

On the Spectra of General Random Graphs

Fan Chung* Mary Radcliffe

University of California, San Diego
La Jolla, CA 92093, U.S.A.

Submitted: Apr 4, 2011; Accepted: Sep 27, 2011; Published: Oct 31, 2011

Mathematics Subject Classification: 05C88

Abstract

We consider random graphs such that each edge is determined by an independent random variable, where the probability of each edge is not assumed to be equal. We use a Chernoff inequality for matrices to show that the eigenvalues of the adjacency matrix and the normalized Laplacian of such a random graph can be approximated by those of the weighted expectation graph, with error bounds dependent upon the minimum and maximum expected degrees. In particular, we use these results to bound the spectra of random graphs with given expected degree sequences, including random power law graphs. Moreover, we prove a similar result giving concentration of the spectrum of a matrix martingale on its expectation.

1 Introduction

The spectra of random matrices and random graphs have been extensively studied in the literature (see, for example, [3], [4], [6], [8], [13]). We here focus on matrices with entries as independent random variables. Throughout, we will consider G to be a random graph, where $\text{pr}(v_i \sim v_j) = p_{ij}$, and each edge independent of each other edge.

For random graphs with such general distributions, we derive several bounds for the spectrum of the corresponding adjacency matrix and (normalized) Laplacian matrix (complete definitions and notation are in section 2). Eigenvalues of the adjacency matrix have many applications in graph theory, such as describing certain topological features of a graph, such as connectivity and enumerating the occurrences of subgraphs [6], [13]. Eigenvalues of the Laplacian matrix provide information about diffusion, and have many applications in studying random walks on graphs and approximation algorithms [6].

Before we proceed to examine eigenvalues of random graphs with given expected degree sequences and random power law graphs, we will first prove the following two general theorems. Previously, Oliveira [18] considered the same problem of approximating the

*Research supported in part by ONR MURI N000140810747, and AFOR's complex networks program.

spectra of the adjacency matrix and the Laplacian of random graphs. The following two theorems improve the results in [18] since the assumptions in our theorems are weaker. In addition, we improve the bound for the eigenvalues of the adjacency matrix by a factor of 2, and we improve those for the Laplacian by a factor of $7/\sqrt{3}$.

Theorem 1. *Let G be a random graph, where $\text{pr}(v_i \sim v_j) = p_{ij}$, and each edge is independent of each other edge. Let A be the adjacency matrix of G , so $A_{ij} = 1$ if $v_i \sim v_j$ and 0 otherwise, and $\bar{A} = E(A)$, so $\bar{A}_{ij} = p_{ij}$. Let Δ denote the maximum expected degree of G . Let $\epsilon > 0$, and suppose that for n sufficiently large, $\Delta > \frac{4}{9} \ln(2n/\epsilon)$. Then with probability at least $1 - \epsilon$, for n sufficiently large, the eigenvalues of A and $\bar{A}$ satisfy*

$$|\lambda_i(A) - \lambda_i(\bar{A})| \leq \sqrt{4\Delta \ln(2n/\epsilon)}$$

for all $1 \leq i \leq n$.

Theorem 2. *Let G be a random graph, where $\text{pr}(v_i \sim v_j) = p_{ij}$, and each edge is independent of each other edge. Let A be the adjacency matrix of G , as in Theorem 1. Let D be the diagonal matrix with $D_{ii} = \deg(v_i)$, and $\bar{D} = E(D)$. Let δ be the minimum expected degree of G , and $L = I - D^{-1/2}AD^{-1/2}$ the (normalized) Laplacian matrix for G . Choose $\epsilon > 0$. Then there exists a constant $k = k(\epsilon)$ such that if $\delta > k \ln n$, then with probability at least $1 - \epsilon$, the eigenvalues of L and $\bar{L}$ satisfy*

$$|\lambda_j(L) - \lambda_j(\bar{L})| \leq 3\sqrt{\frac{3 \ln(4n/\epsilon)}{\delta}}$$

for all $1 \leq j \leq n$, where $\bar{L} = I - \bar{D}^{-1/2}\bar{A}\bar{D}^{-1/2}$.

Here $\ln$ denotes the natural logarithm. We note that in these two theorems, the bound is simultaneously true for all eigenvalues with probability at least $1 - \epsilon$.

As an example, we apply these results to the $G(\mathbf{w})$ model, first introduced in [7], which produces a random graph with a specified expected degree sequence $\mathbf{w} = (w_1, w_2, \dots, w_n)$. The spectrum of the adjacency matrix of this model has been studied in [9]. In that paper, it is proven that if $m = w_{\max}$ is the maximum expected degree, then

$$\tilde{d} - \sqrt{2m^2\rho \ln n} \leq \lambda_{\max}(A) \leq \tilde{d} + \sqrt{6\sqrt{m \ln n}(\tilde{d} + \ln n)} + 3\sqrt{m \ln n}, \quad (1)$$

where $\rho = (\sum w_i)^{-1}$, and $\tilde{d} = \frac{\sum w_i^2}{\sum w_i}$ is the second-order average degree. Using Theorem 1, we prove the following:

Theorem 3. *For the random graph $G(\mathbf{w})$, if the maximum expected degree m satisfies $m > \frac{8}{9} \ln(\sqrt{2}n)$, then with probability at least $1 - 1/n = 1 - o(1)$, we have*

- The largest eigenvalue, $\lambda_{\max}(A)$, of the adjacency matrix of $G(\mathbf{w})$ satisfies

$$\tilde{d} - \sqrt{8m \ln(\sqrt{2}n)} \leq \lambda_{\max}(A) \leq \tilde{d} + \sqrt{8m \ln(\sqrt{2}n)}$$

- For all eigenvalues $\lambda_i(A) < \lambda_{\max}(A)$, we have

$$|\lambda_i(A)| \leq \sqrt{8m \ln(\sqrt{2}n)}$$

In particular, if $\tilde{d} \gg \sqrt{m \ln n}$, then $\lambda_1(A) = (1 + o(1))\tilde{d}$ a.a.s..

While the asymptotics in Theorem 3 for $\lambda_1(A)$ are the same as those in (1), the bounds in Theorem 3 are a significant improvement upon these. Moreover, [9] does not provide bounds for $\lambda_i(A)$ for $i > 1$ other than in the case that $\sqrt{m_i} \gg \tilde{d} \ln^2 n$.

For the Laplacian spectrum of $G(\mathbf{w})$, the best known bound for $\lambda_k(L) > \lambda_{\min}(L) = 0$ is given in [10]. If we take $\bar{w}$ to be the average expected degree, and $g(n)$ a function going to ∞ with n arbitrarily slowly, the result in [10] is that, for $w_{\min} \gg \ln^2 n$,

$$\max_k |1 - \lambda_k(L)| \leq (1 + o(1)) \frac{4}{\sqrt{\bar{w}}} + \frac{g(n) \ln^2 n}{w_{\min}}.$$

We not only improve upon this bound in Theorem 4, but we extend to the case that $w_{\min} \gg \ln n$, rather than $\ln^2 n$.

Theorem 4. *For the random graph $G(\mathbf{w})$, if the minimum expected degree $w_{\min}$ satisfies $w_{\min} \gg \ln n$, then with probability at least $1 - 1/n = 1 - o(1)$, we have that for all eigenvalues $\lambda_k(L) > \lambda_{\min}(L)$ of the Laplacian of $G(\mathbf{w})$,*

$$|\lambda_k(L) - 1| \leq 3 \sqrt{\frac{6 \ln(2n)}{w_{\min}}} = o(1).$$

Using the $G(\mathbf{w})$ model, we can also build random power law graphs in the following way. Given a power law exponent β , maximum degree m , and average degree d , we take $w_i = ci^{-\frac{1}{\beta-1}}$ for each i with $i_0 \leq i < n + i_0$. The values of c and i_0 depend upon β , m and d , in particular, $c = \frac{\beta-2}{\beta-1} dn^{\frac{1}{\beta-1}}$ and $i_0 = n \left(\frac{d(\beta-2)}{m(\beta-1)} \right)^{\beta-1}$. One can easily verify that the number of vertices of degree k is proportional to $k^{-\beta}$ (see [8], [9]). In section 4, we show how Theorems 3 and 4 can be applied in this setting to provide bounds on the spectra of random power law graphs. The remainder of this paper is organized as follows: In section 2, we develop notation and key tools to prove Theorems 1 and 2. Section 3 is devoted to the proofs of these two main theorems. Section 4 contains the details of the proofs of Theorems 3 and 4, applications of these theorems to random power law graphs, as well as a discussion of our main results as applied to $G_{n,p}$.

2 Key Tools

As to notation, throughout the paper, given an $n \times n$ Hermitian matrix A , $\|A\|$ denotes the spectral norm, so $\|A\| = \max |\lambda|$, where the maximum is over all eigenvalues λ of A . We order the eigenvalues $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$. Given two matrices A and B , we say $A \preceq B$

if $B - A$ is positive semidefinite. We refer to this ordering as the semidefinite order. If A is a random $n \times n$ matrix, we write $E(A)$ to denote the coordinate-wise expectation of A , so $E(A)_{ij} = E(A_{ij})$. Similarly, $\text{var}(A) = E((A - E(A))^2)$.

We shall use several applications of functions to matrices. In general, if f is a function with Taylor expansion $f(x) = \sum_{n=0}^{\infty} a_n x^n$, we take $f(A) = \sum_{n=0}^{\infty} a_n A^n$. We note that notions of convergence are as in [16]. In particular, we will often use the matrix exponential, $\exp(A) = \sum_{n=0}^{\infty} \frac{1}{n!} A^n$. We note that $\exp(A)$ is always positive definite when A is Hermitian, and that $\exp(A)$ converges for all choices of A . Moreover, we shall require brief use of the matrix logarithm. In general, if $B = \exp(A)$, we say that A is a logarithm of B . As our matrices will be Hermitian, it is sufficient for uniqueness of this function to require that the logarithm also be Hermitian. Any notation not mentioned here pertaining to matrices is as in [16].

Given a graph G , we will use A to denote the adjacency matrix for G , so $A_{ij} = 1$ if $v_i \sim v_j$ and 0 otherwise. We use D to denote the diagonal matrix with $D_{ii} = \deg_G(v_i)$. If G is a random graph, then $\bar{A}$ denotes the expectation of A , and $\bar{D}$ the expectation of D . The Laplacian of G is denoted by $L = I - D^{-1/2} A D^{-1/2}$, and $\bar{L} = I - \bar{D}^{-1/2} \bar{A} \bar{D}^{-1/2}$ is the Laplacian matrix for the weighted graph whose adjacency matrix is $\bar{A}$. All other notation referring to graph properties is as in [6].

We shall require the following concentration inequality in order to prove our main theorems. Previously, various matrix concentration inequalities have been derived by many authors including Ahlswede-Winter [2], Cristofides-Markström [12], Oliveira [18], Gross [15], Recht [19], and Tropp [20]. Here we give a short proof for a simple version that is particularly suitable for random graphs.

Theorem 5. *Let $X_1, X_2, \dots, X_m$ be independent random $n \times n$ Hermitian matrices. Moreover, assume that $\|X_i - E(X_i)\| \leq M$ for all i , and put $v^2 = \|\sum \text{var}(X_i)\|$. Let $X = \sum X_i$. Then for any $a > 0$,*

$$\text{pr}(\|X - E(X)\| > a) \leq 2n \exp\left(-\frac{a^2}{2v^2 + 2Ma/3}\right).$$

For the proof, we will rely on the following results:

Lemma 1 (see, for example, [20]). *Let $f, g : \mathbb{R} \rightarrow \mathbb{R}$, and suppose there is a subset $S \subseteq \mathbb{R}$ with $f(a) \leq g(a)$ for all $a \in S$. If A is a Hermitian matrix with all eigenvalues contained in S , then $f(A) \preceq g(A)$.*

Lemma 2 ([17]). *Given a fixed Hermitian matrix A , the function $X \mapsto \text{Tr}(\exp(A + \log X))$ is concave on the set of positive definite X .*

We note that any real-valued function that is convex with respect to the semidefinite order admits an operator Jensen's inequality (see, for example, [19]). That is to say, if f is convex with respect to the semidefinite order, then for a random matrix X , $f(E(X)) \leq E(f(X))$. Given a fixed matrix A and a random Hermitian matrix X , we may apply the function in Lemma 2 to e^X . By then applying the operator Jensen's inequality as stated, we obtain the following:

Lemma 3. *If A is a fixed matrix and X is a random Hermitian matrix, then*

$$\mathbb{E}(\text{Tr}(\exp(A + X))) \leq \text{Tr}(\exp[A + \log(\mathbb{E}[\exp X])]). \quad (2)$$

We shall use this result to overcome the difficulties presented by working with the semidefinite order, as opposed to real numbers. The primary problem that must be overcome is that unlike real numbers, the semidefinite order does not respect products.

Proof of Theorem 5. We assume for the sake of the proof that $\mathbb{E}(X_k) = 0$ for all k . Clearly this yields the general case by simply replacing each X_k by $X_k - \mathbb{E}(X_k)$.

Let $g(x) = \frac{2}{x^2}(e^x - x - 1) = 2 \sum_{k=2}^{\infty} \frac{x^{k-2}}{k!}$. Notice that g is increasing, so in particular, if $x \leq M$, $g(x) \leq g(M)$. Given $\theta > 0$, we have that $\|\theta X_k\| \leq \theta M$, and thus $g(\theta X_k) \preceq g(\theta M)I$ by Lemma 1. Therefore,

$$\mathbb{E}(e^{\theta X_k}) = \mathbb{E}(I + \theta X_k + \frac{1}{2}\theta^2 X_k^2 g(\theta X_k)) \quad (3)$$

$$\preceq I + \frac{1}{2}g(\theta M)\theta^2 \mathbb{E}(X_k^2) \quad (4)$$

$$\preceq e^{\frac{1}{2}g(\theta M)\theta^2 \mathbb{E}(X_k^2)}. \quad (5)$$

We use this to prove the following claim:

Claim 1: For matrices X_k as given,

$$\mathbb{E} \left[\text{Tr}(\exp \left(\sum_{k=1}^m \theta X_k \right)) \right] \leq \text{Tr}(\exp \left(\sum_{k=1}^m \frac{1}{2}g(\theta M)\theta^2 \mathbb{E}(X_k^2) \right)) \quad (6)$$

Proof of Claim 1: For a given k , let $\mathbb{E}_k(\cdot) := \mathbb{E}(\cdot | X_1, X_2, \dots, X_k)$. Then we have

$$\mathbb{E} \left[\text{Tr}(\exp \left(\sum_{k=1}^m \theta X_k \right)) \right] = \mathbb{E} \mathbb{E}_1 \mathbb{E}_2 \dots \mathbb{E}_{m-1} \left[\text{Tr}(\exp \left(\sum_{k=1}^{m-1} \theta X_k + \theta X_m \right)) \right]$$

As the X_i are independent, each X_k is fixed with respect to $\mathbb{E}_{m-1}$ except X_m , and $\mathbb{E}_{m-1}(\exp X_m) = \mathbb{E}(\exp X_m)$. Applying inequality (2) from Lemma 3, we have

$$\begin{aligned} \mathbb{E} \mathbb{E}_1 \mathbb{E}_2 \dots \mathbb{E}_{m-1} \left[\text{Tr}(\exp \left(\sum_{k=1}^{m-1} \theta X_k + \theta X_m \right)) \right] &\leq \\ \mathbb{E} \mathbb{E}_1 \mathbb{E}_2 \dots \mathbb{E}_{m-2} \left[\text{Tr}(\exp \left(\sum_{k=1}^{m-1} \theta X_k + \log \mathbb{E}(\exp(\theta X_m)) \right)) \right]. \end{aligned}$$

Iteratively applying this process, we obtain

$$\mathbb{E} \left[\text{Tr}(\exp \left(\sum_{k=1}^m \theta X_k \right)) \right] \leq \text{Tr}(\exp \left(\sum_{k=1}^m \log \mathbb{E}(\exp \theta X_m) \right)).$$

As both $\log(\cdot)$ and $\text{Tr}(\exp(\cdot))$ are monotone with respect to the semidefinite order (these facts can be easily proven with basic manipulations), inequality (5) implies that

$$\begin{aligned} \mathbb{E} \left[\text{Tr}(\exp(\sum_{k=1}^m \theta X_k)) \right] &\leq \text{Tr}(\exp(\sum_{k=1}^m \log e^{\frac{1}{2}g(\theta M)\theta^2 \mathbb{E}(X_k^2)})) \\ &\leq \text{Tr}(\exp(\sum_{k=1}^m \frac{1}{2}g(\theta M)\theta^2 \mathbb{E}(X_k^2))), \end{aligned}$$

as desired. $\square$

Now, given $a > 0$, for all $\theta > 0$ we have

$$\begin{aligned} \text{pr}(\lambda_{\max}(X) \geq a) &\leq e^{-\theta a} \mathbb{E}(\exp(\theta \lambda_{\max}(X))) \\ &\leq e^{-\theta a} \mathbb{E}(\text{Tr}(\exp(\theta X))) \\ &= e^{-\theta a} \mathbb{E} \left[\text{Tr}(\exp(\sum \theta X_k)) \right] \\ &\leq e^{-\theta a} \text{Tr}(\exp(\sum \frac{1}{2}g(\theta M)\theta^2 \mathbb{E}(X_k^2))) \\ &\leq e^{-\theta a} n \lambda_{\max}(\exp(\frac{1}{2}g(\theta M)\theta^2 \sum \mathbb{E}(X_k^2))) \\ &\leq n \exp\left(-\theta a + \frac{1}{2}g(\theta M)\theta^2 v^2\right), \end{aligned}$$

as $v^2 = \|\sum \mathbb{E}(X_k^2)\| \geq \lambda_{\max}(\sum \mathbb{E}(X_k^2))$

Notice that if $x < 3$, we have $g(x) = 2 \sum_{k=2}^{\infty} \frac{x^{k-2}}{k!} \leq \sum_{k=2}^{\infty} \frac{x^{k-2}}{3^{k-2}} = \frac{1}{1-x/3}$. Take $\theta = \frac{a}{v^2 + Ma/3}$. Clearly, $\theta M \leq 3$, and thus we have

$$\begin{aligned} \text{pr}(\lambda_{\max}(X) \geq a) &\leq n \exp\left(-\theta a + \frac{1}{2}g(\theta M)\theta^2 v^2\right) \\ &\leq n \exp\left(-\frac{a^2}{2v^2 + 2Ma/3}\right) \end{aligned}$$

Therefore, $\text{pr}(\|X\| \geq a) \leq 2n \exp\left(-\frac{a^2}{2v^2 + 2Ma/3}\right)$. $\square$

3 Proofs of the Main Theorems

In this section, we provide proofs of the two main theorems using Theorem 5.

Proof of Theorem 1. Let G be a random graph as described in the statement of the theorem, where the edge $v_i v_j$ appears with probability p_{ij} .

Given $1 \leq i, j \leq n$, let A^{ij} be the matrix with a 1 in the ij and ji positions and a 0 everywhere else. Let $h_{ij} = 1$ with probability p_{ij} and 0 otherwise. Take $X_{ij} = h_{ij} A^{ij}$, so

$A = \sum X_{ij}$. Thus, we can apply Theorem 5 to A with $M = 1$. We need first to calculate v^2 .

Now, if $i \neq j$, then

$$\begin{aligned}\text{var}(X_{ij}) &= \mathbb{E}((h_{ij} - p_{ij})^2(A^{ij})^2) \\ &= \text{var}(h_{ij})(A^{ii} + A^{jj}) \\ &= p_{ij}(1 - p_{ij})(A^{ii} + A^{jj})\end{aligned}$$

Similarly, $\text{var}(X_{ii}) = p_{ii}(1 - p_{ii})A^{ii}$. Therefore,

$$\begin{aligned}\left\| \sum \text{var}(X_{ij}) \right\| &= \left\| \sum_{i=1}^n \left(\sum_{j=1}^n p_{ij}(1 - p_{ij}) \right) A^{ii} \right\| \\ &= \max_{i=1, \dots, n} \sum_{j=1}^n p_{ij}(1 - p_{ij}) \\ &\leq \max_{i=1, \dots, n} \sum_{j=1}^n p_{ij} = \Delta.\end{aligned}$$

Take $a = \sqrt{4\Delta \ln(2n/\epsilon)}$. By the assumption on Δ , we have $a < 3\Delta$, and thus we obtain

$$\begin{aligned}\text{pr}(\|A - \bar{A}\| > a) &\leq 2n \exp\left(-\frac{a^2}{2v^2 + 2Ma/3}\right) \\ &\leq 2n \exp\left(\frac{-4\Delta \ln(2n/\epsilon)}{4\Delta}\right) \\ &= \epsilon.\end{aligned}$$

To complete the proof, we recall Weyl's Theorem (see, for example, [16]), which states that for Hermitian matrices M and N , $\max_k |\lambda_k(M) - \lambda_k(N)| \leq \|M - N\|$. Thus, with probability at least $1 - \epsilon$, we have that for all $1 \leq i \leq n$, $|\lambda_i(A) - \lambda_i(\bar{A})| < \sqrt{4\Delta \ln(2n/\epsilon)}$. $\square$

Proof of Theorem 2. We will again use Weyl's Theorem, as in the proof of Theorem 1, so we need only bound $\|L - \bar{L}\|$. For each vertex v_i , put $d_i = \deg(v_i)$ and $t_i = \mathbb{E}(d_i)$, the expected degree of the i th vertex. Let $C = I - \bar{D}^{-1/2}A\bar{D}^{-1/2}$. Then $\|L - \bar{L}\| \leq \|C - \bar{L}\| + \|L - C\|$. We consider each term separately.

Now, $C - \bar{L} = \bar{D}^{-1/2}(A - \bar{A})\bar{D}^{-1/2}$. Using notation as in the proof of Theorem 1, let

$$\begin{aligned}Y_{ij} &= \bar{D}^{-1/2}((h_{ij} - p_{ij})A^{ij})\bar{D}^{-1/2} \\ &= \frac{h_{ij} - p_{ij}}{\sqrt{t_i t_j}} A^{ij}\end{aligned}$$

Then $C - \bar{L} = \sum Y_{ij}$, so we can apply Theorem 5 to bound $\|C - \bar{L}\|$. Notice $\|Y_{ij}\| \leq (t_i t_j)^{-1/2} \leq \frac{1}{\delta}$. Moreover,

$$\mathbb{E}(Y_{ij}^2) = \begin{cases} \frac{1}{t_i t_j} (p_{ij} - p_{ij}^2) (A^{ii} + A^{jj}) & i \neq j \\ \frac{1}{t_i^2} (p_{ii} - p_{ii}^2) A^{ii} & i = j \end{cases}$$

Thus, we obtain

$$\begin{aligned} v^2 = \left\| \sum \mathbb{E}(Y_{ij}^2) \right\| &= \left\| \sum_{i=1}^n \sum_{j=1}^n \frac{1}{t_i t_j} (p_{ij} - p_{ij}^2) A^{ii} \right\| \\ &= \max_{i=1, \dots, n} \left(\sum_{j=1}^n \frac{1}{t_i t_j} p_{ij} - \sum_{j=1}^n \frac{1}{t_i t_j} p_{ij}^2 \right) \\ &\leq \max_{i=1, \dots, n} \left(\frac{1}{\delta} \sum_{j=1}^n \frac{p_{ij}}{t_i} \right) = \frac{1}{\delta} \end{aligned}$$

Take $a = \sqrt{\frac{3 \ln(4n/\epsilon)}{\delta}}$. Take k to be large enough so that $\delta > k \ln n$ implies $a < 1$ (in particular, choosing $k > 3(1 + \ln(\epsilon/4))$ is sufficient). Applying Theorem 5, we have

$$\begin{aligned} \text{pr}(\|C - \bar{L}\| > a) &\leq 2n \exp \left(-\frac{\frac{3 \ln(4n/\epsilon)}{\delta}}{2/\delta + 2a/(3\delta)} \right) \\ &\leq 2n \exp \left(-\frac{3 \ln(4n/\epsilon)}{3} \right) \\ &\leq \epsilon/2 \end{aligned}$$

For the second term, note that by the Chernoff bound (see, for example, [1]), for each i ,

$$\text{pr}(|d_i - t_i| > b t_i) \leq \frac{\epsilon}{2n} \text{ if } b \geq \sqrt{\frac{\ln(4n/\epsilon)}{t_i}}$$

Take $b = \sqrt{\frac{\ln(4n/\epsilon)}{\delta}}$, so that for all i , we have $\text{pr}(|d_i - t_i| > b t_i) \leq \frac{\epsilon}{2n}$. Then we obtain

$$\|\bar{D}^{-1/2} D^{1/2} - I\| = \max_{i=1, \dots, n} \left| \sqrt{\frac{d_i}{t_i}} - 1 \right|.$$

Note that for $0 < x < 1$, we have $|\sqrt{x} - 1| \leq |x - 1|$. Taking $x = \frac{d_i}{t_i} > 0$, we have that with probability at least $1 - \frac{\epsilon}{2}$, this is at most $b = \sqrt{\frac{\ln(4n/\epsilon)}{\delta}} = \frac{1}{\sqrt{3}} a < 1$ for all i . Thus we obtain

$$\|\bar{D}^{-1/2} D^{1/2} - I\| = \max_{i=1, \dots, n} \left| \sqrt{\frac{d_i}{t_i}} - 1 \right| \leq \sqrt{\frac{\ln(4n/\epsilon)}{\delta}}$$

with probability at least $1 - \frac{\epsilon}{2}$.

We note that as the Laplacian spectrum is in $[0, 2]$, we have $\|I - L\| \leq 1$. Therefore, with probability at least $1 - \frac{\epsilon}{2}$, we have

$$\begin{aligned}
\|L - C\| &= \|I - D^{-1/2}AD^{-1/2} - I + \bar{D}^{-1/2}A\bar{D}^{-1/2}\| \\
&= \|D^{-1/2}AD^{-1/2} - \bar{D}^{-1/2}D^{1/2}D^{-1/2}AD^{-1/2}D^{1/2}\bar{D}^{-1/2}\| \\
&= \|(I - L) - (\bar{D}^{-1/2}D^{1/2})(I - L)(D^{1/2}\bar{D}^{-1/2})\| \\
&= \|(\bar{D}^{-1/2}D^{1/2} - I)(I - L)D^{1/2}\bar{D}^{-1/2} + (I - L)(I - D^{1/2}\bar{D}^{-1/2})\| \\
&\leq \|\bar{D}^{-1/2}D^{1/2} - I\| \|D^{1/2}\bar{D}^{-1/2}\| + \|I - D^{1/2}\bar{D}^{-1/2}\| \\
&\leq b(b + 1) + b = b^2 + 2b
\end{aligned}$$

Finally, as $b = \frac{1}{\sqrt{3}}a$ and $a < 1$, we have that with probability at least $1 - \epsilon$,

$$\begin{aligned}
\|L - \bar{L}\| &\leq \|C - \bar{L}\| + \|L - C\| \\
&\leq a + \frac{1}{3}a^2 + \frac{2a}{\sqrt{3}} \leq 3a,
\end{aligned}$$

completing the proof. $\square$

4 Applications to Several Graph Models

The above theorems apply in a very general random graph setting. Here we discuss the applications of Theorems 1 and 2 for the Erdős-Rényi graph and for the $G(\mathbf{w})$ model, as discussed in section 1 above. We begin by examining the Erdős-Rényi graph.

The Erdős-Rényi graph is a well studied random graph (see, for example, [1], [5]) where $p_{ij} = p$ for $i \neq j$ and 0 for $i = j$ for some $p \in (0, 1)$. We denote this graph by $G_{n,p}$. If J represents the $n \times n$ matrix with a 1 in every entry, then for $G_{n,p}$, $\bar{A} = p(J - I)$ and $\bar{D} = (n - 1)pI$. An application of Theorem 1 yields

Theorem 6. *For $G_{n,p}$, if $p > \frac{8}{9n} \ln(\sqrt{2}n)$, then with probability at least $1 - 1/n = 1 - o(1)$, we have*

$$|\lambda_i(A) - \lambda_i(p(J - I))| \leq \sqrt{8np \ln(\sqrt{2}n)}.$$

We note that stronger results for the spectrum of the adjacency matrix of $G_{n,p}$ can be found in [21], [14]. Specifically, in [14], it is shown that for $pn \geq c \ln n$, $\lambda_1(A) = pn + O(\sqrt{pn})$, and all other eigenvalues satisfy $|\lambda_i(A)| = O(\sqrt{pn})$. Here, we are at best able to show that $|\lambda_i(A) - \lambda_i(p(J - I))| \leq O(pn)$. The spectrum of $p(J - I)$ is $\{p(n - 1), -p\}$, where $-p$ has multiplicity $n - 1$, so if $i > 1$, we have only that $\lambda_i(A) = O(pn)$. However, due to the very strong symmetries in $G_{n,p}$, it seems unlikely that the methods used to investigate this graph in detail will extend to general random graphs.

For the Laplacian of $G_{n,p}$, we obtain $\bar{L} = I - \frac{1}{n-1}(J - I)$. An application of Theorem 2 yields

Theorem 7. *If $pn \gg \ln n$, then with probability at least $1 - 1/n = 1 - o(1)$, we have*

$$|\lambda_k(L) - \lambda_k(I - \frac{1}{n-1}(J - I))| \leq 3\sqrt{\frac{6 \ln(2n)}{pn}} = o(1)$$

for all $1 \leq k \leq n$.

The spectrum of $I - \frac{1}{n-1}(J - I)$ is $\{\frac{1}{n-1}, 1 + \frac{1}{n-1}\}$, where $1 + \frac{1}{n-1}$ has multiplicity $n - 1$. Thus, we see that if $pn \gg \ln n$, then w.h.p. L has all eigenvalues other than $\lambda_{\min}(L)$ close to 1. This result is not new (see [10], [11]), and [11] also considers the case where $pn \leq \ln n$.

We now turn to the $G(\mathbf{w})$ model. We begin by precisely defining the model. Given a sequence $\mathbf{w} = (w_1, w_2, \dots, w_n)$, we define the random graph $G(\mathbf{w})$ to have vertex set $\{v_1, \dots, v_n\}$, and edges are independently assigned to each pair (v_i, v_j) with probability $w_i w_j \rho$, where $\rho = \frac{1}{\sum w_i}$. In this way, the expected degree of v_i is w_i for each i . Moreover, the matrix $\bar{A}$ with $\bar{A}_{ij} = \text{pr}(v_i \sim v_j)$ is given by $\bar{A} = \rho \mathbf{w}' \mathbf{w}$, and as such has eigenvalues $\rho \sum w_i^2$ and 0, where 0 has multiplicity $n - 1$. Let $\tilde{d}$ denote the expected second order average degree of $G(\mathbf{w})$, so $\tilde{d} = \frac{\sum w_i^2}{\sum w_i} = \rho \sum w_i^2$. Applying Theorem 1 with $\epsilon = 1/n$, we immediately obtain Theorem 3.

Similarly, we can apply Theorem 2 to obtain concentration results for the spectrum of the Laplacian matrix for $G(\mathbf{w})$. Notice that the value of k given in Theorem 1 will here be $k(\epsilon) = k(1/n) > 3(1 + \ln(1/(4n)))$, so the requirement in Theorem 4 that $w_{\min} \gg \ln n$ is sufficient to give that for n sufficiently large, $w_{\min} > k \ln n$. This theorem improves on the bounds for the eigenvalues of the Laplacian given in [10], as seen in section 1 above.

Let $\mathbf{x} = (w_1^{1/2}, w_2^{1/2}, \dots, w_n^{1/2})$. Then we have

$$(\bar{D}^{-1/2} \bar{A} \bar{D}^{-1/2})_{ij} = \frac{\rho w_i w_j}{\sqrt{w_i w_j}} = \rho \sqrt{w_i w_j},$$

so $\bar{L} = I - \rho \mathbf{x}' \mathbf{x}$. Thus, the eigenvalues of $\bar{L}$ are 0 and 1, where 1 has multiplicity $n - 1$. Applying Theorem 2 with $\epsilon = 1/n$, we obtain Theorem 4.

Recall, as described in section 1, we can use the $G(\mathbf{w})$ model to build random power law graphs. Given a power law exponent β , maximum degree m , and average degree d , we take $w_i = ci^{-\frac{1}{\beta-1}}$ for each i with $i_0 \leq i < n + i_0$, with $c = \frac{\beta-2}{\beta-1} dn^{\frac{1}{\beta-1}}$ and $i_0 = n \left(\frac{d(\beta-2)}{m(\beta-1)} \right)^{\beta-1}$. We obtain the following bounds for $\tilde{d}$ (see [9]):

$$\tilde{d} = \begin{cases} d \frac{(\beta-2)^2}{(\beta-1)(\beta-3)} (1 + o(1)) & \text{if } \beta > 3 \\ \frac{1}{2} d \ln \frac{2m}{d} (1 + o(1)) & \text{if } \beta = 3 \\ d \frac{(\beta-2)^2}{(\beta-1)(3-\beta)} \left(\frac{m(\beta-1)}{d(\beta-2)} \right)^{3-\beta} (1 + o(1)) & \text{if } 2 < \beta < 3 \end{cases} \quad (7)$$

In [10], bounds are given for the largest eigenvalue of the adjacency matrix of a random power law graph as described. In particular, the authors show that for a random power law graph as above with adjacency matrix A ,

1. If $\beta \geq 3$ and $m > d^2 \log^{3+\epsilon} n$, then a.a.s. $\lambda_1(A) = (1 + o(1))\sqrt{m}$.
2. If $2.5 < \beta < 3$ and $m > d^{\frac{\beta-2}{\beta-2.5}} \ln^{\frac{3}{\beta-2.5}} n$, then a.a.s. $\lambda_1(A) = (1 + o(1))\sqrt{m}$.
3. If $2 < \beta < 2.5$ and $m > \ln^{\frac{3}{2.5-\beta}} n$, then a.a.s. $\lambda_1(A) = (1 + o(1))\tilde{d}$.

We note that these theorems require specific relationships between m and d , as noted. Applying Theorem 3 to a random power law graph, we can eliminate such requirements, although the bounds are less clean:

Theorem 8. *Suppose $G(\mathbf{w})$ is a random power law graph as described above. If $m > \frac{8}{9} \ln(\sqrt{2}n)$ then with probability $1 - o(1)$, we have*

- If $\beta > 3$, then

$$\left| \lambda_1(A) - d \frac{(\beta-2)^2}{(\beta-1)(\beta-3)} (1 + o(1)) \right| \leq \sqrt{8m \ln(\sqrt{2}n)}$$

- If $\beta = 3$, then

$$\left| \lambda_1(A) - \frac{1}{2} d \ln \frac{2m}{d} (1 + o(1)) \right| \leq \sqrt{8m \ln(\sqrt{2}n)}$$

- If $2 < \beta < 3$, then

$$\left| \lambda_1(A) - d \frac{(\beta-2)^2}{(\beta-1)(3-\beta)} \left(\frac{m(\beta-1)}{d(\beta-2)} \right)^{3-\beta} (1 + o(1)) \right| \leq \sqrt{8m \ln(\sqrt{2}n)}$$

From these bounds, one might be able to derive specific relationships between m and d that lead to particular values for $\lambda_1(A)$.

Finally, we can apply Theorem 4 to the random power law graph model described here to obtain bounds on the Laplacian spectrum:

Theorem 9. *Suppose $G(\mathbf{w})$ is a random power law graph as described above. If $w_{\min} = c_0^{-\frac{1}{\beta-1}} = mn^{\frac{\beta}{1-\beta}} \gg \ln n$, then with probability $1 - o(1)$, we have that for $\lambda_k(L) > \lambda_{\min}(L)$,*

$$|\lambda_k(L) - 1| \leq 3 \sqrt{\frac{6 \ln(2n)}{mn^{\frac{\beta}{1-\beta}}}} = o(1).$$

5 Concentration of Matrix Martingales

In the proofs of Theorem 1 and 2, the key tool is Theorem 5, a result giving concentration in norm of sums of independent random Hermitian matrices on their expectation. As mentioned in Section 2, these kinds of results have been derived and studied by various authors. Here we provide a generalization of Theorem 5 in the event that the random matrices are not independent, using matrix martingales.

A *matrix martingale* is a series of random $n \times n$ matrices $X_1, X_2, \dots, X_m$, together with a sequence of σ -fields $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \dots \subset \mathcal{F}_{m-1}$, called a *filtration*, such that

1. X_k is $\mathcal{F}_k$ measurable for all k .
2. for all k , $E(X_k|\mathcal{F}_{k-1}) = X_{k-1}$, where $X_0 := E(X_1)$.

We say the martingale $X_1, X_2, \dots, X_m$ is *adapted* to the filtration $\mathcal{F}_0, \mathcal{F}_1, \dots, \mathcal{F}_{m-1}$. We often consider the differences between consecutive matrices $Y_k = X_k - X_{k-1}$. The sequence of such Y_k is called the *associated difference sequence* to the martingale.

Theorem 10. *Let $X_1, X_2, \dots, X_m$ be a martingale of Hermitian matrices adapted to the filtration $\mathcal{F}_0, \mathcal{F}_1, \dots, \mathcal{F}_{m-1}$, so that $E(X_k|\mathcal{F}_{k-1}) = X_{k-1}$. Let $\{Y_k\}$ be the associated difference sequence. Moreover, suppose*

1. $\|Y_i\| \leq M$ for all i
2. $Y_i^2 \preceq A_i^2$, where each A_i is a fixed Hermitian matrix.

Take $c^2 = \|\sum A_i^2\|$. Then for all $a > 0$,

$$\text{pr}(\|X_m - E(X_m)\| > a) \leq 2n \exp\left(-\frac{a^2}{2c^2 + 2Ma/3}\right).$$

We note that the key difference between Theorem 10 and Theorem 5 is the use of the second moment. In the case that the summands are independent, as in Theorem 5, we can use the variance directly. However, for Theorem 10, the use of conditional expectation in the proofs requires a slightly stronger constant. We bound the “worst-case scenario,” that is, the largest possible spectrum of Y_i , represented by A_i , and in that sense bound the second moments without having taken expectations.

Proof. Take $g(x)$ to be as in the proof of Theorem 5. Note that for each $k > 0$,

$$\begin{aligned} E(\exp(\theta Y_k)|\mathcal{F}_{k-1}) &= E\left(I + \theta Y_k + \frac{1}{2}\theta^2 Y_k^2 g(\theta Y_k) \mid \mathcal{F}_{k-1}\right) \\ &= I + E\left(\frac{1}{2}\theta^2 Y_k^2 g(\theta Y_k) \mid \mathcal{F}_{k-1}\right) \end{aligned}$$

Note that the second equality follows as $E(Y_k|\mathcal{F}_{k-1}) = E(X_k - X_{k-1}|\mathcal{F}_{k-1}) = 0$. As in Theorem 5, $g(\theta Y_k) \preceq g(\theta M)I$. As I commutes with all other matrices, we have $Y_k^2 g(\theta Y_k) \preceq Y_k^2 g(\theta M)I \preceq g(\theta M)A_k^2$.

Therefore, we obtain

$$E(\exp(\theta Y_k)|\mathcal{F}_{k-1}) \preceq I + E\left(\frac{1}{2}\theta^2 g(\theta M)A_k^2 \mid \mathcal{F}_{k-1}\right) \quad (8)$$

$$= I + \frac{1}{2}\theta^2 g(\theta M)A_k^2 \quad (9)$$

$$\preceq e^{\frac{1}{2}\theta^2 g(\theta M)A_k^2} \quad (10)$$

We proceed with a claim similar to that in Theorem 5.

Claim 2: For matrices Y_k as given, and any $\theta > 0$,

$$\mathbb{E} \left[\text{Tr}(\exp \left(\sum_{k=1}^m \theta Y_k \right)) \right] \leq \text{Tr}(\exp \left(\sum_{k=1}^m \frac{1}{2} g(\theta M) \theta^2 A_k^2 \right)) \quad (11)$$

Proof of Claim 2: For each $k \geq 1$, let $\mathbb{E}_k(\cdot) = \mathbb{E}(\cdot | \mathcal{F}_k)$. Then we obtain

$$\mathbb{E} \left[\text{Tr}(\exp \left(\sum_{k=1}^m \theta X_k \right)) \right] = \mathbb{E} \mathbb{E}_1 \mathbb{E}_2 \dots \mathbb{E}_{m-1} \left[\text{Tr}(\exp \left(\sum_{k=1}^{m-1} \theta X_k + \theta X_m \right)) \right]$$

Note that each X_k is fixed with respect to $\mathbb{E}_{m-1}$ with the exception of X_m . Applying inequality (2) from Lemma 3, we have

$$\begin{aligned} \mathbb{E} \mathbb{E}_1 \mathbb{E}_2 \dots \mathbb{E}_{m-1} \left[\text{Tr}(\exp \left(\sum_{k=1}^{m-1} \theta X_k + \theta X_m \right)) \right] &\leq \\ \mathbb{E} \mathbb{E}_1 \mathbb{E}_2 \dots \mathbb{E}_{m-2} \left[\text{Tr}(\exp \left(\sum_{k=1}^{m-1} \theta X_k + \log \mathbb{E}(\exp(\theta X_m) | \mathcal{F}_{m-1}) \right)) \right]. \end{aligned}$$

Now, as $\log(\cdot)$, $\text{Tr}(\exp(\cdot))$, and $\mathbb{E}(\cdot)$ are monotone with respect to the semidefinite order, we can apply inequality 10 to obtain

$$\begin{aligned} \mathbb{E} \mathbb{E}_1 \mathbb{E}_2 \dots \mathbb{E}_{m-1} \left[\text{Tr}(\exp \left(\sum_{k=1}^{m-1} \theta X_k + \theta X_m \right)) \right] &\leq \\ \mathbb{E} \mathbb{E}_1 \mathbb{E}_2 \dots \mathbb{E}_{m-2} \left[\text{Tr}(\exp \left(\sum_{k=1}^{m-1} \theta X_k + \frac{1}{2} \theta^2 g(\theta M) A_m^2 \right)) \right]. \end{aligned}$$

Noting that this last term is fixed, we may apply this process iteratively to obtain the desired result. $\square$

Therefore, given $a > 0$, for all $\theta > 0$ we have

$$\begin{aligned} \text{pr}(\lambda_{\max}(X_m - \mathbb{E}(X_m)) \geq a) &= \text{pr}(\lambda_{\max} \left(\sum_{k=1}^m Y_k \right) \geq a) \\ &\leq e^{-\theta a} \mathbb{E}(\exp(\theta \lambda_{\max}(\sum_{k=1}^m Y_k))) \\ &\leq e^{-\theta a} \mathbb{E}(\text{Tr}(\exp(\theta \sum_{k=1}^m Y_k))) \\ &= e^{-\theta a} \mathbb{E} \left[\text{Tr}(\exp \left(\sum_{k=1}^m \theta Y_k \right)) \right] \\ &\leq e^{-\theta a} \text{Tr}(\exp \left(\sum_{k=1}^m \frac{1}{2} g(\theta M) \theta^2 A_k^2 \right)) \\ &\leq e^{-\theta a} n \lambda_{\max}(\exp \left(\frac{1}{2} g(\theta M) \theta^2 \sum_{k=1}^m A_k^2 \right)) \\ &\leq n \exp \left(-\theta a + \frac{1}{2} g(\theta M) \theta^2 c^2 \right). \end{aligned}$$

Take $\theta = \frac{a}{c^2 + Ma/3}$ and proceed as in the proof of Theorem 5 to yield the result. $\square$

References

- [1] N. Alon and J. Spencer, *The Probabilistic Method* (3rd edition), John Wiley & Sons, 2008.
- [2] R. Ahlswede and A. Winter, Strong converse for identification via quantum channels, *IEEE Trans. Inform. Theory*, **48** (3), (2002), 569–579.
- [3] G. W. Anderson, A. Guionnet, and O. Zeitouni, *An Introduction to Random Matrices*, Cambridge University Press, 2010.
- [4] M. Aouchiche and P. Hansen, A survey of automated conjectures in spectral graph theory, *Linear Algebra Appl.* **432**, (2010), no. 9, 2293–2322.
- [5] B. Bollobás, *Random Graphs* (2nd edition), Cambridge University Press, 2001.
- [6] F. Chung, *Spectral Graph Theory*, AMS Publications, 1997.
- [7] F. Chung and L. Lu, Connected components in random graphs with given expected degree sequences. *Ann. Combin.* **6**, (2002) 125–145.
- [8] F. Chung and L. Lu, *Complex Graphs and Networks*, AMS Publications, 2006.
- [9] F. Chung, L. Lu, and V. Vu, Eigenvalues of random power law graphs. *Ann. Combin.* **7**, (2003) 21–33.
- [10] F. Chung, L. Lu, and V. Vu, The spectra of random graphs with given expected degrees. *Internet Math.* **1**, (2004) 257–275.
- [11] A. Coja-Oghlan, On the laplacian eigenvalues of $G_{n,p}$. *Comb. Prob. Comp.* **16**, (2007) 923–946.
- [12] D. Cristofides and K. Markström, Expansion properties of random Cayley graphs and vertex transitive graphs via matrix martingales, *Random Structures Algorithms*, **32**, (2008), 88–100.
- [13] D. M. Cvetković, M. Doob, and H. Sachs, *Spectra of Graphs, Theory and Applications*, Academic Press, 1980.
- [14] U. Feige and E. Ofek, Spectral techniques applied to sparse random graphs. *Random Structures Algorithms*, **27**, (2005) 251–275.
- [15] D. Gross, Recovering low-rank matrices from few coefficients in any basis, *IEEE Trans. Inform. Theory*, **57**, (2011) 1548–1566.
- [16] R. Horn and C. Johnson, *Matrix Analysis*, Cambridge University Press, 1985.
- [17] E. H. Lieb, Convex trace functions and the Wigner-Yanase-Dyson conjecture. *Adv. Math.* **11**, (1973) 267–288.
- [18] R. Oliveira, Concentration of the adjacency matrix and of the Laplacian in random graphs with independent edges, <http://arxiv.org/abs/0911.0600>
- [19] B. Recht, Simpler approach to matrix completion, *J. Mach. Learn. Res.*, to appear.
- [20] J. Tropp, User-Friendly Tail Bounds for Sums of Random Matrices. Available at <http://arxiv.org/abs/1004.4389>
- [21] Z. Füredi and J. Komlós, The eigenvalues of random symmetric matrices. *Combinatorica* **1**, (1981), no. 3, 233–241.