SPARSE REGULAR RANDOM GRAPHS: SPECTRAL DENSITY AND EIGENVECTORS

BY IOANA DUMITRIU[1] AND SOUMIK PAL[2]

University of Washington

We examine the empirical distribution of the eigenvalues and the eigenvectors of adjacency matrices of sparse regular random graphs. We find that when the degree sequence of the graph slowly increases to infinity with the number of vertices, the empirical spectral distribution converges to the semicircle law. Moreover, we prove concentration estimates on the number of eigenvalues over progressively smaller intervals. We also show that, with high probability, all the eigenvectors are delocalized.

1. Introduction. Consider the uniform distribution over the space of all labeled simple graphs on n vertices where every vertex has degree d. We denote a graph randomly selected from this distribution by $G(n, d)$; the vertices of $G(n, d)$ will always be labeled by $\{1, 2, \ldots, n\}$.

Now, consider a sequence of such random graphs $\{G(n, d_n), n \in \mathbb{N}\}$ which are d_n-regular of order n. We assume d_n to be slowly growing with n in a manner which will be made more precise later. Consider the adjacency matrix A_n of $G(n, d_n)$; the (i, j)th element of A_n is one or zero depending on whether there is an edge between vertices i and j in the graph $G(n, d_n)$. The random matrix A_n is always symmetric, and it has n real eigenvalues (perhaps not all distinct) and corresponding real eigenspaces. Under appropriate conditions on the growth of the sequence d_n, we study the following phenomena as n tends to infinity:

(i) Global semicircle law. We prove the convergence of the empirical spectral distribution (ESD) of the scaled adjacency matrix to the probability measure on $[-2, 2]$ with density

$$(1) \qquad f_{\mathrm{sc}}(x) = \frac{1}{2\pi}\sqrt{4 - x^2}, \qquad -2 < x < 2.$$

(ii) Local semicircle law (a.k.a. the semicircle law on short scales). We obtain concentration estimates of the deviation of the number of eigenvalues $\mathcal{N}_I$ that lie in a small interval I from its predicted number $n \int_I f_{\mathrm{sc}}(x)\, dx$. The size of I will be taken to be vanishing at an appropriate rate with increasing n.

(iii) Delocalization of eigenvector coordinates. We obtain probability estimates of the event that, for some eigenvector, a few of the coordinates are significantly larger in magnitude than the rest.

These problems connect two areas of study: Wigner random matrices and spectra of sparse random graphs; since we have already mentioned the latter, we will now talk about the former.

In the last couple of decades there has been an enormous amount of activity in the study of universal properties of random matrices, inspired by their connection to (universal) physical systems. The literature on universality studies in random matrices is vast; we mention here only a few references and ask the reader to look to them for further ones.

For an introduction and motivation to the subject, we recommend Deift's ICM address [17]. Of particular interest are the Wigner matrices (see Bai [3], Soshnikov [54], Bai and Yao [6], Khorunzhy, Khoruzhenko and Pastur [38], Guionnet and Zeitouni [33], Ben Arous and Peche [11], Tao and Vu [55, 56]). Deift and Goiev [18] looked at different potential functions on symmetric, Hermitian and self-dual matrices; Baik and Suidan explored connections to percolation [7] and random walks [8]; β-generalizations of the classical ensembles and universal properties thereof have been explored in Forrester and Baker [9], Johansson [35], Dumitriu and Edelman [20]. Recently, more sophisticated probability tools have been generalized and applied to random matrix theory (e.g., the Lindeberg Principle, by Chatterjee [16] and Tao and Vu [55, 56]).

Most of the focus in universality research has been on proving, under progressively weaker assumptions on the entry distribution, the following:

- convergence of the ESD to the semicircle or Marčenko–Pastur laws and establishing rates of convergence in various ways (large deviations, concentration estimates, central limit theorems). For a comprehensive treatment of the subject, see the books by Bai and Silverstein [5] and by Anderson, Guionnet and Zeitouni [2];
- fluctuations of the spectrum at the edge (the famous Tracy–Widom laws [57–59]) for general Wigner matrices, settled by Tao and Vu [55] and Erdős et al. [24];
- universality of correlation functions in the bulk, under various assumptions (Tao and Vu [56], Erdős et al. [24], Erdős et al. [23], Erdős, Yau and Yin [27]);
- partial or complete delocalization of the eigenvectors (by Erdős, Schlein and Yau [25, 26] and Tao and Vu [55, 56]).

The aim of said research has been to show that the spectral statistics agree in the large n limit to the spectral statistics of the Gaussian Orthogonal Ensemble (GOE) and the Gaussian Unitary Ensemble (GUE), depending on whether the matrices are real symmetric/positive-definite, or complex Hermitian/positive-definite. The Gaussian and Wishart ensembles are some of the most studied and best understood

random matrix models; for an easy introduction to classical random matrix theory, see the books by Mehta [46] and Muirhead [50].

Parallel to these developments, in combinatorics and discrete mathematics, there has always been an interest in studying spectral properties of deterministic and random graphs. There are two matrices of interest in spectral graph theory: the *adjacency* (sometimes called the *incidence*) matrix, which we already defined, and the *Laplacian* matrix. These matrices are the same for regular graphs (although, in general, they can be quite different), and the spectrum of the graph is the spectrum of the matrix.

Among the properties of random graphs that have been the focus of intense research are connectivity, phase transitions and the limiting spectral distribution of random graphs, including trees (McKay [43], Feige and Ofek [28], Mirlin and Fyodorov [48], Bauer and Golinelli [10], Semerjian and Gugliandolo [52], Bordenave and Lelarge [13], Bhamidi, Evans and Sen [12]). Other properties include concentration of eigenvalues (Krivelevich and Sudakov [39], Alon, Krivelevich and Vu [1]), the spectral gap (Fűredi and Komlós [32], Friedman [29] and Friedman and Alon [31], Broder and Shamir [15]).

Another area of recent interest is the study of quasi-random graphs and expanders. These are nonrandom graphs which display properties one expects to hold with high-probability for certain classes of random graph models. For example, expanders are sparse graphs that have high connectivity properties (e.g., a large spectral gap). These graphs are often regular (e.g., the famous Ramanujan graph, described in the seminal articles by Lubotzky, Phillips and Sarnak [41] and Morgenstern [49]). Random d-regular graphs display the same connectivity properties with very high probability, when d is kept fixed and the order is large; this is in essence the Alon conjecture, recently settled by Friedman [31]. Thus a study of random regular graphs suggests possible properties of (deterministic) expanders.

It is easy for a probability audience to appreciate the importance of studying eigenvalues of the graph (e.g., the spectral gap which determines the mixing properties of a random walk), but eigenvectors of graphs are equally important, especially since they are the solutions of various combinatorial optimization problems. Traditionally, there has been much less work on computing the *actual* graph eigenvector distributions, with the notable and recent exception of Wishart-like sample covariance matrices (see Bai, Miao and Pan [4]). Thus, developments in examining properties of the eigenvectors of large random graphs (as in Friedman [30] and Dekel, Lee and Linial [19]) are relatively new, and motivated by the applications of eigenvectors to engineering and computer science. Such applications include the Google page-rank algorithm [14], the Shi–Malik algorithm [53], the Meila–Shi algorithm [47] and other spectral clustering techniques and related segmentation problems (Weiss [61], Pothen, Simon and Liou [51], etc.).

It is probably clear by now that the two fields of research that we have very briefly sketched here (universality studies in random matrix theory and spectra of random graphs) are vast and, by examining the two lists of important problems we

have outlined, one can see that there is a certain amount of overlap. Naturally, this lead to a few papers where the two fields have intersected, despite differences in both the goals and the methodology of each.

A famous such example is McKay's derivation of the limiting empirical spectrum of random d-regular graphs on n vertices, as d is fixed and n grows to infinity [43]. In that case, the empirical spectral distribution converges in probability to what is known as the McKay (or Kesten–McKay) law, which has a density

$$(2) \qquad f_d(x) = \frac{d\sqrt{4(d-1) - x^2}}{2\pi(d^2 - x^2)}, \qquad -2\sqrt{d-1} \le x \le 2\sqrt{d-1}.$$

This density had appeared earlier in Kesten's work on random walks on groups [36]. It can be easily verified that as d grows to infinity, if we normalize the variable x in the above by $\sqrt{d-1}$, the resulting density converges to the semicircle law on $[-2, 2]$.

This naturally raised the question of whether the study of "universal" properties could be pushed into the domain of regular random graphs with increasing degree. The rate of growth of the degree sequence plays an important role, since at both extremes (d fixed and $d = n - 1$) the ESDs do not converge to the semicircle law.

The answer to this question turns out to be difficult. There are a number of major obstacles to developing an applicable universality theory in the spirit of Wigner random matrices to adjacency matrices of random graphs, which are non-Wigner: these matrices are sparse and the entries are not independently distributed.

To see how sparsity affects concentration, consider the question of proper scaling of the adjacency matrix A_n. In the Wigner case, the scaling factor is clearly $1/\sqrt{n}$, which puts *all* of the eigenvalues in $[-2 - \varepsilon, 2 + \varepsilon]$, for any positive ε, with very high probability for a sufficiently large n. One might be tempted then to believe that the proper scaling for adjacency matrices is $1/\sqrt{d_n}$, as this achieves the same kind of finite row-variance as $1/\sqrt{n}$ does in the Wigner case. Unfortunately, it is not known if this scaling will place all the eigenvalues (except the first) in a compact interval, as $d_n \to \infty$.

For the regime when d is fixed, the Alon conjecture states that the second largest eigenvalue (in absolute value) λ_2 has an upper bound $|\lambda_2| \le 2\sqrt{d-1} + \varepsilon$, with very high probability. The well-known lower bound holds for every d-regular graph and we cite it from Friedman [29]: $|\lambda_2| \ge 2\sqrt{d-1} + O(\log d / \log n)$. Unfortunately, when d grows with n, the upper bound is not known to hold outside of a narrow growth regime.[3] Khorunzhy [37] has shown that, for a random matrix model similar to the adjacency matrix of the Erdős–Renýi random graph on n vertices with an expected degree $d_n \gg \log n$, with probability one, the spectral norm

---

[3]More precisely, even under this narrow growth regime the theorem is valid only for Friedman's permutation model. For d fixed, Friedman's model approaches the uniform distribution with increasing n; this is no longer clear once d_n grows with n.

of the adjacency matrix grows faster than $\sqrt{d_n}$. Although this does not necessarily affect convergence of the ESD to the semicircle law, it eliminates the possibility of containing all the eigenvalues of the rescaled centralized adjacency matrix within any compact interval.

Our results investigate the extent to which universality can be extended to the slowly growing d_n case. Our first result is Theorem 1 stated below.

THEOREM 1. *Let d_n satisfy the asymptotic condition*

$$(3) \qquad \lim_{n \to \infty} d_n = \infty, \qquad d_n - 1 = n^{\varepsilon_n} \qquad \text{for some } \varepsilon_n = o(1).$$

Then the ESD of the matrix $(d_n - 1)^{-1/2} A_n$, where A_n denotes the adjacency matrix of G_n, converges in distribution to the semicircle law on $[-2, 2]$ which has a density

$$(4) \qquad f_{sc}(x) := \frac{1}{2\pi}\sqrt{4 - x^2}, \qquad -2 < x < 2.$$

The condition on d_n, for example, includes the logarithmic regime, $d_n = (\log n)^\gamma$ for any positive γ; in which case we can define ε_n as $\gamma \log \log n / \log n$.

Our proof of this result (and the following ones) depends crucially on two facts:

(i) the "locally tree-like" property, which states that with high probability, most vertices in a random regular graph will have a (increasingly larger) neighborhood which is free of any cycles, and

(ii) the fact that d_n grows to infinity, which smooths out irregularities as n tends to infinity.

Our second result is arguably the most important one in this paper.

THEOREM 2. *Fix $\delta > 0$. Let $d_n = (\log n)^\gamma$, where $\gamma > 0$. Let $\eta_n = (r_n - r_n^{-1})/2$ where $r_n = \exp(d_n^{-\alpha})$ for some $0 < \alpha < \min(1, 1/\gamma)$. Then there exists an N large enough such that for all $n \geq N$, for any interval $I \subset \mathbb{R}$ of length $|I| \geq \max\{2\eta_n, \eta_n/(-\delta \log \delta)\}$,*

$$\left| \mathcal{N}_I - n \int_I f_{sc}(x)\,dx \right| < \delta n |I|$$

with probability at least $1 - o(1/n)$. Here $\mathcal{N}_I$ is the number of eigenvalues of $\frac{1}{\sqrt{d_n - 1}} A_n$ in the interval I, and f_{sc} refers to the density of the semicircle law as in (1).

REMARK 1. Note that the shortest length of the interval I that our methods can narrow down to is of length η_n, which is roughly about $1/\log n$, if $d_n \gg \log n$, and $1/d_n$, if $d_n \ll \log n$. For Wigner matrices a far shorter scale can be achieved (effectively poly-log over n in [25]). Such sharp estimates are not to be expected in the graph case, and this again is a consequence of sparsity and lack of concentration estimates.

REMARK 2. A close examination of the proof of Theorem 2 reveals that it can be extended to any deterministic sequence of regular graphs of increasing size and degree, as long as the "locally tree-like" property holds at "most" vertices.

REMARK 3. Since the submission of this paper, significant progress has been made in proving the local semicircle law for random regular graphs in *any* kind of growth regime for d_n (see Tran, Vu and Wang [60]). Their methods rely on proving the local semicircle law first for Erdős–Rényi graphs with suitable parameters, and then using a result by McKay and Wormald [44] about the probability that an Erdős–Rényi graph is regular. Their result subsumes ours (in the sense that the lower bound on the length of the interval I is smaller) for the case when $d_n = \Omega((\log n)^{10})$; when $d_n = o((\log n)^{10})$, our result is slightly stronger in the same sense.

Although these results are similar up to a point to the Wigner matrix results, our methodology is essentially different. Due to sparsity and lack of concentration, we had to adapt a more combinatorial set of tools (in particular, the tree approximation) as well as tools from linear algebra to the Stieltjes transform approach used in [25] and [56].

Several recent articles have done extensive simulations on eigenvalues and eigenvectors of random graphs, with surprising conclusions. For example, Jakobson et al. [34] carries out a numerical study of fluctuations in the spectrum of regular graphs. Their experiments indicate that the level spacing distribution of a generic k-regular graph approaches that of the GOE as we increase the number of vertices. On the eigenvector front, in the article [21] by Elon, the author attempts to characterize the structure of the eigenvectors by suggesting (with numerical observations) that all, except the first, follow approximately a Gaussian distribution. Additionally, the local covariance structure has been conjectured to be given by explicit functions of the Chebyshev polynomials of the second kind. In particular, if two vertices on the graph are at a distance k from each other, it is conjectured that the covariance between the coordinates of any eigenvector at the two vertices decays exponentially in k.

All this empirical data points to universality properties of the adjacency matrices of large, sparse regular graphs; we took here a first step toward proving them.

If the eigenvectors are indeed uniformly distributed over the sphere then they must (with high probability) satisfy delocalization. We give upper bounds on the probability of this phenomenon.

We use the following definition of delocalization, similar to the one used in [25].

DEFINITION 1. Let T be a subset of $\{1, 2, \ldots, n\}$ of size $L \geq 1$. Let $\delta > 0$ be some fixed number. We say that a vector $v = (v(1), \ldots, v(n)) \in \mathbb{R}^n$ with $\mathbb{L}^2$ norm $\|v\|_2 = 1$ exhibits (T, δ) localization if

$$\|v\mid_T\|_2^2 = \sum_{j \in T} |v(j)|^2 \geq 1 - \delta.$$

The vector v is said to be (L, δ) localized there exists some set $T \subset \{1, 2, \ldots, n\}$ such that $|T| = L$ such that v is (T, δ) localized.

Below is our result on eigenvector delocalization.

THEOREM 3. *Assume the set-up of Theorem* 2. *Fix* $\delta > 0$.

(i) *Let* $T_n \subseteq \{1, 2, \ldots, n\}$ *be a deterministic sequence of sets of size* $L_n = o(\eta_n^{-1})$. *Let* $\Omega_1(n)$ *be the event that some normalized eigenvector of the matrix* A_n *is* (T_n, δ) *localized. Then, for all sufficiently large* n,

$$P((\Omega_1(n))^c) \geq e^{-L_n \eta_n / d_n} \left(1 - o\left(\frac{1}{d_n} \right) \right) = 1 - o\left(\frac{1}{d_n} \right).$$

(ii) *Define the sequence*

$$\zeta_n = \frac{1}{4} \frac{\log n}{\log(d_n - 1)} - 4, \qquad n \geq 2.$$

Consider the (*random*) *subset* $J(n)$ *of all vertices in the graph whose* ζ_n-*neighborhood is free of cycles. Then,*

$$P\left(\frac{|J(n)|}{n} \geq 1 - \frac{\eta_n}{d_n} \right) = 1 - o\left(\frac{1}{n} \right).$$

Moreover, there exists an n *large enough such that the event that* $T_n \subset J(n)$ *and some normalized eigenvector is* (T_n, δ) *has probability zero.*

REMARK 4. More progress has been made on the eigenvector delocalization front since the submission of this paper. In their paper [60], the authors prove that the ℓ^∞ norms of all eigenvectors are $o(1)$, regardless of the regime of growth of d_n. Very recently, Erdős et al. posted a paper [22] proving the local semicircle law for Erdős–Rényi graphs with $pn = \Omega(\log n)$ up to a spectral window (an interval I) of size larger than $1/n$; from this, they could deduce that the eigenvectors of such Erdős–Rényi graphs are completely delocalized, that is, that the ℓ^∞ norms of the normalized (unit) eigenvectors are at most of order $1/\sqrt{N}$ with high probability. It would be interesting to see if the methods of [60] (of deducing results for random regular graphs from the same results for Erdős–Rényi) can be combined with the theorems of [22] to obtain complete eigenvector delocalization (and, potentially, a much smaller spectral window) for random regular graphs with $d_n = \Omega(\log n)$.

The bounds in Theorem 3 are not sharp. There are severe technical obstacles in producing sharp bounds by adapting the strategy of Wigner matrices. One such example is eigenvalue collision, that is, the event that A_n does not have n distinct eigenvalues. Since A_n has discrete entries, this event has a positive probability. However, to the best of our knowledge, no good bound on this probability is known.

The paper is organized as follows. In Section 2 we prove the global convergence to the semicircle law (Theorem 1). This is followed by the proof of the local semicircle law (Theorem 2) in Section 3. The eigenvector delocalization is proved in Section 4. Finally, the Appendix contains an exact calculation of eigenvalues and eigenvectors for the random regular (finite) tree, defined in Section 2.

2. Global convergence to the semi-circle law. Recall that $G_n = G(n, d_n)$ denotes a random d_n-regular graph on n vertices whose adjacency matrix is A_n. Recall that d_n satisfies the asymptotic condition

$$(5) \qquad \lim_{n \to \infty} d_n = \infty, \qquad d_n - 1 = n^{\varepsilon_n} \qquad \text{for some } \varepsilon_n = o(1).$$

We prove here Theorem 1, namely, that the empirical spectral distribution (ESD) of the adjacency matrix A_n converges in probability to the semicircle law on $[-2, 2]$ which we recall from (4).

Our main instrument is to use the *moment method*. Our arguments depend crucially on the following local approximation of G_n by a rooted tree. Consider the deterministic rooted tree S_n, which is the infinite regular tree of degree d_n with a distinguished vertex marked as the *root*. For a graph G whose every edge is taken to have unit length, consider the induced metric structure on G. We define the r-neighborhood of the vertex i, to be the subgraph of G whose vertices are at a distance at most r from i, and whose edges are all the edges between those vertices. The following lemma makes precise the idea that, except for a vanishing proportion of the vertices, the r-neighborhood of any vertex is isomorphic to the corresponding neighborhood of the root in the tree S_n.

Recall that a cycle is a sequence of vertices $\{i_1, \ldots, i_k\}$ of a graph such that $i_1 = i_k$, there is no other repeated vertex, and there is an edge between every successive i_j and i_{j+1}. The length of the cycle is the number of vertices except the initial one. A cycle of length k will be called a k-cycle. Finally, a cycle-free or *acyclic* graph is a tree.

LEMMA 4. *Fix a positive integer r. Let $\tau(n)$ be the subset of vertices of G_n which have no cycles in their r-neighborhoods, and let $|\tau(n)|$ denote the size of $\tau(n)$. Then, under the assumptions of (5), we have*

$$P\left(1 - \frac{|\tau(n)|}{n} > n^{-1/4}\right) = o(n^{-5/4}).$$

PROOF.   We use the estimates of McKay, Wormald and Wysocka [45] on the Poisson approximation to the number of short cycles in regular graphs. Let $g(n)$ be a sequence such that

$$(6) \qquad\qquad g(n) \geq 3 \quad \text{and} \quad (d_n - 1)^{2g(n)-1} = o(n).$$

For any $s \leq g$, let M_s denote the number of cycles of length s in the graph G_n. It has been shown in [45] that M_s is approximately distributed as a Poisson random variable and

$$E(M_s) = \mu_s\left(1 + O\left(s(s+d)/n\right)\right) \qquad \text{where } \mu_s = \frac{(d-1)^s}{2s} \quad \text{and}$$

(7)

$$\text{Var}(M_s) = \mu_s + O\left(s(s+d)/n\right)\mu_s^2.$$

Consider now the growth of the degree sequence as in (5). If we choose $g(n)$ such that $2g(n) - 1 = 1/\sqrt{\varepsilon_n}$, it will satisfy

$$(d_n - 1)^{2g(n)-1} = n^{\sqrt{\varepsilon_n}} = o(n).$$

Additionally $g(n)$ grows to infinity with n, since $\varepsilon(n) = o(1)$.

Consider an s-cycle for some $s \leq g(n)$. It has exactly s-vertices. Now, the number of vertices whose r neighborhoods fail to be acyclic because of this s-cycle are precisely those vertices which are at a distance of at most $(2r - s)/2$ from any of the vertices in the s-cycle. The number of such vertices has an easy upper bound of $2(d_n - 1)^{(2r-s)/2}s$, for all large enough d_n. Thus, the total number of vertices whose r neighborhoods are not acyclic can be bounded above by

(8)
$$N_r^* = \sum_{s=3}^{2r} 2s(d_n - 1)^{(2r-s)/2}M_s.$$

Also,

(9)
$$n - |\tau(n)| \leq N_r^*.$$

Taking expectations on both sides of (8), and using formulas (7), we get

$$EN_r^* = \sum_{s=3}^{2r} 2s(d - 1)^{(2r-s)/2}\frac{(d-1)^s}{2s}\left(1 + O\left(s(s+d)/n\right)\right).$$

The quantity $O(s(s+d)/n)$ denotes a function $h(s, d, n)$ such that

$$\frac{n}{s(s+d)}h(s, d, n)$$

remains bounded for all choices of s, d and n. Thus we get

$$EN_r^* = (d-1)^r\sum_{s=3}^{2r}(d-1)^{s/2} + O\left(\frac{1}{n}\sum_{s=3}^{2r}s(s+d)(d-1)^{r+s/2}\right) = O\left((d-1)^{2r}\right).$$

The last equality is true since, by our assumption on d_n, the second term in the sum is $o(1)$.

Similarly, we can compute the second moment. By the Cauchy–Schwarz inequality,

$$\mathrm{Var}(N_r^*) \le 2r \sum_{s=3}^{2r} 4s^2(d-1)^{2r-s} \mathrm{Var}(M_s)$$

$$\le 2r \sum_{s=3}^{2r} 4s^2(d-1)^{2r-s} \big[\mu_s + O\big(s(s+d)/n\big)\mu_s^2\big]$$

$$\le 2r \sum_{s=3}^{2r} 4s^2(d-1)^{2r-s}\mu_s + 2r \sum_{s=3}^{2r} 4s^2(d-1)^{2r-s} O\big(s(s+d)/n\big)\mu_s^2.$$

Plugging in the value of μ_s from (7) we get

$$\mathrm{Var}(N_r^*) \le 4r^2(2r+1)(d-1)^{2r} + 2r(d-1)^{2r} \sum_{s=3}^{2r} (d-1)^s O\big(s(s+d)/n\big).$$

As before, it thus follows that

$$\sum_{s=3}^{2r} (d-1)^s O\big(s(s+d)/n\big) = O\Bigg(\sum_{s=3}^{2r} (d-1)^s s(s+d)/n\Bigg)$$

$$= O\Bigg(n^{-1}(2r+d)(d-1)^{2r} \sum_{s=3}^{2r} s\Bigg)$$

$$= O\big(n^{-1}(2r+d)(d-1)^{2r} r(2r+1)\big).$$

Hence

$$\mathrm{Var}(N_r^*) \le 4r^2(2r+1)(d-1)^{2r} + O\big(r^2(2r+d)(d-1)^{4r}/n\big).$$

Note again that, by our assumption, the quantity $(d-1)^{4r}/n$ is $o(1)$.

We now want to use Markov's inequality to bound the tail probability of the quantity $1 - |\tau(n)|/n$. Fix any $\varepsilon > 0$. Then, by inequality (9), we get

$$P\bigg(1 - \frac{|\tau(n)|}{n} > \varepsilon\bigg) \le P(N_r^* > n\varepsilon) \le \frac{1}{n^2\varepsilon^2} E(N_r^*)^2$$

$$= \frac{1}{n^2\varepsilon^2}\big[\mathrm{Var}(N_r^*) + (E(N_r^*))^2\big]$$

$$\le \frac{1}{n^2\varepsilon^2}\big[4r^2(2r+1)(d-1)^{2r}$$

$$+ r^2(2r+d)o(1) + O\big((d-1)^{4r}\big)\big]$$

$$\le \varepsilon^{-2} O\bigg(\frac{(d-1)^{4r}}{n^2}\bigg)$$

$$= \varepsilon^{-2} O(n^{4r\varepsilon_n - 2})$$

by our choice of the sequence d_n.

Choosing $\varepsilon = n^{-1/4}$ we get

$$P\left(1 - \frac{|\tau(n)|}{n} > n^{-1/4}\right) \leq \sqrt{n} O(n^{4r\varepsilon_n - 2}) = o(n^{-5/4}),$$

since $\varepsilon_n = o(1)$. This completes the proof of the lemma. $\quad\square$

LEMMA 5. *Let $\{\mu_i, i = 1, 2, \ldots\}$ be a sequence of random probability measures on the real line, defined on the same probability space. Let μ be a nonrandom continuous probability measure supported on a compact interval I. Suppose there exits a pair of doubly indexed real-valued sequences $\{a_n(r), b_n(r), r, n \in \mathbb{N}\}$ such that the following hold:*

(1) *For every $r = 1, 2, \ldots$, we have*

$$P\left(\bigcup_{N=1}^{\infty} \bigcap_{n \geq N} \left\{a_n(r) \leq \int x^r d\mu_n(x) \leq b_n(r)\right\}\right) = 1.$$

(2) *For every $r = 1, 2, \ldots$, we have*

$$\lim_{n \to \infty} a_n(r) = \lim_{n \to \infty} b_n(r) = \int x^r d\mu(x) < \infty.$$

Then the sequence of measures $\{\mu_n\}$ converges to μ in probability.

PROOF. Let Ω_r be the event

$$\bigcup_{N=1}^{\infty} \bigcap_{n \geq N} \left\{a_n(r) \leq \int x^r d\mu_n(x) \leq b_n(r)\right\}.$$

Then, from condition (1), it follows that

$$1 - P\left(\bigcap_{r=1}^{\infty} \Omega_r\right) = P\left(\bigcup_{r=1}^{\infty} \Omega_r^c\right) \leq \sum_{r=1}^{\infty} P(\Omega_r^c) = 0.$$

Thus $P(\bigcap_{r=1}^{\infty} \Omega_r) = 1$.

Consider any fixed realization of the sequence $\{\mu_n\} \in \bigcap_{r=1}^{\infty} \Omega_r$. By Helly's selection theorem, this sequence has a limit point ν. Thus, there is a subsequence $\{\mu_{n_k}\}$ that converges to ν in the topology of weak convergence.

Now take r to be a positive integer. We would like to show that

$$\lim_{n_k \to \infty} \int x^r d\mu_{n_k} = \int x^r d\nu.$$

From the standard theory of weak convergence, it follows that this will be true if the function x^r is uniformly integrable under the sequence of measures $\{\mu_{n_k}\}$.

However, uniform integrability follows from the following L^2-boundedness condition:

$$\max_{n_k} \int x^{2r} \, d\mu_{n_k} < \max_{n_k} b_{n_k}(2r) < \infty$$

by conditions (1) and (2).

In particular, from condition (2) we reach the conclusion

$$\int x^r \, d\nu(x) = \int x^r \, d\mu(x), \qquad r = 0, 1, 2, \dots.$$

Since the support of μ is the compact interval I, it follows that the moment problem has a unique solution, and hence, ν must be equal to μ.

This shows that any limit point of any sequence $\{\mu_n\}$ in $\bigcap_{r=1}^{\infty} \Omega_r$ is given by μ. By the usual subsequence argument, this shows that μ_n converges to μ in the set $\bigcap_{r=1}^{\infty} \Omega_r$, and hence with probability one. This proves the result.   □

PROOF OF THEOREM 1.   Consider the random graph sequence $G_n = G(n, d_n)$ as in the statement, and let A_n be the adjacency matrix of G_n.

Let μ_n be the ESD of the matrix $(d_n - 1)^{-1/2} A_n$. Then, for any positive integer r,

$$\int x^r \, d\mu_n(x) = \frac{1}{n} \operatorname{tr}((d_n - 1)^{-r/2} A_n^r) = \frac{(d_n - 1)^{-r/2}}{n} \sum_{i=1}^{n} A_n^r(i, i).$$

Here $A_n^r(i, i)$ is the ith diagonal element of the matrix A_n^r.

Note that $A_n^r(i, i)$ counts the number of paths of length r that start and end at i. Consider the set of vertices in $\tau(n)$, as in Lemma 4, whose $\lceil r/2 \rceil$-neighborhood is acyclic. For any $i \in \tau(n)$, the number of such paths, B_n^r, is equal to the number of paths of size r that start and end at the root of the tree S_n. If $i \notin \tau(n)$, we use the trivial bound $A_n^r(i, i) \leq d_n^r$. Thus

$$(d_n - 1)^{-r/2} \frac{|\tau(n)|}{n} B_n^r \leq \frac{(d_n - 1)^{-r/2}}{n} \sum_{i=1}^{n} A_n^r(i, i)$$

$$\leq (d_n - 1)^{-r/2} \left[B_n^r + \frac{n - |\tau(n)|}{n} d_n^r \right].$$

If we define

(10)
$$a_n(r) = (1 - n^{-1/4})(d_n - 1)^{-r/2} B_n^r,$$
$$b_n(r) = (d_n - 1)^{-r/2} [B_n^r + n^{-1/4} d_n^{r/2}],$$

then from Lemma 4 we get

(11) $$P\left(a_n(r) \leq \int x^r \, d\mu_n(x) \leq b_n(r) \right) \geq 1 - o(n^{-5/4}).$$

In particular, by taking complements of the events above, we get

$$\sum_{n=1}^{\infty} P\left(\left\{a_n(r) \leq \int x^r \, d\mu_n(x) \leq b_n(r)\right\}^c\right) < \sum_{n=1}^{\infty} o(n^{-5/4}) < \infty.$$

Now consider a product probability space on which independent copies of our (countably many) random graphs are defined. Applying the Borel–Cantelli lemma and (11) we get that

$$P\left(\bigcap_{N=1}^{\infty} \bigcup_{n=N}^{\infty} \left\{a_n(r) \leq \int x^r \, d\mu_n(x) \leq b_n(r)\right\}^c\right) = 0.$$

Taking the complements again, we get

$$P\left(\bigcup_{N=1}^{\infty} \bigcap_{n=N}^{\infty} \left\{a_n(r) \leq \int x^r \, d\mu_n(x) \leq b_n(r)\right\}\right) = 1.$$

This satisfies condition (1) in Lemma 5.

Once we show the validity of condition (2) for μ equal to the semicircle law, we will be done by Lemma 5. Clearly, by our choice of d, n as in the statement, and the functions $a_n(r), b_n(r)$ as in (10), this will be true once we establish

$$\lim_{n\to\infty} (d_n - 1)^{-r/2} B_n^r = \int x^r f_{sc}(x) \, dx.$$

We only need to verify above for even r, since for odd r, both sides are zero ($B_n^r = 0$ since in a tree one cannot return to the root in an odd number of steps, and the moment is zero since f_{sc} is a symmetric density).

Now, for an even r, the value of $B_n(r)$ has been computed by McKay in [43] (denoted by $\theta(r)$ in equation (15) on [43]). It is given by

$$B_n(r) = \int_{-2\sqrt{d_n-1}}^{2\sqrt{d_n-1}} x^r f_n(x) \, dx,$$

where $f_n(x)$ is the Kesten–McKay density

$$f_n(x) = \frac{d_n\sqrt{4(d_n-1) - x^2}}{2\pi(d_n^2 - x^2)}, \qquad -2\sqrt{d_n-1} < x < 2\sqrt{d_n-1}.$$

Thus, changing variable to $y = (d_n - 1)^{-1/2} x$, we get

$$\lim_{n\to\infty} (d_n - 1)^{-r/2} B_n^r$$

$$= \lim_{n\to\infty} \frac{1}{(d_n-1)^{r/2}} \int_{-2\sqrt{d_n-1}}^{2\sqrt{d_n-1}} x^r \frac{d_n\sqrt{4(d_n-1) - x^2}}{2\pi(d_n^2 - x^2)} \, dx$$

$$= \lim_{n\to\infty} \int_{-2}^{2} y^r \frac{d_n\sqrt{d_n-1}\sqrt{4 - y^2}}{2\pi(d_n-1)^2((d_n/(d_n-1))^2 - y^2/(d_n-1))} (d_n - 1)^{1/2} \, dy$$

$$= \frac{1}{2\pi} \int_{-2}^{2} y^r f_{sc}(y) \, dy.$$

The last equality follows by the dominated convergence theorem and the fact that $\lim_{n\to\infty} d_n = \infty$. This completes our proof. $\square$

3. Estimating the rate of convergence of the ESD. This is the longest section of the paper, and it is quite technical, so we provide an outline of the proof. The approach we will use is given by the Stieltjes transform of the adjacency matrix of the graph. To estimate how far the Stieltjes transform of the graph is from the Stieltjes transform of the semicircle, we will use as a stepping stone the resolvent of the adjacency matrix of a finite regular tree, which we will show to be very close to both.

The estimation consists of the following steps:

Step 0. Basic definitions and properties of the quantities involved (Section 3.1).

Step 1. Compute the resolvent of the regular tree, and show that, in a certain growth regime for d_n, its (root, root) elements is close to the Stieltjes transform of the semicircle (Section 3.2).

Step 2. Show that, in the same growth regime as before, the (root, root) element of the resolvent of the regular tree is very close to the Stieltjes transform of the regular graph (Section 3.3).

Step 3. Use the estimations from the previous steps to conclude that the Stieltjes transform of the regular graph is close to that of the semicircle, and use the methods of [56] to obtain bounds on the rate of convergence of the ESD (Section 3.4).

3.1. *Basic definitions.*

DEFINITION 2. For a $n \times n$ Hermitian matrix A and a variable $z \in \mathbb{C}$ for which $\text{Im}(z) > 0$ (thus z is not an eigenvalue of A), define the Stieltjes transform to be the function

$$s(A; z) := \frac{1}{n} \text{tr}(A - zI_n)^{-1}$$

$$:= \frac{1}{n} \text{tr}(A - z)^{-1}.$$

Here I_n is the $n \times n$ identity matrix; for convenience, we will drop the identity matrix and use the second notation.

We will also require the notion of Chebyshev orthogonal polynomials of a complex variable. For more details, see the book by Mason and Handscomb [42], page 14.

For a complex number z, define

$$w = z + \sqrt{z^2 - 1}, \qquad z = \tfrac{1}{2}(w + w^{-1}),$$

where the square root of a complex number is taken such that the imaginary part is always positive. It can be verified easily that for any $r > 1$, the set

(12) $$E_r := \{z : |w| = r\}$$

is an ellipse whose foci are at $\{+1, -1\}$.

Note that

$$\max_{z \in E_r} |\Im(z)| = \frac{r - 1/r}{2},$$

and that when $r = 1$, this ellipse degenerates to the interval $[-1, 1]$.

DEFINITION 3. The nth Chebyshev polynomial of the second kind U_n is defined as

(13) $$U_n(z) = \frac{w^{n+1} - w^{-(n+1)}}{w - w^{-1}}, \qquad n = 1, 2, \dots,$$

with $U_0(z) \equiv 1$. It is easy to check that, in addition, $U_n(z)$ satisfies the recursion

(14) $$U_n(z) = 2z U_{n-1}(z) - U_{n-2}(z), \qquad n = 1, 2, \dots,$$

with the initial conditions $U_0(z) = 1$, $U_{-1}(z) = 0$.

REMARK 5. When $r = 1$, the above gives us the traditional orthogonal polynomials for the semicircle law on the interval $[-1, 1]$.

We will need the following bound on U_n which can be found in [42], equations (1.53), (1.55):

(15) $$\frac{r^n - r^{-n}}{r + r^{-1}} \le |U_{n-1}(z)| \le \frac{r^n - r^{-n}}{r - r^{-1}}, \qquad z \in E_r.$$

Finally, we will need the standard formula for inverses of symmetric block matrices, given below.

PROPOSITION 6. Let $\mathbb{A}$ and $\mathbb{D}$ be complex symmetric matrices with sizes $n \times n$, respectively, $m \times m$, and let $\mathbb{B}$ be an $m \times n$ real matrix. Define the $(m + n) \times (m + n)$ complex symmetric matrix

$$\mathbb{M} = \begin{bmatrix} \mathbb{A} & \mathbb{B} \\ \mathbb{B}' & \mathbb{D} \end{bmatrix},$$

where $\mathbb{B}'$ denotes the transpose of $\mathbb{B}$. Then

(16) $$\mathbb{M}^{-1} = \begin{bmatrix} \mathbb{A}^{-1} + \mathbb{A}^{-1}\mathbb{B}\mathbb{F}^{-1}\mathbb{B}'\mathbb{A}^{-1} & -\mathbb{A}^{-1}\mathbb{B}\mathbb{F}^{-1} \\ -\mathbb{F}^{-1}\mathbb{B}'\mathbb{A}^{-1} & \mathbb{F}^{-1} \end{bmatrix}, \qquad \mathbb{F} = \mathbb{D} - \mathbb{B}'\mathbb{A}^{-1}\mathbb{B}.$$

Equivalently, by reversing the roles of the blocks $\mathbb{A}$ and $\mathbb{D}$,

(17) $$\mathbb{M}^{-1} = \begin{bmatrix} \mathbb{G}^{-1} & -\mathbb{G}^{-1}\mathbb{B}\mathbb{D}^{-1} \\ -\mathbb{D}^{-1}\mathbb{B}'\mathbb{G}^{-1} & \mathbb{D}^{-1} + \mathbb{D}^{-1}\mathbb{B}'\mathbb{G}^{-1}\mathbb{B}\mathbb{D}^{-1} \end{bmatrix},$$

$$\mathbb{G} = \mathbb{A} - \mathbb{B}\mathbb{D}^{-1}\mathbb{B}'.$$

These formulas are easy to verify, and their proofs can be found in standard matrix algebra books.

3.2. *Resolvents of regular and almost regular trees.* Fix a positive integer $d \geq 2$. Let $\mathbb{T}$ be a finite ordered rooted tree of depth $\zeta \in \mathbb{N}$ such that every vertex has exactly $(d-1)$ *children.* That is, the root has degree $(d-1)$, and every other vertex, except the leaves, has degree d. Such a tree is almost regular since all vertices, excluding the root and the leaves, have degree d.

In order to define the adjacency matrix of this graph, we must fix a labeling; we will define this labeling recursively down to $\zeta = 0$, in which case all we have is a root vertex which we label 1.

Imagine the tree embedded in the plane. If the depth is zero, the only element is the root, and the adjacency matrix is obvious. If the depth is one, the root has $(d-1)$ children. Consider each child vertex as a tree of depth zero, order their adjacency matrices $H_1, H_2, \ldots, H_{d-1}$ from left to right, and consider a block matrix with these as the diagonal blocks from upper left to bottom right. Finally add a bottom-most row and a rightmost column for the root vertex.

By induction, suppose we have labeled the adjacency matrix for the tree of depth $\zeta - 1$. Consider now the tree of depth ζ. If we remove the root and the edges incident to it, we are left with $(d-1)$ trees of depth $\zeta - 1$ arranged from left to right. We consider their $(d-1)$ adjacency matrices and arrange them as diagonal blocks and add the root as the last element.

Denote by H the adjacency matrix thus obtained.

LEMMA 7. *For any complex number z such that $\Im(z) > 0$, and recall the nth order Chebyshev polynomial, $U_n(z)$. Then the elements of the resolvent of the adjacency matrix H, $(1/\sqrt{d-1}H - z)^{-1}$, have the following properties:*

(i)
$$\left(\frac{1}{\sqrt{d-1}} H - z \right)^{-1}_{\text{root,root}} = \frac{-1}{z+} \frac{-1}{z+} \cdots \frac{-1}{z},$$

where the previous refers to a continued fraction of depth ζ (i.e., $\zeta + 1$ recursions);

(ii) *the above can also be represented as*

$$(18) \qquad \varphi(\zeta) = \left(\frac{1}{\sqrt{d-1}} H - z \right)^{-1}_{\text{root,root}} = -\frac{U_\zeta(z/2)}{U_{\zeta+1}(z/2)};$$

(iii) *furthermore,*

$$(19) \qquad \psi(\zeta) = \left(\frac{1}{\sqrt{d-1}} H - z \right)^{-1}_{\text{root,leaf}} = -\frac{(d-1)^{-\zeta/2}}{U_{\zeta+1}(z/2)},$$

where leaf *represents any leaf of $\mathbb{T}$.*

PROOF. (i) Note that when $\zeta = 0$ (i.e., the tree has only the root vertex), the equality is trivially true. We proceed by induction. Suppose the equality is true until depth $\zeta - 1$. Consider a tree of depth ζ and label the adjacency H matrix as above. Thus

(20)
$$\frac{1}{\sqrt{d-1}}H - z$$
$$= \begin{bmatrix} \frac{1}{\sqrt{d-1}}H_1 - z & & & & \\ & \frac{1}{\sqrt{d-1}}H_2 - z & & & u \\ & & \cdots & & \\ & & & \frac{1}{\sqrt{d-1}}H_{d-1} - z & \\ & & u' & & -z \end{bmatrix}.$$

Here u is the column vector representing the children of the root. Notice that u is $(d-1)^{-1/2}$ exactly at the $(d-1)$ coordinates which are the last elements in each of the block matrices $H_1, \ldots, H_{d-1}$ and zero elsewhere. The vector u' is the transpose of u.

We now use formula (16) treating the the final element $[-z]$ as one block. Thus if $\varphi(\zeta)$ denote the element on the left-hand side of (i) above, we get $\varphi(\zeta) = F^{-1}$, where

(21)
$$F = -z - \frac{1}{d-1} \sum_{i \sim \text{root}} \varphi(\zeta - 1) = -z - \varphi(\zeta - 1).$$

Here "$i \sim$ root" refers to the children of the root which are, in their turn, the roots of trees of depth $\zeta - 1$. The formula now follows by induction.

(ii) We will use the three term recurrence formula for continued fractions which we state below. More details can be found in the excellent book by Lorentzen and Waadeland [40], pages 5 and 6. Given sequences of complex numbers $\{a_n\}$ and $\{b_n\}$ and a complex argument ω, one can define a continued fraction function with argument ω by defining

$$S_n(\omega) = b_0 + \frac{a_1}{b_1+} \frac{a_2}{b_2+} \cdots \frac{a_n}{b_n + \omega}.$$

By Lemma 1.1 in [40] we get the existence of complex sequences $\{A_n\}$ and $\{B_n\}$ such that

$$S_n(\omega) = \frac{A_{n-1}\omega + A_n}{B_{n-1}\omega + B_n} \qquad \text{for } n = 1, 2, \ldots,$$

where

(22)
$$A_n = b_n A_{n-1} + a_n A_{n-2}, \qquad B_n = b_n B_{n-1} + a_n B_{n-2}$$

with initial values $A_{-1} = 1$, $A_0 = b_0$, $B_{-1} = 0$ and $B_0 = 1$.

In our case we will take each $a_i = -1$ and each $b_i = z$, except $b_0 = 0$. The recursions in (22) give us

$$A_n = zA_{n-1} - A_{n-2}, \qquad B_n = zB_{n-1} - B_{n-2}$$

with the initial values $A_{-1} = 1$, $A_0 = 0$, $B_{-1} = 0$ and $B_0 = 1$.

Comparing with the recursions of the Chebyshev polynomials U_n given in (14) we get that

$$B_n(z) = U_n(z/2), \qquad A_n(z) = -U_{n-1}(z/2).$$

Since clearly $\varphi(\zeta) = S_{\zeta+1}(0)$ we get formula (18). This proves part (ii).

(iii) Since there is an obvious isomorphism of the tree that can exchange the labeling of leaves, it is enough to consider the leaf labeled 1 in the adjacency matrix. We will express

$$\psi(\zeta) = \left(\frac{1}{\sqrt{d-1}} H - z \right)^{-1}_{1,\text{root}}$$

in terms of $\psi(\zeta - 1)$.

Let N be the total number of vertices in the tree, and let $\mathcal{A}$ be the diagonal block matrix which is the upper left $(N-1) \times (N-1)$ block in (20). Then from the formula of inverses of block matrices we get

$$\psi(\zeta) = -F^{-1}(\mathcal{A}^{-1}u)_1,$$

where F is defined in (21). But F^{-1} is $\varphi(\zeta)$, and simplifying the elements of $w = \mathcal{A}^{-1}u$, we see that $w(1) = (d-1)^{-1/2}\psi(\zeta - 1)$. In other words, $\psi(\zeta) = -(d-1)^{-1/2}\varphi(\zeta)\psi(\zeta - 1)$. We get by induction

$$\psi(\zeta) = \frac{(-1)^\zeta}{(d-1)^{\zeta/2}} \prod_{i=0}^{\zeta} \varphi(i), \qquad \zeta = 1, 2, \ldots.$$

Now we substitute formula (18) to obtain (19).

$$\psi(\zeta) = \frac{(-1)^\zeta}{(d-1)^{\zeta/2}} \prod_{i=0}^{\zeta} -\frac{U_i(z/2)}{U_{i+1}(z/2)} = -\frac{1}{(d-1)^{\zeta/2}} \frac{U_0(z/2)}{U_{\zeta+1}(z/2)}$$

$$= -\frac{(d-1)^{-\zeta/2}}{U_{\zeta+1}(z/2)}. \qquad \qquad \square$$

Having now calculated the quantities φ and ψ for this "slightly irregular" tree $\mathbb{T}$, let us use them to find the corresponding quantities for the regular one, where the root is adjacent (just like all of the other nonleaf nodes) to precisely d edges (and thus has d children). We consider the same kind of labeling as before.

Lemma 8 below is a variation of Lemma 7 above.

LEMMA 8. *Let $\mathbb{T}_d$ denote a d-regular tree of depth $\zeta \in \mathbb{N}$ such that every vertex has degree d. Let H_d denote the adjacency matrix of the graph. The entries of its resolvent $(1/\sqrt{d-1}H_d - z)^{-1}$ have the following properties*:

(i)

$$\varphi_d(\zeta) := \left(\frac{1}{\sqrt{d-1}}H_d - z\right)^{-1}_{\text{root,root}}$$

(23)

$$= -\frac{U_\zeta(z/2)}{U_{\zeta+1}(z/2) - (d-1)^{-1}U_{\zeta-1}(z/2)};$$

(ii) *the above can also be represented as*

$$\psi_d(\zeta) := \left(\frac{1}{\sqrt{d-1}}H_d - z\right)^{-1}_{\text{root,leaf}}$$

(24)

$$= -\frac{(d-1)^{-\zeta/2}}{U_{\zeta+1}(z/2) - (d-1)^{-1}U_{\zeta-1}(z/2)},$$

where leaf *represents any leaf of $\mathbb{T}_d$*.

PROOF. The proof is identical to that of the last lemma, except that we need to be careful in the first step of the recursion. Since the labeling of the vertices has the same principle as before, we have

(25)
$$\left(\frac{1}{\sqrt{d-1}}H_d - z\right)^{-1}_{\text{root,root}} = \frac{1}{-z - (d/(d-1))\varphi(\zeta-1)},$$

where $\varphi(\cdot)$ has been defined in (21). This reflects the fact that the only change from before is in the number of children on the root (used to be $d-1$, now is d).

Substituting the value of φ from (18) we get

$$\left(\frac{1}{\sqrt{d-1}}H_d - z\right)^{-1}_{\text{root,root}} = \left(-z + \frac{d}{d-1}\frac{U_{\zeta-1}(z/2)}{U_\zeta(z/2)}\right)^{-1}$$

$$= \frac{U_\zeta(z/2)}{-zU_\zeta(z/2) + U_{\zeta-1}(z/2) + (d-1)^{-1}U_{\zeta-1}(z/2)}$$

$$= -\frac{U_\zeta(z/2)}{U_{\zeta+1}(z/2) - (d-1)^{-1}U_{\zeta-1}(z/2)}.$$

The final step above follows from the recursion of the Chebyshev polynomials given in (14). This proves (i). The proof of (ii) follows by a similar argument. □

Recall that we are ultimately interested in how close the Stieltjes transform of the d-regular graph is to the Stieltjes transform of the semicircle, $s(z)$. Toward this goal, we will need some estimates for the functions defined in Lemma 8.

LEMMA 9. *Consider the functions defined in Lemmas 7 and 8. Then for all z such that $z/2 \in E_r = \{y \in \mathbb{C} : |y + \sqrt{y^2 - 1}| = r\}$, for some $r > 1$ such that $r^{-\varsigma} < 1/2$, one has the following estimates*:

(i) *Consider $\varphi_d(\zeta)$ and $s(z) = -(z - \sqrt{z^2 - 4})/2$. We have the following estimate*:

$$(26) \qquad |\varphi_d(\zeta) - s(z)| \leq C_0 \left[\frac{2r^{-2\varsigma}}{1 - r^{-2\varsigma - 2}} + \frac{1}{d - 1} \right],$$

where C_0 is a constant.

(ii) *The following bound on $\psi(\cdot)$ holds*:

$$(27) \qquad |\psi(\zeta)| \leq \frac{r^{-\varsigma - 1}}{(d - 1)^{\varsigma/2}} \frac{2}{1 - r^{-2\varsigma - 4}}.$$

(iii) *Similarly,*

$$(28) \qquad |\psi_d(\zeta)| \leq C_0 \frac{r^{-\varsigma - 1}}{(d - 1)^{\varsigma/2}} \frac{1}{1 - r^{-2\varsigma - 4}}.$$

REMARK 6. The condition $r^{-\varsigma} < 1/2$ is a priori more restrictive than necessary for the purposes of Lemma 9; since we will in fact be interested in the case when $r^{-\varsigma} = o(1)$, this does not matter.

PROOF OF LEMMA 9. (i) Let $\omega = (z + \sqrt{z^2 - 4})/2 = -1/s(z)$. Note that $|\omega| = r > 1$. Recall from (23) and (13) that

$$\varphi_d(\zeta) = \frac{\omega^{\varsigma + 1} - \omega^{-\varsigma - 1}}{\omega^{\varsigma + 2} - \omega^{-\varsigma - 2} - (1/(d - 1))(\omega^{\varsigma} - \omega^{-\varsigma})}$$

$$= \frac{1}{\omega} \frac{1 - \omega^{-\varsigma - 2}}{1 - \omega^{-2\varsigma - 4} - (1/(d - 1))(\omega^{-2} - \omega^{-2\varsigma - 2})}.$$

Thus

$$|\varphi_d(\zeta)| \leq r^{-1} \frac{1 + r^{-\varsigma - 2}}{1 - r^{-2\varsigma - 4} - (d - 1)^{-1}(r^{-2} + r^{-2\varsigma - 2})}$$

as long as the right-hand side above is positive. Since $r > 1$ we get $\varphi_d(\zeta)$ is bounded by an absolute constant C_0 when $r^{-\varsigma} < 1/2$.

We now estimate the quantity $\varphi(\zeta)$. By (18) and (13) we get

$$(29) \qquad \varphi(\zeta) = -\frac{\omega^{\varsigma + 1} - \omega^{-\varsigma - 1}}{\omega^{\varsigma + 2} - \omega^{-\varsigma - 2}} = -\omega^{-1} \frac{1 - \omega^{-2\varsigma - 2}}{1 - \omega^{-2\varsigma - 4}}.$$

Thus

$$
|\varphi(\zeta) + \omega^{-1}| = |\varphi(\zeta) - s(z)| = \left| \omega^{-1} \left(1 - \frac{1 - \omega^{-2\zeta - 2}}{1 - \omega^{-2\zeta - 4}} \right) \right|
$$

$$
= r^{-1} \left| \frac{\omega^{-2\zeta - 2} - \omega^{-2\zeta - 4}}{1 - \omega^{-2\zeta - 4}} \right|
$$

$$
= r^{-2\zeta - 3} \left| \frac{1 - \omega^{-2}}{1 - \omega^{-2\zeta - 4}} \right| \le \frac{r^{-2\zeta - 3}}{1 - r^{-2\zeta - 4}}.
$$

To get to φ_d, consider formula (25). Note that

$$
-\frac{1}{z - s(z)} = -\frac{1}{z/2 + \sqrt{z^2 - 4}/2} = -\omega^{-1} = s(z).
$$

Thus

$$
\varphi_d(\zeta) - s(z) = -\frac{1}{z + (d/(d-1))\varphi(\zeta - 1)} + \frac{1}{z - s(z)}.
$$

Hence

$$
|\varphi_d(\zeta) - s(z)| = \left| s(z) - \frac{d}{d-1}\varphi(\zeta - 1) \right| \left\| \frac{1}{z - s(z)} \right\| \left\| \frac{1}{z + (d/(d-1))\varphi(\zeta - 1)} \right|
$$

$$
= \left| \frac{d}{d-1}(s(z) - \varphi(\zeta - 1)) - \frac{s(z)}{d-1} \right| |\omega^{-1}| |\varphi_d(\zeta)|
$$

$$
\le C_0 \left[\left(\frac{d}{d-1} \right) \frac{r^{-2\zeta - 1}}{1 - r^{-2\zeta - 2}} + \frac{r^{-1}}{d-1} \right] r^{-1}.
$$

Since $r > 1$ this completes the proof.

(ii) We use the estimate on the Chebyshev polynomials given in (15) and our assumption on z to get

$$
|\psi(\zeta)| = \frac{1}{(d-1)^{\zeta/2}} \left| \frac{1}{U_{\zeta+1}(z/2)} \right| \le \frac{1}{(d-1)^{\zeta/2}} \frac{r + r^{-1}}{r^{\zeta+2} - r^{-\zeta - 2}}
$$

$$
= \frac{r^{-\zeta - 1}}{(d-1)^{\zeta/2}} \frac{1 + r^{-2}}{1 - r^{-2\zeta - 4}}.
$$

(iii) This part is similar, since

$$
|\psi_d(\zeta)| = \frac{1}{(d-1)^{\zeta/2}} \left| \frac{1}{U_{\zeta+1}(z/2)} \right| \left| \frac{1}{1 - (d-1)^{-1} U_{\zeta-1}(z/2)/U_{\zeta+1}(z/2)} \right|
$$

$$
\le \frac{r^{-\zeta - 1}}{(d-1)^{\zeta/2}} \frac{1 + r^{-2}}{1 - r^{-2\zeta - 4}} \left\{ \frac{1}{1 - (d-1)^{-1} |U_{\zeta-1}(z/2)/U_{\zeta+1}(z/2)|} \right\}.
$$

Note that, by way of its definition, the constant C_0 in part (i) is an upper bound on the final term, hence the estimate. $\square$

3.3. *From trees to regular graphs.* Consider now a (deterministic) d-regular graph G with a distinguished vertex called the root such that, for some $\zeta \geq 1$, the $(\zeta + 1)$-neighborhood of the root is a tree. That is to say, consider the subgraph consisting of all vertices in G whose distance from the root is at most $\zeta + 1$ and the edges between them; we assume that this subgraph has no cycles.

This gives us a natural partition of the graph. We denote the tree subgraph induced by the root and all vertices of distance at most ζ from the root by $\mathbb{T}_d$. We denote the "boundary" of this graph, that is, the set of vertices that are at distance exactly ζ from the root, by $\partial \mathbb{T}_d$. The subgraph induced by the vertices in the complement of $\mathbb{T}_d$ will be denoted by $\mathbb{T}_d^c$, and its own boundary, that is, the set of vertices at distance exactly $\zeta + 1$ from the root, will be denote by $\partial \mathbb{T}_d^c$. For further clarification, please refer to Figure 1.

Note that all edges between $\mathbb{T}_d$ and $\mathbb{T}_d^c$ are between $\partial \mathbb{T}_d$ and $\partial \mathbb{T}_d^c$.

Additionally we will denote the set of vertices of $\mathbb{T}_d$ (resp., $\mathbb{T}_d^c$) by $V(\mathbb{T}_d)$ [resp., $V(\mathbb{T}_d^c)$].

Let A denote the adjacency matrix of this graph, and let H_d denote the adjacency matrix of the subgraph $\mathbb{T}_d$. Label the vertices of H_d as in Lemmas 7 and 8, and write A in the block matrix form

$$A = \begin{bmatrix} D & B \\ B' & H_d \end{bmatrix}.$$

Here D is the adjacency matrix of $\mathbb{T}_d^c$, the matrix B records only and all the edges between $\partial \mathbb{T}_d$ and $\partial \mathbb{T}_d^c$ and we again use the notation B' for the transpose of B.

We will now proceed to estimate how close the Stieltjes transform of the regular graph is to that of the tree.
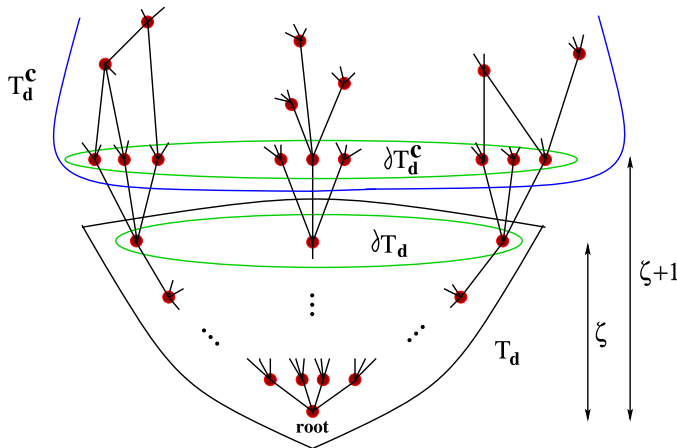

FIG. 1. *Illustration of* $\mathbb{T}_d$, $\mathbb{T}_d^c$, $\partial \mathbb{T}_d$ *and* $\partial \mathbb{T}_d^c$ *starting at a vertex* root, *with maximal tree depth* ζ. *Note that the neighbors in* $\mathbb{T}_d^c$ *of the vertices from* $\partial \mathbb{T}_d^c$ *may, but need not continue the tree-like structure. For this particular tree,* $d = 4$.

LEMMA 10. *Fix a complex number z such that $z \in E_r$ [see (12)] for some $r > 1$. Let ε denote the quantity*

$$\varepsilon = \left(\frac{1}{\sqrt{d-1}}A - z\right)^{-1}_{\text{root,root}} - \left(\frac{1}{\sqrt{d-1}}H_d - z\right)^{-1}_{\text{root,root}}.$$

We have the following bound:

$$|\varepsilon| \leq \left(\frac{2C_0^2}{1 - r^{-2\zeta - 4}}\right)\frac{r^{-2\zeta - 2}}{\Im(z)}.$$

PROOF. Define the vector

$$v = B\left(\frac{1}{\sqrt{d-1}}H_d - z\right)^{-1} e_{\text{root}},$$

where e_{root} is the vector that puts mass at the root vertex and zero elsewhere.

Then by using the formula for the inverse of block matrices (17) we get

$$(30) \qquad \varepsilon = \frac{1}{d-1}\left\{v'\left(\frac{1}{\sqrt{d-1}}D - z - \frac{1}{d-1}B\left(\frac{1}{\sqrt{d-1}}H_d - z\right)^{-1}B'\right)^{-1}v\right\}.$$

Our first job is to estimate the elements of the vector v. From the definition it is clear that the rows of B (and columns of B') are labeled by the vertices of $\mathbb{T}_d^c$. We write B_{i*} to designate the ith row of B.

We obtain

$$v_i = B_{i*}\left(\frac{1}{\sqrt{d-1}}H_d - z\right)^{-1} e_{\text{root}} = \sum_{k \in V(\mathbb{T}_d)} B_{ik}\left(\frac{1}{\sqrt{d-1}}H_d - z\right)^{-1}_{k,\text{root}}$$

(31)

$$= \sum_{k \in V(\mathbb{T}_d^c)} B_{ik}\left(\frac{1}{\sqrt{d-1}}H_d - z\right)^{-1}_{\text{root},k} \qquad \text{by symmetry of } H_d.$$

Note that B_{ik} is positive (i.e., 1) if and only if $i \in \partial\mathbb{T}_d^c$, $k \in \partial\mathbb{T}_d$ and i and k have an edge between them. Thus:

  (i) $v_i = 0$ unless $i \in \partial\mathbb{T}_d^c$.
  (ii) When $i \in \partial\mathbb{T}_d^c$,

$$(32) \qquad v_i = \sum_{k \in \partial\mathbb{T}_d, k \sim i} \left(\frac{1}{\sqrt{d-1}}H_d - z\right)^{-1}_{\text{root},k} = \psi_d(\zeta).$$

Here $\psi_d(\zeta)$ has been defined in (24). The fact that there is exactly one $k \in \partial\mathbb{T}_d$ such that $i \sim k$ follows from our assumption that the $(\zeta + 1)$ neighborhood of the root is a tree (see Figure 1).

(iii) By counting the number of elements in $\partial \mathbb{T}_d^c$, we get

$$(33) \qquad \|v\|^2 = d(d-1)^\zeta |\psi_d(\zeta)|^2 \leq C_0^2 \frac{dr^{-2\zeta-2}}{1 - r^{-2\zeta-4}},$$

where the final estimate is from (28).

Note that the matrix

$$\left(\frac{1}{\sqrt{d-1}} D - z - \frac{1}{d-1} B \left(\frac{1}{\sqrt{d-1}} H_d - z \right)^{-1} B' \right)$$

is precisely the matrix $\mathbb{G}$ appearing in (17). In particular, it is the top left block of the matrix $[(d-1)^{-1/2} A - z]^{-1}$. Hence, by padding the vector v with extra zeros, we get a vector $\bar{v}$ such that

$$\varepsilon = \frac{1}{d-1} \bar{v}' \left(\frac{1}{\sqrt{d-1}} A - z \right)^{-1} \bar{v}.$$

Since the matrix A is real symmetric, it has only real eigenvalues. It follows from spectral decomposition of the real, symmetric matrix A, that for any real vector y

$$(34) \qquad y' \left(\frac{1}{\sqrt{d-1}} A - z \right)^{-1} y \leq \frac{\|y\|^2}{\Im(z)}.$$

Observe from (32) that $v = \psi_d(\zeta) e$, where e is a real vector of ones and zeroes which is one precisely for the labels corresponding to $\partial \mathbb{T}_d$.

Combining our previous observations, we get

$$v' \left(z - \frac{1}{\sqrt{d-1}} D - \frac{1}{d-1} \chi \right)^{-1} v = \psi_d^2(\zeta) e' \left(\frac{1}{\sqrt{d-1}} A - z \right)^{-1} e.$$

Now using the bound from (33) in (30) we get

$$|\varepsilon| \leq \frac{|\psi_d^2(\zeta)|}{d-1} \frac{\|e\|^2}{\Im(z)} = \frac{1}{d-1} \frac{\|v\|^2}{\Im(z)} \leq C_0^2 \frac{d}{d-1} \left(\frac{r^{-2\zeta-2}}{1 - r^{-2\zeta-4}} \right) \frac{1}{\Im(z)}$$

$$\leq \left(\frac{2C_0^2}{1 - r^{-2\zeta-4}} \right) \frac{r^{-2\zeta-2}}{\Im(z)}.$$

This completes the proof of the lemma. $\quad \square$

We now arrive at our main result about *deterministic* regular graphs. Consider the set-up as in Lemma 10. Now a consider a sequence of graphs G_n such that G_n is d_n-regular. Each G_n has a marked vertex called the root such that for some sequence $\{\zeta_n\}$, the $\zeta_n + 1$ neighborhood of the root is acyclic in G_n.

LEMMA 11. *Assume that that sequences $\{d_n\}$ and $\{\zeta_n\}$ both tend to infinity with n with the following restriction. There exists a sequence r_n such that*

$$(35) \qquad r_n = e^{d_n^{-\alpha}} \qquad \text{for some } 0 < \alpha < 1 \quad \text{and} \quad r_n^{-\zeta_n} = o(1/d_n).$$

Let U_n denote all complex numbers z such that $\Im(z) > (r_n - r_n^1)/2$. If A_n denotes the adjacency matrix of the graph G_n, for all d_n such that $d_n \geq (\log 2)^{1/\alpha}$, then for all z, eventually in every U_n, we have

$$(36) \qquad \left| \left(\frac{1}{\sqrt{d_n - 1}} A_n - z \right)^{-1}_{\text{root,root}} - s(z) \right| = O(1/d_n),$$

where $s(z)$ is the Stieltjes transform of the semicircle law $s(z) = -z/2 + \sqrt{z^2 - 4}/2$.

PROOF. Under our assumptions certain simplifications are immediate. The constant C_0 appearing in Lemma 9 (and later) can be taken to be an absolute constant. Since $r_n^{-\zeta_n} = o(1/d_n)$, we can choose C_0 large enough in (26) such that

$$(37) \qquad |\varphi_{d_n}(\zeta_n) - s(z_n)| \leq C_0/d_n.$$

Here z_n is any sequence of complex numbers such that $z_n/2$ belongs to the ellipse E_{r_n} defined in (12).

Note that

$$\varphi_{d_n}(\zeta_n) = \left(\frac{1}{\sqrt{d_n - 1}} A_n - z_n \right)^{-1}_{\text{root,root}}.$$

We now use Lemma 10.

Consider now

$$\Im(z) \geq \frac{r_n - r_n^{-1}}{2} = \frac{1}{2}[e^{d_n^{-\alpha}} - e^{-d_n^{-\alpha}}].$$

One can easily verify the inequality

$$(38) \qquad \frac{e^x - e^{-x}}{2} \geq x/2 \qquad \text{for all } x \leq \log 2.$$

Applying this inequality above, we get $\Im(z) \geq d_n^{-\alpha}/2$ for all d_n as stated in the lemma.

Combining with Lemma 10 we get

$$|\varepsilon| \leq 2C_0^2 \frac{r_n^{-2\zeta_n}}{1 - r_n^{-2\zeta_n - 4}} d_n^\alpha \leq C_1 \frac{1}{d_n^2} d_n^\alpha = o(1/d_n).$$

Combining with (37) this completes the proof of the theorem. $\square$

We will need the following lemma to show that the previous "tree approximation" result about *deterministic* regular graphs can be applied to the random graph by choosing *almost any* vertex as the "root."

LEMMA 12.   *Let $d_n = (\log n)^\gamma$, for some positive γ. Let $\{r_n\}$ be as in Lemma* 11. *Let η_n be defined as $(r_n - r_n^{-1})/2$. For some $\beta > 1$, we define the sequence $\{\zeta_n, n = 1, 2, \ldots\}$ satisfying*

$$\zeta_n = \frac{1}{4} \frac{\log n}{\log(d_n - 1)} - \beta. \tag{39}$$

Let $J(n)$ be the set of vertices in G_n whose ζ_n-neighborhoods are acyclic and let $|J(n)|$ denote its size. Let $\Omega(n)$ be the event

$$|J(n)|/n > 1 - \eta_n/d_n. \tag{40}$$

One can choose β no larger than 4 such that

$$P(\Omega^c(n)) \le o\left(\frac{1}{n}\right). \tag{41}$$

PROOF.   The proof of this result is very similar to the proof of Lemma 4. We again use the estimates of McKay, Wormald and Wysocka [45] on the Poisson approximation to the number of short cycles. Consider a sequence $\{\zeta_n\}$ as in the statement. It is clear from the choice that

$$4\zeta_n \log(d_n - 1) = \log n - \beta \log(d_n - 1) \tag{42}$$

or

$$(d_n - 1)^{4\zeta_n} = \frac{n}{(d_n - 1)^\beta} = o(n). \tag{43}$$

Thus, we can take $g = 2\zeta_n$ in (6).

The argument is essentially the same as the one used in the proof of Lemma 4; rather than repeating it, we choose to only highlight the differences.

The total number of vertices whose ζ_n-neighborhoods are not acyclic can be bounded above by

$$N^*_{\zeta_n} = \sum_{s=3}^{2\zeta_n} s(d_n - 1)^{(2\zeta_n - s)/2} M_s,$$

where M_s is the number of cycles of length s.

Taking expectations and variances above we get

$$E N^*_{\zeta_n} = O\big((d_n - 1)^{2\zeta_n}\big)$$

and

$$\text{Var}(N^*_{\zeta_n}) \le \zeta_n^2 (2\zeta_n + 1)(d_n - 1)^{2\zeta_n} + \zeta_n^2 (2\zeta_n + d_n) o(1).$$

We now want to use Chebyshev's inequality on the quantity $1 - |J(n)|/n$.

$$P\left(1 - \frac{|J(n)|}{n} > \frac{\eta_n}{d_n}\right) \le P(N^*_{\zeta_n} > n\eta_n/d_n) \le \frac{d_n^2}{n^2\eta_n^2}E(N^*_{\zeta_n})^2$$

$$\le \frac{3d_n^2\zeta_n^3(d_n-1)^{2\zeta_n}}{n^2\eta_n^2} + \frac{3\zeta_n^3 d_n^2 o(1)}{n^2\eta_n^2}$$

$$+ \frac{d_n^2}{n\eta_n^2}O\left(\frac{(d_n-1)^{4\zeta_n}}{n}\right).$$

Let us analyze the three terms that appear above. We use inequality (38) to obtain $\eta_n \ge d_n^{-\alpha}/2$. Using (42) we get

$$\frac{3d_n^2\zeta_n^3(d_n-1)^{2\zeta_n}}{n^2\eta_n^2} \le 3d_n^2\zeta_n^3\frac{\sqrt{n}}{(d_n-1)^{\beta/2}}\frac{1}{n^2}4d_n^{2\alpha} = \frac{\text{poly-log }n}{n^{3/2}},$$

$$\frac{3\zeta_n^3 d_n^2 o(1)}{n^2\eta_n^2} \le \frac{12\zeta_n^3 d_n^{2+2\alpha}o(1)}{n^2} = \frac{\text{poly-log }n}{n^2}o(1).$$

The leading term is the last term on the right which is of the order $d_n^{2-\beta}/(n\eta_n^2)$, because by (42) we get $(d_n-1)^{4\zeta_n}/n = (d_n-1)^{-\beta}$.

We now choose $\beta > 2\alpha + 2$ (and thus no larger than 4, since $\alpha < 1$) such that

$$\frac{d_n^{2-\beta}}{\eta_n^2} \le 4d_n^{2-\beta}d_n^{2\alpha} = o(1).$$

This completes the proof of the lemma. $\square$

3.4. *The Stieltjes transform.* We bring now all the results of the previous sections together. Below is the first theorem of this section.

THEOREM 13. *Consider a sequence of d_n-regular graphs on n vertices where $d_n = (\log n)^\gamma$ for some $\gamma > 0$. Let A_n be the adjacency matrix of the graph, and consider the Stieltjes transform*

$$s_n(z) = \text{tr}\left(\frac{1}{\sqrt{d_n-1}}A_n - z\right)^{-1}, \qquad \Im(z) > 0.$$

Let $\eta_n = (r_n - r_n^{-1})/2$ where $r_n = \exp(d_n^{-\alpha})$ for some $0 < \alpha < \min(1, 1/\gamma)$.

Let U_n denote all complex numbers z such that $\Im(z) > \eta_n$. Then there is a large enough constant $C > 0$ such that the Stieltjes transform of the empirical eigenvalue distribution of the $n \times n$ matrix A_n satisfies

$$P\left(\sup_{z\in U_n}|s_n(z) - s(z)| > C/d_n\right) \le o(1/n).$$

Here $s(z)$ refers to the Stieltjes transform of the semicircle law.

PROOF.   We condition on the event $\Omega(n)$ which, from Lemma 12, happens with a probability of at least $1 - o(1/n)$.

Consider the set $J(n)$ from the Lemma 12 and write

$$(44) \quad s_n(z) = \frac{1}{n} \sum_{k \in J(n)} \left(\frac{1}{\sqrt{d_n - 1}} A_n - z \right)^{-1}_{k,k} + \frac{1}{n} \sum_{k \notin J(n)} \left(\frac{1}{\sqrt{d_n - 1}} A_n - z \right)^{-1}_{k,k}.$$

Now, the spectral norm of $(\frac{1}{\sqrt{d_n-1}} A_n - z)^{-1}$ is bounded above by $1/\Im(z)$, which in turn is bounded above by η_n^{-1} for all $z \in U_n$. Thus, for all $z \in U_n$ and all k, we have the obvious bound $(\frac{1}{\sqrt{d_n-1}} A_n - z)^{-1}_{k,k} \leq \eta_n^{-1}$.

Summing up over all $k \notin J(n)$, we get

$$\frac{1}{n} \sum_{k \notin J(n)} \left| \left(\frac{1}{\sqrt{d_n - 1}} A_n - z \right)^{-1}_{k,k} \right| \leq \frac{n - |J(n)|}{n \eta_n} \leq \frac{1}{d_n},$$

where the final inequality holds on the event Ω. Thus, we get

$$s_n(z) = \frac{1}{n} \sum_{k \in J(n)} \left(\frac{1}{\sqrt{d_n - 1}} A_n - z \right)^{-1}_{k,k} + O\left(\frac{1}{d_n} \right).$$

Now, on the event $\Omega(n)$, every $k \in J(n)$ can be considered as the root for a d_n-regular tree of depth ζ_n. We can now apply Lemmas 10 and 11. Note that, technically, the ζ_n in Lemma 12 and the ζ used in Lemma 10 differ by at most one; however, this does not affect the following calculations.

We verify the following assumptions from Lemma 11:

$$\begin{aligned}
r_n^{-\zeta_n} &= \exp\left(-\frac{1}{4} d_n^{-\alpha} \frac{\log n}{\log(d_n - 1)} + \beta d_n^{-\alpha} \right) \\
&= \exp\left(-\frac{1}{4} \frac{(\log n)^{1-\alpha\gamma}}{\log(d_n - 1)} + \frac{\beta}{(\log n)^\alpha} \right) \\
&\leq C \exp\left(-\frac{(\log n)^{1-\alpha\gamma}}{4\gamma \log\log n} \right) \ll \exp(-\gamma \log\log n) = \frac{1}{d_n},
\end{aligned}$$

whenever $\alpha\gamma < 1$.

We now combine our error estimate (36) to see that for any $k \in J(n)$, on the event $\Omega(n)$, and for all $z \in U_n$, we have

$$(45) \qquad \left| \left(\frac{1}{\sqrt{d_n - 1}} A_n - z \right)^{-1} (k, k) - s(z) \right| = O\left(\frac{1}{d_n} \right),$$

where the constants in the $O(\cdot)$ above does not depend on k.

Combining this with decomposition (44) we get

$$(46) \qquad \sup_{z \in U_n} |s_n(z) - s(z)| = O\left(\frac{1}{d_n} \right)$$

on the event $\Omega(n)$, which holds with probability at least $1 - o(1/n)$. This completes the proof of the lemma. $\square$

Recall now the setup of Theorem 2. Fix $\delta > 0$. Let d_n, η_n be as in Theorem 13. Then we will show that there exists an N large enough such that for all $n \geq N$, for any interval $I \subset \mathbb{R}$ of length $|I| \geq \max\{2\eta_n, \eta_n/(-\delta \log \delta)\}$,

$$\left| \mathcal{N}_I - n \int_I f_{sc}(x)\, dx \right| < \delta n |I|$$

with probability at least $1 - o(1/n)$.

Here $\mathcal{N}_I$ is the number of eigenvalues of W_n in the interval I, and f_{sc} refers to the density of the semicircle law as in (1).

PROOF OF THEOREM 2. Theorem 13 leads to Theorem 2 whose proof follows almost identically to Lemma 60 in the article by Tao and Vu [56]. The only major difference between our theorem and Lemma 60 of [56] is the fact that our interval I can lie anywhere on the real line and is not restricted to a subset of $(-2, 2)$ as Lemma 60 requires. We provide an outline of the argument but skip the details.

The idea lies in the observation that a good control over the Stieltjes transform near the real line allows one to invert the transform and have an estimate of the empirical spectral density. This is due to the following inversion formula: if G is a continuous distribution on the real line with Stieltjes transform s_G, one gets

$$G[a, b] = \lim_{\varepsilon \to 0} \frac{1}{\pi} \int_a^b \Im(s_G(x + i\varepsilon))\, dx.$$

Fix an interval $I \subseteq [-2, 2]$ such that $|I| \geq 2\eta$. Define the function

$$F(y) = \frac{1}{\pi} \int_I \frac{\eta_n}{\eta_n^2 + (y - x)^2}\, dx.$$

Then it follows that if λ_i denotes the ith eigenvalues of the matrix $\frac{1}{\sqrt{d_n-1}} A_n$, then

$$\frac{1}{n} \sum_{i=1}^n F(\lambda_i) = \frac{1}{\pi} \int_I \frac{1}{n} \sum_{i=1}^n \frac{\eta_n}{\eta_n^2 + (\lambda_i - x)^2}\, dx$$

$$= \frac{1}{\pi} \int_I \Im(s_n(x + i\eta_n))\, dx.$$

Also, if f_{sc} denotes the density of the semicircle law, we get

$$\int_{-2}^2 F(y) f_{sc}(y)\, dy = \frac{1}{\pi} \int_I \Im(s(x + i\eta_n))\, dx.$$

Using the approximation between s_n and s obtained in Theorem 13, we get

$$\left| \frac{1}{n} \sum_{i=1}^{n} F(\lambda_i) - \int_{-2}^{2} F(y) f_{sc}(y)\, dy \right| \leq \frac{1}{\pi} \int_{I} |s_n(x + i\eta_n) - s(x + i\eta_n)|\, dx$$

$$\leq C \frac{|I|}{d_n} \qquad \text{with probability } 1 - o\!\left(\frac{1}{n}\right).$$

Choose n large enough such that $1/d_n \leq \delta$ for all subsequent n. Thus

$$\frac{1}{n} \sum_{i=1}^{n} F(\lambda_i) = \int_{-2}^{2} F(y) f_{sc}(y)\, dy + O(\delta|I|)$$

with probability $1 - o(1/n)$.

Now following the bounds in [56], page 60, proof of Lemma 64, we obtain the bounds

$$\int_{-2}^{2} F(y) f_{sc}(y)\, dy = \int_{I} f_{sc}\, dy + O\!\left(\eta_n \log \frac{|I|}{\eta_n}\right),$$

$$\frac{1}{n} \sum_{i=1}^{n} F(\lambda_i) = \frac{\mathcal{N}}{n} + O\!\left(\eta_n \log \frac{|I|}{\eta_n}\right).$$

Putting all these together we obtain that with probability $1 - o(1/n)$, one has

$$\mathcal{N}_I - n \int_{I} f_{sc}\, dy = O(n\delta|I|) + O\!\left(n\eta_n \log \frac{|I|}{\eta_n}\right).$$

Finally, as observed in [56], the latter term can be absorbed in the former since $|I| \geq \eta_n/\delta \log(1/\delta)$. This completes the proof of the theorem. $\square$

4. Delocalization of eigenvectors. Closely related to approximation of the empirical spectral distribution is the fact that the $\mathbb{L}^2$-norm of a normalized eigenvector restricted to a *large* subset of the vertices cannot be *small*. We follow Definition 1 and prove Theorem 1.

Recall the set-up of Theorem 2. Fix $\delta > 0$.

Let $T_n \subseteq \{1, 2, \ldots, n\}$ be a sequence of sets of size $L_n = o(\eta_n^{-1})$. Let $\Omega_1(n)$ be the event that some unit-norm eigenvector of the matrix A_n is (T_n, δ) localized; that is,

$$\Omega_1(n) = \{\exists i : \|v_i|_{T_n}\|^2 \geq 1 - \delta, \text{ for some } v_i \text{ such that } A_n v_i = \lambda_i v_i\}.$$

Then, for all sufficiently large n, we will show that

$$P((\Omega_1(n))^c) \geq e^{-L_n \eta_n / d_n}\left(1 - o\!\left(\frac{1}{d_n}\right)\right) = 1 - o\!\left(\frac{1}{d_n}\right).$$

PROOF OF THEOREM 3. Consider the set $J(n)$, defined in Lemma 12, of vertices whose ζ_n-neighborhoods are acyclic. Recall the event $\Omega(n)$, whose probability (as soon as n is large enough) is $1 - o(1/n)$, which is that $|J(n)|/n > 1 - \eta_n/d_n$.

We first prove part (i). The event $\Omega_1(n)$ can be decomposed as two disjoint events depending on whether the set T_n is a subset of $J(n)$ or not. We first examine the event

$$\Omega_1'(n) = \{T_n \subseteq J(n)\} \cap \Omega_1(n)$$

for purposes of exclusion.

Assume $\omega \in \Omega(n) \cap \Omega_1'(n)$, and fix i, v_i depending on ω and n. The matrix A_n has n real eigenvalues and corresponding eigenspaces. The top eigenvalue is d_n with corresponding eigenvector $v_1 = n^{-1/2}\mathbf{1}$, which is completely delocalized. Let us now choose $v_2, v_3, \ldots, v_n$ to be a set of normalized eigenvectors, corresponding, respectively, to the eigenvalues $\lambda_2, \lambda_3, \ldots, \lambda_n$ of $\frac{1}{\sqrt{d_n-1}}A_n$. Fix a subset $T_n \subseteq \{1, 2, \ldots, n\}$ of size L_n as stated in the theorem.

Consider again the Stieltjes transform of the matrix $(d_n - 1)^{-1/2}A_n$ as in the proof of Theorem 13. That is, for $z \in \mathbb{C}$, with $\Im(z) > 0$, consider the matrix $(\frac{1}{\sqrt{d_n-1}}A_n - z)^{-1}$. Then by the spectral representation it follows that for any $1 \leq k \leq n$ we get

$$\left(\frac{1}{\sqrt{d_n-1}}A_n - z\right)^{-1}(k,k) = \sum_{j=1}^{n} \frac{v_j^2(k)}{(\lambda_j - z)}.$$

Summing over the vertices $k \in T_n$, we get

(47) $$\sum_{k \in T_n}\left(\frac{1}{\sqrt{d_n-1}}A_n - z\right)^{-1}(k,k) = \sum_{j=1}^{n} \frac{\sum_{k \in T_n} v_j^2(k)}{(\lambda_j - z)}.$$

Taking the imaginary part on both sides, we get

$$\Im\left(\sum_{k \in T_n}\left(\frac{1}{\sqrt{d_n-1}}A_n - z\right)^{-1}(k,k)\right) = \sum_{j=1}^{n} \sum_{k \in T_n} v_j^2(k)\Im\left(\frac{1}{\lambda_j - z}\right)$$

(48) $$\geq \sum_{k \in T_n} v_i^2(k)\Im(\lambda_i - z)^{-1}$$

$$\geq (1 - \delta)\Im(\lambda_i - z)^{-1}.$$

Now we use the fact that for $\omega \in \Omega(n)$ and for all $k \in J(n)$, we get from equation (45) that

$$\sup_{z \in U_n}\left|\left(\frac{1}{\sqrt{d_n-1}}A_n - z\right)^{-1}(k,k) - s(z)\right| \leq \frac{C}{d_n}$$

for some absolute constant $C > 0$. Recall that $s(z)$, the Stieltjes transform of the semicircle density, is given by

$$s(z) = \tfrac{1}{2}(-z + \sqrt{z^2 - 4}).$$

Thus, summing up over all k in T_n we get

$$\sup_{z \in S_n} \left| \sum_{k \in T_n} \left(\frac{1}{\sqrt{d_n - 1}} A_n - z \right)^{-1} (k, k) - L_n s(z) \right| \le \frac{C L_n}{d_n}.$$

Combining this estimate with (48) for z such that $\Re(z) = \lambda_i$ and $\Im(z) = \eta_n$ we get

$$L_n \Im(s(z)) + C L_n / d_n \ge \eta_n^{-1}(1 - \delta)$$

for all $\omega \in \Omega(n) \cap \Omega_1'(n)$.

Since $\Im(s(z))$ is bounded and $L_n = o(\eta_n^{-1})$, there is a large enough N such that for all $n \ge N$, this inequality will not hold when $\omega \in \Omega_1'(n)$. So we get that

$$(49) \qquad\qquad P\big(\Omega(n) \cap \Omega_1'(n)\big) = 0$$

as soon as n is large enough.

We can now write

$$\Omega_1(n) = \big((\Omega(n))^c \cap \Omega_1(n)\big) \cup \big(\Omega(n) \cap \Omega_1(n)\big);$$

the probability of the first of the two events above is bounded by the probability of $(\Omega(n))^c$ which is $o(1/n)$. Further, by (49), we get that

$$\Omega(n) \cap \Omega_1(n) = \Omega(n) \cap \big(\Omega_1(n) \setminus \Omega_1'(n)\big).$$

Note that the last event is equivalent to saying that $\Omega(n)$ and $\Omega_1(n)$ happen, and that in addition $T_n \not\subseteq J(n)$.

We now bound

$$
\begin{aligned}
P\big[\Omega(n) \cap \big(\Omega_1(n) \setminus \Omega_1'(n)\big)\big] &\le P[\Omega(n) \cap \{T_n \not\subseteq J(n)\}] \\
&= P[\{T_n \not\subseteq J(n)\} \mid \Omega(n)] P[\Omega(n)] \\
&= \left(1 - o\left(\frac{1}{n}\right)\right) P[\{T_n \not\subseteq J(n)\} \mid \Omega(n)].
\end{aligned}
$$

Note now that, given the size $j = |J(n)|$, any set of j labels chosen from $\{1, \ldots, n\}$ is just as likely as any other, and independent of the set T_n. So

$$
\begin{aligned}
P[\{T_n \not\subseteq J(n)\} \mid |J(n)| = j] &= 1 - P[\{T_n \subseteq J(n)\} \mid |J(n)| = j] \\
(50) \qquad\qquad &= 1 - \frac{\binom{n - L_n}{j - L_n}}{\binom{n}{j}} \\
&\le 1 - \frac{(j - L_n)^{L_n}}{n^{L_n}}.
\end{aligned}
$$

Since

$$P[\{T_n \not\subseteq J(n)\} \mid \Omega(n)] = P\left[\{T_n \not\subseteq J(n)\} \,\middle|\, \frac{|J(n)|}{n} > 1 - \frac{\eta_n}{d_n}\right],$$

it follows from (50) that

$$P[\{T_n \not\subseteq J(n)\} \mid \Omega(n)] \le 1 - \left(1 - \frac{\eta_n}{d_n} - \frac{L_n}{n}\right)^{L_n},$$

and since $L_n = o(1/\eta_n)$ and thus $L_n/n = o(\eta_n/d_n)$, it follows that

$$(51) \qquad P[\{T_n \not\subseteq J(n)\} \mid \Omega(n)] \le 1 - e^{-L_n \eta_n/d_n}\left(1 - o\left(\frac{1}{d_n}\right)\right).$$

Putting all of these together, we conclude that

$$P[\Omega_1(n)] \le 1 - e^{-L_n \eta_n/d_n}\left(1 - o\left(\frac{1}{d_n}\right)\right),$$

which completes the proof.

Note that part (ii) is proved in (49). □

APPENDIX: EIGENVALUES AND EIGENVECTORS OF REGULAR TREES

Our main step in the above proofs was to understand the Stieltjes transforms or the resolvent matrix of a finite tree where every nonleaf vertex has $(d-1)$ children. It is quite straightforward to compute all the eigenvalues of such a tree. Explicit eigenvalues of the tree give us ideas about spacing distribution of eigenvalues of the random regular graph. Fix a positive integer $d \ge 2$.

LEMMA 14. *Let $\mathbb{T}$ be a finite ordered rooted tree of depth $\zeta \in \mathbb{N}$ such that every vertex has exactly $(d-1)$ children. That is, the root has degree $(d-1)$, and every other vertex, other than the leaves, has degree d. Let H denote the adjacency matrix of this graph.*

(i) *Then, for any complex number z the characteristic polynomial of H is given by*

$$\Delta(z; \zeta) := \det\left(zI - \frac{1}{\sqrt{d-1}}H\right) = U_{\zeta+1}(z/2) \prod_{i=1}^{\zeta} U_{\zeta+1-i}^{(d-1)^i - (d-1)^{i-1}}(z/2).$$

(ii) *The eigenvalues of the adjacency matrix are given by the following collection. Consider $i = 1, 2 \ldots, \zeta$: then twice the zeros of the Chebyshev polynomial $U_{\zeta+1-i}$ appears with multiplicity $(d-1)^i - (d-1)^{i-1}$. For $i = 0$, the multiplicity is one.*

PROOF. Recall the recursive labeling of vertices as given in Section 3.2.

To prove conclusion (i), note that when $\zeta = 0$ (i.e., the tree has only the root vertex), the equality is trivially true. We proceed by induction. Suppose the equality

is true until depth $\zeta - 1$. Consider a tree of depth ζ, and label the adjacency H matrix as above. Thus

$$
(52) \quad z - \frac{1}{\sqrt{d-1}}H
$$
$$
= \begin{bmatrix} z - \dfrac{1}{\sqrt{d-1}}H_1 & & & & \\ & z - \dfrac{1}{\sqrt{d-1}}H_2 & & & -u \\ & & \ddots & & \\ & & & z - \dfrac{1}{\sqrt{d-1}}H_{d-1} & \\ & & -u' & & z \end{bmatrix}.
$$

Here u is the column vector representing the children of the root. Notice that u is $(d-1)^{-1/2}$ exactly at the $(d-1)$ coordinates which are the last elements in each of the block matrices $H_1, \dots, H_{d-1}$ and zero elsewhere. The vector u' is the transpose of u.

We now use the following well-known formula of determinant of block matrices [akin to (16)]:

$$
\det \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \det(A)\det(D - CA^{-1}B).
$$

We apply this to the matrix $z - (d-1)^{-1/2}H$ treating the the final element $[z]$ as one block: note that, by our labeling, in this case A is a diagonal block matrix, and hence its determinant is a product of the determinants of the individual blocks which are all the same and equal to $\Delta(z; \zeta - 1)$. Thus we get

$$
\Delta(z; \zeta) = \big(\Delta(z; \zeta - 1)\big)^{d-1}(z - u'A^{-1}u).
$$

As shown in Section 3.2, the quantity

$$
(z - u'A^{-1}u) = z + \varphi(\zeta - 1) = z - \frac{U_{\zeta-1}(z/2)}{U_\zeta(z/2)}
$$
$$
= \frac{zU_\zeta - U_{\zeta-1}}{U_\zeta} = \frac{U_{\zeta+1}(z/2)}{U_\zeta(z/2)}.
$$

Here U_n is the Chebyshev polynomial of the second kind. Note the reversal of sign from Section 3.2 which is due to current reversal of sign from the resolvent matrix.

Hence

$$
\Delta(z; \zeta) = \frac{U_{\zeta+1}(z/2)}{U_\zeta(z/2)}\big(\Delta(z; \zeta - 1)\big)^{d-1} = \prod_{i=0}^{\zeta}\left[\frac{U_{\zeta+1-i}(z/2)}{U_{\zeta-i}(z/2)}\right]^{(d-1)^i}.
$$

The last term can be verified from the initial conditions of the Chebyshev polynomials.

Simplifying a bit more, we get

$$\Delta(z;\zeta) = U_{\zeta+1}(z/2) \prod_{i=1}^{\zeta} U_{\zeta+1-i}^{(d-1)^i-(d-1)^{i-1}}(z/2).$$

For part (ii), note from above that the i many zeroes of U_i appear with multiplicity $(d-1)^i - (d-1)^{i-1}$. Hence the total number of eigenvalues are

$$\zeta + 1 + \sum_{i=1}^{\zeta}(\zeta+2-i)[(d-1)^i - (d-1)^{i-1}] = \sum_{i=0}^{\zeta}(d-1)^i,$$

which is the total number of vertices of the tree. $\square$

The zeros of Chebyshev polynomials of order k can be easily shown to be given by

$$\cos\left(\frac{j\pi}{k+1}\right), \qquad j = 1, 2, \ldots, k.$$

Thus an interesting phenomenon transpires in this analysis. If we drop the multiplicities and consider the empirical distribution of the distinct eigenvalues, they are the zeros of the Chebyshev polynomials of increasing order. These zeros are cosine transformations of equidistant points on the unit circle; and hence their empirical distribution converges to the arc-sine law. However, the entire empirical spectral distribution converges (see [13]) to the spectral distribution of the infinite tree which is the semicircle law. The effect of the multiplicities is strong enough to flip the "smile" of the arc-sine law to the "frown" of the semicircle! Also note that the gap of the spectrum from 2 is about twice of $\pi^2/(\zeta+1)^2$, and does not depend on d.

Some facts about eigenvectors of this tree are also easy to derive and might be also worthwhile to look at. For example, by the spectral theorem, one can write

$$(53) \qquad -\frac{U_{\zeta}(z/2)}{U_{\zeta+1}(z/2)} = \varphi(z) = \left(\frac{1}{\sqrt{d-1}}H - z\right)^{-1}_{\text{root,root}} = \sum_i \frac{\|P_i e_{\text{root}}\|^2}{\lambda_i - z}.$$

Here the sum on the right goes over distinct eigenvalues of the adjacency matrix and $P_i e_{\text{root}}$ refers to the projection of the vector e_{root} on the eigenspace corresponding to λ_i.

Notice that the above is a meromorphic function of z. From the leftmost expression in (53), it is obvious that the function has poles at (twice) the zeros of $U_{\zeta+1}$. It follows then that $P_i e_{\text{root}}$ is zero for all eigenvalues except when λ_i is twice of a root of $U_{\zeta+1}$. However, these roots are simple, as we show in the previous lemma. Hence, $\|P_i e_{\text{root}}\|^2$ is precise the square of the "root"-coordinate of the ith eigenvector.

Its value can be easily computed. For any root λ_i of $U_{\zeta+1}$, we get

$$\|P_i e_{\text{root}}\|^2 = \lim_{z \to 2\lambda_i}(z - 2\lambda_i)\frac{U_\zeta(z/2)}{U_{\zeta+1}(z/2)} = \frac{2U_\zeta(\lambda_i)}{U'_{\zeta+1}(\lambda_i)}.$$

Here $U'_{\zeta+1}$ refers to the derivative of the polynomial $U_{\zeta+1}$. Other coordinates can be similarly derived.

REFERENCES

[1] ALON, N., KRIVELEVICH, M. and VU, V. H. (2002). On the concentration of eigenvalues of random symmetric matrices. *Israel J. Math.* **131** 259–267. MR1942311

[2] ANDERSON, G. W., GUIONNET, A. and ZEITOUNI, O. (2010). *An Introduction to Random Matrices. Cambridge Studies in Advanced Mathematics* **118**. Cambridge Univ. Press, Cambridge. MR2760897

[3] BAI, Z. D. (1993). Convergence rate of expected spectral distributions of large random matrices. I. Wigner matrices. *Ann. Probab.* **21** 625–648. MR1217559

[4] BAI, Z. D., MIAO, B. Q. and PAN, G. M. (2007). On asymptotics of eigenvectors of large sample covariance matrix. *Ann. Probab.* **35** 1532–1572. MR2330979

[5] BAI, Z. D. and SILVERSTEIN, J. W. (2006). *Spectral Analysis of Large Dimensional Random Matrices. Mathematics Monograph Series* **2**. Science Press, Beijing.

[6] BAI, Z. D. and YAO, J. (2005). On the convergence of the spectral empirical process of Wigner matrices. *Bernoulli* **11** 1059–1092. MR2189081

[7] BAIK, J. and SUIDAN, T. M. (2005). A GUE central limit theorem and universality of directed first and last passage site percolation. *Int. Math. Res. Not. IMRN* **6** 325–337. MR2131383

[8] BAIK, J. and SUIDAN, T. M. (2007). Random matrix central limit theorems for nonintersecting random walks. *Ann. Probab.* **35** 1807–1834. MR2349576

[9] BAKER, T. H. and FORRESTER, P. J. (1997). The Calogero–Sutherland model and generalized classical polynomials. *Comm. Math. Phys.* **188** 175–216. MR1471336

[10] BAUER, M. and GOLINELLI, O. (2001). Random incidence matrices: Moments of the spectral density. *J. Stat. Phys.* **103** 301–337. MR1828732

[11] BEN AROUS, G. and PÉCHÉ, S. (2005). Universality of local eigenvalue statistics for some sample covariance matrices. *Comm. Pure Appl. Math.* **58** 1316–1357. MR2162782

[12] BHAMIDI, S., EVANS, S. and SEN, A. (2008). Spectra of random large trees. Preprint. Available at arXiv:0903.3589v2.

[13] BORDENAVE, C. and LELARGE, M. (2010). Resolvent of large random graphs. *Random Structures Algorithms* **37** 332–352. MR2724665

[14] BRIN, S. and LAWRENCE, P. (1998). The anatomy of a large-scale hypertextual Web search engine. In *Proceedings of the Seventh International Conference on World Wide Web* 7 107–117. Brisbane, Australia.

[15] BRODER, A. and SHAMIR, E. (1987). On the second eigenvalue of random regular graphs. In 28*th Annual Symposium on Foundations of Computer Science* (*Los Angeles*, 1987) 286–294. IEEE Comput. Soc. Press, Washington, DC.

[16] CHATTERJEE, S. (2006). A generalization of the Lindeberg principle. *Ann. Probab.* **34** 2061–2076. MR2294976

[17] DEIFT, P. (2007). Universality for mathematical and physical systems. In *International Congress of Mathematicians. Vol. I* 125–152. Eur. Math. Soc., Zürich. MR2334189

[18] DEIFT, P. and GIOEV, D. (2007). Universality at the edge of the spectrum for unitary, orthogonal, and symplectic ensembles of random matrices. *Comm. Pure Appl. Math.* **60** 867–910. MR2306224

[19] DEKEL, Y., LEE, J. R. and LINIAL, N. (2011). Eigenvectors of random graphs: Nodal domains. *Random Structures Algorithms* **39** 3958.

[20] DUMITRIU, I. and EDELMAN, A. (2006). Global spectrum fluctuations for the β-Hermite and β-Laguerre ensembles via matrix models. *J. Math. Phys.* **47** 063302, 36. MR2239975

[21] ELON, Y. (2008). Eigenvectors of the discrete Laplacian on regular graphs—a statistical approach. *J. Phys. A* **41** 435203, 17. MR2453165

[22] ERDŐS, L., KNOWLES, A., YAU, H. T. and YIN, J. (2011). Spectral statistics of Erdős–Rényi graphs I: Local semicircle law. Preprint. http://arxiv.org/abs/1103.1919.

[23] ERDŐS, L., PÉCHÉ, S., RAMÍREZ, J. A., SCHLEIN, B. and YAU, H.-T. (2010). Bulk universality for Wigner matrices. *Comm. Pure Appl. Math.* **63** 895–925. MR2662426

[24] ERDŐS, L., RAMÍREZ, J., SCHLEIN, B., TAO, T., VU, V. and YAU, H.-T. (2010). Bulk universality for Wigner Hermitian matrices with subexponential decay. *Math. Res. Lett.* **17** 667–674. MR2661171

[25] ERDŐS, L., SCHLEIN, B. and YAU, H.-T. (2009). Semicircle law on short scales and delocalization of eigenvectors for Wigner random matrices. *Ann. Probab.* **37** 815–852. MR2537522

[26] ERDŐS, L., SCHLEIN, B. and YAU, H.-T. (2009). Local semicircle law and complete delocalization for Wigner random matrices. *Comm. Math. Phys.* **287** 641–655. MR2481753

[27] ERDŐS, L., YAU, H. T. and YIN, J. (2010). Universality for generalized Wigner matrices with Bernoulli distribution. Preprint. Available at arXiv:1003.3813v4.

[28] FEIGE, U. and OFEK, E. (2005). Spectral techniques applied to sparse random graphs. *Random Structures Algorithms* **27** 251–275. MR2155709

[29] FRIEDMAN, J. (1991). On the second eigenvalue and random walks in random *d*-regular graphs. *Combinatorica* **11** 331–362. MR1137767

[30] FRIEDMAN, J. (1993). Some geometric aspects of graphs and their eigenfunctions. *Duke Math. J.* **69** 487–525. MR1208809

[31] FRIEDMAN, J. (2008). A proof of Alon's second eigenvalue conjecture and related problems. *Mem. Amer. Math. Soc.* **195** viii+100. MR2437174

[32] FÜREDI, Z. and KOMLÓS, J. (1981). The eigenvalues of random symmetric matrices. *Combinatorica* **1** 233–241. MR0637828

[33] GUIONNET, A. and ZEITOUNI, O. (2000). Concentration of the spectral measure for large matrices. *Electron. Commun. Probab.* **5** 119–136 (electronic). MR1781846

[34] JAKOBSON, D., MILLER, S. D., RIVIN, I. and RUDNICK, Z. (1999). Eigenvalue spacings for regular graphs. In *Emerging Applications of Number Theory* (*Minneapolis*, *MN*, 1996). *IMA Vol. Math. Appl.* **109** 317–327. Springer, New York. MR1691538

[35] JOHANSSON, K. (1998). On fluctuations of eigenvalues of random Hermitian matrices. *Duke Math. J.* **91** 151–204. MR1487983

[36] KESTEN, H. (1959). Symmetric random walks on groups. *Trans. Amer. Math. Soc.* **92** 336–354. MR0109367

[37] KHORUNZHY, A. (2001). Sparse random matrices: Spectral edge and statistics of rooted trees. *Adv. in Appl. Probab.* **33** 124–140. MR1825319

[38] KHORUNZHY, A. M., KHORUZHENKO, B. A. and PASTUR, L. A. (1996). Asymptotic properties of large random matrices with independent entries. *J. Math. Phys.* **37** 5033–5060. MR1411619

[39] KRIVELEVICH, M. and SUDAKOV, B. (2003). The largest eigenvalue of sparse random graphs. *Combin. Probab. Comput.* **12** 61–72. MR1967486

[40] LORENTZEN, L. and WAADELAND, H. (2008). *Continued Fractions. Vol.* 1. *Convergence Theory*, 2nd ed. *Atlantis Studies in Mathematics for Engineering and Science* **1**. Atlantis Press, Paris. MR2433845

[41] LUBOTZKY, A., PHILLIPS, R. and SARNAK, P. (1988). Ramanujan graphs. *Combinatorica* **8** 261–277. MR0963118

[42] MASON, J. C. and HANDSCOMB, D. C. (2003). *Chebyshev Polynomials*. Chapman & Hall/CRC, Boca Raton, FL. MR1937591

[43] MCKAY, B. D. (1981). The expected eigenvalue distribution of a large regular graph. *Linear Algebra Appl.* **40** 203–216. MR0629617

[44] MCKAY, B. D. and WORMALD, N. C. (1997). The degree sequence of a random graph. I. The models. *Random Structures Algorithms* **11** 97–117. MR1610253

[45] MCKAY, B. D., WORMALD, N. C. and WYSOCKA, B. (2004). Short cycles in random regular graphs. *Electron. J. Combin.* **11** 12 pp. (electronic). MR2097332

[46] MEHTA, M. L. (1991). *Random Matrices*, 2nd ed. Academic Press, Boston, MA. MR1083764

[47] MEILA, M. and SHI, J. (2011). Learning segmentation with random walks. In *Adv. Neural Inf. Process. Syst.* **14** 873–879. MIT Press, Cambridge, MA.

[48] MIRLIN, A. D. and FYODOROV, Y. V. (1991). Universality of level correlation function of sparse random matrices. *J. Phys. A* **24** 2273–2286. MR1118532

[49] MORGENSTERN, M. (1994). Existence and explicit constructions of $q + 1$ regular Ramanujan graphs for every prime power q. *J. Combin. Theory Ser. B* **62** 44–62. MR1290630

[50] MUIRHEAD, R. J. (1982). *Aspects of Multivariate Statistical Theory*. Wiley, New York. MR0652932

[51] POTHEN, A., SIMON, H. D. and LIOU, K.-P. (1990). Partitioning sparse matrices with eigenvectors of graphs. *SIAM J. Matrix Anal. Appl.* **11** 430–452. MR1054210

[52] SEMERJIAN, G. and CUGLIANDOLO, L. F. (2002). Sparse random matrices: The eigenvalue spectrum revisited. *J. Phys. A* **35** 4837–4851. MR1916090

[53] SHI, J. and MALIK, J. (2000). Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **22** 888–905.

[54] SOSHNIKOV, A. (1999). Universality at the edge of the spectrum in Wigner random matrices. *Comm. Math. Phys.* **207** 697–733. MR1727234

[55] TAO, T. and VU, V. (2010). Random matrices: Universality of local eigenvalue statistics up to the edge. *Comm. Math. Phys.* **298** 549–572. MR2669449

[56] TAO, T. and VU, V. (2011). Random matrices: Universality of local eigenvalue statistics. *Acta Math.* To appear. Available at arXiv:0906.0510v10.

[57] TRACY, C. A. and WIDOM, H. (1999). Universality of the distribution functions of random matrix theory. In *Statistical Physics on the Eve of the* 21*st Century. Series on Advances in Statistical Mechanics* **14** 230–239. World Sci. Publishing, River Edge, NJ. MR1704004

[58] TRACY, C. A. and WIDOM, H. (2000). Universality of the distribution functions of random matrix theory. In *Integrable Systems*: *From Classical to Quantum* (*Montréal*, *QC*, 1999). *CRM Proceedings & Lecture Notes* **26** 251–264. Amer. Math. Soc., Providence, RI. MR1791893

[59] TRACY, C. A. and WIDOM, H. (2000). The distribution of the largest eigenvalue in the Gaussian ensembles. In *Calogero–Moser–Sutherland Models* (J. F. vanDiejen and L. Vinet, eds.). *CRM Series in Mathematical Physics* **4** 461–472. Springer, New York.

[60] TRAN, L., VU, V. and WANG, K. (2010). Sparse random graphs: Eigenvalues and eigenvectors. Preprint. Available at http://www.math.rutgers.edu/~linhtran/.

[61] WEISS, Y. (1999). Segmentation using eigenvectors: A unifying view. In *Proc. of the* 7*th IEEE Intern. Conf. on Computer Vision*, *Vol.* 2 975–982. IEEE Press, New York.

UNIVERSITY OF WASHINGTON
C-342 PADELFORD HALL
SEATTLE, WASHINGTON 98195
USA
E-MAIL: dumitriu@math.washington.edu

UNIVERSITY OF WASHINGTON
C-547 PADELFORD HALL
SEATTLE, WASHINGTON 98195
USA
E-MAIL: soumik@math.washington.edu