Empirical Distributions of Laplacian Matrices of Large Dilute Random Graphs

TIEFENG JIANG[1]

Abstract We study the spectral properties of the Laplacian matrices and the normalized Laplacian matrices of the Erdös-Rényi random graph $G(n, p_n)$ for large n. Although the graph is simple, we discover some interesting behaviors of the two Laplacian matrices. In fact, under the dilute case, that is, $p_n \in (0, 1)$ and $np_n \to \infty$, we prove that the empirical distribution of the eigenvalues of the Laplacian matrix converges to a deterministic distribution, which is the free convolution of the semi-circle law and $N(0, 1)$. However, for its normalized version, we prove that the empirical distribution converges to the semi-circle law.

1 Introduction

In graph theory, the Erdös-Rényi model $G(n, p)$, named for Erdös-Rényi [17, 18, 19, 20], is a graph with n vertices; a potential edge between each pair of vertices has probability p, independently of the other edges. Some properties of $G(n, p)$ such as the diameters, the sizes and the giant components are known, see, e.g., [5, 9, 17, 18, 19, 20] for details. For the spectral properties of the graphs, one can see, e.g., [9, 10, 11, 15, 27, 31].

The graph $G(n, p_n)$ corresponds to Bernoulli random variables $\{\xi_{ij}^{(n)}; 1 \le i < j \le n\}$, which are independent random variables with $\xi_{ij}^{(n)} = \xi_{ji}^{(n)}$ and $P(\xi_{ij}^{(n)} = 1) = 1 - P(\xi_{ij}^{(n)} = 0) = p_n$ for all $1 \le i < j \le n$. The adjacency matrix $\mathbf{A}_n$ and Laplacian matrix $\mathbf{\Delta}_n$ are defined as follows:

$$\mathbf{A}_n = \begin{pmatrix} 0 & \xi_{12}^{(n)} & \cdots & \xi_{1n}^{(n)} \\ \xi_{21}^{(n)} & 0 & \cdots & \xi_{2n}^{(n)} \\ \vdots & \vdots & \vdots & \vdots \\ \xi_{n1}^{(n)} & \xi_{n2}^{(n)} & \cdots & 0 \end{pmatrix} \tag{1.1}$$

and

$$\mathbf{\Delta}_n = \begin{pmatrix} \sum_{j\neq 1} \xi_{1j}^{(n)} & -\xi_{12}^{(n)} & \cdots & -\xi_{1n}^{(n)} \\ -\xi_{21}^{(n)} & \sum_{j\neq 2} \xi_{2j}^{(n)} & \cdots & -\xi_{2n}^{(n)} \\ \vdots & \vdots & \vdots & \vdots \\ -\xi_{n1}^{(n)} & -\xi_{n2}^{(n)} & \cdots & \sum_{j\neq n} \xi_{nj}^{(n)} \end{pmatrix}. \tag{1.2}$$

[1]Supported in part by NSF #DMS-0449365, School of Statistics, University of Minnesota, 224 Church Street, MN55455, tjiang@stat.umn.edu.

Key Words: random matrix, random graph, dilute graph, Laplacian matrix, normalized Laplacian matrix, spectral distribution, semi-circle law, free convolution.

AMS (2000) subject classifications: 05C80, 05C50, 15A52, 60B10.

The normalized Laplacian matrix is defined by

$$\mathbf{L}_n = \mathbf{I}_n - \mathbf{D}_n^{-1/2}\mathbf{A}_n\mathbf{D}_n^{-1/2}, \tag{1.3}$$

where $\mathbf{D}_n = (\sum_{j\neq i}^n \xi_{ij}^{(n)})_{1\le i\le n}$ and $\mathbf{D}_n^{-1/2} = ((\sum_{j\neq i}^n \xi_{ij}^{(n)})^{-1/2})_{1\le i\le n}$ with $0^{-1/2} := 0$.

The matrix $\mathbf{\Delta}_n$ is always non-negative definite, and the smallest eigenvalue is zero. The Kirchhoff theorem [30] establishes the relationship between the number of the spanning trees of the graph and the eigenvalues of $\mathbf{\Delta}_n$. The matrix $\mathbf{L}_n$ is a different version of $\mathbf{\Delta}_n$. It looks a bit complicated, however, the spectrum of $\mathbf{L}_n$ is related to the graph discrepancy (or the graph invariant), and the second smallest eigenvalue of $\mathbf{L}_n$ relates to the Cheeger constant and the rate convergence of random walks on the graph, see, e.g., [8, 9].

The spectral properties of the graph-relevant matrices such as $\mathbf{A}_n$, $\mathbf{\Delta}_n$ and $\mathbf{L}_n$ have applications in chemistry [4], where eigenvalues were connected to the stability of molecules. They are also studied and used in theoretical physics and quantum mechanics, e.g., [21, 22, 23, 24, 25, 36, 37].

For graph $G(n, p_n)$ there are two major asymptotic regimes: $np_n \to \infty$ and $np_n \to c \in (0, \infty)$. The former corresponds to the *dilute* model, and the latter corresponds to the *sparse* model. Of course, the dilute model consists of two important cases: $p_n \equiv p$ and $p_n \to 0$ with $np_n \to \infty$.

There are several work on $\mathbf{\Delta}_n$. Bauer and Golinelli [3] simulate the eigenvalues for the Erdös-Rényi random graph with $p_n \equiv p$. Assuming the entries $\{\xi_{ij}^{(n)} \equiv \xi_{ij};\ 1 \le i < j \le n < \infty\}$ in $\mathbf{\Delta}_n$ in (1.2) are independent and identically distributed random variables with mean zero, Bryc, Dembo and Jiang [6] prove that the empirical distribution of the eigenvalues of $\mathbf{\Delta}_n$ converges to the free convolution γ_M of the semi-circular law and $N(0, 1)$. This results does not apply to the Erdös-Rényi random graph directly because the means of ξ_{ij}'s in the graph are not zero. A result from Ding and Jiang [15] complements their result in [6] by showing that, under a general framework including the Erdös-Rényi random graph $G(n, p_n)$ with $p_n \equiv p$, the corresponding empirical distribution converges to γ_M. When $\xi_{ij}^{(n)}$'s in (1.2) are deterministic, some results are given for the normalized Laplacian matrix $\mathbf{L}_n$ in [8].

In this paper, for the Erdös-Rényi random graph $G(n, p_n)$ with $np_n \to \infty$, we study the empirical distributions of the eigenvalues of $\mathbf{\Delta}_n$ and $\mathbf{L}_n$. Before stating the main results, we review some notation.

Given an $n \times n$ symmetric matrix $\mathbf{M}$, let $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_n$ be the eigenvalues of $\mathbf{M}$, we sometimes also write this as $\lambda_1(\mathbf{M}) \ge \lambda_2(\mathbf{M}) \ge \cdots \ge \lambda_n(\mathbf{M})$ for clarity. The notation $\lambda_{max} = \lambda_{max}(\mathbf{M})$, $\lambda_{min} = \lambda_{min}(\mathbf{M})$ stand for the largest eigenvalue and the smallest eigenvalue of $\mathbf{M}$, respectively. The following is the situation we will consider in this paper.

Let $\{\xi_{ij}^{(n)};\, 1 \le i < j < \infty\}$ *be defined on the same probability space with* $\xi_{ii}^{(n)} = 0$ *and* $\xi_{ij}^{(n)} = \xi_{ji}^{(n)}$ *for all* $1 \le i, j \le n < \infty$. *For each* $n \ge 2$, *assume* $\{\xi_{ij}^{(n)};\, 1 \le i < j \le n\}$ *are i.i.d. with*

$$P(\xi_{12}^{(n)} = 1) = 1 - P(\xi_{12}^{(n)} = 0) = p_n \in (0, 1). \tag{1.4}$$

For an $n \times n$ matrix $\mathbf{M}$, we use $\|\mathbf{M}\| = \sup_{\mathbf{x}\in\mathbb{R}^n:\, \|\mathbf{x}\|=1} \|\mathbf{M}\mathbf{x}\|$ to denote its spectral norm, where $\|\mathbf{x}\| = \sqrt{x_1^2 + \cdots + x_n^2}$ for $\mathbf{x} = (x_1, \cdots, x_n)' \in \mathbb{R}^n$. The following are the main results of this paper.

Theorem 1 *Suppose (1.4) holds with* $\alpha_n := (np_n(1-p_n))^{1/2} \to \infty$ *as* $n \to \infty$*, then, almost surely,*

$$F_n(x) := \frac{1}{n}\sum_{i=1}^{n} I\left\{\frac{\lambda_i(\mathbf{\Delta}_n) - np_n}{\sqrt{np_n(1-p_n)}} \le x\right\}, \quad x \in \mathbb{R}, \tag{1.5}$$

converges weakly to the free convolution γ_M *of the semi-circle law and* $N(0,1)$*.*

It is known from [6] that the measure γ_M is a non-random symmetric probability measure with smooth bounded density, doesn't depend on p_n's and has an unbounded support. Theorem 1.3 from [6] and Theorem 2 from [15] deal with similar problems by assuming $\xi_{ij}^{(n)}$'s in (1.2) being arbitrary random variables of certain finite moments. With truncations on $\xi_{ij}^{(n)}$'s, the proofs of the two theorems from [6, 15] are essentially reduced to the case that $\xi_{ij}^{(n)}$'s are bounded. Here the situation is different: we will need to work on $(\xi_{ij}^{(n)} - p_n)/\sqrt{p_n}$'s, which are unbounded since $p_n \to 0$ in one of the dilute cases mentioned earlier. The truncation method used in [6, 15] does not apply here simply because the current random variables just take two values. In fact, more involved analysis and combinatorics are used in the proof of Theorem 1.

If p_n in Theorem 1 is equal to a constant, we get the following result.

Corollary 1.1 *Suppose (1.4) holds with* $p_n \equiv p \in (0,1)$ *for all* $n \ge 2$*. Then, almost surely,*

$$F_n(x) := \frac{1}{n}\sum_{i=1}^{n} I\left\{\frac{\lambda_i(\mathbf{\Delta}_n) - np}{\sqrt{np(1-p)}} \le x\right\}, \quad x \in \mathbb{R},$$

converges weakly to the free convolution γ_M *of the semi-circle law and* $N(0,1)$*.*

Take p_n in Theorem 1 to be a special *dilute* case: $p_n \to 0$ with $np_n \to +\infty$ as $n \to \infty$. Note that, under this condition, $\sqrt{np_n(1-p_n)} \sim \sqrt{np_n}$ as $n \to \infty$, we then easily have the following.

Corollary 1.2 *Suppose (1.4) holds with* $p_n \to 0$ *and* $np_n \to \infty$ *as* $n \to \infty$*, then, almost surely,*

$$F_n(x) := \frac{1}{n}\sum_{i=1}^{n} I\left\{\frac{\lambda_i(\mathbf{\Delta}_n) - np_n}{\sqrt{np_n}} \le x\right\}, \quad x \in \mathbb{R},$$

converges weakly to the free convolution γ_M *of the semi-circle law and* $N(0,1)$*.*

Now we study the normalized Laplacian matrix $\mathbf{L}_n$ in (1.3).

Theorem 2 *Suppose (1.4) holds with* $\sup\{p_n;\, n \ge 2\} < 1$*. If* $np_n/\log n \to \infty$ *as* $n \to \infty$*, then, with probability one,*

$$F_n(x) := \frac{1}{n}\sum_{i=1}^{n} I\left(\sqrt{\frac{np_n}{1-p_n}}(1 - \lambda_i(\mathbf{L}_n)) \le x\right), \quad x \in \mathbb{R},$$

converges weakly to the semi-circle law with density $\frac{1}{2\pi}\sqrt{4-x^2}I(|x| \le 2)$*.*

Fan, Lu and Vu [11] derive a result for a weak form of $\mathbf{L}_n$ when the graphs are of the given expected degrees, of which the Erdös-Rényi model is a special case. The three authors prove that the the empirical distribution of the eigenvalues of $\mathbf{I}_n - \tilde{\mathbf{D}}_n^{-1/2}\mathbf{A}_n\tilde{\mathbf{D}}_n^{-1/2}$ converges to the semi-circle law in probability, where $\tilde{\mathbf{D}}_n$ is the mathematical expectation of $\mathbf{D}_n$.

When p_n does not depend on n, the above theorem obviously implies the following.

Corollary 1.3 *Suppose (1.4) holds with $p_n \equiv p \in (0,1)$ for all $n \geq 2$. Then, with probability one,*

$$F_n(x) := \frac{1}{n}\sum_{i=1}^{n} I\left(\sqrt{\frac{np}{1-p}}(1-\lambda_i(\mathbf{L}_n)) \leq x\right), \quad x \in \mathbb{R},$$

converges weakly to the semi-circle law with density $\frac{1}{2\pi}\sqrt{4-x^2}I(|x|\leq 2)$.

If $p_n \to 0$, then $\sqrt{\frac{np_n}{1-p_n}} \sim \sqrt{np_n}$ as $n \to \infty$. From Theorem 2 we immediately obtain a limiting result for the dilute Erdös-Rényi random graph.

Corollary 1.4 *Suppose (1.4) holds with $p_n \to 0$ and $np_n/\log n \to \infty$ as $n \to \infty$, then, almost surely,*

$$F_n(x) := \frac{1}{n}\sum_{i=1}^{n} I\left(\sqrt{np_n}(1-\lambda_i(\mathbf{L}_n)) \leq x\right), \quad x \in \mathbb{R},$$

converges weakly to the semi-circle law with density $\frac{1}{2\pi}\sqrt{4-x^2}I(|x|\leq 2)$.

In the general context of random graphs, the limits of the empirical distributions of the eigenvalues of $\mathbf{A}_n, \mathbf{\Delta}_n$ and $\mathbf{L}_n$ in (1.1), (1.2) and (1.3) respectively are related to the three distributions: the Wigner's semi-circle law, the free convolution of the semi-circle law and $N(0,1)$, and the Kesten-McKay law. The last is the limit of the empirical distributions of the eigenvalues of $\mathbf{A}_n$ for the random d-regular graphs, see [33].

Remark 1.1 *The above theorems study the limiting spectral distributions of $\mathbf{\Delta}_n$ and $\mathbf{L}_n$ for the dilute case. There are some simulations and theoretical work on $\mathbf{A}_n$ and $\mathbf{\Delta}_n$ for the sparse case in [3, 26, 29, 32, 36, 37, 38], that is, $p_n = c/n$ for all $n \geq 2$ and c is a constant. It seems that the limits of the spectral distributions of $\mathbf{A}_n$ and $\mathbf{\Delta}_n$ are still not identified.*

Remark 1.2 *Theorems 1 and 2 are derived based on the Erdös-Rényi model. The methods of the proofs could be carried to more complex models, for instance, the random geometric graphs [34], the weighted dilute random matrices [28], the power law random graphs, and the random graphs with given expected degrees [9].*

Remark 1.3 *There are a couple of books on the study of the spectral properties of the matrices generated from graphs, see, e.g., [1, 8, 9, 12, 13, 35] for reference.*

The rest of the paper is organized as follows: the proof of Theorem 1 is given in Section 2; the proof of Theorem 2 is given in Section 3.

2 Proof of Theorem 1

Given $n \geq 2$, let $\Gamma_n = \{(i,j);\ 1 \leq j < i \leq n\}$ be a graph. Any element $(i,j) \in \Gamma_n$ is called a vertex, and i and j are called indices. We say vertices $a = (i_1, j_1)$ and $b = (i_2, j_2)$ form an edge and denote it by $a \sim b$, if one of i_1 and j_1 is identical to one of i_2 and j_2. For convenience of notation, from now on, we write $a = (a^+, a^-)$ for any $a \in \Gamma_n$. Of course, $a^+ > a^-$. Given $a, b \in \Gamma_n$, define an $n \times n$ matrix

$$\mathbf{Q}_{a,b}[i,j] = \begin{cases} -1, & \text{if } i = a^+, j = b^+, \text{ or } i = a^-, j = b^-; \\ 1, & \text{if } i = a^+, j = b^-, \text{ or } i = a^-, j = b^+; \\ 0, & \text{otherwise.} \end{cases}$$

Recall (1.4). Set

$$\eta_{ij}^{(n)} = \frac{\xi_{ij}^{(n)} - p_n}{\sqrt{p_n(1-p_n)}} \tag{2.1}$$

for all $1 \leq i \neq j \leq n$ and $n \geq 2$, and

$$\mathbf{\Psi}_n = \begin{pmatrix} \sum_{j\neq 1} \eta_{1j}^{(n)} & -\eta_{12}^{(n)} & \cdots & -\eta_{1n}^{(n)} \\ -\eta_{21}^{(n)} & \sum_{j\neq 2} \eta_{2j}^{(n)} & \cdots & -\eta_{2n}^{(n)} \\ \vdots & \vdots & \vdots & \vdots \\ -\eta_{n1}^{(n)} & -\eta_{n2}^{(n)} & \cdots & \sum_{j\neq n} \eta_{nj}^{(n)} \end{pmatrix}. \tag{2.2}$$

With this and the above notation, we rewrite $\mathbf{\Psi}_n$ as follows.

$$-\mathbf{\Psi}_n = \sum_{a\in\Gamma_n} \eta_a^{(n)} \mathbf{Q}_{a,a}, \tag{2.3}$$

where $\eta_a^{(n)} = \eta_{a^+a^-}^{(n)}$ for $a = (a^+, a^-) \in \Gamma_n$. Let $t_{a,b} = tr(\mathbf{Q}_{a,b})$. We summarize some facts from [6] in the following lemma.

Lemma 2.1 *Let $a, b \in \Gamma_n$. The following assertions hold:*

(i) $t_{a,b} = t_{b,a}$.

(ii) $$t_{a,b} = \begin{cases} -2, & \text{if } a = b; \\ -1, & \text{if } a \neq b \text{ and } a^- = b^- \text{ or } a^+ = b^+; \\ 1, & \text{if } a \neq b \text{ and } a^- = b^+ \text{ or } a^+ = b^-; \\ 0, & \text{otherwise.} \end{cases}$$

(iii) $\mathbf{Q}_{a,b} \times \mathbf{Q}_{c,d} = t_{b,c}\mathbf{Q}_{a,d}$. *Therefore,* $tr(\mathbf{Q}_{a_1,a_1} \times \mathbf{Q}_{a_2,a_2} \times \cdots \times \mathbf{Q}_{a_r,a_r}) = \prod_{j=1}^r t_{a_j,a_{j+1}}$, *where* $a_1, \cdots, a_r \in \Gamma_n$, *and* $a_{r+1} = a_1$.

Given $n \geq 2$, we call $\pi = (a_1, \cdots, a_r)$ a circuit of length r in Γ_n if $a_i \in \Gamma_n$ for $1 \leq i \leq n$ and $a_1 \sim \cdots \sim a_r \sim a_1$. For such a circuit, set

$$\eta_\pi^{(n)} = \Big(\prod_{j=1}^r t_{a_j,a_{j+1}}\Big) \prod_{j=1}^r \eta_{a_j}^{(n)} \tag{2.4}$$

with $a_{r+1} = a_1$. From (2.3) and (iii) of Lemma 2.1, we know

$$tr(\mathbf{\Psi}_n^r) = (-1)^r \sum_{\pi} \eta_\pi^{(n)} \quad \text{and} \quad E\, tr(\mathbf{\Psi}_n^r) = (-1)^r \sum_{\pi} E\eta_\pi^{(n)} \tag{2.5}$$

where the sum is taken over all circuits of length r in Γ_n.

Definition 2.1 *We say that a circuit $\pi = (a_1 \sim \cdots \sim a_r \sim a_1)$ of length r in Γ_n is vertex-matched if for each $i = 1, \cdots, r$ there exists some $j \neq i$ such that $a_j = a_i$; we say π has a match of order 3 if some value is repeated at least three times among $\{a_j, j = 1, \cdots, r\}$.*

Clearly, by independence, the only possible non-zero terms in (2.5) come from vertex-matched circuits. For $x \geq 0$, denote by $\lfloor x \rfloor$ the integer part of x. The following two lemmas will be used later.

Lemma 2.2 *(Proposition 4.10 from [6]). Fix $r \in \mathbb{N}$. Let N denote the number of vertex-matched circuits in Γ_n with vertices having at least one match of order 3. Then $N = O(n^{\lfloor (r+1)/2 \rfloor})$ as $n \to \infty$.*

Recall $\Gamma_n = \{(i,j);\, 1 \leq j < i \leq n\}$ for $n \geq 2$. Let $m, r, s_1, \cdots, s_m$ be positive integers with $r \geq m$ and $s_1 + \cdots + s_m = r$. Define $A_{m,n,r}$ by

$$\{a_1 \sim a_2 \sim \cdots \sim a_r \sim a_1 : a_i \in \Gamma_n \text{ for } 1 \leq i \leq n; \text{there exist pairwise different } a_1', \cdots, a_m' \in \{a_1, \cdots, a_r\} \text{ such that } a_i' \text{ appears } s_i \text{ times in the sequence } a_1 a_2 \cdots a_r \text{ for } 1 \leq i \leq m\}. \tag{2.6}$$

Lemma 2.3 *Let $m, r, s_1, \cdots, s_m$ be positive integers with $r \geq m$ and $s_1 + \cdots + s_m = r$. Then, there exists a constant C_r depending only on r such that $|A_{m,n,r}| \leq C_r \cdot n^{m+1}$ for all $n \geq 2$.*

Proof. First, denote by B the total number of ways to assign m different items, say, $\omega_1, \cdots, \omega_m$, into a sequence of length r such that ω_i appears exactly s_i times for each $1 \leq i \leq m$. Then $B = \binom{r}{s_1, \cdots, s_l} \leq r!$.

Given one of such assignment, recalling (2.6), the number of choices of $a_1 = (a_1^+, a_1^-) \in \Gamma_n$ is at most n^2. Then, each time a new vertex $a_i' = (a_i^+, a_i^-)$ appears in the circuit *sequence*, one index from $\{a_i^+, a_i^-\}$ is free, while the other is identical to one of $\{a^+, a^-\}$'s which appear in the circuit sequence earlier than a_i'. Thus, the total number of the choices of a_i' is bounded by $2^{r-1}n + 2^{r-1}n = 2^r n$. It follows that the total number of the circuits corresponding to the assignment is no more than $n^2 \cdot (2^r n)^{m-1} \leq 2^{r^2} n^{m+1}$. Take $C_r = 2^{r^2} r!$. The conclusion then follows by combining the two steps. ■

Recall vertex $a = (a^+, a^-) \in \Gamma_n$ has two end points a^+ and a^-.

Definition 2.2 *Fix integer $n \geq 2$. Let $\pi_1, \pi_2, \pi_3, \pi_4$ be four circuits of length r in Γ_n, we say they are matched if each vertex of any one of these circuits is either self-matched, that is, there is another vertex of the same circuit with the same end points, or is cross-matched, that is, there is a vertex of the other circuit with the same end points (or both).*

Given $l \in \mathbb{N}$, let Q_l be the number of the above matched quadruples satisfying: (a) for any $i = 1, 2, 3, 4$, there is $j \neq i$ such that π_i and π_j share a common vertex; (b) the total number of different vertices in the four circuits is l.

Lemma 2.4 *For fixed l, $r \in \mathbb{N}$ with $1 \le l \le 2r$, there exists a constant C_r not depending on n, such that $Q_l \le C_r n^{l+2}$ for all $n \ge 2$.*

Proof. Write

$$\begin{aligned} \pi_1 &= a_1 \sim a_2 \sim \cdots \sim a_r \sim a_1; \\ \pi_2 &= b_1 \sim b_2 \sim \cdots \sim b_r \sim b_1; \\ \pi_3 &= c_1 \sim c_2 \sim \cdots \sim c_r \sim c_1; \\ \pi_4 &= d_1 \sim d_2 \sim \cdots \sim d_r \sim d_1. \end{aligned} \tag{2.7}$$

For $i \neq j$, denote by $\pi_i \leftrightarrow \pi_j$ if π_i and π_j have at least one crossed match. Up to the permutations of $\pi_1, \pi_2, \pi_3, \pi_4$, all matched quadruples satisfying (a) above the statement of the lemma can be divided into three cases: (i) $\pi_1 \leftrightarrow \pi_2$ and $\pi_3 \leftrightarrow \pi_4$; (ii) $\pi_1 \leftrightarrow \pi_2$, $\pi_1 \leftrightarrow \pi_3$ and $\pi_2 \leftrightarrow \pi_4$; (iii) $\pi_1 \leftrightarrow \pi_2$, $\pi_1 \leftrightarrow \pi_3$ and $\pi_1 \leftrightarrow \pi_4$. (By "permutations" we mean that the circuits $\pi_1, \pi_2, \pi_3, \pi_4$ in (i), (ii) and (iii) can be $\pi_{i_1}, \pi_{i_2}, \pi_{i_3}, \pi_{i_4}$ for any permutation $\{i_1, i_2, i_3, i_4\}$ of $\{1, 2, 3, 4\}$). Thus, to prove the lemma, it suffices to show that the total number of the circuits in each of the three cases are no more than Cn^{l+2} for some constant C not depending on n. To do so, since each of the four circuits has length r, *it is enough to show the total number of the vertices in the four circuits is bounded by Cn^{l+2} with C not depending on n.*

Case (i): $\pi_1 \leftrightarrow \pi_2$ and $\pi_3 \leftrightarrow \pi_4$. Suppose there are l_1 different vertices among circuits π_1 and π_2, then there are $l_2 = l - l_1 \ge 0$ different vertices in π_3 and π_4 that do not appear in π_1 and π_2.

Since $\pi_1 \leftrightarrow \pi_2$, recalling (2.7), say, $(a_1^+, a_1^-) = a_1 \sim b_1$. Then, there are at most $n(n-1)/2$ ways to assign $\{1, \cdots, n\}$ to a_1^+ and a_1^- with $a_1^+ > a_1^-$. Now assign such numbers to other $l_1 - 1$ vertices, there are at most $2n$ assignment for each of a new vertex since two distinct neighbors are connected. So the total number of vertices in π_1 and π_2 is no more than $n^2 \cdot (2n)^{l_1-1} \le 2^{l_1} n^{l_1+1}$. Similarly, there are at most $2^{l_2} n^{l_2+1}$ different vertices in circuits π_3 and π_4. So the total number of vertices in circuits $\pi_1, \pi_2, \pi_3, \pi_4$ is bounded by $2^{l_1} n^{l_1+1} \cdot 2^{l_2} n^{l_2+1} = 2^l n^{l+2}$.

Cases (ii) and (iii). The difference between the two cases and case (i) is that the four circuits here are all connected, while (π_1, π_2) and (π_3, π_4) in case (i) may not have a common vertex at all. For $i = 1, 2, 3, 4$, let u_i be the number of the distinct vertices in π_i that do not appear in π_j for $j < i$. Then $u_1 + u_2 + u_3 + u_4 = l$. We proceed to assign (i, j)'s with $i > j$ to the vertices in π_1 first, then π_2, π_3, π_4 consecutively. By the same consideration as that in (i), the total number of possible vertices in π_1 is at most n^{u_1+1}. Since for any $i = 2, 3, 4$, there exists $j = 1, 2, 3$ such that $\pi_j \leftrightarrow \pi_i$, then the total number of different vertices in π_k is at most $(2n)^{u_k}$ for $k = 2, 3, 4$. Therefore, the total number of the distinct vertices in the four circuits is at most $n^{u_1+1} \cdot (2n)^{u_2} \cdot (2n)^{u_3} \cdot (2n)^{u_4} \le 2^l n^{l+1}$.
■

Lemma 2.5 *Let m, $n \in \mathbb{N}$. Let $X_1, \cdots, X_n$ be non-negative, independent random variables. Set $g(a_1, \cdots, a_n) = E\left(X_1^{a_1} \cdots X_n^{a_n}\right) < \infty$ for non-negative integers $a_1, \cdots, a_n$ with $0^0 = 1$. Then*

$$g(a_1, \cdots, a_n) \ge \prod_{j=1}^{m} g(a_{1j}, \cdots, a_{nj})$$

where $\{a_{ij}; 1 \le i \le n, 1 \le j \le m\}$ *are nonnegative integers satisfying* $\sum_{j=1}^m a_{ij} = a_i$ *for all* $1 \le i \le n$.

Proof. Given random variable $Y \ge 0$ and non-decreasing functions $f(x)$ and $g(x)$ on $[0, \infty)$. If $E|f(Y)| < \infty$, $E|g(Y)| < \infty$ and $E|f(Y)g(Y)| < \infty$, then, by the covariance inequality, we have that $E(f(Y)g(Y)) \ge Ef(Y) \cdot Eg(Y)$. Thus, for non-negative integers $b, b_1, \cdots, b_m$ with $b = b_1 + \cdots + b_m$, we obtain

$$E(X_i^b) \ge E(X_i^{b_1}) \cdot E(X_i^{b_2+\cdots+b_n}) \ge \cdots \ge \prod_{j=1}^n E(X_i^{b_j}) \tag{2.8}$$

for any $1 \le i \le n$. By independence,

$$g(a_1, \cdots, a_n) = \prod_{i=1}^n E\left(X_i^{a_i}\right) \quad \text{and} \quad \prod_{j=1}^m g(a_{1j}, \cdots, a_{nj}) = \prod_{i=1}^n \prod_{j=1}^m E\left(X_i^{a_{ij}}\right).$$

From (2.8), $E\left(X_i^{a_i}\right) \ge \prod_{j=1}^m E\left(X_i^{a_{ij}}\right)$ since $\sum_{j=1}^m a_{ij} = a_i$ for all $i = 1, 2, \cdots, n$. Then the conclusion follows. ■

Recall from (2.5) that $n^{-r/2-1} E\, tr(\mathbf{\Psi}_n^r) = (-1)^r n^{-r/2-1} \sum_\pi E\eta_\pi^{(n)}$ where the sum is taken over all circuits of length r in Γ_n. We next show that some terms in the sum are negligible.

Lemma 2.6 *Suppose the conditions in Theorem 1 hold. Fix* $r \in \mathbb{N}$. *Let* $W_n = n^{-r/2-1} \sum_{\pi \in \Lambda_n} E\eta_\pi^{(n)}$, *where* Λ_n *is the set of vertex matched circuits of order* 3 *and of length* r *in* Γ_n. *Then* $W_n \to 0$ *as* $n \to \infty$.

Proof. Take a circuit in Λ_n: $a_1 \sim a_2 \sim \cdots \sim a_r \sim a_1$. Let $a_1', \cdots, a_m' \in \{a_1, \cdots, a_r\}$ be pairwise different vertices with a_i' appearing $s_i \ge 2$ times in the sequence $a_1 a_2 \cdots a_r$ for $1 \le i \le m$ with $s_1 + \cdots + s_m = r$ and $\max_{1 \le i \le m} s_i \ge 3$. Evidently, $m \le (r-1)/2$. By (2.4) and independence,

$$|E\eta_\pi^{(n)}| = |\left(\prod_{j=1}^r t_{a_j, a_{j+1}}\right) \cdot E \prod_{i=1}^m (\eta_{a_i'}^{(n)})^{s_i}| \le 2^r \cdot \prod_{i=1}^m E(|\eta_{a_i'}^{(n)}|^{s_i}),$$

where the last inequality is from (ii) of Lemma 2.1. Now, let $Z \sim Ber(p_n)$ and $q_n = 1 - p_n$. Notice $p_n^{s_i} q_n + p_n q_n^{s_i} \le 2p_n q_n$, we obtain from (2.1) that $E(|\eta_{a_i'}^{(n)}|^{s_i})$ is identical to

$$E\left(\frac{|Z - p_n|}{\sqrt{p_n q_n}}\right)^{s_i} = \frac{p_n^{s_i} q_n + p_n q_n^{s_i}}{(p_n q_n)^{s_i/2}} \le \frac{2}{(p_n q_n)^{s_i/2-1}}. \tag{2.9}$$

Using $s_1 + \cdots + s_m = r$ to see that $|E\eta_\pi^{(n)}| \le 2^r \cdot 2^m (p_n q_n)^{m-(r/2)}$. Let $C_{m,r}$ be the number of solutions of the equation $s_1 + \cdots + s_m = r$ with integer $s_i \ge 2$ for each $1 \le i \le m$. Now we classify the sum $\sum_{\pi \in \Lambda_n} E\eta_\pi^{(n)}$ into the sums corresponding to $m = 1, 2, \cdots, \lfloor (r-1)/2 \rfloor$. By Lemma 2.3, the m-th sum is bounded by $C_{m,r} \cdot C_r n^{m+1} \cdot 2^{m+r} (p_n q_n)^{m-(r/2)}$. It follows that

$$\begin{aligned} |W_n| &\le \frac{1}{n^{r/2+1}} \sum_{1 \le m \le (r-1)/2} C_{m,r} \cdot C_r n^{m+1} \cdot 2^{m+r} (p_n q_n)^{m-(r/2)} \\ &\le C \sum_{1 \le m \le (r-1)/2} (n p_n q_n)^{m-(r/2)} = O\left(\frac{1}{\sqrt{n p_n q_n}}\right) \to 0 \end{aligned}$$

as $n \to \infty$ since $np_nq_n \to \infty$ by assumption, where C above is a constant not depending on n. ■

Let $\mathbf{U}_n$ be a symmetric matrix of form

$$\mathbf{U_n} = \begin{pmatrix} \sum_{j\neq 1} Y_{1j} & -Y_{12} & \cdots & -Y_{1n} \\ -Y_{21} & \sum_{j\neq 2} Y_{2j} & \cdots & -Y_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ -Y_{n1} & -Y_{n2} & \cdots & \sum_{j\neq n} Y_{nj} \end{pmatrix} \tag{2.10}$$

where $\{Y_{ij}; 1 \leq i < j < \infty\}$ are i.i.d. standard normal random variables not depending on n.

Lemma 2.7 *Suppose the conditions in Theorem 1 hold. Let* Ψ_n *and* $\mathbf{U}_n$ *be as in (2.2) and (2.10) respectively. Then*

(i) $\lim_{n\to\infty} \frac{1}{n^{k+1/2}} E\, tr(\mathbf{\Psi}_n^{2k-1}) = 0;$

(ii) $\lim_{n\to\infty} \frac{1}{n^{k+1}} \left(E\, tr(\mathbf{\Psi}_n^{2k}) - E\, tr(\mathbf{U}_n^{2k})\right) = 0$

for any integer $k \geq 1$.

Proof. (i) As in (2.5), $E\, tr(\mathbf{\Psi}_n^{2k-1}) = \sum_\pi E\eta_\pi^{(n)}$ where the sum is taken over all circuits of length $2k-1$ in Γ_n in which each vertex appearing at least twice. Notice $r = 2k-1$ is odd, one vertex has to appear at least three times. The result then follows from Lemma 2.6.

(ii) Recall (2.10). For a circuit $a_1 \sim \cdots \sim a_r \sim a_1$ with $a_i \in \Gamma_n$ for all $1 \leq i \leq r$, similar to notation $\eta_\pi^{(n)}$ in (2.4), we define

$$Y_\pi^{(n)} = \Big(\prod_{j=1}^r t_{a_j,a_{j+1}}\Big) \prod_{j=1}^r Y_{a_j}. \tag{2.11}$$

Taking $r = 2k$, we have

$$\begin{aligned} |\, E\, tr(\mathbf{\Psi}_n^{2k}) - E\, tr(\mathbf{U}_n^{2k})\, | &= |\sum_\pi (E\, \eta_\pi^{(n)} - E\, Y_\pi^{(n)})| \\ &\leq |\sum_{\pi\in A_1} E\, \eta_\pi^{(n)}| + |\sum_{\pi\in A_1} E\, Y_\pi^{(n)}| + |\sum_{\pi\in A_2} (E\, \eta_\pi^{(n)} - E\, Y_\pi^{(n)})\, | \\ &:= I_1 + I_2 + I_3, \end{aligned} \tag{2.12}$$

where A_1 denotes the set of the vertex-matched circuits with match of order 3, and A_2 denotes the set of the vertex-matched circuits in Γ_n such that there are exact k distinct matches. Now we analyze the three terms one by one.

By Lemma 2.6, $I_1 = o(n^{k+1})$ as $n \to \infty$. Observe that each vertex of any circuit in A_2 matches exactly two times. From the independence assumption and that $E|\eta_{ij}^{(n)}|^2 = E|Y_{ij}|^2 = 1$ for all $1 \leq i < j \leq n$ and $n \geq 2$, we know $E\, \eta_\pi^{(n)} = E\, Y_\pi^{(n)} = 1$ for all $\pi \in A_2$. This gives $I_3 = 0$. Now we turn to estimate I_2.

Recalling (2.11), for $\pi \in A_1$, $\prod_{j=1}^{2k} Y_{a_j}$ has the form of $\prod_{s=1}^{w} V_s^{l_s}$, where $V_1, \cdots, V_w$ are i.i.d. $N(0,1)$-distributed random variables, and $l_1 \geq 2, \cdots, l_w \geq 2$ with $\sum_{s=1}^{w} l_s = 2k$. Thus, by the Hölder inequality, $E|\prod_{j=1}^{2k} Y_{a_j}| \leq E(|N(0,1)|^{2k})$. From (ii) of Lemma 2.1, $|EY_\pi^{(n)}| \leq 2^{2k}E(|N(0,1)|^{2k})$. It follows that $I_2 \leq 2^{2k}E(|N(0,1)|^{2k}) \cdot |A_1| = O(n^k)$ as $n \to \infty$ by Lemma 2.2.

In summary, $I_1 + I_2 + I_3 = o(n^{k+1})$ as $n \to \infty$, which together with (2.12) concludes (ii). ■

LEMMA 2.8 *Suppose the conditions in Theorem 1 hold. Let $\mathbf{\Psi}_n$ be as in (2.2). Then*

$$\lim_{n\to\infty} \frac{1}{n^{r/2+1}} \left(tr(\mathbf{\Psi}_n^r) - Etr(\mathbf{\Psi}_n^r)\right) = 0 \quad a.s.$$

for any integer $r \geq 1$.

Proof. By (2.4) and (2.5),

$$tr(\mathbf{\Psi}_n^r) - Etr(\mathbf{\Psi}_n^r) = (-1)^r \sum_\pi (\eta_\pi^{(n)} - E\eta_\pi^{(n)})$$

where the sum is taken over all circuits of length r in Γ_n, and

$$\eta_\pi^{(n)} = \prod_{j=1}^r t_{a_j,a_{j+1}} \prod_{j=1}^r \eta_{a_j}^{(n)}$$

for $\pi = (a_1, \cdots, a_r)$. We claim that, to prove the lemma, it is enough to show

$$E\left|tr(\mathbf{\Psi}_n^r) - Etr(\mathbf{\Psi}_n^r)\right|^4 = O(n^{2r+2}) \tag{2.13}$$

as $n \to \infty$. In fact, if this holds, by the Markov inequality,

$$P\left(\frac{1}{n^{r/2+1}}\left|tr(\mathbf{\Psi}_n^r) - Etr(\mathbf{\Psi}_n^r)\right| \geq \epsilon\right) \leq \frac{E\left|tr(\mathbf{\Psi}_n^r) - Etr(\mathbf{\Psi}_n^r)\right|^4}{n^{2r+4}\epsilon^4} = O\left(\frac{1}{n^2}\right)$$

as $n \to \infty$ for any $\epsilon > 0$. Thus, the sum of the left hand side of the above over all $n \geq 2$ is finite, the Borel-Cantelli lemma then gives the desired conclusion. Now,

$$\begin{aligned} & E\left|tr(\mathbf{\Psi}_n^r) - Etr(\mathbf{\Psi}_n^r)\right|^4 \\ = & \sum_{\pi_1,\pi_2,\pi_3,\pi_4} E\left\{(\eta_{\pi_1}^{(n)} - E\eta_{\pi_1}^{(n)})(\eta_{\pi_2}^{(n)} - E\eta_{\pi_2}^{(n)})(\eta_{\pi_3}^{(n)} - E\eta_{\pi_3}^{(n)})(\eta_{\pi_4}^{(n)} - E\eta_{\pi_4}^{(n)})\right\} \end{aligned} \tag{2.14}$$

where the sum is taken over all circuits $\pi_j, j = 1, 2, 3, 4$ of length r each. For every $n \geq 2$, with random variables $\eta_a^{(n)}$'s independent and of mean zero (recalling (2.1)), if a circuit π_k has a vertex that is not matched with any other vertex in the four circuits, then $E\eta_{\pi_k}^{(n)} = 0$ and

$$E\prod_{j=1}^4 (\eta_{\pi_j}^{(n)} - E\eta_{\pi_j}^{(n)}) = E\{\eta_{\pi_k}^{(n)} \cdot \prod_{j\neq k} (\eta_{\pi_j}^{(n)} - E\eta_{\pi_j}^{(n)})\} = 0.$$

Further, if one of the circuits, say π_1, is only self-matched, i.e., has no cross-matched vertex, then obviously

$$E\prod_{j=1}^4 (\eta_{\pi_j}^{(n)} - E\eta_{\pi_j}^{(n)}) = E(\eta_{\pi_1}^{(n)} - E\eta_{\pi_1}^{(n)}) \cdot E\prod_{j=2}^4 (\eta_{\pi_j}^{(n)} - E\eta_{\pi_j}^{(n)}) = 0$$

Therefore, it suffices to take the sum in (2.14) over all matched quadruples of circuits on $\pi_1, \pi_2, \pi_3, \pi_4$ of length r (see Definition 2.2) with the property: for any $i = 1, 2, 3, 4$, there is $j \neq i$ such that π_i and π_j share the same vertex. Recalling the definition of Q_l in Lemma 2.4, since the quadruples are matched, we know $1 \leq l \leq 2r$, and by this lemma, $|Q_l| \leq C_r n^{l+2}$ for all $1 \leq l \leq 2r$, where C_r is a constant depending on r only. Thus, by (2.14),

$$E\big|tr(\mathbf{\Psi}_n^r) - Etr(\mathbf{\Psi}_n^r)\big|^4 \leq \sum_{l=1}^{2r} \sum_{(\pi_1,\pi_2,\pi_3,\pi_4)\in Q_l} |h(\pi_1, \pi_2, \pi_3, \pi_4)|, \tag{2.15}$$

where $h(\pi_1, \pi_2, \pi_3, \pi_4) = E\prod_{j=1}^4 (\eta_{\pi_j}^{(n)} - E\eta_{\pi_j}^{(n)})$. Expand the expectation, there are 16 terms, and the absolute value of each is bounded by $E\prod_{j=1}^4 |\eta_{\pi_j}^{(n)}|$ by Lemma 2.5. Thus, by (2.15),

$$E\big|tr(\mathbf{\Psi}_n^r) - Etr(\mathbf{\Psi}_n^r)\big|^4 \leq 16C_r \sum_{l=1}^{2r} n^{l+2} \max_{(\pi_1,\pi_2,\pi_3,\pi_4)\in Q_l} \left\{ E\prod_{j=1}^4 |\eta_{\pi_j}^{(n)}| \right\}. \tag{2.16}$$

Given $l \geq 1$, let Z be a $Ber(p_n)$-distributed random variable. Then, for each $(\pi_1, \pi_2, \pi_3, \pi_4) \in Q_l$, by independence,

$$E\prod_{j=1}^4 |\eta_{\pi_j}^{(n)}| \leq \max \prod_{i=1}^l E\left(\frac{|Z - p_n|}{\sqrt{p_n q_n}}\right)^{m_i} \tag{2.17}$$

where the maximum is over all $m_1 \geq 1, \cdots, m_l \geq 1$ subject to the constraint $m_1 + \cdots + m_l = 4r$. By (2.9), the above is controlled by

$$\prod_{i=1}^l E\left(\frac{|Z - p_n|}{\sqrt{p_n q_n}}\right)^{m_i} \leq \prod_{i=1}^l \frac{2}{(p_n q_n)^{m_i/2-1}} = \frac{2^l}{(p_n q_n)^{2r-l}}.$$

It follows from (2.16) and (2.17) that

$$\begin{aligned} E\big|tr(\mathbf{\Psi}_n^r) - Etr(\mathbf{\Psi}_n^r)\big|^4 &\leq 16C_r \sum_{l=1}^{2r} \frac{2^l n^{l+2}}{(p_n q_n)^{2r-l}} \\ &\leq (16C_r 2^{2r}) n^{2r+2} \sum_{l=1}^{2r} \frac{1}{(np_n q_n)^{2r-l}} \\ &= O(n^{2r+2}) \end{aligned}$$

as $n \to \infty$ since $np_n q_n \to \infty$ by assumption. This gives (2.13). ■

For an $n \times n$ symmetric matrix $\mathbf{M}$, let $F^{\mathbf{M}}$ be the empirical distribution of the eigenvalues of $\mathbf{M}$.

Lemma 2.9 *Suppose $\alpha_n := (np_n(1-p_n))^{1/2} \to \infty$ as $n \to \infty$. Let $\mathbf{\Psi}_n$ be as in (2.2). Then, as $n \to \infty$, with probability one, $F^{\mathbf{\Psi}_n/\sqrt{n}}$ converges weakly to the free convolution γ_M of the semi-circle law and $N(0,1)$.*

Proof. Proposition A.3 from [6] says that γ_M is a symmetric distribution and uniquely determined by its moments. Thus, to prove the theorem, it is enough to show that

$$\begin{aligned}\frac{1}{n^{r/2+1}} tr(\mathbf{\Psi}_n^r) &= \frac{1}{n} tr(n^{-1/2}\mathbf{\Psi}_n)^r = \int x^r dF^{n^{-1/2}\mathbf{\Psi}_n} \\ &\to \int x^r d\gamma_M \ a.s. \end{aligned} \tag{2.18}$$

as $n \to \infty$ for any integer $r \geq 1$. First, since γ_M is symmetric, we know $\int x^{2k-1}\, d\gamma_M = 0$ for any integer $k \geq 1$. Second, recalling $\mathbf{U}_n$ in (2.10), Proposition 4.13 in [6] says that

$$\frac{1}{n} E\, tr((n^{-1/2}\mathbf{U}_n)^{2k}) \to \int_{\mathbb{R}} x^{2k} d\,\gamma_M$$

as $n \to \infty$ for any $k \geq 1$. Consequently, the two facts together with Lemmas 2.7 and 2.8 yield (2.18). ■

Proof of Theorem 1. Let $\mu_n = p_n$ and $\sigma_n = \sqrt{p_n(1-p_n)}$. Recalling $\mathbf{\Delta}_n$ in (1.2), $\eta_{ij}^{(n)}$ in (2.1) and $\mathbf{\Psi}_n$ in (2.2), we know

$$\mathbf{\Delta}_n = \underbrace{\sigma_n \mathbf{\Psi}_n + (n\mu_n)\mathbf{I}_n}_{\mathbf{\Delta}_{n,1}} - \mu_n \mathbf{J}_n \tag{2.19}$$

where $\mathbf{I}_n$ is the $n \times n$ identity matrix, and $\mathbf{J}_n$ is the $n \times n$ matrix with all of its entries equal to 1. By Lemma 2.9, with probability one,

$$F^{\mathbf{\Psi}_n/\sqrt{n}} \text{ converges weakly to } \gamma_M$$

as $n \to \infty$. Therefore,

$$F^{(\mathbf{\Delta}_{n,1}-n\mu_n\mathbf{I}_n)/\sqrt{n}\sigma_n} = F^{\mathbf{\Psi}_n/\sqrt{n}} \text{ converges weakly to } \gamma_M \tag{2.20}$$

almost surely as $n \to \infty$. By (2.19) and the rank inequality (see Lemma 2.2 from [2]),

$$\begin{aligned} &\|F^{(\mathbf{\Delta}_n-n\mu_n\mathbf{I}_n)/\sqrt{n}\sigma_n} - F^{(\mathbf{\Delta}_{n,1}-n\mu_n\mathbf{I}_n)/\sqrt{n}\sigma_n}\| \leq \frac{1}{n} \cdot \text{rank}\left(\frac{\mathbf{\Delta}_n - \mathbf{\Delta}_{n,1}}{\sqrt{n}\sigma_n}\right) \\ \leq\ & \frac{1}{n} \cdot \text{rank}\,(\mathbf{J}_n) \leq \frac{1}{n} \to 0, \end{aligned} \tag{2.21}$$

as $n \to \infty$, where $\|f\| := \sup_{x\in\mathbb{R}} |f(x)|$ is the supremum norm of a bounded, measurable function $f(x)$ defined on $\mathbb{R}$. Therefore, (2.20) and (2.21) lead to that, as $n \to \infty$,

$$\frac{1}{n}\sum_{i=1}^{n} I\left(\frac{\lambda_i(\mathbf{\Delta}_n) - np_n}{\sqrt{np_n(1-p_n)}} \leq x\right) = F^{(\mathbf{\Delta}_n-n\mu_n\mathbf{I}_n)/\sqrt{n}\sigma_n}(x) \text{ converges weakly to } \gamma_M$$

with probability one. The proof is complete. ■

3 Proof of Theorem 2

For any two probability measures μ and ν on $\mathbb{R}$, define

$$d_{BL}(\mu,\nu) = \sup\{|\int f d\mu - \int f d\nu| : \|f\|_\infty + \|f\|_L \le 1\}, \tag{3.1}$$

where $\|f\|_\infty = \sup_{x\in\mathbb{R}} |f(x)|$, $\|f\|_L = \sup_{x\neq y} |f(x)-f(y)|/|x-y|$. It is well known (see, e.g., Section 11.3 from [16]) that $d_{BL}(\cdot,\cdot)$ is called the bounded Lipschitz metric, which characterizes the weak convergence of probability measures. Given $n\times n$ real and symmetric matrices $\mathbf{M}_1$ and $\mathbf{M}_2$, let $\hat\mu(\mathbf{M}_i)$ be the empirical measure of the eigenvalues (the spectral measure) of $\mathbf{M}_i$ for $i=1,2$, that is,

$$\hat\mu(\mathbf{M}_i) = \frac{1}{n}\sum_{j=1}^n \delta_{\lambda_j(\mathbf{M}_i)}, \quad i=1,2,$$

where δ_x is the point mass probability measure at x. We have (see, e.g., (2.16) from [6])

$$d_{BL}^2(\hat\mu(\mathbf{M}_1),\hat\mu(\mathbf{M}_2)) \le \frac{1}{n} tr((\mathbf{M}_1-\mathbf{M}_2)^2). \tag{3.2}$$

A similar result is the Difference Inequality in Lemma 2.3 from [2]:

$$L^3(F^{\mathbf{M}_1}, F^{\mathbf{M}_2}) \le \frac{1}{n} tr(\mathbf{M}_1-\mathbf{M}_2)^2 \tag{3.3}$$

where $F^{\mathbf{M}_i}$ is the empirical cumulative distribution function of the eigenvalues of $\mathbf{M}_i$ for $i=1,2$. For each $n\ge 2$, recall from (1.4) that $\{\xi_{ij}^{(n)}; 1\le i<j\le n\}$ are i.i.d. $Ber(p_n)$-distributed random variables, and $\xi_{ii}^{(n)}=0$ and $\xi_{ij}^{(n)}=\xi_{ji}^{(n)}$ for all $1\le i\neq j\le n$. Define

$$\begin{aligned} r_{n,i} &= \sum_{j=1}^n \xi_{ij}^{(n)}, \quad \tilde r_{n,i} = r_{n,i} - (n-1)p_n, \\ a_n &= \max_{1\le i\le n}\{|\tilde r_{n,i}|\} \;\text{ and }\; b_n = \min_{1\le i\le n}\{r_{n,i}\}. \end{aligned} \tag{3.4}$$

With this notation, we have the following lemma.

Lemma 3.1 *Let $\mathbf{A}_n$ and $\mathbf{L}_n$ be as in (1.1) and (1.3), respectively. Set $\tilde{\mathbf{L}}_n = \mathbf{I}_n - ((n-1)p_n)^{-1}\mathbf{A}_n$. Then*

$$\frac{1}{n} tr(\tilde{\mathbf{L}}_n - \mathbf{L}_n)^2 \le \frac{2a_n^4 + 8n^2p_n^2a_n^2}{n(n-1)^4p_n^4b_n^2} \sum_{1\le i\neq j\le n} \xi_{ij}^{(n)}$$

for all $n\ge 2$.

Proof. If $b_n=0$, the conclusion holds trivially. Now, without loss of generality, assume that $b_n>0$ for all $n\ge 2$. Note that

$$\begin{aligned} \tilde{\mathbf{L}}_n - \mathbf{L}_n &= -\frac{1}{(n-1)p_n}\mathbf{A}_n + \mathrm{diag}(r_{n,1}^{-1/2},\cdots,r_{n,n}^{-1/2})\,\mathbf{A}_n\,\mathrm{diag}(r_{n,1}^{-1/2},\cdots,r_{n,n}^{-1/2}) \\ &= \left((-\frac{1}{(n-1)p_n} + \frac{1}{\sqrt{r_{n,i}}\sqrt{r_{n,j}}})\xi_{ij}^{(n)}\right)_{n\times n}. \end{aligned}$$

Thus, using $(\xi_{ij}^{(n)})^2 = \xi_{ij}^{(n)}$, we get

$$tr(\tilde{\mathbf{L}}_n - \mathbf{L}_n)^2 = \sum_{1\le i,j\le n}\Big(\frac{1}{(n-1)p_n} - \frac{1}{\sqrt{r_{n,i}}\sqrt{r_{n,j}}}\Big)^2\xi_{ij}^{(n)}. \tag{3.5}$$

Now

$$\Big(\frac{1}{(n-1)p_n} - \frac{1}{\sqrt{r_{n,i}}\sqrt{r_{n,j}}}\Big)^2 = \frac{(\sqrt{r_{n,i}r_{n,j}} - (n-1)p_n)^2}{(n-1)^2p_n^2r_{n,i}r_{n,j}}. \tag{3.6}$$

Use formula $|\sqrt{x} - \sqrt{y}| = |x-y|/(\sqrt{x}+\sqrt{y}) \le |x-y|/\sqrt{y}$ for $x = r_{n,i}r_{n,j}$ and $y = (n-1)^2p_n^2$ to obtain that the right hand side of (3.6) is dominated by

$$\frac{(r_{n,i}r_{n,j} - (n-1)^2p_n^2)^2}{(n-1)^4p_n^4r_{n,i}r_{n,j}} \le \frac{(r_{n,i}r_{n,j} - (n-1)^2p_n^2)^2}{(n-1)^4p_n^4b_n^2}.$$

From the definition of $\tilde{r}_{n,i}$, we see that

$$\begin{aligned}\big|r_{n,i}r_{n,j} - (n-1)^2p_n^2\big| &= \big|(\tilde{r}_{n,i} + (n-1)p_n)(\tilde{r}_{n,j} + (n-1)p_n) - (n-1)^2p_n^2\big| \\ &= |\tilde{r}_{n,i}\cdot\tilde{r}_{n,j} + (n-1)p_n(\tilde{r}_{n,i} + \tilde{r}_{n,j})| \\ &\le a_n^2 + 2(n-1)p_na_n.\end{aligned}$$

It follows from the inequality $(x+y)^2 \le 2x^2 + 2y^2$ for any $x, y \in \mathbb{R}$ that

$$(r_{n,i}r_{n,j} - (n-1)^2p_n^2)^2 \le (a_n^2 + 2(n-1)p_na_n)^2 \le 2a_n^4 + 8n^2p_n^2a_n^2.$$

Thus, collecting all the assertions above, we see that the coefficient of $\xi_{ij}^{(n)}$ in the sum in (3.5) is bounded by

$$\frac{2a_n^4 + 8n^2p_n^2a_n^2}{(n-1)^4p_n^4b_n^2}$$

for all $1 \le i < j \le n$ and $n \ge 2$. This yields the desired inequality. ■

We need the following result.

Lemma 3.2 *Let $\epsilon_1, \cdots, \epsilon_n$ be i.i.d. random variables with $P(\epsilon_1 = 1) = 1 - P(\epsilon_1 = 0) = p \in (0,1)$. Then*

$$P\left(\Big|\frac{\sum_{i=1}^n \epsilon_i}{np} - 1\Big| \ge \delta\right) \le 2\,e^{-(np)\delta^2/4}$$

for all $n \ge 1$ and $\delta \in (0,1)$.

When p is a constant, that is, p does not depend on n, the above is a simple application of the Chernoff bound (see, e.g., p.27 from [14]). In fact, the above result captures the behavior when $p = p_n$ goes to 0.

Lemma 3.3 *Assume (1.4) holds. Let a_n and b_n be as in (3.4). If $np_n/\log n \to \infty$, then*

$$\frac{a_n}{np_n} \to 0 \ \ a.s., \quad \frac{b_n}{np_n} \to 1 \ a.s. \text{ and } \ \frac{1}{n^2p_n}\sum_{1\le i\ne j\le n}\xi_{ij}^{(n)} \to 1 \ a.s.$$

as $n \to \infty$.

Proof. We first prove the last limit. Since $\xi_{ij}^{(n)} = \xi_{ji}^{(n)}$ for all $1 \le i < j \le n$, to show the third assertion in the lemma, it is enough to prove that

$$Z_n := \frac{2}{n(n-1)p_n} \sum_{1\le i<j\le n} \xi_{ij}^{(n)} \to 1 \ \ a.s.$$

as $n \to \infty$. Taking $\delta \in (0,1)$, since $np_n/\log n \to \infty$, we obtain from Lemma 3.2 that $P(|Z_n - 1| \ge \delta) = O(n^{-2})$ as $n \to \infty$. This shows that $\sum_{n\ge 2} P(|Z_n - 1| \ge \delta) < \infty$, which implies $Z_n \to 1 \ \ a.s.$ by the Borel-Cantelli lemma.

Recall that $r_{n,i}$ in (3.4) is a sum of $n-1$ i.i.d. Ber(p_n)-distributed random variables for each $1 \le i \le n$. Also, $\tilde{r}_{n,i} = r_{n,i} - (n-1)p_n$, $a_n = \max_{1\le i\le n}\{|\tilde{r}_{n,i}|\}$ and $b_n = \min_{1\le i\le n}\{r_{n,i}\}$. By Lemma 3.2 again, use condition $np_n/\log n \to \infty$ to get

$$P\left(\frac{a_n}{(n-1)p_n} \ge \delta\right) \le n \cdot P\left(\Big|\frac{\sum_{j=2}^n \xi_{1j}}{(n-1)p_n} - 1\Big| \ge \delta\right) \le \frac{1}{n^2}$$

for any $\delta \in (0,1)$ as n (depending on δ) is sufficiently large. It follows that $\sum_{n\ge 2} P\left(a_n/(n-1)p_n \ge \delta\right)$ is finite for any $\delta \in (0,1)$. By the Borel-Cantelli lemma again, we have

$$\frac{a_n}{np_n} \to 0 \ \ a.s. \tag{3.7}$$

as $n \to \infty$. Finally, from (3.7),

$$\Big|\frac{b_n}{(n-1)p_n} - 1\Big| = \Big|\frac{\min_{1\le i\le n}\{r_{n,i} - (n-1)p_n\}}{(n-1)p_n}\Big| \le \frac{a_n}{(n-1)p_n} \to 0 \ \ a.s.$$

as $n \to \infty$. This leads to $b_n/(np_n) \to 1 \ \ a.s.$ as $n \to \infty$. ■

Proof of Theorem 2. Review from Lemma 3.1 that $\tilde{\mathbf{L}}_n = \mathbf{I}_n - ((n-1)p_n)^{-1}\mathbf{A}_n$, and $\hat{\mu}(\tilde{\mathbf{L}}_n)$ and $\hat{\mu}(\mathbf{L}_n)$ are the empirical measures of the eigenvalues of $\tilde{\mathbf{L}}_n$ and $\mathbf{L}_n$, respectively. By Lemma 3.1,

$$d_{BL}^2(\hat{\mu}(\mathbf{L}_n), \hat{\mu}(\tilde{\mathbf{L}}_n)) \le \frac{1}{n} tr(\mathbf{L}_n - \tilde{\mathbf{L}}_n)^2 \le \frac{2a_n^4 + 8n^2p_n^2a_n^2}{n(n-1)^4p_n^4b_n^2} \sum_{1\le i\ne j\le n} \xi_{ij}^{(n)} \tag{3.8}$$

for $n \ge 2$. Set

$$a_n = (np_n)u_n, \quad b_n = (np_n)v_n \ \text{ and } \sum_{1\le i\ne j\le n} \xi_{ij}^{(n)} = n^2p_nw_n$$

for $n \ge 2$. By Lemma 3.3, $u_n \to 0 \ \ a.s.$, $v_n \to 1 \ \ a.s.$ and $w_n \to 1 \ \ a.s.$ as $n \to \infty$. Therefore, noting $np_n \to \infty$, we obtain

$$\begin{aligned}\frac{2a_n^4 + 8n^2p_n^2a_n^2}{n(n-1)^4p_n^4b_n^2} \sum_{1\le i\ne j\le n} \xi_{ij}^{(n)} &= \frac{2n^4p_n^4u_n^4 + 8n^4p_n^4u_n^2}{n^3(n-1)^4p_n^6v_n^2} \cdot n^2p_nw_n \\ &\sim \frac{(2u_n^4 + 8u_n^2)w_n}{(np_n)v_n^2} = o\left(\frac{1}{np_n}\right) \ \ a.s.\end{aligned}$$

as $n \to \infty$. Therefore,

$$\frac{1}{n} tr(\mathbf{L}_n - \tilde{\mathbf{L}}_n)^2 = o\left(\frac{1}{np_n}\right) \ \ a.s.$$

as $n \to \infty$. It is easy to see from (3.2) that

$$d_{BL}^2\left(\hat{\mu}(\gamma \mathbf{L}_n + \rho \mathbf{I}_n),\, \hat{\mu}(\gamma \tilde{\mathbf{L}}_n + \rho \mathbf{I}_n)\right) \le \gamma^2 \cdot \frac{1}{n} tr(\tilde{\mathbf{L}}_n - \mathbf{L}_n)^2$$

for $n \ge 2$ and any two real numbers γ and ρ. The above two assertions together with (3.8) conclude that

$$d_{BL}^2(\hat{\mu}(\gamma_n \mathbf{L}_n + \rho_n \mathbf{I}_n),\, \hat{\mu}(\gamma_n \tilde{\mathbf{L}}_n + \rho_n \mathbf{I}_n)) = o\left(\frac{\gamma_n^2}{np_n}\right) \quad a.s. \tag{3.9}$$

as $n \to \infty$ for any two sequences of constants $\{\gamma_n,\, n \ge 1\}$ and $\{\rho_n,\, n \ge 1\}$. Notifying $\tilde{\mathbf{L}}_n = \mathbf{I}_n - ((n-1)p_n)^{-1}\mathbf{A}_n$, it is clear that

$$(n-1)p_n(1 - \lambda_i(\tilde{\mathbf{L}}_n)) = \lambda_i(\mathbf{A}_n) \tag{3.10}$$

for $i = 1, 2, \cdots, n$. From the given condition, $\alpha_n := (np_n(1-p_n))^{1/2} \to \infty$. By corollary 1.2 from [15], with probability one, $\hat{\mu}(\mathbf{A}_n/\alpha_n)$ converges weakly to the semi-circle law μ_{cir} with density $\frac{1}{2\pi}\sqrt{4-x^2}I(|x| \le 2)$. Thus, by (3.10), almost surely,

$$\frac{1}{n}\sum_{i=1}^n I\left(\frac{(n-1)p_n}{\alpha_n}(1 - \lambda_i(\tilde{\mathbf{L}}_n)) \le x\right), \quad x \in \mathbb{R},$$

converges weakly to μ_{cir} as $n \to \infty$. Observe that

$$\frac{(n-1)p_n}{\alpha_n} = \frac{(n-1)p_n}{(np_n(1-p_n))^{1/2}} \sim \sqrt{\frac{np_n}{1-p_n}}$$

as $n \to \infty$. By a trivial manipulation, we obtain that, with probability one,

$$\frac{1}{n}\sum_{i=1}^n I\left(\sqrt{\frac{np_n}{1-p_n}}(1 - \lambda_i(\tilde{\mathbf{L}}_n)) \le x\right), \quad x \in \mathbb{R}, \tag{3.11}$$

converges weakly to μ_{cir} as $n \to \infty$. Taking $\gamma_n = -\rho_n = -\sqrt{np_n/(1-p_n)}$, we have

$$\lambda_i(\gamma_n \tilde{\mathbf{L}}_n + \rho_n \mathbf{I}_n) = \sqrt{\frac{np_n}{1-p_n}}(1 - \lambda_i(\tilde{\mathbf{L}}))$$

for $1 \le i \le n$. Thus, by (3.11), with probability one, $\hat{\mu}(\gamma_n \tilde{\mathbf{L}}_n + \rho_n \mathbf{I}_n)$ converges weakly to μ_{cir} as $n \to \infty$. Equivalently,

$$d_{BL}\left(\hat{\mu}(\gamma_n \tilde{\mathbf{L}}_n + \rho_n \mathbf{I}_n),\, \mu_{cir}\right) \to 0 \quad a.s. \tag{3.12}$$

as $n \to \infty$. By the condition that $\sup\{p_n;\, n \ge 2\} < 1$, we know $o(\gamma_n^2/(np_n)) = o(1)$ as $n \to \infty$. Since $d_{BL}(\cdot, \cdot)$ is a metric, by the triangle inequality, we finally conclude from (3.9) and (3.12)that

$$d_{BL}\left(\hat{\mu}(\gamma_n \mathbf{L}_n + \rho_n \mathbf{I}_n), \mu_{cir}\right) \to 0 \quad a.s.$$

as $n \to \infty$. Equivalently, with probability one,

$$\frac{1}{n}\sum_{i=1}^n I\left(\sqrt{\frac{np_n}{1-p_n}}(1 - \lambda_i(\mathbf{L}_n)) \le x\right), \quad x \in \mathbb{R},$$

converges weakly to μ_{cir} as $n \to \infty$. ■

Acknowledgement. The author thanks Dr. Xue Ding for very helpful discussions on the proof of Theorem 1.

References

[1] Z. D. Bai and J. Silverstein, Spectral Analysis of Large Dimensional Random Matrices. *Science Press* (2006).

[2] Z. D. Bai, Methodologies in spectral analysis of large dimensional random matrices, a review. *Statistica Sinica* **9** (1999), 611-677.

[3] M. Bauer and O. Golinelli, Random incidence matrices: moments of the spectral density. *J. Stat. Phys.* **103** (2001), 301-337.

[4] N. L. Biggs, E. K. Lloyd and R. J. Wilson, Graph Theory 1736-1936. *Clarendon Press, Oxford* (1976).

[5] B. Bollobás, Random Graphs. *Academic Press, New York* (1985).

[6] W. Bryc, A. Dembo and T. Jiang, Spectral measure of large random Hankel, Markov and Toeplitz matrices. *Ann. Probab.* **34** (2006), 1-38.

[7] Y. Chow and H. Teicher, Probability Theory, Independence, Interchangeability, Martingales. *Springer and New York* (1998), 2nd Ed.

[8] F. Chung, Spectral Graph Theory (CBMS Regional Conference Series in Mathematics, No. **92**). *American Mathematical Society* (1997).

[9] F. Chung and L. Lu, Complex Graphs and Networks (CBMS Regional Conference Series in Mathematics, No. **107**). *American Mathematical Society* (2006).

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

[11] F. Chung, L. Lu and V. Vu, The spectra of random graphs with given expected degrees. *Proceedings of National Academy of Science*, **100** (2003), 6313-6318.

[12] Y. Colin de Verdière, Spectres de Graphes. *Societe Mathematique De France* (1998).

[13] D. Cvetković, M. Doob and H. Sachs, Spectra of Graphs: Theory and Applications. *Johann Ambrosius Barth, Heidelberg* (1995), 3rd ed.

[14] A. Dembo and O. Zeitouni, Large Deviations Techniques and Applications. *Springer* (1998), 2nd edition.

[15] X. Ding and T. Jiang, Spectral distributions of adjacency and Laplacian matrices of random graphs. *Ann. Appl. Probab.* **20** (2010), 2086-2117.

[16] R. Dudley, Real Analysis and Probability. *Cambridge University Press* (2002).

[17] P. Erdös and A. Rényi, On the strength of connectedness of a random graph. *Acta Math. Acad. Sci. Hungar.* **12** (1961b), 261-267.

[18] P. Erdös and A. Rényi, On the evolution of random graphs. *Bull. Inst. Int. Statist. Tokyo* **38** (1961a), 343-347.

[19] P. Erdös and A. Rényi, On the evolution of random graphs. *Magyar Tud. Akad. Mat. Kuató Int. Közl* **5** (1960), 17-61.

[20] P. Erdös and A. Rényi, On random graphs I. *Publ. Math. Debrecen* **6** (1959), 290-297.

[21] S. N. Evangelou, A numerical studies of sparse random matrices. *J. Stat. Phys.* **69** (1992), 361-383.

[22] S. N. Evangelou, Quantum percolation and Anderson transition in dilute system. *Phys. Rev. B.* **27** (1983), 1397-1400.

[23] S. N. Evangelou and E. N. Economou, Spectral density singularities, level statistics, and localization in sparse random matrices. *Phys. Rev. Lett.* **68** (1992), 361-364.

[24] P. Forrester, Log-Gases and Random Matrices. *Princeton University Press* (2010).

[25] Y. V. Fyodorov and A. D. Mirlin, Universality of the level correlation function of sparse random matrices. *J. Phys. A.* **24** (1991), 2273-2286.

[26] A. Juozulynas, The eigenvalues of very sparse random symmetric matrices. *Lithuanian Math. J.* **44** (2004), 62-70.

[27] L. Khorunzhiy, W. Kirsch and P. Müller, Lifshitz tails for spectra of Erdös-Rényi random graphs. *Ann. Appl. Probab.* **16** (2006), 295-309.

[28] A. Khorunzhy and G. J. Rodgers, Eigenvalue distribution of large dilute random matrices. *J. Math. Phys.* **38** (1997), 3300-3320.

[29] O. Khorunzhy, M. Shcherbina and V. Vengerovsky, Eigenvalue distribution of large weighted random graphs. *J. Math. Phys.* **45** (2004), 1648-1672.

[30] G. Kirchhoff, Über die Auflösung der Gleichungen, auf welche man bei der untersuchung der linearen verteilung galvanischer Ströme geföhrt wird. *Ann. Phys. Chem.* **72** (1847), 497-508.

[31] M. Krivelevich and B. Sudakov, The largest eigenvalue of sparse random graphs. *Combin. Probab. Comput.* **12** (2003), 61-72.

[32] R. Kühn, Spectra of sparse random matrices. *J. Phys. A: Math. Theor.* **41** (2008), 295002.

[33] B. D. Mckay, The expected eigenvalue distribution of a large regular graph. *Linear Algebra and Its Appl.* **40** (1981), 203-216.

[34] M. Penrose, Random Geometric Graphs (Oxford Studies in Probability). *Oxford University Press* (2003).

[35] T. Puppe, Spectral Graph Drawing: A Survey. *VDM Verlag* (2008).

[36] G. J. Rogers and A. J. Bray, Density of states of a sparse random matrix. *Phys. Rev. B*, **37** (1988), 3557-3562.

[37] G. J. Rogers and C. De Dominicis, Density of states of sparse random matrices. *J. Phys. A: Math. Gen.* **23** (1990), 1567-1573.

[38] G. Semerjian and L. F. Cugliandolo, Sparse random matrices: the eigenvalue spectrum revisited. *J. Phys. A: Math. Gen.* **35** (2002), 4837-4851.