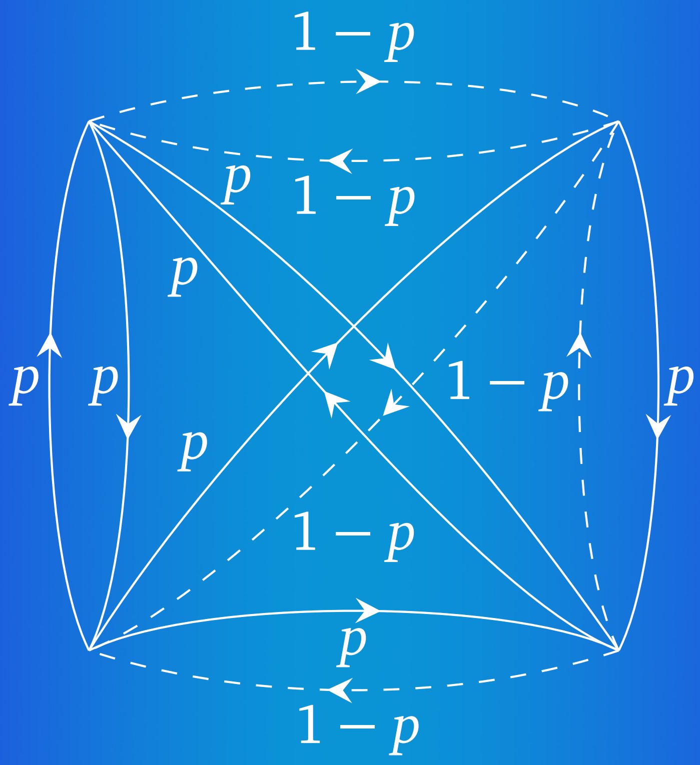
The Spectral Analysis of Random Graph Matrices



Dan Hu

The Spectral Analysis of Random Graph Matrices

Dan Hu

THE SPECTRAL ANALYSIS OF RANDOM GRAPH MATRICES

DISSERTATION

to obtain
the degree of doctor at the University of Twente,
on the authority of the rector magnificus,
prof. dr. T. T. M. Palstra,
on account of the decision of the graduation committee,
to be publicly defended
on Wednesday the 5th of September 2018 at 16.45 hrs

by

Dan Hu

born on the 9th of April 1989
in Yangling, China

This dissertation has been approved by the supervisors
prof. dr. ir. H. J. Broersma and prof. dr. S. Zhang

The research reported in this thesis has been carried out within the framework of the MEMORANDUM OF AGREEMENT FOR A DOUBLE DOCTORATE DEGREE BETWEEN NORTHWESTERN POLYTECHNICAL UNIVERSITY, PEOPLE'S REPUBLIC OF CHINA AND THE UNIVERSITY OF TWENTE, THE NETHERLANDS

DIGITAL SOCIETY INSTITUTE

DSI Ph.D. Thesis Series No. 18-012
Digital Society Institute
P.O. Box 217, 7500 AE Enschede, The Netherlands.

ISBN: 978-90-365-4608-9
ISSN: 2589-7721 (DSI Ph.D. thesis Series No. 18-012)
DOI: 10.3990/1.9789036546089
Available online at `https://doi.org/10.3990/1.9789036546089`

Typeset with LaTeX
Printed by Ipskamp Printing, Enschede
Cover design by Dan Hu

Graduation Committee

Chairman/secretary:

prof. dr. J. N. Kok	University of Twente

Supervisors:

prof. dr. ir. H. J. Broersma	University of Twente
prof. dr. S. Zhang	Northwestern Polytechnical University

Members:

prof. dr. M. Uetz	University of Twente
dr. W. Kern	University of Twente
prof. dr. X. Li	Nankai University
prof. dr. ir. E. R. van Dam	Tilburg University

Preface

The thesis contains six chapters with new results on spectral graph theory (Chapters 2-7), together with an introductory chapter (Chapter 1). Chapters 2 and 3 are mainly based on the research that was done while the author was working as a PhD student at Northwestern Polytechnical University in Xi'an, China; the other chapters are mainly based on the research of the author at the University of Twente, The Netherlands. The purpose of this research was to study the spectra of various matrices and related spectral properties involving several random graph models. The main focus is on analyzing the distributions of the spectra, and estimations of the spectra, as well as on spectral moments, various graph energies, and some other invariants of graphs. This thesis is based on the following papers that have been published in or submitted to scientific journals.

Papers underlying this thesis

[1] The Laplacian energy and Laplacian Estrada index of random multipartite graphs, *Journal of Mathematical Analysis and Applications,* **443** (2016), 675–687 (with X. Li, X. Liu and S. Zhang). (Chapter 2)

[2] The von Neumann entropy of random multipartite graphs, *Discrete Applied Mathematics,* **232** (2017), 201–206 (with X. Li, X. Liu and S. Zhang). (Chapter 2)

[3] The spectral distribution of random mixed graphs, *Linear Algebra and its Applications,* **519** (2017) 343–365 (with X. Li, X. Liu and S. Zhang). (Chapter 3)

[4] The spectra of random mixed graphs, submitted (with H.J. Broersma, J. Hou and S. Zhang). (Chapter 4)

[5] Spectral analysis of normalized Hermitian Laplacian matrices of random mixed graphs, in preparation (with H.J. Broersma, J. Hou and S. Zhang). (Chapter 5)

[6] On the spectra of general random mixed graphs, submitted (with H.J. Broersma, J. Hou and S. Zhang). (Chapter 6)

[7] On the spectra of random oriented graphs, in preparation (with H.J. Broersma, J. Hou and S. Zhang). (Chapter 7)

Contents

Chapter 1

Introduction

Graph theory can be interpreted as the study of binary relations between the elements of a set. In its simplest form, the elements of the set are represented by vertices of the graph, and the binary relation is represented by edges or arcs of the graph: there is an edge or arc in the graph between two vertices if and only if the elements associated with the two vertices are related (If the binary relation is symmetric, this can be represented by an edge; if the binary relation is not symmetric, an arc should be used to indicate the direction of the relation).

Although graph theory is a relatively young area within mathematics, for us mortals it already has a long history, originating with the problem of the *Seven Bridges of Königsberg*, raised by Leonhard Euler in 1735 and solved by him in 1736 [49].

Spectral graph theory is an important study field within graph theory. It mainly focuses on the properties of a graph in relationship to the eigenvalues and eigenvectors of various matrices associated with the graph, as well as on applications. Several different specific matrices can be associated with a given graph, such as its adjacency matrix, its Laplacian matrix, and its normalized Laplacian matrix, to name just a few. The spectra of these matrices, i.e., their (multi)sets of eigenvalues, are called the spectra of the graph. We will study these spectra in detail in this thesis.

The most important themes of spectral graph theory generally include:

relationships between the spectra of graphs and the structure of graphs; estimates, lower and upper bounds for the eigenvalues of graphs; the distribution of the spectra; relations between the spectra of graphs and other invariants of graphs, such as graph energy and spectral moment.

In traditional graph theoretical problems, the graphs are considered to be fixed (deterministic) and their associated matrices contain constant fixed entries. However, for more realistic and complicated network applications containing stochastic elements, the corresponding graphs result in random matrices, and the traditional approaches are no longer feasible. Indeed, the size of such realistic networks typically ranges from hundreds of thousands to billions of vertices, and the corresponding huge and random data poses new difficulties and challenges.

In the 1950s, Erdős and Rényi [48] founded the theory of random graphs. Since then, random graph theory has been one of the fundamental approaches within the research of complex networks. It is an interdisciplinary field between graph theory and probability theory. The simplest random graph model, known as the Erdős-Rényi random graph, was developed by Erdős and Rényi [48] and Gilbert [59]. The Erdős-Rényi random graph $\mathcal{G}_n(p)$ consists of all graphs on n vertices in which the edges are chosen independently with probability p, where $0 < p < 1$. This edge probability can be fixed, or, in more interesting scenarios, a function of n. Random graph theory has developed quickly and considerably in recent years [4, 19, 78], due to its many applications in different real world problems. These include, but are not limited to disperse areas such as telephone and information networks, contact and social networks, and biological networks [90, 94].

A random matrix is a matrix with entries consisting of random values from some specified distribution. Many different random matrices can be associated with a random graph. The spectra of these corresponding matrices are called the spectra of the random graph. The spectra of random graphs are critical to understanding the properties of random graphs. However, there is a relatively small amount of existing literature about the spectral properties of random graphs. This is the main motivation for the work in this thesis.

In the subsequent chapters, we investigate the spectra and spectral properties of several random graph models, such as estimates for the eigenvalues of random graphs, the distribution of their spectra, and relationships between the spectra of random graphs and other invariants. Our researches are mainly based on the following models: the random multipartite graph model, the random mixed graph model, and the random oriented graph model. Among those models, the random mixed graph model is initially proposed and analyzed in this thesis. We finish this section with a short overview of the main contributions of this thesis. In the next section, we give more details, accompanied by the necessary terminology and notation.

1. Random multipartite graph model

The random multipartite graph model can be seen as a generalization of the Erdős-Rényi random graph model. Both models play an important role by serving as relatively simple objects approximating arbitrarily large graphs. Evidently, one can immediately calculate some spectral invariants of a graph by first computing the eigenvalues of the graph. However, it is rather difficult to give an exact expression for the value of the eigenvalues of a large random matrix. In Chapter 2, we estimate the eigenvalues of the Laplacian matrices of random multipartite graphs, and we investigate the relationships between the spectra of these random graphs and other invariants of these graphs, such as the Laplacian energy, the Laplacian Estrada index and the von Neumann entropy.

2. Random mixed graph model

The second part of the thesis consists of Chapters 3, 4, 5 and 6. Results about eigenvalues of digraphs (directed graphs) are sparse. One important reason for this is, that the adjacency matrix of a digraph is usually difficult to work with. In [67], Guo and Mohar showed that mixed graphs are equivalent to digraphs if we regard (replace) each undirected edge as (by) two oppositely directed arcs. A different Hermitian matrix which captures the adjacencies of the digraph is introduced. In this part, motivated by the work of Guo and Mohar, we initially propose a new random graph model – the random mixed graph. Each arc is determined by an independent random variable. More generally, one could have different probabilities assigned to different arcs. We investigate some spectral properties of these random graphs, such as the distributions of the spectra, estimates of the spectra, spectral moments and energies. Moreover, for general random mixed graphs, we estimate the spectrum of the Hermitian adjacency matrix, and we prove a result expressing the concentration of the spectrum of the normalized Hermitian Laplacian matrix.

3. Random oriented graph model

The third part of the thesis is Chapter 7. A natural notion of a random digraph is that of a random orientation of a fixed undirected graph. Starting with a graph, we orient each edge with equal probability for the two possible directions, and independently of all other edges. This model has been studied previously in for instance [2, 64, 87, 95]. In Chapter 7, we investigate the correlation in general random graphs, that is, every edge exists with a different probability, independently of the other edges. From a general random graph, we get a directed graph, which is a random oriented graph, obtained as described above. Eigenvalues of various matrices of random graphs have been related to numerous properties of these graphs. Among these, the spectral radii of different matrices of the graph, i.e., the largest absolute value of eigenvalues of the corresponding matrices, have received the most attention. The investigation on the spectral radii of different matrices of a graph is an important topic in the theory of graph spectra. In Chapter 7, we estimate upper bounds for the spectra radii of the skew adjacency matrix and skew Randić matrix of random oriented graphs.

In the remainder of this chapter, we give a brief account of our main results, and we also formally introduce the three random graph models we consider in this thesis.

1.1 Terminology and notation

This section gives some notations, definitions and preliminary results that we will use throughout the thesis. For terminology and notation not defined here, we refer the reader to [21, 22, 25, 38, 39, 74, 117].

We use $G = (V(G), E(G))$ to denote a graph with vertex set $V(G)$ and edge set $E(G)$. We denote the numbers of vertices and edges in G by $|V(G)|$ and $|E(G)|$, and call these cardinalities the *order* and *size* of G, respectively. A graph is *finite* if its order and size are both finite. For a vertex $v \in V(G)$, we use $N_G(v)$ to denote the *neighborhood* of v, i.e., the set of all vertices adjacent to v. The *degree* of a vertex v in a graph G, denoted by $d_G(v)$, is the number of edges of G incident with v, with each loop counting as two edges. In particular, if G is a simple graph (without loops or multiple edges), $d_G(v) = |N_G(v)|$.

A *complete graph* is a graph in which every pair of distinct vertices is adjacent, and an *edgeless graph* is a graph in which no vertices are adjacent. As usual, we use K_n (respectively, nK_1) to denote the complete graph (respectively, edgeless graph) on n vertices.

A *walk* of length l in G is a sequence $v_0, e_1, v_1, \ldots, v_{l-1}, e_l, v_l$, whose terms are alternately vertices and edges of G (not necessarily distinct), such that $e_i = v_{i-1}v_i \in E(G)$ for all $i \in \{1, 2, \ldots, l\}$. A walk is *closed* if its initial and terminal vertices are identical, and is a *path* if all its vertices and edges are distinct. A closed walk $v_0, e_1, v_1, \ldots, v_{l-1}, e_l, v_l$ of length $l \geq 3$ is a *cycle* if $v_0, e_1, v_1, \ldots, v_{l-1}$ is a path. A graph is said to be *connected* if it contains a path between any pair of distinct vertices, and *disconnected* otherwise. A *tree* is a connected graph without simple cycles.

A graph $G' = (V', E')$ is a *subgraph* of G if $V' \subseteq V(G)$ and $E' \subseteq E(G)$. For a nonempty subset X of $V(G)$, we use $G[X] = (X, E_X)$ to denote the subgraph of G *induced* by X, where $E_X = \{v_iv_j \in E(G) \mid v_i, v_j \in X\}$. A graph G is called a *k-partite graph* if $V(G)$ can be partitioned into k disjoint subsets $V_1, V_2, \ldots, V_k$ such that each $G[V_i]$ is an edgeless graph; such a partition $(V_1, V_2, \ldots, V_k)$ is called a *k-partition* of G, and $V_1, V_2, \ldots, V_k$ its *parts*. In addition, if any two vertices in distinct parts are adjacent in G, then G is said to be a *complete k-partite graph*. As usual, we use $K_{n_1,n_2,\ldots,n_k}$ to denote the complete k-partite graph with $(|V_1|, |V_2|, \ldots, |V_k|) = (n_1, n_2, \ldots, n_k)$.

Let G be a simple undirected graph with vertex set $V_G = \{v_1, v_2, \ldots, v_n\}$ and edge set E_G. The *adjacency matrix* $A(G)$ of G is the symmetric matrix $(A_{ij})_{n\times n}$, where $A_{ij} = A_{ji} = 1$ if vertices v_i and v_j are adjacent, and

$A_{ij} = A_{ji} = 0$ otherwise. We denote by $\lambda_i(A(G))$ the i-th largest eigenvalue of $A(G)$ (multiplicities counted). We use $\{\lambda_1(A(G)), \lambda_2(A(G)), \ldots, \lambda_n(A(G))\}$ to denote the *spectrum* of $A(G)$ in nonincreasing order. The set of these eigenvalues is called the *(adjacency) spectrum* (or *A-spectrum*) of G. Let $d_G(v_i)$ denote the *degree* of the vertex v_i. Denote by $d_G = \Sigma_{v_i \in V_G} d_G(v_i)$ the *degree sum* of G. The *Laplacian matrix* of G is the matrix $L(G) = D(G) - A(G)$, where $D(G)$, called the *degree matrix*, is a diagonal matrix with as diagonal entries the degrees of the vertices of G. We denote by $\mu_i(L(G))$ the i-th largest eigenvalue of $L(G)$ (multiplicities counted). We use $\{\mu_1(L(G)), \mu_2(L(G)), \ldots, \mu_n(L(G))\}$ to denote the spectrum of $L(G)$ in nonincreasing order. The set of these eigenvalues is called the *Laplacian spectrum* of G.

A *Hermitian matrix* (sometimes called *self-adjoint matrix*) is a complex square matrix that is equal to its own *conjugate transpose*, i.e., the (i,j)-th element is equal to the complex conjugate of the (j,i)-th element, for all indices i and j. Hence, a matrix $M = [m_{ij}]$ is Hermitian if for all i,j, we have $m_{ij} = \overline{m_{ji}}$. We let $\mathbb{C}^{n\times n}_{Herm}$ denote the set of $n \times n$ Hermitian matrices, which is a subset of the set $\mathbb{C}^{n\times n}$ of all $n \times n$ matrices with complex entries. For each matrix $M \in \mathbb{C}^{n\times n}$, the *spectral radius* of M is the nonnegative real number $\rho(M) = \max\{|\lambda_i(M)| : 1 \le i \le n\}$, where $\lambda_i(M)$ $(1 \le i \le n)$ are all eigenvalues of M. We use $\lambda_{\max}(M)$ to denote the largest eigenvalue of M. The set $\{\lambda_i(M) : 1 \le i \le n\}$ is called the *spectrum* of M, and denoted by $spec(M)$. The *spectral norm* $\|M\|$ is the largest *singular value* of M, i.e., we have

$$\|M\| = \sqrt{\lambda_{\max}(M^*M)}.$$

Here M^* is the *conjugate transpose* of M. The *Spectral Theorem* for Hermitian matrices states that all $M \in \mathbb{C}^{n\times n}_{Herm}$ have n real eigenvalues (possibly with repetitions) that correspond to an orthonormal set of eigenvectors.

When $M \in \mathbb{C}^{n\times n}_{Herm}$, we have $\|M\| = \max\{|\lambda_i(M)| : 1 \le i \le n\}$. Then $\rho(M) = \|M\| = \max\{\lambda_{\max}(M), \lambda_{\max}(-M)\}$. We use $\mathrm{Tr}(M)$ (the *trace* of M) to denote the sum of the eigenvalues of M.

We say that an event in a probability space holds asymptotically *almost surely* (*a.s.* for short) if its probability goes to one as n tends to infinity. Given a random graph model $\mathscr{G}(n,p)$, we are interested in what properties graphs

$G \in \mathcal{G}(n,p)$ have with high probability. In particular, we say that a property $\mathscr{A}$ holds in $\mathcal{G}(n,p)$ asymptotically *almost surely* (*a.s.* for short), if

$$\lim_{n\to\infty} \Pr(G \in \mathcal{G}(n,p) \text{ has the property } \mathscr{A}) = 1,$$

or we say that *almost all graphs $G \in \mathcal{G}(n,p)$ have property $\mathscr{A}$*, or we say G almost surely (a.s.) satisfies the property $\mathscr{A}$.

We shall use the following standard asymptotic notations throughout. Let $f(n)$, $g(n)$ be two functions of n. Then $f(n) = o(g(n))$ means that $f(n)/g(n) \to 0$, as $n \to \infty$; $f(n) = O(g(n))$ means that there exists a constant C such that $|f(n)| \leq Cg(n)$, as $n \to \infty$; $f(n) = \Omega(g(n))$ means that there exists a constant $C > 0$ such that $f(n) \geq Cg(n)$.

We shall also use standard matrix notation throughout. In particular, the $n \times n$ matrix with every entry equal to 1 will be denoted by J_n, or J if the dimension is understood. The $n \times n$ identity matrix will be denoted by I_n, or I if the dimension is understood.

As we will examine the spectra of random graphs, we will require an understanding of random matrices for several of our main results. A random matrix M is a matrix in which each entry is a random variable. We write $\mathbb{E}(M)$ to denote the coordinate-wise expectation of M, so $\mathbb{E}(M)_{ij} = \mathbb{E}(M_{ij})$. We define the variance matrix in an analogous way to one-dimensional random variables, so $\mathrm{Var}(M) = \mathbb{E}[(M - \mathbb{E}(M))(M - \mathbb{E}(M))^*]$. In particular, for a square Hermitian matrix M, $\mathrm{Var}(M) = \mathbb{E}[(M - \mathbb{E}(M))^2]$.

Other notations and definitions that are not included here will appear at the first place where they are needed in the thesis.

1.2 Random multipartite graphs

We use $K_{n;\beta_1,\ldots,\beta_k}$ to denote the complete k-partite graph of order n, for which the vertex set is the disjoint union of the nonempty parts $V_1, \ldots, V_k$ ($2 \leq k = k(n) \leq n$) satisfying $|V_i| = n\beta_i = n\beta_i(n)$, $i = 1, 2, \ldots, k$. The random k-partite graph model $\mathcal{G}_{n;\beta_1,\ldots,\beta_k}(p)$ consists of all random k-partite graphs in which the edges are chosen independently with probability p from the set of edges of

$K_{n;\beta_1,\ldots,\beta_k}$. We denote by $A_{n,k} := A(G_{n;\beta_1,\ldots,\beta_k}(p)) = (x_{ij})_{n\times n}$ the adjacency matrix of a random k-partite graph $G_{n;\beta_1,\ldots,\beta_k}(p) \in \mathcal{G}_{n;\beta_1,\ldots,\beta_k}(p)$, where x_{ij} is a random indicator variable for v_iv_j being an edge with probability p, for $i \in V_l$ and $j \in V \setminus V_l$, $i \neq j$, $1 \leq l \leq k$. Then $A_{n,k}$ satisfies the following properties:

- x_{ij}'s, $1 \leq i < j \leq n$, are independent random variables with $x_{ij} = x_{ji}$;
- $\Pr(x_{ij} = 1) = 1 - \Pr(x_{ij} = 0) = p$ if $i \in V_l$ and $j \in V \setminus V_l$, while $\Pr(x_{ij} = 0) = 1$ if $i \in V_l$ and $j \in V_l$, $1 \leq l \leq k$.

Note that when $k = n$, then $\mathcal{G}_{n;\beta_1,\ldots,\beta_k} = \mathcal{G}_n(p)$, that is, the random multipartite graph model can be viewed as a generalization of the Erdős-Rényi model.

The *energy* of a graph G of order n is defined as the sum of the absolute values of its eigenvalues. i.e.,

$$\mathcal{E}(G) = \sum_{i=1}^{n} |\lambda_i|.$$

This notion was first introduced by Gutman [68] in 1978. It is a graph parameter that arose from the Hückel molecular orbital approximation for the total π-electron energy [121] from chemistry. Since then, graph energy has been studied extensively by lots of mathematicians and chemists. For results on the study of the energy of graphs, we refer the reader to the book [83] and the more recent book [71].

In 2006, Gutman *et al.* [72] introduced a new matrix $\widetilde{L}(G)$ for a graph G of order n, *i.e.*,

$$\widetilde{L}(G) := L(G) - \sum_{i=1}^{n} \frac{d_G(v_i)}{n} I_n = L(G) - 2\sum_{i=1}^{n}\sum_{i>j} \frac{A_{ij}}{n} I_n.$$

Based on $\widetilde{L}(G)$, they defined the *Laplacian energy* of G as

$$\mathcal{E}_L(G) = \sum_{i=1}^{n} |\mu_i - 2m/n| = \sum_{i=1}^{n} |\xi_i|, \tag{1.1}$$

where m is the number of edges of G, $\mu_1, \mu_2, \ldots, \mu_n$ are the eigenvalues of $L(G)$, and $\xi_1, \xi_2, \ldots, \xi_n$ are the eigenvalues of $\widetilde{L}(G)$. Obviously, the Laplacian energy can be regarded as a variant of the graph energy. Up until now, a lot of results have been obtained on the Laplacian energy. The interested reader is referred to [27, 40, 41, 56, 114, 126]. In [45], Du *et al.* have considered the Laplacian energy of the Erdős-Rényi model $\mathcal{G}_n(p)$. They obtained a lower bound and an upper bound for the Laplacian energy of $\mathcal{G}_n(p)$, and showed that for almost all $G_n(p) \in \mathcal{G}_n(p)$, $\mathcal{E}(G_n(p))$ is no more than $\mathcal{E}_L(G_n(p))$.

In 2009, Fath-Tabar *et al.* [51] first proposed the *Laplacian Estrada index* of graphs. For a graph G of order n, its Laplacian Estrada index is defined as

$$LEE_1(G) = \sum_{i=1}^{n} e^{\mu_i}.$$

Independently, also in 2009, Li *et al.* [84] defined the *Laplacian Estrada index* of G as

$$LEE_2(G) = \sum_{i=1}^{n} e^{\mu_i - 2m/n} = \sum_{i=1}^{n} e^{\xi_i}. \tag{1.2}$$

Clearly, $LEE_1(G) = e^{2m/n} LEE_2(G)$. Thus, these two definitions of the Laplacian Estrada index are essentially equivalent. In this thesis, we adopt Definition (1.2) and denote $LEE_2(G)$ simply by $LEE(G)$ for convenience. For more properties of this index, we refer the interested reader to [15, 42, 51, 77, 84, 127].

The *von Neumann entropy* was originally introduced by von Neumann around 1927 for proving the irreversibility of quantum measurement processes in quantum mechanics [115]. It is defined to be

$$\boldsymbol{S} = -\sum_{i} \zeta_i \log_2 \zeta_i,$$

where ζ_i are the eigenvalues of the density matrix describing the quantum-mechanical system (Normally, a density matrix is a positive semidefinite matrix whose trace is equal to 1). Up until now, there are lots of studies on the von Neumann entropy, and we refer the interested reader to

[5, 6, 9, 85, 96, 99, 100, 104, 110, 115, 125].

In [24], Braunstein *et al.* defined the *density matrix of a graph G* as

$$P_G := \frac{1}{d_G} L(G) = \frac{1}{\mathrm{Tr}(D(G))} L(G),$$

where $d_G = \sum_{v_i \in V_G} d_G(v_i) = \mathrm{Tr}(D(G))$ is the *degree sum* of G, and $\mathrm{Tr}(D(G))$ is the *trace* of $D(G)$. Suppose that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n = 0$ are the eigenvalues of P_G. Then

$$\boldsymbol{S}(G) := -\sum_{i=1}^{n} \lambda_i \log_2 \lambda_i,$$

is called the *von Neumann entropy of the graph G*. By convention, we define $0 \log_2 0 = 0$. It is known that the von Neumann entropy can be interpreted as a measure of the regularity of graphs [101], and also that it can be used as a measure of the graph complexity [73].

Up until now, lots of results on the von Neumann entropy of a graph have been given. For example, Braunstein *et al.* [24] proved that, for a graph G on n vertices, $0 \leq \boldsymbol{S}(G) \leq \log_2(n-1)$, with the left equality holding if and only if G is a graph with only one edge, and the right equality holding if and only if G is the complete graph K_n. In [102], Passerini and Severini showed that the von Neumann entropy of regular graphs with n vertices tends to $\log_2(n-1)$ as n tends to ∞. More interestingly, in [47], Du *et al.* considered the von Neumann entropy of the Erdős-Rényi model $\mathscr{G}_n(p)$. They proved that, for almost all $G_n(p) \in \mathscr{G}_n(p)$, almost surely $\boldsymbol{S}(G_n(p)) = (1+o(1)) \log_2 n$, independently of p.

In Chapter 2, we study the Laplacian energy, the Laplacian Estrada index and the von Neumann entropy for the random k-partite graph model $\mathscr{G}_{n;\beta_1,\ldots,\beta_k}(p)$. In particular, we establish asymptotic lower and upper bounds for $\mathscr{E}_L(G_{n;\beta_1,\ldots,\beta_k}(p))$, $LEE(G_{n;\beta_1,\ldots,\beta_k}(p))$ and $\boldsymbol{S}(G_{n;\beta_1,\ldots,\beta_k})$, respectively, for almost all $G_{n;\beta_1,\ldots,\beta_k}(p) \in \mathscr{G}_{n;\beta_1,\ldots,\beta_k}(p)$, by analyzing the limiting behaviour of the spectra of random symmetric matrices.

1.3 Random mixed graphs

A graph is called a *mixed graph* if it contains both directed and undirected edges. We use $G = (V(G), E_0(G), E_1(G))$ to denote a mixed graph with a set $V(G)$ of vertices, a set $E_0(G)$ of (undirected) edges, and a set $E_1(G)$ of arcs (directed edges). We define the *underlying graph* of G, denoted by $\Gamma(G)$, as the graph with vertex set $V(\Gamma(G)) = V(G)$, and edge set

$$E(\Gamma(G)) = \{v_iv_j \mid v_iv_j \in E_0(G) \text{ or } (v_i, v_j) \in E_1(G) \text{ or } (v_j, v_i) \in E_1(G)\}.$$

We adopt the terminology and notation of Liu and Li in [88], and define the *Hermitian adjacency matrix* of a mixed graph G of order n to be the $n \times n$ matrix $H(G) = (h_{ij})_{n\times n}$, where

$$h_{ij} = \begin{cases} 1, & \text{if } v_iv_j \in E_0(G); \\ \mathbf{i}, & \text{if } (v_i, v_j) \in E_1(G) \text{ and } (v_j, v_i) \notin E_1(G); \\ -\mathbf{i}, & \text{if } (v_i, v_j) \notin E_1(G) \text{ and } (v_j, v_i) \in E_1(G); \\ 0, & \text{otherwise.} \end{cases}$$

Here, $\mathbf{i} = \sqrt{-1}$. This matrix, that is indeed Hermitian, as one easily sees, was also introduced independently by Guo and Mohar in [67]. We denote by $\lambda_i(H(G))$ the i-th largest eigenvalue of $H(G)$ (multiplicities counted). We use $\{\lambda_1(H(G)), \ldots, \lambda_n(H(G))\}$ to denote the *spectrum* of $H(G)$ in nonincreasing order. The set of these eigenvalues is called the *Hermitian adjacency spectrum* (or *H-spectrum*) of G. Let $V(G) = \{v_1, v_2, \ldots, v_n\}$, and let $D(G) = \operatorname{diag}(d_1, d_2, \ldots, d_n)$ be a diagonal matrix, in which d_i is the degree of the vertex v_i in $\Gamma(G)$. Then the matrix $L(G) = D(G) - H(G)$ is called the *Hermitian Laplacian matrix* of G, and the matrix $\mathscr{L}(G) = I - D(G)^{-\frac{1}{2}}H(G)D(G)^{-\frac{1}{2}}$ is called the *normalized Hermitian Laplacian matrix* of G. Here I is the $n \times n$ identity matrix. We denote by $\lambda_i(\mathscr{L}(G))$ the i-th largest eigenvalue of $\mathscr{L}(G)$ (multiplicities counted). We use $\{\lambda_1(\mathscr{L}(G)), \ldots, \lambda_n(\mathscr{L}(G))\}$ to denote the spectrum of $\mathscr{L}(G)$ in nonincreasing order. The set of these eigenvalues is called the *normalized Hermitian Laplacian spectrum* of G.

If we regard (replace) each edge $v_iv_j \in E_0(G)$ in $G = (V(G), E_0(G), E_1(G))$

as (by) two oppositely directed arcs (v_i, v_j) and (v_j, v_i), then G is a directed graph. Throughout the thesis, we regard mixed graphs as directed graphs, in the above sense.

Next, we give the definition of a *general random mixed graph* $\widehat{G}_n(p_{ij})$. Let K_n be a complete graph on n vertices. The *complete directed graph* DK_n is the graph obtained from K_n by replacing each edge of K_n by two oppositely directed arcs. Let p_{ij} be a function of n such that $0 < p_{ij} < 1$ $(i \neq j)$. We always assume that $p_{ii} = 0$ for all indices i. The random mixed graph model $\widehat{\mathcal{G}}_n(p_{ij})$ consists of all random mixed graphs $\widehat{G}_n(p_{ij})$ in which each arc (v_i, v_j) with $i \neq j$ is chosen randomly and independently, with probability p_{ij} from the set of arcs of DK_n, where we let the vertex set be $\{v_1, v_2, \ldots, v_n\}$. Here the probabilities p_{ij} for different arcs are not assumed to be equal, that is, $\widehat{G}_n(p_{ij})$ is an arc-independent random mixed graph of order n. Then the *Hermitian adjacency matrix* of $\widehat{G}_n(p_{ij})$, denoted by $H(\widehat{G}_n(p_{ij})) = (h_{ij})$ (or H_n, for brevity), satisfies that:

- H_n is a random Hermitian matrix, with $h_{ii} = 0$ for $1 \leq i \leq n$;
- the upper-triangular entries h_{ij}, $1 \leq i < j \leq n$ are independent random variables, which take value 1 with probability $p_{ij}p_{ji}$, $\mathbf{i}$ with probability $p_{ij}(1 - p_{ji})$, $-\mathbf{i}$ with probability $(1 - p_{ij})p_{ji}$, and 0 with probability $(1 - p_{ij})(1 - p_{ji})$.

1.3.1 The semicircle law for $\widehat{G}_n(p)$

Let $\{M_n\}_{n=1}^{\infty}$ be a sequence of $n \times n$ random Hermitian matrices. Suppose that $\lambda_1(M_n), \lambda_2(M_n), \ldots, \lambda_n(M_n)$ are the eigenvalues of M_n. The *empirical spectral distribution* (ESD) of M_n is defined as

$$F^{M_n}(x) = \frac{1}{n}\#\{\lambda_i(M_n) \mid \lambda_i(M_n) \leq x, i = 1, 2, \ldots, n\},$$

where $\#\{\cdot\}$ is the cardinality of the set. The distribution to which the ESD of M_n converges as $n \to \infty$ is called the *limiting spectral distribution* (LSD) of $\{M_n\}_{n=1}^{\infty}$.

The ESD of a random Hermitian matrix has a very complicated form when the order of the matrix is large. In particular, it seems very difficult to characterize the LSD of an arbitrary given sequence of random Hermitian matrices. A pioneering work on the spectral distribution of random Hermitian matrices [12,93] we owe to Wigner, is now known as *Wigner's semicircle law* [119, 120]. Wigner's semicircle law characterizes the LSD of a certain type of random Hermitian matrices. This type of random Hermitian matrices is now usually called *Wigner matrices*, denoted by $\boldsymbol{X}_n = (x_{ij})_{n\times n}$, satisfying that

- $\boldsymbol{X}_n$ is an $n \times n$ random Hermitian matrix;
- the upper-triangular entries x_{ij}, $1 \le i < j \le n$, are i.i.d. complex random variables with zero mean and unit variance;
- the diagonal entries x_{ii}, $1 \le i \le n$, are i.i.d. real random variables, independent of the upper-triangular entries, with zero mean; and
- for each positive integer k, $\max\left\{\mathbb{E}(|x_{11}|^k), \mathbb{E}(|x_{12}|^k)\right\} < \infty$.

We state Wigner's semicircle law as follows.

Theorem 1.1. ([120]) *Let $\{\boldsymbol{X}_n\}_{n=1}^{\infty}$ be a sequence of Wigner matrices. Then the ESD of $n^{-1/2}\boldsymbol{X}_n$ converges to the standard semicircle distribution whose density is given by*

$$\phi(x) := \begin{cases} \frac{1}{2\pi}\sqrt{4-x^2}, & \textit{for } |x| \le 2, \\ 0, & \textit{for } |x| > 2. \end{cases}$$

Wigner's semicircle law has been generalized to more general random matrices by lots of researchers, including Arnold [7, 8], Grenander [63], Bai and Yin [10–14, 122], Geman [58], Girko [60–62], Loève [89], and others. More interestingly, it was generalized to random graphs in recent years. Adopting the classical random graphs based on the Erdős-Rényi random graph model $\mathscr{G}_n(p)$, Füredi and Komlós [57] proved that the spectrum of the adjacency matrix follows Wigner's semicircle law. Ding *et al.* [43] considered the spectral distributions of adjacency and Laplacian matrices of random graphs; Du *et al.* [45, 85] considered the spectral distributions of adjacency and Laplacian matrices of the Erdős-Rényi model, and the spectral distribution of adjacency matrices of random multipartite graphs; and Chen *et al.* [29] considered the spectral distribution of skew adjacency matrices of random oriented graphs, and the spectral distribution of adjacency matrices of random regular oriented graphs. Jiang [79] studied the spectral properties of the Laplacian matrices, and the normalized Laplacian matrices of the Erdős-Rényi random graph $G_n(p_n)$ for large n. Under the dilute case, that is, with $p_n \in (0,1)$ and $np_n \to \infty$, Jiang proved that the empirical distribution of the eigenvalues of the Laplacian matrix converges to a deterministic distribution, which is the free convolution of the semicircle law and standard normal distribution $N(0,1)$. However, for its normalized version, Jiang proved that the empirical distribution converges to the semicircle law.

Let $\lambda_1(G), \lambda_2(G), \ldots, \lambda_n(G)$ be the eigenvalues of the Hermitian adjacency matrix of a mixed graph G. The *Hermitian energy* of G was first defined by Liu *et al.* [88] in 2015 as

$$\mathscr{E}_H(G) = \sum_{i=1}^{n} |\lambda_i(G)|,$$

which can be regarded as a variant similar to the graph energy [83, 85]. Up until now, various variants on the graph energy of random graphs have been studied, such as the Laplacian energy [45, 75], the signless Laplacian energy [46], the incidence energy [46], and the distance energy [46]. In [29], Chen *et al.* estimated the skew energy of random oriented graphs. Their results were obtained depending on the LSD of random complex Hermitian matrices.

In Chapter 3 and 5, we respectively characterize the limiting spectral distribution of the Hermitian adjacency matrices and the normalized Hermitian Laplacian matrices of random mixed graphs $\widehat{G}_n(p_{ij})$, where $p_{ij} = p = p(n)$ for any $1 \le i, j \le n$ and 0 for $i = j$, for some $p \in (0,1)$. We denote this graph by $\widehat{G}_n(p)$. We prove that the empirical distribution of the eigenvalues of the Hermitian adjacency matrix converges to Wigner's semicircle law, and also that the empirical distribution of the normalized Hermitian Laplacian matrix converges to Wigner's semicircle law. As an application of the LSD of the Hermitian adjacency matrices, we estimate the Hermitian energy of a random mixed graph.

1.3.2 The spectrum of H_n for $\widehat{G}_n(p)$

The field of spectral graph theory is dedicated to the properties of graph eigenvalues and their applications. Questions about spectra are very important in graph theory, as many important parameters of graphs can be characterized by their spectra, largest eigenvalues and spectral gaps.

Given a graph G of order n, let $\lambda_1(A), \ldots, \lambda_n(A)$ be the eigenvalues of the adjacency matrix A of G in nonincreasing order. Adopting the Erdős-Rényi random graph model $\mathcal{G}_n(p)$, Füredi and Komlós [57] showed that asymptotically almost surely $\lambda_1(A) = (1 + o(1))np$ and $\max\{\lambda_2(A), -\lambda_{n-1}(A)\} \leq (2 + o(1))\sqrt{np(1-p)}$ provided $np(1-p) \gg \ln^6 n$. These results were extended to sparse random graphs [52, 80] and general random symmetric matrices [43, 57].

In Chapter 4, we extend these studies to random mixed graphs. Since we only characterize the limiting spectral distribution of the Hermitian adjacency matrices of random mixed graphs in Chapter 3, the result does not describe the behaviour of the largest eigenvalues of the Hermitian adjacency matrices. The purpose of Chapter 4 is to study the spectrum of the Hermitian adjacency matrix of random mixed graphs.

The *k-th spectral moment* of a graph G of order n with (not necessarily distinct) eigenvalues $\lambda_1(G), \lambda_2(G), \ldots, \lambda_n(G)$ is defined as

$$s_k(G) = \sum_{i=1}^{n} \lambda_i^k(G),$$

where $k \geq 0$ is an integer. Spectral moments are related to many combinatorial properties of graphs. For example, the 4th spectral moment was used in [105] to give an upper bound on the energy of a bipartite graph. The spectral moment is an important algebraic invariant which has found applications in networks. In [28], Chen *et al.* gave an estimate for the spectral moment of random graphs.

As an application of the asymptotic behaviour of the spectrum of the Hermitian adjacency matrix, we estimate the spectral moments of random mixed graphs.

1.3.3 The spectra of H_n and $\mathscr{L}_n$ for $\widehat{G}_n(p_{ij})$

Spectra of the adjacency matrix and the normalized Laplacian matrix of graphs have many applications in graph theory. For example, the spectrum of the adjacency matrix of a graph is related to its connectivity and the number of occurrences of specific subgraphs, and also to its chromatic number and its independence number. The spectrum of the normalized Laplacian matrix is related to diffusion on graphs, random walks on graphs, and the Cheeger constant. For more details on these notions, and for more applications of spectra of the adjacency matrix and the normalized Laplacian matrix, we refer the interested reader to two monographs [31, 38].

Also for random graphs, spectra of their adjacency matrices and their normalized Laplacian matrices are well-studied (See, e.g., [3, 32, 33, 35, 36, 43, 52, 55, 57]). We next present a brief account of some of the results that were obtained for random graphs. We refrain from giving an exhaustive overview, and we refer the reader to the sources for more background, and for terminology and notation.

Tropp [113] determined probability inequalities for sums of independent random self-adjoint matrices. Alon, Krivelevich, and Vu [3] studied the concentration of the s-th largest eigenvalue of a random symmetric matrix with independent random entries of absolute value at most one. Friedman *et al.* [53–55] proved that the second largest eigenvalue (in absolute value) of random d-regular graphs is almost surely $(2+o(1))\sqrt{d-1}$ for any $d \geq 4$. Chung, Lu, and Vu [33] studied spectrum of the adjacency matrix of random power law graphs, and spectrum of the normalized Laplacian matrix of random graphs with given expected degrees. Their results on random graphs with given expected degree sequences were supplemented by Coja-Oghlan *et al.* [35, 36] for sparse random graphs. Lu and Peng [91, 92] studied spectra of the adjacency matrix and the normalized Laplacian matrix of edge-independent random graphs, as well as spectrum of the normalized

Laplacian matrix of random hypergraphs. Oliveira [98] considered the problem of approximating the spectra of the adjacency matrix and the normalized Laplacian matrix of random graphs. His results were improved by Chung and Radcliffe [34].

In Chapter 6, we extend these studies to general random mixed graphs. We study the spectra of the Hermitian adjacency matrix and the normalized Hermitian Laplacian matrix of general random mixed graphs.

1.4 Random oriented graphs

Let G be a simple undirected graph with vertex set $V(G) = \{v_1, v_2, \dots, v_n\}$ and edge set $E(G)$. Let $D(G) = \mathrm{diag}(d_1, d_2, \dots, d_n)$ be a diagonal matrix where d_i is the degree of vertex v_i in G.

In 1975, Randić [106] first proposed a molecular structure descriptor which is defined as the sum of $\frac{1}{\sqrt{d_i d_j}}$ over all (unordered) edges $v_i v_j$ of the underlying (molecular) graph G, i.e., $R = R(G) = \sum_{v_i v_j \in E(G)} \frac{1}{\sqrt{d_i d_j}}$. Nowadays, R is referred to as the *Randić index*. In 1998, Bollobás and Erdős [20] generalized this index by defining $R_\alpha = R_\alpha(G) = \sum_{v_i v_j \in E(G)} (d_i d_j)^\alpha$, and called it the *general Randić index*. The (general) Randić index has many chemical applications, and became a popular topic of research in mathematics and mathematical chemistry. For more details, see [23, 81, 82, 107, 108].

Gutman *et al.* [70] pointed out that for analyzing the Randić index it is useful to associate a matrix of order n with the graph G, named the *Randić matrix* $\boldsymbol{R}(G)$, whose (i, j)-entry is defined as

$$\boldsymbol{R}_{ij} = \begin{cases} 0, & \text{if } i = j; \\ \frac{1}{\sqrt{d_i d_j}}, & \text{if the vertices } v_i \text{ and } v_j \text{ of } G \text{ are adjacent;} \\ 0, & \text{if the vertices } v_i \text{ and } v_j \text{ of } G \text{ are not adjacent.} \end{cases}$$

Let $G^{\sigma} = (V(G), E(G^{\sigma}))$ be an *oriented graph* of G with an orientation $\boldsymbol{\sigma}$, which assigns a direction to each edge of G. So, G^{σ} becomes a directed graph with arc set $E(G^{\sigma})$. In this case, G is called the *underlying graph* of G^{σ}. The *skew adjacency matrix* of G^{σ} is the $n \times n$ real skew symmetric matrix

$S(G^\sigma) = (s_{ij})$, where $s_{ij} = 1 = -s_{ji}$ if $(v_i, v_j) \in E(G^\sigma)$, and $s_{ij} = s_{ji} = 0$ otherwise.

In [66], Gu, Huang and Li defined the *skew Randić matrix* $\boldsymbol{R}_S = \boldsymbol{R}_S(G^\sigma)$ of G^σ, whose (i,j)-th entry is

$$(\boldsymbol{R}_s)_{ij} = \begin{cases} \frac{1}{\sqrt{d_i d_j}}, & \text{if } (v_i, v_j) \in E(G^\sigma)\ ; \\ -\frac{1}{\sqrt{d_i d_j}}, & \text{if } (v_j, v_i) \in E(G^\sigma); \\ 0, & \text{otherwise.} \end{cases}$$

If G does not possess isolated vertices, then it is easy to check that $\boldsymbol{R}_S(G^\sigma) = D(G)^{-\frac{1}{2}} S(G^\sigma) D(G)^{-\frac{1}{2}}$.

The *skew spectrum* of G^σ is defined as the spectrum of $S(G^\sigma)$. As the matrix $S(G^\sigma)$ is real and skew symmetric, the spectrum of $S(G^\sigma)$ consists of only purely imaginary eigenvalues or 0. The *skew spectral radius* of G^σ, denoted by $\rho_S(G^\sigma)$, is defined to be the spectral radius of $S(G^\sigma)$. The *skew Randić spectrum* of G^σ is defined as the spectrum of $\boldsymbol{R}_S(G^\sigma)$. The *skew Randić spectral radius* of G^σ, denoted by $\rho_{\boldsymbol{R}_S}(G^\sigma)$, is defined to be the spectral radius of $\boldsymbol{R}_S(G^\sigma)$.

We next give the definition of a random oriented graph $G_n^\sigma(p_{ij})$. Let p_{ij} be a function of n such that $0 < p_{ij} < 1$. A random oriented graph on n vertices is obtained by drawing an edge between each pair of vertices v_i and v_j, randomly and independently, with probability p_{ij} and then orienting the existing edge $v_i v_j$, randomly and independently, with probability $1/2$. Here $p_{ij} = p_{ji}$ and $\{p_{ij}\}_{1 \le i < j \le n}$ are not assumed to be equal. The random oriented graph model $\mathcal{G}_n^\sigma(p_{ij})$ consists of all random oriented graphs $G_n^\sigma(p_{ij})$. Now, the skew adjacency matrix $S(G_n^\sigma(p_{ij})) = (s_{ij})$ (or S_n, for brevity) of $G_n^\sigma(p_{ij})$ is a random matrix such that

- S_n is skew symmetric, i.e., $s_{ij} = -s_{ji}$ for $1 \le i < j \le n$, and $s_{ii} = 0$ for $1 \le i \le n$;
- the upper-triangular entries s_{ij} , $1 \le i < j \le n$ are i.i.d. random variables such that $s_{ij} = 1$ with probability $\frac{p_{ij}}{2}$, $s_{ij} = -1$ with probability $\frac{p_{ij}}{2}$, and $s_{ij} = 0$ with probability $1 - p_{ij}$.

In Chapter 7, we study the spectra of the skew adjacency matrix and the skew Randić matrix of random oriented graphs. In particular, we apply a probability inequality for sums of independent random matrices to give upper bounds for the skew spectral radius and the skew Randić spectral radius of random oriented graphs.

Chapter 2

The Laplacian energy, Laplacian Estrada index and von Neumann entropy of random multipartite graphs

In this chapter, we study the Laplacian energy, the Laplacian Estrada index and the von Neumann entropy of random multipartite graphs, using the k-partite graph model $\mathcal{G}_{n;\beta_1,\ldots,\beta_k}(p)$. We establish asymptotic lower and upper bounds for $\mathcal{E}_L(G_{n;\beta_1,\ldots,\beta_k}(p))$, $LEE(G_{n;\beta_1,\ldots,\beta_k}(p))$ and $\boldsymbol{S}(G_{n;\beta_1,\ldots,\beta_k})$, respectively, for almost all $G_{n;\beta_1,\ldots,\beta_k}(p) \in \mathcal{G}_{n;\beta_1,\ldots,\beta_k}(p)$, by analyzing the limiting behaviour of the spectra of random symmetric matrices.

2.1 The Laplacian energy

In this section, we establish a lower bound and an upper bound for the Laplacian energy of random multipartite graphs $G_{n;\beta_1,\ldots,\beta_k}(p) \in \mathcal{G}_{n;\beta_1,\ldots,\beta_k}(p)$. Before proceeding, we give some additional essential definitions and present some auxiliary lemmas.

Let M be a real symmetric matrix. Denote by $\mathscr{E}(M)$ the sum of the absolute values of the eigenvalues of M. We are going to use the following inequality.

Lemma 2.1 (Fan [50]). *Let X, Y, and Z be real symmetric matrices of order n such that $X+Y=Z$. Then*

$$\mathscr{E}(X)+\mathscr{E}(Y)\geq \mathscr{E}(Z).$$

We will also use the following result in our proof.

Lemma 2.2 (Shiryaev [112]). *Let $X_1, X_2, \ldots$ be an infinite sequence of i.i.d. random variables with expected value $\mathbb{E}(X_1)=\mathbb{E}(X_2)=\cdots=\mu$, and $\mathbb{E}|X_j|<\infty$. Then*

$$\frac{1}{n}(X_1+X_2+\cdots+X_n)\to\mu \ \ a.s.$$

In [45], Du et al. established the following asymptotic lower and upper bounds for the Laplacian energy of Erdős-Rényi random graphs.

Lemma 2.3 (Du *et al.* [45]). *Almost every random graph $G_n(p)$ satisfies*

$$\left(\frac{2\sqrt{2}}{3}\sqrt{p(1-p)}+o(1)\right)n^{3/2}\leq \mathscr{E}_L(G_n(p))\leq\left(\sqrt{2p-p^2}+o(1)\right)n^{3/2}.$$

We are going to extend this result to random multipartite graphs. Let $G_{n;\beta_1,\ldots,\beta_k}(p)\in\mathscr{G}_{n;\beta_1,\ldots,\beta_k}(p)$ with $\beta_1\geq\beta_2\geq\cdots\geq\beta_k$. Note that $\sum_{l=1}^k\beta_l=1$. Then, we have $\beta_k=\sum_{l=1}^k\beta_k\beta_l\leq\sum_{l=1}^k\beta_l^2\leq\sum_{l=1}^k\beta_1\beta_l=\beta_1$. This implies that we can always find an integer r $(1\leq r\leq k-1)$ such that $\beta_{r+1}\leq\sum_{l=1}^k\beta_l^2\leq\beta_r$. We use this in our first main result, as follows.

Theorem 2.4. *Let $G_{n;\beta_1,\ldots,\beta_k}(p)\in\mathscr{G}_{n;\beta_1,\ldots,\beta_k}(p)$ with $\beta_1\geq\beta_2\geq\cdots\geq\beta_k$ and r $(1\leq r\leq k-1)$ be an integer such that $\beta_{r+1}\leq\sum_{l=1}^k\beta_l^2\leq\beta_r$. Then almost surely, $\mathscr{E}_L(G_{n;\beta_1,\ldots,\beta_k}(p))$ is between*
$2(p+o(1))n^2\left(\sum_{l=1}^r\beta_l^2-\beta_r\sum_{l=1}^r\beta_l\right)-\left[\sqrt{2p-p^2}\left(1+\sum_{i=1}^k\beta_i^{3/2}\right)+o(1)\right]n^{3/2}$ *and*
$2(p+o(1))n^2\left(\sum_{l=1}^r\beta_l^2-\beta_{r+1}\sum_{l=1}^r\beta_l\right)+\left[\sqrt{2p-p^2}\left(1+\sum_{i=1}^k\beta_i^{3/2}\right)+o(1)\right]n^{3/2}.$

Proof. Note that the parts $V_1,\ldots,V_k$ of the random k-partite graph $G_{n;\beta_1,\ldots,\beta_k}(p)$ satisfy $|V_i|=n\beta_i,\ i=1,2,\ldots,k$. Then the adjacency matrix $A_{n,k}$ of $G_{n;\beta_1,\ldots,\beta_k}(p)$

satisfies

$$A_{n,k} + A'_{n,k} = A_n,$$

where

$$A'_{n,k} = \begin{pmatrix} A_{n\beta_1} & & & \\ & A_{n\beta_2} & & \\ & & \ddots & \\ & & & A_{n\beta_k} \end{pmatrix}_{n\times n,}$$

and $A_n := A(G_n(p)),\ A_{n\beta_i} := A(G_{n\beta_i}(p)),\ i = 1, 2, \dots, k.$

The degree matrix $D_{n,k} := D(G_{n;\beta_1,\dots,\beta_k}(p))$ of $G_{n;\beta_1,\dots,\beta_k}(p)$ satisfies

$$D_{n,k} + D'_{n,k} = D_n,$$

where

$$D'_{n,k} = \begin{pmatrix} D_{n\beta_1} & & & \\ & D_{n\beta_2} & & \\ & & \ddots & \\ & & & D_{n\beta_k} \end{pmatrix}_{n\times n,}$$

and $D_n := D(G_n(p)),\ D_{n\beta_i} := D(G_{n\beta_i}(p)),\ i = 1, 2, \dots, k.$

The Laplacian matrix $L_{n,k} := L(G_{n;\beta_1,\dots,\beta_k}(p))$ of $G_{n;\beta_1,\dots,\beta_k}(p)$ satisfies

$$L_{n,k} + L'_{n,k} = L_n,$$

where

$$L'_{n,k} = \begin{pmatrix} L_{n\beta_1} & & & \\ & L_{n\beta_2} & & \\ & & \ddots & \\ & & & L_{n\beta_k} \end{pmatrix}_{n\times n,}$$

and $L_n := L(G_n(p)),\ L_{n\beta_i} := L(G_{n\beta_i}(p)),\ i = 1, 2, \dots, k.$

Note that $L_{n,k} = L_n - L'_{n,k}$, $A_{n,k} = A_n - A'_{n,k}$, and

$$\widetilde{L_n} = L_n - \sum_{i=1}^{n} \frac{d_{G_n(p)}(v_i)}{n} I_n = L_n - 2\sum_{i=1}^{n}\sum_{i>j} \frac{(A_n)_{ij}}{n} I_n.$$

Then

$$
\begin{aligned}
&\widetilde{L_{n,k}}\\
=&L_{n,k} - 2\sum_{i=1}^{n}\sum_{i>j}\frac{(A_{n,k})_{ij}}{n}I_n\\
=&L_n - L'_{n,k} - 2\sum_{i=1}^{n}\sum_{i>j}\frac{(A_n - A'_{n,k})_{ij}}{n}I_n\\
=&L_n - 2\sum_{i=1}^{n}\sum_{i>j}\frac{(A_n)_{ij}}{n}I_n - L'_{n,k} + \frac{2}{n}\sum_{l=1}^{k}\sum_{i=1}^{n\beta_l}\sum_{i>j}(A_{n\beta_l})_{ij}I_n\\
=&\widetilde{L_n} - B_n - C_n, \qquad (2.1)
\end{aligned}
$$

where

$$
B_n = \begin{pmatrix} \widetilde{L_{n\beta_1}} & & \\ & \ddots & \\ & & \widetilde{L_{n\beta_k}} \end{pmatrix}_{n\times n}
$$

with

$$
\widetilde{L_{n\beta_l}} = L_{n\beta_l} - 2\frac{\sum_{i=1}^{n\beta_l}\sum_{i>j}(A_{n\beta_l})_{ij}}{n\beta_l}I_{n\beta_l}, \text{ for } 1\le l\le k,
$$

and

$$
C_n = \begin{pmatrix} C_{n\beta_1} & & \\ & \ddots & \\ & & C_{n\beta_k} \end{pmatrix}_{n\times n}
$$

with

$$
C_{n\beta_l} = \left(2\frac{\sum_{i=1}^{n\beta_l}\sum_{i>j}(A_{n\beta_l})_{ij}}{n\beta_l} - \frac{2}{n}\sum_{l=1}^{k}\sum_{i=1}^{n\beta_l}\sum_{i>j}(A_{n\beta_l})_{ij}\right)I_{n\beta_l}, \text{ for } 1\le l\le k.
$$

By (2.1) and Lemma 2.1, we have

$$
|\mathscr{E}(\widetilde{L_n} - B_n) - \mathscr{E}(C_n)| \le \mathscr{E}(\widetilde{L_{n,k}}) \le \mathscr{E}(\widetilde{L_n}) + \mathscr{E}(B_n) + \mathscr{E}(C_n). \qquad (2.2)
$$

Note that

$$\mathscr{E}_L(G_n(p)) = \sum_{i=1}^{n} \left| \mu(L_n) - \frac{\mathrm{Tr}(D_n)}{n} \right| = \sum_{i=1}^{n} \left| \xi_i(\widetilde{L_n}) \right| = \mathscr{E}(\widetilde{L_n}),$$

and

$$\mathscr{E}_L(G_{n,k}(p)) = \sum_{i=1}^{n} \left| \mu_i(L_{n,k}) - \frac{\mathrm{Tr}(D_{n,k})}{n} \right| = \sum_{i=1}^{n} \left| \xi_i(\widetilde{L_{n,k}}) \right| = \mathscr{E}(\widetilde{L_{n,k}}).$$

Then

$$\mathscr{E}(B_n) = \mathscr{E}(\widetilde{L_{n\beta_1}}) + \cdots + \mathscr{E}(\widetilde{L_{n\beta_k}}) = \mathscr{E}_L(G_{n\beta_1}(p)) + \cdots + \mathscr{E}_L(G_{n\beta_k}(p)).$$

Thus, Lemma 2.3 implies that

$$\begin{aligned}
&\mathscr{E}(\widetilde{L_n}) - \mathscr{E}(B_n) \\
=&\mathscr{E}_L(G_n(p)) - [\mathscr{E}_L(G_{n\beta_1}(p)) + \cdots + \mathscr{E}_L(G_{n\beta_k}(p))] \\
\geq&\left(\frac{2\sqrt{2}}{3}\sqrt{p(1-p)} + o(1) \right) n^{3/2} - \left(\sqrt{2p-p^2} + o(1) \right) n^{3/2} \sum_{i=1}^{k} \beta_i^{3/2} \\
=&\left(\frac{2\sqrt{2}}{3}\sqrt{p(1-p)} - \sqrt{2p-p^2} \sum_{i=1}^{k} \beta_i^{3/2} + o(1) \right) n^{3/2} \;\; a.s., \qquad (2.3)
\end{aligned}$$

and

$$\begin{aligned}
&\mathscr{E}(\widetilde{L_n}) + \mathscr{E}(B_n) \\
=&\mathscr{E}_L(G_n(p)) + [\mathscr{E}_L(G_{n\beta_1}(p)) + \cdots + \mathscr{E}_L(G_{n\beta_k}(p))] \\
\leq&\left(\sqrt{2p-p^2} + o(1) \right) n^{3/2} + \left(\sqrt{2p-p^2} + o(1) \right) n^{3/2} \sum_{i=1}^{k} \beta_i^{3/2} \\
=&\left[\sqrt{2p-p^2} \left(1 + \sum_{i=1}^{k} \beta_i^{3/2} \right) + o(1) \right] n^{3/2} \;\; a.s. \qquad (2.4)
\end{aligned}$$

By Lemma 2.1, we have

$$\mathscr{E}(\widetilde{L_n}) - \mathscr{E}(B_n) \leq \mathscr{E}(\widetilde{L_n} - B_n) \leq \mathscr{E}(\widetilde{L_n}) + \mathscr{E}(B_n). \qquad (2.5)$$

Next, by estimating $\mathscr{E}(C_n)$, we compare $\mathscr{E}(\widetilde{L_n}-B_n)$ and $\mathscr{E}(C_n)$. Since $(A_n)_{ij}(i > j)$ are i.i.d. with mean p and variance $p(1-p)$, it follows from Lemma 2.2 that, with probability 1,

$$\lim_{n\to\infty} \frac{2\sum_{i=1}^{n}\sum_{i>j}(A_n)_{ij}}{n(n-1)} = p.$$

Thus, we have

$$\sum_{i=1}^{n}\sum_{i>j}(A_n)_{ij} = (p/2+o(1))n^2 \;\; a.s. \tag{2.6}$$

Similarly, for $l = 1, 2, \dots, k$,

$$\sum_{i=1}^{n\beta_l}\sum_{i>j}(A_{n\beta_l})_{ij} = (p/2+o(1))n^2\beta_l^2 \;\; a.s. \tag{2.7}$$

Since $\beta_1 \geq \cdots \geq \beta_k$ and $\beta_{r+1} \leq \sum_{l=1}^{k}\beta_l^2 \leq \beta_r$, we have

$$\begin{aligned}
\mathscr{E}(C_n) &= \sum_{l=1}^{k}\left|2\frac{\sum_{i=1}^{n\beta_l}\sum_{i>j}(A_{n\beta_l})_{ij}}{n\beta_l} - \frac{2}{n}\sum_{l=1}^{k}\sum_{i=1}^{n\beta_l}\sum_{i>j}(A_{n\beta_l})_{ij}\right|\cdot n\beta_l \\
&= \sum_{l=1}^{k}\left|(p+o(1))n\beta_l - (p+o(1))n\sum_{i=1}^{k}\beta_i^2\right|\cdot n\beta_l \\
&= (p+o(1))n^2\sum_{l=1}^{k}\left|\beta_l - \sum_{i=1}^{k}\beta_i^2\right|\cdot\beta_l \\
&= 2(p+o(1))n^2\left(\sum_{l=1}^{r}\beta_l^2 - \sum_{l=1}^{k}\beta_l^2\cdot\sum_{l=1}^{r}\beta_l\right) \;\; a.s.
\end{aligned}$$

Note that

$$\sum_{l=1}^{r}\beta_l^2 - \sum_{l=1}^{k}\beta_l^2\cdot\sum_{l=1}^{r}\beta_l \geq \sum_{l=1}^{r}\beta_l^2 - \beta_r\cdot\sum_{l=1}^{r}\beta_l \geq 0.$$

Hence

$$\mathscr{E}(C_n) \geq \mathscr{E}(\widetilde{L_n}-B_n). \tag{2.8}$$

Since $\beta_{r+1} \le \sum_{l=1}^{k} \beta_l^2 \le \beta_r$, we have

$$\begin{aligned}
&2(p+o(1))n^2\left(\sum_{l=1}^{r}\beta_l^2-\beta_r\sum_{l=1}^{r}\beta_l\right)\\
\le&\mathscr{E}(C_n)\\
\le&2(p+o(1))n^2\left(\sum_{l=1}^{r}\beta_l^2-\beta_{r+1}\sum_{l=1}^{r}\beta_l\right).
\end{aligned}\tag{2.9}$$

By (2.2), (2.5) and (2.8), we have

$$\begin{aligned}
&\mathscr{E}(C_n)-\left[\mathscr{E}(\widetilde{L_n})+\mathscr{E}(B_n)\right]\\
\le&\mathscr{E}(C_n)-\mathscr{E}(\widetilde{L_n}-B_n)\\
\le&\mathscr{E}(\widetilde{L_{n,k}})\\
\le&\mathscr{E}(C_n)+\mathscr{E}(\widetilde{L_n})+\mathscr{E}(B_n).
\end{aligned}$$

Then by (2.4) and (2.9), we have

$$\begin{aligned}
&2(p+o(1))n^2\left(\sum_{l=1}^{r}\beta_l^2-\beta_r\sum_{l=1}^{r}\beta_l\right)\\
&-\left[\sqrt{2p-p^2}\left(1+\sum_{i=1}^{k}\beta_i^{3/2}\right)+o(1)\right]n^{3/2}\\
\le&\mathscr{E}(\widetilde{L_{n,k}})\\
\le&2(p+o(1))n^2\left(\sum_{l=1}^{r}\beta_l^2-\beta_{r+1}\sum_{l=1}^{r}\beta_l\right)\\
&+\left[\sqrt{2p-p^2}\left(1+\sum_{i=1}^{k}\beta_i^{3/2}\right)+o(1)\right]n^{3/2}\quad a.s.
\end{aligned}$$

This completes the proof. □

Next, we consider the special case in which each part of $G_{n;\beta_1,\ldots,\beta_k}(p) \in \mathscr{G}_{n;\beta_1,\ldots,\beta_k}(p)$ has the same size as n tends to infinity.

Theorem 2.5. *Let $G_{n;\beta_1,\ldots,\beta_k}(p) \in \mathcal{G}_{n;\beta_1,\ldots,\beta_k}(p)$ satisfy $\lim\limits_{n\to\infty} \frac{\beta_i}{\beta_j} = 1,\ 1 \le i,j \le k$. Then almost surely*

$$\begin{aligned}
&\left(\frac{2\sqrt{2p(1-p)}}{3} - \sqrt{\frac{2p-p^2}{k}} + o(1)\right) n^{3/2} \\
\le &\mathcal{E}_L(G_{n;\beta_1,\ldots,\beta_k}(p)) \\
\le &\left[\sqrt{2p-p^2}\left(1+\frac{1}{\sqrt{k}}\right) + o(1)\right] n^{3/2}.
\end{aligned}$$

Proof. We assume that $\lim\limits_{n\to\infty} \frac{\beta_i}{\beta_j} = 1$, for $1 \le i,j \le k$. Using (2.7), for $l,t = 1,\ldots,k$, we obtain

$$\frac{\sum_{i=1}^{n\beta_l}\sum_{i>j}(A_{n\beta_l})_{ij}}{n\beta_l} = \frac{\sum_{i=1}^{n\beta_t}\sum_{i>j}(A_{n\beta_t})_{ij}}{n\beta_t} = \frac{\sum_{l=1}^{k}\sum_{i=1}^{n\beta_l}\sum_{i>j}(A_{n\beta_l})_{ij}}{n} \quad a.s.$$

Then

$$C_n = 0 \quad a.s.$$

So, by (2.1), we have

$$\widetilde{L_{n,k}} = \widetilde{L_n} - B_n \quad a.s.$$

According to Lemma 2.1, we have

$$\mathcal{E}(\widetilde{L_n}) - \mathcal{E}(B_n) \le \mathcal{E}(\widetilde{L_{n,k}}) \le \mathcal{E}(\widetilde{L_n}) + \mathcal{E}(B_n). \tag{2.10}$$

Note that $\lim\limits_{n\to\infty} \frac{\beta_i}{\beta_j} = 1$ implies that $\lim\limits_{n\to\infty} \beta_i = \frac{1}{k}$, for $1 \le i \le k$. From (2.3) and (2.4), we have

$$\begin{aligned}
&\mathcal{E}(\widetilde{L_n}) - \mathcal{E}(B_n) \\
\ge &\left(\frac{2\sqrt{2}}{3}\sqrt{p(1-p)} - \sqrt{2p-p^2}\sum_{i=1}^{k}\beta_i^{3/2} + o(1)\right) n^{3/2} \\
= &\left(\frac{2\sqrt{2}}{3}\sqrt{p(1-p)} - \sqrt{\frac{2p-p^2}{k}} + o(1)\right) n^{3/2} \quad a.s.,
\end{aligned} \tag{2.11}$$

and

$$\begin{aligned}\mathscr{E}(\widetilde{L_n})+\mathscr{E}(B_n) &\leq \left[\sqrt{2p-p^2}\left(1+\sum_{i=1}^{k}\beta_i^{3/2}\right)+o(1)\right]n^{3/2}\\ &=\left[\sqrt{2p-p^2}\left(1+\frac{1}{\sqrt{k}}\right)+o(1)\right]n^{3/2} \quad a.s. \end{aligned}\tag{2.12}$$

Then (2.10), (2.11) and (2.12) imply that

$$\begin{aligned}&\left(\frac{2\sqrt{2p(1-p)}}{3}-\sqrt{\frac{2p-p^2}{k}}+o(1)\right)n^{3/2}\\ \leq&\mathscr{E}_L(G_{n;\beta_1,\ldots,\beta_k}(p))\\ \leq&\left[\sqrt{2p-p^2}\left(1+\frac{1}{\sqrt{k}}\right)+o(1)\right]n^{3/2}.\end{aligned}$$

This completes the proof. □

2.2 The Laplacian Estrada index

In this section, we will establish a lower bound and an upper bound for $LEE(G_{n;\beta_1,\ldots,\beta_k}(p))$ for almost all $G_{n;\beta_1,\ldots,\beta_k}(p) \in \mathscr{G}_{n;\beta_1,\ldots,\beta_k}(p)$. Recall that we use $A_{n,k}$, $L_{n,k}$ and $\widetilde{L_{n,k}}$ to denote $A(G_{n;\beta_1,\ldots,\beta_k}(p))$, $L(G_{n;\beta_1,\ldots,\beta_k}(p))$ and $\widetilde{L}(G_{n;\beta_1,\ldots,\mu_k}(p))$, respectively.

We need the following two lemmas for the proof of our result.

Lemma 2.6 (Bryc *et al.* [26]). *Let X be a symmetric random matrix satisfying that the entries X_{ij}, $1 \leq i < j \leq n$, are a collection of i.i.d. random variables with $\mathbb{E}(X_{12}) = 0, \mathrm{Var}(X_{12}) = 1$ and $\mathbb{E}(X_{12}^4) < \infty$. Define $T :=$ $\mathrm{diag}\left(\sum_{i\neq j}X_{ij}\right)_{1\leq i\leq n}$, and let $M = T - X$, where $\mathrm{diag}\{\cdot\}$ denotes a diagonal matrix. Denote by $\|M\|$ the spectral radius of M. Then*

$$\lim_{n\to\infty}\frac{\|M\|}{\sqrt{2n\ln n}}=1 \quad a.s.$$

Lemma 2.7 (Weyl [118]). *Let X, Y and Z be $n \times n$ Hermitian matrices such that $X = Y + Z$. Suppose that X, Y, Z have eigenvalues, respectively, $\lambda_1(X) \geq$*

$\cdots \geq \lambda_n(X)$, $\lambda_1(Y) \geq \cdots \geq \lambda_n(Y)$, $\lambda_1(Z) \geq \cdots \geq \lambda_n(Z)$. Then for $i = 1, 2, \ldots, n$ the following inequalities hold:

$$\lambda_i(Y) + \lambda_n(Z) \leq \lambda_i(X) \leq \lambda_i(Y) + \lambda_1(Z).$$

Theorem 2.8. *Let $G_{n;\beta_1,\ldots,\beta_k}(p) \in \mathscr{G}_{n;\beta_1,\ldots,\beta_k}(p)$. Then almost surely*

$$\begin{aligned}
&(n-1+e^{-np})e^{np(\sum_{i=1}^k \beta_i^2 - \max\limits_{1\leq i\leq k}\{\beta_i\}) + o(1)n} \\
\leq & LEE(G_{n;\beta_1,\ldots,\beta_k}(p)) \\
\leq & (n-1+e^{-np})e^{np\sum_{i=1}^k \beta_i^2 + o(1)n}.
\end{aligned}$$

Proof. Define an auxiliary matrix

$$\widehat{L_n} := L_n - p(n-1)I_n + p(J_n - I_n) = [D_n - p(n-1)I_n] - [A_n - p(J_n - I_n)],$$

where J_n is the all-ones matrix. Let

$$T = \frac{1}{\sqrt{p(1-p)}}[D_n - p(n-1)I_n]$$

and

$$X = \frac{1}{\sqrt{p(1-p)}}[A_n - p(J_n - I_n)].$$

Then $\mathbb{E}(X_{12}) = 0, \mathrm{Var}(X_{12}) = 1$, and

$$\mathbb{E}(X_{12}^4) = \frac{1}{p^2(1-p)^2}(p - 4p^2 + 6p^3 - 3p^4) < \infty.$$

By Lemma 2.6, we have

$$\lim_{n\to\infty} \frac{||\widehat{L_n}||}{\sqrt{2p(1-p)n\ln n}} = 1 \quad a.s.$$

Then

$$\lim_{n\to\infty} \frac{||\widehat{L_n}||}{n} = 0 \quad a.s.,$$

i.e.,

$$\|\widehat{L_n}\| = o(1)n \quad a.s.$$

Let $Q_n := p(n-1)I_n - p(J_n - I_n)$. Then $\widehat{L_n} + Q_n = L_n$. Suppose that $L_n, \widehat{L_n}, Q_n$ have eigenvalues, respectively, $\mu_1(L_n) \geq \cdots \geq \mu_n(L_n)$, $\lambda_1(\widehat{L_n}) \geq \cdots \geq \lambda_n(\widehat{L_n})$, $\lambda_1(Q_n) \geq \cdots \geq \lambda_n(Q_n)$. It follows from Lemma 2.7 that

$$\lambda_i(Q_n) + \lambda_n(\widehat{L_n}) \leq \mu_i(L_n) \leq \lambda_i(Q_n) + \lambda_1(\widehat{L_n}), \text{ for } i = 1, 2, \ldots, n.$$

Notice that $\lambda_i(Q_n) = pn$ for $i = 1, 2, \ldots, n-1$ and $\lambda_n(Q_n) = 0$. We have

$$\mu_i(L_n) = (p + o(1))n \;\; a.s., \text{ for } 1 \leq i \leq n-1, \tag{2.13}$$

and

$$\mu_n(L_n) = o(1)n \;\; a.s. \tag{2.14}$$

In the following, we first evaluate the eigenvalues of $L_{n,k}$ according to the spectral distribution of L_n and $L'_{n,k}$.

Since $L_{n,k} = L_n - L'_{n,k}$, Lemma 2.7 implies that for $1 \leq i \leq n$,

$$\mu_i(L_n) + \mu_n(-L'_{n,k}) \leq \mu_i(L_{n,k}) \leq \mu_i(L_n) + \mu_1(-L'_{n,k}), \tag{2.15}$$

where $\mu_n(-L'_{n,k})$ and $\mu_1(-L'_{n,k})$ are the minimum and maximum eigenvalues of $-L'_{n,k}$, respectively. By (2.13), (2.14) and (2.15), we have

$$np(1 - \max_{1\leq i\leq k}\{\beta_i\}) + o(1)n \leq \mu_i(L_{n,k}) \leq np + o(1)n \quad a.s., \text{ for } 1 \leq i \leq n-1, \tag{2.16}$$

and

$$-np \max_{1\leq i\leq k}\{\beta_i\} + o(1)n \leq \mu_n(L_{n,k}) \leq o(1)n \quad a.s. \tag{2.17}$$

Now we consider the trace $\mathrm{Tr}(D_{n,k})$ of $D_{n,k}$. Note that $\mathrm{Tr}(D_{n,k}) = 2\sum_{i>j}(A_{n,k})_{ij}$. Since $(A_n)_{ij} (i > j)$ are i.i.d. with mean p and variance $p(1-p)$, according to

Lemma 2.2, we obtain that with probability 1,

$$\lim_{n\to\infty}\frac{\sum_{i>j}(A_n)_{ij}}{\frac{n(n-1)}{2}}=p,$$

i.e.,

$$\sum_{i>j}(A_n)_{ij}=(p/2+o(1))n^2 \ \ a.s.$$

Then

$$\mathrm{Tr}(D_n)=(p+o(1))n^2 \ \ a.s. \tag{2.18}$$

Similarly, for $i=1,\ldots,k$,

$$\mathrm{Tr}(D_{n\beta_i})=(p+o(1))n^2\beta_i^2 \ \ a.s.$$

Thus,

$$\begin{aligned}
&\mathrm{Tr}(D_{n,k})\\
=&2\sum_{i>j}(A_{n,k})_{ij}\\
=&2\sum_{i>j}(A_n-A'_{n,k})_{ij}\\
=&2\sum_{i>j}(A_n)_{ij}-2\sum_{i>j}(A'_{n,k})_{ij}\\
=&2\sum_{n\geq i>j\geq 1}(A_n)_{ij}-2\left[\sum_{n\beta_1\geq i>j\geq 1}(A_{n\beta_1})_{ij}+\cdots+\sum_{n\beta_k\geq i>j\geq 1}(A_{n\beta_k})_{ij}\right]\\
=&(p+o(1))n^2-\left[(p+o(1))(n\beta_1)^2+\cdots+(p+o(1))(n\beta_k)^2\right]\\
=&p\left(1-\sum_{i=1}^k\beta_i^2\right)n^2+o(1)n^2 \ \ a.s.
\end{aligned} \tag{2.19}$$

Note that $L_{n,k}-\frac{\mathrm{Tr}(D_{n,k})}{n}I_n=\widetilde{L_{n,k}}$. Then $\mu_i(L_{n,k})-\frac{\mathrm{Tr}(D_{n,k})}{n}=\xi_i(\widetilde{L_{n,k}})$, for $i=1,\ldots,n$, where $\mu_i(L_{n,k})$, $\xi_i(\widetilde{L_{n,k}})$ are eigenvalues of $L_{n,k}$ and $\widetilde{L_{n,k}}$, respectively. By (2.16), (2.17) and (2.19), we have for $1 \le i \le n-1$,

$$np\left(\sum_{i=1}^{k}\beta_i^2 - \max_{1\le i\le k}\{\beta_i\}\right) + o(1)n \le \xi_i(\widetilde{L_{n,k}}) \le np\sum_{i=1}^{k}\beta_i^2 + o(1)n \quad a.s., \tag{2.20}$$

and

$$np(\sum_{i=1}^{k}\beta_i^2 - \max_{1\le i\le k}\{\beta_i\} - 1) + o(1)n \le \xi_n(\widetilde{L_{n,k}}) \le np(\sum_{i=1}^{k}\beta_i^2 - 1) + o(1)n \quad a.s. \tag{2.21}$$

Hence, we have

$$(n-1)e^{np(\sum_{i=1}^{k}\beta_i^2 - \max\limits_{1\le i\le k}\{\beta_i\})+o(1)n} \le \sum_{i=1}^{n-1} e^{\xi_i(\widetilde{L_{n,k}})} \le (n-1)e^{np\sum_{i=1}^{k}\beta_i^2+o(1)n} \quad a.s., \tag{2.22}$$

and

$$e^{np(\sum_{i=1}^{k}\beta_i^2 - \max\limits_{1\le i\le k}\{\beta_i\}-1)+o(1)n} \le e^{\xi_n(\widetilde{L_{n,k}})} \le e^{np(\sum_{i=1}^{k}\beta_i^2-1)+o(1)n} \quad a.s. \tag{2.23}$$

Then (2.22) and (2.23) imply that

$$\begin{aligned}
&LEE(G_{n;\mu_1,\dots,\mu_k}(p))\\
=&\sum_{i=1}^{n} e^{\xi_i(\widetilde{L_{n,k}})}\\
\ge&(n-1)e^{np(\sum_{i=1}^{k}\beta_i^2 - \max\limits_{1\le i\le k}\{\beta_i\})+o(1)n} + e^{np(\sum_{i=1}^{k}\beta_i^2 - \max\limits_{1\le i\le k}\{\beta_i\}-1)+o(1)n}\\
=&(n-1+e^{-np})e^{np(\sum_{i=1}^{k}\beta_i^2 - \max\limits_{1\le i\le k}\{\beta_i\})+o(1)n} \quad a.s.,
\end{aligned} \tag{2.24}$$

and

$$\begin{aligned}
&LEE(G_{n;\beta_1,\dots,\beta_k}(p))\\
\le&(n-1)e^{np\sum_{i=1}^{k}\beta_i^2+o(1)n} + e^{np(\sum_{i=1}^{k}\beta_i^2-1)+o(1)n}
\end{aligned}$$

$$=(n-1+e^{-np})e^{np\sum_{i=1}^{k}\beta_i^2+o(1)n} \quad a.s. \tag{2.25}$$

This completes the proof. □

Corollary 2.9. *Let $G_{n;\beta_1,\dots,\beta_k}(p) \in \mathscr{G}_{n;\beta_1,\dots,\beta_k}(p)$. Then*

$$LEE(G_{n;\beta_1,\dots,\beta_k}(p)) = (n-1+e^{-np})e^{o(1)n} \quad a.s. \tag{2.26}$$

if and only if $\max\{\beta_1,\dots,\beta_k\} = o(1)$.

Proof. By (2.20), (2.21), (2.22) and (2.23), we have that (2.26) holds if and only if

$$\xi_i(\widetilde{L_{n,k}}) = o(1)n \quad a.s., \text{for } 1 \le i \le n-1, \tag{2.27}$$

and

$$\xi_n(\widetilde{L_{n,k}}) = -np + o(1)n \quad a.s. \tag{2.28}$$

By (2.16), (2.17) and (2.19), we have that (2.27) and (2.28) hold if and only if $\max\{\beta_1,\dots,\beta_k\} = o(1)$. □

Note that if $k = n$, then $G_{n;\beta_1,\dots,\beta_k}(p) = G_n(p)$, that is, $\beta_i = \frac{1}{n}$, $1 \le i \le n$. Using Corollary 2.9, the following result is immediate.

Corollary 2.10. *Let $G_n(p) \in \mathscr{G}_n(p)$ be a random graph. Then almost surely $LEE(G_n(p)) = (n-1+e^{-np})e^{o(1)n}$.*

Next, we consider two specific families of random k-partite graphs. Let $G_{n;\beta_1,\dots,\beta_k}(p) \in \mathscr{G}_{n;\beta_1,\dots,\beta_k}(p)$ satisfy $\lim_{n\to\infty} \max_{1\le i\le k}\{\beta_i\} > 0$ and $\lim_{n\to\infty} \frac{\beta_i}{\beta_j} = 1$. Then $G_{n;\beta_1,\dots,\beta_k}(p)$ is a balanced k-partite graph. By Theorem 2.8, we have the following result immediately.

Corollary 2.11. *Let $G_{n;\beta_1,\dots,\beta_k}(p) \in \mathscr{G}_{n;\beta_1,\dots,\beta_k}(p)$ and suppose that $G_{n;\beta_1,\dots,\beta_k}(p)$ satisfies $\lim_{n\to\infty} \max\{\beta_1,\beta_2,\dots,\beta_k\} > 0$ and $\lim_{n\to\infty} \frac{\beta_i}{\beta_j} = 1$. Then*

$$(n-1+e^{-np})e^{o(1)n} \le LEE(G_{n;\beta_1,\dots,\beta_k}(p)) \le (n-1+e^{-np})e^{(p/k+o(1))n} \quad a.s.$$

Let $G_{n;\beta_1,\ldots,\beta_k}(p) \in \mathscr{G}_{n;\beta_1,\ldots,\beta_k}(p)$ satisfy $\lim\limits_{n\to\infty} \max\limits_{1\le i\le k}\{\beta_i\} > 0$, and there exist β_i and β_j such that $\lim\limits_{n\to\infty} \frac{\beta_i}{\beta_j} < 1$. Then $G_{n;\beta_1,\ldots,\beta_k}(p)$ is an unbalanced k-partite graph. Since $\sum_{i=1}^k \beta_i = 1$ and $\lim\limits_{n\to\infty} \max\limits_{1\le i\le k}\{\beta_i\} > 0$, there exists at least one $\beta_i = O(1)$ (Otherwise, if $\beta_i = o(1)$ for all $1 \le i \le k$, then $\lim\limits_{n\to\infty} \max\limits_{1\le i\le k}\{\beta_i\} = 0$, a contradiction). Thus, $|V_i| = n\beta_i$ are of order $O(n)$. Without loss of generality, we can find an integer such that $1 \le r \le k$, $|V_1|, \ldots, |V_r|$ are of order $O(n)$ and $|V_{r+1}|, \ldots, |V_k|$ are of order $o(n)$. By Theorem 2.8, we have the following result readily.

Corollary 2.12. *Let $G_{n;\beta_1,\ldots,\beta_k}(p) \in \mathscr{G}_{n;\beta_1,\ldots,\beta_k}(p)$ and suppose that $G_{n;\beta_1,\ldots,\beta_k}(p)$ satisfies $\lim\limits_{n\to\infty} \max\{\beta_1, \beta_2, \ldots, \beta_k\} > 0$, and there exist β_i and β_j such that $\lim\limits_{n\to\infty} \frac{\beta_i}{\beta_j} < 1$, that is, there exists an integer $r \ge 1$ such that $|V_1|, \ldots, |V_r|$ are of order $O(n)$ and $|V_{r+1}|, \ldots, |V_k|$ are of order $o(n)$. Then*

$$\begin{aligned} &(n-1+e^{-np})e^{np(\sum_{i=1}^r \beta_i^2 - \max\limits_{1\le i\le r}\{\beta_i\}) + o(1)n} \\ \le & LEE(G_{n;\beta_1,\ldots,\beta_k}(p)) \\ \le &(n-1+e^{-np})e^{np\sum_{i=1}^r \beta_i^2 + o(1)n} \quad a.s. \end{aligned}$$

2.3 The von Neumann entropy

In this section, we establish a lower and upper bound for $\boldsymbol{S}(G_{n;\beta_1,\ldots,\beta_k})$ for almost all $G_{n;\beta_1,\ldots,\beta_k}(p) \in \mathscr{G}_{n;\beta_1,\ldots,\beta_k}(p)$, by analyzing the limiting behaviour of the spectra of random symmetric matrices. Our main result is stated as follows.

Theorem 2.13. *Let $G_{n;\beta_1,\ldots,\beta_k}(p) \in \mathscr{G}_{n;\beta_1,\ldots,\beta_k}(p)$. Then almost surely*

$$\begin{aligned} &\frac{1+o(1)}{1-\sum\limits_{i=1}^k \beta_i^2} \log_2\left(n\left(1-\sum_{i=1}^k \beta_i^2\right)\right) \\ \le &\boldsymbol{S}(G_{n;\beta_1,\ldots,\beta_k}(p)) \end{aligned}$$

$$\leq \frac{1-\max\limits_{1\leq i\leq k}\{\beta_i\}+o(1)}{1-\sum\limits_{i=1}^{k}\beta_i^2}\log_2\left(\frac{n\left(1-\sum\limits_{i=1}^{k}\beta_i^2\right)}{1-\max\limits_{1\leq i\leq k}\{\beta_i\}}\right),$$

independently of $0<p<1$.

Proof. By (2.16), (2.17) and (2.19), the eigenvalues of $P_{G_{n,k}}=\dfrac{L_{n,k}}{\mathrm{Tr}(D_{n,k})}$ satisfy that, for $1\leq i\leq n-1$,

$$\frac{p\left(1-\max\limits_{1\leq i\leq k}\{\beta_i\}\right)+o(1)}{p\left(1-\sum\limits_{i=1}^{k}\beta_i^2\right)n+o(1)n}\leq\lambda_i(P_{G_{n,k}})\leq\frac{p+o(1)}{p\left(1-\sum\limits_{i=1}^{k}\beta_i^2\right)n+o(1)n}\quad a.s., \tag{2.29}$$

and

$$\frac{-p\max\limits_{1\leq i\leq k}\{\beta_i\}+o(1)}{p\left(1-\sum\limits_{i=1}^{k}\beta_i^2\right)n+o(1)n}\leq\lambda_n(P_{G_{n,k}})\leq\frac{o(1)}{p\left(1-\sum\limits_{i=1}^{k}\beta_i^2\right)n+o(1)n}\quad a.s. \tag{2.30}$$

Then (2.29) and (2.30) imply that

$$\begin{aligned}
&\boldsymbol{S}(G_{n;\beta_1,\ldots,\beta_k}(p))\\
&\geq-\sum_{i=1}^{n-1}\left(\frac{p+o(1)}{p\left(1-\sum\limits_{i=1}^{k}\beta_i^2\right)n+o(1)n}\log_2\left(\frac{p+o(1)}{p\left(1-\sum\limits_{i=1}^{k}\beta_i^2\right)n+o(1)n}\right)\right)\\
&\quad-\frac{o(1)}{p\left(1-\sum\limits_{i=1}^{k}\beta_i^2\right)n+o(1)n}\log_2\left(\frac{o(1)}{p\left(1-\sum\limits_{i=1}^{k}\beta_i^2\right)n+o(1)n}\right)
\end{aligned}$$

$$=\frac{1+o(1)}{1-\sum\limits_{i=1}^{k}\beta_i^2}\log_2\left(n\left(1-\sum_{i=1}^{k}\beta_i^2\right)\right) \tag{2.31}$$

and

$$\begin{aligned}
&\boldsymbol{S}(G_{n;\beta_1,\dots,\beta_k}(p))\\
&\leq -\sum_{i=1}^{n-1}\left(\frac{p\left(1-\max\limits_{1\leq i\leq k}\{\beta_i\}\right)+o(1)}{p\left(1-\sum\limits_{i=1}^{k}\beta_i^2\right)n+o(1)n}\log_2\left(\frac{p\left(1-\max\limits_{1\leq i\leq k}\{\beta_i\}\right)+o(1)}{p\left(1-\sum\limits_{i=1}^{k}\beta_i^2\right)n+o(1)n}\right)\right)\\
&\quad-\frac{-p\max\limits_{1\leq i\leq k}\{\beta_i\}+o(1)}{p\left(1-\sum\limits_{i=1}^{k}\beta_i^2\right)n+o(1)n}\log_2\left(\frac{-p\max\limits_{1\leq i\leq k}\{\beta_i\}+o(1)}{p\left(1-\sum\limits_{i=1}^{k}\beta_i^2\right)n+o(1)n}\right)\\
&=-\frac{1-\max\limits_{1\leq i\leq k}\{\beta_i\}+o(1)}{1-\sum\limits_{i=1}^{k}\beta_i^2}\log_2\left(\frac{1-\max\limits_{1\leq i\leq k}\{\beta_i\}}{n\left(1-\sum\limits_{i=1}^{k}\beta_i^2\right)}\right)\\
&=\frac{1-\max\limits_{1\leq i\leq k}\{\beta_i\}+o(1)}{1-\sum\limits_{i=1}^{k}\beta_i^2}\log_2\left(\frac{n\left(1-\sum\limits_{i=1}^{k}\beta_i^2\right)}{1-\max\limits_{1\leq i\leq k}\{\beta_i\}}\right).
\end{aligned} \tag{2.32}$$

This completes the proof. □

Finally, we present some additional results implied by Theorem 2.13.

Corollary 2.14. *Let* $G_{n;\beta_1,\dots,\beta_k}(p)\in\mathcal{G}_{n;\beta_1,\dots,\beta_k}(p)$. *Then*

$$\boldsymbol{S}(G_{n;\beta_1,\dots,\beta_k}(p))=(1+o(1))\log_2 n\ \ a.s.$$

if and only if $\max\{\beta_1,\dots,\beta_k\}=o(1)$.

Note that if $k=n$, then $G_{n;\beta_1,\dots,\beta_k}(p)=G_n(p)$, that is, $\beta_i=\frac{1}{n}$, $1\leq i\leq n$. By Corollary 2.14, we have the following result immediately.

Corollary 2.15 (Du *et al.* [47])**.** *Let $G_n(p) \in \mathscr{G}_n(p)$ be a random graph. Then almost surely $\boldsymbol{S}(G_n(p)) = (1+o(1))\log_2 n$.*

The following corollaries are also easy to get.

Corollary 2.16. *Let $G_{n;\beta_1,\ldots,\beta_k}(p) \in \mathscr{G}_{n;\beta_1,\ldots,\beta_k}(p)$ satisfy $\lim\limits_{n\to\infty} \max\limits_{1\le i\le k}\{\beta_i\} > 0$ and $\lim\limits_{n\to\infty} \frac{\beta_i}{\beta_j} = 1$. Then*

$$\frac{1+o(1)}{1-\frac{1}{k}}\log_2\left(n\left(1-\frac{1}{k}\right)\right) \le \boldsymbol{S}(G_{n;\beta_1,\ldots,\beta_k}(p)) \le \left(1+\frac{k-1}{k}o(1)\right)\log_2 n.$$

Corollary 2.17. *Let $G_{n;\beta_1,\ldots,\beta_k}(p) \in \mathscr{G}_{n;\beta_1,\ldots,\beta_k}(p)$ satisfy $\lim\limits_{n\to\infty} \max\limits_{1\le i\le k}\{\beta_i\} > 0$, and there exist β_i and β_j such that $\lim\limits_{n\to\infty} \frac{\beta_i}{\beta_j} < 1$, that is, there exists an integer $r \ge 1$ such that $|V_1|, \ldots, |V_r|$ are of order $O(n)$ and $|V_{r+1}|, \ldots, |V_k|$ are of order $o(n)$. Then almost surely*

$$\begin{aligned}
&\frac{1+o(1)}{1-\sum\limits_{i=1}^{r}\beta_i^2}\log_2\left(n\left(1-\sum_{i=1}^{r}\beta_i^2\right)\right)\\
\le&\boldsymbol{S}(G_{n;\beta_1,\ldots,\beta_k}(p))\\
\le&\frac{1-\max\limits_{1\le i\le r}\{\beta_i\}+o(1)}{1-\sum\limits_{i=1}^{r}\beta_i^2}\log_2\left(\frac{n\left(1-\sum\limits_{i=1}^{r}\beta_i^2\right)}{1-\max\limits_{1\le i\le r}\{\beta_i\}}\right).
\end{aligned}$$

Chapter 3

The spectral distribution of random mixed graphs

In this chapter, we characterize the limiting spectral distribution of the Hermitian adjacency matrix of a random mixed graph $\widehat{G}_n(p_{ij})$, where $p_{ij} = p = p(n)$ for any $1 \leq i, j \leq n$ and 0 for $i = j$, for some $p \in (0, 1)$. We denote this graph by $\widehat{G}_n(p)$. We prove that the empirical distribution of the eigenvalues of the Hermitian adjacency matrix converges to Wigner's semicircle law. As an application, we estimate the Hermitian energy of a random mixed graph.

3.1 Preliminaries

Before proceeding, we collect some results that will be used in the sequel of the chapter.

Lemma 3.1 (See [12])**.** *The number of closed walks of length $2s$ which satisfy that each directed edge and its inverse directed edge in the closed walk both appear once and the underlying graph of the closed walk is a tree is $\frac{1}{s+1}\binom{2s}{s}$.*

Lemma 3.2 (See [12]). *Let $\phi(x)$ be as in Theorem 1.1. Then, for $s = 0, 1, 2, \ldots$, we have*

$$\int_{-2}^{2} x^k \phi(x) \mathrm{d}x = \begin{cases} 0, & \text{for } k = 2s+1, \\ \frac{1}{s+1}\binom{2s}{s}, & \text{for } k = 2s. \end{cases}$$

Lemma 3.3 (Cauchy-Schwarz's Inequality). *Let ξ and η be two complex random variables. Then*

$$|\mathbb{E}(\xi\overline{\eta})|^2 \leq \mathbb{E}(|\xi|^2) \cdot \mathbb{E}(|\eta|^2).$$

Proof. For any $t \in \mathbb{C}$, we have

$$\begin{aligned} 0 &\leq \mathbb{E}(t\xi - \eta)(\overline{t\xi - \eta}) \\ &= \mathbb{E}(t\xi - \eta)(\overline{t}\overline{\xi} - \overline{\eta}) \\ &= t\overline{t}\mathbb{E}(\xi\overline{\xi}) - t\mathbb{E}(\xi\overline{\eta}) - \overline{t}\mathbb{E}(\overline{\xi}\eta) + \mathbb{E}(\eta\overline{\eta}). \end{aligned}$$

Let

$$t = \frac{\mathbb{E}(\overline{\xi}\eta)}{\mathbb{E}(\xi\overline{\xi})}.$$

Then

$$\begin{aligned} 0 &\leq -\frac{\mathbb{E}(\xi\overline{\eta})\mathbb{E}(\overline{\xi}\eta)}{\mathbb{E}(\xi\overline{\xi})} + \mathbb{E}(\eta\overline{\eta}) \\ &= -\frac{\mathbb{E}(\xi\overline{\eta})\mathbb{E}(\overline{\xi\overline{\eta}})}{\mathbb{E}(|\xi|^2)} + \mathbb{E}(|\eta|^2) \\ &= -\frac{|\mathbb{E}(\xi\overline{\eta})|^2}{\mathbb{E}(|\xi|^2)} + \mathbb{E}(|\eta|^2). \end{aligned}$$

Hence

$$|\mathbb{E}(\xi\overline{\eta})|^2 \leq \mathbb{E}(|\xi|^2) \cdot \mathbb{E}(|\eta|^2).$$

This completes the proof. □

Lemma 3.4 (Chebyshev's Inequality). *Let X be a random variable. Then for any $\epsilon > 0$, we have*
$$\Pr(|X - \mathbb{E}(X)| \geq \epsilon) \leq \frac{\mathrm{Var}(X)}{\epsilon^2}.$$

Lemma 3.5 (Borel-Cantelli Lemma). *If $\sum_{n=1}^{\infty} \Pr(E_n) < \infty$ and the events $\{E_n\}_{n=1}^{\infty}$ are independent, then $\Pr(\limsup_{n\to\infty} E_n) = 0$.*

Lemma 3.6 (Rank Inequality (See [11])). *Let A and B be two $n \times n$ Hermitian matrices. Then*
$$\|F^A - F^B\| \leq \frac{1}{n}\mathrm{rank}(A - B),$$
where $\|f(x)\| := \sup_x |f(x)|$ for a function $f(x)$, and F^A means the ESD of A.

Lemma 3.7 (Chernoff Bounds (See [30])). *Let $X_1, \ldots, X_n$ be independent random variables with*
$$\Pr(X_i = 1) = p_i \ \ \textit{and} \ \ \Pr(X_i = 0) = 1 - p_i \ \ \textit{for all} \ \ i.$$
Consider the sum $X = \sum_{i=1}^{n} X_i$ with expectation $\mathbb{E}(X) = \sum_{i=1}^{n} p_i$. Then for any $b > 0$,

(i) **Lower tail:** $\Pr(X \leq \mathbb{E}(X) - b) \leq \exp\left(-\frac{b^2}{2\mathbb{E}(X)}\right)$;

(ii) **Upper tail:** $\Pr(X \geq \mathbb{E}(X) + b) \leq \exp\left(-\frac{b^2}{2(\mathbb{E}(X)+b/3)}\right)$.

Definition 3.1 (See [12]). Let A_n be an $n \times n$ Hermitian matrix, and $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of A_n. Then, for any real-valued function f,
$$\int f(x)\mathrm{d}F^{A_n}(x) = \frac{1}{n}\sum_{i=1}^{n} f(\lambda_i)$$
is called the *linear spectral statistics* (LSS) of A_n.

3.2 The LSD of Hermitian adjacency matrices of $\widehat{G}_n(p)$

In this section we characterize the LSD of the Hermitian adjacency matrices of random mixed graphs. We prove that the empirical distribution of the eigenvalues of the Hermitian adjacency matrices converges to Wigner's semicircle law. Our main result is stated as follows.

Theorem 3.8. *Let $\{H_n\}_{n=1}^{\infty}$ be a sequence of Hermitian adjacency matrices of random mixed graphs $\{\widehat{G}_n(p)\}_{n=1}^{\infty}$ with $p = p(n)$, $0 < p < 1$. Define $\sigma = \sqrt{2p - p^2 - p^4}$. Then the ESD of $\frac{1}{\sigma\sqrt{n}}H_n$ converges to the standard semicircle distribution whose density is given by*

$$\phi(x) := \begin{cases} \frac{1}{2\pi}\sqrt{4 - x^2}, & \text{for } |x| \leq 2, \\ 0, & \text{for } |x| > 2. \end{cases}$$

The proof of Theorem 3.8 is postponed until the end of this section. Recall that H_n is a random Hermitian matrix whose upper-triangular entries are i.i.d. copies of a random variable ξ and diagonal entries are 0. Recall also that ξ takes value 1 with probability p^2, $\mathbf{i}$ with probability $p(1-p)$, $-\mathbf{i}$ with probability $p(1-p)$, and 0 with probability $(1-p)^2$. Then

$$\mathbb{E}(\xi) = p^2, \ \ \mathrm{Var}(\xi) = \mathbb{E}[(\xi - \mathbb{E}(\xi))(\overline{\xi - \mathbb{E}(\xi)})] = 2p - p^2 - p^4.$$

Let $f(x) = x^3 + x - 2$. Then $f'(x) = 3x^2 + 1 > 0$. So, $-2 = f(0) < f(p) < f(1) = 0$. Thus $\mathrm{Var}(\xi) = 2p - p^2 - p^4 = p(2 - p - p^3) > 0$.

Let $\sigma = \sqrt{\mathrm{Var}(\xi)} = \sqrt{2p - p^2 - p^4}$, and define

$$M_n = \frac{1}{\sigma}[H_n - p^2(J_n - I_n)] = (\eta_{ij}),$$

where J_n is the all-ones matrix of order n and I_n is the identity matrix of order n. It can be easily verified that

- M_n is a Hermitian matrix;
- the diagonal entries $\eta_{ii} = 0$ and the upper-triangular entries η_{ij}, $1 \leq i < j \leq n$ are i.i.d. copies of random variable η which takes value $\frac{1-p^2}{\sigma}$ with probability p^2, $\frac{\mathbf{i}-p^2}{\sigma}$ with probability $p(1-p)$, $\frac{-\mathbf{i}-p^2}{\sigma}$ with probability $p(1-p)$, and $\frac{-p^2}{\sigma}$ with probability $(1-p)^2$.

We denote the distribution function of η by Φ. Note that the random variable η of M_n has mean 0 and variance 1, that is,

$$\mathbb{E}(\eta) = 0 \quad \text{and} \quad \mathrm{Var}(\eta) = 1.$$

Note also that the expectation

$$\mathbb{E}(|\eta|^s) = \frac{(1-p^2)^s \cdot p^2 + 2(1+p^4)^{s/2} \cdot p(1-p) + p^{2s} \cdot (1-p)^2}{(2p-p^2-p^4)^{s/2}}.$$

It is easy to check that $2p - p^2 - p^4 \to 0$ as $p(n) \to 0$ or $p(n) \to 1$. So, if $\lim_{n\to\infty} p(n) = 0$, then

$$\begin{aligned}\mathbb{E}(|\eta|^s) &\to \frac{2p}{(2p)^{s/2}} \\ &= \frac{1}{(2p)^{s/2-1}}.\end{aligned}$$

This implies that if $p = o(1)$, then M_n is not a Wigner matrix. Thus the LSD of M_n cannot be directly derived by Wigner's semicircle law. In the following, we will use the moment method to prove that the ESD of $\frac{1}{\sqrt{n}}M_n$ converges to the standard semicircle distribution.

Theorem 3.9. *Let $\sigma = \sqrt{2p-p^2-p^4}$, and $M_n = \frac{1}{\sigma}[H_n - p^2(J_n - I_n)]$. Then the ESD of $n^{-1/2}M_n$ converges to the standard semicircle distribution whose density is given by*

$$\phi(x) := \begin{cases} \frac{1}{2\pi}\sqrt{4-x^2}, & \text{for } |x| \le 2, \\ 0, & \text{for } |x| > 2. \end{cases}$$

Proof of Theorem 3.9. Let

$$W_n := \frac{1}{\sqrt{n}} M_n = \left(\frac{\eta_{ij}}{\sqrt{n}}\right).$$

To prove that the ESD of W_n converges to the standard semicircle distribution, it suffices to show that the moments of the ESD converge almost surely to the corresponding moments of the semicircle distribution.

For a positive integer k, by Definition 3.1, the kth moment of the ESD of the matrix W_n is

$$\begin{aligned}
M_{k,n} &= \int x^k \mathrm{d}F^{W_n}(x) \\
&= \frac{1}{n}\sum_{i=1}^{n}(\lambda_i(W_n))^k \\
&= \frac{1}{n}\mathrm{Tr}(W_n^k) \\
&= \frac{1}{n}\mathrm{Tr}\left(\left(\frac{1}{\sqrt{n}}M_n\right)^k\right) \\
&= \frac{1}{n^{1+k/2}}\mathrm{Tr}(M_n^k) \\
&= \frac{1}{n^{1+k/2}}\sum_{1\le i_1,\ldots,i_k\le n}\eta_{i_1i_2}\eta_{i_2i_3}\cdots\eta_{i_ki_1},
\end{aligned} \tag{3.1}$$

where $W := i_1i_2\ldots i_{k-1}i_ki_1$ corresponds to a closed directed walk of length k in the complete directed graph of order n. For each directed edge $(i,j)\in W$, let q_{ij} be the number of occurrences of the directed edge (i,j) in the walk W. Note that all directed edges of a mixed graph are mutually independent. Then we rewrite (3.1) as

$$M_{k,n} = \frac{1}{n^{1+k/2}}\sum_{W}\prod_{i<j}\eta_{ij}^{q_{ij}}\eta_{ji}^{q_{ji}}. \tag{3.2}$$

Then

$$\mathbb{E}(M_{k,n}) = \frac{1}{n^{1+k/2}}\sum_{W}\prod_{i<j}\mathbb{E}\left(\eta_{ij}^{q_{ij}}\eta_{ji}^{q_{ji}}\right). \tag{3.3}$$

Here the summation is taken over all directed closed walks of length k.

To show that $F^{W_n}(x)$ converges to the standard semicircle distribution whose density is $\phi(x)$, by the Moment Convergence Theorem (MCT), it suffices to prove

$$\lim_{n\to\infty} M_{k,n} = \int_{-2}^{2} x^k \phi(x)\mathrm{d}x, \ k = 1, 2, \ldots. \tag{3.4}$$

Define $\widetilde{M_n} = (\eta'_{ij})$, where

$$\eta'_{ij} = \begin{cases} \eta_{ij}, & \text{if } |\eta_{ij}| < \sqrt{n}, \\ 0, & \text{if } |\eta_{ij}| \ge \sqrt{n}. \end{cases}$$

Let

$$\widetilde{W_n} = \frac{1}{\sqrt{n}}\widetilde{M_n} = \left(\frac{\eta'_{ij}}{\sqrt{n}}\right),$$

and let $M'_{k,n}$ be the kth moment of the ESD of the matrix $\widetilde{W_n}$. Similar to (3.1), (3.2) and (3.3), we have

$$M'_{k,n} = \frac{1}{n^{1+k/2}} \sum_{1\le i_1,\ldots,i_k\le n} \eta'_{i_1i_2}\eta'_{i_2i_3}\cdots\eta'_{i_ki_1} = \frac{1}{n^{1+k/2}} \sum_W \prod_{i<j} \eta_{ij}^{\prime q_{ij}} \eta_{ji}^{\prime q_{ji}}, \tag{3.5}$$

and

$$\mathbb{E}(M'_{k,n}) = \frac{1}{n^{1+k/2}} \sum_{1\le i_1,\ldots,i_k\le n} \mathbb{E}(\eta'_{i_1i_2}\eta'_{i_2i_3}\cdot\cdot\eta'_{i_ki_1}) = \frac{1}{n^{1+k/2}} \sum_W \prod_{i<j} \mathbb{E}\left(\eta_{ij}^{\prime q_{ij}} \eta_{ji}^{\prime q_{ji}}\right). \tag{3.6}$$

Now (3.4) can be easily verified by combining Facts 3.1–3.3 below that we are going to prove separately. This completes the proof of Theorem 3.9. □

Fact 3.1. *Let $\phi(x)$ be as in Theorem 3.8, and let $M'_{k,n}$ be as in Eq. (3.5). Then*

$$\lim_{n\to\infty} \mathbb{E}(M'_{k,n}) = \int_{-2}^{2} x^k \phi(x)\mathrm{d}x = \begin{cases} 0, & \textit{for } k = 2s+1, \\ \frac{1}{s+1}\binom{2s}{s}, & \textit{for } k = 2s. \end{cases} \tag{3.7}$$

Fact 3.2. *Let $M'_{k,n}$ be as in Eq. (3.5). Then*

$$\lim_{n\to\infty} M'_{k,n} = \lim_{n\to\infty} \mathbb{E}(M'_{k,n}) \ \ a.s. \tag{3.8}$$

Fact 3.3. *Let $M_{k,n}$ and $M'_{k,n}$ be as in Eqs. (3.2) and (3.5), respectively. Then*

$$\lim_{n\to\infty} M_{k,n} = \lim_{n\to\infty} M'_{k,n} \quad a.s. \tag{3.9}$$

It remains to prove Facts 3.1–3.3.

Proof of Fact 3.1. The second equality of (3.7) follows from Lemma 3.2 straightforwardly. Next, we prove the first equality of (3.7).

Consider the underlying undirected graph $\Gamma(G)$ of the directed graph G. We decompose $\mathbb{E}(M'_{k,n})$ into parts $\mathbb{E}_{m,k,n}, m = 1, 2, \ldots, k$, containing the m-fold sums,

$$\mathbb{E}(M'_{k,n}) = \sum_{m=1}^{k} \mathbb{E}_{m,k,n}, \tag{3.10}$$

where

$$\mathbb{E}_{m,k,n} = \frac{1}{n^{1+k/2}} \sum_{\{W:|E(\Gamma(W))|=m\}} \prod_{i<j} \mathbb{E}\left(\eta'^{q_{ij}}_{ij} \eta'^{q_{ji}}_{ji}\right), \tag{3.11}$$

and $|E(\Gamma(W))| = m$ means the cardinality of the edge set of $\Gamma(W)$ is m. Here the summation in (3.11) is taken over all closed directed walks W of length k.

Recall that $\mathbb{E}(\eta) = 0$, and recall also that q_{ij} denotes the number of occurrences of the directed edge (i, j) in the closed walk W. So, if $q_{ij} + q_{ji} = 1$, that is, $q_{ij} = 1, q_{ji} = 0$ or $q_{ij} = 0, q_{ji} = 1$, then $\prod_{i<j} \mathbb{E}\left(\eta^{q_{ij}}_{ij} \eta^{q_{ji}}_{ji}\right) = 0$ and $\prod_{i<j} \mathbb{E}\left(\eta'^{q_{ij}}_{ij} \eta'^{q_{ji}}_{ji}\right) = 0$. On the other hand, if $m > \frac{k}{2}$ and $q_{ij} + q_{ji} \geq 2$, then $\mathbb{E}_{m,k,n} = 0$. So, in the following, we only consider the case that $m \leq \frac{k}{2}$ and $q_{ij} + q_{ji} \geq 2$.

Case 1. k is odd. Then $m \leq \left\lfloor \frac{k}{2} \right\rfloor$. Note that $|E(\Gamma(W))| = m$, i.e., there are m edges in $\Gamma(W)$. Then there are at most $m + 1$ vertices in $\Gamma(W)$. This shows that the number of such closed walks of length k is at most $n^{m+1} \cdot (m + 1)^k$.

Then

$$\mathbb{E}_{m,k,n} \leq \frac{n^{m+1} \cdot (m+1)^k}{n^{1+k/2}} \prod_{i<j} \mathbb{E}\left(\eta_{ij}'^{q_{ij}} \eta_{ji}'^{q_{ji}}\right) = \frac{(m+1)^k}{n^{k/2-m}} \prod_{i<j} \mathbb{E}\left(\eta_{ij}'^{q_{ij}} \eta_{ji}'^{q_{ji}}\right).$$

Note that $\mathbb{E}\eta = 0$. Then

$$\mathbb{E}(\eta\overline{\eta}) = \mathbb{E}|\eta|^2 = \mathbb{E}[(\eta - \mathbb{E}(\eta))(\overline{\eta - \mathbb{E}(\eta)}] = \mathrm{Var}(\eta) = 1.$$

Recall that the distribution function of η is denoted by Φ. Then

$$\mathbb{E}|\eta|^2 = \int |x|^2 \mathrm{d}\Phi = 1 < \infty.$$

Thus, for any $r \geq 3$,

$$n^{(2-r)/2} \int_{|x|<\sqrt{n}} |x|^r \mathrm{d}\Phi = o(1), \tag{3.12}$$

which follows from the fact (See [7, 8]) that for any distribution function Ψ,

$$\int |x|^t \mathrm{d}\Psi < \infty \Longrightarrow n^{(t-r)/2} \int_{|x|<\sqrt{n}} |x|^r \mathrm{d}\Psi = o(1) \text{ (for any } r \geq t+1).$$

Note that $q_{ij} + q_{ji} \geq 2$ implies that $q_{ij} \geq 1, q_{ji} \geq 1$ or $q_{ij} \geq 2, q_{ji} = 0$ or $q_{ij} = 0, q_{ji} \geq 2$. We consider these three subcases separately.

First assume $q_{ij} \geq 1, q_{ji} \geq 1$. Then we set

$$\begin{aligned}
E_1 &= \{ij \in \Gamma(W) | q_{ij} > 1, q_{ji} > 1\}, \\
E_2 &= \{ij \in \Gamma(W) | q_{ij} > 1, q_{ji} = 1 \text{ or } q_{ij} = 1, q_{ji} > 1\}, \\
E_3 &= \{ij \in \Gamma(W) | q_{ij} = 1, q_{ji} = 1\}.
\end{aligned}$$

Let $m_i = |E_i|$, for $i = 1, 2, 3$. Clearly, $E(\Gamma(W)) = E_1 \cup E_2 \cup E_3$ and $m_1 + m_2 + m_3 = m$. Then, by (3.12) and Lemma 3.3, we have

$$\frac{(m+1)^k}{n^{k/2-m}} \prod_{i<j} \left| \mathbb{E}\left(\eta_{ij}'^{q_{ij}} \eta_{ji}'^{q_{ji}}\right) \right|$$

$$
\begin{aligned}
&\leq \frac{(m+1)^k}{n^{k/2-m}} \prod_{i<j} \sqrt{\mathbb{E}\left|\eta_{ij}^{\prime q_{ij}}\right|^2 \cdot \mathbb{E}\left|\eta_{ji}^{\prime q_{ji}}\right|^2} \\
&= \frac{(m+1)^k}{n^{k/2-m}} \prod_{i<j} \sqrt{\mathbb{E}|\eta'_{ij}|^{2q_{ij}} \cdot \mathbb{E}|\eta'_{ji}|^{2q_{ji}}} \\
&= \frac{(m+1)^k}{n^{k/2-m}} \left(\prod_{E_1} \sqrt{\mathbb{E}|\eta'_{ij}|^{2q_{ij}} \cdot \mathbb{E}|\eta'_{ji}|^{2q_{ji}}}\right) \\
&\quad \cdot \left(\prod_{E_2} \sqrt{\mathbb{E}|\eta'_{ij}|^{2q_{ij}} \cdot \mathbb{E}|\eta'_{ji}|^{2q_{ji}}}\right) \left(\prod_{E_3} \sqrt{\mathbb{E}|\eta'_{ij}|^{2q_{ij}} \cdot \mathbb{E}|\eta'_{ji}|^{2q_{ji}}}\right) \\
&= \frac{(m+1)^k}{n^{k/2-m}} \left(\prod_{E_1} \sqrt{\frac{o(1)}{n^{(2-2q_{ij})/2}} \cdot \frac{o(1)}{n^{(2-2q_{ij})/2}}}\right) \left(\prod_{E_2} \sqrt{\frac{o(1)}{n^{(2-2q_{ij})/2}}}\right) \cdot 1 \\
&= \frac{(m+1)^k}{n^{k/2-m}} \left(\prod_{E_1} \sqrt{\frac{o(1)}{n^{2-q_{ij}-q_{ji}}}}\right) \left(\prod_{E_2} \sqrt{\frac{o(1)}{n^{1-q_{ij}}}}\right) \\
&= \frac{(m+1)^k}{n^{k/2-m}} \sqrt{\frac{o(1)}{n^{2m-k}}} \\
&= (m+1)^k \cdot o(1) \\
&\to 0, \text{ as } n \to \infty.
\end{aligned}
$$

Next assume $q_{ij} \geq 2, q_{ji} = 0$. Then we set

$$
\begin{aligned}
E_4 &= \{ij \in \Gamma(W) | q_{ij} > 2, q_{ji} = 0\}, \\
E_5 &= \{ij \in \Gamma(W) | q_{ij} = 2, q_{ji} = 0\}.
\end{aligned}
$$

Let $m_i = |E_i|$, for $i = 4, 5$. Then $E(\Gamma(W)) = E_4 \cup E_5$ and $m_4 + m_5 = m$. So, we have

$$
\begin{aligned}
&\frac{(m+1)^k}{n^{k/2-m}} \prod_{i<j} \left|\mathbb{E}\left(\eta_{ij}^{\prime q_{ij}} \eta_{ji}^{\prime q_{ji}}\right)\right| \\
\leq &\frac{(m+1)^k}{n^{k/2-m}} \prod_{i<j} \mathbb{E}\left|\eta_{ij}^{\prime q_{ij}}\right|
\end{aligned}
$$

$$
\begin{aligned}
&=\frac{(m+1)^k}{n^{k/2-m}}\prod_{i<j}\mathbb{E}|\eta'_{ij}|^{q_{ij}}\\
&=\frac{(m+1)^k}{n^{k/2-m}}\left(\prod_{E_4}\mathbb{E}|\eta'_{ij}|^{q_{ij}}\right)\left(\prod_{E_5}\mathbb{E}|\eta'_{ij}|^{q_{ij}}\right)\\
&=\frac{(m+1)^k}{n^{k/2-m}}\prod_{E_4}\frac{o(1)}{n^{(2-q_{ij})/2}}\cdot 1\\
&=\frac{(m+1)^k}{n^{k/2-m}}\cdot\frac{o(1)}{n^{(2m-k)/2}}\\
&=(m+1)^k\cdot o(1)\\
&\to 0,\ \text{as } n\to\infty.
\end{aligned}
$$

Finally assume $q_{ij}=0, q_{ji}\geq 2$. Then, by a similar discussion as above, we have

$$\frac{(m+1)^k}{n^{k/2-m}}\prod_{i<j}\left|\mathbb{E}\left(\eta_{ij}^{\prime q_{ij}}\eta_{ji}^{\prime q_{ji}}\right)\right|\to 0,\ \text{as } n\to\infty.$$

Thus, by (3.10), we have

$$\lim_{n\to\infty}\mathbb{E}(M'_{k,n})=0 \text{ for } k \text{ is odd}.$$

Case 2. $k=2s$ $(s=1,2,\ldots)$ is even. Recall that $m\leq\frac{k}{2}=s$ and $q_{ij}+q_{ji}\geq 2$. We distinguish two subcases.

Case 2.1. $m<s=\frac{k}{2}$. Note that $|E(\Gamma(W))|=m$. Then there are at most $m+1$ vertices in $\Gamma(W)$. This shows that the number of such closed walks of length k is at most $n^{m+1}\cdot(m+1)^k$. Then

$$\mathbb{E}_{m,k,n}\leq\frac{n^{m+1}\cdot(m+1)^k}{n^{1+k/2}}\prod_{i<j}\mathbb{E}\left(\eta_{ij}^{\prime q_{ij}}\eta_{ji}^{\prime q_{ji}}\right)=\frac{(m+1)^k}{n^{k/2-m}}\prod_{i<j}\mathbb{E}\left(\eta_{ij}^{\prime q_{ij}}\eta_{ji}^{\prime q_{ji}}\right).$$

Notice that $q_{ij}+q_{ji}\geq 2$. Then $q_{ij}\geq 1, q_{ji}\geq 1$ or $q_{ij}\geq 2, q_{ji}=0$ or

$q_{ij} = 0, q_{ji} \geq 2$. By similar discussions as in Case 1, it can be verified that

$$\frac{(m+1)^k}{n^{k/2-m}} \prod_{i<j} \left| \mathbb{E}\left(\eta_{ij}'^{q_{ij}} \eta_{ji}'^{q_{ji}} \right) \right| \to 0, \text{ as } n \to \infty.$$

Thus, for $m < s$, we have

$$\lim_{n\to\infty} \mathbb{E}_{m,k,n} = 0, \text{ for } k = 2s.$$

Case 2.2. $m = s$. In this case, $q_{ij} + q_{ji} \geq 2$ implies that $q_{ij} = 1, q_{ji} = 1$ (each edge in the closed walk appears only once, and so does its inverse edge) or $q_{ij} = 2, q_{ji} = 0$ or $q_{ij} = 0, q_{ji} = 2$. Consider the following cases.

If $q_{ij} = 1, q_{ji} = 1$, and the underlying graph of the closed walk is a tree (i.e., there are $s+1$ vertices in $\Gamma(W)$), then by Lemma 3.1, the number of closed walks of length $k = 2s$ satisfying $q_{ij} = 1, q_{ji} = 1$ and the underlying graph of the closed walk is a tree is $\frac{1}{s+1}\binom{2s}{s}$. Recall that $\mathbb{E}(\eta\overline{\eta}) = \mathrm{Var}(\eta) = 1$. Then these terms will contribute

$$\begin{aligned}
&\frac{n(n-1)\cdots(n-s)\cdot\frac{1}{s+1}\binom{2s}{s}}{n^{1+k/2}} \prod_{i<j} \mathbb{E}(\eta'_{ij}\eta'_{ji}) \\
=&\frac{n^{1+s}(1+O(n^{-1}))\cdot\frac{1}{s+1}\binom{2s}{s}}{n^{1+s}} \prod_{i<j} \mathbb{E}(\eta'_{ij}\eta'_{ji}) \\
=&(1+O(n^{-1}))\cdot\frac{1}{s+1}\binom{2s}{s}\cdot 1 \\
\to&\frac{1}{s+1}\binom{2s}{s}, \text{ as } n \to \infty.
\end{aligned}$$

If $q_{ij} = 1, q_{ji} = 1$, and the underlying graph of the closed walk is not a tree (i.e., there are at most s vertices in $\Gamma(W)$). It is clear that the number of such closed walks of length k is at most $n^s \cdot s^k$. Recall that $\mathbb{E}(\eta\overline{\eta}) = \mathrm{Var}(\eta) = 1$. Then these terms will contribute at most

$$\frac{n^s \cdot s^k}{n^{1+k/2}} \prod_{i<j} \mathbb{E}(\eta'_{ij}\eta'_{ji}) = \frac{s^k}{n} \to 0, \text{ as } n \to \infty.$$

If $q_{ij} = 2, q_{ji} = 0$, then there are at most s vertices in $\Gamma(W)$. It is clear that the number of such closed walks of length k is at most $n^s \cdot s^k$. Then these terms will contribute at most

$$\frac{n^s \cdot s^k}{n^{1+k/2}} \prod_{i<j} \mathbb{E}\left(\eta_{ij}'^{q_{ij}} \eta_{ji}'^{q_{ji}}\right) = \frac{s^k}{n} \prod_{i<j} \mathbb{E}\left(\eta_{ij}'^{q_{ij}} \eta_{ji}'^{q_{ji}}\right).$$

In addition,

$$\begin{aligned}\frac{s^k}{n} \prod_{i<j} \left|\mathbb{E}\left(\eta_{ij}'^{q_{ij}} \eta_{ji}'^{q_{ji}}\right)\right| &\leq \frac{s^k}{n} \prod_{i<j} \mathbb{E}|\eta_{ij}'^{q_{ij}}| \\ &= \frac{s^k}{n} \prod_{i<j} \mathbb{E}(|\eta_{ij}'|^{q_{ij}}) \\ &= \frac{s^k}{n} \\ &\to 0, \text{ as } n \to \infty.\end{aligned}$$

Hence,

$$\frac{n^s \cdot s^k}{n^{1+k/2}} \prod_{i<j} \mathbb{E}\left(\eta_{ij}'^{q_{ij}} \eta_{ji}'^{q_{ji}}\right) \to 0, \text{ as } n \to \infty.$$

If $q_{ij} = 0, q_{ji} = 2$, by a similar discussion as above, it can be verified that

$$\frac{s^k}{n} \prod_{i<j} \left|\mathbb{E}\left(\eta_{ij}'^{q_{ij}} \eta_{ji}'^{q_{ji}}\right)\right| \to 0, \text{ as } n \to \infty.$$

Thus, for $m = s$, we have

$$\lim_{n\to\infty} \mathbb{E}_{m,k,n} = \frac{1}{s+1}\binom{2s}{s}, \text{ for } k = 2s.$$

Hence, by (3.10), we have

$$\lim_{n\to\infty} \mathbb{E}(M'_{k,n}) = \frac{1}{s+1}\binom{2s}{s}, \text{ for } k = 2s.$$

Therefore, the first equality of (3.7) is proved. This completes the proof of Fact 3.1. □

Proof of Fact 3.2. Note that $|M'_{k,n} - \mathbb{E}(M'_{k,n})|^4$ is a random variable. Suppose that $\{a_i^4\}$ is the set of all values that $|M'_{k,n} - \mathbb{E}(M'_{k,n})|^4$ takes. Then, for any k, n, we have

$$\begin{aligned}
&\mathbb{E}[|M'_{k,n} - \mathbb{E}(M'_{k,n})|^4] \\
=&\sum_i a_i^4 \Pr(|M'_{k,n} - \mathbb{E}(M'_{k,n})|^4 = a_i^4) \\
\geq&\sum_{a_i \geq \epsilon} a_i^4 \Pr(|M'_{k,n} - \mathbb{E}(M'_{k,n})|^4 = a_i^4) \\
\geq&\epsilon^4 \sum_{a_i \geq \epsilon} \Pr(|M'_{k,n} - \mathbb{E}(M'_{k,n})|^4 = a_i^4) \\
=&\epsilon^4 \Pr(|M'_{k,n} - \mathbb{E}(M'_{k,n})|^4 \geq \epsilon^4) \\
=&\epsilon^4 \Pr(|M'_{k,n} - \mathbb{E}(M'_{k,n})| \geq \epsilon).
\end{aligned}$$

Hence,

$$\Pr(|M'_{k,n} - \mathbb{E}(M'_{k,n})| \geq \epsilon) \leq \epsilon^{-4}\mathbb{E}[|M'_{k,n} - \mathbb{E}(M'_{k,n})|^4]. \tag{3.13}$$

Recall that

$$M'_{k,n} = \frac{1}{n^{1+k/2}} \sum_{1 \leq i_1, \ldots, i_k \leq n} \eta'_{i_1 i_2} \eta'_{i_2 i_3} \cdots \eta'_{i_k i_1} := \frac{1}{n^{1+k/2}} \sum_W \eta'(W),$$

where $W := i_1 i_2 \ldots i_{k-1} i_k i_1$ corresponds to a closed directed walk of length k in the complete directed graph of order n. Note (See Bai [11, p.620]) that

$$\mathbb{E}[|M'_{k,n} - \mathbb{E}(M'_{k,n})|^4] = \frac{1}{n^{4+2k}} \sum_{W^1, \ldots, W^4} \mathbb{E}\left\{\prod_{i=1}^{4}[\eta'(W^i) - \mathbb{E}(\eta'(W^i))]\right\}, \tag{3.14}$$

where W^i $(i = 1, \ldots, 4)$ corresponds to a closed directed walk of length k in the complete directed graph of order n.

Set $i_0 \in \{1, 2, 3, 4\}$. If $\Gamma(W^{i_0})$ has no common edge with $\Gamma(\widehat{W} \backslash W^{i_0})$, where $\widehat{W} = W^1 \cup W^2 \cup W^3 \cup W^4$, that is, W^{i_0} is independent of $\widehat{W}\backslash W^{i_0}$, then (3.14) is equal to zero since

$$\begin{aligned}
&\mathbb{E}\left\{\prod_{i=1}^{4}[\eta'(W^i) - \mathbb{E}(\eta'(W^i))]\right\} \\
=&\mathbb{E}\left\{\prod_{\substack{i=1\\ i\neq i_0}}^{4}[\eta'(W^i) - \mathbb{E}(\eta'(W^i))]\right\}\mathbb{E}[\eta'(W^{i_0}) - \mathbb{E}(\eta'(W^{i_0}))] \\
=&0,
\end{aligned}$$

due to the independence.

If there is a directed edge (i_0, j_0) whose number of occurrences in $\widehat{W} = W^1 \cup W^2 \cup W^3 \cup W^4$ is 1 and $(j_0, i_0) \notin \widehat{W}$, without loss of generality, we assume that $(i_0, j_0) \in W^1$, and $(i_0, j_0) \notin W^i$ for $i \in \{2,3,4\}$. Since $\mathbb{E}(\eta') = \mathbb{E}(\eta) = 0$, we have $\mathbb{E}(\eta'(W^1)) = \mathbb{E}(\eta'_{i_0 j_0})\mathbb{E}[\eta'(W^1\backslash\{(i_0, j_0)\})] = 0$. Then

$$\begin{aligned}
&\mathbb{E}\left\{\prod_{i=1}^{4}[\eta'(W^i) - \mathbb{E}(\eta'(W^i))]\right\} \\
=&\mathbb{E}\left\{\eta'(W^1)\prod_{i=2}^{4}[\eta'(W^i) - \mathbb{E}(\eta'(W^i))]\right\} \\
=&\mathbb{E}(\eta'_{i_0 j_0})\mathbb{E}\left\{\eta'(W^1\backslash\{(i_0, j_0)\})\prod_{i=2}^{4}[\eta'(W^i) - \mathbb{E}(\eta'(W^i))]\right\} \\
=&0,
\end{aligned}$$

which implies that (3.14) is also equal to zero.

Next, we consider the case that (3.14) may be nonzero. So, from the cases we already discussed above, we know that, in such a case, there exists no directed edge such that the total number of occurrences of this directed edge and its inverse edge in $\widehat{W}$ is just 1. For $e_i \in E(\Gamma(G))$, define $v_i^{\#}$ to be number of occurrences of the directed edges (x, y) and (y, x) in G such that (x, y) and (y, x) correspond to the edge e_i in $\Gamma(G)$, called the multiplicity of e_i. Assume that $\Gamma(\widehat{W})$ has edges $e_1, e_2, \ldots, e_l$ with multiplicities

$v_1^{\#}, v_2^{\#}, \ldots, v_l^{\#}$. Clearly, $v_i^{\#} \geq 2$ for $i = 1, \ldots, l$, and $v_1^{\#} + v_2^{\#} + \cdots + v_l^{\#} = 4k$. So $l \leq 2k$.

Note that

$$\eta'_{ij} = \begin{cases} \eta_{ij}, & \text{if } |\eta_{ij}| < \sqrt{n}, \\ 0, & \text{if } |\eta_{ij}| \geq \sqrt{n}. \end{cases}$$

Let

$$\tau_{ij} = \frac{\eta'_{ij}}{\sqrt{n}},$$

Then

$$|\tau_{ij}| = \frac{|\eta'_{ij}|}{\sqrt{n}} < 1,$$

and

$$\begin{aligned} M'_{k,n} &= \frac{1}{n^{1+k/2}} \sum_{1 \leq i_1, \ldots, i_k \leq n} \eta'_{i_1 i_2} \eta'_{i_2 i_3} \cdots \eta'_{i_k i_1} \\ &= \frac{1}{n} \sum_{1 \leq i_1, \ldots, i_k \leq n} \tau_{i_1 i_2} \tau_{i_2 i_3} \cdots \tau_{i_k i_1} \\ &:= \frac{1}{n} \sum_{W} \tau(W), \end{aligned}$$

where $W := i_1 i_2 \ldots i_{k-1} i_k i_1$ corresponds to a closed directed walk of length k in the complete directed graph of order n. Then

$$\begin{aligned} &\frac{1}{n^{4+2k}} \sum_{W^1, \ldots, W^4} \left| \mathbb{E} \left\{ \prod_{i=1}^{4} [\eta'(W^i) - \mathbb{E}(\eta'(W^i))] \right\} \right| \\ =& \frac{1}{n^4} \sum_{W^1, \ldots, W^4} \left| \mathbb{E} \left\{ \prod_{i=1}^{4} [\tau(W^i) - \mathbb{E}(\tau(W^i))] \right\} \right| \\ =& \frac{1}{n^4} \sum_{W^1, \ldots, W^4} \Big| \mathbb{E}[\tau(W^1)\tau(W^2)\tau(W^3)\tau(W^4)] \\ & - 4\mathbb{E}(\tau(W^1))\mathbb{E}[\tau(W^2)\tau(W^3)\tau(W^4)] \\ & + 6\mathbb{E}[\tau(W^1)\tau(W^2)]\mathbb{E}(\tau(W^3))\mathbb{E}(\tau(W^4)) \end{aligned}$$

$$
\begin{aligned}
&- 3\mathbb{E}(\tau(W^1))\mathbb{E}(\tau(W^2))\mathbb{E}(\tau(W^3))\mathbb{E}(\tau(W^4))\Big| \\
\leq &\frac{1}{n^4}\sum_{W^1,\ldots,W^4}\Bigg(\Big|\mathbb{E}[\tau(W^1)\tau(W^2)\tau(W^3)\tau(W^4)]\Big| \\
&+4\Big|\mathbb{E}(\tau(W^1))\mathbb{E}[\tau(W^2)\tau(W^3)\tau(W^4)]\Big| \\
&+6\Big|\mathbb{E}[\tau(W^1)\tau(W^2)]\mathbb{E}(\tau(W^3))\mathbb{E}(\tau(W^4))\Big| \\
&+3\Big|\mathbb{E}(\tau(W^1))\mathbb{E}(\tau(W^2))\mathbb{E}(\tau(W^3))\mathbb{E}(\tau(W^4))\Big|\Bigg),
\end{aligned}
\tag{3.15}
$$

where W^i $(i = 1, \ldots, 4)$ corresponds to a closed directed walk of length k in the complete directed graph of order n.

Recall that $\Gamma(\widehat{W})$ has edges $e_1, e_2, \ldots, e_l$ with multiplicities $\upsilon_1^\#, \upsilon_2^\#, \ldots, \upsilon_l^\#$, and $\upsilon_i^\# \geq 2$ for $i = 1, \ldots, l$, and $l \leq 2k$. Without loss of generality, we set $e_h = v_i v_j$. Then $\upsilon_h^\# = q_{ij}^\# + q_{ji}^\#$, where $q_{ij}^\#$ denotes the number of occurrences of the directed edge (i, j) in $\widehat{W}$. Then

$$
\mathbb{E}[\tau(W^1)\tau(W^2)\tau(W^3)\tau(W^4)] = \prod_{\substack{i<j \\ |E(\Gamma(\widehat{W}))|=l}} \mathbb{E}\left(\tau_{ij}^{q_{ij}^\#}\tau_{ji}^{q_{ji}^\#}\right). \tag{3.16}
$$

Next, we will compute $\mathbb{E}\left(\tau_{ij}^{q_{ij}^\#}\tau_{ji}^{q_{ji}^\#}\right)$. Note that $q_{ij}^\# + q_{ji}^\# \geq 2$ implies that $q_{ij}^\# \geq 1, q_{ji}^\# \geq 1$ or $q_{ij}^\# \geq 2, q_{ji}^\# = 0$ or $q_{ij}^\# = 0, q_{ji}^\# \geq 2$, since $|\tau_{ij}| < 1$ and $\mathbb{E}(\tau_{ij}) = 0$. We again consider these three cases.

If $q_{ij}^\# \geq 1, q_{ji}^\# \geq 1$, then we have

$$
\begin{aligned}
\left|\mathbb{E}\left(\tau_{ij}^{q_{ij}^\#}\tau_{ji}^{q_{ji}^\#}\right)\right| &\leq \mathbb{E}\left|\tau_{ij}^{q_{ij}^\#-1}\cdot\tau_{ji}^{q_{ji}^\#-1}\cdot\tau_{ij}\cdot\tau_{ij}\right| \\
&\leq 1^{q_{ij}^\#+q_{ji}^\#-2}\mathbb{E}|\tau_{ij}\tau_{ji}| \\
&= \mathbb{E}|\tau_{ij}|^2 \\
&= \frac{1}{n}\mathbb{E}|\eta'_{ij}|^2
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{1}{n}\int_{|x|<\sqrt{n}} |x|^2 \mathrm{d}\Phi \\
&\leq \frac{1}{n}. \qquad (3.17)
\end{aligned}
$$

If $q_{ij}^{\#} \geq 2, q_{ji}^{\#} = 0$, then we have

$$
\begin{aligned}
\left|\mathbb{E}\left(\tau_{ij}^{q_{ij}^{\#}}\tau_{ji}^{q_{ji}^{\#}}\right)\right| &\leq \mathbb{E}\left|\tau_{ij}^{q_{ij}^{\#}}\right| \\
&\leq 1^{q_{ij}^{\#}-2}\mathbb{E}|\tau_{ij}^2| \\
&\leq \frac{1}{n}. \qquad (3.18)
\end{aligned}
$$

If $q_{ij}^{\#} = 0, q_{ji}^{\#} \geq 2$, by a similar discussion as above, we have

$$
\left|\mathbb{E}\left(\tau_{ij}^{q_{ij}^{\#}}\tau_{ji}^{q_{ji}^{\#}}\right)\right| \leq \frac{1}{n}. \qquad (3.19)
$$

By (3.16), (3.17), (3.18) and (3.19), we have

$$
\left|\mathbb{E}[\tau(W^1)\tau(W^2)\tau(W^3)\tau(W^4)]\right| \leq \frac{1}{n^l}.
$$

If there are l_1 edges in $\Gamma(W^1)$ and there are l_2 edges in $\Gamma(W^2 \cup W^3 \cup W^4)$, then $l_1 + l_2 \geq l$, since $\mathbb{E}(\tau_{ij}) = \mathbb{E}(\frac{\eta'_{ij}}{\sqrt{n}}) = 0$, for all $1 \leq i < j \leq n$. So, $\mathbb{E}(\tau(W^1))$ is nonzero if and only if the total number of occurrences of each directed edge and its inverse edge of DK_n in the directed walk W^1 is at least 2. By (3.17), (3.18) and (3.19), we have

$$
\left|\mathbb{E}(\tau(W^1))\right| \leq \frac{1}{n^{l_1}}.
$$

Similarly, we have

$$
\left|\mathbb{E}[\tau(W^2)\tau(W^3)\tau(W^4)]\right| \leq \frac{1}{n^{l_2}}.
$$

Then

$$\left|\mathbb{E}(\tau(W^1))\mathbb{E}[\tau(W^2)\tau(W^3)\tau(W^4)]\right| \leq \frac{1}{n^{l_1+l_2}} \leq \frac{1}{n^l}.$$

Similarly we have

$$\left|\mathbb{E}(\tau(W^1))\mathbb{E}(\tau(W^2))\mathbb{E}[\tau(W^3)\tau(W^4)]\right| \leq \frac{1}{n^l},$$

and

$$\left|\mathbb{E}(\tau(W^1))\mathbb{E}(\tau(W^2))\mathbb{E}(\tau(W^3))\mathbb{E}(\tau(W^4))\right| \leq \frac{1}{n^l}.$$

Therefore,

$$\frac{1}{n^4}\sum_{W^1,\ldots,W^4}\left|\mathbb{E}\left\{\prod_{i=1}^{4}[\tau(W^i)-\mathbb{E}(\tau(W^i))]\right\}\right| \leq \frac{1}{n^4}\sum_{W^1,\ldots,W^4} 14\cdot\frac{1}{n^l}.$$

Note that there are at most two pieces of connected subgraphs in $\Gamma(\widehat{W})$. Then there are at most $l+2$ vertices in $\Gamma(\widehat{W})$. This shows that the number of such $\widehat{W}$ is at most $n^{l+2}C_{l,k}$, where $C_{l,k}$ is a constant depending on k and l only. Hence

$$\begin{aligned}
&\frac{1}{n^4}\sum_{W^1,\ldots,W^4}\left|\mathbb{E}\left\{\prod_{i=1}^{4}[\tau(W^i)-\mathbb{E}(\tau(W^i))]\right\}\right| \\
\leq&\frac{14}{n^4}\sum_{l=1}^{2k} n^{l+2}C_{l,k}\frac{1}{n^l} \\
=&\frac{14}{n^2}\sum_{l=1}^{2k} C_{l,k},
\end{aligned}$$

By (3.14) and (3.15), we have

$$\mathbb{E}[|M'_{k,n}-\mathbb{E}(M'_{k,n})|^4] = O(n^{-2}),\ k=1,2,\ldots.$$

Then

$$\sum_{n=1}^{\infty}\mathbb{E}[|M'_{k,n}-\mathbb{E}(M'_{k,n})|^4] = \sum_{n=1}^{\infty}O(n^{-2}) < \infty,\ k=1,2,\ldots.$$

By (3.13), we have

$$\sum_{n=1}^{\infty} \Pr(|M'_{k,n} - \mathbb{E}(M'_{k,n})| \geq \epsilon) < \infty, \ k = 1, 2, \ldots.$$

Note that the events $\{|M'_{k,n} - \mathbb{E}(M'_{k,n})| \geq \epsilon\}_{n=1}^{\infty}$ are independent. Then, by Lemma 3.5, we have

$$\Pr(|M'_{k,n} - \mathbb{E}(M'_{k,n})| \geq \epsilon) = 0,$$

i.e.,

$$\lim_{n\to\infty} M'_{k,n} = \lim_{n\to\infty} \mathbb{E}(M'_{k,n}) \ \ a.s.$$

This completes the proof of Fact 3.2. $\square$

Proof of Fact 3.3. Note that

$$M_{k,n} = \int x^k \mathrm{d}F^{W_n}(x) = \int x^k \mathrm{d}F^{n^{-1/2}M_n}(x)$$

and

$$M'_{k,n} = \int x^k \mathrm{d}F^{\widetilde{W_n}}(x) = \int x^k \mathrm{d}F^{n^{-1/2}\widetilde{M_n}}(x).$$

By Lemma 3.6, we have

$$\| F^{W_n} - F^{\widetilde{W_n}} \| = \| F^{n^{-1/2}M_n} - F^{n^{-1/2}\widetilde{M_n}} \| \leq \frac{1}{n} \mathrm{rank}(M_n - \widetilde{M_n}).$$

Notice that $\mathrm{rank}(M_n - \widetilde{M_n}) \leq$ the number of nonzero entries in $(M_n - \widetilde{M_n})$, which is bounded by $\sum_{jk} I_{\{|\eta_{jk}| \geq \sqrt{n}\}}$, where

$$I_{\{|\eta_{jk}| \geq \sqrt{n}\}} = \begin{cases} 0, & \text{if } |\eta_{jk}| < \sqrt{n}, \\ 1, & \text{if } |\eta_{jk}| \geq \sqrt{n}. \end{cases}$$

Then

$$\| F^{W_n} - F^{\widetilde{W_n}} \| \leq \frac{1}{n} \sum_{jk} I_{\{|\eta_{jk}| \geq \sqrt{n}\}}.$$

Let

$$p_{jk} = \Pr(|\eta_{jk}| \geq \sqrt{n}).$$

Since $\mathbb{E}(\eta\overline{\eta}) = \mathbb{E}|\eta|^2 = 1$, we have

$$\sum_{jk} p_{jk} = \sum_{jk} \Pr(|\eta_{jk}| \geq \sqrt{n}) \leq \frac{1}{n}\sum_{jk} \mathbb{E}|\eta_{jk}|^2 I_{\{|\eta_{jk}|\geq\sqrt{n}\}} = O(n).$$

Consider the $n(n-1)/2$ independent terms of $I_{\{|\eta_{jk}|\geq\sqrt{n}\}}$, $(1 \leq j < k \leq n)$, which are independent random variables, with

$$\Pr(I_{\{|\eta_{jk}|\geq\sqrt{n}\}} = 1) = p_{jk}, \quad \Pr(I_{\{|\eta_{jk}|\geq\sqrt{n}\}} = 0) = 1 - p_{jk},$$

and the sum of the $n(n-1)/2$ independent terms of $I_{\{|\eta_{jk}|\geq\sqrt{n}\}}$,

$$\mathbb{E}\left[\sum_{j<k} I_{\{|\eta_{jk}|\geq\sqrt{n}\}}\right] = \sum_{j<k} p_{jk} = \sum_{j<k} \Pr(|\eta_{jk}| \geq \sqrt{n}). \tag{3.20}$$

For any $\epsilon > 0$, applying Lemma 3.7 to (3.20), we have

$$\begin{aligned}
&\Pr\left(\frac{\sum_{j<k} I_{\{|\eta_{jk}|\geq\sqrt{n}\}}}{n} \geq \epsilon\right)\\
=&\Pr\left(\sum_{j<k} I_{\{|\eta_{jk}|\geq\sqrt{n}\}} \geq \epsilon n\right)\\
=&\Pr\left(\sum_{j<k} I_{\{|\eta_{jk}|\geq\sqrt{n}\}} - \mathbb{E}\left[\sum_{j<k} I_{\{|\eta_{jk}|\geq\sqrt{n}\}}\right] \geq \epsilon n - \sum_{j<k} p_{jk}\right)\\
\leq&\exp\left(-\frac{(\epsilon n - \sum_{j<k} p_{jk})^2}{2\left(\sum_{j<k} p_{jk} + \frac{\epsilon n - \sum_{j<k} p_{jk}}{3}\right)}\right)\\
=&\exp\left(-\frac{3(\epsilon n - \sum_{j<k} p_{jk})^2}{2\epsilon n + 5\sum_{j<k} p_{jk}}\right)\\
=&\exp(-bn),
\end{aligned}$$

for some positive constant b. Then, by Lemma 3.5, we have

$$\frac{\sum_{j<k} I_{\{|\eta_{jk}|\geq\sqrt{n}\}}}{n} \to 0 \ \ a.s. \ \ (n\to\infty)$$

Notice that with probability 1, the truncation does not affect the LSD of M_n. So

$$\|F^{n^{-1/2}M_n} - F^{n^{-1/2}\widetilde{M_n}}\| \leq \frac{1}{n}\sum_{jk} I_{\{|\eta_{jk}|\geq\sqrt{n}\}} \to 0.$$

Then we have

$$\lim_{n\to\infty} M_{k,n} = \lim_{n\to\infty} M'_{k,n} \ \ a.s.$$

This completes the proof of Fact 3.3. □

Proof of Theorem 3.8. Recall that

$$W_n = n^{-1/2}M_n = \frac{1}{\sigma\sqrt{n}}[(H_n+p^2I_n)-p^2J_n],$$

and set

$$W_n^0 = \frac{1}{\sigma\sqrt{n}}(H_n+p^2I_n).$$

Then

$$W_n^0 - W_n = \frac{1}{\sigma\sqrt{n}}\cdot p^2J_n.$$

Note that

$$\operatorname{rank}\left(\frac{1}{\sigma\sqrt{n}}\cdot p^2J_n\right) = 1.$$

By Lemma 3.6, we have

$$\|F^{W_n^0}(x) - F^{W_n}(x)\| \leq \frac{1}{n}\cdot 1 = \frac{1}{n}.$$

This implies that the LSDs of W_n^0 and W_n are the same. By Theorem 3.9, we have

$$\lim_{n\to\infty} F^{W_n^0}(x) = \lim_{n\to\infty} F^{W_n}(x) = F(x) := \int_{-\infty}^{x} \phi(x)\mathrm{d}x. \tag{3.21}$$

Consider the matrices $W_n^1 = \frac{1}{\sigma\sqrt{n}} H_n$ and $W_n^0 = \frac{1}{\sigma\sqrt{n}}(H_n + p^2 I_n)$. Note that

$$W_n^0 - W_n^1 = \frac{1}{\sigma\sqrt{n}} \cdot p^2 I_n := \Delta_n I_n,$$

and

$$\Delta_n = \frac{1}{\sigma\sqrt{n}} p^2 \to 0\,(n \to \infty).$$

Note also that λ is an eigenvalue of W_n^1 if and only if $\lambda + \Delta_n$ is an eigenvalue of W_n^0. Then

$$F^{W_n^1}(x) = F^{W_n^0}(x + \Delta_n).$$

On the other hand, $\Delta_n \to 0\,(n \to \infty)$ implies that for any $\epsilon > 0$, there exists an N such that $|\Delta_n| < \epsilon$ for all $n > N$. Since $F^{W_n^0}(x)$ is an increasing function for all $n > N$, we have

$$F^{W_n^0}(x - \epsilon) \leq F^{W_n^0}(x + \Delta_n) \leq F^{W_n^0}(x + \epsilon).$$

Then

$$\begin{aligned} F(x-\epsilon) &= \lim_{n\to\infty} F^{W_n^0}(x - \epsilon) \\ &\leq \lim_{n\to\infty} F^{W_n^0}(x + \Delta_n) \\ &\leq \lim_{n\to\infty} F^{W_n^0}(x + \epsilon) \\ &= F(x+\epsilon) \ \ a.s. \end{aligned}$$

From (3.21), we see that the density of $F(x)$ is smooth. Then $F(x)$ is continuous. By choosing $\epsilon > 0$ as small as possible, we conclude that

$$\lim_{n\to\infty} F^{W_n^1}(x) = \lim_{n\to\infty} F^{W_n^0}(x + \Delta_n) = F(x) \ \ a.s.$$

i.e.,

$$\lim_{n\to\infty} F^{\frac{1}{\sigma\sqrt{n}} H_n}(x) = F(x) \ \ a.s.$$

This completes the proof. □

3.3 The Hermitian energy

In this section we give an estimation of the Hermitian energy for almost all mixed graphs.

Theorem 3.10. *Let $p = p(n)$, $0 < p < 1$. Then the Hermitian energy $\mathscr{E}_H(\widehat{G}_n(p))$ of the random mixed graph $\widehat{G}_n(p)$ satisfies almost surely (a.s.) the following equation:*

$$\mathscr{E}_H(\widehat{G}_n(p)) = n^{3/2}(2p - p^2 - p^4)^{1/2}\left(\frac{8}{3\pi} + o(1)\right),$$

that is, with probability 1, $\mathscr{E}_H(\widehat{G}_n(p))$ satisfies the above equation as $n \to \infty$.

In order to prove the above theorem, we need the following results.

Lemma 3.11 (See [18]). *Let μ be a measure. Suppose that the functions a_n, b_n, and f_n converge almost everywhere to the functions a, b, and f, respectively, and that $a_n \leq f_n \leq b_n$ almost everywhere. If $\int a_n \mathrm{d}\mu \to \int a \mathrm{d}\mu$ and $\int b_n \mathrm{d}\mu \to \int b \mathrm{d}\mu$, then $\int f_n \mathrm{d}\mu \to \int f \mathrm{d}\mu$.*

Theorem 3.12. *Define $\sigma = \sqrt{2p - p^2 - p^4}$. Let H_n be an Hermitian adjacency matrix of a random mixed graph $\widehat{G}_n(p)$ with $p = p(n)$, $0 < p < 1$. Let $\phi(x)$ be as in Theorem 3.8, and $F(x) = \int_{-\infty}^{x} \phi(x)\mathrm{d}x$. Then*

$$\lim_{n\to\infty} \int |x| \mathrm{d}F^{\frac{1}{\sigma\sqrt{n}}H_n}(x) = \int |x| \mathrm{d}F(x) = \int |x|\phi(x)\mathrm{d}x \quad a.s.$$

Proof of Theorem 3.12. Note that $F^{\frac{1}{\sigma\sqrt{n}}H_n}(x) = \int_{-\infty}^{x} \phi^{\frac{1}{\sigma\sqrt{n}}H_n}(x)\mathrm{d}x$ and $F(x) = \int_{-\infty}^{x} \phi(x)\mathrm{d}x$. Note also that

$$\lim_{n\to\infty} F^{\frac{1}{\sigma\sqrt{n}}H_n}(x) = F(x).$$

Then

$$\lim_{n\to\infty} \phi^{\frac{1}{\sigma\sqrt{n}}H_n}(x) = \phi(x).$$

Let $\boldsymbol{I}$ be the interval $[-2,2]$, and $\boldsymbol{I}^C$ the set $\mathbb{R}\backslash\boldsymbol{I}$. Since $\phi(x)$ is bounded on $\boldsymbol{I}$, it follows that with probability 1, $x^2\phi^{\frac{1}{\sigma\sqrt{n}}H_n}(x)$ is bounded almost everywhere on $\boldsymbol{I}$. By the Bounded Convergence Theorem (See [111]), we have

$$\lim_{n\to\infty}\int_{\boldsymbol{I}} x^2 \mathrm{d}F^{\frac{1}{\sigma\sqrt{n}}H_n}(x) = \int_{\boldsymbol{I}} x^2 \mathrm{d}F(x) \;\; a.s.$$

Then

$$\begin{aligned}
\lim_{n\to\infty}\int_{\boldsymbol{I}^C} x^2 \mathrm{d}F^{\frac{1}{\sigma\sqrt{n}}H_n}(x) &= \lim_{n\to\infty}\left(\int x^2 \mathrm{d}F^{\frac{1}{\sigma\sqrt{n}}H_n}(x) - \int_{\boldsymbol{I}} x^2 \mathrm{d}F^{\frac{1}{\sigma\sqrt{n}}H_n}(x)\right) \\
&= \lim_{n\to\infty}\int x^2 \mathrm{d}F^{\frac{1}{\sigma\sqrt{n}}H_n}(x) - \lim_{n\to\infty}\int_{\boldsymbol{I}} x^2 \mathrm{d}F^{\frac{1}{\sigma\sqrt{n}}H_n}(x) \\
&= \int x^2 \mathrm{d}F(x) - \int_{\boldsymbol{I}} x^2 \mathrm{d}F(x) \;\; a.s. \\
&= \int_{\boldsymbol{I}^C} x^2 \mathrm{d}F(x) \;\; a.s. \qquad (3.22)
\end{aligned}$$

Set

$$a_n(x) = 0, \;\; b_n(x) = x^2\phi^{\frac{1}{\sigma\sqrt{n}}H_n}(x), \text{ and } f_n(x) = |x|\phi^{\frac{1}{\sigma\sqrt{n}}H_n}(x).$$

Notice that

$$|x| \le x^2, \quad \text{if } \; x \in \boldsymbol{I}^C.$$

Then

$$a_n(x) \le f_n(x) \le b_n(x), \quad \text{if } \; x \in \boldsymbol{I}^C.$$

By Lemma 3.11 and (3.22), we have

$$\lim_{n\to\infty}\int_{\boldsymbol{I}^C} |x|\phi^{\frac{1}{\sigma\sqrt{n}}H_n}(x)\mathrm{d}x = \int_{\boldsymbol{I}^C} |x|\phi(x)\mathrm{d}x \;\; a.s.,$$

i.e.,

$$\lim_{n\to\infty}\int_{\boldsymbol{I}^C} |x|\mathrm{d}F^{\frac{1}{\sigma\sqrt{n}}H_n}(x) = \int_{\boldsymbol{I}^C} |x|\mathrm{d}F(x) \;\; a.s. \qquad (3.23)$$

Note that with probability 1, $|x|\phi^{\frac{1}{\sigma\sqrt{n}}H_n}(x)$ is bounded almost everywhere on $\boldsymbol{I}$, since $\phi(x)$ is bounded on $\boldsymbol{I}$. Again, by the Bounded Convergence Theorem (See [111]), we have

$$\lim_{n\to\infty}\int_{I}|x|\mathrm{d}F^{\frac{1}{\sigma\sqrt{n}}H_n}(x)=\int_{I}|x|\mathrm{d}F(x)\ \ a.s. \tag{3.24}$$

By (3.23) and (3.24), we have

$$\lim_{n\to\infty}\int|x|\mathrm{d}F^{\frac{1}{\sigma\sqrt{n}}H_n}(x)=\int|x|\mathrm{d}F(x)=\int|x|\phi(x)\mathrm{d}x\ \ a.s.$$

This completes the proof. □

Proof of Theorem 3.10. Recall that $\sigma=\sqrt{2p-p^2-p^4}$, and H_n denotes the Hermitian adjacency matrix of $\widehat{G}_n(p)$. Suppose that $\lambda_1,\lambda_2,\dots,\lambda_n$ and $\lambda_1',\lambda_2',\dots,\lambda_n'$ are the eigenvalues of H_n and $\frac{1}{\sigma\sqrt{n}}H_n$, respectively. By Theorem 3.9, the ESD of $n^{-1/2}M_n$ converges to the standard semicircle distribution whose density is given by

$$\phi(x)=\begin{cases}\frac{1}{2\pi}\sqrt{4-x^2}, & \text{for } |x|\le 2,\\ 0, & \text{for } |x|>2.\end{cases}$$

By Theorem 3.12, we have

$$\begin{aligned}\frac{\mathscr{E}_H(\widehat{G}_n(p))}{\sigma n^{\frac{3}{2}}}&=\frac{1}{\sigma n^{\frac{3}{2}}}\sum_{i=1}^{n}|\lambda_i|\\&=\frac{1}{n}\sum_{i=1}^{n}\left|\frac{1}{\sigma\sqrt{n}}\lambda_i\right|\\&=\frac{1}{n}\sum_{i=1}^{n}|\lambda_i'|\\&=\int|x|\mathrm{d}F^{\frac{1}{\sigma\sqrt{n}}H_n}(x)\\&\to\int|x|\mathrm{d}F(x)(n\to\infty)\end{aligned}$$

$$
\begin{aligned}
&= \int |x|\phi(x)\mathrm{d}x \\
&= \frac{1}{2\pi}\int_{-2}^{2} |x|\sqrt{4-x^2}\mathrm{d}x \\
&= \frac{8}{3\pi}.
\end{aligned}
$$

This completes the proof. $\square$

Chapter 4

The spectrum of H_n for random mixed graphs

Let $\widehat{G}_n(p)$ be a random mixed graph as described in the introduction of Chapter 3. In Chapter 3, we proved that the empirical distribution of the eigenvalues of the Hermitian adjacency matrix of $\widehat{G}_n(p)$ converges to Wigner's semicircle law. Since Theorem 3.8 only characterizes the limiting spectral distribution of the Hermitian adjacency matrix of random mixed graphs, it does not describe the behaviour of the largest eigenvalue of the Hermitian adjacency matrix. In this chapter, we deal with the asymptotic behaviour of the spectrum of the Hermitian adjacency matrix of random mixed graphs. We will present and prove a separation result between the first and the remaining eigenvalues of H_n. As an application of the asymptotic behaviour of the spectrum of the Hermitian adjacency matrix, we estimate the spectral moments of random mixed graphs.

4.1 Preliminaries

We start with some notations and lemmas that we will use throughout the chapter. Let $\lambda_1(G), \lambda_2(G), \ldots, \lambda_n(G)$ be (not necessarily distinct) eigenvalues of the Hermitian adjacency matrix $H(G)$ of a mixed graph G of order n. Recall that the *k-th Hermitian spectral moment* of G is defined as

$$s_k(H(G)) = \sum_{i=1}^{n} \lambda_i^k(G),$$

where $k \geq 0$ is an integer. It is easy to verify that $s_1(H(G)) = \mathrm{Tr}(H(G)) = 0$. As an application of the asymptotic behaviour of the spectrum of the Hermitian adjacency matrix, we estimate the spectral moments of random mixed graphs.

A closed walk W is *good* if each edge in $E(W)$ occurs more than once. Let $\mathscr{G}^0(n,k,m)$ be the set of walks in K_n using k edges and m vertices where each edge in the walk is used at least twice, i.e., let $\mathscr{G}^0(n,k,m)$ be the set of good closed walks in K_n of length k and with m vertices. In [57], Füredi and Komlós proved the following result.

Lemma 4.1 (Füredi and Komlós [57]). *Let $\mathscr{G}^0(n,k,m)$ be the set of good closed walks in K_n of length k and with m vertices. For $m < n+1$,*

$$|\mathscr{G}^0(n,k,m)| \leq n(n-1)\cdots(n-m+1)\frac{1}{m}\binom{2m-2}{m-1}\binom{k}{2m-2}m^{2(k-2m+2)}.$$

Let $\tilde{\mathscr{G}}^0(k,m)$ be the set of good closed walks W of length k in K_m where vertices first appear in W in the order $1,2,\ldots,m$. The main contribution from Vu's paper [116] is the following bound.

Lemma 4.2 (Vu [116]). *Let $\tilde{\mathscr{G}}^0(k,m)$ be the set of good closed walks W of length k in K_m where vertices first appear in W in the order $1,2,\ldots,m$. Then*

$$|\tilde{\mathscr{G}}^0(k,m)| \leq \binom{k}{2m-2}2^{2k-2m+3}m^{k-2m+2}(k-2m+4)^{k-2m+2}.$$

It is easy to check that $|\mathscr{G}^0(n,k,m)| = n(n-1)\cdots(n-m+1)|\tilde{\mathscr{G}}^0(k,m)|$. Thus, this combination with the bound in Lemma 4.2 improves Füredi-Komlós' upper bound.

For a directed edge $e = (v_1, v_2)$, the vertices v_1, v_2 are called the *ends* of e, while v_1 is the *initial* (vertex) of e, and v_2 is the *terminal* (vertex) of e. If two directed edges have the same set of ends, they are said to be *coincident*. If there is no directed edge with the same set of ends as the directed edge (u, v), (u, v) is said to be *single*. Let $\mathscr{I} = (i_1, \ldots, i_k)$ be a vector valued on $\{1, \ldots, n\}^k$. With the vector $\mathscr{I}$, Bai et al. [11, 12] defined a $\mathbf{\Gamma}$-graph as follows. Draw a horizontal line and plot the numbers $i_1, \ldots, i_k$ on it. Consider the distinct numbers as vertices, and draw k directed edges e_j from i_j to i_{j+1}, $j = 1, \ldots, k$, where $i_{k+1} = i_1$ by convention. Denote the number of distinct i_j's by m. Such a graph is called a *$\mathbf{\Gamma}(k, m)$-graph*. An example of a $\mathbf{\Gamma}(7, 4)$-graph is given in Figure 4.1, in which there are 8 vertices (i_1 up to i_8), 4 non-coincident vertices (v_1 up to v_4), 7 edges, a maximum of 4 mutually non-coincident edges (e.g., the non-dashed (solid) edges indicated in Figure 4.1), and 2 single edges (v_4, v_3) and (v_2, v_4). By definition, we can traverse all edges of the $\mathbf{\Gamma}(k, m)$-graph by starting from vertex i_1, and traversing the k directed edges consecutively from i_1 to i_2, i_2 to i_3, etc., and finally returning to i_1 by using the edge from i_k to i_1. That is, a $\mathbf{\Gamma}(k, m)$-graph represents a closed directed walk (possibly containing loops).

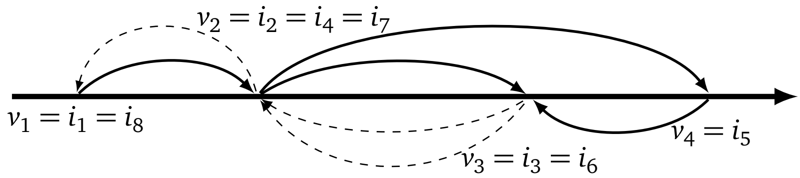


FIGURE 4.1: $\mathbf{\Gamma}(7, 4)$-graph

A closed directed walk W is *good* if the total number of occurrences of each directed edge and its inverse edge in the directed walk W is at least 2. The set of all good closed directed walks of length k in DK_n is denoted by $\mathscr{G}(n, k)$. Let $\mathscr{G}(n, k, t)$ denote the set of closed good directed walks on DK_n of length k using exactly t different vertices.

By definition, a good directed walk in DK_n using k edges and m vertices is indeed equivalent to a $\mathbf{\Gamma}(k, m)$-graph which has no single directed edge. For any $\mathbf{\Gamma}(k, m)$-graph which has no single directed edge, if we ignore the orientation, we will obtain the equivalent of a good walk in K_n using k edges and m vertices. Thus, we have

$$|\mathscr{G}(n, k, m)| \geq |\mathscr{G}^0(n, k, m)|.$$

On the other hand, for any good walk $W^0 = i_1 i_2 \dots i_{k-1} i_k i_1$ in K_n using k edges and m vertices, if we add the orientation (i_j, i_{j+1}) $(1 \le j \le k, i_{k+1} = i_1)$, then we will obtain a good directed walk in DK_n using k edges and m vertices. Thus, we have

$$|\mathcal{G}(n,k,m)| \le |\mathcal{G}^0(n,k,m)|.$$

Hence,

$$|\mathcal{G}(n,k,m)| = |\mathcal{G}^0(n,k,m)|.$$

The following result is immediate.

Lemma 4.3. *Let $\mathcal{G}(n,k,m)$ be the set of good closed directed walks in DK_n using k edges and m vertices. Then $|\mathcal{G}(n,k,m)| \le$*

$$n(n-1)\cdots(n-m+1)\binom{k}{2m-2}2^{2k-2m+3}m^{k-2m+2}(k-2m+4)^{k-2m+2}.$$

4.2 Spectral bounds

In this section, we study the spectrum of the Hermitian adjacency matrix of $\widehat{G}_n(p)$. In Chapter 3, we proved that the empirical distribution of the eigenvalues of the Hermitian adjacency matrix H_n follows Wigner's semicircle law. In particular, for any $c > 2\sigma$, with probability $1-o(1)$ all eigenvalues of H_n except for at most $o(n)$ lie in the interval $\boldsymbol{I} = (-c\sqrt{n}, c\sqrt{n})$ (where $\sigma = \sqrt{2p - p^2 - p^4}$). In this chapter, we show that with probability $1-o(1)$ all eigenvalues except for the largest eigenvalue $\lambda_1(H_n)$ belong to the above interval $\boldsymbol{I}$, that is, only the largest eigenvalue $\lambda_1(H_n)$ (possibly) is outside $\boldsymbol{I}$. Our main result is stated as follows.

Theorem 4.4. *Suppose that C and C' are sufficiently large. Let H_n denote the Hermitian adjacency matrix of $\widehat{G}_n(p)$. Let the eigenvalues of H_n be $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_n$. Define $K = \sqrt{\frac{1+p^4}{2}}$ and $\sigma = \sqrt{2p-p^2-p^4}$. If $\sigma \ge C'Kn^{-\frac{1}{2}}\ln^2 n$, then asymptotically almost surely we have*

(i) $\lambda_1 = (1+o(1))np^2$,

(ii) $\max\limits_{2 \le i \le n} |\lambda_i| \le 2\sigma\sqrt{n} + C\sqrt{K\sigma}n^{\frac{1}{4}}\ln n.$

That $\max\limits_{2\le i\le n}|\lambda_i|$ cannot be much smaller than $2\sigma\sqrt{n}$, is guaranteed by the semicircle law. We postpone the proof of Theorem 4.4. Our proof is based on the following theorem that we prove first. In the following $|x|$ denotes the Euclidean norm of $x \in \mathbf{R}$.

Theorem 4.5. *Let J_n be the all 1's matrix, let H_n be the Hermitian adjacency matrix of $\widehat{G}_n(p)$, and let $U_n = p^2 J_n - H_n$. Suppose that C and C' are sufficiently large. Define $K = \sqrt{\frac{1+p^4}{2}}$ and $\sigma = \sqrt{2p-p^2-p^4}$. If $\sigma \ge C' K n^{-\frac{1}{2}} \ln^2 n$, then asymptotically almost surely*

$$\|U_n\| \le 2\sigma\sqrt{n} + C\sqrt{K\sigma}\, n^{\frac{1}{4}} \ln n.$$

Here $\|U_n\| = \max\limits_{|x|=1}|U_n x| = \max\{|\lambda_1(U_n)|, |\lambda_n(U_n)|\}$.

Before presenting the proof of Theorem 4.5, we recall the following well-known result that will be used in the sequel of the chapter.

Lemma 4.6 (Markov's Inequality [112])**.** *Let X be a nonnegative, real-valued random variable and $a > 0$. Then*

$$\Pr(X > a) \le \frac{\mathbb{E}(X)}{a}.$$

Proof of Theorem 4.5. We rely on Wigner's high moment method. We define

$$\widehat{M_n} = \frac{1}{\sqrt{2}}(U_n - p^2 I_n),$$

where I_n denotes the $n \times n$ identity matrix. This implies that

$$\|U_n\| \le \sqrt{2}\|\widehat{M_n}\| + \|p^2 I_n\| = \sqrt{2}\|\widehat{M_n}\| + p^2.$$

So, it remains to establish the right upper bound for $\|\widehat{M_n}\|$.

Recall that $H_n = (h_{ij})_{n\times n}$ is the Hermitian adjacency matrix of $\widehat{G}_n(p)$. Then h_{ij} $(1 \le i < j \le n)$ are independent random variables with the following properties:

- $\mathbb{E}(h_{ij}) = p^2$;

- $\mathrm{Var}(h_{ij}) = 2p - p^2 - p^4 = \sigma^2$;
- $h_{ij}, h_{i'j'}$ are independent, unless $(i,j) = (j',i')$. If $i > j$, we have $\overline{h_{ji}} = h_{ij}$, i.e., h_{ji} is the complex conjugate of h_{ij};
- $|h_{ij}| \le 1$.

Let m_{ij} denote the (i,j)-th entry of $\widehat{M_n}$. Using the definitions and the above properties, we easily deduce that m_{ij} $(1 \le i < j \le n)$ are independent random variables with the following properties:

- $\mathbb{E}(m_{ij}) = 0$;
- $\mathrm{Var}(m_{ij}) = \frac{\mathrm{Var}(h_{ij})}{2} = \frac{2p-p^2-p^4}{2} = \frac{\sigma^2}{2}$;
- $m_{ij}, m_{i'j'}$ are independent, unless $(i,j) = (j',i')$. If $i > j$, we have $\overline{m_{ji}} = m_{ij}$;
- $|m_{ij}| \le \sqrt{\frac{1+p^4}{2}} = K \le 1$.

Now let $k \ge 2$ be an even integer. We estimate

$$\begin{aligned}\mathrm{Tr}(\widehat{M_n}^k) =& \sum_{i=1}^{n} \lambda_i(\widehat{M_n})^k \\ \ge& \max\{\lambda_1(\widehat{M_n})^k, \lambda_n(\widehat{M_n})^k\} \\ =& \|\widehat{M_n}\|^k.\end{aligned}$$

A standard fact in linear algebra tells us that for any positive integer k,

$$\mathrm{Tr}(\widehat{M_n}^k) = \sum_{i_1,\ldots,i_k \in [n]} m_{i_1i_2} m_{i_2i_3} \cdots m_{i_ki_1}, \tag{4.1}$$

where $[n] = \{1, 2, \ldots, n\}$.

Let us now take a closer look at $\mathrm{Tr}(\widehat{M_n}^k)$. This is a sum where a typical term is $m_{i_1i_2} m_{i_2i_3} \ldots m_{i_{k-1}i_k} m_{i_ki_1}$, where $W := i_1 i_2 \ldots i_{k-1} i_k i_1$ corresponds to a closed directed walk of length k in the complete directed graph DK_n of order n. In other words, each term corresponds to a closed walk of length k (containing k, not necessarily distinct, directed edges) of the complete directed graph DK_n on $[n]$. For each directed edge $(i,j) \in W$, let q_{ij} be the number of occurrences of the directed edge (i,j) in the walk W. Note that all directed edges of a mixed graph are mutually independent. We rewrite (4.1) as

$$\mathrm{Tr}(\widehat{M_n}^k) = \sum_W \prod_{i<j} m_{ij}^{q_{ij}} m_{ji}^{q_{ji}}. \tag{4.2}$$

Then

$$\mathbb{E}(\mathrm{Tr}(\widehat{M_n}^k)) = \mathbb{E}\left(\sum_W \prod_{i<j} m_{ij}^{q_{ij}} m_{ji}^{q_{ji}}\right) = \sum_W \prod_{i<j} \mathbb{E}\left(m_{ij}^{q_{ij}} m_{ji}^{q_{ji}}\right),$$

where the summation is taken over all directed closed walks of length k.

We decompose $\mathbb{E}(\mathrm{Tr}(\widehat{M_n}^k))$ into parts $\mathbb{E}_{n,k,t}$, $t = 2, \ldots, k$, containing the t-fold sums, as follows:

$$\mathbb{E}(\mathrm{Tr}(\widehat{M_n}^k)) = \sum_{t=2}^{k} \mathbb{E}_{n,k,t}, \tag{4.3}$$

where

$$\mathbb{E}_{n,k,t} = \sum_{\{W:|V(W)|=t\}} \prod_{i<j} \mathbb{E}\left(m_{ij}^{q_{ij}} m_{ji}^{q_{ji}}\right), \tag{4.4}$$

and $|V(W)| = t$ means the cardinality of the vertex set of W is t. (Note that as $m_{ii} = 0$ by construction of $\widehat{M_n}$ we have that $\mathbb{E}_{n,k,1} = 0$.) So, the summation in (4.4) is taken over all closed directed walks W of length k using exactly t different vertices.

Recall that the entries m_{ij} of $\widehat{M_n}$ are independent random variables with mean zero, i.e., $\mathbb{E}(m_{ij}) = 0$, for all $1 \le i < j \le n$, and recall also that q_{ij} denotes the number of occurrences of the directed edge (i,j) in the closed walk W. So, if $q_{ij} + q_{ji} = 1$, that is, $q_{ij} = 1, q_{ji} = 0$ or $q_{ij} = 0, q_{ji} = 1$, then $\prod_{i<j} \mathbb{E}\left(m_{ij}^{q_{ij}} m_{ji}^{q_{ji}}\right) = 0$. Thus, the expectation of a term is nonzero if and only if the total number of occurrences of each directed edge and its inverse edge of DK_n in the directed walk W is at least 2, i.e., we only need to consider the case that $q_{ij} + q_{ji} \ge 2$. In other words, we only need to consider good directed walks. The set of all good closed directed walks of length k in DK_n is denoted by $\mathscr{G}(n,k)$. Considering a good closed directed

walk W, the underlying graph $\Gamma(W)$ of W uses l different edges $e_1, \dots, e_l$, i.e., $|E(\Gamma(W))| = l$, with corresponding multiplicities $s_1, \dots, s_l$ (where the s_hs are positive integers at least 2 summing up to k). Without loss of generality, we set $e_h = v_i v_j$, so that $s_h = q_{ij} + q_{ji}$. The (expected) contribution of the term defined by this directed walk to $\mathbb{E}(\mathrm{Tr}(\widehat{M_n}^k))$ is

$$\prod_{\substack{i<j \\ |E(\Gamma(W))|=l}} \mathbb{E}\left(m_{ij}^{q_{ij}} m_{ji}^{q_{ji}}\right). \tag{4.5}$$

Next, we will compute $\mathbb{E}\left(m_{ij}^{q_{ij}} m_{ji}^{q_{ji}}\right)$. Note that $q_{ij} + q_{ji} \geq 2$ implies that $q_{ij} \geq 1, q_{ji} \geq 1$ or $q_{ij} \geq 2, q_{ji} = 0$ or $q_{ij} = 0, q_{ji} \geq 2$, since $|m_{ij}| \leq K \leq 1$ and $\mathbb{E}(m_{ij}) = 0$. We consider these three cases separately.
If $q_{ij} \geq 1, q_{ji} \geq 1$, then we have

$$\begin{aligned}
\left|\mathbb{E}\left(m_{ij}^{q_{ij}} m_{ji}^{q_{ji}}\right)\right| &\leq \mathbb{E}\left|m_{ij}^{q_{ij}-1} \cdot m_{ji}^{q_{ji}-1} \cdot m_{ij} \cdot m_{ij}\right| \\
&\leq K^{q_{ij}+q_{ji}-2}\mathbb{E}|m_{ij}m_{ji}| \\
&= K^{q_{ij}+q_{ji}-2}\mathbb{E}|m_{ij}|^2 \\
&= K^{q_{ij}+q_{ji}-2}\mathbb{E}(m_{ij}\overline{m_{ij}}) \\
&= K^{q_{ij}+q_{ji}-2}\mathrm{Var}(m_{ij}) \\
&= \frac{1}{2}K^{q_{ij}+q_{ji}-2}\sigma^2.
\end{aligned} \tag{4.6}$$

If $q_{ij} \geq 2, q_{ji} = 0$, then we have

$$\begin{aligned}
\left|\mathbb{E}\left(m_{ij}^{q_{ij}} m_{ji}^{q_{ji}}\right)\right| &\leq \mathbb{E}\left|m_{ij}^{q_{ij}}\right| \\
&\leq K^{q_{ij}-2}\mathbb{E}|m_{ij}^2| \\
&= \frac{1}{2}K^{q_{ij}-2}\sigma^2.
\end{aligned} \tag{4.7}$$

If $q_{ij} = 0, q_{ji} \geq 2$, then similarly, we have

$$\left|\mathbb{E}\left(m_{ij}^{q_{ij}} m_{ji}^{q_{ji}}\right)\right| \leq \frac{1}{2}K^{q_{ji}-2}\sigma^2. \tag{4.8}$$

Let $\mathscr{G}(n,k,t)$ denote the set of good closed directed walks on DK_n of length

k using exactly t different vertices. Notice that for each directed walk W in $\mathscr{G}(n,k,l+1)$, the underlying graph $\Gamma(W)$ of W must have at least l different edges. By (4.5)-(4.8), the contribution of a term corresponding to such a good directed walk to $\mathbb{E}(\mathrm{Tr}(\widehat{M_n}^k))$ is at most

$$\frac{1}{2^l}K^{k-2l}\sigma^{2l}.$$

By the pigeon hole principle, if $l+1 > \frac{k}{2}+1$, then there must be a directed edge (i,j) such that the total number of occurrences of this directed edge and its inverse edge of DK_n in the directed walk W is 1, i.e., $q_{ij}+q_{ji}=1$. As we argued before, this implies $\mathbb{E}_{n,k,l+1}=0$ for $l > \frac{k}{2}$.

So, in the following, we only consider the case that $l \leq \frac{k}{2}$ and $q_{ij}+q_{ji} \geq 2$. Using Lemma 4.3, we have

$$\begin{aligned}
&\mathbb{E}(\mathrm{Tr}(\widehat{M_n}^k)) \\
\leq &\sum_{l=1}^{\frac{k}{2}} |\mathscr{G}(n,k,l+1)| \frac{1}{2^l}K^{k-2l}\sigma^{2l} \\
= &\sum_{m=2}^{\frac{k}{2}+1} |\mathscr{G}(n,k,m)| \frac{1}{2^{m-1}}K^{k-2(m-1)}\sigma^{2(m-1)} \\
\leq &\sum_{m=2}^{\frac{k}{2}+1} \frac{K^{k-2(m-1)}\sigma^{2(m-1)}}{2^{m-1}} n^m \binom{k}{2m-2} 2^{2k-2m+3} m^{k-2m+2}(k-2m+4)^{k-2m+2} \\
= &\sum_{m=2}^{\frac{k}{2}+1} S(n,k,m), \qquad (4.9)
\end{aligned}$$

where the final equality defines $S(n,k,m)$. Now fix $k = g(n)\ln n$, where $g(n)$ tends to infinity (with n) arbitrarily slowly. Let us consider the ratio $S(n,k,m-1)/S(n,k,m)$ for some $m \leq \frac{k}{2}+1$:

$$\frac{S(n,k,m-1)}{S(n,k,m)}$$

$$
\begin{aligned}
&=\frac{\frac{K^{k-2(m-2)}\sigma^{2(m-2)}}{2^{m-2}}n^{m-1}\binom{k}{2m-4}2^{2k-2m+5}(m-1)^{k-2m+4}(k-2m+6)^{k-2m+4}}{\frac{K^{k-2(m-1)}\sigma^{2(m-1)}}{2^{m-1}}n^{m}\binom{k}{2m-2}2^{2k-2m+3}m^{k-2m+2}(k-2m+4)^{k-2m+2}}\\
&=\frac{K^2(2m-2)(2m-3)2^2(m-1)^{k-2m+4}(k-2m+6)^{k-2m+4}}{\frac{\sigma^2}{2}n(k-2m+4)(k-2m+3)m^{k-2m+2}(k-2m+4)^{k-2m+2}}\\
&\leq\frac{K^2 4(m-1)^2 2^2(m-1)^{k-2m+4}(k-2m+6)^{k-2m+4}}{\frac{\sigma^2}{2}n(k-2m+4)(k-2m+3)m^{k-2m+2}(k-2m+4)^{k-2m+2}}\\
&=\frac{16K^2(m-1)^{k-2m+6}(k-2m+6)^{k-2m+4}}{\frac{\sigma^2}{2}n(k-2m+4)(k-2m+3)m^{k-2m+2}(k-2m+4)^{k-2m+2}}\\
&\leq\frac{16K^2m^{k-2m+6}(k-2m+6)^{k-2m+4}}{\frac{\sigma^2}{2}n(k-2m+4)(k-2m+3)m^{k-2m+2}(k-2m+4)^{k-2m+2}}\\
&\leq\frac{32K^2m^4(k-2m+6)^{k-2m+4}}{\sigma^2n(k-2m+3)(k-2m+4)^{k-2m+3}}\\
&\leq\frac{32K^2m^4(k-2m+6)^{k-2m+4}}{\sigma^2n(k-2m+3)^{k-2m+4}}\\
&\to\frac{32C_0K^2m^4}{\sigma^2n}\\
&\leq\frac{32C_0K^2k^4}{\sigma^2n}
\end{aligned}
$$

for some constant C_0 independent of σ and K. This implies that

$$S(n,k,m-1)\leq\frac{32C_0K^2k^4}{\sigma^2n}S(n,k,m).$$

With a proper choice of $g(n)$ guaranteeing that $k^4\leq\frac{\sigma^2n}{64C_0K^2}$, we have

$$S(n,k,m-1)\leq\frac{1}{2}S(n,k,m).$$

Then

$$\mathbb{E}(\mathrm{Tr}(\widehat{M_n}^k))\leq\sum_{m=2}^{\frac{k}{2}+1}S(n,k,m)$$

$$
\begin{aligned}
&= S\left(n, k, \frac{k}{2}+1\right) \sum_{m=2}^{\frac{k}{2}+1} \left(\frac{1}{2}\right)^{\frac{k}{2}+1-m} \\
&\le 2S\left(n, k, \frac{k}{2}+1\right) \\
&= 2\frac{\sigma^k}{2^{\frac{k}{2}}} n^{\frac{k}{2}+1} 2^{k+1} \\
&= 4n(\sigma\sqrt{2n})^k.
\end{aligned}
$$

Then

$$\mathbb{E}(\|\widehat{M_n}^k\|) \le \mathbb{E}(\mathrm{Tr}(\widehat{M_n}^k)) \le 4n(\sigma\sqrt{2n})^k.$$

Using Lemma 4.6, we get

$$
\begin{aligned}
&\Pr(\|\widehat{M_n}\| \ge \sigma\sqrt{2n} + C''\sqrt{K\sigma} n^{\frac{1}{4}} \ln n) \\
=&\Pr(\|\widehat{M_n}\|^k \ge (\sigma\sqrt{2n} + C''\sqrt{K\sigma} n^{\frac{1}{4}} \ln n)^k) \\
\le&\frac{\mathbb{E}(\|\widehat{M_n}^k\|)}{(\sigma\sqrt{2n} + C''\sqrt{K\sigma} n^{\frac{1}{4}} \ln n)^k} \\
\le&\frac{4n(\sigma\sqrt{2n})^k}{(\sigma\sqrt{2n} + C''\sqrt{K\sigma} n^{\frac{1}{4}} \ln n)^k} \\
=&4n\left(\frac{\sigma\sqrt{2n}}{\sigma\sqrt{2n} + C''\sqrt{K\sigma} n^{\frac{1}{4}} \ln n}\right)^k \\
=&4n\left(1 - \frac{C''\sqrt{K\sigma} n^{\frac{1}{4}} \ln n}{\sigma\sqrt{2n} + C''\sqrt{K\sigma} n^{\frac{1}{4}} \ln n}\right)^k \\
<&4ne^{-\frac{C''\sqrt{K\sigma} k n^{\frac{1}{4}} \ln n}{\sigma\sqrt{2n} + C''\sqrt{K\sigma} n^{\frac{1}{4}} \ln n}} \\
=&4ne^{-\frac{C''\sqrt{K\sigma} k n^{-\frac{1}{4}} \ln n}{\sigma\sqrt{2} + C''\sqrt{K\sigma} n^{-\frac{1}{4}} \ln n}}.
\end{aligned}
$$

Now let k be an even integer closest to (and at most) $\left(\frac{\sigma^2 n}{64 C_0 K^2}\right)^{\frac{1}{4}}$. By the assumption that $\sigma \ge C' K n^{-\frac{1}{2}} \ln^2 n$, we get

$$\Pr(\|\widehat{M_n}\| \ge \sigma\sqrt{2n} + C''\sqrt{K\sigma} n^{\frac{1}{4}} \ln n)$$

$$
\begin{aligned}
&\le 4ne^{-(1+o(1))\frac{C''\sqrt{K\sigma}}{\sigma\sqrt{2}}kn^{-\frac{1}{4}}\ln n}\\
&\le 4ne^{-(1+o(1))\frac{C''\sqrt{K\sigma}}{\sigma\sqrt{2}}(\frac{\sigma^2 n}{64C_0K^2})^{\frac{1}{4}}n^{-\frac{1}{4}}\ln n}\\
&=4ne^{-(1+o(1))\frac{C''}{4C_0^{1/4}}\ln n}\\
&=4n^{1-(1+o(1))\frac{C''}{4C_0^{1/4}}}\\
&=o(1)
\end{aligned}
$$

for sufficiently large C''.

Recall that $\widehat{M_n} = \frac{1}{\sqrt{2}}(U_n - p^2 I_n)$. Then asymptotically almost surely we have

$$
\begin{aligned}
\|U_n\| &\le \sqrt{2}\|\widehat{M_n}\| + p^2\\
&\le 2\sigma\sqrt{n} + \sqrt{2}C''\sqrt{K\sigma}n^{\frac{1}{4}}\ln n\\
&= 2\sigma\sqrt{n} + C\sqrt{K\sigma}n^{\frac{1}{4}}\ln n
\end{aligned}
$$

for sufficiently large $C = \sqrt{2}C''$. This completes the proof. □

Finally, to complete this section, we will provide our proof of Theorem 4.4, using Theorem 4.5 and the following min-max result due to Courant-Fischer (Theorem 4.2.11 in [74]).

Lemma 4.7 (Courant-Fischer [74]). *Let A be an $n \times n$ Hermitian matrix with eigenvalues $\lambda_1 \le \lambda_2 \le \cdots \le \lambda_n$, and let k be an integer with $1 \le k \le n$. Then*

$$
\lambda_k = \min_{w_1,w_2,\ldots,w_{n-k}\in\mathbb{C}^n} \max_{\substack{x\neq 0, x\in\mathbb{C}^n\\ x\perp w_1,w_2,\ldots,w_{n-k}}} \frac{x^*Ax}{x^*x},
$$

and

$$
\lambda_k = \max_{w_1,w_2,\ldots,w_{k-1}\in\mathbb{C}^n} \min_{\substack{x\neq 0, x\in\mathbb{C}^n\\ x\perp w_1,w_2,\ldots,w_{k-1}}} \frac{x^*Ax}{x^*x}.
$$

Proof of Theorem 4.4. Let $\mathbf{e}$ denote the all 1's vector. Suppose that $|\xi| = 1$ and $\xi \perp \mathbf{e}$. Then $J_n\xi = 0$. Since $U_n = p^2 J_n - H_n$, using Theorem 4.5, we get

$$
|H_n\xi| = |U_n\xi| \le \|U_n\| \le 2\sigma\sqrt{n} + C\sqrt{K\sigma}n^{\frac{1}{4}}\ln n.
$$

Now let $|x| = 1$ and $x = \alpha y + \beta z$, where $y = \frac{1}{\sqrt{n}}\mathbf{e}$, $z \perp \mathbf{e}$ and $|z| = 1$ and $\alpha^2 + \beta^2 = 1$. Then $J_n z = 0$, and

$$|H_n x| \leq |\alpha||H_n y| + |\beta||H_n z|.$$

Since $H_n = p^2 J_n - U_n$, we have

$$\begin{aligned}
|H_n y| &= \frac{1}{\sqrt{n}}|H_n \mathbf{e}| \\
&\leq \frac{1}{\sqrt{n}}(np^2|\mathbf{e}| + \|U_n\||\mathbf{e}|) \\
&= np^2 + \|U_n\| \\
&\leq np^2 + 2\sigma\sqrt{n} + C\sqrt{K\sigma}n^{\frac{1}{4}}\ln n,
\end{aligned}$$

and

$$\begin{aligned}
|H_n z| &= |(p^2 J_n - U_n)z| \\
&= |U_n z| \\
&\leq \|U_n\||z| \\
&= \|U_n\| \\
&\leq 2\sigma\sqrt{n} + C\sqrt{K\sigma}n^{\frac{1}{4}}\ln n.
\end{aligned}$$

Since $0 \leq \alpha, \beta \leq 1$ and $\alpha\beta \leq 1/2$ (because $\alpha^2 + \beta^2 = 1$), we have $|\alpha| + |\beta| \leq \sqrt{(|\alpha| + |\beta|)^2} \leq \sqrt{2}$. Thus

$$\begin{aligned}
|H_n x| \leq& |\alpha|\left[np^2 + 2\sigma\sqrt{n} + C\sqrt{K\sigma}n^{\frac{1}{4}}\ln n\right] + |\beta|\left[2\sigma\sqrt{n} + C\sqrt{K\sigma}n^{\frac{1}{4}}\ln n\right] \\
=& |\alpha|np^2 + (|\alpha| + |\beta|)\left[2\sigma\sqrt{n} + C\sqrt{K\sigma}n^{\frac{1}{4}}\ln n\right] \\
\leq& np^2 + \sqrt{2}\left[2\sigma\sqrt{n} + C\sqrt{K\sigma}n^{\frac{1}{4}}\ln n\right].
\end{aligned}$$

This implies that $\lambda_1 \leq (1 + o(1))np^2$. However,

$$\begin{aligned}
|H_n y| &\geq |(H_n + U_n)y| - |U_n y| \\
&= |p^2 J_n y| - |U_n y|
\end{aligned}$$

$$\geq np^2 - (2\sigma\sqrt{n} + C\sqrt{K\sigma}n^{\frac{1}{4}}\ln n),$$

implying $\lambda_1 \geq (1+o(1))np^2$. Hence $\lambda_1 = (1+o(1))np^2$, which completes the proof of (i).

By Lemma 4.7, we have

$$\begin{aligned}
\lambda_2 &= \min_{\eta\in\mathbb{C}^n}\max_{0\neq\xi\perp\eta}\frac{\xi^*H_n\xi}{\xi^*\xi}\\
&\leq \max_{0\neq\xi\perp y}\frac{\xi^*H_n\xi}{\xi^*\xi}\\
&= \max_{0\neq\xi\perp y}\frac{\xi^*(p^2J_n - U_n)\xi}{\xi^*\xi}\\
&= \max_{0\neq\xi\perp y}\frac{\xi^*(-U_n)\xi}{\xi^*\xi}\\
&\leq \|U_n\|\\
&\leq 2\sigma\sqrt{n} + C\sqrt{K\sigma}n^{\frac{1}{4}}\ln n,
\end{aligned}$$

and

$$\begin{aligned}
\lambda_n &= \min_{|\xi|=1}\xi^*H_n\xi\\
&\geq \min_{|\xi|=1}(\xi^*H_n\xi - p^2\xi^*J_n\xi)\\
&= \min_{|\xi|=1}(-\xi^*U_n\xi)\\
&= -\max_{|\xi|=1}\xi^*U_n\xi\\
&= -\lambda_1(U_n)\\
&\geq -\|U_n\|\\
&\geq -(2\sigma\sqrt{n} + C\sqrt{K\sigma}n^{\frac{1}{4}}\ln n),
\end{aligned}$$

where we use that $\|U_n\| = \max\{|\lambda_1(U_n)|, |\lambda_n(U_n)|\}$ and $U_n = p^2J_n - H_n$, and that $\lambda_1(U_n) > 0$ and $\|U_n\| \geq \lambda_1(U_n)$. So

$$|\lambda_i| \leq 2\sqrt{n(2p - p^2 - p^4)} + C\sqrt{K\sigma}n^{\frac{1}{4}}\ln n$$

for $2 \le i \le n$. This completes the proof of (ii).

4.3 Spectral moments of random mixed graphs

In this section, as an application of Theorem 4.4, we give an estimation of the Hermitian spectral moment for random mixed graphs. The result is stated as follows.

Theorem 4.8. *Let $p = p(n)$, $0 < p < 1$. Let H_n denote the Hermitian adjacency matrix of $\widehat{G}_n(p)$. Then almost surely*

$$s_2(H_n) = (2p - p^2 + o(1))n^2.$$

Suppose C' is sufficiently large. Define $K = \sqrt{\frac{1+p^4}{2}}$ and $\sigma = \sqrt{2p - p^2 - p^4}$. If $\sigma \ge C'Kn^{-\frac{1}{2}} \ln^2 n$, then for $k > 2$, the k-th Hermitian spectral moment $s_k(H_n)$ of the random mixed graph $\widehat{G}_n(p)$ almost surely satisfies the following equation:

$$s_k(H_n) = (p^{2k} + o(1))n^k.$$

Proof of Theorem 4.8. Let $\widehat{G}_n(p)$ and $H_n = (h_{ij})_{n\times n}$ be defined as above. For $k = 2$,

$$s_2(H_n) = \mathrm{Tr}(H_n^2) = \sum_{i\ne j} h_{ij}h_{ji} = \sum_{i\ne j} |h_{ij}|^2 = 2 \sum_{1\le i<j\le n} |h_{ij}|^2.$$

Since $|h_{ij}|^2 (i > j)$ are i.i.d. with mean $2p - p^2$, it follows from Lemma 2.2 that, with probability 1,

$$\lim_{n\to\infty} \frac{\sum_{i=1}^{n} \sum_{i>j} |h_{ij}|^2}{\frac{n(n-1)}{2}} = 2p - p^2.$$

Then

$$\lim_{n\to\infty} \frac{s_2(H_n)}{n^2} = \lim_{n\to\infty} \frac{2\sum_{i=1}^{n} \sum_{i>j} |h_{ij}|^2}{n^2} = 2p - p^2,$$

i.e.,

$$s_2(H_n) = (2p - p^2 + o(1))n^2 \quad a.s.$$

Suppose that $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$ are the eigenvalues of H_n. By Theorem 4.4, we have

$$\lim_{n\to\infty} \frac{\lambda_1}{n} = p^2 \quad a.s., \tag{4.6}$$

and for any $\epsilon > \frac{1}{2}$,

$$\lim_{n\to\infty} \frac{\lambda_i}{n^\epsilon} = 0 \quad a.s., \quad i = 2, \ldots, n. \tag{4.7}$$

With the definition of the k-th Hermitian spectral moment, one can deduce that for $k > 2$,

$$(\frac{\lambda_1}{n})^k + (n-1) \min_{2\leq i\leq n} \{(\frac{\lambda_i}{n})^k\} \leq \frac{s_k(H_n)}{n^k} \leq (\frac{\lambda_1}{n})^k + (n-1) \max_{2\leq i\leq n} \{(\frac{\lambda_i}{n})^k\},$$

or, equivalently

$$(\frac{\lambda_1}{n})^k + \frac{n-1}{n} \min_{2\leq i\leq n} \{(\frac{\lambda_i}{n^{1-\frac{1}{k}}})^k\} \leq \frac{s_k(H_n)}{n^k} \leq (\frac{\lambda_1}{n})^k + \frac{n-1}{n} \max_{2\leq i\leq n} \{(\frac{\lambda_i}{n^{1-\frac{1}{k}}})^k\}.$$

Since $1 - \frac{1}{k} > \frac{1}{2}$, using (4.6) and (4.7), we obtain that

$$\lim_{n\to\infty} \frac{s_k(H_n)}{n^k} = p^{2k},$$

i.e.,

$$s_k(H_n) = (p^{2k} + o(1))n^k \quad a.s.$$

This completes the proof. □

Chapter 5

The spectral analysis of $\mathscr{L}_n$ for random mixed graphs

In this chapter, we again consider random mixed graphs $\widehat{G}_n(p)$ as described in the introduction of Chapter 3. Let $\widehat{G}_n(p)$ be a random mixed graph on the vertex set $\{1, 2, \ldots, n\}$. We study the spectral properties of the normalized Hermitian Laplacian matrix of $\widehat{G}_n(p)$ for large n, and characterize the limiting spectral distribution in case $p \in (0,1)$ and $n(2p - p^2 - p^4)/\ln^4 n \to \infty$. Our main result is stated as follows.

Theorem 5.1. *Let $\{\mathscr{L}_n\}_{n=1}^{\infty}$ be a sequence of normalized Hermitian Laplacian matrices of random mixed graphs $\{\widehat{G}_n(p)\}_{n=1}^{\infty}$ with $p = p(n)$, $0 < p < 1$. Let $\sigma = \sqrt{2p - p^2 - p^4}$, and $\delta = (n-1)(2p - p^2)$. If $n\sigma^2/\ln^4 n \to \infty$ as $n \to \infty$, then the ESD of $\frac{\delta}{\sigma\sqrt{n}}(I_n - \mathscr{L}_n)$ converges to the standard semicircle distribution whose density is given by*

$$\phi(x) := \begin{cases} \frac{1}{2\pi}\sqrt{4 - x^2}, & \text{for } |x| \le 2, \\ 0, & \text{for } |x| > 2. \end{cases}$$

We are going to present our proof of the above theorem in the final section of this chapter. We first concentrate on the spectral properties of several different matrices related to the normalized Hermitian Laplacian matrix $\mathscr{L}_n$.

5.1 The spectral properties of random matrices

The normalized Hermitian Laplacian matrix of $\widehat{G}_n(p)$ is defined as

$$\mathscr{L}_n = I_n - D_n^{-\frac{1}{2}} H_n D_n^{-\frac{1}{2}},$$

where I_n is the identity matrix, H_n is the Hermitian adjacency matrix of $\widehat{G}_n(p)$, and D_n denotes the diagonal degree matrix of the underlying graph $\Gamma(\widehat{G}_n(p))$. We can rewrite $\mathscr{L}_n$ as

$$\mathscr{L}_n = I_n - [D_n^{-\frac{1}{2}} H_n D_n^{-\frac{1}{2}} - D_n^{-\frac{1}{2}} \mathbb{E}H_n D_n^{-\frac{1}{2}}] - D_n^{-\frac{1}{2}} \mathbb{E}H_n D_n^{-\frac{1}{2}}.$$

Now, we let

$$C_n = D_n^{-\frac{1}{2}} H_n D^{-\frac{1}{2}} - D_n^{-\frac{1}{2}} \mathbb{E}H_n D_n^{-\frac{1}{2}}.$$

Instead of directly dealing with C_n, we first consider the related matrix

$$R_n = (\mathbb{E}D_n)^{-\frac{1}{2}} H_n (\mathbb{E}D_n)^{-\frac{1}{2}} - (\mathbb{E}D_n)^{-\frac{1}{2}} \mathbb{E}H_n (\mathbb{E}D_n)^{-\frac{1}{2}}.$$

Similar to our proof of Theorem 4.5 in Chapter 4, we can derive the next theorem by using Lemma 4.3. We will use the conclusion of this result near the end of the proof of Theorem 5.1.

Theorem 5.2. *Let $\sigma = \sqrt{2p - p^2 - p^4}$ and $R_n = (\mathbb{E}D_n)^{-\frac{1}{2}}(H_n - \mathbb{E}H_n)(\mathbb{E}D_n)^{-\frac{1}{2}}$. Assume that $\delta = (n-1)(2p-p^2)$. If $n\sigma^2/\ln^4 n \to \infty$ as $n \to \infty$, then we have*

$$\|R_n\| \le (1 + o(1))\frac{2\sigma}{\delta}\sqrt{n}.$$

Proof of Theorem 5.2. We rely on Wigner's high moment method. Recall that $H_n = (h_{ij})_{n\times n}$, where h_{ij} $(1 \le i < j \le n)$ are independent random variables with the following properties:

- $\mathbb{E}(h_{ij}) = p^2$;
- $\mathrm{Var}(h_{ij}) = 2p - p^2 - p^4 = \sigma^2 < 2$;
- $h_{ij}, h_{i'j'}$ are independent, unless $(i,j) = (j',i')$. If $i > j$, we have $\overline{h_{ji}} = h_{ij}$, i.e., h_{ji} is the complex conjugate of h_{ij};

- $|h_{ij}| \le 1$.

Let $\mathbb{E}D_n = \operatorname{diag}(t_1, t_2, \ldots, t_n)$, where $t_i = \mathbb{E}(d_i) = \sum_{j=1}^{n} \mathbb{E}|h_{ij}| = (n-1)(2p-p^2) = \delta$ for all $1 \le i \le n$. Let r_{ij} denote the (i,j)-th entry of R_n. Then r_{ij} $(1 \le i < j \le n)$ are independent random variables with the following properties:

- $\mathbb{E}(r_{ij}) = 0$;
- $\operatorname{Var}(r_{ij}) = \frac{1}{t_i t_j}\operatorname{Var}(h_{ij}) = \frac{1}{t_i t_j}(2p - p^2 - p^4) = \frac{\sigma^2}{\delta^2} < \frac{2}{\delta^2} \le 1$;
- $r_{ij}, r_{i'j'}$ are independent, unless $(i,j) = (j',i')$. If $i > j$, we have $\overline{r_{ji}} = r_{ij}$;
- $|r_{ij}| \le \sqrt{\frac{1+p^4}{t_i t_j}} \le \frac{\sqrt{2}}{\delta} \le 1$.

Now let $k \ge 2$ be an even integer. We estimate

$$\begin{aligned}
\operatorname{Tr}(R_n^k) &= \sum_{i=1}^{n} \lambda_i(R_n)^k \\
&\ge \max\{\lambda_1(R_n)^k, \lambda_n(R_n)^k\} \\
&= \|R_n\|^k.
\end{aligned}$$

A standard fact in linear algebra tells us that for any positive integer k,

$$\operatorname{Tr}(R_n^k) = \sum_{i_1,\ldots,i_k \in [n]} r_{i_1 i_2} r_{i_2 i_3} \cdots r_{i_k i_1}, \tag{5.1}$$

where $[n] = \{1, 2, \ldots, n\}$.

Let us now take a closer look at $\operatorname{Tr}(R_n^k)$. This is a sum where a typical term is $r_{i_1 i_2} r_{i_2 i_3} \ldots r_{i_{k-1} i_k} r_{i_k i_1}$, where $W := i_1 i_2 \ldots i_{k-1} i_k i_1$ corresponds to a closed directed walk of length k in the complete directed graph DK_n of order n. In other words, each term corresponds to a closed walk of length k (containing k, not necessarily distinct, directed edges) of the complete directed graph DK_n on $[n]$. For each directed edge $(i,j) \in W$, let q_{ij} be the number of occurrences of the directed edge (i,j) in the walk W. Note that all directed edges of a mixed graph are mutually independent. Now, we rewrite (5.1) as

$$\mathrm{Tr}(R_n^k) = \sum_W \prod_{i<j} r_{ij}^{q_{ij}} r_{ji}^{q_{ji}}. \tag{5.2}$$

Then

$$\mathbb{E}(\mathrm{Tr}(R_n^k)) = \mathbb{E}\left(\sum_W \prod_{i<j} r_{ij}^{q_{ij}} r_{ji}^{q_{ji}}\right) = \sum_W \prod_{i<j} \mathbb{E}\left(r_{ij}^{q_{ij}} r_{ji}^{q_{ji}}\right),$$

where the summation is taken over all directed closed walks of length k.

We decompose $\mathbb{E}(\mathrm{Tr}(R_n^k))$ into parts $\mathbb{E}_{n,k,t}, t = 2, \ldots, k$, containing the t-fold sums, as follows:

$$\mathbb{E}(\mathrm{Tr}(R_n^k)) = \sum_{t=2}^{k} \mathbb{E}_{n,k,t}, \tag{5.3}$$

where

$$\mathbb{E}_{n,k,t} = \sum_{\{W:|V(W)|=t\}} \prod_{i<j} \mathbb{E}\left(r_{ij}^{q_{ij}} r_{ji}^{q_{ji}}\right), \tag{5.4}$$

and $|V(W)| = t$ means the cardinality of the vertex set of W is t. (Note that as $r_{ii} = 0$ by construction of R_n we have that $\mathbb{E}_{n,k,1} = 0$.) Here the summation in (5.4) is taken over all closed directed walks W of length k using exactly t different vertices.

Recall that the entries r_{ij} of R_n are independent random variables with zero mean, i.e., $\mathbb{E}(r_{ij}) = 0$, for all $1 \le i < j \le n$. Recall also that q_{ij} denotes the number of occurrences of the directed edge (i, j) in the closed walk W. So, if $q_{ij} + q_{ji} = 1$, that is, $q_{ij} = 1, q_{ji} = 0$ or $q_{ij} = 0, q_{ji} = 1$, then $\prod_{i<j} \mathbb{E}\left(r_{ij}^{q_{ij}} r_{ji}^{q_{ji}}\right) = 0$. Thus, the expectation of a term is nonzero if and only if the total number of occurrences of each directed edge and its inverse edge of DK_n in the directed walk W is at least 2. So, we only need to consider the case that $q_{ij} + q_{ji} \ge 2$. Note that such a closed directed walk is a good directed walk, and the set of all good closed directed walks of length k in DK_n is denoted by $\mathscr{G}(n, k)$. Considering a good closed directed walk W, the underlying graph $\Gamma(W)$ of W uses l distinct edges $e_1, \ldots, e_l$, i.e., $|E(\Gamma(W))| = l$, with corresponding multiplicities $s_1, \ldots, s_l$ (the s_hs are positive integers at least 2 summing up to k). Without loss of generality, we set $e_h = v_i v_j$ and

then $s_h = q_{ij} + q_{ji}$. The (expected) contribution of the term defined by this directed walk to $\mathbb{E}(\mathrm{Tr}(R_n^k))$ is

$$\prod_{\substack{i<j \\ |E(\Gamma(W))|=l}} \mathbb{E}\left(r_{ij}^{q_{ij}} r_{ji}^{q_{ji}}\right). \tag{5.5}$$

Next, we will compute $\mathbb{E}\left(r_{ij}^{q_{ij}} r_{ji}^{q_{ji}}\right)$. Note that $q_{ij} + q_{ji} \geq 2$ implies that $q_{ij} \geq 1, q_{ji} \geq 1$ or $q_{ij} \geq 2, q_{ji} = 0$ or $q_{ij} = 0, q_{ji} \geq 2$, since $|r_{ij}| \leq \frac{\sqrt{2}}{\delta} \leq 1$ and $\mathbb{E}(r_{ij}) = 0$. We consider these three cases separately.

If $q_{ij} \geq 1, q_{ji} \geq 1$, then we have

$$\begin{aligned}
\left|\mathbb{E}\left(r_{ij}^{q_{ij}} r_{ji}^{q_{ji}}\right)\right| &\leq \mathbb{E}\left|r_{ij}^{q_{ij}-1} \cdot r_{ji}^{q_{ji}-1} \cdot r_{ij} \cdot r_{ij}\right| \\
&\leq \left(\frac{\sqrt{2}}{\delta}\right)^{q_{ij}+q_{ji}-2} \mathbb{E}|r_{ij} r_{ji}| \\
&= \left(\frac{\sqrt{2}}{\delta}\right)^{q_{ij}+q_{ji}-2} \mathbb{E}|r_{ij}|^2 \\
&= \left(\frac{\sqrt{2}}{\delta}\right)^{q_{ij}+q_{ji}-2} \mathbb{E}(r_{ij}\overline{r_{ij}}) \\
&= \left(\frac{\sqrt{2}}{\delta}\right)^{q_{ij}+q_{ji}-2} \mathrm{Var}(r_{ij}) \\
&= \left(\frac{\sqrt{2}}{\delta}\right)^{q_{ij}+q_{ji}-2} \frac{\sigma^2}{\delta^2}.
\end{aligned} \tag{5.6}$$

If $q_{ij} \geq 2, q_{ji} = 0$, then we have

$$\begin{aligned}
\left|\mathbb{E}\left(r_{ij}^{q_{ij}} r_{ji}^{q_{ji}}\right)\right| &= \mathbb{E}\left|r_{ij}^{q_{ij}}\right| \\
&\leq \left(\frac{\sqrt{2}}{\delta}\right)^{q_{ij}-2} \mathbb{E}|r_{ij}^2| \\
&= \left(\frac{\sqrt{2}}{\delta}\right)^{q_{ij}-2} \frac{\sigma^2}{\delta^2}.
\end{aligned} \tag{5.7}$$

If $q_{ij} = 0, q_{ji} \geq 2$, then similarly, we have

$$\left|\mathbb{E}\left(r_{ij}^{q_{ij}} r_{ji}^{q_{ji}}\right)\right| \leq \left(\frac{\sqrt{2}}{\delta}\right)^{q_{ji}-2} \frac{\sigma^2}{\delta^2}. \tag{5.8}$$

Recall that $\mathscr{G}(n,k,t)$ denotes the set of good closed directed walks on DK_n of length k using exactly t different vertices. Notice that for each directed walk W in $\mathscr{G}(n,k,l+1)$, the underlying graph $\Gamma(W)$ of W must have at least l different edges. By (5.5)-(5.8), the contribution of a term corresponding to such a good directed walk to $\mathbb{E}(\mathrm{Tr}(R_n^k))$ is at most

$$\left(\frac{\sqrt{2}}{\delta}\right)^{k-2l} \frac{\sigma^{2l}}{\delta^{2l}} = \frac{\sqrt{2}^{k-2l}\sigma^{2l}}{\delta^k}.$$

By the pigeon hole principle, if $l+1 > \frac{k}{2}+1$, then there must be a directed edge (i,j) such that the total number of occurrences of this directed edge and its inverse edge of DK_n in the directed walk W is 1, i.e., $q_{ij}+q_{ji} = 1$. As we argued before, this implies $\mathbb{E}_{n,k,l+1} = 0$ for $l > \frac{k}{2}$.

So, in the following, we only consider the case that $l \leq \frac{k}{2}$ and $q_{ij}+q_{ji} \geq 2$. By Lemma 4.3, we have

$$\begin{aligned}
&\mathbb{E}(\mathrm{Tr}(R_n^k)) \\
&\leq \sum_{l=1}^{\frac{k}{2}} |\mathscr{G}(n,k,l+1)| \frac{\sqrt{2}^{k-2l}\sigma^{2l}}{\delta^k} \\
&= \sum_{m=2}^{\frac{k}{2}+1} |\mathscr{G}(n,k,m)| \frac{\sqrt{2}^{k-2(m-1)}\sigma^{2(m-1)}}{\delta^k} \\
&\leq \sum_{m=2}^{\frac{k}{2}+1} \frac{\sqrt{2}^{k-2(m-1)}\sigma^{2(m-1)}}{\delta^k} n^m \binom{k}{2m-2} 2^{2k-2m+3} m^{k-2m+2}(k-2m+4)^{k-2m+2} \\
&= \sum_{m=2}^{\frac{k}{2}+1} S(n,k,m),
\end{aligned} \tag{5.9}$$

where the final equality defines $S(n,k,m)$. Now fix $k = g(n)\ln n$, where

$g(n)$ tends to infinity (with n) arbitrarily slowly. Let us consider the ratio $S(n,k,m-1)/S(n,k,m)$ for some $m \leq \frac{k}{2}+1$:

$$
\begin{aligned}
&\frac{S(n,k,m-1)}{S(n,k,m)} \\
=&\frac{\frac{\sqrt{2}^{k-2(m-2)}\sigma^{2(m-2)}}{\delta^k} n^{m-1}\binom{k}{2m-4} 2^{2k-2m+5}(m-1)^{k-2m+4}(k-2m+6)^{k-2m+4}}{\frac{\sqrt{2}^{k-2(m-1)}\sigma^{2(m-1)}}{\delta^k} n^{m}\binom{k}{2m-2} 2^{2k-2m+3}m^{k-2m+2}(k-2m+4)^{k-2m+2}} \\
=&\frac{2(2m-2)(2m-3)2^2(m-1)^{k-2m+4}(k-2m+6)^{k-2m+4}}{\sigma^2 n(k-2m+4)(k-2m+3)m^{k-2m+2}(k-2m+4)^{k-2m+2}} \\
\leq&\frac{32m^{k-2m+6}(k-2m+6)^{k-2m+4}}{\sigma^2 n(k-2m+4)(k-2m+3)m^{k-2m+2}(k-2m+4)^{k-2m+2}} \\
=&\frac{32m^4(k-2m+6)^{k-2m+4}}{\sigma^2 n(k-2m+3)(k-2m+4)^{k-2m+3}} \\
\leq&\frac{32m^4(k-2m+6)^{k-2m+4}}{\sigma^2 n(k-2m+3)^{k-2m+4}} \\
\to&\frac{32C_0 m^4}{\sigma^2 n} \\
\leq&\frac{32C_0 k^4}{\sigma^2 n}
\end{aligned}
$$

for some constant C_0 independent of σ. This implies that

$$S(n,k,m-1) \leq \frac{32C_0k^4}{\sigma^2 n} S(n,k,m).$$

With a proper choice of $g(n)$ guaranteeing that $k^4 \leq \frac{\sigma^2 n}{64C_0}$, we have

$$S(n,k,m-1) \leq \frac{1}{2} S(n,k,m).$$

Then

$$\mathbb{E}(\mathrm{Tr}(R_n^k)) \leq \sum_{m=2}^{\frac{k}{2}+1} S(n,k,m)$$

$$
\begin{aligned}
&= S\left(n, k, \frac{k}{2}+1\right) \sum_{m=2}^{\frac{k}{2}+1} \left(\frac{1}{2}\right)^{\frac{k}{2}+1-m} \\
&\le 2S\left(n, k, \frac{k}{2}+1\right) \\
&= 2\frac{\sigma^k}{\delta^k} n^{\frac{k}{2}+1} 2^{k+1} \\
&= 4n\left(\frac{2\sigma}{\delta}\sqrt{n}\right)^k.
\end{aligned}
$$

Then

$$
\begin{aligned}
\mathbb{E}(\|R_n^k\|) &\le \mathbb{E}(\mathrm{Tr}(R_n^k)) \\
&\le 4n\left(\frac{2\sigma}{\delta}\sqrt{n}\right)^k.
\end{aligned}
$$

Using Markov's inequality we get

$$
\begin{aligned}
\Pr\left(\|R_n\| \ge (1+\epsilon)\frac{2\sigma}{\delta}\sqrt{n}\right) &= \Pr\left(\|R_n\|^k \ge \left((1+\epsilon)\frac{2\sigma}{\delta}\sqrt{n}\right)^k\right) \\
&\le \frac{\mathbb{E}(\|R_n^k\|)}{((1+\epsilon)\frac{2\sigma}{\delta}\sqrt{n})^k} \\
&\le \frac{4n(\frac{2\sigma}{\delta}\sqrt{n})^k}{((1+\epsilon)\frac{2\sigma}{\delta}\sqrt{n})^k} \\
&= \frac{4n}{(1+\epsilon)^k}.
\end{aligned}
$$

Since $k = \Omega(\ln n)$, we can find an $\epsilon = \epsilon(n)$ tending to 0 when n tends to infinity, so that $\frac{n}{(1+\epsilon)^k} = o(1)$. Thus, we get

$$
\|R_n\| \le (1+o(1))\frac{2\sigma}{\delta}\sqrt{n}.
$$

This completes the proof. □

5.2 The LSD of $\mathscr{L}_n$

In this section we characterize the LSD of the normalized Hermitian Laplacian matrix $\mathscr{L}_n$ by proving Theorem 5.1. For our proof, we will rely on the following known result. Here, let $L(F,G)$ denote the Levy distance between distribution functions F and G, defined by

$$L(F,G) = \inf\{\epsilon \mid F(x-\epsilon) \le G(x) \le F(x+\epsilon)\},$$

which characterizes the weak convergence of probability distributions. Then the following holds for the ESD of Hermitian matrices.

Lemma 5.3 (Norm Inequality (See [11])). *Let A and B be two $n \times n$ Hermitian matrices. Then*

$$L(F^A, F^B) \le \|A - B\|,$$

where F^A denotes the ESD of A.

We also use the following concentration result. It involves a variation on the Chernoff bound, and can, e.g., be found as Lemma A in [32].

Lemma 5.4. *Let $X_1, X_2, \ldots, X_m$ be independent random variables satisfying $|X_i| \le c$ for all i. Let $X = \sum_{i=1}^m X_i$. Then for any $a > 0$,*

$$\Pr(|X - \mathbb{E}(X)| \ge a) \le \exp\Big(-\frac{a^2}{2\sum_{i=1}^m \mathrm{Var}(X_i) + 2ac/3}\Big).$$

We now have all the ingredients to present our proof of Theorem 5.1.

Proof of Theorem 5.1. Recall that

$$\begin{aligned} R_n &= (\mathbb{E}D_n)^{-\frac{1}{2}} H_n (\mathbb{E}D_n)^{-\frac{1}{2}} - (\mathbb{E}D_n)^{-\frac{1}{2}} \mathbb{E}H_n (\mathbb{E}D_n)^{-\frac{1}{2}} \\ &= \frac{1}{\delta}[H_n - p^2(J_n - I_n)], \end{aligned}$$

where J_n is the all 1's matrix. Set

$$M_n = \frac{1}{\sigma}[H_n - p^2(J_n - I_n)].$$

It is clear that

$$\frac{\delta}{\sigma}\lambda_i(R_n) = \lambda_i(M_n)$$

for $i = 1, 2, \ldots, n$. Thus by Theorem 3.9, we have that almost surely, the empirical distribution $F^{\frac{\delta}{\sigma\sqrt{n}}R_n}(x)$ of $\frac{\delta}{\sigma\sqrt{n}}R_n$ converges to the standard semicircle distribution $F(x)$ with density $\phi(x)$ as $n \to \infty$. Recall that

$$C_n = D_n^{-\frac{1}{2}} H_n D^{-\frac{1}{2}} - D_n^{-\frac{1}{2}} \mathbb{E} H_n D_n^{-\frac{1}{2}}.$$

We rewrite C_n as follows:

$$C_n = R_n + B_n,$$

where

$$B_n = D_n^{-\frac{1}{2}}(H_n - \mathbb{E}H_n)D_n^{-\frac{1}{2}} - (\mathbb{E}D_n)^{-\frac{1}{2}}(H_n - \mathbb{E}H_n)(\mathbb{E}D_n)^{-\frac{1}{2}},$$

and

$$R_n = (\mathbb{E}D_n)^{-\frac{1}{2}}(H_n - \mathbb{E}H_n)(\mathbb{E}D_n)^{-\frac{1}{2}}.$$

Let b_{ij} denote the (i, j)-th entry of B_n, and let r_{ij} denote the (i, j)-th entry of R_n. To bound $\|B_n\|$, we have that almost surely

$$\begin{aligned}
\|B_n\| &= \sup_{|x|=1} |x^* B_n x| \\
&= \sup_{|x|=1} \left| \sum_{i,j} x_i^* b_{ij} x_j \right| \\
&= \sup_{|x|=1} \left| \sum_{i,j} x_i^* r_{ij} x_j \frac{\sqrt{t_i t_j} - \sqrt{d_i d_j}}{\sqrt{d_i d_j}} \right| \\
&\le \sup_{|x|=1} \left(\left| \sum_{i,j} x_i^* r_{ij} x_j \frac{\sqrt{t_j} - \sqrt{d_j}}{\sqrt{d_j}} \right| + \left| \sum_{i,j} x_i^* \frac{\sqrt{t_i} - \sqrt{d_i}}{\sqrt{d_i}} r_{ij} x_j \frac{\sqrt{t_j}}{\sqrt{d_j}} \right| \right) \\
&=: \sup_{|x|=1} (|x^* R_n y| + |y^* R_n z|),
\end{aligned}$$

where $y = \left(x_1 \frac{\sqrt{t_1} - \sqrt{d_1}}{\sqrt{d_1}}, \ldots, x_n \frac{\sqrt{t_n} - \sqrt{d_n}}{\sqrt{d_n}}\right)^T$, and $z = \left(x_1 \frac{\sqrt{t_1}}{\sqrt{d_1}}, \ldots, x_n \frac{\sqrt{t_n}}{\sqrt{d_n}}\right)^T$.

Then we have

$$\begin{aligned}\|B_n\| \leq & \sup_{|x|=1} (\|R_n\||y| + \|R_n\||y||z|) \\ =& \|R_n\| \sup_{|x|=1} (|y| + |y||z|),\end{aligned}$$

where $|y|^2 = \sum_{i=1}^{n} |x_i|^2 \left(\frac{\sqrt{t_i} - \sqrt{d_i}}{\sqrt{d_i}}\right)^2, |z|^2 = \sum_{i=1}^{n} |x_i|^2 \frac{t_i}{d_i}$.

Next, we are going to obtain upper bounds for $|y|$ and $|z|$. For this, we will apply Lemma 5.4 to the random variables $|h_{ij}|$ (in the role of X_i), and using the observations that $d_i = \sum_{j=1}^{n} |h_{ij}|$, and $t_i = \mathbb{E}(d_i) = \sum_{j=1}^{n} \mathbb{E}|h_{ij}| = (n-1)(2p-p^2) = \delta$. We first need some preparation in order to obtain upper bounds for $|h_{ij}|$ and $\sum_{j=1}^{n} \mathrm{Var}(|h_{ij}|)$.

Obviously, $|h_{ij}| \leq 1$, so we can take $c = 1$, and

$$\begin{aligned}\sum_{j=1}^{n} \mathrm{Var}(|h_{ij}|) =& \sum_{j=1}^{n} [\mathbb{E}(|h_{ij}|^2) - (\mathbb{E}|h_{ij}|)^2] \\ \leq & \sum_{j=1}^{n} \mathbb{E}(|h_{ij}|^2) \\ =& (n-1)(2p - p^2) \\ =& \delta.\end{aligned}$$

We choose $a = 3\sqrt{t_i \ln n}$. Then, the assumption that $\delta = t_i = (n-1)(2p - p^2) \gg \ln n$ implies that $a = 3\sqrt{t_i \ln n} \leq 3t_i$. Applying Lemma 5.4, we have for all i,

$$\Pr(|d_i - t_i| \geq a) \leq e^{-\frac{a^2}{2(t_i + a/3)}} \leq \frac{1}{n^{9/4}}.$$

Thus asymptotically almost surely, for all i we have $|d_i - t_i| \leq a = 3\sqrt{t_i \ln n}$.

Note that

$$\Pr\left(\max_{1\leq i\leq n} \frac{t_i}{d_i} > (1+\epsilon)\right) \leq n \cdot \max_{1\leq i\leq n} \Pr\left(\frac{t_i}{d_i} > (1+\epsilon)\right).$$

This inequality holds since $\Pr(\bigcup_i A_i) \leq \sum_i \Pr(A_i)$. Choose $0 < b = 3\sqrt{\frac{\ln n}{t_i}} <$

1 such that $\frac{1}{1-b} < 1+\epsilon$. Then

$$
\begin{aligned}
&\Pr\left(\frac{t_i}{d_i} > (1+\epsilon)\right)\\
=&\Pr\left(\{\frac{t_i}{d_i} > (1+\epsilon)\}\middle|\{|d_i - t_i| < bt_i\}\right)\Pr\left(|d_i - t_i| < bt_i\right)\\
&+\Pr\left(\{\frac{t_i}{d_i} > (1+\epsilon)\}\middle|\{|d_i - t_i| \geq bt_i\}\right)\Pr\left(|d_i - t_i| \geq bt_i\right)\\
=&\Pr\left(\{\frac{t_i}{d_i} > (1+\epsilon)\}\cap\{|d_i - t_i| < bt_i\}\right)\\
&+\Pr\left(\{\frac{t_i}{d_i} > (1+\epsilon)\}\cap\{|d_i - t_i| \geq bt_i\}\right)\\
\leq&\Pr\left(\{\frac{t_i}{d_i} > (1+\epsilon)\}\cap\{|d_i - t_i| < bt_i\}\right)+\Pr\left(|d_i - t_i| \geq bt_i\right)\\
=&0+\Pr\left(|d_i - t_i| \geq bt_i\right)\\
\leq&e^{-\frac{(bt_i)^2}{2(t_i+bt_i/3)}}\\
\leq&e^{-\frac{(bt_i)^2}{2(t_i+t_i)}}\\
=&e^{-\frac{b^2t_i}{4}}\\
=&e^{-\frac{9\ln n}{4}}\\
=&\frac{1}{n^{9/4}}.
\end{aligned}
$$

So

$$\Pr\left(\max_{1\leq i\leq n}\frac{t_i}{d_i} > (1+\epsilon)\right) \leq \frac{1}{n^{5/4}}.$$

Then we have

$$\sum_{n=1}^{\infty}\Pr\left(\max_{1\leq i\leq n}\frac{t_i}{d_i} > (1+\epsilon)\right) < \infty.$$

By Lemma 3.5, we have

$$\Pr\left(\limsup_{n\to\infty}\max_{1\leq i\leq n}\frac{t_i}{d_i} > (1+\epsilon)\right) = 0.$$

i.e.,

$$\limsup_{n\to\infty} \frac{t_i}{d_i} \le 1.$$

Then we have

$$|z| = \left(\sum_{i=1}^{n} |x_i|^2 \frac{t_i}{d_i}\right)^{\frac{1}{2}} \le \max_{1\le i\le n} \left(\frac{t_i}{d_i}\right)^{\frac{1}{2}} \cdot \left(\sum_{i=1}^{n} |x_i|^2\right)^{\frac{1}{2}} = \max_{1\le i\le n} \left(\frac{t_i}{d_i}\right)^{\frac{1}{2}} \le 1,$$

and

$$\begin{aligned}
|y| =& \left(\sum_{i=1}^{n} |x_i|^2 \left(\frac{\sqrt{t_i}-\sqrt{d_i}}{\sqrt{d_i}}\right)^2\right)^{\frac{1}{2}} \\
\le& \max_{1\le i\le n} \left(\frac{(\sqrt{t_i}-\sqrt{d_i})^2}{d_i}\right)^{\frac{1}{2}} \cdot \left(\sum_{i=1}^{n} |x_i|^2\right)^{\frac{1}{2}} \\
=& \max_{1\le i\le n} \left(\frac{(t_i-d_i)^2}{d_i(\sqrt{t_i}+\sqrt{d_i})^2}\right)^{\frac{1}{2}} \\
\le& \max_{1\le i\le n} \left(\frac{(t_i-d_i)^2}{t_i(\sqrt{t_i}+\sqrt{t_i})^2}\right)^{\frac{1}{2}} \\
=& \max_{1\le i\le n} \left(\frac{(t_i-d_i)^2}{4t_i^2}\right)^{\frac{1}{2}} \\
=& \max_{1\le i\le n} \frac{|t_i-d_i|}{2t_i} \\
\le& \max_{1\le i\le n} \frac{3\sqrt{t_i \ln n}}{2t_i} \\
=& \max_{1\le i\le n} \frac{3}{2}\sqrt{\frac{\ln n}{t_i}} \\
=& o(1),
\end{aligned}$$

where the final equality holds since $\frac{t_i}{\ln n} = (n-1)(2p-p^2)/\ln n \to \infty$. Hence, using Theorem 5.2, we obtain

$$\|B_n\| \le \|R_n\| \sup_{|x|=1} (|y| + |y||z|)$$

$$\begin{aligned}&\leq o(\|R_n\|)\\&\leq o\Big((1+o(1))\frac{2\sigma}{\delta}\sqrt{n}\Big).\end{aligned}$$

Recall that almost surely $F^{\frac{\delta}{\sigma\sqrt{n}}R_n}(x)$ converges to the standard semicircle distribution $F(x)$ with density $\phi(x)$ as $n\to\infty$. Recall that

$$C_n = R_n + B_n.$$

Then by Lemma 5.3, we have

$$L\Big(F^{\frac{\delta}{\sigma\sqrt{n}}C_n}, F^{\frac{\delta}{\sigma\sqrt{n}}R_n}\Big) \leq \frac{\delta}{\sigma\sqrt{n}}\|B_n\| \leq \frac{\delta}{\sigma\sqrt{n}}o\Big((1+o(1))\frac{2\sigma}{\delta}\sqrt{n}\Big) \to 0.$$

This implies that the LSDs of $\frac{\delta}{\sigma\sqrt{n}}C_n$ and $\frac{\delta}{\sigma\sqrt{n}}R_n$ are the same. Thus, by Theorem 3.9, almost surely, $F^{\frac{\delta}{\sigma\sqrt{n}}C_n}(x)$ converges weakly to the standard semicircle distribution $F(x)$ with density $\phi(x)$ as $n\to\infty$. Recall that

$$\begin{aligned}\mathscr{L}_n =& I_n - C_n - D_n^{-\frac{1}{2}}\mathbb{E}H_n D_n^{-\frac{1}{2}}\\=& I_n - C_n - D_n^{-\frac{1}{2}}p^2(J_n - I_n)D_n^{-\frac{1}{2}}.\end{aligned}$$

By Lemma 3.6, we have

$$\begin{aligned}&\Big\|F^{\frac{\delta}{\sigma\sqrt{n}}(I_n-\mathscr{L}_n+D_n^{-\frac{1}{2}}p^2I_nD_n^{-\frac{1}{2}})} - F^{\frac{\delta}{\sigma\sqrt{n}}C_n}\Big\|\\=&\Big\|F^{\frac{\delta}{\sigma\sqrt{n}}(I_n-\mathscr{L}_n+p^2D_n^{-1})} - F^{\frac{\delta}{\sigma\sqrt{n}}C_n}\Big\|\\\leq&\frac{1}{n}\mathrm{rank}\Big(\frac{\delta}{\sigma\sqrt{n}}D_n^{-\frac{1}{2}}p^2J_nD_n^{-\frac{1}{2}}\Big)\\\leq&\frac{1}{n}\mathrm{rank}(J_n)\\=&\frac{1}{n}\\\to&0.\end{aligned}$$

This implies that the LSDs of $\frac{\delta}{\sigma\sqrt{n}}(I_n - \mathscr{L}_n + p^2D_n^{-1})$ and $\frac{\delta}{\sigma\sqrt{n}}C_n$ are the same.

By Lemma 5.3, we have

$$
\begin{aligned}
&L\left(F^{\frac{\delta}{\sigma\sqrt{n}}\left(I_n-\mathscr{L}_n+D_n^{-\frac{1}{2}}p^2I_nD_n^{-\frac{1}{2}}\right)}, F^{\frac{\delta}{\sigma\sqrt{n}}\left(I_n-\mathscr{L}_n+(\mathbb{E}D_n)^{-\frac{1}{2}}p^2I_n(\mathbb{E}D_n)^{-\frac{1}{2}}\right)}\right)\\
\leq&\frac{\delta}{\sigma\sqrt{n}}p^2\|D_n^{-1}-(\mathbb{E}D_n)^{-1}\|\\
=&\frac{\delta}{\sigma\sqrt{n}}p^2\max_{1\leq i\leq n}\left|\frac{1}{d_i}-\frac{1}{t_i}\right|\\
=&\frac{\delta}{\sigma\sqrt{n}}p^2\max_{1\leq i\leq n}\frac{|t_i-d_i|}{t_id_i}\\
\leq&\frac{1}{\sigma\sqrt{n}}p^2\max_{1\leq i\leq n}\frac{|t_i-d_i|}{t_i}\\
\leq&\frac{1}{\sigma\sqrt{n}}p^2\max_{1\leq i\leq n}\frac{3\sqrt{t_i\ln n}}{t_i}\\
=&\frac{1}{\sigma\sqrt{n}}p^2\max_{1\leq i\leq n}3\sqrt{\frac{\ln n}{t_i}}\\
=&o(1),
\end{aligned}
$$

where the final equality holds since $\frac{t_i}{\ln n}=(n-1)(2p-p^2)/\ln n\to\infty$. This implies that the LSDs of $\frac{\delta}{\sigma\sqrt{n}}(I_n-\mathscr{L}_n+p^2D_n^{-1})$, $\frac{\delta}{\sigma\sqrt{n}}(I_n-\mathscr{L}_n+p^2(\mathbb{E}D_n)^{-1})$, and $\frac{\delta}{\sigma\sqrt{n}}C_n$ are the same.

By Lemma 5.3, we have

$$
\begin{aligned}
&L\left(F^{\frac{\delta}{\sigma\sqrt{n}}(I_n-\mathscr{L}_n+p^2(\mathbb{E}D_n)^{-1})}, F^{\frac{\delta}{\sigma\sqrt{n}}(I_n-\mathscr{L}_n)}\right)\\
\leq&\frac{\delta}{\sigma\sqrt{n}}p^2\|(\mathbb{E}D_n)^{-1}\|\\
=&\frac{\delta}{\sigma\sqrt{n}}p^2\frac{1}{\delta}\\
=&\frac{p^2}{\sigma\sqrt{n}}\\
\to&0.
\end{aligned}
$$

This implies that the LSDs of $\frac{\delta}{\sigma\sqrt{n}}(I_n-\mathscr{L}_n+p^2(\mathbb{E}D_n)^{-1})$ and $\frac{\delta}{\sigma\sqrt{n}}(I_n-\mathscr{L}_n)$

are the same.

So the LSDs of $\frac{\delta}{\sigma\sqrt{n}}(I_n - \mathscr{L}_n)$ and $\frac{\delta}{\sigma\sqrt{n}}C_n$ are the same. Equivalently, almost surely $F^{\frac{\delta}{\sigma\sqrt{n}}(I_n - \mathscr{L}_n)}(x)$ converges weakly to the standard semicircle distribution $F(x)$ with density $\phi(x)$ as $n \to \infty$. This completes the proof. $\square$

Chapter 6

The spectra of H_n and $\mathscr{L}_n$ for general random mixed graphs

In this chapter, we study the spectra of the Hermitian adjacency matrix and the normalized Hermitian Laplacian matrix of general random mixed graphs, i.e., in which all arcs are chosen independently with different probabilities (and an edge is regarded as two oppositely oriented arcs joining the same pair of vertices). For our first main result, we derive a new probability inequality and apply it to obtain an upper bound on the eigenvalues of the Hermitian adjacency matrix. Our second main result shows that the eigenvalues of the normalized Hermitian Laplacian matrix can be approximated by the eigenvalues of a closely related weighted expectation matrix, with error bounds depending on the minimum expected degree of the underlying undirected graph.

6.1 Preliminaries and auxiliary results

We start with some additional terminology and notation that we will use throughout the chapter.

6.1.1 Additional terminology and notation

We will use the notation $A \succeq \mathbf{0}$ to indicate that A is *positive semidefinite*, i.e., $A \in \mathbb{C}^{n\times n}_{Herm}$ and its eigenvalues are nonnegative, and use the notation $A \succ \mathbf{0}$ to indicate that A is *positive definite*, i.e., $A \in \mathbb{C}^{n\times n}_{Herm}$ and its eigenvalues are positive, where $\mathbf{0}$ is the zero matrix of the same size as A. With $\preceq$ we denote the *positive semidefinite order* on Hermitian matrices, as follows. Given two Hermitian matrices A and B, we use $A \preceq B$ or $B \succeq A$ to indicate that $B - A \succeq \mathbf{0}$.

Let $f : \mathbb{C} \to \mathbb{C}$ be an entire analytic function with a power-series representation $f(x) \equiv \sum_{n=0}^{\infty} a_n x^n$ $(x \in \mathbb{C})$. If all a_n are real, the expression:

$$f(A) \equiv \sum_{n=0}^{\infty} a_n A^n \ (A \in \mathbb{C}^{d\times d}_{Herm})$$

corresponds to a mapping from $\mathbb{C}^{d\times d}_{Herm}$ to itself. We note that notions of convergence are as in [74]. The *Spectral Mapping Theorem* states that each eigenvalue of $f(A)$ is equal to $f(\lambda)$ with $\lambda \in spec(A)$, i.e.,

$$spec(f(A)) = f(spec(A)). \tag{6.1}$$

In the sequel, we use the following lemma applied to the matrix exponential, to be defined shortly.

Lemma 6.1 (Lieb [113])**.** *Let $f, g : \mathbb{R} \to \mathbb{R}$, 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 $f(A) \preceq g(A)$.*

In our proofs, we make use of the *matrix exponential*, defined as $\exp(A) = \sum_{n=0}^{\infty} \frac{1}{n!} A^n$. From the Spectral Mapping Theorem we know that $\exp(A)$ is always positive definite when A is Hermitian, and that $\exp(A)$ converges for all choices of A. By Lemma 6.1, we have:

$$\text{for any } A \in \mathbb{C}^{d\times d}_{Herm},\ I + A \preceq e^{A}. \tag{6.2}$$

Moreover, we shall require brief use of the matrix logarithm. The *matrix logarithm* is defined as the functional inverse of the matrix exponential:

$$\text{for any } A \in \mathbb{C}^{d\times d}_{Herm},\ \log(e^A) := A. \tag{6.3}$$

This formula defines the logarithm of a positive definite matrix. In general, if $B = \exp(A)$, we say that A is the 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. The matrix logarithm is monotone with respect to the positive semidefinite order (See [17]):

$$\text{for any } A,B \in \mathbb{C}^{d\times d}_{Herm},\ \text{if } A \succ \mathbf{0}, B \succ \mathbf{0} \text{ and } A \preceq B,\ \text{then } \log(A) \preceq \log(B). \tag{6.4}$$

Any notation not mentioned here pertaining to matrices is as in [74].

6.1.2 Auxiliary concentration results

We shall require the following concentration inequalities in order to prove our main theorems. Various matrix concentration inequalities have been derived by many authors, including Ahlswede and Winter [1], Cristofides and Markström [37], Oliveira [98], Gross [65], Recht [109], Tropp [113], and Chung and Radcliffe [34]. In [34], Chung and Radcliffe give a short proof for the following relatively simple version that is particularly suitable for random graphs.

Theorem 6.2 ([34]). *Let $X_1, X_2, \ldots, X_m$ be independent random $n \times n$ Hermitian matrices. Moreover, assume that $\|X_i - \mathbb{E}(X_i)\| \le c$ for all i. Let $X = \sum_{i=1}^m X_i$. Then for any $a > 0$,*

$$\Pr(\|X - \mathbb{E}(X)\| \ge a) \le 2n\ \exp\Big(-\frac{a^2}{2\|\sum_{i=1}^m \mathrm{Var}(X_i)\| + 2ac/3}\Big).$$

A strengthened version of Theorem 6.2 that we need for our proof in Section 6.2, is as follows.

Theorem 6.3. *Let $X_1, X_2, \ldots, X_m$ be independent random $n \times n$ Hermitian matrices. Moreover, assume that $\|X_i\| \leq c$ for all i. Let $X = \sum_{i=1}^m X_i$. Then, for $a > \|\mathbb{E}(X)\|$:*

$$\Pr(\lambda_{\max}(X) \geq a) \leq n \ \exp\left(- \frac{(a - \|\mathbb{E}(X)\|)^2}{2\|\sum_{i=1}^m \mathbb{E}(X_i^2)\| + \frac{2c}{3}(a - \|\mathbb{E}(X)\|)} \right).$$

In particular, for $a > \|\mathbb{E}(X)\|$:

$$\Pr(\|X\| \geq a) \leq 2n \ \exp\left(- \frac{(a - \|\mathbb{E}(X)\|)^2}{2\|\sum_{i=1}^m \mathbb{E}(X_i^2)\| + \frac{2c}{3}(a - \|\mathbb{E}(X)\|)} \right). \tag{6.5}$$

Before we present our proof of Theorem 6.3, we will first show that Theorem 6.3 implies Theorem 6.2. For this purpose, let X_i $(1 \leq i \leq m)$ be as in Theorem 6.2. Let $X_i' = X_i - \mathbb{E}(X_i)$ and $X' = \sum_{i=1}^m X_i' = X - \mathbb{E}(X)$. Then $\mathbb{E}(X') = \mathbf{0}$. From the hypothesis of Theorem 6.2, we see that

$$\|X_i'\| \leq c \ \text{ for all } i \in \{1, \ldots, m\}.$$

We also have

$$\begin{aligned} \left\| \sum_{i=1}^m \mathbb{E}(X_i'^2) \right\| &= \left\| \sum_{i=1}^m \mathbb{E}(X_i - \mathbb{E}(X_i))^2 \right\| \\ &= \left\| \sum_{i=1}^m \mathrm{Var}(X_i) \right\|. \end{aligned}$$

Applying Theorem 6.3, we get that for $a > 0 = \|\mathbb{E}(X')\|$,

$$\begin{aligned} \Pr(\|X - \mathbb{E}(X)\| \geq a) &= \Pr(\|X'\| \geq a) \\ &\leq 2n \ \exp\left(- \frac{(a - \|\mathbb{E}(X')\|)^2}{2\|\sum_{i=1}^m \mathbb{E}(X_i'^2)\| + \frac{2c}{3}(a - \|\mathbb{E}(X')\|)} \right) \\ &= 2n \ \exp\left(- \frac{a^2}{2\|\sum_{i=1}^m \mathrm{Var}(X_i)\| + 2ac/3} \right). \end{aligned}$$

This shows that Theorem 6.3 implies Theorem 6.2.

6.1.3 The proof of Theorem 6.3

We are now going to prove Theorem 6.3. For our proof, we will rely on Lemma 6.1 and the following known result.

Lemma 6.4 ([113]). *Consider a finite sequence $\{X_i\}_{i=1}^m$ of independent, random, Hermitian matrices. Then*

$$\mathbb{E}\left[\operatorname{Tr}\left(\exp\left(\sum_{i=1}^m \theta X_i\right)\right)\right] \leq \operatorname{Tr}\left(\exp\left(\sum_{i=1}^m \log \mathbb{E}(\exp(\theta X_i))\right)\right)$$

for any constant $\theta \in \mathbb{R}$.

Proof of Theorem 6.3. We define

$$g(x) = 2\sum_{k=2}^{\infty} \frac{x^{k-2}}{k!} = \frac{2(e^x - 1 - x)}{x^2},$$

and first show the following facts about g, followed by short justifications for the statements.

- $g(0) = 1$. In fact,

$$g(0) = \lim_{x \to 0} \frac{2(e^x - 1 - x)}{x^2} = \lim_{x \to 0} \frac{2(e^x - 1)}{2x} = \lim_{x \to 0} \frac{e^x}{1} = 1.$$

- $g(x)$ is monotone increasing for $x \geq 0$. Note that for $x \neq 0$, $g'(x) = 2x^{-3}((x-2)e^x + x + 2)$, and so it suffices to show that $h(x) = (x-2)e^x + x + 2$ satisfies $h(x) \geq 0$ for all $x \in \mathbb{R}$. Clearly, $h(0) = 0$ and $h'(x) = (x-1)e^x + 1$. Hence, $h'(0) = 0$ and $h''(x) = xe^x$, so $h''(x) < 0$ for $x < 0$ and $h''(x) > 0$ for $x > 0$. Therefore, $h'(x)$ is monotone decreasing in $x \in (-\infty, 0]$ and $h'(x)$ is monotone increasing in $x \in (0, +\infty)$. So, $h'(x) \geq h'(0) = 0$ for all $x \in \mathbb{R}$. Thus, $h(x)$ is monotone increasing for all $x \in \mathbb{R}$. Indeed, $h(x) \geq h(0) = 0$ for all $x \in \mathbb{R}$, as required.

- $g(x) \leq 1$ for $x < 0$. In fact, $g'(x) = 2x^{-3}h(x) \leq 0$ if $x < 0$. So, the function g is decreasing for $x < 0$. Thus, $g(x) \leq g(0) = 1$ for $x < 0$.

- for $x < 3$, using $k! \geq 2 \cdot 3^{k-2}$, we obtain

$$g(x) = 2\sum_{k=2}^{\infty} \frac{x^{k-2}}{k!} \leq \sum_{k=2}^{\infty} \frac{x^{k-2}}{3^{k-2}} = \frac{1}{1 - x/3}. \tag{6.6}$$

Recalling that $g(x)$ is monotone increasing for $x \geq 0$, for $0 < x \leq c$, we get $g(x) \leq g(c)$. Now let X_i $(1 \leq i \leq m)$ be as in the hypothesis of Theorem 6.3.

Given a real constant $\theta > 0$, we have $\|\theta X_i\| \leq \theta c$. Applying Lemma 6.1, we obtain that $g(\theta X_i) \preceq g(\theta c)I$. Therefore, noting that $e^x = 1+x+\frac{1}{2}x^2 g(x)$, we have

$$\begin{aligned} e^{\theta X_i} &= I + \theta X_i + \frac{1}{2}\theta^2 g(\theta X_i) X_i^2 \\ &\preceq I + \theta X_i + \frac{1}{2}\theta^2 g(\theta c) X_i^2. \end{aligned} \tag{6.7}$$

We now use that the expectation respects the positive semidefinite order (See [113]), i.e.,

$$\text{for any } A, B \in \mathbb{C}^{d\times d}_{Herm},\ A \preceq B \text{ almost surely implies } \mathbb{E}A \preceq \mathbb{E}B. \tag{6.8}$$

Using (6.2), (6.7), and (6.8), we obtain

$$\begin{aligned} \mathbb{E}(e^{\theta X_i}) &\preceq \mathbb{E}(I + \theta X_i + \frac{1}{2}\theta^2 g(\theta c) X_i^2) \\ &= I + \theta \mathbb{E}(X_i) + \frac{1}{2}\theta^2 g(\theta c)\mathbb{E}(X_i^2) \\ &\preceq e^{\theta \mathbb{E}(X_i) + \frac{1}{2}\theta^2 g(\theta c)\mathbb{E}(X_i^2)}. \end{aligned} \tag{6.9}$$

Next, we prove the following claim related to the trace of the matrix exponential.

Claim 1. For the given matrices X_i,

$$\mathbb{E}\left[\mathrm{Tr}\left(\exp\left(\sum_{i=1}^{m} \theta X_i\right)\right)\right] \leq \mathrm{Tr}\left(\exp\left(\theta \mathbb{E}(X) + \frac{1}{2}\theta^2 g(\theta c)\sum_{i=1}^{m} \mathbb{E}(X_i^2)\right)\right). \tag{6.10}$$

Proof of Claim 1. Here we work with the trace of the matrix exponential, $\mathrm{Tr}(\exp) : A \mapsto \mathrm{Tr}(\exp(A))$. This trace exponential function is monotone with respect to the positive semidefinite order, i.e.,

$$\forall A, B \in \mathbb{C}^{d\times d}_{Herm},\ A \preceq B \text{ implies } \mathrm{Tr}(\exp(A)) \leq \mathrm{Tr}(\exp(B)). \tag{6.11}$$

See, e.g., [103], Section 2 for a short proof of this fact. Now, using Lemma 6.4, (6.3), (6.4), (6.9) and (6.11), we obtain

$$\begin{aligned}
&\mathbb{E}\left[\mathrm{Tr}\left(\exp\left(\sum_{i=1}^{m}\theta X_i\right)\right)\right]\\
\leq &\mathrm{Tr}\left(\exp\left(\sum_{i=1}^{m}\log \mathbb{E}(\exp(\theta X_i))\right)\right)\\
\leq &\mathrm{Tr}\left(\exp\left(\sum_{i=1}^{m}\log e^{\theta\mathbb{E}(X_i)+\frac{1}{2}\theta^2 g(\theta c)\mathbb{E}(X_i^2)}\right)\right)\\
= &\mathrm{Tr}\left(\exp\left(\sum_{i=1}^{m}\left(\theta\mathbb{E}(X_i)+\frac{1}{2}\theta^2 g(\theta c)\mathbb{E}(X_i^2)\right)\right)\right)\\
= &\mathrm{Tr}\left(\exp\left(\theta\mathbb{E}(X)+\frac{1}{2}\theta^2 g(\theta c)\sum_{i=1}^{m}\mathbb{E}(X_i^2)\right)\right),
\end{aligned}$$

as required. □

Note that $\exp(\theta\lambda_{\max}(X))$ is a random variable. Suppose that $\{e^{\theta a_i}\}$ is the set of all values that $\exp(\theta\lambda_{\max}(X))$ can take. Then, for any real constant $a > 0$,

$$\begin{aligned}
\mathbb{E}\Big(\exp(\theta\lambda_{\max}(X))\Big) &= \sum_i e^{\theta a_i}\,\Pr(\exp(\theta\lambda_{\max}(X)) = e^{\theta a_i})\\
&= \sum_i e^{\theta a_i}\,\Pr(\lambda_{\max}(X) = a_i)\\
&\geq \sum_{a_i\geq a} e^{\theta a_i}\,\Pr(\lambda_{\max}(X) = a_i)\\
&\geq e^{\theta a}\sum_{a_i\geq a}\Pr(\lambda_{\max}(X) = a_i)\\
&\geq e^{\theta a}\,\Pr(\lambda_{\max}(X) \geq a).
\end{aligned} \tag{6.12}$$

By (6.1), for any $s \geq 0$, and for any $A \in \mathbb{C}^{d\times d}_{Herm}$, the largest eigenvalue of e^{sA} is $e^{s\lambda_{max}(A)}$ and all eigenvalues of e^{sA} are nonnegative. Hence,

$$\exp(s\lambda_{\max}(A)) = \lambda_{\max}(\exp(sA)) \leq \mathrm{Tr}(\exp(sA)). \tag{6.13}$$

We need two more inequalities from matrix analysis, where the first one is usually referred to as the *Golden-Thompson Inequality* (See, e.g., [16]), and the second one can be found, e.g., in [124].

$$\forall d \in \{1,2,3,\dots\}, \text{ and any } A,B \in \mathbb{C}^{d\times d}_{Herm}, \mathrm{Tr}(e^{A+B}) \leq \mathrm{Tr}(e^A e^B). \tag{6.14}$$

If A and B are $n \times n$ positive semidefinite Hermitian matrices, then

$$0 \leq \mathrm{Tr}(A \cdot B) \leq \mathrm{Tr}(A) \cdot \lambda_{\max}(B) \leq \mathrm{Tr}(A) \cdot \mathrm{Tr}(B). \tag{6.15}$$

Now, given a real constant $a > \|\mathbb{E}(X)\|$, for every real constant $\theta > 0$, using (6.10), (6.12), (6.13), (6.14), and (6.15), we obtain

$$\begin{aligned}
&\Pr(\lambda_{\max}(X) \geq a) \\
\leq& e^{-\theta a}\mathbb{E}\left(e^{\theta\lambda_{\max}(X)}\right) \\
\leq& e^{-\theta a}\mathbb{E}\Big(\mathrm{Tr}(\exp(\theta X)\Big) \\
=& e^{-\theta a}\mathbb{E}\left[\mathrm{Tr}\left(\exp\left(\sum_{i=1}^{m}\theta X_i\right)\right)\right] \\
\leq& e^{-\theta a}\mathrm{Tr}\left[\exp\left(\theta\mathbb{E}(X) + \frac{1}{2}\theta^2 g(\theta c)\sum_{i=1}^{m}\mathbb{E}(X_i^2)\right)\right] \\
\leq& e^{-\theta a}\mathrm{Tr}\left[\exp\Big(\theta\mathbb{E}(X)\Big)\cdot\exp\left(\frac{1}{2}\theta^2 g(\theta c)\sum_{i=1}^{m}\mathbb{E}(X_i^2)\right)\right] \\
\leq& e^{-\theta a}\lambda_{\max}\Big(\exp\Big(\theta\mathbb{E}(X)\Big)\Big)\cdot\mathrm{Tr}\left[\exp\left(\frac{1}{2}\theta^2 g(\theta c)\sum_{i=1}^{m}\mathbb{E}(X_i^2)\right)\right] \\
\leq& e^{-\theta a}\lambda_{\max}\Big(\exp\Big(\theta\mathbb{E}(X)\Big)\Big)\cdot n\lambda_{\max}\left[\exp\left(\frac{1}{2}\theta^2 g(\theta c)\sum_{i=1}^{m}\mathbb{E}(X_i^2)\right)\right]
\end{aligned}$$

$$
\begin{aligned}
=& ne^{-\theta a}\Big(\exp\Big(\theta\lambda_{\max}(\mathbb{E}(X))\Big)\Big)\cdot\exp\left(\frac{1}{2}\theta^2 g(\theta c)\lambda_{\max}\left(\sum_{i=1}^m \mathbb{E}(X_i^2)\right)\right)\\
=& n\ \exp\left(-\theta a+\theta\lambda_{\max}(\mathbb{E}(X))+\frac{1}{2}\theta^2 g(\theta c)\lambda_{\max}\left(\sum_{i=1}^m \mathbb{E}(X_i^2)\right)\right)\\
\le& n\ \exp\left(-\theta a+\theta\|\mathbb{E}(X)\|+\frac{1}{2}\theta^2 g(\theta c)\left\|\sum_{i=1}^m \mathbb{E}(X_i^2)\right\|\right).
\end{aligned}
$$

The final inequality holds since $\|\mathbb{E}(X)\| \ge \lambda_{\max}(\mathbb{E}(X))$ and $\left\|\sum_{i=1}^m \mathbb{E}(X_i^2)\right\| \ge \lambda_{\max}(\sum_{i=1}^m \mathbb{E}(X_i^2))$.

Recall that we assume $a > \|\mathbb{E}(X)\|$. Now take $\theta = \frac{a-\|\mathbb{E}(X)\|}{\|\sum_{i=1}^m \mathbb{E}(X_i^2)\|+\frac{c}{3}(a-\|\mathbb{E}(X)\|)}$. Then, clearly $\theta c < 3$. Using (6.6), we obtain

$$
\begin{aligned}
&\Pr(\lambda_{\max}(X)\ge a)\\
\le& n\ \exp\left(-\theta a+\theta\|\mathbb{E}(X)\|+\frac{1}{2}\theta^2 g(\theta c)\left\|\sum_{i=1}^m \mathbb{E}(X_i^2)\right\|\right)\\
\le& n\ \exp\left(-\theta a+\theta\|\mathbb{E}(X)\|+\frac{\theta^2\|\sum_{i=1}^m \mathbb{E}(X_i^2)\|}{2(1-\frac{\theta c}{3})}\right)\\
=& n\ \exp\left(-\theta\left[a-\|\mathbb{E}(X)\|-\frac{\theta\|\sum_{i=1}^m \mathbb{E}(X_i^2)\|}{2(1-\frac{\theta c}{3})}\right]\right)\\
=& n\ \exp\left(-\theta\left[a-\|\mathbb{E}(X)\|-\frac{\theta\|\sum_{i=1}^m \mathbb{E}(X_i^2)\|}{2\Big(1-\frac{a-\|\mathbb{E}(X)\|}{\|\sum_{i=1}^m \mathbb{E}(X_i^2)\|+\frac{c}{3}(a-\|\mathbb{E}(X)\|)}\cdot\frac{c}{3}\Big)}\right]\right)\\
=& n\ \exp\left(-\theta\left[a-\|\mathbb{E}(X)\|-\frac{\theta\|\sum_{i=1}^m \mathbb{E}(X_i^2)\|}{\frac{2\|\sum_{i=1}^m \mathbb{E}(X_i^2)\|}{\|\sum_{i=1}^m \mathbb{E}(X_i^2)\|+\frac{c}{3}(a-\|\mathbb{E}(X)\|)}}\right]\right)\\
=& n\ \exp\left(-\theta\left[a-\|\mathbb{E}(X)\|-\frac{\theta}{2}\left(\left\|\sum_{i=1}^m \mathbb{E}(X_i^2)\right\|+\frac{c}{3}(a-\|\mathbb{E}(X)\|)\right)\right]\right)\\
=& n\ \exp\left(-\theta\left[a-\|\mathbb{E}(X)\|-\frac{1}{2}(a-\|\mathbb{E}(X)\|)\right]\right)\\
=& n\ \exp\left(-\frac{\theta}{2}(a-\|\mathbb{E}(X)\|)\right)
\end{aligned}
$$

$$=n\ \exp\left(-\frac{(a-\|\mathbb{E}(X)\|)^2}{2\|\sum_{i=1}^m \mathbb{E}(X_i^2)\|+\frac{2c}{3}(a-\|\mathbb{E}(X)\|)}\right). \tag{6.16}$$

This proves the first statement of Theorem 6.3. To obtain the norm bound (6.5) in the second statement of Theorem 6.3, recall that for any $Y \in \mathbb{C}^{n\times n}_{Herm}$,

$$\|Y\| = \max\{\lambda_{\max}(Y), -\lambda_{\min}(Y)\} = \max\{\lambda_{\max}(Y), \lambda_{\max}(-Y)\}.$$

Using this, we next apply the inequality (6.16) to the sequence $\{-X_i\}$, i.e., we replace the sequence $\{X_i\}$ by the sequence $\{-X_i\}$ in the above inequality (6.16). We obtain

$$\begin{aligned}\Pr(\lambda_{\max}(-X)\geq a) &\leq n\ \exp\left(-\frac{(a-\|\mathbb{E}(-X)\|)^2}{2\|\sum_{i=1}^m \mathbb{E}((-X_i)^2)\|+\frac{2c}{3}(a-\|\mathbb{E}(-X)\|)}\right)\\ &= n\ \exp\left(-\frac{(a-\|\mathbb{E}(X)\|)^2}{2\|\sum_{i=1}^m \mathbb{E}(X_i^2)\|+\frac{2c}{3}(a-\|\mathbb{E}(X)\|)}\right).\end{aligned}$$

Applying the union bound to the estimates for $\lambda_{\max}(X)$ and $-\lambda_{\min}(X)$, we obtain

$$\Pr(\|X\|\geq a)\leq 2n\ \exp\left(-\frac{(a-\|\mathbb{E}(X)\|)^2}{2\|\sum_{i=1}^m \mathbb{E}(X_i^2)\|+\frac{2c}{3}(a-\|\mathbb{E}(X)\|)}\right).$$

This completes the proof of Theorem 6.3. □

6.2 The spectrum of H_n

In this section, we give an upper bound on the eigenvalues of the Hermitian adjacency matrix for general random mixed graphs. We use $\Delta(\Gamma(\widehat{G}_n(p_{ij})))$ to denote the maximum expected degree of the underlying graph of $\widehat{G}_n(p_{ij})$. Hence, by straightforward calculations, we obtain the following expression: $\Delta(\Gamma(\widehat{G}_n(p_{ij}))) = \max_{1\leq i\leq n}\sum_{j=1}^n (p_{ij}+p_{ji}-p_{ij}p_{ji})$. We can thus apply Theorem 6.3 to obtain the following result.

Theorem 6.5. *Let $\widehat{G}_n(p_{ij})$ and $H_n=(h_{ij})$ be defined as in Section 1.3, and let $\Delta=\Delta(\Gamma(\widehat{G}_n(p_{ij})))$. Let $\epsilon>0$ be an arbitrarily small constant, chosen such* *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 H_n satisfy*

$$|\lambda_i(H_n)| \le \max_{1\le i\le n} \sum_{j=1}^{n} \sqrt{p_{ij}^2 p_{ji}^2 + (p_{ij}-p_{ji})^2} + 2\sqrt{\Delta \ln(2n/\epsilon)}$$

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

Before presenting our proof of Theorem 6.5, we recall one more known result that will be used in the sequel of the chapter.

Lemma 6.6 ([74]). *Let $M = (m_{ij})$ be an $n \times n$ matrix. Then*

$$\rho(M) \le \min\left\{ \max_{1\le i\le n} \sum_{j=1}^{n} |m_{ij}|, \max_{1\le j\le n} \sum_{i=1}^{n} |m_{ij}| \right\}.$$

We use $\mathbb{E}H_n$ as shorthand for $\mathbb{E}(H_n)$, and note that it is obvious that $(\mathbb{E}H_n)_{ij} = \mathbb{E}(h_{ij}) = p_{ij}p_{ji} + \mathbf{i}(p_{ij} - p_{ji})$.

Proof of Theorem 6.5. Let $\widehat{G}_n(p_{ij})$ and $H_n = (h_{ij})$ be defined as in Section 1.3, and let $\Delta = \Delta(\Gamma(\widehat{G}_n(p_{ij}))) = \max_{1\le i\le n} \sum_{j=1}^{n} (p_{ij} + p_{ji} - p_{ij}p_{ji})$.

For the indices i and j with $1 \le i, j \le n$, let H^{ij} be the $n \times n$ matrix with a 1 in the (i, j)-th position and a 0 everywhere else. Recall that h_{ij} takes value 1 with probability $p_{ij}p_{ji}$, value $\mathbf{i}$ with probability $p_{ij}(1-p_{ji})$, value $-\mathbf{i}$ with probability $(1-p_{ij})p_{ji}$, and value 0 with probability $(1-p_{ij})(1-p_{ji})$. So, $h_{ji} = \overline{h_{ij}}$, i.e., h_{ji} is the complex conjugate of h_{ij}. Take $X_{ij} = h_{ij}H^{ij} + h_{ji}H^{ji} = h_{ij}H^{ij} + \overline{h_{ij}}H^{ji}$. Then, $H_n = \sum_{1\le i<j\le n} X_{ij}$. Now, we can apply Theorem 6.3 to H_n if we derive a suitable upper bound c on $\|X_{ij}\|$. Note that X_{ij} $(1 \le i < j \le n)$ are independent random $n \times n$ Hermitian matrices, and that, with the choice $c = 1$,

$$\|X_{ij}\| = \|h_{ij}H^{ij} + \overline{h_{ij}}H^{ji}\| = |h_{ij}| < 1 = c.$$

Before applying Theorem 6.3, we first perform some additional calculations in order to obtain upper bounds for $\|\sum_{1\le i<j\le n} \mathbb{E}(X_{ij}^2)\|$ and $\|\mathbb{E}H_n\|$.

For all $1 \le i < j \le n$, we have

$$\begin{aligned}
\mathbb{E}(X_{ij}^2) &= \mathbb{E}(h_{ij}H^{ij} + \overline{h_{ij}}H^{ji})^2 \\
&= \mathbb{E}[h_{ij} \cdot \overline{h_{ij}}](H^{ii} + H^{jj}) \\
&= \mathbb{E}[|h_{ij}|^2](H^{ii} + H^{jj}) \\
&= (p_{ij} + p_{ji} - p_{ij}p_{ji})(H^{ii} + H^{jj}).
\end{aligned}$$

We set $p_{ii} = 0$. Then,

$$\begin{aligned}
\left\| \sum_{1 \le i < j \le n} \mathbb{E}(X_{ij}^2) \right\| &= \left\| \sum_{i=1}^{n} \left(\sum_{j=1}^{n} (p_{ij} + p_{ji} - p_{ij}p_{ji}) \right) H^{ii} \right\| \\
&= \max_{i=1,\dots,n} \sum_{j=1}^{n} (p_{ij} + p_{ji} - p_{ij}p_{ji}) \\
&= \Delta.
\end{aligned}$$

Recall that $(\mathbb{E}H_n)_{ij} = \mathbb{E}h_{ij} = p_{ij}p_{ji} + \mathbf{i}(p_{ij} - p_{ji})$, and in particular, $\mathbb{E}H_n$ is a Hermitian matrix. So, $\|\mathbb{E}H_n\| = \rho(\mathbb{E}H_n)$. By Lemma 6.6, we have

$$\begin{aligned}
\|\mathbb{E}H_n\| &= \rho(\mathbb{E}H_n) \\
&\le \min\left\{ \max_{1 \le i \le n} \sum_{j=1}^{n} |\mathbb{E}h_{ij}|, \max_{1 \le j \le n} \sum_{i=1}^{n} |\mathbb{E}h_{ij}| \right\} \\
&\le \max_{1 \le i \le n} \sum_{j=1}^{n} \sqrt{p_{ij}^2 p_{ji}^2 + (p_{ij} - p_{ji})^2}.
\end{aligned}$$

Now, we take $a = \|\mathbb{E}H_n\| + \sqrt{4\Delta \ln(2n/\epsilon)}$. By the assumption that $\Delta > \frac{4}{9}\ln(2n/\epsilon)$, we obtain that $a - \|\mathbb{E}H_n\| < 3\Delta$. Applying Theorem 6.3, and using $c = 1$, we get

$$\begin{aligned}
\Pr(\|H_n\| \ge a) &\le 2n \ \exp\left(- \frac{(a - \|\mathbb{E}H_n\|)^2}{2\|\sum_{1 \le i < j \le n} \mathbb{E}(X_{ij}^2)\| + \frac{2c}{3}(a - \|\mathbb{E}H_n\|)} \right) \\
&\le 2n \ \exp\left(- \frac{4\Delta \ln(2n/\epsilon)}{4\Delta} \right)
\end{aligned}$$

$$= \epsilon.$$

Thus, with probability at least $1-\epsilon$, we have that for all $1 \le i \le n$,

$$\begin{aligned}|\lambda_i(H_n)| &\le \|H_n\| \\ &\le a \\ &= \|\mathbb{E}H_n\| + \sqrt{4\Delta \ln(2n/\epsilon)}.\end{aligned}$$

This completes the proof of Theorem 6.5. □

6.3 The spectrum of $\mathscr{L}_n$

In this section, we study the spectrum of the normalized Hermitian Laplacian matrix of general random mixed graphs. We assume that $V(\widehat{G}_n(p_{ij})) = \{v_1, v_2, \ldots, v_n\}$, and we let $D_n = \mathrm{diag}(d_1, d_2, \ldots, d_n)$ denote the diagonal matrix in which d_i is the degree of the vertex v_i in the underlying graph of $\widehat{G}_n(p_{ij})$. We let $\mathbb{E}D_n$ denote the coordinate-wise expectation of D_n. Recall that $\mathscr{L}_n = I_n - D_n^{-1/2} H_n D_n^{-1/2}$ denotes the normalized Hermitian Laplacian matrix of $\widehat{G}_n(p_{ij})$, where I_n denotes the $n \times n$ identity matrix. We let $\delta(\Gamma(\widehat{G}_n(p_{ij})))$ denote the minimum expected degree of the underlying graph of $\widehat{G}_n(p_{ij})$. Hence, $\delta(\Gamma(\widehat{G}_n(p_{ij}))) = \min_{1\le i\le n} \sum_{j=1}^{n} (p_{ij} + p_{ji} - p_{ij}p_{ji})$. Our result can be stated as follows.

Theorem 6.7. *Let $\widehat{G}_n(p_{ij})$, H_n, D_n and $\mathscr{L}_n$ be defined as above, and let $\delta = \delta(\Gamma(\widehat{G}_n(p_{ij})))$. Let $\epsilon > 0$ be an arbitrarily small constant. 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 $\mathscr{L}_n$ and $\widetilde{\mathscr{L}_n}$ satisfy*

$$|\lambda_i(\mathscr{L}_n) - \lambda_i(\widetilde{\mathscr{L}_n})| \le 7\sqrt{\frac{\ln(4n/\epsilon)}{\delta}}$$

for all $1 \le i \le n$, where $\widetilde{\mathscr{L}_n} = I_n - (\mathbb{E}D_n)^{-1/2}(\mathbb{E}H_n)(\mathbb{E}D_n)^{-1/2}$.

Let $G = (V(G), E_0(G), E_1(G))$ be a mixed graph of order n. For brevity, we write D for $D(G)$, L for $L(G)$ and $\mathscr{L}$ for $\mathscr{L}(G)$. Hence, $\mathscr{L} = I - D^{-\frac{1}{2}} H D^{-\frac{1}{2}} =$

$D^{-1/2}LD^{-1/2}$. We are first going to show that $\mathscr{L}$ is positive semidefinite, by deriving an alternative expression for $\frac{x^*\mathscr{L}x}{x^*x}$ for an arbitrary nonzero complex $n \times 1$ column vector x. After that, we are going to expand the alternative expression in order to obtain an upper bound for the eigenvalues of $\mathscr{L}$, using Lemma 4.3.

In the following expansion, $y = D^{-1/2}x$, $N(v_i)$ denotes the neighborhood of v_i in the underlying graph $\Gamma(G)$, and $\sum_{e=v_iv_j}$ denotes the sum over all unordered pairs $\{v_i, v_j\}$ for which v_i and v_j are adjacent in $\Gamma(G)$.

$$
\begin{aligned}
\frac{x^*\mathscr{L}x}{x^*x} &= \frac{x^*D^{-1/2}LD^{-1/2}x}{x^*x} \\
&= \frac{y^*Ly}{(D^{1/2}y)^*(D^{1/2}y)} \\
&= \frac{y^*Ly}{y^*Dy} \\
&= \frac{(y_1^*, y_2^*, \dots, y_n^*)\begin{pmatrix} d_1 & -h_{12} & \cdots & -h_{1n} \\ -h_{21} & d_2 & \cdots & -h_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -h_{n1} & -h_{n2} & \cdots & d_n \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}}{(y_1^*, y_2^*, \dots, y_n^*)\begin{pmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}} \\
&= \frac{\sum_{v_i \in V(G)} d_i |y_i|^2 - \sum_{v_i \neq v_j} h_{ij} y_i^* y_j}{\sum_{v_i \in V(G)} d_i |y_i|^2} \\
&= \frac{\sum_{v_i} \sum_{v_j \in N(v_i)} |y_j|^2 - \sum_{v_i \neq v_j} h_{ij} y_i^* y_j}{\sum_{v_i \in V(G)} d_i |y_i|^2} \\
&= \frac{\sum_{e=v_iv_j} (|y_i|^2 + |y_j|^2) - \sum_{e=v_iv_j} (h_{ij} y_i^* y_j + h_{ji} y_i y_j^*)}{\sum_{v_i \in V(G)} d_i |y_i|^2} \\
&= \frac{\sum_{e=v_iv_j} (|y_i|^2 + |y_j|^2) - \sum_{e=v_iv_j} (h_{ij} y_i^* y_j + h_{ij}^* y_i y_j^*)}{\sum_{v_i \in V(G)} d_i |y_i|^2}
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{\sum_{e=v_iv_j}(y_i - h_{ij}y_j)(y_i^* - h_{ij}^* y_j^*)}{\sum_{v_i\in V(G)} d_i|y_i|^2}\\
&= \frac{\sum_{e=v_iv_j}(y_i - h_{ij}y_j)(y_i - h_{ij}y_j)^*}{\sum_{v_i\in V(G)} d_i|y_i|^2}\\
&= \frac{\sum_{e=v_iv_j}|y_i - h_{ij}y_j|^2}{\sum_{v_i\in V(G)} d_i|y_i|^2}.
\end{aligned}
$$

Before we continue our calculations, we note that the derived expression for $\frac{x^*\mathscr{L}x}{x^*x}$ implies that $\mathscr{L}$ is positive semidefinite. Next, we are going to expand the obtained expression for $\frac{x^*\mathscr{L}x}{x^*x}$, using the known fact that

$$
|f(x)-f(y)|^2 \le 2(|f(x)|^2 + |f(y)|^2), \tag{6.16}
$$

where equality holds if and only if $f(x) = -f(y)$.

We split $\sum_{e=v_iv_j}$ in the above expression by distinguishing undirected edges (or pairs of oppositely oriented arcs), denoted as $v_i \longleftrightarrow v_j$, and arcs, denoted as $v_i \longrightarrow v_j$ if the orientation is from v_i to v_j, and as $v_i \longleftarrow v_j$ if the orientation is from v_j to v_i. Adopting this notation, and using (6.16), we obtain

$$
\begin{aligned}
\sum_{e=v_iv_j}|y_i - h_{ij}y_j|^2 &= \sum_{v_i\longleftrightarrow v_j}|y_i - y_j|^2 + \sum_{v_i\rightarrow v_j \text{ or } v_i\leftarrow v_j}|y_i - h_{ij}y_j|^2\\
&\le \sum_{v_i\longleftrightarrow v_j} 2(|y_i|^2 + |y_j|^2) + \sum_{v_i\rightarrow v_j \text{ or } v_i\leftarrow v_j} 2(|y_i|^2 + |h_{ij}y_j|^2)\\
&= \sum_{v_i\longleftrightarrow v_j} 2(|y_i|^2 + |y_j|^2) + \sum_{v_i\rightarrow v_j \text{ or } v_i\leftarrow v_j} 2(|y_i|^2 + |y_j|^2)\\
&= 2\left(\sum_{v_i\longleftrightarrow v_j}(|y_i|^2 + |y_j|^2) + \sum_{v_i\rightarrow v_j \text{ or } v_i\leftarrow v_j}(|y_i|^2 + |y_j|^2)\right).
\end{aligned}
$$

We also obtain

$$
\sum_{v_i\in V(G)} d_i|y_i|^2 = \sum_{v_i}\sum_{v_j\in N(v_i)}|y_j|^2
$$

$$
\begin{aligned}
&= \sum_{e=v_iv_j} (|y_i|^2 + |y_j|^2) \\
&= \sum_{v_i \leftrightarrow v_j} (|y_i|^2 + |y_j|^2) + \sum_{v_i \to v_j \text{ or } v_i \leftarrow v_j} (|y_i|^2 + |y_j|^2).
\end{aligned}
$$

Therefore, using the latter two expressions and applying Lemma 4.3, we get the following upper bound on the eigenvalues of $\mathscr{L}$.

$$
\lambda_i(\mathscr{L}) \leq \sup_x \frac{\sum_{e=v_iv_j} |y_i - h_{ij}y_j|^2}{\sum_{v_i \in V(G)} d_i |y_i|^2} \leq 2.
$$

This shows that the normalized Hermitian Laplacian spectrum is in $[0,2]$, and hence that $\|I - \mathscr{L}\| \leq 1$. We will use this conclusion near the end of the proof of Theorem 6.7. We now have all the ingredients to present our proof of Theorem 6.7.

Proof of Theorem 6.7. Let $\widehat{G}_n(p_{ij})$ and $H_n = (h_{ij})$ be defined as in Section 1.3, and let $\delta = \delta(\Gamma(\widehat{G}_n(p_{ij}))) = \min_{1 \leq i \leq n} \sum_{j=1}^{n} (p_{ij} + p_{ji} - p_{ij}p_{ji})$.

For each vertex v_i of $\widehat{G}_n(p_{ij})$, we let d_i denote the degree of v_i in the underlying graph $\Gamma(\widehat{G}_n(p_{ij}))$, and we use $t_i = \mathbb{E}(d_i)$ to denote the expected degree of v_i, so $\mathbb{E}D_n = \operatorname{diag}(\mathbb{E}(d_1), \ldots, \mathbb{E}(d_n)) = \operatorname{diag}(t_1, \ldots, t_n)$. This means that the matrix $\widetilde{\mathscr{L}_n} = I_n - (\mathbb{E}D_n)^{-1/2}(\mathbb{E}H_n)(\mathbb{E}D_n)^{-1/2}$ can be seen as the "expected Laplacian matrix" of $\widehat{G}_n(p_{ij})$. Let $C_n = I_n - (\mathbb{E}D_n)^{-1/2}H_n(\mathbb{E}D_n)^{-1/2}$. Then, clearly

$$
\|\mathscr{L}_n - \widetilde{\mathscr{L}_n}\| \leq \|C_n - \widetilde{\mathscr{L}_n}\| + \|\mathscr{L}_n - C_n\| = \|\widetilde{\mathscr{L}_n} - C_n\| + \|C_n - \mathscr{L}_n\|.
$$

In the next stages, we derive bounds for each of the last two terms separately.

We first consider $\widetilde{\mathscr{L}_n} - C_n = (\mathbb{E}D_n)^{-1/2}(H_n - \mathbb{E}H_n)(\mathbb{E}D_n)^{-1/2}$. Let

$$
\begin{aligned}
Y_{ij} &= (\mathbb{E}D_n)^{-1/2}[(h_{ij} - \mathbb{E}h_{ij})H^{ij} + (h_{ji} - \mathbb{E}h_{ji})H^{ji}](\mathbb{E}D_n)^{-1/2} \\
&= \frac{(h_{ij} - \mathbb{E}h_{ij})H^{ij} + (\overline{h_{ij}} - \mathbb{E}\overline{h_{ij}})H^{ji}}{\sqrt{t_i t_j}}.
\end{aligned}
$$

Then, $\widetilde{\mathscr{L}_n} - C_n = \sum_{1 \leq i < j \leq n} Y_{ij}$. We are going to apply Theorem 6.2 to obtain an upper bound for $\|\widetilde{\mathscr{L}_n} - C_n\|$. Before we can do so, we have to perform some preliminary calculations in order to obtain an upper bound c_0 for $\|Y_{ij} - \mathbb{E}(Y_{ij})\|$, and a suitable upper bound for $\|\sum_{1\le i<j\le n} \mathrm{Var}(Y_{ij})\|$. First of all, note that for all $1 \le i < j \le n$,

$$\mathbb{E}(Y_{ij}) = \mathbb{E}\left[\frac{(h_{ij} - \mathbb{E}h_{ij})H^{ij} + (\overline{h_{ij}} - \mathbb{E}\overline{h_{ij}})H^{ji}}{\sqrt{t_i t_j}}\right] = \mathbf{0}.$$

We set $\mathbb{E}(Y_{ii}) = \mathbf{0}$. Then,

$$\begin{aligned}
\|Y_{ij} - \mathbb{E}(Y_{ij})\| &= \|Y_{ij}\| \\
&= \frac{\|(h_{ij} - \mathbb{E}h_{ij})H^{ij} + (\overline{h_{ij}} - \mathbb{E}\overline{h_{ij}})H^{ji}\|}{\sqrt{t_i t_j}} \\
&= \frac{\|(h_{ij} - \mathbb{E}h_{ij})H^{ij} + (\overline{h_{ij} - \mathbb{E}h_{ij}})H^{ji}\|}{\sqrt{t_i t_j}} \\
&= \frac{|h_{ij} - \mathbb{E}h_{ij}|}{\sqrt{t_i t_j}} \\
&= \frac{|h_{ij} - [p_{ij}p_{ji} + \mathbf{i}(p_{ij} - p_{ji})]|}{\sqrt{t_i t_j}} \\
&= \begin{cases}
\frac{\sqrt{(1-p_{ij}p_{ji})^2 + (p_{ij}-p_{ji})^2}}{\sqrt{t_i t_j}}, & \text{for } h_{ij} = 1, \\
\frac{\sqrt{(p_{ij}p_{ji})^2 + (1-(p_{ij}-p_{ji}))^2}}{\sqrt{t_i t_j}}, & \text{for } h_{ij} = \mathbf{i}, \\
\frac{\sqrt{(p_{ij}p_{ji})^2 + (1+(p_{ij}-p_{ji}))^2}}{\sqrt{t_i t_j}}, & \text{for } h_{ij} = -\mathbf{i}, \\
\frac{\sqrt{(p_{ij}p_{ji})^2 + (p_{ij}-p_{ji})^2}}{\sqrt{t_i t_j}}, & \text{for } h_{ij} = 0.
\end{cases} \\
&\le \begin{cases}
\frac{\sqrt{2}}{\sqrt{t_i t_j}}, & \text{for } h_{ij} = 1, \\
\frac{\sqrt{4}}{\sqrt{t_i t_j}}, & \text{for } h_{ij} = \mathbf{i}, \\
\frac{\sqrt{4}}{\sqrt{t_i t_j}}, & \text{for } h_{ij} = -\mathbf{i}, \\
\frac{1}{\sqrt{t_i t_j}}, & \text{for } h_{ij} = 0.
\end{cases} \\
&\le \frac{2}{\sqrt{t_i t_j}}
\end{aligned}$$

$$\leq \frac{2}{\delta}.$$

So, we are going to use $c_0 = \frac{2}{\delta}$. Next, we consider $\mathrm{Var}(Y_{ij})$ for all $1 \leq i < j \leq n$, and obtain

$$\begin{aligned}
\mathrm{Var}(Y_{ij}) &= \mathbb{E}((Y_{ij} - \mathbb{E}(Y_{ij}))^2 \\
&= \mathbb{E}Y_{ij}^2 \\
&= \frac{\mathbb{E}[(h_{ij} - \mathbb{E}h_{ij})H^{ij} + (\overline{h_{ij}} - \mathbb{E}\overline{h_{ij}})H^{ji}]^2}{t_i t_j} \\
&= \frac{\mathbb{E}[(h_{ij} - \mathbb{E}h_{ij})(\overline{h_{ij}} - \mathbb{E}\overline{h_{ij}})](H^{ii} + H^{jj})}{t_i t_j} \\
&= \frac{\mathbb{E}[(h_{ij} - \mathbb{E}h_{ij})(\overline{h_{ij} - \mathbb{E}h_{ij}})](H^{ii} + H^{jj})}{t_i t_j} \\
&= \frac{\mathrm{Var}(h_{ij})(H^{ii} + H^{jj})}{t_i t_j} \\
&= \frac{(p_{ij} + p_{ji} + p_{ij}p_{ji} - p_{ij}^2 - p_{ji}^2 - p_{ij}^2 p_{ji}^2)(H^{ii} + H^{jj})}{t_i t_j}.
\end{aligned}$$

We also have $\mathrm{Var}(Y_{ii}) = \mathbb{E}Y_{ii}^2 = \mathbf{0}$ as $p_{ii} = 0$. Therefore,

$$\begin{aligned}
\left\| \sum_{1 \leq i < j \leq n} \mathrm{Var}(Y_{ij}) \right\| &= \left\| \sum_{1 \leq i < j \leq n} \mathbb{E}Y_{ij}^2 \right\| \\
&= \left\| \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{(p_{ij} + p_{ji} + p_{ij}p_{ji} - p_{ij}^2 - p_{ji}^2 - p_{ij}^2 p_{ji}^2) H^{ii}}{t_i t_j} \right\| \\
&= \max_{i=1,\ldots,n} \sum_{j=1}^{n} \frac{p_{ij} + p_{ji} + p_{ij}p_{ji} - p_{ij}^2 - p_{ji}^2 - p_{ij}^2 p_{ji}^2}{t_i t_j} \\
&= \max_{i=1,\ldots,n} \sum_{j=1}^{n} \frac{p_{ij} + p_{ji} - p_{ij}p_{ji} + 2p_{ij}p_{ji} - p_{ij}^2 - p_{ji}^2 - p_{ij}^2 p_{ji}^2}{t_i t_j} \\
&= \max_{i=1,\ldots,n} \sum_{j=1}^{n} \frac{p_{ij} + p_{ji} - p_{ij}p_{ji} - (p_{ij} - p_{ji})^2 - p_{ij}^2 p_{ji}^2}{t_i t_j}
\end{aligned}$$

$$
\begin{aligned}
&\leq \max_{i=1,\ldots,n} \sum_{j=1}^{n} \frac{p_{ij}+p_{ji}-p_{ij}p_{ji}}{t_i t_j} \\
&\leq \max_{i=1,\ldots,n} \frac{1}{\delta} \sum_{j=1}^{n} \frac{p_{ij}+p_{ji}-p_{ij}p_{ji}}{t_i} \\
&= \frac{1}{\delta}.
\end{aligned}
$$

For the final equality, note that $d_i = \sum_{j=1}^n |h_{ij}|$, so $t_i = \mathbb{E}(d_i) = \sum_{j=1}^n \mathbb{E}|h_{ij}| = \sum_{j=1}^n [p_{ij}p_{ji} + p_{ij}(1-p_{ji}) + (1-p_{ij})p_{ji}] = \sum_{j=1}^n (p_{ij}+p_{ji}-p_{ij}p_{ji})$. Now, in order to apply Theorem 6.2, we take $a = \sqrt{\frac{4\ln(4n/\epsilon)}{\delta}}$, and we let k be large enough so that $\delta > k\ln n$ implies $a < 1$ (in particular, choosing $k > 4(1+\ln(4/\epsilon))$ is sufficient). Now, noting that $\mathbb{E}(\widetilde{\mathscr{L}_n} - C_n) = \mathbf{0}$, applying Theorem 6.2, we obtain

$$
\begin{aligned}
\Pr(\|\widetilde{\mathscr{L}_n} - C_n\| \geq a) &\leq 2n\ \exp\left(-\frac{a^2}{2\|\sum_{1\leq i<j\leq n}\mathrm{Var}(Y_{ij})\| + 2c_0 a/3}\right) \\
&\leq 2n\ \exp\left(-\frac{\frac{4\ln(4n/\epsilon)}{\delta}}{2/\delta + 4a/3\delta}\right) \\
&= 2n\ \exp\left(-\frac{4\ln(4n/\epsilon)}{2+4a/3}\right) \\
&\leq 2n\ \exp\left(-\frac{4\ln(4n/\epsilon)}{4}\right) \\
&= \frac{\epsilon}{2}.
\end{aligned}
$$

So, with probability at least $1-\frac{\epsilon}{2}$, $\|\widetilde{\mathscr{L}_n} - C_n\| \leq a$. For the second term, we first rewrite $C_n - \mathscr{L}_n$, as follows.

$$
\begin{aligned}
&C_n - \mathscr{L}_n \\
=&I_n - (\mathbb{E}D_n)^{-1/2}H_n(\mathbb{E}D_n)^{-1/2} - I_n + D_n^{-1/2}H_nD_n^{-1/2} \\
=&D_n^{-1/2}H_nD_n^{-1/2} - (\mathbb{E}D_n)^{-1/2}D_n^{1/2}D_n^{-1/2}H_nD_n^{-1/2}D_n^{1/2}(\mathbb{E}D_n)^{-1/2} \\
=&I_n - \mathscr{L}_n - (\mathbb{E}D_n)^{-1/2}D_n^{1/2}(I_n - \mathscr{L}_n)D_n^{1/2}(\mathbb{E}D_n)^{-1/2} \\
=&(I_n - \mathscr{L}_n) - (I_n - \mathscr{L}_n)D_n^{1/2}(\mathbb{E}D_n)^{-1/2}
\end{aligned}
$$

$$
\begin{aligned}
&- (\mathbb{E}D_n)^{-1/2}D_n^{1/2}(I_n - \mathscr{L}_n)D_n^{1/2}(\mathbb{E}D_n)^{-1/2} + (I_n - \mathscr{L}_n)D_n^{1/2}(\mathbb{E}D_n)^{-1/2} \\
=&(I_n - \mathscr{L}_n)[I_n - D_n^{1/2}(\mathbb{E}D_n)^{-1/2}]+ \\
&+ [I_n - (\mathbb{E}D_n)^{-1/2}D_n^{1/2}](I_n - \mathscr{L}_n)D_n^{1/2}(\mathbb{E}D_n)^{-1/2}.
\end{aligned}
$$

Recalling that $\|I_n - \mathscr{L}_n\| \leq 1$, we obtain the following expression for $\|C_n - \mathscr{L}_n\|$.

$$
\begin{aligned}
\|C_n - \mathscr{L}_n\| \leq& \|I_n - \mathscr{L}_n\|\|I_n - D_n^{1/2}(\mathbb{E}D_n)^{-1/2}\| \\
&+ \|I_n - (\mathbb{E}D_n)^{-1/2}D_n^{1/2}\|\|I_n - \mathscr{L}_n\|\|D_n^{1/2}(\mathbb{E}D_n)^{-1/2}\| \\
\leq& \|I_n - D_n^{1/2}(\mathbb{E}D_n)^{-1/2}\| + \|I_n - (\mathbb{E}D_n)^{-1/2}D_n^{1/2}\|\|D_n^{1/2}(\mathbb{E}D_n)^{-1/2}\|.
\end{aligned}
$$

Next, we are going to obtain an upper bound for $\|I_n - D_n^{1/2}(\mathbb{E}D_n)^{-1/2}\|$. For this, we will apply Lemma 5.4 to the random variables $|h_{ij}|$ (in the role of X_i), and using the observations that $d_i = \sum_{j=1}^n |h_{ij}|$, and $t_i = \mathbb{E}(d_i) = \sum_{j=1}^n \mathbb{E}|h_{ij}|$. We first need some preparation in order to obtain upper bounds for $|h_{ij}|$ and $\sum_{j=1}^n \mathrm{Var}(|h_{ij}|)$.

Obviously, $|h_{ij}| \leq 1$, so we can take $c = 1$, and

$$
\begin{aligned}
\sum_{j=1}^n \mathrm{Var}(|h_{ij}|) &= \sum_{j=1}^n [\mathbb{E}(|h_{ij}|^2) - (\mathbb{E}(|h_{ij}|))^2] \\
&= \sum_{j=1}^n [p_{ij} + p_{ji} - p_{ij}p_{ji} - (p_{ij} + p_{ji} - p_{ij}p_{ji})^2] \\
&\leq \sum_{j=1}^n (p_{ij} + p_{ji} - p_{ij}p_{ji}) \\
&= \sum_{j=1}^n \mathbb{E}|h_{ij}| \\
&= t_i.
\end{aligned}
$$

Since we already used a above, with $a = \sqrt{\frac{4\ln(4n/\epsilon)}{\delta}} < 1$, we are going to use a b instead of an a when applying Lemma 5.4. We choose $b = \sqrt{3t_i \ln(4n/\epsilon)}$. Then, since $a = \sqrt{\frac{4\ln(4n/\epsilon)}{\delta}} < 1$, $t_i \geq \delta > 4\ln(4n/\epsilon)$, implying that $b =$

$\sqrt{3t_i \ln(4n/\epsilon)} < t_i$. Applying Lemma 5.4, we have for all i,

$$\Pr(|d_i - t_i| \geq b) \leq e^{-\frac{b^2}{2(t_i+b/3)}} < e^{-\frac{3t_i \ln(4n/\epsilon)}{3t_i}} = \frac{\epsilon}{4n}.$$

This implies that with probability at least $1-o(1)$, for all $1 \leq i \leq n$, $|d_i - t_i| \leq b = \sqrt{3t_i \ln(4n/\epsilon)}$.

Next, we use the known fact that $|\sqrt{x}-1| \leq |x-1|$ for any real number $x > 0$. Taking $x = \frac{d_i}{t_i} > 0$, we obtain that with probability at least $1-\frac{\epsilon}{2}$,

$$\left|\sqrt{\frac{d_i}{t_i}} - 1\right| \leq \left|\frac{d_i}{t_i} - 1\right| = \frac{|d_i - t_i|}{t_i} \leq \frac{\sqrt{3t_i \ln(4n/\epsilon)}}{t_i} \leq \sqrt{\frac{3\ln(4n/\epsilon)}{\delta}} = \frac{\sqrt{3}}{2}a.$$

Thus, we obtain

$$\|I_n - D_n^{1/2}(\mathbb{E}D_n)^{-1/2}\| = \max_{i=1,2,\ldots,n} \left|\sqrt{\frac{d_i}{t_i}} - 1\right| \leq \frac{\sqrt{3}}{2}a$$

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

$$\begin{aligned}
\|C_n - \mathscr{L}_n\| &\leq \|I_n - D_n^{1/2}(\mathbb{E}D_n)^{-1/2}\| + \|I_n - (\mathbb{E}D_n)^{-1/2}D_n^{1/2}\|\|D_n^{1/2}(\mathbb{E}D_n)^{-1/2}\| \\
&\leq \frac{\sqrt{3}}{2}a + \frac{\sqrt{3}}{2}a\left(\frac{\sqrt{3}}{2}a + 1\right) \\
&= \frac{3}{4}a^2 + \sqrt{3}a.
\end{aligned}$$

Combining the above bound with the bound we obtained for $\|C_n - \widetilde{\mathscr{L}_n}\|$, and using that $a < 1$, we conclude that with probability at least $1-\epsilon$,

$$\begin{aligned}
\|\mathscr{L}_n - \widetilde{\mathscr{L}_n}\| &\leq \|C_n - \widetilde{\mathscr{L}_n}\| + \|C_n - \mathscr{L}_n\| \\
&\leq a + \frac{3}{4}a^2 + \sqrt{3}a \\
&\leq \frac{7}{2}a \\
&= \frac{7}{2}\sqrt{\frac{4\ln(4n/\epsilon)}{\delta}}
\end{aligned}$$

$$= 7\sqrt{\frac{\ln(4n/\epsilon)}{\delta}}.$$

For the final step in our proof, we use Lemma 2.7, 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 $1 \leq i \leq n$,

$$|\lambda_i(\mathscr{L}_n) - \lambda_i(\widetilde{\mathscr{L}_n})| \leq \|\mathscr{L}_n - \widetilde{\mathscr{L}_n}\| \leq 7\sqrt{\frac{\ln(4n/\epsilon)}{\delta}}.$$

This completes the proof of Theorem 6.7. $\square$

Chapter 7

The spectra of S_n and R_S for random oriented graphs

In this chapter we study the spectra of the skew adjacency matrix and the skew Randić matrix for random oriented graphs. In particular, we apply a probability inequality to deduce upper bounds for the skew spectral radius and skew Randić spectral radius of random oriented graphs.

7.1 Preliminaries

Previously, various matrix concentration inequalities for random matrices have been derived by Tropp [113]. Here, we only need to recall the following inequality in order to prove our main theorems.

Theorem 7.1 (Tropp [113]). *Let $\{Z_k\}_{k=1}^m$ be a finite sequence of independent, random matrices with dimensions $n_1 \times n_2$. Assume that each random matrix satisfies*

$$\mathbb{E}(Z_k) = \mathbf{0} \ \ \textit{and} \ \ \|Z_k\| \leq c \ \ \textit{almost surely.}$$

Define

$$\omega^2 = \max\left\{ \left\| \sum_k \mathbb{E}(Z_k Z_k^*) \right\|, \left\| \sum_k \mathbb{E}(Z_k^* Z_k) \right\| \right\}.$$

Then for any $a \geq 0$,

$$
\begin{aligned}
\Pr\Big(\Big\|\sum_k Z_k\Big\| \geq a\Big) &\leq (n_1+n_2)\cdot \exp\Big(-\frac{\omega^2}{c^2}\cdot h\Big(\frac{ac}{\omega^2}\Big)\Big) \\
&\leq (n_1+n_2)\cdot \exp\Big(-\frac{a^2/2}{\omega^2+ac/3}\Big) \\
&\leq \begin{cases} (n_1+n_2)\cdot \exp\Big(-\frac{3a^2}{8\omega^2}\Big), & \textit{for } a \leq \omega^2/c; \\ \\ (n_1+n_2)\cdot \exp\Big(-\frac{3a}{8c}\Big), & \textit{for } a \geq \omega^2/c. \end{cases}
\end{aligned}
$$

Here, the function $h(u) := (1+u)\ln(1+u) - u$ for $u \geq 0$.

7.2 The spectrum of S_n

In this section we study the spectrum of the skew adjacency matrix for random oriented graphs. In particular, we derive an upper bound for their skew spectral radius.

Let $G_n^{\sigma}(p_{ij})$ be a random oriented graph of order n, and let $S_n = (s_{ij})_{n\times n}$ be the skew adjacency matrix of $G_n^{\sigma}(p_{ij})$ as described in Section 1.4, where p_{ij} is a function of n such that $0 < p_{ij} < 1$. We use $\Delta(\Gamma(G_n^{\sigma}(p_{ij})))$ to denote the maximum expected degree of the underlying graph $\Gamma(G_n^{\sigma}(p_{ij}))$ of $G_n^{\sigma}(p_{ij})$, Hence, by straightforward calculations, we obtain the following expression: $\Delta(\Gamma(G_n^{\sigma}(p_{ij}))) = \max_{i=1,\dots,n}\sum_{j=1}^{n} p_{ij}$. We can apply Theorem 7.1 to obtain the following result.

Theorem 7.2. *Let $G_n^{\sigma}(p_{ij})$ and $S_n = (s_{ij})$ be defined as in Section 1.4, and let $\Delta = \Delta(\Gamma(G_n^{\sigma}(p_{ij})))$. Let $\epsilon > 0$ be an arbitrarily small constant, chosen such that for n sufficiently large, $\Delta > \frac{4}{9}\ln(2n/\epsilon)$. Then with probability at least $1-\epsilon$, for n sufficiently large, the skew spectral radius of $G_n^{\sigma}(p_{ij})$ satisfies*

$$\rho(S_n) \leq 2\sqrt{\Delta \ln(2n/\epsilon)}.$$

Proof of Theorem 7.2. Let $G_n^{\sigma}(p_{ij})$ and $S_n = (s_{ij})$ be defined as in Section 1.4, and let $\Delta = \Delta(\Gamma(G_n^{\sigma}(p_{ij}))) = \max_{i=1,\dots,n}\sum_{j=1}^{n} p_{ij}$.

For the indices i and j with $1 \le i \ne j \le n$, let S^{ij} be the $n \times n$ matrix with a 1 in the (i,j)-th position, a -1 in the (j,i)-th position, and a 0 everywhere else. Recall that s_{ij} takes value 1 with probability $\frac{p_{ij}}{2}$, -1 with probability $\frac{p_{ij}}{2}$, and 0 with probability $1 - p_{ij}$. Take $X_{ij} = s_{ij}S^{ij}$. Then $S_n = \sum_{1 \le i<j \le n} X_{ij}$. Note that X_{ij} $(1 \le i < j \le n)$ are independent random $n \times n$ matrices, with

$$\mathbb{E}(X_{ij}) = \mathbb{E}(s_{ij})S^{ij} = \mathbf{0}.$$

Now, we can apply Theorem 7.1 to S_n if we derive an upper bound c on $\|X_{ij}\|$. With the choice $c = 1$,

$$\begin{aligned}
\|X_{ij}\| &= \|s_{ij}S^{ij}\| \\
&= |s_{ij}|\|S^{ij}\| \\
&\le \|S^{ij}\| \\
&= \sqrt{\lambda_{\max}((S^{ij})^* S^{ij})} \\
&= \sqrt{\lambda_{\max}((S^{ij})^T S^{ij})} \\
&= \sqrt{\lambda_{\max}(E^i + E^j)} \\
&= 1 \\
&= c,
\end{aligned}$$

where $(S^{ij})^T$ is the transpose of S^{ij}, and E^i is the matrix with a 1 in the (i,i)-th position, and a 0 everywhere else. Before applying Theorem 7.1, we first perform some additional calculations in order to obtain upper bounds for $\|\sum_{1\le i<j\le n} \mathbb{E}[X_{ij}(X_{ij})^*]\|$ and $\|\sum_{1\le i<j\le n} \mathbb{E}[X_{ij}^*(X_{ij})]\|$.

For all $1 \le i < j \le n$, we have

$$\begin{aligned}
\mathbb{E}[(X_{ij})^* X_{ij}] &= \mathbb{E}[(s_{ij}S^{ij})^* s_{ij}S^{ij}] \\
&= \mathbb{E}[s_{ij}^2 (S^{ij})^* S^{ij}] \\
&= \mathbb{E}[s_{ij}^2 (S^{ij})^T S^{ij}] \\
&= \mathbb{E}(s_{ij}^2) \cdot (S^{ij})^T S^{ij} \\
&= p_{ij}(E^i + E^j).
\end{aligned}$$

Similarly, we have

$$\mathbb{E}[X_{ij}(X_{ij})^*] = p_{ij}(E^i + E^j).$$

We set $p_{ii} = 0$. Then,

$$\begin{aligned}
\omega^2 &= \max\left\{\left\|\sum_{1\le i<j\le n}\mathbb{E}[(X_{ij})^*X_{ij}]\right\|, \left\|\sum_{1\le i<j\le n}\mathbb{E}[X_{ij}(X_{ij})^*]\right\|\right\} \\
&= \left\|\sum_{i=1}^{n}\left(\sum_{j=1}^{n}p_{ij}\right)E^i\right\| \\
&\le \max_{i=1,\ldots,n}\sum_{j=1}^{n}p_{ij} \\
&= \Delta.
\end{aligned}$$

Now, we take $a = \sqrt{4\Delta\ln(2n/\epsilon)}$. By the assumption that $\Delta > \frac{4}{9}\ln(2n/\epsilon)$, we obtain that $a < 3\Delta$. Applying Theorem 7.1, and using $c = 1$, we get

$$\begin{aligned}
\Pr(\|S_n\| \ge a) &\le 2n\ \exp\left(-\frac{a^2}{2\Delta + 2a/3}\right) \\
&\le 2n\ \exp\left(-\frac{4\Delta\ln(2n/\epsilon)}{4\Delta}\right) \\
&= \epsilon.
\end{aligned}$$

Thus, with probability at least $1-\epsilon$, we have that for all $1 \le i \le n$,

$$\|S_n\| \le a = 2\sqrt{\Delta\ln(2n/\epsilon)}.$$

It is well known that all the eigenvalues of S_n are purely imaginary numbers. Assume that $\lambda_1 = \mathbf{i}\mu_1, \lambda_2 = \mathbf{i}\mu_2, \ldots, \lambda_n = \mathbf{i}\mu_n$ are all the eigenvalues of S_n, where every μ_k $(1 \le k \le n)$ is a real number and $\mathbf{i}$ is the imaginary unit. Let $\widetilde{S_n} = (-\mathbf{i})S_n$. Then $\widetilde{S_n}$ is an Hermitian matrix with eigenvalues exactly $\mu_1, \mu_2, \ldots, \mu_n$. Therefore,

$$\begin{aligned}
\|S_n\| =& \|\widetilde{S_n}\| \\
=& \rho(\widetilde{S_n})
\end{aligned}$$

$$
\begin{aligned}
&= \max_{1\le i\le n}\{|\mu_i(\widetilde{S_n})|\} \\
&= \max_{1\le i\le n}\{|-\mathbf{i}\lambda_i(S_n)|\} \\
&= \max_{1\le i\le n}\{|\lambda_i(S_n)|\} \\
&=\rho(S_n). \qquad (7.1)
\end{aligned}
$$

Thus, with probability at least $1-\epsilon$, we have

$$\rho(S_n) = \|S_n\| \le 2\sqrt{\Delta \ln(2n/\epsilon)}.$$

This completes the proof. □

7.3 The spectrum of $\boldsymbol{R}_S$

In this section we study the spectrum of the skew Randić matrix for random oriented graphs. In particular, we derive an upper bound for their skew Randić spectral radius.

Let $V(G_n^{\sigma}(p_{ij})) = \{v_1, v_2, \ldots, v_n\}$, and we let $D_n = \mathrm{diag}(d_1, d_2, \ldots, d_n)$ denote the diagonal matrix in which d_i is the degree of the vertex v_i in the underlying graph of $G_n^{\sigma}(p_{ij})$. Recall that $\boldsymbol{R}_S = D_n^{-\frac{1}{2}} S_n D_n^{-\frac{1}{2}}$ denotes the skew Randić matrix of $G_n^{\sigma}(p_{ij})$. We let $\delta(\Gamma(G_n^{\sigma}(p_{ij})))$ denote the minimum expected degree of the underlying graph of $G_n^{\sigma}(p_{ij})$. Hence, $\delta(\Gamma(G_n^{\sigma}(p_{ij}))) = \min_{i=1,\ldots,n} \sum_{j=1}^{n} p_{ij}$. Our result is stated as follows.

Theorem 7.3. *Let $G_n^{\sigma}(p_{ij}), S_n, D_n$ and $\boldsymbol{R}_S$ be defined as above, and let $\delta = \delta(\Gamma(G_n^{\sigma}(p_{ij})))$. Let $\epsilon > 0$ be an arbitrarily small constant. Then there exists a constant $k = k(\epsilon)$ such that if $\delta > k \ln n$, then with probability at least $1-\epsilon$, the skew Randić spectral radius of $\boldsymbol{R}_S$ satisfies*

$$\rho(\boldsymbol{R}_S) \le \frac{9}{4}\sqrt{\frac{3\ln(4n/\epsilon)}{\delta}}.$$

Proof of Theorem 7.3. Let $G_n^{\sigma}(p_{ij})$ and $S_n = (s_{ij})$ be defined as in Section 1.4, and let $\delta = \delta(\Gamma(G_n^{\sigma}(p_{ij}))) = \min_{i=1,\ldots,n} \sum_{j=1}^{n} p_{ij}$.

For each vertex v_i of $G_n^{\sigma}(p_{ij})$, we let d_i denote the degree of v_i in the underlying graph $\Gamma(G_n^{\sigma}(p_{ij}))$, and we use $t_i = \mathbb{E}(d_i)$ to denote the expected degree of v_i, so $\mathbb{E}D_n = \operatorname{diag}(\mathbb{E}(d_1), \mathbb{E}(d_2), \ldots, \mathbb{E}(d_n)) = \operatorname{diag}(t_1, t_2, \ldots, t_n)$. Let $\boldsymbol{B}_n = (\mathbb{E}D_n)^{-1/2} S_n (\mathbb{E}D_n)^{-1/2}$. Then, clearly

$$\|\boldsymbol{R}_S\| = \|\boldsymbol{R}_S - \boldsymbol{B}_n + \boldsymbol{B}_n\| \leq \|\boldsymbol{R}_S - \boldsymbol{B}_n\| + \|\boldsymbol{B}_n\|.$$

In the next stages, we derive bounds for each of the last two terms separately.

We first consider $\|\boldsymbol{B}_n\|$. Let

$$\begin{aligned} Y_{ij} &= (\mathbb{E}D_n)^{-1/2}(s_{ij}S^{ij})(\mathbb{E}D_n)^{-1/2} \\ &= \frac{s_{ij}S^{ij}}{\sqrt{t_i t_j}}. \end{aligned}$$

Then, $\boldsymbol{B}_n = \sum_{1 \leq i < j \leq n} Y_{ij}$. We are going to apply Theorem 7.1 to obtain an upper bound for $\|\boldsymbol{B}_n\|$. Before we can do so, we have to perform some preliminary calculations in order to obtain an upper bound c_0 for $\|Y_{ij}\|$, and suitable upper bounds for $\left\|\sum_{1 \leq i < j \leq n} \mathbb{E}[(Y_{ij})^* Y_{ij}]\right\|$ and $\left\|\sum_{1 \leq i < j \leq n} \mathbb{E}[Y_{ij}(Y_{ij})^*]\right\|$. First of all, note that for all $1 \leq i < j \leq n$,

$$\begin{aligned} \mathbb{E}Y_{ij} &= \mathbb{E}[(\mathbb{E}D_n)^{-1/2}(s_{ij}S^{ij})(\mathbb{E}D_n)^{-1/2}] \\ &= (\mathbb{E}D_n)^{-1/2}\mathbb{E}(s_{ij}S^{ij})(\mathbb{E}D_n)^{-1/2} \\ &= \mathbb{E}(s_{ij})(\mathbb{E}D_n)^{-1/2}S^{ij}(\mathbb{E}D_n)^{-1/2} \\ &= \mathbf{0}. \end{aligned}$$

We set $\mathbb{E}(Y_{ii}) = \mathbf{0}$. Then,

$$\begin{aligned} \|Y_{ij}\| &= \frac{\|s_{ij}S^{ij}\|}{\sqrt{t_i t_j}} \\ &\leq \frac{1}{\sqrt{t_i t_j}} \\ &\leq \frac{1}{\delta}. \end{aligned}$$

So, we can use $c_0 = \frac{1}{\delta}$. Next, we consider $\mathbb{E}[(Y_{ij})^* Y_{ij}]$ and $\mathbb{E}[Y_{ij}(Y_{ij})^*]$ for $1 \le i < j \le n$, and obtain

$$\begin{aligned}
\mathbb{E}[(Y_{ij})^* Y_{ij}] &= \mathbb{E}\left[\left(\frac{s_{ij}S^{ij}}{\sqrt{t_i t_j}}\right)^* \frac{s_{ij}S^{ij}}{\sqrt{t_i t_j}}\right] \\
&= \mathbb{E}\left[\frac{s_{ij}^2}{t_i t_j}(S^{ij})^* S^{ij}\right] \\
&= \mathbb{E}\left[\frac{s_{ij}^2}{t_i t_j}(S^{ij})^T S^{ij}\right] \\
&= \frac{1}{t_i t_j}\mathbb{E}[s_{ij}^2] \cdot (S^{ij})^T S^{ij} \\
&= \frac{p_{ij}}{t_i t_j}(E^i + E^j).
\end{aligned}$$

Similarly, we have

$$\mathbb{E}[Y_{ij}(Y_{ij})^*] = \frac{p_{ij}}{t_i t_j}(E^i + E^j).$$

We also have $\mathbb{E}[(Y_{ii})^* Y_{ii}] = \mathbf{0}$ and $\mathbb{E}[Y_{ii}(Y_{ii})^*] = \mathbf{0}$ as $p_{ii} = 0$. Therefore,

$$\begin{aligned}
\omega^2 &= \max\left\{\left\|\sum_{1\le i<j\le n}\mathbb{E}[(Y_{ij})^* Y_{ij}]\right\|, \left\|\sum_{1\le i<j\le n}\mathbb{E}[Y_{ij}(Y_{ij})^*]\right\|\right\} \\
&= \left\|\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{p_{ij}E^i}{t_i t_j}\right\| \\
&= \max_{i=1,\dots,n}\sum_{j=1}^{n}\frac{p_{ij}}{t_i t_j} \\
&\le \max_{i=1,\dots,n}\frac{1}{\delta}\sum_{j=1}^{n}\frac{p_{ij}}{t_i} \\
&= \frac{1}{\delta}.
\end{aligned}$$

For the final equality, we used that $d_i = \sum_{j=1}^n |s_{ij}|$, so $t_i = \mathbb{E}(d_i) = \sum_{j=1}^n \mathbb{E}|s_{ij}| = \sum_{j=1}^n p_{ij}$. Now, in order to apply Theorem 7.1, we take $a = \sqrt{\frac{3\ln(4n/\epsilon)}{\delta}}$ and we let k be large enough so that $\delta > k \ln n$ implies $a < 1$ (in particular, choosing $k > 3(1+\ln(4/\epsilon))$ is sufficient). Now, noting that $\mathbb{E}(\boldsymbol{B}_n) = \mathbf{0}$, applying Theorem 7.1, we obtain

$$\begin{aligned}
\Pr(\|\boldsymbol{B}_n\| > a) &\leq 2n \ \exp\left(-\frac{\frac{3\ln(4n/\epsilon)}{\delta}}{\frac{2}{\delta} + 2a/(3\delta)} \right) \\
&\leq 2n \ \exp\left(\frac{-3\ln(4n/\epsilon)}{3} \right) \\
&= \frac{\epsilon}{2}.
\end{aligned}$$

So, with probability at least $\geq 1 - \frac{\epsilon}{2}$,

$$\|\boldsymbol{B}_n\| \leq a = \sqrt{\frac{3\ln(4n/\epsilon)}{\delta}}.$$

For the second term, we first rewrite $\boldsymbol{R}_S - \boldsymbol{B}_n$, as follows.

$$\begin{aligned}
\boldsymbol{R}_S - \boldsymbol{B}_n =& D_n^{-1/2} S_n D_n^{-1/2} - (\mathbb{E}D_n)^{-1/2} S_n (\mathbb{E}D_n)^{-1/2} \\
=& D_n^{-1/2}(\mathbb{E}D_n)^{1/2}(\mathbb{E}D_n)^{-1/2} S_n (\mathbb{E}D_n)^{-1/2}(\mathbb{E}D_n)^{1/2} D_n^{-1/2} - \boldsymbol{B}_n \\
=& D_n^{-1/2}(\mathbb{E}D_n)^{1/2} \boldsymbol{B}_n (\mathbb{E}D_n)^{1/2} D_n^{-1/2} - \boldsymbol{B}_n \\
=& D_n^{-1/2}(\mathbb{E}D_n)^{1/2} \boldsymbol{B}_n (\mathbb{E}D_n)^{1/2} D_n^{-1/2} - \boldsymbol{B}_n (\mathbb{E}D_n)^{1/2} D_n^{-1/2} - \boldsymbol{B}_n \\
& + \boldsymbol{B}_n (\mathbb{E}D_n)^{1/2} D_n^{-1/2} \\
=& [D_n^{-1/2}(\mathbb{E}D_n)^{1/2} - I_n] \boldsymbol{B}_n (\mathbb{E}D_n)^{1/2} D_n^{-1/2} + \boldsymbol{B}_n [(\mathbb{E}D_n)^{1/2} D_n^{-1/2} - I_n].
\end{aligned}$$

Then we have

$$\begin{aligned}
\|\boldsymbol{R}_S - \boldsymbol{B}_n\| \leq& \|D_n^{-1/2}(\mathbb{E}D_n)^{1/2} - I_n\| \|\boldsymbol{B}_n\| \|(\mathbb{E}D_n)^{1/2} D_n^{-1/2}\| + \\
& + \|\boldsymbol{B}_n\| \|(\mathbb{E}D_n)^{1/2} D_n^{-1/2} - I_n\|.
\end{aligned}$$

Next, we are going to obtain an upper bound for $\|(\mathbb{E}D_n)^{1/2} D_n^{-1/2} - I_n\|$. For this, we will apply Lemma 5.4 to the random variables $|s_{ij}|$ (in the role of X_i), and using the observation that $d_i = \sum_{j=1}^n |s_{ij}|$, so $t_i = \mathbb{E}(d_i) = \sum_{j=1}^n \mathbb{E}|s_{ij}|$. We first need some preparation in order to obtain upper bounds for $|s_{ij}|$ and

$\sum_{j=1}^{n} \mathrm{Var}(|s_{ij}|)$.

Obviously, $|s_{ij}| \leq 1$, so we take $c = 1$, and

$$\begin{aligned}\sum_{j=1}^{n} \mathrm{Var}(|s_{ij}|) &= \sum_{j=1}^{n} [\mathbb{E}(|s_{ij}|^2) - (\mathbb{E}(|s_{ij}|))^2] \\ &= \sum_{j=1}^{n} p_{ij}(1 - p_{ij}) \\ &\leq \sum_{j=1}^{n} p_{ij} \\ &= t_i.\end{aligned}$$

Since we already used a above, with $a = \sqrt{\frac{3\ln(4n/\epsilon)}{\delta}} < 1$, we are going to use a b instead of an a when applying Lemma 5.4. We choose $b = \sqrt{3t_i \ln(4n/\epsilon)}$. Then, since $a = \sqrt{\frac{3\ln(4n/\epsilon)}{\delta}} < 1$, $t_i \geq \delta > 3\ln(4n/\epsilon)$), implying that $b = \sqrt{3t_i \ln(4n/\epsilon)} < t_i$. Applying Lemma 5.4, we have for all i,

$$\Pr(|d_i - t_i| \geq b) \leq e^{-\frac{b^2}{2(t_i + b/3)}} < e^{-\frac{3t_i \ln(4n/\epsilon)}{3t_i}} = \frac{\varepsilon}{4n}.$$

This implies that with probability at least $1 - o(1)$, for all $1 \leq i \leq n$,

$$|d_i - t_i| \leq b = \sqrt{3t_i \ln(4n/\epsilon)}.$$

Next, we choose $0 < a_0 = \sqrt{\frac{3\ln(4n/\epsilon)}{t_i}} < 1$ such that $\frac{1}{1-a_0} < 1 + \epsilon$.

$$\begin{aligned}&\Pr\left(\frac{t_i}{d_i} > (1+\epsilon)\right) \\ =&\Pr\left(\{\frac{t_i}{d_i} > (1+\epsilon)\} \cap \{|d_i - t_i| < a_0 t_i\}\right) + \\ &+ \Pr\left(\{\frac{t_i}{d_i} > (1+\epsilon)\} \cap \{|d_i - t_i| \geq a_0 t_i\}\right) \\ \leq&\Pr\left(\{\frac{t_i}{d_i} > (1+\epsilon)\} \cap \{|d_i - t_i| < a_0 t_i\}\right) + \Pr\left(|d_i - t_i| \geq a_0 t_i\right)\end{aligned}$$

$$
\begin{aligned}
&= 0 + \Pr\left(|d_i - t_i| \geq a_0 t_i\right) \\
&= 0 + \Pr\left(|d_i - t_i| \geq b\right) \\
&\leq \frac{\epsilon}{4n}.
\end{aligned}
$$

Hence, with probability at least $1 - o(1)$, for all $1 \leq i \leq n$, $\frac{t_i}{d_i} \leq 1$. Then

$$
\begin{aligned}
\left|\sqrt{\frac{t_i}{d_i}} - 1\right| \leq& \left|\frac{t_i - d_i}{\sqrt{d_i}(\sqrt{t_i} + \sqrt{d_i})}\right| \\
\leq& \frac{1}{2}\sqrt{\frac{3\ln(4n/\epsilon)}{t_i}} \\
\leq& \frac{1}{2}\sqrt{\frac{3\ln(4n/\epsilon)}{\delta}} \\
=& \frac{a}{2}.
\end{aligned}
$$

Thus, we obtain

$$
\|D_n^{-1/2}(\mathbb{E}D_n)^{1/2} - I_n\| = \max_{i=1,2,\ldots,n}\left|\sqrt{\frac{t_i}{d_i}} - 1\right| \leq \frac{a}{2}
$$

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

$$
\begin{aligned}
&\|\boldsymbol{R}_S - \boldsymbol{B}_n\| \\
\leq& \|D_n^{-1/2}(\mathbb{E}D_n)^{1/2} - I_n\|\|\boldsymbol{B}_n\|\|(\mathbb{E}D_n)^{1/2}D_n^{-1/2}\| + \|\boldsymbol{B}_n\|\|(\mathbb{E}D_n)^{1/2}D_n^{-1/2} - I_n\| \\
\leq& \frac{a^2}{2}\left(\frac{a}{2} + 1\right) + \frac{a^2}{2} \\
\leq& \frac{5}{4}a.
\end{aligned}
$$

Combining the above bound with the bound we obtained for $\|\boldsymbol{B}_n\|$, and using that $a < 1$, we conclude that with probability at least $1 - \epsilon$,

$$
\|\boldsymbol{R}_S\| \leq \|\boldsymbol{R}_S - \boldsymbol{B}_n\| + \|\boldsymbol{B}_n\|
$$

$$
\begin{aligned}
&\leq \frac{5}{4}a + a \\
&= \frac{9}{4}a.
\end{aligned}
$$

Since $\boldsymbol{R}_S = D_n^{-1/2} S_n D_n^{-1/2}$ is a skew matrix, by similar arguments as those that led to (7.1), we have $\|\boldsymbol{R}_S\| = \rho(\boldsymbol{R}_S)$. Then we have

$$
\rho(\boldsymbol{R}_S) = \|\boldsymbol{R}_S\| \leq \frac{9}{4}a = \frac{9}{4}\sqrt{\frac{3\ln(4n/\epsilon)}{\delta}}.
$$

This completes the proof. $\square$

Summary

This thesis contains a number of results on the spectra and related spectral properties of several random graph models.

In Chapter 1, we briefly present the background, some history as well as the main ideas behind our work. Apart from the introduction in Chapter 1, the first part of the main body of the thesis is Chapter 2. In this part we estimate the eigenvalues of the Laplacian matrix of random multipartite graphs. We also investigate some spectral properties of random multipartite graphs, such as the Laplacian energy, the Laplacian Estrada index, and the von Neumann entropy.

The second part consists of Chapters 3, 4, 5 and 6. In [67], Guo and Mohar showed that mixed graphs are equivalent to digraphs if we regard (replace) each undirected edge as (by) two oppositely directed arcs. Motivated by the work of Guo and Mohar, we initially propose a new random graph model – the random mixed graph. Each arc is determined by an independent random variable. The main themes of the second part are the spectra and related spectral properties of random mixed graphs.

In Chapter 3, we prove that the empirical distribution of the eigenvalues of the Hermitian adjacency matrix converges to Wigner's semicircle law. As an application of the LSD of Hermitian adjacency matrices, we estimate the Hermitian energy of a random mixed graph.

In Chapter 4, we deal with the asymptotic behaviour of the spectrum of the Hermitian adjacency matrix of random mixed graphs. We derive a separation result between the first and the remaining eigenvalues of H_n. As an

application of the asymptotic behaviour of the spectrum of the Hermitian adjacency matrix, we estimate the spectral moments of random mixed graphs.

In Chapter 5, we prove that the empirical distribution of the eigenvalues of the normalized Hermitian Laplacian matrix converges to Wigner's semicircle law.

Moreover, in Chapter 6, we provide several results on the spectra of general random mixed graphs. In particular, we present a new probability inequality for sums of independent, random, self-adjoint matrices, and then apply this probability inequality to matrices arising from the study of random mixed graphs. We prove a concentration result involving the spectral norm of a random matrix and that of its expectation. Assuming that the probabilities of all the arcs of the mixed graph are mutually independent, we write the Hermitian adjacency matrix as a sum of random self-adjoint matrices. Using this, we estimate the spectrum of the Hermitian adjacency matrix, and prove a concentration result involving the spectrum of the normalized Hermitian Laplacian matrix and its expectation.

Finally, in Chapter 7, we estimate upper bounds for the spectral radii of the skew adjacency matrix and skew Randić matrix of random oriented graphs.

Samenvatting

Dit proefschrift bevat een aantal resultaten op het gebied van de spectra van verschillende typen randomgrafen en daarmee verwante spectrale eigenschappen.

In Hoofdstuk 1 wordt de achtergrond van het onderzoek geschetst, met een historische perspectief, alsmede de belangrijkste ideeën die ten grondslag liggen aan het gepresenteerde werk in dit proefschrift.

Naast de inleiding in Hoofdstuk 1, bestaat het eerste deel van de technische inhoud van dit proefschrift uit de resultaten uit Hoofdstuk 2. In dit gedeelte worden benaderingen gegeven voor de eigenwaarden van de Laplacian matrix van random multipartiete grafen. Tevens worden enkele spectrale eigenschappen van deze random multipartiete grafen bestudeerd, waaronder de Laplacian energie, de Laplacian Estrada index, en de von Neumann entropie.

Het tweede deel van het proefschrift bestaat uit de Hoofdstukken 3, 4, 5 en 6. In [67] tonen Guo and Mohar aan dat gemengde grafen equivalent zijn aan gerichte grafen, als de ongerichte lijnen worden opgevat als paren van twee tegengesteld gerichte pijlen. Gemotiveerd door het werk van Guo and Mohar beschouwen we een nieuw randomgraaf model, te weten het model van de random gemengde graaf. Hierin is elke pijl aanwezig met een zekere waarschijnlijkheid, onafhankelijk van de andere pijlen. De belangrijkste thema's in dit tweede gedeelte van het proefschrift zijn de spectra en gerelateerde spectrale eigenschappen van deze random gemengde grafen.

In Hoofdstuk 3 wordt bewezen dat de empirische verdeling van de eigenwaarden van de Hermitische buurmatrix van deze grafen convergeert naar een verdeling die bekend staat als Wigner's semicircle law. Als een toepassing geven we een benadering voor de Hermitische energie van een random gemengde graaf.

In Hoofdstuk 4 wordt ingegaan op het asymptotische gedrag van het spectrum van de Hermitische buurmatrix van random gemengde grafen. Er wordt een resultaat afgeleid voor de scheiding van de eerste en de overige eigenwaarden. Als een toepassing van het asymptotische gedrag van het spectrum van de Hermitische buurmatrix geven we een benadering voor de spectrale momenten van random gemengde grafen.

In Hoofdstuk 5 wordt bewezen dat de empirische verdeling van de eigenwaarden van de genormaliseerde Hermitische Laplacian matrix van deze grafen ook convergeert naar Wigner's semicircle law.

Meer aanvullende resultaten met betrekking tot het spectrum van random gemengde grafen worden beschreven in Hoofdstuk 6. Allereerst presenteren we daar een nieuwe ongelijkheid voor de waarschijnlijkheid van de som van onafhankelijke, random, zelf-adjuncte matrices. Die ongelijkheid passen we vervolgens toe op matrices die gerelateerd zijn aan de studie naar random gemengde grafen. We bewijzen een concentratieresultaat met betrekking tot de spectrale norm van een random matrix en die van zijn verwachtingswaarde. We schrijven vervolgens de Hermitische buurmatrix als een som van random zelf-adjuncte matrices en geven een benadering voor het spectrum van de Hermitische buurmatrix. Bovendien bewijzen we nog een concentratieresultaat met betrekking tot het spectrum van de genormaliseerde Hermitische Laplacian matrix en zijn verwachtingswaarde.

Tenslotte geven we in Hoofdstuk 7 benaderingen van bovengrenzen voor de spectrale radius van de scheve buurmatrix en de scheve Randić matrix van random georiënteerde grafen.

Bibliography

[1] R. Ahlswede and A. Winter, Strong converse for identification via quantum channels, *IEEE Trans. Inform. Theory*, **48 (3)** (2002), 569–579.

[2] S.E. Alm and S. Linusson, A counter-intuitive correlation in a random tournament, *Combin. Probab. Comput.*, **20** (2011), 1–9.

[3] N. Alon, M. Krivelevich and V.H. Vu, Concentration of eigenvalues of random matrices, *Israel Math. J.,* **131** (2002), 259–267.

[4] N. Alon and J. Spencer, *The Probabilistic Method*, Wiley-Interscience, 2000.

[5] K. Anand and G. Bianconi, Entropy measures for networks: Toward an information theory of complex topologies, *Phys. Rev. E,* **80** (2009), 045102(R).

[6] K. Anand, G. Bianconi and S. Severini, Shannon and von Neumann entropy of random networks with heterogeneous expected degree, *Phys. Rev. E,* **83** (2011), 036109.

[7] L. Arnold, On the asymptotic distribution of the eigenvalues of random matrices, *J. Math. Anal. Appl.*, **20** (1967), 262–268.

[8] L. Arnold, On Wigner's semicircle law for the eigenvalues of random matrices, *Z. Wahrsch. Verw. Gebiete,* **19** (1971), 191–198.

[9] K.M.R. Audenaert, A Sharp Fannes-type Inequality for the von Neumann Entropy, *J. Phys. A,* **40(28)** (2007), 8127-8136.

[10] Z. Bai, Circular law, *Ann. Probab.*, **25** (1997), 494–529.

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

[12] Z. Bai and J.W. Silverstein, *Spectral Analysis of Large Dimensional Random Matrices*, second ed., Springer Series in Statistics, Springer, 2010.

[13] Z. Bai and Y. Yin, Convergence to the semicircle law, *Ann. Probab.*, **16** (1988a), 863–875.

[14] Z. Bai and Y. Yin, Necessary and sufficient conditions for the almost sure convergence of the largest eigenvalue of a Wigner matrix, *Ann. Probab.*, **16** (1988b), 1729–1741.

[15] H. Bamdad, F. Ashraf and I. Gutman, Lower bounds for Estrada index and Laplacian Estrada index, *Appl. Math. Lett.*, **23** (2010), 739–742.

[16] R. Bhatia, *Matrix Analysis*, Graduate Texts in Mathematics, vol. 169, Springer, Berlin, 1997, p.10.

[17] R. Bhatia, *Positive Definite Matrices*, Princeton Univ. Press, Princeton, 2007.

[18] P. Billingsley, *Probability and Measure*, third ed., John Wiley & Sons, Inc., 1995.

[19] B. Bollobás, *Random Graphs*, Second Edition, Cambridge University Press, 2001.

[20] B. Bollobás and P. Erdős, Graphs of extremal weights, *Ars Combin.*, 50 (1998), 225–233.

[21] J.A. Bondy and U.S.R. Murty, *Graph Theory with Application*, Macmillan London and Elsevier, New York, 1976.

[22] J.A. Bondy and U.S.R. Murty, *Graph Theory*, Springer Graduate Texts in Mathematics, vol. 244 (2008).

[23] Ş.B. Bozkurt, A.D. Göngör, I. Gutman and A.S. Çevik, Randić Matrix and Randić Energy, *MATCH Commun. Math. Comput. Chem.*, 64 (2010), 239–250.

[24] S.L. Braunstein, S. Ghosh and S. Severini, The Laplacian of a graph as a density matrix: a basic combinatorial approach to separability of mixed states, *Ann. Comb.*, **10(3)** (2006), 291–317.

[25] A.E. Brouwer and W.H. Haemers, Spectra of Graphs. Springer, 2012, available online at http://homepages.cwi.nl/ aeb/math/ipm/.

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

[27] A. Chang and B. Deng, On the Laplacian energy of trees with perfect matchings, *MATCH Commun. Math. Comput. Chem.*, **68** (2012), 767–776.

[28] Z. Chen, Y. Fan and W. Du, Spectral moment of random graphs, *Math. Appl.*, **24** (2011), 851–857.

[29] X. Chen, X. Li and H. Lian, The skew energy of random oriented graphs, *Linear Algebra Appl.*, **438** (2013), 4547–4556.

[30] H. Chernoff, A note on an inequality involving the normal distribution, *Ann. Probab.*, **9** (1981), 533–535.

[31] F. Chung, *Spectral graph theory*, AMS publications, 1997.

[32] F. Chung, L. Lu and V.H. Vu, Eigenvalues of random power law graphs, *Ann. Combin.*, **7** (2003), 21–33.

[33] F. Chung, L. Lu and V.H. Vu, Spectra of random graphs with given expected degrees, *Proc. Nat. Acad. Sci. USA*, **100(11)** (2003), 6313–6318.

[34] F. Chung and M. Radcliffe, On the Spectra of General Random Graphs, *Electron. J. Combin.*, **18** (2011), P215, 14 pp.

[35] A. Coja-Oghlan, On the Laplacian eigenvalues of $G(n,p)$, *Combin. Probab. Comput.*, **16** (2007), 923–946.

[36] A. Coja-Oghlan and A. Lanka, The spectral gap of random graphs with given expected degrees, *Electron. J. Combin.*, **16** (2009), R138.

[37] D. Cristofides and K. Markström, Expansion properties of random Cayley graphs and vertex transitive graphs via matrix martingales, *Random Struct. Alg.*, **32** (2008), 88–100.

[38] D.M. Cvetković, M. Doob and H. Sachs, *Spectra of Graphs: Theory and Applications*, III revised and enlarged edition. Johan Ambrosius Barth Verlag, Heidelberg-Leipzig, 1995.

[39] D.M. Cvetković, P. Rowlinson and S. K. Simić, *An Introduction to the Theory of Graph Spectra*, Cambridge University Press, Cambridge, 2010.

[40] K.C. Das and S.A. Mojallal, On Laplacian energy of graphs, *Discrete Math.*, **325** (2014), 52–64.

[41] K.C. Das, S.A. Mojallala and I. Gutman, On Laplacian energy in terms of graph invariants, *Appl. Math. Comput.*, **268** (2015), 83–92.

[42] H. Deng and J. Zhang, A note on the Laplacian Estrada index of trees, *MATCH Commun. Math. Comput. Chem.*, **63** (2010), 777–782.

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

[44] W. Du, X. Li and Y. Li, The energy of random graphs, *Linear Algebra Appl.*, **435** (2011), 2334–2346.

[45] W. Du, X. Li and Y. Li, The Laplacian energy of random graphs, *J. Math. Anal. Appl.*, **368** (2010), 311–319.

[46] W. Du, X. Li and Y. Li, Various energies of random graphs, *MATCH Commun. Math. Comput. Chem.*, **64** (2010), 251–260.

[47] W. Du, X. Li, Y. Li and S. Severini, A note on the von Neumann entropy of random graphs, *Linear Algebra Appl.*, **433** (2010), 1722–1725.

[48] P. Erdős and A. Rényi, On random graphs I, *Publ. Math. Debrecen.*, **6** (1959), 290–297.

[49] L. Euler. Solvtio problematis ad geometriam sitvs pertinentis. *Comment. Acad. Sci. U. Petrop.*, **8** (1736), 128–140.

[50] K. Fan, Maximum properties and inequalities for the eigenvalues of completely continuous operators, *Proc. Natl. Acad. Sci. USA*, **37** (1951), 760–766.

[51] G.H. Fath-Tabar, A.R. Ashrafi and I. Gutman, Note on Estrada and L-Estrada indices of graphs, *Bull. Cl. Sci. Math. Nat. Sci. Math.*, **139** (2009), 1–16.

[52] U. Feige and E. Ofek, Spectral techniques applied to sparse random graphs, *Random Struct. Alg.*, **27(2)** (2005), 251–275.

[53] J. Friedman, *A Proof of Alon's Second Eigenvalue Conjecture and Related Problem*, Memoirs of the American Mathematical Society 2008, 100 pp.

[54] J. Friedman, On the second eigenvalue and random walks in random d-regular graphs, *Combinatorica*, **11(4)** (1991), 331–362.

[55] J. Friedman, J. Kahn, E. Szemerédi, On the second eigenvalue in random regular graphs, in *Proc. 21st ACM Symp. Theory of Computing*, 1989, 587–598.

[56] E. Fritscher, C. Hoppen, I. Rocha and V. Trevisan, On the sum of the Laplacian eigenvalues of a tree, *Linear Algebra Appl.*, **435 (2)** (2011), 371–399.

[57] Z. Füredi and J. Komlós, The eigenvalues of random symmetric matrices, *Combinatorica*, **1(3)** 1981, 233–241.

[58] S. Geman, The spectral radius of large random matrices, *Ann. Probab.*, **14** (1986), 1318–1328.

[59] E.N. Gilbert, Random graphs, *Ann. Math. Stat.*, **30** (1959), 1141–1144.

[60] V.L. Girko, Circle law, *Theory Probab. Appl.*, **4** (1984a), 694–706.

[61] V.L. Girko, On the circle law, *Theory Probab. Math. Statist.*, **28** (1984b), 15–23.

[62] V.L. Girko, W. Kirsch and A. Kutzelnigg, A necessary and sufficient condition for the semicircular law, *Random Oper. Stoch. Equ.*, **2** (1994), 195–202.

[63] U. Grenander, *Probabilities on Algebraic Structures*, John Wiley, New York-London. 1963.

[64] G.R. Grimmett, Infinite paths in randomly oriented lattices, *Random Struct. Alg.*, **18** (2001), 257–266.

[65] D. Gross, Recovering low-rank matrices from few coefficients in any basis, *IEEE Trans. Inform. Theory*, **57** (2011), 1548–1566.

[66] R. Gu, F. Huang, X. Li, Skew Randić matrix and skew Randić energy, *Trans.Comb.*, 5(1) (2016), 1–14.

[67] K. Guo and B. Mohar, Hermitian adjacency matrix of digraphs and mixed graphs, *J. Graph Theory*, **85** (2017), no. 1, 217–248.

[68] I. Gutman, The energy of a graph, *Ber. Math. Statist. Sekt. Forschungsz. Graz.*, **103** (1978), 1–22.

[69] I. Gutman and B. Furtula (Eds.), *Recent Results in the Theory of Randić Index*, Univ. Kragujevac, Kragujevac, 2008.

[70] I. Gutman, B. Furtula and Ş. B. Bozkurt, On Randić energy, *Linear Algebra Appl.*, 442 (2014), 50–57.

[71] I. Gutman and X. Li, *Energies of Graphs–Theory and Applications*, Mathematical Chemistry Monographs No.17, Kragujevac, 2016, pp.III+290. ISBN: 978-86-6009-033-3.

[72] I. Gutman and B. Zhou, Laplacian energy of a graph, *Linear Algebra Appl.*, **414** (2006), 29–37.

[73] L. Han, F. Escolano and E.R. Hancock, Graph characterizations from von Neumann entropy, *Pattern Recognit. Lett.*, **33(15)** (2012), 1958–1967.

[74] R.A. Horn and C.R. Johnson, *Matrix Analysis*, 2nd, Cambridge University Press, 2012.

[75] D. Hu, X. Li, X. Liu and S. Zhang, The Laplacian energy and Laplacian Estrada index of random multipartite graphs, *J. Math. Anal. Appl.*, **443** (2016), 675–687.

[76] D. Hu, X. Li, X. Liu and S. Zhang, The spectral distribution of random mixed graphs, *Linear Algebra Appl.*, **519** (2017), 343–365.

[77] A. Ilić and B. Zhou, Laplacian Estrada index of trees, *MATCH Commun. Math. Comput. Chem.*, **63** (2010), 769–776.

[78] S. Janson, T. Łuczak and A. Ruczynski, *Random Graphs*, Wiley, 2000.

[79] T. Jiang, Empirical distributions of Laplacian matrices of large dilute random graphs, *Random Matrices Theory Appl.,* **1** (2012), no. 3, 1250004, 20 pp.

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

[81] X. Li and I. Gutman, *Mathematical Aspects of Randić-type Molecular Structure Descriptors*, Univ. Kragujevac, Kragujevac, 2006.

[82] X. Li and Y. Shi, A survey on the Randić index, *MATCH Commun. Math. Comput. Chem.*, 59 (2008), 127–156.

[83] X. Li, Y. Shi and I. Gutman, *Graph Energy*, Springer, New York, 2012.

[84] J. Li, W.C. Shiu and A. Chang, On the Laplacian Estrada index of a graph, *Appl. Anal. Discrete Math.*, **3** (2009), 147–156.

[85] Y. Li and Y. Wang, Further results on entropy and separability, *J. Phys. A: Math. Theor.*, **45** (2012), 385305.

[86] E.H. Lieb, Convex trace functions and the Wigner-Yanase-Dyson conjecture, *Adv. Math.*, **11** (1973), 267–288.

[87] S. Linusson, A note on correlations in randomly oriented graphs, arXiv:0905.2881.

[88] J. Liu and X. Li, Hermitian-adjacency matrices and Hermitian energies of mixed graphs, *Linear Algebra Appl.*, **466** (2015), 182–207.

[89] M. Loève, *Probability Theory*, fourth ed., Springer-Verlag, New York, 1977.

[90] László Lovász, *Large Networks and Graph Limits*, American Mathematical Society Colloquium Publications 60. Amer. Math. Soc., Providence, RI, 2012.

[91] L. Lu and X. Peng, Loose Laplacian spectra of random hypergraphs. *Random Struct. Alg.*, **41** (2012), no. 4, 521–545.

[92] L. Lu and X. Peng, Spectra of edge-independent random graphs, *Electron. J. Combin.*, **20** (2013), Paper 27, 18 pp.

[93] M.L. Mehta, *Random Matrices*, second ed., Academic Press, 1991.

[94] M. Mesbahi and M. Egerstedt, *Graph Theoretic Methods in Multiagent Networks*, 1st edition (Princeton Series in Applied Mathematics). Princeton, NJ: Princeton University Press, 2010.

[95] B. Narayanan, Connections in randomly oriented graphs, *Combin. Probab. Comput.*, (2016), pages 1–5.

[96] T. Nie, Z. Guo, K. Zhao and Z. Lu, Using mapping entropy to identify node centrality in complex networks, *Phys. A*, **453** (2016), 290–297.

[97] V. Nikiforov, The energy of graphs and matrices, *J. Math. Anal. Appl.*, **320** (2007), 1472–1475.

[98] R. Oliveira, Concentration of the adjacency matrix and of the Laplacian in random graphs with independent edges, http://arxiv.org/abs/0911.0600.

[99] O. Osenda, F.M. Pont, A. Okopińska and P. Serra, Exact finite reduced density matrix and von Neumann entropy for the Calogero model, *J. Phys. A*, **48(48)** (2015), 485301.

[100] O. Osenda and P. Serra, Scaling of the von Neumann entropy in a two-electron system near the ionization threshold, *Phys. Rev. A*, **75(4)** (2007), 810–814.

[101] F. Passerini and S. Severini, Quantifying complexity in networks: the von Neumann entropy, *Int. J. Agent Technol. Syst.*, **1(4)** (2009), 58–67.

[102] F. Passerini and S. Severini, The von Neumann entropy of networks, December 14 2008. Available at SSRN: `http://ssrn.com/abstract=1382662` or `http://dx.doi.org/10.2139/ssrn.1382662`.]

[103] D. Petz, A survey of certain trace inequalities, in *Functional Analysis and Operator Theory*. Banach Center Publications, vol. 30(Polish Acad. Sci., Warsaw, 1994), pp. 287–298.

[104] D. Petz, *Quantum Information Theory and Quantum Statistics*, Springer, Berlin, 2008.

[105] J. Rada, A. Tineo. Upper and lower bounds for the energy of bipartite graphs. *J. Math. Anal. Appl.*, **289(2)** (2004), 446–455.

[106] M. Randić, On characterization of molecular branching, *J. Amer. Chem. Soc.*, 97 (1975), 660–6615.

[107] M. Randić, On history of the Randić index and emerging hostility toward chemical graph theory, *MATCH Commun. Math. Comput. Chem.*, 59 (2008), 5–124.

[108] M. Randić, The connectivity index 25 years after, *J. Mol. Graph. Model.*, 20 (2001), 19–35.

[109] B. Recht, Simpler approach to matrix completion, *J. Mach. Learn. Res.*, **12** (2011), 3413–3430.

[110] C. Rovelli and F. Vidotto, Single particle in quantum gravity and Braunstein-Ghosh-Severini entropy of a spin network, *Phys. Rev. D*, **81** (2010), 044038.

[111] H.L. Royden, *Real Analysis*, Prentice Hall. 1988.

[112] A.N. Shiryaev, *Probability*, 2nd edition, New York: Springer-Verlag, 1996.

[113] J. Tropp, User-Friendly Tail Bounds for Sunms of Random Matrices, *Found. Comput. Math.*, **12** (2012), 389–434.

[114] C.T.M. Vinagre, R.R.Del-Vecchio, D.A.R. Justo and V. Trevisan, Maximum Laplacian energy among threshold graphs, *Linear Algebra Appl.*, **439** (2013), 1479–1495.

[115] J. von Neumann, *Mathematische Grundlagen der Quantenmechanik*, Berlin, 1932; English translation by R. T. Beyer, *Mathematical Foundations of Quantum Mechanics*, Princeton, 1955.

[116] V.H. Vu, Spectral norm of random matrices, *Combinatorica*, **27(6)** (2007), 721–736.

[117] D.B. West, *Introduction to Graph Theory*, volume 2. Prentice Hall Englewood Cliffs, 2001.

[118] H. Weyl, Das asymptotische Verteilungsgesetz der Eigenwerte linearer partieller Differentialgleichungen, *Math. Ann.*, **71** (2010), 441–479.

[119] E.P. Wigner, Characteristic vectors of bordered matrices with infinite dimensions, *Ann. Math.*, **62** (1955), 548–564.

[120] E.P. Wigner, On the distribution of the roots of certain symmetric matrices, *Ann. Math.*, **67** (1958), 325–327.

[121] K. Yates, *Hückel Molecular Orbital Theory*, Academic Press, New York, 1978.

[122] Y. Yin, LSD' for a class of random matrices, *J. Multivariate Anal.*, **20** (1986), 50–68.

[123] G. Yu and H. Qu, Hermitian Laplacian matrix and positive of mixed graphs, *Appl. Math. Comput.*, **269** (2015), 70–76.

[124] F. Zhang, *Matrix theory*, New York, Springer-Verlag, 1999.

[125] L. Zhang and J. Wu, Von Neumann entropy-preserving quantum operation, *Phys. Lett. A*, **375** (2011), 4163-4165.

[126] B. Zhou, On Laplacian eigenvalues of a graph, *Z. Naturforsch.*, **59a** (2004), 181–184.

[127] B. Zhou and I. Gutman, More on the Laplacian Estrada index, *Appl. Anal. Discrete Math.*, **3** (2009), 371–378.

Acknowledgements

First and foremost, I would like to thank Professor Xueliang Li, for giving me the opportunity to pursue this research that led to a PhD, and for his encouragement. He is always positive and creative, and he introduced me to many research directions and interesting problems. His extensive knowledge and good academic taste influenced me a lot.

I am also grateful to Professor Shenggui Zhang of NPU, for his immense amount of patience and support in shaping my academic skills. He supported me to visit the UT in order to enlarge my insights and enrich my experience. Certainly, I would not be where I am today without his guidance, for which I am eternally grateful.

I would like to express my gratitude to Professor Hajo Broersma, my supervisor at the UT. Without his guidance and encouragement, I could not have achieved what I have done in this thesis. The experience of doing a PhD with him is exciting, unforgettable and greatly beneficial. I have to say that I am lucky to be his student, and I will thank him forever.

I also wish to thank my co-authors, especially Xiaogang Liu, for his role, not only in working with me, but educating me in ways of which he may not even be aware. He was always generous in providing me with many helpful discussions and kind advices. I also wish to thank Chuandong Xu and Ruonan Li, for their helpful discussions and kind advices. I have learned a great deal from all of you.

I would also like to thank all my NPU colleagues, for the pleasant time on study and living. I have made many friends here in UT who have all contributed in a large way to my successful work. I also owe special thanks to all FMT members at UT, for their many kinds of help.

I am deeply indebted to my beloved husband Jiangyong Hou, my parents, and Jiangyong's parents, for their love, understanding and support. They have never doubted me, and their faith and love have kept me going when even I thought I could not.

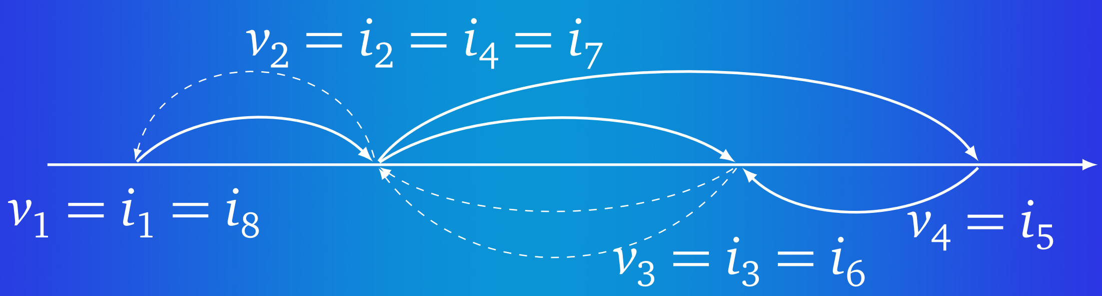
Dan Hu

June 2018, Enschede

About the Author

Dan Hu was born on April 9, 1989 in Yangling City, Shaanxi Province, P.R. of China. From 1995 to 2006, Dan Hu attended her elementary and secondary education in her hometown.

In September 2007, she started her study at Qingdao University, where she obtained her Bachelor degree in June 2011. After the national postgraduate entrance examination in 2011, she started her study as a Master student at the Ocean University of China, in Qingdao. After obtaining her Master degree in June 2014, she was selected to become a PhD candidate under the supervision of Professor Xueliang Li and Professor Shenggui Zhang, at Northwestern Polytechnical University in Xi'an. Starting from October 2017, she visited the group of Formal Methods and Tools, University of Twente, as a joint PhD student to perform research on random spectral graph theory under the supervision of Professor Hajo Broersma. The research has been sponsored by the China Scholarship Council.



$v_2 = i_2 = i_4 = i_7$
$v_1 = i_1 = i_8$
$v_4 = i_5$
$v_3 = i_3 = i_6$