The spectra of random graphs with given expected degrees

Fan Chung[*†] Linyuan Lu[*†] Van Vu [‡]

Abstract

In the study of the spectra of power law graphs, there are basically two competing approaches. One is to prove analogues of Wigner's semi-circle law while the other predicts that the eigenvalues follow a power law distributions. Although the semi-circle law and the power law have nothing in common, we will show that both approaches are essentially correct if one considers the appropriate matrices. We will show that (under certain conditions) the eigenvalues of the (normalized) Laplacian of a random power law graph follow the semi-circle law while the spectrum of the adjacency matrix of a power law graph obeys the power law. Our results are based on the analysis of random graphs with given expected degrees and their relations to several key invariants. Of interest are a number of (new) values for the exponent β where phase transitions for eigenvalue distributions occur. The spectrum distributions have direct implications to numerous graph algorithms such as randomized algorithms that involve rapidly mixing Markov chains, for example. [1]

1 Introduction

Eigenvalues of graphs are useful for controlling many graph properties and consequently have numerous algorithmic applications including low rank approximations [4], information retrieval [22] and computer vision [16]. Of particular interest is the study of eigenvalues for graphs with power law degree distributions (i.e., the number of vertices of degree j is proportional to $j^{-\beta}$ for some exponent β). It has been observed by many research groups [2, 3, 5, 14, 20, 23, 24] that many realistic massive graphs including Internet graphs, telephone call graphs and various social and biological networks have power law degree distributions.

For the classical random graphs based on the Erdős-Rényi's model, it has been proved by Füredi and Komlós that the spectrum of the adjacency matrix follows Wigner's semi-circle law [19]. Wigner's theorem [28] and its extensions have long been used for the stochastic treatment of complex quantum systems that lie beyond the reach of exact methods. The semi-circle law has extensive applications in statistical physics and solid state physics [11, 18].

In the 1999 paper by Faloutsos et al. [14] on Internet topology, several power law examples of Internet topology are given and the eigenvalues of the adjacency matrices are plotted which does not follow the semi-circle law. It is conjectured that the eigenvalues of the adjacency matrices have a power law distribution with its own exponent different from the exponent of the graph. Farkas et. al. [15] looked beyond the semi-circle law and described a 'triangular-like' shape distribution (also see [17]). Recently, Mihail and Papadimitriou [26] showed that the eigenvalues of the adjacency matrix of a power law graphs with exponent β are distributed according to a power law, for $\beta > 3$.

[*]Research supported in part by NSF Grants DMS 0100472 and ITR 0205061. fan@ucsd.edu, llu@math.ucsd.edu

[†]Department of Mathematics, UCSD, La Jolla, CA 92093.

[‡]Research supported in part by NSF grant DMS-0200357 and by an A. Sloan fellowship. vanvu@ucsd.edu

[1]A short version of this paper without all the proofs has appeared in the *Proceedings of National Academy of Sciences,* **100**, no. 11, (2003), 6313-6318.

Here we intend to reconcile these two schools of thoughts on eigenvalue distributions. To begin with, there are in fact several ways to associate a matrix to a graph. The usual adjacency matrix A associated with a (simple) graph has eigenvalues quite sensitive to the maximum degree (which is a *local* property). The combinatorial Laplacian $D-A$ with D denoting the diagonal degree matrix is a major tool for enumerating spanning trees and has numerous applications [6, 21]. Another matrix associated with a graph is the (normalized) Laplacian $L = I - D^{-1/2}AD^{-1/2}$ which controls the expansion/isoperimetrical properties (which are *global*) and essentially determines the mixing rate of a random walk on the graph. The traditional random matrices and random graphs are regular or almost regular so the spectra of all the above three matrices are basically the same (with possibly a scaling factor or a linear shift). However, for graphs with uneven degrees, the above three matrices can have very different distributions.

In this paper, we will consider random graphs with a general given expected degree distribution and we examine the spectra for both the adjacency matrix and the Laplacian. We will first establish bounds for eigenvalues for graphs with a general degree distribution from which the results on random power law graphs then follow. Here is a summary of our results:

1. The largest eigenvalue of the adjacency matrix of a random graph with a given expected degree sequence is determined by m, the maximum degree, and $\tilde{d}$, the weighted average of the squares of the expected degrees. We show that the largest eigenvalue of the adjacency matrix is almost surely $(1+o(1))\max\{\tilde{d}, \sqrt{m}\}$ provided some minor conditions are satisfied. In addition, suppose that the k^{th} largest expected degree m_k is significantly larger than $\tilde{d}^2$. Then the k^{th} largest eigenvalue of the adjacency matrix is almost surely $(1+o(1))\sqrt{m_k}$.

2. For a random power law graph with exponent $\beta > 2.5$, the largest eigenvalue of a random power law graph is almost surely $(1+o(1))\sqrt{m}$ where m is the maximum degree. Moreover, the k largest eigenvalues of a random power law graph with exponent β have power law distribution with exponent $2\beta-1$ if the maximum degree is sufficiently large and k is bounded above by a function depending on β, m and d, the average degree. When $2 < \beta < 2.5$, the largest eigenvalue is heavily concentrated at $cm^{3-\beta}$ for some constant c depending on β and the average degree.

3. We will show that the eigenvalues of the Laplacian satisfy the semi-circle law under the condition that the minimum expected degree is relatively large ($\gg$ the square root of the expected average degree). This condition contains the basic case when all degrees are equal (the Erdös-Rényi model). If we weaken the condition on the minimum expected degree, we can still have the following strong bound for the eigenvalues of the Laplacian which implies strong expansion rates for rapidly mixing,

$$\max_{i\neq 0} |1-\lambda_i| \leq (1+o(1))\frac{4}{\sqrt{\bar{w}}} + \frac{g(n)\log^2 n}{w_{\min}}$$

   where $\bar{w}$ is the expected average degree, $w_{\min}$ is the minimum expected degree and $g(n)$ is any slow growing function of n.

In applications, it usually suffices to have the λ_i's ($i > 0$) bounded away from zero. Our result shows that (under some mild conditions) these eigenvalues are actually very close to 1.

The rest of the paper has two parts. In Section 2, we present our model and the results concerning the spectrum of the adjacency matrix. Section 3 deals with the Laplacian.

2 The spectra of the adjacency matrix

2.1 The random graph model

The primary model for classical random graphs is the Erdős-Rényi model $\mathcal{G}_p$, in which each edge is independently chosen with the probability p for some given $p > 0$ (see [13]). In such random graphs the degrees (the number of neighbors) of vertices all have the same expected value. Here we consider the following extended random graph model for a general degree distribution.

For a sequence $\mathbf{w} = (w_1, w_2, \ldots, w_n)$, we consider random graphs $G(\mathbf{w})$ in which edges are independently assigned to each pair of vertices (i,j) with probability $w_i w_j \rho$, where $\rho = \frac{1}{\sum_{i=1}^n w_i}$. Notice that we allow loops in our model (for computational convenience) but their presence does not play any essential role. It is easy to verify that the expected degree of i is w_i.

To this end, we assume that $\max_i w_i^2 < \sum_k w_k$, so that $p_{ij} \le 1$ for all i and j. This assumption insures that the sequence w_i is graphical (in the sense that it satisfies the necessary and sufficient condition for a sequence to be realized by a graph [12]) except that we do not require the w_i's to be integers). We will use d_i to denote the actual degree of v_i in a random graph G in $G(\mathbf{w})$ where the weight w_i denotes the expected degree.

For a subset S of vertices, the volume $\mathrm{Vol}(S)$ is defined as the sum of weights in S and $\mathrm{vol}(S)$ is the sum of the (actual) degrees of vertices in S. That is, $\mathrm{Vol}(S) = \sum_{i \in S} w_i$ and $\mathrm{vol}(S) = \sum_{i \in S} d_i$. In particular, we have $\mathrm{Vol}(G) = \sum_i w_i$, and we denote $\rho = \frac{1}{\mathrm{Vol}(G)}$. The induced subgraph on S is a random graph $G(\mathbf{w}')$ where the weight sequence is given by $w_i' = w_i \mathrm{Vol}(S) \rho$ for all $i \in S$. The expected average degree is $\bar{w} = \sum_{i=1}^n w_i / n = 1/(\rho n)$. The second order average degree of $G(\mathbf{w}')$ is $\tilde{d} = \frac{\sum_{i \in S} w_i^2}{\sum_{i=1}^n w_i} = \sum_{i \in S} w_i^2 \rho$. The maximum expected degree is denoted by m.

The classical random graph $G(n,p)$ can be viewed as a special case of $G(\mathbf{w})$ by taking $\mathbf{w}$ to be $(pn, pn, \ldots, pn)$. In this special case, we have $\tilde{d} = \bar{w} = m = np$. It is well known that the largest eigenvalue of the adjacency matrix of $G(n,p)$ is almost surely $(1+o(1))np$ provided that $np \gg \log n$.

The asymptotic notation is used under the assumption that n, the number of vertices, tends to infinity. All logarithms have the natural base.

2.2 The spectra of the adjacency matrix of random graphs with given degree distribution

For random graphs with given expected degrees $w_1, w_2, \ldots, w_n$, there are two easy lower bounds for the largest eigenvalue $\|A\|$ of the adjacency matrix A, namely, $(1+o(1))\tilde{d}$ and $(1+o(1))\sqrt{m}$.

In [10], the present authors proved that the maximum of the above two lower bounds is essentially an upper bound.

Theorem 1 *If $\tilde{d} > \sqrt{m} \log n$, then the largest eigenvalue of a random graph in $G(\mathbf{w})$ is almost surely $(1+o(1))\tilde{d}$.*

Theorem 2 *If $\sqrt{m} > \tilde{d} \log^2 n$, then almost surely the largest eigenvalue of a random graph in $G(\mathbf{w})$ is $(1+o(1))\sqrt{m}$.*

If the k-th largest expected degree m_k satisfies $\sqrt{m_k} > \tilde{d} \log^2 n$ and $m_k^2 \gg m\tilde{d}$, then almost surely the largest k eigenvalues of a random graph in $G(\mathbf{w})$ is $(1+o(1))\sqrt{m_k}$.

Theorem 3 *The largest eigenvalue of a random graph in $G(\mathbf{w})$ is almost surely at most*

$$7\sqrt{\log n} \cdot \max\{\sqrt{m}, \tilde{d}\}.$$

We remark that the largest eigenvalue $\|A\|$ of the adjacency matrix of a random graph is almost surely $(1+o(1))\sqrt{m}$ if $\sqrt{m}$ is greater than $\tilde{d}$ by a factor of $\log^2 n$, and $\|A\|$ is almost surely $(1+o(1))\tilde{d}$ if $\sqrt{m}$ is smaller than $\tilde{d}$ by a factor of $\log n$. In other words, $\|A\|$ is (asymptotically) the maximum of $\sqrt{m}$ and $\tilde{d}$ if the two values of $\sqrt{m}$ and $\tilde{d}$ are far apart (by a power of $\log n$). One might be tempted to conjecture that

$$\|A\| = (1+o(1))\max\{\sqrt{m}, \tilde{d}\}.$$

This, however, is not true as shown by a counterexample given in [10].

We also note that with a more careful analysis the factor of $\log n$ in Theorem 1 can be replaced by $(\log n)^{1/2+\epsilon}$ and the factor of $\log^2 n$ can be replaced by $(\log n)^{3/2+\epsilon}$ for any positive ϵ provided that n is sufficiently large. We remark that the constant "7" in Theorem 3 can be improved. We made no effort to get the best constant coefficient here.

2.3 The eigenvalues of the adjacency matrix of power law graphs

In this section, we consider random graphs with power law degree distribution with exponent β. We want to show that the largest eigenvalue of the adjacency matrix of a random power law graph is almost surely approximately the square root of the maximum degree m if $\beta > 2.5$, and is almost surely approximately $cm^{3-\beta}$ if $2 < \beta < 2.5$. A phase transition occurs at $\beta = 2.5$. This result for power law graphs is an immediate consequence of a general result for eigenvalues of random graphs with arbitrary degree sequences.

We choose the degree sequence $\mathbf{w} = (w_1, w_2, \ldots, w_n)$ satisfying $w_i = ci^{-\frac{1}{\beta-1}}$ for $i_0 \le i \le n+i_0$. Here c is determined by the average degree and i_0 depends on the maximum degree m, namely, $c = \frac{\beta-2}{\beta-1} dn^{\frac{1}{\beta-1}}, i_0 = n(\frac{d(\beta-2)}{m(\beta-1)})^{\beta-1}$. It is easy to verify that the number of vertices of degree k is proportional to $k^{-\beta}$.

The second order average degree $\tilde{d}$ can be computed as follows:

$$\tilde{d} = \begin{cases} d\frac{(\beta-2)^2}{(\beta-1)(\beta-3)}(1+o(1)) & \text{if } \beta > 3. \\ \frac{1}{2} d \ln \frac{2m}{d}(1+o(1)). & \text{if } \beta = 3. \\ d\frac{(\beta-2)^2}{(\beta-1)(3-\beta)}(\frac{(\beta-1)m}{d(\beta-2)})^{3-\beta}(1+o(1)). & \text{if } 2<\beta<3. \end{cases}$$

We remark that for $\beta > 3$, the second order average degree is independent of the maximum degree. Consequently, the power law graphs with $\beta > 3$ are much easier to deal with. However, many massive graphs are power law graphs with $2 < \beta < 3$, in particular, Internet graphs [23] have exponents between 2.1 and 2.4 while the Hollywood graph [5] has exponent $\beta \sim 2.3$. In these cases, it is $\tilde{d}$ which determines the first eigenvalue. The following theorem is a consequence of Theorems 1 and 2. When $\beta > 2.5$, we have

$$\lambda_i \approx \sqrt{m_i} \propto (i+i_0-1)^{-1/((2\beta-1)-1)},$$

for λ_i sufficiently large. These large eigenvalues follows the power law distribution with exponent $2\beta-1$. (The exponent is different from one in Mihail and Papadimitriou's paper [26] because they use a different definition for power law.)

Theorem 4 *1. For $\beta \ge 3$ and $m > d^2 \log^{3+\epsilon} n$, almost surely the largest eigenvalue of the random power law graph G is $(1+o(1))\sqrt{m}$.*

2. For $2.5 < \beta < 3$ and $m > d^{\frac{\beta-2}{\beta-2.5}} \log^{\frac{3}{\beta-2.5}} n$, almost surely the largest eigenvalue of the random power law graph G is $(1+o(1))\sqrt{m}$.

3. *For $2 < \beta < 2.5$ and $m > \log^{\frac{3}{2.5-\beta}} n$, almost surely the largest eigenvalue is $(1+o(1))\tilde{d}$.*

4. *For $k < (\frac{d}{m \log n})^{\beta-1} n$ and $\beta > 2.5$, almost surely the k largest eigenvalues of the random power law graph G with exponent β have power law distribution with exponent $2\beta-1$, provided that m is large enough (satisfying the inequalities in 1, 2).*

3 The spectrum of the Laplacian

Suppose G is a graph that does not contain any isolated vertices. The Laplacian L is defined to be the matrix $L = I - D^{-1/2} A D^{-1/2}$ where I is the identity matrix, A is the adjacency matrix of G and D denotes the diagonal degree matrix. The eigenvalues of L are all non-negative between 0 and 2 (see [7]). We denote the eigenvalues of L by $0 = \lambda_0 \leq \lambda_1 \leq \ldots \lambda_{n-1}$. For each i, let ϕ_i denote an orthonormal eigenvectors associated with λ_i. We can write L as

$$L = \sum_i \lambda_i P_i,$$

where P_i denotes the i-projection into the eigenspace associated with eigenvalue λ_i. We consider

$$\begin{aligned} M &= I - L - P_0 \\ &= \sum_{i \neq 0} (1 - \lambda_i) P_i. \end{aligned}$$

For any positive integer k, we have

$$\mathrm{Trace}(M^{2k}) = \sum_{i \neq 0} (1 - \lambda_i)^{2k}$$

Lemma 3.1 *For any positive integer k, we have*

$$\max_{i \neq 0} |1 - \lambda_i| \leq \|M\| \leq (\mathit{Trace}(M^{2k})^{1/(2k)}.$$

The matrix M can be written as

$$\begin{aligned} M &= D^{-1/2} A D^{-1/2} - P_0 \\ &= D^{-1/2} A D^{-1/2} - \phi_0^* \phi_0 \\ &= D^{-1/2} A D^{-1/2} - \frac{1}{\mathrm{vol}(G)} D^{1/2} K D^{1/2} \end{aligned}$$

where ϕ_0 is regarded as a row vector $(\sqrt{d_1/\mathrm{vol}(G)}, \ldots, \sqrt{d_n/\mathrm{vol}(G)})$, ϕ_0^* is the transpose of ϕ_0 and K is the all 1's matrix.

Let W denote the diagonal matrix with the (i,i)-entry having value w_i, the expected degree of the i-th vertex. We will approximate M by

$$\begin{aligned} C &= W^{-1/2} A W^{-1/2} - \frac{1}{\mathrm{Vol}(G)} W^{1/2} K W^{1/2} \\ &= W^{-1/2} A W^{-1/2} - \chi^* \chi \end{aligned}$$

where χ is a row vector $(\sqrt{w_1 \rho}, \ldots, \sqrt{w_n \rho})$. We note that $\|\chi^* \chi - \phi^* \phi\|$ is strongly concentrated at 0 for random graphs with given expected degree w_i. C can be seen as the *expectation* of M and we shall consider the spectrum of C carefully.

3.1 A sharp bound for random graphs with relatively large minimum expected degree

In this section we consider the case when the minimum of the expected degrees is not too small compared to the mean. In this case, we are able to prove a sharp bound on the largest eigenvalue of C.

Theorem 5 *For a random graph with given expected degrees $w_1,\dots,w_n$ where $w_{\min} \gg \sqrt{\bar{w}}\log^3 n$, we have almost surely*

$$\|C\| = (1+o(1))\frac{2}{\sqrt{\bar{w}}}.$$

Proof. We rely on Wigner's high moment method. For any positive integer k and any symmetric matrix C

$$\mathrm{Trace}(C^{2k}) = \lambda_1(C)^{2k}+\cdots+\lambda_n(C)^{2k},$$

which implies

$$\mathbf{E}(\lambda_1(C)^{2k}) \le \mathbf{E}(\mathrm{Trace}(C^{2k})),$$

where λ_1 is the eigenvalue with maximum absolute value: $|\lambda_1| = \|C\|$.

If we can bound $\mathbf{E}(\mathrm{Trace}(C^{2k}))$ from above, then we have an upper bound for $\mathbf{E}(\lambda_1(C)^{2k})$. The latter would imply an upper bound (almost surely) on $|\lambda_1(C)|$ via Markov's inequality, provided that k is sufficiently large.

Let us now take a closer look at $\mathrm{Trace}(C^{2k})$. This is a sum where a typical term is $c_{i_1i_2}c_{i_2i_3}\dots c_{i_{2k-1}i_{2k}}c_{i_{2k}i_1}$. In other words, each term corresponds to a closed walk of length $2k$ (containing $2k$, not necessarily different, edges) of the complete graph K_n on $\{1,\dots,n\}$ (K_n has a loop at every vertex). On the other hand, the entries c_{ij} of C are independent random variables with mean zero. Thus, the expectation of a term is non-zero if and only if each edge of K_n appears in the walk at least twice. To this end, we call such a walk a *good* walk. Consider a closed good walk which uses l different edges $e_1,\dots,e_l$ with corresponding multiplicities $m_1,\dots,m_l$ (the m_h's are positive integers at least 2 summing up to $2k$). The (expected) contribution of the term defined by this walk in $\mathbf{E}(\mathrm{Trace}(C^{2k})$ is

$$\prod_{h=1}^{l}\mathbf{E}(c_{e_h}^{m_h}). \tag{1}$$

In order to compute $\mathbf{E}(c_{ij}^m)$, let us first describe the distribution of c_{ij}: $c_{ij} = \frac{1}{\sqrt{w_iw_j}} - \sqrt{w_iw_j}\rho = \frac{q_{ij}}{\sqrt{w_iw_j}}$ with probability $p_{ij} = w_iw_j\rho$ and $c_{ij} = -\sqrt{w_iw_j}\rho = -\frac{p_{ij}}{\sqrt{w_iw_j}}$ with probability $q_{ij} = 1-p_{ij}$. This implies that for any $m \ge 2$

$$|\mathbf{E}(c_{ij}^m)| \le \frac{q_{ij}^m p_{ij}+(-p_{ij})^m q_{ij}}{(w_iw_j)^{m/2}} \le \frac{p_{ij}}{(w_iw_j)^{m/2}} = \frac{\rho}{(w_iw_j)^{m/2-1}} \le \frac{\rho}{w_{\min}^{m-2}}. \tag{2}$$

Here we used the fact that $q_{ij}^m p_{ij}+(-p_{ij})^m q_{ij} \le p_{ij}$ in the first inequality (the reader can consider this fact an easy exercise) and the definition $p_{ij} = w_iw_j\rho$ in the second equality.

Let $W_{l,k}$ denote the set of closed good walks on K_n of length $2k$ using exactly $l+1$ different vertices. Notice that each walk in $W_{l,k}$ must have at least l different edges. By (1) and (2), the contribution of a term corresponding to such a walk towards $\mathbf{E}(\mathrm{Trace}(C^{2k}))$ is at most

$$\frac{\rho^l}{w_{\min}^{2k-2l}}.$$

It follows that

$$\mathbf{E}(\mathrm{Trace}(C^{2k})) \leq \sum_{l=0}^{k} |W_{l,k}| \frac{\rho^l}{w_{\min}^{2k-2l}}. \tag{3}$$

In order to bound the last sum, we need the following result of Füredi and Komlós [19].

Lemma 3.2 *For all $l < n$*

$$|W_{l,k}| \leq n(n-1)\dots(n-l)\binom{2k}{2l}\binom{2l}{l}\frac{1}{l+1}(l+1)^{4(k-l)}. \tag{4}$$

In order to prove our theorem, it is more convenient to use the following cleaner bound, which is a direct corollary of (4)

$$|W_{l,k}| \leq n^{l+1}4^l\binom{2k}{2l}(l+1)^{4(k-l)}. \tag{5}$$

Substituting (5) into (3) yields

$$\mathbf{E}(\mathrm{Trace}(C^{2k})) \leq \sum_{l=0}^{k} \frac{\rho^l}{w_{\min}^{2k-2l}} n^{l+1}4^l\binom{2k}{2l}(l+1)^{4(k-l)} = \sum_{l=0}^{k} s_{l,k}. \tag{6}$$

Now fix $k = g(n)\log n$, where $g(n)$ tends to infinity (with n) arbitrarily slowly. With this k and the assumption about the degree sequence, the last sum in (6) is dominated by its highest term. To see this, let us consider the ratio $s_{k,k}/s_{l,k}$ for some $l \leq k-1$:

$$\frac{s_{k,k}}{s_{l,k}} = \frac{((4\rho n)w_{\min}^2)^{k-l}}{\binom{2k}{2l}(l+1)^{4(k-l)}} \geq \frac{((4\rho n)w_{\min}^2)^{k-l}}{2k^{2(k-l)}k^{4(k-l)}} \geq \frac{1}{2}\Big(\frac{4\rho n w_{\min}^2}{k^6}\Big)^{k-l},$$

where in the first inequality we used the simple fact that $\binom{2k}{2l} \leq \frac{(2k)^{2(k-l)}}{2(k-l)!} \leq 2k^{2(k-l)}$. With a proper choice of $g(n)$, the assumption $w_{\min} = \Omega(\log^3 n)\sqrt{\bar{w}}$ guarantees that $\frac{4\rho n w_{\min}^2}{k^6} = \Omega(1)$, where $\Omega(1)$ tends to infinity with n. This implies $s_{k,k}/s_{l,k} \geq (\Omega(1))^{k-l}$. Consequently,

$$\mathbf{E}(\mathrm{Trace}(C^{2k})) \leq \sum_{l=0}^{k} s_{l,k} \leq (1+o(1))s_{k,k} = (1+o(1))\rho^k n^{k+1}4^k = (1+o(1))n(4\rho n)^k.$$

Since $\mathbf{E}(\lambda_1(C)^{2k}) \leq \mathbf{E}(\mathrm{Trace}(C^{2k}))$ and $\rho n = \frac{1}{\bar{w}}$, we have

$$\mathbf{E}(\lambda_1(C)^{2k}) \leq (1+o(1))n(\frac{2}{\sqrt{\bar{w}}})^{2k}. \tag{7}$$

By (7) and Markov's equality

$$\begin{aligned}\mathbf{P}\Big(|\lambda_1(C)| \geq (1+\epsilon)\frac{2}{\sqrt{\bar{w}}}\Big) &= \mathbf{P}\Big(\lambda_1(C)^{2k} \geq (1+\epsilon)^{2k}(\frac{2}{\sqrt{\bar{w}}})^{2k}\Big) \\ &\leq \frac{\mathbf{E}(\lambda_1(C)^k)}{(1+\epsilon)^{2k}(\frac{2}{\sqrt{\bar{w}}})^{2k}} \leq \frac{(1+o(1))n(\frac{2}{\sqrt{\bar{w}}})^{2k}}{(1+\epsilon)^{2k}(\frac{2}{\sqrt{\bar{w}}})^{2k}} \\ &= \frac{(1+o(1))n}{(1+\epsilon)^{2k}}.\end{aligned}$$

Since $k = \Omega(\log n)$, we can find an $\epsilon = \epsilon(n)$ tending to 0 with n so that $\frac{n}{(1+\epsilon)^{2k}} = o(1)$. This implies that almost surely $|\lambda_1(C)| \leq (1+o(1))\frac{2}{\sqrt{\bar{w}}}$, as desired. The lower bound on $|\lambda_1(C)|$ follows from the semi-circle law proved in the next section. □

3.2 The semi-circle law.

We show that if the minimum expected degree is relatively large then the eigenvalues of C satisfy the semi-circle law with respect to the circle of radius $r = \frac{2}{\sqrt{\bar{w}}}$ centered at 0. Let W be an absolute continuous distribution function with (semi-circle) density $w(x) = \frac{2}{\pi}\sqrt{1-x^2}$ for $|x| \leq 1$ and $w(x) = 0$ for $|x| > 1$. For the purpose of normalization, consider $C_{\text{nor}} = (\frac{2}{\sqrt{\bar{w}}})^{-1}C$. Let $N(x)$ be the number of eigenvalues of C_{nor} less than x and $W_n(x) = n^{-1}N(x)$.

Theorem 6 *For random graphs with a degree sequence satisfying $w_{\min} \gg \sqrt{\bar{w}}$, $W_n(x)$ tends to $W(x)$ in probability as n tends to infinity.*

Remark. The assumption here is weaker than that of Theorem 5, due to the fact that we only need to consider moments of constant order.

Proof. As convergence in probability is entailed by the convergence of moments, to prove this theorem, we need to show that for any fixed s, the s^{th} moment of $W_n(x)$ (with n tending to infinity) is asymptotically the s^{th} moment of $W(x)$. The s^{th} moment of $W_n(x)$ equals $\frac{1}{n}\mathbf{E}(\text{Trace}(C^s_{\text{nor}}))$. For s even, $s = 2k$, the s^{th} moment of W_x is $\frac{(2k)!}{2^{2k}k!(k+1)!}$ (see [28]). For s odd, the s^{th} moment of W_x is 0 by symmetry.

In order to verify Theorem 5, we need to show that for any fixed k

$$\frac{1}{n}\mathbf{E}(\text{Trace}(C^{2k}_{\text{nor}})) = (1+o(1))\frac{(2k)!}{2^{2k}k!(k+1)!}, \tag{8}$$

and

$$\frac{1}{n}\mathbf{E}(\text{Trace}(C^{2k+1}_{\text{nor}})) = o(1). \tag{9}$$

We first consider (8). Let us go back to (3). Now we need to use the more accurate estimate of $|W_{l,k}|$ given in (4), instead of the weaker but cleaner one in (5). Define $s'_{l,k} = \frac{\rho^l}{w_{\min}^{2k-2l}}n(n-1)\ldots(n-l)\binom{2k}{2l}\binom{2l}{l}\frac{1}{l+1}(l+1)^{4(k-l)}$. One can check, with a more tedious computation, that the sum $\sum_{l=0}^{k} s'_{l,k}$ is still dominated by the last term, namely

$$\sum_{l=0}^{k} s'_{l,k} = (1+o(1))s'_{k,k}.$$

It follows that $\mathbf{E}(\mathrm{Trace}(C^{2k})) \le (1+o(1))s'_{k,k}$. On the other hand, $\mathbf{E}(\mathrm{Trace}(C^{2k})) \ge |W_{k,k}|\rho^k$. Now comes the important point, for $l = k$, $|W_{l,k}|$ is not only upper bounded by, but in fact equals, the right hand side of (4). Therefore,

$$\mathbf{E}(\mathrm{Trace}(C^{2k})) = (1+o(1))s'_{k,k}.$$

It follows that

$$\mathbf{E}(\mathrm{Trace}(C\mathrm{nor}^{2k})) = (1+o(1))(\frac{2}{\sqrt{\bar{w}}})^{-2k} s'_{k,k} = (1+o(1))n\frac{(2k)!}{2^{2k}k!(k+1)!},$$

which implies (8).

Now we turn to (9). Consider a term in $\mathrm{Trace}(C^{2k+1})$. If the closed walk corresponding to this term has at least $k+1$ different edges, then there should be an edge with multiplicity one, and the expectation of the term is 0. Therefore, we only have to look at terms whose walks have at most k different edges (and at most $k+1$ different vertices). It is easy to see that the number of closed good walks of length $2k+1$ with exactly $l+1$ different vertices is at most $O(n^{l+1})$. The constant in O depends on k and l (recall that now k is a constant) but for the current task we do not need to estimate this constant. The contribution of a term corresponding to a walk with at most $l+1$ different edges is bounded by

$$\frac{\rho^l}{w_{\min}^{2k+1-2l}}.$$

Thus $|\mathbf{E}(\mathrm{Trace}(C^{2k+1}))|$ is upper bounded by

$$\sum_{l=0}^{k} c\frac{\rho^l}{w_{\min}^{2k+1-2l}} n^{l+1}, \tag{10}$$

for some constant c. To compute the $(2k+1)^{th}$ moment of $W_n(x)$, we need to multiply $\mathbf{E}(\mathrm{Trace}(C^{2k+1}))$ by the normalizing factor $\frac{1}{n}(\frac{1}{2\sqrt{n\rho}})^{2k+1}$. It follows from (10) that the absolute value of the $(2k+1)^{th}$ moment of $W_n(x)$ is upper bounded by

$$\sum_{l=0}^{k} \frac{1}{n}(\frac{1}{2\sqrt{n\rho}})^{2k+1}\frac{\rho^l}{w_{min}^{2k+1-2l}} n^{l+1} \le \sum_{l=0}^{k}(\frac{1}{2\sqrt{n\rho}w_{\min}})^{2k+1-2l}. \tag{11}$$

Under the assumption of the theorem $\frac{1}{2\sqrt{n\rho}w_{\min}} = o(1)$. Thus, the last sum in (11) is $o(1)$, completing the proof. $\square$

3.3 An upper bound on the spectral norm of the Laplacian

In this section, we assume that $w_{\min} \gg \log^2 n$ and we will show the following.,

Theorem 7 *For a random graph with given expected degrees, if the minimal expected degree $w_{\min}$ satisfies $w_{\min} \gg \log^2 n$, then almost surely the eigenvalues of the Laplacian L satisfy*

$$\max_{i \neq 0} |1-\lambda_i| \leq (1+o(1))\frac{4}{\sqrt{\bar{w}}} + \frac{g(n)\log^2 n}{w_{\min}}$$

where $\bar{w} = \frac{\sum_{i=1}^n w_i}{n}$ is the average expected degree and $g(n)$ is a function tending to infinity (with n) arbitrarily slowly.

To proof Theorem 7, we recall that eigenvalues of the Laplacian satisfy

$$\max_{i \neq 0} |1-\lambda_i| = \|M\|$$

where $M = D^{-1/2}AD^{-1/2} - \frac{1}{\mathrm{vol}(G)}D^{1/2}KD^{1/2}$. We rewrite M as follows:

$$M = B+C+R+S \quad \text{where} \quad \begin{aligned} b_{i,j} &= (a_{i,j} - w_i w_j \rho)\left(\frac{1}{\sqrt{d_i d_j}} - \frac{1}{\sqrt{w_i w_j}}\right) \\ r_{i,j} &= \rho \frac{w_i w_j - d_i d_j}{\sqrt{d_i d_j}} \\ s_{i,j} &= \left(\frac{1}{\mathrm{Vol}(G)} - \frac{1}{\mathrm{vol}(G)}\right)\sqrt{d_i d_j} \end{aligned}$$

and C is as defined in the previous section. Clearly,

$$\|M\| \leq \|B\| + \|C\| + \|R\| + \|S\|.$$

It suffices to establish upper bounds for the norms of B,C, E and F separately. To do so, we will use the following concentration inequality for a sum of independent random variables (see [8, 25]).

Let X_i $(1 \leq i \leq n)$ be independent random variables satisfying $|X_i| \leq M$. Let $X = \sum_i X_i$. Then we have

$$\mathbf{P}(|X - \mathbf{E}(X)| > a) \leq e^{-\frac{a^2}{2(Var(X)+Ma/3)}}. \tag{12}$$

For each fixed i, we consider the degree d_i as a sum of random indicator variables $d_i = \sum_j a_{ij}$. Since $Var(d_j) \leq w_j$, we then have

$$\mathbf{P}(|d_i - w_i| > a) \leq e^{-a^2/(w_i + a/3)}. \tag{13}$$

By the assumption that $w_{\min} \gg \log^2 n$, we have almost surely

$$|d_i - w_i| < \epsilon w_i \tag{14}$$

for all i where ϵ is any fixed (small) positive value.

Similarly, by considering the volume $\mathrm{vol}(G)$ as $\mathrm{vol}(G) = \sum_i \sum_j a_{ij}$, we have almost surely

$$|\mathrm{vol}(G) - \mathrm{Vol}(G)| < 2\sqrt{\mathrm{Vol}(G)}g(n) \tag{15}$$

for any slow growing function $g(n)$.

We will use the following lemma which will be proved later.

Lemma 3.3 *Suppose that* $w_{\min} \gg \log n$. *Almost surely the vector* χ *with* $\chi(i) = (d_i - w_i)/\sqrt{w_i}$ *satisfies*

$$\|\chi\|^2 \leq (1+o(1))n.$$

Proof of Theorem 7: To establish an upper bound for $\|C\|$, we follow the proof of Theorem 5. The following inequality can be derived from (6).

$$\begin{aligned}
\mathbf{E}(\mathrm{Trace}(C^{2k})) &\leq \sum_{l=0}^{k} \frac{\rho^l}{w_{\min}^{2k-2l}} n^{l+1} 4^l \binom{2k}{2l} (l+1)^{4(k-l)} \\
&\leq \sum_{l=0}^{k} \frac{\rho^l}{w_{\min}^{2k-2l}} n^{l+1} 4^l \binom{2k}{2l} (k+1)^{4(k-l)} \\
&\leq (1+o(1))n(\frac{2}{\sqrt{\bar{w}}} + \frac{(k+1)^2}{w_{\min}})^{2k}
\end{aligned}$$

By choosing $k = \sqrt{g(n)}\log n$, we have

$$(\mathbf{E}(\mathrm{Trace}(C^{2k})))^{1/(2k)} \leq n^{1/(2k)}(\frac{2}{\sqrt{\bar{w}}} + \frac{g(n)\log^2 n}{w_{\min}}).$$

Thus, by similar arguments, almost surely we have

$$\|C\| \leq \frac{2}{\sqrt{\bar{w}}} + \frac{g(n)\log^2 n}{w_{\min}}.$$

To bound $\|R\|$, we have almost surely

$$\begin{aligned}
\|R\| &= \max_{\|y\|=1} \langle y, Ry \rangle \\
&\leq \max_{\|y\|=1} \rho \sum_{ij} y_i y_j \frac{d_i(d_j - w_j) + (d_i - w_i)w_j}{\sqrt{d_i d_j}} \\
&\leq \rho \max_{\|y\|=1} \left\{ \sum_i \sqrt{d_i} y_i \sum_j \frac{(d_j - w_j)y_j}{\sqrt{d_j}} + \sum_i \frac{d_i - w_i)y_i}{\sqrt{d_i}} \sum_j \frac{w_j y_j}{\sqrt{d_j}} \right\} \\
&\leq \rho \max_{\|y\|=1} \left\{ (\sum_i d_i)^{1/2} \|y\| \cdot \left(\sum_j \frac{(d_j - w_j)^2}{d_j} \right)^{1/2} \|y\| + \left(\sum_i \frac{(d_i - w_i)^2}{d_i} \right)^{1/2} \|y\| \cdot (\sum_j \frac{w_j^2}{d_j})^{1/2} \|y\| \right\} \\
&\leq (2+o(1))\sqrt{\rho n} \\
&= (1+o(1))\frac{2}{\sqrt{\bar{w}}}
\end{aligned}$$

by using (12), Lemma 3.3 and the Cauchy-Schwartz inequality.

To bound $\|S\|$, we have

$$\begin{aligned}
\|S\| &= \max_{\|y\|=1}\langle y, Sy\rangle = \max_{\|y\|=1}\sum_{ij}|y_i y_j(\frac{1}{\mathrm{Vol}(G)}-\frac{1}{\mathrm{vol}(G)})|\sqrt{d_i d_j} \\
&\le (\frac{1}{\mathrm{Vol}(G)}-\frac{1}{\mathrm{vol}(G)})|\max_{\|y\|=1}\sum_{ij}|y_i\sqrt{d_i}||y_j\sqrt{d_j}| \\
&\le \frac{2\sqrt{\mathrm{Vol}(G)\log n}}{\mathrm{vol}(G)\mathrm{Vol}(G)}(\sum_i |y_i\sqrt{d_i}|)^2 \\
&= o(\sqrt{\rho\log n}\|y\|^2) \\
&= o(\frac{1}{\sqrt{\bar{w}}})
\end{aligned}$$

almost surely by using (15).

It remains to bound $\|B\|$. We note that

$$\begin{aligned}
b_{ij} &= (a_{ij}-w_i w_j\rho)(\frac{1}{\sqrt{d_i d_j}}-\frac{1}{\sqrt{w_i w_j}}) \\
&= c_{ij}\frac{\sqrt{w_i w_j}-\sqrt{d_i d_j}}{\sqrt{d_i d_j}}.
\end{aligned}$$

Thus, we have

$$\begin{aligned}
\|B\| &= \max_{\|y\|=1}\langle y, By\rangle \\
&\le \max_{\|y\|=1}\sum_{ij} y_i y_j c_{ij}\frac{\sqrt{d_i}(\sqrt{d_j}-\sqrt{w_j})+(\sqrt{d_i}-\sqrt{w_i})\sqrt{w_j}}{\sqrt{d_i d_j}}.
\end{aligned}$$

We define $y_i' = y_i(\sqrt{d_i}-\sqrt{w_i})/\sqrt{d_i}$ and $y_i'' = y_i\sqrt{w_i}/\sqrt{d_i}$. Then we have almost surely

$$\begin{aligned}
\|B\| &\le \max_{\|y\|=1}\langle y, Cy'\rangle + \langle y', Cy''\rangle \\
&\le \max_{\|y\|=1}\|C\|\|y'\| + \|C\|\|y'\|\|y''\| \\
&\le o(\|C\|)
\end{aligned}$$

since $\|y'\|^2 = \sum_i y_i^2(\sqrt{d_i}-\sqrt{w_i})^2/d_i = o(\sum_i y^2) = o(1)$ and $\|y''\| = (1+o(1))\|y\|$.

Together we have

$$\begin{aligned}
\max_{i\neq 0}|1-\lambda_i| &\le \|M\| \\
&\le \|B\|+\|C\|+\|R\|+\|S\| \\
&\le (1+o(1))(\frac{4}{\sqrt{\bar{w}}}+\frac{g(n)\log^2 n}{w_{\min}}).
\end{aligned}$$

The proof is complete. $\square$

Proof of Lemma 3.3: Let $X_i = (d_i - w_i)^2$, $X = \sum_{i=1}^n \frac{1}{w_i} X_i$, and $x_{ij} = a_{ij} - w_i w_j \rho$. We have

$$\begin{aligned}
\mathbf{E}(X_i) &= \mathrm{Var}(d_i) = \mathbf{E}(\sum_{j=1}^n x_{ij}^2) < w_i \\
\mathbf{E}(X_i^2) &= \mathbf{E}((d_i - w_i)^4) = \mathbf{E}((\sum_{j=1}^n x_{ij})^4) \\
&= \sum_{j=1}^n \mathbf{E}(x_{ij}^4) + 6 \sum_{j_1=j_2, \neq j_3=j_4} \mathbf{E}(x_{ij_1} x_{ij_2} x_{ij_3} x_{ij_4}) \\
&\leq w_i + 6w_i^2
\end{aligned}$$

since $\mathbf{E}(x_{ij}) = 0$. For $i \neq j$, we have

$$\begin{aligned}
\mathbf{E}(X_i X_j) &= \mathbf{E}(d_i - w_i)^2 (d_j - w_j)^2 \\
&= \mathbf{E}((\sum_{k=1}^n x_{ik})^2)((\sum_{l=1}^n x_{il})^2) \\
&= \mathbf{E}(X_i)\mathbf{E}(X_j) + \mathbf{E}(x_{ij}^4) - (\mathbf{E}(x_{ij}^2))^2 \\
&\leq w_i w_j + w_i w_j \rho
\end{aligned}$$

Thus,

$$\begin{aligned}
\mathrm{Var}(X_i) &\leq w_i + 5w_i^2, \\
\mathrm{coVar}(X_i, X_j) &\leq w_i w_j \rho.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\mathbf{E}(X) &= \sum_{i=1}^n \frac{1}{w_i} \mathbf{E}(X_i) < n \\
\mathrm{Var}(X) &= \sum_{i=1}^n \frac{1}{w_i^2} \mathrm{Var}(X_i) + 2 \sum_{i<j}^n \frac{1}{w_i w_j} \mathrm{coVar}(X_i, X_j) \\
&\leq (5 + \frac{1}{w_{min}} + \frac{1}{\bar{w}})n \\
&= (5 + o(1))n
\end{aligned}$$

Using the Chebyshev inequality, we have

$$\mathbf{P}(|X - \mathbf{E}(X)| > a) \leq \frac{a^2}{\mathrm{Var}(X)}.$$

By choosing $a = \sqrt{n}\, g(n)$, where $g(n)$ is an arbitrarily slow growing function, almost surely, we have $X = (1 + o(1))n$. Thus, we have almost surely

$$\|\chi\|^2 \leq (1 + o(1))n$$

as desired. □.

References

[1] W. Aiello, F. Chung and L. Lu, A random graph model for massive graphs, *STOC* 2001, 171-180. The paper version appeared in *Experimental Math.* **10**, (2001), 53-66.

[2] W. Aiello, F. Chung and L. Lu, Random evolution in massive graphs, *Handbook of Massive Data Sets*, Volume 2, (Eds. J. Abello et al.), Kluwer Academic Publishers, (2002), 97-122. Extended abstract appeared in FOCS 2001, 510-519.

[3] R. Albert, H. Jeong and A. Barabási, Diameter of the world Wide Web, *Nature* **401** (1999), 130-131.

[4] D. Achlioptas and F. McSherry, Fast computation of low-ranked matrix approximation, *STOC* (2001), 611-618.

[5] Albert-László Barabási and Réka Albert, Emergence of scaling in random networks, Science **286** (1999) 509-512.

[6] N. Biggs, Algebraic Graph Theory, Cambridge University Press, 1993.

[7] F. Chung, *Spectral Graph Theory*, AMS Publications, 1997.

[8] F. Chung and L. Lu, Connected components in random graphs with given degree sequences, *Annals of Combinatorics*, **6** (2002), 125-145.

[9] F. Chung and L. Lu, The average distances in random graphs with given expected degrees, *Proc. Nat. Acad. Sci.*, **99**, no. 25 (2002), 15879-15882.

[10] F. Chung, L. Lu and Van Vu, Eigenvalues of random power law graphs, *Annals of Combinatorics*, **7** (2003), 21-33.

[11] A. Crisanti, G. Paladin and A. Vulpiani, *Products of Random Matrices in Statistical Physics*, Springer Series in Solid-State Sciences 104 (Springer, Berlin, 1993).

[12] P. Erdős and T. Gallai, Gráfok elóírt fokú pontokkal (Graphs with points of prescribed degrees, in Hungarian), *Mat. Lapok* **11** (1961), 264-274.

[13] P. Erdős and A. Rényi, On random graphs. I, *Publ. Math. Debrecen* **6** (1959), 290-291.

[14] M. Faloutsos, P. Faloutsos and C. Faloutsos, On power-law relationships of the Internet topology, ACM SIG-COMM '99; *Comput. Commun. Rev.* **29** (1999) 251-263.

[15] I. J. Farkas, I. Derényi, A.-L. Barabási and T. Vicsek, Spectra of "Real-World" graphs: Beyond the semi-circle law, *Phy. Rev. E* **64**, 026704 (2001), cond-mat/0102335.

[16] Charless Fowlkes, S. Belongie, F. Chung and J. Malik, Spectral Partitioning with Indefinite Kernels using the Nyström Extension,
European Conference on Computer Vision 2002, Copenhagen.

[17] K.-I. Goh, B. Kahng and D. Kim, Spectra and eigenvectors of scale-free networks, *Phys. Rev. E* **64**, 051903, (2001), cond-mat/0103337.

[18] T. Guhr, A. Müller-Groeling and H. A. Weidenmüller, Random-matrix theories in quantum physics: common concepts, *Phys. Rep.* **2:99** (1998) 189-425.

[19] Z. Füredi and J. Komlós, The eigenvalues of random symmetric matrices, *Combinatorica* **1** (1981), no. 3, 233–241.

[20] H. Jeong, B. Tomber, R. Albert, Z. Oltvai and A. L. Barábasi, The large-scale organization of metabolic networks, *Nature*, **407** (2000), 378-382.

[21] F. Kirchhoff, Über die Auflösung der Gleichungen, auf welche man bei der Untersuchung der linearen Verteilung galvanischer Ströme geführt wird, *Ann. Phys. chem.* 72 (1847) 497-508.

[22] J. Kleinberg, Authoritative sources in a hyperlinked environment, *Proc. 9th ACM-SIAM Symposium on Discrete Algorithms* 1998. Extended version in Journal of the ACM 46 (1999), 604-632.

[23] J. Kleinberg, S. R. Kumar, P. Raphavan, S. Rajagopalan and A. Tomkins, The web as a graph: Measurements, models and methods, *Proceedings of the International Conference on Combinatorics and Computing*, 1999, 26-28.

[24] L. Lu, The Diameter of Random Massive Graphs, *Proceedings of the Twelfth ACM-SIAM Symposium on Discrete Algorithms*, (2001) 912-921.

[25] C. McDiarmid, Concentration. Probabilistic methods for algorithmic discrete mathematics, *Algorithms Combin.*, **16**, Springer, Berlin, (1998) 195–248.

[26] M. Mihail and C. Papadimitriou, On the eigenvalue power law, preprint.

[27] O. Perron, Theorie der algebraischen Gleichungen, II (zweite Auflage), de Gruyter, Berlin (1933).

[28] E. P. Wigner, On the distribution of the roots of certain symmetric matrices, *The Annals of Mathematics* **67** (1958) 325-327.