SIAM J. COMPUT.
Vol. 47, No. 6, pp. 2488–2509

INTERLACING FAMILIES IV: BIPARTITE RAMANUJAN GRAPHS OF ALL SIZES*

ADAM W. MARCUS†, DANIEL A. SPIELMAN‡, AND NIKHIL SRIVASTAVA§

Abstract. We prove that there exist bipartite Ramanujan graphs of every degree and every number of vertices. The proof is based on an analysis of 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.

Key words. expander graphs, free probability, interlacing, random graphs, random matrices

AMS subject classifications. 68R10, 60B20

DOI. 10.1137/16M106176X

1. Introduction.

Ramanujan graphs are undirected regular graphs whose nontrivial adjacency matrix eigenvalues are as small as possible; that is, they are spectrally optimal 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 [15] 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 [15]. 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 technical contributions of this paper are 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 [14]. This latter technique is inspired by ideas in free probability theory [19, 24], 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.

*Received by the editors February 16, 2016; accepted for publication (in revised form) April 30, 2018; published electronically December 18, 2018. An earlier version of this paper appeared in the Proceedings of the 2015 IEEE Symposium on Foundations of Computer Science.
http://www.siam.org/journals/sicomp/47-6/M106176.html
Funding: This research was partially supported by NSF grant CCF-1111257, NSF Mathematical Sciences Postdoctoral Research Fellowship grant DMS-0902962, a Simons Investigator Award to Daniel Spielman, and a MacArthur Fellowship.
†Department of Mathematics, Princeton University, Princeton, NJ 08544 (adam.marcus@princeton.edu).
‡Department of Computer Science, Yale University, New Haven, CT 06520 (spielman@cs.yale.edu).
§Department of Mathematics, UC Berkeley, Berkeley, CA 94720 (nikhil@math.berkeley.edu).

1.1. Summary of results. Recall that the adjacency matrix A of an *m-regular graph on d vertices*[1] has largest eigenvalue $\lambda_1(A) = m$ and smallest eigenvalue $\lambda_d(A) = -m$ when the graph is bipartite. Following Friedman [9], we will refer to these as the *trivial* eigenvalues of A, and we will call a graph *Ramanujan* if all of its nontrivial 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 in [20] tells us that for every $\epsilon > 0$, every infinite sequence of m-regular graphs must contain a graph with a nontrivial 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.

THEOREM 1.1. *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 nonbipartite version of this theorem, regarding a union of m random perfect matchings on d vertices (not bipartite), with d even.

THEOREM 1.2. *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 essentially the same ideas. Note that Theorem 1.2 does not produce Ramanujan graphs in the nonbipartite case because it provides no control over the least eigenvalue λ_d.

As the graphs we produce are unions of independent matchings, they may have multiple edges between two vertices. Thus, they are (strictly speaking) multigraphs. One would expect that it should be more difficult to construct Ramanujan graphs with multiedges than without, but we do not presently know how to prove a theorem to this effect.

As in our previous work [15], 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 polynomials used in [15], which are $\#P$ hard to compute, it has been shown by Michael Cohen that the expected characteristic polynomials in this paper can be computed in polynomial time [5], yielding a polynomial time construction of Ramanujan graphs of all sizes and degrees.

1.2. Related work and context. Infinite families of Ramanujan graphs were first shown to exist for $m = p + 1$, p a prime, in the seminal works of Margulis

[1] In order to be consistent with our companion paper [14], we will, unconventionally, use m to denote the degree of a graph and d to denote its number of vertices.

and Lubotzky, Phillips and Sarnak [17, 12]. The graphs they produce are Cayley graphs and can be constructed very efficiently. Friedman [9] showed that a random m-regular graph is almost Ramanujan: specifically, that with high probability all of the nontrivial eigenvalues of the union of m random perfect matchings have absolute value at most $2\sqrt{m-1}+\epsilon$ for every $\epsilon > 0$.

More recently, in [15], 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 [11] and [15] for a more detailed discussion of expander graphs, Ramanujan graphs, and 2-lifts. Building on the present paper, Hall, Puder, and Sawin [10] have used related techniques to show that every m-regular graph has a k-lift which is Ramanujan. Their result subsumes both the results of the present paper and [15].

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 [18] 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 supported on the interval $[-2\sqrt{m-1}, 2\sqrt{m-1}]$ known as the *Kesten–McKay law*. 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 nonasymptotic 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. Our result shows that the largest root of the expected characteristic polynomial of a random m-regular graph of size d lies inside the support of the Kesten–McKay law $\mu_m(x)$. On one hand, the support of this limiting measure can be calculated using techniques from free probability, and on the other, our bound on the largest root of the finite polynomials is calculated using analogous tools that we call "finite free probability" (see [13]), which mimick free probability but operate on finite-dimensional polynomials. Since these bounds coincide, and it is impossible to construct a sequence of graphs with eigenvalues strictly smaller than the support o $\mu_m(x)$, the graphs produced by our method are necessarily optimal.

We remark that all comments regarding the relationship between polynomial convolutions and free probability in the present paper are intended to draw attention to conceptual parallels only, and that the results themselves do not require any formal knowledge of free probability. For those interested in such a relationship, we refer the reader to [13], where a more formal connection is established.

1.3. 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 nonbipartite 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 form "interlacing families" (reviewed in section 2.1). By an argument introduced in [15], this reduces the problem of proving the existence of Ramanujan graphs to an analysis of the roots of these polynomials. The required real-rootedness and interlacing properties are established in section 3, by decomposing the random permutations used to generate these expected polynomials into random swaps acting on two coordinates at a time and showing that such random swaps preserve real-rootedness. Theorem 3.3, which may be of independent interest, says that if A and B are any symmetric matrices, then the expected characteristic polynomial of $A+PBP^T$ is real-rooted for a uniformly random permutation matrix P. We remark that this argument is completely elementary and self-contained and, unlike [15, 16], does not appeal to any results from the theory of real stable or hyperbolic polynomials.

Next, in section 4, we derive a useful closed-form formula for the expected characteristic polynomial of a sum of randomly permuted regular graphs, which includes our random m-regular graphs as a special case. We begin by proving that, in the case of adjacency matrices, the expected characteristic polynomials taken over random permutations match the expected characteristic polynomials taken 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.

To obtain the formula, we appeal to machinery developed in our companion paper [14], which studies the structure of expected characteristic polynomials over random orthogonal matrices. In particular, it is shown there that 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 expected characteristic polynomial of a union of m random matchings decouples as an m-fold convolution of the characteristic polynomial of a single matching, yielding the formula.

Finally, we apply new bounds derived in [14] on the roots of such convolutions to obtain the desired Ramanujan bound of $2\sqrt{m-1}$. The requisite material regarding free convolutions is introduced in sections 2.2 and 2.3. These ingredients are combined in section 5 to complete the proofs of Theorems 1.1 and 1.2.

2. Preliminaries.

2.1. Interlacing families. Showing that a random matrix has all 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 all 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 [16], stated here in the slightly different language of product distributions.

THEOREM 2.1 (interlacing families). *Suppose $\{f_\omega(x)\}_{\omega\in\{0,1\}^m}$ is a family of real-rooted polynomials of the same degree d 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 $k = 1, \ldots, d$ 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.

The above theorem is relevant to this paper because our random graphs are generated from independent random permutations, which are in turn generated by independent random swaps, yielding product distributions in a natural way.

In order to apply Theorem 2.1 in our setting, we will need to appeal to interlacing properties of the polynomials on hand. Recall that real-rooted polynomials $f = \prod_{i=1}^d (x-\lambda_i)$ and $g = \prod_{i=1}^{d-1}(x-\mu_i)$ *interlace* if $\lambda_1 \leq \mu_1 \leq \lambda_2 \leq \cdots \leq \mu_{d-1} \leq \lambda_d$. There are two ways that two real-rooted polynomials f and g of the same degree can interlace; 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 well-known relationship between interlacing and real-rootedness of convex combinations, which appears as Theorem 2.1 of Dedieu [6], as Theorem 2′ of Fell [7], and as a special case of Theorem 3.6 of Chudnovsky and Seymour [4].

LEMMA 2.2. *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 interlacer (a third polynomial interlacing both of them).*

We will also use the following elementary facts about interlacing and real-rootedness, which may be found in [8].

LEMMA 2.3. *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 coefficients, then

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

We refer the interested reader to [15] for a more thorough introduction to interlacing families. Part 1 of the above lemma is known as Obreschkoff's theorem [21], and a proof written in English may be found in [22, Theorem 6.3.8].

2.2. 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 [14]. 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 special sums of independent random matrices, prototypically of type $A + QBQ^T$ where Q is a random orthogonal matrix,[2] 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.

We remark that the finite free convolution was inspired by Voiculescu's free convolution [24] in free probability theory (hence the name). The connection lies in the fact that free probability provides precise descriptions of the the *limiting* spectral distribution of random matrix ensembles of type $\{A_d + Q_d B_d Q_d^T\}_{d=1}^{\infty}$, where the Q_d are random orthogonal matrices as the dimension d tends to infinity. In particular, if the limiting spectral distributions of the $\{A_d\}$ and $\{B_d\}$ are μ_A and μ_B, then the limiting spectral distribution of this model is given by $\mu_A \boxplus \mu_B$, where $\boxplus$ is an operation on measures called the free convolution.

Our corresponding operation on polynomials will mimic this setup in finite dimensions. We denote the characteristic polynomial of a matrix A by

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

Definition 2.4 (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 orthogonal matrices.

Note that this is a well-defined operation on polynomials because, as is shown in [14], 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.

In the case that the matrices of interest are not symmetric (as will happen in bipartite adjacency matrices), we will require the following two-sided variant of the above, which yields singular values rather than eigenvalues.[3]

Definition 2.5 (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

[2]More generally, those for which the eigenvectors of the two matrices in the sum are independent and "as random as possible."

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

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, closed-form 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 [14]. The results in this paper only require the following important consequences of these formulas, also established in [14]. We will occasionally drop the subscripts in $\boxplus_d$ and $\boxplus\!\!\boxplus_d$ when they are clear from context.

Lemma 2.6 (properties of $\boxplus$ and $\boxplus\!\!\boxplus$). *Let p and q be degree-d polynomials.*

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 [14], and bilinearity follows immediately from Theorems 1.1 and 1.3 of [14]. 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 unambiguously 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 $A_1, \ldots, A_m$.

2.3. 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, 15, 16]. The methods below are different from those used to study the mixed characteristic polynomials of [16], 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 about $\sqrt{2}$ in this setting).

DEFINITION 2.7 (Cauchy transform). *The* Cauchy transform *of a polynomial* $p(x)$ *with roots* $\lambda_1, \ldots, \lambda_d$ *is defined to be the rational 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*

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

where $w > 0$ *is a real parameter.*

Note that the Cauchy transform has poles at the roots λ_i of p, and when all of the roots 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 [14].

THEOREM 2.8 (see [14, Theorem 1.7]). *For real-rooted degree* d *polynomials* p *and* q *and* $w > 0$,

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

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 1 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}. This 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 2.8 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 [14, 13] for more details. To analyze the case of bipartite random graphs, we will need the corresponding inequality for the asymmetric convolution.

THEOREM 2.9 (see [14, Theorem 1.8]). *For degree* d *polynomials* p *and* q *having only nonnegative real roots,*

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

3. 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 3.1 (random swap). *A* random swap *is a matrix-valued random variable which is equal to a transposition of two (fixed) indices a, b with probability α and equal to the identity with probability $(1-\alpha)$ for some $\alpha \in [0,1]$.*

Definition 3.2 (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 3.5 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 [15, 16].

Theorem 3.3. *Let $A_1, \ldots, A_m$ be symmetric $d \times d$ matrices, and let $\{S_{ij}\}_{i \le m, j \le N}$ 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 3.3 and 2.1, applied to the family of polynomials indexed by all possible values of the swaps S_{ij}, is the following existence result.

Theorem 3.4 (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 $k \le d$,*

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

with nonzero probability.

Theorem 3.4 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 3.5. *Let P and R be uniformly random $d \times d$ permutation matrices. Both P and $P \oplus R$ are realizable by swaps, where $P \oplus R = \left(\begin{smallmatrix} P & 0 \\ 0 & R \end{smallmatrix}\right)$ is the direct sum of P and R.*

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

The rest of this section is devoted to proving Theorem 3.3. The proof we present here is a simplification of our original proof inspired by the proof of [10, Lemma 4.3]. Similar techniques may be used to prove the real-rootedness of the mixed characteristic polynomials that appeared in [15, 16] (see [23]).

Definition 3.6 (real-rooted distributions). *We say that a tuple of independent (not necessarily identically distributed) d-dimensional random permutation matrices $(P_1,\ldots,P_m)$ is real-rooted if for all $d \times d$ symmetric matrices $A_1,\ldots,A_m$, the polynomial*

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

is real rooted.

For example, if each P_i is always the identity, then $(P_1,\ldots,P_m)$ is real-rooted. Given this, Theorem 3.3 follows immediately from the following lemma.

Lemma 3.7. *If S is a random swap and if $(P_1,\ldots,P_m)$ is real-rooted, then $(P_1S,\ldots,P_m)$ is also real-rooted.*

The proof of Lemma 3.7 will rely on two auxiliary lemmas.

Lemma 3.8. *If σ is a transposition and A is symmetric, then $A - \sigma A\sigma^T$ has rank 2 and trace 0. Thus, we can write $\sigma A \sigma^T = A - uu^T + vv^T$ for some vectors u and v.*

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

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

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

Lemma 3.9. *Let A be a d-dimensional symmetric matrix, and let v be a vector. Let*

$$p_t(x) = \chi_x(A + tvv^T).$$

Then there is a degree $d-1$ polynomial with positive leading coefficient $q(x)$ so that

$$p_t(x) = \chi_x(A) - tq(x).$$

Proof. Consider the case in which v is the elementary unit vector in the first coordinate. It suffices to consider this case as determinants, and thus characteristic polynomials, are unchanged by multiplication by rotation matrices. The matrix tvv^T is zeros everywhere except for the element t in the upper left entry. So,

$$\begin{aligned}\chi_x(A + tvv^T) &= \det(xI - A - tvv^T) = \det(xI - A) - t\det(xI^{(1)} - A^{(1)}) \\ &= \chi_x(A) - t\chi_x(A^{(1)}),\end{aligned}$$

where $A^{(1)}$ is the submatrix of A obtained by removing its first row and column. ∎

Proof of Lemma 3.7. Let S be equal to the transposition σ with probability α and the identity with probabilty $1-\alpha$. Let $A_1, \ldots, A_m$ be arbitrary d-dimensional symmetric matrices. Lemma 3.8 tells us that there exist vectors u and v so that

$$\sigma A_1 \sigma^T = A_1 - uu^T + vv^T.$$

We will now show that

$$\mathbb{E}\chi_x\left(\sum_{i=1}^m P_i A_i P_i^T\right) \to \mathbb{E}\chi_x\left(P_1(A_1 + vv^T)P_1^T + \sum_{i=2}^m P_i A_i P_i^T\right). \tag{6}$$

Lemma 3.9 tells us that for every fixed tuple of permutation matrices $(Q_1, \ldots, Q_m)$, there exists a degree $d-1$ polynomial $q(x)$ with positive leading coefficient so that

$$\chi_x\left(Q_1(A_i + tvv^T)Q_1^T + \sum_{i=2}^m Q_i A_i Q_i^T\right) = \chi_x\left(\sum_{i=1}^m Q_i A_i Q_i^T\right) - tq(x).$$

Thus, averaging over all $(Q_1, \ldots, Q_m)$ in the support of $(P_1, \ldots, P_m)$, there exists a degree $d-1$ polynomial $p(x)$ with positive leading coefficient so that

$$p_t(x) \stackrel{\text{def}}{=} \mathbb{E}\chi_x\left(P_1(A_1 + tvv^T)P_1^T + \sum_{i=2}^m P_i A_i P_i^T\right) = \mathbb{E}\chi_x\left(\sum_{i=1}^m P_i A_i P_i^T\right) - tp(x).$$

As $(P_1, \ldots, P_m)$ is real-rooted and $A + tvv^T$ is symmetric, the polynomial $p_t(x)$ is real-rooted for every t. By parts 1 and 2 of Lemma 2.3, this implies $p_0(x) \to p_1(x)$, a statement that is identical to (6). The same argument implies that

$$\begin{aligned}&\mathbb{E}\chi_x\left(P_1(A_1 + vv^T - uu^T)P_1^T + \sum_{i=2}^m P_i A_i P_i^T\right) \\ &\quad\to \mathbb{E}\chi_x\left(P_1(A_1 + vv^T)P_1^T + \sum_{i=2}^m P_i A_i P_i^T\right).\end{aligned}$$

Thus, both the polynomials $\mathbb{E}\chi_x\left(\sum_{i=1}^m P_i A_i P_i^T\right)$ and

$$\mathbb{E}\chi_x\left(P_1(A_1 + vv^T - uu^T)P_1^T + \sum_{i=2}^m P_i A_i P_i^T\right) = \mathbb{E}\chi_x\left(P_1 \sigma A_1 \sigma^T P_1^T + \sum_{i=2}^m P_i A_i P_i^T\right)$$

interlace

$$\mathbb{E}\chi_x\left(P_1(A_1 + vv^T)P_1^T + \sum_{i=2}^m P_i A_i P_i^T\right).$$

So, we may use Lemma 2.2 to conclude that

$$\begin{aligned}\mathbb{E}\chi_x&\left(P_1SA_1S^TP_1^T+\sum_{i=2}^{m}P_iA_iP_i^T\right)\\&=\alpha\mathbb{E}\chi_x\left(P_1\sigma A_1\sigma^TP_1^T+\sum_{i=2}^{m}P_iA_iP_i^T\right)+(1-\alpha)\mathbb{E}\chi_x\left(\sum_{i=1}^{m}P_iA_iP_i^T\right)\end{aligned}$$

is real-rooted. ◻

4. Quadrature. In this section, we show how finite free convolutions (which are expectations over orthogonal matrices) can be applied to polynomials from the previous section (which are expectations over permutation matrices). In short, we will see that the expected characteristic polynomials we are interested in will become finite free convolutions when the original matrices are projected orthogonally to the all-ones vector. This will give us explicit formulas for these polynomials and, more importantly, a way to bound for 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 4.2 we derive the result for the bipartite case as a corollary of the result for the symmetric case.

4.1. 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.

We will use $\mathbf{1}$ to denote the all-ones vector.

THEOREM 4.1. *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 that simultaneously block diagonalizes all of them (for concreteness, we use the one mapping $\mathbf{1}$ to the standard basis vector e_n):

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

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 $\hat{e_1}, \ldots, \hat{e_d}$, 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 will interpret each $\hat{P}$ as an element 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 4.2 (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 4.3 (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 4.2 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 4.3 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 [14].

LEMMA 4.4 (invariance for $\mathcal{A}_2$). *If A and B are symmetric $d \times d$ matrices, then for every $Q_0 \in \mathcal{O}(2)$,*

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

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

$$\begin{aligned} &\sum_{\tau \in T} \det(A + (R_\tau R_\theta \oplus I)B(R_\tau R_\theta \oplus I)^T) \\ &= \sum_{\tau \in T} \det(A + (R_{\tau+\theta} \oplus I)B(R_{\tau+\theta} \oplus I)^T) \\ &= \sum_{\tau \in T} \sum_{k=-2}^{2} c_k e^{ik(\tau+\theta)} \quad \text{for some coefficients } c_k\text{, by Lemma 4.5} \\ &= \sum_{k=-2}^{2} c_k e^{ik\theta} \left(e^{ik0} + e^{ik2\pi/3} + e^{ik4\pi/3}\right) \\ &= 3c_0 \quad \text{since the terms with } |k| = 1, 2 \text{ vanish.} \end{aligned}$$

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

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

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 4.5 (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. 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 4.6 (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 4.4 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 4.7 ($\mathcal{O}_{i,j,k}$ generates $\mathcal{O}(d)$). *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 $\bigcup_{i,j,k\leq h} \mathcal{O}_{i,j,k}$. Let $\hat{e}_0, \ldots, \hat{e}_d$ be the vertices of the regular simplex. For $1 \leq h \leq d$, 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 $1 \leq i \leq h$.

We first consider the case in which $w_h^T \hat{e}_h \geq 0$. Let F_h denote the two-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 4.8 (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 4.6 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 4.7, 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

$$Q_0 = Q_1 Q_2 \cdots Q_m.$$

By applying the previous equality m times, we obtain

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

Proof of Theorem 4.2. The proof follows from Theorem 4.8 and Lemma 4.3. $\square$

Proof of Theorem 4.1. The proof follows from Theorem 4.2, (8), and (9). $\square$

We conclude the section by recording the obvious extension of Theorem 4.1 to sums of m matrices.

COROLLARY 4.9. *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,*

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

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 4.2 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. $\square$

4.2. Quadrature for bipartite matrices.

THEOREM 4.10. *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,*

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

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

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

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

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 $\hat{\cdot}$s as before to ease notation, we see that the conclusion (17) of Theorem 4.10 is implied by the following more general quadrature statement.

Theorem 4.11. *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 4.2.

Corollary 4.12. *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 4.2, 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 4.6 as before reveals that for every $i, j, k \leq d$ and every $Q_2 \in \mathcal{O}_{i,j,k}$,

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

Since an arbitrary $Q_0 \in \mathcal{O}(d)$ is a finite product of such Q_2 by Lemma 4.7, 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 4.3 completes the proof. ∎

Proof of Theorem 4.11. Since P and S are independent, we have

$$
\begin{aligned}
&\mathop{\mathbb{E}}_{P,S\in\mathcal{A}_d} \chi_x \left(C + (P\oplus S)D(P\oplus S)^T\right)\\
&= \mathop{\mathbb{E}}_{S\in\mathcal{A}_d} \mathop{\mathbb{E}}_{P\in\mathcal{A}_d} \det(xI + C + (P\oplus I)(I\oplus S)D(I\oplus S)^T(P\oplus I)^T)\\
&= \mathop{\mathbb{E}}_{S\in\mathcal{A}_d} \mathop{\mathbb{E}}_{Q\in\mathcal{O}(d)} \det(xI + C + (Q\oplus I)(I\oplus S)D(I\oplus S)^T(Q\oplus I)^T) \quad \text{by Corollary 4.12}\\
&= \mathop{\mathbb{E}}_{Q\in\mathcal{O}(d)} \mathop{\mathbb{E}}_{S\in\mathcal{A}_d} \det(xI + (Q\oplus I)^T C(Q\oplus I) + (I\oplus S)D(I\oplus S)^T)\\
&= \mathop{\mathbb{E}}_{Q\in\mathcal{O}(d)} \mathop{\mathbb{E}}_{R\in\mathcal{O}(d)} \det(xI + (Q\oplus I)^T C(Q\oplus I) + (I\oplus R)D(I\oplus R)^T) \quad \text{by Corollary 4.12}\\
&= \mathop{\mathbb{E}}_{Q,R\in\mathcal{O}(d)} \det(xI + C + (Q\oplus R)D(Q\oplus R)^T),
\end{aligned}
$$

as desired. ∎

Proof of Theorem 4.10. The proof follows from Theorem 4.11, (19), and (20). ∎

Like Theorem 4.1, Theorem 4.10 extends effortlessly to the case of many matrices.

COROLLARY 4.13. *If $A_1, \ldots, A_m$ 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*

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

where the P_i and S_i are independent uniformly random permutations.

We omit the proof, which is identical to the proof of Corollary 4.9.

5. Ramanujan graphs. We now combine the Cauchy transform, interlacing, and quadrature results of the previous sections to establish Theorems 1.1 and 1.2. We start with the (less complicated) Theorem 1.2.

Proof of Theorem 1.2. 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 3.5), Theorem 3.4 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 4.9 reveals that the polynomial in the right-hand expression may be written as an m-wise symmetric additive convolution

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

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 $p(x) \boxplus_{d-1} \cdots \boxplus_{d-1} p(x)$, which is the second largest root of $E(x)$. We do this using the inverse Cauchy transform described in section 2.3. 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 2.8 $(m-1)$ times yields the following upper bound on the inverse Cauchy transform of the m-wise convolution of interest.

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

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 1.1. 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 3.5), Theorem 3.4 implies that

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

with nonzero probability. Since $I\mathbf{1} = \mathbf{1}$, Corollary 4.13 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 2.9:

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

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. ∎

REFERENCES

[1] G. W. Anderson, A. Guionnet, and O. Zeitouni, *An Introduction to Random Matrices*, 118, Cambridge University Press, Cambridge Stud. Adv. Math Cambridge, UK, 2010.

[2] J. Batson, D. A. Spielman, and N. Srivastava, *Twice-Ramanujan sparsifiers*, SIAM J. Comput., 41 (2012), pp. 1704–1721.

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

[4] M. Chudnovsky and P. Seymour, *The roots of the independence polynomial of a clawfree graph*, J. Combin. Theory Ser. B, 97 (2007), pp. 350–357.

[5] M. B. Cohen, *Ramanujan graphs in polynomial time*, in Proceedings of the Symposium on Foundations of Computer Science, IEEE, 2016, pp. 276–281.

[6] J. P. Dedieu, *Obreschkoff's theorem revisited: What convex sets are contained in the set of hyperbolic polynomials?*, J. Pure Appl. Algebra, 81 (1992), pp. 269–278, https://doi.org/10.1016/0022-4049(92)90060-S.

[7] H. J. Fell, *Zeros of convex combinations of polynomials*, Pacific J. Math., 89 (1980), pp. 43–50.

[8] S. Fisk, *Polynomials, Roots, and Interlacing*, preprint, arXiv:0612833 [math.CA], 2008.

[9] J. Friedman, *A Proof of Alon's Second Eigenvalue Conjecture and Related Problems*, Mem. Amer. Math. Soc. 910, American Mathematical Society, Providence, RI, 2008.

[10] C. Hall, D. Puder, and W. F. Sawin, *Ramanujan Coverings of Graphs*, arXiv:1506.02335 [math.co], 2015.

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

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

[13] A. W. Marcus, *Polynomial Convolutions and (Finite) Free Probability*, preprint, 2015.

[14] A. W. Marcus, D. A. Spielman, and N. Srivastava, *Finite Free Convolutions of Polynomials*, arXiv:1504.00350 [math.Co], 2015.

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

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

[17] G. A. Margulis, *Explicit group theoretical constructions of combinatorial schemes and their application to the design of expanders and concentrators*, Probl. Inf. Transm., 24 (1988), pp. 39–46.

[18] B. D. McKay, *The expected eigenvalue distribution of a large regular graph*, Linear Algebra Appl., 40 (1981), pp. 203–216.

[19] A. Nica and R. Speicher, *Lectures on the Combinatorics of Free Probability*, Cambridge University Press, Cambridge, UK, 2006.

[20] A. Nilli, *On the second eigenvalue of a graph*, Discrete Math., 91 (1991), pp. 207–210.

[21] N. Obreschkoff, *Verteilung und berechnung der Nullstellen reeller Polynome*, Deutscher Verlag der Wissenschaften, Berlin, 1963.

[22] Q. I. Rahman and G. Schmeisser, *Analytic Theory of Polynomials*, Oxford University Press, Cambridge, UK, 2002.

[23] D. A. Spielman, *Bipartite Ramanujan graphs of every degree*, Spectral Graph Theory Lecture 26, 2015, http://www.cs.yale.edu/homes/spielman/561/lect26-15.pdf.

[24] D. V. Voiculescu, *Free Probability Theory*, American Mathematical Society, Providence, RI, 1997.