2015 IEEE 56th Annual Symposium on Foundations of Computer Science

Interlacing Families IV: Bipartite Ramanujan Graphs of All Sizes

Adam W. Marcus
Princeton University

Daniel A. Spielman
Yale University

Nikhil Srivastava
UC Berkeley

Abstract

We prove that there exist bipartite Ramanujan graphs of every degree and every number of vertices. The proof is based on analyzing the expected characteristic polynomial of a union of random perfect matchings, and involves three ingredients: (1) a formula for the expected characteristic polynomial of the sum of a regular graph with a random permutation of another regular graph, (2) a proof that this expected polynomial is real rooted and that the family of polynomials considered in this sum is an interlacing family, and (3) strong bounds on the roots of the expected characteristic polynomial of a union of random perfect matchings, established using the framework of finite free convolutions introduced recently by the authors.

I. Introduction

Ramanujan graphs are undirected regular graphs whose nontrivial adjacency matrix eigenvalues are as small as possible; in other words, they are the optimal spectral expander graphs. In this paper, we prove the existence of bipartite Ramanujan graphs of every degree and every size. We do this by showing that a random m-regular bipartite graph, obtained as a union of m random perfect matchings across a bipartition of an even number of vertices, is Ramanujan with nonzero probability. Infinite families of bipartite Ramanujan graphs were shown in [11] to exist for every degree $m \geq 3$, but it was not known whether they exist for every number of vertices.

Our proof is based on the method of *interlacing families* of polynomials, introduced in [11]. This method allows one to control the eigenvalues of a random matrix by controlling the roots of its *expected characteristic polynomial*, and its name refers to the chain of intermediate polynomials whose interlacing properties provide the relationship between the two. The main content of this paper is the construction of an interlacing family for the adjacency matrix of a random regular graph, and the derivation of an explicit formula for its expected characteristic polynomial. The roots of the expected polynomial are then analyzed using a tool which we call the *finite free convolution*, developed in our companion paper [10]. This latter technique is inspired by ideas in Free Probability theory [15], [18], an area that originally grew out of operator algebras but which has a number of applications to asymptotic random matrix theory [1]. This allows us to obtain the optimal "Ramanujan" bound of $2\sqrt{d-1}$ from completely generic considerations involving random matrices (in particular making no use of results from algebraic graph theory or number theory as in previous work).

A. Summary of Results

Recall that the adjacency matrix A of an $\mathbf{m-}$**regular graph on** $\mathbf{d}$ **vertices**[1] has largest eigenvalue $\lambda_1(A) = m$, and smallest eigenvalue $\lambda_d(A) = -m$ when the graph is bipartite. Following Friedman [6], we will refer to these as the *trivial* eigenvalues of A, and we will call a graph *Ramanujan* if all of its non-trivial eigenvalues have absolute value at most $2\sqrt{m-1}$. Such graphs are asymptotically best possible in the sense that a theorem of Alon and Boppana [16] tells us that for every $\epsilon > 0$, every infinite sequence of m-regular graphs must contain a graph with a non-trivial eigenvalue of absolute value at least $2\sqrt{m-1} - \epsilon$.

Our main theorem is that a union of m random perfect matchings across a bipartition of $2d$ vertices is Ramanujan with nonzero probability.

[1]In order to be consistent with our companion paper [10], we will, unconventionally, use m to denote the degree of a graph and d to denote its number of vertices.

0272-5428/15 $31.00 © 2015 IEEE
DOI 10.1109/FOCS.2015.87

Theorem I.1. *For $m \geq 3$, let $P_1, \ldots, P_m$ be independent uniformly random $d \times d$ permutation matrices. Then, with nonzero probability the nontrivial eigenvalues of*

$$A = \sum_{i=1}^{m} \begin{bmatrix} 0 & P_i \\ P_i^T & 0 \end{bmatrix}$$

are all less than $2\sqrt{m-1}$ in absolute value.

We also prove the following non-bipartite version of this theorem, regarding a union of m random perfect matchings on d vertices (not bipartite), with d even.

Theorem I.2. *Let d be even and let M be the adjacency matrix of any fixed perfect matching on d vertices. For $m \geq 3$, let $P_1, \ldots, P_m$ be independent uniformly random $d \times d$ permutation matrices. Then with nonzero probability:*

$$\lambda_2 \left(\sum_{i=1}^{m} P_i M P_i^T \right) < 2\sqrt{m-1}.$$

Since we only prove nonzero bounds on the probabilities, the nonbipartite theorem is a logical consequence of the bipartite one. We describe it here because its proof is substantially easier and contains most of the main ideas. Note that Theorem I.2 does not produce Ramanujan graphs because it does not guarantee any control of the least eigenvalue λ_d.

We remark that as they are unions of independent matchings, the graphs we produce may have multiple edges between two vertices. Thus, they are (strictly speaking) multigraphs, and do not subsume the previous results if one insists on simple graphs. However, one would think that it should be more difficult to construct Ramanujan graphs with multiedges than without, so understanding the extent to which this can actually happen would be interesting.

As in our previous work [11], the fact that we can only produce bipartite graphs is a consequence of the fact that the method of interlacing families can only control one eigenvalue at a time; in the bipartite case, this automatically yields both upper and lower bounds, since the eigenvalues are symmetric about zero. In contrast with the technique used in [11], there is no obvious obstruction to being able to compute the expected characteristic polynomials in this paper in polynomial time [4].

B. Related Work and Context

Infinite families of Ramanujan graphs were first shown to exist for $m = p+1$, p a prime, in the seminal work of Margulis and Lubotzky, Phillips and Sarnak [13], [9]. The graphs they produce are Cayley graphs and can be constructed very efficiently, and their analysis relies on deep results from number theory, which is responsible for the "Ramanujan" nomenclature. Friedman [6] showed that a random m-regular graph is almost Ramanujan: specifically, that a union of m perfect matchings has non-trivial eigenvalues bounded by $2\sqrt{m-1}+\epsilon$ with high probability, for every $\epsilon > 0$.

More recently, in [11], we proved the existence of infinite families of m-regular bipartite Ramanujan graphs for every $m \geq 3$ by proving (part of) a conjecture of Bilu and Linial [3] regarding the existence of good 2-lifts of regular graphs. Prior to the present paper, it was not known if there are Ramanujan graphs of every number of vertices. We refer the reader to [8] and [11] for a more detailed discussion of expander graphs, Ramanujan graphs, and 2-lifts. Subsequent to the release of this paper, Hall et al. [7] have used related techniques to show that every m-regular graph has a k-lift which is Ramanujan.

In a different vein, it has been known for much longer that the eigenvalue distributions of random m-regular graphs converge weakly *in the limit* to the spectrum of the infinite m-regular tree. In particular, McKay showed in 1981 [14] that for every *fixed* p, the normalized pth moments of a sequence $\{A_d\}$ of random m-regular graphs of increasing size $d \to \infty$ satisfy:

$$\lim_{d \to \infty} \mathbb{E} \frac{1}{d} \mathrm{tr}(A_d^p) = \int_{-\infty}^{\infty} x^p d\mu_m(x), \tag{1}$$

where $\mu_m(x)$ is a density (known as the Kesten–McKay law) supported on the interval $[-2\sqrt{m-1}, 2\sqrt{m-1}]$. Notice that this notion of convergence is too weak to yield information about the extreme eigenvalues of A_d for any fixed d. We remark that Friedman's result is based on a much more delicate calculation which controls the $p = O(\log d)$th moment.

The present work may be seen as connecting the non-asymptotic and asymptotic (i.e., finite d vs. large d limit) sides of the above story, with expected characteristic polynomials playing the mediating role. In particular, by the method of interlacing families, we first reduce the existence of Ramanujan graphs for any fixed size and degree to an analysis of the roots of a single expected characteristic polynomial. We believe that many of the ideas in this paper lie in a more general strategy of deriving a relationship between finite extrema and asymptotic extrema. Our result shows that the largest root of some expected characteristic polynomial of degree d lies inside the spectral radius of the limiting eigenvalue distribution of random graphs as $d \to \infty$. In our case, the most precise techniques for computing these limits come from Free Probability theory, so by forming our expected characteristic polynomials in a way that mimics the operations of Free Probability, we are able to achieve identical bounds (in this case, $2\sqrt{m-1}$ from (1)). Since the relevant operations in Free Probability are known as *free convolutions*, we refer to the corresponding operations on polynomials as "finite free convolutions."

C. Outline of the Paper

The proofs of both of our theorems follow the same strategy and consist of three steps. In each step we present the simpler non-bipartite case first, and then indicate the modifications required for the bipartite case.

First, we show that the expected characteristic polynomials of the random graphs we are interested in are real rooted and come from interlacing families (reviewed in Section II-A), which reduces our existence theorems to analyzing the roots of these polynomials. This is achieved in Section III by decomposing the random permutations used to generate these expected polynomials into swaps acting on two vertices at a time, and showing that such swaps correspond to linear transformations which preserve real-rootedness properties of certain multivariate polynomials. Theorem III.3 implies that if A and B are symmetric matrices, then the expected characteristic polynomial of $A + PBP^T$ is real rooted for a random permutation matrix P. We remark that this argument is completely elementary and self-contained, and unlike [11], [12] does not appeal to any results from the theory of real stable or hyperbolic polynomials. In the process, we introduce a class of "determinant-like" polynomials which may be of independent interest.

Next, in Section IV we derive a closed-form formula for the expected characteristic polynomial of a sum of randomly permuted regular graphs. We begin by proving that the expected characteristic polynomials over random permutations can be replaced by expected characteristic polynomials over random orthogonal matrices. This may be seen as a "quadrature" (or derandomization) statement, which says that these characteristic polynomials are not able to distinguish between the set of permutation matrices and the set of orthogonal matrices; essentially this happens because determinants are multilinear, which causes certain restrictions of them to have very low degree Fourier coefficients. This component of the proof may also be of independent interest.

Finally, we appeal to machinery developed in our companion paper [10], which studies the structure of expected characteristic polynomials over random orthogonal matrices. In particular, such polynomials may be expressed in terms of a simple (and explicitly computable) convolution operation on characteristic polynomials, which we call the finite free additive convolution. In this framework, the characteristic polynomial of a union of m random matchings is simply the m-wise convolution of the characteristic polynomial of a single matching. By applying strong bounds on the roots of these convolutions derived in [10], we obtain the desired Ramanujan bound of $2\sqrt{m-1}$. The requisite material regarding free convolutions is introduced in Sections II-B and II-C.

These three ingredients are combined in Section V to complete the proofs of Theorems I.1 and I.2.

II. Preliminaries

A. Interlacing Families

Showing that a random matrix has small eigenvalues with nonzero probability is a special case (by considering characteristic polynomials) of the more generic problem of showing that some polynomial from a collection must have small roots. The method of interlacing families is a device which allows one to reach the latter conclusion by studying the roots of the average of the polynomials in such a collection. The power of the method stems from the fact that averaging the coefficients is easier and amenable to different algebraic tools than averaging the roots (which are highly nonlinear in the coefficients) directly, and sometimes yields significantly sharper bounds.

The known sufficient conditions for the method to apply all involve real-rootedness properties of certain convex combinations of the polynomials under consideration. We recall the following theorem from [12], stated here in the slightly different language of product distributions.

Theorem II.1 (Interlacing Families)**.** *Suppose $\{f_\omega(x)\}_{\omega\in\{0,1\}^m}$ is a family of real-rooted polynomials of the same degree n with positive leading coefficient, such that*

$$E_\mu(x) := \mathbb{E}_{\omega\sim\mu} f_\omega(x)$$

is real-rooted for every product distribution $\mu = \mu_1 \otimes \cdots \otimes \mu_m$ on $\Omega = \{0,1\}^m$. Then for every $k = 1, \ldots, n$ and every such μ, there is some $\omega_k \in \Omega$ such that

$$\lambda_k(f_{\omega_k}) \leq \lambda_k(E_\mu),$$

where λ_k denotes the kth largest root of a real-rooted polynomial.

For real rooted polynomials f and g, we write $g \longrightarrow f$ if the roots of f and g interlace and the largest root of f is at least as big as the largest root of g. We will use the following elementary facts about interlacing and real-rootedness, which may be found in [5].

Lemma II.2. *If g has degree one less than f and both are real-rooted, then*

1 *$g \longrightarrow f$ if and only if $f + \alpha g$ is real-rooted for all $\alpha \in \mathbb{R}$.*

If, in addition, both f and g have positive leading coefficient, then

2 *$g \longrightarrow f$ implies that $f \longrightarrow f - g$.*

If f_1 and f_2 are real-rooted of the same degree and have positive leading coefficients, then $f_1 + \alpha f_2$ is real-rooted for all $\alpha \geq 0$ if and only if f_1 and f_2 have a common interlacing — i.e. , there is a third polynomial that interlaces both of them.

We refer the interested reader to [11], [12] for a more thorough introduction to interlacing families.

B. Finite Free Convolutions of Polynomials

To analyze the expected characteristic polynomials of the random graphs we consider, we will need the notion of a *finite free convolution* of two polynomials, developed in our companion paper [10].

One way to motivate this notion is the following. Recall that the distribution of the sum of two independent scalar random variables $X + Y$ is the convolution of the individual distributions. Similarly, one can ask about the eigenvalue distribution of a sum of independent random *matrices* $A + B$; the latter problem does not have a simple answer in general, since the eigenvalues of a sum of matrices depend in a nonlinear way on the relative positions of their eigenvectors. The critical observation in our context is that nonetheless, the expected characteristic polynomials of certain sums of independent random matrices depend *linearly* on the expected characteristic polynomials of their summands, in a way that is not that different from the convolution of scalar random variables. The finite free convolution is the bilinear operation that implements this fact.

The finite free convolution was inspired by Voiculescu's free convolution [18] in Free Probability theory (hence the name). The connection lies in the fact that the spectral distribution μ_{A+B} of the sum of two "freely independent" operators A, B depends on the individual spectral distributions μ_A and μ_B but *not* their eigenvectors.

We denote the characteristic polynomial of a matrix by:

$$\chi_x(A) := \det(xI - A).$$

Definition II.3 (Symmetric Additive Convolution)**.** Let $p(x) = \chi_x(A)$ and $q(x) = \chi_x(B)$ be two real-rooted polynomials, for some symmetric $d \times d$ matrices A and B. The *symmetric additive convolution* of p and q is defined as:

$$p(x) \boxplus_d q(x) = \mathop{\mathbb{E}}_{Q} \chi_x\left(A + QBQ^T\right),$$

where the expectation is taken over random orthogonal matrices Q sampled according to the Haar measure on $\mathcal{O}(d)$, the group of d-dimensional orthonormal matrices.

Note that this is a well-defined operation on polynomials because the distribution of the eigenvalues of $A + QBQ^T$ depends only on the eigenvalues of A and the eigenvalues of B, which are the roots of p and q.

To handle the bipartite case, we will require the following two-sided variant of the above, which yields singular values rather than eigenvalues[2].

Definition II.4 (Asymmetric Additive Convolution)**.** Let $p(x) = \chi_x\left(AA^T\right)$ and $q(x) = \chi_x\left(BB^T\right)$ be two real-rooted polynomials with nonnegative roots, for some arbitrary (not necessarily symmetric) $d \times d$ matrices A and B. The *asymmetric additive convolution* of p and q is defined as

$$p(x) \boxplus\!\!\boxplus_d q(x) = \mathop{\mathbb{E}}_{Q,R} \chi_x\left((A + QBR^T)(A + QBR^T)^T\right),$$

where Q and R are independent random orthogonal matrices sampled uniformly from $\mathcal{O}(d)$.

When dealing with a possibly asymmetric $d \times d$ matrix M, we will frequently consider the *dilation*

$$\begin{bmatrix} 0 & M \\ M^T & 0 \end{bmatrix},$$

which is by construction a symmetric $2d \times 2d$ matrix. We will refer to a matrix of this type as a *bipartite* matrix. It is easy to see that its eigenvalues are symmetric about 0 and are equal to $\pm\lambda_1(MM^T)^{1/2}, \ldots, \pm\lambda_d(MM^T)^{1/2}$, i.e., in absolute value to the singular values of M. This correspondence also gives the useful identity

$$\mathbb{S}\chi_x\left(MM^T\right) = \chi_x\left(\begin{bmatrix} 0 & M \\ M^T & 0 \end{bmatrix}\right), \tag{2}$$

where the operator $\mathbb{S}$ is defined by

$$(\mathbb{S}p)(x) := p(x^2).$$

With this notation in hand, we can alternately express the asymmetric additive convolution as

$$\mathbb{S}(p(x) \boxplus\!\!\boxplus_d q(x)) = \mathop{\mathbb{E}}_{Q,R} \chi_x\left(\begin{bmatrix} 0 & A \\ A^T & 0 \end{bmatrix} + \begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix}\begin{bmatrix} 0 & B \\ B^T & 0 \end{bmatrix}\begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix}^T\right). \tag{3}$$

Explicit, polynomial time computable formulas for the additive convolutions in terms of the coefficients of p and q may be found in Theorems 1.1 and 1.3 of [10]. For this work, we only require the following important consequences of these formulas, also established in [10]. We will occasionally drop the subscripts in $\boxplus_d$ and $\boxplus\!\!\boxplus_d$ when they are clear from context.

[2]The asymmetric additive convolution can be used with rectangular matrices as well, but we will not need such generality in this paper.

Lemma II.5 (Properties of $\boxplus$ and $\boxplus\!\!\boxplus$)**.** 1) *If $p(x)$ and $q(x)$ are real-rooted then $p(x) \boxplus_d q(x)$ is also real-rooted.*

2) *If $p(x)$ and $q(x)$ are real-rooted with all roots nonnegative, then $p(x) \boxplus\!\!\boxplus_d q(x)$ is also real-rooted with all roots nonnegative.*

3) *The operations $\boxplus_d$ and $\boxplus\!\!\boxplus_d$ are bilinear (in the coefficients of the polynomials on which they operate) and associative.*

Proof: (1) and (2) are Theorems 1.2 and 1.4 of [10], and bilinearity follows immediately from Theorems 1.1 and 1.3 of [10]. To see associativity, let $p(x) = \chi_x(A), q(x) = \chi_x(B)$ and $r(x) = \chi_x(C)$, and observe that

$$\begin{aligned}(p(x) \boxplus q(x)) \boxplus r(x) &= \left(\mathop{\mathbb{E}}_{Q}\mathop{\mathbb{E}}_{R} \chi_x\left(QAQ^T + RBR^T\right)\right) \boxplus \chi_x(C) \\ &= \mathop{\mathbb{E}}_{Q}\mathop{\mathbb{E}}_{R}\left(\chi_x\left(QAQ + RBR^T\right) \boxplus \chi_x(C)\right) \quad \text{by bilinearity} \\ &= \mathop{\mathbb{E}}_{Q}\mathop{\mathbb{E}}_{R}\mathop{\mathbb{E}}_{W} \chi_x\left(QAQ + RBR^T + WCW^T\right),\end{aligned}$$

for random orthogonal matrices Q, R, W. The same argument shows that this is also equal to $p(x) \boxplus (q(x) \boxplus r(x))$.

An analogous argument using the formula (3) shows that $\boxplus\!\!\boxplus$ is also associative. ∎

Applying the above lemma inductively allows one to write

$$\mathop{\mathbb{E}}_{Q_1,\ldots,Q_m} \chi_x\left(\sum_{i=1}^{m} Q_i A_i Q_i^T\right) = \chi_x(A_1) \boxplus \chi_x(A_2) \boxplus \cdots \boxplus \chi_x(A_m) \tag{4}$$

for $m \geq 3$ matrices $A_1, \ldots, A_m$.

C. Cauchy Transforms

The device that we use to analyze the roots of finite free convolutions of polynomials is the Cauchy transform. This is the same (up to normalization) as the Stieltjes transform and the "barrier function" of [2], [11], [12]. The methods below are different from those used to study the mixed characteristic polynomials of [12], and the bounds we obtain are strictly stronger than those produced by the original "barrier method" argument introduced in [2] (which is off by a factor of two in this setting).

Definition II.6 (Cauchy Transform)**.** The *Cauchy transform* of a polynomial $p(x)$ with roots $\lambda_1, \ldots, \lambda_d$ is defined to be the function

$$\mathcal{G}_p(x) = \frac{1}{d}\sum_{i=1}^{d} \frac{1}{x - \lambda_i} = \frac{1}{d}\frac{p'(x)}{p(x)}.$$

We define the *inverse Cauchy transform* of p to be

$$\mathcal{K}_p(w) = \max\{x : \mathcal{G}_p(x) = w\}.$$

Note that the Cauchy transform has poles at the roots of p, and when all the roots λ_i of p are real, $\mathcal{G}_p(x)$ is monotone decreasing for x greater than the largest root. Thus, $\mathcal{K}_p(w)$ is the unique value of x that is larger than all the λ_i for which $\mathcal{G}_p(x) = w$. In particular, it is an upper bound on the largest root of p, and approaches the largest root as $w \to \infty$.

Our bounds on the expected characteristic polynomials of random graphs are a consequence of the following two theorems, which are proved in [10].

Theorem II.7 (Theorem 1.7 of [10])**.** *For real-rooted degree d polynomials p and q and $w > 0$,*

$$\mathcal{K}_{p \boxplus_d q}(w) \leq \mathcal{K}_p(w) + \mathcal{K}_q(w) - 1/w.$$

The above theorem is a strengthening of the univariate barrier function argument for characteristic polynomials introduced in [2]. This may be seen by taking $q(x) = \chi_x(B) = x^{d-1}(x-d)$, which corresponds to a rank one matrix $B = vv^T$ with trace equal to d. It is easy to check that in this case $p(x) \boxplus q(x) = p(x) - p'(x)$. We remark that bounds of this type are generally much better than the trivial triangle inequality on λ_{max}. The is due to the fact that $\mathcal{K}_p(w)$ takes into account the location of all roots (rather than just the largest one), creating what can be viewed as a "soft maximum" on the roots.

We remark that the inequality in Theorem II.7 is inspired by an *equality* regarding inverse Cauchy transforms of limiting spectral distributions of certain random matrix models arising in Free Probability theory; we refer the interested reader to [10] for a more detailed discussion. To analyze the case of bipartite random graphs, we will need the corresponding inequality for the asymmetric convolution.

Theorem II.8 (Theorem 1.8 of [10])**.** *For degree d polynomials p and q having only nonnegative real roots,*

$$\mathcal{K}_{\mathbb{S}(p \boxplus_d q)}(w) \leq \mathcal{K}_{\mathbb{S}p}(w) + \mathcal{K}_{\mathbb{S}q}(w) - 1/w.$$

III. Interlacing for Permutations

In this section, we show that the expected characteristic polynomials obtained by averaging over certain random permutation matrices form interlacing families. The class of random permutations which have this property are those that are products of independent random swaps, which we now formally define.

Definition III.1 (Random Swap)**.** A *random swap* is a matrix-valued random variable which is equal to a transposition of two (fixed) indices s, t with probability α and equal to the identity with probability $(1-\alpha)$, for some $\alpha \in [0,1]$.

Definition III.2 (Realizability by Swaps)**.** A matrix-valued random variable P supported on permutation matrices is *realizable by swaps* if there are random swaps $S_1, \ldots, S_N$ such that the distribution of P is the same as the distribution of the product $S_N S_{N-1} \ldots S_2 S_1$.

For example, we show in Lemma III.5 below that a uniformly random permutation matrix is realizable by swaps.

The main result of this section is that expected characteristic polynomials over products of random swaps are always real-rooted. These polynomials play a role analogous to that of mixed characteristic polynomials in [11], [12].

Theorem III.3. *Let $A_1, \ldots, A_m$ be symmetric $d \times d$ matrices and let $\{S_{ij}\}_{i \leq m, j \leq N}$ be independent (not necessarily identical) random swaps. Then the expected characteristic polynomial*

$$\mathbb{E} \det \left(tI - \sum_{i=1}^{m} \left(\prod_{j=N}^{1} S_{ij} \right) A_i \left(\prod_{j=1}^{N} S_{ij}^T \right) \right) \tag{5}$$

is real-rooted.

An immediate consequence of Theorems III.3 and II.1, applied to the family of polynomials indexed by all possible values of the swaps S_{ij}, is the following existence result.

Theorem III.4 (Interlacing Families for Permutations)**.** *Suppose $A_1, \ldots, A_m$ are symmetric $d \times d$ matrices, and $P_1, \ldots, P_m$ are independent random permutations realizable by swaps. Then, for every $k \leq d$:*

$$\lambda_k \left(\sum_{i=1}^{m} P_i A_i P_i^T \right) \leq \lambda_k \left(\mathbb{E} \chi_x \left(\sum_{i=1}^{m} P_i A_i P_i^T \right) \right),$$

with nonzero probability.

Theorem III.4 is useful because the uniform distribution on permutations and its bipartite version, which we use to generate our random graphs, are realizable by swaps.

Lemma III.5. *Let P and S be uniformly random $d \times d$ permutation matrices. Both P and $P \oplus S$ are realizable by swaps, where $P \oplus S = \begin{pmatrix} P & 0 \\ 0 & S \end{pmatrix}$ is the direct sum of P and S.*

Proof: We will establish the claim for P first. We proceed inductively. Let M_2 be a random swap which swaps e_1 and e_2 with probability $1/2$, and for $k > 2$ let

$$M_k = M_{k-1} S_{1k} M_{k-1},$$

where S_{1k} swaps e_1 and e_k with probability $1/k$.

Let $v = (1, 2, 3, \ldots, d)^T$. By induction, assume that the first $k-1$ coordinates of $M_{k-1}v$ are in uniformly random order; in particular, that $(M_{k-1}v)(1)$ is a random element of $\{1, \ldots, k-1\}$. This means that:

- With probability $1/k$: $(M_{k-1}S_{1k}M_{k-1}v)(k) = k$ and the remaining indices contain a random permutation of $\{1, \ldots, k-1\}$.
- With probability $1 - 1/k$: $(M_{k-1}S_{1k}M_{k-1}v)(k)$ is a uniformly random element $j \in \{1, \ldots, k-1\}$ and the remaining indices contain a random permutation of $\{1, \ldots, k\} \setminus \{j\}$.

Thus, M_k is uniformly random on $\{1, \ldots, k\}$, and by induction $M_d = P$.

For $P \oplus S$, we use the above argument to realize $P \oplus I$ and $I \oplus S$ separately and then multiply them. ∎

The rest of this section is devoted to proving Theorem III.3. This is achieved by showing that the polynomials in (5) are univariate restrictions of certain nice multivariate polynomials. The relevant notion is the following.

Definition III.6 (Determinant-like Polynomials)**.** A homogeneous polynomial $H(X_1, \ldots, X_m)$ of degree d in the entries of m symmetric $d \times d$ matrices $X_1, \ldots, X_m$ is called *determinant-like* if it has the following two properties.

Hyperbolicity. The univariate restrictions

$$q(t) = H(tI - A_1, \ldots, tI - A_m)$$

have positive leading coefficient and are real-rooted for all symmetric $A_1, \ldots, A_m$.

This condition is known as *hyperbolicity* of the polynomial $H(X_1, \ldots, X_m)$ with respect to the point $(I, I, \ldots, I)$. We do not discuss the notion of hyperbolicity further, since the self-contained definition above suffices for this paper. We point the interested reader to [17] for a detailed discussion of the theory.

Rank-1 Linearity. For every vector v, index $i \leq m$, and real number s, we have

$$H(X_1, X_2, \ldots, X_i + svv^T, \ldots, X_m) = H(X_1, \ldots, X_m) + sD_{i,vv^T}H(X_1, \ldots, X_m)$$

where

$$D_{i,vv^T}H(X_1, \ldots, X_m) = \left(\frac{\partial}{\partial s} H(X_1, \ldots, X_i + svv^T, \ldots, X_m) \right) \Big|_{s=0}$$

is the directional derivative of H in direction $(0, \ldots, vv^T, \ldots, 0)$, where vv^T appears in the ith position. Note that $D_{i,vv^T}H(X_1, \ldots, X_m)$ is homogeneous of degree $d-1$.

An important example of a determinant-like polynomial is the determinant of a sum of matrices:

$$H(X_1, \ldots, X_m) = \det(X_1 + \cdots + X_m).$$

Hyperbolicity follows from the fact that

$$H(tI - A_1, \ldots, tI - A_m) = \det(mtI - A_1 - \cdots - A_m)$$

is the characteristic polynomial of a symmetric matrix. Rank-1 linearity can be seen to follow from the invariance of the determinant under change of basis and its linearity with respect to matrix entries. Alternatively, one can prove it by using the matrix determinant lemma, which tells us

$$\det(X_1 + svv^T + \cdots + X_m) = \det(X_1 + \cdots + X_m) + s\langle vv^T, \det(X_1 + \cdots + X_m)(X_1 + \cdots + X_m)^{-1}\rangle.$$

The crux of the proof of Theorem III.3 lies in the fact that random swaps define linear operators which preserve the property of being determinant-like.

Lemma III.7 (Random swaps preserve determinant-likeness)**.** *If $H(X_1, \ldots, X_m)$ is determinant-like, then for every $i \leq m$ and random swap S, the polynomial*

$$\mathbb{E}_S H(X_1, \ldots, SX_iS^T, \ldots, X_m)$$

is determinant-like.

Before proving this lemma, we record some preliminary facts about determinant-like polynomials.

Lemma III.8 (Rank-1 updates interlace)**.** *Suppose $H(X_1, \ldots, X_m)$ is determinant-like. Then for every vector v and symmetric matrices $A_1, \ldots, A_m$ we have*

$$H(tI - A_1, \ldots, tI - A_m) \longrightarrow H(tI - A_1, \ldots, tI - A_i - vv^T, \ldots tI - A_m),$$

where $\longrightarrow$ denotes interlacing, pointing to the polynomial with the largest root.

Proof: Assume without loss of generality that $i = 1$. By rank-1 linearity, for every $s \in \mathbb{R}$

$$H(tI - A_1 - svv^T, \ldots, tI - A_m) = H(tI - A_1, \ldots, tI - A_m) - sD_{vv^T}H(tI - A_1, \ldots, tI - A_m).$$

By the hyperbolicity of H, we know that this is real rooted when viewed as a univariate polynomial in t. Since $D_{1,vv^T}H$ is of degree one less than H, the first part of Lemma II.2 implies that

$$D_{1,vv^T}H(tI - A_1, \ldots, tI - A_m) \longrightarrow H(tI - A_1, \ldots, tI - A_m),$$

which in turn by the second part of Lemma II.2 gives

$$\begin{aligned} H(tI - A_1, \ldots, tI - A_m) &\longrightarrow H(tI - A_1, \ldots, tI - A_m) - D_{1,vv^T}H(A_1 - tI, \ldots, A_m - tI) \\ &= H(tI - A_1 - vv^T, \ldots, tI - A_m), \end{aligned}$$

as desired. ∎

Lemma III.9 (Permutations preserve rank-1 linearity)**.** *(1) If Π is a permutation matrix and $H(X_1, \ldots, X_m)$ is rank-1 linear then $H(\Pi X_1 \Pi^T, X_2, \ldots, X_m)$ is also rank-1 linear. (2) If H and F are rank-1 linear then so is $H + F$.*

Proof: (1) is true because the set of rank one matrices is invariant under conjugation by permutations. (2) holds because D_{i,vv^T} is a linear operator. ∎

We will also need the following elementary observation, which says that random swaps correspond to trace zero rank two updates. This is the structural property which causes interlacing to occur.

Lemma III.10. *If σ is a transposition and A is symmetric then $A - \sigma A \sigma^T$ has rank 2 and trace 0.*

Proof: Assume without loss of generality that σ swaps the first two coordinates. Then by symmetry the difference $A - \sigma A \sigma^T$ has entries

$$\begin{bmatrix} a_{11} - a_{22} & a_{12} - a_{21} & a_{13} - a_{23} & a_{14} - a_{24} & \ldots \\ a_{21} - a_{12} & a_{22} - a_{11} & a_{23} - a_{13} & a_{24} - a_{14} & \ldots \\ a_{31} - a_{32} & a_{32} - a_{31} & 0 & \ldots & \\ a_{41} - a_{42} & a_{42} - a_{41} & 0 & \ldots & \\ \ldots & & & & \end{bmatrix} = \begin{bmatrix} \alpha & \beta & v^T \\ -\beta & -\alpha & -v^T \\ v & -v & 0_{n-2} \end{bmatrix}$$

for some numbers α, β and some column vector v of length $d-2$. If $\alpha \neq \beta$ then the sum of the first two rows is equal to $(c, -c, 0, \ldots, 0)$ for some $c \neq 0$, and every other row is a scalar multiple of this. On the other hand, if $\alpha = \beta$ then the first two rows are linearly dependent, and all of the other rows are multiples of $(1, -1, 0, \ldots, 0)$. ∎

We can now complete the proof of Lemma III.7

Proof of Lemma III.7: Assume H is determinant-like, and let S be a random swap, equal to some transposition σ with probability α and the identity with probability $(1-\alpha)$. We will show that

$$F(X_1, \ldots, X_m) = (1-\alpha)H(X_1, \ldots, X_m) + \alpha H(X_1, \ldots, \sigma X_i \sigma^T, \ldots, X_m),$$

is hyperbolic and rank-1 linear. It is clear that $F(X_1, \ldots, X_m)$ is homogeneous since swaps and convex combinations preserve homogeneity. Lemma III.9 implies that rank-1 linearity is also preserved, so all that remains is hyperbolicity. Assume without loss of generality that $i = 1$ and consider any univariate restriction along $(I, I, \ldots, I)$:

$$F(tI - A_1, \ldots, tI - A_m) = (1-\alpha)H(tI - A_1, \ldots, tI - A_m) + \alpha H(tI - \sigma A_1 \sigma^T, \ldots, tI - A_m). \quad (6)$$

We need to show that this has all real roots. Observe that the second polynomial may be written as

$$H(tI - A_1 - aa^T + bb^T, \ldots, tI - A_m),$$

for some vectors a and b , since $\sigma A_1 \sigma^T - A_1$ is rank two and trace zero by Lemma III.10. Since H is determinant-like, Lemma III.8 tells us that

$$H(tI - A_1 + bb^T, \ldots, tI - A_m) \longrightarrow H(tI - A_1 - aa^T + bb^T, \ldots, tI - A_m)$$

and

$$H(tI - A_1 + bb^T, \ldots, tI - A_m) \longrightarrow H(tI - A_1, \ldots, tI - A_m),$$

whence the two polynomials on the right hand side of (6) have a common interlacing. Lemma II.2 then implies that their convex combination must be real-rooted, and the claim is proved. ∎

Applying Lemma III.7 inductively yields Theorem III.3.

Proof of Theorem III.3: Applying Lemma III.7 nN times (once for every swap S_{ij}) starting with $H(X_1, \ldots, X_m) = \det(\sum_i X_i)$ tells us that

$$\mathbb{E}_{S_{1N}} \ldots \mathbb{E}_{S_{11}} \mathbb{E}_{S_{2N}} \ldots \mathbb{E}_{S_{n1}} \det \left(\sum_{i=1}^{n} \left(\prod_{j=N}^{1} S_{ij} \right) X_i \left(\prod_{j=1}^{N} S_{ij}^T \right) \right)$$

is determinant-like. Considering the restriction $X_i = (t/m)I - A_i$ finishes the proof.

∎

IV. Quadrature

In this section, we show that the expected characteristic polynomials we are interested in are free convolutions of the characteristic polynomials of perfect matchings, after the trivial eigenvalues corresponding to the all ones vector are removed. This gives us explicit formulas for these polynomials, and more importantly (since we understand the behavior of roots under free convolutions) a way of bounding their roots. We begin by showing how to do this for the symmetric case, which is more transparent and contains all the main ideas. In Section IV-B we derive the result for the bipartite case as a corollary of the result for the symmetric case.

A. Quadrature for Symmetric Matrices

The following theorem gives an explicit formula for the expected characteristic polynomial of the sum of two symmetric matrices with constant row sums when the rows and columns of one of the matrices is randomly permuted. This can be used to compute the expected characteristic polynomial of the Laplacian matrix of the sum of two graphs when one is randomly permuted. In this paper, we use the result to compute the expected characteristic polynomial of the adjacency matrix when both graphs are regular.

Theorem IV.1. *Suppose A and B are symmetric $d \times d$ matrices with $A\mathbf{1} = a\mathbf{1}$ and $B\mathbf{1} = b\mathbf{1}$. Let $\chi_x(A) = (x-a)p(x)$ and $\chi_x(B) = (x-b)q(x)$. Then,*

$$\mathbb{E}_P \chi_x\left(A + PBP^T\right) = (x-(a+b))p(x) \boxplus_{d-1} q(x), \tag{7}$$

where P is a uniformly random permutation.

We begin by writing (7) in a more concrete form. Observe that all of the matrices A, B, P have $\mathbf{1}$ as a left and right eigenvector, which means that there is an orthogonal change of basis V (for concreteness, mapping $\mathbf{1}$ to the standard basis vector e_n) that simultaneously block diagonalizes all of them:

$$VAV^T = \hat{A} \oplus a, \quad VBV^T = \hat{B} \oplus b, \quad VPV^T = \hat{P} \oplus 1, \tag{8}$$

where $\hat{A} \oplus a$ denotes the direct sum

$$\begin{bmatrix} \hat{A} & 0 \\ 0 & a \end{bmatrix}.$$

Since the determinant is invariant under change of basis, we may write

$$\begin{aligned} \mathbb{E}_P \det(xI - A - PBP^T) &= \mathbb{E}_P \det(xI - VAV^T - (VPV^T)(VBV^T)(VP^TV^T)) \\ &= \mathbb{E}_{\hat{P}} \det(xI - (\hat{A} \oplus a) - (\hat{P} \oplus 1)(\hat{B} \oplus b)(\hat{P}^T \oplus 1)) \\ &= (x-a-b)\mathbb{E}_{\hat{P}} \det(xI - \hat{A} - \hat{P}\hat{B}\hat{P}^T). \end{aligned} \tag{9}$$

Notice also that $p(x) = \chi_x\left(\hat{A}\right)$ and $q(x) = \chi_x\left(\hat{B}\right)$, so

$$p(x) \boxplus_{d-1} q(x) = \mathbb{E}_Q \det(xI - \hat{A} - Q\hat{B}Q^T),$$

where Q is a (Haar) random $(d-1) \times (d-1)$ orthogonal matrix. Thus, (7) is equivalent to showing that

$$\mathbb{E}_{\hat{P}} \det(xI - \hat{A} - \hat{P}\hat{B}\hat{P}^T) = \mathbb{E}_Q \det(xI - \hat{A} - Q\hat{B}Q^T), \tag{10}$$

for all $(d-1)\times(d-1)$ symmetric matrices $\hat{A}, \hat{B}$. Note that for any permutation P, the orthogonal transformation $\hat{P}$ correspondingly permutes $\hat{e_1}, \ldots, \hat{e_n}$, the projections orthogonal to $\mathbf{1}$ of the standard basis vectors $e_1, \ldots, e_d$, embedded in $\mathbb{R}^{d-1}$. Since these are the vertices of a regular simplex with d vertices in $\mathbb{R}^{d-1}$ centered at the origin, we interpret the $\hat{P}$ as elements of the symmetry group of this simplex. We denote this subgroup of $\mathcal{O}(d-1)$ by $\mathcal{A}_{d-1}$.

Since there is no longer any assumption on $\hat{A}$ or $\hat{B}$ other than symmetry, we may absorb the xI term into $\hat{A}$ in (10), and we see that it is sufficient to establish the following.

Theorem IV.2 (Quadrature Theorem)**.** *For symmetric $d \times d$ matrices A and B,*

$$\mathop{\mathbb{E}}_{P \in \mathcal{A}_d} \det(A + PBP^T) = \mathop{\mathbb{E}}_{Q \in \mathcal{O}(d)} \det(A + QBQ^T). \tag{11}$$

It is easy to see that the theorem will follow if we can show that the left hand side of (11) is invariant under right multiplication of P by orthogonal matrices.

Lemma IV.3 (Invariance Implies Quadrature)**.** *Let f be a function from $\mathcal{O}(d)$ to $\mathbb{R}$ and let H be a finite subgroup of $\mathcal{O}(d)$. If*

$$\mathop{\mathbb{E}}_{P\in H} f(P) = \mathop{\mathbb{E}}_{P\in H} f(PQ_0). \tag{12}$$

for all $Q_0 \in \mathcal{O}(d)$, then

$$\mathop{\mathbb{E}}_{P\in H} f(P) = \mathop{\mathbb{E}}_{Q\in\mathcal{O}(d)} f(Q), \tag{13}$$

where Q is chosen according to Haar measure and P is uniform on H.

Proof:

$$\begin{aligned}\mathop{\mathbb{E}}_{Q\in\mathcal{O}(d)} f(Q) &= \mathop{\mathbb{E}}_{Q\in\mathcal{O}(d)} \mathop{\mathbb{E}}_{P\in H} f(PQ) = \mathop{\mathbb{E}}_{P\in H} \mathop{\mathbb{E}}_{Q\in\mathcal{O}(d)} f(PQ)\\ &= \mathop{\mathbb{E}}_{P\in H} \mathop{\mathbb{E}}_{Q\in\mathcal{O}(d)} f(P) = \mathop{\mathbb{E}}_{P\in H} f(P),\end{aligned}$$

as desired. ∎

We will prove Theorem IV.2 by showing that $f(P) = \det(A + PBP^T)$ satisfies (12). We will achieve this by demonstrating that f is invariant under certain elementary orthogonal transformations acting on 3-faces of the regular simplex, which generate all orthogonal transformations. Let us fix some notation to precisely describe these elementary transformations.

Given three vertices $\hat{e_i}, \hat{e_j}, \hat{e_k}$ of the regular simplex, let $\mathcal{A}_{i,j,k}$ denote the subgroup of $\mathcal{A}_d$ consisting of permutations of $\hat{e_i}, \hat{e_j}, \hat{e_k}$ which leave all of the other vertices fixed. Let $\mathcal{O}_{i,j,k}$ denote the subgroup of $\mathcal{O}(d)$ acting on the two dimensional linear subspace parallel to the affine subspace through these three vertices, and leaving the orthogonal subspace fixed. Note that $\mathcal{A}_{i,j,k}$ is a subgroup of $\mathcal{O}_{i,j,k}$, and that these groups are isomorphic to $\mathcal{A}_2$ and $\mathcal{O}(2)$, respectively.

The heart of the proof lies in the following lemma, which implies by Lemma IV.3 that the polynomials we are interested in are not able to distinguish between the uniform distributions on $\mathcal{A}_2$ and $\mathcal{O}(2)$. The reason for this is that these polynomials have very low degree (at most two) in the entries of any orthogonal matrix Q acting on a two-dimensional subspace, a fact which is essentially a consequence of the multilinearity of the determinant. The argument below is similar to the proof of Lemma 2.7 in [10].

Lemma IV.4 (Invariance for $\mathcal{A}_2$)**.** *If A and B are symmetric $d \times d$ matrices, then for every $Q_0 \in \mathcal{O}(2)$,*

$$\mathop{\mathbb{E}}_{P\in\mathcal{A}_2} \det(A + (P \oplus I_{d-2})B(P \oplus I_{d-2})^T) = \mathop{\mathbb{E}}_{P\in\mathcal{A}_2} \det(A + (PQ_0 \oplus I_{d-2})B(PQ_0 \oplus I_{d-2})^T). \tag{14}$$

Proof: Let $S\mathcal{O}(2)$ be the subgroup of $\mathcal{O}(2)$ consisting of rotation matrices

$$R_\theta = \begin{bmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{bmatrix},$$

and let Z_3 be the subgroup of $\mathcal{A}_2$ consisting of the three rotations $R_\tau, \tau \in T := \{0, 2\pi/3, 4\pi/3\}$. We begin by showing that

$$\mathop{\mathbb{E}}_{P\in Z_3} \det(A + (P \oplus I)B(P \oplus I)^T) = \mathop{\mathbb{E}}_{P\in Z_3} \det(A + (PR_\theta \oplus I)B(PR_\theta \oplus I)^T), \tag{15}$$

for every θ, where I is the $(d-2)$-dimensional identity. Since the elements of Z_3 are themselves rotations, we

can rewrite thrice the right hand side of (15) as

$$\begin{aligned}
&\sum_{\tau\in T}\det(A+(R_\tau R_\theta\oplus I)B(R_\tau R_\theta\oplus I)^T)\\
&=\sum_{\tau\in T}\det(A+(R_{\tau+\theta}\oplus I)B(R_{\tau+\theta}\oplus I)^T)\\
&=\sum_{\tau\in T}\sum_{k=-2}^{2}c_k e^{ik(\tau+\theta)}\quad\text{for some coefficients } c_k\text{, by Lemma IV.5}\\
&=\sum_{k=-2}^{2}c_k e^{ik\theta}\left(e^{ik0}+e^{ik2\pi/3}+e^{ik4\pi/3}\right)\\
&=3c_0\quad\text{since the terms with } |k|=1,2 \text{ vanish.}
\end{aligned}$$

As this quantity is independent of θ, we can assume $\theta=0$, which gives the left hand side of (15).

To finish the proof, we observe that

$$\begin{aligned}
\mathop{\mathbb{E}}_{P\in\mathcal{A}_2}\det(A+(P\oplus I_{d-2})B(P\oplus I_{d-2})^T)&=\mathop{\mathbb{E}}_{D\in F}\mathop{\mathbb{E}}_{P\in Z_3}\det(A+(PD\oplus I)B(PD\oplus I)^T)\\
&=\mathop{\mathbb{E}}_{D\in F}\mathop{\mathbb{E}}_{P\in Z_3}\det(A+(P\oplus I)(D\oplus I)B(D\oplus I)^T(P\oplus I)^T),
\end{aligned}$$

where F consists of the identity and the reflection across the horizontal axis:

$$F:=\left\{\begin{bmatrix}1&0\\0&1\end{bmatrix},\begin{bmatrix}1&0\\0&-1\end{bmatrix}\right\},$$

and D is chosen uniformly from F.

Thus, the left hand side of (14) is invariant under conjugation of B with the matrices $D\oplus I, D\in F$. Since every $Q_0\in\mathcal{O}(2)$ can be written as $R_\theta D$ for some $D\in F$, and we have already established invariance under $R_\theta\oplus I$ in (15), the lemma is proved.

■

Lemma IV.5 (Determinants are Low Degree in Rank 2 Rotations)**.** *Let A,B be $d\times d$ symmetric matrices. Then there are numbers c_k for $k\in\{-2,-1,0,1,2\}$ so that*

$$\det\left(A+(R_\theta\oplus I_{d-2})B(R_\theta\oplus I_{d-2})^T\right)=\sum_k c_k e^{ik\theta}.$$

Proof: Recall that all 2×2 rotations may be diagonalized as

$$R_\theta=\begin{bmatrix}\cos\theta&\sin\theta\\-\sin\theta&\cos\theta\end{bmatrix}=U\begin{bmatrix}e^{i\theta}&0\\0&e^{-i\theta}\end{bmatrix}U^\dagger,$$

where

$$U=\frac{1}{\sqrt{2}}\begin{bmatrix}1&1\\i&-i\end{bmatrix}$$

is independent of θ. This implies that $(R_\theta\oplus I_{d-2})=VDV^\dagger$ for diagonal D containing $e^{i\theta}$ and $e^{-i\theta}$ in the upper right 2×2 block and ones elsewhere, with V independent of θ. Thus, we see that

$$\begin{aligned}
\det\left(A+(R_\theta\oplus I_{d-2})B(R_\theta\oplus I_{d-2})^T\right)&=\det\left(A(R_\theta\oplus I_{d-2})+(R_\theta\oplus I_{d-2})B\right)\\
&=\det\left(AVDV^\dagger+VDV^\dagger B\right)\\
&=\det\left(V^\dagger AVD+DV^\dagger BV\right)
\end{aligned}$$

Notice that the matrix $M = V^\dagger AVD + DV^\dagger BV$ depends linearly on $e^{i\theta}, e^{-i\theta}$, and that the $e^{i\theta}$ (resp. $e^{-i\theta}$) terms appear only in the first (resp. second) row and column of M, respectively. Since each monomial in the expansion of the determinant contains at most one entry from each row and each column and $e^{i\theta} \cdot e^{-i\theta} = 1$, this implies that no terms of degree higher than two in $e^{i\theta}$ or $e^{-i\theta}$ appear. ∎

Corollary IV.6 (Invariance for $\mathcal{A}_{i,j,k}$)**.** *For every i, j and k,*

$$\mathop{\mathbb{E}}_{P\in\mathcal{A}_{i,j,k}} \det(A + PBP^T) = \mathop{\mathbb{E}}_{Q\in\mathcal{O}_{i,j,k}} \det(A + QBQ^T).$$

Proof: Let V be the orthogonal transformation that maps the affine subspace spanned by the vertices $\hat{e}_i, \hat{e}_j, \hat{e}_k$ to the first two coordinates of $\mathbb{R}^2$, with any one vertex mapped to a multiple of e_1. Conjugation by V maps $\mathcal{A}_{i,j,k}$ to $\mathcal{A}_2 \oplus I_{d-2}$ and $\mathcal{O}_{i,j,k}$ to $\mathcal{O}(2) \oplus I_{d-2}$, abusing notation slightly in the natural way. Since the determinant is invariant under change of basis, Lemma IV.4 tells us that

$$\begin{aligned}
\mathop{\mathbb{E}}_{P\in\mathcal{A}_{i,j,k}} \det(A + PBP^T) &= \mathop{\mathbb{E}}_{P_2\in\mathcal{A}_2} \det(VAV^T + (P_2 \oplus I)VBV^T(P_2 \oplus I)^T) \\
&= \mathop{\mathbb{E}}_{Q_2\in\mathcal{O}(2)} \det(VAV^T + (Q_2 \oplus I)VBV^T(Q_2 \oplus I)^T) \\
&= \mathop{\mathbb{E}}_{Q\in\mathcal{O}_{i,j,k}} \det(A + QBQ^T),
\end{aligned}$$

as desired. ∎

Lemma IV.7 ($\mathcal{O}_{i,j,k}$ generate $\mathcal{O}(d)$)**.** *Given a regular simplex in $\mathbb{R}^d$, the union over i, j, and k of $\mathcal{O}_{i,j,k}$ generates $\mathcal{O}(d)$. In particular, every matrix in $\mathcal{O}(d)$ may be written as a product of a finite number of these matrices.*

Proof: Let Γ_h be the subgroup of $\mathcal{O}(d)$ generated by $\bigcup_{i,j,k\le h} \mathcal{O}_{i,j,k}$. Let $\hat{e}_0, \ldots, \hat{e}_d$ be the vertices of the regular simplex. For $1 \le h \le d$, let E_h be the linear subspace parallel to the affine subspace through $\hat{e}_0, \ldots, \hat{e}_h$. We will prove by induction on h that Γ_h contains the action of the orthogonal group on E_h. The base case is $h = 2$, for which $\mathcal{O}_{0,1,2}$ is precisely the action of the orthogonal group on E_2.

Assuming that we have proved this result for $h-1$, we now prove it for h. To this end, let $u_h = \hat{e}_h$, and let $u_1, \ldots, u_{h-1}$ be arbitrary orthonormal vectors in E_h that are orthogonal to u_h. We will prove that for every orthonormal basis $w_1, \ldots, w_h$ of E_h, there is a $Q \in \Gamma_h$ such that $Qw_i = u_i$ for $1 \le i \le h$.

We first consider the case in which $w_h^T \hat{e}_h \ge 0$. Let F_h denote the 2-dimensional affine subspace spanned by $\{\hat{e}_h, \hat{e}_{h-1}, \hat{e}_{h-2}\}$, and observe that there must be a unit vector $p \in E_h \cap F_h$ with $p^T \hat{e}_h = w_h^T \hat{e}_h$. This follows because the intersection of F_h with the unit sphere in E_h is a circle containing $\{\hat{e}_h, \hat{e}_{h-1}, \hat{e}_{h-2}\}$, $p \mapsto p^T \hat{e}_h$ is a continuous function, and we have $\hat{e}_h^T \hat{e}_h = 1$ and $\hat{e}_{h-1}^T \hat{e}_h = \hat{e}_{h-2}^T \hat{e}_h < 0$. As $\hat{e}_h$ is orthogonal to E_{h-1} and $\hat{e}_h$ is invariant under Γ_{h-1}, the induction hypothesis implies that there must be a $T \in \Gamma_{h-1}$ so that $Tw_h = p$. Moreover, there is an element T_2 of $\mathcal{O}_{h-2,h-1,h}$ that maps p to $\hat{e}_h$. So, their composition $W = T_2 T$ sends w_h to $\hat{e}_h$. Since W is orthogonal, it must send $w_1, \ldots, w_{h-1}$ to E_{h-1}, and so by induction may be composed with a map in Γ_{h-1} that sends $Ww_1, \ldots, Ww_{h-1}$ to $u_1, \ldots, u_{h-1}$ without moving $\hat{e}_h$. The resulting map is the desired Q.

In the case that $w_h^T \hat{e}_h < 0$, we begin by applying a map in Γ_h that sends w_h to a vector that is orthogonal to $\hat{e}_h$ so that we can then apply the previous argument. For example, we can do this by defining p to be one of the two unit vectors in F_h with $p^T \hat{e}_h = -w_h^T \hat{e}_h$. We then apply a map in Γ_{h-1} that sends w_h to $-p$, and then a map in $\mathcal{O}_{h-2,h-1,h}$ that maps p, and thus also $-p$, to a vector orthogonal to $\hat{e}_h$. ∎

Theorem IV.8 (Invariance for $\mathcal{A}_d$)**.** *Let A and B be $d \times d$ matrices, and let*

$$f_{A,B}(Q) = \det\left(A + QBQ^T\right).$$

Then, for all $Q_0 \in \mathcal{O}(d)$,

$$\mathop{\mathbb{E}}_{P\in\mathcal{A}_d} f_{A,B}(P) = \mathop{\mathbb{E}}_{P\in\mathcal{A}_d} f_{A,B}(PQ_0).$$

Proof: We will use the fact that

$$\mathop{\mathbb{E}}_{P\in\mathcal{A}_d} f_{A,B}(P) = \mathop{\mathbb{E}}_{P\in\mathcal{A}_d}\mathop{\mathbb{E}}_{P_2\in\mathcal{A}_{i,j,k}} f_{A,B}(PP_2) = \mathop{\mathbb{E}}_{P\in\mathcal{A}_d}\mathop{\mathbb{E}}_{P_2\in\mathcal{A}_{i,j,k}} f_{P^TAP,B}(P_2).$$

Applying Corollary IV.6 reveals that for every $Q_2 \in \mathcal{O}_{i,j,k}$,

$$\begin{aligned}\mathop{\mathbb{E}}_{P\in\mathcal{A}_d}\mathop{\mathbb{E}}_{P_2\in\mathcal{A}_{i,j,k}} f_{P^TAP,B}(P_2) &= \mathop{\mathbb{E}}_{P\in\mathcal{A}_d}\mathop{\mathbb{E}}_{P_2\in\mathcal{A}_{i,j,k}} f_{P^TAP,B}(P_2Q_2).\\ &= \mathop{\mathbb{E}}_{P\in\mathcal{A}_d}\mathop{\mathbb{E}}_{P_2\in\mathcal{A}_{i,j,k}} f_{A,B}(PP_2Q_2).\\ &= \mathop{\mathbb{E}}_{P\in\mathcal{A}_d} f_{A,B}(PQ_2).\end{aligned}$$

Thus, we conclude that

$$\mathop{\mathbb{E}}_{P\in\mathcal{A}_d} f_{A,B}(P) = \mathop{\mathbb{E}}_{P\in\mathcal{A}_d} f_{A,B}(PQ_2)$$

for every $Q_2 \in \mathcal{O}_{i,j,k}$, for every i, j, k.

Let $Q_0 \in \mathcal{O}(d)$. By Lemma IV.7, there is a sequence of matrices $Q_1, \ldots, Q_m$, each of which is in $\mathcal{O}_{i,j,k}$ for some i, j and k, so that

$$Q_0 = Q_1Q_2\cdots Q_m.$$

By applying the previous equality m times, we obtain

$$\mathop{\mathbb{E}}_{P\in\mathcal{A}_d} f(PQ_0) = \mathop{\mathbb{E}}_{P\in\mathcal{A}_d} f(PQ_1\cdots Q_m) = \mathop{\mathbb{E}}_{P\in\mathcal{A}_d} f(P).$$

■

Proof of Theorem IV.2: Follows from Theorem IV.8 and Lemma IV.3. ■

Proof of Theorem IV.1: Follows from Theorem IV.2, (8), and (9). ■

We conclude the section by recording the obvious extension of Theorem IV.1 to sums of m matrices.

Corollary IV.9. *Let $A_1, \ldots, A_m$ be symmetric $d \times d$ matrices with $A_i\mathbf{1} = a_i\mathbf{1}$ and $\chi_x(A_i) = (x - a_i)p_i(x)$. Then,*

$$\mathop{\mathbb{E}}_{P_1,\ldots,P_m} \chi_x\left(\sum_{i=1}^m P_iA_iP_i^T\right) = \left(x - \sum_{i=1}^m a_i\right) p_1(x) \boxplus \cdots \boxplus p_m(x), \tag{16}$$

where $P_1, \ldots, P_m$ are independent uniformly random permutation matrices.

Proof: Apply a change of basis so that each $A_i = \hat{A}_i \oplus a_i$, divide out the $(x - \sum_{i=1}^m a_i)$ term as in (9), and apply Theorem IV.2 inductively $(m-1)$ times, replacing each $\hat{P}_i$ with a random orthogonal Q_i (this requires conditioning on the other $\hat{P}_j$ and Q_j, but by independence the distribution of each $\hat{P}_i$ is still uniform on $\mathcal{A}_d$). Finally, appeal to the identity (4) to write this as an m-wise additive convolution. ■

B. Quadrature for Bipartite Matrices

Theorem IV.10. *Suppose A and B are (not necessarily symmetric) $d \times d$ matrices such that $A\mathbf{1} = A^T\mathbf{1} = a\mathbf{1}$ and $B\mathbf{1} = B^T\mathbf{1} = b\mathbf{1}$. Let $\chi_x\left(AA^T\right) = (x - a^2)p(x)$ and $\chi_x\left(BB^T\right) = (x - b^2)q(x)$. Then,*

$$\mathop{\mathbb{E}}_{P,S} \chi_x\left(\begin{bmatrix}0 & A\\ A^T & 0\end{bmatrix} + (P \oplus S)\begin{bmatrix}0 & B\\ B^T & 0\end{bmatrix}(P \oplus S)^T\right) = \mathbb{S}\left((x - (a+b)^2)p(x) \boxplus_{d-1} q(x)\right) \tag{17}$$

$$= (x^2 - (a+b)^2)\mathbb{S}\left(p(x) \boxplus_{d-1} q(x)\right), \tag{18}$$

where P and S are independent uniform random permutation matrices.

As in the nonbipartite case, we begin by applying a change of basis V that isolates the common all ones eigenvector and block diagonalizes our matrices as:

$$VAV^T = \hat{A} \oplus a, \quad VBV^T = \hat{B} \oplus b, \quad VPV^T = \hat{P} \oplus 1, \quad VSV^T = \hat{S} \oplus 1. \tag{19}$$

Conjugating the left hand side of (17) by $(V \oplus V)$, we see that it is the same as

$$\begin{aligned}
&\mathop{\mathbb{E}}_{P,S} \chi_x \left(\begin{bmatrix} 0 & (\hat{A} \oplus a) \\ (\hat{A} \oplus a)^T & 0 \end{bmatrix} + ((\hat{P} \oplus 1) \oplus (\hat{S} \oplus 1)) \begin{bmatrix} 0 & (\hat{B} \oplus b) \\ (\hat{B} \oplus b)^T & 0 \end{bmatrix} ((\hat{P} \oplus 1) \oplus (\hat{S} \oplus 1))^T \right) \\
&= \mathop{\mathbb{E}}_{P,S} \chi_x \left(\begin{bmatrix} 0 & (\hat{A} + \hat{P}\hat{B}\hat{S}^T \oplus (a+b)) \\ (\hat{A} + \hat{P}\hat{B}\hat{S}^T \oplus (a+b))^T & 0 \end{bmatrix} \right) \\
&= \mathop{\mathbb{E}}_{P,S} \mathbb{S}\chi_x \left((\hat{A} + \hat{P}\hat{B}\hat{S}^T \oplus (a+b))(\hat{A} + \hat{P}\hat{B}\hat{S}^T \oplus (a+b))^T \right) \\
&= (x^2 - (a+b)^2) \mathop{\mathbb{E}}_{P,S} \mathbb{S}\chi_x \left((\hat{A} + \hat{P}\hat{B}\hat{S}^T)(\hat{A} + \hat{P}\hat{B}\hat{S}^T)^T \right) \\
&= (x^2 - (a+b)^2) \mathop{\mathbb{E}}_{P,S} \chi_x \left(\begin{bmatrix} 0 & \hat{A} \\ \hat{A}^T & 0 \end{bmatrix} + (\hat{P} \oplus \hat{S}) \begin{bmatrix} 0 & \hat{B} \\ \hat{B}^T & 0 \end{bmatrix} (\hat{P} \oplus \hat{S})^T \right). \qquad (20)
\end{aligned}$$

As in the previous section, the matrices $\hat{P}$ and $\hat{S}$ are random elements of the group $\mathcal{A}_{d-1}$. Observe that

$$p(x) = \chi_x \left(\hat{A}\hat{A}^T \right) \quad \text{and} \quad q(x) = \chi_x \left(\hat{B}\hat{B}^T \right).$$

Recalling from (3) that

$$\mathbb{S}\left(p(x) \boxplus_{d-1} q(x)\right) = \mathbb{E}_{Q,R \in \mathcal{O}(d-1)} \chi_x \left(\begin{bmatrix} 0 & A \\ A^T & 0 \end{bmatrix} + (Q \oplus R) \begin{bmatrix} 0 & B \\ B^T & 0 \end{bmatrix} (Q \oplus R)^T \right)$$

and removing all the $\hat{}$s as before to ease notation, we see that the conclusion (17) of Theorem IV.10 is implied by the following more general quadrature statement.

Theorem IV.11. *For all symmetric $2d \times 2d$ matrices C and D:*

$$\mathop{\mathbb{E}}_{P,S \in \mathcal{A}_d} \chi_x \left(C + (P \oplus S) D (P \oplus S)^T \right) = \mathop{\mathbb{E}}_{Q,R \in \mathcal{O}(d)} \chi_x \left(C + (Q \oplus R) D (Q \oplus R)^T \right). \qquad (21)$$

This theorem is an immediate consequence of two applications of the following corollary of Theorem IV.2.

Corollary IV.12. *If C and D are symmetric $2d \times 2d$ matrices,*

$$\mathop{\mathbb{E}}_{P \in \mathcal{A}_d} \det(C + (P \oplus I) D (P \oplus I)^T) = \mathop{\mathbb{E}}_{Q \in \mathcal{O}(d)} \det(C + (Q \oplus I) D (Q \oplus I)^T).$$

Proof: The proof is identical to the proof of Theorem IV.2, except we replace $P \in \mathcal{A}_d$ with $P \oplus I$ and $Q \in \mathcal{O}(d)$ with $Q \oplus I$ at each step.

Specifically, let

$$f_{C,D}(Q) := \det(C + (Q \oplus I) D (Q \oplus I)^T).$$

Applying Corollary IV.6 as before reveals that for every $i, j, k \leq d$ and every $Q_2 \in \mathcal{O}_{i,j,k}$,

$$\mathop{\mathbb{E}}_{P \in \mathcal{A}_d} f_{C,D}(P) = \mathop{\mathbb{E}}_{P \in \mathcal{A}_d} \mathop{\mathbb{E}}_{P_2 \in \mathcal{A}_{i,j,k}} f_{C,D}(PP_2) = \mathop{\mathbb{E}}_{P \in \mathcal{A}_d} \mathop{\mathbb{E}}_{P_2 \in \mathcal{A}_{i,j,k}} f_{C,D}(PP_2Q_2) = \mathop{\mathbb{E}}_{P \in \mathcal{A}_d} f_{C,D}(PQ_2).$$

Since an arbitrary $Q_0 \in \mathcal{O}(d)$ is a finite product of such Q_2 by Lemma IV.7, this means that

$$\mathop{\mathbb{E}}_{P \in \mathcal{A}_d} f_{C,D}(PQ_0) = \mathop{\mathbb{E}}_{P \in \mathcal{A}_d} f_{C,D}(P)$$

for all $Q_0 \in \mathcal{O}(d)$. Lemma IV.3 completes the proof. ∎

Proof of Theorem IV.11: Since P and S are independent, we have

$$\begin{aligned}
&\mathop{\mathbb{E}}_{P,S\in\mathcal{A}_d} \chi_x\left(C+(P\oplus S)D(P\oplus S)^T\right)\\
&=\mathop{\mathbb{E}}_{S\in\mathcal{A}_d}\mathop{\mathbb{E}}_{P\in\mathcal{A}_d}\det(xI+C+(P\oplus I)(I\oplus S)D(I\oplus S)^T(P\oplus I)^T)\\
&=\mathop{\mathbb{E}}_{S\in\mathcal{A}_d}\mathop{\mathbb{E}}_{Q\in\mathcal{O}(d)}\det(xI+C+(Q\oplus I)(I\oplus S)D(I\oplus S)^T(Q\oplus I)^T) \quad \text{by Corollary IV.12}\\
&=\mathop{\mathbb{E}}_{Q\in\mathcal{O}(d)}\mathop{\mathbb{E}}_{S\in\mathcal{A}_d}\det(xI+(Q\oplus I)^TC(Q\oplus I)+(I\oplus S)D(I\oplus S)^T)\\
&=\mathop{\mathbb{E}}_{Q\in\mathcal{O}(d)}\mathop{\mathbb{E}}_{R\in\mathcal{O}(d)}\det(xI+(Q\oplus I)^TC(Q\oplus I)+(I\oplus R)D(I\oplus R)^T) \quad \text{by Corollary IV.12}\\
&=\mathop{\mathbb{E}}_{Q,R\in\mathcal{O}(d)}\det(xI+C+(Q\oplus R)D(Q\oplus R)^T),
\end{aligned}$$

as desired. ∎

Proof of Theorem IV.10: Follows from Theorem IV.11, (19), and (20). ∎

Like Theorem IV.1, Theorem IV.10 extends effortlessly to the case of many matrices.

Corollary IV.13. *If $A_1,\dots,A_m$ are matrices with $A_i\mathbf{1}=A_i^T\mathbf{1}=a_i$ and $\chi_x\left(A_iA_i^T\right)=(x-a_i^2)p_i(x)$, then*

$$\begin{aligned}
&\mathop{\mathbb{E}}_{P_1,\dots,P_m,S_1,\dots,S_m}\chi_x\left(\sum_{i=1}^m(P_i\oplus S_i)\begin{bmatrix}0 & A_i\\ A_i^T & 0\end{bmatrix}(P_i\oplus S_i)^T\right)\\
&=\left(x^2-\left(\sum_i a_i\right)^2\right)\mathbb{S}\left[p_1(x)\boxplus\cdots\boxplus p_m(x)\right],
\end{aligned}$$

where the P_i and S_i are independent uniformly random permutations.

We omit the proof, which is identical to the proof of Corollary IV.9.

V. Ramanujan Graphs

In this section, we combine the Cauchy transform, interlacing, and quadrature results of the previous sections to establish our main Theorems I.1 and I.2

Proof of Theorem I.2: Let M be the adjacency matrix of a fixed perfect matching on d vertices, with d even. Since the uniform distribution on permutations is realizable by swaps (Lemma III.5), Theorem III.4 tells us that with nonzero probability:

$$\lambda_2\left(\sum_{i=1}^d P_iMP_i^T\right)\le\lambda_2\left(\mathbb{E}\chi_x\left(\sum_{i=1}^m P_iMP_i^T\right)\right).$$

Corollary IV.9 reveals that the polynomial in the right-hand expression may be written as an m-wise symmetric additive convolution

$$E(x):=\mathop{\mathbb{E}}_{P_1,\dots,P_m}\chi_x\left(\sum_{i=1}^m P_iA_iP_i^T\right)=(x-m)[\underbrace{p\boxplus_{d-1}\cdots\boxplus_{d-1}p}_{m\text{ times}}](x),$$

where

$$p(x)=\frac{\chi_M(x)}{x-1}=(x-1)^{d/2-1}(x+1)^{d/2},$$

is the characteristic polynomial of a single matching with the trivial root at 1 removed. Our goal is therefore to bound the largest root of $p(x)\boxplus\cdots\boxplus p(x)$, which is the second largest root of $E(x)$.

We will do this using the inverse Cauchy transform described in Section II-C. The Cauchy transform of $p(x)$ is given by

$$\mathcal{G}_p(x) = \frac{d/2-1}{d-1}\frac{1}{x-1} + \frac{d/2}{d-1}\frac{1}{x+1}.$$

Notice that for every $x > 1$, putting the trivial root at 1 back only increases the Cauchy transform:

$$\mathcal{G}_p(x) < \frac{d/2}{d}\frac{1}{x-1} + \frac{d/2}{d}\frac{1}{x+1} = \frac{x}{x^2-1} = \mathcal{G}_{\chi(M)}(x). \tag{22}$$

Since both functions are decreasing for $x > 1$, this implies that the inverse Cauchy transform of p is upper bounded by that of $\chi(M)$:

$$\mathcal{K}_p(w) < \mathcal{K}_{\chi(M)}(w),$$

for every $w > 0$.

Applying the convolution inequality in Theorem II.7 $(m-1)$ times yields the following upper bound on the inverse Cauchy transform of the m-wise convolution of interest.

$$\mathcal{K}_{p \boxplus \cdots \boxplus p}(w) \le m \cdot \mathcal{K}_p(w) - \frac{m-1}{w} < m \cdot \mathcal{K}_{\chi(M)}(w) - \frac{m-1}{w}. \tag{23}$$

Recalling from (22) that

$$\mathcal{K}_{\chi(M)}(w) = x \iff w = \frac{x}{x^2-1},$$

the right hand side of (23) may be written as

$$mx - \frac{m-1}{w} = mx - \frac{(m-1)(x^2-1)}{x} = \frac{x^2+(m-1)}{x},$$

which is easily seen to be minimized at $x = \sqrt{m-1}$ with value $2\sqrt{m-1}$. Thus, the second largest root of $E(x)$ is at most $2\sqrt{m-1}$. ∎

Proof of Theorem I.1: Let

$$M = \begin{bmatrix} 0 & I \\ I^T & 0 \end{bmatrix}$$

be the adjacency matrix of a perfect matching on $2d$ vertices, across the natural bipartition. Then, for independent uniformly random $d \times d$ permutation matrices $P_1, \ldots, P_m, S_1, \ldots, S_m$, the random matrix

$$A = \sum_{i=1}^{m} (P_i \oplus S_i) M (P_i \oplus S_i)^T = \sum_{i=1}^{m} \begin{bmatrix} 0 & (P_i S_i^T) \\ (P_i S_i^T)^T & 0 \end{bmatrix}$$

is the adjacency matrix of a union of m random matchings across the same bipartition. Since the distribution of the $(P_i \oplus S_i)$ is realizable by swaps (Lemma III.5), Theorem III.4 implies that

$$\lambda_2(A) \le \lambda_2 \left(\mathbb{E}\, \chi_x \left(\sum_{i=1}^{m} (P_i \oplus S_i) M (P_i \oplus S_i) \right) \right),$$

with nonzero probability. Since $I\mathbf{1} = \mathbf{1}$, Corollary IV.13 implies that the polynomial on the right hand side is equal to

$$(x^2 - m^2)\mathbb{S}[\underbrace{p \boxplus_{d-1} \cdots \boxplus_{d-1} p}_{m \text{ times}}](x),$$

where

$$p(x) = \chi_x\left(I_{d-1} I_{d-1}^T\right) = (x-1)^{d-1}.$$

We upper bound the inverse Cauchy transform of this m-wise convolution using Theorem II.8:

$$\mathcal{K}_{\mathbb{S}(p \boxplus \cdots \boxplus p)}(w) \leq m \cdot \mathcal{K}_{\mathbb{S}p}(w) - \frac{m-1}{w} = m \cdot \mathcal{K}_{(x^2-1)^{d-1}}(w) - \frac{m-1}{w}.$$

Since

$$\mathcal{G}_{(x^2-1)^{d-1}}(w) = \frac{x}{x^2-1},$$

this is now identical to the calculation (23), so we obtain again the bound $2\sqrt{m-1}$. Thus, we conclude that $\lambda_2(A) \leq 2\sqrt{m-1}$ with nonzero probability. Since A is bipartite, its spectrum is symmetric about zero, so we must also have $\lambda_{d-1}(A) \geq -2\sqrt{m-1}$, whence A is Ramanujan. ■

VI. Acknowledgment

This research was partially supported by NSF grant CCF-1111257, an NSF Mathematical Sciences Postdoctoral Research Fellowship, Grant No. DMS-0902962, a Simons Investigator Award to Daniel Spielman, and a MacArthur Fellowship.

References

[1] G. W. Anderson, A. Guionnet, and O. Zeitouni. *An introduction to random matrices*. Number 118. Cambridge University Press, 2010.

[2] J. Batson, D. A. Spielman, and N. Srivastava. Twice-Ramanujan sparsifiers. *SIAM Journal on Computing*, 41(6):1704–1721, 2012.

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

[4] M. Cohen. private communication, 2015.

[5] S. Fisk. *Polynomials, roots, and interlacing*. arXiv:math/0612833 [math.CA], 2008.

[6] J. Friedman. *A Proof of Alon's Second Eigenvalue Conjecture and Related Problems*. Number 910 in Memoirs of the American Mathematical Society. American Mathematical Society, 2008.

[7] C. Hall, D. Puder, and W. F. Sawin. Ramanujan coverings of graphs. *arXiv preprint arXiv:1506.02335*, 2015.

[8] S. Hoory, N. Linial, and A. Wigderson. Expander graphs and their applications. *Bulletin of the American Mathematical Society*, 43(4):439–561, 2006.

[9] A. Lubotzky, R. Phillips, and P. Sarnak. Ramanujan graphs. *Combinatorica*, 8(3):261–277, 1988.

[10] A. W. Marcus, D. A. Spielman, and N. Srivastava. Finite free convolutions of polynomials. *arXiv preprint arXiv:1504.00350*, Apr. 2015.

[11] A. W. Marcus, D. A. Spielman, and N. Srivastava. Interlacing families I: Bipartite Ramanujan graphs of all degrees. *Ann. of Math.*, 182-1:307–325, 2015.

[12] A. W. Marcus, D. A. Spielman, and N. Srivastava. Interlacing families II: Mixed characteristic polynomials and the Kadison-Singer problem. *Ann. of Math.*, 182-1:327–350, 2015.

[13] G. A. Margulis. Explicit group theoretical constructions of combinatorial schemes and their application to the design of expanders and concentrators. *Problems of Information Transmission*, 24(1):39–46, July 1988.

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

[15] A. Nica and R. Speicher. *Lectures on the combinatorics of free probability*, volume 13. Cambridge University Press, 2006.

[16] A. Nilli. On the second eigenvalue of a graph. *Discrete Math*, 91:207–210, 1991.

[17] R. Pemantle. Hyperbolicity and stable polynomials in combinatorics and probability. *arXiv preprint arXiv:1210.3231*, 2012.

[18] D. V. Voiculescu. *Free probability theory*. American Mathematical Society, 1997.