2(2d)^\ell \leq 4k^2(2d)^k \cdot \sum_{\ell=0}^{k} \binom{2k}{k-\ell} \leq 4k^2(8d)^k$$

ways to choose ω such that the reduced word ω' is bad. Since we can choose ω in $(2d)^{2k}$ ways, it follows that the probability that ω reduces to a bad word is less than $4k^2(8d)^k/(2d)^{2k} = 4k^2\,(2/d)^k$. □

Lemma 5.6. *Let $k = o(n)$ and let ω be a word of length $2k$ over Σ satisfying that the corresponding reduced word, ω', is good. Let G be drawn from $\mathcal{G}_{n,2d}$. The probability that ω corresponds to a closed walk on G from vertex 1 to itself is at most $\frac{(2k)^4}{(n-2k)^2} + \frac{1}{n-2k}$. i.e.,*

$$P(\omega(1) = 1) \leq \frac{(2k)^4}{(n-2k)^2} + \frac{1}{n-2k}. \tag{24}$$

Proof. The walk corresponding to ω in G is closed if and only if the walk corresponding to ω' is closed. Hence, let the reduced word $\omega' = \sigma_1\sigma_2\ldots\sigma_s, \sigma_i \in \Sigma$, have length s and let us consider a walk as a sequence of random variables $v_0, v_1, \ldots, v_s$ given by $v_0 = 1$ and $v_i = \sigma_i(v_{i-1})$ for $i \in [s]$. We shall view this process as one in which every value $\sigma(a)$ for $\sigma \in \Sigma$ and $a \in [n]$ is undecided at the beginning. In each step, going from v_{i-1} to v_i reveals the value of σ_i on v_{i-1} and σ_i^{-1} on v_i but nothing else. A step from v_{i-1} to v_i is called *forced* if $\sigma(v_{i-1})$ has already been fixed by the preceding steps. Otherwise, we say that the step is *free*. If a free step from v_{i-1} to v_i coincides with a previously visited vertex, i.e., if $v_i = v_j, j < i$, we call the step a *coincidence*.

The probability that a coincidence occurs at step i is at most $\frac{i}{n-i}$ since the random choice of the value $\sigma(v_{i-1}) \in [n]$ must take a value in $\{v_0, \ldots, v_{i-1}\}$ and at most i values of σ can have been fixed before that choice. Thus, for distinct $i, j \in [s]$, the probability that both step i and step j are coincidences is less than $s^2/(n-s)^2$. Since there are less than s^2 ways to choose the pair of indices, i and j, the probability that two or more coincidences occur is less than $s^4/(n-s)^2$ by a union bound.

Since ω' is good, some coincidence must occur for the walk $v_0, v_1, \ldots, v_s$ to be closed and return to $v_0 = 1$. We will bound the probability that exactly one coincidence occurs and that the walk returns to its origin. Suppose that $v_0, \ldots, v_s$ is an instance of exactly such an outcome and suppose that the step from v_{i-1} to v_i is where the coincidence occurs. Then since ω' is reduced, the walk $v_0, \ldots, v_{i-1}$ is a simple path, i.e., it is a path with no repeating vertices. Now, the coincidence at step i turns the path into a cycle with a (possibly empty) tail. We must have $v_i = v_j$ for some $j < i$ and then $v_j, v_{j+1}, \ldots, v_i$ is a cycle while the tail consists of $v_0, \ldots, v_{j-1}$. The only coincidence was at step i, so all remaining steps are forced. This means that the walk will continue solely on already traversed edges. Since the walk $v_0, \ldots, v_s$ is closed and ω' is reduced, the only option for the walk is to continue around the cycle $t \in \mathbb{N}_0$ times and then traverse back along the tail. Thus, the walk

will have the form $v_0 v_1 \dots v_{j-1}(v_j \dots v_{i-1})^{t+1} v_j v_{j-1} \dots v_0$. Given that every move after the ith is forced, this means that ω' can be written as $\omega_1 \omega_2^t \omega_1^{-1}$ where $\omega_1 = \sigma_1 \dots \sigma_j$ and $\omega_2 = \sigma_{j+1} \dots \sigma_i$. Note that ω_2 cannot be the empty string because ω' is reduced. Furthermore, $\sigma_{j+1} \neq \sigma_i^{-1}$ since then step i would not be a free move. It follows that writing $\omega' = \omega_1 \omega_2^{t+1} \omega_1^{-1}$ is the unique way to write ω' in that form with the length of ω_1 maximised. Since ω' is good, we furthermore have $t = 0$. In order to conclude, we observe that the lengths of ω_1 and ω_2 were not dependent on G. Their lengths were deduced merely based on the fact that exactly one coincidence occurred. Thus, said coincidence must occur at step $|\omega_1| + |\omega_2|$ and is a free move to the specific vertex v_j. The probability of this event is at most $\frac{1}{n-s}$ since at most s values of σ_i can have been fixed at that point.

In conclusion, the probability that the walk returns to its origin is bounded by the sum of the probability that more than one coincidence occurs and the probability that the walk is closed with exactly one coincidence. Thus,

$$P(\omega(1) = 1) \leq \frac{(s)^4}{(n-s)^2} + \frac{1}{n-s} \leq \frac{(2k)^4}{(n-2k)^2} + \frac{1}{n-2k}$$

as we set out to prove. □

Combining the two lemmas with a union bound, we get the following corollary.

Corollary 5.7. *The probability that a randomly chosen walk on G from a fixed starting vertex is closed is less than $4k^2 (2/d)^k + (2k)^4/(n-2k)^2 + 1/(n-2k)$.*

We are now ready to prove the main result of the subsection.

Proof of Theorem 5.4. Let $\varepsilon > 0$ and set $k = \left\lfloor 2 \log_{d/2} n \right\rfloor$. The corollary above now states that the probability that a random walk on G sampled from $\mathcal{G}_{n,2d}$ of length $2k$ returns to its origin is less than $C \cdot \log^4 n/n^2 + 1/(n-2k)$, for some constant $C > 0$ independent of n. By (23) and the remark following it,

$$\begin{aligned}
\mathbb{E}\left[\lambda(G)\right] &\leq d \cdot \left(\mathbb{E}\left[\mathrm{Tr}(M^{2k})\right] - 1\right)^{\frac{1}{2k}} \\
&\leq d \cdot \left(n \cdot \left(\frac{C \cdot \log^4 n}{n^2} + \frac{1}{n-2k}\right) - 1\right)^{\frac{1}{2k}} \\
&= d \cdot \left(\frac{C \cdot \log^4 n}{n} + \frac{2k}{n-2k}\right)^{\frac{1}{2k}} \\
&\leq d \cdot \left(\frac{C' \cdot \log^4 n}{n}\right)^{\frac{1}{2k}}
\end{aligned}$$

for some new constant, $C' > 0$, independent of n. Now, $n^{-\frac{1}{2k}} \to (d/2)^{-1/4}$ as $n \to \infty$, while

$$(C' \cdot \log^4 n)^{\frac{1}{2k}} = e^{O\left(\frac{\log(C \cdot \log^4 n)}{\log n}\right)} \to 1$$

as $n \to \infty$. It follows that $d \cdot \left(C' \cdot \log^4 n/n\right)^{1/(2k)} \to \sqrt[4]{2} \cdot d^{3/4}$ as $n \to \infty$, such that, indeed, $\mathbb{E}\left[\lambda(G)\right] \leq \sqrt[4]{2} \cdot d^{3/4} \cdot (1 + o(1))$.

□

Combining the result above with our concentration bound, we find that the spectral expansion of a random graph is very unlikely to be significantly larger than $O(d^{3/4})$.

Corollary 5.8. *Let $d \geq 2$ be given. There exists a constant $N \in \mathbb{N}$ such that for every $n \geq N$, a graph G drawn from the distribution $\mathcal{G}_{n,2d}$ satisfies*

$$P\left(\lambda(G) \leq (2+\gamma)d^{3/4}\right) \geq 1 - 2\exp\left(-\frac{\gamma^2 d^{1/2}}{32}\right).$$

Proof. This follows directly from Proposition 5.2 and Theorem 5.4 since for sufficiently large n, $\mathbb{E}\left[\lambda(G)\right] \leq 2d^{3/4}$. □

6 Spectral Expanders from Cayley Graphs

Given a group $(\Gamma, \cdot)$ one may define a graph structure on the group elements in the following way.

Definition 6.1. [Cayley Graph] Let Γ be a group and $S \subset \Gamma$ be a symmetric subset of Γ, i.e., $g^{-1} \in S$ whenever $g \in S$. The *Cayley graph* of Γ and S is denoted by $\mathrm{Cay}\,(\Gamma; S) = (V, E)$. It has vertex set $V = \Gamma$ and edge set $E = \{\{g, gs\} \mid g \in \Gamma, s \in S\}$.

In other words, $\mathrm{Cay}\,(\Gamma; S)$ is a graph on the elements of Γ satisfying that the neighbours of any $g \in \Gamma$ are exactly the group elements $g \cdot S$. Cayley graphs are studied within many parts of mathematics. In expander theory, they are of great interest since their algebraic nature allows us to translate graph theoretic properties to group properties, which are often better understood.

In this section, we shall explore some very basic results and present a construction of a graph with spectral expansion $O(\sqrt{d})$. The first subsection loosely follows Section 4.3 of a book by Krebs and Shaheen [7] while the two last subsections are based on chapter 5 of lecture notes by Luca Trevisan [12].

6.1 Negative Results

We start out by proving a negative result, showing that the construction of good expanders from groups is not straight-forward. More specifically, we show that the expansion of d-regular Cayley graphs of Abelian groups with n vertices is close to d when d is kept constant and n approaches infinity. Our proof relies on the fact that the diameter of the Cayley graph of an Abelian group is large, while the diameter of an expander graph is small. The following lemma provides a bound to the effect of the latter.

Lemma 6.2. *Let G be a connected d-regular graph on n vertices and $h(G)$ be the Cheeger constant. Then* $\mathrm{diam}(G) \leq 2\log n / \log\left(1 + h(G)/d\right)$.

Note: The denominator of the right hand side of the inequality is well-defined since the Cheeger constant is strictly positive for connected graphs.

Proof. Let $u, v \in V$ be vertices and denote by $r_1, r_2 \in \mathbb{N}$ the smallest integers satisfying that $|B_{r_1}(v)|$ and $|B_{r_2}(u)|$ are both strictly larger than $n/2$. Focusing on v, note that $B_0(v) = 1$ and for every $s < r_1$, $B_s(v) \leq n/2$. Since every vertex of G has d neighbours and $B_{s+1}(v) = B_s(v) \cup \partial B_s(v)$,

$$|B_{s+1}(v)| \geq |B_s(v)| + \frac{|\partial B_s(v)|}{d} \geq |B_s(v)| \cdot \left(1 + \frac{h(G)}{d}\right)$$

for every $s < r_1$. Thus, $n \geq |B_{r_1}(v)| \geq (1 + h(G)/d)^{r_1}$, implying $r_1 \leq \log n / \log(1 + h(G)/d)$. The conclusion now follows, as $\mathrm{dist}(v, u) \leq r_1 + r_2$ and the choice of u and v was arbitrary. □

Proposition 6.3. *Let Γ be an Abelian group with $|\Gamma| = n$ and $S \subset \Gamma$ a symmetric subset with $|S| = d > 1$. The spectral expansion of the Cayley graph $\mathrm{Cay}(\Gamma; S)$ is bounded by*

$$|d - \lambda(\mathrm{Cay}(\Gamma; S))| \leq 2\left(\exp\left(\frac{4\log n}{n^{\frac{1}{d}} - e}\right) - 1\right).$$

Proof. Let $g \in \Gamma$ be given and enumerate the elements of S as $s_1, \ldots, s_d$. We shall bound the size of the ball $B_r(g)$ in $\mathrm{Cay}(\Gamma; S)$. For every $h \in B_r(g)$ there exists a path of length $t \leq r$ from g to h in $\mathrm{Cay}(\Gamma; S)$. Thus, we can write $h \cdot g^{-1}$ as a product of t elements of S, $h \cdot g^{-1} = a_1 a_2 \ldots a_t, a_i \in S$. Now, as Γ is Abelian, we can write $h \cdot g^{-1} = s_1^{\alpha_1} \ldots s_d^{\alpha_d}$ with $\sum_{i=1}^{d} \alpha_i = t$. Adding a dummy variable $\alpha_0 = r - t$, we get $\sum_{i=0}^{d} \alpha_i = r$. It is well-known that the number of ways to express a positive integer, r, as a sum of $d + 1$ non-negative integers is $\binom{r+d}{d}$. Thus, the number of such elements, $h \in B_r(g)$, can at most be $\binom{r+d}{d}$. In conclusion, $|B_r(g)| \leq \binom{r+d}{d} \leq (e(r + d)/d)^d$, by an elementary bound on the binomial coefficient.

Denote by D the diameter $D = \mathrm{diam}(\mathrm{Cay}(\Gamma; S))$ and by C the quantity $C = 1 + h(\mathrm{Cay}(\Gamma; S))$. We invoke Lemma 6.2 and the bound $2d > e$, to find that

$$n = B_D(g) \leq \left(\frac{e(D + d)}{d}\right)^d \leq \left(\frac{e\left(\frac{2\log n}{\log C} + d\right)}{d}\right)^d \leq \left(\frac{4\log n}{\log C} + e\right)^d,$$

which implies $1 + h(\mathrm{Cay}(\Gamma; S)) = C \leq \exp\left(4\log n/(n^{1/d} - 2d)\right)$. By the Cheeger inequality,

$$0 \leq d - \lambda(\mathrm{Cay}(\Gamma; S)) \leq 2h(\mathrm{Cay}(\Gamma; S)) \leq 2\left(\exp\left(\frac{4\log n}{n^{\frac{1}{d}} - e}\right) - 1\right),$$

from which the conclusion follows. □

In conclusion, for constant $d > 1$, it is not possible to have a family of spectral expander graphs of increasing size, where each graph is a d-regular Cayley graph generated by an Abelian group.

Corollary 6.4. *Fix an integer $d > 1$ and let $(G_n)_{n \in \mathbb{N}}$ be a family of d-regular Cayley graphs satisfying $|V(G_n)| \to \infty$ as $n \to \infty$. If each G_n is the Cayley graph of an Abelian group, then $(G_n)_{n \in \mathbb{N}}$ is not a λ-spectral expander family for any $\lambda < d$.*

Proof. This follows directly from the fact that by Proposition 6.3,

$$|d - \lambda(G_n)| \leq \exp\left(\frac{4\log|V(G_n)|}{|V(G_n)|^{1/d} - e}\right) - 1 \to 0,$$

as $n \to \infty$. □

6.2 Linear Characters and Eigenvalues of Cayley Graphs

Despite the negative result of the previous subsection, we shall nonetheless analyse the Eigenvalues of Cayley graphs of Abelian groups. The theory we develop here shall serve in the construction of spectral expander graph families of non-constant degree in the next subsection.

We first introduce the notion of a linear character of a group.

Definition 6.5. [Linear Character] Let Γ be a group. A *linear character* of Γ is a group homomorphism $\chi\colon \Gamma \to \mathbb{C}$ from Γ into the multiplicative group of $\mathbb{C}$.

We shall only be working with finite groups in this exposition. For finite groups, linear characters have a particularly nice form.

Lemma 6.6. *Let Γ be a finite group with $n = |\Gamma|$. For any $g \in \Gamma$ and any linear character χ of Γ, $\chi(g)$ is an nth root of unity.*

Proof. This follows trivially by Lagrange's theorem as $\chi(g)^n = \chi(g^n) = 1$. □

One can consider linear characters of a finite group Γ to be elements of $\ell^2(\Gamma)$. In this light, we get the following lemma.

Lemma 6.7. *Let $\chi_1, \chi_2 \in \ell^2(\Gamma)$ be distinct linear characters of the finite group Γ. Then χ_1 and χ_2 are orthogonal.*

Proof. By assumption there exists $g \in \Gamma$ such that $\chi_1(g) \neq \chi_2(g)$. Now, the left action of multiplication by g permutes Γ yielding

$$\begin{aligned}
\langle \chi_1, \chi_2 \rangle &= \sum_{a \in \Gamma} \chi_1(a)\overline{\chi_2(a)} \\
&= \sum_{a \in \Gamma} \chi_1(g \cdot a)\overline{\chi_2(g \cdot a)} \\
&= \chi_1(g)\overline{\chi_2(g)} \cdot \sum_{a \in \Gamma} \chi_1(a)\overline{\chi_2(a)} \\
&= \chi_1(g)\overline{\chi_2(g)} \cdot \langle \chi_1, \chi_2 \rangle .
\end{aligned}$$

Since $\chi_1(g)$ and $\chi_2(g)$ are distinct roots of unity, $\chi_1(g)\overline{\chi_2(g)} \neq 1$. Hence, $\langle \chi_1, \chi_2 \rangle = 0$. □

Proposition 6.8. *The number of distinct linear characters of a finite Abelian group, Γ, is $|\Gamma|$.*

Proof. First, Γ can have at most $|\Gamma|$ distinct linear characters since they are non-zero orthogonal vectors of $\ell^2(\Gamma)$, a vector space of dimension $|\Gamma|$.

Second, note that the group $(\mathbb{Z}_n, +)$ has n distinct linear characters, $\chi_1, \ldots, \chi_n$, given by $\chi_m(k) = e^{2\pi i \cdot mk/n}$ for $m \in [n]$. Furthermore, suppose that the groups Γ_1 and Γ_2 have n_1 and n_2 distinct linear characters $\chi_1, \ldots, \chi_{n_1}$ and $\eta_1, \ldots, \eta_{n_2}$, respectively. Then $\Gamma_1 \times \Gamma_2$ have the $n_1 n_2$ distinct linear characters, $\{\kappa_{i,j}\}_{i \in [n_1], j \in [n_2]}$, given by $\kappa_{i,j}((g,h)) = \chi_i(g)\eta_j(h)$. It is easily verified that these are indeed distinct well-defined linear characters. The conclusion now follows by the fundamental theorem of finite Abelian groups by which Γ can be written as a direct product $\Gamma = \mathbb{Z}_{n_1} \times \cdots \times \mathbb{Z}_{n_k}$ of cyclic groups with $\prod_{i=1}^{k} n_i = |\Gamma|$. □

The eigenvalues of the adjacency matrix of a Cayley graph of an Abelian group are very accurately described by the linear characters of the group. This is the content of the following theorem.

Theorem 6.9. *Let Γ be an Abelian group with $|\Gamma| = n$ and let $\chi_1, \ldots, \chi_n$ be the linear characters of Γ. For any symmetric subset $S \subset \Gamma$, the eigenvalues of the adjacency matrix of* $\mathrm{Cay}\,(\Gamma; S)$ *are given by $\lambda_i = \sum_{a \in S} \chi_i(a), i \in [n]$.*

Proof. Let A be the adjacency matrix of $\mathrm{Cay}\,(\Gamma; S)$. Each χ_i is an eigenvector of A with corresponding eigenvalue $\sum_{a \in S} \chi_i(a)$. To see this, observe that for every $g \in \Gamma$,

$$(A\chi_i)_g = \sum_{h \in g \cdot S} \chi_i(h) = \sum_{h \in S} \chi_i(g \cdot h) = \chi_i(g) \sum_{h \in S} \chi_i(h).$$

It follows that indeed $A\chi_i = \left(\sum_{h \in S} \chi_i(h)\right) \cdot \chi_i$. This proves the theorem as the linear characters of Γ form a basis for $\ell^2(\Gamma)$ by Lemma 6.7 and Proposition 6.8. □

6.3 Non-Constant Degree Spectral Expanders

We are now ready for a construction of spectral expanders from Cayley graphs. We shall jump right into it. Denote by $\mathbb{F}_p$ the finite field of order p.

Theorem 6.10. *Let p be a prime, let $t > 2$ be an integer, let Γ be the group $\Gamma = (\mathbb{F}_p^t, +)$, and let $S = \{(a, ab, \ldots, ab^{t-1}) \in \mathbb{F}_p^t \mid a, b \in \mathbb{F}_p\}$. Then* $\mathrm{Cay}\,(\Gamma; S)$ *is a p^2-regular graph with p^t vertices and spectral expansion $(t-1)p$.*

Proof. The following facts are immediate: S is a symmetric set with $|S| = p^2$ and $|\Gamma| = p^t$. So $\mathrm{Cay}\,(\Gamma; S)$ is indeed a p^2-regular graph on p^t vertices.

Let A be the adjacency matrix of $\mathrm{Cay}\,(\Gamma; S)$. Each linear characters, χ, of $(\mathbb{F}_p^t, +)$ is of the form $\chi(x_1, \ldots, x_t) = e^{\frac{2i\pi}{p} \sum_{k=1}^t x_k c_k}$ for some $c_1, \ldots, c_t \in [p]$. To see this, one verifies that each such function is indeed a linear character and that we have identified all p^t linear characters since each choice of $c_1, \ldots, c_t \in [p]$ yields a different linear character. By Theorem 6.9, the eigenvalue of A corresponding to χ is then

$$\lambda = \sum_{b \in [p]} \sum_{a \in [p]} e^{\frac{2i\pi}{p} \sum_{k=1}^t c_k a b^{k-1}} = \sum_{b \in [p]} \sum_{a \in [p]} \left(e^{\frac{2i\pi}{p} \sum_{k=1}^t c_k b^{k-1}}\right)^a = \sum_{b \in [p]} \left(p \cdot \left[\sum_{k=1}^t c_k b^{k-1} = 0\right]\right),$$

where the last equality follows since for every pth root of unity, ξ, the sum $\sum_{a \in [p]} \xi^a$ is p if $\xi = 1$ and 0 otherwise. If some c_k is non-zero then $|\lambda| \leq (t-1)p$ since the polynomial $P(b) = \sum_{k=1}^t c_k b^{k-1}$ is of degree $< t$ and thus has at most $t-1$ roots. The only linear character of $\mathbb{F}_p^t$ not satisfying this is the trivial character, $\chi = 1$. Thus, every eigenvalue of A except one has absolute value $\leq (t-1)p$. The conclusion follows. □

Setting $t = 3$ and $d = p^2$, we get a d-regular graph G on $n = p^3$ vertices with spectral expansion $(t-1)p = 2\sqrt{d}$. This is pretty close to being a Ramanujan graph! One problem, though, is that the graph is not very sparse, something that is required for most applications of expanders. For instance, the member search and average estimation applications of Section 3.3, which are very typical of computer science applications of expander graphs, crucially rely on spectral expanders of

constant or at least sub-polynomial degree to reduce the number of random bits used. In our case, the degree of the graph is $d = n^{2/3}$, which is very much not sub-polynomial. For some applications, however, such graphs can still be useful. For instance, a recent paper co-authored by the author (see [10]) crucially relies on graphs of high degree with strong spectral expansion. In such a case, the above graph is perfect, since it is very easy to compute the neighbours of a vertex and to decide whether a given pair of vertices are neighbours.

7 Constant Degree Expanders from the Zig-Zag Product

The zig-zag product construction of constant degree spectral expanders is another gem of expander theory. For all but the most theoretical applications in computer science, expander graphs are required to have an explicit, efficient description to be of use. By explicit and efficient, we mean that given a vertex of the graph (represented in binary), we may find the neighbours of the vertex in time polynomial in the size of the representation of the vertex. Before the paper of Reingold, Vadhan, and Wigderson [11] that this section is based on, there were no combinatorial constructions of spectral expander graphs satisfying this.

The construction of Reingold et al. start from an expander graph and recursively extend it by taking a graph product called the zig-zag product. We shall first describe this product and its properties. Then we describe the recursive construction of constant degree spectral expander families.

7.1 The Zig-Zag Product

Essential to the zig-zag product is the idea of an enumeration or labelling of the edges adjacent to a vertex of a graph. Let $G = (V, E)$ be a d-regular graph. We may fix an *enumeration of the incident edges* of the vertices of G as follows. For each vertex $v \in V$, enumerate the edges incident to it by $\{e_v^1, \ldots, e_v^d\} \subset E$. Note that for any neighbour $u \in V$ of v there are indices $k, \ell \in [d]$ with $e_v^k = e_u^\ell$.

In the following we shall be comparing the spectral expansion of regular graphs of different degrees. Thus, the following "normalised" way of parameterising a graph will come in handy.

Definition 7.1. Leg $G = (V, E)$ be a d-regular graph on $n = |V|$ vertices. We say that G is an (n, d, γ)-spectral expander graph whenever $\gamma \geq \lambda(G)/d$.

With this definition in place, we are ready to define the zig-zag product of two graphs.

Definition 7.2. [Zig-Zag Product] Let $G = (V_1, E_1)$ be a d_1-regular graph on n vertices and let $H = ([d_1], E_2)$ be a d_2-regular graph on d_1 vertices. Fix an enumeration of the edges of G. Then the *zig-zag product* of G and H, $G \textcircled{z} H$, is a d_2^2-regular graph on $n \cdot d_1$ vertices with vertex set $V_1 \times [d_2]$ and edge set given as follows. The vertices $(u, k), (v, \ell) \in V_1 \times [d_1]$ are connected by an edge if and only if there exists $i, j \in [d_1]$ such that

1. The vertices k and i are adjacent in H, i.e., $(k, i) \in E_2$.
2. Edge i incident with $u \in V_1$ is identical to edge j incident with $v \in V_1$, i.e., $e_u^i = e_v^j$.
3. The vertices ℓ and j are adjacent in H, i.e., $(\ell, j) \in E_2$.

Remark 7.3. We shall call each vertex set of the form $\{u\} \times [d_1]$ for $u \in V_1$ a *cloud* of the zig-zag product. In this manner, we consider $G \textcircled{z} H$ to consist of the graph G (the *big graph*)with a copy of H sitting at each vertex. With this terminology, each edge of $G \textcircled{z} H$ is found by starting at a vertex (u, k) of a cloud, taking a step in the cloud to (u, i), jumping across clouds in the big graph to (v, j), and finally taking a step in the destination cloud, ending up at (v, ℓ).

Remark 7.4. The adjacency matrix of $G \textcircled{z} H$, which we denote by $Z \in \mathbb{R}^{(V_1 \times [d_1]) \times (V_1 \times [d_1])}$ may be expressed as follows. Let $P \in \mathbb{R}^{(V_1 \times [d_1]) \times (V_1 \times [d_1])}$ be the permutation matrix satisfying that $P_{(u,k),(v,\ell)} = 1$ if and only if $e_u^k = e_v^\ell$, and let B be the adjacency matrix of H. Denote by $\tilde{B}$ the matrix $I_n \otimes B \in \mathbb{R}^{(V_1 \times [d_1]) \times (V_1 \times [d_1])}$, the tensor product of the $n \times n$ identity matrix and B. The matrix P corresponds to a step on G between clouds in the big graph, while $\tilde{B}$ corresponds to taking a step on H within a cloud. Having defined this, we may express Z as $Z = \tilde{B} P \tilde{B}$.

Furthermore, we define the square of a graph.

Definition 7.5. [Square of Graph] Let G be a d-regular graph on n vertices. If A is the adjacency matrix of G, then the *square of* G is the d^2-regular graph G^2 on n vertices with adjacency matrix A^2.

We shall construct expanders using these two graph operations. To that end, we prove the following two statements.

Proposition 7.6. *If G is an (n, d, γ)-spectral expander, then G^2 is an (n, d^2, γ^2)-spectral expander.*

Proof. Suppose that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ are the eigenvalues of the adjacency matrix A of G. Then $\lambda_1^2, \ldots, \lambda_n^2$ are the eigenvalues of A^2 since A is real symmetric. It follows that $\lambda(G^2) = \lambda(G)^2$ and since G^2 is d^2-regular, $\gamma^2 \geq \lambda(G)^2/d^2 = \lambda(G^2)/d^2$. □

Theorem 7.7. *If $G = (V, E_1)$ is an (n, d_1, γ_1)-spectral expander and $H = ([d_1], E_2)$ is an (d_1, d_2, γ_2)-spectral expander. Then $G \textcircled{z} H$ is a $(nd_1, d_2^2, \gamma_1 + \gamma_2 + \gamma_2^2)$-spectral expander.*

Proof. Initially, we note that $G \textcircled{z} H$ is indeed a d_2^2-regular graph on nd_1 vertices. Thus, all that remains is the claim regarding the spectral expansion.

Let $Z = \tilde{B} P \tilde{B}$ be the adjacency matrix of $G \textcircled{z} H$ as described in Remark 7.4. Let $f \in \ell^2(V \times [d_1])$ be a vector of length $||f||_2 = 1$ orthogonal to $1_{V \times [d_1]}$. We may decompose f as $f = f^{||} + f^{\perp}$, where $f^{||}$ is given by $f^{||}(u, i) = \frac{1}{d_1} \sum_{k \in [d_1]} f(u, k)$. In other words, $f^{||}$ is constant on each cloud. Note that now, $f^{||} \perp 1_{V \times [d_1]}$ since $\sum_{u \in V, k \in [d_1]} f^{||}(u, k) = \sum_{u \in V, k \in [d_1]} f(u, k) = 0$. It follows that also $f^{\perp} \perp 1_{V \times [d_1]}$.

Towards applying the min-max theorem (Lemma 2.9), we bound the inner product $\langle Zf, f \rangle$ by

$$|\langle Zf, f \rangle| \leq \underbrace{\left| \left\langle \tilde{B} P \tilde{B} f^{||}, f^{||} \right\rangle \right|}_{T_1} + 2 \underbrace{\left| \left\langle \tilde{B} P \tilde{B} f^{||}, f^{\perp} \right\rangle \right|}_{T_2} + \underbrace{\left| \left\langle \tilde{B} P \tilde{B} f^{\perp}, f^{\perp} \right\rangle \right|}_{T_3}.$$

We bound each term individually.

First, since $\tilde{B} f^{||} = d_2 f^{||}$ and because $\tilde{B}$ is symmetric,

$$T_1 = \left| \left\langle P \tilde{B} f^{||}, \tilde{B} f^{||} \right\rangle \right| = d_2^2 \left| \left\langle P f^{||}, f^{||} \right\rangle \right|.$$

Let $g \in \ell^2(V)$ be given by $g(v) = f^{||}(v,i)$ for $v \in V$ and $i \in [d_1]$ arbitrary. This definition is independent of the choice of i because $f^{||}$ is constant on each cloud. Furthermore, we must have $g \perp 1_V$ as $f^{||} \perp 1_{V\times[d_1]}$. Denoting by A the adjacency matrix of G, it holds that $\langle Ag, g\rangle = \left\langle Pf^{||}, f^{||}\right\rangle$. Thus, $T_1 \leq d_2^2(\gamma_1 d_1)\left|\left|g\right|\right|_2^2 = \gamma_1 d_2^2 \left|\left|f^{||}\right|\right|_2^2$.

Second, we may write $T_2 = d_2 \left|\left\langle Pf^{||}, \tilde{B}f^{\perp}\right\rangle\right|$. Then, as P is a permutation,

$$T_2 \leq d_2 \left|\left|Pf^{||}\right|\right|_2 \cdot \left|\left|\tilde{B}f^{\perp}\right|\right|_2 = d_2 \left|\left|f^{||}\right|\right|_2 \cdot \left|\left|\tilde{B}f^{\perp}\right|\right|_2.$$

Now, as on each cloud, $f^{\perp}\restriction_{\{u\}\times[d_1]}$ is orthogonal to $1_{\{u\}\times[d_1]}$ and $\tilde{B} = I_n \otimes B$, we may consider B to be an operator on $\ell^2(\{u\}\times[d_1])$ for each $u \in V$ so that

$$\left|\left|\tilde{B}f^{\perp}\right|\right|_2 = \sqrt{\sum_{u\in V}\left|\left|Bf^{\perp}\restriction_{\{u\}\times[d_1]}\right|\right|_2^2} \leq \sqrt{\sum_{u\in V} d_2^2\gamma_2^2\left|\left|f^{\perp}\restriction_{\{u\}\times[d_1]}\right|\right|_2^2} = d_2\gamma_2\left|\left|f^{\perp}\right|\right|_2.$$

Thus, $T_2 \leq d_2^2\gamma_2 \left|\left|f^{||}\right|\right|_2 \cdot \left|\left|f^{\perp}\right|\right|_2$.

Third, we find that $T_3 \leq \left|\left|\tilde{B}f^{\perp}\right|\right|_2^2 \leq d_2^2\gamma_2^2\left|\left|f^{\perp}\right|\right|_2^2$.

Combining these three bounds and noting that $\left|\left|f^{\perp}\right|\right|_2^2 + \left|\left|f^{||}\right|\right|_2^2 = 1$, we conclude that

$$|\langle Zf, f\rangle| \leq d_2^2\left(\gamma_1\left|\left|f^{||}\right|\right|_2^2 + 2\gamma_2\left|\left|f^{||}\right|\right|_2 \cdot \left|\left|f^{\perp}\right|\right|_2 + \gamma_2^2\left|\left|f^{\perp}\right|\right|_2^2\right) \leq d_2^2(\gamma_1+\gamma_2+\gamma_2^2). \tag{25}$$

Thus, $\lambda(G \textcircled{z} H) \leq d_2^2(\gamma_1+\gamma_2+\gamma_2^2)$, from which the conclusion follows. □

7.2 Expander Construction

Having established the above properties, we proceed to construct constant degree expanders. The construction requires an instantiation using a (d^4, d, γ)-spectral expander for some $d > 1$ and $\gamma \leq 1/8$. Such an expander may either be found using an exhaustive search on d-regular graphs on d^4 vertices or using the construction of Section 6.3.

Theorem 7.8. *Let H be a (d^4, d, γ)-spectral expander for some $d > 1$ and $\gamma \leq 1/8$. Recursively define a series of graphs $G_1 = H^2$ and for $n \geq 1$,*

$$G_{n+1} = (G_n)^2 \textcircled{z} H. \tag{26}$$

Then for every $n \in \mathbb{N}$, G_n is a $(d^{4n}, d^2, \gamma + 3\gamma^2)$-spectral.expander

Proof. We proceed by induction on n. The claim is clearly true for $n = 1$. Now, suppose it is true for some $n \in \mathbb{N}$. Then $(G_n)^2$ is a $(d^{4n}, d^4, (\gamma+3\gamma^2)^2)$-spectral expander by Proposition 7.6. It follows now by Theorem 7.7 that G_{n+1} is a $(d^{4n}, d^2, (\gamma+3\gamma^2)^2+\gamma+\gamma^2)$-spectral expander. Since $\gamma \leq 1/8$,

$$(\gamma+3\gamma^2)^2+\gamma+\gamma^2 = \gamma+2\gamma^2+6\gamma^3+9\gamma^4 \leq \gamma+3\gamma^2, \tag{27}$$

and the conclusion follows. □

It is worth noting that suppose we enumerate the edges of the d-regular graph $H = (V, E)$ and there is an efficient algorithm to determine given $v \in V$ and $i \in [d]$, $u \in V$ and $j \in [d]$ with $u \neq v$ such that $e_u^j = e_v^i$. In other words, an efficient algorithm to walk on the graph and know the enumeration of the edges. It is then possible to construct an algorithm to efficiently walk on $G_1 = H^2$ and $G_2 = G_1^2 ⓩ H$. Recursively, one may construct such an algorithm for any G_n using $\theta(rn)$ steps, where r is the number of steps the original algorithm would take to find u and j. Thus, the zig-zag product construction satisfies the initially mention requirement of being efficiently computable. The enumeration is fairly straightforward when using for instance the expander of Section 6.3.

References

[1] Bilu, Y., and Linial, N. Lifts, discrepancy and nearly optimal spectral gap*. *Combinatorica 26*, 5 (Oct 2006), 495–519.

[2] Broder, A., and Shamir, E. On the second eigenvalue of random regular graphs. In *28th Annual Symposium on Foundations of Computer Science (sfcs 1987)* (Oct 1987), pp. 286–294.

[3] Davidoff, G., Sarnak, P., and Valette, A. *Elementary Number Theory, Group Theory and Ramanujan Graphs*. London Mathematical Society Student Texts. Cambridge University Press, 2003.

[4] Friedman, J. On the second eigenvalue and random walks in randomd-regular graphs. *Combinatorica 11* (1991), 331–362.

[5] Friedman, J. A proof of Alon's second eigenvalue conjecture. In *Proceedings of the Thirty-fifth Annual ACM Symposium on Theory of Computing* (New York, NY, USA, 2003), STOC '03, ACM, pp. 720–724.

[6] Hoory, S., Linial, N., and Wigderson, A. Expander graphs and their applications. *Bull. Amer. Math. Soc. 43* (2006), 439–561.

[7] Krebs, M., and Shaheen, A. *Expander Families and Cayley Graphs a beginner's guide*. Oxford University Press, 2011.

[8] Marcus, A., Spielman, D. A., and Srivastava, N. Interlacing families i: Bipartite ramanujan graphs of all degrees. In *Proceedings of the 2013 IEEE 54th Annual Symposium on Foundations of Computer Science* (Washington, DC, USA, 2013), FOCS '13, IEEE Computer Society, pp. 529–537.

[9] Marcus, A. W., Spielman, D. A., and Srivastava, N. Interlacing families iv: Bipartite ramanujan graphs of all sizes. In *2015 IEEE 56th Annual Symposium on Foundations of Computer Science* (Oct 2015), pp. 1358–1377.

[10] Rasmussen, P. M. R., and Sahai, A. Expander graphs are non-malleable codes, 2018.

[11] Reingold, O., Vadhan, S., and Wigderson, A. Entropy waves, the zig-zag graph product, and new constant-degree expanders and extractors. In *Proceedings of the 41st Annual Symposium on Foundations of Computer Science* (Washington, DC, USA, 2000), FOCS '00, IEEE Computer Society, pp. 3–.

[12] Trevisan, L. Lecture notes on expansion, sparsest cut, and spectral graph theory, August 2014.

[13] Vadhan, S. P. Pseudorandomness. *Foundations and Trends in Theoretical Computer Science 7*, 1–3 (2012), 1–336.