Title

Pseudospectra of matrices and Point spectra of infinite graphs

Author

Mukherjee, Satyaki

Publication Date

2021

Peer reviewed|Thesis/dissertation

Pseudospectra of matrices and Pointspectra of infinite graphs

by

Satyaki Mukherjee

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Mathematics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor James W. Pitman, Co-chair
Associate Professor Nikhil Srivastava, Co-chair
Professor Satish Rao

Spring 2021

Pseudospectra of matrices and Pointspectra of infinite graphs

Abstract

Pseudospectra of matrices and Pointspectra of infinite graphs

by

Satyaki Mukherjee

Doctor of Philosophy in Mathematics

University of California, Berkeley

Professor James W. Pitman, Co-chair

Associate Professor Nikhil Srivastava, Co-chair

In this thesis we look at various tools to analyse eigenvalues and eigenvectors and use them to prove the following main results.

1. We show that given any matrix A, there is a small perturbation of the matrix such that post perturbation, the matrix is almost normal. In particular there exists E, with $||E|| \leq \delta||A||$ such that $A + E$ is diagonalizable and its eigenvector matrix has polynomially (in $1/\delta$ and n) bounded condition number.
2. We prove a necessary and sufficient condition for any local periodic operator on the universal cover of a finite graph to have a point spectrum. In particular we show that for λ to be in the point spectrum, the base graph must admit an induced forest with a very specific combinatorial structure and that the induced operator on it must also have λ as an eigenvalue.

To prove the first result we study the volume of the pseudospectrum with the help of some tools from stochastic calculus. Along the way we also see why it implies a conjecture by Sankar, Spielman and Teng on the optimal constant for smoothed analysis of condition numbers.

For the second result we show that a condition conjectured by Aomoto to be necessary and sufficient for the existence of point spectrum of certain operators on periodic trees is indeed so. Aomoto had already shown why the condition was necessary. We give a more intuitive proof of it and along the way also show sufficiency.

Kichu kichu bostu ache suru tei sesh
Kichu kichu bostu ache seshe jar suru
Aar kichu kichu bostu ache surur sesher baire jemon..

Contents

Contents ii

List of Figures iv

1 A brief overview 1
- 1.1 Davies' Conjecture . . . 1
- 1.2 Point spectrum of the universal cover . . . 3
- 1.3 Other short stories . . . 5
- 1.4 Bibliographic note . . . 7

2 Davies' Conjecture 8
- 2.1 Introduction . . . 8
- 2.2 Tools from Random Matrix Theory . . . 13
- 2.3 Proof of Theorems 2.1.1 and 2.1.5 . . . 15
- 2.4 Optimality of the Bounds . . . 19
- 2.5 Conclusion and Discussion . . . 20
- 2.6 Appendix : Proof of Theorem 2.2.4 . . . 21

3 Point Spectrum of the Universal Cover 24
- 3.1 Introduction . . . 24
- 3.2 Preliminaries . . . 26
- 3.3 Main Results . . . 31
- 3.4 Acyclic Nature of Aomoto Sets . . . 35
- 3.5 Aomoto's Index Formula . . . 37
- 3.6 A Generalized Converse to Aomoto's Theorem . . . 42
- 3.7 Spectral Delocalization for $A_{\mathcal{T}}$ . . . 45

4 Other short stories 49
- 4.1 Introduction . . . 49
- 4.2 Yet another class of real rooted polynomials . . . 52
- 4.3 A special case of the existential version of the Non Commutative Khintchine inequality . . . 55

Bibliography **61**

List of Figures

2.1 T is a sample of an upper triangular 10×10 Toeplitz matrix with zeros on the diagonal and independent (modulo the Toeplitz structure) standard real Gaussian entries above the diagonal. Pictured is the boundary of the ϵ-pseudospectrum of T (left) and $T + 10^{-6}G$ (right) for $\epsilon = 10^{-5}$, $\epsilon = 10^{-5.5}$, and $\epsilon = 10^{-6}$, along with the spectra. These plots were generated with the MATLAB package EigTool [WT02]. 12

3.1 On the left, a finite graph G, where the vertices in $X_1(G) = X_{-1}(G)$ are colored in red. In the middle and on the right, we show the two distinct spectral measures of $A_{\mathcal{T}}$ associated to the vertices of G, the center one having atoms at ± 1. . . . 31

3.2 Two distinct sets of vertices (in red and blue respectively) of a graph G are shown. If $a \equiv 1$ and $b \equiv 0$, it is easy to show from Corollary 3.3.4 that $I_0(G) = 1$. Then both the red and the blue vertex set belong to $\mathcal{A}_0(G)$. It will follow from Observation 3.4.1 below that $X_0(G)$ is precisely the set indicated by the red vertices. 33

3.3 As in Section 3.3, for each graph above, a combination of Observation 3.4.1 and Theorem 3.3.3 reveals the red vertices as the Aomoto set associated to 0. In both cases $I_0(G) = 1$ while $|V(G)|$ is even. 36

3.4 On the left an example of a graph G with Aomoto trees in red. On the right its auxiliary graph G', where each tree T_i has been contracted into a vertex t_i and the blue edges have been removed. 38

Acknowledgments

Firstly I would like to thank my advisor Nikhil, without whom this phd would have been impossible. Next I would like to thank my friends some but not all of whom and in alphabetical order are Anamika, Milind, Pinaki, Soumendu, Sujayam for putting up with my terrible jokes. In the same order I would like to thank my collaborators Archit, Jess, Jorge, Nick, Tarun who did most of the work, while I derped around. Lastly I would like to thank the various professors that I have met and had interesting discussions with in ISI and UC Berkeley which include Anirudh N., Anita N., Jaydeb S., James W.P., Luca T., Siva A. to name a few.

Chapter 1

A brief overview

1.1 Davies' Conjecture

Matrices are used in representing all sorts of objects in mathematics. Among the various classes of matrices, one important subclass is that of the normal matrices. The principle reason for this is the following fundamental theorem.

Theorem 1.1.1. *Given any $n \times n$ normal matrix M, there exists an orthonormal basis $v_1, ..., v_n$ such that*

$$M = \sum_{i=1}^{n} d_i v_i v_i^* = VDV^*,$$

where V is the unitary matrix whose columns are the eigenvectors v_i, and D is a diagonal matrix with entries as the eigenvalues d_i.

This property is useful for a variety of numerical and algebraic reasons. For instance this allows for fast computations of various matrix functions. Lets quickly look at a simple numerical example.

Example 1.1.2. *Given a diagonalizable matrix M compute M^n.*

The text-book answer to this is to write $M = VDV^{-1}$. Then

$$M^n = VD^nV^{-1}.$$

In practice however one issue crops up. What if V is almost singular? Then V^{-1} would contain very large entries, and would be numerically infeasible and unstable to compute.

Now imagine that the matrix were to be normal. Then we could write $M = VDV^*$, and $M^n = VD^nV^*$. Here we would only need to find V, i.e. the eigenvectors, there would be no issue with computing V^*

Unfortunately however a lot of matrices that one encounters are non-normal. Thus a natural question to ask is given a matrix M is there always a normal matrix somewhat close to it? Unfortunately no. However as Mathematicians often do we rearrange the words to end up in a meaningful statement. Given a matrix M there is a *somewhat* normal matrix close to it. To make things rigorous we now need to make sense of the idea of *somewhat* normal.

As the set of matrices with all distinct eigenvalues is dense, we can simply restrict our definitions to such matrices. As noted the principle property of normality we are interested in is the fact that the matrix of its eigenvectors form an orthonormal matrix. By stipulating that any eigenvector we choose must have norm 1, and by restricting our matrices to those with distinct eigenvalues, we now get a unique (upto permutation) matrix V of eigenvectors associated to an arbitrary given matrix M. The question of *how normal* M is thus has now been outsourced to the question of *how orthonormal* V is.

To build such a yardstick, we remind ourselves that each orthonormal matrix[1] defines an isometry i.e. all of its singular values are 1. Thus one measure of how far a matrix is from being an isometry would be to look at the range of its singular values. A very large singular value or a very small singular value would suggest that the matrix is highly non-isometric. This idea is rigorised by the quantity called condition number :

$$\kappa(V) = \frac{\sigma_{max}}{\sigma_{min}} = ||V||_2||V^{-1}||_2,$$

where σ_{max} and σ_{min} are the largest and smallest singular values of V respectively.

Armed with such notions E.B.Davies proved in 2007 the following theorem

Theorem 1.1.3. *Suppose $A \in \mathbb{C}^{n\times n}$ and $\delta \in (0,1)$. Then there exists an $E \in \mathbb{C}^{n\times n}$ such that $||E|| \leq \delta||A||$ and*

$$\kappa_V(A+E) \leq \left(\frac{n}{\delta}\right)^{(n-1)/2}.$$

Here $\kappa_V(M)$ is the condition number of the eigenvector matrix of M.

This bound however is exponential. Thus a natural follow up question, one that Davies posited is whether there is an E such that the eigenvector condition number of $A+E$ is bounded polynomially in δ

In Chapter 2 we give a positive answer to such a question in the form of the following theorem.

Theorem 1.1.4. *Suppose $A \in \mathbb{C}^{n\times n}$ and $\delta \in (0,1)$. Then there exists an $E \in \mathbb{C}^{n\times n}$ such that $||E|| \leq \delta||A||$ and*

$$\kappa_V(A+E) \leq 4n^{3/2}\left(1+\frac{1}{\delta}\right).$$

Here $\kappa_V(M)$ is the condition number of the eigenvector matrix of M.

[1] also known as orthogonal matrix

We also show in proposition 2.4.1 that this inequality is tight in δ. To do this we study the volume of the $\epsilon-$*pseudospectrum* of a matrix M, defined for $\epsilon > 0$ as:

$$\Lambda_\epsilon(M) \triangleq \{z \in \mathbb{C} : z \in \Lambda(M+E) \text{ for some } \|E\| < \epsilon\}.$$

We use results from the theory of *epsilon*$-$ pseudospectrum to turn our original question into one about the tail distribution of singular values of a matrix with a complex Ginibre perturbation. Lastly to complete the proof we use the following theorem of Śniady[Śni02]. It allows us to compare the singular values of two matrices performing Brownian motion as long as they are coupled in a very particular way.

Theorem 1.1.5 (Śniady). *Let A_1 and A_2 be $n \times n$ complex matrices such that $\sigma_i(A_1) \leq \sigma_i(A_2)$ for all $1 \leq i \leq n$. Assume further that $\sigma_i(A_1) \neq \sigma_j(A_1)$ and $\sigma_i(A_2) \neq \sigma_j(A_2)$ for all $i \neq j$. Then for every $t \geq 0$, there exists a joint distribution on pairs of $n \times n$ complex matrices (G_1, G_2) such that*

(i) the marginals G_1 and G_2 are distributed as (normalized) complex Ginibre matrices G_n, and

(ii) almost surely $\sigma_i(A_1 + \sqrt{t}G_1) \leq \sigma_i(A_2 + \sqrt{t}G_2)$ for every i.

This also allows us to prove the real version of a conjecture by Sankar, Spielman and Teng stated as Conjecture 1 in [SST06]. In particular we prove :

Proposition 1.1.6. *Let G be an $n \times n$ matrix with i.i.d. real $N(0,1)$ entries, and A be any $n \times n$ matrix with real entries. Then*

$$\mathcal{P}[\sigma_n(A+G) < \epsilon] \leq \epsilon\sqrt{n}.$$

1.2 Point spectrum of the universal cover

Given a graph G with vertex set $V(G)$ and edge set $E(G)$, its adjacency matrix A is a $|V(G)| \times |V(G)|$ matrix whose rows and columns are indexed by $V(G)$ such that $A_{u,v} = 1$ iff $(u,v) \in E(G)$ and 0 otherwise.

A graph H is a lift of a graph G if there exists a covering map from H to G under the standard graph topology. Using standard theory from topology, one can consider an object called the universal cover, $\mathcal{T}_G$. This is an infinite graph such that whenever any H is a lift of G, $\mathcal{T}_G$ is a lift of H.

It might be worthwhile for the discussion of lifts on graphs to look at a definition with less jargon. For simplicity let us restrict our attention to simple graphs only. In this case the idea is that H is a lift of G if there exists a surjective map $\phi : H \to G$ such that for any vertex u, ϕ is a bijection from the neighbours of u to those of $\phi(u)$.

To give a simple definition of the universal cover we need to first define a non-backtracking walk. A non-backtracking walk $v_0, v_1, ..., v_k$ on a (simple)graph G is a walk such that for any i, $v_i \neq v_{i+2}$. The universal cover of G is then the infinite tree T_G constructed in the following fashion:

Fix any vertex v_0 in G. Then vertex set of $\mathcal{T}_G$ is the collection of all non-backtracking walks in G starting from v_0. We then draw an edge between two walks, p_1 and p_2 if p_2 is p_1 followed by another vertex.

Let us look at an example. Consider a d-regular graph G. Then its universal cover is the infinite d-ary tree. But why is this useful? Principally the idea is that many asymptotic properties of a graph come solely from its local structure. Thus such properties can often be derived by studying the universal cover. The universal cover by its very definition after all is the "biggest" graph with a given local structure. Consider the following famous result commonly called as the Alon-Boppana bound[Nil91] :

Theorem 1.2.1. *Let $G_1, G_2, ...$ be a sequence of connected d-regular simple graphs with number of vertices v_n going to infinity. Let $\lambda_{2,n}$ be the second largest eigenvalue in modulus of the Adjacency operator of G_n. Then* $\lim_{n\to\infty} |\lambda_{2,n}| \geq 2\sqrt{d-1}$

A standard proof sketch of this goes as follows :

Proof. Since every walk in the Universal Covering tree, $\mathcal{T}$ gives a walk in any base graph G_n (via the covering map), we note that the number of closed walks in G_n is atleast as many as those in the d-ary tree T. Let A_n be the adjacency matrix of G_n. Given any walk on the T write a sequence of $+1$ and -1, where you write $+1$ if in the i'th step of the walk you move away from the starting vertex and -1 otherwise. Since all the partial sums of this sequence are non-negative and there are always atleast $d-1$ ways of moving away from the starting vertex, it is not hard to see that the number of closed walks of length $2k$ on the tree starting from a given vertex is atleast $C_k(d-1)^k$, where C_k is the k'th Catalan number. Thus we have the following inequality for any k and n:

$$\begin{aligned} d^{2k} + (v_n - 1)|\lambda_{2,n}|^{2k} &\geq \operatorname{tr}(A_n^{2k}) \geq v_n C_k (d-1)^k \\ \implies |\lambda_{2,n}| &\geq \sqrt[2k]{\frac{v_n}{v_n - 1} C_k (d-1)^k - \frac{d^{2k}}{v_n - 1}} \\ \implies \lim_{n\to\infty} |\lambda_{2,n}| &\geq \sqrt[2k]{C_k (d-1)^k}. \end{aligned}$$

Here the last line follows by using the assumption that the number of vertices v_n goes to ∞. Finally we send k to ∞ and use that in limit $\sqrt[2k]{C_k}$ goes to 2 to complete the proof.

□

In this proof we already see why studying the universal cover might shed light on properties of the base graph. But there is an alternate proof that couples the two even better. To do so let us first define the spectrum of an infinite operator. Given a infinite operator X, define its spectrum as the set $\operatorname{Spec}(X)$ such that

$$\lambda \in \operatorname{Spec}(X) \iff X - \lambda I \textit{ is not invertible.}$$

It is well known that for the infinite d-ary tree, the spectrum of the adjacency matrix of the universal cover is $[-2\sqrt{d-1}, 2\sqrt{d-1}]$ ([Kes59], [McK81]). The spectral radius i.e. the maximum absolute value in the the spectrum of the regular d-ary tree is thus $2\sqrt{d-1}$. We can now couple this fact with a theorem by Greenberg :

Theorem 1.2.2. *Let $\mathcal{T}$ be an infinite connected graph with finite maximum degree and spectral radius $\rho(\mathcal{T})$. Then given any $\epsilon > 0$, there exists a constant $c = c(\mathcal{T}, \epsilon)$ such that for any G which is covered by $\mathcal{T}$, atleast $cV(G)$ many eigenvalues of A_G are greater in modulus than $\rho(\mathcal{T}) - \epsilon$.*

The full statement and a proof of this can be found in this paper by S.M.Cioabă [Cio06]. With this discussion we see how it might be of considerable interest to study infinite trees or universal covers in general and the spectra of their adjacency operators in particular. Unfortunately not much is known about this spectrum when the graph, and hence its cover, is irregular.

In 1991 Aomoto came up with a result [Aom91] which sheds some light on this mysterious animal. To discuss the result let us first define the point spectrum of an infinite operator. We say λ is in the point spectrum of an infinite operator X, if there is a vector v, with finite L_2 norm, such that $Xv = \lambda v$. Aomoto derived a set of necessary conditions for the presence of point spectrum of $A_{\mathcal{T}}$ and used it to show that the d-ary tree has no point spectrum. Given any λ in the point spectrum and the base graph G, Aomoto shows the existence of a particular subset of the set of vertices of G whose induced graph is acyclic. We call this the Aomoto set $X_\lambda(G)$. Intuitively this set is sort of the support of the eigenvector corresponding to λ. It was also conjectured that these acyclic structures should not only be necessary but also sufficient for the existence of point spectrum. We show precisely that in Theorem 3.3.3. We also give a simplified and more combinatorial proof of Aomoto's original result in Theorem 3.3.1. The principle idea here that we exploit is a relation between the number of connected components of the set $X_\lambda(G)$ and the number of vertices outside of it which are still neighbours to atleast one vertex in X_λ. Essentially the former being greater than the latter allows us to construct an eigenvector of $A_{\mathcal{T}}$ with eigenvalue λ.

This in addition to completely classifying the point spectrum of an irregular tree also gives us a finite time algorithm to determine whether any given λ is in the point spectrum of adjacency operator. Also of interest is the fact that our results are true not just for the adjacency operator but for the more general class of "local periodic" operators, which we define in chapter 3 when discussing our results.

1.3 Other short stories

We end by discussing a grab bag of various small problems involving covers. We saw in the previous section that an "increasing" sequence of d regular graphs has the property that their second largest eigenvalue in modulus is asymptotically greater than or equal to $2\sqrt{d-1}$. This poses a natural question. Can we get a sequence of d-regular graphs whose second largest eigenvalue is never greater than $2\sqrt{d-1}$ and thus in limit is $2\sqrt{d-1}$? It was shown by Marcus, Spielman and Srivastava in [MSS14] that you can indeed get such a sequence of bipartite graphs. In particular they proved the following theorem :

Theorem 1.3.1. *A d-regular graph G is called Ramanujan if all the eigenvalues of its adjacency matrix except one lies in the interval $[-2\sqrt{d-1}, 2\sqrt{d-1}]$. Then if $d > 2$ there exists a sequence of connected ramanujan bipartite graphs with increasing vertex size.*

One interesting way to show this is via [MSS15]. Here they show that for any d-regular graph G (with $d > 2$), there exists a connected 2-lift H i.e. H is a lift of G with exactly twice the number of vertices such that if G is ramanujan then so is H. In reality if the adjacency matrix of G and H are A_G and A_H respectively, then it turns out that all the eigenvalues of A_G are also eigenvalues of A_H. So all one would need to show that there is a connected 2-lift whose "new" eigenvalues are in the interval $[-2\sqrt{d-1}, 2\sqrt{d-1}]$.

One might then wonder if there is anything special about 2-lift. And indeed, in [HPS18], it is shown that given a *simple* graph G the above theorem generalizes to arbitrary r-lifts. In particular they show the following theorem :

Theorem 1.3.2. *Every connected, loopless d-regular graph has a one-sided Ramanujan r-covering. If the graph is bipartite then there is a a two-sided Ramanujan r-covering.*

There is one small hiccup however. They need the graph to be loopless. To remedy this they make the following conjecture in their paper which would extend the result to any connected d-regular graph. We see in Chapter 4 why this small result about real rooted-ness is true:

Theorem 1.3.3. *Let $A_1, \ldots, A_n$ be any real matrices, then the following polynomial is real rooted*

$$\mathbb{E}_{P_i,Q_i}\chi(\sum_i P_i A_i Q_i^T + Q_i A_i^T P_i^T)$$

where P_i, Q_i is any distribution of permutations realized by swaps.

The proof of this result largely follows the techniques in [MSS15] with some suitable changes. This gives us a new family of real rooted polynomials.

Surprisingly enough, these techniques of real rooted-ness can sometimes be used to get seemingly unrelated results in functional analysis. Consider the following theorem proved by Françoise Lust-Piquard in 1997 [LP97]

Theorem 1.3.4. *For every matrix $A = (a_{ij})$ such that A and A^* are bounded in $l^\infty(l^2)$ norm, there exists a matrix $B = (b_{ij})$ defining a bounded operator: $l^2(\mathbb{C}) \to l^2(\mathbb{C})$ such that*
(i) $|B|_{2\to 2} \leq K max\{|A|_{l^\infty(l^2)}, |A^*|_{l^\infty(l^2)}\}$
(ii) $\forall i, j \in \mathbb{N}$, $|b_{ij}| \geq |a_{ij}|$,
where K is an absolute constant and $|A|_{l^\infty(l^2)} := \max_j \sqrt{\sum_i a_{ij}^2}$.

For real matrices using the Heilman-Lieb theorem [HL72] not only can we prove this very easily, we also get 2 as K which is a large improvement over the constant listed in previous paper. This is a special case of the following more general conjecture :

Conjecture 1.3.5. *Given any $n \times n$ symmetric matrices $A_1, ..., A_n$, there exists $\epsilon_1, ..., \epsilon_n \in \{1, -1\}$ such that*

$$||\sum_i \epsilon_i A_i|| \leq K\sqrt{||\sum_i A_i^2||},$$

where K is an absolute constant.

The above conjecture in turn is related to the Khintchine-Kahane inequality :

Theorem 1.3.6. *Let $A_1, ..., A_n$ be $d \times d$ symmetric matrices. Let $\epsilon_1, ..., \epsilon_n$ be random variable taking values 1 or -1 with equal probability. Then there exists a constants K and K' such that*

$$K\sqrt{||\sum_i A_i^2||} \leq \mathbb{E}[||\sum_i \epsilon_i A_i||] \leq K'\sqrt{\log d}\sqrt{||\sum_i A_i^2||}.$$

Thus while the Khintchine-Kahane inequality gives information about the expected norm, the conjecture asks if we can lose $\sqrt{\log d}$ by moving from expectation to mere existence.

1.4 Bibliographic note

Many of the results presented in this dissertation have been published or submitted for publication and a lot of the presentation is joint work. Chapter 2 is joint work with Archit Kulkarni, Jess Banks and Nikhil Srivastava and have been submitted for publication. Chapter 3 is joint work with Jess Banks and Jorge Garza Vargas and have been submitted for publication. A part of chapter 4, namely 4.2 is joint work with Nick Ryder and Nikhil Srivastava

Chapter 2

Davies' Conjecture

2.1 Introduction

A matrix $A \in \mathbb{C}^{n \times n}$ is diagonalizable if it can be written as $A = VDV^{-1}$, where D is diagonal and V is a matrix consisting of linearly independent eigenvectors of A. Further, A is normal if and only if $V^{-1} = V^*$, or in other words if the eigenvectors can be chosen to be orthogonal. A fundamental quantity capturing the nonnormality of a matrix is the *eigenvector condition number*

$$\kappa_V(A) \triangleq \inf_{V : A = VDV^{-1}} \|V\|\|V^{-1}\|,$$

which ranges between 1 and ∞ when A is normal and nondiagonalizable respectively, where $\|\cdot\|$ denotes the operator norm. Matrices with small κ_V enjoy many of the desirable properties of normal matrices, such as stability of the spectrum under small perturbations (this is the content of the Bauer-Fike theorem [BF60]). In this paper we study the following question posed by E. B. Davies in [Dav07]:

> *How well can an arbitrary matrix be approximated by one with a small eigenvector condition number?*

Our main theorem is as follows.

Theorem 2.1.1. *Suppose $A \in \mathbb{C}^{n \times n}$ and $\delta \in (0, 1)$. Then there is a matrix $E \in \mathbb{C}^{n \times n}$ such that $\|E\| \leq \delta\|A\|$ and*

$$\kappa_V(A + E) \leq 4n^{3/2}\left(1 + \frac{1}{\delta}\right).$$

In other words, every matrix is at most inverse polynomially close to a matrix whose eigenvectors have condition number at most polynomial in the dimension. The previously best known general bound in such a result was [Dav07, Theorem 3.8]:

$$\kappa_V(A + E) \leq \left(\frac{n}{\delta}\right)^{(n-1)/2}, \tag{2.1}$$

so our theorem constitutes an exponential improvement in the dependence on both δ and n. We show in Proposition 2.4.1 that the $1/\delta$-dependence in Theorem 2.1.1 cannot be improved beyond $1/\delta^{1-1/n}$, so our bound is essentially optimal in δ for large n.

Davies' Conjecture

Theorem 2.1.1 implies a positive resolution to a conjecture of Davies [Dav07].

Conjecture 2.1.2. *For every positive integer n there is a constant c_n such that for every $A \in \mathbb{C}^{n\times n}$ with $\|A\| \le 1$ and $\epsilon \in (0,1)$:*

$$\inf_{E\in\mathbb{C}^{n\times n}} (\kappa_V(A+E)\epsilon + \|E\|) \le c_n\sqrt{\epsilon}. \tag{2.2}$$

Proof of Conjecture 2.1.2. Given $\epsilon > 0$, set $\delta = d_n\sqrt{\epsilon}$ for some $d_n > 0$ and apply Theorem 2.1.1. This yields $c_n = 4n^{3/2} + 4n^{3/2}/d_n + d_n$. This is minimized at $d_n = 2n^{3/4}$, which yields $c_n = 4n^{3/2} + 4n^{3/4} \le 8n^{3/2}$. □

The phrasing of Conjecture 2.1.2 is motivated by a particular application in numerical analysis. Suppose one wants to evaluate analytic functions $f(A)$ of a given matrix A, which may be nonnormal. If A is diagonalizable, one can use the formula $f(A) = Vf(D)V^{-1}$, where $f(D)$ means the function is applied to the scalar diagonal entries of D. However, this may be numerically infeasible if $\kappa_V(A)$ is very large: if all computations are carried to precision ϵ, the result may be off by an error of $\kappa_V(A)\epsilon$. Davies' idea was to replace A by a perturbation $A+E$ with a much smaller $\kappa_V(A+E)$, and compute $f(A+E)$ instead. In [Dav07, Theorem 2.4], he showed that the net error incurred by this scheme for a given $\epsilon > 0$ and sufficiently regular f is controlled by:

$$\kappa_V(A+E)\epsilon + \|E\|,$$

which is the quantity appearing in (2.2). The key desirable feature of (2.2) is the dimension-independent fractional power of ϵ on the right-hand side, which shows that the total error scales slowly.

Davies proved his conjecture in the special case of upper triangular Toeplitz matrices, in dimension $n = 3$ with the constant $c_n = 2$, as well as in the general case with the weaker dimension-dependent and nonconstructive bound $(n+1)\epsilon^{2/(n+1)}$. This last result corresponds to (2.1) above. He also speculated that a *random* regularizing perturbation E suffices to prove Conjecture 2.1.2, and presented empirical evidence to that effect. Our proof of Theorem 2.1.1 below indeed follows this strategy.

Gaussian Regularization

Theorem 2.1.1 follows from a probabilistic result concerning complex Gaussian perturbations of a given matrix A. To state our result, we recall two standard notions.

Definition 2.1.3. *A* complex Ginibre matrix *is an* $n \times n$ *random matrix* $G_n = (g_{ij})$ *with i.i.d complex Gaussian entries* $g_{ij} \sim N(0, 1_{\mathbb{C}}/n)$*, by which we mean* $\mathbb{E} g_{ij} = 0$ *and* $\mathbb{E}|g_{ij}|^2 = 1/n$*. Equivalently the real and imaginary parts of each* g_{ij} *are independent* $N(0, 1/2n)$ *random variables.*

Definition 2.1.4. *Let* $M \in \mathbb{C}^{n \times n}$ *have distinct eigenvalues* $\lambda_1, \ldots, \lambda_n$*, and spectral expansion*

$$M = \sum_{i=1}^{n} \lambda_i v_i w_i^* = V D V^{-1},$$

where the right and left eigenvectors v_i *and* w_i^* *are the columns and rows of* V *and* V^{-1}*, respectively, normalized so that* $w_i^* v_i = 1$*. The* eigenvalue condition number of λ_i *is defined as:*

$$\kappa(\lambda_i) \triangleq \|v_i w_i^*\| = \|v_i\| \|w_i\|.$$

The $\kappa(\lambda_i)$'s are called condition numbers because they determine the sensitivity of the λ_i to perturbations of the matrix. We show that adding a small Ginibre perturbation regularizes the eigenvalue condition numbers of any matrix in the following averaged sense.

Theorem 2.1.5. *Suppose* $A \in \mathbb{C}^{n \times n}$ *with* $\|A\| \le 1$ *and* $\delta \in (0,1)$*. Let* G_n *be a complex Ginibre matrix, and let* $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ *be the (random) eigenvalues of* $A + \delta G_n$*. Then for every measurable open set* $B \subset \mathbb{C}$*,*

$$\mathbb{E} \sum_{\lambda_i \in B} \kappa(\lambda_i)^2 \le \frac{n^2}{\pi \delta^2} \mathrm{vol}(B).$$

Note that the $\kappa(\lambda_i)$ appearing above are well-defined because $A + \delta G_n$ has distinct eigenvalues with probability one.

Related Work

Random Matrix Theory. There have been numerous studies of the eigenvalue condition numbers $\kappa(\lambda_i)^2$, sometimes called eigenvector *overlaps* in the random matrix theory and mathematical physics literature, for non-Hermitian random matrix models of type $A + \delta G_n$. In the centered case $A = 0$ and $\delta = 1$ of a standard complex Ginibre matrix, the seminal work of Chalker and Mehlig [CM98] calculated the large-n limit of the conditional expectations

$$\mathbb{E}[\kappa(\lambda)^2 | \lambda = z] \underset{n \to \infty}{\sim} n(1 - |z|^2),$$

whenever $|z| < 1$. Recent works by Bourgade and Dubach [BD18] and Fyodorov [Fyo18] improved on this substantially by giving exact nonasymptotic formulas for the distribution of $\kappa(\lambda)^2$ conditional on the location of the eigenvalue λ, as well as concise descriptions of the scaling limits for these formulas. The paper [BGZ18] proved (in the more general setup of invariant ensembles) that the angles between the right eigenvectors $(v_i^* v_j)/\|v_i\|\|v_j\|$ have subgaussian tails, which has some bearing on κ_V (for instance, a small angle between unit eigenvectors causes $\|V^{-1}\|$ and therefore κ_V to blow up.)

In the non-centered case, Davies and Hager [DH09] showed that if A is a Jordan block and $\delta = n^{-\alpha}$ for some appropriate α, then almost all of the eigenvalues of $A + \delta G_n$ lie near a circle of radius $\delta^{1/n}$ with probability $1 - o_n(1)$. Basak, Paquette, and Zeitouni [BPZ18, BPZ19] showed that for a sequence of banded Toeplitz matrices A_n with a finite symbol, the spectral measures of $A_n + n^{-\alpha} G_n$ converge weakly in probability, as $n \to \infty$, to a predictable density determined by the symbol. Both of the above results were recently and substantially improved by Sjöstrand and Vogel [SV19a, SV19b] who proved that for any Toeplitz A, almost all of the eigenvalues of $A + n^{-\alpha} G_n$ are close to the symbol curve of A with exponentially good probability in n. Note that none of the results mentioned in this paragraph explicitly discuss the $\kappa(\lambda_i)$; however, they do deal qualitatively with related phenomena surrounding spectral instability of non-Hermitian matrices.

The idea of managing spectral instability by adding a random perturbation can be traced back to the influential papers of Haagerup and Larsen [HL00] and Śniady [Śni02] (see also [GWZ14, FPZ14]), who used it to study convergence of the eigenvalues of certain non-Hermitian random matrices to a limiting Brown measure, in the context of free probability theory.

There are three notable differences between Theorem 2.1.5 and the results mentioned above:

1. Our result is much coarser, and only guarantees an upper bound on the $\mathbb{E}\kappa(\lambda_i)^2$, rather than a precise description of any distribution, limiting or not.

2. It applies to any $A \in \mathbb{C}^{n \times n}$ and $\delta \in (0, 1)$.

3. It is completely nonasymptotic and does not require $n \to \infty$ or even sufficiently large n.

Numerical Analysis. In the numerical linear algebra literature, several works have analyzed the condition numbers of Gaussian matrices (notably the seminal results of Demmel [Dem83] and Edelman [Ede88]) as well as perturbations of arbitrary matrices by Gaussian matrices (beginning with [SST06]) in the nonasymptotic regime. In contrast, this paper studies the condition numbers of the *eigenvectors* of such matrices, rather than of the matrices themselves.

The idea of approximating matrix functions by adding a regularizing perturbation was introduced in [Dav07] and has since appeared in several works regarding numerical computation of the matrix logarithm, sine, cosine, and related functions [AMHR13, HL13, AMHR15, NH18, DIP$^+$19].

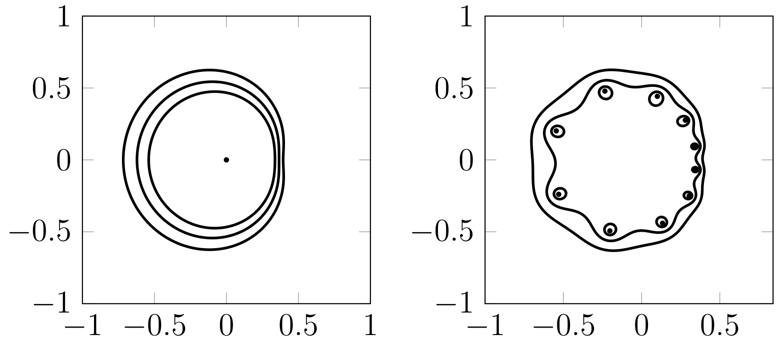

Figure 2.1: T is a sample of an upper triangular 10×10 Toeplitz matrix with zeros on the diagonal and independent (modulo the Toeplitz structure) standard real Gaussian entries above the diagonal. Pictured is the boundary of the ϵ-pseudospectrum of T (left) and $T + 10^{-6}G$ (right) for $\epsilon = 10^{-5}$, $\epsilon = 10^{-5.5}$, and $\epsilon = 10^{-6}$, along with the spectra. These plots were generated with the MATLAB package EigTool [WT02].

Techniques and Organization

The proofs of Theorems 2.1.1 and 2.1.5 are quite simple and rely on an interplay between various notions of spectral stability. In addition to κ_V and the $\kappa(\lambda_i)$, we will heavily use the notion of the $\epsilon-$*pseudospectrum* of a matrix M, defined for $\epsilon > 0$ as:

$$\begin{aligned}
\Lambda_\epsilon(M) &\triangleq \{z \in \mathbb{C} : z \in \Lambda(M + E) \text{ for some } \|E\| < \epsilon\} && (2.3)\\
&= \left\{z \in \mathbb{C} : \|(zI - M)^{-1}\| > 1/\epsilon\right\} && (2.4)\\
&= \{z \in \mathbb{C} : \sigma_n(zI - M) < \epsilon\}\,, && (2.5)
\end{aligned}$$

where $\Lambda(M)$ denotes the spectrum M. For a proof of the equivalence of these three sets and a comprehensive treatment of pseudospectra, see the beautiful book by Trefethen and Embree [TE05]. Note that for a normal matrix, we have

$$\Lambda_\epsilon(M) = \Lambda(M) + \bigcup_i D(\lambda_i, \epsilon),$$

whereas for a nonnormal matrix such as a Jordan block, Λ_ϵ can be much larger. Figure 1 illustrates the regularizing effect of a small complex Gaussian perturbation on the pseudospectrum of a nondiagonalizable matrix, which is the key phenomenon underlying our results.

To analyze this phenomenon, we first collect some tools from random matrix theory in Section 2.2, along the way proving a conjecture of Sankar, Spielman, and Teng [SST06] regarding the optimal constant in their smoothed analysis of condition numbers of matrices under *real* Gaussian perturbations in Section 2.2. Section 2.3 contains the proofs of our main results, Theorems 2.1.1 and 2.1.5. In Section 2.4, we prove optimality of the $1/\delta$-dependence in Theorem 2.1.1 as discussed above, and show that Theorem 2.1.5 is sharp up to a small constant factor. We conclude with a discussion of some open problems in Section 2.5.

Notation

We denote the singular values of an $n \times n$ matrix by $\sigma_1(M) \geq \ldots \geq \sigma_n(M)$, its operator and Frobenius (Hilbert-Schmidt) norms by $\|M\|$ and $\|M\|_F$, and its condition number by $\kappa(M) \triangleq \sigma_1(M)/\sigma_n(M)$. Open disks in the complex plane will be written as $D(z_0, r) \triangleq \{z \in \mathbb{C} : |z - z_0| < r\}$. We will often write G for a standard complex Gaussian matrix with $N(0, 1_{\mathbb{C}})$ entries, and $G_n = n^{-1/2}G$ for a (normalized) Ginibre matrix.

2.2 Tools from Random Matrix Theory

Nonasymptotic Extreme Singular Value Estimates

Let us record some standard non-asymptotic estimates for the extreme singular values of complex Ginibre matrices. The lower tail behavior of the smallest singular value of a Ginibre matrix was worked out by Edelman in the unnormalized scaling of i.i.d. $N(0, 1_{\mathbb{C}})$ entries [Ede88, Chapter 5], and in our setting it translates to:

Theorem 2.2.1. *For a complex Ginibre matrix G_n,*

$$\mathbb{P}[\sigma_n(G_n) < \epsilon] = 1 - e^{-\epsilon^2 n^2} \leq \epsilon^2 n^2.$$

We will also require a cruder tail estimate on the largest singular value. We believe the lemma holds with a constant 2 instead of $2\sqrt{2}$, but did not find a reference to a nonasymptotic result to this effect; since the difference is not very consequential in this context, we reduce to the real case.

Lemma 2.2.2. *For a complex Ginibre matrix G_n,*

$$\mathbb{P}[\sigma_1(G_n) > 2\sqrt{2} + t] \leq 2\exp(-nt^2).$$

Proof. We can write $G_n = \frac{1}{\sqrt{2}}(X + iY)$ where X and Y are independent with i.i.d. *real* $N(0, 1/n)$ entries. It is well-known (e.g. [DS01, Theorem II.11]) that:

$$\mathbb{E}\sigma_1(G_n) \leq \frac{2}{\sqrt{2}}\mathbb{E}\|X\| \leq 2\sqrt{2}.$$

Lipschitz concentration of functions of real Gaussian random variables yields the result. $\square$

Śniady's Comparison Theorem

To bound the least singular value of noncentered Gaussian matrices, we will lean on a remarkable theorem of Śniady [Śni02].

Theorem 2.2.3 (Śniady). *Let A_1 and A_2 be $n \times n$ complex matrices such that $\sigma_i(A_1) \leq \sigma_i(A_2)$ for all $1 \leq i \leq n$. Assume further that $\sigma_i(A_1) \neq \sigma_j(A_1)$ and $\sigma_i(A_2) \neq \sigma_j(A_2)$ for all $i \neq j$. Then for every $t \geq 0$, there exists a joint distribution on pairs of $n \times n$ complex matrices (G_1, G_2) such that*

(i) the marginals G_1 and G_2 are distributed as (normalized) complex Ginibre matrices G_n, and

(ii) almost surely $\sigma_i(A_1 + \sqrt{t}G_1) \leq \sigma_i(A_2 + \sqrt{t}G_2)$ for every i.

We will briefly sketch the proof of this theorem for the reader's benefit, since it is quite beautiful and we will need to perform a slight modification to prove the conjecture of Sankar-Spielman-Teng in the next subsection.

Sketch of proof. The key insight of the proof is that it is possible to couple the distributions of G_1 and G_2 through their singular values. To do so, one first derives a stochastic differential equation satisfied by the singular values $s_1, ..., s_n$ of a matrix Brownian motion (i.e., a matrix whose entries are independent complex Brownian motions):

$$ds_i = \frac{1}{\sqrt{2n}} dB_i + \frac{dt}{2s_i}\left(1 - \frac{1}{2n} + \sum_{j \neq i} \frac{s_i^2 + s_j^2}{n(s_i^2 - s_j^2)}\right), \tag{2.6}$$

where the B_i are independent standard real Brownian motions. Next, one uses a single n-tuple of real Brownian motions $B_1, ..., B_n$ to drive two processes $(s_1^{(1)}, \ldots, s_n^{(1)})$ and $(s_1^{(2)}, \ldots, s_n^{(2)})$ according to (2.6), with initial conditions $s_i^{(1)}(0) = \sigma_i(A_1)$ and $s_i^{(2)}(0) = \sigma_i(A_2)$ for all i. (To do this rigorously, one needs existence and uniqueness of strong solutions to the above SDE; this is shown in [KO$^+$01] under the hypothesis $s_i(0) \neq s_j(0)$ for all $i \neq j$.)

Things have been arranged so that the joint distribution of $(s_1^{(j)}, \ldots, s_n^{(j)})$ at time t matches the joint distribution of the singular values of $A_j + \sqrt{t}G_j$ for each $j = 1, 2$. One can then sample unitaries U_j and V_j from the distribution arising from the singular value decomposition $A_j + \sqrt{t}G_j = U_j D_j V_j^*$, conditioned on $D_j = \operatorname{diag}(s_1^{(j)}, \ldots, s_n^{(j)})$. Thus each G_j is separately Ginibre-distributed. However, $A_1 + \sqrt{t}G_1$ and $A_2 + \sqrt{t}G_2$ are coupled through the shared randomness driving the evolution of their singular values. In particular, since the same B_i were used for both processes, from (2.6) one can verify that the n differences $s_i^{(2)} - s_i^{(1)}$ are C^1 in t. By taking derivatives, one can then show the desired monotonicity property: if $s_i^{(2)} - s_i^{(1)} \geq 0$ holds for all i at $t = 0$, it must hold for all $t \geq 0$. $\square$

Sankar-Spielman-Teng Conjecture

The proof technique of Śniady can be adapted to prove a counterpart of Theorem 2.2.3 for *real* Ginibre perturbations (by this we mean matrices with i.i.d. real $N(0, 1/n)$ entries). Because a rigorous proof requires some stochastic analysis, we defer the proof and discussion of the following theorem to the appendix.

Theorem 2.2.4. *Let A_1 and A_2 be $n \times n$ real matrices such that $\sigma_i(A_1) \le \sigma_i(A_2)$ for all $1 \le i \le n$. Assume further that $\sigma_i(A_1) \neq \sigma_j(A_1)$ and $\sigma_i(A_2) \neq \sigma_j(A_2)$ for all $i \neq j$. Then for every $t \ge 0$, there exists a joint distribution on pairs of real $n \times n$ matrices (G_1, G_2) such that*

(i) the marginals G_1 and G_2 are distributed as real Ginibre matrices (with i.i.d. $N(0, 1/n)$ entries), and

(ii) almost surely $\sigma_i(A_1 + \sqrt{t}G_1) \le \sigma_i(A_2 + \sqrt{t}G_2)$ for every i.

This resolves Conjecture 1 in [SST06], which we restate below as a proposition:

Proposition 2.2.5. *Let G be an $n \times n$ matrix with i.i.d. real $N(0, 1)$ entries, and A be any $n \times n$ matrix with real entries. Then*

$$\mathcal{P}[\sigma_n(A + G) < \epsilon] \le \epsilon\sqrt{n}.$$

Proof. The case $A = 0$ is a result of Edelman [Ede88]. The proposition for general A would then follow from Theorem 2.2.4 with $A_1 = 0$ and $A_2 = A$ if not for the hypothesis $\sigma_i(A_1) \neq \sigma_j(A_1)$ and $\sigma_i(A_2) \neq \sigma_j(A_2)$ for all $i \neq j$. So we approach 0 and A by matrices satisfying this hypothesis, apply Theorem 2.2.4, and take limits, using the continuous mapping theorem and continuity of $\sigma_n(\cdot)$. □

2.3 Proof of Theorems 2.1.1 and 2.1.5

We begin with a lemma relating the eigenvector and eigenvalue condition numbers. For related results, including an extension of this lemma to the more general context of block diagonalization, see the thesis of Demmel [Dem83, Equation 3.6].

Lemma 2.3.1. *Let M be an $n \times n$ matrix with distinct eigenvalues, and let V be the matrix whose columns are the eigenvectors of M normalized to have unit norm. Then*

$$\kappa(V) \le \sqrt{n \sum_{i=1}^{n} \kappa(\lambda_i)^2}.$$

Proof. Note that the left eigenvectors w_i are the rows of V^{-1}. Then $\|V\|_F^2 = n$ and $\|V^{-1}\|_F^2 = \sum_{i=1}^n \|w_i\|^2 = \sum_{i=1}^n \kappa(\lambda_i)^2$, so

$$\kappa(V) = \|V\|\|V^{-1}\| \leq \|V\|_F\|V^{-1}\|_F = \sqrt{n\sum_{i=1}^n \kappa(\lambda_i)^2}.$$

□

We can now prove the main theorem.

Proof of Theorem 2.1.1 given Theorem 2.1.5. Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of the random matrix $A + \delta G_n$, and $t > 2\sqrt{2}$ and $s > 1$ be parameters to be optimized later. Davies' original bound (2.1) implies our bound for $n \leq 3$, so assume $n \geq 4$. Then Lemma 2.2.2 tells us that

$$\mathbb{P}[\|\delta G_n\| \geq t\delta] \leq 2e^{-4(t-2\sqrt{2})^2}. \tag{2.7}$$

Letting $B = D(0, \|A\| + t\delta)$, we have

$$\mathbb{P}\left[\sum_{\lambda_i \in B} \kappa(\lambda_i)^2 \neq \sum_{i \leq n} \kappa(\lambda_i)^2\right] \leq \mathbb{P}[\|\delta G_n\| \geq t\delta] \leq 2e^{-4(t-2\sqrt{2})^2} \tag{2.8}$$

since $\max_{i \leq n} |\lambda_i| \leq \|A\| + \|\delta G_n\|$. On the other hand, by Theorem 2.1.5 applied to B and Markov's inequality:

$$\mathbb{P}\left[\sum_{\lambda_i \in B} \kappa(\lambda_i)^2 \geq s\frac{n^2 \mathrm{vol}(B)}{\delta^2 \pi}\right] \leq \frac{1}{s}. \tag{2.9}$$

By the union bound, if we choose s and t such that

$$2e^{-4(t-2\sqrt{2})^2} + \frac{1}{s} < 1 \tag{2.10}$$

then there exists a choice of G_n such that neither of the events (2.8), (2.9) occurs. Letting $E = \delta G_n$ for this choice, we have

$$\sum_{i=1}^n \kappa(\lambda_i)^2 = \sum_{i \in B} \kappa(\lambda_i)^2 \leq s\frac{n^2 \mathrm{vol}(B)}{\pi\delta^2}.$$

Taking a square root and applying Lemma 2.3.1, we have

$$\kappa_V(A + E) \leq \frac{\sqrt{s}n^{3/2}}{\delta}(\|A\| + t\delta) \leq \frac{\sqrt{s}n^{3/2}\|A\|}{\delta} + t\sqrt{s}n^{3/2}.$$

Because $\|E\| \leq t\delta$ and not δ, replacing δ by δ/t yields the bound

$$\kappa_V(A + E) \leq \frac{t\sqrt{s}n^{3/2}\|A\|}{\delta} + t\sqrt{s}n^{3/2}.$$

To get the best bound, we must minimize $t\sqrt{s}$ subject to the constraints (2.10), $t > 2\sqrt{2}$ and $s > 1$. Solving for s this becomes a univariate optimization problem, and one can check numerically that the optimum is achieved at $t \approx 3.7487$ and $t\sqrt{s} \approx 3.8822 < 4$, as advertised. □

We begin the proof of Theorem 2.1.5 by relating the eigenvalue condition numbers of a matrix to the rate at which its pseudospectrum Λ_ϵ shrinks as a function of the parameter $\epsilon > 0$. The following proposition is not new; the proof essentially appears for example in Section 3.6 of [BD18], but we include it for completeness since it is critical to our argument.

Lemma 2.3.2 (Limiting Area of the Pseudospectrum). *Let M be an $n \times n$ matrix with n distinct eigenvalues $\lambda_1, \ldots, \lambda_n$ and let $B \subset \mathbb{C}$ be a measurable open set. Then*

$$\lim_{\epsilon \to 0} \frac{\mathrm{vol}(\Lambda_\epsilon(M) \cap B)}{\epsilon^2} = \pi \sum_{\lambda_i \in B}^{n} \kappa(\lambda_i)^2.$$

Proof. Write the spectral decomposition

$$(zI - M)^{-1} = \sum_{i=1}^{n} \frac{v_i w_i^*}{z - \lambda_i},$$

where the v_i and w_i^* are right and left eigenvectors of M, respectively. Since the λ_i are distinct, we may choose $\epsilon_0 > 0$ sufficiently small to guarantee that there exists a constant $C > 0$ satisfying (1) the disks $D(\lambda_i, \epsilon_0)$ are disjoint; (2) for every $\lambda_i \in B$ the disk $D(\lambda_i, \epsilon_0)$ is contained in B; and (3) whenever $z \in D(\lambda_i, \epsilon_0)$ for some i,

$$\|(zI - M)^{-1}\| \geq \frac{\|v_i w_i^*\|}{|z - \lambda_i|} - C = \frac{\kappa(\lambda_i)}{|z - \lambda_i|} - C. \tag{2.11}$$

Using the definition of the ϵ-pseudospectrum in (2.4), we rearrange (2.11) to obtain

$$\Lambda_\epsilon(M) \cap B \supset \left\{ z : |z - \lambda_i| \leq \min\left\{ \epsilon_0, \frac{\kappa(\lambda_i)\epsilon}{1 + \epsilon C} \right\}, \text{ for some } \lambda_i \in B \right\}.$$

Thus, taking ϵ small enough, we have

$$\liminf_{\epsilon \to 0} \frac{\mathrm{vol}(\Lambda_\epsilon(M) \cap B)}{\epsilon^2} \geq \pi \sum_{i=1}^{n} \kappa(\lambda_i)^2.$$

For the opposite inequality, Theorem 52.1 of [TE05] states that the ϵ-pseudospectrum is contained in disks around the eigenvalues λ_i of radii $\epsilon\kappa(\lambda_i) + O(\epsilon^2)$. Choosing ϵ small enough so that for $\lambda_i \in B$ these disks are entirely contained in B:

$$\mathrm{vol}(\Lambda_\epsilon \cap B) \leq \sum_{\lambda_i \in B} \pi(\epsilon\kappa(\lambda_i) + O(\epsilon^2))^2 \Rightarrow \limsup_{\epsilon \to 0} \frac{\mathrm{vol}(\Lambda_\epsilon \cap B)}{\epsilon^2} \leq \sum_{\lambda_i \in B} \pi\kappa(\lambda_i)^2.$$

□

Next, we show that for fixed $\epsilon > 0$, any particular point $z \in \mathbb{C}$ is unlikely to be in $\Lambda_\epsilon(A + \delta G_n)$. This is based on the following singular value estimate, which generalizes Theorem 2.2.1.

Lemma 2.3.3 (Small Ball Estimate for σ_n)**.** *Let M be an $n \times n$ matrix with complex entries, and G be drawn from the Ginibre ensemble. Then for all $\delta > 0$ and $\epsilon > 0$*

$$\mathbb{P}\left[\sigma_n(M + \delta G_n) < \epsilon\right] \leq n^2 \frac{\epsilon^2}{\delta^2}.$$

Proof. Repeat the proof of Proposition 2.2.5 using instead Theorems 2.2.1 and 2.2.3. □

Remark 2.3.4. *If one is willing to lose a small constant factor in the bound, Lemma 2.3.3 has an elementary geometric proof (which avoids stochastic calculus), essentially identical to the proof of Sankar-Spielman-Teng [SST06, Theorem 3.1] in the case of real Ginibre perturbations. Note however that it is crucial to use a* complex *Gaussian for our purposes. A real Gaussian would yield a small ball estimate of order ϵ (see Proposition 2.2.5) rather than ϵ^2, which is not good enough to take the limit below.*

Proof of Theorem 2.1.5. For every $z \in \mathbb{C}$ we have the upper bound

$$\mathbb{P}[z \in \Lambda_\epsilon(A + \delta G_n)] = \mathbb{P}[\sigma_n(zI - (A + \delta G_n)) < \epsilon] \leq n^2 \frac{\epsilon^2}{\delta^2}, \tag{2.12}$$

by applying Lemma 2.3.3 to $M = zI - A$ and noting that G and $-G$ have the same distribution.

Fix a measurable open set $B \subset \mathbb{C}$. Then

$$\begin{aligned}
\mathbb{E}\,\mathrm{vol}(\Lambda_\epsilon(A + \delta G_n) \cap B) &= \mathbb{E} \int_B \mathbb{1}_{\{z \in \Lambda_\epsilon(A+\delta G_n)\}}\, dz \\
&= \int_B \mathbb{E}\{z \in \Lambda_\epsilon(A + \delta G_n)\}\, dz && \text{by Fubini} \\
&\leq \int_B n^2 \frac{\epsilon^2}{\delta^2}\, dz && \text{by (2.12)} \\
&= n^2 \frac{\epsilon^2}{\delta^2} \mathrm{vol}(B) && (2.13)
\end{aligned}$$

where the integrals are with respect to Lebesgue measure on $\mathbb{C}$. Finally, taking a limit as $\epsilon \to 0$ yields the desired bound:

$$\begin{aligned}
\mathbb{E} \sum_{\lambda_i \in B} \kappa(\lambda_i^2) &= \mathbb{E} \liminf_{\epsilon \to 0} \frac{\mathrm{vol}(\Lambda_\epsilon(A + \delta G_n) \cap B)}{\pi \epsilon^2} && \text{by Lemma 2.3.2} \\
&\leq \liminf_{\epsilon \to 0} \mathbb{E} \frac{\mathrm{vol}(\Lambda_\epsilon(A + \delta G_n) \cap B)}{\pi \epsilon^2} && \text{by Fatou's Lemma} \\
&\leq \frac{n^2 \mathrm{vol}(B)}{\pi \delta^2} && \text{by (2.13).}
\end{aligned}$$

□

2.4 Optimality of the Bounds

We first show that Theorem 2.1.1 has essentially the optimal dependence on δ for n large. The example which requires this dependence is simply a Jordan block J, for which Davies [Dav07] established the upper bound $\kappa_V(J+\delta E) \leq 2/\delta^{1-1/n}$, for some E with $||E|| < 1$.

Proposition 2.4.1. *Fix $n > 0$ and let $J \in \mathbb{C}^{n\times n}$ be the upper triangular Jordan block with ones on the superdiagonal and zeros everywhere else. Then there exist $c_n > 0$ and $\delta_n > 0$ such that for all $E \in \mathbb{C}^{n\times n}$ with $\|E\| \leq 1$ and all $\delta < \delta_n$, we have*

$$\kappa_V(J+\delta E) \geq \frac{c_n}{\delta^{1-1/n}}.$$

Proof. As a warm-up, we'll need the following bound on the pseudospectrum of J. Let λ be an eigenvalue of $J+\delta E$, with v its associated right eigenvector; then $(J+\delta E)^n v = \lambda^n v$ and, accordingly, $|\lambda|^n \leq \|(J+\delta E)^n\|$. Expanding, using nilpotence of J, $\|J\| = 1$, and submultiplicativity of the operator norm, we get

$$|\lambda|^n \leq \|(J+\delta E)^n\| \leq (1+\delta)^n - 1 = O(\delta) \tag{2.14}$$

where the big-O refers to the limit $\delta \to 0$ (recall n is fixed).

Writing $J+\delta E = V^{-1}DV$, we want to lower bound the condition number of V. As above, let λ be an eigenvalue of $J+\delta E$, now writing w^* and v for its left and right eigenvectors. We'll use the lower bound

$$\kappa(V) = \|V^{-1}\|\|V\| \geq \frac{\|w^*\|\|v\|}{|w^*v|} = \kappa(\lambda).$$

Since the formula above is agnostic to the scaling of the left and right eigenvectors, we'll assume that both have unit length and show that $|w^*v|$ is small.

Let $0 \leq k \leq n$. Then $\|(J+\delta E)^k v\| = |\lambda|^k$, and analogously to (2.14),

$$\|(J+\delta E)^k - J^k\| \leq (1+\delta)^k - 1 = O(\delta).$$

Since J acts on the left as a left shift,

$$\begin{aligned}\left(\sum_{i=k+1}^{n} |v_i|^2\right)^{1/2} &= \|J^k v\| \\ &\leq \|(J+\delta E)^k v\| + \|(J^k - (J+\delta E)^k)v\| \\ &\leq |\lambda|^k + O(\delta) \\ &= O(\delta^{k/n}),\end{aligned}$$

where the final line follows from (2.14). Similarly,

$$\left(\sum_{i=1}^{n-k} |w_i|^2\right)^{1/2} = \|w^* J^k\| = O(\delta^{k/n}).$$

Finally, we have $\kappa(V)^{-1} = |w^* v| \leq \sum_{j=1}^{n} |w_j||v_j|$, which in turn is at most

$$\sum_{j=1}^{n}\left(\sum_{i=1}^{j}|w_i|^2\right)^{1/2}\left(\sum_{i=j}^{n}|v_i|^2\right)^{1/2} = O(\delta^{(n-j)/n}\delta^{(j-1)/n}) = O(\delta^{1-1/n}).$$

□

We end by showing that the dependence on n in Theorem 2.1.5 cannot be improved.

Proposition 2.4.2. *There exists $c > 0$ such that for all n,*

$$\mathbb{E}\sum_{i\in[n]}\kappa^2\left(\lambda_i(G_n)\right) \geq cn^2.$$

Proof. Bourgade and Dubach [BD18, Theorem 1.1, Equation 1.8] show that eigenvalue condition numbers in the bulk of the spectrum of complex Ginibre matrices are of order $\sqrt{n}$. Precisely,

$$\lim_{n\to\infty}\frac{\mathbb{E}[\kappa(\lambda_i)^2|\lambda_i = z]}{n} = 1 - |z|^2$$

uniformly for (say) $z \in D(0,r)$ for any $r < 1$. The classical *circular law* for the limiting spectral distribution of Ginibre matrices ensures that

$$\lim_{n\to\infty}\frac{\mathbb{E}|\Lambda(G_n)\cap D(0,r)|}{n} = \frac{\mathrm{vol}(D(0,r))}{\mathrm{vol}(D(0,1))} = r^2.$$

Thus,

$$\liminf_{n\to\infty}\frac{\mathbb{E}\sum_{i\in[n]}\kappa\left(\lambda_i(G_n)\right)^2}{n^2} \geq r^2(1-r^2) > 0.$$

□

2.5 Conclusion and Discussion

A key theme in our work is the interplay between the related notions of eigenvector condition number κ_V, eigenvalue condition number $\kappa(\lambda_i)$ and pseudospectrum Λ_ϵ. Equally important is the fact that global objects such as κ_V and Λ_ϵ can be controlled by local quantities, specifically the least singular values of shifts $\sigma_n(zI - M)$ for each $z \in \mathbb{C}$. The proof also heavily exploits the left and right unitary invariance of the Ginibre ensemble (via Theorem 2.2.3, due to Śniady) as well as anticoncentration of the complex Gaussian.

One natural question is whether similar results hold if one replaces Gaussian perturbations with a different class of random perturbations G'. To apply the approach in this paper, the key difficulty would be obtaining suitable bounds for the least singular value of

$z - A - \delta G'$. Davies [Dav07] presents experimental evidence that Theorem 2.1.1 holds for random real rank-one perturbations and random real Gaussian perturbations, but a proof (or disproof) remains to be found. See Remark 2.3.4 for a discussion of why our proof does not extend to the case of real Gaussian perturbations.

One may also ask if Theorem 2.1.1 can be derandomized; that is, if the regularizing perturbation E can be chosen by a deterministic algorithm given A as input. One natural choice would be to perturb in the direction of the nearest normal matrix in either operator or Frobenius norm, the latter of which can be written as a certain optimization problem over unitary matrices [Ruh87].

Proposition 2.4.1 shows that the upper bound in Theorem 2.1.1 is tight in the perturbation size δ. Now, let c_n be the smallest constant such that Theorem 2.1.1 holds with an upper bound of c_n/δ. Theorem 2.1.1 implies that $c_n \leq 8n^{3/2}$, and since $\kappa_V = \|V\|\|V^{-1}\| \geq 1$ for any matrix, we have $c_n \geq 1$. It would be interesting to determine the correct asymptotic behavior of c_n. Davidson, Herrero, and Salinas asked in 1989 [DHS89] whether the statement of Theorem 2.1.1 is possible with $\kappa_V(A + E)$ depending only on δ and not on n. In the present context, we can ask the more refined question: does Theorem 2.1.1 hold with bounded c_n, or must c_n go to infinity with n?

2.6 Appendix : Proof of Theorem 2.2.4

The goal of this appendix is to adapt Śniady's [Śni02] proof of Theorem 2.2.3, as outlined below the statement of Theorem 2.2.3, to the case of real matrices with real Ginibre perturbations.

The stochastic differential equation satisfied by the squared singular values of a real matrix Brownian motion was derived by Bru in her work on Wishart processes [Bru89, Bru91] and independently by Le in her work on shape theory [Le94, Le99]. The equation reads as follows:

$$d\lambda_i = \frac{2\sqrt{\lambda_i}}{n}\, dB_i + \left(1 + \sum_{j \neq i} \frac{\lambda_i + \lambda_j}{\lambda_i - \lambda_j}\right) dt, \qquad 1 \leq i \leq n. \tag{2.15}$$

The proof strategy of Śniady crucially relies on the existence and uniqueness of strong solutions to the singular value SDE. This is needed in order to obtain two solutions driven by the same Brownian motion, and to assert that the law of each solution indeed matches the law of the singular values of a noncentered Ginibre matrix. See [AGZ10] for a definition of strong solution and a rigorous proof of existence and uniqueness of strong solutions for Dyson Brownian motion, the Hermitian analogue of the Ginibre singular values process.

Fortunately, such results are known for the SDE (2.15). Let Λ denote the domain

$$\Lambda \in \mathbb{R}^n := \{\lambda : 0 \leq \lambda_n < \cdots < \lambda_1\}.$$

For any initial data $\lambda(0)$ lying in the closure $\overline{\Lambda}$, it is known that strong solutions to (2.15) exist, are unique, and lie in Λ for all $t > 0$, almost surely [GM14, Corollary 6.5]. Combining this with [Bru89, Theorem 1], we have that for initial data $\lambda(0)$ lying in Λ, the law of the strong solutions to (2.15) matches the law of the squared singular values process of $A+M/\sqrt{n}$, where M is a matrix of i.i.d. standard real Brownian motions and A has squared singular values $\lambda(0)$. (It should be possible to extend this last statement for initial data in $\overline{\Lambda}$, but the proof may be somewhat involved—cf. [AGZ10], which contains a proof of the corresponding extension for Dyson Brownian motion.)

Let $a_i(\lambda) = 1 + \sum_{j \neq i} \frac{\lambda_i + \lambda_j}{\lambda_i - \lambda_j}$ denote the drift coefficient in (2.15). As in Śniady's proof for the complex Ginibre case (Theorem 2.2.3), the key property of a allowing for the comparison theorem is the so-called *quasi-monotonicity* (see [DW98]) or *Kamke–Ważewski condition* [MPF91, §XI.13] from differential inequalities, which is simply that

$$\text{for all } i,\ a_i(\lambda^{(1)}) \leq a_i(\lambda^{(2)}) \text{ whenever } \lambda_i^{(1)} = \lambda_i^{(2)} \text{ and } \lambda_j^{(1)} \leq \lambda_j^{(2)} \text{ for all } j \neq i. \tag{2.16}$$

One easily checks that a satisfies this condition on the domain Λ.

The nonconstant (indeed, non-Lipschitz) diffusion coefficient $2\sqrt{\lambda_i}/n$ in (2.15) is a technical obstacle which does not appear in the SDE (2.6) for the complex case. Consequently, the final step of Śniady's proof as sketched below Theorem 2.2.3 cannot be repeated naively, because taking the difference of two solutions no longer cancels out the diffusion terms. Fortunately, theory has been developed to handle Hölder-1/2 diffusion coefficients; see [RY99, §IX.3] for exposition of the one-dimensional case and see [Kra10] for a survey of comparison theorems for SDEs in general.

Quasi-monotonicity and the one-dimensional Hölder-1/2 comparison theory are combined in a rather general multidimensional comparison theorem of Geiß and Manthey [GM94, Theorem 1.2]. Applied to the SDE (2.15), this theorem provides exactly the right conclusion to replace the final step of Śniady's proof. We state the relevant special case of their theorem below:

Theorem 2.6.1 (Geiß-Manthey). *Consider the SDE*

$$dX_i = \sigma_i(X)\, dB_i + a_i(X)\, dt, \qquad 1 \leq i \leq n,$$

where the B_i are independent standard real Brownian motions, and $\sigma_i, a_i : \mathbb{R}^n \to \mathbb{R}$ are continuous. Suppose the following conditions are satisfied:

1. *the drift coefficient a satisfies the quasi-monotonicity condition (2.16)*
2. *there exists $\rho : \mathbb{R}_+ \to \mathbb{R}_+$ increasing with $\int_0^\epsilon \rho^{-2}(u)\, du = \infty$ for some $\epsilon > 0$, such that $|\sigma_i(x) - \sigma_i(y)| \leq \rho(|x_i - y_i|)$ for all i and all $x, y \in \mathbb{R}^n$*
3. *strong solutions for the SDE exist for all time and are unique.*

Suppose initial conditions $X^{(1)}(0)$ and $X^{(2)}(0)$ satisfy the inequality $X_i^{(1)}(0) \leq X_i^{(2)}(0)$ for all i. Then almost surely, $X_i^{(1)}(t) \leq X_i^{(2)}(t)$ for all i and for all $t > 0$.

Setting $\rho(u) := \sqrt{u}$, the SDE (2.15) satisfies the conditions of the Geiß-Manthey theorem, except that our domain for both a_i and σ_i is Λ, not $\mathbb{R}^n$. We address these two coefficients in turn.

First we deal with the drift coefficient a_i, using a standard localization argument already implicit in the proof of Geiß and Manthey. They (implicitly) define the stopping time ϑ_N to be the first time $\|X^{(1)}\| \geq N$ or $\|X^{(2)}\| \geq N$, and use the fact that a is Lipschitz on the restricted domain $\|X\| \leq N$ to show that

$$\mathbb{P}\left[X_i^{(1)}(t) \leq X_i^{(2)}(t) \text{ for all } 0 \leq t \leq \vartheta_N\right] = 1.$$

Since strong solutions exist for all time, we have $\vartheta_N \to \infty$ as $N \to \infty$ almost surely, which proves the theorem. We modify this strategy for our SDE (2.15) in the standard way: Define the stopping time $\tau_{1/m}$ to be the first time either $\lambda^{(1)}$ or $\lambda^{(2)}$ leaves the set

$$\Lambda_{1/m} := \{\lambda \in \Lambda : |\lambda_i - \lambda_{i+1}| > 1/m \text{ for all } 1 \leq i \leq n-1.\}.$$

Since strong solutions starting in Λ stay in Λ for all $t \geq 0$ and are continuous, we have $\tau_{1/m} \to \infty$ as $m \to \infty$ almost surely. Since our a is Lipschitz on $\Lambda_{1/m}$, the proof of Theorem 2.6.1 shows that

$$\mathbb{P}\left[\lambda_i^{(1)}(t) \leq \lambda_i^{(2)}(t) \text{ for all } 0 \leq t \leq \tau_{1/m}\right] = 1$$

for all m. Taking $m \to \infty$, the result follows.

Finally, we address the diffusion coefficient $\sigma_i(\lambda) = 2\sqrt{\lambda_i}/n$. The standard fix is to first modify the SDE to have diffusion coefficients $2\sqrt{|\lambda_i|}/n$ for all i, so that the domain of σ_i is enlarged to $\mathbb{R}^n$ and Theorem 2.6.1 may be applied. For this modified SDE, note that the constant zero function $\lambda^{(1)}(t) = 0$ is a strong solution. Now let $\lambda^{(2)}$ be any solution with $\lambda_i^{(2)}(0) \geq 0$ for all i. Applying Theorem 2.6.1 to $\lambda^{(1)}$ and $\lambda^{(2)}$, we conclude that in fact, $\lambda^{(2)}(t) \geq 0$ for all $t \geq 0$. Thus, the absolute value bars in the modified SDE can be removed a posteriori. This argument is used, for example, when setting up the SDE for the so-called *Bessel process*, which shares this square-root diffusion coefficient—see [RY99, §XI.1] for details.

Chapter 3

Point Spectrum of the Universal Cover

3.1 Introduction

Consider a finite graph $G = (V, E)$ and its universal cover $\mathcal{T} = (\mathcal{V}, \mathcal{E})$, together with a covering map $\Xi : \mathcal{T} \to G$. The purpose of this paper is to relate the point spectrum of certain local, periodic, self-adjoint operators on $\ell^2(\mathcal{V})$ to the combinatorial structure of G.

Precise definitions, notation and assumptions will be discussed below in Section 3.2, but for now we give a high-level overview of the problem setting. By endowing G with edge weights and a potential on its vertex set, we obtain a natural self-adjoint operator A_G on $\ell^2(V)$. This framework encompasses Schrödinger operators on graphs, weighted adjacency matrices, graph Laplacians and transition matrices for random walks, and the corresponding pull-back of the weights and potential to $\mathcal{T}$ induces an analogous periodic, self-adjoint operator $A_{\mathcal{T}}$ on $\ell^2(\mathcal{V})$.

The class of operators $A_{\mathcal{T}}$ obtained in this way contains, but is richer than, the periodic Schrödinger operators in one dimension, which are of great relevance to spectral theory and the theory of orthogonal polynomials. The spectra of these $A_{\mathcal{T}}$ are additionally crucial to the study of relative expanders [F$^+$03] and, as shown in [BC19], control in a strong sense the spectra of large random lifts of a fixed base graph. However, despite the many advances in functional analysis, operator algebras and operator theory, many natural questions regarding the spectral properties of $A_{\mathcal{T}}$ remain unanswered and seem inaccessible with current techniques. We direct the reader to [ABS20] for a survey of both periodic Jacobi matrices and the difficulty in generalizing to the more generic case considered here.

In this paper we will be concerned with the spectrum of $A_{\mathcal{T}}$, which we denote by $\operatorname{Spec} A_{\mathcal{T}}$, the density of states μ (a natural and canonical measure on $\operatorname{Spec} A_{\mathcal{T}}$), and most importantly those $\lambda \in \operatorname{Spec} A_{\mathcal{T}}$ for which there exists a corresponding ℓ^2 eigenvector—in other words, the point spectrum $\operatorname{Spec}_p A_{\mathcal{T}}$.

Our main result is a set of necessary and sufficient condition on G (including its edge weights and potential) for $\mathrm{Spec}_p A_{\mathcal{T}}$ to be non-empty. This gives a finite algorithm to compute $\mathrm{Spec}_p A_{\mathcal{T}}$ from G, and extends the work of Aomoto, who has already shown the necessary half of our result in [Aom91]. However, our new and elementary argument is essentially different from his, and we build on it to show, surprisingly, that $\mathrm{Spec}_p A_{\mathcal{T}} \subset \mathrm{Spec}\, A_G$, and to give a lower bound for the multiplicity of each eigenvalue of A_G arising in this way. Finally, we prove that the set of edge weights and potentials for which $A_{\mathcal{T}}$ has point spectrum is a closed semialgebraic set of large codimension, which in particular implies that the set has Lebesgue measure zero. This may be regarded as a spectral delocalization result of the kind long-studied in mathematical physics [Ana18]; see [AISW20] for recent and analogous work in the context of quantum graphs. In particular, our result implies that even when $\mathrm{Spec}_p A_{\mathcal{T}}$ has an isolated point, there are arbitrarily small perturbations of $A_{\mathcal{T}}$ with no point spectrum at all. In view of the Kato-Rellich theorem on stability of the discrete point spectrum, this is a strong manifestation of the fact that the eigenspaces of $A_{\mathcal{T}}$ are infinite-dimensional.

Related Work

The operators $A_{\mathcal{T}}$ defined here have been studied by several authors with different motivations and levels of generality, and are variously referred to as *operators of nearest-neighbor type* [Aom91], *connected, local, pull-back operators* [AFH15] or *periodic Jacobi operators* [ABS20]; we will use the latter terminology. When G is an unweighted d-regular graph (making A_G is its adjacency matrix), classical work of Kesten in the context of Cayley graphs [Kes59], and McKay in the context of random graphs [McK81], proved that $\mathrm{Spec}\, A_{\mathcal{T}} = [-2\sqrt{d-1}, 2\sqrt{d-1}]$ and that μ follows what is now called the Kesten-Mackay distribution with parameter d. When G is an unweighted (c,d)-bireguar bipartite graph with $c < d$, Godsil and Mohar showed that $\mathrm{Spec}\, A_{\mathcal{T}} = \{0\} \cup \{\lambda \in \mathbb{R} : \sqrt{d-1} - \sqrt{c-1} \leq |\lambda| \leq \sqrt{c-1} + \sqrt{d-1}\}$ and that $\mu\{0\} = \frac{d-c}{d+c}$ [GM88]. These results imply that in the unweighted case, when G is d-regular $A_{\mathcal{T}}$ has no point spectrum, while when G is (c,d)-biregular and bipartite, $\mathrm{Spec}_p A_{\mathcal{T}} = \{0\}$.

Subsequent work focused on the properties of $\mathrm{Spec}\, A_{\mathcal{T}}$ and μ, and their relation to A_G, without making any assumptions on G; see [A$^+$88, Aom91, Sun92, SS92] as well the more recent [AFH15, BC19, ABS20, GVK19]. Of relevance for the current paper is a result of Avni, Breuer and Simon in [ABS20], which states that for any G, any edge weights, and any potential, the operator $A_{\mathcal{T}}$ has no singular continuous spectrum. As a corollary one can deduce that the continuous part of $\mathrm{Spec}\, A_{\mathcal{T}}$ always consists of a finite union of closed non-degenerate intervals, and its singular part is the finite set of eigenvalues $\mathrm{Spec}_p A_{\mathcal{T}}$. Equivalently, μ can be decomposed into a measure that is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}$ and a finite sum of atomic measures.

The most noteworthy prior result regarding the point spectrum of $A_{\mathcal{T}}$ is the aforementioned work of Aomoto, who in addition to deriving necessary conditions for the presence of point spectrum of $A_{\mathcal{T}}$ deduced a remarkable formula relating $\mu\{\lambda\}$ to the combinatorial structure of G for every $\lambda \in \mathrm{Spec}_p A_{\mathcal{T}}$. He then used these results to show that when G is a d-regular graph, regardless of the edge weights and potential, $A_{\mathcal{T}}$ has no point spectrum.

This generalizes the case when G is a cycle, which was established by different authors in the mathematical physics literature; see Section 2 of [ABS20] for a discussion and survey. In a different context, Keller, Lenz and Warzel [KLW13] showed that adjacency matrices of certain trees have no point spectrum and that this property is stable under small perturbations of the potential. For our setting, their results imply that if G has a loop at every vertex and A_G is the adjacency matrix of G, then $A_{\mathcal{T}}$ has no point spectrum.

Our results, stated in Section 3.3 after the preliminary material in Section 3.2, recover many of the ones above and provide a pleasant unification and generalization of the literature on point spectra.

3.2 Preliminaries

Graphs and Covers

We will work in the general setting of weighted graphs with self-loops and multi-edges. In this setup we will regard a graph as a tuple $G = (V, E, a, b)$, consisting of vertices, edges, edge-weights $a : E \to \mathbb{C}$ and a potential $b : V \to \mathbb{C}$. When it is not clear from the context, we will write $V(G)$ and $E(G)$ to emphasize that we are referring to the set of vertices and edges of G. It will be convenient to regard E as a set of directed edges, equipped with a direction-reversing involution $e \mapsto \check{e}$ with no fixed points, as well as *source* and *terminal* maps $\sigma, \tau : E \to V$ so that $\sigma(e) = \tau(\check{e})$ for every $e \in E$. An edge for which $\sigma(e) = \tau(e)$ and $e = \check{e}$ is a *self-loop*, and we refer to the remainder as *proper edges*.[1] We will also abuse notation and write $\sigma(u)$ and $\tau(u)$ for the sets of directed edges whose source and terminal, respectively, are the vertex $u \in V$.

We say that a graph H *covers* G if there exists a *covering map* $\xi : H \to G$, namely a map of vertices and edges which is compatible with the source and terminal maps, preserves potential and edge weights, and is an isomorphism on $\sigma(u)$ and $\tau(u)$ for each vertex u. If both are finite, then $|V(G)|$ necessarily divides $|V(H)|$, and we call their ratio n the degree of the cover; equivalently we say that H is an *n-lift* of G. Each n-lift H may be expressed explicitly by an assignment of permutations to edges $\pi : E \to \mathfrak{S}_n$, with the property that $\pi_e^{-1} = \pi_{\check{e}}$ for each edge $e \in E$. Then $V(H) = V(G) \times [n]$—throughout the paper we will use the notation $[n] = \{1, ..., n\}$—and for every $e \in E(G)$ and each $i \in [n]$, we include an edge $\tilde{e} \in E(H)$ with $\sigma(\tilde{e}) = (\sigma(e), i)$ and $\tau(\tilde{e}) = (\tau(e), \pi_e(i))$.

The *universal cover* of a connected graph G is the unique (up to isomorphism) infinite tree $\mathcal{T} = (\mathcal{V}, \mathcal{E}, \mathscr{a}, \mathscr{b})$ that covers every other cover of G. It can be constructed directly in terms of *non-backtracking walks* on G, which are sequences of edges $e_1, e_2, ...e_\ell$ such that, for every $s \in [\ell - 1]$, $\tau(e_s) = \sigma(e_{s+1})$ and $e_s \neq \check{e}_{s+1}$. If we choose a root vertex $u \in V$, then we may set the vertex set $\mathcal{V}$ of $\mathcal{T}$ to be the set of non-backtracking walks on G starting at u, with directed edges $\mathcal{E}$ whenever one walk is an immediate prefix or suffix of another, and

[1] Some authors additionally include so-called *half-loops*, which are edges e with $\sigma(e) = \tau(e)$ and $e = \check{e}$; see [Fri93]. Our results easily extend to this case, but for simplicity we will not consider it here.

edge weights and potential inherited from the final edge and vertex of the walk, respectively. Up to isomorphism $\mathcal{T}$ is independent of the root choice, and is manifestly a cover of G; we will call the covering map Ξ. Finally, we note that $\mathcal{T}$ is finite if and only if G is *acyclic*—that is, if it does not contain a closed non-backtracking walk. In this case $G = \mathcal{T}$.

Given G and a universal cover $\mathcal{T}$, the latter is endowed with a set of symmetries which act transitively on $\mathcal{V}$ by simultaneously permuting the fibres over every vertex.

Jacobi Operators, Spectra, and the Density of States

Following the convention introduced in [ABS20], the *Jacobi operator* associated to $G = (V, E, a, b)$ acts on $\eta \in \ell^2(V) \simeq \mathbb{C}^{|V|}$ as:

$$(A_G \eta)(u) = b_u \eta(u) + \sum_{e \in \tau(u)} a_e \eta(\sigma(e)). \tag{3.1}$$

Throughout the paper, we will assume that the edge weights satisfy a conjugate symmetry condition $\overline{a_e} = a_{\check{e}}$ and that the potential b is real—these ensure that A_G is Hermitian, and we will accordingly call such edge weights and potential Hermitian as well.

When H is an n-lift of G, we will always think of A_H as acting on $\ell^2(V) \otimes \mathbb{C}^n$, regarded as the set of $\mathbb{C}^n$-valued functions on V. A standard result characterizes the spectrum of A_H. Let $\pi : E \to \mathfrak{S}_n$ be the set of permutations which define H, and overload notation to write π_e as well for the unitary operator which acts by accordingly permuting the coordinates of $\mathbb{C}^n$. Then A_H acts on $\eta \in \ell^2(V) \otimes \mathbb{C}^n$ as

$$(A_H \eta)(u) = b_u \eta(u) + \sum_{e \in \tau(u)} a_e \pi_e \eta(\sigma(e)) \in \mathbb{C}^n.$$

By writing $\mathbb{C}^n = \mathbb{C}^n_1 \oplus \mathbb{C}^n_0$, where $\mathbb{C}^n_1 \simeq \mathbb{C}$ is the span of the all-ones vector and $\mathbb{C}^n_0 \simeq \mathbb{C}^{n-1}$ is its orthogonal complement, we can simultaneously decompose every edge permutation as $\pi_e = 1 \oplus \rho(\pi_e)$; the latter unitary operator on $\mathbb{C}^n_0$ is the *regular representation* of π_e. Thus we may write

$$A_H = A_G \oplus A_{H/G}, \tag{3.2}$$

where the second acts on $\eta \in \ell^2(V) \otimes \mathbb{C}^n_0$ in the natural way:

$$(A_{H/G} \eta)(u) = b_u \eta(u) + \sum_{e \in \tau(u)} a_e \rho(\pi_e) \eta(\sigma(e)) \in \mathbb{C}^n_0. \tag{3.3}$$

In other words,

$$\operatorname{Spec} A_H = \operatorname{Spec} A_G \sqcup \operatorname{Spec} A_{H/G}, \tag{3.4}$$

and we refer to $\operatorname{Spec} A_{H/G}$ as the *new eigenvalues* of H.

Once again writing $\mathcal{T} = (\mathcal{V}, \mathcal{E}, \mathscr{a}, \mathscr{b})$ for the universal cover of G, we will call the analogous operator on $\ell^2(\mathcal{V})$ the *periodic Jacobi operator* of $\mathcal{T}$. Since the edge weights $\mathscr{a}$ and potential

$\mathcal{b}$ are related to those of G by $\mathcal{a}_{\tilde{e}} = a_{\Xi(\tilde{e})}$ for every $\tilde{e} \in \mathcal{E}$ and $\mathcal{b}_{\tilde{v}} = b_{\Xi(\tilde{v})}$ for every $v \in \mathcal{V}$, finiteness of G, a, and b ensure that $A_{\mathcal{T}}$ belongs to the set $\mathcal{B}(\ell^2(\mathcal{V}))$ of bounded operators on $\ell^2(\mathcal{V})$, and the inherited conjugate symmetry condition $\overline{\mathcal{a}_e} = \mathcal{a}_{\check{e}}$ guarantees that it is Hermitian. We use $\operatorname{Spec} A_{\mathcal{T}}$ to denote the *spectrum* of the periodic Jacobi operator, that is

$$\operatorname{Spec} A_{\mathcal{T}} = \left\{ \lambda \in \mathbb{R} : (\lambda - A_{\mathcal{T}})^{-1} \notin \mathcal{B}(\ell^2(\mathcal{V})) \right\}. \tag{3.5}$$

We remind the reader that, unlike in the finite dimensional case, $\lambda \in \operatorname{Spec} A_{\mathcal{T}}$ does not guarantee an ℓ^2 eigenvector for λ. The subset of the spectrum with this additional property—the *point spectrum*—will be our primary concern in this work. We will return to it below.

For any $u \in V(G)$, the quantities $\langle \delta_{\tilde{u}}, A_{\mathcal{T}}^{\ell} \delta_{\tilde{u}} \rangle$ for $\ell \in \mathbb{N}$ are real and constant over all $\tilde{u}$ in the fibre over u, and a routine application of the Riesz representation theorem guarantees an accompanying *spectral measure* μ_u on $\operatorname{Spec} A_{\mathcal{T}}$ associated to each u, satisfying

$$\langle \delta_{\tilde{u}}, f(A_{\mathcal{T}}) \delta_{\tilde{u}} \rangle = \int_{\operatorname{Spec} A_{\mathcal{T}}} f(x) \, \mathrm{d}\mu_u(x) \qquad \forall \tilde{u} \in \Xi^{-1}(u) \tag{3.6}$$

for every bounded measurable function $f : \operatorname{Spec} A_{\mathcal{T}} \to \mathbb{C}$, where $f(A_{\mathcal{T}})$ is defined via the Borel functional calculus. The *density of states* of $A_{\mathcal{T}}$ is the unique measure obtained by averaging these spectral measures over $u \in V(G)$:

$$\mu = \frac{1}{|V(G)|} \sum_{u \in V(G)} \mu_u. \tag{3.7}$$

It is typical in the literature to work with real positive edge weights instead of Hermitian ones. The latter choice will make some of our proofs more convenient, but from the perspective of $\operatorname{Spec} A_{\mathcal{T}}$ it does not add any generality. In particular, $A_{\mathcal{T}}$ is *gauge invariant* in the following sense.

Lemma 3.2.1. *Let $G = (V, E, a, b)$ be a graph with Hermitian edge weights and real potential, and let $G' = (V, E, a', b)$, where $a'_e = |a_e|$; write $\mathcal{T}$ and $\mathcal{T}'$ for their respective covers. Then $A_{\mathcal{T}} = U^* A_{\mathcal{T}'} U$, where U is a diagonal unitary operator.*

One may prove Lemma 3.2.1 by choosing a root r for $\mathcal{T}$ and letting $U_{v,v}$ be the product of edge weight arguments along the unique shortest path connecting r and v. An immediate corollary is that *the spectrum and density of states of $A_{\mathcal{T}}$ depend only on the moduli of the edge weights*. In the case $b \equiv 0$, note that the above implies that $A_{\mathcal{T}}$ and $-A_{\mathcal{T}}$ are unitarily equivalent, which has the following consequence.

Lemma 3.2.2. *Let G be a finite graph, $\mathcal{T}$ its universal cover. If $b_v = 0$ for all $v \in V(G)$ then the spectrum and density of states of $A_{\mathcal{T}}$ are symmetric about zero.*

On several occasions we will use the following well-known facts to relate the empirical spectral measures of finite graphs G to the densities of states of their universal covers.

Lemma 3.2.3. *Let G be a finite graph, $\mathcal{T}$ its universal cover, and μ the density of states of $A_{\mathcal{T}}$. There exists a sequence of covers G_n of G whose girth[2] diverges as n goes to infinity.*

[2] The girth of a graph is the length of its shortest cycle.

Moreover, for this sequence, the empirical spectral measures μ_n of A_{G_n} converges weakly to μ.

This lemma follows directly from results stated in [ABS20, Section 4] whose proofs will appear in [ABKSon]; their approach is to construct the sequence of lifts from a tower of normal subgroups of the group of deck transformations of $\mathcal{T}$. For completeness, below we provide a purely combinatorial proof.

Proof of Lemma 3.2.3. By induction, it suffices to show that for every finite graph $H = (V, E)$ with girth$(H) = p$ and $|E| = 2m$, there exists a finite lift L of H whose girth is strictly larger (the weights and potential are irrelevant here, and we will suppress them). We will construct L as a 2^{m+1}-lift of H, with the following set of permutations $\pi : E \to \mathfrak{S}_{2^m+1}$. Group the edges in to pairs $(e, \check{e})$ consisting of an edge and its reversal, and order these $(e_1, \check{e}_1), ..., (e_m, \check{e}_m)$. Now let π_{e_i} be the permutation that maps $j \mapsto j + 2^i (\mod 2^{m+1})$ for every $j \in [2^{m+1}]$, and let $\pi_{\check{e}_i} = \pi_{e_i}^{-1}$ as required.

Since $\operatorname{girth} L > \operatorname{girth} H$, we need only to show that L contains no cycle of length p. Seeking contradiction, assume instead that $e_{i_1}, ..., e_{i_p}$ is a sequence of p directed edges forming a cycle in L. Writing $\xi : L \to H$ for the covering map, the edges $\xi(e_{i_1}), ..., \xi(e_{i_p})$ form a cycle in H with length p, and since $\operatorname{girth} H = p$, they are distinct. The vertices of L are $V \times [2^{m+1}]$, which we regard as a set of 2^{m+1} 'layers;' assume $\sigma(e_{i_1})$ is in the tth one. Because of how we have arranged the permutations, $\tau(e_{i_p})$ is in layer $t \pm 2^{i_1} \pm \cdots \pm 2^{i_p} \neq t (\mod 2^{m+1})$, because the $i_1, ..., i_p$ are distinct and smaller than $m + 1$. Thus $\tau(e_{i_p}) \neq \sigma(e_{i_1})$, a contradiction.

We finally show that, given such a sequence G_n with diverging girth, the empirical spectral measures μ_n converge weakly to μ. For every fixed positive integer k and each vertex u of G_n, the quantity $\langle \delta_u, A^k_{G_n} \delta_u \rangle$ is a weighted count of length-k closed walks in G_n starting and ending at u. Since the G_n have diverging girth, for n sufficiently large the depth-k neighborhood of u in G_n is identical to that of every $\tilde{u} \in \Xi^{-1}(u)$, and thus this count is eventually constant and equal to $\langle \delta_{\tilde{u}}, A^k_{\mathcal{T}} \delta_{\tilde{u}} \rangle$. Finally, as the kth moments of the empirical spectral measures μ_n are given by normalized traces of $A^k_{G_n}$, the method of moments gives us weak convergence to the density of states. □

Substantially stronger versions of this result are known but will not be necessary for us; we direct the reader for instance to the recent work of Bordenave and Collins [BC19].

Point Spectrum and the Aomoto Sets

We will denote the point spectrum of $A_{\mathcal{T}}$ as

$$\operatorname{Spec}_p A_{\mathcal{T}} = \{\lambda \in \mathbb{R} : \operatorname{Ker}(\lambda - A_{\mathcal{T}}) \neq \{0\}\}. \tag{3.8}$$

The following proposition collates several equivalent characterizations of $\operatorname{Spec}_p A_{\mathcal{T}}$.

Proposition 3.2.4. *Let G be a finite graph with at least one cycle, $\mathcal{T}$ its universal cover. Then $\lambda \in \operatorname{Spec}_p A_{\mathcal{T}}$ if and only if any of the following hold:*

(i) $\dim \operatorname{Ker}(\lambda - A_{\mathcal{T}}) = \infty$

(ii) λ *is an atom of* μ

(iii) *For some* $u \in V(G)$, λ *is an atom of* μ_u.

(iv) *For some* $u \in V(G)$, *the Cauchy transform*

$$S_u(z) = \int_{\operatorname{Spec} A_{\mathcal{T}}} (z-x)^{-1} \, \mathrm{d}\mu_u(x)$$

has a pole at λ.

(v) *For some* $u \in V(G)$, *and every* $\tilde{u} \in \Xi^{-1}(u)$, *there exists* $\zeta \in \operatorname{Ker}(\lambda - A_{\mathcal{T}})$ *with* $\zeta(\tilde{u}) \neq 0$.

Moreover, the vertices satisfying (iii),(iv), and (v) coincide.

Proof. In Section 8 of [ABS20] it is proven that $\lambda \in \operatorname{Spec}_p A_{\mathcal{T}}$ implies (i). Intuitively, if $\xi \in \operatorname{Ker}(\lambda - A_{\mathcal{T}})$, then precomposing ξ with deck transformations of $\mathcal{T}$ gives rise to many linearly independent eigenvectors of $A_{\mathcal{T}}$ for λ.

The rest of the claims are standard and hold for general bounded, self-adjoint operators. Equivalence of $\lambda \in \operatorname{Spec}_p A_{\mathcal{T}}$, (ii), (iii), and (v) follows directly from the definition of μ and the forthcoming Lemma 3.2.7. On the other hand (iii) and (iv) are clearly equivalent. □

By way of a complicated set of coupled equations satisfied by the Cauchy transforms S_u, Aomoto identified a set of vertices of G whose combinatorial structure is instrumental in understanding $\operatorname{Spec}_p A_{\mathcal{T}}$ and will be the focus of much of this paper.

Definition 3.2.5 (The Aomoto set)**.** *Let* G *be a finite graph and assume that* $\lambda \in \operatorname{Spec}_p A_{\mathcal{T}}$. *The Aomoto set of* G *associated to* λ *consists of those vertices in* $V(G)$ *that satisfy the equivalent conditions (iii-v) in Proposition 3.2.4. This set will be denoted by* $X_\lambda(G)$.[3]

Example 3.2.6. *In Figure 3.2,* G *is a finite graph with* $a_e \equiv 1$ *and* $b_v \equiv 0$ *for every* $e \in E(G)$ *and* $v \in V(G)$. *By the symmetries in* G, *the spectral measures corresponding to the vertices* u_1, u_2, u_4 *and* u_5 *are equal. Hence,* $A_{\mathcal{T}}$ *has only two distinct spectral measures associated to the vertices of* G, *sketched in Figure 3.2. These sketches were generated by taking a random lift of* G *of degree 1200 and by plotting the weighted histogram for the corresponding discrete spectral measures. As the figures show, the spectral measures corresponding to* u_1, u_2, u_4 *and* u_5 *have atoms at* -1 *and 1, while the spectral measure corresponding to* u_3 *is absolutely continuous with respect to the Lebesgue measure on* $\mathbb{R}$. *Hence,* $X_{-1}(G) = X_1(G) = \{u_1, u_2, u_4, u_5\}$. *Note that the subgraph induced by* $X_1(G)$ *consists of two disconnected trees; later in Theorem 3.3.1 we will show that this a general property of Aomoto sets.*

[3]This set was referred to as $X_\lambda^{(1)}(G)$ in [Aom91] and [ABS20]; we have dropped the superscript to lighten notation, and because we will not consider the sets $X_\lambda^{(\alpha)}(G)$ for $\alpha \neq 1$ which appear in that work.

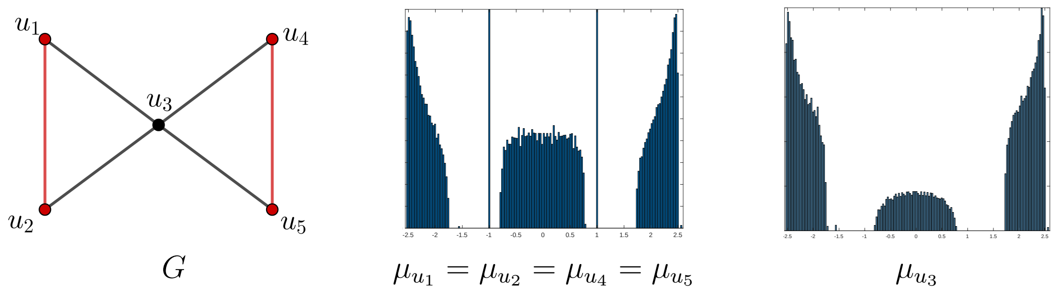

Figure 3.1: On the left, a finite graph G, where the vertices in $X_1(G) = X_{-1}(G)$ are colored in red. In the middle and on the right, we show the two distinct spectral measures of $A_\mathcal{T}$ associated to the vertices of G, the center one having atoms at ± 1.

We will use repeatedly an equivalent form of Proposition 3.2.4 (v) above: *if $u \notin X_\lambda(G)$, then any eigenvector $\eta \in \mathrm{Ker}(\lambda - A_\mathcal{T})$ is identically zero on the fibre over u.* We will also require a standard identity expressing the mass assigned to $\lambda \in \mathrm{Spec}_p\, A_\mathcal{T}$ by the spectral measure μ_u.

Lemma 3.2.7. *Let G be a finite graph, $\mathcal{T}$ its universal cover, and $\lambda \in \mathrm{Spec}_p\, A_\mathcal{T}$. Then if $\mathfrak{B}$ is any orthonormal basis for $\mathrm{Ker}(\lambda - A_\mathcal{T})$, for any $u \in V$ and $\tilde{u} \in \Xi^{-1}(u)$,*

$$\mu_u\{\lambda\} = \sum_{\eta \in \mathfrak{B}} |\eta(\tilde{u})|^2. \tag{3.9}$$

Equation (3.9) follows from a standard application of the Borel functional calculus, where the key observation is that if $f_\lambda : \mathrm{Spec}\, A_\mathcal{T} \to \mathbb{R}$ is the indicator function of the singleton $\{\lambda\}$ then $f_\lambda(A_\mathcal{T})$ is the orthogonal projection onto $\ker(\lambda - A_\mathcal{T})$.

3.3 Main Results

Our first contribution is to strengthen a result of Aomoto [Aom91], by way of a new and more conceptual proof. This result characterizes the induced subgraph on $X_\lambda(G)$ for any $\lambda \in \mathrm{Spec}_p\, A_\mathcal{T}$, and relate the mass $\mu\{\lambda\}$ to the local structure of this subgraph and neighboring vertices. Let us write $\partial X_\lambda(G)$ for the set of vertices outside the Aomoto set but connected to it by an edge, $\mathsf{cc}X_\lambda(G)$ for the number of connected components of the subgraph induced by $X_\lambda(G)$, and define the *index* of λ as

$$I_\lambda(G) = \mathsf{cc}X_\lambda(G) - |\partial X_\lambda(G)|. \tag{3.10}$$

Recall that for us a graph $G = (V, E, a, b)$ contains vertices V, directed edges E, Hermitian edge weights $a : E \to \mathbb{C}$ satisfying $\overline{a_e} = a_{\check{e}}$ and real potential $b : V \to \mathbb{R}$.

Theorem 3.3.1. *Let G be a finite graph, $\mathcal{T}$ its universal cover, and $\lambda \in \mathrm{Spec}_p\, A_\mathcal{T}$. Then:*

(i) The subgraph induced by $X_\lambda(G)$ is acyclic,

(ii) λ is an eigenvalue, with multiplicity one, of the induced Jacobi operator of each connected component of this subgraph, and

(iii) The density of states of $A_{\mathcal{T}}$ satisfies

$$\mu\{\lambda\} = \frac{I_\lambda(G)}{|V(G)|} \tag{3.11}$$

Assertion (i), claimed without proof in [ABS20], clarifies an ambiguity in Aomoto's result, which did not rule out self-loops or multi-edges in the subgraph induced by $X_\lambda(G)$; (ii) is a new observation, and (iii) is due to Aomoto. Our new proof is combinatorial and linear algebraic, using properties of eigenvectors of Jacobi operators on finite and infinite trees; the question of finding an alternative to Aomoto's original proof explaining the significance of the quantity $I_\lambda(G)$, was posed in [ABS20, Problem 8.1]. The proofs of (i) and (ii) may be found in Section 3.4, and that of (iii) in Section 3.5.

We then build on the proof of Theorem 3.3.1 to prove a number of novel results. First, we show that for any graph G, the point spectrum of the periodic Jacobi operator on its unversal cover is contained in $\operatorname{Spec} A_G$—with multiplicity bounded in terms of the index $I_\lambda(G)$. In fact, we can further refine this result for the Jacobi operator of any cover H of G.

Theorem 3.3.2. *Let G be a finite graph, H an n-lift of G, and $\mathcal{T}$ their common universal cover. If $\lambda \in \operatorname{Spec}_p A_{\mathcal{T}}$, then*

(i) $\lambda \in \operatorname{Spec} A_G$ with multiplicity at least $|V(G)| \cdot \mu\{\lambda\}$, and

(ii) $\lambda \in \operatorname{Spec} A_{H/G}$ with multiplicity at least $(n-1)|V(G)| \cdot \mu\{\lambda\}$,

so that the multiplicity of $\lambda \in \operatorname{Spec} A_H$ is at least $n|V(G)| \cdot \mu\{\lambda\}$.

We will show at the end of Section 3.3 that these lower bounds on multiplicity need not be tight.

Additionally, we prove a converse to Theorem 3.3.1, namely that if a graph has a set replicating the structure of the Aomoto set for some λ, then its universal cover has λ in its point spectrum. To be precise, let us extend the notation ∂ and cc to apply to any $X \subset V(G)$. For each $\lambda \in \mathbb{R}$, let $\mathcal{A}_\lambda(G)$ be the set of all subsets $X \subset V(G)$ inducing an acyclic subgraph, each connected component of which has λ in the spectrum of its induced Jacobi operator and such that $\mathsf{cc}(X) - |\partial X| > 0$.

Theorem 3.3.3. *Let G be a finite graph, and $\mathcal{T}$ its universal cover. For any $\lambda \in \mathbb{R}$ and $X \in \mathcal{A}_\lambda(G)$,*

(i) $\lambda \in \operatorname{Spec} A_G$, with multiplicity at least $\mathsf{cc}(X) - |\partial X|$.

(ii) $\lambda \in \operatorname{Spec}_p A_{\mathcal{T}}$, and $|V(G)| \cdot \mu\{\lambda\} \geq \mathsf{cc}(X) - |\partial X|$.

The lower bounds in (i) and (ii) need not be tight. For (i), this is shown at the end of Section 3.3; for (ii), when $I_\lambda(G) > 1$, one may choose X to contain only a subset of the trees in the true Aomoto set for some $\lambda \in \mathrm{Spec}_p A_{\mathcal{T}}$. Furtheremore, there are cases where the inequality in (ii) is strict for elements of $\mathcal{A}_\lambda(G)$ that are not subsets of the Aomoto set.

The proofs of both Theorem 3.3.2 and Theorem 3.3.3 follow from a generalization of the latter, Theorem 3.6.1, which we state and verify in Section 3.6. The argument proceeds, roughly, by patching together and extending the λ-eigenvectors on each component of the Aomoto set promised by Theorem 3.3.1 to global eigenvectors of A_G and $A_{H/G}$. This has the interesting consequence that if $\lambda \in \mathrm{Spec}_p A_{\mathcal{T}}$ and $G_1, G_2, \dots$ is a sequence of lifts of G, there is a constant fraction of $|V(G_n)|$ of λ-eigenvectors of A_{G_n} that are localized. In contrast, under some technical assumptions, quantum ergodicity results [AS19] imply that if λ is in the absolutely continuous part of the spectrum of $A_{\mathcal{T}}$, the number of localized λ-eigenvectors of A_{G_n} is sublinear in and $|V(G_n)|$ as $n \to \infty$.

Combining Theorems 3.3.2 and 3.3.3, we find the following corollary. Note that since there are finitely many induced subgraphs, in finite time we can find every $\lambda \in \mathbb{R}$ for which $\mathcal{A}_\lambda(G)$ is nonempty.

Corollary 3.3.4. *Let G be a finite graph, $\mathcal{T}$ it's universal cover. Then for each $\lambda \in \mathbb{R}$,*

$$\mu\{\lambda\} = \frac{1}{|V(G)|} \max_{X \in \mathcal{A}_\lambda(G)} \big(\mathsf{cc}(X) - |\partial(X)|\big). \tag{3.12}$$

Moreover, $\mathrm{Spec}_p A_{\mathcal{T}}$ *may be computed from G in finite time.*

Although Theorem 3.3.1 implies that the Aomoto set $X_\lambda(G)$ is an element in $A_\lambda(G)$ maximizing the quantity on the right side of (3.12), it can happen that there is not a unique maximizer. Figure 3.2 gives such an example.

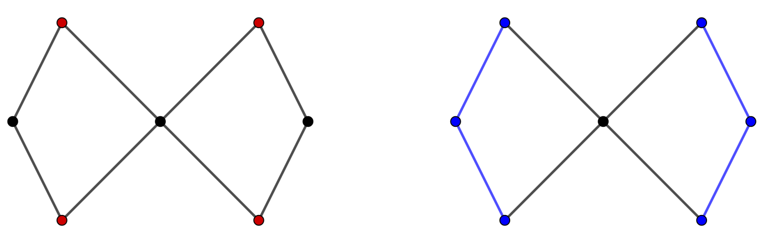

Figure 3.2: Two distinct sets of vertices (in red and blue respectively) of a graph G are shown. If $a \equiv 1$ and $b \equiv 0$, it is easy to show from Corollary 3.3.4 that $I_0(G) = 1$. Then both the red and the blue vertex set belong to $\mathcal{A}_0(G)$. It will follow from Observation 3.4.1 below that $X_0(G)$ is precisely the set indicated by the red vertices.

Finally, we use Theorem 3.3.1 and Theorem 3.3.3 to argue that point spectrum is rare in a certain sense. To make this precise fix $G = (V, E)$ and think of the set of possible Hermitian edge weights $a = (a_e)_{e \in E}$ and vertex potentials $b = (b_v)_{v \in V}$ as $\mathbb{C}^{|E|/2} \oplus \mathbb{R}^{|V|} \cong \mathbb{R}^{|E|+|V|}$.

Theorem 3.3.5. *Let $G = (V, E)$ be a finite graph with at least one cycle and $\mathcal{T}$ be its universal cover. Assume that every vertex in G has at least $d_{\min}$ distinct neighborhs. Leaving V and E fixed, let $\mathcal{P} \subset \mathbb{R}^{|E|+|V|}$ be the set of Hermitian edge weights and potentials for which $\operatorname{Spec}_p A_{\mathcal{T}} \neq \emptyset$. Then, $\mathcal{P}$ is a semialgebraic closed set of codimension at least $\max\{d_{\min} - 1, 1\}$.*[4]

Remark 3.3.6. *Even if the bound $\operatorname{codim}(\mathcal{P}) \geq \max\{d_{\min} - 1, 1\}$ is tight in general, for many specific instances a stronger bound can be obtained. We refer the reader to the discussion in Section 3.7 for a stronger bound that depends in a more complicated way on the combinatorial structure of G.*

Theorem 3.3.5 will be proved in Section 3.7 and resolves [Aom91, Question 2], which speculated that the existence of point spectrum was dependent on the combinatorial structure of G and not on the edge weights and potential. Results in a similar direction were obtained in [KLW12] and [KLW13]. Their results are less general in the sense that they require G to have edge weights $a \equiv 1$ and a loop at every vertex. However, they allow for more general potentials on the more general class of trees with finite cone type.

Theorem 3.3.5 implies in particular that $\mathcal{P}^c$ is an open dense set and hence that the point spectrum of $A_{\mathcal{T}}$ can be destroyed by adding arbitrarily small perturbations, even when $A_{\mathcal{T}}$ has isolated eigenvalues. This is surprising, given a result of Kato (see [RS78, Section XII.2]) that if H is a bounded self-adjoint operator and $\lambda \in \operatorname{Spec}_p H$ is isolated with $\dim \ker(\lambda - H) < \infty$, then every sufficiently small perturbation of H has non-empty point spectrum. Of course, this does not contradict our result, since Proposition 3.2.4 ensures that every $\lambda \in \operatorname{Spec}_p A_{\mathcal{T}}$ has an infinite-dimensional eigenspace. However, it is not the case that infinite-dimensional eigenspaces are unstable in general, and in many cases the phenomenon implied by Kato's result is still present.

Furthermore, it is an immediate consequence of Theorem 3.3.5 that $\mathcal{P}$ has Lebesgue measure zero. This can interpreted as an almost sure spectral delocalization result, since it implies that under a random absolutely continuous perturbation (with respect to the Lebesgue measure) of the edge weights and potential of G, the spectrum of $A_{\mathcal{T}}$ becomes purely absolutely continuous.

We conclude this section by giving some applications of the results presented above.

Point Spectrum of Biregular Trees

Let G be the complete bipartite graph $K_{c,d}$ for some integers $d > c$, and denote by V_c and V_d the vertex components of G having c and d vertices respectively. We will first analyze the case when $b \equiv 0$ and a is any Hermitian edge weighting. It is easy to see that V_d is a set satisfying the conditions of Theorem 3.3.3 for $\lambda = 0$, and hence that $\mu\{0\} \geq \frac{d-c}{d+c}$. Then

[4] In a previous version of this paper we only proved that $\mathcal{P}$ is a closed set of Lebesgue measure 0 by showing that $\mathcal{P}$ was contained in an algebraic set of codimension 1. We thank Barry Simon for pointing out this stronger version of the theorem and suggesting a sketch of the proof.

by Theorem 3.3.1, $X_0(G) \neq \emptyset$ and moreover $I_0(G) \geq d - c$. In the forthcoming Observation 3.4.1, we will show that that when $b \equiv 0$ (regardless of the structure of G) the set $X_0(G)$ is an independent set in G, which together with the previous observations implies that in fact $X_0(G) = V_d$. So by Theorem 3.3.1, when $b \equiv 0$, $\mu\{0\} = \frac{d-c}{d+c}$ and the vectors in $\ker(A_{\mathcal{T}})$ are supported on the fibre of V_d, which extends a result of Godsil and Mohar [GM88].

In the case of arbitrary real potential b, the existence and location of eigenvalues of $A_{\mathcal{T}}$ depend on the particular choice of b, and moreover by Theorem 3.3.5 we know that one may choose b such that $A_{\mathcal{T}}$ has no point spectrum. This discussion resolves Problems 8.6 and 8.7 posed in [ABS20]. Finally, we note in passing that in this case, when $b \equiv 0$ we have $I_0(G) = d-c$ but the multiplicity of zero in $\operatorname{Spec} A_G$ is $d+c-2$. This shows that the bounds on the multiplicity given in Theorems 3.3.2 and 3.3.3 may not be tight.

Non-isolated Point Spectrum

Let G be a finite graph and let $n = |V(G)|$. Sunada's gap labeling theorem (see[ABS20, Theorem 5.1] or [GVK19, Theorem 1.8]) states that $\operatorname{Spec} A_{\mathcal{T}}$ is a disjoint union of at most n (possibly degenerate) closed intervals typically called *bands*, and that if B is one of these bands then $\mu(B) = j/n$ for some $j \in [n]$. If $\lambda \in \operatorname{Spec}_p A_{\mathcal{T}}$ is isolated then $\{\lambda\}$ is a (degenerate) band of $\operatorname{Spec} A_{\mathcal{T}}$. This is the case, for example, when $\lambda = 0$, G is a bipartite biregular graph with components of different sizes, $a \equiv 1$ and $b \equiv 0$. Here we will show that it is possible for $0 \in \operatorname{Spec}_p A_{\mathcal{T}}$ to lie inside a non-degenerate band of $\operatorname{Spec} A_{\mathcal{T}}$. Our argument is similar in spirit to the one used at the end of Section 5 of [ABS20] for an unrelated purpose.

Set $b \equiv 0$, noting that μ is symmetric about zero from Lemma 3.2.2, and assume that $0 \in \operatorname{Spec}_p A_{\mathcal{T}}$ and that $n - I_0(G)$ is an odd number. Figure 3.3 shows two instances where these conditions are met. If $\{0\}$ is an isolated point of $\operatorname{Spec} A_{\mathcal{T}}$ then any band in $\operatorname{Spec} A_{\mathcal{T}}$ is either disjoint from $(0, \infty)$ or fully contained in this infinite interval. Hence, by Sunada's gap labeling theorem $\mu(0, \infty) = j/n$ for some integer $j \in [n]$. On the other hand, since μ is symmetric, $\mu(-\infty, 0) = j/n$. Finally, by Theorem 3.3.1 $\mu\{0\} = I_0(G)/n$. Putting all these observations together we get $1 = \mu(\mathbb{R}) = \frac{2j+I_0(G)}{n}$, which contradicts the assumption that $n - I_0(G)$ is odd.

3.4 Acyclic Nature of Aomoto Sets

In this section we will prove the first two assertions of Theorem 3.3.1, namely that if $\lambda \in \operatorname{Spec}_p A_{\mathcal{T}}$, then the Aomoto set $X_\lambda(G)$ is acyclic, and λ is an eigenvalue of the induced Jacobi operator on each of its connected components. We begin by generalizing to the infinite case a result of Fielder regarding eigenvectors of finite trees [Fie75, Proposition 1].

Lemma 3.4.1. *Let T be a locally finite tree with Hermitian edge weights a, potential b, and Jacobi operator A_T. If $\eta \in \operatorname{Ker}(\lambda - A_t)$ and $\eta(v) \neq 0$ for every vertex $v \in V(T)$, then* $\dim \operatorname{Ker}(\lambda - A_T) = 1$.

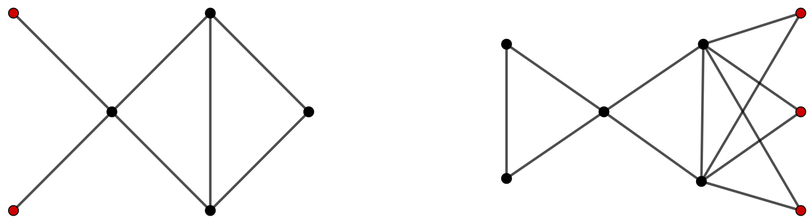

Figure 3.3: As in Section 3.3, for each graph above, a combination of Observation 3.4.1 and Theorem 3.3.3 reveals the red vertices as the Aomoto set associated to 0. In both cases $I_0(G) = 1$ while $|V(G)|$ is even.

Proof. Choose a root r for T, and for each vertex v, write $p(v)$ for its unique parent, T_v the infinite sub-tree emanating from v away from its parent and $\eta_{\geq v}$ for the restriction of η to the subtree T_v. As T is acyclic, it has no multi-edges or self-loops, and there is no ambiguity in writing $a_{v\leftarrow u}$ for the weight of the unique edge with source u and terminal v. Since $\eta_{\geq v}$ is identically zero on vertices above v, and locally satisfies the eigenvector equation at each vertex below v, we find that

$$\begin{aligned}(\lambda - A_T)\eta_{\geq v} &= \sum_{u\in V(T_v)} \lambda\eta(u)\delta_u - \sum_{u\in V(T_v)\setminus\{v\}} (A_T\eta)(u)\delta_u \\ &\qquad - \left(b_v\eta(v) + \sum_{x:p(x)=v} a_{v\leftarrow x}\eta(x)\right)\delta_v - a_{p(v)\leftarrow v}\eta(v)\delta_{p(v)} \\ &= \lambda\eta(v)\delta_v - \big((A_T\eta)(v) - a_{v\leftarrow p(v)}\eta(p(v))\big)\,\delta_v - a_{p(v)\leftarrow v)}\eta(v)\delta_{p(v)} \\ &= a_{v\leftarrow p(v)}\eta(p(v))\delta_v - a_{p(v)\leftarrow v}\eta(v)\delta_{p(v)}\end{aligned}$$

for any $v \neq r$, where δ_v as usual denotes the function which is one at v and zero elsewhere. Now, let $\zeta \in \operatorname{Ker}(\lambda - A_T)$. As A_T is self-adjoint and λ real,

$$0 = \langle \zeta, (\lambda - A_T)\eta_{\geq v}\rangle = a_{v\leftarrow p(v)}\eta(p(v))\overline{\zeta(v)} - a_{p(v)\leftarrow v}\eta(v)\overline{\zeta(p(v))},$$

which implies

$$\frac{\zeta(v)}{\zeta(p(v))} = \overline{\frac{a_{p(v)\leftarrow v}}{a_{v\leftarrow p(v)}}\frac{\eta(v)}{\eta(p(v))}}.$$

This identity holds for every $\zeta \in \operatorname{Ker}(\lambda - A_{\mathcal{T}})$, including η itself, so we obtain

$$\frac{\zeta(v)}{\zeta(p(v))} = \overline{\frac{a_{p(v)\leftarrow v}}{a_{v\leftarrow p(v)}}\frac{\eta(v)}{\eta(p(v))}} = \overline{\frac{a_{p(v)\leftarrow v}}{a_{v\leftarrow p(v)}}\overline{\frac{a_{p(v)\leftarrow v}}{a_{v\leftarrow p(v)}}\frac{\eta(v)}{\eta(p(v))}}} = \frac{|a_{p(v)\leftarrow v}|^2}{|a_{v\leftarrow p(v)}|^2}\frac{\eta(v)}{\eta(p(v))} = \frac{\eta(v)}{\eta(p(v))};$$

in the final equality we have used conjugate symmetry of the edge weights. Since $\eta|_{\geq r} = \eta \in \operatorname{Ker}(\lambda - A_{\mathcal{T}})$, ζ is unconstrained at the root, and the above equation propagates a condition down the tree that $\zeta = \eta \cdot \zeta(r)/\eta(r)$. $\square$

We now prove that the subgraph of G induced by $X_\lambda(G)$ is a forest, and that λ is an eigenvalue, with multiplicity one, of the induced Jacobi operator of each of its connected components.

Proof of Theorem 3.3.1(i-ii). Assume λ is in the point spectrum of $A_\mathcal{T}$, and let G' be a connected component of the subgraph induced by $X_\lambda(G)$. Let $\mathcal{T}'$ be the universal cover of G'. If we view $\mathcal{T}'$ as a subgraph of $\mathcal{T}$ then any vector in $\operatorname{Ker}(\lambda - A_\mathcal{T})$ vanishes on the boundary of $\mathcal{T}'$ in $\mathcal{T}$, and thus restricts to a λ-eigenvector of $\mathcal{T}'$. Hence $X_\lambda(G') = V(G')$ by Proposition 3.2.4(v). In this case, it is a consequence of Zorn's lemma that there exists $\eta \in \operatorname{Ker}(\lambda - A_{\mathcal{T}'})$ satisfying $\eta(u) \neq 0$ for every $u \in V(\mathcal{T})$; this fact appeared as [Nyl98, Lemma 7]. Combining this fact with Lemma 3.4.1, we conclude finally that $\dim \operatorname{Ker}(\lambda - A_{\mathcal{T}'}) = 1$, and thus, by Proposition 3.2.4(i), that G' is acyclic. This further implies that $\mathcal{T}' = G'$, which proves the second assertion. $\square$

In the course of the proof above we showed the following fact, which will be of repeated use throughout the paper.

Lemma 3.4.2. *Let G be a finite graph with Hermitian edge weights and potential, with $\lambda \in \operatorname{Spec}_p A_\mathcal{T}$ and $T_1, \ldots, T_p$ the Aomoto trees of G associated to λ. Then, for every $i \in [p]$ there is a unique (up to phase) unit vector $\zeta_i \in \operatorname{Ker}(\lambda - A_{T_i})$ satisfying $\zeta(u) \neq 0$ for every $u \in V(T_i)$.*

For use in the next section, we record one consequence of the above lemma.

Observation 3.4.1. Let G be a graph with $b \equiv 0$, $\mathcal{T}$ its universal cover, and assume $0 \in \operatorname{Spec}_p A_\mathcal{T}$. Then $X_0(G)$ is an independent set in G.[5]

Proof. By Lemma 3.4.2, each Aomoto tree of G must have a unique, everywhere nonzero eigenvector in the kernel of its Jacobi operator. On the other hand, a vector in the kernel of a Jacobi operator for a tree with potential zero cannot be nonzero at the parent of a leaf. Thus every Aomoto tree of G is an isolated vertex as desired. $\square$

3.5 Aomoto's Index Formula

In this section we complete the proof of Theorem 3.3.1 by verifying the formula in equation (3.11): if $\lambda \in \operatorname{Spec}_p A_\mathcal{T}$, then

$$|V(G)| \cdot \mu\{\lambda\} = I_\lambda(G).$$

Our strategy will be to reduce the problem to the proof of an analogous result on an auxiliary bipartite graph G'.

[5]By an independent set in G we mean a set $X \subset V(G)$ which induces a subgraph with no edges.

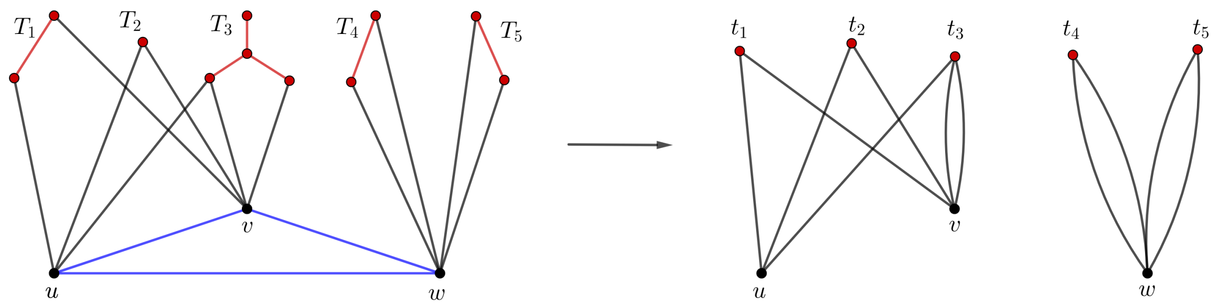

Figure 3.4: On the left an example of a graph G with Aomoto trees in red. On the right its auxiliary graph G', where each tree T_i has been contracted into a vertex t_i and the blue edges have been removed.

Constructing the Auxilliary Graph

Let $T_1, \ldots, T_p$ be the Aomoto trees of $G = (V, E, a, b)$ associated to λ, write $\mathcal{F}_i$ for the set of disjoint copies of T_i in $\mathcal{T} = (\mathcal{V}, \mathcal{E}, \mathscr{a}, \mathscr{b})$ obtained by lifting T_i, and let $\mathcal{F} = \bigcup_{i=1}^{p} \mathcal{F}_i$. Note that $\mathcal{F}$ is a subforest of $\mathcal{T}$ and all of its subtrees are isomorphic to some Aomoto tree of G.

By Lemma 3.4.2, for each T_i there is a unique (up to phase) vector $\zeta_i \in \mathrm{Ker}(\lambda - A_{T_i})$ with unit norm and nonzero entries. Take any $\eta \in \mathrm{Ker}(\lambda - A_{\mathcal{T}})$. For every $S \in \mathcal{F}$, by definition of the Aomoto set it holds that η is zero on all vertices in $\partial V(S)$. Hence, the restriction of η to any $S \in \mathcal{F}_i$ induces an eigenvector of A_{T_i}. This implies that η can be decomposed as

$$\eta = \sum_{S \in \mathcal{F}} \alpha_S \zeta_S, \tag{3.13}$$

where $\alpha_S \in \mathbb{C}$ are coefficients and the $\zeta_S \in \ell^2(V(\mathcal{T}))$ are inclusions of the λ-eigenvectors of each Aomoto tree:

$$\zeta_S(v) = \begin{cases} \zeta_i(\Xi(v)) & \text{if } v \in V(S) \text{ and } S \in \mathcal{F}_i \\ 0 & \text{otherwise} \end{cases}.$$

We now construct the auxiliary graph $G' = (V', E', a', b')$; the process is summarized in Figure 3.4. First, V' is obtained from V by deleting every vertex outside $X_\lambda(G) \cup \partial X_\lambda(G)$, and contracting each Aomoto tree T_i to a single vertex t_i. Write $\{t_1, ..., t_p\} = U \subset V'$, and identify $\partial X_\lambda(G)$ with ∂U. Now, for each $v \in \partial U = \partial X_\lambda(G)$ and each edge $e \in \tau(v) \subset E$ whose source is in a tree T_i, include an edge $e' \in E'$ with $\tau(e') = v$ and $\sigma(e') = t_i$, and set its weight as

$$a'_{e'} = a_e \zeta_i(\sigma(e)). \tag{3.14}$$

This process is mirrored to construct an edge $f' \in E'$ from any $f \in \sigma(v) \subset E$ whose terminal is in T_i; no other edges are included. Finally, the potential b' is identically zero.

We have arranged things so that G' is bipartite, with connected components $G'_1, ..., G'_m$, whose respective covers we will denote $\mathcal{T}'_1, ..., \mathcal{T}'_m$. We may also construct a new infinite graph $\mathcal{T}' = (\mathcal{V}', \mathcal{E}', a', b')$ from $\mathcal{T}$ analogously to the construction of G' from G: by deleting the fibres over any vertex outside $X_\lambda(G) \cup \partial X_\lambda(G)$, contracting each tree $S \in \mathcal{F} \subset \mathcal{T}$ into a single vertex u_S, including for each $e \in \mathcal{E}$ with ends in $\Xi^{-1}(X_\lambda(G))$ and $\Xi^{-1}(\partial X_\lambda(G))$ a corresponding edge $e' \in \mathcal{E}'$ with ends in the contraction if $\Xi^{-1}(X_\lambda(G))$ and its boundary, and reweighting any such edge according to (3.14). This $\mathcal{T}'$ consists of countably many copies of each $\mathcal{T}'_j$, and is a cover of G' via a map Ξ'; note that $(\Xi')^{-1}(U) = \{u_S : S \in \mathcal{F}\}$, the contraction of $\Xi^{-1}(X_\lambda(G))$. With this setup and the decomposition in (3.13), any $\eta \in \operatorname{Ker}(\lambda - A_\mathcal{T})$ gives rise to a vector $\eta' \in \ell^2(V(\mathcal{T}'))$ in a natural way:

$$\eta = \sum_{S \in \mathcal{F}} \alpha_S \zeta_S \mapsto \eta' = \sum_{S \in \mathcal{F}} \alpha_S \delta_{u_S}. \tag{3.15}$$

Observation 3.5.1. The map $\eta \mapsto \eta'$ is an isometric inclusion of $\operatorname{Ker}(\lambda - A_\mathcal{T})$ in $\operatorname{Ker} A_{\mathcal{T}'}$.

Proof. Preservation of norm is immediate since ζ_S is a unit vector, and since the map is an isometry it is injective; it remains only to show that $\operatorname{Ker}(\lambda - A_\mathcal{T})$ is mapped to $\operatorname{Ker} A_{\mathcal{T}'}$. The vector η' is identically zero on the fibre over ∂U and thus $(A_{\mathcal{T}'}\eta')(u) = 0$ for any $u \in (\Xi')^{-1}(U)$. It remains only to consider $v \in (\Xi')^{-1}(\partial U)$, which as above we may identify with $\Xi^{-1}(\partial X_\lambda(G)) \subset \mathcal{V}$. For each edge $e \in \tau(v) \subset \mathcal{E}$ write S_e for the tree in $\mathcal{F}$ to which $\sigma(e)$ belongs, so that the reweighting in (3.14) gives $a'_{e'} = a_e \zeta_{S_e}(\sigma(e))$. As the potential b' is identically zero and η and η' vanish outside the fibres over $X_\lambda(G)$ and U respectively, we have

$$\begin{aligned}
(A_{\mathcal{T}'}\eta')(v') &= b_{v'}\eta'(v') + \sum_{e' \in \tau(v') \subset \mathcal{E}'} a'_{e'}\eta'(\sigma(e')) \\
&= \sum_{\substack{e \in \tau(v) \subset \mathcal{E}: \\ \sigma(e) \in \Xi^{-1}(X_\lambda(G))}} a_e \zeta_{S_e}(\sigma(e))\alpha_{S_e} \\
&= b_v \eta(v) + \sum_{e \in \tau(v) \subset \mathcal{E}} a_e \eta(\sigma(e)) \\
&= (A_\mathcal{T}\eta)(v) \\
&= \lambda\eta(v) = 0.
\end{aligned}$$

In the third line, note that some edges in $\tau(v) \subset \mathcal{E}$ have a source outside of the fibre over $X_\lambda(G)$, but that η is identically zero there. □

Immediately from this observation, we can conclude that $0 \in \operatorname{Spec}_p A_{\mathcal{T}'}$. Moreover, as $\mathcal{T}'$ is comprised of disjoint copies of the $\mathcal{T}'_j$'s, $A_{\mathcal{T}'}$ restricts to $A_{\mathcal{T}'_j}$ on each one, and thus $0 \in \operatorname{Spec}_p A_{\mathcal{T}'_j}$ for at least one $\mathcal{T}'_j$. Our next observation characterizes the associated Aomoto set on G'_j. Recall that G' is bipartite with vertex classes U and ∂U, and let us write U_j and ∂U_j for the corresponding classes of vertices in each connected component G'_j.

Observation 3.5.2. $X_0(G'_j) = U_j$.

Proof. By Proposition 3.2.4, the definition of the map $\eta \mapsto \eta'$, and Observation 3.5.1, we immediately have the inclusion $U_j \subset X_0(G'_j)$, since any $\eta \in \operatorname{Ker}(\lambda - A_{\mathcal{T}})$ maps to η' supported only on the fibre over U_j. On the other hand, from Lemma 3.4.1, we know that $X_0(G'_j)$ is an independent set, and U_j is a maximal independent set in G'_j by definition of ∂U_j. □

We can now strengthen Observation 3.5.1

Observation 3.5.3. The map $\eta \mapsto \eta'$ gives an isomorphism between $\operatorname{Ker}(\lambda - A_{\mathcal{T}})$ and $\operatorname{Ker} A_{\mathcal{T}'}$.

Proof. We noted above that $A_{\mathcal{T}'}$ decomposes as a direct sum of the Jacobi operators on the copies of $\mathcal{T}'_j$ comprising $\mathcal{T}'$. By applying Observation 3.5.2 separately to each copy, any $\theta \in \operatorname{Ker} A_{\mathcal{T}'}$ is supported only on $(\Xi')^{-1}(U)$. Thus the adjoint of the map $\eta \mapsto \eta'$ takes any vector $\theta \in \operatorname{Ker} A_{\mathcal{T}'}$ to one in $\ell^2(V(\mathcal{T}))$:

$$\theta = \sum_{S \in \mathcal{F}} \alpha_S \delta_{u_S} \mapsto \sum_{S \in \mathcal{F}} \alpha_S \zeta_S.$$

Once again this is clearly injective and norm-preserving, and a parallel argument to Observation 3.5.1 shows that it takes $\operatorname{Ker} A_{\mathcal{T}'}$ into $\operatorname{Ker}(\lambda - A_{\mathcal{T}})$. □

Finally, we relate the spectral measures of $A_{\mathcal{T}}$ to the density of states of $A_{\mathcal{T}'_j}$.

Observation 3.5.4. Let T be an Aomoto tree in G, and $t \in V(G'_j)$ its contraction in a component G'_j of G'. Writing μ'_t for the spectral measure of t in $A_{\mathcal{T}'_j}$, and for μ_v for the spectral measure of $A_{\mathcal{T}}$ for each $v \in V(T) \subset V(G)$, we have

$$\sum_{v \in V(T)} \mu_v\{\lambda\} = \mu'_t\{0\}.$$

Proof. Choose a copy $\tilde{T}$ of T in its fibre in $\mathcal{T}$, and let $\tilde{t}$ be the contraction of $\tilde{T}$ in $\mathcal{T}'_j \subset \mathcal{T}'$. By construction, for each $\eta \in \ker(\lambda - A_{\mathcal{T}})$,

$$\eta'(\tilde{t})^2 = \sum_{\tilde{v} \in V(\tilde{T})} \eta(\tilde{v})^2.$$

Now, let $\mathfrak{B}'_j$ be an orthonormal basis of $\operatorname{Ker} A_{\mathcal{T}'_j}$. By Observation 3.5.3 this is the image of some orthonormal set $\mathfrak{B}_j$ in $\operatorname{Ker}(\lambda - A_{\mathcal{T}})$. In particular, recalling our construction of $\mathcal{T}'$ from $\mathcal{T}$ by deleting vertices and contracting Aomoto trees, our chosen copy of $\mathcal{T}'_j$ in $\mathcal{T}'$ pulls back to a subtree $\mathcal{T}_j$ of $\mathcal{T}$ containing $\tilde{T}$. Moreover, $\mathfrak{B}_j$ is an orthonormal basis for the orthogonal projection of $\operatorname{Ker}(\lambda - A_{\mathcal{T}})$ to the subspace of $\ell^2(\mathcal{V})$ supported on the vertices of $\mathcal{T}_j$, and we can therefore augment $\mathfrak{B}_j$ to an orthonormal basis $\mathfrak{B}$ of $\operatorname{Ker}(\lambda - A_{\mathcal{T}})$, whose additional vectors vanish on $\mathcal{T}_j$.

We now use Lemma 3.2.7 to compute

$$\sum_{v \in V(T)} \mu_v\{\lambda\} = \sum_{\tilde{v} \in V(\tilde{T})} \sum_{\eta \in \mathfrak{B}} \eta(\tilde{v})^2 = \sum_{\tilde{v} \in V(\tilde{T})} \sum_{\eta \in \mathfrak{B}_j} \eta(\tilde{v})^2 = \sum_{\eta \in \mathfrak{B}_j} \eta'(\tilde{t})^2 = \sum_{\eta' \in \mathfrak{B}'_j} \eta'(\tilde{t})^2 = \mu'_t\{0\}.$$

□

Analyzing the Auxiliary Graph

This section is devoted to the final observation of our proof:

Observation 3.5.5. Fix $j \in [m]$ and assume that $0 \in \operatorname{Spec}_p A_{\mathcal{T}'_j}$. Then

$$\sum_{t \in U_j} \mu'_t\{0\} = I_0(G'_j).$$

This will finish the proof, as combining Observations 3.5.4 and 3.5.5 and recalling the construction of G' gives

$$\begin{aligned} |V(G)| \cdot \mu\{\lambda\} &= \sum_{u \in X_\lambda(G)} \mu_u\{\lambda\} = \sum_{t \in U} \mu'_t\{0\} \\ &= \sum_{j \in [m]} I_0(G'_j) = \sum_{j \in [m]} |U_j| - |\partial U_j| \\ &= |U| - |\partial U| = \mathsf{cc} X_\lambda(G) - |\partial X_\lambda(G)| \\ &= I_\lambda(G). \end{aligned}$$

Proof of Observation 3.5.5. Let μ' be the density of states of $A_{\mathcal{T}'_j}$. Let $L_1, L_2, \dots$ be a sequence of finite lifts of G'_j with covering maps $\xi_n : L_n \to G'_j$. By Lemma 3.2.3 we may choose the L_n with girth going to infinity. Since G'_j is bipartite with zero potential, the Jacobi matrices A_{L_n} and $A_{\mathcal{T}'_j}$ have the following block structure

$$A_{L_n} = \begin{pmatrix} 0 & Z_n^T \\ Z_n & 0 \end{pmatrix} \quad \text{and} \quad A_{\mathcal{T}'_j} = \begin{pmatrix} 0 & Z_\infty^T \\ Z_\infty & 0 \end{pmatrix},$$

where for $n \in \mathbb{N} \cup \{\infty\}$ the domain and range of Z_n are supported on the fibres over ∂U_j and U_j respectively. Note that $A_{L_n}^2 = Z_n^T Z_n \oplus Z_n Z_n^T$ and $A_{\mathcal{T}'_j}^2 = Z_\infty^T Z_\infty \oplus Z_\infty Z_\infty^T$.

Let $\mu_{Z_n Z_n^T}$ and $\mu_{Z_n^T Z_n}$ be the empirical spectral distributions of $Z_n Z_n^T$ and $Z_n^T Z_n$ respectively. Fix a positive integer k and note that, since L_n is bipartite, the terms in $\operatorname{tr}(Z_n^T Z_n)^k$ are in one-to-one correspondence with the closed walks of length $2k$ in L_n that start and end at the same vertex in $\xi_n^{-1}(\partial U_j)$. Moreover, by the girth assumption, for large enough n it holds that the value of the diagonal entries of $A_{L_n}^{2k}$ are constant on each fibre $\xi_n^{-1}(v)$ for every $v \in V(G'_j)$ and coincide with the respective diagonal entries of $A_{\mathcal{T}'_j}^{2k}$. Hence, if we write

ν_v for the spectral measure of u for the operator $A^2_{\mathcal{T}'_j}$, then by the method of moments $\mu_{Z_n^T Z_n}$ and $\mu_{Z_n Z_n^T}$ converge weakly to

$$\nu_{\partial U_j} = \frac{1}{|\partial U_j|} \sum_{v \in \partial U_j} \nu_v \qquad \text{and} \qquad \nu_{U_j} = \frac{1}{|U_j|} \sum_{v \in U_j} \nu_v. \tag{3.16}$$

Since $X_0(G'_j) = U_j$, equation (3.16) implies $\nu_{\partial U_j}\{0\} = 0$. If it were the case that $|\partial U_j| > |U_j|$, we would have by standard properties of matrices that the spectrum of $Z_n^T Z_n$ is equal to that of $Z_n Z_n^T$, plus an eigenvalue at zero with multiplicity at least $|\partial U_j| - |U_j|$, and thus that

$$\frac{|\partial U_j| - |U_j|}{|\partial U_j|} \le \limsup_{n\to\infty} \mu_{Z_n Z_n^T}\{0\} \le \nu_{\partial U_j}\{0\} = 0,$$

a contradiction. Therefore $I_0(G'_j) \ge 0$. Applying the same matrix property a second time, we have

$$\mu_{Z_n Z_n^T} = \left(1 - \frac{I_0(G'_j)}{|U_j|}\right) \mu_{Z_n^T Z_n} + \frac{I_0(G'_j)}{|U_j|} \delta_0,$$

where by δ_0 we mean an atomic measure at 0. By weak convergence, and as compact measures are determined by their moments,

$$\nu_{U_j} = \left(1 - \frac{I_0(G'_j)}{|U_j|}\right) \nu_{\partial U_j} + \frac{I_0(G'_j)}{|U_j|} \delta_0,$$

and thus

$$\sum_{t \in U_j} \mu'_t\{0\} = |U_j| \nu_{U_j}\{0\} = \big(|U_j| - I_0(G'_j)\big)\, \nu_{\partial U_j}\{0\} + I_0(G'_j)\delta_0\{0\} = I_0(G'_j).$$

□

3.6 A Generalized Converse to Aomoto's Theorem

In this section we will prove the following generalization of Theorem 3.3.3, and use it to prove Theorem 3.3.2. Recall that we defined the set $\mathcal{A}(G)$ to contain those subsets $X \subset V(G)$ with $\mathsf{cc}(X) - |\partial X| > 0$, and which induce acyclic subgraphs of G, each of whose components has λ as an eigenvalue of its induced Jacobi operator.

Theorem 3.6.1. *Let G be a fintite graph, $\mathcal{T}$ its universal cover, $U : E(G) \to \mathrm{U}(n)$ a set of unitary-valued edge weights satisfying $U_e^* = U_{\check{e}}$ for every $e \in E(G)$, and $A_{G,U}$ the* unitary-weighted Jacobi operator *acting on $\eta \in \ell^2(V(G)) \otimes \mathbb{C}^n$ as*

$$(A_{G,U}\eta)(v) = b_v \eta(v) + \sum_{e \in \tau(v)} a_e U_e \eta(\sigma(e)) \in \mathbb{C}^n.$$

For every $X \in \mathcal{A}(G)$, $\lambda \in \operatorname{Spec} A_{G,U}$ with multiplicity at least $n(\mathsf{cc}(X) - |\partial X|)$.

We begin with a lemma regarding unitary-weighted Jacobi operators of finite trees.

Lemma 3.6.2. *Let $T = (V, E, a, b)$ be a finite tree, $U : E(T) \to \mathrm{U}(n)$ a set of unitary-valued edge weights satisfying $U_e^* = U_{\check{e}}$ for every $e \in E(T)$, and $A_{T,U}$ the associated unitary-weighted Jacobi operator. If $\lambda \in \operatorname{Spec} A_T$, then $\lambda \in \operatorname{Spec} A_{T,U}$ with multiplicity at least n.*

Proof. As in the proof of Lemma 3.4.1, we will choose a root r of T, for each vertex v write $p(v)$ for its unique parent and $c(v)$ for its set of children, and, since T is acyclic, write $v \leftarrow u$ for the unique edge with source u and terminal v. By absorbing λ into the potential, it suffices to study the case when $\lambda = 0$. So, let $\eta \in \operatorname{Ker} A_T$; we will produce a subspace of dimension n contained in $\operatorname{Ker} A_{T,U}$.

Fix a vector $\zeta_0 \in \mathbb{C}^n$ and set $\zeta(r) = \zeta_0$. For each vertex $v \in V(T)$, letting γ_v denote the directed edges in the unique shortest path from v to r, set

$$\zeta(v) = \prod_{e \in \gamma_v} U_e^* \cdot \eta(v) \cdot \zeta_0$$

We claim that $\zeta \in \operatorname{Ker} A_{T,U}$; since ζ_0 was arbitrary, this will complete the proof.

At the root, we have

$$(A_{T,U}\zeta)(r) = b_r \eta(r)\zeta_0 + \sum_{u \in c(r)} a_{r\leftarrow u} U_{r \leftarrow u} U_{u \leftarrow r} \eta(u)\zeta_0 = \left(b_r \eta(r) + \sum_{u \in c(r)} a_{r \leftarrow u}\eta(u) \right) \zeta_0 = 0,$$

since $U_{r\leftarrow u}U_{u\leftarrow r} = 1$ and $\eta \in \operatorname{Ker} A_T$. Similarly, for any other vertex $v \in V(T)$, conjugate symmetry of the unitary weights gives us

$$\begin{aligned}(A_{T,U}\zeta)(v) &= b_v \prod_{e\in\gamma_v} U_e^* \eta(v)\zeta_0 + a_{v\leftarrow p(v)} U_{v \leftarrow p(v)} \prod_{e \in \gamma p(v)} U_e^* \eta(p(v))\zeta_0 + \sum_{u \in c(v)} a_{v\leftarrow u} U_{v \leftarrow u} \prod_{e \in \gamma_u} U_e^* \eta(u)\zeta_0 \\ &= \left(b_v + a_{v \leftarrow p(v)} \eta(p(v)) + \sum_{u \in c(v)} a_{v \leftarrow u} \eta(u) \right) \prod_{e \in \gamma_v} U_e^* \zeta_0 \\ &= 0.\end{aligned}$$

□

We can now proceed with the proof.

Proof of Theorem 3.6.1. For any Aomoto tree T of G, the induced Jacobi operator A_T has λ in its spectrum. By Lemma 3.6.2, the induced unitary-weighted Jacobi operator $A_{T,U}$ thus satisfies $\dim \operatorname{Ker}(\lambda - A_{T,U}) \geq n$, and therefore the space

$$\bigoplus_{T \subset X_\lambda(G)} \operatorname{Ker}(\lambda - A_{T,U}) \subset \ell^2(X_\lambda(G)) \otimes \mathbb{C}^n \subset \ell^2(V) \otimes \mathbb{C}^n$$

has dimension $n\,\mathsf{cc}X_\lambda(G)$. We will show that it contains a subspace of dimension $nI_\lambda(G)$ which is itself contained in $\operatorname{Ker}(\lambda - A_G U)$.

For each $v \in X_\lambda(G)$, let $\Pi_v : \ell^2(V) \otimes \mathbb{C}^n \to \ell^2(v) \simeq \mathbb{C}^n$ be the orthogonal projection to the $\mathbb{C}^n$-valued functions in $\ell^2(V) \otimes \mathbb{C}^n$ supported on v. For each $u \in \partial X_\lambda(G)$, there is an operator

$$\phi_u = \sum_{\substack{e \in \tau(u) \\ \sigma(e) \in X_\lambda(G)}} a_e U_e \Pi_{\sigma(e)} : \bigoplus_{T \subset X_\lambda(G)} \operatorname{Ker}(\lambda - A_{T,U}) \to \ell^2(v) \simeq \mathbb{C}^n$$

and we define

$$\phi = \bigoplus_{u \in \partial X_\lambda(G)} \phi_u : \bigoplus_{T \subset X_\lambda(G)} \operatorname{Ker}(\lambda - A_{T,U}) \to \ell^2(\partial X_\lambda(G)) \simeq \mathbb{C}^{n|\partial X_\lambda(G)|}.$$

Counting dimensions, $\dim \operatorname{Ker}\phi \geq nI_\lambda(G)$, and we will show that $\operatorname{Ker}\phi \subset \operatorname{Ker}(\lambda - A_{G,U})$.

Let $\zeta \in \operatorname{Ker}\phi$; since the latter is a subspace of $\ell^2(X_\lambda(G)) \otimes \mathbb{C}^n \subset \ell^2(V) \otimes \mathbb{C}^n$, we have $\zeta(u) = 0$ for every $u \notin X_\lambda(G)$. This immediately gives $((\lambda - A_{G,U})\zeta)(u) = 0$ for any u outside the Aomoto set and its boundary, as ζ is identically zero on u and its neighbors. On the other hand, if u belongs to some tree T in the Aomoto set, then because $\operatorname{Ker}\phi \subset \bigoplus_{T \subset X_\lambda(G)} \operatorname{Ker}(\lambda - A_{T,U})$ and ζ vanishes on $\partial X_\lambda(G)$, we have

$$\begin{aligned}\big((\lambda - A_{G,U})\zeta\big)(u) &= \lambda\zeta(u) - b_u\zeta(u) - \sum_{e \in \tau(u)} a_e U_e \zeta(\sigma(e)) \\ &= \lambda\zeta(u) - b_u\zeta(u) - \sum_{\substack{e \in \tau(u) \\ \sigma(e) \in T}} a_e U_e \zeta(\sigma(e)) = \big((\lambda - A_{T,U})\zeta\big)(u) = 0\end{aligned}$$

It remains to check that $((\lambda - A_{G,U})\zeta)(u) = 0$ when $u \in \partial X_\lambda(G)$, which will follow from $\zeta \in \operatorname{Ker}\phi$. In particular, using a final time that ζ is supported only on the Aomoto set, if $u \in \partial X_\lambda(G)$ we have

$$\begin{aligned}\big((\lambda - A_{G,U})\zeta\big)(u) &= \lambda\zeta(u) - b_u\zeta(u) - \sum_{e \in \tau(u)} a_e U_e \zeta(\sigma(e)) \\ &= -\sum_{\substack{e \in \tau(u) \\ \sigma(e) \in X_\lambda(G)}} a_e U_e \zeta(\sigma(e)) = -\big(\phi\zeta\big)(u) = 0.\end{aligned}$$

□

Theorems 3.3.2 and 3.3.3 now follow easily.

Proof of Theorem 3.3.2. By Theorem 3.3.1, the Aomoto set satisfies the hypotheses of Theorem 3.6.1, and if H is an n-lift of G, both A_G and $A_{H/G}$ are unitary-weighted Jacobi operators for G—the former with weights taking values in $U(1)$ and the latter in $U(n-1)$ by the discussion in Section 3.2. Thus $\lambda \in \operatorname{Spec} A_G$ with multiplicity at least

$$\mathsf{cc}(X_\lambda(G)) - |\partial X_\lambda(G)| = I_\lambda(G) = |V(G)| \cdot \mu\{\lambda\}$$

and similarly $\lambda \in \operatorname{Spec} A_{H/G}$ with multiplicity at least $(n-1)|V(G)| \cdot \mu\{\lambda\}$, as desired. □

Proof of Theorem 3.3.3. Assertion (i) is a special case of Theorem 3.6.1. For (ii), let G_n be the sequence of lifts of G promised in Lemma 3.2.3, whose empirical spectral measures μ_{G_n} converge weakly to the density of states μ. Applying Theorem 3.6.1 to each A_{G_n}, viewed again as a unitary-weighted Jacobi operator on G, the empirical spectral measures μ_{G_n} satisfy $\mu_{G_n}\{\lambda\} \geq \frac{\mathsf{cc}(X)-|\partial X|}{V(G)}$. As these converge weakly to μ, we have

$$\mu\{\lambda\} \geq \frac{\mathsf{cc}(X) - |\partial X|}{V(G)}.$$

□

3.7 Spectral Delocalization for $A_{\mathcal{T}}$

In this section we will prove Theorem 3.3.5. Let $G = (V, E)$ be a fixed finite and unweighted graph, for which we will vary the weights a and potential b. In what follows we will identify $\mathbb{C}^{|E|/2}$ with $\mathbb{R}^{|E|}$ so that the parameter space for the a_e and the b_v is a subset of $\mathbb{R}^{|E|+|V|}$, and we will denote elements of $\mathbb{R}^{|E|+|V|}$ by (a, b), where $a = (a_e)_{e\in E}$ and $b = (b_v)_{v\in V}$. To ease notation define $m = |E|+|V|$. Theorem 3.3.5 asserts that the set $\mathcal{P} \subset \mathbb{R}^{|E|+|V|}$ of parameters for which $A_{\mathcal{T}}$ has point spectrum is closed, with large codimension. The former fact will follow from Theorem 3.3.3, the latter from Theorem 3.3.1.

Let $\mathcal{A}(G)$ be the family of vertex sets $X \subset V$ that induce an acyclic subgraph of G with the property that $\mathsf{cc}(X) - |\partial X| > 0$. For $X \in \mathcal{A}(G)$ let $\mathcal{P}_X \subset \mathbb{R}^m$ be the set of parameters for which all the Jacobi matrices of the trees induced by X have a common eigenvalue. Note that Theorems 3.3.1 (ii) and 3.3.3 (ii) imply

$$\mathcal{P} = \bigcup_{X\in\mathcal{A}(G)} \mathcal{P}_X. \tag{3.17}$$

To compute the dimension of $\mathcal{P}$ we will analyze each $\mathcal{P}_X$ individually. This will require basic techniques and concepts from real algebraic geometry, which we condense below. The experienced reader may proceed directly to Section 3.7.

Real Algebraic Geometry Preliminaries

We will need some elementary facts about algebraic and semialgebraic sets, as well as appropriate notions of dimension for each of these. A thorough introduction can be found, for instance, in Sections 2 and 3 of [Cos00].

An *algebraic set* (or, more formally, a *real affine algebraic set*) is a subset of $\mathbb{R}^n$ defined as the zero set of a family of polynomials with real coefficients. It is easy to see from the definition that any finite union or finite intersection of algebraic sets is still an algebraic set. Similarly, a *semialgebraic set* is a subset of $\mathbb{R}^n$ defined by a family of polynomial inequalities. Any algebraic set is semialgebraic, but the reverse need not be true.

An algebraic set $\mathcal{X}$ is *irreducible* if it cannot be expressed as a disjoint union of two algebraic sets strictly contained in $\mathcal{X}$. It well known [BCR13, Section 2.8] that any algebraic set $\mathcal{X}$ admits a unique decomposition of the form $\mathcal{X} = \bigcup_{i=1}^k \mathcal{X}_i$ where each $\mathcal{X}_i$ is an irreducible algebraic set and such that for no $i \neq j$ is $\mathcal{X}_i$ contained in $\mathcal{X}_j$. If $\mathcal{X}$ is an irreducible algebraic set we define the *algebraic dimension* of $\mathcal{X}$, denoted $\dim \mathcal{X}$, as the maximum integer d such that there exists a chain of the form $\mathcal{X}_0 \subset \mathcal{X}_1 \subset \cdots \subset \mathcal{X}_d = \mathcal{X}$, where each $\mathcal{X}_i$ is an irreducible algebraic set and each containment is strict. If $\mathcal{X}$ is any algebraic set and $\mathcal{X} = \bigcup_{i=1}^k \mathcal{X}_i$ is its decomposition into irreducible sets, we define the algebraic dimension of $\mathcal{X}$ as $\dim \mathcal{X} = \max_{i\in[k]} \dim \mathcal{X}_i$. It follows from these definitions that if $\mathcal{X}$ and $\mathcal{Y}$ are two algebraic sets, with $\mathcal{X}$ irreducible, and $\mathcal{X}$ is not contained in $\mathcal{Y}$, then $\dim \mathcal{X} \cap \mathcal{Y} < \dim \mathcal{X}$.

The notion of algebraic dimension for algebraic sets can be extended to a notion of dimension for semialgebraic sets via the *cylindrical algebraic decomposition*, which we describe here. Any semialgebraic set admits a decomposition of the form $\mathcal{S} = \bigcup_{i=1}^k \mathcal{C}_i$, where the $\mathcal{C}_i$ are disjoint semialgebraic subsets dieffeomorphic to the open hypercube $(0,1)^{d_i}$ for some nonnegative integer d_i [Cos00, Corollary 3.8]. With this setup we define the dimension of $\mathcal{S}$ as $\dim \mathcal{S} = \max_{i\in[k]} d_i$. We also remind the reader that the Hausdorff dimension of a semialgebraic set coincides with the notion of dimension described here.

Finally, we will use a fundamental fact about projections of semialgebraic sets. If $\mathcal{S} \subset \mathbb{R}^{n+1}$ is semialgebraic and $\Pi : \mathbb{R}^{n+1} \to \mathbb{R}^n$ is the projection onto the first n coordinates, then $\Pi\mathcal{S}$ is a semialgebraic subset of $\mathbb{R}^n$ [Cos00, Theorem 2.3], and moreover $\dim \Pi(\mathcal{S}) \leq \dim \mathcal{S}$ [Cos00, Lemma 3.17].

The Dimension of $\mathcal{P}$

We begin by proving the main technical result of this section.

Proposition 3.7.1. *For any $X \in \mathcal{A}(G)$, $\mathcal{P}_X$ is a semialgebraic set of dimension at most $m - \mathsf{cc}X + 1$.*

Proof. Let $p = \mathsf{cc}X$ and $T_1, \ldots, T_p$ be the trees induced by X. For any Hermitian edge weights $a : E \to \mathbb{C}$, let $x = \Re a$ and $y = \Im a$, that is, for every $e \in E$ we write $a_e = x_e + iy_e$, with $x_e, y_e \in \mathbb{R}$. View the characteristic polynomials of the T_i as polynomials in the x_e, y_e, b_v and z, namely, define $P_i(x, y, b, z) = \det(z - A_{T_i})$. We will first show that each $P_i(x, y, b, z)$ is a polynomial with real coefficients. Remember that $\Im P_i(x, y, b, z)$ is a polynomial with real coefficients in the aformentioned variables. Now, since A_{T_i} is Hermitian, for any choice of $x, y \in \mathbb{R}^{|E|/2}$ and $b \in \mathbb{R}^{|V|}$, we have that $\det(z - A_{T_i}) \in \mathbb{R}$, so $\Im P_i \equiv 0$ on $\mathbb{R}^{m+1}$ and hence $\Im P_i$ is the zero polynomial. It then follows that $P_i = \Re P_i$, which means that $P_i \in \mathbb{R}[x, y, b, z]$.

Now, for $i, r \in [p]$ define the algebraic sets

$$\mathcal{X}_i = \{(x, y, b, z) \in \mathbb{R}^{m+1} : P_i(x, y, b, z) = 0\} \qquad \text{and} \qquad \mathcal{X}_{\leq r} = \bigcap_{i=1}^{r} \mathcal{X}_i.$$

We will now show that $\mathcal{X}_{\leq r}$ has codimension at least r for all $r \in [p]$, which implies in particular that that $\mathcal{X}_{\leq p}$ has algebraic dimension at most $m + 1 - p$. For the base case note that since P_1 is not the zero polynomial, $\mathcal{X}_1$ is a proper algebraic subset of $\mathbb{R}^{m+1}$, and since $\mathbb{R}^{m+1}$ is irreducible $\dim(\mathcal{X}_1) < \dim(\mathbb{R}^{m+1}) = m + 1$. Now, for $1 \leq r \leq p - 1$ assume that $\dim(\mathcal{X}_{\leq r}) \leq m - r + 1$, define the subspace

$$W_r = \left\{ (x, y, b, z) \in \mathbb{R}^{m+1} : b_v = 0 \text{ for } v \in V(G) \setminus \bigcup_{i=1}^{r} V(T_i) \right\},$$

and let $\mathcal{Y} = \mathcal{X}_{\leq r} \cap W_r$.

The set $\mathcal{Y}$ is itself algebraic. Moreover, since for each $i \leq r$ the set P_i depends only on those b_v's with $v \in \bigcup_{i=1}^{r} V(T_i)$, we have that $\mathcal{X}_{\leq r} = \mathcal{Y} \times W_r^{\perp}$. Let $\bigcup_{i=1}^{n_1} \mathcal{Y}_i$ be the decomposition of $\mathcal{Y}$ into irreducible components. Since $\mathcal{Y}_i$ and $W_r^{\perp}$ are both irreducible $\mathcal{Y}_i \times W_r^{\perp}$ is as well, and hence $\bigcup_{i \in [n_1]} (\mathcal{Y}_i \times W_r^{\perp})$ is in fact the decomposition of $\mathcal{X}_{\leq r}$ into irreducible components. On the other hand

$$\mathcal{X}_{\leq r+1} = \mathcal{X}_{\leq r} \cap \mathcal{X}_{r+1} = \bigcup_{i=1}^{n_1} (\mathcal{Y}_i \times W_r^{\perp}) \cap \mathcal{X}_{r+1}.$$

We will now show that, for every $i \in [n_1]$, $\mathcal{Y}_i \times W_r^{\perp}$ is not contained in $\mathcal{X}_{r+1}$.

Indeed, fix $(x, y, b, z) \in \mathcal{Y}_i \times W_r^{\perp}$. Adding a constant $c \in \mathbb{R}$ to the b_v with $v \in V(T_{r+1})$ has the effect of shifting the spectrum of $A_{T_{r+1}}$ by c. We can then find $b' \in \mathbb{R}^{|V|}$ with the property that $b'_v = b_v$ for all $v \in V(G) \setminus V(T_{r+1})$ and such that $P_{r+1}(x, y, b', z) \neq 0$. By construction we have $(x, y, b', z) \in \mathcal{Y}_i \cap W_r^{\perp}$ and $(x, y, b', z) \notin \mathcal{X}_{r+1}$ as we wanted to show. This implies that $(\mathcal{Y}_i \times W_r^{\perp}) \cap \mathcal{X}_{r+1}$ is a proper subset of $\mathcal{Y}_i \times W_r^{\perp}$, and since the latter is irreducible we get

$$\dim \left((\mathcal{Y}_i \times W_r^{\perp}) \cap \mathcal{X}_{r+1} \right) < \dim(\mathcal{Y}_i \times W_r^{\perp}) \leq \dim(\mathcal{X}_{\leq r}) \leq m + 1 - r,$$

concluding the inductive step. Finally, let $\Pi : \mathbb{R}^{m+1} \to \mathbb{R}^m$ be the projection defined by $\Pi(x, y, b, z) = (x, y, b)$, and note that $\Pi \mathcal{X}_{\leq p} = \mathcal{P}_X$. From the results mentioned in Section 3.7, $\Pi \mathcal{X}_{\leq p}$ is a semialgebraic set whose dimension is less or equal to that of $\mathcal{X}_{\leq p}$, and in turn $\dim \mathcal{X}_{\leq p} \leq m - p + 1$. □

We are now ready to prove Theorem 3.3.5.

Proof of Theorem 3.3.5. By Proposition 3.7.1 and Equation (3.17), $\mathcal{P}$ is semialgebraic with

$$\operatorname{codim} \mathcal{P} \geq \min_{X \in \mathcal{A}(G)} \mathsf{cc}X - 1 \geq \min_{X \in \mathcal{A}(G)} \partial X,$$

and we want to further lower bound the latter quantity. As G has at least one cycle, $\partial X \neq \emptyset$ for all $X \in \mathcal{A}(G)$, so $\mathcal{P}$ has codimension at least 1. Now, if $d_{\min} \geq 2$ take any tree T induced by X. Any vertex v of T must be connected to at least $d_{\min} - 1$ distinct vertices in ∂X, and hence $\mathsf{cc}X - 1 \geq \partial X \geq d_{\min} - 1$. This proves the bound $\operatorname{codim} \mathcal{P} \geq \max\{d_{\min} - 1, 1\}$.

We show finally that $\mathcal{P}^c$ is open. For every $X \in \mathcal{A}(G)$ denote the forest induced by X by $\mathcal{F}_X$. Fix $(a, b) \in \mathcal{P}^c$. By Theorem 3.3.3, for every $X \in \mathcal{A}(G)$ the Jacobi matrices of the trees in $\mathcal{F}_X$ (with weights and potentials given by a and b), do not have a common eigenvalue. Now, define

$$S = \bigcup_{X \in \mathcal{A}(G)} \bigcup_{T \in \mathcal{F}_X} \operatorname{Spec} A_T.$$

As $\mathcal{F}_X$ is finite for each of the finitely many $X \in \mathcal{A}(G)$, we may safely define $\Delta > 0$ to be the smallest distance between two distinct points in S. We will show that if $(a', b') \in \mathbb{R}^m$ satisfies $\|(a, b) - (a', b')\|_2 < \Delta/2$ then $(a', b') \in \mathcal{P}^c$.

Assume otherwise. Then there exists an $X \in \mathcal{A}(G)$ such that the Jacobi matrices with parameters in (a', b') of the trees in $\mathcal{F}_X$ have a common eigenvalue λ. Let $T_1, \ldots, T_p$ be the trees in $\mathcal{F}_X$ with parameters in (a, b) and let $T_1', \ldots, T_p'$ denote the same trees but with parameters in (a', b'). For every i let λ_i be the closest point in $\operatorname{Spec} A_{T_i}$ to λ. Since $(a, b) \in \mathcal{P}^c$ we have $\lambda_i \neq \lambda_j$ for some i, j. On the other hand since $\|A_{T_i} - A_{T_i'}\| \leq \|A_{T_i} - A_{T_i'}\|_F \leq \|(a, b)\|_2 < \Delta/2$ and similarly $\|A_{T_j} - A_{T_j'}\| < \Delta/2$, the triangle inequality and Weyl's inequality together imply

$$|\lambda_i - \lambda_j| \leq |\lambda_i - \lambda| + |\lambda_j - \lambda| < \Delta/2 + \Delta/2 = \Delta,$$

contradicting the definition of Δ. $\square$

Chapter 4

Other short stories

4.1 Introduction

In this chapter we will look at two problems which have both been solved using techniques from geometry of polynomials. To begin let us first give a brief introduction to the subject matter at hand.

A polynomial $p : \mathbb{R} \to \mathbb{R}$ with degree n is called **real-rooted** if all its n roots are real. One way real-rooted polynomials appear naturally throughout graph theory is via the characteristic polynomials of various symmetric matrices related to graphs. For instance given a graph G with adjacency matrix A, its characteristic polynomial $\det(xI - A)$ is real-rooted.

Next we define interlacing between two polynomials. Its is crucial to note that in this definition interlacing has a direction and is not symmetric. One could have symmetric versions of this definition, but the asymmetry of this definition is what often allows one to tease out important information.

Definition 4.1.1. *A monic real-rooted polynomial of degree $n-1$, $g = \prod_{i=1}^{n-1}(x - \beta_i)$ is said to interlace a monic real-rooted polynomial of degree n, $f = \prod_{i=1}^{n}(x - \alpha_i)$ if*

$$\alpha_1 \leq \beta_1 \leq \alpha_2 \leq \beta_2 ... \leq \beta_{n-1} \leq \alpha_n.$$

Definition 4.1.2. *A set of real rooted monic polynomials $f_1, ..., f_k$ have a common interlacing if there exists a polynomial g such that g interlaces f_i for every $1 \leq i \leq k$.*

Note that f_1 and f_2 have a common interlacing **does not** mean that they themselves have alternating roots. This is where the asymmetry of our definition comes into play. Of particular importance is the following theorem about polynomials with common interlacing. This theorem which occurs in [Fel80] and [CS07] provides alternate characterizations of what it means to have a common interlacing.

Theorem 4.1.3. *If $f_1, ..., f_k$ are real-rooted monic polynomials of degree n then the following are equivalent :*

1. *$f_1, ..., f_k$ have a common interlacing.*
2. *There exists disjoint real intervals $I_1, ...I_n$ such that every polynomial f_i , $1 \leq i \leq k$ has **exactly** one root in every interval I_j, $1 \leq j \leq n$.*
3. *There exists disjoint real intervals $I_1, ..., I_n$ such that every convex combination of the polynomials $\sum_i a_i f_i$, $\sum_i a_i = 1; a_i \geq 0$ has **exactly** one root in every interval I_j, $1 \leq j \leq n$.*
4. *Every convex combination of the polynomials $\sum_i a_i f_i$, $\sum_i a_i = 1; a_i \geq 0$ is real-rooted.*

Additionally when any of the above statements are true, the same intervals $I_1, ..., I_n$ validate each of the statements.

Typical proofs of this are largely a matter of carefully applying Lagrange Mean Value Theorem and observing the signs of the polynomials after each root. We omit a proof of this theorem here. One crucial corollary that would be remiss of me to not note is the following :

Corollary 4.1.4. *If $f_1, ..., f_k$ is a set of polynomials with common interlacing and if all the roots of all the polynomial $\sum_i f_i$ are bounded above by m then there exists a polynomial f_j such that all of its roots are bounded above by m.*

Let us look at a couple of critical examples :

Example 4.1.5. ***Rank 1 updates have a common interlacing :***

Let A be a real symmetric matrix. Then the polynomials $\det(xI - A)$ and $\det(xI - A - vv^T)$ have a common interlacing.

Proof. We will show the result by showing that every convex combination of the two polynomials is real-rooted. To begin with let us use the matrix determinant lemma to get

$$\det(xI - A - vv^T) = (1 - v^T(xI - A)^{-1}v)\det(xI - A).$$

Now consider any convex combination of the two polynomials,

$$\begin{aligned} p\det(xI - A) + (1-p)\det(xI - A - vv^T) &= p\det(xI - A) + (1-p)(1 - v^T(xI - A)^{-1}v)\det(xI - A) \\ &= \det(xI - A) - (1-p)(1 - v^T(xI - A)^{-1}v)\det(xI - A) \\ &= \det(xI - A - (1-p)vv^T), \end{aligned}$$

where the last line is by reusing matrix determinant lemma. Thus as $A + (1-p)vv^T$ is symmetric, its characteristic polynomial is real-rooted. $\square$

Example 4.1.6. ***Sometimes rank 2 updates also have a common interlacing :***
Let A be a real symmetric matrix. Then the polynomials $\det(xI - A)$ *and* $\det(xI - A - vv^T + uu^T)$ *have a common interlacing.*

Proof. The crucial idea here is that both of the matrices A and $A + vv^T - uu^T$ are "positive" rank 1 updates of the same matrix, namely $A - uu^T$. This means that the updates pushes all the eigenvalues in the same "direction". Thus it is enough to show that if $B = A - uu^T$ is a symmetric matrix, then the two polynomials $f = \det(xI - (B + uu^T))$ and $g = \det(xI - (B + vv^T))$ have a common interlacing. Since B is symmetric it admits an orthonormal eigenbasis. Let us now write $\lambda_1, ..., \lambda_n$ as the eigenvalues and $w_1, ..., w_n$ as the orthonormal eigenvectors of B. Then we have that

$$(xI - B)^{-1} = \sum_{i=1}^{n} \frac{w_i w_i^T}{x - \lambda_i}.$$

As in the previous example we will show that every convex combination of the two polynomials is real-rooted. To that end let $p \det(xI - B - uu^T) + (1 - p)(\det(xI - B - vv^T)$ be an arbitrary convex combination. Then again using matrix determinant lemma gives us

$$\begin{aligned} & p \det(xI - B - uu^T) + (1 - p)(\det(xI - B - vv^T) \\ = & \det(xI - B) - pu^T(xI - B)^{-1}u - (1 - p)v^T(xI - B)^{-1}v \\ = & \det(xI - B) - \sum_{i=1}^{n} \frac{p(u \cdot w_i)^2 + (1 - p)(v \cdot w_i)^2}{x - \lambda_i}. \end{aligned}$$

Now note that as w_i is an orthonormal basis and $p(u.w_i)^2 + (1-p)(v.w_i)^2$ is a positive quantity, there exists a vector γ_p such that for every $1 \le i \le n$, $\gamma_p.w_i = \sqrt{p(u \cdot w_i)^2 + (1 - p)(v \cdot w_i)^2}$. Plugging this back into the equation gives us

$$\begin{aligned} & p \det(xI - B - uu^T) + (1 - p)(\det(xI - B - vv^T) \\ = & \det(xI - B) - \sum_{i=1}^{n} \frac{(\gamma_p.w_i)^2}{x - \lambda_i} \\ = & \det(xI - B) - \gamma_p^T(xI - B)^{-1}\gamma_p \\ = & \det(xI - B - \gamma_p\gamma_p^T). \end{aligned}$$

Again as $B + \gamma_p\gamma_p^T$ is symmetric, our polynomial is real-rooted. □

Let us finally take a look at interlacing families.

Definition 4.1.7. *Let $S_1, ..., S_m$ be finite indexing sets i.e. for every $s_1, ..., s_m \in S_1 \times ... \times S_m$ there is a monic real rooted polynomial of degree n $f_{s_1,...,s_m}$. For any $k < m$ and $s_1, ..., s_k \in$*

$S_1 \times \ldots \times S_k$ define

$$f_{s_1,\ldots,s_k} := \sum_{s_{k+1}\in S_{k+1}} f_{s_1,\ldots,s_{k+1}} = \sum_{s_{k+1},\ldots,s_m \in S_{k+1}\times\ldots\times S_m} f_{s_1,\ldots,s_{k+1},\ldots,s_m}.$$

Define

$$f_\emptyset = \sum_{s_1,\ldots,s_m\in S_1\times\ldots\times S_m} f_{s_1,\ldots,s_m}.$$

We say that the polynomials $\{f_{s_1,\ldots,s_m} : s_1,\ldots,s_m \in S_1 \times \ldots \times S_m\}$ form an interlacing family if for every $k \leq m$ and every $s_1,\ldots,s_{k-1} \in S_1 \times \ldots \times S_{k-1}$ the set of polynomials $\{f_{s_1,\ldots,s_k} : s_k \in S_k\}$ have a common interlacing.

Combining this definition with Corollary 4.1.4 and using induction, gives us the following theorem.

Theorem 4.1.8. *Let $\{f_{s_1,\ldots,s_m} : s_1,\ldots,s_m \in S_1 \times \ldots \times S_m\}$ be an interlacing family, so that all roots of $f_\emptyset$ are smaller than m, then there exists a $t_1,\ldots,t_k \in S_1 \times \ldots \times S_m$ such that all roots of $f_{t_1,\ldots,t_k}$ are smaller than m.*

The following is a theorem from [MSS14]

Theorem 4.1.9. *Let A be a real symmetric matrix and $S_i = \{1, -1\}$ for $1 \leq i \leq m = n(n+1)/2$. Then we note that each element of $S_1 \times \ldots \times S_m$ corresponds to a symmetric signed matrix $S_{s_1,\ldots,s_m}$. Thus we can define the polynomial $f_{s_1,\ldots,s_m}(x) = \det(xI - A \circ S_{s_1,\ldots,s_m})$. Then $f_{s_1,\ldots,s_m}(x)$ forms an interlacing family.*

The crux of the proof is same as the ideas in Example 4.1.6. In short the idea is that when we flip the signs of a symmetric matrix say at position $1,2$ and $2,1$, we essentially make a rank two update, and this update has exactly one positive and one negative eigenvalue. While in [MSS14], the matrix, A, is an adjacency matrix of graphs, we note that like in Example 4.1.6, the matrix only needs to be a real symmetric one for the proof to hold. In fact the ideas can be further generalised to the ideas of determinant-like polynomials which we use later, developed in [MSS15]. For further reading it is recommended to read the afore-mentioned papers.

4.2 Yet another class of real rooted polynomials

Given a matrix A, let us denote its characteristic polynomial by $\chi(A) = det(xI - A)$. Then in [MSS15] the following expected polynomials are shown to be real-rooted.

Theorem 4.2.1. *Given two symmetric matrices A and B, the following polynomials are real rooted*

1. $\mathbb{E}_Q \chi(A + QBQ^T)$

2. $\mathbb{E}_{P,Q}\chi\left((A+PBQ^T)(A+PBQ^T)^T\right)$,

where P, Q is any distribution of permutations realized by swaps. Here the second polynomial is real rooted even if the matrices are not symmetric.

We look at a a different expected polynomial arising out of asymmetric matrices.

Theorem 4.2.2. *Let $A_1, \ldots, A_n$ be any matrices, then the following polynomial is real rooted*

$$\mathbb{E}_{P_i,Q_i}\chi(\sum_i P_iA_iQ_i^T + Q_iA_i^TP_i^T)$$

where P_i, Q_i is any distribution of permutations realized by swaps.

The motivation for this is found as a conjecture in [HPS18]. The motivation is that when proved, this conjecture removes the loopless assumption of the following theorem that they prove:

Theorem 4.2.3. *Every connected, loopless d-regular graph has a one-sided Ramanujan r-covering. If the graph is bipartite then there is a a two-sided Ramanujan r-covering.*

The conjecture is also of independent interest as it can be viewed as another kind on convolution of two characteristic polynomials, similar to symmetric and assymetric additive convolutions.

Preliminaries and Notation

Let U be the set of vectors in $\mathbb{R}^{2n}$ whose first n co-ordinates is same as its last n i.e.

$$U = \{\begin{bmatrix} x \\ x \end{bmatrix}\} \subset \mathbb{R}^{2n}.$$

Let Π be a scalar multiple of the projection onto U. In particular

$$\Pi = \frac{1}{\sqrt{2}}\begin{bmatrix} I & I \\ I & I \end{bmatrix}.$$

Lastly Given a $n \times n$ matrix A define $A^{\bowtie} = \begin{bmatrix} 0 & A \\ A^T & 0 \end{bmatrix}$.

Given a linear map $T : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ which fixes U, $T(U) \subset U$ we can consider the U restriction of $T, T|_U : U \to U$. We abuse notation and define the U-determinant of T to be $\det|_U(T) = \det(T|_U)$. We similarly define $\chi|_U(T) = \det|_U(xI - T)$.

Proposition 4.2.4. $\chi|_U(\Pi A^{\bowtie}\Pi) = \chi(A + A^T)$

Proof.

$$\begin{aligned}\Pi A^{\bowtie}\Pi &= \frac{1}{2}\begin{bmatrix} I & I \\ I & I \end{bmatrix}\begin{bmatrix} 0 & A \\ A^T & 0 \end{bmatrix}\begin{bmatrix} I & I \\ I & I \end{bmatrix} \\ &= \frac{1}{2}\begin{bmatrix} A+A^T & A+A^T \\ A+A^T & A+A^T \end{bmatrix}\end{aligned}$$

If $(A+A^T)x = \lambda x$ we have $\Pi A^{\bowtie}\Pi \begin{bmatrix} x \\ x \end{bmatrix} = \lambda \begin{bmatrix} x \\ x \end{bmatrix}$. This shows the spectrum of the restriction of $\Pi A^{\bowtie}\Pi$ to U is the same as the spectrum as $A + A^T$, as desired. □

It is important to notice how certain permutations act on $A^{\bowtie}$

$$(P \oplus Q)A^{\bowtie}(P \oplus Q)^T = \begin{bmatrix} P & 0 \\ 0 & Q \end{bmatrix}\begin{bmatrix} 0 & A \\ A^T & 0 \end{bmatrix}\begin{bmatrix} P^T & 0 \\ 0 & Q^T \end{bmatrix} = \begin{bmatrix} 0 & PAQ^T \\ QA^TP^T & 0 \end{bmatrix} = (PAQ^T)^{\bowtie}$$

We follow the framework of [MSS15].

Lemma 4.2.5 (Interlacing IV, Lemma 3.5)**.** *If P, Q are uniformly random permutation then the following is realizable by swaps:*

$$\begin{bmatrix} P & 0 \\ 0 & Q \end{bmatrix}$$

[MSS15] introduces objects called determinant-like polynomials. A homogeneous polynomial $P(X_1, ..., X_m)$ of degree d, with entries being m symmetric matrices is *determinant-like* if it has the following properties.

- Hyperbolicity : The univariate restrictions $q(t) = P(tI - A_1, ..., tI - A_m)$ are real rooted for all symmetric $A_1, ..., A_m$.
- Rank one Linearity : For every vector v, index $i \leq m$, and real number s, we have

  $$P(X_1, X_2, ..., X_i + svv^T, ..., X_m) = P(X_1, ..., X_m) + sD_{i,vv^T}P(X_1, ..., X_m)$$

  with

  $$D_{i,vv^T}P(X_1, ..., X_m) = (\frac{\partial}{\partial s}P(X_1, ..., X_i + svv^T, ..., X_m))|_{s=0}$$

  is the directional derivative of P in the direction $(0, ..., vv^T, ..., 0)$, where vv^T appears in the ith position.

We proceed by showing the following polynomial (which takes as input $2n \times 2n$ symmetric matrices) is determinant-like:

$$P(X_1, \ldots, X_n) = \det|_U(\Pi \sum X_i \Pi)$$

- **H**yperbolicity: $P(tI - X_1, \ldots, tI - X_n) = \det|_U(\Pi tI\Pi - \sum \Pi X_i\Pi)$. Note that the linear transformation preserves U, and the restriction of any self adjoint map is self adjoint, so we are guaranteed that $\sum \Pi X_i \Pi|_U$ is self adjoint. Since Π^2 restricted to U is $2I$, we get a stretch by 2 of the characteristic polynomial of a self adjoint operator. We conclude the polynomials are real rooted.

- **R**ank-one Linearity: $P(X_1, \ldots, X_i + svv^T, \ldots, X_n) = \det|_U(\Pi \sum X_i\Pi + s\Pi(vv^T)\Pi) = \det|_U(\Pi \sum X_i\Pi + s\Pi(vv^T)\Pi) = \det|_U(\Pi \sum X_i\Pi + s(\Pi v)(\Pi v)^T = \det(\Pi \sum X_i\Pi|_U + s(\Pi v)(\Pi v)^T|_U)$. The restriction of a rank one map is rank one, so the rank one linearity argument proceeds as in [MSS15].

Completing the proof

Since determinant-like polynomials are closed under random swaps, we get $P((P_i \oplus Q_i)X_i(P_i \oplus Q_i)^T)$ is determinant-like. Namely it is hyperbolic in direction $(I, \ldots, I)$, and we can restrict our inputs so the following is real rooted:

Proof.

$$\begin{aligned}
P(\frac{t}{2m}I - (P_i \oplus Q_i)A_i^{\bowtie}(P_i \oplus Q_i)^T) &= P(\frac{t}{2m}I - (P_iA_iQ_i^T)^{\bowtie}) \\
&= \det(\Pi\frac{t}{2}\Pi|_U - \sum((P_iA_iQ_i^T)^{\bowtie})|_U) \\
&= \det(tI|_U - \sum((P_iA_iQ_i^T)^{\bowtie})|_U) \\
&= \chi|_U((\sum_i P_iA_iQ_i^T)^{\bowtie}) \\
&= \chi(\sum_i(P_iA_iQ_i^T + Q_iA_i^TP_i^T))
\end{aligned}$$

□

4.3 A special case of the existential version of the Non Commutative Khintchine inequality

The following theorem found in [LO94],[Tro12] is known as the Non-commutative Khintchine or Khintchine-Kahane inequality.

Theorem 4.3.1. *Let $A_1, ..., A_n$ be $d \times d$ symmetric matrices. Let $e_1, ..., e_n$ be random variable taking values 1 or -1 with equal probability. Then there exists constants K, K' such that*

$$K\sqrt{||\sum_i A_i^2||} \leq \mathbb{E}[||\sum_i e_iA_i||] \leq K'\sqrt{\log d}\sqrt{||\sum_i A_i^2||}.$$

For further reading on this inequality we refer to the works of [BVH+16], [LvHY18]. To ask a sort of existential version of this question let us look at the following theorem by Spencer [Spe85].

Theorem 4.3.2. *Let $L_1, ..., L_n$ be linear functionals, $L_i(x_1, ..., x_n) = \sum_j a_{ij}x_j$. such that all $|a_{ij}| \leq 1$. Then there exists $\epsilon_1, ..., \epsilon_n \in \{1, -1\}$ such that for all $1 \leq i \leq n$,*

$$|L(\epsilon_1, ..., \epsilon_n)| \leq K\sqrt{n},$$

where K is an absolute constant.

Thus we could ask if an existential version of these Khintchine inequality exists. In particular we ask the following question.

Does there exist a constant, K independent of d, such that given $A_1, ..., A_n$ symmetric $d \times d$ matrices, is it true that there exists a signing, as in a sequence of 1 and -1, $\epsilon_1, ..., \epsilon_n$, such that

$$||\sum_i \epsilon_i A_i|| \leq K\sqrt{||\sum_i A_i^2||}?$$

Here we prove a special case of the above. Namely we consider the case when the matrices $A_1, ..., A_n$ are linearly independent and have exactly one non-zero entry.

Theorem 4.3.3. *Let $A = \{a_{ij}\}_{i,j\in\mathbb{N}}$ be a bounded operator. Then there exists a signing of A such that*

$$||A \circ S||_2 < 2||A||_{l_\infty(l_2)},$$

where $A \circ S$ denotes the matrix generated by the entry-wise product of A and S.

A similar result was proved in 1997 by Françoise Lust-Piquard [LP97].

Theorem 4.3.4. *For every matrix $A = (a_{ij})$ such that A and A^* are bounded in $l^\infty(l^2)$ norm, there exists a matrix $B = (b_{ij})$ defining a bounded operator: $l^2(\mathbb{C}) \to l^2(\mathbb{C})$ such that*

(i) $|B|_{2\to2} \leq K max\{|A|_{l^\infty(l^2)}, |A^*|_{l^\infty(l^2)}\}$

(ii) $\forall i, j \in \mathbb{N}$, $|b_{ij}| \geq |a_{ij}|$,

where K is an absolute constant and $|A|_{l^\infty(l^2)} := \max_j \sqrt{\sum_i a_{ij}^2}$.

Our theorem is an improvement of this result in two ways. Firstly we show that there exists a signing of the matrix A which satisfies the above theorem. (A signing is a matrix B such that $|b_{ij}| = |a_{ij}|$). Secondly we get that the constant K as 2 suffices. In fact for the restricted class of symmetric signings the constant 2 is tight as we show in Theorem 4.3.14.

Similar results but for different norms have been proved by Pisier [Pis77]. In particular they prove that given a matrix A, there exists a signing, B such that

$$||B||_{\infty\to1} \leq K||A||_{l_1(l_2)}.$$

Notation and definitions

Given a $n \times n$ matrix $A = \{a_{ij}\}$, denote by

$$|A|_2 = \max_{x \in \mathbb{R}^n} \frac{||Ax||_2}{||x||_2}.$$

And denote

$$|A|_{l_\infty(l_2)} = \max_j \sqrt{\sum_i a_{ij}^2}.$$

Definition 4.3.5. *Given two $n \times n$ matrices $A = (a_{ij})$ and $B = (b_{ij})$, define their Schur product, $A \circ B$ to be the matrix whose (i,j)'th entry is $a_{ij}b_{ij}$.*

Definition 4.3.6. ***(Signings)*** *A sign matrix is a $n \times n$ matrix, S all of whose entries are 1 or -1. A symmetric sign matrix is, as the name suggests, a symmetric matrix which is also a sign matrix. Let $\mathcal{S}$ be the collection of all symmetric sign matrices of size n.*

Given any matrix A and a sign matrix S, a signing of A by S is simply the matrix $A \circ S$.

Definition 4.3.7. *A dimer arrangement D of size d on the set $\{1, 2, ..., n\}$ is a set of tuples $\{(i_1, j_1), ..., (i_d, j_d)\}$ such that all the the i's and j's are distinct from one another. The size of the dimer arrangement D, denoted by $|D|$ is the number of tuples, d. Let $\mathcal{D}$ be the set of all dimer arrangements of size d.*

The canonical weight of a dimer arrangement on a matrix A is defined to be $W_A(D) = \Pi_{(i,j)\in D} a_{ij}$.

Again given a matrix A, define the dimer partition function as

$$Z_d(A) = \sum_{|D|=d} W_A(D),$$

where the sum runs over all possible dimer arrangements of size d.

Definition 4.3.8. *Finally given a $n \times n$ matrix A, the matching polynomial of A is then defined to be*

$$\mu_A(x) = \sum_{i=0}^{n/2} (-1)^d Z_d(A) x^{n-2d}.$$

Preliminaries

The following is a trivial modification of theorem 3.6 in [MSS14].

Theorem 4.3.9. *Let A be a symmetric matrix. As previously defined let $\mathcal{S}$ be the set of all symmetric signing matrices. Let S be a random signing chosen uniformly from $\mathcal{S}$. Then*

$$\mathbb{E}_S[\det(xI - A \circ S)] = \mu_{A \circ A}(x).$$

Proof. Let $Sym(T)$ denote the set of permutations of a set T. Let $|\sigma|$ denote the entropy or the number of inversions of a permutation σ. Then

$$\begin{aligned}\mathbb{E}_S[\det(xI - A\circ S)] &= \mathbb{E}_S\left[\sum_{\sigma\in Sym([n])}(-1)^{|\sigma|}\prod_{i=1}^{n}(xI - A\circ S)_{i,\sigma(i)}\right]\\ &= \mathbb{E}_S\left[\sum_{k=0}^{n}x^{n-k}\sum_{T\subset[n];|T|=k}\sum_{\sigma\in Sym(T)}(-1)^{|\sigma|}\prod_{i=1}^{k}(-A\circ S)_{i,\sigma(i)}\right]\\ &= \sum_{k=0}^{n}x^{n-k}\sum_{T\subset[n];|T|=k}\sum_{\sigma\in Sym(T)}(-1)^{|\sigma|}\mathbb{E}_S[\prod_{i=1}^{k}-a_{i,\sigma(i)}s_{i,\sigma(i)}].\end{aligned}$$

But the $s_{i,j}$ are all independent excepting $s_{i,j} = s_{j,i}$, with expectation, $\mathbb{E}(s_{i,j}) = 0$. Thus only even powers of $s_{i,j}$ survive the expectation. So we may only consider permutations which only have orbits of size 2. These are just the perfect matchings on S or alternatively exactly all the dimer arrangements of size $|S|$. There are no such matchings when $|S|$ is odd. Otherwise its entropy is $|S|/2$. And since

$$\mathbb{E}[(-a_{i,j}s_{i,j})^2] = a_{i,j}^2,$$

we get

$$\mathbb{E}_S[\det(xI - A\circ S)] = \sum_{k=0}^{n/2}x^{n-2k}\sum_{|D|=k;D\in\mathcal{D}}(-1)^k\prod_{(i,j)\in D}a_{i,j}^2 = \mu_{A\circ A}(x).$$

□

The next theorem is the famous Heilman-Leib theorem which proves that the matching polynomial is real rooted and gives a bound for the maximum root of the matching polynomial of a matrix. It can be found in [HL72] as theorem 4.2 and 4.3.

Theorem 4.3.10. *Let A be a symmetric matrix with real positive entries. Let b be the maximum row sum of A i.e. $b = \max_{i\in[n]}\{\sum_j a_{i,j}\}$. Then $\mu_A(x)$ is real rooted and any root λ satisfies, $\lambda < 2\sqrt{b}$.*

An immediate corollary of the above theorem is the following.

Corollary 4.3.11. *Let A be any symmetric matrix. Let $r_1, ..., r_n$ be the rows of A. Let $||r_i||_2$ be the L2 norm of the vector r_i. Let $|A|_{l_\infty(l_2)} = \max_{i\in[n]}||r_i||_2 = \max\frac{||Ax||_\infty}{||Ax||_2}$.*

Then every root λ of $\mu_{A\circ A}(x)$ satisfies $|\lambda| < 2|A|_{l_\infty(l_2)}$.

Statement and Proof of main Theorem

Now we proceed to proving our main theorem.

Theorem 4.3.12. *Let A be any $n \times n$ matrix. Then there exists a signing matrix not necessarily symmetric such that*

$$||A \circ S||_2 \leq 2||A||_{l_\infty(l_2)}.$$

Proof. Given A, define the dilation A_D to be the $2n \times 2n$ matrix,

$$A_D = \begin{bmatrix} 0 & A \\ A^T & 0 \end{bmatrix},$$

where A^T denotes the transpose of A.

Note that A_D is a symmetric matrix. Let $\mathcal{S}$ be the set of all $2n \times 2n$ sign matrices. Let S be a sign matrix chosen uniformly from $\mathcal{S}$.

Then by Theorem 4.3.9,

$$\mathbb{E}_S[\det(xI - A_D \circ S)] = \mu_{A_D \circ A_D}(x).$$

But by Theorem 4.1.9, the polynomials in the left hand side of the above equation form an interlacing family. Therefore by Theorem 4.1.8, there exists some signing matrix S' such that, the largest root of $\det(xI - A_D \circ S')$ is less than or equal to the largest root of $\mu_{A_D \circ A_D}(x)$.

But using Corollary 4.3.11, every root of $\mu_{A_D \circ A_D}(x)$ is in modulus smaller than $2|A_D|_{l_\infty(l_2)}$.

Combining all this we have a $2n \times 2n$ sign matrix S' such that the largest eigenvalue of $A_D \circ S'$ is less $2|A_D|_{l_\infty(l_2)}$. Let $S' = \begin{bmatrix} S_1 & S_2 \\ S_2^T & S_4 \end{bmatrix}$ Then using Schur complements,

$$\det(xI - A_D \circ S) = x^n \det(x - x^{-1}(A \circ S_2)(A^T \circ S_2^T)) = \det(x^2 - (A \circ S_2)(A \circ S_2)^T).$$

Thus the largest eigenvalue of $A_D \circ S'$ is simply the largest singular value or the L2 norm of $A \circ S_2$.

So we have a signing matrix S_2, with

$$||A \circ S_2||_2 < 2||A \circ S_2||_{l_\infty(l_2)} = ||A||_{l_\infty(l_2)}.$$

$\square$

Theorem 4.3.13. *(Extension to infinite dimensions). Let $A = \{a_{ij}\}_{i,j \in \mathbb{N}}$ be a bounded infinite dimensional operator. Then there exists a signing of A such that*

$$||A \circ S||_2 < 2||A \circ S||_{l_\infty(l_2)}.$$

Proof. For any integer n, let A_n be the operator constructed from A by taking the upper $n \times n$ part of A and filling everything else with 0. Then by our previous result, there exists a signing S_n such that

$$||A_n \circ S_n||_2 < 2||A_n \circ S_n||_{l_\infty(l_2)} = 2||A_n||_{l_\infty(l_2)} \leq 2||A||_{l_\infty(l_2)}.$$

Thus as the sequence $\{A_n \circ S_n\}$ is uniformly bounded, by using sequential Banach Alaoglu, there is a subsequence k_n such that $A_{k_n} \circ S_{k_n}$ converges weakly to some matrix B. Note that k_n approaches infinity, thus eventually every i, j position of this subsequence is either a_{ij} or $-a_{ij}$. Thus the weak limit is also a signing of A. Denote $B_n = A_{k_n} \circ S_{k_n}$.

Thus $B_n^* B_n$ also converges weakly to $B^* B$. Then for any x, we have that $\langle x, B_n^* B_n x\rangle$ converges to $\langle x, B^* B x\rangle = ||Bx||_2$. Thus we have that for any x, such that $||x||_2 = 1$,

$$||Bx||_2 < 2||A||_{2,\infty}.$$

□

Heuristics of tightness

Let us briefly see why 2 is the best possible constant at least if we restrict our matrices to be chosen only from symmetric signings.

Theorem 4.3.14. *Let K be any constant strictly less than 2. Then the following statement is **false** : Let A be any symmetric $n \times n$ matrix. Then there exists a symmetric signing matrix such that*

$$||A \circ S||_2 \leq K||A||_{l_\infty(l_2)}.$$

Proof. We prove by contradiction. Suppose such a K exists. Then there exists a integer d such that $K\sqrt{d} < 2\sqrt{d-1}$. Now to construct a counterexample we first consider the d-regular complete graph K_{d+1}. Let A_1 be its adjacency matrix. We note that A_1 has exactly one eigenvalue at d and the rest at -1. Thus the second largest eigenvalue in modulus of A_1 is smaller than $K\sqrt{d}$ (by choosing a suitably large d.) Then by the assumption, there exists a signing S_1 of A_1 such that

$$||A_1 \circ S_1||_2 \leq K||A||_{l_\infty(l_2)} = K\sqrt{d}.$$

Note that as the largest eigenvalue of A_1 is d, S_1 is a non-trivial signing. As each non-trivial signing gives a connected 2-lift of G, let G_2 be the connected two lift whose adjacency matrix A_2 is similar to $A_1 \bigoplus (A_1 \circ S_1)$. Thus we note that A_2 preserves the property that its largest eigenvalue is d and all other eigenvalues are smaller than $K\sqrt{d}$.

Now repeating the previous argument we can get a A_3, A_4 and so on with the properties that each A_i is an adjacency matrix of a connected d-regular graph such that the second largest eigenvalue in modulus is smaller than $K\sqrt{d}$. As we chose d large enough such that $K\sqrt{d} < 2\sqrt{d-1}$, this is impossible by Alon-Boppana. □

Bibliography

[A+88] Kazuhiko Aomoto et al. Algebraic equations for green fkernel on a tree. *Proceedings of the Japan Academy, Series A, Mathematical Sciences*, 64(4):123–125, 1988.

[ABKSon] N. Avni, J. Breuer, G. Kilai, and B. Simon. Periodic boundary conditions for periodic jacobi matrices on trees, paper in preparation.

[ABS20] Nir Avni, Jonathan Breuer, and Barry Simon. Periodic Jacobi matrices on trees. *Advances in Mathematics*, 370:107241, 2020.

[AFH15] Omer Angel, Joel Friedman, and Shlomo Hoory. The non-backtracking spectrum of the universal cover of a graph. *Transactions of the American Mathematical Society*, 367(6):4287–4318, 2015.

[AGZ10] Greg W Anderson, Alice Guionnet, and Ofer Zeitouni. *An introduction to random matrices*, volume 118. Cambridge university press, 2010.

[AISW20] Nalini Anantharaman, Maxime Ingremeau, Mostafa Sabri, and Brian Winn. Absolutely continuous spectrum for quantum trees. *arXiv preprint arXiv:2003.12765*, 2020.

[AMHR13] Awad H. Al-Mohy, Nicholas J. Higham, and Samuel D. Relton. Computing the Fréchet derivative of the matrix logarithm and estimating the condition number. *SIAM Journal on Scientific Computing*, 35(4):C394–C410, 2013.

[AMHR15] Awad H. Al-Mohy, Nicholas J. Higham, and Samuel D. Relton. New algorithms for computing the matrix sine and cosine separately or simultaneously. *SIAM Journal on Scientific Computing*, 37(1):A456–A487, 2015.

[Ana18] Nalini Anantharaman. Delocalization of schrödinger eigenfunctions. *Proceedings of the ICM, Rio de Janeiro*, 2018.

[Aom91] K Aomoto. Point spectrum on a quasihomogeneous tree. *Pacific Journal of Mathematics*, 147(2):231–242, 1991.

[AS19] Nalini Anantharaman and Mostafa Sabri. Quantum ergodicity on graphs: from spectral to spatial delocalization. *Annals of Mathematics*, 189(3):753–835, 2019.

[BC19] Charles Bordenave and Benoît Collins. Eigenvalues of random lifts and polynomials of random permutation matrices. *Annals of Mathematics*, 190(3):811–875, 2019.

[BCR13] Jacek Bochnak, Michel Coste, and Marie-Françoise Roy. *Real algebraic geometry*, volume 36. Springer Science & Business Media, 2013.

[BD18] Paul Bourgade and Guillaume Dubach. The distribution of overlaps between eigenvectors of Ginibre matrices. *arXiv preprint arXiv:1801.01219*, 2018.

[BF60] Friedrich L. Bauer and Charles T. Fike. Norms and exclusion theorems. *Numerische Mathematik*, 2(1):137–141, 1960.

[BGZ18] Florent Benaych-Georges and Ofer Zeitouni. Eigenvectors of non normal random matrices. *Electronic Communications in Probability*, 23, 2018.

[BPZ18] Anirban Basak, Elliot Paquette, and Ofer Zeitouni. Spectrum of random perturbations of Toeplitz matrices with finite symbols. *arXiv preprint arXiv:1812.06207*, 2018.

[BPZ19] Anirban Basak, Elliot Paquette, and Ofer Zeitouni. Regularization of non-normal matrices by Gaussian noise - the banded Toeplitz and twisted Toeplitz cases. In *Forum of Mathematics, Sigma*, volume 7. Cambridge University Press, 2019.

[Bru89] Marie-France Bru. Diffusions of perturbed principal component analysis. *Journal of multivariate analysis*, 29(1):127–136, 1989.

[Bru91] Marie-France Bru. Wishart processes. *Journal of Theoretical Probability*, 4(4):725–751, 1991.

[BVH+16] Afonso S Bandeira, Ramon Van Handel, et al. Sharp nonasymptotic bounds on the norm of random matrices with independent entries. *The Annals of Probability*, 44(4):2479–2506, 2016.

[Cio06] Sebastian M Cioabă. Eigenvalues of graphs and a simple proof of a theorem of greenberg. *Linear algebra and its applications*, 416(2-3):776–782, 2006.

[CM98] John T. Chalker and Bernhard Mehlig. Eigenvector statistics in non-Hermitian random matrix ensembles. *Physical review letters*, 81(16):3367, 1998.

[Cos00] Michel Coste. An introduction to semialgebraic geometry, 2000.

[CS07] Maria Chudnovsky and Paul Seymour. The roots of the independence polynomial of a clawfree graph. *Journal of Combinatorial Theory, Series B*, 97(3):350–357, 2007.

[Dav07] E. Brian Davies. Approximate diagonalization. *SIAM Journal on Matrix Analysis and Applications*, 29(4):1051–1064, 2007.

[Dem83] James Weldon Demmel. A numerical analyst's Jordan canonical form. Technical report, UC Berkeley Center for Pure and Applied Mathematics, 1983.

[DH09] E. Brian Davies and Mildred Hager. Perturbations of Jordan matrices. *Journal of Approximation Theory*, 156(1):82–94, 2009.

[DHS89] K. R. Davidson, D. A. Herrero, and N Salinas. Quasidiagonal operators, approximation, and C*–algebras. *Indiana University Mathematics Journal*, 38(4):973–998, 1989.

[DIP$^+$19] Emilio Defez, Javier Ibáñez, Jesús Peinado, Jorge Sastre, and Pedro Alonso-Jordá. An efficient and accurate algorithm for computing the matrix cosine based on new Hermite approximations. *Journal of Computational and Applied Mathematics*, 348:1–13, 2019.

[DS01] Kenneth R. Davidson and Stanisław J. Szarek. Local operator theory, random matrices and Banach spaces. *Handbook of the geometry of Banach spaces*, 1(317-366):131, 2001.

[DW98] Xiaodong Ding and Rangquan Wu. A new proof for comparison theorems for stochastic differential inequalities with respect to semimartingales. *Stochastic Processes and their applications*, 78(2):155–171, 1998.

[Ede88] Alan Edelman. Eigenvalues and condition numbers of random matrices. *SIAM Journal on Matrix Analysis and Applications*, 9(4):543–560, 1988.

[F$^+$03] Joel Friedman et al. Relative expanders or weakly relatively ramanujan graphs. *Duke Mathematical Journal*, 118(1):19–35, 2003.

[Fel80] H. J. Fell. On the zeros of convex combinations of polynomials. *Pacific Journal of Mathematics*, 89(1):43 – 50, 1980.

[Fie75] Miroslav Fiedler. Eigenvectors of acyclic matrices. *Czechoslovak Mathematical Journal*, 25(4):607–618, 1975.

[FPZ14] Ohad N. Feldheim, Elliot Paquette, and Ofer Zeitouni. Regularization of non-normal matrices by Gaussian noise. *International Mathematics Research Notices*, 2015(18):8724–8751, 2014.

[Fri93] Joel Friedman. Some geometric aspects of graphs and their eigenfunctions. *Duke Math. J.*, 69(3):487–525, 03 1993.

[Fyo18] Yan V. Fyodorov. On statistics of bi-orthogonal eigenvectors in real and complex Ginibre ensembles: combining partial Schur decomposition with supersymmetry. *Communications in Mathematical Physics*, 363(2):579–603, 2018.

[GM88] Chris D. Godsil and Bojan Mohar. Walk generating functions and spectral measures of infinite graphs. *Linear Algebra and its Applications*, 107:191–206, 1988.

[GM94] Christel Geiß and Ralf Manthey. Comparison theorems for stochastic differential equations in finite and infinite dimensions. *Stochastic processes and their applications*, 53(1):23–35, 1994.

[GM14] Piotr Graczyk and Jacek Małecki. Strong solutions of non-colliding particle systems. *Electronic Journal of Probability*, 19, 2014.

[GVK19] Jorge Garza-Vargas and Archit Kulkarni. Spectra of infinite graphs via freeness with amalgamation. *arXiv preprint arXiv:1912.10137*, 2019.

[GWZ14] Alice Guionnet, Philip Wood, and Ofer Zeitouni. Convergence of the spectral measure of non-normal matrices. *Proceedings of the American Mathematical Society*, 142(2):667–679, 2014.

[HL72] Ole J Heilmann and Elliott H Lieb. Theory of monomer-dimer systems. In *Statistical Mechanics*, pages 45–87. Springer, 1972.

[HL00] Uffe Haagerup and Flemming Larsen. Brown's spectral distribution measure for R-diagonal elements in finite von Neumann algebras. *Journal of Functional Analysis*, 176(2):331–367, 2000.

[HL13] Nicholas J. Higham and Lijing Lin. An improved Schur–Padé algorithm for fractional powers of a matrix and their Fréchet derivatives. *SIAM Journal on Matrix Analysis and Applications*, 34(3):1341–1360, 2013.

[HPS18] Chris Hall, Doron Puder, and William F. Sawin. Ramanujan coverings of graphs. *Advances in Mathematics*, 323:367–410, Jan 2018.

[Kes59] Harry Kesten. Symmetric random walks on groups. *Transactions of the American Mathematical Society*, 92(2):336–354, 1959.

[KLW12] Matthias Keller, Daniel Lenz, and Simone Warzel. Absolutely continuous spectrum for random operators on trees of finite cone type. *Journal d'Analyse Mathématique*, 118(1):363–396, 2012.

[KLW13] Matthias Keller, Daniel Lenz, and Simone Warzel. On the spectral theory of trees with finite cone type. *Israel Journal of Mathematics*, 194(1):107–135, 2013.

[KO+01] Wolfgang König, Neil O'Connell, et al. Eigenvalues of the Laguerre process as non-colliding squared Bessel processes. *Electronic Communications in Probability*, 6:107–114, 2001.

[Kra10] Vladislav Krasin. *Comparison theorem and its applications to finance*. PhD thesis, University of Alberta, Edmonton, Alberta, 2010.

[Le94] Huiling Le. Brownian motions on shape and size-and-shape spaces. *Journal of applied probability*, 31(1):101–113, 1994.

[Le99] Huiling Le. Singular-values of matrix-valued Ornstein–Uhlenbeck processes. *Stochastic processes and their applications*, 82(1):53–60, 1999.

[LO94] Rafał Latała and Krzysztof Oleszkiewicz. On the best constant in the khinchin-kahane inequality. *Studia Mathematica*, 109(1):101–104, 1994.

[LP97] Françoise Lust-Piquard. On the coefficient problem: a version of the kahane–katznelson–de leeuw theorem for spaces of matrices. *journal of functional analysis*, 149(2):352–376, 1997.

[LvHY18] Rafał Latała, Ramon van Handel, and Pierre Youssef. The dimension-free structure of nonhomogeneous random matrices. *Inventiones mathematicae*, 214(3):1031–1080, 2018.

[McK81] Brendan D. McKay. The expected eigenvalue distribution of a large regular graph. *Linear Algebra and its Applications*, 40:203–216, 1981.

[MPF91] Dragoslav S. Mitrinovic, Josip Pecaric, and Arlington M Fink. *Inequalities involving functions and their integrals and derivatives*, volume 53. Springer Science & Business Media, 1991.

[MSS14] Adam Marcus, Daniel A. Spielman, and Nikhil Srivastava. Interlacing families i: Bipartite ramanujan graphs of all degrees, 2014.

[MSS15] Adam W. Marcus, Nikhil Srivastava, and Daniel A. Spielman. Interlacing families iv: Bipartite ramanujan graphs of all sizes, 2015.

[NH18] Prashanth Nadukandi and Nicholas J. Higham. Computing the wave-kernel matrix functions. *SIAM Journal on Scientific Computing*, 40(6):A4060–A4082, 2018.

[Nil91] Alon Nilli. On the second eigenvalue of a graph. *Discrete Mathematics*, 91(2):207–210, 1991.

[Nyl98] Peter Nylen. Null space structure of tree-patterned matrices. *Linear algebra and its applications*, 279(1-3):153–161, 1998.

[Pis77] Gilles Pisier. Sur l'espace de banach des séries de fourier aléatoires presque sûrement continues. *Séminaire Analyse fonctionnelle (dit" Maurey-Schwartz")*, pages 1–33, 1977.

[RS78] Michael Reed and Barry Simon. *Methods of Modern Mathematical Physics, IV: Analysis of Operators*. Academic Press, New York, 1978.

[Ruh87] Axel Ruhe. Closest normal matrix finally found! *BIT Numerical Mathematics*, 27(4):585–598, 1987.

[RY99] Daniel Revuz and Marc Yor. *Continuous martingales and Brownian motion*, volume 293. Springer Science & Business Media, 1999.

[Śni02] Piotr Śniady. Random regularization of Brown spectral measure. *Journal of Functional Analysis*, 193(2):291–313, 2002.

[Spe85] Joel Spencer. Six standard deviations suffice. *Transactions of the American mathematical society*, 289(2):679–706, 1985.

[SS92] Polly Wee Sy and Toshikazu Sunada. Discrete Schrödinger operators on a graph. *Nagoya Mathematical Journal*, 125:141–150, 1992.

[SST06] Arvind Sankar, Daniel A. Spielman, and Shang-Hua Teng. Smoothed analysis of the condition numbers and growth factors of matrices. *SIAM Journal on Matrix Analysis and Applications*, 28(2):446–476, 2006.

[Sun92] Toshikazu Sunada. Group $C*$-algebras and the spectrum of a periodic Schrödinger operator on a manifold. *Canadian journal of mathematics*, 44(1):180–193, 1992.

[SV19a] Johannes Sjöstrand and Martin Vogel. General Toeplitz matrices subject to gaussian perturbations. *arXiv preprint arXiv:1905.10265*, 2019.

[SV19b] Johannes Sjöstrand and Martin Vogel. Toeplitz band matrices with small random perturbations. *arXiv preprint arXiv:1901.08982*, 2019.

[TE05] Lloyd N. Trefethen and Mark Embree. *Spectra and pseudospectra: the behavior of nonnormal matrices and operators*. Princeton University Press, 2005.

[Tro12] Joel A Tropp. User-friendly tail bounds for sums of random matrices. *Foundations of computational mathematics*, 12(4):389–434, 2012.

[WT02] Thomas G. Wright and Lloyd N. Trefethen. EigTool. 2002. Software available at `http://www.comlab.ox.ac.uk/pseudospectra/eigtool`.