Spectral analysis of inhomogeneous network models

Malhotra, N.

Spectral Analysis of Inhomogeneous Network Models

Proefschrift

ter verkrijging van
de graad van doctor aan de Universiteit Leiden,
op gezag van rector magnificus prof. dr. ir. H. Bijl,
volgens besluit van het college voor promoties
te verdedigen op donderdag 20 november 2025
klokke 11:30 uur

door

Nandan Malhotra

geboren te New Delhi, India
in 1998

Promotores:

Prof. dr. W.T.F. den Hollander

Co-promotores:

Dr. R.S. Hazra
Dr. L. Avena (Università degli Studi di Firenze)

Promotiecommissie:

Prof. dr. ir. G.L.A. Derks
Prof. dr. A. Doelman
Prof. dr. M. Löwe (University of Münster)
Prof. dr. R. Speicher (Saarland University)
Dr. N.S. Müller (Eindhoven University of Technology)

Cover: designed by Aishwarya R. Viswamitra, based on designs from the text *Spectrum analysis* by H.E.Roscoe and Sir A.Schuster, Macmillan and Co., 1885, and *Spectrum of various sources of light* by Ainsworth Potter and Co., 1882. Images by Nandan Malhotra.
Chapter title pages: Background image downloaded from rawpixel.com
Bookmark: desgined by Paolo Bordignon

The research in this thesis was supported in part by the Netherlands Organisation for Scientific Research (NWO) through the Gravitation NETWORKS grant 024.002.003, and further supported by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement no. 945045.

Contents

CHAPTER 1
Introduction

The broad goal of this thesis is to study the graph spectrum of various inhomogeneous random graph models, in particular, to characterise the eigenvalue distributions of random matrices associated with these random graphs. The first three chapters cover the graph *adjacency matrix*, whereas the fourth chapter is dedicated to the *Laplacian matrix*.

§1.1 Background

Throughout the history of mathematics, complex challenges have often driven the evolution of new areas of research. In the 20^{th} century, there were major developments in several related disciplines, such as in the natural sciences (physics, chemistry, and biology), as well as in computer science, the social sciences, and medicine. Consequently, there was increasing interest in analysing data and describing phenomena observed through experimental methods, which in turn pushed the boundaries of mathematics. There was a need for precise mathematical frameworks to capture complex phenomena, giving rise to entirely new branches that are now fundamental in modern mathematics.

Typically, *complex systems* such as social networks, biological networks, and atomic nuclei, are difficult to analyse directly, even in the era of supercomputers and increasingly efficient algorithms. Mathematical models provide a reasonable approximation of such systems, and are built up over years of research. They often begin with deceptively simple "toy models", and are subsequently generalised to more "realistic models" where the analysis can be challenging. Naturally, this also gives rise to several interesting questions in mathematics itself from a more abstract point of view. Moreover, while these branches of mathematics originate from distinct problems, as is the case for *random matrix theory* and *random graphs*, they cross paths frequently, yet continue to exist as independent research topics in their own right.

This chapter will serve as a preface to the material that will follow in the rest of the thesis. We dive into *spectral graph theory*, a topic that emerged in the 1950s and serves as the backbone of this thesis. We describe graphs and their matrices, namely the adjacency matrix and the Laplacian matrix, and give a brief overview of the relation between graph properties and the spectrum of their associated matrices.

Transitioning to the world of probability, we move on to *random graphs*, which were introduced in the mid 20^{th} century. Over the years, a wide range of systems have been studied as complex networks, in particular biological and social networks. The explosive growth of these networks in the digital age and their increasing complexity underscore the need for robust mathematical models, which led to further development of the subject in the late 20^{th} and early 21^{st} century. Random graphs are graph-valued probabilistic objects and are essential in modelling real-world networks. We will present a brief overview of a toy model and various *graph regimes*, before proceeding with more general models.

We proceed with another kind of probabilistic object: a *random matrix*, which is a matrix with random entries. This thesis focuses primarily on the eigenvalue distribution of random matrix models, that are associated with random graph models. It is important to note that although the main motivation comes from the study of random graph models, the essential tools of the trade come from random matrix theory, thereby also making the study of the spectrum relevant from a random matrix perspective.

The above naturally eases us into a more abstract theory of random variables. Abstraction is a fundamental aspect of mathematics, giving rise to areas such as *free probability* where one abstracts the notion of random variables and moves away from an underlying *sample space*. Despite this abstraction, a link with reality remains. Random matrix theory connects with free probability, and was born out of applications in statistics, operator algebras, and quantum physics.

With these notions well established, we proceed with a literature overview of *spectral theory for random graphs*, in particular for the Erdős-Rényi random graph. We conclude with an outline of the thesis and technical results that are used in later chapters, as well as a short discussion and concluding remarks.

§1.2 Spectral Graph Theory

Spectral graph theory is the study of the relation between geometric properties of graphs and the eigenvalues and eigenvectors of the associated graph matrices. Motivated by applications in quantum physics and chemistry, the theory is now used in various areas of mathematics, such as discrete mathematics and combinatorics, statistics, and probability, while also playing a crucial role in statistical physics and computer science. There are various references on the subject. We refer to Chung [1997] for an introduction, and to Spielman [2012] for a modern approach to the subject.

§1.2.1 Graphs and matrices

Graphs can be defined set-theoretically as a collection of two sets: a vertex set, and an edge set that indicates connections between the vertices. A *self-loop* is an edge from a vertex to itself. *Simple graphs* are graphs with no self-loops, and *at most* one edge between two vertices. Figure 1.1 illustrates a few special examples. For instance, a *tree* is a graph with no cycles: there is *exactly one* path from any vertex to any other vertex. On the other hand, a *clique* has an edge between every vertex. Figure 1.1 showcases *simple undirected graphs*, that is, the edges have no orientation (or direction). This thesis does not cover graphs that are directed, nor does it consider graphs that can have multiple edges between two vertices.

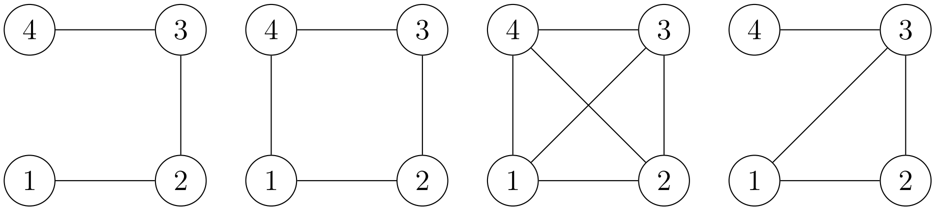


Figure 1.1: Some graphs on 4 vertices. The first three graphs are a tree, *a* cycle, *a* clique *respectively.*

A graph can be represented through its *adjacency matrix*. Let $G := (V, E)$ be the graph, with V being the vertex set and E the edge set. The adjacency matrix of G is defined as the matrix A with entries

$$A(i,j) := \begin{cases} 1\,, & \text{if } (i,j) \in E, \\ 0\,, & \text{if } (i,j) \notin E, \end{cases}$$

for all $i, j \in V$. For example, the cyclic graph in Figure 1.1 has the representation $G = (\{1,2,3,4\}, \{(1,2),(2,3),(3,4),(1,4)\})$. The corresponding adjacency matrix is

$$A = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}.$$

We notice that A is symmetric. In fact, all undirected graphs have symmetric adjacency matrices, that is, $A(i,j) = A(j,i)$ for all $i, j \in V$. Moreover, A is zero on the diagonal, since G has no self-loops.

Another important graph matrix is the *graph Laplacian*. Let D denote the diagonal matrix with entries

$$D(i,j) = \begin{cases} \sum_{k \neq i, k \in V} A(i,k)\,, & \text{if } i = j, \\ 0\,, & \text{if } i \neq j. \end{cases}$$

The *combinatorial graph Laplacian* L is defined as $\mathrm{L} := A - D$. The *normalised Laplacian* is defined as $\mathcal{L} = \mathrm{I} - D^{-1/2} A D^{-1/2}$, where $D^{-1/2}$ is the diagonal matrix defined as

$$D^{-1/2}(i,i) := \begin{cases} \frac{1}{\sqrt{D(i,i)}} & \text{if } D(i,i) \neq 0, \\ 0 & \text{if } D(i,i) = 0. \end{cases}$$

This thesis only covers the combinatorial graph Laplacian, which henceforth will be referred to as the Laplacian matrix. Note that if A is symmetric, then so are the graph Laplacians.

The Laplacian matrix gets its name from the fact that it can be viewed as the matrix form of the *discrete Laplacian operator*, which approximates the *continuous Laplacian operator* through a finite difference method (LeVeque [2007]). This can be illustrated by the discrete heat equation as follows: Let ϕ be a distribution across a graph $G = (V, E)$, with $\phi(i)$ being the temperature at a vertex $i \in V$. If $(i, j) \in E$, then the heat transfer between i and j is proportional to $\phi(i) - \phi(j)$. In particular, one obtains a matrix-vector ordinary differential equation of the form

$$\frac{\mathrm{d}\,\phi}{\mathrm{d}\,t} = k\mathrm{L}\phi\,, \tag{1.1}$$

where $\mathrm{L} = A - D$ is the graph Laplacian matrix and k is the thermal conductivity. This is analogous to the classical heat equation, hence the name "graph Laplacian". The solution to (1.1) and its stability properties are obtained by analysing the eigenvalues of L.

§1.2.2 Spectral theory

Spectral theory traces its origins back to the works of David Hilbert in the early 20^{th} century. He referred to the theory as *spectral analysis*. The name proved prophetic: a key result in the field, known as the *spectral theorem*, was later found to be useful in explaining atomic spectra in quantum mechanics.

Finite undirected graphs have adjacency (and Laplacian) matrices that are symmetric, which are diagonalisable and have real eigenvalues. For a finite graph G on N vertices, the spectrum of its adjacency matrix A (or its Laplacian matrix L) is the set of eigenvalues of A (or L, respectively). *Linear algebra* provides the necessary framework to study the eigenvalues and eigenvectors of graph matrices. For example, for a clique on N vertices, the spectrum of A is the eigenvalue $N - 1$ with multiplicity 1 and the eigenvalue -1 with multiplicity $N - 1$. The Laplacian, on the other hand, has eigenvalues 0 (once) and N (with multiplicity $N - 1$). Another powerful result, the *Perron-Frobenius theorem*, states that if G is connected, then the following is true (Spielman [2012, Theorem 4.5.1]): For the eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_N$ of the adjacency matrix of G, we have

- $\lambda_1 > \lambda_2$, and $\lambda_1 \geq -\lambda_N$, and
- λ_1 has a strictly positive eigenvector.

Although linear algebra provides a solid foundation, we outgrow it as N grows to infinity, and lean on spectral theory for our analysis. A *locally finite graph* $G = (V, E)$, that is, a graph where every vertex has finite degree, can be identified with a linear operator A on the Hilbert space $(\ell^2(V), \langle\cdot,\cdot\rangle)$, where $\langle\cdot,\cdot\rangle$ is the canonical inner product $\langle\phi,\psi\rangle := \sum_{i\in V} \overline{\psi(i)}\phi(i)$, and A acts on the canonical basis $(\delta_i)_{i\in V}$ as

$$A\delta_i = \sum_{j:(i,j)\in E} \delta_j \,.$$

We call the operator A the *adjacency operator*, and the Laplacian operator is defined similarly on $\ell^2(V)$. Note that in the finite-dimensional setting, we get back the graph matrices, and so we use the same notation for the matrix and the operator.

For an infinite graph, the spectrum of the adjacency operator is the set

$$\operatorname{spec}(A) = \{\lambda \in \mathbb{C} : A - \lambda\,\mathrm{I} \text{ is not invertible}\}\,,$$

and we define spec(L) in a similar fashion. The spectrum of an operator on a finite-dimensional space is the set of eigenvalues. However, it consists of more components in the infinite-dimensional setting, and may have no eigenvalues or no point spectrum. In some instances, for a locally finite undirected graph, the operators A and L are self-adjoint, and $\operatorname{spec}(A)$ and spec(L) are contained in $\mathbb{R}$.

If A (and L) are self-adjoint operators defined over the Hilbert space $\ell^2(V)$, then the spectral theorem guarantees that these operators are in some sense "diagonalisable". For example, consider the adjacency A. If A is a self-adjoint matrix, then the spectral theorem yields a *spectral decomposition* for A of the form

$$A = U\Lambda U^*\,,$$

where Λ is the diagonal matrix of eigenvalues of A, which are real, and U is a unitary matrix with columns as the orthonormal eigenvectors of A. If A is a self-adjoint operator, then we can still have a spectral decomposition in terms of a *spectral measure* (Rudin [1991]). The spectral theorem now gives us a *projection-valued measure* Λ on the spectrum $\operatorname{spec}(A) \subset \mathbb{R}$ such that

$$A = \int_{\operatorname{spec}(A)} \lambda \,\mathrm{d}\,\Lambda(\lambda)\,.$$

This formulation will later allow us to associate a probability measure to a certain class of "nice" operators.

Let us go back to a finite simple graph G. The largest eigenvalue λ_1 of the adjacency A is bounded above by the maximal degree of G. Additionally, λ_1

plays another crucial role in the study of the spread of epidemics on a graph. Following the ideas of Pastor-Satorras et al. [2015], we consider the following illustration.

Consider a community with (finite) vertex set V of size N and edge set E. Individuals are either *infected* or *susceptible*. Once an individual recovers, it becomes susceptible again. Any susceptible individual $x \in V$ gets infected at a rate β, and any infected individual recovers at a rate δ. If $p_x(t)$ denotes the probability that x is infected at time t, then,

$$p_x(t+\Delta t) = p_x(t)(1-\delta\Delta t) + (1-p_x(t))\beta\Delta t \sum_{y\in V} A(x,y)p_y(t)\,.$$

Using this reasoning, one can derive an ODE for the dynamics of the spread as

$$\frac{\mathrm{d}\,p_x(t)}{\mathrm{d}\,t} = -\delta p_x(t) + \beta \sum_{y\in V} A(x,y)(1-p_x(t))p_y(t)\,.$$

To simplify the ODE, we can perform a linearisation trick by taking $1-p_x(t) \approx 1$. This step is justified heuristically when the $p_x(t)$ is small for any x, that is, the epidemic spread is in the early stage. Let $p(t) = (p_1(t),\dots,p_N(t))^{\mathrm{T}}$ be the vector of probabilities. We obtain a system of linearised ODEs given by

$$\frac{\mathrm{d}\,p(t)}{\mathrm{d}\,t} = (\beta A - \delta)p(t)\,.$$

Spectral analysis of the solution tells us that the equilibrium state is stable if

$$\frac{\beta}{\delta} < \frac{1}{\lambda_1}\,,$$

that is, the infection dies out below this threshold. Heuristically, a large λ_1 indicates that nodes with many connections aid the spread of the disease.

The adjacency spectra have further applications. For instance, if the graph G has d_{max} as the maximal degree and d_{av} as the average degree, and λ_1 is the largest eigenvalue of the adjacency of G, then, by Spielman [2012, Lemma 4.2.1], we have

$$d_{av} \leq \lambda_1 \leq d_{max}\,.$$

If λ_1 and λ_N are the extremal eigenvalues, and if G is connected, then $\lambda_1 = -\lambda_N$ if and only if G is bipartite (Spielman [2012, Proposition 4.5.3]). Moreover, if $\alpha(G)$ is the chromatic number of a $k-$regular graph G on N vertices, then

$$\alpha(G) \leq \frac{-N\lambda_N}{k-\lambda_N}\,.$$

The above inequality is known as the Hoffman bound (Haemers [2021]). These results show that eigenvalues of the adjacency matrix can be used to study various properties of the graph.

The spectrum of the Laplacian can provide further insight into the graph structure. When the entries of the matrix are not restricted to 0 or 1, the matrix is also referred to as the *Markov matrix* (Bryc et al. [2006], Bordenave et al. [2014]). The graph Laplacian is essential in diffusion theory and network flow analysis, as it can be seen as the negative of the infinitesimal generator of a continuous-time random walk associated with a graph, and its spectral properties are useful in the analysis of mixing times and relaxation times of the random walk. It has several other key applications. The Kirchhoff Matrix-Tree Theorem relates the determinant of the Laplacian to the count of spanning trees in a graph (Chung [1997]), and the multiplicity of the zero eigenvalue indicates the number of connected components (Chung [1997]). The second-smallest eigenvalue, known as the Fiedler value or the algebraic connectivity, measures the graph's connectivity; higher values signify stronger connectivity De Abreu [2007].

In modern machine learning, spectral techniques are pivotal in spectral clustering algorithms, where the techniques use the Laplacian eigenvalues and eigenvectors for dimensionality reduction before applying algorithms like k-means clustering (see Abbe et al. [2020], Abbe [2017]). These algorithms are particularly effective for detecting clusters that are not linearly separable. Recent advancements integrate spectral clustering with graph neural networks to enhance graph pooling operations (Bianchi et al. [2020]). Spectral algorithms are also crucial for identifying communities within networks by analysing the spectral properties of the graph (Chung [1997]).

The normalised graph Laplacian, just like the graph Laplacian, is the negative of the infinitesimal generator of another continuous-time random walk associated with the graph. It has applications in studying the so-called Cheeger constant, as well as the diameter of the graph, but we will not be studying this matrix in this thesis.

§1.3 Random Graphs

How likely is it that you and another individual have a mutual friend? Will a disease spread in a community rapidly, or will it be restricted to isolated groups? How does one model social networks? Where are you likely to be if you walk randomly on the streets of Amsterdam? How likely are oil particles to percolate through a rock?

These questions only begin to scratch the surface of *random graph theory*. Random graphs first appeared in the context of sociology in the early 1900s. They reappeared in the context of mathematical biology, before the pioneering works of Paul Erdős and Alfred Rényi in 1959, which laid the foundation of the most elementary random graph model: The Erdős-Rényi random graph (ERRG). Thereafter, the interest in the topic grew rapidly, fuelled by the boom of computer science and the increasing interest in modelling complex networks.

§1.3.1 Erdős-Rényi random graphs

There are two models typically referred to as the Erdős-Rényi random graph. The first, introduced in Erdős and Rényi [1959], is a simple graph chosen *uniformly at random* from the set of all graphs on N vertices and m edges, and is parametrised by the tuple (N, m). The second model, also called the *Gilbert-Erdős-Rényi* model, was introduced in Gilbert [1959] as a *percolation model* on the complete graph K_N on N vertices, where edges are kept with probability p and discarded with probability $1-p$, for some $p \in [0,1]$, and is parametrised by the tuple (N, p). The two models are quite close. The latter will be the model used throughout this thesis and will be denoted as $\mathrm{ER}_N(p)$ and abbreviated as ERRG. There are various texts on random graphs, and in particular on the ERRG. We refer to the monographs van der Hofstad [2017], van der Hofstad [2024] for an exposition of the topic.

In the setting where $p := \lambda/N$, the random graph is usually classified into three regimes:

- *Subcritical regime:* When $\lambda < 1$, the graph consists of small components that are tree-like. In particular, the graph is a *forest*, with the size of the largest component of the order $\mathrm{O}_{N,\mathbb{P}}(\log N)$ (where O_N is the Landau notation, and the additional subscript $\mathbb{P}$ indicates the statement holds with high probability).

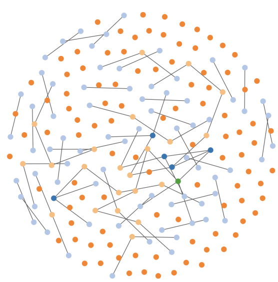
- *Critical regime:* When $\lambda = 1$, the graph exhibits a so-called *phase transition*. The largest component in the graph is now $\mathrm{O}_{N,\mathbb{P}}(N^{2/3})$. This regime is the most delicate of the three, and we refer to Janson et al. [1993], Aldous [1997] for further analysis.

- *Supercritical regime:* When $\lambda > 1$, the graph has a *unique giant component* of size $\mathrm{O}_N(N)$ with high probability, and other components of size $\mathrm{O}_N(\log N)$. We have a further sub-regime in this regime:
  - *Connectivity regime:* If $p \gg \frac{\log N}{N}$, then with high probability the graph is *connected*, that is, there is only one component.

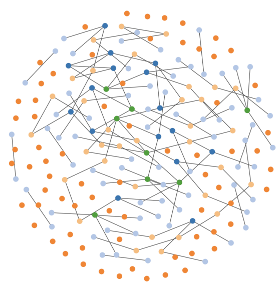
In Figure 1.2, we see graph realisations for the three regimes.

Another classification for random graph models is *sparsity*. In particular, for this thesis, we say that the graph is *sparse* when the average degree of the graph is bounded. On the other hand, we say that the graph is *dense* if the average degree grows with N. For $\mathrm{ER}_N(p)$, we have the following:

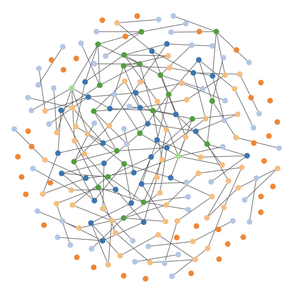
- *Dense regime:* $p := \varepsilon_N$ such that $\varepsilon_N \to 0$ and $N\varepsilon_N \to \infty$. In the literature, the dense regime is characterised by $\varepsilon_N \equiv$ constant, but this regime will not be covered in this thesis, and hence we abuse terminology.
- *Sparse regime:* $p := \varepsilon_N$ such that $\varepsilon_N \to 0$ and $N\varepsilon_N \to \lambda \in (0, \infty)$.



(a) $\lambda = 0.7$.



(b) $\lambda = 1$.



(c) $\lambda = 1.8$.

Figure 1.2: Realisations of $\mathrm{ER}_N(p)$ *for three regimes, with* $p = \lambda/N$*, and* $N = 200$*. Simulated on https://www.networkpages.nl.*

Local weak convergence

The theory of local weak convergence builds on Aldous and Lyons [2007], Benjamini and Schramm [2001]. Since random graphs are essentially graph-valued random variables, this theory describes a framework to analyse the "limits" of sparse random graphs, by providing a natural topology to understand convergence. Consequently, any graph parameter that is continuous with respect to the local topology converges to the graph parameter of the limiting object, akin to how several functionals of random walks converge to functionals of Brownian motion in the appropriate topology. Local weak convergence is a remarkable tool, since in many instances the limit is easier to analyse than the prelimit. Before looking at formal details, we give a heuristic description:

Consider a uniformly chosen vertex of the random graph, say o_N*, where* N *is the number of vertices. If the graph has no underlying geometry, as is the case for* $\mathrm{ER}_N(\cdot)$*, we say that edges have length one. Fix a positive radius* $r \in \mathbb{R}_+$*, and from the vertex* o_N*, observe the graph up to the radius* r*. So, if* $r \in [0, 1)$*, observe only* o_N *and nothing more, if* $r \in [1, 2)$*, observe the immediate neighbours of*

o_N, and so on. The local weak limit *is in some sense what the graph "looks like" up to any finite radius r.*

We now state the above formally. A *rooted graph* $(\mathbb{G}, o)$ is a graph $\mathbb{G}$ with a specified root o. Let $\mathcal{G}^*$ denote the set of *locally finite connected rooted graphs* up to equivalence $\equiv$, where $\equiv$ denotes graph isomorphism. Given $r \in \mathbb{N}$, let $[\mathbb{G}, o]_r$ denote the finite rooted subgraph obtained from $(\mathbb{G}, o)$ by keeping vertices that are up to a distance r from o, including edges. We say that a sequence $(\mathbb{G}_N, o_N)_{N\in\mathbb{N}}$ converges *locally* to $(\mathbb{G}, o)$ if for each $r \in \mathbb{N}$ there exists an $n_r \in \mathbb{N}$ such that for all $N > n_r$, we have

$$[\mathbb{G}_N, o_N]_r \equiv [\mathbb{G}, o]_r \,.$$

If we define d_{LW} as

$$d_{LW} : (\mathbb{G}, o), (\mathbb{G}', o') \mapsto 1/\sup\{r \in \mathbb{N} : [\mathbb{G}, o]_r \equiv [\mathbb{G}', o']_r\},$$

then $(\mathcal{G}^*, d_{LW})$ becomes a complete separable metric space. We can endow this space with its Borel $\sigma-$algebra, and consider the complete separable metric space of probability measures $\mathcal{P}(\mathcal{G}^*)$ on $\mathcal{G}^*$.

Definition 1.3.1 (Local weak convergence).
Let $(\mathbb{G}_N)_{N\geq 1}$ denote a sequence of (possibly disconnected) random graphs. If o_N is a uniformly chosen vertex (restricted to the connected component of $\mathbb{G}_N$), then we say that $(\mathbb{G}_N, o_N)$ converges locally weakly *to $(\mathbb{G}, o)$ having law $\mathcal{L} \in \mathcal{P}(\mathcal{G}^*)$ if, for any bounded and continuous function $h : \mathcal{G}^* \to \mathbb{R}$, we have*

$$\mathbb{E}[h(\mathbb{G}_N, o_N)] \to \mathbb{E}_{\mathcal{L}}[h(\mathbb{G}, o)]$$

as $N \to \infty$, where $\mathbb{E}$ is with respect to the law of the random graph and the root o_N.

As an example, consider the graph $\mathrm{ER}_N(p)$, with $p = \lambda/N$ for a fixed λ. This graph converges locally weakly to a *Galton-Watson tree* (or a branching process) with offspring distribution $\mathrm{Poi}(\lambda)$, that is, a process starting with a single vertex, giving birth to progeny that are distributed as $\mathrm{Poi}(\lambda)$, and repeating this for each offspring.

Local weak convergence provides a powerful framework for the analysis of graph properties that are *local*, that is, continuous with respect to the local topology (see van der Hofstad [2024], Salez [2011]), as well as dynamics on the graph (see for example Avena et al. [2024] for *interacting particle systems*, and Hupkes et al. [2023] for a discussion on PDEs). Later, we will see how local weak convergence can be related to the spectrum of random graphs via the Stieltjes transform.

Inhomogeneous Erdős-Rényi random graphs
ERRG serve as the basis for many mathematical theories in random graphs. Real-world networks are highly inhomogeneous and have a far more complex structure. Various attempts have been made to generalise them to other kinds of random graph models. One of the successful extensions is the inhomogeneous Erdős-Rényi random graph model introduced by Bollobás et al. [2007]. This graph has N vertices labelled by $[N] = 1, ..., N$, and edges are present independently with probability p_{ij} given by $p_{ij} = \frac{f(x_i, x_j)}{N} \wedge 1$, where f is a symmetric kernel on a state space $S \times S$, and x_i are certain attributes associated with vertex i belonging to S. If f is bounded, then the graph is a sparse random graph. In this thesis, we study a variant of the above inhomogeneous random graph, namely, the vertex set remains the same, but the connection probabilities are given by

$$p_{ij} = \varepsilon_N f(w_i, w_j) \wedge 1\,,$$

where ε_N is a tuning parameter, (w_i) is a sequence of deterministic weights, and f is a symmetric bounded function on $[0, \infty)^2$. The weights can signify a property of vertex i. They can also be taken random, but are not considered to be so in this thesis for this model. Note that when $N\varepsilon_N \to \infty$, the average degree is unbounded, and when $N\varepsilon_N = \mathrm{O}(1)$, the average degree is bounded. In the sparse case, the properties of the connected components were studied in Bollobás et al. [2007], which focused on the properties of the connected components and their relationship with the branching process. It was shown that the largest component of the graph has a size of order N if the operator norm of the kernel operator corresponding to f is strictly greater than 1 (see also [van der Hofstad, 2024, Theorem 3.9]). In the subcritical case, the sizes of the largest connected components can exhibit different behaviour compared to the ERRG. The study of the largest connected components in various inhomogeneous random graphs has attracted a lot of attention (see, for example, Bhamidi et al. [2010], Broutin et al. [2021], Bet et al. [2023], van der Hofstad [2013], Devroye and Fraiman [2014]). We abbreviate the inhomogeneous Erdős-Rényi random graph as IER.

§1.3.2 Kernel-based random graphs

In recent years, many random graph models have been proposed in an attempt to model real-life networks. These models aim to capture three key properties that real-world networks exhibit: scale-free nature of the degree distribution, small-world property, and high clustering coefficients [van der Hofstad, 2024]. It is generally difficult to find random graph models that incorporate all three features. Classical random graph models typically fail to capture scale-freeness, small-world behaviour, and high clustering simultaneously. For instance, the Erdős-Rényi model only exhibits the small-world property, while models like Chung-Lu, Norros-Reittu, and preferential attachment models are scale-free (Chung and Lu [2002], Barabási and Albert [1999]) and small-world, but have clustering coefficients that vanish as the network grows. In contrast, regular lattices have high clustering but large typical distances. The Watts-Strogatz model (Watts and Strogatz [1998]) was an early attempt to create a network with high clustering and small-world features, but it does not produce scale-free degree distributions.

Scale-free percolation, introduced in Deijfen et al. [2013], blends ideas from long-range percolation (see e.g. Berger [2002]) with inhomogeneous random graphs, such as the Norros–Reittu model. In this framework, vertices are positioned on $\mathbb{Z}^d$, and each vertex x is independently assigned a random weight W_x. These weights follow a power-law distribution:

$$\mathbb{P}(W > w) = w^{-(\tau-1)} L(w),$$

where $\tau > 1$ and $L(w)$ is a slowly varying function at infinity.

Edges between pairs of vertices x and y are added independently, with a probability that increases with the product of their weights and decreases with their Euclidean distance. The edge probability is given by

$$p_{xy} = 1 - \exp\left(-\lambda \frac{W_x W_y}{\|x-y\|^\alpha}\right), \tag{1.2}$$

where $\lambda, \alpha > 0$ are model parameters and $\|\cdot\|$ denotes the Euclidean norm. This model has been proposed as a suitable representation for certain real-world systems, such as interbank networks, where both spatial structure and heavy-tailed connectivity distributions are relevant (Deprez et al. [2015]). Various properties of the model are now well known, and we refer to the articles by Jorritsma et al. [2024], Cipriani and Salvi [2024], Cipriani et al. [2025], Heydenreich et al. [2017], Dalmau and Salvi [2021], van der Hofstad et al. [2024] for further references.

In recent times, there has been a lot of interest in models that have connection probabilities similar to (1.2). Kernel-based spatial random graphs encompass a wide variety of classical random graph models where vertices are embedded in some metric space. In their simplest form (see Jorritsma et al. [2023] for a more complete exposition) they can be defined as follows: Let V be the vertex set of the graph and, sample a collection of weights $(W_i)_{i\in V}$ that are independent and identically distributed (i.i.d.), serving as marks on the vertices. Conditionally on the weights, two vertices i and j are connected by an undirected edge with probability

$$\mathbb{P}\left(i \leftrightarrow j \mid W_i, W_j\right) = \kappa(W_i, W_j)\|i-j\|^{-\alpha} \wedge 1\,,$$

where κ is a symmetric kernel, $\|i-j\|$ denotes the distance between vertices i and j in the underlying metric space and $\alpha > 0$ is a constant.

Common choices for κ include:

$$\kappa_{\text{triv}}(w,\,v) \equiv 1, \qquad \kappa_{\text{strong}}(w,\,v) = w \vee v,$$
$$\kappa_{\text{prod}}(w, v) = w\,v, \qquad \kappa_{\text{pa}}(w, v) = (w \vee v)(w \wedge v)^{\sigma_{\text{pa}}}.$$

In the above, $\sigma_{\text{pa}} = \alpha(\tau-1)/d - 1$, where $\tau - 1$ is the exponent of the tail distribution of the weights, so that the kernel κ_{pa} mimics the form that appears in preferential attachment models [Jorritsma et al., 2023]. While these models are well-studied from a random graph perspective, there is minimal literature on their spectral properties.

§1.4 Random Matrix Theory and Free Probability

§1.4.1 Random Matrices

Random matrices are matrix-valued random variables where each entry of the matrix is a classical random variable. They are of significant interest, not only from the point of view of modern probability and statistical physics, but also because they connect to various areas. First appearing in 1928 in the work of Wishart (Wishart [1928]) in the context of statistics and multivariate data analysis, the topic was further researched from a spectral analysis point of view in the pioneering work of Wigner (Wigner [1955]). There are now several connections with other areas. For instance, a connection with *number theory* was established when eigenvalues of certain random matrices were used to model the distribution of zeroes of the Riemann zeta function (Montgomery [1973]). There are also connections with dynamical systems, in particular with Painleve's ordinary differential equations (Tracy and Widom [1994]), as well as with the Dyson Brownian motion (Dyson [1962]). There are several applications in numerical linear algebra, computer science, and statistics (see Johnstone [2001], or the textbook Tropp [2015]).

Quantum mechanics tells us that energy levels of large nuclei correspond to the eigenvalues of some Hermitian operator. Wigner chose to model this operator by using *Wigner matrix ensembles*, wherein he ignored all physical aspects of the system except symmetry. The reason to do so was the observation that gaps in energy levels of large nuclei followed similar patterns regardless of the material chosen. Systems with time-reversal symmetry were modelled by using real symmetric random matrices with Gaussian entries, known as the *Gaussian orthogonal ensemble* (GOE), and those without were modelled by using complex Hermitian matrices with complex Gaussian entries, known as the *Gaussian unitary ensemble* (GUE).

Spectral analysis of random matrices is a broad subject, with a vast literature focusing on the distribution of eigenvalues of the matrix, the largest eigenvalue (or more generally, the k largest eigenvalues for some $k \in \mathbb{N}$), and the eigenvectors. One of the key statistics that we focus on in this thesis is the empirical spectral distribution, defined below.

Definition 1.4.1 (Empirical Spectral Distribution).
The empirical measure that assigns mass $1/N$ to each eigenvalue of random matrix $\mathbf{M}_N$ is called the Empirical Spectral Distribution (ESD)*, and is defined as*

$$\mathrm{ESD}(\mathbf{M}_N)(\cdot) = \frac{1}{N}\sum_{i=1}^{N} \delta_{\lambda_i}(\cdot)\,,$$

where $\lambda_i := \lambda_i(\mathbf{M}_N)$ is the $i-$th eigenvalue of $\mathbf{M}_N$ and, for any x, $\delta_x(\cdot)$ is the Dirac delta mass at the value x.

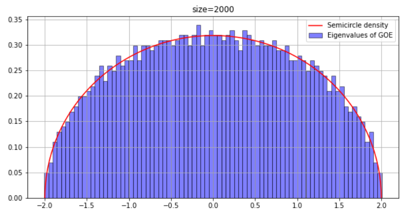
Notice that since the eigenvalues are random, $\mathrm{ESD}(\mathbf{M}_N)$ is a *random measure*. The *bulk distribution* of eigenvalues refers to the distribution of the non-extremal eigenvalues of the random matrix. The ESD is a central object of interest in studying the bulk of the eigenvalue distribution, so it would be heresy not to ask about its limiting behaviour as $N \to \infty$. The work of Wigner [1958] showed that for the GOE and GUE models, as well as for a large class of other random matrix models, under appropriate scaling of the entries the $\mathrm{ESD}(\mathbf{M}_N)$ converges *weakly almost-surely* to a (deterministic) measure μ_{sc}, where μ_{sc} is the semicircle law with density

$$\mu_{sc}(\mathrm{d}\,x) := \frac{1}{2\pi}\sqrt{4 - x^2}\,\mathbf{1}_{|x|\le 2}\,\mathrm{d}\,x\,.$$

For instance, consider the symmetric matrix $\mathbf{A}_N$ with entries

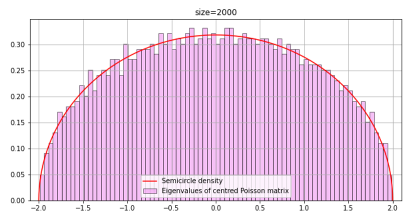
$$\mathbf{A}_N(i,j) \overset{d}{=} \frac{1}{\sqrt{p_N(1-p_N)}}\,\mathrm{Ber}(\varepsilon_N)\,,$$

such that $N\varepsilon_N \to \infty$. This is the adjacency matrix of the Erdős-Rényi random graph $\mathrm{ER}_N(\varepsilon_N)$. It is known that, as $N \to \infty$, $\mathrm{ESD}(\mathbf{A}_N)$ converges to μ_{sc} in probability (see for example Jung and Lee [2018], Tran et al. [2013]). The fact that different empirical distributions converge to the same limit sparked the idea of *universality*, and we will later see how μ_{sc} is a universal limit in this area.



(a) GOE *ensemble, with entries distributed as* $\mathrm{N}(0,1)$.

(b) Symmetric and centred Bernoulli matrix, with entries distributed as $\mathrm{Ber}(0.5) - 0.5$.



(c) Symmetric and centred Poisson matrix, with entries distributed as $\mathrm{Poi}(1) - 1$.

Figure 1.3: Eigenvalue distributions of some random matrix models.

The above has been generalised beyond the original Wigner matrices. In particular, some works consider matrix entries that are not i.i.d. and have a variance profile that is not constant. The study of the bulk blends flavours from various areas of mathematics. In favourable scenarios, the problem can be analysed within the framework of universality classes. Typically, in these cases, the matrix may have entries drawn from any distribution, but one can implement *Gaussianisation*, that is, replace them with Gaussian entries with the same mean and variance profile, without affecting the ESD too much (in probability). This technique follows ideas of Chatterjee [2005].

Of course, it would be naïve to assume that every random matrix model can be Gaussianised. There are numerous concrete examples, particularly in random graph theory, where this step fails. In such cases, one falls back on explicitly working with the ESD. One approach is through the *method of moments* (see Bordenave [2019]), which computes the moments of the ESD and find the limiting moments. By the spectral theorem, for any random matrix $\mathbf{M}_N$,

$$\int_{\mathbb{R}} x^k \,\mathrm{ESD}(\mathbf{M}_N)(\mathrm{d}\, x) = \frac{1}{N}\sum_{i=1}^{k} \lambda_i^k = \frac{1}{N}\,\mathrm{Tr}(\mathbf{M}_N^k)\,.$$

The "standard procedure" is to begin by computing expected moments and see if they concentrate, that is, the target is to show that

$$\lim_{N\to\infty}\int_{\mathbb{R}} x^k \,\mathrm{ESD}(\mathbf{M}_N)(\mathrm{d}\,x) = M_k \quad \text{in } \mathbb{P}\text{−probability}\,,$$

where $\mathbb{P}$ is the underlying law of the matrix entries. To guarantee the existence of a limiting measure, we have to check if the moments satisfy one of *Carleman's conditions*, that is, the moments uniquely determine a limiting measure if

$$\sum_{k\geq 0} M_{2k}^{-1/2k} = \infty\,, \quad \text{or equivalently,} \quad \limsup_{k\to\infty} M_{2k}^{1/2k} < \infty\,.$$

This approach has a strong combinatorial flavour due to computation of the combinatorial expression $\mathbb{E}[\mathrm{tr}(\mathbf{M}_N^k)]$, where $\mathrm{tr} := N^{-1}\,\mathrm{Tr}$ is the normalised trace. Naturally, there are examples where the moments do not exist, and this approach then fails. This brings us to another classical approach, namely, the *Stieltjes transform approach*, where one translates the measure-theoretic problem into an analytic problem on the upper-half complex plane $\mathbb{C}^+ := \{z \in \mathbb{C} : \Im(z) > 0\}$ (see Bordenave [2019], Mingo and Speicher [2017], Anderson et al. [2010]).

For any complex number $z \in \mathbb{C}^+$, we define the *resolvent* of a random matrix $\mathbf{M}_N$ as

$$\mathrm{R}_{\mathbf{M}_N}(z) := (\mathbf{M}_N - z\,\mathrm{I}_N)^{-1}\,,$$

where I_N is the $N \times N$ identity matrix. For any measure μ and $z \in \mathbb{C}^+$, its *Stieltjes transform* is defined as

$$\mathrm{S}_\mu(z) := \int_{\mathbb{R}} \frac{1}{x-z}\mu(\mathrm{d}\,x)\,.$$

So, we have that

$$\mathrm{S}_{\mathrm{ESD}(\mathbf{M}_N)}(z) = \int_{\mathbb{R}} \frac{1}{x-z}\,\mathrm{ESD}(\mathbf{M}_N)(\mathrm{d}\,x) = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{\lambda_i - z} = \mathrm{tr}(\mathrm{R}_{\mathbf{M}_N}(z))\,,$$

where $\mathrm{tr} := N^{-1}\,\mathrm{Tr}$ is the normalised trace operator. The Stieltjes transform allows us to work with analytic tools from complex analysis and functional analysis to deduce properties of the measure itself. For instance, if the Stieltjes transform is uniformly bounded (in z), then the measure has an absolutely continuous component (Sen and Virág [2011]). In some cases, we can also derive the exact density of the measure from its Stieltjes transform by using an inversion formula (Bai and Silverstein [2010]). Notable works show that one can bound the distance between two measures in terms of the distance between their respective Stieltjes transforms (Bai and Silverstein [2010], Augeri [2025]). In particular

Augeri [2025] defines a distance d_{S} compatible with the weak topology on the space $\mathcal{P}(\mathbb{R})$ of probability measures on the real line as follows:

$$d_{\mathrm{S}}(\mu,\nu) := \sup\{|\,\mathrm{S}_\mu(z) - \mathrm{S}_\nu(z)| : \Im(z) \geq 2, z \in \mathbb{C}^+\}, \quad \mu,\nu \in \mathcal{P}(\mathbb{R})\,.$$

If $d_{KL}(\cdot,\cdot)$ denotes the Kolmogorov-Smirnov distance and $\mathcal{W}^p(\cdot,\cdot)$ the L^p-Wasserstein distance for any $p \geq 1$, then

$$d_{\mathrm{S}}(\mu,\nu) \leq d_{KL}(\mu,\nu) \wedge \mathcal{W}^p(\mu,\nu), \quad \mu,\nu \in \mathcal{P}(\mathbb{R})\,.$$

If a measure μ is compactly supported in $[-R,R]$ for some $R>0$ with moments $\{M_k\}_{k\geq 1}$, then the Stieltjes transform can be related to the moments as:

$$\mathrm{S}_\mu(z) = -\sum_{k\geq 0} \frac{M_k}{z^{k+1}}\,, \tag{1.3}$$

where the Laurent series on the right-hand size converges when $|z|>R$.

§1.4.2 An illustration

Let us next see a heuristic for the two methods of analysis. We begin with the moment method, following ideas from Speicher [2024].

Consider an i.i.d. sequence $\{G_{i,j} : N \geq i > j\}$ of random variables distributed as $N^{-1/2}\mathrm{N}(0,1)$, where $\mathrm{N}(0,1)$ is the standard Gaussian random variable with mean 0 and variance 1. Take an $N \times N$ Wigner matrix G, with entries $\mathrm{G}(i,j) = G_{i\wedge j, i\vee j}$, where $G_{ii} = 0$. By trace expansion, we have

$$\mathbb{E}[\mathrm{tr}(\mathrm{G}^k)] = \frac{1}{N^{1+\frac{k}{2}}} \sum_{i_1,\ldots,i_k=1}^{N} \mathbb{E}[G(i_1,i_2)G(i_2,i_3)\ldots G(i_k,i_1)]\,.$$

To compute this sum, we use the following well-known result (see for example Speicher [2024]).

Lemma 1.4.2 (Wick's formula).
Let $(X_1,\ldots,X_n)$ be a real Gaussian vector and $\mathcal{P}_2(k)$ the set of pair partitions of $[k]$. Then, for any $1 \leq k \leq n$,

$$\mathbb{E}[X_{i_1}\cdots X_{i_k}] = \sum_{\pi\in\mathcal{P}_2(k)} \prod_{(r,s)\in\pi} \mathbb{E}[X_{i_r}X_{i_s}]\,, \tag{1.4}$$

where $(r,s)\in\pi$ indicates a pair (r,s) *that is in the pair partition π.*

This expression already tells us that the odd moments are identically zero, since one cannot construct pair partitions for a tuple $[k]$ if k is odd. So, we only need to compute the even moments. Thus, for any $k \in \mathbb{N}$, we have

$$\begin{aligned}\mathbb{E}[\mathrm{tr}(\mathrm{G}^{2k})] &= \frac{1}{N^{k+1}} \sum_{i_1,\ldots,i_{2k}=1}^{N} \sum_{\pi\in\mathcal{P}_2(2k)} \prod_{(r,s)\in\pi} \mathbb{E}[G(i_r, i_{r+1})G(i_s, i_{s+1})] \\ &= \frac{1}{N^{k+1}} \sum_{i_1,\ldots,i_{2k}=1}^{N} \sum_{\pi\in\mathcal{P}_2(2k)} \prod_{(r,s)\in\pi} \mathbf{1}_{(i_r,i_{r+1})=(i_s,i_{s+1})} \,.\end{aligned}$$

While there are two cases where the indicator is in force, namely $i_r = i_s$ or $i_r = i_{s+1}$, it turns out that the latter is the contributing factor in the limit. In particular, we get $i_r = i_{\pi(r)+1} = i_{\gamma\pi(r)}$, where $\gamma := (1, 2, \ldots, 2k)$ is the shift by 1 modulo $2k$ permutation and, for any partition π, $\gamma\pi$ is read as a composition of two permutations by reading π as a permutation. Thus, we have that $\mathbf{i} := \{i_1, \ldots, i_{2k}\}$ is constant on the cycles of $\gamma\pi$. We skip some technical steps, which involve the interchange of summands, and obtain the expression

$$\mathbb{E}[\mathrm{tr}(\mathrm{G}^{2k})] \sim \frac{1}{N^k} \sum_{\pi\in\mathcal{P}_2(2k)} N^{\#\gamma\pi} \,,$$

where $\#\gamma\pi$ is the number of blocks in $\gamma\pi$, and $\sim$ means asymptotic. The contributing partitions are the *non-crossing pair partitions* $NC_2(2k)$, where we have that, for any $\pi \in NC_2(2k)$, $\#\gamma\pi = k + 1$, and, for $\pi \in \mathcal{P}_2(2k) \setminus NC_2(2k)$, $\#\gamma\pi \leq k$. Figure 1.4 illustrates some partitions of $\{1, 2, 3, 4\}$, with π_1 and π_2 as non-crossing pair partitions. This combinatorial approach yields that, for any even moments, we have

$$\lim_{N\to\infty} \mathbb{E}[\mathrm{tr}(\mathrm{G}^{2k})] = |NC_2(2k)| = C_k \,,$$

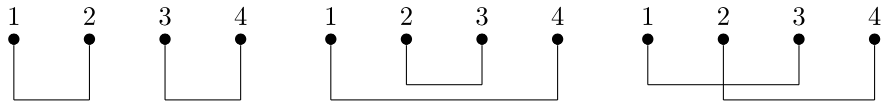
where C_k is the $k-$th Catalan number defined as

$$C_k = \frac{1}{k+1}\binom{2k}{k} \,.$$

The Catalan numbers are the even moments of the semicircle law μ_{sc}, which also has odd moments identically 0.

Let us proceed with the Stieltjes transform method. We begin by fixing $z \in \mathbb{C}^+$. Let μ_{sc} be the limiting measure of the ESD of G, which we a priori know is the semicircular law. To derive a recursive expression for $\mathrm{S}_{\mu_{sc}}(z)$, one can use the moment relation (1.3), along with the following relation for Catalan numbers:

$$C_{k+1} = \sum_{i=0}^{k} C_i C_{k-i}$$



(a) π_1 (b) π_2 (c) π_3

Figure 1.4: Pair partitions of $\{1, 2, 3, 4\}$.

for any $k \in \mathbb{N}$. Manipulating the terms, we get

$$\mathrm{S}_{\mu_{sc}}(z) = -\frac{1}{z + \mathrm{S}_{\mu_{sc}}(z)}\,, \tag{1.5}$$

which is the unique analytic equation that characterises μ_{sc} (Bai and Silverstein [2010]). It is known that pointwise convergence of Stieltjes transforms (in z) implies weak convergence of measures (and vice versa). To that end, we analyse the resolvent matrix.

Let $r_{ij} := \mathrm{R}_{\mathrm{G}}(z)(i, j)$. It is well-known that $r_{ij} \in \mathbb{C}^+$ for all $i, j \in [N]$ (Bai and Silverstein [2010]). One approach to prove that $\mathrm{S}_{\mathrm{ESD(G)}}(z)$ converges to $\mathrm{S}_{\mu_{sc}}(z)$ for each $z \in \mathbb{C}^+$ would be to use the resolvent identities (see Bordenave [2019]). However, there is a different approach that is used later on in the thesis. For any $z \in \mathbb{C}^+$, the following is a fact from complex analysis:

$$z = i\int_0^\infty \mathrm{e}^{-itz^{-1}}\,\mathrm{d}\,t\,.$$

So, we have for any $k \in [N]$

$$r_{kk} = i\int_0^\infty \mathrm{e}^{-itr_{kk}^{-1}}\,\mathrm{d}\,t\,.$$

From Bordenave [2019], we use the *Schur complement formula*, which gives us

$$r_{kk} = -\frac{1}{z + \sum_{i,j\neq k} \tilde{r}_{ij} G(i,k)G(j,k)}\,,$$

where $\tilde{r}_{ij}$ is the $(i, j)-$th entry of the matrix $\mathrm{R}_{\mathrm{G}^{(k)}}(z) := (\mathrm{G}^{(k)} - z\,\mathrm{I})^{-1}$, and $\mathrm{G}^{(k)}$ is the matrix G with the $k-$th row and column deleted. We will not spell out details here, but rather give a heuristic of what the computation looks like.

Namely,

$$
\begin{aligned}
\mathbb{E}[r_{kk}] &= i\mathbb{E}\left[\int_0^\infty \mathrm{e}^{itz}\mathrm{e}^{it\sum_{i,j\neq k}^N \tilde{r}_{ij}G(i,k)G(j,k)}\,\mathrm{d}\,t\right] \\
&\approx i\mathbb{E}\left[\int_0^\infty \mathrm{e}^{itz}\mathrm{e}^{it\sum_{j\neq k}^N \tilde{r}_{jj}G(j,k)^2}\,\mathrm{d}\,t\right] \\
&\approx i\int_0^\infty \mathrm{e}^{itz}\exp\left\{it\mathbb{E}\left[\sum_{j\neq k}^N r_{jj}G(j,k)^2\right]\right\}\mathrm{d}\,t \\
&\approx i\int_0^\infty \exp\left\{it(z+\mathbb{E}[\mathrm{tr}(\mathrm{R}_{\mathrm{G}}(z)])\right\}\mathrm{d}\,t = -\frac{1}{z+\mathbb{E}[\mathrm{tr}(\mathrm{R}_{\mathrm{G}}(z))]}\,,
\end{aligned}
$$

where each approximation requires justification, and becomes an equality in the limit $N \to \infty$. Summing over i on both sides, scaling by N and taking the limit $N \to \infty$ gives us (1.5) by the relation between the Stieltjes transform and the matrix resolvent. Using this approach has some advantages, particularly when dealing with random matrices with heavy-tailed entries (Benaych-Georges et al. [2014]).

Note that in both approaches, we only illustrate the convergence of the expected empirical measure. However, there are concentration results in both approaches that yield convergence in probability or almost surely (Bordenave [2019], Speicher [2024]).

Both approaches offer new insights into the problem as well as establish mysterious connections with numerous other areas, making random matrix theory an ideal playground for modern mathematics. The analytic approach has opened up the area of understanding local weak convergence, and many open problems were resolved in the last few years (see Erdős and Yau [2017] for more details).

§1.4.3 Free Probability

Consider two random variables X_1 *and* X_2*, where* X_1 *takes values* -1 *or* 1 *with probability* $\frac{1}{2}$*, and* X_2 *takes values 0 or 1* independently *with probabilities* $\frac{1}{3}$ *and* $\frac{2}{3}$*, respectively. Then, the* distribution *of the random variable* $Y = X_1X_2$ *is the same as that of* $\tilde{Y} = X_2X_1$*, and we write* $X_1X_2 \overset{d}{=} X_2X_1$*. Indeed, the probability that* X_1X_2 *takes a value, for example 1, is the same as the probability that* X_2X_1 *takes that value, which in our example is* $\frac{1}{2} \times \frac{2}{3} = \frac{1}{3}$*. We say that* X_1 *and* X_2 commute*. We ask out of curiosity:*

Are there instances when X_1 *and* X_2 do not *commute? For instance, what if* X_1 *and* X_2 *are not real-valued, but are matrix-valued?*

Classical probability studies random variables that commute, and a crucial concept in classical probability is that of *independence*: The outcome of X_1 does not affect X_2 and vice-versa. How do we abstract to a non-commutative setting? Does the concept of independence extend as well?

In the 1980s, the concept of *freeness*, or *free independence*, was studied by Dan Voiculescu in the context of operator algebras (Voiculescu [1985]). The generalisation of classic random variables to a non-commutative setting was through this very notion, which is a non-commutative analogue of (classical) independence. The combinatorial aspects are summarised in the classical text by Nica and Speicher [2006]. We now begin with some technical definitions.

Definition 1.4.3 (Non-commutative probability space).
A non-commutative probability space $(\mathcal{A}, \varphi)$ *consists of a unital (associative) algebra* $\mathcal{A}$ *over* $\mathbb{C}$ *equipped with a linear functional* $\varphi : \mathcal{A} \to \mathbb{C}$ *such that* $\varphi(1) = 1$.

Let us fix an index set I. Elements of the space $(\mathcal{A}, \varphi)$ are called *non-commutative random variables*, and for any $a \in \mathcal{A}$, $\{\varphi(a^n)\}_{n\in\mathbb{N}}$ are the *moments* of a. The *joint distribution* of $a_1, \dots, a_k \in \mathcal{A}$ for any $k \in \mathbb{N}$ is the collection of mixed moments $\varphi(a_{i_1}, \dots, a_{i_\ell})$ for each $\ell \in \mathbb{N}$ and $i_1, \dots, i_\ell \in [k]$.

Definition 1.4.4 (Freeness).
Let $(\mathcal{A}_i)_{i\in I}$ *be the unital subalgebras of* $\mathcal{A}$. *These are said to be* free *if, for any* $k \in \mathbb{N}$, $\varphi(a_1 \dots a_k) = 0$ *whenever:*

- *For* $a_j \in \mathcal{A}_{i_j}$ *with* $i_j \in I$, $\varphi(a_j) = 0$ *for all* $j \in [k]$;
- $i_1 \neq i_2, i_2 \neq i_3, \dots, i_{k-1} \neq i_1$.

Recall that in classical probability theory, one studies random variables over a (classical) probability space of the form $(\Omega, \mathcal{F}, \mathbb{P})$. The generalisation to the non-commutative setting deviates from the notion of an underlying event space and law, and is instead developed over the notion of a non-commutative algebra of random variables and their "expectations". In fact, the functional φ is the non-commutative analogue of the classical notion of expectation. Similar to how classical random variables $(X_i)_{i\in I} \in (\Omega, \mathcal{F}, \mathbb{P})$ are said to be independent if the sigma-fields $(\mathcal{F}_i)_{i\in I}$ generated by them are independent, we say that random variables $(a_i)_{i\in I} \in (\mathcal{A}, \varphi)$ are said to be free if their generated unital subalgebras $(\mathcal{A}_i)_{i\in I}$ are free. This abstraction allows one to study a larger variety of objects, such as random matrices or random operators, as well as objects in other areas, notably in quantum mechanics.

Recovering classical probability is fairly straightforward, and we illustrate it for bounded random variables as follows: Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a classical probability space. Set $\mathcal{A} := L^\infty(\Omega, \mathbb{P})$ as the unital algebra of bounded measurable functions ("random variables") $X : \Omega \to \mathbb{C}$, and set φ to be the unital linear functional on $\mathcal{A}$ as the "expectation" with respect to $\mathbb{P}$, that is,

$$\varphi(X) := \mathbb{E}[X], \quad X \in \mathcal{A}.$$

Note that $\varphi(1)$ corresponds to $\mathbb{P}(\Omega) = 1$.

Voiculescu, in Voiculescu [1985], studied the notion of freeness in the context of *Von-Neumann algebras* (also called W^*-algebras). In particular, if G is *group*, then saying that its subgroups $(G_i)_{i\in I}$ are free is equivalent to saying that the subalgebras $(\mathbb{C}G_i)_{i\in I}$ are free in the space $(\mathbb{C}G, \varphi_G)$ (Speicher [2011, Proposition 1.3]), where $\mathbb{C}G$ is the group algebra of G and $\varphi_G : \mathbb{C}G \to \mathbb{C}$ is a unital functional. Certain $*-$algebras, in particular C^*-algebras, are of particular interest.

Definition 1.4.5 (C^*-algebras).
A C^*-algebra is a Banach algebra $\mathcal{A}$ *over* $\mathbb{C}$ *such that it is a $^*-$algebra possessing the involution $* : \mathcal{A} \to \mathcal{A}$ satisfying $\|xx^*\| = \|x\|^2$ for each $x \in \mathcal{A}$.*

Any C^*-algebra is isomorphic to a C^*-subalgebra of $\mathcal{B}(\mathcal{H})$, the space of bounded linear operators on $\mathcal{H}$, for some Hilbert space $\mathcal{H}$.

Definition 1.4.6 (W^*-algebras).
A W^*-algebra, or a von Neumann algebra $\mathcal{A} \subseteq \mathcal{B}(\mathcal{H})$*, is a C^*-algebra that is closed under the weak operator topology, that is, if any net $A_\alpha \in \mathcal{A}$ converges to $A \in \mathcal{B}(\mathcal{H})$ in the weak operator topology, then $A \in \mathcal{A}$.*

The respective non-commutative probability spaces are called a C^*-probability space and a W^*-probability space.

Definition 1.4.7 (Tracial, state, and faithful functionals).
Let $(\mathcal{A}, \varphi)$ be a C^*-probability space.

- *If $\varphi(a^*a) \geq 0$ for all $a \in \mathcal{A}$, then φ is a* state.
- *We say φ is* tracial *if $\varphi(ab) = \varphi(ba)$ for all $a, b \in \mathcal{A}$.*
- *We say φ is* faithful *if for all $a \in \mathcal{A}$, $\varphi(a^*a) = 0$ implies $a = 0$.*

Naturally, one would wonder if there are universality results in free probability, as is the case for classical probability. In particular, a natural question would be regarding a generalisation of the classical Central Limit Theorem, where sums of independent and identically distributed (i.i.d.) centred random variables having a finite variance converge to the *standard normal distribution* under appropriate scaling. In free probability, such a result does exist, but the limiting law is not the normal distribution; rather, it is the non-commutative analogue of the normal distribution (see Speicher [2011]).

Theorem 1.4.8 (Free Central Limit Theorem).
Let $(\mathcal{A}, \varphi)$ be a non-commutative probability space and let $(a_i)_{i\in I} \in \mathcal{A}$ be a family of free random variables such that $\varphi(a_i) = 0$ and $\varphi(a_i^2) = 1$ for each $i \in I$. Further, assume that $(a_i)_{i\in I}$ are identically distributed, in the sense that $\varphi(a_i^r) = \varphi(a_j^r)$ for any $r \in \mathbb{N}$ and all $i, j \in I$. Then, if $S_n = \sum_{i=1}^n a_i$, then, for any $k \in \mathbb{N}$,

$$\lim_{n\to\infty} \varphi(n^{-k/2} S_n^k) = \varphi(s^k)\,,$$

where s is the semicircle variable, or the semicircle element, *with*

$$\varphi(s^k) = \begin{cases} 0, & k \text{ is odd,} \\ C_m = \frac{1}{m+1}\binom{2m}{m}, & k = 2m \text{ for some } m \in \mathbb{N}. \end{cases}$$

In the literature, the limiting law is called the semicircle law, or the *Wigner semicircle law*, named after the theoretical physicist Eugene Wigner, whose pioneering work in the 1950s on the study of eigenvalue statistics of random matrices led to the foundation of *random matrix theory*.

Connections between random matrices and free probability were established in 1991 in the seminal work of Voiculescu (Voiculescu [1991]), where it was shown that random matrix models exhibit *asymptotic freeness*. This allows one to exploit tools from free probability to analyse various random matrix problems.

For any $N \in \mathbb{N}$, a $^*-$probability space of random $N \times N$ matrices is just $(M_N(L^{\infty-}(\Omega, \mathbb{P})), \operatorname{tr} \otimes \mathbb{E})$, where $(\Omega, \mathbb{P})$ is a classical probability space, and

$$L^{\infty-}(\Omega, \mathbb{P}) := \bigcap_{1\le p<\infty} L^p((\Omega, \mathbb{P}))\,,$$

and for any complex algebra $\mathcal{A}$, $M_N(\mathcal{A}) \cong M_N(\mathbb{C}) \otimes \mathcal{A}$ is the space of $N \times N$ matrices with entries drawn from $\mathcal{A}$. Moreover, $\mathbb{E}$ is the expectation with respect to the law $\mathbb{P}$. Recall that, for any (random) matrix $\mathbf{M}_N$, we have

$$\operatorname{Tr}(\mathbf{M}_N) := \sum_{i=1}^N \mathbf{M}_N(i,i) = \sum_{i=1}^N \lambda_i(\mathbf{M}_N)\,,$$

where $(\lambda_i(\mathbf{M}_N))_{i=1}^N$ are the *eigenvalues* of $\mathbf{M}_N$. The following result, which is an extension of the original work by Voiculescu, shows asymptotic freeness of Gaussian ensembles and deterministic matrices (see Speicher [2011, Theorem 6.14]).

Theorem 1.4.9 (Asymptotic freeness in matrix ensembles).
For $t \in \mathbb{N}$ and $N \in \mathbb{N}$, let $\mathrm{G}^{(1)}, \ldots, \mathrm{G}^{(t)}$ be t independent $N\times N$ Gaussian unitary ensembles GUE(N). *Let $X_N \in M_N(\mathbb{C})$ be a deterministic $N \times N$ matrix such*

that $\sup_N \|X_N\| \leq C$ *for some* $C > 0$ *(where* $\|.\|$ *denotes the Hilbert-Schmidt norm) and* $X_N \xrightarrow{d} x$ *in the space* $(\mathcal{A}, \varphi)$*, that is,*

$$\lim_{N\to\infty} \operatorname{tr}(X_N^k) = \varphi(x^k)$$

for each $k \in \mathbb{N}$*. Then*

$$(\mathrm{G}^{(1)}, \ldots, \mathrm{G}^{(t)}, X_N) \xrightarrow{d} (s_1, \ldots, s_t, x)\,,$$

where $s_1, \ldots, s_t$ *are semicircle elements in* $(\mathcal{A}, \phi)$ *and* $s_1, \ldots, s_t, x$ *are free, that is, for all* $m \in \mathbb{N}$*,* $q : m \to \mathbb{N}_0$*, and* $p : [m] \to [t]$*,*

$$\lim_{N\to\infty} \mathbb{E}\left[\operatorname{tr}\left(\mathrm{G}^{(p(1))} X_N^{q(1)} \ldots \mathrm{G}^{(p(m))} X_N^{q(m)}\right)\right] = \varphi\left(s_1^{(p(1))} x^{q(1)} \ldots s_t^{(p(m))} x^{q(m)}\right)\,.$$

The above theorem shows that $(\{\mathrm{G}^{(i)}\}_{1\leq i\leq t}, X_N)$ are asymptotically free, allowing us to conclude results about sums and products of random matrices.

The remaining technical details are quoted from Anderson et al. [2010], Hazra and Maulik [2013].

Definition 1.4.10 (Affiliated operators).
A self-adjoint operator X *is said to be* affiliated *to a* W^*-*algebra* $\mathcal{A}$*, if* $f(X) \in \mathcal{A}$ *for an bounded Borel function* f *on* $\mathbb{R}$*.*

We call self-adjoint operators associated to $\mathcal{A}$ *random elements* of $\mathcal{A}$. For any affiliated random element X, the algebra generated by X is defined as $\mathcal{A}_X := \{f(X) : f \text{ bounded measurable}\}$. Naturally, $X_1, X_2 \in \mathcal{A}$ are free if $\mathcal{A}_{X_1}, \mathcal{A}_{X_2}$ are free, as in the following definition.

Definition 1.4.11 (Free operators).
Self-adjoint operators $(X_i)_{i\in I}$ *affiliated with a* W^*-*algebra* $\mathcal{A}$ *are said to be* free *if and only if the algebras generated by* $\{f(X_i) : f \text{ bounded measurable}\}_{i\in I}$ *are free.*

Definition 1.4.12 (Law of an operator).
For a self-adjoint operator (or a random element) X *affiliated to a* W^*-*algebra* $\mathcal{A}$*, and the probability space* $(\mathcal{A}, \varphi)$*, the* law *of* X *is the unique probability measure* μ_X *on* $\mathbb{R}$ *satisfying*

$$\varphi(f(X)) = \int_{\mathbb{R}} f(t)\mu_X(\mathrm{d}\,x)$$

for every bounded Borel function f *on* $\mathbb{R}$*.*

If Λ is the projection-valued spectral measure associated with X (which is guaranteed by the spectral theorem), with Λ_A denoting the measure evaluated at a set A, then

$$\mu_X(-\infty, x] = \varphi(\Lambda_{(-\infty,x]}(X)) .$$

The following is quoted from Anderson et al. [2010, Proposition 5.3.34].

Proposition 1.4.13.
Let $\mu_1, \dots, \mu_p$ be probability measures on $\mathbb{R}$. Then there exists a W^*-probability space $(\mathcal{A}, \varphi)$ with φ a normal faithful tracial state*, and self-adjoint operators $(X_i)_{1\leq i\leq p}$ affiliated with $\mathcal{A}$, with laws $(\mu_i)_{1\leq i\leq p}$ that are free.*

From Anderson et al. [2010, Property 5.3.34, Corollary 5.3.35], one can always construct a Hilbert space $\mathcal{H}$, a tracial state φ, and two free variables X_1 and X_2 with laws μ_1 and μ_2, respectively, affiliated with the space $\mathcal{B}(\mathcal{H})$ of bounded linear operators on $\mathcal{H}$. Then, *free additive convolution* of μ_1 and μ_2, denoted as $\mu_1 \boxplus \mu_2$, is the law of $X_1 + X_2$. Additionally, if either X_1 or X_2 is non-negative, then the *free multiplicative convolution* $\mu_1 \boxtimes \mu_2$ is the law of $X_1 X_2$. The extension of free convolutions to unbounded measures can be done in the context of finite von Neumann algebras. Assume that $\mathcal{A}$ is a finite von Neumann algebra with a normal faithful tracial state φ, that is, $(\mathcal{A}, \varphi)$ is a tracial W^*-probability space and $\mathcal{A}$ is acting on a Hilbert space $\mathcal{H}$. A closed, densely defined operator T on $\mathcal{H}$ is *affiliated* with $\mathcal{A}$ if its polar decomposition $T = uX$ has the property that $u \in \mathcal{A}$ and X is affiliated with $\mathcal{A}$. Let $\tilde{\mathcal{A}}$ denote the set of all operators on $\mathcal{H}$ that are affiliated with $\mathcal{A}$. Then, $\tilde{\mathcal{A}}$ is an algebra, that is, if $X, Y \in \tilde{\mathcal{A}}$, then $X + Y$ and XY are densely defined, closable, and their closures are in $\tilde{A}$. See Bercovici and Voiculescu [1993] for further details.

§1.5 Spectral approach to random graphs

Spectral analysis of random graph models studies the limiting spectral distribution of the associated random matrices. The analysis follows a similar structure as in random matrix theory, where we begin with the ESD of the matrix of the finite graph and study its behaviour asymptotically as the size tends to infinity.

Results on the bulk distribution in random matrix theory and spectral theory of random graphs are *CLT-type*, that is, they have the same flavour as the free central limit theorem. In particular, if $\mathbf{M}_N$ is some Hermitian random matrix with entries having law $\mathbb{P}$, which could also be an adjacency or a Laplacian matrix, then the main question is as follows: Does there exist a (possibly random) measure μ_0 such that

$$\mathrm{ESD}\left(\frac{\mathbf{M}_N - \mathbb{E}[\mathbf{M}_N]}{c_N}\right) \xrightarrow{*} \mu_0\,?$$

Here, $*$ denotes that this convergence could be (weakly, in the measure-theoretic sense) in distribution, in $\mathbb{P}-$probability, or $\mathbb{P}-$almost surely, and c_N is a scaling that is of the order of the variance of the entries, given by

$$c_N = \mathbb{E}[\mathrm{Tr}(\mathbf{M}_N^2)] = \sum_{i,j=1}^{N} \mathbb{E}[\mathbf{M}_N(i,j)^2]\,.$$

For random graph models, this scaling also turns out to be the expected degree of a uniformly chosen vertex.

§1.5.1 Revisiting the Erdős-Rényi random graph

In the case of the homogeneous ERRG(N, ε_N), it is known that in the dense case the empirical distribution converges to the semicircle law after an appropriate scaling (Jung and Lee [2018], Tran et al. [2013]). The Laplacian spectrum for the dense case was studied in Ding and Jiang [2010], Jiang [2012].

In the sparse case, the spectra converge to limiting measures that depend on the parameter $\lambda := \lim_{N\to\infty} N\varepsilon_N$. The behaviour is much more complicated in this setting. Various interesting properties for spectra of the adjacency matrix were predicted by Bauer and Golinelli [2001]. The existence of the limiting distribution was proved by Khorunzhy et al. [2004], who study both the adjacency and the Laplacian matrices, and also show some interesting properties of the moments and the limiting Stieltjes transform. The local geometric behaviour of sparse random graphs can be studied using the theory of local weak convergence (LWC), which builds on the works Aldous and Lyons [2007] and Benjamini and Schramm [2001]. LWC describes how a graph looks like around a uniformly chosen vertex in the limit as the size of the graph tends to infinity. For a detailed review of LWC and various other applications, see van der Hofstad [2024]. In a remarkable work by Bordenave and Lelarge [2010], where the authors study the adjacency and the Laplacian matrices, it was proved that if a graph with N vertices converges locally weakly to a Galton-Watson tree, then the Stieltjes transform of the empirical spectral distribution converges in L^1 to the Stieltjes transform of the spectral measure of the tree, and satisfies a recursive distributional equation. The example of a homogeneous ERRG was treated in [Bordenave and Lelarge, 2010, Example 2].

The limiting measure of the adjacency matrix of the sparse ERRG depends on λ and is still very non-explicit. It was proved by Bordenave et al. [2017], Arras and Bordenave [2023] that the measure has an absolutely continuous component if and only if $\lambda > 1$. The size of the atom at the origin was computed by Bordenave et al. [2011], and the nature of the atomic part of the measure was studied in Salez [2020], where it was shown that the set of atoms is dense when $\lambda < 1$, and is linked with a countable dense ring $\mathbb{A}$ of *totally real algebraic integers*. The study of so-called extended states at the origin was initiated in Coste and Salez [2021], and it was shown that for $\lambda < e$ there were no extended states, while for $\lambda > e$, there are extended states.

All these results were conjectured in Bauer and Golinelli [2001]. Most results on local limits show that properties are generally true for unimodular Galton-Watson trees. In the simulations of Bauer and Golinelli [2001], it is clear that when λ is slightly larger than 1, the limiting measure already starts taking the shape of the semicircle law. It was shown in Jung and Lee [2018] that, indeed, if $\lambda \to \infty$, then the limiting measure converges to the semicircle law. Some key questions still remain open for the sparse ERRG, such as the following:

- What are the explicit moments of μ_λ?
- How "close" is the measure μ_λ to μ_{sc}? Is there a way to quantify the distance between the two measures?

§1.5.2 Local weak convergence

The seminal work of Bordenave and Lelarge [2010] characterises the limiting spectral distribution for locally tree-like graphs. In particular, consider $\mathbf{A}_N$ and $\mathbf{\Delta}_N$ to be the scaled adjacency and Laplacian matrices, respectively, of a random graph model $\mathbb{G}_N$, such that the following hold:

- The sequence of random graphs $\{\mathbb{G}_N\}_{N\geq 1}$ has a weak limit $\mathbb{G}$.
- For a uniformly chosen root $o_N \in \mathbb{G}_N$, the degree sequence of the rooted graph $(\deg(\mathbb{G}_N, o_N))_{N\geq 1}$ is uniformly integrable.
- Let $\mathcal{G}^*$ denote the set of rooted isomorphism classes of rooted connected locally finite graphs, and let $U_2(\mathbb{G})$ be the distribution on $\mathcal{G}^* \times \mathcal{G}^*$ of the pair of rooted graphs $((\mathbb{G}, o_1), (\mathbb{G}, o_2))$, where o_1, o_2 are uniformly chosen roots of G. Then, $U_2(\mathbb{G}_N)$ converges weakly to $\mathbb{G} \otimes \mathbb{G}$, that is, to two independent and identical copies of $\mathbb{G}$.

Under the above conditions, there exists unique probability measures μ_λ and ν_λ on $\mathbb{R}$ such that $\lim_{N\to\infty} \mathrm{ESD}(\mathbf{A}_N) = \mu_\lambda$ and $\lim_{N\to\infty} \mathrm{ESD}(\mathbf{\Delta}_N) = \nu_\lambda$ weakly in probability. Furthermore, if $\mathbb{G}_N$ is the graph $\mathrm{ER}_N(\varepsilon_N)$, and $\mathbf{A}_{\mathbb{G}_N}$ is the adjacency matrix of the graph, then $\mathbf{A}_N := \lambda^{-1/2}\mathbf{A}_{\mathbb{G}_N}$, and the measure μ_λ represents the expected spectral measure associated with the root of a Galton-Watson tree with offspring distribution $\mathrm{Poi}(\lambda)$ and weights $1/\sqrt{\lambda}$. This result comes from the theory of local weak convergence (see Benjamini and Schramm

[2001], van der Hofstad [2024]), which is a powerful tool to study spectral measures associated with many sparse random graph models.

In particular, consider the adjacency matrix (though a similar result holds for the Laplacian matrix). Consider the space $\mathbb{H}$ of holomorphic functions $f : \mathbb{C}^+ \to \mathbb{C}^+$, equipped with the topology induced by uniform convergence on compact sets. Then, $\mathbb{H}$ is a complete separable metrizable compact space. The *resolvent* of the adjacency matrix is given as

$$\mathrm{R}_{\mathbf{A}_N}(z) = (\mathbf{A}_N - zI)^{-1}$$

for each $z \in \mathbb{C}^+$. The map $z \mapsto \mathrm{R}_{\mathbf{A}_N}(z)(i,i)$ is in $\mathbb{H}$, and the Stieltjes transform of $\mathrm{ESD}(\mathbf{A}_N)$ is given by $\operatorname{tr} \mathrm{R}_{\mathbf{A}_N}(z)$, where $\operatorname{tr} = N^{-1}\operatorname{Tr}$ denotes the normalised trace operator. Let $\mathcal{G}^*$ denote the set of rooted isomorphism classes of rooted connected locally finite graphs. Assume that the random graph sequence $(\mathbb{G}_N)_{N\geq 1}$ has the random local limit $\mathbb{G} \in \mathcal{G}^*$, and assume further that $\mathbb{G}$ is a *Galton-Watson Tree* with degree distribution F_*, that is, a rooted random tree obtained from a Galton-Watson process with root having offspring distribution F_* and all children having a distribution F (which may or may not be the same as F_*).

Let $\mathrm{S}_{\mathbf{A}_N}(z)$ denote the Stieltjes transform of the empirical measure $\mathrm{ESD}(\mathbf{A}_N)$. It was shown in [Bordenave and Lelarge, 2010, Theorem 2] that there exists a unique probability measure Q on $\mathbb{H}$ such that, for each $z \in \mathbb{C}^+$,

$$Y(z) \stackrel{d}{=} \left(z + \sum_{i=1}^{P} Y_i(z)\right)^{-1}$$

where P has distribution F and $Y, \{Y_i\}_{i\geq 1}$ are i.i.d. with law Q and independent of P. Moreover,

$$\lim_{N\to\infty} \mathrm{S}_{\mathbf{A}_N}(z) = \mathbb{E}X(z) \text{ in } L^1,$$

where $X(z)$ is such that:

$$X(z) \stackrel{d}{=} -\left(z + \sum_{i=1}^{P_*} Y_i(z)\right)^{-1},$$

where $\{Y_i\}_{i\geq 1}$ are i.i.d. copies with law Q, and P_* is a random variable independent of $\{Y_i\}_{i\geq 1}$ having distribution F_*.

The analysis and expressions are similar for $\mathrm{S}_{\mathbf{\Delta}_N}$, as illustrated in Bordenave and Lelarge [2010].

§1.5.3 Further literature

Adjacency matrix

In recent years, there has been significant research on inhomogeneous Erdős–Rényi random graphs, which can be equivalently modelled by Wigner matrices with a variance profile. The limiting spectral distribution of the adjacency matrix of such graphs has been studied in Chakrabarty et al. [2021b], Zhu [2020], Bose et al. [2022], while local eigenvalue statistics have been analysed in Dumitriu and Zhu [2019], Ajanki et al. [2019]. Zhu and Zhu [2024] studies the fluctuations of the linear eigenvalue statistics for a wide range of such inhomogeneous graphs. Additionally, various properties of the largest eigenvalue have been investigated in Cheliotis and Louvaris [2024], Husson [2022], Chakrabarty et al. [2022], Ducatez et al. [2024], Dionigi et al. [2023]. One of the most significant properties of the limiting spectral measure for random graphs is its absolute continuity with respect to the Lebesgue measure, which is closely tied to the concept of mean quantum percolation [Bordenave et al., 2017, Anantharaman et al., 2021, Arras and Bordenave, 2023]. Quantum percolation investigates whether the limiting measure has a non-trivial absolutely continuous spectrum. Recently, it was shown in Arras and Bordenave [2023] that the adjacency operator of a supercritical Poisson Galton-Watson tree has a non-trivial absolutely continuous part when the average degree is sufficiently large. Additionally, Bordenave et al. [2017] demonstrated that supercritical bond percolation on $\mathbb{Z}^d$ has a non-trivial absolutely continuous part for $d = 2$. These results motivate similar questions for *kernel-based random graphs* and other percolation models. In Bhamidi et al. [2012] the spectra of the adjacency matrix of random trees are studied, including the *preferential attachment tree.* Spectral analysis of weighted adjacency matrices has also been used in hidden clique problems (see Chatterjee et al. [2025]).

Laplacian Matrix

Bryc et al. [2006] established that, for large symmetric matrices with i.i.d. entries, the empirical spectral distribution (ESD) of the corresponding Laplacian matrix converges to the free convolution of the semicircle law and the standard Gaussian distribution. In the context of sparse Erdős–Rényi random graphs, Huang and Landon [2020] studied the local law of the ESD of the Laplacian matrix. They demonstrated that the Stieltjes transform of the ESD closely approximates that of the free convolution of the semicircle law and a standard Gaussian distribution, down to scale N^{-1}. Additionally, they showed that the gap statistics and averaged correlation functions align with those of the Gaussian Orthogonal Ensemble in the bulk. Ding and Jiang [2010] investigated the spectral distributions of adjacency and Laplacian matrices of random graphs, assuming that the variance of the entries depend only on N. They established the convergence of the ESD of these matrices under such conditions. The results for the Erdős-Rényi random graphs were extended to the inhomogeneous setting by Chakrabarty et al. [2021b]. In a recent work, Chatterjee and Hazra [2022] derived a combinatorial way to describe the limiting moments for a wide variety of random matrix models with a variance profile.

§1.6 Outline of the thesis

The three main chapters of this thesis are based on three papers on spectral properties of inhomogeneous random graph models.

Chapter 2

In Chapter 2, we study the inhomogeneous Erdős-Rényi random graph model on N vertices in the sparse setting, where vertices have deterministic weights and edges are added between two vertices independently with a probability that is proportional to a function of their two weights, scaled by a factor of N. We take the vertex set $[N]$, and consider a sequence of *deterministic* weights $(w_i)_{i=1}^N$, such that if o_N is a uniform random variable on $[N]$, then there exists a limiting random variable W with law μ_w such that $w_{o_N} \xrightarrow{d} W$. We add edges independently with probability

$$p_{ij} := \varepsilon_N f(w_i, w_j)\,, i, j \in [N]\,,$$

where ε_N is a sparsity parameter such that $N\varepsilon_N \to \lambda \in (0, \infty)$, and f is a bounded continuous function.

We study the scaled adjacency matrix $\mathbf{A}_N$ of the random graph, with entries given by

$$\mathbf{A}_N(i,j) = \mathbf{A}_N(j,i) \stackrel{d}{=} \frac{1}{\sqrt{\lambda}} \mathrm{Ber}(p_{ij})\,.$$

In Theorem 2.3.7, we find that there exists a deterministic non-degenerate limiting measure μ_λ such that $\lim_{N\to\infty} \mathrm{ESD}(\mathbf{A}_N) = \mu_\lambda$ in probability, and the moments of μ_λ are given by

$$\int x^k \mu_\lambda(\mathrm{d}\, x) = \begin{cases} 0, & k \text{ is odd,} \\ \displaystyle\sum_{l=2}^{k/2+1} \sum_{\substack{\pi \in SS(k): \\ |\gamma\pi| = l}} \lambda^{l-1-\frac{k}{2}}\, t(G_{\gamma\pi}, f, \mu_w), & k \text{ is even,} \end{cases}$$

where $SS(k)$ is the set of Simple Symmetric partitions of $[k]$, as in Bose et al. [2022], $G_{\gamma\pi}$ is a graph associated to a partition π that is described later, and

$t(\cdot,\cdot,\cdot)$ is a generalisation of the graph homomorphism density that appears in graphon theory in Lovász and Szegedy [2006]. We further find that $\lim_{\lambda\to\infty}\mu_\lambda = \mu_f$, where μ_f is the measure in the dense regime that appears in Chakrabarty et al. [2021b], Zhu [2020], which extends the results of Jung and Lee [2018].

In Theorem 2.3.9, under the assumption that f is Lipschitz in one coordinate, we show that, in an appropriate Banach space $\mathcal{B}$, there exists a functional $\phi_z^* \in \mathcal{B}$ that is the unique solution to a fixed-point equation in $\mathcal{B}$, such that

$$\mathrm{S}_{\mu_\lambda}(z) = i\int_0^\infty \mathrm{e}^{-\lambda d_f(y)} \int_0^\infty \mathrm{e}^{ivz}\mathrm{e}^{\lambda\phi_z^*(y,\frac{v}{\lambda})}\,\mathrm{d}\,v\;\;\mu_w(\mathrm{d}\,y),\;\; z\in\mathbb{C}^+,$$

where $d_f(y) = \int f(x,y)\mu_w(\mathrm{d}\,x)$. This chapter is based on the paper Avena et al. [2023].

Chapter 3

In Chapter 3, we study a model with spatial geometry. We consider a kernel-based random graph model on a $d-$dimensional discrete torus $\mathbf{V}_N$, which serves as the vertex set of the random graph. Each vertex $i\in\mathbf{V}_N$ has a random weight W_i, where $(W_i)_{i\in\mathbf{V}_N}$ are i.i.d. random variables sampled from a Pareto distribution W (whose law is denoted by $\mathbf{P}$ and measure μ_W) with parameter $\tau-1$, where $\tau>1$, that is,

$$\mathbf{P}(W>t) = t^{-(\tau-1)}\mathbf{1}_{\{t\geq 1\}} + \mathbf{1}_{\{t<1\}}.$$

Conditionally on the weights, edges are added independently with probability

$$p_{ij} := P^W(i\leftrightarrow j) = \frac{\kappa(W_i,W_j)}{\|i-j\|^\alpha}\wedge 1\,,$$

where $\|\cdot\|$ is the torus distance, $\alpha\in(0,d)$ is a parameter of choice, and κ is a kernel that has the form $\kappa(x,y) := (x\vee y)(x\wedge y)^\sigma$ for some $0<\sigma<\tau-1$, as in Jorritsma et al. [2023].

We consider the scaled adjacency matrix of this graph, which is a symmetric random matrix with entries

$$\mathbf{A}_N(i,j) = \mathbf{A}_N(j,i) \stackrel{d}{=} c_N^{-1/2}\,\mathrm{Ber}(p_{ij})\,,$$

where $c_N = N^{1-\alpha}$. For $\tau>2$, Theorems 3.2.1 and 3.2.3 show that there exists a deterministic non-degenerate limiting measure $\mu_{\sigma,\tau}$ with finite second moment such that

$$\lim_{N\to\infty}\mathrm{ESD}(\mathbf{A}_N) = \mu_{\sigma,\tau}\,,\quad \text{in } \mathbb{P}-\text{probability},$$

where $\mathbb{P} = \mathbf{P}\otimes P^W$ is the joint law.

Theorem 3.2.4 shows that $\mu_{\sigma,\tau}$ is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}$. Theorem 3.2.5 shows that, when $\tau>3$ and $\sigma<\tau-2$, in an appropriate Banach space $\mathcal{B}$ there exists a unique analytic solution $a^* \in \mathcal{B}$ to a fixed-equation in $\mathcal{B}$, such that

$$\mathrm{S}_{\mu_{\sigma,\tau}}(z) = \int_1^\infty a^*(z,x)\mu_W(\mathrm{d}\,x)\,, \quad z \in \mathbb{C}^+.$$

When $\sigma = 1$, there is an explicit description of the measure. In particular, Theorem 3.2.2 tells us that $\mu_{1,\tau} = \mu_{sc} \boxtimes \mu_W$, with tail asymptotic $\mu_{1,\tau}(x,\infty) \sim C_\tau x^{-2(\tau-1)}$ as $x \to \infty$, for some $\tau-$dependent constant $C_\tau < \infty$. Here, $\boxtimes$ is the *free multiplicative convolution* of measures. This chapter is based on the paper Cipriani et al. [2025].

Chapter 4

In Chapter 4, we take the model from Chapter 3 with $\sigma = 1$ and $\tau > 3$, that is, weights with finite variance. This model is called the scale-free percolation model. We begin with the scaled adjacency $\mathbf{A}_N$ as in Chapter 3, and define the corresponding Laplacian as $\mathbf{\Delta}_N = \mathbf{A}_N - \mathbf{D}_N$. We study the centred Laplacian $\mathbf{\Delta}_N^\circ := \mathbf{\Delta}_N - \mathbb{E}[\mathbf{\Delta}_N]$. Theorem 4.2.1 shows that there exists a deterministic limiting measure ν_τ such that

$$\lim_{N\to\infty} \mathrm{ESD}(\mathbf{\Delta}_N^\circ) = \nu_\tau \quad \text{in } \mathbb{P}-\text{probability}\,.$$

Theorem 4.2.5 identifies ν_τ in terms of the spectral distribution of some non-commutative operators. Heuristically, ν_τ has (in an operator sense) the law given by the spectral law of

$$W^{1/2}SW^{1/2} + m_1 W^{1/4}GW^{1/4},$$

where W is an unbounded operator with spectral law given by the Pareto distribution, S is a bounded compact operator whose spectral law is the semicircle law, and G is an unbounded operator whose law is given by the Gaussian distribution. We will see a more formal description of this later on in Chapter 4. This chapter is based on the paper Hazra and Malhotra [2025].

Chapter 5

In Chapter 5, we show some further simulations of the above models, and conclude with a short discussion on open problems.

§1.7 Concluding remarks

The thesis gives a spectral perspective to some inhomogeneous random graph problems. The results mainly describe properties of the bulk distribution. There are many other interesting features, and we hope that this thesis will form a baseline for future research.

CHAPTER 2

Limiting spectra of inhomogeneous random graphs

This chapter is based on:
L. Avena, R.S. Hazra, N. Malhotra. *Limiting spectra of inhomogeneous random graphs.* [*arxiv:2312.02805*], 2023.

Abstract

We consider sparse inhomogeneous Erdős-Rényi random graph ensembles where edges are connected independently with probability p_{ij}. We assume that $p_{ij} = \varepsilon_N f(w_i, w_j)$ where $(w_i)_{i\geq 1}$ is a sequence of deterministic weights, f is a bounded function and $N\varepsilon_N \to \lambda \in (0,\infty)$. We characterise the limiting moments in terms of graph homomorphisms and also classify the contributing partitions. We present an analytic way to determine the Stieltjes transform of the limiting measure. The convergence of the empirical distribution function follows from the theory of local weak convergence in many examples but we do not rely on this theory and exploit combinatorial and analytic techniques to derive some interesting properties of the limit. We extend the methods of Khorunzhy et al. [2004] and show that a fixed point equation determines the limiting measure. The limiting measure crucially depends on λ and it is known that in the homogeneous case, if $\lambda \to \infty$, the measure converges weakly to the semicircular law (Jung and Lee [2018]). We extend this result of interpolating between the sparse and dense regimes to the inhomogeneous setting and show that as $\lambda \to \infty$, the measure converges weakly to a measure which is known as the operator-valued semicircular law.

§2.1 Introduction

Homogeneous Erdős-Rényi Random Graphs (ERRG) serve as the basis for many mathematical theories in random graphs. Real-world networks are highly inhomogeneous and have a far more complex structure. Various attempts have been made to generalise this to other kinds of random graph models. One of the successful extensions is the inhomogeneous Erdős-Rényi random graph model introduced by Bollobás et al. [2007]. This graph has N vertices labelled by $[N] = 1, ..., N$, and edges are present independently with probability p_{ij} given by $p_{ij} = \frac{f(x_i,x_j)}{N} \wedge 1$, where f is a nice symmetric kernel on a state space $S \times S$, and x_i are certain attributes associated with vertex i belonging to S. If f is bounded, the graph is a sparse random graph. To introduce the non-sparse regime, in this article, we consider a small variant of the above inhomogeneous random graph. The vertex set remains the same, but the connection probabilities are given by

$$p_{ij} = \varepsilon_N f(w_i, w_j) \wedge 1, \tag{2.1}$$

where ε_N is a tuning parameter, (w_i) is a sequence of deterministic weights, and f is a symmetric, bounded function on $[0, \infty)^2$. The weights can also be generally random, but we do not consider this case. Note that when $N\varepsilon_N \to \infty$, the average degree is unbounded, and when $N\varepsilon_N = \mathrm{O}(1)$, the average degree is bounded. We call the former case *dense* and the latter case *sparse*. In the sparse case, the properties of the connected components were studied in Bollobás et al. [2007]. They studied the properties of the connected components and their relationship with the branching process. It was shown that the largest component of the graph has a size of order N if the operator norm of the kernel operator corresponding to f is strictly greater than 1 (see also [van der Hofstad, 2024, Theorem 3.9]). In the subcritical case, the sizes of the largest connected components can exhibit different behaviour compared to the ERRG. The study of the largest connected components in various inhomogeneous random graphs has attracted a lot of attention (see, for example, Bhamidi et al. [2010], Broutin et al. [2021], Bet et al. [2023], van der Hofstad [2013], Devroye and Fraiman [2014]). In this chapter, we are interested in the empirical distribution of the eigenvalues of the adjacency matrix of the graph and how the transition occurs from the sparse to the dense case in terms of the limiting spectral distribution. There hasn't been much literature in this area, even though various specific graphs have been studied. For example, the largest eigenvalue of the sparse Chung-Lu random graph was studied in Chung et al. [2003], and this was extended to an inhomogeneous setting by Benaych-Georges et al. [2020, 2019]. The bulk of the spectrum of sparse graphs is mainly studied through local weak convergence. Here, we present a unifying approach to understanding both the sparse and the dense cases, allowing us to interpolate between the two regimes.

In the case of homogeneous ERRG, it is known that in the dense case, the empirical distribution converges to the semicircle law after an appropriate scaling (Tran et al. [2013]). In the sparse case, it converges to a measure that depends on the parameter $N\varepsilon_N \to \lambda$. The behaviour is much more complicated in the sparse case. Various interesting properties were predicted by Bauer and Golinelli [2001]. The existence of the limiting distribution was proved by Khorunzhy et al. [2004], who also showed some interesting properties of the moments and the limiting Stieltjes transform. The local geometric behaviour of sparse random graphs can be well studied using the theory of local weak convergence (LWC), which builds on the works Aldous and Lyons [2007] and Benjamini and Schramm [2001]. It roughly describes how a graph looks like in the limit around a uniformly chosen vertex. For a detailed review of LWC and various other applications, see van der Hofstad [2024]. In a remarkable work by Bordenave and Lelarge [2010], it was proved that if a graph with N vertices converges locally weakly to a Galton-Watson tree, then the Stieltjes transform of the empirical spectral distribution converges in L^1 to the Stieltjes transform of the spectral measure of the tree, and it satisfies a recursive distributional equation. The example of homogeneous ERRG was treated in [Bordenave and Lelarge, 2010, Example 2]. The limiting measure of sparse ERRG depends on λ and is still very non-explicit. It was proved by Bordenave et al. [2017], Arras and Bordenave [2023] that the measure has an absolutely continuous component if and only if $\lambda > 1$. The size of the atom at the origin was shown by Bordenave et al. [2011], and the nature of the atomic part of the measure was studied in the same article. The study of so-called extended states at origin was initiated in Coste and Salez [2021], and it was shown that for $\lambda < e$, there were no extended states, and for $\lambda > e$, it has extended states. All these results were conjectured in Bauer and Golinelli [2001]. Most of these results on local limits show that properties are generally true for unimodular Galton-Watson trees.

In the simulations of Bauer and Golinelli [2001], it is clear that when λ is slightly larger than 1, the limiting measure already starts taking the shape of the semicircle law. It was shown in Jung and Lee [2018] that indeed, if $\lambda \to \infty$, then the limiting measure converges to the semicircle law. In the general case, the moments of the limiting measure depend on certain kinds of graph homomorphism counts, which also appeared in the works of Zhu [2020]. Although the theory of local weak convergence is very useful, we do not know if it can be used to derive the moments of the limiting measure. In Chakrabarty et al. [2021b], they considered IER to have weights $w_i = i/N$, and $N\varepsilon_N \to \infty$. This result can be extended to general deterministic weights without significant effort, and we state this general result in Section 2.2. The limiting measure is well-known in the free probability literature and appears as a universal object in many inhomogeneous systems, referred to as the *operator-valued semicircle law* [Speicher, 2011, Theorem 22.7.2]. The Stieltjes transform satisfies a recursive analytic equation.

Our contribution

As mentioned earlier, although the convergence of the empirical spectral distribution of graphs with a local-weak limit follows from the general result in Bordenave and Lelarge [2010], the limiting moments and contributing partitions are not known in full generality. It is also unclear how closely the limiting measures align in the sparse and dense regimes. Our main motivation for the work comes from [Jung and Lee, 2018, Theorem 1], which addresses these issues in the case of ERRG. We extend the results from ERRG to inhomogeneous models. We explicitly derive the moments of the limiting measure for the inhomogeneous setting, extending the works of Khorunzhy et al. [2004], albeit with a different proof. We also study the Stieltjes transform of the limiting measure, following the idea of Khorunzhy et al. [2004], and attempt an expansion of it for λ large enough. This has also gained attention in the physics literature, see references in Akara-pipattana and Evnin [2023]. We show that when $\lambda \gg 1$, the limiting moments closely resemble those of the IER, as derived in Chakrabarty et al. [2021b] and also implied by the work of Zhu [2020]. We derive the Stieltjes transform in the sparse setting using a fixed-point equation. The fixed point is simpler in the case of homogeneous ERRG, but in the inhomogeneous case, it becomes more complex. We explicitly characterise this fixed-point equation. We believe that in the future, this will aid in determining the rate of convergence of the empirical spectral distribution, which can be precisely quantified in terms of λ and N. The rates of convergence in the free central limit theorem were recently explored in Banna and Mai [2023], but these results are not directly applicable to our setting. We leave this as an open problem. Obtaining an explicit rate of convergence will provide an exact explanation of why the limiting measure in the sparse setting is very close to the non-sparse setting for relatively small $\lambda > 1$. We believe that the methods used in this article will be applicable in a setting even when the local limits of the graphs are not tree-like.

Brief summary of the results

The two main results of this work aim to characterise the limiting spectral measure of inhomogeneous Erdős-Rényi random graphs. Our first result, Theorem 2.3.7, gives a characterisation of the moments of this measure, where the k^{th} moment for any $k \geq 0$ is described in terms of homomorphism densities of the inhomogeneity function f and special classes of partitions of the tuple $[k]$. We can recover the moments of the dense regime asymptotically (as $\lambda \to \infty$) using this result. The second result, Theorem 2.3.9, provides an analytic characterisation of the measure. In particular, we provide an analytic characterisation of a functional of the resolvent of the adjacency matrix in terms of a fixed-point equation. As a consequence, in Corollaries 2.3.10 and 2.3.11, we obtain the Stieltjes transform of the sparse and dense limiting measures. The form of the limiting Stieltjes transform can be seen as an alternative description of the form obtained through local weak convergence (whenever it applies).

Outline

We begin Section 2.2 by describing the model and stating the results of the dense regime. We state the assumptions on the sparse setting more explicitly and proceed by stating our main results for this setting. We then describe a relationship with local weak convergence and also give some examples of popular random graph models. We show that the sparse Chung-Lu type model falls into our setting, and while the Norros-Reittu model and the Generalised Random Graph model do not directly fall into our setting, we show that asymptotically the three models have the same spectral distribution, which has a free-multiplicative part that can be seen from our main results.

In Section 2.4 we prove our first main result, which takes a combinatorial approach, and we set up all the necessary tools used in proving the result. We identify the moments of the limiting spectral measure in terms of partitions of a tuple and graph-homomorphism densities. We provide a characterisation of the partitions and explicit expressions for the moments that are given by homomorphism densities defined based on these partitions. We further identify a leading order of the moments and a polynomial in λ^{-1}, which was also seen for the homogeneous setting in Jung and Lee [2018].

In Section 2.5 we prove our second main result, which in contrast has an analytic flavour. We set up the relevant analytic structures, and instead of working directly with the Stieltjes Transform, we work with a functional of the resolvent of the adjacency matrix, which was introduced in Khorunzhy et al. [2004]. We borrow both fundamental and advanced tools from analysis to provide an exact analytic characterisation of the limiting spectral measure. We conclude with the Appendix as Section 2.6 where we state the key analytic tools we use in Section 2.5.

§2.2 Setting

§2.2.1 Model

We consider the inhomogeneous Erdős-Rényi random graph (IER) $\mathbb{G}_N$ on the vertex set $[N] = \{1, \ldots, N\}$ where edges are added independently with probability p_{ij}. As mentioned before, we will assume that p_{ij} has a special form as

$$p_{ij} = \varepsilon_N f(w_i, w_j) \wedge 1\,,$$

where ε_N is a tuning parameter such that $\varepsilon_N \to 0$, $(w_i)_{i\geq 1}$ is a sequence of deterministic non-negative weights and $f : [0,\infty)^2 \to [0,\infty)$ is bounded and continuous. We will use $\mathbb{P}_N$ to denote the law of this random graph, and we will drop the subscript N for notational convenience, and $\mathbb{E}$ will be the expectation with respect to the law $\mathbb{P}$. We will always assume that N is large enough and hence ε_N is small enough to make $p_{ij} \leq 1$ since f is bounded.

Let $\mathbf{M}_N$ denote the adjacency matrix of the graph $\mathbb{G}_N$, that is, the (i,j)-th entry is 1 if i shares an edge with j, and 0 otherwise. So $\mathbf{M}_N$ is a symmetric matrix, where any entry $\mathbf{M}_N(i,j)$ is distributed as Bernoulli random variable with parameter p_{ij} as in (2.1) and $\{\mathbf{M}_N(i,j), i \geq j\}$ is an independent collection. Instead of studying the adjacency matrix $\mathbf{M}_N$ we will study the scaled adjacency matrix. In particular, we do a CLT-type scaling by the variance of the entries, that is, we study the matrix

$$\frac{1}{\sqrt{N\varepsilon_N(1-\varepsilon_N)}}\mathbf{M}_N. \tag{2.2}$$

The empirical measure which puts mass $1/N$ on each eigenvalue of an $N \times N$ random matrix $\mathbf{A}_N$ is called the *Empirical Spectral Distribution* of $\mathbf{A}_N$, and is denoted by

$$\mathrm{ESD}(\mathbf{A}_N) := \frac{1}{N}\sum_{i=1}^{N}\delta_{\lambda_i}\,. \tag{2.3}$$

We are interested in studying the following object:

$$\mathrm{ESD}\left(\frac{\mathbf{M}_N}{\sqrt{N\varepsilon_N(1-\varepsilon_N)}}\right) = \frac{1}{N}\sum_{i=1}^{N}\delta_{\lambda_i}\,,$$

where $\lambda_1, \ldots, \lambda_N$ are the eigenvalues of $(N\varepsilon_N(1-\varepsilon_N))^{-1/2}\mathbf{M}_N$.

We are interested in the weak convergence (in probability) of the above measure and the limiting measure is called the *Limiting Spectral Distribution* (LSD). The limiting measure depends on the following two geometric regimes in random graphs and its properties differ in the two cases:

- **Dense Regime**: $\varepsilon_N \to 0$ and $N\varepsilon_N \to \infty$. The connectivity regime with $N\varepsilon_N \gg C\log N$ falls in this regime.

- **Sparse Regime** : $\varepsilon_N \to 0$ and $N\varepsilon_N \to \lambda \in (0,\infty)$.

Dense regime

In literature, the dense regime is characterised by $\varepsilon_N \equiv$ constant but we will not use the features of dense graphs in this article and hence by abuse of terminology, we say that a graph is dense when it is not sparse. Let us now recall briefly what happens in the dense regime. The following result was proved in Chakrabarty et al. [2021b] and can also be obtained from Zhu [2020].

Theorem 2.2.1 (ESD in the dense case).
Consider the IER graph with p_{ij} as in (2.1) with $\varepsilon_N \to 0$ and $N\varepsilon_N \to \infty$. Suppose the deterministic weights satisfy the following assumption:

Let o_N be an uniform random variable on $[N]$ and let $W_N = w_{o_N}$. We assume that there exists a W with law μ_w such that

$$W_N \xrightarrow{d} W.$$

Then there exists a measure μ_f which is compactly supported such that

$$\lim_{N\to\infty} \operatorname{ESD}\left(\frac{\mathbf{M}_N}{\sqrt{N\varepsilon_N(1-\varepsilon_N)}}\right) = \mu_f \text{ weakly in probability.}$$

Many interesting properties of this limiting measure are known. To define the moments we need a quantity which is similar to the homomorphism density of graphons. Define

$$t(H_k, f, \mu_w) := \int_{[0,\infty)^k} \prod_{\{a,b\}\in E(H_k)} f(w_a, w_b)\mu_w^{\bigotimes k}(\mathrm{d}\,\mathbf{w})\,, \tag{2.4}$$

where H_k is a simple graph on k vertices with the edge set $E(H_k)$, $\mu_w^{\bigotimes k}(\cdot)$ is the k-fold product measure of $\mu_w(\cdot)$, and $\mathbf{w} = (w_1, ..., w_k)$. If we restrict the range of f to $[0,1]$ and take $\mu_w(\cdot)$ as the Lebesgue measure on $[0,1]$, then this quantity is the standard graph homomorphism density (see Lovász and Szegedy [2006]).

The *rooted planar tree* is a planar graph with no cycles, with one distinguished vertex as a root, and with a choice of ordering at each vertex. The ordering defines a way to explore the tree starting at the root. One of the algorithms used for traversing the rooted planar trees is *depth-first search*. An enumeration of the vertices of a tree is said to have depth-first search order if it is the output of the depth-first search.

We now recall the definition of a Stieltjes transform of a measure μ on $\mathbb{R}$. For $z \in \mathbb{C}^+$, where $\mathbb{C}^+$ is the upper half complex plane, the Stieltjes Transform of a measure μ is given by

$$\mathrm{S}_\mu(z) = \int_{\mathbb{R}} \frac{1}{x-z}\mu(\mathrm{d}\,x).$$

The following proposition gives the properties of the measure μ_f which appears in Theorem 2.2.1.

Proposition 2.2.2.

(a) [**Moments**] *The measure μ_f is the unique probability measure identified by the following moments:*

$$\int x^{2k}\mu_f(\mathrm{d}\,x) = \sum_{j=1}^{C_k} t(T_j^{k+1}, f, \mu_w), \quad \int x^{2k+1}\mu_f(\mathrm{d}\,x) = 0,\ k \geq 0, \qquad (2.5)$$

where T_j^{k+1} is the j^{th} rooted planar tree with $k+1$ vertices and C_k is the k^{th} Catalan number.

(b) [**Stieltjes transform**] *There exists an unique analytic function $\mathcal{H}$ defined on $\mathbb{C}^+ \times [0,\infty)$ such that*

$$\mathrm{S}_{\mu_f}(z) = \int_0^\infty \mathcal{H}(z,x)\mu_w(\mathrm{d}\,x),$$

and $\mathcal{H}(z,x)$ satisfies the integral equation

$$z\mathcal{H}(z,x) = 1 + \mathcal{H}(z,x)\int_0^\infty \mathcal{H}(z,y)f(x,y)\mu_w(\mathrm{d}\,y), \quad x \geq 0. \qquad (2.6)$$

Example 2.2.3 (Rank 1).
One special case which arises in many examples of random graphs, and will be discussed later is when f has a multiplicative structure, that is, $f(x,y) = r(x)r(y)$, where $r : [0,\infty) \to [0,\infty)$ is a bounded continuous function. In this case, the measure

$$\mu_f = \mu_s \boxtimes \mu_{r(W)}$$

where μ_{sc} is the standard semicircle law and $\mu_{r(W)}$ is the law of $r(W)$ and $\boxtimes$ is the free multiplicative convolution of the two measures. When r is identically equal to 1 then $\mu_f = \mu_s$, the standard semicircle law. We refer to [Chakrabarty et al., 2021b, Theorem 1.3] for details.

Sparse regime

The seminal work of Bordenave and Lelarge [2010] characterises the limiting spectral distribution for locally tree-like graphs. In particular, if one takes $\mathbf{A}_N$ to be the scaled adjacency matrix as given in (2.2) of a random graph $\mathbb{G}_N$, they show that if the following hold:

- The sequence of random graphs $\{\mathbb{G}_N\}_{N\geq1}$ have a weak limit $\mathbb{G}$;
- For a uniformly chosen root $o_N \in \mathbb{G}_N$, the degree sequence of the rooted graph $(\deg(\mathbb{G}_N, o_N))_{N\geq1}$ is uniformly integrable;

- Let $\mathcal{G}^*$ denote the set of rooted isomorphism classes of rooted connected locally finite graphs, and let $U_2(\mathbb{G})$ be the distribution on $\mathcal{G}^* \times \mathcal{G}^*$ of the pair of rooted graphs $((\mathbb{G}, o_1), (\mathbb{G}, o_2))$, where o_1, o_2 are uniformly chosen roots of G. Then, $U_2(\mathbb{G}_N)$ converges weakly to $\mathbb{G} \otimes \mathbb{G}$, that is, to two independent and identical copies of $\mathbb{G}$;

then, there exists a unique probability measure μ_λ on $\mathbb{R}$ such that $\mathrm{ESD}(\mathbf{A}_N) \implies \mu_\lambda$ weakly in probability as $N \to \infty$. Furthermore, it is shown that when $f \equiv 1$, the measure μ_λ represents the expected spectral measure associated with the root of a Galton-Watson tree with an offspring distribution of $\mathrm{Poi}(\lambda)$ and weights $1/\sqrt{\lambda}$. This result comes from the theory of local weak convergence, also known as Benjamini-Schramm convergence (see van der Hofstad [2024], Benjamini and Schramm [2001]), which is a powerful tool to study spectral measures associated with many sparse random graph models.

In particular, consider the space $\mathbb{H}$ of holomorphic functions $f : \mathbb{C}^+ \to \mathbb{C}^+$, equipped with the topology induced by uniform convergence on compact sets. Then, this is a complete separable metrizable compact space. The *resolvent* of the adjacency operator is given as

$$\mathrm{R}_{\mathbf{A}_N}(z) = (\mathbf{A}_N - zI)^{-1}$$

for each $z \in \mathbb{C}^+$. The map $z \mapsto \mathrm{R}_{\mathbf{A}_N}(z)(i,i)$ is in $\mathbb{H}$, and the Stieltjes transform of $\mathrm{ESD}(\mathbf{A}_N)$ is given by $\mathrm{tr}\, \mathrm{R}_{\mathbf{A}_N}(z)$, where $\mathrm{tr} = N^{-1}\,\mathrm{Tr}$ denotes the normalised trace operator. Let $\mathcal{G}^*$ denote the set of rooted isomorphism classes of rooted connected locally finite graphs. Assume that the random graph sequence $(\mathbb{G}_N)_{N\geq 1}$ has the random local limit $\mathbb{G} \in \mathcal{G}^*$, and further that $\mathbb{G}$ is a *Galton Watson Tree* with degree distribution F_*, that is, a rooted random tree obtained from a Galton-Watson process with root having offspring distribution F_* and all children having a distribution F (which may or may not be the same as F_*).

Let $\mathrm{S}_{\mathbf{A}_N}(z)$ denote the Stieltjes transform of the empirical measure $\mathrm{ESD}(\mathbf{A}_N)$. It was shown in [Bordenave and Lelarge, 2010, Theorem 2] that there exists a unique probability measure Q on $\mathbb{H}$, such that for each $z \in \mathbb{C}^+$

$$Y(z) \stackrel{d}{=} \left(z + \sum_{i=1}^{P} Y_i(z) \right)^{-1}$$

where P has distribution F and $Y, \{Y_i\}_{i\geq 1}$ are i.i.d. with law Q and independent of P. Moreover

$$\lim_{N\to\infty} \mathrm{S}_{\mathbf{A}_N}(z) = \mathbb{E}X(z) \text{ in } L^1,$$

where $X(z)$ is such that:

$$X(z) \stackrel{d}{=} -\left(z + \sum_{i=1}^{P_*} Y_i(z)\right)^{-1}, \tag{2.7}$$

where $\{Y_i\}_{i\geq 1}$ are i.i.d. copies with law Q, and P_* is a random variable independent of $\{Y_i\}_{i\geq 1}$ having distribution F_*.

In [Bordenave and Lelarge, 2010, Example 2], we see that the sparse Erdős-Rényi random graph with $p = \frac{\lambda}{N}$ falls in their setup, and in particular, P is distributed as $\mathrm{Poi}(\lambda)$. For a general f, [Bordenave and Lelarge, 2010, Theorem 1] still guarantees the existence of μ_λ, since the graphs we will consider will have a local weak limit known as the *multi-type branching process* (see [van der Hofstad, 2024, Chapter 3] for more details). As f is bounded, we get that the degree sequence will still remain uniformly integrable. As mentioned before we will not follow this well-known route of local weak convergence. Instead, we show the above convergence through albeit classical methods. We now introduce the conditions under which we will work. We will have the following sparsity assumption on ε_N and a regularity assumption on the function f and the weights:

A.1 **Connectivity function:** Let $f : [0,\infty)^2 \to [0,\infty)$ be a bounded, continuous function, with $|f| \leq C_f \in (0,\infty)$,

A.2 **Sparsity assumption :** $N\varepsilon_N \to \lambda \in (0,\infty)$,

A.3 **Assumption on weights:** Let o_N be an uniform random variable on $[N]$ and let $W_N = w_{o_N}$. We assume that there exists a W with law μ_w such that

$$W_N \xrightarrow{d} W.$$

We make some preliminary remarks about the assumptions. Since f is bounded, we can easily see that f is μ_w-integrable. In the sparse setting, in most important examples, the graph is locally tree-like and this can be seen from the theory of local weak convergence.

Note that the limit $\lambda \to \infty$ recovers the dense regime. By this choice, we can see that $1-\varepsilon_N \approx 1$ as N becomes very large, and $N\varepsilon_N(1-\varepsilon_N) \to \lambda$. Thus, our matrix of interest is a scaled adjacency matrix now defined as follows:

$$\mathbf{A}_N = \frac{1}{\sqrt{\lambda}}\mathbf{M}_N\,. \tag{2.8}$$

§2.3 Main Results

In this subsection, we state the main results of this article. As mentioned before in the introduction, we would like to understand first the limiting empirical distribution of the sparse inhomogeneous Erdős Rényi (IER) random graph and also study the behaviour of the measure when the sparsity parameter increases. Recall that the adjacency matrix is defined in (2.8) and the empirical spectral distribution is denoted by $\mathrm{ESD}(\mathbf{A}_N)$ (see (2.3)). In what follows, we will see that

$$\lim_{N\to\infty} \mathrm{ESD}(\mathbf{A}_N) = \mu_\lambda \text{ weakly in probability} \tag{2.9}$$

and $\mu_\lambda \Rightarrow \mu_f$ where μ_f is as in Theorem 2.2.1. For the homogeneous case, where $f \equiv 1$, we get the final limit as the classical Wigner's semicircle law, that is, $\mu_f = \mu_s$. These iterated limits were studied in Jung and Lee [2018]. An interesting open question is how close μ_λ is to μ_f. Although we do not manage to give an explicit estimate, through the moment method we show that it is very close and the structure of the moments of μ_f is hidden inside the structure of the moments of μ_λ. This will be our first result. To describe the moments we need to introduce some notation.

§2.3.1 Method of moments: Combinatorial Approach

We first define the *Special Symmetric Partitions* which was introduced in Bose et al. [2022]. Let $\mathcal{P}(k)$ denote the set of partitions of k and $\mathcal{P}_2(k)$ be the set of pair partitions where each block has size 2. Let $NC(k)$ be the set of non-crossing partitions of $[k]$ and $NC_2(k)$ be the set of non-crossing pair partitions of $[k]$. Note that $|NC_2(2k)| = \frac{1}{k+1}\binom{2k}{k}$ and these are known as the Catalan numbers and represent the even moments of the semicircle distribution.

Partition terminology. Let π be a partition of a tuple $[k]$. Let π consist of disjoint blocks $V_1, V_2, \ldots, V_m$, for some $1 \le m \le k$. We arrange the blocks in the ascending order of their smallest element. For any block V_i, a *sub-block* is defined to be a subset of *consecutive integers* in the block. Two elements j and k in a block V_i are said to be *successive* if for all a between j and k, $a \notin V_i$.

Definition 2.3.1 (Special Symmetric Partition).
A partition π of a tuple $[k] = \{1, 2, ..., k\}$ is said to be a Special Symmetric partition if it satisfies the following:

- *All blocks of π are of even size.*

- *Let $V \in \pi$ be any arbitrary block, and let $a, b \in V$ be two successive elements in V with $b > a$. Then, either of the following is true:*

  1. *$b = a + 1$, or,*
  2. *between a and b there are sub-blocks of even size.* *In other words, there are blocks $V_1, V_2, \ldots, V_\ell$, such that there exist elements $\{a_{i_1}, a_{i_1+1}, \ldots, a_{i_1+k_1}\} \in V_1$, $\{a_{i_2}, \ldots, a_{i_2+k_2}\} \in V_2$, $\ldots, \{a_{i_l}, \ldots, a_{i_\ell+k_\ell}\} \in V_\ell$, with $a = a_{i_1} - 1$ and $b = a_{i_\ell+k_\ell} + 1$, such that $k_1, k_2, \ldots, k_\ell$ are even.*

We denote the class of Special Symmetric partitions as $SS(k)$. Note that for k odd, $SS(k) = \emptyset$. For example, take $\pi = \{\{1, 4, 5, 8\}, \{2, 3, 6, 7\}, \{9, 10\}\} \in SS(10)$. Note here that between 4 and 5 in the first block, there are no elements from the other blocks, and between 5 and 8, there is the sub-block $\{6, 7\}$ that is of even size.

In Bose et al. [2022] a more elaborate definition was given and this is useful in computations. Later, it was shown by [Pernici, 2021, Section 3] that the definition in Bose et al. [2022] is equivalent to the above one. In Pernici [2021], the set $SS(2k)$ is denoted by $P_2^{(2)}(k)$, a special subclass of k-divisible partitions. These partitions appeared as "Clickable Partitions" in Ryan [1998], where they were introduced to describe the limit distribution of dense random matrix models, and in the same spirit, they were also used for sparse random graphs in the paper Male [2017].

Remark 2.3.2.
We note down some important properties of $SS(k)$:

1. *If k is even, then*

$$\{\pi \in SS(k) : |\pi| = k/2\} = \{\pi \in NC_2(k)\}.$$

2. *$SS(2k) = NC(2k)$ for $1 \leq k \leq 3$. When $k \geq 4$, there are partitions $\pi \in SS(2k)$ that are either crossing or non-paired. For example, for $k = 8$, $\{\{1, 2, 5, 6\}, \{3, 4, 7, 8\}\}$ is a Special Symmetric partition. In particular, crossings start appearing when there are at least two or more blocks in a partition having 4 or more elements.*

3. *The set of Special Symmetric partitions are in one-to-one correspondence with coloured rooted trees (see [Bose et al., 2022, Lemma 5.1]) and these trees appeared first in the analysis in the works of Bauer and Golinelli [2001].*

Any partition $\pi \in \mathcal{P}(k)$ can be realized as a permutation of $[k]$, that is, a mapping from $[k] \to [k]$. Let S_k denote the set of permutations on k elements. Let $\gamma = (1, 2, \dots, k) \in S_k$ be the shift by 1 modulo k. We will be interested in the compositions of the two permutations γ and π, denoted by $\gamma\pi$, and this will be seen below as a partition.

Remark 2.3.3.
While π is a partition *and γ is a* permutation*, we do a composition in the permutation sense. We read the partition π as a permutation, compose it with the permutation γ, and finally read $\gamma\pi$ as a partition. As an example, consider $\pi = \{\{1, 2\}, \{3, 4\}\}$ and $\gamma = (1, 2, 3, 4)$. To compute $\gamma\pi$, we read π as $(1, 2)(3, 4)$, and compute $\gamma\pi = (1, 3)(2)(4)$. We finally read $\gamma\pi$ as $\{\{1, 3\}, \{2\}, \{4\}\}$.*

Definition 2.3.4 (Graph associated to a partition).
For a fixed $k \geq 1$, let γ denote the cyclic permutation $(1, 2, \dots, k)$. For a partition π, we define $G_{\gamma\pi} = (V_{\gamma\pi}, E_{\gamma\pi})$ as a rooted, labelled graph associated with any partition π of $[k]$, constructed as follows.

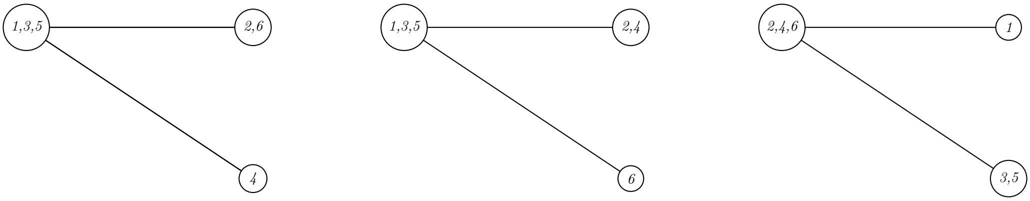
- *Initially consider the vertex set $V_{\gamma\pi} = [k]$ and perform a closed walk on $[k]$ as $1 \to 2 \to 3 \to \cdots \to k \to 1$ and with each step of the walk, add an edge.*
- *Evaluate $\gamma\pi$, which will be of the form $\gamma\pi = \{V_1, V_2, \dots, V_m\}$ for some $m \geq 1$ where $\{V_i\}_i$ are disjoint blocks. Then, collapse vertices in $V_{\gamma\pi}$ to a single vertex if they belong to the same block in $\gamma\pi$, and collapse the corresponding edges. Thus, $V_{\gamma\pi} = \{V_1, \dots, V_m\}$.*
- *Finally root and label the graph as follows.*
  - Root: *We always assume that the first element of the closed walk (in this case '1') is in V_1, and we fix the block V_1 as the root.*
  - Label: *Each vertex V_i gets labelled with the elements belonging to the corresponding block in $\gamma\pi$.*

Example 2.3.5.
Consider for example partitions of $k = 6$ and reading the partitions as permutations and evaluating their composition with γ gives us:

(a) $\pi_1 = \{\{1, 2, 5, 6\}, \{3, 4\}\}$,
(b) $\pi_2 = \{\{1, 2, 3, 4\}, \{5, 6\}\}$,
(c) $\pi_3 = \{\{1, 6\}, \{2, 3, 4, 5\}\}$.

(a) $\gamma\pi_1 = \{\{1, 3, 5\}, \{2, 6\}, \{4\}\}$,
(b) $\gamma\pi_2 = \{\{1, 3, 5\}, \{2, 4\}, \{6\}\}$,
(c) $\gamma\pi_3 = \{\{1\}, \{2, 4, 6\}, \{3, 5\}\}$.

The corresponding graphs $G_{\gamma\pi_1}, G_{\gamma\pi_2}$ and $G_{\gamma\pi_3}$ are as follows:



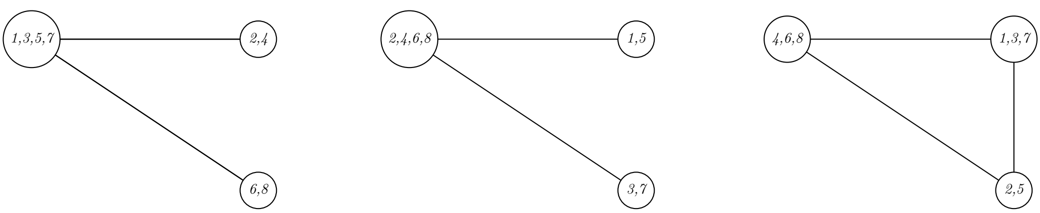
One can see that structurally the three graphs are the same. However, if we root them on V_1*, then the first two graphs are different from the third. Further, if we label the vertices as shown, all three graphs become distinct.*

Example 2.3.6.
Here, we illustrate the type of graph structures that can occur for $\pi \in SS(k)$*. Consider* $k = 8$*, and the following three partitions.*

(a) $\pi_1 = \{\{1,2,3,4\},\{5,6,7,8\}\}$.
(b) $\pi_2 = \{\{1,4,5,8\},\{2,3,6,7\}\}$.
(c) $\pi_3 = \{\{1,2,4,5\},\{3,6,7,8\}\}$.

(a) $\gamma\pi_1 = \{\{1,3,5,7\},\{2,4\},\{6,8\}\}$,
(b) $\gamma\pi_2 = \{\{(1,5\},\{2,4,6,8\},\{3,7\}\}$,
(c) $\gamma\pi_3 = \{\{1,3,7\},\{2,5\},\{4,6,8\}\}$.

Then, $\pi_1, \pi_2 \in SS(8)$ *but* $\pi_3 \notin SS(8)$*. Moreover,* π_1 *is non-crossing whereas* π_2 *has 2 crossings. The corresponding graphs are as below.*



The following result is the first main result of the article. This is an extension of the results obtained recently in Bose et al. [2022] and the homogeneous case obtained in Jung and Lee [2018].

Theorem 2.3.7 (Identification of moments).
(a) Let $\mathbf{A}_N$ *be the adjacency matrix of the sparse IER random graph as defined in* (2.8) *satisfying assumptions* ***A.1*–*A.3****. Then there exists a deterministic measure* μ_λ *such that*

$$\lim_{N\to\infty} \mathrm{ESD}(\mathbf{A}_N) = \mu_\lambda \text{ weakly in probability.}$$

Moreover, μ_λ is uniquely determined by its moments, which are given as follows:

$$m_k(\mu_\lambda) = \int x^k \mu_\lambda(\mathrm{d}\, x) = \begin{cases} 0, & k \text{ is odd,} \\ \sum_{l=2}^{k/2+1} \sum_{\substack{\pi \in SS(k): \\ |\gamma\pi| = l}} \lambda^{l-1-\frac{k}{2}}\, t(G_{\gamma\pi}, f, \mu_w), & k \text{ is even,} \end{cases} \tag{2.10}$$

where $SS(k)$ is the set of all Special Symmetric partitions of $[k]$ as defined in Definition 2.3.1, $G_{\gamma\pi}$ is the graph associated to a partition π as defined in Definition 2.3.4, and t is the homomorphism density as in (2.4)*.*

(b) As $\lambda \to \infty$,

$$\mu_\lambda \Rightarrow \mu_f,$$

where μ_f is the measure described in Theorem 2.2.1.

Remark 2.3.8.
Note that limiting second moment is given by $m_2 = t(G_{\gamma\pi}, f, \mu_w)$ where $\pi = \{1, 2\}$ and $\gamma\pi = \{\{1\}, \{2\}\}$. Hence $G_{\gamma\pi}$ is the graph with 2 vertices and 1 edge. Therefore

$$m_2(\mu_\lambda) = \int_{(0,\infty]^2} f(x, y) \mu_w(\mathrm{d}\, x) \mu_w(\mathrm{d}\, y)\,,$$

and hence the measure is non-degenerate.

§2.3.2 Stieltjes transform: Analytic approach

It is well-known that μ_λ can be characterised by its Stieltjes transform, which, in turn, can be characterised by a random recursive equation. Local weak convergence is a powerful tool for studying the Stieltjes transform of spectral measures associated with sparse random graphs. However, it becomes challenging to provide accurate estimates on the Stieltjes transform to study local laws and extreme values. Therefore, we present an alternative approach to studying the Stieltjes transform of the spectral measure of IER graphs. The ideas used here originate from the works of Khorunzhy et al. [2004].

We denote the upper half complex plane by

$$\mathbb{C}^+ = \{z \in \mathbb{C} : \, z = \zeta + \iota\eta, \, \eta > 0\}.$$

For an analytic approach to the problem, we analyse the *resolvent* of this matrix, defined as

$$\mathrm{R}_{\mathbf{A}_N}(z) := (\mathbf{A}_N - zI)^{-1}, \, z \in \mathbb{C}^+.$$

The Stieltjes transform of the empirical spectral distribution of $\mathbf{A}_N$ is given by

$$\mathrm{S}_{\mathbf{A}_N}(z) = \int_{\mathbb{R}} \frac{1}{x-z} \,\mathrm{ESD}(\mathbf{A}_N)(\mathrm{d}\, x) = \mathrm{tr}(\mathrm{R}_{\mathbf{A}_N}(z)),$$

where tr denotes the normalised trace. To get more refined estimates we need an additional assumption on the connectivity function:

A.4 $f : [0,\infty)^2 \to [0,\infty)$ is symmetric and bounded by a constant C_f. Moreover, f is Lipschitz in one coordinate, that is, for all $x_1, x_2, y \in [0,\infty)$,

$$|f(x_1,y) - f(x_2,y)| \le C_L |x_1 - x_2|$$

where C_L is the Lipschitz constant for f.

To state the result we will need a Banach space of analytic functions. Consider the space $\mathcal{B}$ defined by

$$\mathcal{B} = \left\{ \phi : [0,\infty) \times [0,\infty) \to \mathbb{C} \;\text{ analytic} \;\middle|\; \sup_{x,y\ge 0} \frac{|\phi(x,y)|}{\sqrt{1+y}} < \infty \right\} \tag{2.11}$$

and take the norm

$$\|\phi\|_{\mathcal{B}} = \sup_{x,y\ge 0} \frac{|\phi(x,y)|}{\sqrt{1+y}}.$$

Then, $(\mathcal{B}, \|\cdot\|_{\mathcal{B}})$ is a Banach space. We defer the proof of this in Proposition 2.6 in the appendix.

Consider the function $G_N : [0,\infty) \times \mathbb{C}^+$ given by

$$G_N(u,z) := \frac{1}{N} \sum_{i=1}^{N} \mathrm{e}^{iur_{ii}^N(z)} \tag{2.12}$$

where $r_{ii}^N(z) = \mathrm{R}_{\mathbf{A}_N}(z)(i,i)$, the i^{th} diagonal element of the resolvent of $\mathbf{A}_N$. It turns out that

$$\left.\frac{\partial G_N(u,z)}{\partial u}\right|_{u=0} = \mathrm{S}_{\mathbf{A}_N}(z)$$

and hence one can derive a form of the limiting Stieltjes transform.

Theorem 2.3.9 (Analytic functional of the resolvent).
*Let $\mathbf{A}_N$ be the adjacency of the IER random graph as defined in (2.8) and satisfying assumptions (**A.2**)–(**A.4**). Further, consider G_N as defined in (2.12). Define the function $d_f(x)$ as*

$$d_f(y) = \int_0^\infty f(x,y) \mu_w(\mathrm{d}\, x). \tag{2.13}$$

Then, for $z \in \mathbb{C}^+$ there exists a function $\phi^*(x,u) := \phi_z^*(x,u) \in \mathcal{B}$ such that for each $z \in \mathbb{C}^+$ and uniformly in $u \in (0,1]$ we have

$$\begin{aligned}&\lim_{N\to\infty} \mathbb{E}[G_N(u,z)] \\ &= 1 - \sqrt{u}\int_0^\infty \mathrm{e}^{-\lambda d_f(y)} \int_0^\infty \frac{\mathrm{J}_1(2\sqrt{uv})}{\sqrt{v}} \mathrm{e}^{ivz}\mathrm{e}^{\lambda\phi^*(y,v/\lambda)}\,\mathrm{d}\,v\,\mu_w(\mathrm{d}\,y)\end{aligned} \tag{2.14}$$

and

$$\mathrm{Var}[G_N(u,z)] \to 0.$$

Here, $\phi^* := \phi_z^*$ is a unique analytic solution (in the space $\mathcal{B}$) for the fixed point equation:

$$\begin{aligned}&\phi^*(x,u) \\ &= F_z(\phi^*)(x,u) \\ &= d_f(x) - \int_0^\infty f(x,y)\mathrm{e}^{-\lambda d_f(y)} \left(\sqrt{u}\int_0^\infty \frac{\mathrm{J}_1(2\sqrt{uv})}{\sqrt{v}} \mathrm{e}^{ivz}\mathrm{e}^{\lambda\phi^*(y,\frac{v}{\lambda})}\,\mathrm{d}\,v\right)\mu_w(\mathrm{d}\,y),\end{aligned} \tag{2.15}$$

where J_1 is the Bessel function of the first order of the first kind defined as

$$\mathrm{J}_1(x) = \frac{x}{2}\sum_{k=0}^{\infty} \frac{(-1)^k (x^2/4)^k}{k!(k+1)!}. \tag{2.16}$$

Observe that there is a slight difference in the right-hand sides of (2.14) and (2.15) but in the case $f \equiv 1$ both are the same. The next corollary describes the convergence of the Stieltjes transform.

Corollary 2.3.10 (Identification of the Stieltjes Transform).
Under the assumptions of the above theorem, we have that any $z \in \mathbb{C}^+$,

$$\mathrm{S}_{\mathbf{A}_N}(z) \to \mathrm{S}_{\mu_\lambda}(z) \text{ in probability,}$$

where μ_λ is as in Theorem 2.3.7. The $\mathrm{S}_{\mu_\lambda}(\cdot)$ satisfies the following equation:

$$\mathrm{S}_{\mu_\lambda}(z) = i\int_0^\infty \mathrm{e}^{-\lambda d_f(y)} \int_0^\infty \mathrm{e}^{ivz}\mathrm{e}^{\lambda\phi_z^*(y,\frac{v}{\lambda})}\,\mathrm{d}\,v\;\;\mu_w(\mathrm{d}\,y),\;\; z \in \mathbb{C}^+. \tag{2.17}$$

To recover the dense regime, we study the asymptotic $\lambda \to \infty$ as in the next corollary.

Corollary 2.3.11 (Stieltjes Transform as $\lambda \to \infty$).
For $\lambda \to \infty$, we have that

$$\lim_{\lambda \to \infty} \mathrm{S}_{\mu_\lambda}(z) = \mathrm{S}_{\mu_f}(z) \tag{2.18}$$

for each $z \in \mathbb{C}^+$, where $\mathrm{S}_{\mu_f}(z)$ satisfies an integral equation given by

$$\mathrm{S}_{\mu_f}(z) := \int_0^\infty \mathcal{H}(z,x)\mu_w(\mathrm{d}\,x)\,, \tag{2.19}$$

where $\mathcal{H}(z,x)$ satisfies the f dependent fixed point equation (2.6).

Remark 2.3.12 (The case $f \equiv 1$).
In the case when $f \equiv 1$, we recover the homogeneous setting. We know ϕ_z^* satisfies the fixed point equation (2.15). If we substitute $f \equiv 1$ in (2.15) we get

$$\phi^*(x,u) = 1 - \sqrt{u}\int_0^\infty \mathrm{e}^{-\lambda}\left(\sqrt{u}\int_0^\infty \frac{\mathrm{J}_1(2\sqrt{uv})}{\sqrt{v}}\mathrm{e}^{ivz}\mathrm{e}^{\lambda\phi^*(y,\frac{v}{\lambda})}\,\mathrm{d}\,v\right)\mu_w(\mathrm{d}\,y)\,.$$

We see that the right-hand side has no dependency on the parameter x, and so, we have a unique analytical functional $\widetilde{\phi^*}(u) = \phi^*(x,u)$ that satisfies the fixed point equation

$$\widetilde{\phi^*}(u) = 1 - \mathrm{e}^{-\lambda}\sqrt{u}\int_0^\infty \frac{\mathrm{J}_1(2\sqrt{uv})}{\sqrt{v}}\mathrm{e}^{ivz}\mathrm{e}^{\lambda\widetilde{\phi^*}(v/\lambda)}\,\mathrm{d}\,v\,. \tag{2.20}$$

This matches the result of Khorunzhy et al. [2004].

From Example 2 of Bordenave and Lelarge [2010], we have that $\widetilde{\phi_z^*}$ has the form $\widetilde{\phi_z^*}(u) = \mathbb{E}[\mathrm{e}^{iuX(z)}]$ for each $z \in \mathbb{C}^+$, where $X(z)$ has the law Q as described in (2.7). So, for any $z \in \mathbb{C}^+$, we have

$$\mathrm{S}_{\mu_\lambda}(z) = i\int_0^\infty \mathrm{e}^{ivz}e^{-\lambda+\lambda\mathbb{E}\left[\mathrm{e}^{i\frac{v}{\lambda}X(z)}\right])}\,\mathrm{d}\,v = i\int_0^\infty \mathrm{e}^{ivz}\varphi_P\left(\mathbb{E}\left[\mathrm{e}^{i\frac{v}{\lambda}X(z)}\right]\right)\mathrm{d}\,v\,,$$

where

$$\varphi_P(z) = \mathbb{E}[z^P] = \mathrm{e}^{\lambda(z-1)},\ ,P \sim \mathrm{Poi}(\lambda)\,.$$

§2.3.3 Examples

We now list out a few examples of the model that can be approached by our methods.

Example 1: Homogeneous Erdős-Rényi Random Graph. When we have $f \equiv 1$, the model reduces to the standard homogeneous Erdős-Rényi graph with edge probability $p = \lambda/N$. As discussed, in this case the moments of μ_λ can be computed. In particular, we have $t(G_{\gamma\pi}, f, \mu_w) = 1$ for all π. Hence we have

$$\begin{aligned} m_{2k}(\mu_\lambda) &= \sum_{l=1}^{k} \lambda^{l-k} |\{\pi \in SS(2k) : |\pi| = l\}| \\ &= |NC_2(2k)| + \sum_{l=1}^{k-1} \lambda^{l-k} |\{\pi \in SS(2k) : |\pi| = l\}| \,. \end{aligned}$$

Since the (even) moments of the semicircle law are given by the Catalan numbers, it is immediate that

$$\lim_{\lambda\to\infty} m_{2k}(\mu_\lambda) = m_{2k}(\mu_s).$$

Hence Theorem 2.3.7(b) is true in this special case. It is known that μ_λ has an absolutely continuous spectrum when $\lambda > 1$ (see Bordenave et al. [2017], Arras and Bordenave [2023]). In this case, the Stieltjes transform is given by

$$\mathrm{S}_{\mu_\lambda}(z) = -i \int_0^\infty \mathrm{e}^{ivz} \mathrm{e}^{-\lambda + \lambda \widetilde{\phi^*}(v/\lambda)} \,\mathrm{d}\, v \,,$$

and $\widetilde{\phi^*}(v/\lambda)$ satisfies the equation (2.20). What is interesting and cannot be immediately derived from our results is the rate of convergence of the measure μ_λ to μ_s as λ becomes large. In the simulation below we consider the $\lambda = 10$ and the simulation already suggests the appearance of semicircle law. We believe the representation above of the Stieltjes transform as in Corollary 2.3.10 can be used to prove the rate of convergence as done in the classical Wigner case in Bai [1999].

Example 2: Chung-Lu Random Graph. Let $(d_i)_{i\in[n]}$ be a graphical sequence and denote by $m_1 = \sum_i d_i$ and $m_\infty = \max_i d_i$, the total and the maximum degree, respectively. Let f be defined on $[0,1]^2$ as

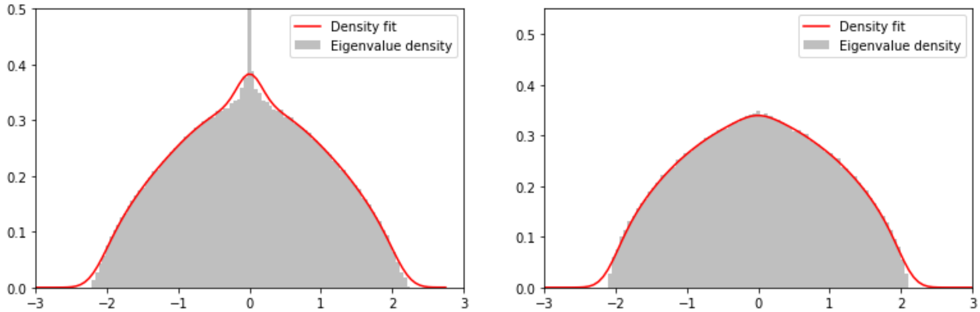
$$f(x,y) = xy \wedge 1$$

and

$$w_i = \frac{d_i}{m_\infty}, \quad \varepsilon_N = \frac{m_\infty^2}{m_1}.$$

We can choose an appropriate degree sequence $(d_i)_{i\geq 1}$ such that $m_\infty = \mathrm{o}(\sqrt{m_1})$ and $N\varepsilon_N \to \lambda$. The connection probabilities will be given by

$$p_{ij}^{\mathrm{cl}} = \varepsilon_N \left(\frac{d_i d_j}{m_\infty^2} \wedge 1 \right) = \frac{d_i d_j}{m_1}.$$



(a) $\lambda = 5$. *(b)* $\lambda = 10$.

Figure 2.1: The homogeneous Erdős-Rényi Random Graph on 10,000 vertices.

Let o_N be a uniformly chosen vertex and d_{o_N} be the degree of this vertex. We assume that

$$\frac{d_{o_N}}{m_\infty} \xrightarrow{d} W$$

where W has law μ_w which is compactly supported. Then the conditions of Theorem 2.3.7 are satisfied. Hence there exists a limiting spectral distribution which we call $\mu_{CL,\lambda}$ and the even moments can identified in the following way.

Let $SS_\ell(2k)$ be the set of Special Symmetric partitions with ℓ blocks. Then,

$$\begin{aligned}\int_{\mathbb{R}} x^{2k} \mu_{CL,\lambda}(\mathrm{d}\, x) &= \sum_{\ell=1}^{k} \sum_{\pi \in SS_\ell(2k)} \lambda^{\ell-k} t(G_{\gamma\pi}, f, \mu_w) \\ &= \sum_{\ell=1}^{2k} \sum_{\pi \in SS_\ell(2k)} \lambda^{\ell-k} \prod_{j=1}^{|\gamma\pi|} \int_{\mathbb{R}} x^{b_j(\gamma\pi)} \mu_w(\mathrm{d}\, x)\,,\end{aligned}$$

where $b_1(\sigma), \cdots, b_{\#\sigma}(\sigma)$ denotes the size of the blocks of a partition σ. For $\sigma \in NC_2(k)$, its *Kreweras complement* $K(\sigma)$ is the maximal non-crossing partition $\bar{\sigma}$ of $\{\bar{1}, \ldots, \bar{k}\}$, such that $\sigma \cup \bar{\sigma}$ is a non-crossing partition of $\{\bar{1}, 1, \ldots, \bar{k}, k\}$. For example,

$$\begin{aligned} K\left(\{\{1,2\},\{3,4\},\{5,6\}\}\right) &= \{\{1\},\{2,4,6\},\{3\},\{5\}\}, \\ K\left(\{(\{1,2\},\{3,6\},\{4,5\},\{7,8\}\}\right) &= \{\{1,3,7\},\{4,6\},\{2\},\{5\},\{8\}\}. \end{aligned}$$

Note that this slightly differs from the standard notation of Kreweras complement in Nica and Speicher [2006] but for pairings, the π and π^{-1} coincide. It follows easily that when $\pi \in NC_2(2k)$, $\gamma\pi$ can be replaced by $K(\pi)$. The benefit of this representation is the following. It follows from [Nica and Speicher, 2006, Page 228] that

$$\int_{\mathbb{R}} x^{2k}(\mu_w \boxtimes \mu_s)(\mathrm{d}\, x) = \sum_{\pi \in NC_2(2k)} \prod_{j=1}^{k+1} \int_{\mathbb{R}} x^{b_j(K(\pi))} \mu_w(\mathrm{d}\, x),$$

where $\mu_w \boxtimes \mu_s$ is the free multiplicative convolution of the measures μ_w and semicircle law μ_s. Hence the moments of $\mu_{CL,\lambda}$ can be written as

$$\int_{\mathbb{R}} x^{2k} \mu_{CL,\lambda}(\mathrm{d}\, x) = \int_{\mathbb{R}} x^{2k}(\mu_w \boxtimes \mu_s)(\mathrm{d}\, x) \\ + \sum_{\ell=1}^{k-1} \sum_{\pi \in SS_\ell(2k)} \lambda^{\ell-k} \prod_{j=1}^{|\gamma\pi|} \int_{\mathbb{R}} x^{b_j(\gamma\pi)} \mu_w(\mathrm{d}\, x)\,.$$

This also shows that

$$\lim_{\lambda\to\infty} \int_{\mathbb{R}} x^{2k} \mu_{CL,\lambda}(\mathrm{d}\, x) = \int_{\mathbb{R}} x^{2k}(\mu_w \boxtimes \mu_s)(\mathrm{d}\, x),$$

and consequently, μ_f is of the form $\mu_w \boxtimes \mu_s$.

Remark 2.3.13.
We want to add a remark about heavy-tailed degrees. Our conditions are not satisfied when the degree sequence follows a power-law distribution. In that case, the w_i need to be scaled differently, and the limiting W will not have a compact support. For further discussion on inhomogeneous random graphs with heavy tails, we refer to [van der Hofstad, 2017, Chapter 6].

Example 3: Generalised random graph. Again, let (d_i) be as above. Let $f(x, y) = \frac{xy}{1+xy}$ and $w_i = d_1/\sqrt{m_1}$. Then,

$$p_{ij}^{\mathrm{grg}} = \frac{d_i d_j}{m_1 + d_i d_j}.$$

Although the above example does not directly fall in our set-up (due to lack of ε_N), one can still derive the limiting spectral distribution using the Chung-Lu model. We will use the following two facts. The first is a fact, which is the Hoffman-Wielandt inequality from [Bai, 1999, Corollary A.41].

Fact 2.3.14.
If d_L denotes the Lévy distance between two probability measures, then for $N \times N$ symmetric matrices A and B,

$$d_L^3\left(\mathrm{ESD}(A), \mathrm{ESD}(B)\right) \leq \frac{1}{N} \operatorname{Tr}\left((A-B)^2\right)\,.$$

The following is a fact about the coupling of two Bernoulli random variables with parameters p and q (see [van der Hofstad, 2024, Theorem 2.9])

Fact 2.3.15.
There exits a coupling between $X \sim \mathrm{Ber}(p)$ and $Y \sim \mathrm{Ber}(q)$ such that

$$\mathbb{P}(X \neq Y) = |p - q|.$$

Using the above coupling, we can construct a sequence of independent Bernoulli random variables (b_{ij}) and (c_{ij}) with parameters p_{ij}^{cl} and q_{ij}^{grg}, respectively. Let $\mathbf{M}_N^{\mathrm{cl}}$ and $\mathbf{M}_N^{\mathrm{grg}}$ be the adjacency matrices of Chung-Lu and generalised random graph models, respectively, with the above coupled Bernoulli random variables. Suppose the sequence $(d_i)_{i\in[n]}$ satisfies the assumptions described in Example 2 and let $N\varepsilon_N \to \lambda$ and $\mathbf{A}_N^{\mathrm{cl}} = \lambda^{-1/2}\mathbf{M}_N^{\mathrm{cl}}$ and $\mathbf{A}_N^{\mathrm{grg}} = \lambda^{-1/2}\mathbf{M}_N^{\mathrm{grg}}$. Then,

$$\begin{aligned}
\mathbb{E}\left[d_L^3\left(\mathrm{ESD}(\mathbf{A}_N^{\mathrm{cl}}), \mathrm{ESD}(\mathbf{A}_N^{\mathrm{grg}})\right)\right] &\le \frac{1}{N}\mathbb{E}\left[\mathrm{Tr}(\mathbf{A}_N^{\mathrm{cl}} - \mathbf{A}_N^{\mathrm{grg}})^2\right] \\
&= \frac{1}{N\lambda}\mathbb{E}\left[\sum_{i,j=1}^{N}(b_{ij} - c_{ij})^2\right] \\
&= \frac{1}{\lambda N}\mathbb{E}\left[\sum_{i,j=1}^{N}(b_{ij} - c_{ij})^2 \mathbb{1}_{\{b_{ij}\neq c_{ij}\}}\right] \\
&\le \frac{1}{\lambda N}\sum_{i,j=1}^{N}\mathbb{P}(b_{ij} \neq c_{ij}) \le \frac{1}{\lambda N}\sum_{i,j=1}^{N}\left|p_{ij}^{\mathrm{cl}} - p_{ij}^{\mathrm{grg}}\right|,
\end{aligned}$$

since $(b_{ij} - c_{ij})^2$ can be trivially bounded by 1. Using $x - \frac{x}{1+x} \le \frac{x^2}{1+x} \le x^2$ for any $x > 0$, we have

$$p_{ij}^{\mathrm{cl}} - p_{ij}^{\mathrm{grg}} = \frac{d_i d_j}{m_1} - \frac{d_i d_j}{m_1 + d_i d_j} \le \frac{d_i^2 d_j^2}{m_1^2} \le \frac{m_\infty^4}{m_1^2}.$$

Therefore

$$\begin{aligned}
\mathbb{E}\left[d_L^3\left(\mathrm{ESD}(\mathbf{A}_N^{\mathrm{cl}}), \mathrm{ESD}(\mathbf{A}_N^{\mathrm{grg}})\right)\right] &\le \frac{C}{\lambda N}\sum_{i,j=1}^{N}\frac{m_\infty^4}{m_1^2} \\
&= \frac{C}{\lambda N}N^2\frac{m_\infty^4}{m_1^2} \le \mathrm{O}\left(\frac{N m_\infty^4}{m_1^2}\right).
\end{aligned}$$

If we consider $m_\infty = o(m_1^{1/4})$, then the empirical distribution functions are close. Now using Markov inequality and the fact that $\mathrm{ESD}(\mathbf{A}_N^{\mathrm{cl}})$ converges weakly in probability to $\mu_{CL,\lambda}$ it follows that

$$\lim_{N\to\infty}\mathrm{ESD}(\mathbf{A}_N^{\mathrm{grg}}) = \mu_{CL,\lambda} \text{ weakly in probability.}$$

Example 4: Norros-Reittu. Let $(d_i)_i$ be a given sequence and $w_i = \frac{d_i}{\sqrt{m_1}}$. Take $f(x, y) = 1 - \exp(-xy)$. Then,

$$p_{ij}^{\mathrm{nr}} = 1 - \exp\left(-\frac{d_i d_j}{m_1}\right).$$

Again, the form of the above connection probability does not fall directly in our set-up, but we can show that Norros-Reittu model is close to the generalised random graph models. Let $\mathbf{A}_N^{\mathrm{nr}} = \lambda^{-1/2}\mathbf{M}_N^{\mathrm{nr}}$ where $\mathbf{M}_N^{\mathrm{nr}}$ is the adjacency of the Norros-Reittu model. Without loss of generality, we assume that we can couple Bernoulli random variable c_{ij} and d_{ij} with parameters p_{ij}^{grg} and p_{ij}^{nr} using Fact 2.3.15. Just as in the previous example, it follows using Fact 2.3.14 that

$$\mathbb{E}\left[d_L^3\left(\mathrm{ESD}(\mathbf{A}_N^{\mathrm{grg}}), \mathrm{ESD}(\mathbf{A}_N^{\mathrm{nr}})\right)\right] \le \frac{1}{\lambda N}\sum_{i,j=1}^{N} \mathbb{E}\left[(c_{ij} - d_{ij})^2 \mathbf{1}_{\{c_{ij} \neq d_{ij}\}}\right].$$

We bound trivially $(c_{ij} - d_{ij})^2$ by a constant $C_1 > 0$ and hence we get that

$$\begin{aligned}\mathbb{E}\left[d_L^3\left(\mathrm{ESD}(\mathbf{A}_N^{\mathrm{grg}}), \mathrm{ESD}(\mathbf{A}_N^{\mathrm{nr}})\right)\right] &\le \frac{C_1}{\lambda N}\sum_{i,j=1}^{N} \mathbb{P}\left(c_{ij} \neq d_{ij}\right) \\ &= \frac{C_1}{\lambda N}\sum_{i\neq j}(p_{ij}^{\mathrm{nr}} - p_{ij}^{\mathrm{grg}}).\end{aligned}$$

Now, for $i \neq j$,

$$\begin{aligned}p_{ij}^{\mathrm{nr}} - p_{ij}^{\mathrm{grg}} &= \left(1 - \exp\left(-\frac{d_i d_j}{m_1}\right) - \frac{d_i d_j}{m_1 + d_i d_j}\right) \\ &= \left(\frac{d_i^2 d_j^2}{m_1^2 + m_1 d_i d_j}\right) + \frac{\lambda}{N}\mathrm{O}\left(\frac{d_i^2 d_j^2}{m_1^2}\right) \\ &\le C' \frac{d_i^2 d_j^2}{m_1^2},\end{aligned}$$

for some constant $C' > 0$. Therefore, for some new constant $C_1' > 0$,

$$\mathbb{E}\left[d_L^3\left(\mathrm{ESD}(\mathbf{A}_N^{\mathrm{grg}}), \mathrm{ESD}(\mathbf{A}_N^{\mathrm{nr}})\right)\right] \le \frac{C_1'}{\lambda N}\frac{m_2^2}{m_1^2} \tag{2.21}$$

where $m_2 = \sum_{i=1}^{N} d_i^2$. Since W has compact support, we have that $\frac{m_2}{N m_\infty} \to \mathbb{E}[W^2]$ and $\frac{m_1}{N m_\infty} \to \mathbb{E}[W]$. So $\frac{m_2^2}{m_1^2}$ is bounded for large N and hence the right hand side of (2.21) goes to 0. This shows that

$$\lim_{N\to\infty} \mathrm{ESD}(\mathbf{A}_N^{\mathrm{nr}}) = \mu_{CL,\lambda} \text{ weakly in probability.}$$

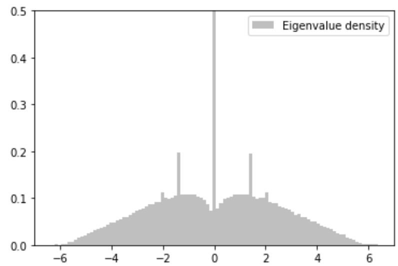
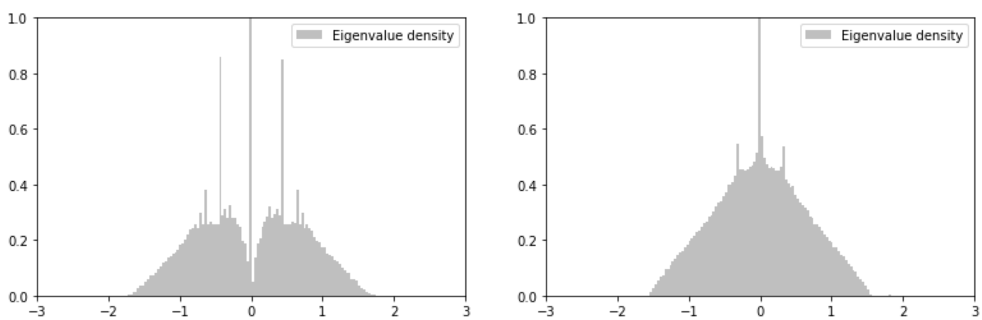


Figure 2.2: Spectral distributions for the Chung-Lu random graph, the generalised random graph, and the Norros-Reittu random graph on 10,000 vertices with $\{d_i\}_i$ *uniformly generated integers in* $[1, 5]$

Example 5: Inhomogeneous Random Graphs. Let $w_i = \frac{i}{N}$ and $f : [0,1]^2 \to [0,1]$ be any continuous function. Then,

$$p_{ij} = \varepsilon_N f\left(\frac{i}{N}, \frac{j}{N}\right).$$

This is a case which falls directly into our set-up if we assume $N\varepsilon_N \to \lambda$ and the measure μ_w is the Lebesgue measure. The other examples considered in this section are mostly of the rank-1 type but through this example, one can achieve limiting measures which are of a wide variety.



(a) $\lambda = 5$. *(b)* $\lambda = 10$.

Figure 2.3: The Inhomogeneous Random Graph on 10,000 vertices, with the inhomogeneity function $f(x, y) = r_1(x)r_1(y) + r_2(x)r_2(y)$*, where* $r_1(x) = \frac{x}{1+x}$ *and* $r_2(x) = x$.

We note that in van der Hofstad [2024], inhomogeneous random graphs are introduced in a much more abstract setting, following the works of Bollobás et al. [2007]. The connectivity function f is generally continuous and also satisfies reducibility properties. The above examples also fall under the setup described there.

§2.4 Existence, uniqueness, and moments

In this section we will prove the main result Theorem 2.3.7 using the method of moments.

We begin with a small observation. Recall from Assumption **A.3** that if o_N is an uniformly chosen vertex and $W_N = w_{o_N}$ and we assume $W_N \xrightarrow{\text{d}} W$. This means that W_N has a distribution function $F_N(x)$ given by

$$F_N(x) = \frac{1}{N} \sum_{i=1}^{N} \mathbf{1}_{\{w_i \le x\}}$$

and if we denote by F the distribution of W then for any continuity point x of F we have

$$F_N(x) \to F(x).$$

Also for any bounded continuous function g, we have $\mathbb{E}[g(W_N)] \to \mathbb{E}[g(W)]$. Let $o_1, \ldots, o_k$ be i.i.d. Uniform random variables on $[N]$. Let $W_{N,i} = w_{o_i}$ for $i = 1, \ldots, k$. Then

$$(W_{N,1}, ..., W_{N,k}) \xrightarrow{\text{d}} (W_1, W_2, ..., W_k)$$

where $W_1, \ldots, W_k$ are k independent copies of the limiting variable W. Hence for any bounded continuous g in k-variables we have

$$\mathbb{E}\left[g(W_{N,1}, \ldots, W_{N,k})\right] \to \mathbb{E}\left[g(W_1, \ldots, W_k)\right]. \tag{2.22}$$

In our model, we can allow self-loops as we are not imposing that $f(x, x) = 0$ but the presence of self-loops does not affect the ESD. The following lemma shows that we can remove the self-loops.

Lemma 2.4.1 (Diagonal contribution).
Let $\widetilde{\mathbf{A}}_N$ be the matrix $\mathbf{A}_N$ with zero on the diagonal, and let d_L denote the Lévy distance. Then,

$$d_L\left(\mathrm{ESD}(\widetilde{\mathbf{A}}_N), \mathrm{ESD}(\mathbf{A}_N)\right) \xrightarrow{\mathbb{P}} 0\,.$$

In particular, if $\mathrm{ESD}(\mathbf{A}_N)$ *converges weakly in probability to* μ_λ*, then so will* $\mathrm{ESD}(\widetilde{\mathbf{A}}_N)$ *and visa-versa.*

Proof. Let D_N denote the diagonal of $\mathbf{A}_N$. Then, $D_N = \mathbf{A}_N - \tilde{\mathbf{A}}_N$. Using Fact 2.3.14 we have

$$d_L^3\left(\mathrm{ESD}(\widetilde{\mathbf{A}}_N), \mathrm{ESD}(\mathbf{A}_N)\right) \le \frac{1}{N}\,\mathrm{Tr}(D_N^2) = \frac{1}{N\lambda} \sum_{1 \le i \le N} a_{ii}^2\,.$$

Chapter Two

Hence we have

$$\begin{aligned}\mathbb{E}\left[d_L^3\left(\mathrm{ESD}(\widetilde{\mathbf{A}}_N),\mathrm{ESD}(\mathbf{A}_N)\right)\right] &\leq \frac{\sqrt{\lambda}}{N^2}\sum_{1\leq i\leq N} f\left(w_i,w_i\right)\\ &\leq C_f\frac{\sqrt{\lambda}}{N},\end{aligned}$$

for some constant C_f, which comes from the fact that f is bounded. The result follows using Markov's inequality. □

We are now ready to begin with the proofs of the main results.

§2.4.1 Expected Moments

We split up the proof into three parts. To ease the notation we abbreviate the empirical spectral distribution and its expectation as

$$\mu_{N,\lambda}(\cdot)=\mathrm{ESD}(\mathbf{A}_N)(\cdot) \quad \text{and} \quad \bar{\mu}_{N,\lambda}(\cdot)=\mathbb{E}[\mathrm{ESD}[\mathbf{A}_N]](\cdot)=\frac{1}{N}\sum_{i=1}^{N}\mathbb{P}(\lambda_i\in\cdot)\,. \tag{2.23}$$

Note that $\bar{\mu}_{N,\lambda}$ is now a deterministic measure, for which we compute the moments as

$$\int x^k\bar{\mu}_{N,\lambda}(\mathrm{d}\,x)=\frac{1}{N}\sum_{i=1}^{N}\int_{\mathbb{R}}x^k\mathbb{P}(\lambda_i\in\mathrm{d}\,x)=\frac{1}{N}\sum_{i=1}^{N}\mathbb{E}[\lambda_i^k]=\mathbb{E}[\mathrm{tr}(\mathbf{A}_N^k)]\,,$$

where tr denotes the normalised trace. Using the trace formula it follows that

$$\mathbb{E}[\mathrm{tr}(\mathbf{A}_N^k)]=\frac{1}{N}\mathbb{E}[\mathrm{Tr}(\mathbf{A}_N^k)]=\frac{1}{N\lambda^{k/2}}\sum_{1\leq i_1,i_2,\ldots,i_k\leq N}\mathbb{E}[a_{i_1i_2}a_{i_2i_3}\ldots a_{i_ki_1}]\,, \tag{2.24}$$

where a_{ij} are entries of the adjacency matrix $\mathbf{M}$. We compute the expected moments and demonstrate that they are finite. Subsequently, we establish a concentration result to show that the moments of the empirical measure converge to m_k in probability. Next, we prove that the sequence m_k satisfies Carleman's condition, thereby uniquely determining the limiting measure.

Let $SS(k)$ be the set of Special Symmetric partitions, and $\gamma=(1,2,\ldots,k)$ be the cyclic permutation. For the following computations, one has to read the partition π as a permutation, with elements of a block in the partition set in an ascending manner in the permutation. That is, if $\pi=\{\{1,2,5,6\},\{3,4\}\}$, then the corresponding permutation is $(1,2,5,6)(3,4)$.

Lemma 2.4.2 (Expected moments).
Let $\mu_{N,\lambda}$ be the ESD *of $\mathbf{A}_N$ and $\bar{\mu}_{N,\lambda} = \mathbb{E}\mu_{N,\lambda}$. Let $\gamma\pi$ be decomposed into blocks of the form*

$$\gamma\pi = \{V_1, V_2, \ldots, V_m\}.$$

where $m = |\gamma\pi|$ be the number of blocks. Define $\mathcal{F}_{\gamma\pi}$ as

$$\mathcal{F}_{\gamma\pi} := \{\mathbf{i} \in \mathbb{N}^k \mid i_j = i_{j'} \text{ if and only if there exists } l \in [m] \text{ s.t. } j, j' \in V_l\}. \tag{2.25}$$

Then,

$$\int x^k \bar{\mu}_{N,\lambda}(\mathrm{d}\, x) = \tag{2.26}$$

$$\begin{cases} \mathrm{O}(\lambda^{k/2}N^{-1}), & k \text{ odd} \\ \sum\limits_{\pi \in SS(k)} \lambda^{|\gamma\pi|-1-k/2} \sum\limits_{\mathbf{i} \in \mathcal{F}_{\gamma\pi}} \frac{1}{N^{|\gamma\pi|}} \prod\limits_{(a,b) \in E_{\gamma\pi}} f(w_{i_a}, w_{i_b}) + \mathrm{O}(\lambda^{k/2}N^{-1}), & k \text{ even} \end{cases}. \tag{2.27}$$

Example 2.4.3.
For $k = 4$, take $\pi = \{\{1,2\},\{3,4\}\}$. Then, $\gamma\pi = \{\{1,3\},\{2\},\{4\}\}$. We see that tuples of the form $(1,2,1,3)$ and $(2,3,2,4)$ belong in $\mathcal{F}_{\gamma\pi}$.

Proof of Lemma 2.4.2. Recall from (2.24) that

$$\frac{1}{N\lambda^{k/2}} \mathbb{E}[\mathrm{Tr}(\mathbf{A}_N^k)] = \frac{1}{N\lambda^{k/2}} \sum_{\mathbf{i} \in N^k} \mathbb{E}[a_{i_1 i_2} a_{i_2 i_3} \ldots a_{i_k i_1}],$$

where $\mathbf{i} = (i_1, \ldots, i_k)$. The term $a_{i_1 i_2} a_{i_2 i_3} \ldots a_{i_k i_1}$ is associated with the closed walk $i_1 i_2 \ldots i_k i_1$. Let the set of distinct vertices and edges along a closed walk correspond to a k-tuple $\mathbf{i}$ be denoted by $V(\mathbf{i})$ and $E(\mathbf{i})$, respectively. An edge that connects vertices i_j and i_{j+1}, will be denoted by $e = (i_j, i_{j+1})$. Without loss of generality, we assume that in $V(\mathbf{i})$ we assign the positions where the first of distinct indices appear in $\mathbf{i}$.

For example, for the 4-tuple $\mathbf{i} = (1,2,1,3)$, we have $V(\mathbf{i}) = \{1,2,4\}$. So, $E(\mathbf{i}) = \{(1,2),(1,4)\}$. Since

$$a_{i_1 i_2} a_{i_2 i_3} \ldots a_{i_k i_1} = 1 \text{ if and only } a_l a_{l+1} = 1 \text{ for all } (l, l+1) \in E(\mathbf{i})$$

we can rewrite (2.24) as

$$\frac{1}{N\lambda^{k/2}} \mathbb{E}[\mathrm{Tr}(\mathbf{A}_N^k)] = \frac{1}{N\lambda^{k/2}} \sum_{1 \le i_j \le N : j \in V(\mathbf{i})} \left(\frac{\lambda}{N}\right)^{|E(\mathbf{i})|} \prod_{(a,b) \in E(\mathbf{i})} f(w_{i_a}, w_{i_b}). \tag{2.28}$$

Let π be a partition of $[k] := \{1, 2, \ldots, k\}$ and $\gamma\pi = \{V_1, V_2, \ldots, V_m\}$, where $m = |\gamma\pi|$. Recall the definition of $\mathcal{F}_{\gamma\pi}$ as in (2.25) and also the graph $G_{\gamma\pi}$ corresponding to $\gamma\pi$ as in Definition 2.3.4. Note that for a fixed $\mathbf{i} \in \mathcal{F}_{\gamma\pi}$, $V(\mathbf{i}) = V_{\gamma\pi}$ and $E(\mathbf{i}) = E_{\gamma\pi}$. Moreover, if $\mathbf{i}, \mathbf{i}' \in \mathcal{F}_{\gamma\pi}$, then $V(\mathbf{i}) = V(\mathbf{i}')$ and $E(\mathbf{i}) = E(\mathbf{i}')$. Using this formulation, we can rewrite our summation in (2.28) once again as

$$\frac{1}{N\lambda^{k/2}}\mathbb{E}[\mathrm{Tr}(\mathbf{A}_N^k)] = \frac{1}{N\lambda^{k/2}} \sum_{\pi\in\mathcal{P}(k)} \sum_{\mathbf{i}\in\mathcal{F}_{\gamma\pi}} \left(\frac{\lambda}{N}\right)^{|E_{\gamma\pi}|} \prod_{(a,b)\in E_{\gamma\pi}} f\left(w_{i_a}, w_{i_b}\right).$$

Since $|\gamma\pi| = |V(\mathbf{i})|$, we can multiply and divide by $N^{|\gamma\pi|}$ to get

$$\begin{aligned}&\frac{1}{N\lambda^{k/2}}\mathbb{E}[\mathrm{Tr}(\mathbf{A}_N^k)] \\ &= \sum_{\pi\in\mathcal{P}(k)} \frac{1}{N^{|\gamma\pi|}} \sum_{\mathbf{i}\in\mathcal{F}_{\gamma\pi}} \lambda^{|E_{\gamma\pi}|-k/2} N^{|\gamma\pi|-|E_{\gamma\pi}|-1} \prod_{(a,b)\in E_{\gamma\pi}} f\left(w_{i_a}, w_{i_b}\right).\end{aligned}$$

Note that since f is bounded, then the product is bounded. For a fixed k and a partition π of $[k]$, $|E_{\gamma\pi}| \le k$. One can also see that $|\mathcal{F}_{\gamma\pi}| \sim N^{|\gamma\pi|}$. We thus focus only on $\lambda^{|E_{\gamma\pi}|-k/2}N^{|\gamma\pi|-|E_{\gamma\pi}|-1}$. For this to contribute, a tuple $\mathbf{i}$ must yield a tree structure in $G_{\gamma\pi}$, this will give us $|V(\mathbf{i})| = |E(\mathbf{i})| + 1$, which would imply $|\gamma\pi| = |E_{\gamma\pi}| + 1$. In particular, all tuples $\mathbf{i} \in \mathcal{F}_{\gamma\pi}$ such that $G_{\gamma\pi}$ is a coloured rooted tree as defined in Definition 2.3.4 contribute to the summation.

For other graphs with $|V(\mathbf{i})| < |E(\mathbf{i})| + 1$, the leading error would be of the order $\mathrm{O}(N^{-1})$. The leading order error is given when $G_{\gamma\pi}$ is a k-cycle and hence the error is of the order of $\lambda^{k/2}N^{-1}$. Thus, our sum reduces to

$$\begin{aligned}&\frac{1}{N\lambda^{k/2}}\mathbb{E}[\mathrm{Tr}(\mathbf{A}_N^k)] \\ &= \sum_{\substack{\pi\in\mathcal{P}(k): \\ G_{\gamma\pi} \text{ is a} \\ \text{rooted labelled tree}}} \sum_{\mathbf{i}\in\mathcal{F}_{\gamma\pi}} \lambda^{|E_{\gamma\pi}|-k/2} \frac{1}{N^{|\gamma\pi|}} \prod_{(a,b)\in E_{\gamma\pi}} f\left(w_{i_a}, w_{i_b}\right) + \mathrm{O}(\lambda^{k/2}N^{-1}).\end{aligned}$$

Thus rewriting the expression with $|E_{\gamma\pi}| = |\gamma\pi| + 1$ we get,

$$\begin{aligned}&\frac{1}{N\lambda^{k/2}}\mathbb{E}[\mathrm{Tr}(\mathbf{A}_N^k)] && (2.29)\\ &= \sum_{\substack{\pi\in\mathcal{P}(k): \\ G_{\gamma\pi} \text{ is a} \\ \text{rooted labelled tree}}} \lambda^{|\gamma\pi|+1-k/2} \sum_{\mathbf{i}\in\mathcal{F}_{\gamma\pi}} \frac{1}{N^{|\gamma\pi|}} \prod_{(a,b)\in E_{\gamma\pi}} f\left(w_{i_a}, w_{i_b}\right) + \mathrm{O}(\lambda^{k/2}N^{-1}). && (2.30)\end{aligned}$$

Remark 2.4.4.
We would like to remark here that if there exists an edge e, such that it is traversed only once in the closed walk, then the graph cannot be a tree. Consider, without loss of generality, that this edge e is $(1, 2)$, with $1 \in V_1$ and $2 \in V_2$, as in figure 2.4, where $V_1, V_2 \in \gamma\pi$. Here C_1 and C_2 are the remaining components of the graph $G_{\gamma\pi}$.

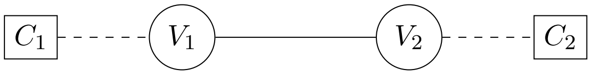


Figure 2.4: Graph associated to $\gamma\pi$ having blocks V_1 and V_2 with the edge between them traversed only once.

Thus, since the closed walk $1 \to 2, 2 \to 3, \ldots k \to 1$ has to return back to V_1, it has to do so via C_1 since the edge e cannot be traversed again. Clearly, this will form a cycle in the graph. Thus, every edge must be traversed at least twice.

It is well-known (see Nica and Speicher [2006]) that for $\pi \in NC_2(k)$ if and only if $|\gamma\pi| = 1 + k/2$, but in the above setting we shall see that other partitions will also contribute as $|\mathcal{F}_{\gamma\pi}| \sim N^{|\gamma\pi|}$. In particular, we need to sum over only those π that give rise to a tree structure. We show in a series of characterizations that the resulting partitions are $SS(k)$.

Characterising partitions

Recall from Definition 2.3.4 that to construct a graph $G_{\gamma\pi}$ associated with a partition π of $[k]$, we need to evaluate $\gamma\pi$ to construct the vertex set and then perform a closed walk. We prove a property that will play a key role in characterising partitions in the proof of Theorem 2.3.7.

Property 1: Block characterisation. For $\pi \in \mathcal{P}(k)$ with $\gamma\pi = \{V_1, \ldots, V_l\}$, if $G_{\gamma\pi}$ has a tree structure, then all elements of a block V_j, $\forall\, 1 \leq j \leq l$, have either all odd elements or all even elements.

Proof of Property 1. For simplicity, we show that the first block has this property. Assume that V_1 has all odd elements except one special element $a \in [k]$. We assume that element '1' belongs to V_1.

Recall from the definition of $G_{\gamma\pi}$ that we first perform a closed walk on $[k]$ as $1 \to 2 \to 3 \to \ldots \to k \to 1$, and then collapse elements of the same block of $\gamma\pi$ into a single vertex. Thus, if $a - 1$ (or $a + 1$) belongs to V_1, then we get a self-loop since $a - 1$ and a collapse to the same vertex and the edge $a - 1 \to a$ (or $a \to a + 1$) forms a loop, which does not give a tree structure. Hence $a - 1$ (respectively $a + 1$) is not in V_1.

Now, suppose $a-1 \in V_j$ for some $j \neq 1$. Then, there exists a path from V_1 to V_j of length $t > 1$, since if $t = 1$, the closed walk $1 \to 2 \to \ldots$ would imply that $a-2 \in V_1$, which contradicts our claim. Now, if $t > 1$, the next edge $\{a-1 \to a\}$ from the closed walk will be from V_j to V_1, leading to a cycle in the graph. Thus, violating property 1 yields a graph that is not a tree.

Property 2: Initial characterisation of π. If $\pi \in \mathcal{P}(k)$ then in any block of π, no two consecutive elements can either be both odd or both even.

Proof of Property 2. Suppose a_1 and a_2 belong in the same block of π with no elements between them, and $a_1 < a_2$, either both even or both odd. Then in $\gamma\pi$, a_1 and a_2+1 belong in the same block, which contradicts Property 1. □

Property 3: Diagonal terms. If π is a contributing partition, then for any $\mathbf{i} = (i_1, \ldots, i_k)$ in $\mathcal{F}_{\gamma\pi}$, each element of $\mathbf{i}$ must be pairwise distinct, that is, $i_1 \neq i_2, i_2 \neq i_3, \ldots, i_{k-1} \neq i_k$.

Proof of Property 3. Suppose not, and assume $i_a = i_{a+1}$ for some $1 \leq a \leq k-1$. Then, in $\gamma\pi$, 'a' and '$a+1$' belong to the same block. This contradicts Property 1. □

We now use the above properties for further characterisation of the partitions.

Lemma 2.4.5.
Every block in π must be of even size.

Proof of Lemma 2.4.5. We prove this by contradiction. Consider an odd-sized block $V = \{l_1, \ldots, l_r\} \in \pi$ with $l_1 < l_2 < \cdots < l_r$. Assume that l_1 is odd. By Property 2, l_2 must be even, and by continuing the argument, we have that at every even position, the element is even, and at odd positions, it is odd. Since r is odd, and l_r is in the r^{th} position, which is an odd position, l_r must be odd. Then, in $\gamma\pi$, the element l_r will map to the element l_1+1 which is even, which contradicts Property 1. A similar argument holds when l_1 is taken to be even. This proves the result. □

Corollary 2.4.6 (Vanishing odd moments).
The odd moments vanish as $N \to \infty$.

Proof of Corollary 2.4.6. Recall that partitions whose graphs do not yield a tree structure contribute to the error term with leading order $\mathrm{O}(N^{-1})$. For k odd, every $\pi \in SS(k)$ must have at least one block of odd size. Therefore, Lemma 2.4.5 is violated, and consequently, the odd moments vanish asymptotically. □

Proposition 2.4.7.
Let $\pi \in \mathcal{P}(k)$ such that $G_{\gamma\pi}$ is a rooted labelled tree. Then π must satisfy the following properties.

- *All blocks of the partition must be of even size.*
- *Between any two successive elements of a block, there are sub-blocks of even sizes.*

Proof of Proposition 2.4.7. The first condition is already proved using Lemma 2.4.5. For the second condition, begin by considering a block B that is of the form

$$B = \{\dots, a_1, a_1+1, \dots, a_1+e, a_2, \dots\}$$

with $a_1 - 1 \notin B$, and there doesn't exist any element a' such that $a_1 + e < a' < a_2$ and $a' \in B$. The sub-block here of interest is $\{a_1, a_1+1, \dots, a_1+e\}$. We claim that this sub-block has an odd number of elements, or equivalently, e is an even number. We can also assume, without loss of generality, that a_1 is an odd number. As a consequence of Property 2, a_2 must be even. If we now evaluate $\gamma\pi$ using the above information, we have that $\gamma\pi$ contains the following three (and possibly more) blocks.

$$\begin{aligned}
V_1 &= \{\dots, a_1, a_1+2, \dots, a_1+e, a_2+1, \dots\},\\
V_2 &= \{\dots, a_1+1, a_1+3, \dots, a_1+e-1, a_1+e+1, \dots\},\\
V_3 &= \{\dots, a_2, \dots\}.
\end{aligned}$$

Thus, the graph associated with $\gamma\pi$ will be as shown in Figure 2.5, where C_1, C_2, and C_3 are the remaining components of the graph.

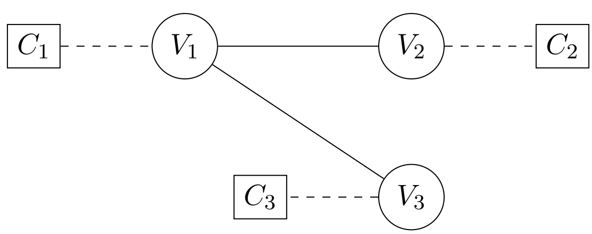


Figure 2.5: Graph associated to $\gamma\pi$ having blocks V_1, V_2 and V_3.

We now focus on the closed walk that occurs on the tuple $[k]$. Since this is a closed walk, it does not matter if instead of beginning at 1, we begin at an arbitrary element $k_1 \in [k]$ and perform $\{k_1 \to k_1+1, \dots, k \to 1, 1 \to 2, \dots, k_1 - 1 \to k_1\}$. So, we pick a_1 as the starting point and consequently, without loss of generality, we assume the walk begins at V_1.

The walk will immediately proceed to move back and forth between V_1 and V_2 due to the path $\{a_1 \to a_1+1, a_1+1 \to a_1+2, \ldots, a_1+e \to a_1+e+1\}$, and will eventually end at V_2.

Now, the walk will jump from V_2 into the component C_2. On the other hand, when the walk eventually enters V_3, it will move at least once to V_1, due to the path $\{a_2 \to a_2+1\}$. So, to preserve the tree structure, the walk must first come back to V_2 and then proceed to V_3 via V_1. Thus, there is an element a' such that $a' \in V_2$ and $a'+1 \in V_1$, where $a' > a_1+e$ and $a' < a_2$. Therefore, in $\gamma\pi$, a_1+e maps to $a'+1$. This implies that a_1+e and a' belong to the same block in π, and thus, $a' \in B$. This contradicts our construction, and therefore, the walk must form a cycle from V_2 or C_2 to either C_1, C_3 or V_3. □

Recall the definition of *Special Symmetric Partitions* as provided in Definition 2.3.1, where the two properties outlined in Proposition 2.4.7 are the main characteristics. As a result, we have demonstrated (2.26), leading us to the conclusion of the proof of Lemma 2.4.2. □

We would now like to take limits in (2.26) and finally get the expression for the moments. The following lemma is an easy consequence of Lemma 2.22 and the fact that $|\mathcal{F}_{\gamma\pi}| \sim N^{|\gamma\pi|}$.

Lemma 2.4.8.
Let $\pi \in SS(k)$ and $\mathcal{F}_{\gamma\pi}$ be as in Lemma 2.4.2. Also, $G_{\gamma\pi} = (V_{\gamma\pi}, E_{\gamma\pi})$ be the graph as in Definition 2.3.4.

$$\begin{aligned}&\lim_{N\to\infty} \sum_{\mathbf{i}\in\mathcal{F}_{\gamma\pi}} \frac{1}{N^{|\gamma\pi|}} \prod_{(a,b)\in E_{\gamma\pi}} f\left(w_{i_a}, w_{i_b}\right)\\ &= \int_{[0,\infty)^{|\gamma\pi|}} \prod_{(a,b)\in E_{\gamma\pi}} f(w_a, w_b)\mu_w^{\bigotimes|\gamma\pi|}(\mathrm{d}\,\mathbf{w}).\end{aligned} \tag{2.31}$$

Now, going back to equation (2.29) and taking limits gives us

$$\lim_{N\to\infty} \mathbb{E}[\operatorname{tr}(\mathbf{A}_N^k)] = \begin{cases} 0, & k \text{ odd} \\ \sum\limits_{\pi\in SS(k)} \lambda^{|\gamma\pi|-1-k/2} t(G_{\gamma\pi}, f, \mu_w), & k \text{ even} \end{cases}. \tag{2.32}$$

Now, the sum over $SS(k)$ can be further split up as the sum over $NC_2(k)$ and the remaining partitions. Moreover, for $\pi \in SS(k)$, we have $|V_{\gamma\pi}| = |\gamma\pi| \in \{2, 3, \ldots, k/2+1\}$. In particular, for $\pi \in NC_2(k)$, $|\gamma\pi| = k/2+1$, and when π

is the full partition $\{\{1, 2, \ldots, k\}\}$, $|\gamma\pi| = 2$. So, we can write

$$\lim_{N\to\infty} \mathbb{E}[\mathrm{tr}(\mathbf{A}_N^k)] = \begin{cases} 0, & k \text{ odd} \\ \sum\limits_{\pi\in NC_2(k)} t(G_{\gamma\pi}, f, \mu_w) + \sum\limits_{l=2}^{k/2} \sum\limits_{\substack{\pi\in SS(k)\setminus NC_2(k): \\ |\gamma\pi|=l}} \lambda^{l-1-k/2} t(G_{\gamma\pi}, f, \mu_w), & k \text{ even} \end{cases} . \tag{2.33}$$

§2.4.2 Concentration and uniqueness

We now show a concentration result to obtain convergence in probability.

Lemma 2.4.9 (Concentration of trace).
For all $k \geq 0$, we have that

$$\mathrm{Var}\left[\mathrm{tr}(\mathbf{A}_N^k)\right] = \mathrm{O}_N((\lambda N)^{-1}) .$$

Proof. We shall proceed to compute the variance

$$\mathrm{Var}\left[\mathrm{tr}(\mathbf{A}_N^k)\right] .$$

Let $\mathbf{i}$ and $\mathbf{i}'$ denote the tuples

$$\mathbf{i} = \{i_1, \ldots, i_k\}, \ \ \mathbf{i}' = \{i_{k+1}, \ldots, i_{2k}\}$$

and denote by $P(\mathbf{i})$ the expectation

$$P(\mathbf{i}) = \mathbb{E}[a_{i_1 i_2} a_{i_2 i_3} \ldots a_{i_k i_1}] .$$

Similarly, we have

$$P(\mathbf{i}') = \mathbb{E}[a_{i_{k+1} i_{k+2}} a_{i_{k+2} i_{k+3}} \ldots a_{i_{2k} i_{k+1}}] .$$

For the tuple $\mathbf{i}$, we can define a closed walk as in the proof of Lemma 2.4.2 to get a graph $G(\mathbf{i}) := (V(\mathbf{i}), E(\mathbf{i}))$. In the same spirit, one can define $G(\mathbf{i}, \mathbf{i}') = (V(\mathbf{i}, \mathbf{i}'), E(\mathbf{i}, \mathbf{i}'))$, with the closed walk now performed as

$$1 \to 2 \to \ldots k \to 1, k+1 \to k+2 \to \ldots 2k \to k+1 ,$$

where the jump from 1 to $k+1$ is without an edge. Then, we can define

$$P(\mathbf{i}, \mathbf{i}') = \mathbb{E}[a_{i_1 i_2} a_{i_2 i_3} \ldots a_{i_k i_1} a_{i_{k+1} i_{k+2}} \ldots a_{i_{2k} i_{k+1}}] .$$

With this notation set up, one can see that

$$\begin{aligned}\mathrm{Var}\left[\mathrm{tr}(\mathbf{A}_N^k)\right] &= \frac{1}{N^2}\left[\mathbb{E}[(\mathrm{Tr}(\mathbf{A}_N^k)^2] - (\mathbb{E}[\mathrm{Tr}(\mathbf{A}_N^k)])^2\right] \\ &= \frac{1}{N^2\lambda^k}\sum_{i_1,i_2,\ldots,i_k,i_{k+1},\ldots,i_{2k}=1}^{N} P(\mathbf{i},\mathbf{i}') - P(\mathbf{i})P(\mathbf{i}')\,.\end{aligned} \tag{2.34}$$

We remark here that the construction of the graph $G(\mathbf{i},\mathbf{i}')$ is similar to how we did in Lemma 2.4.2, with the essential difference being the closed walk structure over two separate $k-$tuples.

Suppose that $E(\mathbf{i}) \cap E(\mathbf{i}') = \phi$. Then by independence, (2.34) becomes 0. Thus, we must have $E(\mathbf{i}) \cap E(\mathbf{i}') \neq \phi$. Moreover, due to remark 2.4.4, each term must appear at least twice in $P(\mathbf{i},\mathbf{i}')$, that is, each edge in $E(\mathbf{i},\mathbf{i}')$ is traversed at least twice. This implies that the maximum number of edges our graph can have is k.

Next, note that the only way the graph $G(\mathbf{i},\mathbf{i}')$ will be disconnected is when the closed walk over the two $k-$ tuples yields two disjoint graphs, and thus we once again obtain $P(\mathbf{i},\mathbf{i}') = P(\mathbf{i})P(\mathbf{i}')$.

Thus, our computation boils down to the case where $G(\mathbf{i},\mathbf{i}')$ is a connected graph, with each edge appearing at least twice, and $E(\mathbf{i}) \cap E(\mathbf{i}') \neq \phi$. Note that one can have $G(\mathbf{i},\mathbf{i}')$ to be connected and still have $E(\mathbf{i}) \cap E(\mathbf{i}') = \phi$, for example when i_1 and i_{k+1} are collapsed into the same vertex. This gives us that $|V(\mathbf{i},\mathbf{i}')| \leq |E(\mathbf{i},\mathbf{i}')| + 1 \leq k+1$. Using $|f| \leq C_f$ gives us that

$$\mathrm{Var}\left[\mathrm{tr}(\mathbf{A}_N^k)\right] \leq C_f \frac{1}{N^2\lambda^k} N^{|V|}\left(\frac{\lambda}{N}\right)^{|E|} = C_f \lambda^{|E|-k} N^{|V|-|E|-2} = \mathrm{O}_N(N^{-1})\,.$$

This completes the proof. □

An immediate consequence from Chebychev's inequality is that the moments concentrate around their mean as $N \to \infty$. In other words, for all $k \geq 1$,

$$\lim_{N\to\infty} \mathrm{tr}(\mathbf{A}_N^k) = m_k(\mu_\lambda) \;\text{ in probability},$$

where $m_k(\mu_\lambda)$ are as in (2.10). To conclude Theorem 2.3.7, we now further analyse the sequence $\{m_k\}_{k\geq 0}$, and show that it is unique for the measure μ_λ. A measure μ is said to be uniquely determined by its moment sequence $\{m_k\}_{k\geq 0}$ if the following holds (Carleman's condition):

$$\sum_{k\geq 0} m_{2k}^{-1/2k} = \infty. \tag{2.35}$$

Lemma 2.4.10 (Uniqueness of moments).
For λ bounded away from 0, that is, $\lambda > 0$, the moments uniquely determine the limiting spectral measure.

Proof. Let m_k denote the k^{th} moment. Since f is bounded, we have

$$\begin{aligned} m_{2k} &= \sum_{\pi \in SS(2k)} \lambda^{|\gamma\pi|-1-k} \int_{[0,1]^{|\gamma\pi|}} \prod_{(ab)\in E_{\gamma\pi}} f(x_a, x_b) \prod_{i=1}^{|\gamma\pi|} \mu_w(\mathrm{d}\, x_i) \\ &\leq \sum_{\pi \in SS(2k)} C_f^{|\gamma\pi|} \lambda^{|\gamma\pi|-1-k} \\ &= \sum_{l=2}^{k+1} \sum_{\pi \in SS(2k):|\gamma\pi|=l} C_f^{l} \lambda^{l-1-k}, \end{aligned}$$

Let A_k be defined as

$$A_k = \begin{cases} 1, & \text{if } \lambda \geq 1, \\ \lambda^{-k}, & \text{if } 1 > \lambda > 0. \end{cases}$$

Then,

$$\begin{aligned} m_{2k} &\leq C_f^{k+1} A_k \sum_{l=2}^{k+1} |\{\pi \in SS(2k) : |\gamma\pi| = l\}| \\ &\leq A_k C_f^{k+1} |\{SS(2k)\}| \\ &\leq A_k C_f^{k+1} (2k)^{2k}, \end{aligned}$$

where the last inequality follows since $SS(2k) \subset P(2k)$ and $|P(2k)|$ is bounded by $2k^{2k}$. Thus,

$$m_{2k}^{-1/2k} \geq \frac{1}{2k\sqrt{C_f}} . \frac{1}{(A_k C_f)^{\frac{1}{2k}}}$$

So, we have the series $\sum_{k\geq 1} m_{2k}^{-1/2k}$ to be lower bounded by $\sum_{k\geq 1} a_k$, where

$$a_k = \frac{1}{2k\sqrt{C_f}} . \frac{1}{(A_k C_f)^{\frac{1}{2k}}} = \frac{1}{C_1 k \mathrm{e}^{\frac{1}{2k}\log(A_k C_f)}} .$$

Thus,

$$a_k = \begin{cases} \frac{\mathrm{e}^{-C_2/2k}}{C_1 k}, & \text{for } \lambda \geq 1\,, \\ \frac{\sqrt{\lambda}\mathrm{e}^{-C_2/2k}}{C_1 k}, & \text{for } \lambda < 1\,. \end{cases}$$

Since $\mathrm{e}^{-x} > 1 - x$, we see that the series $\sum_{k\geq 1} a_k$ diverges, and consequently,

$$\sum_{k\geq 0} m_{2k}^{-1/2k} = \infty.$$

§2.5 Stieltjes Transform and analytic description

§2.5.1 Resolvent and Stieltjes Transform

We fix a $z \in \mathbb{C}^+$ throughout this argument, with $\Im(z) = \eta > 0$. Recall that the resolvent is given by

$$\mathrm{R}_{\mathbf{A}_N}(z) := (\mathbf{A}_N - zI)^{-1},\ z \in \mathbb{C}^+.$$

The Stieltjes transform of the empirical spectral distribution of $\mathbf{A}_N$ is given by

$$\mathrm{S}_{\mathbf{A}_N}(z) = \int_{\mathbb{R}} \frac{1}{x-z} \operatorname{ESD}(\mathbf{A}_N)(\mathrm{d}\,x) = \operatorname{tr}(\mathrm{R}_{\mathbf{A}_N}(z)), \tag{2.36}$$

where tr denotes the normalised trace.

Lemma 2.5.1 (Resolvent Properties).
For any $z \in \mathbb{C}^+, 1 \le i, j \le N$, the following properties are well-known for the resolvent $\mathrm{R}_{\mathbf{A}}$ of an $N \times N$ matrix $\mathbf{A}$.

(i) ***Analytic:*** *$z \mapsto \mathrm{R}_{\mathbf{A}}(z)(i,j)$ is an analytic function on $\mathbb{C}^+ \to \mathbb{C}^+$.*

(ii) ***Bounded :*** *$\|\mathrm{R}_{\mathbf{A}}(z)\|_{op} \le \Im(z)^{-1}$, where $\|\cdot\|_{op}$ denotes the operator norm.*

(iii) ***Normal :*** *$\mathrm{R}_{\mathbf{A}}(z)\,\mathrm{R}_{\mathbf{A}}(z)^* = \mathrm{R}_{A}(z)^*\,\mathrm{R}_{A}(z)$.*

(iv) ***Diagonals are bounded:*** *$|\mathrm{R}_{\mathbf{A}}(z)(i,j)| \le \Im(z)^{-1}$.*

(v) ***Trace bounded:*** *$|\operatorname{tr}(\mathrm{R}_{\mathbf{A}}(z))| \le \Im(z)^{-1}$. In particular,*

$$\left|\operatorname{tr}(\mathrm{R}^p_{\mathbf{A}}(z))\right| \le \Im(z)^{-p},\ \textit{for any } p \ge 1.$$

For the first three properties see [Bordenave, 2019, Chapter 3]. Note that the property (iv) follows from (iii) by the following argument:

$$|\mathrm{R}_{\mathbf{A}}(z)(i,j)| \le |\langle \delta_i,\, \mathrm{R}_{\mathbf{A}}(z)\delta_j\rangle| \le \sup_{v:\|v\|=1} |\langle \delta_i, \mathrm{R}_{\mathbf{A}}(z)\delta_j\rangle| = \|\mathrm{R}_{\mathbf{A}}(z)\|_{\mathrm{op}}.$$

The last property (v) follows from (iv). We now state the Ward's identity, for which we refer the reader to [Erdős and Yau, 2017, Lemma 8.3].

Lemma 2.5.2 (Ward's identity).
Let $\mathbf{A}$ be a Hermitian matrix and $\mathrm{R}_{\mathbf{A}}$ be the resolvent. Let $z \in \mathbb{C}^+$. Then for any fixed k, we have

$$\sum_{l \ne k} |\mathrm{R}_{\mathbf{A}}(l,k)|^2 = \frac{1}{\eta}\Im(\mathrm{R}_{\mathbf{A}}(k,k)).$$

Since we have already shown in the previous section $\lim_{n\to\infty} \mathrm{ESD}(\mathbf{A}_N) = \mu_\lambda$ weakly in probability and hence it follows that for any $z \in \mathbb{C}^+$

$$\lim_{N\to\infty} \mathrm{S}_{\mathbf{A}_N}(z) \to S_{\mu_\lambda}(z).$$

Due to the involved structure of the moments, it is not immediately evident what the limiting Stieltjes transform looks like.

Recall the notation of *expected* empirical spectral distribution of $\mathbf{A}_N$ from (2.23). Let $\bar{\mathrm{S}}_{\mathbf{A}_N}(z)$ denote the Stieltjes transform of $\bar{\mu}_{N,\lambda}$. Notice that $\bar{\mathrm{S}}_{\mathbf{A}_N}(z) = \mathbb{E}[\mathrm{S}_{\mathbf{A}_N}(z)]$. It is known that if a measure μ_N converges weakly in probability to a measure μ, then the corresponding Stieltjes transforms converge. In particular, we have the following lemma.

Lemma 2.5.3.
Anderson et al. [2010, Theorem 2.4.4] A sequence of measures μ_N converge weakly in probability to a measure μ if and only if $\mathrm{S}_{\mu_N}(z)$ converges in probability to $\mathrm{S}_\mu(z)$ for each $z \in \mathbb{C}^+$.

Thus, we compute an expression for the expected Stieltjes transform $\mathrm{S}_{\bar{\mathbf{A}}_N}$, and using convergence in probability from Theorem 2.3.7, we can claim that the Stieltjes transform $\mathrm{S}_{\mathbf{A}_N}(z)$ converges in probability to the same expression. For ease of notation we shall denote by $r^N_{kk}(z) := \mathrm{R}_{\mathbf{A}_N}(z)(k,k)$ for $1 \le k \le N$.

The following identity can be found in Abramowitz and Stegun [1964]. For any complex number $z \in \mathbb{C}^+$, we have for all $u \ge 0$,

$$\mathrm{e}^{iuz} = 1 - \sqrt{u}\int_0^\infty \frac{\mathrm{J}_1(2\sqrt{uv})}{\sqrt{v}} \mathrm{e}^{-ivz^{-1}}\,\mathrm{d}\,v, \tag{2.37}$$

where $\mathrm{J}_1(x)$ is the first-order Bessel function of the first kind given by (2.16). Note that for all $x \ge 0$, $|\,\mathrm{J}_1(x)| \le 1$ (see [Abramowitz and Stegun, 1964, Chapter 9]). We know that the resolvent maps the upper half complex plane to the upper half complex plane. Thus, we begin by fixing $r^N_{jj}(z)$, the j^{th} diagonal entry of the $N \times N$ resolvent matrix, as our complex variable in $\mathbb{C}^+$. So we can get

$$\mathrm{e}^{iur^N_{jj}(z)} = 1 - \sqrt{u}\int_0^\infty \frac{\mathrm{J}_1(2\sqrt{uv})}{\sqrt{v}} \mathrm{e}^{-iv(r^N_{jj})^{-1}}\,\mathrm{d}\,v\,. \tag{2.38}$$

If we look at $\sum_{j=1}^N \mathrm{e}^{iur^N_{jj}(z)}$ then the relation between the Stieltjes transform and the above equation becomes apparent. It turns out that

$$\mathrm{S}_{\mathbf{A}_N}(z) = \frac{\partial}{\partial u}\frac{1}{N}\sum_{j=1}^N \mathrm{e}^{iur^N_{jj}(z)}\Bigg|_{u=0}. \tag{2.39}$$

To understand the Stieltjes transform we will first try to understand the behaviour of (2.38). We will adapt the approach of Khorunzhy et al. [2004]. For ease of notation, for what follows, $\|\cdot\|$ will denote the norm $\|\cdot\|_{\mathcal{B}}$ as defined in (2.11), unless stated otherwise.

Proposition 2.5.4.
Let $r_{jj}^N := r_{jj}^N(z)$ denote the j^{th} diagonal entry of the resolvent $\mathrm{R}_{\mathbf{A}_N}(z)$. Let

$$d_j = \frac{1}{N}\sum_{k=1}^{N} f(w_j, w_k) \tag{2.40}$$

and for any $b > 0$ define the function $g_N : (0,\infty)\times(0,\infty)\times\mathbb{C}^+ \to \mathbb{C}$ as follows

$$g_N(x,b,z) := \frac{1}{N}\sum_{k=1}^{N} f(x,w_k)\mathrm{e}^{ibr_{kk}^N(z)}\,. \tag{2.41}$$

Then, for any $z \in \mathbb{C}^+$,

$$\mathbb{E}[\mathrm{e}^{r_{jj}^N}] = 1 - \mathrm{e}^{-\lambda d_j}\sqrt{u}\int_0^\infty \frac{\mathrm{J}_1(2\sqrt{uv})}{\sqrt{v}}\mathrm{e}^{ivz}\mathbb{E}\left[\mathrm{e}^{\lambda g_N\left(w_j,\frac{v}{\lambda},z\right)}\right]\mathrm{d}\,v + q_{N,\lambda}(u,z)\,, \tag{2.42}$$

where $q_{N,\lambda}(u,z) = \mathrm{O}\left(\frac{\lambda\sqrt{u}}{\eta^{5/2}\sqrt{N}}\right)$.

We begin by stating two results we use in this proof. Note that we conveniently drop the dependence on z for $r_{jj}^N(z)$, since we fix $z \in \mathbb{C}^+$ throughout and hence just use the notation r_{jj}^N.

Fact 2.5.5 (Exponential Inequalities).
The following holds for any real numbers $a, b \in \mathbb{R}$ and complex numbers $z_1, z_2 \in \mathbb{C}^+$.

$$|e^{aiz_1} - \mathrm{e}^{aiz_2}| \le |a||z_1 - z_2| \tag{2.43}$$

$$|e^a - \mathrm{e}^b| \le |a-b|e^{|a|+|b|} \tag{2.44}$$

Proof of Proposition 2.5.4. For the resolvent of a matrix with zero diagonal, we have the relation

$$r_{jj}^N = -\left(z + \sum_{k,l\neq j}\tilde{r}_{kl}^{N-1}a_{kj}a_{lj}\right)^{-1},$$

for any diagonal element r^N_{jj} of the resolvent $\mathrm{R}_{\mathbf{A}_N}(z)$, where $\tilde{r}^{N-1}_{kl} := \tilde{r}^{N-1}_{kl}(z)$ are the entries of the resolvent of $\mathbf{A}^{(j)}_{N-1}$ in $z \in \mathbb{C}^+$, which is the adjacency matrix with deleted j^{th} row and column. Plugging into (2.37) yields

$$\mathrm{e}^{iur^N_{jj}} = 1 - \sqrt{u}\int_0^\infty \frac{\mathrm{J}_1(2\sqrt{uv})}{\sqrt{v}}\mathrm{e}^{ivz}\exp\left(iv\sum_{k,l\neq j}\tilde{r}^{N-1}_{kl}a_{kj}a_{lj}\right)\mathrm{d}\,v\,. \tag{2.45}$$

Adding and subtracting the appropriate exponential to (2.45) yields

$$\mathrm{e}^{iur^N_{jj}} = 1 - \sqrt{u}\int_0^\infty \frac{\mathrm{J}_1(2\sqrt{uv})}{\sqrt{v}}\mathrm{e}^{ivz}\exp\left(iv\sum_{k\neq j}\tilde{r}^{N-1}_{kk}a^2_{kj}\right)\mathrm{d}\,v + E_1\,, \tag{2.46}$$

where E_1 is an error term given by

$$E_1 =$$
$$\sqrt{u}\int_0^\infty \frac{\mathrm{J}_1(2\sqrt{uv})}{\sqrt{v}}\mathrm{e}^{ivz}\left(\exp\left(iv\sum_{k,l\neq j}\tilde{r}^{N-1}_{kl}a_{kj}a_{lj}\right) - \exp\left(iv\sum_{k\neq j}\tilde{r}^{N-1}_{kk}a^2_{kj}\right)\right)\mathrm{d}\,v\,.$$

It is easy to see that for $z \in \mathbb{C}^+$ with $\Re(z) = \zeta \in \mathbb{R}$ and $\Im(z) = \eta > 0$, we have $|\mathrm{e}^{ivz}| = |\mathrm{e}^{i\zeta v}\mathrm{e}^{-\eta v}| \leq \mathrm{e}^{-\eta v}$. Thus,

$$\begin{aligned}
&|E_1| = \\
&\left|\sqrt{u}\int_0^\infty \frac{\mathrm{J}_1(2\sqrt{uv})}{\sqrt{v}}\mathrm{e}^{ivz}\left(\exp\left(iv\sum_{k,l\neq j}\tilde{r}^{N-1}_{kl}a_{kj}a_{lj}\right) - \exp\left(iv\sum_{k\neq j}\tilde{r}^{N-1}_{kk}a^2_{kj}\right)\right)\mathrm{d}\,v\right| \\
&\leq \sqrt{u}\int_0^\infty \frac{v\mathrm{e}^{-\eta v}}{\sqrt{v}}\sum_{k\leq N}\sum_{l\neq k}|\tilde{r}^{N-1}_{kl}|a_{kj}a_{lj}\,\mathrm{d}\,v \\
&= \left(\sqrt{u}\int_0^\infty \sqrt{v}\mathrm{e}^{-\eta v}\,\mathrm{d}\,v\right)\sum_{k\leq N}\sum_{l\neq k}|\tilde{r}^{N-1}_{kl}|a_{kj}a_{lj}
\end{aligned} \tag{2.47}$$

where in the last step, we use inequality (2.43) and the bound $|\,\mathrm{J}_1(x)| \leq 1$ for $x \geq 0$. Note that in the last sum in (2.47), the entries a_{kj} and a_{lj} are independent of one another, and of $\tilde{r}^{N-1}_{kl}$. Thus, since f is bounded by a constant C_f, taking expectation on the summation gives us

$$\mathbb{E}\left[\sum_{l\neq k}|\tilde{r}^{N-1}_{kl}|a_{kj}a_{lj}\right] \leq \frac{\lambda C_f^2}{N^2}\sum_{l\neq k}|\tilde{r}^{N-1}_{kl}| \tag{2.48}$$

since a_{ij} are distributed as Bernoulli random variables with parameter p_{ij}, and are scaled by a factor $\lambda^{-1/2}$. Using (2.48) and taking expectation in (2.47) gives us

$$\begin{aligned}
\mathbb{E}\left[|E_1|\right] &\leq C_f^2\sqrt{u}\int_0^\infty \frac{\sqrt{v}\mathrm{e}^{-\eta v}\lambda}{N^2}\sum_{k\leq N}\sum_{l\neq k}\mathbb{E}[|\tilde{r}_{kl}^{N-1}|]\,\mathrm{d}\,v \\
&\leq C_f^2\sqrt{u}\int_0^\infty \frac{\sqrt{v}\mathrm{e}^{-\eta v}\lambda}{N\sqrt{N}}\mathbb{E}\left[\sum_{k\leq N}\left(\sum_{l\neq k}|\tilde{r}_{kl}^{N-1}|^2\right)^{\frac{1}{2}}\right]\mathrm{d}\,v \quad \text{(Cauchy-Schwarz)} \\
&\leq C_f^2\sqrt{u}\int_0^\infty \frac{\sqrt{v}\mathrm{e}^{-\eta v}\lambda}{N\sqrt{N\eta}}\mathbb{E}\left[\sum_{k\leq N}(\Im(\tilde{r}_{kk}^{N-1}))^{\frac{1}{2}}\right]\mathrm{d}\,v \;\; \text{(using Lemma 2.5.2)} \\
&\leq C_f^2\sqrt{u}\int_0^\infty \frac{\sqrt{v}\mathrm{e}^{-\eta v}\lambda}{\sqrt{N\eta}}\,\mathrm{d}\,v \quad \text{(using property (iv) from Lemma 2.5.1)} \\
&= C_f^2\frac{\sqrt{u}\lambda}{\eta^{5/2}\sqrt{N}}\int_0^\infty \sqrt{\eta v}\mathrm{e}^{-\eta v}\,\mathrm{d}\,(\eta v) = \mathrm{O}\left(\frac{\lambda\sqrt{u}}{\eta^{5/2}\sqrt{N}}\right),
\end{aligned}$$

where in the last step we do a change of variable $\eta v = v'$ to show the integral is finite. So, if we now take an expectation in (2.46), we get

$$\mathbb{E}[\mathrm{e}^{iur_{jj}^N}] = 1-\sqrt{u}\int_0^\infty \frac{\mathrm{J}_1(2\sqrt{uv})}{\sqrt{v}}\mathrm{e}^{ivz}\mathbb{E}\left[\exp\left(iv\sum_{k\neq j}\tilde{r}_{kk}^{N-1}a_{kj}^2\right)\right]\mathrm{d}\,v+q_{N,\lambda}(u,z), \tag{2.49}$$

where $q_{N,\lambda}(u,z) = \mathrm{O}\left(\frac{\lambda\sqrt{u}}{\eta^{5/2}\sqrt{N}}\right)$. Note that the expectation could be pulled inside the integral in (2.46) using Fubini's Theorem since the integral is bounded above by a constant. To evaluate the expectation inside (2.49), we use a conditioning argument as follows. We have

$$\mathbb{E}\left[\exp\left(iv\sum_{k\neq j}\tilde{r}_{kk}^{N-1}a_{kj}^2\right)\right] = \mathbb{E}\left[\mathbb{E}\left[\exp\left(iv\sum_{k\neq j}\tilde{r}_{kk}^{N-1}a_{kj}^2\right)\middle|\mathbf{A}_{N-1}^{(j)}\right]\right].$$

Evaluating the conditional expectation yields

$$\begin{aligned}
&\mathbb{E}\left[\exp\left(iv\sum_{k\neq j}\tilde{r}_{kk}^{N-1}a_{kj}^2\right)\right]\\
&=\mathbb{E}\left[\prod_{k=1}^{N}\left(1-\frac{\lambda}{N}f(w_k,w_j)+\frac{\lambda}{N}f(w_k,w_j)\mathrm{e}^{iv\tilde{r}_{kk}^{N-1}/\lambda}\right)\right]\\
&=\mathbb{E}\left[\prod_{k=1}^{N}\left(1+\frac{\lambda}{N}f(w_k,w_j)\left(\mathrm{e}^{iv\tilde{r}_{kk}^{N-1}/\lambda}-1\right)\right)\right]\\
&=\mathbb{E}\left[\prod_{k=1}^{N}\left(\exp\left(\frac{\lambda}{N}f(w_k,w_j)\left(\mathrm{e}^{iv\tilde{r}_{kk}^{N-1}/\lambda}-1\right)\right)+q_k'(N,\lambda)\right)\right], \qquad (2.50)
\end{aligned}$$

where $q_k'(N,\lambda)$ is an error given by

$$\begin{aligned}
&q_k'(N,\lambda)\\
&=1+\frac{\lambda}{N}f(w_k,w_j)\left(\mathrm{e}^{iv\tilde{r}_{kk}^{N-1}/\lambda}-1\right)-\exp\left(\frac{\lambda}{N}f(w_k,w_j)\left(\mathrm{e}^{iv\tilde{r}_{kk}^{N-1}/\lambda}-1\right)\right).
\end{aligned}$$

Since $|\mathrm{e}^{iv\tilde{r}_{kk}^{N-1}/\lambda}-1|\leq 2$, doing a Taylor expansion for the exponential term in $q_k'(N,\lambda)$ gives us

$$|q_k'(N,\lambda)|\leq\frac{4C_f^2\lambda^2}{N^2}=\mathrm{O}\left(\frac{\lambda^2}{N^2}\right). \qquad (2.51)$$

We can write

$$\begin{aligned}
&\mathbb{E}\left[\exp\left(iv\sum_{k\neq j}\tilde{r}_{kk}^{N-1}a_{kj}^2\right)\right]\\
&=\mathbb{E}\left[\prod_{k=1}^{N}\left(\exp\left(\frac{\lambda}{N}f(w_k,w_j)\left(\mathrm{e}^{iv\tilde{r}_{kk}^{N-1}/\lambda}-1\right)\right)\right)\right]+\mathbb{E}[E_2], \qquad (2.52)
\end{aligned}$$

where E_2 is an expression involving all the other terms of the product in (2.50). To get the order of E_2, we take a supremum over k in (2.50) and compute the binomial expansion of the form $(a+b)^N$ modulo the leading term a^N. In particular, since $|\mathrm{e}^{iv\tilde{r}_{kk}^{N-1}/\lambda}-1|\leq 2$, and again using (2.51), we have

$$|E_2|\leq\sum_{j=1}^{N}\binom{N}{j}\left(\mathrm{e}^{\frac{2\lambda C_f}{N}}\right)^{N-j}\left(\frac{4C_f^2\lambda^2}{N^2}\right)^{j},$$

which for some constant $C_a > 0$ and N large enough further simplifies to

$$|E_2| \leq C_a \sum_{j=1}^{N} (2C_f\lambda)^{2j} N^j \mathrm{e}^{-\frac{2j\lambda C_f}{N}} N^{-2j}$$

$$= C_a \sum_{j=1}^{N} (2C_f\lambda)^{2j} N^{-j} \mathrm{e}^{-\frac{2j\lambda C_f}{N}} = C_a \frac{4C_f\lambda^2 N^{-1} \mathrm{e}^{-\frac{2\lambda C_f}{N}}}{1 - 4C_f\lambda^2 N^{-1} \mathrm{e}^{-\frac{2\lambda C_f}{N}}} ,$$

where the last equality is due to the sum being a geometric series. Thus,

$$|E_2| = \mathrm{O}\left(\frac{\lambda^2}{N}\right) , \tag{2.53}$$

which is a faster error than $q_{N,\lambda}(u, z)$ so we can later absorb it into the existing error of (2.49). Thus, using (2.53), we can rewrite (2.52) as

$$\mathbb{E}\left[\exp\left(iv \sum_{k\neq j} \tilde{r}_{kk}^{N-1} a_{kj}^2\right)\right] = \mathbb{E}\left[\mathrm{e}^{-\lambda d_j} \exp\left(\lambda \tilde{g}_{N-1}\left(w_j, \frac{v}{\lambda}, z\right)\right)\right] + \mathrm{O}\left(\frac{\lambda^2}{N}\right) \tag{2.54}$$

where

$$d_j = \frac{1}{N}\sum_{k=1}^{N} f(w_j, w_k) \quad \text{and} \quad \tilde{g}_{N-1}(w_j, b, z) = \sum_{k=1}^{N} f(w_j, w_k) \mathrm{e}^{ib\tilde{r}_{kk}^{N-1}} . \tag{2.55}$$

Note that $\tilde{g}_N$ is a bounded function and is bounded above by C_f. To get the error down from the exponent, we again use inequality (2.44).

To conclude the proof of the proposition, we need to return to an expression involving terms of the form r_{kk}^N of the original resolvent. To do so, we do an interpolation argument. Let $0 \leq t \leq 1$ and define $\mathbf{A}_N^t = (1-t)\mathbf{A}_N + t\mathbf{A}_{N-1}^{(j)}$ with the resolvent $\mathrm{R}_{\mathbf{A}_N^t}(z)$, whose entries we denote by $r_{kl}^N(t) := r_{kl}^N(z, t)$, that also implicitly depends on z but we drop that for convenience of notation. Also, define

$$\mathbf{g}_N^t(w_j, b, z) = \frac{1}{N}\sum_{i=1}^{N} f(w_i, w_j) \mathrm{e}^{ibr_{kk}^N(t)} .$$

We remark using property (i) from Lemma 2.5.1 that $\mathbf{g}_N^t$ is also bounded above by C_f for all values of t, since the complex exponential $\mathrm{e}^{ibr_{kk}^N(t)}$ is bounded by 1 for any $b \geq 0$ and $1 \leq k \leq N$. In particular, we have that $|g_N(x, b, z)| \leq C_f$ for all $x, b \geq 0$.

Our target function is $g_N(w_j,b,z) = \frac{1}{N}\sum_{i=1}^{N} f(w_i,w_j)\mathrm{e}^{ibr_{kk}^N}$. By the fundamental theorem of calculus,

$$\begin{aligned}|g_N(w_j,b,z)-\tilde{g}_{N-1}(w_j,b,z)| &= \left|\mathbf{g}_N^0(w_j,b,z)-\mathbf{g}_N^1(w_j,b,z)\right| \\ &= \left|\int_0^1 \frac{\partial}{\partial t}\mathbf{g}_N^t \,\mathrm{d}\, t\right| = \left|\int_0^1 \frac{b}{N}\sum_{k=1}^{N} \mathrm{e}^{ibr_{kk}^N(t)}\frac{\partial}{\partial t} r_{kk}^N(t)\right|.\end{aligned}$$

Now, $\mathrm{R}_{\mathbf{A}_N^t}(z) = (\mathbf{A}_N^t - zI)^{-1}$ and thus, $\frac{\mathrm{d}}{\mathrm{d}\, t}\mathrm{R}_{\mathbf{A}_N^t}(z) = -\mathrm{R}_{\mathbf{A}_N^t}(z)\frac{\mathrm{d}\,\mathbf{A}_N^t}{\mathrm{d}\, t}\mathrm{R}_{\mathbf{A}_N^t}(z)$.

Note that $\frac{\mathrm{d}\,\mathbf{A}_N^t}{\mathrm{d}\, t} = -J_N$, where J_N is given by

$$J_N(k,l) = \begin{cases} 0, & \text{if } k,l \neq j \\ a_{kl}, & \text{if } k=j \text{ or } l=j\,. \end{cases}$$

Thus,

$$\begin{aligned}&|g_N(w_j,b,z)-\tilde{g}_{N-1}(w_j,b,z)| \\ &= \left|\int_0^1 \frac{b}{N}\sum_{k=1}^{N}\mathrm{e}^{ibr_{kk}^N(t)}\sum_{m,n=1}^{N} r_{km}^N(t)\frac{\partial a_{mn}^t}{\partial t} r_{nk}^N(t)\right| \\ &= \left|\int_0^1 \frac{b}{N}\sum_{k=1}^{N}\mathrm{e}^{ibr_{kk}^N(t)}\sum_{m=1}^{N} r_{km}^N(t)a_{mj}r_{jk}^N(t)\,\mathrm{d}\, t\right| \\ &\leq \int_0^1 \frac{b}{N}\sum_{k=1}^{N}\sum_{m=1}^{N}|r_{km}^N(t)a_{mj}r_{jk}^N(t)|\,\mathrm{d}\, t \qquad (2.56)\end{aligned}$$

since the complex exponential $\mathrm{e}^{ibr_{kk}^N(t)}$ is trivially bounded by 1 as $r_{kk}^N(t) \in \mathbb{C}^+$. Then, using Cauchy-Schwarz and Lemma 2.5.2 in (2.56), we have

$$\begin{aligned}&|g_N(w_j,b,z)-\tilde{g}_{N-1}(w_j,b,z)| \\ &\leq \int_0^1 \frac{b}{N}\sum_{k=1}^{N}|r_{jk}^N(t)|\left(\frac{\Im(r_{kk}^N(t))}{\eta}\right)^{1/2}\left(\sum_{m=1}^{N}a_{mj}^2\right)^{1/2}\mathrm{d}\, t\,.\end{aligned}$$

Bounding $\Im(r_{kk}^N(t))$ by $1/\eta$ (Property (iv) of Lemma 2.5.1) and taking expectation, we get

$$\mathbb{E}[|g_N(w_j,b,z)-\tilde{g}_{N-1}(w_j,b,z)|] \leq \int_0^1 \frac{b}{N\eta}\mathbb{E}\left[\sum_{k=1}^{N}|r_{jk}^N(t)|\left(\sum_{m=1}^{N}a_{mj}^2\right)^{1/2}\right]\mathrm{d}\, t\,. \qquad (2.57)$$

Now, again using Cauchy-Schwarz and Lemma 2.5.2, we have for some constant C' that

$$\sum_{k=1}^{N} |r_{jk}^N(t)| \le \sqrt{N}\left(\sum_{k=1}^{N} |r_{jk}^N(t)|^2\right)^{1/2} \le C'\frac{\sqrt{N}}{\sqrt{\eta}}\,. \tag{2.58}$$

Thus, using (2.58) and Jensen's inequality on the function $\sqrt{X}$ in (2.57), we get

$$\begin{aligned}\mathbb{E}[|g_N(w_j,b,z) - \tilde{g}_{N-1}(w_j,b,z)|] &\le \int_0^1 \frac{b}{N\eta}\mathbb{E}\left[\frac{C'\sqrt{N}}{\sqrt{\eta}}\left(\sum_{m=1}^{N} a_{mj}^2\right)^{1/2}\right]\mathrm{d}\,t \\ &\le C'\int_0^1 \frac{b}{\sqrt{N}\eta^{3/2}}\left(\mathbb{E}\left[\sum_{m=1}^{N} a_{mj}^2\right]\right)^{1/2}\mathrm{d}\,t\,.\end{aligned}$$

Since f is bounded, we have for some new constant C_f' that

$$\mathbb{E}[|g_N(w_j,b,z) - \tilde{g}_{N-1}(w_j,b,z)|] \le \frac{C_f' b\sqrt{\lambda}}{\eta^{3/2}\sqrt{N}}\,.$$

Using the fact that $\mathbf{g}_N^t$ is bounded by C_f for all t, we get

$$\mathbb{E}[|\mathrm{e}^{\lambda\tilde{g}_{N-1}} - \mathrm{e}^{\lambda g_N}|] \le \mathbb{E}[|\tilde{g}_{N-1} - g_N|]\mathrm{e}^{2C_f\lambda} = \mathrm{O}\left(\frac{\sqrt{\lambda}}{\eta^{3/2}\sqrt{N}}\right)\,.$$

Since this is an error of the same order as $q_{N,\lambda}(u,z)$, we can absorb it into the existing error $q_{N,\lambda}$. Finally, using (2.54) and the interpolation argument allows us to write (2.49) as

$$\mathbb{E}[\mathrm{e}^{iur_{jj}^N}] = 1 - \mathrm{e}^{-\lambda d_j}\sqrt{u}\int_0^\infty \frac{\mathrm{J}_1(2\sqrt{uv})}{\sqrt{v}}\mathrm{e}^{ivz}\mathbb{E}\left[\mathrm{e}^{\lambda g_N(w_j,\frac{v}{\lambda},z)}\right]\mathrm{d}\,v + q_{N,\lambda}(u,z)\,,$$

which proves the proposition. □

Now, consider the expression (2.42) from the Proposition 2.5.4. If we multiply throughout by $f(x,w_j)$ and then sum over j, and finally scale by N, we get

$$\begin{aligned}\mathbb{E}[g_N(x,u,z)] = {} & \frac{1}{N}\sum_{j=1}^{N} f(x,w_j) \\ & - \frac{1}{N}\sum_{j=1}^{N} f(x,w_j)\mathrm{e}^{-\lambda d_j}\sqrt{u}\int_0^\infty \frac{\mathrm{J}_1(2\sqrt{uv})}{\sqrt{v}}\mathrm{e}^{ivz}\mathbb{E}\left[\mathrm{e}^{\lambda g_N(w_j,\frac{v}{\lambda},z)}\right]\mathrm{d}\,v \\ & + q_{N,\lambda}(u,z)\,.\end{aligned} \tag{2.59}$$

Consider the space of Lipschitz functions Lip($\mathbb{R}$) defined as

$$\mathrm{Lip}(\mathbb{R}) = \left\{ h \in C_b(\mathbb{R}) : \sup_x |h(x)| \leq 1, \sup_{x \neq y} \frac{|h(x) - h(y)|}{|x-y|} \leq C_L, 0 < C_L < \infty \right\}.$$

Now, under the bounded Lipshitz metric $d_{BL}(\cdot,\cdot)$ given by

$$d_{BL}(\mu,\nu) = \sup_{h \in \mathrm{Lip}(\mathbb{R})} \left\{ \left| \int h \,\mathrm{d}\,\mu - \int h \,\mathrm{d}\,\nu \right| \right\},$$

we have

$$\mu_{W_N} \implies \mu_w \quad \text{if and only if} \quad d_{BL}(\mu_{W_N}, \mu_w) \to 0,$$

where $W_N = w_{o_N}$ for a uniformly chosen vertex o_N. So, taking f to be Lipschitz in one coordinate (and since we already have that f is bounded), the first term in the RHS of (2.59) becomes

$$\frac{1}{N} \sum_{j=1}^{N} f(x, w_j) = \int f(x,y) \mu_{W_N}(\mathrm{d}\,y) \leq d_f(x) + E_N, \tag{2.60}$$

where $E_N = d_{BL}(\mu_{W_N}, \mu_w)$.

Recall from (2.13) that we have

$$d_f(w_j) := \int f(x, w_j) \mu_w(\mathrm{d}\,x).$$

Then, one simply gets

$$|\mathrm{e}^{-\lambda d_j} - \mathrm{e}^{-\lambda d_f(w_j)}| \leq \lambda E_N \mathrm{e}^{2\lambda}. \tag{2.61}$$

Thus, using (2.60) and (2.61) in (2.42) gives us

$$\begin{aligned} &\mathbb{E}[g_N(x,u,z)] \\ &= d_f(x) - \frac{1}{N} \sum_{j=1}^{N} f(x, w_j) \mathrm{e}^{-\lambda d_f(w_j)} \left(\sqrt{u} \int_0^\infty \frac{\mathrm{J}_1(2\sqrt{uv})}{\sqrt{v}} \mathrm{e}^{ivz} \mathbb{E}\left[\mathrm{e}^{\lambda g_N\left(w_j, \frac{v}{\lambda}, z\right)} \right] \mathrm{d}\,v \right) \\ &+ \tilde{q}_{N,\lambda}(u,z), \end{aligned} \tag{2.62}$$

where

$$\tilde{q}_{N,\lambda}(u,z) = q_{N,\lambda}(u,z) + \mathrm{O}(E_N).$$

Finally, for a fixed $x \in [0,\infty)$, define

$$\mathbf{I}_g(y) = f(x,y) \mathrm{e}^{-\lambda d_f(y)} \left(\sqrt{u} \int_0^\infty \frac{\mathrm{J}_1(2\sqrt{uv})}{\sqrt{v}} \mathrm{e}^{ivz} \mathbb{E}\left[\mathrm{e}^{\lambda g_N\left(y, \frac{v}{\lambda}, z\right)} \right] \mathrm{d}\,v \right).$$

Then, we have the following lemma.

Lemma 2.5.6.
$\mathbf{I}_g(y)$ is Lipschitz.

Proof. Consider $\mathbf{I}_g(y)$ as defined. Then,

$$
\begin{aligned}
&|\partial_y \mathbf{I}_g(y)| \\
&\leq \left| \partial_y f(x,y) \mathrm{e}^{-\lambda d_f(y)} \left(\sqrt{u} \int_0^\infty \frac{\mathrm{J}_1(2\sqrt{uv})}{\sqrt{v}} \mathrm{e}^{ivz} \mathbb{E}\left[\mathrm{e}^{\lambda g_N\left(y, \frac{v}{\lambda}, z\right)} \right] \mathrm{d}\, v \right) \right| \\
&+ \left| f(x,y) \mathrm{e}^{-\lambda d_f(y)} \partial_y d_f(y) \left(\sqrt{u} \int_0^\infty \frac{\mathrm{J}_1(2\sqrt{uv})}{\sqrt{v}} \mathrm{e}^{ivz} \mathbb{E}\left[\mathrm{e}^{\lambda g_N\left(y, \frac{v}{\lambda}, z\right)} \right] \mathrm{d}\, v \right) \right| \\
&+ \left| f(x,y) \mathrm{e}^{-\lambda d_f(y)} \left(\sqrt{u} \int_0^\infty \frac{\mathrm{J}_1(2\sqrt{uv})}{\sqrt{v}} \mathrm{e}^{ivz} \mathbb{E}\left[\mathrm{e}^{\lambda g_N\left(y, \frac{v}{\lambda}, z\right)} \right] \partial_y g_N(y, v/\lambda, z) \, \mathrm{d}\, v \right) \right| .
\end{aligned}
\tag{2.63}
$$

Recall that a function is Lipschitz if and only if it has a bounded derivative. Thus, if f is Lipschitz in y, the first term in (2.63) is uniformly bounded in y. Moreover, this makes the second term in (2.63) bounded as well since

$$
|\partial_y d_f(y)| \leq \int_0^\infty |\partial_y f(x,y)| \mu_w(\mathrm{d}\, x) \tag{2.64}
$$

is bounded. To justify interchanging the derivative and the integral in (2.64), we have to utilise Theorem 2.6.2 for which we need to verify the following conditions.

- $f(x,y)$ is μ_w-integrable for each y and the map $y \mapsto f(x,y)$ is continuous for each x.
- For each x, the derivative $\partial_y f(x,y)$ exists.
- For each y, there is a μ_w-integrable function $\Psi_y(x)$ and a neighbourhood U_y containing y, such that for all $y' \in U_y$, $|\partial_{y'} f(x,y')| \leq \Psi_y(x)$.

The first and second are trivial to check, and by Lipschitz property, since $\partial_y f(x,y) \equiv \mathit{const.}$, we have $\Psi_y(x) \equiv \mathit{const}$, which is integrable on $[0,\infty)$ since μ_w is a probability measure.

Finally, for notational convenience, let $h(y,v)$ be denote

$$
h(y,v) = \frac{\mathrm{J}_1(2\sqrt{uv})}{\sqrt{v}} \mathrm{e}^{ivz} \mathbb{E}\left[\mathrm{e}^{\lambda g_N(y,v,z)} \right] .
$$

Once again, we need to verify the three conditions as above to apply Theorem 2.6.2. Note that $h(y,v)$ is integrable with respect to v. Moreover,

$$
\partial_y h(y,v) = h(y,v) \partial_y g_N(y,v,z)
$$

where one can compute

$$\partial_y g_N(y,v,z) = \frac{1}{N}\sum_{k=1}^{N} \partial_y f(w_k,y)\mathrm{e}^{ivr_{kk}}\,,$$

which again is bounded. Thus, $\partial_y h(y,v)$ exists, and is bounded above by $C_0 v^{-\frac{1}{2}}\mathrm{e}^{-\eta v}$, which is integrable with respect to v. This verifies the three conditions and allows us to pull the derivative inside the third term in (2.63), and also makes that term bounded. Thus, $\mathbf{I}_g(y)$ is Lipschitz. □

Since $\mathbf{I}_g(y)$ is Lipschitz, we can exploit the weak convergence of μ_w under the Lipschitz metric d_{BL} in (2.62) to give us

$$\begin{aligned}
&\mathbb{E}[g_N(x,u,z)] \\
&\quad = d_f(x) \\
&\quad - \int_0^\infty f(x,y)\mathrm{e}^{-\lambda d_f(y)}\left(\sqrt{u}\int_0^\infty \frac{\mathrm{J}_1(2\sqrt{uv})}{\sqrt{v}}\mathrm{e}^{ivz}\mathbb{E}\left[\mathrm{e}^{\lambda g_N\left(y,\frac{v}{\lambda},z\right)}\right]\mathrm{d}\,v\right)\mu_w(\mathrm{d}\,y) \qquad (2.65)\\
&\quad + \tilde{q}_{N,\lambda}(u,z)\,.
\end{aligned}$$

Recall the Banach space as defined in (2.11), and consider $\phi\in(\mathcal{B},\|\cdot\|)$. In this space, consider the map

$$\begin{aligned}
&F_z(\phi)(x,u) \\
&= d_f(x) - \sqrt{u}\int_0^\infty f(x,y)\mathrm{e}^{-\lambda d_f(y)}\left(\sqrt{u}\int_0^\infty \frac{\mathrm{J}_1(2\sqrt{uv})}{\sqrt{v}}\mathrm{e}^{ivz}\mathrm{e}^{\lambda\phi\left(y,\frac{v}{\lambda},z\right)}\,\mathrm{d}\,v\right)\mu_w(\mathrm{d}\,y).
\end{aligned} \tag{2.66}$$

Note that ϕ also implicitly depends on z but we drop that for notational purposes since we fix z throughout.

Take $\phi_1,\phi_2\in(\mathcal{B},\|\cdot\|)$ such that $\|\phi_1\|,\|\phi_2\|\le C_f$. Then, using the norm we defined in (2.11) and inequality 2.44, from (2.66) we get

$$
\begin{aligned}
&\|F_z(\phi_1) - F_z(\phi_2)\| \\
&\leq \sup_{x,u\geq 0} \sqrt{\frac{1}{1+u}} \bigg| \int_0^\infty f(x,y)\mathrm{e}^{-\lambda d_f(y)} \\
&\qquad \times \left(\sqrt{u}\int_0^\infty \frac{\mathrm{J}_1(2\sqrt{uv})}{\sqrt{v}} \mathrm{e}^{ivz}\left(\mathrm{e}^{\lambda\phi_1(y,\frac{v}{\lambda})} - \mathrm{e}^{\lambda\phi_2(y,\frac{v}{\lambda})}\right)\mathrm{d}\, v\right)\mu_w(\mathrm{d}\, y)\bigg| \\
&\leq \sup_{u\geq 0}\sqrt{\frac{1}{1+u}}\int_0^\infty\int_0^\infty \frac{\lambda}{\sqrt{v}}\mathrm{e}^{-\eta v}\left|\phi_1\left(y,\frac{v}{\lambda}\right) - \phi_2\left(y,\frac{v}{\lambda}\right)\right| \\
&\qquad\qquad \times \mathrm{e}^{\lambda|\phi_1(y,\frac{v}{\lambda})|+\lambda|\phi_2(y,\frac{v}{\lambda})|}\,\mathrm{d}\, v\ \ \mu_w(\mathrm{d}\, y) \\
&\leq \lambda\|\phi_1-\phi_2\|\int_0^\infty\int_0^\infty \frac{\lambda}{\sqrt{v}}\mathrm{e}^{-\eta v}\sup_{y,v\geq 0}\frac{\sqrt{1+v/\lambda}}{\sqrt{1+v/\lambda}}\mathrm{e}^{\lambda|\phi_1(y,\frac{v}{\lambda})|+\lambda|\phi_2(y,\frac{v}{\lambda})|}\,\mathrm{d}\, v\ \ \mu_w(\mathrm{d}\, y) \\
&\leq \lambda\|\phi_1-\phi_2\|\int_0^\infty\int_0^\infty \frac{\sqrt{1+v/\lambda}}{\sqrt{v}}\mathrm{e}^{-\eta v}\exp\left(\lambda\sqrt{1+v/\lambda}(\|\phi_1\|+\|\phi_2\|)\right)\mathrm{d}\, v\ \ \mu_w(\mathrm{d}\, y) \\
&\leq \|\phi_1-\phi_2\|\int_0^\infty\left(\frac{\mathrm{e}^{-\eta v}}{\sqrt{v}} + \frac{\mathrm{e}^{-\eta v}}{\sqrt{\lambda}}\right)\mathrm{e}^{2C_f\sqrt{\lambda v}}\,\mathrm{d}\, v \leq \frac{C_1}{\eta^{5/2}}\|\phi_1-\phi_2\|\,,
\end{aligned}
$$

where C_1 is the constant upper bound to the integral of the form

$$\int_0^\infty c_1\mathrm{e}^{-c_2x+c_3\sqrt{x}}\,\mathrm{d}\, x$$

for some $c_3 > 0$, and is finite. Taking $\eta > 0$ sufficiently large, we get that F_z is a contraction in an open ball $B \subset \mathcal{B}$ of radius $C_f < \infty$, and thus, by the Banach Fixed Point Theorem, there exists a unique ϕ^* such that $\phi^* = F_z(\phi^*)$ for $F_z : B \to B$.

We are now ready to prove a concentration result. Recall the function $G_N(u)$ defined in (2.12) as

$$G_N(u) = \frac{1}{N}\sum_{i=1}^N \mathrm{e}^{iur_{ii}^N}.$$

If we now define a new function $\tilde{G}_N(x,u)$ that acts identically on the first coordinate as

$$\tilde{G}_N(x,u) := G_N(u),$$

then one can see that $\sup_{x,u}\frac{1}{\sqrt{1+u}}\tilde{G}_N(x,u) < \infty$, and so $\tilde{G}_N(x,u) \in \mathcal{B}$, and consequently, a concentration result for $\tilde{G}_N$ would imply concentration for G_N.

Proposition 2.5.7 (Concentration and convergence).
For any $z \in \mathbb{C}^+$ and $x \in [0,\infty)$, and uniformly over u in $[0,1]$, we have $\mathbb{E}[g_N(x,u,z)] \xrightarrow{N\to\infty} \phi^*(x,u)$. Further, we have

$$\mathbb{E}\left[\|g_N - \mathbb{E}[g_N]\|^2\right] = o(1), \quad \textit{and}$$
$$\mathbb{E}\left[\left\|\tilde{G}_N - \mathbb{E}[\tilde{G}_N]\right\|^2\right] = o(1)\,.$$

Proof of Proposition 2.5.7. Let $\delta_N(x,u,z)$ denote the error

$$\delta_N(x,u,z) := \mathrm{e}^{\lambda g_N(x,u,z)} - \mathrm{e}^{\lambda \mathbb{E}[g_N(x,u,z)]}.$$

Let $1 \le k \ne l \le N$ and consider the covariance

$$A_{k,l} := \mathbb{E}[\mathrm{e}^{iur^N_{kk}}\mathrm{e}^{iur^N_{ll}}] - \mathbb{E}[\mathrm{e}^{iur^N_{kk}}]\mathbb{E}[\mathrm{e}^{iur^N_{kk}}]. \tag{2.67}$$

Using (2.46) for the first term and Proposition 2.5.4 for the second term, we get

$$\begin{aligned}
&A_{k,l} = \\
&-\mathbb{E}[T_j] - \mathbb{E}[T_k] \\
&+ u\int\int \frac{\mathrm{J}_1(2\sqrt{uv_1})}{\sqrt{v_1}}\frac{\mathrm{J}_1(2\sqrt{uv_2})}{\sqrt{v_2}}\mathrm{e}^{i(v_1+v_2)z}\mathbb{E}\left[\mathrm{e}^{iv_1\sum\limits_{l\ne j}\tilde{r}^{N-1}_{ll}a^2_{jl} + iv_2\sum\limits_{l\ne k}\tilde{r}^{N-1}_{ll}a^2_{kl}}\right]\mathrm{d}\,v_1\,\mathrm{d}\,v_2 \\
&+ \mathbb{E}[\tilde{T}_j] + \mathbb{E}[\tilde{T}_k] \\
&- u\int\int \frac{\mathrm{J}_1(2\sqrt{uv_1})}{\sqrt{v_1}}\frac{\mathrm{J}_1(2\sqrt{uv_2})}{\sqrt{v_2}}\mathrm{e}^{i(v_1+v_2)z}\mathbb{E}\left[\mathrm{e}^{\lambda g_N(w_j,\frac{v_1}{\lambda},z)+\lambda g_N(w_k,\frac{v_2}{\lambda},z)}\right]\mathrm{d}\,v_1\,\mathrm{d}\,v_2\,,
\end{aligned} \tag{2.68}$$

where T_i and $\tilde{T}_i$ are the RHS of equations (2.46) and (2.42) respectively, and differ by the error $q_{N,\lambda}(u,z)$ in expectation. In the first double integral of (2.68), one can do the interpolation argument term-wise, and obtain the error $C_I q^2_{N,\lambda}(u,z) + q^2_{N,\lambda}(u,z)$ by making a difference with the second double integral in (2.68), where C_I is the constant upper bound to $\tilde{T}_k$ for any k. Thus, we have that

$$|A_{k,l}| \le C'_I q_{N,\lambda}(u,z) + q^2_{N,\lambda}(u,z). \tag{2.69}$$

Using inequality 2.44 on $\delta_N(x,u,z)$ gives us

$$\begin{aligned}
&\mathbb{E}[|\delta_N(x,u,z)|^2] \\
&= \mathbb{E}\left[\left|\mathrm{e}^{\lambda g_N(x,u,z)} - \mathrm{e}^{\lambda\mathbb{E}[g_N(x,u,z)]}\right|^2\right] \le C_1\mathbb{E}\left[|g_N(x,u,z) - \mathbb{E}[g_N(x,u,z)]|^2\right].
\end{aligned}$$

since $|g_N(x,v,z)| \leq C_f$ and $C_1 = \mathrm{e}^{2\lambda C_f}$. We can now bound this by using the definition of g_N to get

$$\begin{aligned}&\mathbb{E}[|\delta_N(x,u,z)|^2]\\&\leq \frac{C_1}{N^2}\left|\sum_{k,l=1}^{N}\mathbb{E}[f(x,w_k)\mathrm{e}^{iur^N_{kk}}f(x,w_l)\mathrm{e}^{iur^N_{ll}}]-\mathbb{E}[f(x,w_k)\mathrm{e}^{iur^N_{kk}}]\mathbb{E}[f(x,w_l)\mathrm{e}^{iur^N_{ll}}]\right|.\end{aligned} \tag{2.70}$$

Since f is deterministic, we can pull it out of the expectation and take it common, giving us

$$\mathbb{E}[|\delta_N(x,u,z)|^2] \leq \frac{C_1}{N^2}\left|\sum_{k,l=1}^{N} f(x,w_k)f(x,w_l)A_{k,l}\right|,$$

where $A_{k,l}$ is as in (2.67). We can conclude using the triangle inequality that

$$\mathbb{E}[|\delta_N(x,u,z)|^2] \leq C_1C_f^2\sup_{k,l}|A_{k,l}| = \mathrm{O}\left(\frac{\lambda\sqrt{u}}{\eta^{5/2}\sqrt{N}}\right). \tag{2.71}$$

For $\eta > 0$ sufficiently large, taking the norm, we get

$$\mathbb{E}\left[\left\|\mathrm{e}^{\lambda g_N} - \mathrm{e}^{\lambda\mathbb{E}[g_N]}\right\|^2\right] = o(1). \tag{2.72}$$

However, δ_N is a bounded analytic function in $[0,\infty)^2 \times \mathbb{C}^+$. Using the identity theorem from complex analysis, which states that if two holomorphic functions agree in an open set of the domain then they must agree everywhere on the domain, we have that since $\delta_N \to 0$ on an open set of the upper-half complex plane, it must approach 0 everywhere on the upper-half plane. Since the error in (2.71) can be absorbed in $\tilde{q}_{N,\lambda}(u,z)$, using 2.44 gives us

$$\begin{aligned}&\mathbb{E}[g_N(x,u,z)]\\&= d_f(x) - \int_0^\infty f(x,y)\mathrm{e}^{-\lambda d_f(y)}\left(\sqrt{u}\int_0^\infty \frac{\mathrm{J}_1(2\sqrt{uv})}{\sqrt{v}}\mathrm{e}^{ivz}\mathrm{e}^{\lambda\mathbb{E}[g_N(y,\frac{v}{\lambda},z)]}\,\mathrm{d}\,v\right)\mu_w(\mathrm{d}\,y)\\&+\tilde{q}_{N,\lambda}(u,z),\end{aligned} \tag{2.73}$$

where the error vanishes in the norm as

$$\|\tilde{q}_{N,\lambda}\| = \|q_{N,\lambda}(u,z) + \mathrm{O}(E_N)\| \leq \sup_{x,u\geq 0}\left|C\frac{\lambda\sqrt{u}}{\eta^{5/2}\sqrt{N}}\right| + E_N = o(1).$$

Now, consider the function $\tilde{G}_N(x,u)$ and the error

$$\Delta_N(u) := \tilde{G}_N(x,u) - \mathbb{E}[\tilde{G}_N(x,u)].$$

By definition of $\tilde{G}_N$, one can see that expanding $\Delta_N(u)$ will yield an expression similar to (2.70) modulo f, and so, using (2.69) again, we get that

$$\mathbb{E}[|\Delta_N|^2] \leq C_1 C_f^2 \sup_{k,l} |A_{k,l}| = \mathrm{O}\left(\frac{\lambda\sqrt{u}}{\eta^{5/2}\sqrt{N}}\right).$$

By taking the norm and again using the identity theorem, we get that Δ_N vanishes in $[0,\infty)^2 \times \mathbb{C}^+$ and thus

$$\mathbb{E}\left[\left\|\tilde{G}_N - \mathbb{E}[\tilde{G}_N]\right\|^2\right] = o(1). \tag{2.74}$$

A quick inspection of (2.70) shows that in fact we also have the concentration for g_N, since the RHS is precisely the upper bound on

$$\mathbb{E}[|g_N(x,u,z) - \mathbb{E}[g_N(x,u,z)]|^2],$$

and so,

$$\mathbb{E}\left[\|g_N - \mathbb{E}[g_N]\|^2\right] = o(1). \tag{2.75}$$

Finally, comparing (2.73) with the contraction mapping (2.66), we have the following:

$$\begin{aligned} \mathbb{E}[g_N(x,u,z)] &= F_z(\mathbb{E}[g_N(x,u,z)]) + \tilde{q}_{N,\lambda}(u,z), \\ \phi^*(x,u) &= F_z(\phi^*(x,u)). \end{aligned}$$

So, with $\eta > 0$ large enough and F_z being a contraction on $B \subset \mathcal{B}$ of radius C_f, we have

$$\|\mathbb{E}[g_N] - \phi^*\| \leq \|F_z(\mathbb{E}[g_N]) - F_z(\phi^*)\| + \|\tilde{q}_N\|,$$

and consequently,

$$\frac{1}{2}\|\mathbb{E}[g_N] - \phi^*\| \leq \|\tilde{q}_N\|.$$

Thus, since $\|\mathbb{E}g_N\| \leq C_f$,

$$\|\mathbb{E}[g_N] - \phi^*\| \xrightarrow{N\to\infty} 0.$$

As a quick remark, notice that

$$\|\phi^*\| \leq C_f, \tag{2.76}$$

since g_N is bounded.

Now, since $\mathbb{E}[g_N(x,u,z)]$ is an analytic function on $[0,\infty)^2 \times \mathbb{C}^+$, we have $\lim_{N\to\infty} \mathbb{E}[g_N(x,u,z)]$ is an analytic function. Again from the identity theorem of complex analysis, since $\lim_{N\to\infty} \mathbb{E}[g_N]$ and ϕ^* are analytic and agree on an open set of $[0,\infty)^2 \times \mathbb{C}^+$, they agree everywhere in the complex domain $[0,\infty)^2 \times \mathbb{C}^+$, and thus the convergence holds for any $z \in \mathbb{C}^+$. Note that *for a fixed* $z \in \mathbb{C}^+$, although both the functionals $\mathbb{E}[g_N]$ and ϕ^* live in $(\mathcal{B}, \|\cdot\|_{\mathcal{B}})$, the domain of ϕ^* is $[0,\infty)^2 \times \mathbb{C}^+$ since $\mathbb{E}[g_N]$ has the domain $[0,\infty)^2 \times \mathbb{C}^+$. Now, for each $z \in \mathbb{C}^+$, fixing u in the compact set $[0,1]$ gives us that for each $x \in [0,\infty)$ and uniformly over $u \in [0,1]$,

$$\mathbb{E}[g_N(x,u,z)] \xrightarrow{N\to\infty} \phi^*(x,u) \tag{2.77}$$

We can now prove Theorem 2.3.9.

Proof of Theorem 2.3.9. Equation (2.74) proves the concentration statement of Theorem 2.3.9. Recall that we had shown that

$$\mathbb{E}\left[\mathrm{e}^{iur_{jj}^N}\right] = 1 - \mathrm{e}^{-\lambda d_j}\sqrt{u}\int_0^\infty \frac{\mathrm{J}_1(2\sqrt{uv})}{\sqrt{v}}\mathrm{e}^{ivz}\mathbb{E}\left[\mathrm{e}^{\lambda g_N\left(w_j,\frac{v}{\lambda},z\right)}\right]\mathrm{d}\,v + q_{N,\lambda}(u,z)\,,$$

and so,

$$\begin{aligned}\mathbb{E}[G_N(u,z)] &= \frac{1}{N}\sum_{j=1}^N \mathbb{E}[\mathrm{e}^{iur_{jj}^N}]\\ &= 1 - \frac{1}{N}\sum_{j=1}^N \mathrm{e}^{-\lambda d_j}\sqrt{u}\int_0^\infty \frac{\mathrm{J}_1(2\sqrt{uv})}{\sqrt{v}}\mathrm{e}^{ivz}\mathbb{E}\left[\mathrm{e}^{\lambda g_N\left(w_j,\frac{v}{\lambda},z\right)}\right]\mathrm{d}\,v + q_{N,\lambda}(u,z)\,.\end{aligned} \tag{2.78}$$

Next, we see that the function

$$\tilde{\mathbf{I}}_g(y) = \mathrm{e}^{-\lambda d_f(y)}\sqrt{u}\int_0^\infty \frac{\mathrm{J}_1(2\sqrt{uv})}{\sqrt{v}}\mathrm{e}^{ivz}\mathrm{e}^{\lambda\phi^*(y,v/\lambda)}\,\mathrm{d}\,v$$

is Lipschitz by using an argument similar to Lemma 2.5.6. Thus, we get

$$\mathbb{E}[G_N(u,z)] = \\ 1 - \int_0^\infty \mathrm{e}^{-\lambda d_f(y)}\left(\sqrt{u}\int_0^\infty \frac{\mathrm{J}_1(2\sqrt{uv})}{\sqrt{v}}\mathrm{e}^{ivz}\mathbb{E}\left[\mathrm{e}^{\lambda g_N\left(y,\frac{v}{\lambda},z\right)}\right]\mathrm{d}\,v\right)\mu_w(\mathrm{d}\,y)\ + \tilde{q}_{N,\lambda}(u,z)\,.$$

Since from Proposition 2.5.7 we have concentration for g_N, using inequality (2.44) we have that

$$\mathbb{E}[G_N(u,z)] = \\ 1 - \int_0^\infty \mathrm{e}^{-\lambda d_f(y)}\sqrt{u}\int_0^\infty \frac{\mathrm{J}_1(2\sqrt{uv})}{\sqrt{v}}\mathrm{e}^{ivz}\mathrm{e}^{\lambda\mathbb{E}\left[g_N\left(y,\frac{v}{\lambda},z\right)\right]}\,\mathrm{d}\,v\ \ \mu_w(\mathrm{d}\,y) + \tilde{q}_{N,\lambda}(u,z)\,.$$

Finally, taking the limit $N \to \infty$ gives us

$$\lim_{N\to\infty} \mathbb{E}[G_N(u,z)] =$$
$$1 - \int_0^\infty \mathrm{e}^{-\lambda d_f(y)} \sqrt{u} \int_0^\infty \frac{\mathrm{J}_1(2\sqrt{uv})}{\sqrt{v}} \mathrm{e}^{ivz} \mathrm{e}^{\lambda\phi^*(y,v/\lambda)} \,\mathrm{d}\, v \ \ \mu_w(\mathrm{d}\, y)\,, \tag{2.79}$$

completing the proof of Theorem 2.3.9. □

§2.5.2 Deriving the expression for the Stieltjes Transform

Since we took u to be in $[0,1]$, we can take a derivative with respect to u and evaluate it at $u=0$. Recall from equation (2.78) that we have

$$\mathbb{E}[G_N(u,z)] = \frac{1}{N}\mathbb{E}\sum_{j=1}^N \mathrm{e}^{iur_{jj}^N}$$
$$= 1 - \frac{1}{N}\sum_{j=1}^N \mathrm{e}^{-\lambda d_j}\sqrt{u}\int_0^\infty \frac{\mathrm{J}_1(2\sqrt{uv})}{\sqrt{v}} \mathrm{e}^{ivz}\mathbb{E}\left[\mathrm{e}^{\lambda g_N(w_j,\frac{v}{\lambda},z)}\right] \mathrm{d}\, v + q_{N,\lambda}(u,z)\,.$$

Note that by definition, $G_N(u,z)$ is a bounded function, and thus by DCT, limit operations can be interchanged with expectation. We would like to take a derivative with respect to u and evaluate at $u=0$ to extract out $\mathrm{tr}(\mathrm{R}_{\mathbf{A}_N}(z))$ from the LHS of (2.78). On the other hand, we would first like to take $N\to\infty$ for the RHS to remove the error term. To interchange these operations, we have the following result.

Proposition 2.5.8.
Both the limits $\lim_{N\to\infty} \frac{\partial}{\partial u}\mathbb{E}[G_N(u,z)]\big|_{u=0}$ *and* $\frac{\partial}{\partial u}\lim_{N\to\infty}\mathbb{E}[G_N(u,z)]\big|_{u=0}$ *exist and are equal.*

Proof. We fix a $z\in\mathbb{C}^+$. Now, $\lim_{N\to\infty}\mathbb{E}[G_N(u,z)]$ exists due to the RHS of (2.78), which we denote by $G(u,z)$. If we define $H_N(u,z)$ and $H(u,z)$ as

$$H_N(u,z) = \frac{\mathbb{E}[G_N(u,z)] - \mathbb{E}[G_N(0,z)]}{u}\,,$$
$$H(u,z) = \frac{G(u,z) - G(0,z)}{u}\,.$$

Then,

$$\lim_{u\to 0} H_N(u,z) = \frac{\partial}{\partial u}\mathbb{E}[G_N(u,z)]\bigg|_{u=0}\,,$$
$$\lim_{u\to 0} H(u,z) = \frac{\partial}{\partial u}G(u,z)\bigg|_{u=0}\,.$$

We would like to claim

$$\lim_{N\to\infty} \frac{\partial}{\partial u}\mathbb{E}[G_N(u,z)]\bigg|_{u=0} = \frac{\partial}{\partial u}G(u,z)\bigg|_{u=0}.$$

Thus, we want to interchange the order of limits. Note that

$$\lim_{N\to\infty} H_N(u,z) = H(u,z)$$

uniformly in $u \in (0,1]$, and

$$\lim_{u\to 0} H_N(u,z) = \frac{\partial}{\partial u}\mathbb{E}[G_N(u,z)]\bigg|_{u=0} = \mathbb{E}[\operatorname{tr}(\mathrm{R}_{\mathbf{A}_N}(z))]$$

for each N, where the limit can be taken inside the expectation using dominated convergence. Thus, using [Rudin, 1976, Theorem 7.11], we have that the limits $\lim_{u\to 0} H(u,z)$ and $\lim_{N\to\infty}\mathbb{E}[\operatorname{tr}(\mathrm{R}_{\mathbf{A}_N}(z))]$ exist and are equal. □

We are now ready to prove Corollary 2.3.10.

Proof of Corollary 2.3.10. We now do precisely as we stated before Proposition 2.5.8. We evaluate the derivative at $u = 0$ and then take $N \to \infty$ on the LHS of (2.78), and we do the reverse for the RHS of (2.78). Note that since $\lim_{N\to\infty}\mu_{N,\lambda} = \mu_\lambda$ in probability, $\mathrm{S}_{\mathbf{A}_N}(z) \to S_{\mu_\lambda}(z)$ and also $\bar{\mathrm{S}}_{\mathbf{A}_N}(z) \to S_{\mu_\lambda}(z)$ as $N \to \infty$ for all $z \in \mathbb{C}^+$. Thus, we then obtain using Proposition 2.5.8

$$\begin{aligned}
&i\,\mathrm{S}_{\mu_\lambda}(z)\\
&= i\lim_{N\to\infty}\bar{\mathrm{S}}_{\mathbf{A}_N}(z) \overset{(2.36)}{=} i\lim_{N\to\infty}\mathbb{E}\operatorname{tr}(\mathrm{R}_{\mathbf{A}_N}(z)) \overset{(2.39)}{=} \lim_{N\to\infty}\frac{\partial}{\partial u}\mathbb{E}[G_N(u,z)]\bigg|_{u=0}\\
&= \frac{\partial}{\partial u}\lim_{N\to\infty}\mathbb{E}[G_N(u,z)]\bigg|_{u=0}\\
&\overset{(2.79)}{=} -\frac{\partial}{\partial u}\int_0^\infty \mathrm{e}^{-\lambda d_f(y)}\sqrt{u}\int_0^\infty \frac{\mathrm{J}_1(2\sqrt{uv})}{\sqrt{v}}\mathrm{e}^{ivz}\mathrm{e}^{\lambda\phi_z^*(y,\frac{v}{\lambda})}\,\mathrm{d}\,v\;\;\mu_w(\mathrm{d}\,y)\bigg|_{u=0}\\
&= -\int_0^\infty \mathrm{e}^{-\lambda d_f(y)}\frac{\partial}{\partial u}\sqrt{u}\int_0^\infty \frac{\mathrm{J}_1(2\sqrt{uv})}{\sqrt{v}}\mathrm{e}^{ivz}\mathrm{e}^{\lambda\phi_z^*(y,\frac{v}{\lambda})}\,\mathrm{d}\,v\;\;\mu_w(\mathrm{d}\,y)\bigg|_{u=0}.
\end{aligned} \tag{2.80}$$

We now wish to evaluate the derivative on the RHS of (2.80). Let $K(u)$ denote

$$K(u) := \sqrt{u}\int_0^\infty \frac{\mathrm{J}_1(2\sqrt{uv})}{\sqrt{v}}\mathrm{e}^{ivz}\mathrm{e}^{\lambda\phi_z^*(y,\frac{v}{\lambda})}\,\mathrm{d}\,v. \tag{2.81}$$

Observe that

$$\sum_{k\geq 0}\int_0^\infty \frac{v^k}{k!(k+1)!}\mathrm{e}^{-\eta v}\,\mathrm{d}\,v = \sum_{k\geq 0}\frac{\Gamma(k+1)}{k!(k+1)!\eta^k} \leq \mathrm{e}^{1/\eta} \tag{2.82}$$

for $\eta > 0$ by a change of variables. If we expand the Bessel function as defined in (2.16) in equation (2.81) and take the absolute value, we observe using (2.82) and using $|\phi^*(x,u)| \leq C_f$ (from (2.76)), that we can use Fubini's Theorem to interchange the integral with the summand. Thus, we have

$$\begin{aligned} K(u) &= \sqrt{u} \int_0^\infty \frac{\mathrm{J}_1(2\sqrt{uv})}{\sqrt{v}} \mathrm{e}^{ivz} \mathrm{e}^{\lambda \phi_z^*(y,\frac{v}{\lambda})} \,\mathrm{d}\, v \\ &= \sqrt{u} \int_0^\infty \frac{1}{\sqrt{v}} \sum_{k=0}^\infty \frac{(-1)^k (\sqrt{uv})^{2k+1}}{k!(k+1)!} \mathrm{e}^{ivz} \mathrm{e}^{\lambda \phi_z^*(y,\frac{v}{\lambda})} \,\mathrm{d}\, v \\ &= \sum_{k=0}^\infty \frac{(-1)^k u^{k+1}}{k!(k+1)!} \int_0^\infty v^k \mathrm{e}^{ivz} \mathrm{e}^{\lambda \phi_z^*(y,\frac{v}{\lambda})} \,\mathrm{d}\, v \,. \end{aligned}$$

Denote by $I_k(y)$ the integral

$$I_k(y) := \int v^k \mathrm{e}^{ivz} \mathrm{e}^{\lambda \phi_z^*(y,\frac{v}{\lambda})} \,\mathrm{d}\, v \,.$$

Therefore,

$$\frac{K(u)}{u} = \sum_{k=0}^\infty \frac{(-1)^k u^k}{k!(k+1)!} I_k(y) = I_0(y) + \sum_{k\geq 1} \frac{(-1)^k u^k}{k!(k+1)!} I_k(y) =: I_0(y) + \sum_{k=1}^\infty a_k(u) \,, \tag{2.83}$$

where $a_k(u)$ denotes

$$a_k(u) := \frac{(-1)^k u^k I_k(y)}{k!(k+1)!}.$$

Note that for any k, we have that $I_k(y)$ is finite since

$$|I_k(y)| \leq \int_0^\infty v^k \mathrm{e}^{-\eta v} \mathrm{e}^{C_f \lambda} \,\mathrm{d}\, v = \frac{\mathrm{e}^{C_f \lambda}}{\eta^{k+1}} \Gamma(k+1) \,.$$

Since $K(0) = 0$ and by (2.83) it follows that

$$\left. \frac{\partial}{\partial u} K(u) \right|_{u=0} = \lim_{u \to 0} \frac{K(u)}{u} = I_0(y) + \lim_{u\to 0} \sum_{k\geq 1} a_k(u), \tag{2.84}$$

Therefore we would like to evaluate $\lim_{u\to 0} \sum_{k\geq 1} a_k(u)$. Note that

$$|a_k(u)| \leq \frac{\mathrm{e}^{C_f \lambda} \Gamma(k+1)}{\eta^{k+1} k!(k+1)!}$$

, as u is bounded by 1. Note that the series

$$\sum_{k\geq 1} \frac{\Gamma(k+1) \mathrm{e}^{C_f \lambda}}{k!(k+1)! \eta^{k+1}} = \frac{\mathrm{e}^{C_f \lambda}}{\eta^2} \sum_{k\geq 0} \frac{1}{\eta^k (k+2)!} \leq \frac{\mathrm{e}^{C_f \lambda} \mathrm{e}^{\frac{1}{\eta}}}{\eta^2}$$

converges, and consequently by the dominated convergence theorem, we have

$$\lim_{u\to 0}\sum_{k\geq 1} a_k(u) = \sum_{k\geq 1}\lim_{u\to 0} a_k(u) = 0.$$

Thus by (2.84) we have

$$\lim_{u\to 0}\frac{K(u)}{u} = I_0(y).$$

Therefore we get

$$i\,\mathrm{S}_{\mu_\lambda}(z) = -\int_0^\infty \mathrm{e}^{-\lambda d_f(y)} I_0(y)\mu_w(\mathrm{d}\,y) - \int_0^\infty \mathrm{e}^{-\lambda d_f(y)}\int_0^\infty \mathrm{e}^{ivz}e^{\lambda\phi_z^*(y,\frac{v}{\lambda})}\,\mathrm{d}\,v\ \ \mu_w(\mathrm{d}\,y).$$

To conclude the argument, we use Lemma 2.5.3 with Theorem 2.3.7 to state that $\mathrm{S}_{\mathbf{A}_N}(z)$ converges in probability to $\mathrm{S}_{\mu_\lambda}(z)$ for each $z\in\mathbb{C}^+$. □

We conclude with the proof of Corollary 2.3.11

Proof of Corollary 2.3.11. From Corollary 2.3.10, we have

$$\mathrm{S}_{\mu_\lambda}(z) = i\int_0^\infty\int_0^\infty \mathrm{e}^{ivz}\mathrm{e}^{-\lambda d_f(y)+\lambda\phi^*(y,v/\lambda)}\,\mathrm{d}\,v\ \ \mu_w(\mathrm{d}\,y).$$

Recall that

$$\begin{aligned}&\phi^*(x,u)\\ &= d_f(x) - \int_0^\infty f(x,y)\mathrm{e}^{-\lambda d_f(y)}\left(\sqrt{u}\int_0^\infty \frac{\mathrm{J}_1(2\sqrt{uv})}{\sqrt{v}}\mathrm{e}^{ivz}\mathrm{e}^{\lambda\phi^*(y,\frac{v}{\lambda})}\,\mathrm{d}\,v\right)\mu_w(\mathrm{d}\,y)\end{aligned} \tag{2.85}$$

is the unique analytical solution of the fixed point equation as in (2.66). Expanding the Bessel function $\mathrm{J}_1(x)$ in (2.85) using (2.16) gives

$$\begin{aligned}&\phi^*(x,u)\\ &= d_f(x) - \int_0^\infty f(x,y)\mathrm{e}^{-\lambda d_f(y)}\left(\int_0^\infty\sum_{k\geq 0}\frac{(-1)^k u^{k+1}v^k}{k!(k+1)!}\mathrm{e}^{ivz}\mathrm{e}^{\lambda\phi^*(y,\frac{v}{\lambda})}\,\mathrm{d}\,v\right)\mu_w(\mathrm{d}\,y).\end{aligned} \tag{2.86}$$

We would like to interchange the summand and integral with respect to v in (2.86). Using the $z=\zeta+i\eta$ for some $\zeta\in\mathbb{R}$ and $\eta>0$, we have that

$$\begin{aligned}&\sum_{k\geq 0}\int_0^\infty\left|\frac{(-1)^k u^{k+1}v^k}{k!(k+1)!}\mathrm{e}^{ivz}\mathrm{e}^{-\lambda d_f(y)+\lambda\phi^*(y,v/\lambda)}\right|\mathrm{d}\,v\\ &\leq \mathrm{e}^{C_f\lambda-\lambda d_f(y)}\sum_{k\geq 0}\frac{u^{k+1}\Gamma(k+1)}{k!(k+1)!\eta^{k+1}} \leq \frac{u}{\eta}\mathrm{e}^{C_f\lambda-\lambda d_f(y)}\mathrm{e}^{u/\eta}.\end{aligned}$$

Thus, by Fubini's Theorem, we can interchange the summand with the integral with respect to v, giving us

$$\phi^*(x,u) = \\ d_f(x) - \int_0^\infty f(x,y)\mathrm{e}^{-\lambda d_f(y)} \left(\sum_{k\geq 0} \frac{(-1)^k u^{k+1}}{k!(k+1)!} \int_0^\infty v^k \mathrm{e}^{ivz} \mathrm{e}^{\lambda\phi^*\left(y,\frac{v}{\lambda}\right)} \,\mathrm{d}\,v \right) \mu_w(\mathrm{d}\,y). \tag{2.87}$$

Now, denote by $\mathcal{H}^\lambda(z,y)$ the function

$$\mathcal{H}^\lambda(z,y) := i\int_0^\infty \mathrm{e}^{ivz}\mathrm{e}^{-\lambda d_f(y)+\lambda\phi^*(y,v/\lambda)}\,\mathrm{d}\,v. \tag{2.88}$$

Then, by Corollary 2.3.10, we can see that $\mathrm{S}_{\mu_\lambda}(z) = \int_0^\infty \mathcal{H}^\lambda(z,y)\mu_w(\mathrm{d}\,y)$. From (2.87) we get that

$$\begin{aligned}\phi^*(x,u) = d_f(x) &- u\int_0^\infty f(x,y)\int_0^\infty \mathrm{e}^{ivz}\mathrm{e}^{-\lambda d_f(y)+\lambda\phi^*(y,v/\lambda)}\,\mathrm{d}\,v\;\;\mu_w(\mathrm{d}\,y) \\ &- \int_0^\infty f(x,y)\sum_{k\geq 1}\int_0^\infty \frac{(-1)^k u^{k+1}v^k}{k!(k+1)!}\mathrm{e}^{ivz}\mathrm{e}^{-\lambda d_f(y)+\lambda\phi^*(y,v/\lambda)}\,\mathrm{d}\,v\;\;\mu_w(\mathrm{d}\,y),\end{aligned}$$

and so, we can write

$$\phi^*(x,u) = d_f(x) + iu\int_0^\infty f(x,y)\mathcal{H}^\lambda(z,y)\mu_w(\mathrm{d}\,y) + T(x,u,\lambda,z) \tag{2.89}$$

where

$$T(x,u,\lambda,z) \\ := -\int_0^\infty f(x,y)\sum_{k\geq 1}\int_0^\infty \frac{(-1)^k u^{k+1}v^k}{k!(k+1)!}\mathrm{e}^{ivz}\mathrm{e}^{-\lambda d_f(y)+\lambda\phi^*(y,v/\lambda)}\,\mathrm{d}\,v\;\;\mu_w(\mathrm{d}\,y). \tag{2.90}$$

Substituting $u = v/\lambda$ for $v \in \mathbb{R}_+$ in (2.89) and multiplying throughout by λ, we have

$$-\lambda d_f(x) + \lambda\phi^*(x,v/\lambda) = iv\int_0^\infty f(x,y)\mathcal{H}^\lambda(z,y)\mu_w(\mathrm{d}\,y) + \lambda T(x,v/\lambda,\lambda,z).$$

We begin by claiming the following:

Claim 2.5.9.
For any $x,u \geq 0$, we have

$$|\mathrm{e}^{-\lambda d_f(x)+\lambda\phi^*(x,u)}| \leq 1. \tag{2.91}$$

Then, one can see that

$$
\begin{aligned}
|T(x, v/\lambda, \lambda, z)| &\le \int_0^\infty f(x,y)\frac{v}{\lambda\eta}\left(\sum_{k\ge 1}\frac{v^k\Gamma(k+1)}{\eta^k\lambda^k k!(k+1)!}\right)\mu_w(\mathrm{d}\,y) \\
&\le \frac{v^2}{\lambda^2\eta^2}\mathrm{e}^{\frac{v}{\eta\lambda}}d_f(x),
\end{aligned}
$$

and so for each $v \in (0,\infty)$

$$
\lim_{\lambda\to\infty}\lambda|T(x, v/\lambda, \lambda, z)| \to 0.
$$

Thus, from (2.89), for any v we have

$$
\lim_{\lambda\to\infty}(-\lambda d_f(x) + \lambda\phi^*(x, v/\lambda)) = iv\lim_{\lambda\to\infty}\int_0^\infty f(x,y)\mathcal{H}^\lambda(z,y)\mu_w(\mathrm{d}\,y). \tag{2.92}
$$

What remains now is to justify Claim 2.5.9, and taking the limit $\lambda\to\infty$ inside the integral in (2.92).

First we consider the homogeneous case when $f\equiv 1$. Recall from Remark 2.3.12, that due to the lack of dependency of one coordinate, we denote $\widetilde{\phi^*}(u) = \phi^*(x, v/\lambda)$ Then,

$$
\widetilde{\phi^*}(u) = 1 - \sqrt{u}\int_0^\infty \frac{\mathrm{J}_1(2\sqrt{uv})}{\sqrt{v}}\mathrm{e}^{ivz}\mathrm{e}^{-\lambda+\lambda\widetilde{\phi^*}(v/\lambda)}\,\mathrm{d}\,v,
$$

and from (2.92) we have $\lim_{\lambda\to\infty}(-\lambda+\lambda\widetilde{\phi^*}(v/\lambda)) = iv\,\mathrm{S}_{\mu_f}(z)$. Moreover, from Corollary 2.3.10, we have

$$
\mathrm{S}_{\mu_\lambda}(z) = i\int_0^\infty \mathrm{e}^{ivz}\mathrm{e}^{-\lambda+\lambda\widetilde{\phi^*}(v/\lambda)}\,\mathrm{d}\,v.
$$

Since $f\equiv 1$, from (2.76) we have that $C_f = 1$ and $|\widetilde{\phi^*}|\le 1$. Then, $|\mathrm{e}^{-\lambda+\lambda\widetilde{\phi^*}}|\le 1$, justifying Claim 2.5.9. Thus, the expression inside the integral is uniformly bounded by $\mathrm{e}^{-\eta v}$. Using dominated convergence, we can pull the limit $\lambda\to\infty$ inside the integral to obtain

$$
\mathrm{S}_{\mu_f}(z) = i\int_0^\infty \mathrm{e}^{ivz}\mathrm{e}^{iv\,\mathrm{S}_{\mu_f}(z)}\,\mathrm{d}\,v = -\frac{1}{z+\mathrm{S}_{\mu_f}(z)},
$$

which is precisely the Stieltjes transform of the semicircle law.

In the case of general f, recall from (2.77) that for any x and u,

$$
\phi^*(x,u) = \lim_{N\to\infty}\frac{1}{N}\mathbb{E}\left[\sum_{i=1}^N f(x,w_i)\mathrm{e}^{iur_{ii}^N}\right].
$$

Now, for any N, by trivially bounding the complex exponential $\mathrm{e}^{iur^N_{ii}}$ by 1 for any i, we have that

$$\left|\frac{1}{N}\mathbb{E}\sum_{i=1}^{N} f(x,w_i)\mathrm{e}^{iur^N_{ii}}\right| \leq \frac{1}{N}\sum_{i=1}^{N}|f(x,w_i)| = \frac{1}{N}\sum_{i=1}^{N} f(x,w_i).$$

Thus, by triangle inequality, we have that

$$|\phi^*(x,u)| \leq |\phi^*(x,u) - \mathbb{E}[g_N(x,u,z)]| + \frac{1}{N}\sum_{i=1}^{N} f(x,w_i).$$

Thus, we have that

$$\begin{aligned}
-\frac{\lambda}{N}\sum_{i=1}^{N} f(x,w_i) + \lambda|\phi^*(x,u)| &\leq \lambda\sqrt{1+u}\frac{1}{\sqrt{1+u}}\,|\phi^*(x,u) - \mathbb{E}[g_N(x,u,z)]| \\
&\leq \lambda\sqrt{1+u}\,\|\phi^* - \mathbb{E}g_N\|_{\mathcal{B}}. \qquad (2.93)
\end{aligned}$$

Taking $N \to \infty$ on both sides in (2.93) yields that

$$-\lambda d_f(x) + \lambda|\phi^*(x,u)| \leq 0.$$

Using this, we conclude that

$$\left|\mathrm{e}^{-\lambda d_f(x)+\lambda\phi^*(x,u)}\right| \leq \mathrm{e}^{-\lambda d_f(x)}\mathrm{e}^{\lambda|\phi^*(x,u)|} \leq 1 \qquad (2.94)$$

for any x and u, proving Claim 2.5.9. Now, to evaluate $\lim_{\lambda\to\infty}\mathrm{S}_{\mu_\lambda}(z)$, we take the limit inside the integral in the RHS of (2.17) using DCT, which we can use from (2.94). This gives us

$$\mathrm{S}_{\mu_f}(z) = \lim_{\lambda\to\infty}\mathrm{S}_{\mu_\lambda}(z) = i\int_0^\infty\int_0^\infty \mathrm{e}^{ivz}\lim_{\lambda\to\infty}\left(\mathrm{e}^{-\lambda d_f(y)+\lambda\phi^*(y,v/\lambda)}\right)\mathrm{d}\,v\;\;\mu_w(\mathrm{d}\,y).$$

and so, using (2.92), we get

$$\mathrm{S}_{\mu_f}(z) = i\int_0^\infty\int_0^\infty \mathrm{e}^{ivz}\lim_{\lambda\to\infty}\mathrm{e}^{iv\int_0^\infty f(x,y)\mathcal{H}^\lambda(z,x)\mu_w(\mathrm{d}\,x)}\,\mathrm{d}\,v\;\;\mu_w(\mathrm{d}\,y). \qquad (2.95)$$

Recall from (2.88) that

$$\mathcal{H}^\lambda(z,y) = i\int_0^\infty \mathrm{e}^{ivz}\mathrm{e}^{-\lambda d_f(y)+\lambda\phi^*(y,v/\lambda)}\,\mathrm{d}\,v.$$

Again using (2.94), we have that the integral is bounded in absolute value, and so, using DCT allows us to define

$$\mathcal{H}(z,y) := \lim_{\lambda\to\infty}\mathcal{H}^\lambda(z,y)$$

where $\int_0^\infty \mathcal{H}(z,y)\mu_w(\mathrm{d}\,y) = \mathrm{S}_{\mu_f}(z)$. Moreover, since $|\mathcal{H}^\lambda(z,y)|$ is bounded by a constant, and μ_w is a probability measure, we use DCT once again to take the limit $\lambda \to \infty$ inside $\int_0^\infty f(x,y)\mathcal{H}^\lambda(z,x)\mu_w(\mathrm{d}\,x)$. Thus, we obtain

$$\begin{aligned}\mathrm{S}_{\mu_f}(z) &= i\int_0^\infty \int_0^\infty \mathrm{e}^{ivz}\mathrm{e}^{iv\int_0^\infty f(x,y)\mathcal{H}(z,x)\mu_w(\mathrm{d}\,x)}\,\mathrm{d}\,v\;\;\mu_w(\mathrm{d}\,y)\\ &= -\int_0^\infty \frac{\mu_w(\mathrm{d}\,y)}{z+\int_0^\infty f(x,y)\mathcal{H}(z,x)\mu_w(\mathrm{d}\,x)}.\end{aligned}$$

The proof follows by observing that $\mathcal{H}(z,x)$ satisfies the analytic equation defined in (2.6). □

§2.6 Appendix

Proposition 2.6.1 (Banach Space).
Let $X = [0,\infty)^2$ and consider the space $\mathcal{B}$ defined by

$$\mathcal{B} = \left\{\phi : X \to \mathbb{C}\;\; analytic \;\middle|\; \sup_{x,y\geq 0} \frac{|\phi(x,y)|}{\sqrt{1+y}} < \infty\right\}$$

and consider the norm

$$\|\phi\|_{\mathcal{B}} = \sup_{x,y\geq 0}\frac{|\phi(x,y)|}{\sqrt{1+y}}.$$

Then, $(\mathcal{B}, \|\cdot\|_{\mathcal{B}})$ is a Banach space.

Proof of Proposition 2.6. For ease of notation, throughout this argument, $\|\cdot\| := \|\cdot\|_{\mathcal{B}}$. Clearly $\|\cdot\|$ is a norm, and thus, $(\mathcal{B}, \|\cdot\|_{\mathcal{B}})$ is a normed vector space.

Let $\{\phi_n\}_n$ be a Cauchy sequence in $(\mathcal{B}, \|\cdot\|_{\mathcal{B}})$. Thus, for all $\epsilon > 0$, there is an $N_\varepsilon \in \mathbb{N}$ such that for all $m, n > N_\varepsilon$,

$$\|\phi_m - \phi_n\| < \varepsilon.$$

Let μ be the Lebesgue measure on X. Define

$$E_{mn} = \{(x,y) \in X : |\phi_n(x,y) - \phi_m(x,y)| > \|\phi_n - \phi_m\|\sqrt{1+y}\}.$$

Then, $\mu(E_{mn}) = 0$. Let $E = \bigcup\limits_{m,n} E_{mn}$ and $F = E^c$. Then, $\mu(E) = 0$, and

$$F = \{(x,y) \in X : |\phi_n(x,y) - \phi_m(x,y)| \leq \|\phi_n - \phi_m\|\sqrt{1+y}\}.$$

So, for all $\varepsilon > 0$, we have an N_ε such that for all $(x, y) \in F$ and $m, n > N_\varepsilon$,

$$|\phi_n(x, y) - \phi_m(x, y)| < \varepsilon\sqrt{1 + y}.$$

Let $\psi_m(x, y) := \frac{\phi_m(x,y)}{\sqrt{1+y}}$. Then, we have for all $(x, y) \in F$ and $m, n > N_\varepsilon$

$$|\psi_n(x, y) - \psi_m(x, y)| < \varepsilon.$$

In other words, for all $(x, y) \in F$, denoting $a_n = \psi_n(x, y)$ gives us that $\{a_n\}_n$ is a Cauchy sequence in the metric space $(\mathbb{C}, |\cdot|)$. Since $\mathbb{C}$ is a complete metric space, for all $(x, y) \in F$, there exists a limit $a := \lim_n a_n$, that is, for all $(x, y) \in F$, there exists a ψ such that

$$\psi(x, y) := \lim_{n\to\infty} \psi_n(x, y).$$

For $(x, y) \in E$ with $\mu(E) = 0$, $\psi(x, y) = 0$. This is a well-defined limit. Note that since ϕ_n lives in $(\mathcal{B}, \|\cdot\|_{\mathcal{B}})$, ψ_n lives in $(L^\infty(X), \|\cdot\|_\infty)$, and we thus conclude that

$$\|\psi_n - \psi_m\|_\infty < \varepsilon.$$

Passing the limit through m, we have

$$\|\psi_n - \psi\|_\infty < \varepsilon.$$

For all $(x, y) \in X$, define

$$\phi(x, y) = \psi(x, y)\sqrt{1 + y}.$$

One can see that $\|\phi_n - \phi\| = \|\psi_n - \psi\|_\infty$. Use triangle inequality to conclude $\phi \in (\mathcal{B}, \|\cdot\|_{\mathcal{B}})$ □

For the next theorem, we refer the reader to [Billingsley, 2012, Theorem 16.8].

Theorem 2.6.2 (Interchanging derivative and integral).
Consider the measure space $(\Omega, \mathcal{F}, \mu)$ and an open set $A \subset \mathbb{R}$. Let $f : A\times\Omega \to \mathbb{C}$ be such that for each $x \in A$, $\omega \mapsto f(x, \omega)$ is $\mu-$integrable, and moreover for $\mu-$a.e. ω, $x \mapsto f(x, \omega)$ is continuous. Consider the function $g : A \to \mathbb{C}$ defined by

$$g(x) = \int_\Omega f(x, \omega)\mu(\mathrm{d}\,\omega).$$

Suppose that for each ω the partial derivative $\partial_x f(x, \omega)$ of f with respect to x exists. Then, if for every x, there is a non-negative $\mu-$integrable function $h_x : \Omega \to \mathbb{C}$ and a neighbourhood U_x containing x such that for all $x' \in U_x$, $|\partial_{x'} f(x', \omega)| \leq h_x(\omega)$, then, $g(x)$ is continuously differentiable and

$$\partial_x g(x) = \int_\Omega \partial_x f(x, \omega)\mu(\mathrm{d}\,\omega).$$

CHAPTER 3

Adjacency spectra of kernel-based random graphs

This chapter is based on:
A. Cipriani, R.S. Hazra, N. Malhotra, M. Salvi. *Spectrum of dense kernel-based random graphs.* [*arxiv:2502:09415*], 2025.

Abstract

Kernel-based random graphs (KBRGs) are a broad class of random graph models that account for inhomogeneity among vertices. We consider KBRGs on a discrete $d-$dimensional torus $\mathbf{V}_N$ of size N^d. Conditionally on an i.i.d. sequence of Pareto weights $(W_i)_{i\in\mathbf{V}_N}$ with tail exponent $\tau-1>0$, we connect any two points i and j on the torus with probability

$$p_{ij} = \frac{\kappa_\sigma(W_i, W_j)}{\|i-j\|^\alpha} \wedge 1$$

for some parameter $\alpha > 0$ and $\kappa_\sigma(u,v) = (u \vee v)(u \wedge v)^\sigma$ for some $\sigma \in (0, \tau - 1)$. We focus on the adjacency operator of this random graph and study its empirical spectral distribution. For $\alpha < d$ and $\tau > 2$, we show that a non-trivial limiting distribution exists as $N \to \infty$ and that the corresponding measure $\mu_{\sigma,\tau}$ is absolutely continuous with respect to the Lebesgue measure. $\mu_{\sigma,\tau}$ is given by an operator-valued semicircle law, whose Stieltjes transform is characterised by a fixed point equation in an appropriate Banach space. We analyse the moments of $\mu_{\sigma,\tau}$ and prove that the second moment is finite even when the weights have infinite variance. In the case $\sigma = 1$, corresponding to the so-called scale-free percolation random graph, we can explicitly describe the limiting measure and study its tail.

§3.1 Introduction

Kernel-based spatial random graphs encompass a wide variety of classical random graph models where vertices are embedded in some metric space. In their simplest form (see Jorritsma et al. [2023] for a more complete exposition) they can be defined as follows. Let V be the vertex set of the graph and sample a collection of weights $(W_i)_{i\in V}$, which are independent and identically distributed (i.i.d.), serving as marks on the vertices. Conditionally on the weights, two vertices i and j are connected by an undirected edge with probability

$$\mathbb{P}\left(i \leftrightarrow j \mid W_i, W_j\right) = \kappa(W_i, W_j)\|i-j\|^{-\alpha} \wedge 1\,, \tag{3.1}$$

where κ is a symmetric kernel, $\|i-j\|$ denotes the distance between the two vertices in the underlying metric space and $\alpha > 0$ is a constant parameter. Common choices for κ include:

$$\kappa_{\text{triv}}(w,\, v) \equiv 1, \qquad \kappa_{\text{strong}}(w,\, v) = w \vee v,$$
$$\kappa_{\text{prod}}(w, v) = w\, v, \qquad \kappa_{\text{pa}}(w, v) = (w \vee v)(w \wedge v)^{\sigma_{\text{pa}}}.$$

In the above $\sigma_{\text{pa}} = \alpha(\tau-1)/d - 1$, where $\tau - 1$ is the exponent of the tail distribution of the weights, such that the kernel κ_{pa} mimics the form that appears in preferential attachment models [Jorritsma et al., 2023], while the trivial kernel κ_{triv} corresponds to the classical long-range percolation model [Schulman, 1983, Newman and Schulman, 1986]. The kernel κ_{prod} yields a model which is substantially equivalent to scale-free percolation, introduced in Deijfen et al. [2013], which has connection probabilities of the form

$$1 - \exp\left(-W_i W_j \|i-j\|^{-\alpha}\right).$$

Various percolation properties for kernel-based spatial random graphs are known on $\mathbb{Z}^d$ and beyond (Deprez et al. [2015], Hao and Heydenreich [2023], van der Hofstad and Komjáthy [2017], Gracar et al. [2021], Jorritsma et al. [2024], see also Deprez and Wüthrich [2019], Dalmau and Salvi [2021] for a version of the same in the continuum) as well as the behaviour of interacting particle systems on them [Berger, 2002, Heydenreich et al., 2017, Komjáthy and Lodewijks, 2020, Cipriani and Salvi, 2024, Gracar and Grauer, 2024, Bansaye and Salvi, 2024, Komjáthy et al., 2023]. In contrast, their spectral properties, to the best of the authors' knowledge, have received less attention.

As a branch of random matrix theory, the study of the spectrum of random graphs has wide applications ranging from the study of random Schrödinger operators [Carmona and Lacroix, 2012, Geisinger, 2015] and quantum chaos in physics, to the analysis of community structures [Bordenave et al., 2015] and diffusion processes in network science, to the problems of spectral clustering [Champion et al., 2020] and graph embeddings [Gallagher et al., 2024] in data science. Many challenges remain unsolved in this area, even for the simplest models. As a prominent example, for bond percolation on $\mathbb{Z}^2$ it is known that the expected spectral measure has a continuous component if and only if $p > p_c$, but this result has not yet been established in higher dimensions [Bordenave et al., 2017]. In this chapter, we begin the study of spectral properties of spatial inhomogeneous random graphs, which in turn have been proposed as models for several real-world networks (see e.g. Dalmau and Salvi [2021]).

We will work with KBRGs in the typical setting where the weights (W_i) have support in $[1, \infty)$ and the kernel κ is an increasing function of the weights. Let us recall that in this case the vertices of KBRG random graphs on $\mathbb{Z}^d$ have almost surely infinite degree as soon as $\alpha < d$. Thus, as it happens in many percolation problems, the regime $\alpha > d$ would be the most appealing (and the toughest to tackle). In the present work we will focus instead on the dense case $\alpha < d$. We consider the discrete torus with N^d vertices equipped with the torus distance $\|\cdot\|$. The weights are sampled independently from a Pareto distribution with parameter $\tau - 1$ with $\tau > 2$. Conditionally on the weights, vertices i and j are connected independently from other pairs with probability given by (3.1) with a kernel of the form $\kappa_\sigma(w, v) := (w \vee v)(w \wedge v)^\sigma$. It is worth noting a difference between our connection probability and that studied recently in Jorritsma et al. [2023], van der Hofstad et al. [2023], where the connection probabilities are given by

$$\mathbb{P}\left(i \leftrightarrow j \mid W_i, W_j\right) = \left(\kappa_\sigma(W_i, W_j)\|i - j\|^{-d} \wedge 1\right)^\alpha .$$

The two forms can be made equivalent through a simple modification of the weights and an appropriate choice of α.

We call $\mathbb{G}_N$ the random graph obtained with this procedure and study the empirical spectral distribution of its adjacency matrix, appropriately scaled. Note that when $\alpha = 0$ we recover the (inhomogeneous) Erdős–Rényi random graph (modulo a tweak inserting a suitable tuning parameter ε_N) . In recent years, there has been significant research on inhomogeneous Erdős–Rényi random graphs, which can be equivalently modelled by Wigner matrices with a variance profile. The limiting spectral distribution of the adjacency matrix of such graphs has been studied in Chakrabarty et al. [2021b], Zhu [2020], Bose et al. [2022], while local eigenvalue statistics have been analysed in Dumitriu and Zhu [2019], Ajanki et al. [2019]. Zhu and Zhu [2024] studies the fluctuations of the linear eigenvalue statistics for a wide range of such inhomogeneous graphs. Additionally, various properties of the largest eigenvalue have been investigated in Cheliotis and Louvaris [2024], Husson [2022], Chakrabarty et al. [2022], Ducatez et al. [2024]. One of the most significant properties of the limiting spectral measure for random graphs is its absolute continuity with respect to the Lebesgue measure, which is closely tied to the concept of mean quantum percolation [Bordenave et al., 2017, Anantharaman et al., 2021, Arras and Bordenave, 2023]. Quantum percolation investigates whether the limiting measure has a non-trivial absolutely continuous spectrum. Recently, it was shown in Arras and Bordenave [2023] that the adjacency operator of a supercritical Poisson Galton-Watson tree has a non-trivial absolutely continuous part when the average degree is sufficiently large. Additionally, Bordenave et al. [2017] demonstrated that supercritical bond percolation on $\mathbb{Z}^d$ has a non-trivial absolutely continuous part for $d = 2$. These results motivate similar questions for KBRGs.

Our contributions: Results and proofs

Here below we showcase our main results and the novelties of our proofs Recall that we work in the regime $\alpha < d$ and $\tau > 2$. We also restrict to values of σ in $(0, \tau - 1)$.

(a) In Theorem 3.2.1 we show that, after scaling the adjacency matrix of $\mathbb{G}_N$ by $c_0 N^{(d-\alpha)/2}$, the empirical spectral distribution converges weakly in probability to a deterministic measure $\mu_{\sigma,\tau}$. The classical approach to proving the convergence of the empirical distribution is generally through either the method of moments or the Stieltjes transform. However, the limiting measure is expected to be heavy-tailed (see Figure 3.3) and so it is not determined by its moments. As a consequence, we cannot directly apply the method of moments. To overcome this issue, we pass through a truncation argument where we impose a maximal value to the weights, reducing the problem to well-behaved measures. To simplify the method of moments, we further reduce the model by substituting the adjacency matrix of $\mathbb{G}_N$ with a Gaussian matrix whose entries are centred and have roughly the same variance as before. This is made possible by a classical result of Chatterjee [2005]. Once we have shifted our attention to this simpler Gaussianised matrix with bounded weights, we can use the classical method of moments using finding its moments is made possible by a combinatorial argument on partitions and their graphical representation. Finally we remove the truncation effect.

(b) In Theorem 3.2.2 we investigate the graph corresponding to κ_{prod}, that is, when $\sigma = 1$. In this case we can explicitly identify $\mu_{1,\tau}$ as the free multiplicative convolution of the semicircle law and the measure of the weight distribution. In the $\sigma = 1$ case the moment expression derived in Theorem 3.2.1

simplifies, so the challenge is to recover the limiting measure from those moments. This is made possible thanks to the extension of the free multiplicative convolution to measures with unbounded support by Arizmendi and Pérez-Abreu [2009]. Furthermore, we show that $\mu_{1,\tau}$ has power-law tails with exponent $2(\tau - 1)$. This is based on a Breiman-type argument for free multiplicative convolutions [Kołodziejek and Szpojankowski, 2022].

(c) In Theorem 3.2.3 we explicitly derive the second moment of $\mu_{\sigma,\tau}$ and prove that it is finite and non-degenerate. The proof is based on the ideas of Chakrabarty et al. [2016, Theorem 2.2]. This result is noteworthy because our weight distribution may exhibit infinite variance in the chosen range of parameters. To show that the second moment is finite, we need to establish the uniform integrability of a sequence of measures converging to the limiting measure. This is achieved through an extension of Skorohod's representation theorem for measures that converge weakly in probability.

(d) In Theorem 3.2.4 we prove that $\mu_{\sigma,\tau}$ is absolutely continuous. What makes the result possible is that we are able to split the original matrix as a free sum of a standard Wigner matrix and another Wigner matrix with a carefully chosen variance profile (yielding, as a by-product, another characterisation of the limit measure $\mu_{\sigma,\tau}$). Once this is established, the result is a consequence of Biane [1997].

(e) In Theorem 3.2.5 we provide an analytical description of $\mu_{\sigma,\tau}$ when $\tau > 3$ and $\sigma < \tau - 2$. Removing the truncation in the method of moments proof of Theorem 3.2.1 does not yield an explicit characterisation of the limiting measure. On the other hand, certain moment recursions for the truncated Gaussian matrix that appear in the proof can be used to derive properties of $\mu_{\sigma,\tau}$ through the Stieltjes transform. When the weights are bounded, the limiting measure corresponds to the operator-valued semicircle law (Speicher [2011]). Its transform can be expressed in terms of functions solving an analytic recursive equation (see Avena et al. [2023], Zhu [2020] for similar results in other random graph ensembles). In our case, when the weights are heavy-tailed, this is no longer possible. We achieve instead the convergence of the analytic recursive equation by constructing a suitable Banach space and demonstrating that it forms a contractive mapping.

Outline of the article.
In Section 3.2 we will define the model and state precisely the main results. In Section 3.3 we will give some auxiliary results which will be used to prove the main theorems in the rest of the article. More precisely, in Section 3.4 we will prove the existence of the limiting ESD, and in Section 3.5 we will give estimates

on its tail behaviour. In Section 3.6 we will prove the non-degeneracy of the limiting measure and in Section 3.7 we will show its absolute continuity. Finally, Section 3.8 is devoted to describing the Stieltjes transform of the limiting ESD.

§3.2 Set-up and main results

§3.2.1 Random graph models

To introduce our models, we use $a \wedge b$ to denote the minimum of two real numbers a and b, and $a \vee b$ to denote their maximum.

(a) **Vertex set:** the vertex set is $\mathbf{V}_N := \{1, 2, \dots, N\}^d$. The vertex set is equipped with torus the distance $\|i - j\|$, where

$$\|i - j\| = \sum_{\ell=1}^{d} |i_\ell - j_\ell| \wedge (N - |i_\ell - j_\ell|).$$

(b) **Weights:** the weights $(W_i)_{i \in \mathbf{V}_N}$ are i.i.d. random variables sampled from a Pareto distribution W (whose law we denote by $\mathbf{P}$) with parameter $\tau - 1$, where $\tau > 1$. That is,

$$\mathbf{P}(W > t) = t^{-(\tau-1)} \mathbf{1}_{\{t \geq 1\}} + \mathbf{1}_{\{t < 1\}}. \tag{3.2}$$

(c) **Kernel:** the kernel function $\kappa_\sigma : [0, \infty) \times [0, \infty) \to [0, \infty)$ determines how the weights interact. In this article, we focus on kernel functions of the form

$$\kappa_\sigma(w, v) := (w \vee v)(w \wedge v)^\sigma, \tag{3.3}$$

where $\sigma \geq 0$.

(d) **Long-range parameter:** $\alpha > 0$ tunes the influence of the distance between vertices on their connection probability.

(e) **Connectivity function:** conditional on the weights, each pair of distinct vertices i and j is connected independently with probability $P^W(i \leftrightarrow j)$ given by

$$P^W(i \leftrightarrow j) := \mathbb{P}(i \leftrightarrow j \mid W_i, W_j) = \frac{\kappa_\sigma(W_i, W_j)}{\|i - j\|^\alpha} \wedge 1. \tag{3.4}$$

We will be using the short-hand notation $p_{ij} := \mathbb{P}(i \leftrightarrow j \mid W_i, W_j)$ for convenience. Note that the graph does not have self-loops (see Remark 3.4.1).

The associated graph is connected, as nearest neighbours with respect to the torus distance are always linked.

§3.2.2 Spectrum of a random graph

Let us denote the random graph generated by our choice of edge probabilities by $\mathbb{G}_N$. Let $\mathbb{A}_{\mathbb{G}_N}$ denote the adjacency matrix (operator) associated with this random graph, defined as

$$\mathbb{A}_{\mathbb{G}_N}(i,j) = \begin{cases} 1 & \text{if } i \leftrightarrow j, \\ 0 & \text{otherwise.} \end{cases}$$

Since the graph is finite, the adjacency matrix is always self-adjoint and has real eigenvalues. For $\alpha < d$, the eigenvalues require a scaling, which turns out to be independent of the kernel in our setup. Here we assume $\sigma \in (0, \tau - 1)$ and $\tau > 2$, ensuring that the vertex weights $(W_i)_{i \in \mathbf{V}_N}$ have finite mean. We define the scaling factor as

$$c_N = \frac{1}{N^d} \sum_{i \neq j \in \mathbf{V}_N} \frac{1}{\|i-j\|^\alpha} \sim c_0 N^{d-\alpha}, \tag{3.5}$$

where c_0 is a constant depending on α and d, and for two functions $f(\cdot)$ and $g(\cdot)$ we use $f(t) \sim g(t)$ to indicate that their quotient $f(t)/g(t)$ tends to one as t tends to infinity. The scaled adjacency matrix is then defined as

$$\mathbf{A}_N := \frac{\mathbb{A}_{\mathbb{G}_N}}{\sqrt{c_N}}. \tag{3.6}$$

The empirical measure that assigns a mass of $1/N^d$ to each eigenvalue of the $N^d \times N^d$ random matrix $\mathbf{A}_N$ is called the Empirical Spectral Distribution (ESD) of $\mathbf{A}_N$, denoted as

$$\mathrm{ESD}\,(\mathbf{A}_N) := \frac{1}{N^d} \sum_{i=1}^{N^d} \delta_{\lambda_i},$$

where $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_{N^d}$ are the eigenvalues of $\mathbf{A}_N$.

§3.2.3 Main results

We are now ready to state the main result of this article. Let μ_W denote the law of W. Here onwards, let $\mathbb{P} = \mathbf{P} \otimes P^W$ represent the joint law of the weights and the edge variables. Note that $\mathbb{P}$ depends on N, but we omit this dependence for simplicity. Let $\mathbb{E}, \mathbf{E}$, and E^W denote the expectation with respect to $\mathbb{P}, \mathbf{P}$, and P^W respectively. Furthermore, if $(\mu_N)_{N \geq 0}$ is a sequence of probability measures, we write $\lim_{N \to \infty} \mu_N = \mu_0$ to denote that μ_0 is the weak limit of the measures μ_N. Since the empirical spectral distribution is a random probability measure, we require the notion of convergence in probability in the context of weak convergence.

The Lévy-Prokhorov distance $d_L : \mathcal{P}(\mathbb{R})^2 \to [0, +\infty)$ between two probability measures μ and ν on $\mathbb{R}$ is defined as

$$d_L(\mu, \nu) := \inf\left\{\varepsilon > 0 \mid \mu(A) \le \nu\left(A^\varepsilon\right) + \varepsilon \text{ and } \nu(A) \le \mu\left(A^\varepsilon\right) + \varepsilon \quad \forall\, A \in \mathcal{B}(\mathbb{R})\right\},$$

where $\mathcal{B}(\mathbb{R})$ denotes the Borel σ-algebra on $\mathbb{R}$, and A^ε is the ε-neighbourhood of A. For a sequence of random probability measures $(\mu_N)_{N\ge 0}$, we say that

$$\lim_{N\to\infty} \mu_N = \mu_0 \text{ in } \mathbb{P}\text{-probability}$$

if, for every $\varepsilon > 0$,

$$\lim_{N\to\infty} \mathbb{P}(d_L(\mu_N, \mu_0) > \varepsilon) = 0.$$

The first result states the existence of the limiting spectral distribution of the scaled adjacency matrix.

Theorem 3.2.1 (Limiting spectral distribution).
Consider the random graph $\mathbb{G}_N$ on $\mathbf{V}_N$ with connection probabilities given by (3.4) *with parameters $\tau > 2$, $0 < \alpha < d$ and $\sigma \in (0, \tau - 1)$. Let $\mathrm{ESD}(\mathbf{A}_N)$ be the empirical spectral distribution of $\mathbf{A}_N$ defined in* (3.6)*. Then there exists a deterministic measure $\mu_{\sigma,\tau}$ on $\mathbb{R}$ such that*

$$\lim_{N\to\infty} \mathrm{ESD}(\mathbf{A}_N) = \mu_{\sigma,\tau} \qquad \textit{in } \mathbb{P}\textit{–probability}.$$

The remaining results focus of the properties of the limiting measure. First we note that when we set $\sigma = 1$ we can explicitly identify the limiting measure in terms of free multiplicative convolution. We refer the reader to Anderson et al. [2010, Section 5.2.3] for an exposition on free multiplicative and additive convolutions.

For two probability measures μ and ν the free multiplicative convolution $\mu \boxtimes \nu$ of the two measures is defined as the law of the product ab of free, random, non-commutative operators a and b, with laws μ and ν respectively. The free multiplicative convolution for two non-negatively supported measures was introduced in Bercovici and Voiculescu [1993]. Note that the semicircle law is not non-negatively supported and hence we use the extended definition of Arizmendi and Pérez-Abreu [2009] for the multiplicative convolution.

Theorem 3.2.2 (Limiting ESD for $\boldsymbol{\sigma = 1}$).
Consider the KBRG for $\sigma = 1$, while α, τ are as in the assumptions of Theorem 3.2.1. The the limiting spectral distribution $\mu_{1,\tau}$ is given by

$$\mu_{1,\tau} = \mu_{sc} \boxtimes \mu_W\,,$$

where μ_{sc} is the semicircle law

$$\mu_{sc}(\mathrm{d}\,x) = \frac{1}{2\pi}\sqrt{4-x^2}\,\boldsymbol{1}_{|x|\le 2}\,\mathrm{d}\,x$$

and $\boxtimes$ is the free multiplicative convolution of the two measures. Moreover, the limiting measure $\mu_{1,\tau}$ has a power-law tail, that is,

$$\mu_{1,\tau}(x,\infty) \sim \frac{1}{2}\big(m_1(\mu_W)\big)^{\tau-1} x^{-2(\tau-1)} \quad \textit{as } x \to \infty,$$

where $m_1(\nu)$ denotes the first moment of the probability measure ν.

In the general case, it is hard to explicitly identify the limiting measure, so we present some characterisations of it. Since we do not impose that $\tau > 3$ and consequently the weights can have infinite variance, it is not immediate if the second moment of the limiting measure is non-degenerate and finite. We prove this in the following result.

Theorem 3.2.3 (Non-degeneracy of the limiting measure).
Under the assumptions of Theorem 3.2.1, the second moment of the limiting measure $\mu_{\sigma,\tau}$ is given by

$$\int_{\mathbb{R}} x^2 \mu_{\sigma,\tau}(\mathrm{d}\,x) = (\tau-1)^2 \int_1^\infty \int_1^\infty \frac{1}{(x\wedge y)^{\tau-\sigma}(x\vee y)^{\tau-1}}\,\mathrm{d}\,x\,\mathrm{d}\,y \in (0,\infty).$$

Moreover, for $p \in \mathbb{N}$ and $p < (\tau-1)/(\sigma\vee 1)$, we have $\int_{\mathbb{R}} |x|^{2p}\,\mu_{\sigma,\tau}(\mathrm{d}\;x) < \infty$.

We state the following result as an independent theorem as the absolute continuity of the KBRG model deserves to be treated separately.

Theorem 3.2.4 (Absolute continuity).
Let $\tau > 2$ and $\sigma \in (0,\tau-1)$, then $\mu_{\sigma,\tau}$ is symmetric and absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}$.

We conclude the main results by providing an analytic description of the limiting measure in terms of its Stieltjes transform when we slightly restrict our parameters. Recall that, for $z \in \mathbb{C}^+$, where $\mathbb{C}^+$ denotes the upper half-plane of the complex plane, the Stieltjes transform of a measure μ on $\mathbb{R}$ is given by

$$\mathrm{S}_\mu(z) = \int_{\mathbb{R}} \frac{1}{x-z}\mu(\mathrm{d}x)\,. \tag{3.7}$$

Theorem 3.2.5 (Stieltjes transform).
Let $0 < \alpha < d$, $\tau > 3$ and $\sigma < \tau - 2$. Then there exists a unique analytic function a^* on $\mathbb{C}^+ \times [1,\infty)$ such that

$$\mathrm{S}_{\mu_{\sigma,\tau}}(z) = \int_1^\infty a^*(z,x)\mu_W(\mathrm{d}x),$$

where we recall that μ_W is the law of the random variable W.

The function a^* in the above theorem turns out to be a fixed point of a contraction mapping on an appropriate Banach space. The equation above shares similarities with the quadratic vector equations introduced and studied in Ajanki et al. [2019], although in our setting the measures have unbounded support. The properties and the proof of Theorem 3.2.5 are discussed in Section 3.8.

Remark 3.2.6 (Higher dimensions).
While we have presented our results for $0 < \alpha < d$, our proofs are worked out in the $d = 1$ setup. This is in order to avoid notational complications that would especially affect the clarity of Theorem 3.2.1. The limiting spectral distribution and its properties remain unchanged for $d > 1$.

§3.2.4 Examples, simulations and discussion

Firstly, in Figure 3.1 we plot the eigenvalue distribution of the adjacency matrix of two realisations of kernel-based graphs with different parameters, indicated at the top of the image. Secondly, in Figure 3.2 we sample 10 realisations of scale-

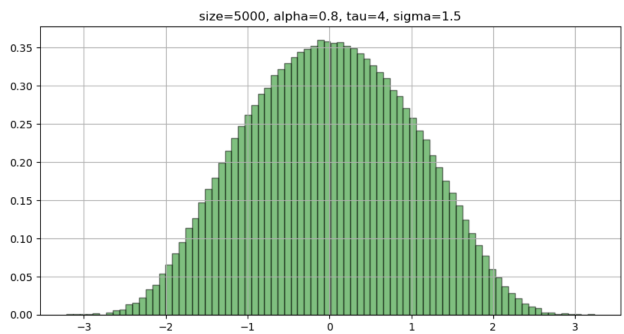


Figure 3.1: Eigenvalue distribution a KBRG realisation.

free percolation adjacency matrices of size 4000×4000 with $\sigma = 1$ and plot their eigenvalues (in green). We superpose on them the eigenvalues of the product $P_N G_N P_N$ of a GUE matrix G_N with a diagonal matrix P_N with i.i.d. entries distributed as $\sqrt{Pareto(\tau)}$ (in blue). Note that by Nica and Speicher [2006, Remark 14.2], Chakrabarty et al. [2021a, Remark 4.3], the a.s. limiting ESD of $P_N G_N P_N$ is $\mu_{sc} \boxtimes \mu_W$. All matrices are centred and rescaled by the sample second moment. Thirdly, to elucidate the tail behaviour of the limiting ESD when $\sigma = 1$ (Theorem 3.2.2) we draw in Figure 3.3 the empirical survival function of the eigenvalues of a matrix of size 7000×7000 in $x \geq 1.5$.

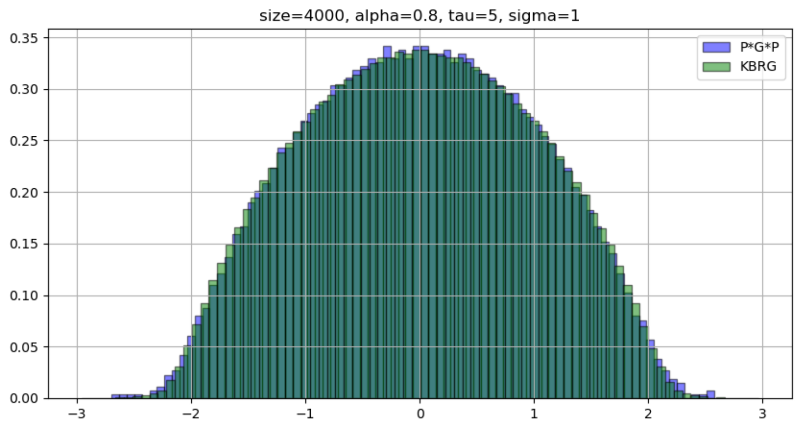


Figure 3.2: KBRG eigenvalue distribution and $P_N G_N P_N$ distribution.

Finally, we provide in Figure 3.4 a simulation of the eigenvalues of the Gaussian matrix $\tilde{\mathbf{A}}_{N,m,g}$ (see (3.24)) when $\alpha = 0$ and $N = 6000$. We compare this picture with the right-hand side of Figure 3.1, which has a small α. We conjecture that the atom appearing in the latter is due to high connectivity of the kernel-based realisation (if $\alpha = 0$, for all i, j we have that p_{ij} is identically one in (3.4)), whilst in the Gaussian setup this trivialization does not arise.

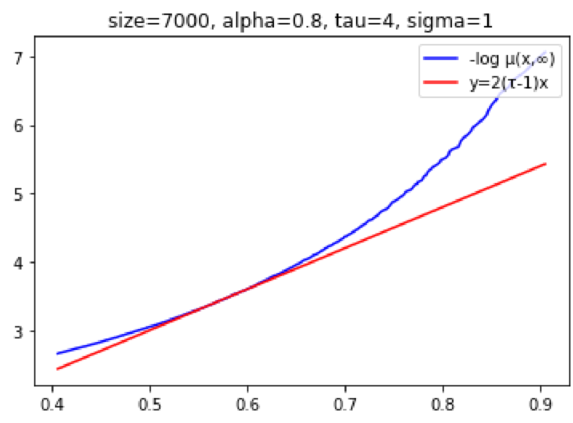


Figure 3.3: Negative of the log-empirical survival function and tails of Theorem 3.2.2 for $x \geq 1.5$.

Remark 3.2.7 (Sparse case).
We expect the case $\alpha > d$ to be very different due to the sparse nature of the graph. There has been a significant development in the area of spectral prop-

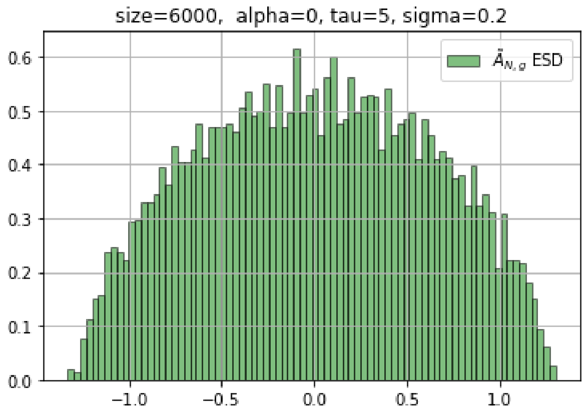


Figure 3.4: ESD for $\tilde{\mathbf{A}}_{N,m,g}$.

erties of sparse random graphs using the techniques of local weak convergence [Bordenave and Lelarge, 2010, Bordenave et al., 2017, 2011]. However, it is not immediately clear whether these techniques can be employed in our framework in order to determine the properties of the limiting measure: the underlying random graph generated in our model will not be tree-like to begin with. We plan to address this case in a future work.

§3.3 Notation and preliminary lemmas

In this section, we fix some notation and collect some technical lemmas that will be used in the proofs of our main results.

§3.3.1 Notation

We will use the Landau notation o_N, O_N indicating in the subscript the variable under which we take the asymptotic (typically this variable will grow to infinity unless otherwise specified). Universal positive constants are denoted as $c, c_1, \ldots,$ and their value may change with each occurrence. For an $N \times N$ matrix $A = (a_{ij})_{i,j=1}^N$ we use $\mathrm{Tr}(A) := \sum_{i=1}^N a_{ii}$ for the trace and $\mathrm{tr}(A) := N^{-1}\,\mathrm{Tr}(A)$ for the normalised trace. When $n \in \mathbb{N}$ we write $[n] := \{1, 2, \ldots, n\}$. We denote the cardinality of a set A as $\#A$, and, with a slight abuse of notation, $\#\sigma$ also denotes the number of cycles in a permutation σ.

§3.3.2 Technical lemmas

The following proposition, known as the *Hoffman-Wielandt inequality*, follows from Bai and Silverstein [2010, Corollary A.41].

Proposition 3.3.1 (Hoffman-Wielandt inequality).
Let $\mathbf{A}$ *and* $\mathbf{B}$ *be two* $N \times N$ *normal matrices and let* $\mathrm{ESD}(\mathbf{A})$ *and* $\mathrm{ESD}(\mathbf{B})$ *be their ESDs, respectively. Then,*

$$d_L\left(\mathrm{ESD}(\mathbf{A}), \mathrm{ESD}(\mathbf{B})\right)^3 \le \frac{1}{N}\operatorname{Tr}\left[(\mathbf{A}-\mathbf{B})(\mathbf{A}-\mathbf{B})^*\right]. \tag{3.8}$$

Here $\mathbf{A}^*$ *denotes the conjugate transpose of* $\mathbf{A}$*. Moreover, if* $\mathbf{A}$ *and* $\mathbf{B}$ *are two Hermitian matrices of size* $N \times N$*, then*

$$\sum_{i=1}^{N}\left(\lambda_i(\mathbf{A})-\lambda_i(\mathbf{B})\right)^2 \le \operatorname{Tr}[(\mathbf{A}-\mathbf{B})^2]. \tag{3.9}$$

The next two straightforward lemmas control the tail of the product of two Pareto random variables and the expectation of a truncated Pareto.

Lemma 3.3.2.
Let X *and* Y *be two independent Pareto r.v.'s with parameters* β_1 *and* β_2 *respectively, with* $\beta_1 \le \beta_2$*. There exist constants* $c_1 = c_1(\beta_1, \beta_2) > 0$ *and* $c_2 = c_2(\beta_1) > 0$ *such that*

$$\mathbf{P}(XY > t) = \begin{cases} c_1 t^{-\beta_1} & \text{if } \beta_1 < \beta_2 \\ c_2 t^{-\beta_1}\log t & \text{if } \beta_1 = \beta_2. \end{cases}$$

Lemma 3.3.3.
Let X *be a Pareto random variable with law* $\mathbf{P}$ *and parameter* $\beta > 1$*. For any* $m > 0$ *it holds*

$$\mathbf{E}\left[X \mathbb{1}_{X \ge m}\right] = \frac{\beta}{(\beta-1)} m^{1-\beta}.$$

We state one final auxiliary lemma related to the approximation of sums by integrals.

Lemma 3.3.4.
Let $\beta \in (0, 1]$*. Then there exists a constant* $c_1 = c_1(\beta) > 0$ *such that*

$$\frac{1}{N}\sum_{i \ne j \in \mathbf{V}_N}\frac{1}{\|i-j\|^\beta} \sim c_1 \max\{N^{1-\beta}, \log N\}. \tag{3.10}$$

If instead $\beta > 1$*, there exists a constant* $c_2 > 0$ *such that*

$$\frac{1}{N}\sum_{i \ne j \in \mathbf{V}_N}\frac{1}{\|i-j\|^\beta} \sim c_2\,.$$

We end this section by quoting, for the reader's convenience, the following lemma from Chakrabarty et al. [2016, Fact 4.3].

Lemma 3.3.5.
Let (Σ, d) be a complete metric space, and let $(\Omega, \mathcal{A}, P)$ be a probability space. Suppose that $\left(X_{mn} : (m,n) \in \{1,2,\ldots,\infty\}^2 \backslash \{\infty,\infty\}\right)$ is a family of random elements in Σ, that is, measurable maps from Ω to Σ, the latter being equipped with the Borel σ-field induced by d. Assume that

(1) for all fixed $1 \leq m < \infty$

$$\lim_{n\to\infty} d\left(X_{mn}, X_{m\infty}\right) = 0 \textit{ in } P\textit{-probability.}$$

(2) For all $\varepsilon > 0$,

$$\lim_{m\to\infty} \limsup_{n\to\infty} P\left(d\left(X_{mn}, X_{\infty n}\right) > \varepsilon\right) = 0.$$

Then, there exists a random element $X_{\infty\infty}$ of Σ such that

$$\lim_{m\to\infty} d\left(X_{m\infty}, X_{\infty\infty}\right) = 0 \textit{ in } P\textit{-probability} \tag{3.11}$$

and

$$\lim_{n\to\infty} d\left(X_{\infty n}, X_{\infty\infty}\right) = 0 \textit{ in } P\textit{-probability.}$$

Furthermore, if $X_{m\infty}$ is deterministic for all m, then so is $X_{\infty\infty}$, and (3.11) simplifies to

$$\lim_{m\to\infty} d\left(X_{m\infty}, X_{\infty\infty}\right) = 0. \tag{3.12}$$

§3.4 Existence and Uniqueness

The proof of Theorem 3.2.1 is split into several parts and we will now briefly sketch them.

(1) **Truncation**: The first part of the proof is a truncation argument on the unbounded weights $(W_i)_{i\in\mathbf{V}_N}$. We construct a new sequence $(W_i^m)_{i\in\mathbf{V}_N}$ that is obtained by truncating the original weights at a value $m > 1$. We construct another scaled adjacency matrix $\mathbf{A}_{N,m}$, with entries $\mathbf{A}_{N,m}(i,j)$ distributed as Bernoulli random variables with parameter p_{ij}^m given by (3.4) with the weights substituted by the truncated ones. We then show (see Lemma 3.4.2) that the empirical measure $\mathrm{ESD}(\mathbf{A}_N)$ is well approximated by $\mathrm{ESD}(\mathbf{A}_{N,m})$, that is, their Lévy distance vanishes in probability in the limit $m \to \infty$.

(2) **Gaussianisation**: In the second part, we aim to Gaussianise $\mathbf{A}_{N,m}$ using the ideas of Chatterjee [2005]. We begin with the construction of a centred matrix $\overline{\mathbf{A}}_{N,m}$, that is obtained by subtracting out the expectation from each entry of $\mathbf{A}_{N,m}$. We then Gaussianise $\overline{\mathbf{A}}_{N,m}$, that is, we pass to another matrix $\mathbf{A}_{N,g}$ with each entry $\mathbf{A}_{N,g}(i,j)$ being a normal random variable with mean 0 and the same variance $p_{ij}^m(1-p_{ij}^m)$ as the corresponding entry of $\overline{\mathbf{A}}_{N,m}$. Lastly, we tweak the variances of $\mathbf{A}_{N,g}$ to obtain a Gaussian random matrix $\tilde{\mathbf{A}}_{N,m,g}$ with entries $\tilde{\mathbf{A}}_{N,m,g}(i,j)$ having mean 0 and variance equal to r_{ij}^m, the "unbounded version" of p_{ij}^m (see (3.13)). Thanks to (3.8), we can show (Lemma 3.4.3, Lemma 3.4.4 and Lemma 3.4.6) that in this whole process we did not lose too much: the Lévy distance between the empirical measures $\mathrm{ESD}(\mathbf{A}_{N,m})$ and $\mathrm{ESD}(\tilde{\mathbf{A}}_{N,m,g})$ is small in probability. We remark here that the order of the errors in Lemmas 3.4.3 and 3.4.6 is $N^{-\alpha}$, and these steps fail for $\alpha = 0$.

(3) **Identification of the limit**: We then proceed to analyse the limit of the measure $\mathrm{ESD}(\tilde{\mathbf{A}}_{N,m,g})$ as N goes to infinity. We use Wick's formula to compute its expected moments and use a concentration argument to show the existence of a unique limiting measure

$$\mu_{\sigma,\tau,m} := \lim_{N\to\infty} \mathrm{ESD}(\tilde{\mathbf{A}}_{N,m,g})$$

using Proposition 3.4.9. We conclude the proof of Theorem 3.2.1 by letting the truncation m go to infinity: using Lemma 3.3.5 we can show that there is a unique limiting measure $\mu_{\sigma,\tau}$ such that $\mu_{\sigma,\tau} := \lim_{m\to\infty}\mu_{\sigma,\tau,m}$. In the case $\sigma = 1$ calculations become explicit.

Remark 3.4.1 (Self-loops).
We can use Proposition 3.3.1 to show that having self-loops in the model will not affect the limiting spectral distribution. Let $\mathbf{A}_N$ be the scaled adjacency matrix of the model as defined in (3.6)*. Now, consider*

$$D_N = c_N^{-1/2}\,\mathrm{Diag}(1,\dots,1)$$

to be the $N\times N$ diagonal matrix with all diagonal entries "1", scaled by a factor of $\sqrt{c_N}$, and $\mathbf{A}_{N,SL} = \mathbf{A}_N + D_N$. If we extend the definition of p_{ij} for the case $i = j$ as $p_{ii} = 1$, then $\mathbf{A}_{N,SL}$ will be the scaled adjacency of the random graph with self-loops. Using (3.8)*, we get*

$$d_L^3(\mu_{\mathbf{A}_N},\mu_{\mathbf{A}_{N,SL}}) \le \frac{1}{N}\,\mathrm{Tr}[(\mathbf{A}_N-\mathbf{A}_{N,SL})^2] = \frac{1}{N}\,\mathrm{Tr}[D_N^2] = \frac{N}{Nc_N} = \mathrm{O}(c_N^{-1}).$$

§3.4.1 Truncation

Now we show that for our analysis the weights can be truncated. More precisely, let $m > 1$ be a truncation threshold and define $W_i^m = W_i \mathbf{1}_{W_i \leq m}$ for any $i \in \mathbf{V}_N$. For all $N \in \mathbb{N}$, we define a new random graph with vertex set $\mathbf{V}_N$ and connection probability as follows: conditional on the weights $(W_i^m)_{i \in \mathbf{V}_N}$ we connect $i, j \in \mathbf{V}_N$ with probability

$$p_{ij}^m = r_{ij}^m \wedge 1 \qquad \text{with} \quad r_{ij}^m = \frac{(W_i^m \vee W_j^m)(W_i^m \wedge W_j^m)^\sigma}{\|i - j\|^\alpha} \qquad i \neq j \in \mathbf{V}_N \,. \tag{3.13}$$

Let $\mathbf{A}_{N,m}$ be the corresponding adjacency matrix scaled by $\sqrt{c_N}$ and let its ESD be denotes by $\mathrm{ESD}(\mathbf{A}_{N,m})$.

It will be useful later to have the two following easy bounds (following from Lemma 3.3.4):

$$\sum_{i \neq j \in \mathbf{V}_N} r_{ij}^m \leq m^{1+\sigma} N c_N \,, \qquad \sum_{i \neq j \in \mathbf{V}_N} (r_{ij}^m)^t \leq c\, m^{2+2\sigma} \max\{N^{1-t\alpha}, \log N\} \,, \tag{3.14}$$

for some constant $c > 0$ and $t > 1$ a real number. The second bound is not optimal, since for some $t > 1$ such that $t\alpha > 1$, the upper bound will just be a constant depending on t and α. However, for our computations, this bound suffices.

Lemma 3.4.2 (Truncation).
For every $\delta > 0$ one has

$$\limsup_{m \to \infty} \lim_{N \to \infty} \mathbb{P}\left(d_L(\mathrm{ESD}(\mathbf{A}_N), \mathrm{ESD}(\mathbf{A}_{N,m})) > \delta\right) = 0 \,.$$

Proof. By (3.8) we have that

$$\begin{aligned}
&\mathbb{E}\left[d_L^3\left(\mathrm{ESD}(\mathbf{A}_N),\, \mathrm{ESD}(\mathbf{A}_{N,m})\right)\right] \\
&\leq \frac{1}{N c_N} \mathbb{E}\left[\mathrm{Tr}\left((\mathbf{A}_N - \mathbf{A}_{N,m})^2\right)\right] \\
&= \frac{1}{N c_N} \sum_{i \neq j \in \mathbf{V}_N} \mathbb{E}\left[(\mathbf{A}_N(i,j) - \mathbf{A}_{N,m}(i,j))^2 \mathbf{1}_{\mathbf{A}_N(i,j) \neq \mathbf{A}_{N,m}(i,j)}\right] \\
&\leq \frac{1}{N c_N} \sum_{i \neq j \in \mathbf{V}_N} \mathbb{P}\left(\mathbf{A}_N(i,j) \neq \mathbf{A}_{N,m}(i,j)\right) .
\end{aligned} \tag{3.15}$$

For fixed i, j we will analyse $\mathbb{P}\left(\mathbf{A}_N(i,j) \neq \mathbf{A}_{N,m}(i,j)\right)$ as follows. We notice that $\mathbf{A}_N(i,j) \neq \mathbf{A}_{N,m}(i,j)$ can occur only if one between W_i and W_j exceeds m. Calling

$$A = \{W_i \geq m > W_j\} \quad \text{and} \quad B = \{W_i \geq W_j \geq m\} \tag{3.16}$$

we have, by symmetry of W_i and W_j, that $\mathbb{P}\left(\mathbf{A}_N(i,j) \neq \mathbf{A}_{N,m}(i,j)\right)$ equals

$$2\mathbb{P}\left(\{\mathbf{A}_N(i,j) \neq \mathbf{A}_{N,m}(i,j)\} \cap A\right) + 2\mathbb{P}\left(\{\mathbf{A}_N(i,j) \neq \mathbf{A}_{N,m}(i,j)\} \cap B\right).$$

Notice that on the events A and B the variable $\mathbf{A}_{N,m}(i,j)$ is always 0. So we can bound

$$\begin{aligned}
&\mathbb{P}\left(\{\mathbf{A}_N(i,j) \neq \mathbf{A}_{N,m}(i,j)\} \cap A\right) \\
&= \mathbb{P}\left(\{\mathbf{A}_N(i,j) = 1\} \cap A\right) \\
&\leq \mathbf{E}\Big[\frac{\kappa_\sigma(W_i, W_j)}{\|i-j\|^\alpha} \mathbf{1}_A\Big] \leq \frac{\mathbf{E}[W_i \mathbf{1}_{W_i \geq m}]\mathbf{E}[W_j^\sigma]}{\|i-j\|^\alpha} \leq c\frac{m^{2-\tau}}{\|i-j\|^\alpha}
\end{aligned}$$

for some constant $c > 0$, where we have used Lemma 3.3.3 and the fact that $\mathbb{E}[W_j^\sigma] < \infty$. Analogously we can bound the second summand by

$$\begin{aligned}
&\mathbb{P}\left(\{\mathbf{A}_N(i,j) \neq \mathbf{A}_{N,m}(i,j)\} \cap B\right) \\
&\leq \mathbf{E}\left[\frac{W_i W_j^\sigma}{\|i-j\|^\alpha} \mathbf{1}_B\right] \leq \frac{\mathbf{E}[W_i \mathbf{1}_{W_i \geq m}]\mathbf{E}[W_j^\sigma]}{\|i-j\|^\alpha} \\
&\leq c\frac{m^{2-\tau}}{\|i-j\|^\alpha}.
\end{aligned}$$

Plugging these estimates back into (3.15) we obtain

$$\mathbb{E}\left[d_L^3\left(\mathrm{ESD}(\mathbf{A}_N), \mathrm{ESD}(\mathbf{A}_{N,m})\right)\right] \leq \frac{4c}{Nc_N} \sum_{i \neq j \in \mathbf{V}_N} \frac{m^{2-\tau}}{\|i-j\|^\alpha} = 4cm^{2-\tau}.$$

We can then conclude by applying Markov's inequality:

$$\begin{aligned}
&\limsup_{m\to\infty} \lim_{N\to\infty} \mathbb{P}\left(d_L\left(\mathrm{ESD}(\mathbf{A}_N), \mathrm{ESD}(\mathbf{A}_{N,m})\right) > \delta\right) \\
&\leq \limsup_{m\to\infty} \lim_{N\to\infty} \frac{\mathbb{E}\left[d_L^3\left(\mathrm{ESD}(\mathbf{A}_N), \mathrm{ESD}(\mathbf{A}_{N,m})\right)\right]}{\delta^3} \\
&= 0
\end{aligned}$$

since $\tau > 2$. □

§3.4.2 Centring

Let $1 < m \leq \infty$ and $\overline{\mathbf{A}}_{N,m}$ be the centred and rescaled truncated adjacency matrix, i.e. the matrix defined as

$$\overline{\mathbf{A}}_{N,m}(i,j) = \mathbf{A}_{N,m}(i,j) - E^W[\mathbf{A}_{N,m}(i,j)], \quad i \neq j \in \mathbf{V}_N. \tag{3.17}$$

Note that here $m = \infty$ corresponds to the matrix with non-truncated weights. The following lemma says that the centring does not affect the limiting spectral distribution.

Lemma 3.4.3 (Centring).
For any $m \in (1, \infty]$, under the conditions in Theorem 3.2.1, we have, for all $\delta > 0$,

$$\lim_{N\to\infty} \mathbb{P}\left(d_L\left(\mathrm{ESD}(\mathbf{A}_{N,m}),\, \mathrm{ESD}(\overline{\mathbf{A}}_{N,m})\right) > \delta\right) = 0\,,$$

where $\mathrm{ESD}(\overline{\mathbf{A}}_{N,m})$ is the empirical spectral distribution of $\overline{\mathbf{A}}_{N,m}$.

Proof. By (3.8) we have

$$\begin{aligned}
\mathbb{E}\left[d_L^3\left(\mathrm{ESD}(\mathbf{A}_{N,m}),\, \mathrm{ESD}(\overline{\mathbf{A}}_{N,m})\right)\right] &\leq \frac{1}{N}\mathbb{E}\left[\mathrm{Tr}(E^W[\mathbf{A}_{N,m}]^2)\right] \\
&= \frac{1}{Nc_N}\sum_{i\neq j\in\mathbf{V}_N} \mathbf{E}[p_{ij}^m]^2 \\
&\leq \frac{1}{Nc_N}\sum_{i\neq j\in\mathbf{V}_N} \frac{\mathbf{E}\left[(W_i\vee W_j)(W_i\wedge W_j)^\sigma\right]^2}{\|i-j\|^{2\alpha}} \\
&\leq \frac{c}{Nc_N}\max\{N^{1-2\alpha}, \log N\}. \qquad (3.18)
\end{aligned}$$

Here c is some constant as for $\tau > 2$ and $\sigma < \tau - 1$ we have

$$\mathbf{E}\left[(W_i\vee W_j)(W_i\wedge W_j)^\sigma\right] = 2\mathbf{E}\left[W_iW_j^\sigma \mathbf{1}_{W_i>W_j}\right] \leq 2\mathbf{E}[W_i]\mathbf{E}[W_j^\sigma] < \infty.$$

In the last inequality we used Lemma 3.3.4. The result follows by applying Markov's inequality. □

§3.4.3 Gaussianisation

Let $\{G_{i,j}, 1 \leq i \leq j\}$ be a family of i.i.d. standard Gaussian random variables, independent of the weights and the graph. Define a symmetric $N\times N$ matrix $\mathbf{A}_{N,m,g}$ by

$$\mathbf{A}_{N,m,g}(i,j) = \begin{cases} \frac{\sqrt{p_{ij}^m(1-p_{ij}^m)}}{\sqrt{c_N}} G_{i\wedge j, i\vee j} & \text{for } 1 \leq i \neq j \leq N \\ 0 & \text{for } i = j. \end{cases} \qquad (3.19)$$

Notice that the entries of $\mathbf{A}_{N,m,g}$ have the same mean and variance of the corresponding entries of $\overline{\mathbf{A}}_{N,m}$. Consider a three-times continuously differentiable function $h : \mathbb{R} \to \mathbb{R}$ such that

$$\max_{0\leq k\leq 3}\sup_{x\in\mathbb{R}}\left|h^{(k)}(x)\right| < \infty$$

where $h^{(k)}$ denotes the k-th derivative. For an $N\times N$ real symmetric matrix $\mathbf{M}_N$ define the *resolvent* of $\mathbf{M}_N$ as

$$R_{M_N}(z) = (\mathbf{M}_N - z\,\mathrm{I}_N)^{-1}\,, \qquad z\in\mathbb{C}^+,$$

where I_N is the $N \times N$ identity matrix. In particular, if $\mu := \mu_{\mathbf{M}_N}$ is the ESD of $\mathbf{M}_N$, the relation between the Stieltjes transform $\mathrm{S}_{\mathbf{M}_N}$ of $\mu_{\mathbf{M}_N}$ and resolvent can be expressed as

$$H(\mathbf{M}_N) := \mathrm{S}_{\mathbf{M}_N}(z) = \mathrm{tr}(R_{M_N}(z)),\ z \in \mathbb{C}^+ \tag{3.20}$$

[Bai and Silverstein, 2010, Section 1.3.2]. The next result shows that the real and imaginary parts of the Stieltjes transform of $\mu_{\overline{\mathbf{A}}_{N,m}}$ are close to those of $\mu_{\mathbf{A}_{N,m,g}}$. Since one knows that the convergence of the ESD is equivalent to showing the convergence of the corresponding Stieltjes transform, one can shift the problem to the Gaussianised setup and work with the matrix $\mathbf{A}_{N,m,g}$.

Lemma 3.4.4 (Gaussianisation).
Consider the matrix $\overline{\mathbf{A}}_{N,m}$ defined in Subsection 3.4.1 and the matrix $\mathbf{A}_{N,m,g}$ defined in (3.19). *For any three-times continuously differentiable function $h : \mathbb{R} \to \mathbb{R}$ such that*

$$\max_{0\le k\le 3} \sup_{x\in\mathbb{R}} \left| h^{(k)}(x) \right| < \infty$$

we have

$$\lim_{N\to\infty} \left| \mathbb{E}\left[h\left(\Re H\left(\mathbf{A}_{N,m,g}\right)\right)\right] - \mathbb{E}\left[h\left(\Re H\left(\overline{\mathbf{A}}_{N,m}\right)\right)\right] \right| = 0,$$
$$\lim_{N\to\infty} \left| \mathbb{E}\left[h\left(\Im H\left(\mathbf{A}_{N,m,g}\right)\right)\right] - \mathbb{E}\left[h\left(\Im H\left(\overline{\mathbf{A}}_{N,m}\right)\right)\right] \right| = 0\,,$$

where $\Re$ and $\Im$ denote the real and imaginary parts respectively and $h^{(k)}$ denotes the k-th derivative of h.

To prove the above lemma, we will need the following result from Chatterjee [2005].

Theorem 3.4.5 (Chatterjee [2005, Theorem 1.1]).
Let $\mathbf{X} = (X_1, \ldots, X_n)$ and $\mathbf{Y} = (Y_1, \ldots, Y_n)$ be two vectors of independent random variables with finite second moments, taking values in some open interval I and satisfying, for each $i, \mathbb{E}X_i = \mathbb{E}Y_i$ and $\mathbb{E}X_i^2 = \mathbb{E}Y_i^2$. Let $f : I^n \to \mathbb{R}$ be three-times differentiable in each argument. If we set $U = f(\mathbf{X})$ and $V = f(\mathbf{Y})$, then for any thrice differentiable $h : \mathbb{R} \to \mathbb{R}$ and any $K > 0$,

$$|\mathbb{E}h(U) - \mathbb{E}h(V)| \le C_1(h)\lambda_2(f) \sum_{i=1}^{n} \left[\mathbb{E}\left[X_i^2 \mathbf{1}_{|X_i|>K}\right] + \mathbb{E}\left[Y_i^2 \mathbf{1}_{|Y_i|>K}\right]\right]$$
$$+ C_2(h)\lambda_3(f) \sum_{i=1}^{n} \left[\mathbb{E}\left[|X_i|^3 \mathbf{1}_{|X_i|\le K}\right] + \mathbb{E}\left[|Y_i|^3 \mathbf{1}_{|Y_i|\le K}\right]\right]$$

where $C_1(h) = \|h'\|_\infty + \|h''\|_\infty$, $C_2(h) = \frac{1}{6}\|h'\|_\infty + \frac{1}{2}\|h''\|_\infty + \frac{1}{6}\|h'''\|_\infty$ and

$$\lambda_s(f) \coloneqq \sup\left\{|\partial_i^q f(x)|^{\frac{s}{q}} : 1 \le i \le n, 1 \le q \le s, x \in I^n\right\},$$

where ∂_i^q denotes q-fold differentiation with respect to the i-th coordinate.

Proof of Lemma 3.4.4. We prove this for the real part of the Stieltjes transform. The bounds for the imaginary part remain the same. We fix a complex number $z \in \mathbb{C}^+$, given by $z = \Re(z) + i\eta$ with $\eta > 0$.

Let $n = N(N-1)/2$ and $\mathbf{x} = (x_{ij})_{1 \le i < j \le N} \in \mathbb{R}^n$. Define $\mathrm{R}(\mathbf{x})$ to be the matrix-valued differentiable function given by

$$\mathrm{R}(\mathbf{x}) \coloneqq (\mathbf{M}_N(\mathbf{x}) - z\,\mathrm{I}_N)^{-1},$$

where $\mathbf{M}_N(\cdot)$ is the matrix-valued differentiable function that maps a vector in $\mathbb{R}^n$ to the space of $N \times N$ Hermitian matrices, given by

$$\mathbf{M}_N(\mathbf{x})_{ij} = \begin{cases} c_N^{-1/2} x_{ij} & \text{if } i < j, \\ c_N^{-1/2} x_{ji} & \text{if } i > j, \\ 0 & \text{if } i = j. \end{cases}$$

Since $\mathbf{M}_N$ is symmetric, it has all real eigenvalues. The function $H(\mathbf{M}_N(\mathbf{x}))$ admits partial derivatives of all orders. In particular, we denote for any $\mathbf{u} \in \{(i,j)\}_{1 \le j < i \le n}$ the partial derivative as $\partial H / \partial x_{\mathbf{u}}$. For any $\mathbf{u} \in \{(i,j)\}_{1 \le j < i \le n}$, using the identity $(\mathbf{M}_N(\mathbf{x}) - z\,\mathrm{I})\mathrm{R}(\mathbf{x}) = \mathrm{I}_N$ we have

$$\frac{\partial \mathrm{R}(\mathbf{x})}{\partial x_{\mathbf{u}}} = -\mathrm{R}(\mathbf{x})(\partial_{\mathbf{u}} \mathbf{M}_N)\mathrm{R}(\mathbf{x}).$$

By iterative application of derivatives, three identities were derived in Chatterjee [2005]:

$$\begin{aligned}
\frac{\partial H}{\partial x_{\mathbf{u}}} &= -\frac{1}{N}\operatorname{Tr}\left(\frac{\partial \mathbf{M}_N(\mathbf{x})}{\partial x_{\mathbf{u}}}\mathrm{R}(\mathbf{x})^2\right), \\
\frac{\partial^2 H}{\partial x_{\mathbf{u}}^2} &= \frac{2}{N}\operatorname{Tr}\left(\frac{\partial \mathbf{M}_N(\mathbf{x})}{\partial x_{\mathbf{u}}}\mathrm{R}(\mathbf{x})\frac{\partial \mathbf{M}_N(\mathbf{x})}{\partial x_{\mathbf{u}}}\mathrm{R}(\mathbf{x})^2\right), \\
\frac{\partial^3 H}{\partial x_{\mathbf{u}}^3} &= -\frac{6}{N}\operatorname{Tr}\left(\frac{\partial \mathbf{M}_N(\mathbf{x})}{\partial x_{\mathbf{u}}}\mathrm{R}(\mathbf{x})\frac{\partial \mathbf{M}_N(\mathbf{x})}{\partial x_{\mathbf{u}}}\mathrm{R}(\mathbf{x})\frac{\partial \mathbf{M}_N(\mathbf{x})}{\partial x_{\mathbf{u}}}\mathrm{R}(\mathbf{x})^2\right).
\end{aligned}$$

Note that $\partial_{ij}\mathbf{M}_N(\mathbf{x})$ is a matrix with $c_N^{-1/2}$ at the $(i,j)^{\text{th}}$ and $(j,i)^{\text{th}}$ entry, and 0 everywhere else. Using the bounds on Hilbert-Schmidt norms and following the exact argument regarding the bounds in equations (4), (5) and (6) in Chatterjee [2005] we get that

$$\left\|\frac{\partial H}{\partial x_{\mathbf{u}}}\right\|_\infty \le \frac{2}{\eta N \sqrt{c_N}}, \quad \left\|\frac{\partial^2 H}{\partial x_{\mathbf{u}}^2}\right\|_\infty \le \frac{4}{\eta^3 N c_N}, \quad \left\|\frac{\partial^3 H}{\partial x_{\mathbf{u}}^3}\right\|_\infty \le \frac{12}{\eta^4 N c_N^{3/2}}.$$

Hence

$$\lambda_2(H) \leq 4 \max\left\{\frac{1}{\eta^4}, \frac{1}{\eta^3}\right\} \frac{1}{Nc_N}$$

and

$$\lambda_3(H) \leq 12 \max\left\{\frac{1}{\eta^6}, \frac{1}{\eta^{9/2}}, \frac{1}{\eta^4}\right\} \frac{1}{Nc_N^{3/2}}.$$

Conditional on the weights $(W_i)_{i\geq 1}$, consider the following sequence of independent random variables. Let $\mathbf{X}_b = (X^b_{ij})_{1\leq i<j\leq N}$ be a vector with $X^b_{ij} \sim \mathrm{Ber}(p^m_{ij}) - p^m_{ij}$. Similarly, take another vector $\mathbf{X}_g = (X^g_{ij})_{1\leq i<j\leq N}$ with $X^g_{ij} \sim \mathcal{N}\left(0, p^m_{ij}(1-p^m_{ij})\right)$. Then,

$$\overline{\mathbf{A}}_{N,m} = \mathbf{M}_N(\mathbf{X}_b) \text{ and } \mathbf{A}_{N,g} = \mathbf{M}_N(\mathbf{X}_g)$$

in law. We have that

$$\begin{aligned}
&\left|\mathbb{E}\left[h\left(\Re H_z\left(\mathbf{A}_{N,m,g}\right)\right) - h\left(\Re H_z\left(\overline{\mathbf{A}}_{N,m}\right)\right)\right]\right| \\
&= \left|\mathbf{E}\left[E^W\left[h\left(\Re H_z\left(\mathbf{A}_{N,m,g}\right)\right) - h\left(\Re H_z\left(\overline{\mathbf{A}}_{N,m}\right)\right)\right]\right]\right|.
\end{aligned}$$

Conditionally on the weights, the sequences $\mathbf{X}_g$ and $\mathbf{X}_b$ form two vectors of independent random variables, with $E^W[X^b_{ij}] = E^W[X^g_{ij}]$ and $E^W[(X^b_{ij})^2] = E^W[(X^g_{ij})^2]$. Then, using Theorem 3.4.5 on the conditional expectation

$$E^W[h\left(\Re H_z\left(\mathbf{A}_{N,m,g}\right)\right) - h\left(\Re H_z\left(\overline{\mathbf{A}}_{N,m}\right)\right)],$$

we have that

$$\begin{aligned}
&\left|\mathbf{E}\left[E^W[h\left(\Re H_z\left(\mathbf{A}_{N,m,g}\right)\right) - h\left(\Re H_z\left(\overline{\mathbf{A}}_{N,m}\right)\right)]\right]\right| \\
&\leq C_1(h)\lambda_2(H) \sum_{1\leq i<j\leq N} \mathbb{E}[(X^b_{ij})^2 \mathbf{1}_{|X^b_{ij}|>K_N}] + \mathbb{E}[(X^g_{ij})^2 \mathbf{1}_{|X^g_{ij}|>K_N}] \qquad (3.21)\\
&+ C_2(h)\lambda_3(H) \sum_{1\leq i<j\leq N} \mathbb{E}[(X^b_{ij})^3 \mathbf{1}_{|X^b_{ij}|\leq K_N}] + \mathbb{E}[(X^g_{ij})^3 \mathbf{1}_{|X^g_{ij}|\leq K_N}], \qquad (3.22)
\end{aligned}$$

where K_N is a (possibly) $N-$dependent truncation and where we have used that $|\partial^p_{\mathbf{u}} \Re H| = |\Re \partial^p_{\mathbf{u}} H| \leq |\partial^p_{\mathbf{u}} H|$. Now using the fact that $r/p > 0$ we have $|\partial^p_{\mathbf{u}} \Re H|^{\frac{r}{p}} \leq |\partial^p_{\mathbf{u}} H|^{\frac{r}{p}}$, and therefore

$$\lambda_r(\Re H) \leq \lambda_r(H).$$

We begin by evaluating (3.21). To compute the Bernoulli term, notice that X^b_{ij} are uniformly bounded by 1, so, for any $K_N > 1$, we automatically have that

$$\sum_{1\leq i<j\leq N} \mathbb{E}[(X^b_{ij})^2 \mathbf{1}_{|X^b_{ij}|>K_N}] = 0.$$

For the Gaussian term, we apply the Cauchy-Schwarz inequality (with respect to $\mathbb{E}$). Using also the trivial bound $p_{ij}^m \le r_{ij}^m$ and Markov's inequality, we obtain

$$\begin{aligned}\sum_{1\le i<j\le N} \mathbb{E}[(X_{ij}^g)^2 \mathbf{1}_{|X_{ij}^g|>K_N}] &\le \sum_{1\le i<j\le N} \mathbb{E}[(X_{ij}^g)^4]^{1/2}\mathbb{P}(|X_{ij}^g|>K_N)^{1/2} \\ &\le 3\sum_{1\le i<j\le N} \mathbf{E}[(r_{ij}^m)^2]^{1/2}\,\frac{\mathbb{E}[(X_{ij}^g)^2]^{1/2}}{K_N} \le 3\sum_{1\le i<j\le N} \mathbf{E}[(r_{ij}^m)^2]^{1/2}\,\frac{\mathbf{E}[r_{ij}^m]^{1/2}}{K_N} \\ &\overset{(3.14)}{=} \mathrm{O}_N(N\cdot K_N^{-1}\max\{N^{1-3\alpha/2},\,\log N\}).\end{aligned}$$

We thus conclude that (3.21) is of order

$$(3.21) = \mathrm{O}_N(c_N^{-1}K_N^{-1}\max\{N^{1-3\alpha/2},\,\log N\}).$$

For (3.22), we use that for any random variable X we have the bound

$$\mathbb{E}[|X|^3\mathbf{1}_{|X|\le K}] \le K\mathbb{E}[X^2]\,.$$

Hence we can bound

$$\begin{aligned}&\sum_{1\le i<j\le N} \mathbb{E}[(X_{ij}^b)^3\mathbf{1}_{|X_{ij}^b|\le K_N} + (X_{ij}^g)^3\mathbf{1}_{|X_{ij}^g|\le K_N}] \\ &\le K_N \sum_{1\le i<j\le N} \mathbb{E}[(X_{ij}^b)^2 + (X_{ij}^g)^2] \\ &\le 2K_N \sum_{1\le i<j\le N} \mathbf{E}[r_{ij}^m] \overset{(3.14)}{=} \mathrm{O}_N(K_N N c_N)\,.\end{aligned}$$

This yields that (3.22) is of order $\mathrm{O}_N(K_N c_N^{-1/2})$. Choosing $K_N = \mathrm{O}_N 1$ gives us that

$$\left|\mathbb{E}\left[h\left(\Re H\left(\mathbf{A}_{N,m,g}\right)\right)\right] - \mathbb{E}\left[h\left(\Re H\left(\overline{\mathbf{A}}_{N,m}\right)\right)\right]\right| = o_N(1)\,. \tag{3.23}$$

A similar argument holds for the imaginary part $\Im(H)$ and this completes the proof. □

Simplification of the variance structure

To conclude Gaussianisation, we would like to construct a final matrix $\tilde{\mathbf{A}}_{N,m,g}$ with a simpler variance structure than that of $\mathbf{A}_{N,m,g}$. We let its entries be

$$\tilde{\mathbf{A}}_{N,m,g}(i,j) = \frac{\sqrt{r_{ij}^m}}{\sqrt{c_N}} G_{i\wedge j, i\vee j} \quad 1\le i,j\le N \tag{3.24}$$

where r_{ij}^m is as in (3.13) and the $\{G_{i,j} : i \ge j\}$ are the i.i.d. collection of Gaussian variables used in (3.19). We need to prove that the ESD of this matrix gives asymptotically a good approximation of the ESD of $\mathbf{A}_{N,m,g}$.

Lemma 3.4.6 (Simplification of variance).
For any $\delta > 0$

$$\lim_{N\to\infty} \mathbb{P}\left(d_L(\mathrm{ESD}(\mathbf{A}_{N,m,g}), \mathrm{ESD}(\tilde{\mathbf{A}}_{N,m,g})) > \delta\right) = 0\,.$$

Proof. Construct a matrix $L_{N,g}$ with entries

$$L_{N,g}(i,j) = \begin{cases} \frac{\sqrt{p^m_{ij}}}{\sqrt{c_N}} G_{i\wedge j, i\vee j} & 1 \le i \ne j \le N \\ 0 & 1 \le i = j \le N \end{cases}$$

where $p^m_{ij} = r^m_{ij} \wedge 1$. By (3.8), we have that

$$\begin{aligned}
\mathbb{E}[d^3_L(\mathrm{ESD}(\mathbf{A}_{N,m,g}), \mathrm{ESD}(L_{N,g}))] &\le \frac{1}{Nc_N} \sum_{i\ne j\in \mathbf{V}_N} \mathbb{E}\left[G^2_{i,j} p^m_{ij}\left(\sqrt{1-p^m_{ij}} - 1\right)^2\right] \\
&\le \frac{1}{Nc_N} \sum_{i\ne j\in \mathbf{V}_N} \mathbf{E}[p^m_{ij}|(1-p^m_{ij}) - 1|] \\
&\le \frac{1}{Nc_N} \sum_{i\ne j\in \mathbf{V}_N} \mathbf{E}[(r^m_{ij})^2] \overset{(3.14)}{=} o_N(1).
\end{aligned}$$

For $i \ne j \in \mathbf{V}_N$ define the events $\mathcal{A}_{ij} = \{r^m_{ij} \le 1\}$. Construct yet another matrix $\tilde{L}_{N,g}$ as

$$\tilde{L}_{N,g}(i,j) = L_{N,g}(i,j)\mathbf{1}_{\mathcal{A}_{ij}} + \frac{X_{ij}}{\sqrt{c_N}}\mathbf{1}_{\mathcal{A}^c_{ij}}$$

where, conditional on the weights, $X_{ij} \sim \mathcal{N}\left(0, r^m_{ij}\right)$ are mutually independent and independent of the $\{G_{i,j}\}_{i>j}$. It is easy to see that $\tilde{L}_{N,g} = \tilde{\mathbf{A}}_{N,m,g}$ in distribution. So, comparing $L_{N,g}$ with $\tilde{L}_{N,g}$, using (3.8) we get

$$\begin{aligned}
&\mathbb{E}[d^3_L(\mathrm{ESD}(\tilde{L}_{N,g}), \mathrm{ESD}(L_{N,g}))] \le \frac{1}{N} \sum_{i\ne j\in \mathbf{V}_N} \mathbb{E}[(L_{N,g}(i,j) - \tilde{L}_{N,g}(i,j))^2] \\
&= \frac{1}{N} \sum_{i\ne j\in \mathbf{V}_N}^{N} \mathbb{E}[(L_{N,g}(i,j) - \tilde{L}_{N,g}(i,j))^2 \mathbf{1}_{\mathcal{A}^c_{ij}}] \\
&= \frac{1}{N} \sum_{i\ne j\in \mathbf{V}_N}^{N} \mathbb{E}\left[\left(\frac{\sqrt{p^m_{ij}}}{\sqrt{c_N}} G_{i\wedge j, i\vee j} - \frac{X_{ij}}{\sqrt{c_N}}\right)^2 \mathbf{1}_{\mathcal{A}^c_{ij}}\right].
\end{aligned}$$

Using that the $G_{i,j}$ are centred and independent of the weights, and the Cauchy-Schwarz inequality, we can develop the square to obtain a further upper bound of the form

$$\begin{aligned}
&\frac{1}{Nc_N}\sum_{i\neq j\in \mathbf{V}_N}^{N} \mathbf{E}[G^2_{i\wedge j, i\vee j}\mathbf{1}_{\mathcal{A}^c_{ij}}]+\mathbf{E}[X^2_{ij}\mathbf{1}_{\mathcal{A}^c_{ij}}]\\
&\leq \frac{1}{Nc_N}\sum_{i\neq j\in \mathbf{V}_N}^{N} \mathbf{P}(\mathcal{A}^c_{ij})+\mathbf{E}[X^4_{ij}]^{1/2}\mathbf{P}(\mathcal{A}^c_{ij})^{1/2}\\
&\leq \frac{1}{Nc_N}\sum_{i\neq j\in \mathbf{V}_N}^{N} \mathbf{P}(\mathcal{A}^c_{ij})+\frac{3\mathbf{E}[(W_i^m\vee W_j^m)^2(W_i^m\wedge W_j^m)^{2\sigma}]^{1/2}}{\|i-j\|^\alpha}\mathbf{P}(\mathcal{A}^c_{ij})^{1/2}\\
&= o_N(1)
\end{aligned}$$

since

$$\mathbf{P}(\mathcal{A}^c_{ij})\leq \mathbf{P}\left(W_iW_j^\sigma\geq \|i-j\|^\alpha\right)\leq \frac{c}{\|i-j\|^{\alpha\left((\tau-1)\wedge\frac{\tau-1}{\sigma}\right)}}.$$

Using the triangle inequality, we get

$$\mathbb{E}[d_L^3(\mathrm{ESD}(\mathbf{A}_{N,m,g}),\mathrm{ESD}(\tilde{\mathbf{A}}_{N,m,g}))]=o_N(1)\,.$$

We conclude the proof using Markov's inequality. □

§3.4.4 Moment method

Preliminary results: combinatorial setup

We will recall here the combinatorics features of partitions we need in the chapter, and refer the reader for a detailed exposition to Nica and Speicher [2006, Chapter 9].

For $k\geq 1$, denote by $\mathcal{P}(2k)$ the set of partitions of $[2k]$, and by $NC(2k):=NC([2k])$ the set of non-crossing partitions of $\{1,2,\dots,2k\}$. When we write a partition, we order its blocks in such a way that the first block always contains 1, and the $(i+1)$th block contains the smallest element not belonging to any of the previous i blocks.

In what follows, we shall use Wick's formula. Let $(X_1,\dots,X_n)$ be a real Gaussian vector, then

$$\mathbb{E}[X_{i_1}\cdots X_{i_k}]=\sum_{\pi\in\mathcal{P}_2(2k)}\prod_{(r,s)\in\pi}\mathbb{E}[X_{i_r}X_{i_s}], \tag{3.25}$$

where $\mathcal{P}_2(2k)$ denotes the pair partitions of $[2k]$.

Any partition $\pi\in\mathcal{P}(k)$ can be realised as a *permutation* of $[k]$, that is, a bijective mapping $[k]\to[k]$. Let S_k denote the set of permutations on k

elements. Let $\gamma = (1, 2, \ldots, k) \in S_k$ be the shift by 1 modulo k. We will be interested in the composition of two permutations γ and π, denoted by $\gamma\pi$, which will be seen below as a partition.

As an example, consider $\pi = \{\{1, 2\}, \{3, 4\}\}$ and $\gamma = (1, 2, 3, 4)$. To compute $\gamma\pi$, we read π as $(1, 2)(3, 4)$, and compute $\gamma\pi = (1, 3)(2)(4)$. We finally read $\gamma\pi$ as $\{\{1, 3\}, \{2\}, \{4\}\}$. We now define a graph associated to a partition, borrowing the definition from Avena et al. [2023, Definition 2.3].

Definition 3.4.7 (Graph associated to a partition).
For a fixed $k \geq 1$, let γ denote the cyclic permutation $(1, 2, \ldots, k)$. For a partition π, we define $G_{\gamma\pi} = (\mathbf{V}_{\gamma\pi}, E_{\gamma\pi})$ as a rooted, labelled directed graph associated with any partition π of $[k]$, constructed as follows.

- *Initially consider the vertex set $\mathbf{V}_{\gamma\pi} = [k]$ and perform a closed walk on $[k]$ as $1 \to 2 \to 3 \to \cdots \to k \to 1$ and with each step of the walk, add an edge.*
- *Evaluate $\gamma\pi$, which will be of the form $\gamma\pi = \{V_1, V_2, \ldots, V_m\}$ for some $m \geq 1$ where $\{V_i\}_{1 \leq i \leq m}$ are disjoint blocks. Then, collapse vertices in $\mathbf{V}_{\gamma\pi}$ to a single vertex if they belong to the same block in $\gamma\pi$, and collapse the corresponding edges. Thus, $\mathbf{V}_{\gamma\pi} = \{V_1, \ldots, V_m\}$.*
- *Finally root and label the graph as follows.*
  - Root: *we always assume that the first element of the closed walk (in this case '1') is in V_1, and we fix the block V_1 as the root.*
  - Label: *each vertex V_i gets labelled with the elements belonging to the corresponding block in $\gamma\pi$.*

For the partitions $\pi = \{\{1, 2\}, \{3, 4\}\}$, $\gamma\pi = \{\{1, 3\}, \{2\}, \{4\}\}$, Figure 3.5 illustrates this procedure.

The following lemma is an exercise in Nica and Speicher [2006, Exercise 22.15] and explains also why non-crossing pair partitions will have the dominant role in the computations that follow. We will denote as $NC_2(2k)$ the set of non-crossing pair partitions of $[2k]$. For a partition π we let $\#\pi$ the number of its blocks.

Lemma 3.4.8.
Given $\pi \in \mathcal{P}_2(2k)$, one has $\#\gamma\pi \leq k + 1$ and the equality holds if and only $\pi \in NC_2(2k)$. If $\pi \in NC_2(2k)$, the graph $G_{\gamma\pi}$ is a rooted tree.

Finally, given $\pi \in NC_2(2k)$, we define the map $\mathcal{T} = \mathcal{T}_\pi : [2k] \to [k + 1]$ as follows. By Lemma 3.4.8, we know that $\#\gamma\pi = k + 1$ and let $\gamma\pi = \{V_1, V_2, \ldots, V_{k+1}\}$. Define

$$\mathcal{T}_\pi(i) = j \quad \text{if} \quad i \in V_j. \tag{3.26}$$

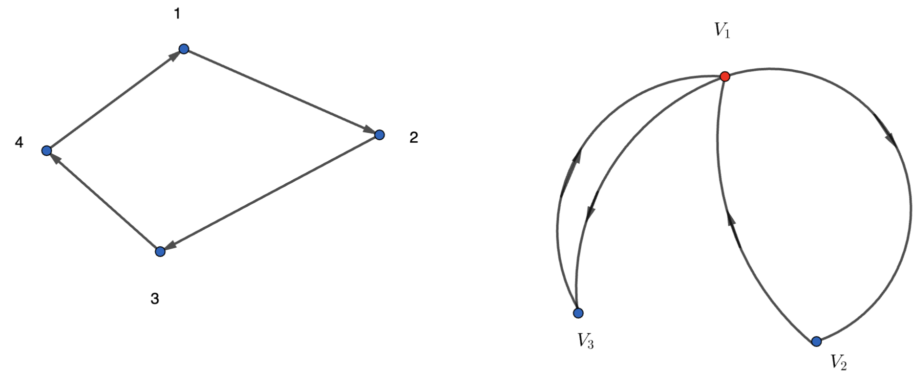


Figure 3.5: Left: closed walk on $[4]$. Right: graph associated to $\gamma\pi = \{\{1,3\},\{2\},\{4\}\}$. The root is in red.

Moment characterisation

We are now ready to give the proofs on Gaussianisation leading to the main result of this subsection, the proof of Theorem 3.2.1.

Proposition 3.4.9.
Let $\tilde{\mathbf{A}}_{N,m,g}$ be defined as in (3.24). Let $\mathrm{ESD}(\tilde{\mathbf{A}}_{N,m,g})$ be its empirical spectral distribution. Then, for $k \in \mathbb{N}$, one has

$$\lim_{N\to\infty} \mathbb{E}\left[\int_{\mathbb{R}} x^{2k}\, \mathrm{ESD}(\tilde{\mathbf{A}}_{N,m,g})(\mathrm{d}\, x)\right] = M_{2k} \tag{3.27}$$

and odd moments are zero. Moreover,

$$\lim_{N\to\infty} \mathrm{Var}\left(\int_{\mathbb{R}} x^{2k}\, \mathrm{ESD}(\tilde{\mathbf{A}}_{N,m,g})(\mathrm{d}\, x)\right) = 0, \tag{3.28}$$

where

$$M_{2k} = \sum_{\pi \in NC_2(2k)} \mathbf{E}\left[\prod_{(u,v)\in E(G_{\gamma\pi})} \kappa_\sigma(W_u^m, W_v^m)\right] < \infty, \tag{3.29}$$

where κ_σ is as in (3.3) and $E(G_{\gamma\pi})$ is the edge set of the tree $G_{\gamma\pi}$. Moreover, there exists a unique compactly supported symmetric and deterministic measure $\mu_{\sigma,\tau,m}$ characterised by the moment sequence $\{M_{2k}\}_{k\in\mathbb{N}}$ such that

$$\lim_{N\to\infty} \mathrm{ESD}(\tilde{\mathbf{A}}_{N,m,g}) = \mu_{\sigma,\tau,m} \quad \textit{in } \mathbb{P}\textit{-probability.} \tag{3.30}$$

Proof. Let $\{G_{i,j} : 1 \leq i < j \leq N\}$ be a sequence of standard independent centred Gaussian random variables as in (3.24) which is also independent of $(W_i)_{i\in[N]}$. Let $\mathcal{G}$ be the matrix

$$\mathcal{G}(i,j) = \begin{cases} \|i-j\|^{-\alpha/2} G_{i\wedge j, i\vee j} & i \neq j \\ 0 & i = j \end{cases} \tag{3.31}$$

Observe that

$$\tilde{\mathbf{A}}_{N,m,g} \stackrel{d}{=} \Upsilon_{\sigma,m} \circ \mathcal{G},$$

where $\Upsilon_{\sigma,m}$ is the matrix with elements

$$\Upsilon_{\sigma,m}(i,j) = \sqrt{\frac{\kappa_\sigma(W_i^m, W_j^m)}{c_N}}$$

and $\circ$ denotes the Hadamard product. Using Wick's formula (3.25) we have

$$\begin{aligned}
\mathbb{E}\left[\operatorname{tr}\left(\tilde{\mathbf{A}}^{2k}_{N,m,g}\right)\right] &= \frac{1}{Nc_N^k} \sideset{}{'}\sum_{1\leq i_1,\ldots,i_{2k}\leq N} \mathbb{E}\left[\prod_{\ell=1}^{2k} \Upsilon_{\sigma,m}(i_\ell, i_{\ell+1}) \prod_{\ell=1}^{2k} \mathcal{G}(i_\ell, i_{\ell+1})\right] \\
&= \frac{1}{Nc_N^k} \sideset{}{'}\sum_{1\leq i_1,\ldots,i_{2k}\leq N} \mathbb{E}\left[\prod_{\ell=1}^{2k} \kappa_\sigma^{1/2}(W_{i_\ell}^m, W_{i_{\ell+1}}^m)\right] \\
&\qquad \times \sum_{\pi\in\mathcal{P}_2(2k)} \prod_{(r,s)\in\pi} \mathbb{E}\left[\mathcal{G}(i_r, i_{r+1})\mathcal{G}(i_s, i_{s+1})\right] \\
&= \frac{1}{Nc_N^k} \sideset{}{'}\sum_{1\leq i_1,\ldots,i_{2k}\leq N} \mathbb{E}\left[\prod_{\ell=1}^{2k} \kappa_\sigma^{1/2}(W_{i_\ell}^m, W_{i_{\ell+1}}^m)\right] \\
&\qquad \times \sum_{\pi\in\mathcal{P}_2(2k)} \prod_{(r,s)\in\pi} \frac{1}{\|i_r - i_{r+1}\|^\alpha} \mathbf{1}_{\{i_r, i_{r+1}\} = \{i_s, i_{s+1}\}},
\end{aligned} \tag{3.32}$$

where we set $i_{2k+1} = i_1$ to ease notation, and $(r,s) \in \pi$ means $\pi(r) = s$ and $\pi(s) = r$. Here the $\sum'$ indicates the sum over all the indices $(i_1, \ldots, i_{2k})$ such that $i_\ell \neq i_{\ell+1}$ for $\ell \in [2k]$. The condition $\{i_r, i_{r+1}\} = \{i_s, i_{s+1}\}$ is satisfied in two cases:

C1) $i_r = i_{s+1}$ and $i_s = i_{r+1}$, that is, $i_r = i_{\gamma\pi(r)}$ and $i_s = i_{\gamma\pi(s)}$, or

C2) $i_r = i_s$ and $i_{r+1} = i_{s+1}$, that is, $i_r = i_{\pi(r)}$ and $i_{r+1} = i_{\pi(r)+1}$.

As we are going to show, the limit of (3.32) will be supported on permutations $\pi \in NC_2(2k)$ and such that Case 1) is true for all $(r,s) \in \pi$. To prove this, let us define

$$\operatorname{Cat}_{\pi,k} = \{\mathbf{i} = (i_1, \ldots, i_{2k}) \in [N]^{2k} : i_r \neq i_{r+1},\, i_r = i_{\gamma\pi(r)}\ \forall\, r \in [2k]\}.$$

When the condition $i_r = i_{\gamma\pi(r)}$ holds for all r, we see that $\mathbf{i}$ is constant on the blocks of $\gamma\pi$. We construct a graph $G(\mathbf{i})$ associated to $\mathbf{i} \in \mathrm{Cat}_{\pi,k}$ by performing a closed walk $i_1 \to i_2 \to \dots i_{2k} \to i_1$, and then collapsing elements i_r, i_s into the same vertex if r, s belong to the same block in $\gamma\pi$. We then collapse multiple edges. After this, we see that $G(\mathbf{i}) = G_{\gamma\pi}$. Thus, when we sum over $\mathbf{i} \in \mathrm{Cat}_{\pi,k}$, the count is over $\#\gamma\pi$ many indices.

We split the summation in (3.32) into two parts: a first sum over the non-crossing pairings and $\mathbf{i} \in \mathrm{Cat}_{\pi,k}$, and a second part with all the other terms, that we call $\mathcal{R}_1$. Since we take $\mathbf{i} \in \mathrm{Cat}_{\pi,k}$, $\mathbf{i}$ is constant on the blocks of $\gamma\pi$. Using this property, we obtain

$$
\begin{aligned}
&\mathbb{E}\left[\mathrm{tr}\left(\tilde{\mathbf{A}}^{2k}_{N,m,g}\right)\right] \\
&= \sum_{\pi \in NC_2(2k)} \frac{1}{N c_N^k} \sum_{\mathbf{i} \in \mathrm{Cat}_{\pi,k}} \mathbb{E}\left[\prod_{j=1}^{2k} \kappa_\sigma^{1/2}(W^m_{i_j}, W^m_{i_{j+1}})\right] \prod_{(r,s)\in\pi} \frac{1}{\|i_r - i_{r+1}\|^\alpha} + \mathcal{R}_1 \\
&= \sum_{\pi \in NC_2(2k)} \frac{1}{N c_N^k} \sum_{\mathbf{i} \in \mathrm{Cat}_{\pi,k}} \mathbb{E}\left[\prod_{(u,v)\in E(G_{\gamma\pi})} \kappa_\sigma(W^m_u, W^m_v)\right] \prod_{(r,s)\in\pi} \frac{1}{\|i_r - i_{r+1}\|^\alpha} + \mathcal{R}_1
\end{aligned}
$$

where in the last line we have used that $\mathbf{i}$ is constant on the blocks of $\gamma\pi$. Since the inner expectation no longer depends on $\mathbf{i}$, we get that

$$
\begin{aligned}
&\mathbb{E}\left[\mathrm{tr}\left(\tilde{\mathbf{A}}^{2k}_{N,m,g}\right)\right] \\
&= \sum_{\pi \in NC_2(2k)} \mathbb{E}\left[\prod_{(u,v)\in E(G_{\gamma\pi})} \kappa_\sigma(W^m_u, W^m_v)\right] \frac{1}{N c_N^k} \sum_{\mathbf{i} \in \mathrm{Cat}_{\pi,k}} \prod_{(r,s)\in\pi} \frac{1}{\|i_r - i_{r+1}\|^\alpha} \\
&+ \mathcal{R}_1.
\end{aligned}
$$

Now we make the following two claims which will finish the proof.

Claim 3.4.10.
The following hold.

a) For any $\pi \in NC_2(2k)$,

$$\lim_{N\to\infty} \frac{1}{N c_N^k} \sum_{\mathbf{i} \in \mathrm{Cat}_{\pi,k}} \prod_{(r,s)\in\pi} \frac{1}{\|i_r - i_{r+1}\|^\alpha} = 1.$$

b) We have that $\lim_{N\to\infty} \mathcal{R}_1 = 0$.

With the above claim, whose proof is deferred to page 131, we have that (3.27) holds. Moreover, the odd moments are identically 0, since there are no non-crossing pair partitions for tuples of the form $\{1, 2, \dots, 2k+1\}, k \in \mathbb{N}$. We now need to now show that (3.28) holds.

We introduce some new notation to prove (3.28). Let $\mathbf{j} = (j_1, \dots, j_{2k})$. Let $P(\mathbf{i})$ denote the expectation

$$P(\mathbf{i}) \overset{(3.31)}{:=} \mathbb{E}\left[\prod_{\ell=1}^{2k} \kappa_\sigma^{1/2}(W_{i_\ell}^m, W_{i_{\ell+1}}^m)\mathcal{G}(i_\ell, i_{\ell+1})\right],$$

and $P(\mathbf{i}, \mathbf{j})$ be

$$P(\mathbf{i}, \mathbf{j}) := \mathbb{E}\left[\prod_{\ell=1}^{2k} \kappa_\sigma^{1/2}(W_{i_\ell}^m, W_{i_{\ell+1}}^m)\mathcal{G}(i_\ell, i_{\ell+1}) \prod_{p=1}^{2k} \kappa_\sigma^{1/2}(W_{i_p}^m, W_{i_{p+1}}^m)\mathcal{G}(i_p, i_{p+1})\right]$$

(with the usual cyclic convention that $2k+1$ equals 1 for subscripts of indices). We can then see that

$$\operatorname{Var}\left(\int_{\mathbb{R}} x^{2k}\, \mathrm{ESD}(\tilde{\mathbf{A}}_{N,m,g})(\mathrm{d}\, x)\right) = \frac{1}{N^2 c_N^{2k}} \sum_{\mathbf{i},\mathbf{j}:[2k]\to[N]} \left[P(\mathbf{i}, \mathbf{j}) - P(\mathbf{i})P(\mathbf{j})\right]. \tag{3.33}$$

Note that if the terms involving $\mathbf{i}$ and $\mathbf{j}$ are completely different, that is, if the product of the terms $\mathcal{G}(i_1, i_2)\cdots\mathcal{G}(i_{2k}, i_1)$ is independent of $\mathcal{G}(j_1, j_2)\cdots\mathcal{G}(j_{2k}, j_1)$, then $P(\mathbf{i}, \mathbf{j}) = P(\mathbf{i})P(\mathbf{j})$, and (3.33) becomes identically 0. Hence, we have

$$\operatorname{Var}\left(\int_{\mathbb{R}} x^{2k} \mu_{\tilde{\mathbf{A}}_{N,m,g}}(\mathrm{d}\, x)\right) = \frac{1}{N^2 c_N^{2k}} \sum_{\mathbf{i},\mathbf{j}:[2k]\to[N]}^{(\geq 1)} P(\mathbf{i}, \mathbf{j}), \tag{3.34}$$

where $\sum^{(\geq 1)}$ is over $\mathbf{i}, \mathbf{j}$ such that there is *at least* one matching of the form $\tilde{\mathbf{A}}_{N,m,g}(i_r, i_{r+1}) = \tilde{\mathbf{A}}_{N,m,g}(j_s, j_{s+1})$ for some $1 \leq r,\ s \leq 2k-1$. If there is only one entry of $\mathbf{i}$, say i_1, equal to only one entry of $\mathbf{j}$, say j_1, then we still have

$$E^W\left[\prod_{\ell=1}^{2k} \mathcal{G}(i_\ell, i_{\ell+1})\mathcal{G}(j_\ell, j_{\ell+1})\right] = 0$$

since all entries $\mathcal{G}(i_\ell, i_{\ell+1})$ are independent (even if $i_1 = j_1$) and centred. All the more, $P(\mathbf{i}, \mathbf{j}) = 0$, so let us pass to having two equal indices, that is, a matching.

Let us consider the case when there is *exactly one* matching. Since both indices in $\mathbf{i}$ and $\mathbf{j}$ can be reordered without affecting the variance,without loss of generality we can assume that the matching is $(i_1, i_2) = (j_1, j_2)$, and the rest of the indices of $\mathbf{i}$ are different from the ones in $\mathbf{j}$. One now has $\mathbf{i}' = (i_3, \dots, i_{2k})$

and $\mathbf{j}' = (j_3, \ldots, j_{2k})$ with $2k-2$ indices each, and so we can construct partitions π, π' for each of them independently.

For the ease of notation, let

$$a_{i,j} := \kappa_\sigma^{1/2}(W_i^m, W_j^m)\mathcal{G}(i, j)$$

and let $\sum^{(1)}$ be the sum over $\mathbf{i}, \mathbf{j}$ such that there is exactly one matching between $\mathbf{i}$ and $\mathbf{j}$. Using Wick's formula in the second equality, we have

$$\begin{aligned}
&\frac{1}{N^2 c_N^{2k}} \sum_{\mathbf{i},\mathbf{j}:[2k]\to[N]}^{(1)} P(\mathbf{i}, \mathbf{j}) \\
&= \frac{1}{N^2 c_N^{2k}} \sum_{\mathbf{i},\mathbf{j}:[2k]\to[N]}^{(1)} \mathbf{E}\left[E^W\left[\prod_{\ell=1}^{2k} a_{i_\ell, i_{\ell+1}} a_{j_\ell, j_{\ell+1}}\right]\right] \\
&= \frac{1}{N^2 c_N^{2k}} \sum_{\mathbf{i},\mathbf{j}:[2k]\to[N]} \mathbf{E}\Bigg[E^W[a_{i_1,i_2}^2] \sum_{\pi,\pi'\in\mathcal{P}_2(\{3,\ldots,2k\})} \prod_{(r,s)\in\pi} E^W[a_{i_r, i_{r+1}} a_{i_s, i_{s+1}}] \\
&\qquad \times \prod_{(r',s')\in\pi'} E^W[a_{j_{r'}, j_{r'+1}} a_{j_{s'}, j_{s'+1}}]\Bigg].
\end{aligned} \tag{3.35}$$

Following the idea of the proof for (3.27), we assume Claim 3.4.10 to be true to obtain the optimal order. We will consider $\mathbf{i}', \mathbf{j}' \in \mathrm{Cat}_{\pi,k-1}$, and notice that

$$E^W[a_{\ell,\ell'}^2] \le \frac{m^{1+\sigma}}{\|\ell - \ell'\|^\alpha}. \tag{3.36}$$

Interchanging summands, we obtain

$$\begin{aligned}
(3.35) &= \frac{1}{N^2 c_N^{2k}} \mathbf{E}\Bigg[\sum_{\pi,\pi'\in\mathcal{P}_2(\{3,\ldots,2k\})} \sum_{\substack{\mathbf{i}',\mathbf{j}'\in\mathrm{Cat}_{\pi,k-1},\\ i_1\ne i_2\in[N]}} E^W\left[a_{i_1,i_2}^2\right] \prod_{(r,s)\in\pi} E^W\left[a_{i_r i_{\gamma\pi(r)}}^2\right] \\
&\qquad \times \prod_{(r',s')\in\pi'} E^W\left[a_{j_{r'} j_{\gamma\pi(r')}}^2\right]\Bigg] + \mathcal{R}_1' \\
&\overset{(3.36)}{\le} \frac{1}{N^2 c_N^{2k}} \sum_{\pi,\pi'\in\mathcal{P}_2(\{3,\ldots,2k\})} \sum_{\substack{\mathbf{i}',\mathbf{j}'\in\mathrm{Cat}_{\pi,k-1},\\ i_1\ne i_2\in[N]}} \frac{m^{1+\sigma}}{\|i_1 - i_2\|^\alpha} \prod_{(r,s)\in\pi} \frac{m^{1+\sigma}}{\|i_r - i_{\gamma\pi(r)}\|^\alpha} \\
&\qquad \times \prod_{(r',s')\in\pi'} \frac{m^{1+\sigma}}{\|j_{r'} - j_{\gamma\pi(r')}\|^\alpha} + \mathcal{R}_1',
\end{aligned} \tag{3.37}$$

where $\mathcal{R}_1'$ is an error term such that $\lim_{N\to\infty}\mathcal{R}_1' = 0$, which follows from Claim 3.4.10. The contributing terms of the right-hand side of (3.37) can be upper-bounded by

$$\begin{aligned}
&\frac{1}{N^2 c_N^{2k}} \sum_{\pi,\pi'\in\mathcal{P}_2(\{3,\dots,2k\})} \sum_{\substack{\mathbf{i}:\mathbf{i}'\in\mathrm{Cat}_{\pi,k-1},\\ i_1\neq i_2}} \frac{m^{1+\sigma}}{\|i_1-i_2\|^\alpha} \prod_{(r,s)\in\pi} \frac{m^{1+\sigma}}{\|i_r - i_{\gamma\pi(r)}\|^\alpha} \\
&\qquad\times \sum_{\mathbf{j}'\in\mathrm{Cat}_{\pi',k-1}} \prod_{(r',s')\in\pi'} \frac{m^{1+\sigma}}{\|j_{r'} - j_{\gamma\pi(r')}\|^\alpha} \\
&= \frac{1}{N^2 c_N^{2k}} \sum_{\pi,\pi'\in\mathcal{P}_2(\{3,\dots,2k\})} \sum_{\substack{\mathbf{i}:\mathbf{i}'\in\mathrm{Cat}_{\pi,k-1},\\ i_1\neq i_2}} \frac{m^{1+\sigma}}{\|i_1-i_2\|^\alpha} \prod_{(r,s)\in\pi} \frac{m^{1+\sigma}}{\|i_r - i_{\gamma\pi(r)}\|^\alpha} O_N(Nc_k^{k-1}).
\end{aligned}$$

Analogously, the sum over $\mathbf{i}$ conditioned on $\mathbf{i}'\in\mathrm{Cat}_{\pi,k-1}$ will be at most of order Nc_N^k. Since the sum over partitions is finite and independent of N, we obtain

$$\frac{1}{N^2 c_N^{2k}} \sum_{\mathbf{i},\mathbf{j}:[2k]\to[N]}^{(1)} P(\mathbf{i},\mathbf{j}) = \mathrm{O}_N(c_N^{-1}).$$

More generally, if one has t pairings of the form $(i_1,i_2) = (j_1,j_2),\dots,(i_{t-1},i_t) = (j_{t-1},j_t)$, one can use the same argument and instead obtain a faster error of the order of c_N^{-t+1}, simply due to the set $(j_{t+1},j_2,\dots,j_{2k})$ now having only $2k-t$ independent indices from $\mathbf{i}$. Thus, we conclude

$$\mathrm{Var}\left(\int_{\mathbb{R}} x^{2k}\mu_{\tilde{\mathbf{A}}_{N,m,g}}(\mathrm{d}\,x)\right) = \mathrm{O}_N(c_N^{-1}). \tag{3.38}$$

This proves (3.28).

To conclude, one can see that

$$M_{2k} \le (m^{1+\sigma})^k C_k, \tag{3.39}$$

where C_k is the k^{th} Catalan number. Since $\sum_{k\ge1} C_k^{-1/2k} = \infty$, so Carleman's condition implies that $\{M_{2k}\}_{k\ge1}$ uniquely determine the limiting measure. Therefore we can find C, $R>0$ such that for all $k\ge1$ we have $M_{2k}\le CR^{2k}$. In turn, it is a straightforward exercise to show that this implies that $\mu_{\tau,\sigma,m}$ is compactly supported, and since it has odd moments equal to zero it is symmetric. To conclude the proof of Proposition 3.4.9 we use for example Tao [2012, pg. 134].

Proof of Claim 3.4.10. We first show a). Fix $\pi\in NC_2(2k)$. Recall that $\mathbf{i}\in\mathrm{Cat}_{\pi,k}$ is constant on the blocks of $\gamma\pi$. Therefore the number of free indices over which we can construct $\mathbf{i}$ is $\#\gamma\pi = k+1$ (Lemma 3.4.8).

Chapter Three

For any $\pi \in NC_2(2k)$, there exists *at least* one block of the form $(r, r+1) \in \pi$, where $1 \leq r \leq 2k$, and $2k+1$ is identified with "1". Then, $\{r+1\} \in \gamma\pi$ is a singleton, and consequently, i_{r+1} is a free index under $\gamma\pi$, that is, under the summation over indices $i_1, \dots, i_{2k}$, i_{r+1} runs from 1 to N independent of other indices. Moreover, as $\mathbf{i} \in \mathrm{Cat}_{\pi,k}$, we have $i_r = i_{r+2}$. If we remove the block $(r, r+1)$ from π, we obtain $\pi' \in NC_2(2k-2)$ as a new partition on $\{1, 2, \dots, r-1, r+2, \dots 2k\}$. Let $\mathbf{i}'$ be the tuple $(i_1, i_2, \dots, i_{r-1}, i_{r+2}, \dots, i_{2k})$. We then have $\mathbf{i}' \in \mathrm{Cat}_{\pi',k-1}$. So, we can write

$$\frac{1}{Nc_N^k} \sum_{\mathbf{i} \in \mathrm{Cat}_{\pi,k}} \prod_{(r,s) \in \pi} \frac{1}{\|i_r - i_s\|^\alpha}$$

$$= \frac{1}{Nc_N^k} \sum_{\mathbf{i}' \in \mathrm{Cat}_{\pi',k-1}} \left(\prod_{(r,s) \in \pi'} \frac{1}{\|i_r - i_s\|^\alpha} \right) \left(\sum_{i_{r+1}=1}^{N} \frac{1}{\|i_{r+1} - i_{r+2}\|^\alpha} \right). \tag{3.40}$$

We now proceed inductively. For $k = 1$ the result is given by (3.5). Assume now that we have shown, for some $k - 1 \geq 0$ and any $\pi' \in NC_2(2(k-1))$, that

$$\lim_{N \to \infty} \frac{1}{Nc_N^{k-1}} \sum_{\mathbf{i}' \in \mathrm{Cat}_{\pi',k-1}} \prod_{(r,s) \in \pi'} \frac{1}{\|i_r - i_s\|^\alpha} = 1. \tag{3.41}$$

We need to show the same statement holds for k, which is precisely Claim 3.4.10a). Now, we have that

$$(3.40) = \frac{1}{Nc_N^{k-1}} \sum_{\mathbf{i}' \in \mathrm{Cat}_{\pi',k-1}} \left(\prod_{(r,s) \in \pi'} \frac{1}{\|i_r - i_s\|^\alpha} \right) \left(\frac{1}{c_N} \sum_{i_{r+1}=1}^{N} \frac{1}{\|i_{r+2} - i_{r+1}\|^\alpha} \right). \tag{3.42}$$

Taking the limit $N \to \infty$, we have that the second factor in brackets above by (3.5), and then the remaining expression equals 1 by the induction hypothesis (3.41). This proves a).

To show b), we now analyse $\mathcal{R}_1$ explicitly. We have to deal with two cases:

b.1) $\pi \in \mathcal{P}_2(2k)$ and $\mathbf{i} \notin \mathrm{Cat}_{\pi,k}$.

b.2) $\pi \in \mathcal{P}_2(2k) \setminus NC_2(k)$ and $\mathbf{i} \in \mathrm{Cat}_{\pi,k}$.

Note that for both cases the following factor involving the weights will not play any role:

$$\mathbb{E}\left[\prod_{j=1}^{2k} \kappa_\sigma^{1/2}(W_{i_j}^m, W_{i_{j+1}}^m) \right] \leq m^{k(1+\sigma)}.$$

We first deal with Case b.2). From Lemma 3.4.8 we have $\#\gamma\pi \leq k$ and hence

$$\sum_{\pi \in \mathcal{P}_2(2k) \setminus NC_2(2k)} \frac{1}{Nc_N^k} \sum_{\mathbf{i} \in \mathrm{Cat}_{\pi,k}} \mathbb{E}\left[\prod_{j=1}^{2k} \kappa_\sigma^{1/2}(W_{i_j}^m, W_{i_{j+1}}^m)\right] \prod_{(r,s)\in\pi} \frac{1}{\|i_r - i_{r+1}\|^\alpha}$$
$$\leq m^{k(1+\sigma)} \sum_{\pi \in \mathcal{P}_2(2k) \setminus NC_2(2k)} \frac{1}{Nc_N^k} \sum_{i_1 \in [N]} \sum_{i_2,\ldots,i_k \in [N]} \frac{1}{\|i_2\|^\alpha \ldots \|i_k\|^\alpha}, \tag{3.43}$$

where (3.43) follows from $\mathbf{i}$ being constant on the cycles of $\gamma\pi$. Thus, we get that the terms involved in Case b.2) give a contribution of the order

$$(3.43) \leq c m^{k(1+\sigma)} \sum_{\pi \in \mathcal{P}_2(2k) \setminus NC_2(2k)} \frac{1}{N^{1+k(1-\alpha)}} N^{1+(k-1)(1-\alpha)} = \mathrm{O}_N \frac{1}{N^{1-\alpha}} = o_N(1)\,. \tag{3.44}$$

We now show that the contribution from b.1) is also negligible. Begin by fixing a partition π. For any tuple $\mathbf{i}$, we construct a corresponding graph $G(\mathbf{i})$ (recall that when $\mathbf{i} \in \mathrm{Cat}_{\pi,k}$ we ended up with $G(\mathbf{i}) = G_{\gamma\pi}$). For $\mathbf{i} \notin \mathrm{Cat}_{\pi,k}$, $G(\mathbf{i})$ is constructed by a closed walk $i_1 \to i_2 \to \ldots i_{2k} \to i_1$, thereby adding the edges $(i_p, i_{p+1})_{p=1}^{2k}$ with $i_{2k+1} = i_1$. We then collapse indices i_r, i_s into the same vertex when $\{i_r, i_{r+1}\} = \{i_s, i_{s+1}\}$, which can be justified by (3.32). We then proceed by collapsing the multiple edges and looking at the skeleton graph $G(\mathbf{i})$, with vertex set $V(\mathbf{i})$. Hence, we see that

$$\begin{aligned}
&\sum_{\pi \in \mathcal{P}_2(2k)} \frac{1}{Nc_N^k} \sum_{\mathbf{i}:[2k]\to[N]}{}^{'} \mathbb{E}\left[\prod_{j=1}^{2k} \kappa_\sigma^{1/2}(W_{i_j}^m, W_{i_{j+1}}^m)\right] \prod_{(r,s)\in\pi} \frac{1}{\|i_r - i_{r+1}\|^\alpha} \\
&\leq m^{k(1+\sigma)} \sum_{\pi \in \mathcal{P}_2(2k)} \frac{1}{Nc_N^k} N^{1+(\#V(\mathbf{i})-1)(1-\alpha)} \\
&\leq \mathrm{O} N^{(\#V(\mathbf{i})-k-1)(1-\alpha)}.
\end{aligned} \tag{3.45}$$

since $m > 1$ is fixed and the sum over the set $\mathcal{P}_2(2k)$ is finite. We see that the only non-trivial contribution comes when $\#V(\mathbf{i}) = k+1$, which signifies that $G(\mathbf{i})$ is a tree. Now we claim that for any $\pi \in \mathcal{P}_2(2k)$ and $\mathbf{i} \notin \mathrm{Cat}_{\pi,k}$ we have $\#V(\mathbf{i}) < k+1$.

When $\mathbf{i} \notin \mathrm{Cat}_{\pi,k}$, it implies that there exists at least one $(r,s) \in \pi$, such that $i_r = i_s$ and $i_{r+1} = i_{s+1}$. Let us begin by assuming that there exists *exactly one* such pair. Observe that due to the restrictions in $\sum'$, no pair-wise indices are same, hence s can neither be $r+1$, nor $r-1$. Now consider the reduced partition $\pi' = \pi \setminus (r, s)$. Observe that $\pi' \in \mathcal{P}_2(2k)(\{1, \dots, r-1, r+1, \dots, s-1, s+1, \dots, 2k\})$. Note that now $\mathbf{i}' \in \mathrm{Cat}_{\pi', k-1}$, so its contribution to (3.37) is of the order of $N^{1+(k-1)(1-\alpha)}$, which comes from the tree $G(\mathbf{i}')$ on k vertices, and where $\mathbf{i}'$ are the $(2k-2)$ indices which are obtained by removal of (i_r, i_{r+1}). So, all we are left to show is that due to Case 2), i_r and i_s will not give rise to a new vertex in $G(\mathbf{i})$.

Now, there exists an $r < e < s-1$ such that $(e, s-1) \in \pi$. Due to Case 2), we have that $i_r = i_s$ contribute to the same vertex in $G(\mathbf{i})$. Also $i_e = i_s$ and $i_{e+1} = i_{s-1}$ due to Case 1). This implies that $i_r = i_s = i_e$, where i_e is already a contributing index in $G(\mathbf{i}')$. This implies that $G(\mathbf{i})$ is a tree on at most k vertices, and hence $\#V(\mathbf{i}) \leq k$. This shows that the contribution in (3.45) goes to 0.

The case for which there is more than one pair breaking the constraint in $\mathrm{Cat}_{\pi, k}$ leads to an even smaller order. When none of the pairs satisfy the constraint then $i_r = i_{\pi(r)}$ for all r and hence i is constant on the blocks of π. So $\#V(\mathbf{i}) \leq k$ and again the contribution in (3.45) goes to 0, thus proving the claim. □

We wish to highlight that Proposition 3.4.9 is in fact more general, and works beyond the kernels κ_σ defined in (3.3).

Remark 3.4.11.
The statement of Proposition 3.4.9 holds when we replace the entries of $\tilde{\mathbf{A}}_{N,m,g}$ in (3.24) by

$$\sqrt{\frac{\kappa(W_i, W_j)}{c_N \left\|i-j\right\|^\alpha}} G_{i\wedge j, i\vee j} \quad 1 \leq i, j \leq N$$

for any function $\kappa : [1, \infty)^2 \to [0, \infty)$ which is symmetric and such that, for all $k \in \mathbb{N}$,

$$\mathbb{E}\left[\prod_{j=1}^{2k} \sqrt{\kappa(X_j, X_{j+1})}\right] < \infty \tag{3.46}$$

where $X_1, \dots, X_{2k}$ are i.i.d. random variables in $[1, \infty)$.

In our case the kernels $\kappa(x, y) := \kappa_\sigma(x, y)\mathbf{1}_{x,y\leq m}$ satisfy (3.46).

Proof of Theorem 3.2.1. To prove the final result, we shall use Lemma 3.3.5 with the complete metric space $\Sigma = \mathcal{P}(\mathbb{R})$ and metric d_L. Recall also the definition of $\tilde{\mathbf{A}}_{N,m,g}$ resp. $\overline{\mathbf{A}}_{N,m}$ of (3.24) resp. (3.17). In Proposition 3.4.9 we have shown that there exists a (deterministic) measure $\mu_{\sigma,\tau,m}$ such that, for every $m > 0$,

$$\lim_{N\to\infty} \mathrm{ESD}(\tilde{\mathbf{A}}_{N,m,g}) = \mu_{\sigma,\tau,m} \text{ in } \mathbb{P}\text{–probability.}$$

Hence for any h satisfying the assumptions of Lemma 3.4.4 and H as in (3.20) it follows that

$$\lim_{N\to\infty} \mathbb{E}\left[h\left(\Re H\left(\tilde{\mathbf{A}}_{N,m,g}\right)\right)\right] = h\left(\Re\, \mathrm{S}_{\mu_{\sigma,\tau,m}}(z)\right).$$

and thus, by means of Lemma 3.4.4 and Lemma 3.4.6,

$$\lim_{N\to\infty} \mathbb{E}\left[h\left(\Re H\left(\overline{\mathbf{A}}_{N,m}\right)\right)\right] = h\left(\Re\, \mathrm{S}_{\mu_{\sigma,\tau,m}}(z)\right).$$

Since the above holds true for any h satisfying the assumptions of Lemma 3.4.4 and $\mu_{\sigma,\tau,m}$ is deterministic, it follows that

$$\lim_{N\to\infty} \Re H\left(\overline{\mathbf{A}}_{N,m}\right) = \Re\, \mathrm{S}_{\mu_{\sigma,\tau,m}}(z) \text{ in } \mathbb{P}\text{–probability.}$$

A similar argument for the imaginary part shows that

$$\lim_{N\to\infty} \Im H\left(\overline{\mathbf{A}}_{N,m}\right) = \Im\, \mathrm{S}_{\mu_{\sigma,\tau,m}}(z) \text{ in } \mathbb{P}\text{–probability.}$$

Combining the real and imaginary parts, we have, for any $z \in \mathbb{C}^+$,

$$\lim_{N\to\infty} \mathrm{S}_{\mathrm{ESD}(\overline{\mathbf{A}}_{N,m})}(z) = \mathrm{S}_{\mu_{\sigma,\tau,m}}(z) \;\text{ in } \mathbb{P}\text{–probability.}$$

Since the convergence of the Stieltjes transform characterises weak convergence, we have

$$\lim_{N\to\infty} \mathrm{ESD}(\overline{\mathbf{A}}_{N,m}) = \mu_{\sigma,\tau,m} \;\text{ in } \mathbb{P}\text{–probability.}$$

From Lemma 3.4.6 and Lemma 3.4.3, it also follows that, for every $\delta > 0$ and $m > 0$,

$$\limsup_{N\to\infty} \mathbb{P}(d_L(\mu_{\mathbf{A}_{N,m}}, \mu_{\sigma,\tau,m}) > \delta) = 0.$$

This shows condition (1) of Lemma 3.3.5. Condition (2) follows from Lemma 3.4.2 where we have proved that

$$\limsup_{m\to\infty} \lim_{N\to\infty} \mathbb{P}\left(d_L(\mu_{\mathbf{A}_{N,m}}, \mu_{\mathbf{A}_N}) > \delta\right) = 0.$$

Thus, it follows from Lemma 3.3.5 that there exists a deterministic measure $\mu_{\sigma,\tau}$ such that

$$\lim_{m\to\infty} d_L(\mu_{\sigma,\tau,m}, \mu_{\sigma,\tau}) = 0, \tag{3.47}$$

and hence using the triangle inequality the result follows.

§3.5 Scale-Free Percolation: A special case

Proof of Theorem 3.2.2. Step 1: identification. We are now dealing with the special case of $\sigma = 1$. We go back to the moments of $\mu_{\sigma,\tau,m}$. Let $\gamma\pi = (V_1, \ldots, V_{k+1})$ and let $\ell_i = \#V_i$ (with a slight abuse of notation, we are viewing here V_i as a set rather than a cycle). Since $\sigma = 1$, $\kappa_\sigma(W_u^m, W_v^m) = W_u^m W_v^m$. It follows that

$$\begin{aligned} M_{2k} &= \sum_{\pi \in NC_2(2k)} \mathbb{E}\left[\prod_{(u,v)\in E(G_{\gamma\pi})} W_u^m W_v^m\right] \\ &= \sum_{\pi \in NC_2(2k)} \prod_{i=1}^{k+1} \mathbb{E}[(W_1^m)^{\ell_i}] \\ &= \int_{\mathbb{R}} x^{2k} \mu_{sc} \boxtimes \mu_{W,m}(\mathrm{d}\, x). \end{aligned}$$

The last equality follows from the combinatorial expression of the moments of the free multiplicative convolution of the semicircle element with an element whose law is given by $\mu_{W,m}$ (see Nica and Speicher [2006, Theorem 14.4]). Consider the map $x \mapsto x^2$ from $\mathbb{R} \to [0, \infty)$ and let μ^2 be the push-forward of a probability measure μ under this mapping, so that μ_{sc} is pushed forward to μ_{sc}^2. Then by Bercovici and Voiculescu [1993, Corollary 6.7] it follows that

$$\lim_{m\to\infty} \mu_{W,m} \boxtimes \mu_{sc}^2 \boxtimes \mu_{W,m} = \mu_W \boxtimes \mu_{sc}^2 \boxtimes \mu_W.$$

A consequence of Arizmendi and Pérez-Abreu [2009, Lemma 8] is that

$$\mu_{W,m} \boxtimes \mu_{sc}^2 \boxtimes \mu_{W,m} = (\mu_{sc} \boxtimes \mu_{W,m})^2$$

and

$$\mu_W \boxtimes \mu_{sc}^2 \boxtimes \mu_W = (\mu_{sc} \boxtimes \mu_W)^2. \tag{3.48}$$

Thus

$$\lim_{m\to\infty} (\mu_{sc} \boxtimes \mu_{W,m})^2 = (\mu_{sc} \boxtimes \mu_W)^2.$$

Observe that $\mu_{sc} \boxtimes \mu_{W,m}$ and $\mu_{sc} \boxtimes \mu_W$ are symmetric around the origin [Arizmendi and Pérez-Abreu, 2009, Theorem 7], hence we have that

$$\lim_{m\to\infty} d_L(\mu_{sc} \boxtimes \mu_W, \mu_{sc} \boxtimes \mu_{W,m}) = \lim_{m\to\infty} d_L(\mu_{sc} \boxtimes \mu_W,\, \mu_{1,\tau,m},) = 0.$$

Theorem 3.2.1 then implies that the ESD($\mathbf{A}_N$) converges to $\mu_{sc} \boxtimes \mu_W$ weakly in probability.

Step 2: tail asymptotics. In the following we use the recent results of Kołodziejek and Szpojankowski [2022, Lemma 7.2] from which we also borrow the notation. The free probability analogue of the classical Breiman's lemma is as follows: let μ, ν be probability measures and

$$\mu(x,\infty) \sim x^{-\beta} L(x) \tag{3.49}$$

with $L(\cdot)$ a slowly varying function [Kołodziejek and Szpojankowski, 2022, Definition 1.1]. Assume furthermore that the $\lfloor \beta+1 \rfloor$-th moment of ν exists:

$$m_{\lfloor \beta+1 \rfloor}(\nu) < \infty.$$

Then

$$\mu \boxtimes \nu(x,\infty) \sim m_1^{\beta}(\nu)\mu(x,\infty)$$

with $m_1(\nu)$ the first moment of ν.

Since $\mu_W \boxtimes \mu_{sc}$ is a symmetric measure we have, using Kołodziejek and Szpojankowski [2022, equation (7.3)] and (3.48),

$$\mu_W \boxtimes \mu_{sc}(x,\infty) = \frac{1}{2}(\mu_W \boxtimes \mu_{sc})^2(x^2,\infty) = \frac{1}{2}\mu_W \boxtimes \mu_{sc}^2 \boxtimes \mu_W(x^2,\infty). \tag{3.50}$$

By the commutativity and associativity of the free multiplicative convolution [Nica and Speicher, 2006, Remark 14.2] we have $\mu_W \boxtimes \mu_{sc}^2 \boxtimes \mu_W = \mu_{sc}^2 \boxtimes \mu_W \boxtimes \mu_W$. Let $\nu_W := \mu_W \boxtimes \mu_W$. Then a consequence of Kołodziejek and Szpojankowski [2022, Theorem 1.3(iv)] is that

$$\nu_W(x,\infty) \sim (m_1(\mu_W))^{\tau-1}\, \mu_W(x,\infty). \tag{3.51}$$

Therefore ν_W satisfies (3.49) with $\beta := \tau - 1$, and clearly $m_{\lfloor \tau \rfloor}(\mu_{sc}^2) < \infty$. Thus, applying Kołodziejek and Szpojankowski [2022, Lemma 7.2],

$$\begin{aligned}
(\mu_{sc} \boxtimes \nu_W)(x,\infty) &\overset{(3.50)}{=} \frac{1}{2}\mu_W \boxtimes \mu_{sc}^2 \boxtimes \mu_W(x^2,\infty) \\
&\sim \frac{1}{2}\left(m_1(\mu_{sc}^2)\right)^{\tau-1} \nu_W(x^2,\infty) \\
&\overset{(3.51)}{\sim} \frac{1}{2}\left(m_1(\mu_{sc}^2)\right)^{\tau-1} (m_1(\mu_W))^{\tau-1}\, \mu_W(x^2,\infty) \\
&\sim \frac{1}{2}\left(m_1(\mu_{sc}^2)\right)^{\tau-1} (m_1(\mu_W))^{\tau-1}\, x^{-2(\tau-1)}.
\end{aligned}$$

We can conclude noting that $m_1(\mu_W)$ is finite since $\tau > 2$ and $m_1(\mu_{sc}^2) = m_2(\mu_{sc}) = 1$ [Arizmendi and Pérez-Abreu, 2009, Proposition 5 a)]. $\square$

Chapter Three

§3.6 Non-degeneracy of the limiting measure

The proof of Theorem 3.2.3 follows the arguments in Chakrabarty et al. [2016, Theorem 2.2]. A key observation is that the limiting measure $\mu_{\sigma,\tau}$ does not depend on the parameter α. This will allow us to deal with an easier model, formally corresponding to the case $\alpha = 0$, that does not feel the influence of the torus' geometry. The lack of geometry also allows us to work on a unique probability space. More precisely, let $(G_{i,j})_{i,j\geq 1}$ be an i.i.d. sequence of $\mathcal{N}(0,1)$ random variables, and let $(W_i)_{i\geq 1}$ be an i.i.d. sequence of Pareto-distributed random variables with parameter $\tau - 1$. Assume they are defined on the same probability space $(\Omega, \mathcal{F}, \mathbf{P})$. Define the $N \times N$ matrix

$$B_{N,m} = N^{-1/2}\sqrt{\kappa_\sigma(W_i^m, W_j^m)}\, G_{i\wedge j, i\vee j}\,.$$

Let $B_{N,\infty}$ denote the matrix with non-truncated weights. The following result can be proven exactly as in Proposition 3.4.9.

Proposition 3.6.1.
Let $\mathrm{ESD}(B_{N,m})$ *be the empirical spectral distribution of* $B_{N,m}$. *Then for all* $m \geq 1$,

$$\lim_{N\to\infty} \mathrm{ESD}(B_{N,m}) = \mu_{\sigma,\tau,m} \qquad \text{in } \mathbf{P}\text{-probability.}$$

Moreover,

$$\lim_{N\to\infty} \mathrm{ESD}(B_{N,\infty}) = \mu_{\sigma,\tau} \qquad \text{in } \mathbf{P}\text{-probability.}$$

We use this result to prove Theorem 3.2.3. Recall that, for a distribution function F, the generalised inverse is given by

$$F^{\leftarrow}(y) := \inf\{x \in \mathbb{R} : F(x) \geq y\}, \quad 0 < y < 1.$$

Proof of Theorem 3.2.3. From Proposition 3.6.1, it follows that there exists a subsequence $(N_k)_{k\geq 1}$ such that $\mu_{N_k,m}$ converges weakly almost surely to $\mu_{\sigma,\tau,m}$; that is,

$$\lim_{k\to\infty} d_L(\mathrm{ESD}(B_{N_k,m}), \mu_{\sigma,\tau,m}) = 0 \quad \mathbf{P}\text{-almost surely.} \tag{3.52}$$

For a $n \times n$ matrix A, let us denote by $\lambda_1(A) \leq \lambda_2(A) \leq \cdots \leq \lambda_n(A)$ its eigenvalues. For fixed integers $1 \leq k < \infty$, $1 < m < \infty$, define the following random variables on the probability space $(\Omega \times (0,1), \mathcal{F}\otimes\mathcal{B}(0,1), \mathsf{P} = \mathbf{P}\times \mathrm{Leb})$:

$$Z_{k,m}(\omega, x) = \lambda_{\lceil N_k x\rceil}\big(B_{N_k,m}(\omega)\big), \quad \omega \in \Omega,\ x \in (0,1),$$

and

$$Z_{k,\infty}(\omega, x) := \lambda_{\lceil N_k x \rceil}\big(B_{N_k,\infty}(\omega)\big), \quad \omega \in \Omega,\ x \in (0,1).$$

Let F_m be the distribution function of $\mu_{\sigma,\tau,m}$ (we suppress the dependence on σ and τ in F_m for ease of notation), and define

$$Z_{\infty,m}(\omega, x) := F_m^{\leftarrow}(x), \quad \omega \in \Omega,\ x \in (0,1).$$

Now consider $L^2(\Omega \times (0,1))$ with the P measure. This is a complete metric space, with $d(X,Y) = \mathsf{E}[(X-Y)^2]$. Our aim is to use Lemma 3.3.5 applied to the sequence of random variables $Z_{k,m}$. We proceed therefore to check assumptions (1) and (2) of the lemma. These will directly follow if we prove that

$$\lim_{k\to\infty} \mathsf{E}\left[(Z_{k,m} - Z_{\infty,m})^2\right] = 0 \tag{3.53}$$

and

$$\lim_{m\to\infty}\lim_{k\to\infty} \mathsf{E}\left[(Z_{k,m} - Z_{k,\infty})^2\right] = 0. \tag{3.54}$$

We start by (3.53). First of all we show that

$$\lim_{k\to\infty} Z_{k,m} = Z_{\infty,m} \quad \mathsf{P}\text{-almost surely.} \tag{3.55}$$

Define

$$\begin{aligned} A &:= A' \times (0,\,1) \\ &:= \left\{\omega \in \Omega : \lim_{k\to\infty} d_L(\mathrm{ESD}(B_{N_k,m}), \mu_{\sigma,\tau,m}) = 0,\ \forall m > 1\right\} \times (0,1)\,. \end{aligned}$$

Observe that $\mathsf{P}(A) = 1$ due to (3.52) and $\mathrm{Leb}(0,1) = 1$. To prove (3.55), it suffices to show that, for all $\omega \in A'$,

$$\lim_{k\to\infty} Z_{k,m}(\omega, x) = Z_{\infty,m}(\omega, x), \quad x \in (0,1). \tag{3.56}$$

Let $F_{k,m}(\omega,\cdot)$ be the distribution function of $\mathrm{ESD}(B_{N_k,m}(\omega))$. On A, we have $F_{k,m}(\omega, x) \to F_m(x)$ for all x at which F_m is continuous. Note that

$$Z_{k,m}(\omega, x) = F_{k,m}^{\leftarrow}(\omega, x).$$

It then follows from Resnick [2008, Proposition 0.1] that for all $x \in (0,1)$

$$\lim_{k\to\infty} F_{k,m}^{\leftarrow}(x) = F_m^{\leftarrow}(x).$$

Thus, we have proved (3.55).

Chapter Three

Next, we show that for all $m \geq 1$,

$$\{Z_{k,m}^2 : 1 \leq k < \infty\} \text{ is uniformly integrable.} \tag{3.57}$$

It suffices to show that $\sup_{k\geq 1} \mathsf{E}[Z_{k,m}^4] < \infty$. Since $\lceil N_k x\rceil$ is constant on intervals of length $1/N_k$, it easily follows that

$$\begin{aligned}\lim_{k\to\infty} \mathsf{E}[Z_{k,m}^4] &= \lim_{k\to\infty} \frac{1}{N_k} \mathbf{E}\left[\sum_{i=1}^{N_k} \lambda_i(B_{N_k,m})^4\right] \\ &= \lim_{k\to\infty} \frac{1}{N_k} \mathbf{E}\, \mathrm{Tr}(B_{N_k,m}^4) = \int_{\mathbb{R}} x^4\, \mu_{\sigma,\tau,m}(\mathrm{d}\, x) < \infty\end{aligned}$$

using (3.27) and (3.29), hence (3.57) is proven. Using this and (3.55), we obtain (3.53).

We move to (3.54). To prove this note that

$$\begin{aligned}\mathsf{E}\left[(Z_{k,m} - Z_{k,\infty})^2\right] &= \frac{1}{N_k}\mathbf{E}\left[\sum_{j=1}^{N_k} \big(\lambda_j(B_{N_k,m}) - \lambda_j(B_{N_k,\infty})\big)^2\right] \\ &\overset{(3.9)}{\leq} \frac{1}{N_k}\mathbf{E}\left[\mathrm{Tr}\,\big((B_{N_k,m} - B_{N_k,\infty})^2\big)\right] \\ &= \frac{1}{N_k}\mathbf{E}\left[\sum_{i,j=1}^{N_k} \big(B_{N_k,m}(i,j) - B_{N_k,\infty}(i,j)\big)^2\right].\end{aligned}$$

Reasoning as in the proof of Lemma 3.4.2, it follows that

$$\begin{aligned}&\frac{1}{N_k}\mathbf{E}\left[\sum_{i,j=1}^{N_k} (B_{N_k,m}(i,j) - B_{N_k,\infty}(i,j))^2\right] \\ &= \frac{1}{N_k^2}\sum_{i,j=1}^{N_k} \mathbf{E}\left[\left(\sqrt{\kappa_\sigma(W_i^m, W_j^m)} - \sqrt{\kappa_\sigma(W_i, W_j)}\right)^2\right] \\ &\leq \frac{2}{N_k^2}\sum_{i,j=1}^{N_k} \mathbf{E}\left[\kappa_\sigma(W_i, W_j)\mathbf{1}_{W_j < m < W_i}\right] \\ &+ \frac{2}{N_k^2}\sum_{i,j=1}^{N_k} \mathbf{E}\left[\kappa_\sigma(W_i, W_j)\mathbf{1}_{W_i \geq W_j > m}\right].\end{aligned}$$

We can use similar bounds as for Lemma 3.4.2, which yield that both summands have order at most $m^{2-\tau}$. Hence (3.54) follows, since $\tau > 2$.

Since we have now checked assumptions (1) and (2) of Lemma 3.3.5, it follows that there exists $Z_\infty \in L^2(\Omega \times (0,1))$ such that

$$\lim_{m\to\infty} \mathsf{E}\left[(Z_{\infty,m} - Z_\infty)^2\right] = 0.$$

Let U be a uniform random variable on $(0,1)$. Then $F_m^{\leftarrow}(U)$ has the same distribution as $\mu_{\sigma,\tau,m}$. Since $\mu_{\sigma,\tau,m}$ converges weakly to $\mu_{\sigma,\tau}$ by (3.47), Z_∞ has law $\mu_{\sigma,\tau}$. Hence

$$\lim_{m\to\infty} \mathsf{E}[Z_{\infty,m}^2] = \lim_{m\to\infty} \int_{\mathbb{R}} x^2\, \mu_{\sigma,\tau,m}(\mathrm{d}\, x) = \int_{\mathbb{R}} x^2\, \mu_{\sigma,\tau}(\mathrm{d}\, x),$$

and

$$\lim_{m\to\infty} \int_{\mathbb{R}} x^2\, \mu_{\sigma,\tau,m}(\mathrm{d}\, x) = (\tau-1)^2 \int_1^\infty \int_1^\infty \frac{1}{(x\wedge y)^{\tau-\sigma}(x\vee y)^{\tau-1}}\, \mathrm{d}\, x\, \mathrm{d}\, y$$

which can be easily obtained from (3.29) with $k=1$. This completes the proof of the first part.

Since $\lim_{m\to\infty} \mu_{\sigma,\tau,m} = \mu_{\sigma,\tau}$ weakly, we apply Fatou's lemma to obtain

$$\int x^{2p}\, \mu_{\sigma,\tau}(\mathrm{d}\, x) \le \liminf_{m\to\infty} \int x^{2p}\, \mu_{\sigma,\tau,m}(\mathrm{d}\, x) = \lim_{m\to\infty} M_{2p},$$

where, recalling (3.29),

$$M_{2p} = \sum_{\pi\in NC_2(2p)} \mathbf{E}\left[\prod_{(u,v)\in E(G_{\gamma\pi})} \kappa_\sigma(W_u^m, W_v^m)\right].$$

For $\sigma > 0$, we observe that $(x\wedge y)^\sigma (x\vee y) \le (xy)^{\sigma\vee 1}$. Thus,

$$M_{2p} \le \sum_{\pi\in NC_2(2p)} \prod_{i=1}^{p+1} \mathbb{E}\left[(W_i^m)^{(\sigma\vee 1)\#V_i}\right], \tag{3.58}$$

where $\{V_1,\dots,V_{p+1}\}$ are the blocks of $\gamma\pi$. Due to Lemma 3.4.8, it follows that $\max_{1\le i\le p+1} \#V_i \le p$, typically achieved by partitions π such that

$$\gamma\pi = \{(1,3,\dots,2p-1),(2),(4),\dots,(2p)\}.$$

This shows that the maximum moment bound required for the right-hand side of (3.58) to remain finite is $\mathbb{E}[(W_i)^{p(\sigma\vee 1)}]$. Since W_i has a tail index of $\tau-1$, if $p(\sigma\vee 1) < \tau-1$, then $\mathbb{E}[(W_i)^{p(\sigma\vee 1)}] < \infty$. Therefore, M_{2p} is uniformly bounded in m, completing the proof of the theorem. $\square$

§3.7 Absolute continuity and symmetry

We begin by showing absolute continuity. We shall use the following fact from Chakrabarty and Hazra [2016, Fact 2.1], which follows from Nica and Speicher [2006, Proposition 22.32].

Lemma 3.7.1.
Assume that, for each N, A_N is a $N \times N$ Gaussian Wigner matrix scaled by $\sqrt{N}$, that is, $(A_N(i,j) : 1 \le i \le j \le N)$ are i.i.d. normal random variables with mean zero and variance $1/N$, and $A_N(j,i) = A_N(i,j)$. Suppose that B_N is a $N \times N$ random matrix, such that for all $k \ge 1$

$$\lim_{N\to\infty} \frac{1}{N} \operatorname{Tr}\left(B_N^k\right) = \int_{\mathbb{R}} x^k \mu(\mathrm{d}\, x)$$

in probability, for some compactly supported (deterministic) probability measure μ. Furthermore, let the families $(A_N : N \ge 1)$ and $(B_N : N \ge 1)$ be independent. Then for all $k \ge 1$

$$\lim_{N\to\infty} \frac{1}{N} \mathrm{E}_{\mathcal{F}} \operatorname{Tr}\left[(A_N + B_N)^k\right] = \int_{\mathbb{R}} x^k \mu \boxplus \mu_{sc}(\mathrm{d}\, x)$$

in probability, where $\mathcal{F} := \sigma(B_N : N \ge 1)$ and $\mathrm{E}_{\mathcal{F}}$ denotes the conditional expectation with respect to $\mathcal{F}$.

Proof of Theorem 3.2.4. We consider the truncated weights $(W_i^m)_{i\ge 1}$. Let Γ_m be an $N \times N$ matrix with entries given by

$$\Gamma_m(i,j) = \sqrt{\kappa_\sigma(W_i^m, W_j^m)}.$$

Given $\delta \in (0,1)$, define the function $g_{\delta,m}$ such that

$$g_{\delta,m}(W_i^m, W_j^m)^2 = \left(\sqrt{\kappa_\sigma(W_i^m, W_j^m)} - \delta\right)^2 + 2\delta\left(\sqrt{\kappa_\sigma(W_i^m, W_j^m)} - \delta\right).$$

As a consequence

$$g_{\delta,m}(W_i^m, W_j^m)^2 + \delta^2 = \kappa_\sigma(W_i^m, W_j^m)\,. \tag{3.59}$$

Define the matrix $\Gamma_{g_{\delta,m}}(i,j) = g_{\delta,m}(W_i^m, W_j^m)$. Let $\{G_{i,j}\}_{1\le i,j\le N}$ be i.i.d. standard Gaussian random variables, independent of the sequence $(W_i)_{i\ge 1}$. Denote by $\mathfrak{G}_N$ the matrix with entries

$$\mathfrak{G}_N(i,j) = \frac{1}{\sqrt{N}} G_{i\wedge j, i\vee j}\,.$$

Define
$$\mathbf{B}^{(1)}_{N,m} = \Gamma_m \circ \mathfrak{G}_N \,.$$
Similarly, define
$$\mathbf{B}^{(2)}_{N,m} = \Gamma_{g_{\delta,m}} \circ \mathfrak{G}_N \,.$$
Lastly, consider a sequence of i.i.d. standard Gaussian random variables $(G'_{i,j})_{1\le i,j\le N}$, independent of the sigma field $\mathcal{F}$ generated by $(W_i)_{i\ge 1}, (G_{i,j})_{i,j\ge 1}$. Define a matrix $\mathbf{B}^{(3)}_{N,m}$ with entries

$$\mathbf{B}^{(3)}_{N,m}(i,j) = \frac{1}{\sqrt{N}} G'_{i\wedge j, i\vee j} \,.$$

We claim that, conditionally on $(W_i)_{i\in[N]}$,

$$\mathbf{B}^{(1)}_{N,m} \stackrel{d}{=} \mathbf{B}^{(2)}_{N,m} + \delta \mathbf{B}^{(3)}_{N,m}. \tag{3.60}$$

Indeed, conditionally on $(W_i)_{i\in[N]}$, the entries of $\mathbf{B}^{(1)}_{N,m}$, $\mathbf{B}^{(2)}_{N,m}$, and $\mathbf{B}^{(3)}_{N,m}$ are normally distributed. Thus, it is sufficient to compare the mean and variance of the entries. All the variables in question have mean zero and the variances match, too, due to (3.59). Following Proposition 3.6.1, there exists a measure $\mu_{g_{\delta,m}}$ such that

$$\lim_{N\to\infty} \frac{1}{N} \operatorname{Tr}\left((\mathbf{B}^{(2)}_{N,m})^k \right) = \int_{\mathbb{R}} x^k \, \mu_{g_{\delta,m}}(\mathrm{d}\, x)$$

in probability. In particular, we recall the expression for the even moments of $\mu_{g_{\delta,m}}$ given in (3.29):

$$M_{2k} = \sum_{\pi\in NC_2(2k)} \mathbb{E}\left[\prod_{(u,v)\in E(G_{\gamma\pi})} g^2_{\delta,m}(W^m_u, W^m_v) \right].$$

Since $g^2_{\delta,m}(W^m_u, W^m_v) \le \kappa_\sigma(W^m_u, W^m_v)$, it follows that $\mu_{g_{\delta,m}}$ is uniquely determined by its moments, and is also compactly supported (Corollary 3.4.11). This verifies the first condition of Lemma 3.7.1. Since $\mathbf{B}^{(3)}_{N,m}$ is a standard Wigner matrix, it follows from Lemma 3.7.1 that

$$\lim_{N\to\infty} \frac{1}{N} \mathbb{E}_{\mathcal{F}} \left[\operatorname{Tr}\left((\mathbf{B}^{(2)}_{N,m} + \delta \mathbf{B}^{(3)}_{N,m})^k \right) \right] = \int_{\mathbb{R}} x^k \, (\mu_{g_{\delta,m}} \boxplus \mu_{sc,\delta})(\mathrm{d}\, x),$$

where $\mu_{sc,\delta}$ is the semicircular law with variance δ^2 and density

$$\mu_{sc,\delta}(\mathrm{d}\, x) = \frac{1}{2\pi\delta} \sqrt{4 - \left(\frac{x}{\delta}\right)^2} \mathbf{1}_{|x|\le 2\delta} \, \mathrm{d}\, x, \quad x \in \mathbb{R}.$$

Since both $\mu_{g_{\delta,m}}$ and $\mu_{sc,\delta}$ are compactly supported, so is $\mu_{g_{\delta,m}} \boxplus \mu_{sc,\delta}$, and thus the measure is completely determined by its moments.

From Proposition 3.4.9 we have

$$\lim_{N\to\infty} \mathbb{E}\left[\frac{1}{N}\mathbb{E}_{\mathcal{F}}[\mathrm{Tr}(\mathbf{B}_{N,m}^{(1)})^k]\right] = \int_{\mathbb{R}} x^k \, \mu_{\sigma,\tau,m}(\mathrm{d}\, x)$$

and

$$\lim_{N\to\infty} \mathrm{Var}\left(\frac{1}{N}\mathbb{E}_{\mathcal{F}}[\mathrm{Tr}((\mathbf{B}_{N,m}^{(1)})^k)]\right) \le \lim_{N\to\infty} \mathrm{Var}\left(\frac{1}{N}\,\mathrm{Tr}((\mathbf{B}_{N,m}^{(1)})^k)\right) = 0.$$

Thus,

$$\lim_{N\to\infty} \frac{1}{N}\mathbb{E}_{\mathcal{F}}\left[\mathrm{Tr}(\mathbf{B}_{N,m}^{(1)})^k\right] = \int_{\mathbb{R}} x^k \, \mu_{\sigma,\tau,m}(\mathrm{d}\, x)$$

in probability. Since the measures are uniquely determined by their moments, this shows that

$$\mu_{\sigma,\tau,m} = \mu_{g_{\delta,m}} \boxplus \mu_{sc,\delta}. \tag{3.61}$$

We show that there exists μ_{g_δ} such that

$$\lim_{m\to\infty} d_L(\mu_{g_{\delta,m}}, \mu_{g_\delta}) = 0. \tag{3.62}$$

If we can prove this, using Bercovici and Voiculescu [1993, Proposition 4.13] it will follow that

$$\lim_{m\to\infty} d_L(\mu_{g_{\delta,m}} \boxplus \mu_{sc,\delta}, \mu_{g_\delta} \boxplus \mu_{sc,\delta}) \le \lim_{m\to\infty} d_L(\mu_{g_{\delta,m}}, \mu_{g_\delta}) = 0. \tag{3.63}$$

To show (3.62), we employ Lemma 3.3.5. Note that, from Remark 3.4.11, we get that for any fixed $m \ge 1$ one has

$$\lim_{N\to\infty} d_L\big(\mu_{\mathbf{B}_{N,m}^{(2)}}, \mu_{g_{\delta,m}}\big) = 0 \qquad \text{in } \mathbb{P}\text{-probability}$$

where $\mu_{\mathbf{B}_{N,m}^{(2)}}$ is the empirical spectral distribution of $\mathbf{B}_{N,m}^{(2)}$.

This establishes condition (1) of Lemma 3.3.5. To complete the proof, we need to verify condition (2), namely,

$$\lim_{m\to\infty} \limsup_{N\to\infty} \mathbb{P}\big(d_L(\mathrm{ESD}(\mathbf{B}_{N,m}^{(2)}), \mathrm{ESD}(\mathbf{B}_N^{(2)})) > \varepsilon\big) = 0. \tag{3.64}$$

Here $\mathbf{B}_N^{(2)}$ is defined as $\mathbf{B}_{N,\infty}^{(2)}$ with $m = \infty$. From Proposition 3.3.1 we see that

$$\begin{aligned} d_L\left(\mathrm{ESD}(\mathbf{B}_{N,m}^{(2)}), \mathrm{ESD}(\mathbf{B}_N^{(2)})\right)^3 &\le \frac{1}{N}\,\mathrm{Tr}\left((\mathbf{B}_{N,m}^{(2)} - \mathbf{B}_N)^2\right) \\ &= \frac{1}{N^2}\sum_{i,j=1}^{N} \left(\Gamma_{g_{\delta,m}}(i,j) - \Gamma_{g_{\delta,\infty}}(i,j)\right)^2 G_{i\wedge j, i\vee j}^2 \,. \end{aligned}$$

Hence we have

$$\mathbb{E}\left[d_L\big(\operatorname{ESD}(\mathbf{B}_{N,m}^{(2)}),\operatorname{ESD}(\mathbf{B}_N^{(2)})\big)^3\right] \leq \frac{1}{N^2}\sum_{i\neq j=1}^{N}\mathbf{E}[\big(\Gamma_{g_{\delta,m}}(i,j)-\Gamma_{g_{\delta,\infty}}(i,j)\big)^2]$$
$$\leq \frac{2}{N^2}\sum_{i\neq j=1}^{N}\mathbf{E}\left[g_{\delta,\infty}(W_i,\,W_j)^2\left(\mathbf{1}_{W_j<m<W_i}+\mathbf{1}_{W_i>W_j>m}\right)\right]$$
$$\leq \frac{2}{N^2}\sum_{i\neq j=1}^{N}\mathbf{E}\left[\kappa_\sigma(W_i,W_j)\left(\mathbf{1}_{W_j<m<W_i}+\mathbf{1}_{W_i>W_j>m}\right)\right].$$

Just as in the proof of (3.54), it follows that the last term is bounded by $Cm^{2-\tau}$. Thus, using Markov's inequality, condition (2) of Lemma 3.3.5 holds, too. In conclusion, we can show that there exists μ_{g_δ} such that

$$\lim_{m\to\infty} d_L(\mu_{g_{\delta,m}}\boxplus\mu_{sc,\delta},\mu_{\sigma,\tau}) \overset{(3.61)}{=} \lim_{m\to\infty} d_L(\mu_{\sigma,\tau,m},\mu_{\sigma,\tau}) \overset{(3.47)}{=} 0$$
$$\overset{(3.63)}{=} \lim_{m\to\infty} d_L(\mu_{g_{\delta,m}}\boxplus\mu_{sc,\delta},\mu_{g_\delta}\boxplus\mu_{sc,\delta}).$$

Therefore it must be that $\mu_{\sigma,\tau}=\mu_{g_\delta}\boxplus\mu_{sc,\delta}$. The right-hand side is absolutely continuous, as shown by Biane [1997, Corollary 2].

Finally, to show symmetry, we see that $\mu_{\sigma,\tau}$ does not give weight to singletons by absolute continuity. Therefore, in light of the weak convergence stated in (3.47),

$$\mu_{\sigma,\tau}(-\infty,\,-x) = \lim_{m\to\infty}\mu_{\sigma,\tau,m}(-\infty,\,-x)$$
$$= \lim_{m\to\infty}\mu_{\sigma,\tau,m}(x,\,+\infty) = \mu_{\sigma,\tau}(x,\,+\infty)$$

for all $x\geq 0$. This completes the proof. □

§3.8 Stieltjes transform of the limiting measure

To prove Theorem 3.2.5, we first identify the Stieltjes transform for the measure $\mu_{\sigma,\tau,m}$. We then proceed to take the limit $m\to\infty$, which requires a functional analytic approach. Throughout this section, we fix $z\in\mathbb{C}^+$, given as $z=\xi+i\eta$ with $\eta>0$. If μ is a probability measure having all its moments $\{m_k\}_{k\geq 1}$, it follows from the definition of Stieltjes transform (3.7) that, for any $z\in\mathbb{C}^+$,

$$\mathrm{S}_\mu(z) = -\sum_{k\geq 0}\frac{m_k}{z^{k+1}}, \tag{3.65}$$

where the Laurent series on the right-hand side of (3.65) converges for $|z|>R>0$, with $\operatorname{supp}(\mu)=[-R,R]$.

§3.8.1 Stieltjes transform for truncated weights

To derive a characterisation of the limiting measure $\mu_{\sigma,\tau}$, we need to first study the truncated version $\mu_{\sigma,\tau,m}$. We borrow ideas from the proof of Chakrabarty et al. [2015, Theorem 4.1]. The main result of this subsection will be Proposition 3.8.1, which requires a few technical lemmas to prove. The results in this subsection hold for the regime $\tau > 2$ and $\sigma < \tau - 1$, as before.

We have that the (even) moments for the measure $\mu_{\sigma,\tau,m}$ are given by (3.29). Using these, we derive a representation of $\mathrm{S}_{\mu_{\sigma,\tau,m}}(z)$.

Proposition 3.8.1.
For $\tau > 2$ and $\sigma \in (0, \tau-1)$ there exists a function $a(z,x) = a_m(z,x)$ defined on $\mathbb{C}^+ \times [1,\infty)$ such that

$$\mathrm{S}_{\mu_{\sigma,\tau,m}}(z) = \int_1^\infty a(z,x)\mu_{W,m}(\mathrm{d}\,x)\,,$$

where $\mu_{W,m}$ is the law of the truncated weights (W_i^m). Moreover, $a(z,x)$ satisfies the following recursive equation:

$$a(z,x)\left(z + \int_1^\infty a(z,y)\kappa_\sigma(x,y)\mu_{W,m}(\mathrm{d}\,y)\right) = -1\,. \tag{3.66}$$

Before tackling the proof of the proposition, we lay the ground with two auxiliary results. For any $k \geq 1$ and $\pi \in NC_2(2k)$, recall the map $\mathcal{T}_\pi$ of (3.26), where $\gamma\pi = \{V_1, \dots, V_{k+1}\}$. Consider the mapping $L_\pi : [1,\infty)^{k+1} \to \mathbb{R}$ defined as

$$L_\pi(\mathbf{x}) = \kappa_\sigma^{1/2}(x_{\mathcal{T}_\pi(1)}, x_{\mathcal{T}_\pi(2)})\kappa_\sigma^{1/2}(x_{\mathcal{T}_\pi(2)}, x_{\mathcal{T}_\pi(3)}) \dots \kappa_\sigma^{1/2}(x_{\mathcal{T}_\pi(2k)}, x_{\mathcal{T}_\pi(1)}) \tag{3.67}$$

and the function $H_\pi : \mathbb{R} \to \mathbb{R}^+$ given as

$$H_\pi(y) = \int_{[1,\infty)^k} L_\pi(y,\, x_2, \dots,\, x_{k+1})\mu_{W,m}^{\otimes k}(\mathrm{d}\,\mathbf{x}')\,, \tag{3.68}$$

where we are integrating over $\mathbf{x}' = (x_2, \dots, x_{k+1}) \in [1,\infty)^k$.

Lemma 3.8.2.
Let $\{M_{2k}\}_{k\geq 1}$ be as in (3.29). Then

$$M_{2k} = \sum_{\pi \in NC_2(2k)} \int_1^\infty H_\pi(y)\mu_{W,m}(\mathrm{d}\,y).$$

Proof of Lemma 3.8.2. We begin by evaluating the integral on the right-hand side. We have

$$\begin{aligned}
&\int_1^\infty H_\pi(y)\mu_{W,m}(\mathrm{d}\, y) \\
&= \int_1^\infty \int_{[1,\infty)^k} L_\pi(y, x_2, \ldots, x_{k+1})\mu_{W,m}^{\otimes k}(\mathrm{d}\,\mathbf{x}')\mu_{W,m}(\mathrm{d}\, y) \\
&= \int_{[1,\infty)^{k+1}} \kappa_\sigma^{1/2}(x_{\mathcal{T}_\pi(1)}, x_{\mathcal{T}_\pi(2)}) \cdots \kappa_\sigma^{1/2}(x_{\mathcal{T}_\pi(2k)}, x_{\mathcal{T}_\pi(1)})\mu_{W,m}^{\otimes k+1}(\mathrm{d}\,\mathbf{x}).
\end{aligned}$$

We know that, for $\pi \in NC_2(2k)$, $\#\gamma\pi = k+1$ and so the graph $G_{\gamma\pi}$ has $k+1$ vertices. Furthermore, when we perform a closed walk of the form $1 \to 2 \to \ldots \to 2k \to 1$ on the (unoriented) graph $G_{\gamma\pi}$, we traverse each edge *exactly* twice. In particular, the product $\kappa_\sigma^{1/2}(x_{\mathcal{T}_\pi(1)}, x_{\mathcal{T}_\pi(2)}) \cdots \kappa_\sigma^{1/2}(x_{\mathcal{T}_\pi(2k)}, x_{\mathcal{T}_\pi(1)})$ has $2k$ terms with k matchings, and so

$$\kappa_\sigma^{1/2}(x_{\mathcal{T}_\pi(1)}, x_{\mathcal{T}_\pi(2)}) \cdots \kappa_\sigma^{1/2}(x_{\mathcal{T}_\pi(2k)}, x_{\mathcal{T}_\pi(1)}) = \prod_{(u,v)\in E(G_{\gamma\pi})} \kappa_\sigma(x_u, x_v)\,.$$

We then have that

$$\begin{aligned}
\int_1^\infty H_\pi(y)\mu_{W,m}(\mathrm{d}\, y) &= \int_{[1,\infty)^{k+1}} \prod_{(u,v)\in E(G_{\gamma\pi})} \kappa_\sigma(x_u, x_v)\mu_{W,m}^{\otimes k+1}(\mathrm{d}\,\mathbf{x}) \\
&= \mathbb{E}\left[\prod_{(u,v)\in E(G_{\gamma\pi})} \kappa_\sigma(W_u^m, W_v^m)\right],
\end{aligned}$$

which concludes the proof. $\square$

We show now some properties of H_π that will help us in the upcoming computations.

Lemma 3.8.3.
Let $k \geq 1$ and let H_π be as defined in (3.68). *Let $\pi \in NC_2(2k)$. Then,*

(1) If $\pi = (1, 2k) \cup \pi_1$, where π_1 is a non-crossing pair partition of $\{2, \ldots, 2k-1\}$, then,

$$H_\pi(y) = \int_1^\infty H_{\pi_1}(x)\kappa_\sigma(x, y)\mu_{W,m}(\mathrm{d}\, x). \tag{3.69}$$

(2) If $\pi = \pi_1 \cup \pi_2$, then $H_\pi(\cdot) = H_{\pi_1}(\cdot)H_{\pi_2}(\cdot)$.

Proof of Lemma 3.8.3. We first prove property (1). Let $\pi = (1, 2k) \cup \pi_1$. Then, $\gamma\pi = \{(1), V_2, \ldots, V_{k+1}\}$. We know that $2 \in V_2$ and then $\gamma\pi(2k) = 2 \in V_2$. Now, fix $x_1 = y$. Then

$$
\begin{aligned}
&H_\pi(y) \\
&= \int_{[1,\infty)^k} L_\pi(y, x_2, \ldots, x_{k+1}) \mu_{W,m}^{\otimes k}(\mathrm{d}\,\mathbf{x}') \\
&= \int_{[1,\infty)^k} \kappa_\sigma^{1/2}(y, x_2)\kappa_\sigma^{1/2}(x_2, x_{\mathcal{T}_\pi(3)}) \ldots \kappa_\sigma^{1/2}(x_{\mathcal{T}_\pi(2k-1)}, x_2)\kappa_\sigma^{1/2}(x_2, y)\mu_{W,m}^{\otimes k}(\mathrm{d}\,\mathbf{x}') \\
&= \int_1^\infty \kappa_\sigma(y, x_2) \\
&\qquad \times \int_{[1,\infty)^{k-1}} \kappa_\sigma^{1/2}(x_2, x_{\mathcal{T}_\pi(3)}) \ldots \kappa_\sigma^{1/2}(x_{\mathcal{T}_\pi(2k-1)}, x_2)\mu_{W,m}^{\otimes k-1}(\mathrm{d}\,\mathbf{x}'')\mu_{W,m}(\mathrm{d}\,x_2) \\
&= \int_1^\infty \kappa_\sigma(y, x_2) H_{\pi_1}(x_2)\mu_{W,m}(\mathrm{d}\,x_2),
\end{aligned}
$$

which is what we desired.

For property (2), let $\pi = \pi_1 \cup \pi_2$, with $\pi_1 \in NC_2(\{1, 2, \ldots, 2r\})$ and $\pi_2 \in NC_2(\{2r+1, \ldots, 2k\})$ and let us consider the function $H_\pi(y)$ with $y = x_1 = x_{\mathcal{T}_\pi(1)}$. Then,

$$
\begin{aligned}
&H_\pi(y) = \\
&\int_{[1,\infty)^k} \kappa_\sigma^{1/2}(y, x_{\mathcal{T}_\pi(2)}) \ldots \kappa_\sigma^{1/2}(x_{\mathcal{T}_\pi(2r)}, x_{\mathcal{T}_\pi(2r+1)}) \ldots \kappa_\sigma^{1/2}(x_{\mathcal{T}_\pi(2k)}, y)\mu_{W,m}^{\otimes k}(\mathrm{d}\,\mathbf{x}').
\end{aligned}
$$

We now claim that this integral can be split up into two integrals. First, consider the element $x_{\mathcal{T}_\pi(1)}$. Since we assume that '1' maps to $V_1 \in \gamma\pi$, all elements of V_1 are mapped to y. To understand where other elements are mapped, we will state a claim and see its consequences to this proof, and then prove it on 149.

Claim 3.8.4.
Under $\gamma\pi$, the elements $\{2, \ldots, 2r\}$ are mapped to the blocks

$$V_1 \cup \{V_2, \ldots, V_{r'}\} \subset \gamma\pi\,,$$

and the elements $\{2r+1, \ldots, 2k\}$ are mapped to the blocks

$$V_1 \cup \{V_{r'+1}, \ldots, V_{k+1}\} \subset \gamma\pi\,,$$

where $r' < k+1$ is some index. In particular $\gamma\pi(2r+1) \in V_1$.

From this claim we have that

$$
\begin{aligned}
&H_\pi(y)\\
&= \int_{[1,\infty)^k} \kappa_\sigma^{1/2}(y, x_{\mathcal{T}_\pi(2)}) \dots \kappa_\sigma^{1/2}(x_{\mathcal{T}_\pi(2r)}, x_{\mathcal{T}_\pi(2r+1)}) \dots \kappa_\sigma^{1/2}(x_{\mathcal{T}_\pi(2k)}, y) \mu_{W,m}^{\otimes k}(\mathrm{d}\,\mathbf{x}')\\
&= \int_{[1,\infty)^k} \kappa_\sigma^{1/2}(y, x_{\mathcal{T}_\pi(2)}) \dots \kappa_\sigma^{1/2}(x_{\mathcal{T}_\pi(2r)}, y) \dots \kappa_\sigma^{1/2}(x_{\mathcal{T}_\pi(2k)}, y) \mu_{W,m}^{\otimes k}(\mathrm{d}\,\mathbf{x}')\\
&= \int_{[1,\infty)^{r'}} \kappa_\sigma^{1/2}(y, x_{\mathcal{T}_\pi(2)}) \dots \kappa_\sigma^{1/2}(x_{\mathcal{T}_\pi(2r)}, y) \mu_{W,m}^{\otimes r'}(\mathrm{d}\,\mathbf{x}^{(r')})\\
&\times \int_{[1,\infty)^{k-r'}} \kappa_\sigma^{1/2}(y, x_{\mathcal{T}_\pi(2r+2)}) \dots \kappa_\sigma^{1/2}(x_{\mathcal{T}_\pi(2k)}, y) \mu_{W,m}^{\otimes (k-r')}(\mathrm{d}\,\mathbf{x}^{(k-r')})\\
&= H_{\pi_1}(y) H_{\pi_2}(y).
\end{aligned}
$$

This concludes the proof. $\square$

Proof of Claim 3.8.4. Let γ_1 resp. γ_2 be the shift by one on $[2r]$ resp. $\{2r+1, \dots, 2k\}$. To prove this claim, it suffices to analyse the special indices $\{1, 2r, 2r+1, 2k\}$, since γ_1 and γ_2 are cyclic permutations on $[2r]$ and $\{2r+1, \dots, 2k\}$, respectively. We will be using the fact that all elements in a block of $\gamma\pi$ must be either all odd or all even [Avena et al., 2023, Property 1], and that any pairing in π must have one element odd and the other even [Avena et al., 2023, Property 2].

(a) We already have $1 \in V_1$. Now, let $(o_1, 2k) \in \pi_2$, for some o_1 such that $o_1 \geq 2r+1$. Then, o_1 must be odd. Now, $o_1 + 1$ is even, and cannot belong to V_1. Thus $\gamma\pi(2k) = o_1 + 1 \in \{V_{r'+1}, \dots, V_{2k}\}$. This takes care of the index $2k$.

(b) Let us continue with $(o_2, 2r) \in \pi_1$ for some o_2. We know that o_2 must be odd. Thus, $\gamma\pi(2r) = o_2 + 1 \in \{V_2, \dots, V_{r'}\} =: \gamma_1\pi_1 \setminus V_1$. This resolves the case of $2r$.

(c) Lastly, by construction, $\gamma\pi(o_2) = 2r+1$, which brings us to the last special element. Since o_2 and $2r+1$ belong to the same block in $\gamma\pi$, it suffices to show that this block is V_1, that is, the block to which element 1 belongs. Now, if $(1, o_2 - 1) \in \pi_1$, we are done, since $\gamma\pi(1) = o_2$. Suppose not, and let $(1, e_1) \in \pi_1$ for some even integer e_1. Similarly as before, if now $(e_1 + 1, o_2 - 1) \in \pi_1$, we are done. Since π_1 and π_2 act on the first $2r$ elements and the remaining $2k - 2r$ elements respectively, then, by the non-crossing nature, there is a sequence of even integers $\{e_i\}_{i=1}^t$ such that $(1, e_1), (e_1 + 1, e_2), \dots, (e_t + 1, o_2 - 1) \in \pi_1$. Computing $\gamma\pi$ recursively gives us that $\gamma\pi(1) = o_2$, and so $\gamma\pi(2r+1) \in V_1$.

This proves the claim. $\square$

We are now ready to prove Proposition 3.8.1.

Proof of Proposition 3.8.1. We now derive the Stieltjes transform of the measure $\mu_{\sigma,\tau,m}$. Using (3.65) and Proposition 3.4.9, we have that

$$\mathrm{S}_{\mu_{\sigma,\tau,m}}(z) = -\sum_{k\geq 0} \frac{M_{2k}}{z^{2k+1}}.$$

Using Lemma 3.8.2 we substitute the expression for M_{2k}. We have

$$\begin{aligned}\mathrm{S}_{\mu_{\sigma,\tau,m}}(z) &= -\sum_{k\geq 0} \frac{1}{z^{2k+1}} \int_1^\infty \sum_{\pi\in NC_2(2k)} H_\pi(x)\mu_{W,m}(\mathrm{d}\, x) \\ &= -\int_1^\infty \sum_{k\geq 0} \sum_{\pi\in NC_2(2k)} \frac{H_\pi(x)}{z^{2k+1}} \mu_{W,m}(\mathrm{d}\, x), \qquad (3.70)\end{aligned}$$

where we could interchange the integral and the sum by Fubini's theorem. Now, we define the function $a(z,x)$ as

$$a(z,x) := -\sum_{k\geq 0} \sum_{\pi\in NC_2(2k)} \frac{H_\pi(x)}{z^{2k+1}}. \qquad (3.71)$$

Then using (3.70) we have

$$\mathrm{S}_{\mu_{\sigma,\tau,m}}(z) = \int_1^\infty a(z,x)\mu_{W,m}(\mathrm{d}\, x).$$

We now state some properties of the function $a(z,x)$. Firstly, for any $z\in\mathbb{C}^+$ the map $x\mapsto a(z,x)$ is in $L^\infty([1,\infty),\mu_{W,m})$ as H_π is bounded . Secondly, for any $x\in[1,\infty)$, the map $z\mapsto a(z,x)$ is analytic in $\mathbb{C}$, which follows from the Laurent series expansion. Finally we see that $a(z,x)$ lies in $\mathbb{C}^+$, for any $z\in\mathbb{C}^+$ and $x>1$. Indeed, for any $\Im(z)>0$, the expansion on the right-hand side of (3.71) will always have a non-trivial imaginary part. Thus, since $a(\cdot,\cdot)$ is analytic, it will either lie completely in $\mathbb{C}^-$ or $\mathbb{C}^+$, since it can never take values in $\mathbb{R}$. However, $\mathrm{S}_{\mu_{W,m}}(z)\in\mathbb{C}^+$, and thus, $a(z,x)\in\mathbb{C}^+$ for any $z\in\mathbb{C}^+$ and $x>1$.

To write down a functional recursion for $a(\cdot,\cdot)$ it is convenient to use the notion of words. Any partition π can be associated to a *word* w, with any elements in $i,j\in[2k]$ being associated with the same letter in w if i,j are in the same block of π. For example, $\pi=\{\{1,2\},\{3,4\}\}$ can be written as $w=aabb$. In particular, any partition $\pi\in NC_2(2k)$ can be associated to a word w of the form $w = aw_1aw_2$, where w_1, w_2 are words that can be empty. For any word w associated to a partition π, let $H_\pi = H_w$. Furthermore, for $w \in NC_2(2k)$ we mean a word w whose associated partition π is in $NC_2(2k)$. Then we have, using Lemma 3.8.3 in the third equality,

$$\begin{aligned}
a(z,x) &= -\sum_{k\geq 0}\sum_{w\in NC_2(2k)} \frac{H_w(x)}{z^{2k+1}} \\
&= -\frac{1}{z} - \sum_{k\geq 1}\sum_{\substack{w\in NC_2(2k)\\ w=aw_1aw_2}} \frac{H_{aw_1aw_2}(x)}{z^{2k+1}} \\
&= -\frac{1}{z} - \sum_{k\geq 1}\sum_{\substack{w\in NC_2(2k)\\ w=aw_1aw_2}} \frac{H_{aw_1a}(x)H_{w_2}(x)}{z^{2k+1}} \\
&= -\frac{1}{z} - \frac{1}{z}\sum_{k\geq 1}\sum_{\ell=1}^{k}\sum_{w_1\in NC_2(2\ell-2)} \frac{H_{aw_1a}(x)}{z^{2\ell-2+1}} \sum_{w_2\in NC_2(2k-2\ell)} \frac{H_{w_2}(x)}{z^{2k-2\ell+1}}.
\end{aligned} \tag{3.72}$$

One can see that the word aw_1a has as corresponding partition $(1, 2\ell)\cup\pi_1$, with $\pi_1 \in NC_2(2\ell-2)$. Using (3.69) from Lemma 3.8.3, we have

$$\begin{aligned}
&a(z,x) \\
&= -\frac{1}{z} - \frac{1}{z}\sum_{k\geq 1}\sum_{\ell=1}^{k}\sum_{w_1\in NC_2(2\ell-2)} \frac{1}{z^{2\ell-1}}\int_1^\infty H_{w_1}(y)\kappa_\sigma(x,y)\mu_{W,m}(\mathrm{d}\,y) \\
&\qquad\qquad \times \sum_{w_2\in NC_2(2k-2\ell)} \frac{H_{w_2}(x)}{z^{2k-2\ell+1}} \\
&= -\frac{1}{z} - \frac{1}{z}\sum_{k\geq 1}\sum_{\ell=1}^{k}\sum_{\pi_2\in NC_2(2k-2\ell)} \frac{H_{\pi_2}(x)}{z^{2k-2\ell+1}} \\
&\qquad\qquad \times \sum_{\pi_1\in NC_2(2\ell-2)} \frac{1}{z^{2\ell-1}}\int_1^\infty H_{\pi_1}(y)\kappa_\sigma(x,y)\mu_{W,m}(\mathrm{d}\,y) \\
&= -\frac{1}{z} - \frac{a(z,x)}{z}\int_1^\infty a(z,y)\kappa_\sigma(x,y)\mu_{W,m}(\mathrm{d}\,y).
\end{aligned}$$

Thus, we have (3.66), which completes the proof of Proposition 3.8.1. □

Remark 3.8.5.
Equation (3.66) gives an analytic description of a in terms of the recursive equation. Now, for any $z \in \mathbb{C}^+$, we have that

$$z = i\int_0^\infty \mathrm{e}^{-itz^{-1}}\,\mathrm{d}\,t. \tag{3.73}$$

Chapter Three

Since $a(z,x) \in \mathbb{C}^+$ *for any fixed* $x \in [1,\infty)$*, applying* (3.73) *to* $a(z,x)$ *and using* (3.66) *gives us that*

$$a(z,x) = i \int_0^\infty \mathrm{e}^{itz} \exp\left\{ it \int_1^\infty a(z,y)\kappa_\sigma(x,y)\mu_{W,m}(\mathrm{d}\,y) \right\} \mathrm{d}\,t. \tag{3.74}$$

An immediate consequence of (3.74) *is that* $a(z,x)$ *is uniformly bounded in* x *and* m*. Indeed, if we take* $z = \xi + i\eta$ *with* $\eta > 0$*, we have that*

$$\begin{aligned} |a(z,x)| &\leq \int_0^\infty \mathrm{e}^{-\eta t} \left| \exp\left\{ it \int_1^\infty a(z,y)\kappa_\sigma(x,y)\mu_{W,m}(\mathrm{d}\,y) \right\} \right| \mathrm{d}\,t \\ &\leq \int_0^\infty \mathrm{e}^{-\eta t}\,\mathrm{d}\,t = \frac{1}{\eta}. \end{aligned} \tag{3.75}$$

The bound in the second line holds since $a(z,x) \in \mathbb{C}^+$*, and so*

$$\int_1^\infty a(z,y)\kappa_\sigma(x,y)\mu_{W,m}(\mathrm{d}\,y) \in \mathbb{C}^+$$

as $\kappa_\sigma \geq 1$*.*

§3.8.2 Limiting Stieltjes transform

We now set up the framework required to prove Theorem 3.2.5. For the remainder of this section, denote $a_z(x) := a(z,x)$, which implicitly depends on m. We wish to extend Proposition 3.8.1 to the measure $\mu_{\sigma,\tau}$ by passing to the limit $m \to \infty$. We have a natural candidate for the function a^* in Theorem 3.2.5, which should be the limit of $a(\cdot,\cdot)$ as m tends to infinity. We now formalise this idea through a series of lemmas.

Since our goal now is to show Theorem 3.2.5 we are going to work for the remainder of this section with the following parameters:

(a) $\tau > 3$,

(b) $\sigma < \tau - 2$, and

(c) a parameter β such that $2 \vee 1 + \sigma < \beta < \tau - 1$.

Let $\overline{\mathbb{C}}^+ = \mathbb{C}^+ \cup \mathbb{R}$ be the closure of $\mathbb{C}^+$, and let ν be the measure defined as

$$\nu(\mathrm{d}\,x) = x^{-\beta}\,\mathrm{d}\,x. \tag{3.76}$$

Consider the space $L^1([1,\infty),\nu)$ of all functions $f : [1,\infty) \to \overline{\mathbb{C}}^+$ that are L^1–integrable with respect to ν.

Definition 3.8.6.
Let $\mathcal{B}$ denote the Banach space $\mathcal{B} := (L^1([1,\infty),\nu), \|\cdot\|_1)$, where the norm $\|\cdot\|_1$ is the L^1 norm with respect to ν as in (3.76)*, which is defined for $f \in L^1([1,\infty),\nu)$ as*

$$\|f\|_1 := \int_1^\infty |f(x)| x^{-\beta}\,\mathrm{d}\,x. \tag{3.77}$$

Recall that $\mu_{W,m}$ denotes the law of the truncated weights $(W_x^m)_x$, given as

$$\mu_{W,m}(\cdot) = c_m^{-1}\mu_w(\cdot)\mathbb{1}_{\{\cdot\le m\}},$$

where $c_m = 1 - m^{-(\tau-1)}$ is a normalizing constant converging to 1 as m tends to infinity, and μ_W is the Pareto law defined in (3.2). For $z \in \mathbb{C}^+$, let T_z denote the map

$$T_z f(\cdot) = i\int_0^\infty \mathrm{e}^{itz}\exp\left\{it\int_1^\infty f(y)\kappa_\sigma(\cdot,\, y)\mu_W(\mathrm{d}\,y)\right\}\mathrm{d}\,t. \tag{3.78}$$

Then, we have the following result.

Lemma 3.8.7.
There exists a constant $\tilde{c} = \tilde{c}(\tau,\sigma,\beta)$ such that, for all $z \in \mathbb{C}^+$ with $\Im(z)^2 = \eta^2 > \tilde{c}$ $T_z : \mathcal{B} \to \mathcal{B}$ is a contraction mapping, with a contraction constant $\tilde{c}\eta^{-2}$.

Proof of Lemma 3.8.7. We first need to show that, for any $f \in \mathcal{B}$, one has $T_z f \in \mathcal{B}$. Indeed, for $x \ge 1$ it holds that

$$\left|T_z f(x)\right| \le \int_0^\infty \mathrm{e}^{-\eta t}\left|\exp\left\{it\int_1^\infty f(y)\kappa_\sigma(x,y)\mu_W(\mathrm{d}\,y)\right\}\right|\mathrm{d}\,t \le \frac{1}{\eta},$$

where the last inequality holds as $f(y) \in \overline{\mathbb{C}^+}$ for any $y \ge 1$, and thus the second complex exponential is bounded by 1. Since $|T_z f(\cdot)|$ is uniformly bounded, it is L^1-integrable with respect to ν, and so $T_z(\mathcal{B}) \subseteq \mathcal{B}$.

Now, we wish to show T_z is a contraction. Let us take $f_1, f_2 \in \mathcal{B}$. Recall that for any $z_1, z_2 \in \overline{\mathbb{C}^+}$ and $t > 0$, we have

$$|\mathrm{e}^{itz_1} - \mathrm{e}^{itz_2}| \le t|z_1 - z_2|. \tag{3.79}$$

Then, for any $x \in [1,\infty)$ we have that

$$\begin{aligned}
&|T_z f_1(x) - T_z f_2(x)| \\
&= \left|i\int_0^\infty \mathrm{e}^{itz}\left(\mathrm{e}^{it\int_1^\infty f_1(y)\kappa_\sigma(x,y)\mu_W(\mathrm{d}\,y)} - \mathrm{e}^{it\int_1^\infty f_2(y)\kappa_\sigma(x,y)\mu_W(\mathrm{d}\,y)}\right)\mathrm{d}\,t\right| \\
&\le \int_0^\infty \mathrm{e}^{-\eta t}\left|\mathrm{e}^{it\int_1^\infty f_1(y)\kappa_\sigma(x,y)\mu_W(\mathrm{d}\,y)} - \mathrm{e}^{it\int_1^\infty f_2(y)\kappa_\sigma(x,y)\mu_W(\mathrm{d}\,y)}\right|\mathrm{d}\,t \\
&\le \int_0^\infty \mathrm{e}^{-\eta t}t\left|\int_1^\infty (f_1(y) - f_2(y))\,\kappa_\sigma(x,y)\mu_W(\mathrm{d}\,y)\right|\mathrm{d}\,t,
\end{aligned} \tag{3.80}$$

Chapter Three

where in (3.80) we use (3.79). Now, evaluating the integral over t in (3.80), we obtain

$$\begin{aligned}
&|T_z f_1(x) - T_z f_2(x)| \\
&\leq \frac{(\tau-1)}{\eta^2} \int_1^\infty |f_1(y) - f_2(y)| \kappa_\sigma(x,y) y^{-\tau} \,\mathrm{d}\, y,
\end{aligned} \tag{3.81}$$

where we explicitly write down the Pareto law $\mu_W(\mathrm{d}\, y) := (\tau-1)y^{-\tau}\,\mathrm{d}\, y$. Recall that $\kappa_\sigma(x,y) = (x \wedge y)(x \vee y)^\sigma$. Thus, (3.81) becomes

$$\begin{aligned}
&|T_z f_1(x) - T_z f_2(x)| \\
&\leq \frac{\tau-1}{\eta^2}\Big(\int_1^x |f_1(y) - f_2(y)| x y^{\sigma-\tau} \,\mathrm{d}\, y + \int_x^\infty |f_1(y) - f_2(y)| x^\sigma y^{1-\tau} \,\mathrm{d}\, y \Big).
\end{aligned}$$

Integrating with respect to ν gives us

$$\begin{aligned}
&\|T_z f_1 - T_z f_2\|_1 \\
&\leq \frac{\tau-1}{\eta^2} \int_1^\infty \left(x \int_1^x |f_1(y) - f_2(y)| y^{\sigma-\tau} \,\mathrm{d}\, y \right) x^{-\beta} \,\mathrm{d}\, x \\
&\qquad + \frac{\tau-1}{\eta^2} \int_1^\infty \left(x^\sigma \int_x^\infty |f_1(y) - f_2(y)| y^{1-\tau} \,\mathrm{d}\, y \right) x^{-\beta} \,\mathrm{d}\, x \\
&= \frac{\tau-1}{\eta^2} \Bigg(\int_1^\infty |f_1(y) - f_2(y)| y^{\sigma-\tau} \int_y^\infty x^{1-\beta} \,\mathrm{d}\, x \,\mathrm{d}\, y \\
&\qquad\qquad + \int_1^\infty |f_1(y) - f_2(y)| y^{1-\tau} \int_1^y x^{\sigma-\beta} \,\mathrm{d}\, x \,\mathrm{d}\, y \Bigg) .
\end{aligned} \tag{3.82}$$

Using $\beta > 2$, the first integral in (3.82) can be bounded by

$$\begin{aligned}
&\int_1^\infty |f_1(y) - f_2(y)| y^{\sigma-\tau} \int_y^\infty x^{1-\beta} \,\mathrm{d}\, x \,\mathrm{d}\, y \\
&= c_1 \int_1^\infty |f_1(y) - f_2(y)| y^{-\beta} y^{2+\sigma-\tau} \,\mathrm{d}\, y \leq c_1 \|f_1 - f_2\|_1 ,
\end{aligned} \tag{3.83}$$

since $y^{2+\sigma-\tau} \leq 1$ and $c_1 = 1/(\beta-2)$. Similarly, the second integral in (3.82) gives us

$$\begin{aligned}
\int_1^\infty |f_1(y) - f_2(y)| y^{1-\tau} \int_1^y x^{\sigma-\beta} \,\mathrm{d}\, x \,\mathrm{d}\, y &\leq c_2 \int_1^\infty |f_1(y) - f_2(y)| y^{1-\tau} \,\mathrm{d}\, y \\
&\leq c_2 \|f_1 - f_2\|_1,
\end{aligned} \tag{3.84}$$

with $c_2 = 1/(\beta-1-\sigma)$, where for the last line we have used $1-\tau < -\beta$. Combining (3.83) and (3.84) in (3.82) gives us that

$$\|T_z f_1 - T_z f_2\|_1 \leq \frac{\tilde{c}}{\eta^2} \|f_1 - f_2\|_1 , \tag{3.85}$$

where $\tilde{c}$ is a constant depending on τ, σ and β. Thus, taking $\eta > 0$ to be sufficiently large such that $\eta > \sqrt{\tilde{c}}$ gives us that T_z is a contraction mapping on $\mathcal{B}$, hence proving the result. $\square$

The following corollary is immediate from the Banach fixed-point theorem for contraction mappings.

Corollary 3.8.8.
Let $T_z : \mathcal{B} \to \mathcal{B}$ be the contraction map given in (3.78). *Then, there exists a unique analytic function $a_z^* \in \mathcal{B}$ such that $T_z(a_z^*) = a_z^*$.*

We know from (3.74) that

$$a_z(x) = i \int_0^\infty \mathrm{e}^{itz} \exp\left\{ it \int_1^\infty c_m^{-1} a_z(y) \kappa_\sigma(x,y) \mathbb{1}_{\{y \le m\}} \mu_W(\mathrm{d}\, y) \right\} \mathrm{d}\, t\,. \tag{3.86}$$

Define $\tilde{a}_z$ as

$$\tilde{a}_z(x) = i \int_0^\infty \mathrm{e}^{itz} \exp\left\{ it \int_1^\infty c_m^{-1} a_z(y) \kappa_\sigma(x,y) \mu_W(\mathrm{d}\, y) \right\} \mathrm{d}\, t\,. \tag{3.87}$$

Then, $\tilde{a}_z = T_z(c_m^{-1} a_z)$. We now have the following lemma.

Lemma 3.8.9.
Let a_z and $\tilde{a}_z$ be as in (3.86) *and* (3.87), *respectively. Then,*

$$\|a_z - \tilde{a}_z\|_1 \le \frac{C(m)}{\eta^3}\,,$$

where $C(m)$ is a constant depending on m such that $\lim_{m\to\infty} C(m) = 0$.

Proof of Lemma 3.8.9. Since $a_z \in \mathcal{B}$, we again use (3.79) to get

$$\begin{aligned} |a_z(x) - \tilde{a}_z(x)| &\le \int_0^\infty \mathrm{e}^{-\eta t} t \left| \int_m^\infty c_m^{-1} a_z(y) \kappa_\sigma(x,y) \mu_W(\mathrm{d}\, y) \right| \mathrm{d}\, t \\ &\le \frac{\tau - 1}{c_m \eta^2} \int_m^\infty |a_z(y)| \kappa_\sigma(x,y) y^{-\tau} \,\mathrm{d}\, y\,, \end{aligned} \tag{3.88}$$

where we evaluate the integral over t to get the factor of η^{-2} in (3.88). Recall that $c_m = 1 - m^{-(\tau-1)}$. Using (3.75), we have that

$$|a_z(x) - \tilde{a}_z(x)| \le \frac{\tau - 1}{c_m \eta^3} \int_m^\infty \kappa_\sigma(x,y) y^{-\tau} \,\mathrm{d}\, y. \tag{3.89}$$

Since $\kappa_\sigma(x,y) \le (xy)^{1\vee\sigma}$, we have

$$|a_z(x) - \tilde{a}_z(x)| \le \frac{\tau - 1}{c_m \eta^3} x^{1\vee\sigma} \int_m^\infty y^{(1\vee\sigma)-\tau} \,\mathrm{d}\, y = \frac{(\tau-1) m^{(1\vee\sigma)-(\tau-1)}}{c_m((\tau-1)-(1\vee\sigma))\eta^3} x^{1\vee\sigma}, \tag{3.90}$$

Chapter Three

where we use the fact that $\tau > \max(2, 1+\sigma)$, and so the integral evaluated in (3.90) is finite. Define

$$c(m) := \frac{(\tau-1)c_m^{-1} m^{(1\vee\sigma)-(\tau-1)}}{(\tau-1)-(1\vee\sigma)}.$$

Since c_m tends to one, and $m^{(1\vee\sigma)-(\tau-1)}$ tends to zero we have $c(m) = o_m(1)$. Now, integrating both sides of (3.90) against $x^{-\beta}\,\mathrm{d}\,x$ gives us

$$\|a_z - \tilde{a}_z\|_1 \leq \frac{c(m)}{\eta^3}\int_1^\infty x^{1\vee\sigma-\beta}\,\mathrm{d}\,x = \frac{C(m)}{\eta^3}\,, \tag{3.91}$$

since $\beta > 2\vee 1+\sigma$, and where $C(m) = o_m(1)$, completing the proof. □

We are now at the penultimate step, where we have the necessary tools to show the convergence of a_z to a_z^* in the space $\mathcal{B}$.

Lemma 3.8.10.
Let a_z^* be the unique fixed point of the contraction map T_z defined in (3.78). *Then, we have that*

$$\lim_{m\to\infty}\|a_z - a_z^*\|_1 = 0\,. \tag{3.92}$$

Proof of Lemma 3.8.10. We have, using Lemma 3.8.9 and the fact that T_z is a contraction, that

$$\begin{aligned}
\|a_z - a_z^*\|_1 &\leq \|a_z - \tilde{a}_z\|_1 + \|\tilde{a}_z - a_z^*\|_1 \\
&\leq C(m)\eta^{-3} + \|T_z(c_m^{-1}a_z) - T_z(a_z^*)\|_1 \\
&\leq C(m)\eta^{-3} + \tilde{c}\eta^{-2}\|c_m^{-1}a_z - a_z^*\|_1 \\
&\leq C(m)\eta^{-3} + \tilde{c}\eta^{-2}c_m^{-1}\|a_z - a_z^*\|_1 + \tilde{c}\eta^{-2}\|a_z^*\|_1|c_m^{-1} - 1|\,.
\end{aligned}$$

Thus, choosing $\eta > 0$ such that $0 < 1 - \tilde{c}c_m^{-1}\eta^{-2} < 1$, we have that

$$\|a_z - a_z^*\|_1 \leq \frac{1}{1 - C_\tau c_m^{-1}\eta^{-2}}\left(C(m)\eta^{-3} + C_\tau\eta^{-2}\|a_z^*\|_1|c_m^{-1} - 1|\right)\,. \tag{3.93}$$

Now, as $m\to\infty$, we have that $C(m)\to 0$, and $c_m \to 1$. Since $\|a_z^*\| < \infty$, we have that the right-hand side of (3.93) goes to 0 as $m\to\infty$. Thus, $\|a_z - a_z^*\|_1 \to 0$ as $m\to\infty$ for z in an appropriate domain $D_\eta \subset \mathbb{C}^+$. However, in the complex variable z, the domains of a_z and a_z^* are $\mathbb{C}^+$. Since the convergence holds for an open set of this domain (that is, in $D_\eta \subset \mathbb{C}^+$), by the identity theorem of complex analysis, the convergence holds everywhere in $\mathbb{C}^+$, that is, for each $z \in \mathbb{C}^+$. □

We now proceed towards a proof of Theorem 3.2.5, and to achieve this we wish to take the limit $m \to \infty$ to characterise $\mathrm{S}_{\mu_{\sigma,\tau}}$. We know that since $\lim_{m\to\infty} \mu_{\sigma,\tau,m} = \mu_{\sigma,\tau}$, then for each $z \in \mathbb{C}^+$, $\lim_{m\to\infty} \mathrm{S}_{\mu_{\sigma,\tau,m}}(z) = \mathrm{S}_{\mu_{\sigma,\tau}}(z)$.

Proof of Theorem 3.2.5. Let a_z^* be the unique fixed point of the contraction mapping T_z as in Corollary 3.8.8, and let $\mathrm{S}_{\mu_{\sigma,\tau}}(z)$ be the Stieltjes transform of $\mu_{\sigma,\tau}$ for any $z \in \mathbb{C}^+$. We wish to show that

$$\mathrm{S}_{\mu_{\sigma,\tau}}(z) = \int_1^\infty a_z^*(x)\mu_W(\mathrm{d}\, x).$$

We have that

$$\begin{aligned}
&\left|\int_1^\infty a_z(x)\mu_{W,m}(\mathrm{d}\, x) - \int_1^\infty a_z^*(x)\mu_W(\mathrm{d}\, x)\right| \\
&\le \left|\int_1^\infty a_z(x)\mu_{W,m}(\mathrm{d}\, x) - \int_1^\infty a_z^*(x)\mu_{W,m}(\mathrm{d}\, x)\right| \\
&\quad + \left|\int_1^\infty a_z^*(x)\mu_{W,m}(\mathrm{d}\, x) - \int_1^\infty a_z^*(x)\mu_W(\mathrm{d}\, x)\right|. \qquad (3.94)
\end{aligned}$$

The first term in (3.94) can be evaluated as

$$\begin{aligned}
&\left|\int_1^\infty a_z(x)\mu_{W,m}(\mathrm{d}\, x) - \int_1^\infty a_z^*(x)\mu_{W,m}(\mathrm{d}\, x)\right| \\
&\le (\tau-1)c_m^{-1}\int_1^m |a_z(x) - a_z^*(x)|x^{-\tau}\,\mathrm{d}\, x \\
&\le (\tau-1)c_m^{-1}\int_1^\infty |a_z(x) - a_z^*(x)|x^{-\beta}x^{\beta-\tau}\,\mathrm{d}\, x \\
&\le (\tau-1)c_m^{-1}\|a_z - a_z^*\|_1 = o_m(1), \qquad (3.95)
\end{aligned}$$

as $x^{\beta-\tau} \le 1$, and $\|a_z - a_z^*\|_1 = o_m(1)$ from Lemma 3.8.10. The second term of (3.94) can be evaluated as

$$\begin{aligned}
&\left|\int_1^\infty a_z^*(x)\mu_{W,m}(\mathrm{d}\, x) - \int_1^\infty a_z^*(x)\mu_W(\mathrm{d}\, x)\right| \\
&\le c_m^{-1}\left|\int_1^m a_z^*(x)\mu_W(\mathrm{d}\, x) - \int_1^\infty a_z^*(x)\mu_W(\mathrm{d}\, x)\right| + \left|\int_1^\infty a_z^*(x)\mu_W(\mathrm{d}\, x)\right| |c_m^{-1} - 1| \\
&\le \frac{(\tau-1)}{c_m\eta}\int_m^\infty x^{-\tau}\,\mathrm{d}\, x + \frac{|c_m^{-1}-1|}{\eta} = \frac{(\tau-1)m^{1-\tau}}{c_m\eta} + \frac{|c_m^{-1}-1|}{\eta} = o_m(1), \qquad (3.96)
\end{aligned}$$

since $|a_z^*| \le \eta^{-1}$. Combining (3.95) and (3.96) completes the proof of the theorem. $\square$

Chapter Three

CHAPTER 4

Scale-free percolation: The graph Laplacian

This chapter is based on:
R.S. Hazra, N. Malhotra. *Spectral properties of the Laplacian of scale-free percolation models.* [*arxiv:2504.17552*], 2025.

Abstract

We consider scale free percolation on a discrete torus $\mathbf{V}_N$ of size N. Conditionally on an i.i.d. sequence of Pareto weights $(W_i)_{i \in \mathbf{V}_N}$ with tail exponent $\tau - 1 > 0$, we connect any two points i and j on the torus with probability

$$p_{ij} = \frac{W_i W_j}{\|i - j\|^{\alpha}} \wedge 1$$

for some parameter $\alpha > 0$. We focus on the (centered) Laplacian operator of this random graph and study its empirical spectral distribution. We explicitly identify the limiting distribution when $\alpha < 1$ and $\tau > 3$, in terms of the spectral distribution of some non-commutative unbounded operators.

§4.1 Introduction

In recent years, many random graph models have been proposed to model real-life networks. These models aim to capture three key properties that real-world networks exhibit: scale-free nature of the degree distribution, small-world property, and high clustering coefficients [van der Hofstad, 2024]. It is generally difficult to find random graph models which incorporate all three features. Classical random graph models typically fail to capture scale-freeness, small-world behaviour, and high clustering simultaneously. For instance, the Erdős-Rényi model only exhibits the small-world property, while models like Chung-Lu, Norros-Reittu, and preferential attachment models are scale-free (Chung and Lu [2002], Barabási and Albert [1999] and small-world but have clustering coefficients that vanish as the network grows. In contrast, regular lattices have high clustering but large typical distances. The Watts-Strogatz model (Watts and Strogatz [1998]) was an early attempt to create a network with high clustering and small-world features, but it does not produce scale-free degree distributions.

Scale-free percolation, introduced in Deijfen et al. [2013], blends ideas from long-range percolation (see e.g. Berger [2002]) with inhomogeneous random graphs such as the Norros–Reittu model. In this framework, vertices are positioned on the $\mathbb{Z}^d$ lattice, and each vertex x is independently assigned a random weight W_x. These weights follow a power-law distribution:

$$\mathbb{P}(W > w) = w^{-(\tau-1)} L(w),$$

where $\tau > 1$ and $L(w)$ is a slowly varying function at infinity.

Edges between pairs of vertices x and y are added independently, with a probability that increases with the product of their weights and decreases with their Euclidean distance. The edge probability is given by

$$p_{xy} = 1 - \exp\left(-\lambda \frac{W_x W_y}{\|x-y\|^\alpha}\right), \tag{4.1}$$

where $\lambda, \alpha > 0$ are model parameters and $\|\cdot\|$ denotes the Euclidean norm. This model has been proposed as a suitable representation for certain real-world systems, such as interbank networks, where both spatial structure and heavy-tailed connectivity distributions are relevant (Deprez et al. [2015]). Various properties of the model are now well known and we refer to the articles by Jorritsma et al. [2024], Cipriani and Salvi [2024], Cipriani et al. [2025], Heydenreich et al. [2017], Dalmau and Salvi [2021] for further references.

In recent times, there has been a lot of interest in the models which have connection probabilities similar to (4.1). Kernel-based spatial random graphs encompass a wide variety of classical random graph models where vertices are embedded in some metric space. In their simplest form (see Jorritsma et al. [2023] for a more complete exposition) they can be defined as follows. Let V be the vertex set of the graph and sample a collection of weights $(W_i)_{i\in V}$, which are independent and identically distributed (i.i.d.), serving as marks on the vertices. Conditionally on the weights, two vertices i and j are connected by an undirected edge with probability

$$\mathbb{P}\left(i \leftrightarrow j \mid W_i, W_j\right) = \kappa(W_i, W_j)\|i-j\|^{-\alpha} \wedge 1\,, \tag{4.2}$$

where κ is a symmetric kernel, $\|i-j\|$ denotes the distance between the two vertices in the underlying metric space and $\alpha > 0$ is a constant parameter. In a recent article, the present authors with A. Cipriani and M. Salvi (Cipriani et al. [2025]) proved the spectral properties of the adjacency matrix when $\alpha < d$ and the weights have a finite mean. In the above setting, the model was considered on a torus of side length N so that the adjacency operator as a matrix was easier to describe. In this article, we aim to extend this study to the case of a Laplacian matrix. Although our approach would extend to general kernel-based models, we shall stick to the product form kernel, that is, $\kappa(x,y) = xy$, so that the ideas can be clearly presented. It is one of the few cases where the limiting distribution can be explicitly described.

The Laplacian of a graph with N vertices is defined as $A_N - D_N$ where A_N is the adjacency matrix and D_N is the diagonal matrix where the i-th diagonal entry corresponds to the i-th degree. When the entries of the matrix are not restricted to 0 or 1, the matrix is also referred to as the *Markov matrix* (Bryc et al. [2006], Bordenave et al. [2014]). The graph Laplacian serves as the discrete analogue of the continuous Laplacian, essential in diffusion theory and network flow analysis. The Laplacian matrix has several key applications: The Kirchhoff Matrix-Tree Theorem relates the determinant of the Laplacian to the count of spanning trees in a graph (Chung [1997]), the multiplicity of the zero eigenvalue indicates the number of connected components (Chung [1997]), the second-smallest eigenvalue, known as the Fiedler value or the algebraic connectivity, measures the graph's connectivity; higher values signify stronger connectivity De Abreu [2007]. For a comprehensive overview of spectral methods in graph theory, refer to Chung's monograph Chung [1997] and Spielman's book Spielman [2012]. In modern machine learning, spectral techniques are pivotal in spectral clustering algorithms, where the techniques use the Laplacian eigenvalues and eigenvectors for dimensionality reduction before applying clustering algorithms like k-means (Abbe et al. [2020], Abbe [2017]). It is particularly effective for detecting clusters that are not linearly separable. Recent advancements integrate spectral clustering with graph neural networks to enhance graph pooling operations (Bianchi et al. [2020]). Spectral algorithms are also crucial for identifying communities within networks by analysing the spectral properties of the graph (Chung [1997]).

Bryc et al. [2006] established that for large symmetric matrices with independent and identically distributed entries, the empirical spectral distribution (ESD) of the corresponding Laplacian matrix converges to the free convolution of the semicircle law and the standard Gaussian distribution. In the context of sparse Erdős–Rényi graphs, Huang and Landon [2020] studied the local law of the ESD of the Laplacian matrix. They demonstrated that the Stieltjes transform of the ESD closely approximates that of the free convolution of the semicircle law and a standard Gaussian distribution down to the scale N^{-1}. Additionally, they showed that the gap statistics and averaged correlation functions align with those of the Gaussian Orthogonal Ensemble in the bulk. Ding and Jiang [2010] investigated the spectral distributions of adjacency and Laplacian matrices of random graphs, assuming that the variance of the entries of an $N \times N$ adjacency matrix depends only on N. They established the convergence of the ESD of these matrices under such conditions. These results of the Erdős-Rényi random graphs were extended to the inhomogeneous setting by Chakrabarty et al. [2021b]. In a recent work, Chatterjee and Hazra [2022] derived a combinatorial way to describe the limiting moments for a wide variety of random matrix models with a variance profile.

Our contribution

The empirical spectral distribution of the (centred) Laplacian of a graph that incorporates spatial distance has not been studied before. For example, we are not aware of a result that describes the spectral properties of the Laplacian for the long-range percolation model. Our main contribution is that we establish this result for the scale-free percolation model on the torus. We restrict ourselves to the dense regime, that is, $\alpha < 1$. We show that under mild assumptions on the weights (having finite variance), we establish the existence of the limiting distribution. It turns out to be a distribution that involves the Gaussian, the semicircle, and the Pareto distribution. In a symbolic (and rather informal) way, it is given by the spectral law of

$$W^{1/2} S W^{1/2} + m_1 W^{1/4} G W^{1/4},$$

where W is an unbounded operator with spectral law given by the Pareto distribution, S is a bounded compact operator whose spectral law is the semicircle law, and G is an unbounded operator whose law is given by the Gaussian distribution. Finally, m_1 is the first moment of W. The interaction between these operators comes from the fact that in the non-commutative space, $\{W, G\}$ is a commutative algebra, freely independent of S. Similar results have been established when the weights are bounded and degenerate, and no spatial distances are involved (Chatterjee and Hazra [2022] and Chakrabarty et al. [2021b]). The present work extends the results to settings that involve random heavy-tailed weights and spatial distances.

Outline of the article

In section 4.2 we explicitly describe the setup of the model and state our main results. In Theorem 4.2.1 we show the existence of the limiting spectral distribution, and in Theorem 4.2.5, we identify the measure and state some of its properties. In Section 4.3 we first introduce a Gaussianised version of the matrix, and this helps us to simplify the variance profile. We then truncate the weights and decouple the diagonal, which allows us to apply the moment method. In Section 4.4 we identify the limiting moments of the decoupled Laplacian and show that it does not depend on the spatial parameter $\alpha > 0$, which is crucial in the identification of the limiting measure of the original Laplacian. Finally, in Section 4.5 we identify the limiting measure using results from free probability. In Appendix 4.6 we provide references to some of the results we use in our proofs, which are collections of results from other articles and are stated here for completeness.

§4.2 Setup and main results

In this section we describe the setup of the model and also state the main results.

§4.2.1 Setup

(a) **V**ertex set: the vertex set is $\mathbf{V}_N := \{1, 2, \ldots, N\}$. The vertex set is equipped with torus the distance $\|i-j\|$, where

$$\|i-j\| = |i-j| \wedge (N - |i-j|).$$

(b) **Weights:** the weights $(W_i)_{i\in\mathbf{V}_N}$ are i.i.d. random variables sampled from a Pareto distribution W (whose law we denote by $\mathbf{P}$) with parameter $\tau - 1$, where $\tau > 1$. That is,

$$\mathbf{P}(W > t) = t^{-(\tau-1)}\mathbf{1}_{\{t\geq 1\}} + \mathbf{1}_{\{t<1\}}. \tag{4.3}$$

(c) **Long-range parameter:** $\alpha > 0$ is a parameter which controls the influence of the distance between vertices on their connection probability.

(d) **Connectivity function:** conditional on the weights, each pair of distinct vertices i and j is connected independently with probability $P^W(i \leftrightarrow j)$ given by

$$P^W(i \leftrightarrow j) := \mathbb{P}(i \leftrightarrow j \mid W_i, W_j) = \frac{W_i W_j}{\|i-j\|^\alpha} \wedge 1. \tag{4.4}$$

Chapter Four

We will be using the short-hand notation $p_{ij} := \mathbb{P}(i \leftrightarrow j \mid W_i, W_j)$ for convenience. Note that the graph does not have self-loops.

In what follows, we denote by $\mathbb{P} = \mathbf{P} \otimes P^W$ the joint law of the weights and the edge variables. Note that $\mathbb{P}$ depends on N, but we will omit this dependence for simplicity. Let $\mathbb{E}, \mathbf{E}$, and E^W denote the expectation with respect to $\mathbb{P}, \mathbf{P}$, and P^W respectively.

The associated graph is connected, as nearest neighbours with respect to the torus distance are always linked. Let us denote the random graph generated by our choice of edge probabilities by $\mathbb{G}_N$. Let $\mathbb{A}_{\mathbb{G}_N}$ denote the adjacency matrix (operator) associated with this random graph, defined as

$$\mathbb{A}_{\mathbb{G}_N}(i,j) = \begin{cases} 1 & \text{if } i \leftrightarrow j, \\ 0 & \text{otherwise.} \end{cases}$$

Since the graph is finite and undirected, the adjacency matrix is always self-adjoint and has real eigenvalues. Let

$$\mathbb{D}_{\mathbb{G}_N} = \mathrm{Diag}(d_1, \cdots, d_N)$$

where d_i denotes the degree of the vertex i and in this case given by

$$d_i = \sum_{j \neq i} \mathbb{A}_{\mathbb{G}_N}(i,j).$$

We will consider the Laplacian of the matrix, which is denoted as follows:

$$\Delta_{\mathbb{G}_N} = \mathbb{A}_{\mathbb{G}_N} - \mathbb{D}_{\mathbb{G}_N}.$$

In general, when $\alpha < 1$, the eigenvalue distribution requires scaling in order to observe meaningful limiting behaviour. In Cipriani et al. [2025], it was shown that an appropriate scaling of the adjacency matrix, under which the convergence of the bulk eigenvalue distribution can be studied, is given by

$$c_N = \frac{1}{N} \sum_{i \neq j \in \mathbf{V}_N} \frac{1}{\|i-j\|^\alpha} \sim c_0 N^{1-\alpha}, \tag{4.5}$$

where c_0 is a positive constant. The scaled adjacency matrix is then defined as

$$\mathbf{A}_N := \frac{\mathbb{A}_{\mathbb{G}_N}}{\sqrt{c_N}}. \tag{4.6}$$

We define the corresponding (scaled) Laplacian as

$$\mathbf{\Delta}_N = \mathbf{A}_N - \mathbf{D}_N,$$

where $\mathbf{D}_N$ is given by $\mathbf{D}_N = \mathrm{Diag}(\mathbf{d}_1, \cdots, \mathbf{d}_N)$ with

$$\mathbf{d}_i = \sum_{k \neq i} \mathbf{A}_N(i,k).$$

The empirical measure that assigns a mass of $1/N$ to each eigenvalue of the $N \times N$ random matrix $\mathbf{M}_N$ is called the Empirical Spectral Distribution (ESD) of $\mathbf{M}_N$, denoted as

$$\mathrm{ESD}\,(\mathbf{M}_N) := \frac{1}{N} \sum_{i=1}^{N} \delta_{\lambda_i},$$

where $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_N$ are the eigenvalues of $\mathbf{M}_N$. We are interested in the centred Laplacian matrix for the bulk distribution. So define

$$\mathbf{\Delta}_N^\circ = \mathbf{\Delta}_N - \mathbb{E}[\mathbf{\Delta}_N] \tag{4.7}$$

where $\mathbb{E}[\mathbf{\Delta}_N](i,j) = \mathbb{E}[\mathbf{\Delta}_N(i,j)]$. If we define $\mathbf{A}_N^\circ = \mathbf{A}_N - \mathbb{E}[\mathbf{A}_N]$ and $\mathbf{D}_N^\circ$ is the diagonal matrix $\mathrm{Diag}(\mathbf{d}_1^\circ, \ldots, \mathbf{d}_N^\circ)$ where $\mathbf{d}_i^\circ = \sum_{k \neq i} \mathbf{A}_N^\circ(i,k)$, then it is easy to see that

$$\mathbf{\Delta}_N^\circ = \mathbf{A}_N^\circ - \mathbf{D}_N^\circ.$$

In this article we will be interested in understanding the behaviour of $\mathrm{ESD}(\mathbf{\Delta}_N^\circ)$ as $N \to \infty$.

§4.2.2 Main Results

The Lévy-Prokhorov distance $d_L : \mathcal{P}(\mathbb{R})^2 \to [0, +\infty)$ between two probability measures μ and ν on $\mathbb{R}$ is defined as

$$d_L(\mu, \nu) := \inf \big\{ \varepsilon > 0 \mid \mu(A) \leq \nu\,(A^\varepsilon) + \varepsilon \text{ and } \nu(A) \leq \mu\,(A^\varepsilon) + \varepsilon \quad \forall\, A \in \mathcal{B}(\mathbb{R}) \big\},$$

where $\mathcal{B}(\mathbb{R})$ denotes the Borel σ-algebra on $\mathbb{R}$, and A^ε is the ε-neighbourhood of A. For a sequence of random probability measures $(\mu_N)_{N \geq 0}$, we say that

$$\lim_{N \to \infty} \mu_N = \mu_0 \text{ in } \mathbb{P}\text{-probability}$$

if, for every $\varepsilon > 0$,

$$\lim_{N \to \infty} \mathbb{P}(d_L(\mu_N, \mu_0) > \varepsilon) = 0.$$

Our first main result is existential and is as follows.

Theorem 4.2.1.
Consider the random graph $\mathbb{G}_N$ on $\mathbf{V}_N$ with connection probabilities given by (4.4) with parameters $\tau > 3$ and $0 < \alpha < 1$. Let $\mathrm{ESD}(\mathbf{\Delta}_N^\circ)$ be the empirical spectral distribution of $\mathbf{\Delta}_N^\circ$ defined in (4.7). Then there exists a deterministic measure ν_τ on $\mathbb{R}$ such that

$$\lim_{N \to \infty} \mathrm{ESD}(\mathbf{\Delta}_N^o) = \nu_\tau \qquad \textit{in } \mathbb{P}\textit{–probability}.$$

Chapter Four

The characterisation of ν_τ is achieved by results from the theory of free probability. For convenience, we state some technical definitions. We refer the readers to [Anderson et al., 2010, Chapter 5] for further details.

For the following definitions, we refer the reader to Mingo and Speicher [2017], and recall from Chapter 1 that a W^*-algebra is a C^*-algebra of bounded operators on a Hilbert space closed in the weak operator topology.

Definition 4.2.2.
Let $(\mathcal{A}, \varphi)$ be a W^*-probability space, where $\mathcal{A}$ is a W^*-algebra, and φ is a faithful, tracial state. *A densely defined operator T is said to be affiliated with $\mathcal{A}$ if for every bounded measurable function h, we have $h(T) \in \mathcal{A}$. The law (or distribution) $\mathcal{L}(T)$ of such an affiliated operator T is the unique probability measure on $\mathbb{R}$ satisfying*

$$\varphi(h(T)) = \int_{\mathbb{R}} h(x)\, d\mathcal{L}(T)(x).$$

For a collection of self-adjoint operators $T_1, \ldots, T_n$, their joint distribution is described by specifying

$$\varphi(h_1(T_{i_1}) \ldots h_k(T_{i_k})),$$

for all $k \geq 1$, all index sequences $i_1, \ldots, i_k \in \{1, \ldots, n\}$, and all bounded measurable functions $h_1, \ldots, h_k : \mathbb{R} \to \mathbb{R}$.

Definition 4.2.3.
Let $(\mathcal{A}, \varphi)$ be a W^*-probability space, and suppose $a_1, a_2 \in \mathcal{A}$. Then a_1 and a_2 are said to be freely independent if

$$\varphi(p_1(a_{i_1}) \ldots p_n(a_{i_n})) = 0,$$

for every $n \geq 1$, every sequence $i_1, \ldots, i_n \in \{1, 2\}$ with $i_j \neq i_{j+1}$ for all $j = 1, \ldots, n-1$, and all polynomials $p_1, \ldots, p_n$ in one variable satisfying

$$\varphi(p_j(a_{i_j})) = 0, \quad \text{for all } j = 1, \ldots, n.$$

Definition 4.2.4.
Let $a_1, \ldots, a_k$ and $b_1, \ldots, b_m$ be operators affiliated with $\mathcal{A}$. The families $(a_1, \ldots, a_k)$ and $(b_1, \ldots, b_m)$ are freely independent if and only if

$$p(h_1(a_1) \ldots h_k(a_k)) \quad \text{and} \quad q(g_1(b_1) \ldots g_m(b_m))$$

are freely independent for all bounded measurable functions $h_1, \ldots, h_k$ and $g_1, \ldots, g_m$, and for all polynomials p and q in k and m non-commutative variables, respectively.

We are now ready to state our second main result.

Theorem 4.2.5.
Under the assumptions of Theorem 4.2.1, the limiting measure ν_τ can be identified as

$$\nu_\tau = \mathcal{L}\left(T_W^{1/2} T_s T_W^{1/2} + \mathbf{E}[W] T_W^{1/4} T_g T_W^{1/4}\right).$$

Here, T_g and T_W are commuting self-adjoint operators affiliated with a W^*-probability space $(\mathcal{A}, \varphi)$ such that, for bounded measurable functions h_1, h_2 from $\mathbb{R}$ to itself,

$$\varphi\left(h_1\left(T_g\right) h_2\left(T_W\right)\right) = \left(\int_{-\infty}^{\infty} h_1(x)\phi(x)dx\right)\left(\int_1^{\infty} h_2(u)(\tau-1)u^{-\tau}du\right)$$

with ϕ the standard normal density. Furthermore, T_s has a standard semicircle law and is freely independent of (T_g, T_W).

In particular, when W is degenerate, say $W \equiv 1$, then ν_τ is given by the free additive convolution of semicircle and Gaussian law.

§4.2.3 Discussion and simulations

(a) We now briefly describe the main steps of the proof.

1. Gaussianisation: In the first step, we show that replacing the Bernoulli entries with Gaussian entries having the same mean and variance results in empirical spectral distributions that are close.
2. Simplification of the variance profile: In this step, we show that the variance profile can be simplified to $W_i W_j / \|i-j\|^\alpha$, effectively removing the truncation at 1.
3. Truncation: Here, we show that in the Gaussian matrix, the weights W_i can be replaced by the truncated weights $W_i^m = W_i \mathbf{1}_{W_i \le m}$.
4. Decoupling the diagonal: In this step, we show that the Laplacian can be viewed as the sum of two independent random matrices (conditionally on the weights). Thus, we replace the diagonal matrix D_N with an independent copy Y_N, which has the same variance profile.
5. Moment method: With truncated weights and decoupled matrices, we apply the moment method to show convergence of the empirical spectral distribution and identify the limiting moments. A key observation is that the limiting measure and the method are independent of α, so the results remain valid even when $\alpha = 0$.

6. Identification of the limiting measure: Finally, we first identify the limiting measure in the case of truncated weights. These are typically associated with bounded operators (except in the Gaussian case). We then use techniques from Bercovici and Voiculescu [1993] to remove the truncation and identify the limiting measure in the general case.

(b) We now present some simulations that illustrate how the proof outline aligns with a specific value of α. In Figure 4.1, we plot the eigenvalue distribution of the centred Laplacian matrix, with the parameter range $N = 6000$, $\alpha = 0.5$ and $\tau = 4.1$. A crucial step in the proof of Theorem 4.2.1 requires us to replace the Bernoulli entries with Gaussian entries with the same variance profile. Also in the Gaussian case, we can simplify the variance to the following form:

$$\frac{W_i W_j}{\|i-j\|^{\alpha}}$$

for any $(i,j)^{\text{th}}$ entry. We compare the two spectra in Figure 4.2. We also consider the Gaussianised Laplacian matrix with a decoupled diagonal, and in Section 4.5, we apply an idea used in Cipriani et al. [2025], where we take $\alpha = 0$. We also compare the spectrum of this matrix to the original centred Laplacian in Figure 4.2. We see that the spectra are quite similar.

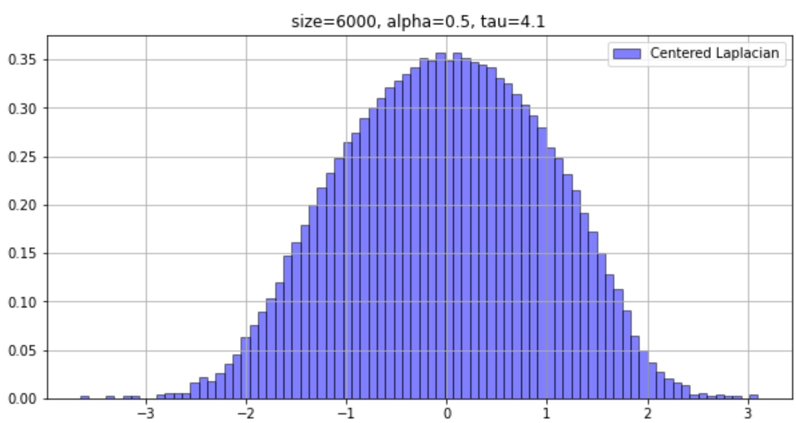


Figure 4.1: Spectrum of the centred Laplacian matrix .

(c) We remark that our results can be extended in two directions. Although we state and prove them for the case $d = 1$ and $\alpha < 1$, they naturally generalise to any dimension $d \geq 1$ and $\alpha < d$. In that case, the scaling

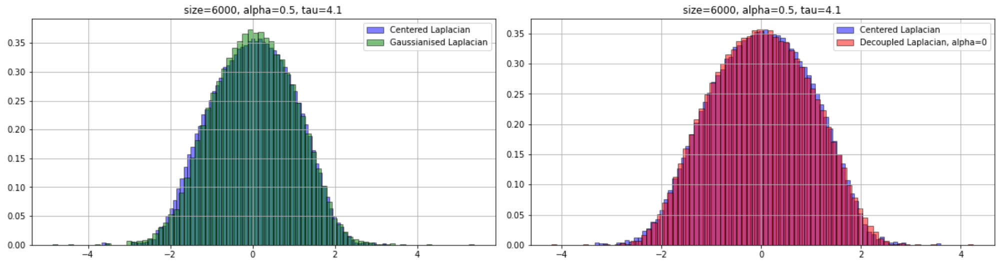


Figure 4.2: Comparing the spectrum of the centred Laplacian with the Gaussianised and the decoupled case.

constant requires an adjustment, with $c_N \sim c_0(d)\, N^{d-\alpha}$. For ease of presentation, we restrict ourselves to $d = 1$ in this work.

Another possible extension of our first result involves modifying the connection probabilities between vertices i and j to

$$p_{ij} = \frac{\kappa_\sigma(W_i, W_j)}{\|i-j\|^\alpha} \wedge 1,$$

where $\kappa_\sigma(x, y) = (x \vee y)(x \wedge y)^\sigma$. In this setting, we additionally assume $0 < \sigma < (\tau - 1)$. Such extensions have been studied in the context of adjacency matrices in Cipriani et al. [2025]. We strongly believe that in this case the limiting spectral distribution will exist, but it would be challenging to identify the limiting measure.

§4.2.4 Notation

We will use the Landau notation o_N, O_N indicating in the subscript the variable under which we take the asymptotic (typically this variable will grow to infinity unless otherwise specified). Universal positive constants are denoted as c, c_1, ..., and their value may change with each occurrence. For an $N \times N$ matrix $\mathbf{A} = (a_{ij})_{i,j=1}^N$ we use $\mathrm{Tr}(\mathbf{A}) := \sum_{i=1}^N a_{ii}$ for the trace and $\mathrm{tr}(\mathbf{A}) := N^{-1}\,\mathrm{Tr}(\mathbf{A})$ for the normalised trace. When $n \in \mathbb{N}$ we write $[n] := \{1, 2, \ldots, n\}$. We denote the cardinality of a set A as $\#A$, and, with a slight abuse of notation, $\#\sigma$ also denotes the number of cycles in a permutation σ.

§4.3 Gaussianisation and setup for main proofs

To prove Theorem 4.2.1, we construct a Laplacian matrix with truncated weights and a simplified variance profile, with the diagonal decoupled from the adjacency matrix. We follow the ideas of Cipriani et al. [2025], albeit with a slightly modified approach, as follows:

(a) We begin by Gaussianising the matrix $\mathbf{\Delta}_N^o$ to obtain a matrix $\bar{\mathbf{\Delta}}_N$, using the ideas of Chatterjee [2005]. Since we have $\tau > 3$, the proof proceeds without the need to truncate the weight sequence $\{W_i\}_{i\in\mathbf{V}_N}$.

(b) We then tweak the entries of $\bar{\mathbf{\Delta}}_N$ further through a series of lemmas to obtain the Laplacian matrix $\widehat{\mathbf{\Delta}}_{N,g}$, whose corresponding adjacency has mean-zero Gaussian entries and a simplified variance profile.

(c) Next, we truncate the weights $\{W_i\}_{i\in\mathbf{V}_N}$ at $m \geq 1$, and construct the corresponding matrix $\mathbf{\Delta}_{N,g,m}$. We show that, in $\mathbb{P}$-probability, the Lévy distance vanishes in the iterated limit $m \to \infty$ and $N \to \infty$.

(d) We conclude by decoupling the diagonal of the matrix $\mathbf{\Delta}_{N,g,m}$ from the off-diagonal terms. This follows from classical results used in studying the spectrum of Laplacian matrices.

§4.3.1 Gaussianisation

Suppose $(G_{i,j})_{i>j}$ is a sequence of i.i.d. $\mathrm{N}(0,1)$ random variables and independent of the sequence $(W_i)_{i\in\mathbf{V}_N}$. Define

$$\bar{\mathbf{A}}_N = \begin{cases} \frac{\sqrt{p_{ij}(1-p_{ij})}}{\sqrt{c_N}} G_{i\wedge j, i\vee j} + \frac{\mu_{ij}}{\sqrt{c_N}} & i \neq j \\ 0 & i = j, \end{cases}$$

where $\mu_{ij} = p_{ij} - \mathbb{E}[p_{ij}]$. Let $\bar{\mathbf{\Delta}}_N$ be the corresponding Laplacian of the matrix $\bar{\mathbf{A}}_N$. Let h be a 3 times differentiable function on $\mathbb{R}$ such that

$$\max_{0\leq k\leq 3} \sup_{x\in\mathbb{R}} |h^{(k)}(x)| < \infty \,,$$

where $h^{(k)}$ is the k-th derivative of h. Define the resolvent of the $N \times N$ matrix $\mathbf{M}_N$ as

$$R_{M_N}(z) = \left(\mathbf{M}_N - z\,\mathrm{I}_N\right)^{-1}, \qquad z \in \mathbb{C}^+,$$

where I_N is the $N \times N$ identity matrix and $\mathbb{C}^+$ is the upper-half complex plane. Further, define $H_z(\mathbf{M}_N) = \mathrm{S}_{M_N}(z) = \mathrm{tr}(R_{M_N}(z))$ for $z \in \mathbb{C}^+$.

Lemma 4.3.1 (Gaussianisation of $\mathbf{\Delta}_N$).
Consider $\bar{\mathbf{\Delta}}_N$ and $\mathbf{\Delta}_N^o$ defined as above. Then for any h as above,

$$\lim_{N\to\infty} \left|\mathbb{E}[h(\Re H_z(\bar{\mathbf{\Delta}}_N))] - \mathbb{E}[h(\Re H_z(\mathbf{\Delta}_N^o))]\right| = 0 \,,$$

and

$$\lim_{N\to\infty} \left|\mathbb{E}[h(\Im H_z(\bar{\mathbf{\Delta}}_N))] - \mathbb{E}[h(\Im H_z(\mathbf{\Delta}_N^o))]\right| = 0\,.$$

The proof is very similar to the one presented in Chatterjee and Hazra [2022] and is modified along the lines of Cipriani et al. [2025]. It uses the classical result of Chatterjee [2005], and we only give a brief sketch by showing the estimates of the error probabilities in this setting. In Cipriani et al. [2025], the Gaussianisation was done with truncated weights, but here we will not need that.

Proof. Following the proof of Cipriani et al. [2025] for the Laplacian, we define, conditional on the weights $(W_i)_{i\in\mathbf{V}_N}$, a sequence of independent random variables. Let $\mathbf{X}_b = (X_{ij}^b)_{1\le i<j\le N}$ be a vector with $X_{ij}^b \sim \mathrm{Ber}(p_{ij}) - \mathbb{E}[p_{ij}]$. Similarly, take another vector $\mathbf{X}_g = (X_{ij}^g)_{1\le i<j\le N}$ with $X_{ij}^g \sim \mathrm{N}\left(\mu_{ij}\,, p_{ij}(1-p_{ij})\right)$.

Let $n = N(N-1)/2$ and $\mathbf{x} = (x_{ij})_{1\le i<j\le N} \in \mathbb{R}^n$. Define $\mathrm{R}(\mathbf{x})$ to be the matrix-valued differentiable function given by

$$\mathrm{R}(\mathbf{x}) \coloneqq (\mathbf{M}_N(\mathbf{x}) - z\,\mathrm{I}_N)^{-1},$$

where $\mathbf{M}_N(\cdot)$ is the matrix-valued differentiable function that maps a vector in $\mathbb{R}^n$ to the space of $N\times N$ Hermitian matrices, given by

$$\mathbf{M}_N(\mathbf{x})_{ij} = \begin{cases} c_N^{-1/2} x_{ij} & \text{if } i<j, \\ c_N^{-1/2} x_{ji} & \text{if } i>j, \\ -c_N^{-1/2}\sum_{k\neq i} x_{ik} & \text{if } i=j. \end{cases}$$

Then, we see that $\mathbf{\Delta}_N^o = \mathbf{M}_N(\mathbf{X}_b)$ and $\bar{\mathbf{\Delta}}_N = \mathbf{M}_N(\mathbf{X}_g)$. Note that

$$E^W[X_{ij}^b] = E^W[X_{ij}^g] = \mu_{ij},$$

and

$$E^W[(X_{ij}^b)^2] = E^W[(X_{ij}^g)^2] = p_{ij} + \mathbb{E}[p_{ij}]^2 - 2p_{ij}\mathbb{E}[p_{ij}].$$

Consequently, using [Chatterjee, 2005, Theorem 1.1] we have that

$$\begin{aligned}
&\left|\mathbb{E}[h(\Re H_z(\bar{\mathbf{\Delta}}_N))] - \mathbb{E}[h(\Re H_z(\mathbf{\Delta}_N^o))]\right| \\
&= \left|\mathbf{E}\left[E^W[h(\Re H_z(\bar{\mathbf{\Delta}}_N)) - h(\Re H_z(\mathbf{\Delta}_N^o))]\right]\right| \\
&\le C_1(h)\lambda_2(H) \sum_{1\le i<j\le N} \mathbb{E}[(X_{ij}^b)^2 \mathbf{1}_{|X_{ij}^b|>K_N}] + \mathbb{E}[(X_{ij}^g)^2 \mathbf{1}_{|X_{ij}^g|>K_N}] && (4.8) \\
&+ C_2(h)\lambda_3(H) \sum_{1\le i<j\le N} \mathbb{E}[(X_{ij}^b)^3 \mathbf{1}_{|X_{ij}^b|\le K_N}] + \mathbb{E}[(X_{ij}^g)^3 \mathbf{1}_{|X_{ij}^g|\le K_N}] && (4.9)
\end{aligned}$$

Chapter Four

where $\lambda_2(H) \le C_2(\Im z)\frac{1}{Nc_N}$ and $\lambda_3(H) \le C_3(\Im z)\frac{1}{Nc_N^{3/2}}$.

We first deal with the terms in (4.8). Note that since $p_{ij} \le 1$, we have $|X^b_{ij}| \le 1$, and as a consequence, for any $K_N \ge 2$, the first term in (4.8) is zero. For the Gaussian term, applying the Cauchy-Schwarz inequality followed by the second-moment Markov inequality yields

$$\mathbb{E}[(X^g_{ij})^2\mathbf{1}_{|X^g_{ij}|>K_N}] \le \mathbb{E}[(X^g_{ij})^4]^{1/2}\mathbb{P}(X^g_{ij} > K_N)^{1/2} \le K_N^{-1}\mathbb{E}[(X^g_{ij})^4]^{1/2}\mathbb{E}[(X^g_{ij})^2]^{1/2}.$$

Since $\mathbb{E}[(X^g_{ij})^2] = \mathbf{E}[p_{ij}+\mathbb{E}[p_{ij}]^2-2p_{ij}\mathbb{E}[p_{ij}]] \le \mathbf{E}[p_{ij}]$, and similarly, $\mathbb{E}[(X^g_{ij})^4] \le \mathbf{E}[p^2_{ij}]$, we have

$$\begin{aligned}
&\lambda_2(H) \sum_{1\le i<j\le N} \mathbb{E}[(X^g_{ij})^2\mathbf{1}_{|X^g_{ij}|>K_N}] \\
&\le \frac{\lambda_2(H)}{K_N} \sum_{1\le i<j\le N} \frac{\mathbf{E}[W_i^2]^{1/2}\mathbf{E}[W_j^2]^{1/2}}{\|i-j\|^{\alpha}} \frac{\mathbf{E}[W_i]^{1/2}\mathbf{E}[W_j]^{1/2}}{\|i-j\|^{\alpha/2}} \\
&\le \frac{\lambda_2(H)}{K_N}\mathbf{E}[W_1]\mathbf{E}[W_1^2]N^{2-\frac{3\alpha}{2}} \\
&\le \frac{\tilde{c_2}\mathbf{E}[W_1]\mathbf{E}[W_1^2]N^{2-\frac{3\alpha}{2}}}{K_N N^{2-\alpha}} = \mathrm{O}_N(N^{-\alpha/2}K_N^{-1}),
\end{aligned}$$

where the last equality follows as $\tau > 3$ and $\tilde{c_2}$ is a constant depending on $\Im(z)$ only. For the term containing the third moments, we see that

$$\begin{aligned}
&\lambda_3(H) \sum_{1\le i<j\le N} \mathbb{E}[(X^b_{ij})^3\mathbf{1}_{|X^b_{ij}|\le K_N}] + \mathbb{E}[(X^g_{ij})^3\mathbf{1}_{|X^g_{ij}|\le K_N}] \\
&\le \lambda_3(H)K_N \sum_{1\le i\le j\le N} \mathbb{E}[(X^b_{ij})^2] + \mathbb{E}[(X^g_{ij})^2] \\
&\le \lambda_3(H)K_N 2\mathbf{E}[W_1]^2 \sum_{1\le i\le j\le N} \frac{1}{\|i-j\|^{\alpha}} \\
&\le \frac{c_3(\Im z)}{Nc_N^{3/2}}K_N\mathbf{E}[W_1]^2 Nc_N \le \tilde{c}_3 K_N c_N^{-1/2}.
\end{aligned}$$

Here $\tilde{c}_3$ is a constant depending on $\Im(z)$. Choosing any $2 \le K_N \ll c_N^{1/2}$, both terms go to zero. This completes the proof of the Gaussianisation. □

§4.3.2 Simplification of the variance profile

We now proceed with a series of lemmas to simplify the variance profile of our Gaussianised matrix. First, we construct a new matrix $\mathbf{\Delta}_{N,g}$ as the Laplacian corresponding to the matrix $\mathbf{A}_{N,g}$, defined as follows:

Suppose $(G_{i,j})_{i>j}$ is a sequence of i.i.d. $\mathrm{N}(0,1)$ random variables as before, and independent of the sequence $(W_i)_{i\in\mathbf{V}_N}$. Define

$$\mathbf{A}_{N,g}=\begin{cases}\frac{\sqrt{p_{ij}(1-p_{ij})}}{\sqrt{c_N}}G_{i\wedge j,i\vee j} & i\neq j\\ 0 & i=j.\end{cases}$$

We now have the following result.

Lemma 4.3.2.
Let $\bar{\mathbf{\Delta}}_N$ and $\mathbf{\Delta}_{N,g}$ be as defined above. Then,

$$\lim_{N\to\infty}\mathbb{P}(d_L(\mathrm{ESD}(\bar{\mathbf{\Delta}}_N),\mathrm{ESD}(\mathbf{\Delta}_{N,g}))>\varepsilon)=0.$$

Proof. The proof follows using Proposition 4.6.1. Taking expectation on the d_L distance, we have

$$\begin{aligned}\mathbb{E}\left[d_L^3(\mathrm{ESD}(\mathbf{\Delta}_{N,g},\mathrm{ESD}(\bar{\mathbf{\Delta}}_N)\right]&\le\frac{1}{N}\mathbb{E}\,\mathrm{Tr}\left((\mathbf{\Delta}_{N,g}-\bar{\mathbf{\Delta}}_N)^2\right)\\&=\frac{1}{N}\mathbb{E}\left[\sum_{1\le i,j\le N}\left(\mathbf{\Delta}_{N,g}(i,j)-\bar{\mathbf{\Delta}}_N(i,j)\right)^2\right]\\&=\frac{1}{N}\sum_{1\le i\neq j\le N}\mathbb{E}\left[(\mathbf{A}_{N,g}(i,j)-\bar{\mathbf{A}}_N(i,j))^2\right]\\&\quad+\frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\left[\left(\sum_{k\neq i}\mathbf{A}_{N,g}(i,k)-\bar{\mathbf{A}}_N(i,k)\right)^2\right].\end{aligned}$$

We deal with the last two terms separately. The first term is bounded above by

$$\frac{1}{Nc_N}\sum_{i\neq j}\mathbb{E}[\mu_{ij}^2]\le\frac{1}{Nc_N}\sum_{i\neq j}\frac{\mathbf{E}[W_1^2]^2}{\|i-j\|^{2\alpha}}\approx\frac{N^{2-2\alpha}}{N^{2-\alpha}}=N^{-\alpha}\to 0.$$

Next, we have that

$$\sum_{k\neq i}\mathbf{A}_{N,g}(i,k)-\bar{\mathbf{A}}_N(i,k)\le\frac{1}{\sqrt{c_N}}\sum_{k\neq i}p_{ik}=c_N^{-1/2}.$$

This makes the second term of the order $\mathrm{o}_N(c_N)$. We conclude the proof using Markov's inequality. $\square$

Define for $i\neq j$

$$\tilde{\mathbf{A}}_{N,g}(i,j)=\frac{\sqrt{p_{ij}}}{\sqrt{c_N}}G_{i\wedge j,i\vee j}$$

and put zero on the diagonal. Here $(G_{i,j})_{i\geq j}$ are the i.i.d. $\mathrm{N}(0,1)$ random variables used in the previous result. Let $\tilde{\mathbf{\Delta}}_{N,g}$ be analogously defined. The next lemma shows that $\mathbf{\Delta}_{N,g}$ and $\tilde{\mathbf{\Delta}}_{N,g}$ have asymptotically the same spectrum.

Lemma 4.3.3.

$$\lim_{N\to\infty} \mathbb{P}\left(d_L(\mathrm{ESD}(\tilde{\mathbf{\Delta}}_{N,g}), \mathrm{ESD}(\mathbf{\Delta}_{N,g})) > \varepsilon\right) = 0.$$

Proof. Again using Proposition 4.6.1, we have that

$$\begin{aligned}
&\mathbb{E}\left[d_L^3(\mathrm{ESD}(\tilde{\mathbf{\Delta}}_{N,g}), \mathrm{ESD}(\mathbf{\Delta}_{N,g}))\right] \\
&\le \frac{1}{N}\mathbb{E}\,\mathrm{Tr}\left((\mathbf{\Delta}_{N,g} - \tilde{\mathbf{\Delta}}_{N,g})^2\right) \\
&= \frac{1}{N}\mathbb{E}\left[\sum_{1\le i,j\le N}\left(\mathbf{\Delta}_{N,g}(i,j) - \tilde{\mathbf{\Delta}}_{N,g}(i,j)\right)^2\right] \\
&= \frac{1}{N}\sum_{1\le i\neq j\le N}\mathbb{E}\left[(\mathbf{A}_{N,g}(i,j) - \tilde{\mathbf{A}}_{N,g})^2\right] \\
&\qquad + \frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\left[\left(\sum_{k\neq i}\mathbf{A}_{N,g}(i,k) - \tilde{\mathbf{A}}_{N,g}(i,k)\right)^2\right]
\end{aligned}$$

Dealing with the last two terms separately as before, we proceed by bounding the first term by

$$\frac{1}{Nc_N}\sum_{i\neq j}\frac{\mathbf{E}[W_1^2]^2}{\|i-j\|^{2\alpha}} \approx \frac{N^{2-2\alpha}}{N^{2-\alpha}} = N^{-\alpha} \to 0.$$

Expanding the square in the second term, we have

$$\begin{aligned}
&\frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\left[\left(\sum_{k\neq i}\mathbf{A}_{N,g}(i,k) - \tilde{\mathbf{A}}_{N,g}(i,k)\right)^2\right] \\
&= \frac{1}{N}\sum_{i=1}^{N}\sum_{k\neq i}\mathbb{E}\left[\left(\mathbf{A}_{N,g}(i,k) - \tilde{\mathbf{A}}_{N,g}(i,k)\right)^2\right] \\
&\quad + \frac{1}{N}\sum_{i=1}^{N}\sum_{k\neq i}\sum_{\ell\neq i,k}\mathbb{E}\left[\left(\mathbf{A}_{N,g}(i,k) - \tilde{\mathbf{A}}_{N,g}(i,k)\right)\left(\mathbf{A}_{N,g}(i,\ell) - \tilde{\mathbf{A}}_{N,g}(i,\ell)\right)\right].
\end{aligned}$$

Again, the first term in above sum is of the order $N^{-\alpha}$ and the expectation in the second term is zero. Indeed, using the independence between $(W_i)_{i\in\mathbf{V}_N}$ and $G_{i,j}$ we have for $k\neq\ell$,

$$\begin{aligned}
&\mathbb{E}\left[\left(\mathbf{A}_{N,g}(i,k) - \tilde{\mathbf{A}}_{N,g}(i,k)\right)\left(\mathbf{A}_{N,g}(i,\ell) - \tilde{\mathbf{A}}_{N,g}(i,\ell)\right)\right] \\
&= \mathbb{E}\left[(\sqrt{p_{ik}(1-p_{ik})} - \sqrt{p_{ik}})(\sqrt{p_{i\ell}(1-p_{i\ell})} - \sqrt{p_{i\ell}})\right]\mathbb{E}[G_{i,k}G_{i,\ell}] = 0.
\end{aligned}$$

This completes the proof of the lemma. □

We conclude this subsection with one final simplification. For any $i \neq j$, let

$$r_{ij} = \frac{W_i W_j}{\|i-j\|^\alpha},$$

and let $r_{ii} = 0$. Define the matrix $\widehat{\mathbf{A}}_{N,g}$ as follows: for $i \neq j$,

$$\widehat{\mathbf{A}}_{N,g}(i,j) = \frac{\sqrt{r_{i\wedge j, i\vee j}}}{\sqrt{c_N}} G_{i\wedge j, i\vee j}$$

and put 0 on the diagonal. Define Laplacian matrix $\widehat{\mathbf{\Delta}}_{N,g}$ accordingly with $\widehat{\mathbf{A}}_{N,g}$.

Lemma 4.3.4.

$$\lim_{N\to\infty} \mathbb{P}\left(d_L(\mathrm{ESD}(\tilde{\mathbf{\Delta}}_{N,g}), \mathrm{ESD}(\widehat{\mathbf{\Delta}}_{N,g})) > \varepsilon\right) = 0.$$

Proof. For any $1 \leq i \neq j \leq N$, define the set $\mathcal{C}_{ij} = \{r_{ij} < 1\}$. Let $(X_{i,j})_{i\geq j}$ be defined as follows

$$X_{ij} = \frac{\sqrt{r_{ij}}}{\sqrt{c_N}} G'_{ij},$$

where $(G'_{ij})_{i\geq j}$ be a sequence of independent $\mathrm{N}(0,1)$ random variables, independent of the previously defined (G_{ij}) and $(W_i)_{i\in\mathbf{V}_N}$. Define a symmetric matrix $L_{N,g}$ as follows: for $1 \leq i < j \leq N$,

$$L_{N,g}(i,j) = \tilde{\mathbf{A}}_{N,g}(i,j)\mathbf{1}_{\mathcal{C}_{ij}} + X_{ij}\mathbf{1}_{\mathcal{C}_{ij}^c}.$$

We put zero on the diagonal and consider the $\mathbf{\Delta}_L$ as the Laplacian matrix corresponding to $L_{N,g}$. Note that $L_{N,g}$ has the same distribution as $\widehat{\mathbf{A}}_{N,g}$ and hence the $\mathbf{\Delta}_L$ has the same distribution as $\widehat{\mathbf{\Delta}}_{N,g}$.

By Proposition 4.6.1, we again have

$$\begin{aligned}
\mathbb{E}\left[d^3\left(\mathrm{ESD}(\mathbf{\Delta}_L), \mathrm{ESD}(\tilde{\mathbf{\Delta}}_{N,g})\right)\right] &\leq \frac{1}{N}\mathbb{E}\left[\sum_{1\leq i,j\leq N}\left(\mathbf{\Delta}_L(i,j) - \tilde{\mathbf{\Delta}}_{N,g}(i,j)\right)^2\right] \\
&= \frac{1}{N}\sum_{1\leq i\neq j\leq N}\mathbb{E}\left[(L_{N,g}(i,j) - \tilde{\mathbf{A}}_{N,g}(i,j))^2\right] \\
&\quad + \frac{1}{N}\sum_{i=1}^N \mathbb{E}\left[\left(\sum_{k\neq i} L_{N,g}(i,k) - \tilde{\mathbf{A}}_{N,g}(i,k)\right)^2\right].
\end{aligned}$$

Chapter Four

Expanding terms on the right-hand side, we obtain

$$
\begin{aligned}
&\mathbb{E}\left[d^3\left(\mathrm{ESD}(\boldsymbol{\Delta}_L),\mathrm{ESD}(\tilde{\boldsymbol{\Delta}}_{N,g})\right)\right]\\
&\le \frac{4}{N}\sum_{1\le i\neq j\le N}\mathbb{E}\left[(L_{N,g}(i,j)-\tilde{\mathbf{A}}_{N,g}(i,j))^2\right]\\
&\le +\frac{1}{N}\sum_{i=1}^{N}\sum_{k\neq i}\sum_{\ell\neq i,k}\mathbb{E}\left[\left(L_{N,g}(i,k)-\tilde{\mathbf{A}}_{N,g}(i,k)\right)\left(L_{N,g}(i,\ell)-\tilde{\mathbf{A}}_{N,g}(i,\ell)\right)\right]
\end{aligned}
$$

Again, we deal with the two sums separately. The first sum can be bounded above as follows:

$$
\begin{aligned}
&\frac{4}{N}\sum_{1\le i\neq j\le N}\mathbb{E}\left[(L_{N,g}(i,j)-\tilde{\mathbf{A}}_{N,g}(i,j))^2\right]\\
&\le \frac{4}{N}\sum_{1\le i\neq j\le N}\mathbb{E}\left[(\tilde{\mathbf{A}}_{N,g}(i,j)-X_{ij})^2\mathbf{1}_{\mathcal{C}_{ij}}\right]\\
&\le \frac{8}{N}\sum_{1\le i\neq j\le N}\mathbb{E}\left[\tilde{\mathbf{A}}_{N,g}(i,j)^2\mathbf{1}_{\mathcal{C}_{ij}}\right]+\mathbb{E}\left[X_{ij}^2\mathbf{1}_{\mathcal{C}_{ij}}\right]\\
&\le \frac{1}{Nc_N}\sum_{i\neq j\in\mathbf{V}_N}^{N}\mathbf{E}[G^2_{i\wedge j,i\vee j}\mathbf{1}_{\mathcal{C}_{ij}^c}]+\mathbf{E}[X_{ij}^2\mathbf{1}_{\mathcal{C}_{ij}^c}]\\
&\le \frac{1}{Nc_N}\sum_{i\neq j\in\mathbf{V}_N}^{N}\mathbf{P}(\mathcal{C}_{ij}^c)+\mathbf{E}[X_{ij}^4]^{1/2}\mathbf{P}(\mathcal{C}_{ij}^c)^{1/2}\\
&\le \frac{1}{Nc_N}\sum_{i\neq j\in\mathbf{V}_N}^{N}\mathbf{P}(\mathcal{C}_{ij}^c)+\frac{3\mathbf{E}[W_i^2W_j^2]^{1/2}}{\|i-j\|^{\alpha}}\mathbf{P}(\mathcal{C}_{ij}^c)^{1/2}\\
&\le C(N^{-\alpha(\tau-2)}+N^{-\frac{\alpha}{2}(\tau-1)})=\mathrm{o}_N(1),
\end{aligned}
$$

where we have used in the last line the following estimate:

$$
\mathbf{P}(\mathcal{C}_{ij}^c)\le\mathbf{P}\left(W_iW_j\ge\|i-j\|^{\alpha}\right)\le\frac{c}{\|i-j\|^{\alpha(\tau-1)}}
$$

which follows from Lemma 4.6.2. For the second term note that

$$
\begin{aligned}
&\mathbb{E}\left[\left(L_{N,g}(i,k)-\tilde{\mathbf{A}}_{N,g}(i,k)\right)\left(L_{N,g}(i,\ell)-\tilde{\mathbf{A}}_{N,g}(i,\ell)\right)\right]\\
&=\frac{1}{c_N}\mathbf{E}[\sqrt{p_{ik}}\sqrt{p_{i\ell}}\mathbf{1}_{\mathcal{C}_{ij}^c}\mathbf{1}_{\mathcal{C}_{i\ell}^c}]\mathbb{E}[G_{ik}G_{i\ell}]\\
&\quad-\frac{1}{c_N}\mathbf{E}[\sqrt{p_{ik}}\sqrt{r_{i\ell}}\mathbf{1}_{\mathcal{C}_{ij}^c}\mathbf{1}_{\mathcal{C}_{i\ell}^c}]\mathbb{E}[G_{ik}G'_{i\ell}]-\frac{1}{c_N}\mathbf{E}[\sqrt{r_{ik}}\sqrt{p_{i\ell}}\mathbf{1}_{C_{ij}^c}\mathbf{1}_{C_{i\ell}^c}]\mathbb{E}[G'_{ik}G_{i\ell}]\\
&\quad+\frac{1}{c_N}\mathbf{E}[\sqrt{r_{ik}}\sqrt{r_{i\ell}}\mathbf{1}_{C_{ij}^c}\mathbf{1}_{C_{i\ell}^c}]\mathbb{E}[G'_{ik}G'_{i\ell}],
\end{aligned}
$$

Chapter Four

and since $k \neq \ell$, all the above terms are zero. Thus the proof follows. $\square$

§4.3.3 Truncation

Let $m > 1$ be a truncation threshold and define $W_i^m = W_i \mathbf{1}_{W_i \leq m}$ for any $i \in \mathbf{V}_N$. For all $N \in \mathbb{N}$, we define a new random matrix as follows: Let

$$r_{ij}^m = \frac{W_i^m W_j^m}{\|i-j\|^\alpha} \qquad i \neq j \in \mathbf{V}_N\,,$$

and let $\mathbf{A}_{N,g,m}$ be defined for $i \neq j$ as

$$\mathbf{A}_{N,g,m}(i,j) = \frac{\sqrt{r_{ij}^m}}{\sqrt{c_N}} G_{i\wedge j, i\vee j},$$

and put 0 on the diagonal. Analogously define $\mathbf{\Delta}_{N,g,m}$.

Lemma 4.3.5 (Truncation).
For every $\delta > 0$ one has

$$\limsup_{m\to\infty} \lim_{N\to\infty} \mathbb{P}\left(d_L(\mathrm{ESD}(\mathbf{\Delta}_{N,g,m}), \mathrm{ESD}(\tilde{\mathbf{\Delta}}_{N,g})) > \delta\right) = 0\,.$$

Proof. The proof follows the same idea as the previous lemmas. Recall that

$$\widehat{\mathbf{A}}_{N,g}(i,j) = \frac{\sqrt{r_{ij}}}{\sqrt{c_N}} G_{i\wedge j, i\vee j}$$

for all $i \neq j$, with 0 on the diagonal, and $\widehat{\mathbf{\Delta}}_{N,g}$ is the corresponding Laplacian. Once again, we have

$$\begin{aligned}
&\mathbb{E}\left[d^3\left(\mathrm{ESD}(\mathbf{\Delta}_{N,g,m}), \mathrm{ESD}(\widehat{\mathbf{\Delta}}_{N,g})\right)\right] \\
&\leq \frac{1}{N}\mathbb{E}\left[\sum_{1\leq i,j\leq N}\left(\mathbf{\Delta}_{N,g,m}(i,j) - \widehat{\mathbf{\Delta}}_{N,g}(i,j)\right)^2\right] \\
&= \frac{1}{N}\sum_{1\leq i\neq j\leq N}\mathbb{E}\left[(\mathbf{A}_{N,g,m}(i,j) - \widehat{\mathbf{A}}_{N,g}(i,j))^2\right] \\
&\qquad + \frac{1}{N}\sum_{i=1}^N \mathbb{E}\left[\left(\sum_{k\neq i}\mathbf{A}_{N,g,m}(i,k) - \widehat{\mathbf{A}}_{N,g}(i,k)\right)^2\right] \\
&\leq \frac{4}{N}\sum_{1\leq i\neq j\leq N}\mathbb{E}\left[(\mathbf{A}_{N,g,m}(i,j) - \widehat{\mathbf{A}}_{N,g}(i,j))^2\right] \\
&+ \frac{1}{N}\sum_{i=1}^N\sum_{k\neq i}\sum_{\ell\neq i,k}\mathbb{E}\left[\left(\mathbf{A}_{N,g,m}(i,k) - \widehat{\mathbf{A}}_{N,g}(i,k)\right)\left(\mathbf{A}_{N,g,m}(i,\ell) - \widehat{\mathbf{A}}_{N,g}(i,\ell)\right)\right].
\end{aligned}$$

Chapter Four

The proof of Lemma 4.3.4 aids us by taking care of the second factor in the last line, which turns out to be equal to 0 by the independence of Gaussian terms. For the first term, the common Gaussian factor pulls out by independence, yielding the upper bound

$$
\begin{aligned}
&\frac{4}{Nc_N}\sum_{1\le i\neq j\le N}\frac{\mathbf{E}\left[\left(\sqrt{W_iW_j}-\sqrt{W_i^mW_j^m}\right)^2\right]}{\|i-j\|^\alpha}\\
&\le \frac{4}{Nc_N}\sum_{1\le i\neq j\le N}\frac{\mathbf{E}[W_iW_j-W_i^mW_j^m]}{\|i-j\|^\alpha},
\end{aligned}
$$

where the inequality follows by using the identity $(a-b)^2\le|a^2-b^2|$ for any $a,b\ge 1$. Adding and subtracting the term $W_iW_j^m$ inside the expectation gives us that

$$
\begin{aligned}
&\frac{4}{N}\sum_{1\le i\neq j\le N}\mathbb{E}\left[(\mathbf{A}_{N,g,m}(i,j)-\widehat{\mathbf{A}}_{N,g}(i,j))^2\right]\\
&\le \frac{4}{Nc_N}\sum_{1\le i\neq j\le N}\frac{\mathbf{E}[W_i]\mathbf{E}[W_j\mathbf{1}_{\{W_j>m\}}]+\mathbf{E}[W_j^m]\mathbf{E}[W_i\mathbf{1}_{\{W_i>m\}}]}{\|i-j\|^\alpha}\\
&\le \frac{C_\tau}{Nc_N}\sum_{1\le i\neq j\le N}\frac{m^{2-\tau}}{\|i-j\|^\alpha}=\mathrm{O}_m(m^{2-\tau}),
\end{aligned}
$$

where the last inequality follows from Lemma 4.6.3, with C_τ a $\tau-$dependent constant. Markov inequality concludes the proof. □

§4.3.4 Decoupling

Since we now have bounded weights, the decoupling result follows from the arguments from [Bryc et al., 2006, Lemma 4.12]. See also the proof of [Chakrabarty et al., 2021b, Lemma 4.2] for the inhomogeneous extension.

Lemma 4.3.6.
Let $(Z_i : i\ge 1)$ be a family of i.i.d. standard normal random variables, independent of $(G_{i,j} : 1\le i\le j)$. Define a diagonal matrix Y_N of order N by

$$Y_N(i,i)=Z_i\sqrt{\frac{\sum_{k\neq i}r_{ik}^m}{c_N}},\quad 1\le i\le N.$$

and let

$$\mathbf{\Delta}_{N,g,c}=\mathbf{A}_{N,g,m}+Y_N. \tag{4.10}$$

Chapter Four

Then for every $m > 1$, and for any $k \in \mathbb{N}$,

$$\lim_{N\to\infty} \frac{1}{N}\mathbb{E}\left(\mathrm{Tr}\left[(\mathbf{\Delta}_{N,g,c})^{2k} - (\mathbf{\Delta}_{N,g,m})^{2k}\right]\right) = 0.$$

and

$$\lim_{N\to\infty} \frac{1}{N^2}\mathbb{E}\left(\mathrm{Tr}^2\left[(\mathbf{\Delta}_{N,g,c})^{k}\right] - \mathrm{Tr}^2\left[(\mathbf{\Delta}_{N,g,m})^{k}\right]\right) = 0.$$

§4.4 Moment method: Existence and uniqueness of the limit

We begin by stating a key proposition that describes the limit of the empirical spectral distribution of $\mathbf{\Delta}_{N,g,c}$. The majority of this section will be devoted to the proof of this proposition, and so, we defer the proof of the proposition to page 180.

Proposition 4.4.1.
Let $\mathrm{ESD}(\mathbf{\Delta}_{N,g,c})$ *be the empirical spectral distribution of* $\mathbf{\Delta}_{N,g,c}$ *defined in* (4.10). *Then there exists a deterministic measure* ν_τ *on* $\mathbb{R}$ *such that*

$$\lim_{N\to\infty} \mathrm{ESD}(\mathbf{\Delta}_{N,g,c}) = \nu_{\tau,m} \qquad \textit{in } \mathbb{P}\textit{–probability}.$$

We now use Proposition 4.4.1 and tools from Appendix 4.6 and Section 4.3 to prove Theorem 4.2.1.

Proof of Theorem 4.2.1. Combining Proposition 4.4.1 with Lemma 4.3.6 gives us that

$$\lim_{N\to\infty} \mathrm{ESD}(\mathbf{\Delta}_{N,g,m}) = \nu_{\tau,m} \qquad \text{in } \mathbb{P}\text{–probability}. \tag{4.11}$$

To show the existence of the limit $\nu_\tau := \lim_{m\to\infty} \nu_{\tau,m}$, we wish to apply Lemma 4.6.5. Equation (4.11) satisfies Condition (1) of Lemma 4.6.5. Moreover, Condition (2) can be easily verified by Lemma 4.3.5. Thus, there exists a unique limit ν_τ such that

$$\lim_{N\to\infty} \mathrm{ESD}(\tilde{\mathbf{\Delta}}_{N,g}) = \nu_{\tau} \qquad \text{in } \mathbb{P}\text{–probability}. \tag{4.12}$$

Combining equation (4.12) with Lemma 4.3.4, and subsequently with Lemma 4.3.3 and Lemma 4.3.2 yields

$$\lim_{N\to\infty} \mathrm{ESD}(\bar{\mathbf{\Delta}}_{N}) = \nu_{\tau} \qquad \text{in } \mathbb{P}\text{–probability}. \tag{4.13}$$

We now wish to show that the limiting empirical spectral distribution for $\mathbf{\Delta}_N^\circ$ is ν_τ in $\mathbb{P}$–probability. To this end, note that for any h satisfying conditions of

Lemma 4.3.1, and H_z as in subsection 4.3.1, we have by the means of Lemma 4.3.1 that

$$\lim_{N\to\infty} h\left(\Re(H_z(\mathbf{\Delta}_N^\circ))\right) = h\left(\Re\,\mathrm{S}_{\nu_\tau}(z)\right).$$

The above characterises convergence in law. However, since ν_τ is a deterministic measure, the above convergence holds in $\mathbb{P}$–probability, and analogously for $\Im(H_z(\mathbf{\Delta}_N^\circ))$. This gives us that

$$\lim_{N\to\infty} \mathrm{S}_{\mathrm{ESD}(\mathbf{\Delta}_N^\circ)}(z) = \mathrm{S}_{\nu_\tau}(z) \qquad \text{in } \mathbb{P}\text{–probability}.$$

Since convergence of Stieltjes transforms characterises weak convergence, we obtain

$$\lim_{N\to\infty} \mathrm{ESD}(\mathbf{\Delta}_N^\circ) = \nu_\tau \qquad \text{in } \mathbb{P}\text{–probability},$$

completing the proof. □

We now provide the proof of Proposition 4.4.1. We borrow the main ideas of Chatterjee and Hazra [2022, Section 5.2.1, 5.2.2], and adapt them to our setting using the results of Cipriani et al. [2025, Section 4.4].

Proof of Proposition 4.4.1. The proof of the moment method is valid when the weights are bounded, and so for notational convenience, in this proof we will drop the dependence on m from $\{r_{ij}^m\}_{i,j\in\mathbf{V}_N}$. Thus, for the remainder of the proof, we have that

$$r_{ij} = \frac{W_i^m W_j^m}{\|i-j\|^\alpha}.$$

We apply the method of moments to show the convergence to the law $\nu_{\tau,m}$. The proof is split up into three parts as follows:

(a) For any $k \geq 1$, we compute the expected moment

$$\mathbb{E}\int_1^\infty x^k\,\mathrm{ESD}(\mathbf{\Delta}_{N,g,c})(\mathrm{d}\,x),$$

and show that as $N \to \infty$, the above quantity converges to a value $0 < M_k < \infty$ for k even, and 0 otherwise.

(b) We then show concentration by proving (under the law $\mathbb{P}$) that

$$\mathrm{Var}\left(\int_1^\infty x^k\,\mathrm{ESD}(\mathbf{\Delta}_{N,g,c})(\mathrm{d}\,x)\right) \to 0 \quad \text{as } N \to \infty.$$

(c) Lastly, we show that the sequence $\{M_k\}_{k\geq1}$ uniquely determines a limiting measure.

Step 1. We begin by considering that k is even. By using the expansion for $(a+b)^k$, it is easy to see that

$$\begin{aligned}\mathbb{E}\int_1^\infty x^k\,\mathrm{ESD}(\mathbf{\Delta}_{N,g,c})(\mathrm{d}\,x) &= \frac{1}{N}\mathbb{E}\left[\mathrm{Tr}\left(\mathbf{\Delta}_{N,g,c}^k\right)\right] \\ &= \frac{1}{N}\sum_{\substack{m_1,\ldots,m_k,\\ n_1,\ldots,n_k}}\mathbb{E}\left[\mathrm{Tr}\left(\mathbf{A}_{N,g,m}^{m_1}Y_N^{n_1}\ldots\mathbf{A}_{N,g,m}^{m_k}Y_N^{n_k}\right)\right],\end{aligned}$$

where $\mathbf{A}_{N,g,m}$ and Y_N are as in Lemma 4.3.6, and $\{m_i,n_i\}_{1\le i\le k}$ are such that $\sum_{i=1}^k m_i+n_i=k$.
Let $M(p)$ and $N(p)$ be defined as

$$M(p)=\sum_{i=1}^p m_i,\quad N(p)=\sum_{i=1}^p n_i$$

for any $1\le p\le k$. To expand the trace term, we sum over all $\mathbf{i}=(i_1,\ldots,i_{M(k)+N(k)+1})\in[N]^{M(k)+N(k)+1}$, where $[p]:=\{1,2,\ldots,p\}$, and we identify $i_{M(k)+N(k)+1}\equiv i_1$. Then, from Chatterjee and Hazra [2022, Eq. 5.2.2], we have

$$\begin{aligned}&\frac{1}{N}\mathrm{Tr}\left(\mathbf{A}_{N,g,m}^{m_1}Y_N^{n_1}\ldots\mathbf{A}_{N,g,m}^{m_k}Y_N^{n_k}\right)\\ &=\frac{1}{N}\sum_{i_1,\ldots,i_{M(k)}=1}^{N}\prod_{j=1}^{M(k)}G_{i_j\wedge i_{j+1},i_j\vee i_{j+1}}\prod_{j=1}^{M(k)}\frac{\sqrt{r_{i_ji_{j+1}}}}{\sqrt{c_N}}\prod_{j=1}^{k}\left(\frac{1}{c_N}\sum_{t=1}^{N}r_{i_{1+M(j)}t}\right)^{\frac{n_j}{2}}\prod_{j=1}^{k}Z_{i_{1+M(j)}}^{n_j},\end{aligned}\tag{4.14}$$

where also in (4.14) we identify $i_{M(k)+1}\equiv i_1$. Taking expectation in (4.14), we have that

$$\begin{aligned}&\mathbb{E}\left[\mathrm{tr}\left(\mathbf{A}_{N,g,m}^{m_1}Y_N^{n_1}\ldots\mathbf{A}_{N,g,m}^{m_k}Y_N^{n_k}\right)\right]\\ &=\frac{1}{N}\sum_{i_1,\ldots,i_{M(k)}}\mathbb{E}\left[\prod_{j=1}^{M(k)}G_{i_j\wedge i_{j+1},i_j\vee i_{j+1}}\right]\\ &\qquad\times\mathbb{E}\left[\prod_{j=1}^{M(k)}\frac{\sqrt{r_{i_ji_{j+1}}}}{\sqrt{c_N}}\prod_{j=1}^{k}\left(\frac{1}{c_N}\sum_{t=1}^{N}r_{i_{1+M(j)}t}\right)^{\frac{n_j}{2}}\right]\mathbb{E}\left[\prod_{j=1}^{k}Z_{i_{1+M(j)}}^{n_j}\right].\end{aligned}\tag{4.15}$$

It is well known that the expectation over a product of independent Gaussian random variables is simplified using the Wick's formula (see Lemma 4.6.7). In particular, if one were to partition the tuple $\{1,\ldots,K\}$ for some non-negative integer K, the contributing partitions are typically non-crossing pair partitions (Nica and Speicher [2006]).

We now introduce some notation from Cipriani et al. [2025]. For any fixed non-negative even integer K, let $\mathcal{P}_2(K)$ and $NC_2(K)$ be the set of all pair partitions and the set of all non-crossing pair partitions of $[K]$, respectively. Let $\gamma = (1, \ldots, K) \in S_K$ be the right-shift permutation (modulo K), and for any π which is a pair-partition, we identify it as a permutation of $[K]$, and read $\gamma\pi$ as a composition of permutations. Further, for any $\pi \in \mathcal{P}_2(K)$, let Cat_π denote the set

$$\mathrm{Cat}_\pi := \mathrm{Cat}_\pi(K, N) = \{\mathbf{i} \in [N]^K : i_r = i_{\gamma\pi(r)} \text{ for all } r \in [K]\}.$$

Let $C(K, N) = \mathrm{Cat}_\pi^c$, the complement of Cat_π, wherein we have $i_r = i_{\pi(r)}$ for any r. By Wick's formula for the Gaussian terms $\{G_{i,j}\}$, since the the sum over tuples $\mathbf{i}$ would be reduced to the sum over pair partitions $\pi \in \mathcal{P}_2(K)$ and the associated tuples $\mathbf{i} \in \mathrm{Cat}_\pi \cup C(K, N)$, we can write

$$\sum_{\mathbf{i}\in[K]^N} = \sum_{\pi\in\mathcal{P}_2(K)} \sum_{\mathbf{i}\in C(K,N)} + \sum_{\pi\in NC_2(K)} \sum_{\mathbf{i}\in\mathrm{Cat}_\pi} + \sum_{\pi\in\mathcal{P}_2(K)\setminus NC_2(K)} \sum_{\mathbf{i}\in\mathrm{Cat}_\pi} . \tag{4.16}$$

To analyse further, we use a key tool in the proof which is the following fact (Cipriani et al. [2025, Claim 4.10]).

Fact 4.4.2.
Let K be an even non-negative integer. Then, we have the following to be true:

(a) *For any $\pi \in NC_2(K)$, we have*

$$\lim_{N\to\infty} \frac{1}{Nc_N^{K/2}} \sum_{\mathbf{i}\in\mathrm{Cat}_\pi} \prod_{(r,s)\in\pi} \frac{1}{\|i_r - i_{r+1}\|^\alpha} = 1\,.$$

(b) *For any pair partition π, if $\mathbf{i} \in C(K, N)$, then,*

$$\lim_{N\to\infty} \frac{1}{Nc_N^{K/2}} \sum_{\mathbf{i}\in C(K,N)} \prod_{(r,s)\in\pi} \frac{1}{\|i_r - i_{r+1}\|^\alpha} = 0\,.$$

(c) *For a partition $\pi \in \mathcal{P}_2(K) \setminus NC_2(K)$, we have*

$$\lim_{N\to\infty} \frac{1}{Nc_N^{K/2}} \sum_{\mathbf{i}\in\mathrm{Cat}_\pi\cup C(K,N)} \prod_{(r,s)\in\pi} \frac{1}{\|i_r - i_{r+1}\|^\alpha} = 0\,.$$

Let $\tilde{\pi} := \gamma\pi$ for any choice of π. From Chatterjee [2005, Eq. 5.2.5], we have that

$$\mathcal{E}(\tilde{\pi}) := \mathbb{E}\left[\prod_{j=1}^{k} Z_{i_{1+M(j)}}^{n_j}\right] = \prod_{u\in\tilde{\pi}} \mathbb{E}\left[\prod_{\substack{j\in[k]:\\ 1+M(j)\in u}} Z_{\ell_u}^{n_j}\right] < \infty, \tag{4.17}$$

where u is a block in $\tilde{\pi}$ and ℓ_u its representative element. Note that this does not depend on the choice of $\mathbf{i}$, and to obtain a non-zero contribution, we must have that for all $u \in \tilde{\pi}$,

$$\sum_{j\in[k]:1+M(j)\in u} n_j \equiv 0 \pmod 2. \tag{4.18}$$

Observe that $\mathbb{E}\left[\prod_{j=1}^{M(k)} G_{i_j\wedge i_{j+1}, i_j\vee i_{j+1}}\right]$ depends only on $\tilde{\pi}$ and not the choice of $\mathbf{i}$, and as a consequence, we can define

$$\Phi(\tilde{\pi}) := \mathbb{E}\left[\prod_{j=1}^{M(k)} G_{i_j\wedge i_{j+1}, i_j\vee i_{j+1}}\right] < \infty\,. \tag{4.19}$$

Next, note that the sum

$$\frac{1}{c_N}\sum_{t=1}^{N} r_{i_{1+M(j)}t} = \mathrm{O}_N(1) \tag{4.20}$$

by definition of c_N (and the weights are uniformly bounded). Finally, if we look at the terms

$$\mathbb{E}\left[\left(\prod_{j=1}^{M(k)} \frac{r_{i_j i_{j+1}}}{c_N}\right)^{1/2}\right], \tag{4.21}$$

we can again bound the weights above by m. Recall that Wick's formula on the Gaussian terms imposes the restriction on choices of $\mathbf{i}$. Using these facts, in combination with (4.17), (4.19), and (4.20), we have that (4.15) gets bounded by

$$(4.15) \le \frac{C}{Nc_N}\sum_{\pi\in\mathcal{P}_2(M(k))}\sum_{\mathbf{i}\in\mathrm{Cat}_\pi\cup C(M(k),N)} \Phi(\tilde{\pi})\mathcal{E}(\tilde{\pi})\prod_{(r,s)\in\pi}\frac{m^2}{\|i_r - i_{r+1}\|^\alpha}\,. \tag{4.22}$$

If we split (4.22) as (4.16), then using Fact 4.4.2, we see that in the cases when $\pi \in \mathcal{P}_2(M(k))$ and $\mathbf{i} \in C(M(k), N)$, and when $\pi \in \mathcal{P}_2(M(k)) \setminus NC_2(M(k))$ for all $\mathbf{i}$, the contribution in the limit $N \to \infty$ is 0.

We are now in the setting where we take $\pi \in NC_2(M(k))$ and $\tilde{\pi} := \gamma\pi$, and $\mathbf{i} \in \mathrm{Cat}_\pi$. First, note that $\tilde{\pi}$ is a partition of $[M(k)]$. We remark that if $M(k) \equiv 1 \pmod 2$ then $NC_2(M(k)) = \emptyset$, and so, $M(k)$ must be even.

Next, we focus on analysing the product $\prod_{j=1}^{M(k)}\sqrt{r_{i_j i_{j+1}}^m}$ appearing in (4.15). We wish to show that this depends only on π, and not on the choice of $\mathbf{i}$. We follow the idea of Cipriani et al. [2025], wherein one constructs a graph associated to a chosen partition π, and any tuple $\mathbf{i} \in \mathrm{Cat}_\pi$ is equivalent to a tuple $\tilde{\mathbf{i}}$ with as many distinct indices as the number of vertices in the constructed graph. First, note that the coordinates are pairwise distinct (we take $r_{ii} = 0$ for all i). Next, we construct a preliminary graph from the closed walk $i_1 \to i_2 \to \dots i_{M(k)} \to i_1$. Lastly, we collapse vertices and edges that are matched in Cat_π, and we denote the resulting graph as $G_{\tilde{\pi}}$, since it does not depend on the choice of $\mathbf{i}$ but rather the choice of π itself. The resulting graph $G_{\tilde{\pi}}$ is the graph associated to the partition π, and we refer the reader to Definition 4.6.8 for a formal description. For clarity, consider the following example:

Let $M(k) = 4$, and let $\pi = \{\{1, 2\}, \{3, 4\}\}$. Then, $\tilde{\pi} = \{\{1, 3\}, \{2\}, \{4\}\}$. For any $\mathbf{i} \in \mathrm{Cat}_\pi$, we see that $i_1 = i_3$, and i_2, i_4 are independent indices. Now, $G_{\tilde{\pi}}$ is a graph on 3 vertices, which are labelled as $\{\{1, 3\}\}$, $\{2\}$ and $\{4\}$, and so its corresponding tuple $\tilde{\mathbf{i}}$ is exactly the same as $\mathbf{i}$.

We then have, from Chatterjee and Hazra [2022, Eq. 5.2.12], that

$$\prod_{j=1}^{M(k)} \sqrt{r^m_{i_j i_{j+1}}} = \prod_{e \in E_{\tilde{\pi}}} r_e^{t_e/2}, \tag{4.23}$$

where $E_{\tilde{\pi}}$ is the edge set of $G_{\tilde{\pi}}$ and t_e is the number of times an edge e is traversed in the closed walk on $G_{\tilde{\pi}}$. Also observe

$$\Phi(\tilde{\pi}) = \mathbb{E}\left[\prod_{e \in E_{\tilde{\pi}}} G_e^{t_e}\right].$$

Consequently, we must have that t_e to be even for all e, since the Gaussian terms are independent and mean 0. We claim that $t_e = 2$ for all $e \in E_{\tilde{\pi}}$. Indeed, if for all e, $t_e \geq 2$ with at least one e' such that $t_{e'} > 2$, then, $\sum_{e \in E_{\tilde{\pi}}} t_e > 2|E_{\tilde{\pi}}|$. Since $G_{\tilde{\pi}}$ is connected, $|E_{\tilde{\pi}}| \geq |V_{\tilde{\pi}}| - 1 = M(k)/2$, where $V_{\tilde{\pi}}$ is the vertex set. Thus, $\sum_{e \in E_{\tilde{\pi}}} t_e > M(k)$. But, $\sum_e t_e = M(k)$, gives a contradiction. We conclude that $t_e = 2$ for all $e \in E_{\tilde{\pi}}$.

A similar contradiction arises when we assume that there exists a self-loop in $G_{\tilde{\pi}}$. Thus $G_{\tilde{\pi}}$ is a tree on $\frac{M(k)}{2} + 1$ vertices with each edge traversed twice in the closed walk. As a consequence, every Gaussian term in $\Phi(\tilde{\pi})$ appears exactly twice, and so, $\Phi(\tilde{\pi}) = 1$.

Let b_s be the s^{th} block of $\tilde{\pi}$ and let ℓ_s its representative element. Define

$$\gamma_s := \#\{1 \leq j \leq k : 1 + M(j) \in b_s\},$$

and

$$\{s_1, s_2, \dots, s_{\gamma_s}\} = \{1 \leq j \leq k : 1 + M(j) \in b_s\}.$$

We then have

$$\prod_{j=1}^{k}\left(\frac{1}{c_N}\sum_{t=1}^{N} r_{i_{1+M(j)t}}\right)^{\frac{n_j}{2}} = \prod_{s=1}^{\frac{M(k)}{2}+1}\left(\frac{1}{c_N}\sum_{t=1}^{N} r_{\ell_s t}\right)^{\sum_{j=1}^{\gamma_s} n_{s_j}/2}. \tag{4.24}$$

Note that

$$\sum_{j=1}^{\gamma_s} n_{s_j} = \sum_{j\in[k]:1+M(j)\in b_s} n_j.$$

Let us define $\tilde{n}_s := \sum_{j=1}^{\gamma_s} n_{s_j}/2$. Then,

$$\sum_{s:b_s\in\tilde{\pi}} \tilde{n}_s = \frac{N(k)}{2}. \tag{4.25}$$

Using Chatterjee and Hazra [2022, Eq. 5.2.16], we obtain

$$\begin{aligned}
&\frac{1}{N c_N^{\frac{M(k)}{2}}}\sum_{\mathbf{i}\in\mathrm{Cat}_\pi}\prod_{j=1}^{M(k)}\sqrt{r_{i_j i_{j+1}}}\prod_{j=1}^{k}\left(\frac{1}{c_N}\sum_{t=1}^{N} r_{i_{1+M(j)t}}\right)^{\frac{n_j}{2}}\\
&= \frac{1}{N c_N^{\frac{M(k)+N(k)}{2}}}\sum_{\substack{\ell_1\neq\ldots\neq\ell_{M(k)/2+1},\\ p_{(s,1)},\ldots,p_{(s,\tilde{n}_s)}:s\in\left[\frac{M(k)}{2}+1\right]}}\prod_{(u,v)\in E_{\tilde{\pi}}} r_{\ell_u\ell_v}\prod_{s=1}^{\frac{M(k)}{2}+1}\prod_{t=1}^{\tilde{n}_s} r_{\ell_s p_{(s,t)}},
\end{aligned} \tag{4.26}$$

where for any two blocks b_{s_1} and b_{s_2}, $\{p(s_1,1),p(s_1,2),\ldots\}$ and $\{p(s_2,1),p(s_2,2),\ldots\}$ are non-intersecting sets of indices $\{p_1,p_2,\ldots,p_{\tilde{n}_{s_1}}\}$ and $\{p'_1,p'_2,\ldots,p'_{\tilde{n}_{s_2}}\}$. Note that for $(u,v)\in E_{\tilde{\pi}}$, $r_{\ell_u\ell_v}=r_{uv}$ as before, but we rewrite in terms of representative elements to indicate common factors with the terms $r_{\ell_s p_{(s,t)}}$. Taking expectation in (4.26) gives us

$$\begin{aligned}
&\mathbb{E}[(4.26)]\\
&= \frac{1}{N c_N^{\frac{M(k)+N(k)}{2}}}\sum_{\ell_1\neq\ldots\neq\ell_{\frac{M(k)}{2}+1}}\mathbb{E}\Bigg[\prod_{(u,v)\in E_{\tilde{\pi}}}\frac{W_{\ell_u}^m W_{\ell_v}^m}{\|\ell_u-\ell_v\|^\alpha}\\
&\qquad\times\sum_{\substack{p_{(s,1)},\ldots,p_{(s,\tilde{n}_s)}:\\ s\in\left[\frac{M(k)}{2}+1\right]}}\prod_{s=1}^{\frac{M(k)}{2}+1}\prod_{t=1}^{\tilde{n}_s}\frac{W_{\ell_s}^m W_{p_{(s,t)}}^m}{\|\ell_s-p_{(s,t)}\|^\alpha}\Bigg].
\end{aligned} \tag{4.27}$$

The vertex set $V_{\tilde{\pi}}$ of the graph $G_{\tilde{\pi}}$ yields $M(k)/2+1$ distinct indices, due to the tree structure. Using Fact 4.4.2, we see that the factor of

$$\sum_{\ell_1,\ldots,\ell_{\frac{M(k)}{2}+1}} \prod_{(u,v)\in E_{\tilde{\pi}}} \frac{1}{\|\ell_u-\ell_v\|^\alpha}$$

is of the order of $\mathrm{O}_N\left(c_N^{\frac{M(k)}{2}}\right)$ since the weights are uniformly bounded in the range $[1,m]$. For the second summand in (4.27), the index ℓ_s already appears in the graph $G_{\tilde{\pi}}$, and for any s, we have $\tilde{n}_s$ many distinct indices from the sequence $\{p_{s,t}\}$, and summing over all s yields $N(k)/2$ many distinct indices due to (4.25). The second summation is therefore of the order of $\mathrm{O}_N\left(c_N^{\frac{N(k)}{2}}\right)$.

We claim that as $N\to\infty$, (4.27) converges to the limit

$$\mathbb{E}\left[\prod_{(u,v)\in E_{\tilde{\pi}}} W_{\ell_u}^m W_{\ell_v}^m \prod_{s=1}^{\frac{M(k)}{2}+1}\prod_{t=1}^{\tilde{n}_s} W_{\ell_s}^m W_{p_{(s,t)}}\right].$$

First, note that the weights are bounded, and so, (4.27) is bounded above and below. Next, we note that with the scaling of $Nc_N^{M(k)/2}$, we have

$$\lim_{N\to\infty}\frac{1}{Nc_N^{M(k)/2}}\sum_{\ell_1\neq\ldots\neq\ell_{\frac{M(k)}{2}+1}}\mathbb{E}\left[\prod_{(u,v)\in E_{\tilde{\pi}}}\frac{W_{\ell_u}^m W_{\ell_v}^m}{\|\ell_u-\ell_v\|^\alpha}\right]=\mathbb{E}\left[\prod_{(u,v)\in E_{\tilde{\pi}}} W_{\ell_u}^m W_{\ell_v}^m\right],$$

which is the moments of the adjacency matrix of the model as in Cipriani et al. [2025]. Thus, combinatorially, the first summand in (4.27) corresponds with the graph $G_{\tilde{\pi}}$, as defined in Definition 4.5. Now, consider a modification of the graph as follows: For each vertex s in $G_{\tilde{\pi}}$, attach $\tilde{n}_s$ many independent leaves, and call the new graph $\tilde{G}_{\tilde{\pi}}$. We refer to Chatterjee and Hazra [2022] for a detailed description, and Figure 4.3 for a visual representation.

The second summand over the sequence $\{p_{s,t}\}$ for each s corresponds to the added leaves, since the only common index with the original graph is the index ℓ_s for each s. Keeping the index ℓ_s fixed (since it is summed out in the first summand involving the indices $\ell_1\neq\ldots\neq\ell_{\frac{M(k)}{2}+1}$) , we see that with the

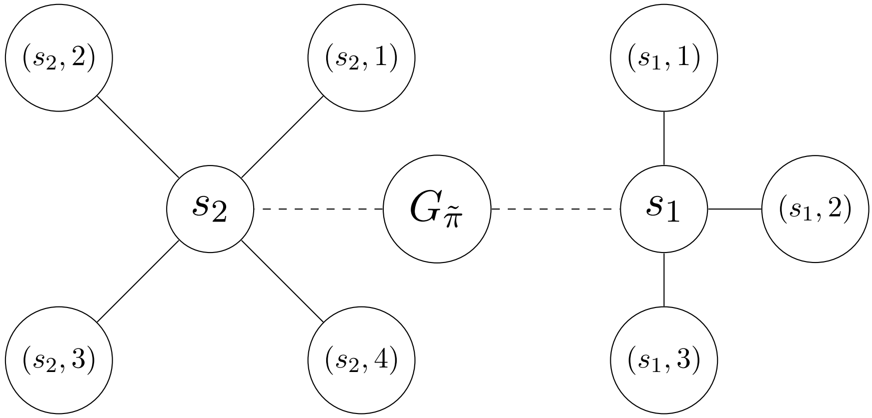


Figure 4.3: Modifying the graph $G_{\tilde{\pi}}$ to construct $\tilde{G}_{\tilde{\pi}}$. Here, we pick two vertices $s_1, s_2 \in V_{\tilde{\pi}}$, with $\tilde{n}_{s_1} = 3, \tilde{n}_{s_2} = 4$.

scaling $c_N^{N(k)/2}$ we have

$$\lim_{N\to\infty} \frac{1}{c_N^{N(k)/2}} \mathbb{E}\left[\sum_{p_{(s,1)},\dots,p_{(s,\tilde{n}_s)}} \prod_{s=1}^{\frac{M(k)}{2}+1} \prod_{t=1}^{\tilde{n}_s} \frac{W^m_{\ell_s} W^m_{p_{(s,t)}}}{\|\ell_s - p_{(s,t)}\|^\alpha} \middle| W^m_{\ell_s} \right]$$
$$= \mathbb{E}\left[\prod_{s=1}^{\frac{M(k)}{2}+1} \prod_{t=1}^{\tilde{n}_s} W^m_{\ell_s} W^m_{p_{(s,t)}} \middle| W^m_{\ell_s} \right].$$

Due to the compact support of the weights, it is now easy to conclude that

$$\lim_{N\to\infty} (4.27) = \mathbb{E}\left[\prod_{(u,v)\in E_{\tilde{\pi}}} W^m_{\ell_u} W^m_{\ell_v} \prod_{s=1}^{\frac{M(k)}{2}+1} \prod_{t=1}^{\tilde{n}_s} W^m_{\ell_s} W_{p_{(s,t)}} \right] =: t(\tilde{G}_{\tilde{\pi}}, W^m) \tag{4.28}$$

where $W^m = (W^m_1, W^m_2, \dots)$ and $\tilde{G}_{\tilde{\pi}}$ is the modified graph as described above and illustrated in Figure 4.3.

We can therefore conclude that for all even k,

$$\lim_{N\to\infty} \frac{1}{N} \mathbb{E}\left[\mathrm{tr}(\boldsymbol{\Delta}^k_{N,g,c}) \right] = \sum_{\substack{m_1,\dots,m_k,\\ n_1,\dots,n_k}} \sum_{\pi \in NC_2(M(k))} \mathcal{E}(\tilde{\pi}) t(\tilde{G}_{\tilde{\pi}}, W^m). \tag{4.29}$$

Now, consider the case when k is odd. Due to (4.18), we have that $M(k)$ must be odd. Thus, π cannot be a pair partition, and in particular, $\pi \notin$

$NC_2(M(k))$. Consider the term $\Phi(\tilde{\pi})$ in (4.19), and notice that by Wick's formula, this term is identically 0 if $M(k)$ is odd. Since the other expectations in (4.15) are of order $\mathrm{O}_N(1)$, we conclude that the odd moments are 0 in expectation.

Step 2. We now wish to show the concentration of the moments. Define

$$P(\mathbf{i}) := \mathbb{E}\left[\prod_{j=1}^{M(k)} G_{i_j\wedge i_{j+1}, i_j\vee i_{j+1}} \prod_{j=1}^{M(k)} \frac{\sqrt{r_{i_j i_{j+1}}}}{\sqrt{c_N}} \prod_{j=1}^{k}\left(\frac{1}{c_N}\sum_{t=1}^{N} r_{i_{1+M(j)}t}\right)^{\frac{n_j}{2}} \prod_{j=1}^{k} Z^{n_j}_{i_{1+M(j)}}\right],$$

and

$$\begin{aligned} P(\mathbf{i},\mathbf{i}') := \mathbb{E}\Bigg[&\prod_{j=1}^{M(k)} G_{i_j\wedge i_{j+1}, i_j\vee i_{j+1}} \prod_{j=1}^{M(k)} \frac{\sqrt{r_{i_j i_{j+1}}}}{\sqrt{c_N}} \prod_{j=1}^{k}\left(\frac{1}{c_N}\sum_{t=1}^{N} r_{i_{1+M(j)}t}\right)^{\frac{n_j}{2}} \prod_{j=1}^{k} Z^{n_j}_{i_{1+M(j)}} \\ &\times \prod_{j=1}^{M(k)} G_{i'_j\wedge i'_{j+1}, i'_j\vee i'_{j+1}} \prod_{j=1}^{M(k)} \frac{\sqrt{r_{i'_j i'_{j+1}}}}{\sqrt{c_N}} \prod_{j=1}^{k}\left(\frac{1}{c_N}\sum_{t=1}^{N} r_{i'_{1+M(j)}t}\right)^{\frac{n_j}{2}} \prod_{j=1}^{k} Z^{n_j}_{i'_{1+M(j)}}\Bigg].\end{aligned}$$

Then,

$$\begin{aligned} &\mathrm{Var}\left(\int_{\mathbb{R}} x^k \,\mathrm{ESD}(\boldsymbol{\Delta}_{N,g,c})(\mathrm{d}\,x)\right) \\ &= \frac{1}{N^2}\sum_{\substack{m_1,\dots,m_k,\\ n_1,\dots,n_k}}\ \sum_{\mathbf{i},\mathbf{i}':[M(k)]\to[N]} \left[P(\mathbf{i},\mathbf{i}') - P(\mathbf{i})P(\mathbf{i}')\right], \end{aligned} \tag{4.30}$$

and we would like to show (4.30) $\to 0$. If $\mathbf{i}$ and $\mathbf{i}'$ have no common indices, then $P(\mathbf{i},\mathbf{i}') = P(\mathbf{i})P(\mathbf{i}')$ by independence. If there is *exactly* one common index, say $i_1 = i'_1$, then by independence of Gaussian terms, the factors $\mathbb{E}[G_{i_1,i_2}]$ and $\mathbb{E}[G_{i_1,i'_2}]$ would pull out, causing (4.30) to be identically 0. Thus, we have *at least* one matching of the form $(i_1,i_2) = (i'_1,i'_2)$.

Let us begin by taking k to be even. Consider exactly one matching, which we take to be $(i_1,i_2) = (i'_1,i'_2)$ without loss of generality. Let π,π' be partitions of $\{1,2,\dots,M(k)\},\{1',2',\dots,M(k)'\}$ respectively. Let $\sum^{(1)}$ denote the sum over index sets $\mathbf{i},\mathbf{i}'$ with exactly one matching. Then, we have by an extension of the previous argument

$$\frac{1}{N^2}\sum^{(1)}_{\mathbf{i},\mathbf{i}':[M(k)]\to[N]} P(\mathbf{i},\mathbf{i}')$$
$$\leq \frac{1}{N^2 c_N^{M(k)}}\sum_{\pi,\pi'}\Phi(\tilde{\pi})\mathcal{E}(\tilde{\pi})\Phi(\tilde{\pi}')\mathcal{E}(\tilde{\pi}')\sum_{\mathbf{i},\mathbf{i}'}\mathbb{E}\left[r_{i_1 i_2}\prod_{j=2}^{M(k)}\sqrt{r_{i_j i_{j+1}}}\prod_{j=2}^{M(k)}\sqrt{r_{i'_j i'_{j+1}}}\right]. \tag{4.31}$$

Expanding the expression for r_{ij} and using the fact that $W_i^m \leq m$ gives us that (4.31) is bounded above by

$$\frac{m^{2M(k)}}{N^2 c_N^{M(k)}}\sum_{\pi}\Phi(\tilde{\pi})\mathcal{E}(\tilde{\pi})\Phi(\tilde{\pi}')\mathcal{E}(\tilde{\pi}')\sum_{\mathbf{i},\mathbf{i}'}\frac{1}{\|i_1-i_2\|^{\alpha}}\prod_{j=1}^{M(k)}\frac{1}{\|i_j-i_{j+1}\|^{\alpha/2}}\frac{1}{\|i'_j-i'_{j+1}\|^{\alpha/2}}. \tag{4.32}$$

We are now precisely in the setting of Cipriani et al. [2025], and in particular, following the ideas from Cipriani et al. [2025, p24] and using Fact 4.4.2, we obtain that the right-hand side of (4.32) is of order $\mathrm{O}_N(c_N^{-1})$. For t matchings in $\mathbf{i},\mathbf{i}'$, the order is $\mathrm{O}(c_N^{-t})$, giving us that (4.26) is of order $\mathrm{O}(c_N^{-1})$ when k is even.

The argument for the case where k is odd is similar. Since the optimal order is achieved when we take $\mathbf{i}\setminus\{i_1,i_2\}\in \mathrm{Cat}_{\pi}$ and $\mathbf{i}'\setminus\{i'_1,i'_2\}\in\mathrm{Cat}_{\pi'}$, with $\pi,\pi'\in NC_2(M(k))$, one cannot construct these partitions with k being odd with the restriction from (4.18) imposing that $M(k)$ must be odd. Consequently, we have convergence in $\mathbb{P}$–probability of the moments of $\mathrm{ESD}(\mathbf{\Delta}_{N,g,c})$. Thus, we conclude that

$$\lim_{N\to\infty}\mathrm{tr}(\mathbf{\Delta}^k_{N,g,c}) = M_k \quad \text{in } \mathbb{P}\text{–probability}\,,$$

where

$$M_k = \begin{cases}\sum_{\mathcal{M}(k)}\sum_{\pi\in NC_2(M(k))} t(\tilde{G}_{\tilde{\pi}},W^m)\mathcal{E}(\tilde{\pi})\,, & k \text{ even},\\ 0\,, & k \text{ odd}\,,\end{cases} \tag{4.33}$$

where $\mathcal{M}(k)$ is the multiset of all numbers $(m_1,\dots,m_k,n_1,\dots,n_k)$ that appear in the expansion $(a+b)^k$ for two non-commutative variables a and b.

Step 3. We are now left to show that these moments uniquely determine a limiting measure. This follows from Chatterjee and Hazra [2022, Section 5.2.2], but we show the bounds for the sake of completeness.

First, from Chatterjee and Hazra [2022, Section 5.2.2], we have that $\mathcal{E}(\tilde{\pi}) \leq 2^k k!$. Next, observe from (4.28) that $|t(\tilde{G}_{\tilde{\pi}}, W^m)| \leq (m^2)^{\frac{k}{2}} = m^2$, since $W_i \leq m$ for all i and $\tilde{G}_{\tilde{\pi}}$ is a graph on $\frac{k}{2}+1$ vertices with $\frac{k}{2}$ edges. Lastly, $|NC_2(M(k)| \leq |NC_2(k)| = C_k$, where C_k is the k^{th} Catalan number, and moreover, $|\mathcal{M}(k)| \leq 2^k$. Combining these, we have

$$\beta_k := |M_k| \leq 2^k.C_k.m^k.2^k k! = (4m)^k C_k k!\,.$$

Using Sterling's approximation, we have

$$\frac{1}{k}\beta_k^{\frac{1}{k}} \leq \frac{4m}{(k+1)^{\frac{1}{k}}}.\frac{4\mathrm{e}^{-\left(1+\frac{1}{k}\right)}}{\pi^{\frac{1}{k}}}\,,$$

where π here is now the usual constant, and subsequently, we have

$$\limsup_{k\to\infty} \frac{1}{2k}\beta_{2k}^{\frac{1}{2k}} < \infty\,. \tag{4.34}$$

Equation (4.34) is a well-known criteria to show that the moments uniquely determine the limiting measure (see Lin [2017, Theorem 1]). This completes the argument. □

§4.5 Identification of the limit

§4.5.1 Removing geometry

In Section 4.4, we show the existence of a unique limiting measure ν_τ such that

$$\lim_{N\to\infty} \mathrm{ESD}(\mathbf{\Delta}_N^\circ) = \nu_\tau \qquad \text{in } \mathbb{P}\text{–probability}\,.$$

We have also shown that ν_τ is the limiting measure for the ESD of the Laplacian matrix $\widehat{\mathbf{\Delta}}_{N,g}$. In particular, through the proof of Proposition 4.4.1, we show that the limit $\nu_{\tau,m}$ is independent of the choice of α, and consequently, ν_τ is α–independent. We then use the idea of substituting $\alpha = 0$ from Cipriani et al. [2025, Section 6] in the matrix $\widehat{\mathbf{\Delta}}_{N,g}$, to obtain the Laplacian matrix $\mathbf{\Delta}_{N,g}^\circ$, which corresponds to the adjacency matrix $\mathbf{A}_{N,g}^\circ$ with entries given by

$$\mathbf{A}_{N,g}^\circ(i,j) = \begin{cases} \frac{\sqrt{W_i W_j}}{\sqrt{N}} G_{i\wedge j, i\vee j}, & i \neq j \\ 0, & i = j. \end{cases}$$

Then, $\lim_{N\to\infty} \mathrm{ESD}(\mathbf{\Delta}_{N,g}^\circ) = \nu_\tau$ in $\mathbb{P}$–probability. Recall that for all $1 \leq i \leq N$, $W_i^m := W_i \mathbf{1}_{W_i \leq m}$ for any $m \geq 1$. We can now apply Lemmas 4.3.5 and 4.3.6 to contruct a matrix $\mathbf{\Delta}_{N,g,c}^\circ = \mathbf{A}_{N,g,m}^\circ + Y_N^\circ$ such that

$$\limsup_{m\to\infty} \lim_{N\to\infty} \mathbb{P}\left(d_L(\mathrm{ESD}(\mathbf{\Delta}_{N,g}^\circ), \mathrm{ESD}(\mathbf{\Delta}_{N,g,c}^\circ)) > \delta\right) = 0\,,$$

where

$$\mathbf{A}^{\circ}_{N,g,m}(i,j) = \begin{cases} \frac{\sqrt{W_i^m W_j^m}}{\sqrt{N}} G_{i\wedge j, i\vee j}, & i \neq j \\ 0, & i = j, \end{cases}$$

and Y°_N is a diagonal matrix with entries

$$Y^{\circ}_N(i,i) = Z_i \sqrt{\frac{\sum_{k\neq i} W_i^m W_k^m}{N}}.$$

By Proposition 4.4.1, we have that $\lim_{N\to\infty} \mathrm{ESD}(\boldsymbol{\Delta}^{\circ}_{N,g,c}) = \nu_{\tau,m}$ in $\mathbb{P}$–probability. Thus, we begin by identifying $\nu_{\tau,m}$. To that end, consider the matrix $\widehat{\boldsymbol{\Delta}}^{\circ}_{N,g,c} := \mathbf{A}_{N,g,c} + \widehat{Y}^{\circ}_N$, with $\mathbf{A}_{N,g,c}$ as before, and $\widehat{Y}^{\circ}_N$ a diagonal matrix with entries

$$\widehat{Y}^{\circ}_N(i,i) = Z_i \sqrt{W_1^m} \sqrt{\mathbf{E}[W_1^m]}.$$

We now have the following lemma.

Lemma 4.5.1.
Let $\boldsymbol{\Delta}^{\circ}_{N,g,c}$ *and* $\widehat{\boldsymbol{\Delta}}^{\circ}_{N,g,c}$ *be as defined above. Then,*

$$\lim_{N\to\infty} \mathbb{P}\left(d_L(\mathrm{ESD}(\boldsymbol{\Delta}^{\circ}_{N,g,c}), \mathrm{ESD}(\widehat{\boldsymbol{\Delta}}^{\circ}_{N,g,c})) > \delta \right) = 0.$$

Proof. We apply Proposition 4.6.1 to obtain

$$\begin{aligned} &\mathbb{E}\left[d_L(\mathrm{ESD}(\boldsymbol{\Delta}^{\circ}_{N,g,c}), \mathrm{ESD}(\widehat{\boldsymbol{\Delta}}^{\circ}_{N,g,c}))^3 \right] \\ &\leq \frac{1}{N} \mathbb{E} \sum_{i=1}^N \left(Y^{\circ}_N(i,i) - \widehat{Y}^{\circ}_N(i,i) \right)^2 \\ &\leq \frac{1}{N} \mathbb{E}[Z_1^2] \mathbf{E}[W_1^m] \sum_{i=1}^N \mathbf{E}\left[\left(\frac{\sqrt{\sum_{k\neq i} W_k^m}}{\sqrt{N}} - \sqrt{\mathbf{E}[W_1^m]} \right)^2 \right] \\ &\leq \frac{m}{N} \sum_{i=1}^N \mathbf{E}\left[\left| \frac{\sum_{k=1}^N W_k^m}{N} - \mathbf{E}[W_1^m] \right| \right]. \end{aligned} \tag{4.35}$$

We have that $(W_i^m)_{i\in\mathbf{V}_N}$ is a bounded sequence of i.i.d. random variables, and in particular have finite variance. By the strong law of large numbers, we have that

$$\lim_{N\to\infty} \frac{\sum_{k=1}^N W_k^m}{N} = \mathbf{E}[W_1^m] \quad \mathbb{P}\text{–almost surely}.$$

However, by the boundedness of the weights, we have that $N^{-1}\sum_{i=1}^N W_i^m$ is uniformly bounded by m, which is integrable (with respect to $\mathbb{E}$). By the dominated convergence theorem, we have convergence in L^1, and consequently, (4.35) goes to 0 as $N \to \infty$. We conclude with Markov's inequality. □

We can now conclude that $\nu_{\tau,m}$ is the limiting measure of the ESD of the matrix $\widehat{\mathbf{\Delta}}^{\circ}_{N,g,c}$.

§4.5.2 Identification of the truncated measure

We have that

$$\lim_{N\to\infty} \mathrm{ESD}(\widehat{\mathbf{\Delta}}^{\circ}_{N,g,c}) = \nu_{\tau,m} \quad \text{in } \mathbb{P}\text{–probability}\,.$$

Notice that $\widehat{\mathbf{\Delta}}^{\circ}_{N,g,c}$ can be written as

$$\begin{aligned}\widehat{\mathbf{\Delta}}^{\circ}_{N,g,c} &= \mathbf{A}^{\circ}_{N,g,m} + \widehat{Y}^{\circ}_{N}\\ &= \mathrm{W}_m^{1/2}\left(\frac{1}{\sqrt{N}}\mathrm{G}\right)\mathrm{W}_m^{1/2} + \sqrt{\mathbf{E}[W_1^m]}\mathrm{W}_m^{1/4}\mathrm{Z}\mathrm{W}_m^{1/4}\,,\end{aligned}$$

where $\mathrm{W}_m = \mathrm{Diag}(W_1^m, \ldots, W_N^m)$, G is a standard Wigner matrix with i. i. d $\mathrm{N}(0,1)$ entries above the diagonal and 0 on the diagonal, and Z is a diagonal matrix with i.i.d. $\mathrm{N}(0,1)$ entries.

First, we need to show that

$$\begin{aligned}&\lim_{N\to\infty} \mathrm{ESD}\left(\mathrm{W}_m^{1/2}\left(\frac{1}{\sqrt{N}}\mathrm{G}\right)\mathrm{W}_m^{1/2} + \sqrt{\mathbf{E}[W_1^m]}\mathrm{W}_m^{1/4}\left(\frac{1}{\sqrt{N}}\mathrm{Z}\right)\mathrm{W}_m^{1/4}\right)\\ &= \mathcal{L}\left(T_{W_m}^{1/2}T_sT_{W_m}^{1/2} + \sqrt{\mathbf{E}[W_m]}T_{W_m}^{1/4}T_gT_{W_m}^{1/4}\right) \text{ weakly in probability}\,.\end{aligned}$$

This easily follows by retracing the arguments in the proof of [Chakrabarty et al., 2021b, Theorem 1.3] and using the Lemma 4.6.6 presented in the appendix. This shows that

$$\nu_{\tau,m} = \mathcal{L}\left(T_{W_m}^{1/2}T_sT_{W_m}^{1/2} + \sqrt{\mathbf{E}[W_m]}T_{W_m}^{1/4}T_gT_{W_m}^{1/4}\right)\,.$$

§4.5.3 Identification of the limiting measure

We now conclude with the proof of Theorem 4.2.5.

Proof of Theorem 4.2.5. Consider the measure μ_{W^m} and μ_W which are laws of $W^m = W\mathbf{1}_{W\leq m}$ and W respectively. Also consider μ_g and μ_s to be the laws of the standard Gaussian and semicircle law, respectively. We have for all $t\in\mathbb{R}$,

$$|F_{\mu_{W^m}}(t) - F_{\mu_W}(t)| \leq \varepsilon \tag{4.36}$$

for m large enough. Hence from [Bercovici and Voiculescu, 1993, Theorem 3.9] there exists a W^* probability space (A,φ) and self-adjoint operators T_{W^m}, T_W, T_g

and T_s affiliated to (A, φ) and projection $p \in A$ such that $pT_{W^m}p = pT_W p$ and $\varphi(p) \geq 1-\varepsilon$. Also the spectral laws of T_{W^m}, T_W, T_g and T_s are given respectively by μ_{W^m}, μ_W, μ_g and μ_s respectively.

We can consider the commutative subalgebra generated by $\{T_{W^m}, T_g\}$. Then using [Bercovici and Voiculescu, 1993, Proposition 4.1], it is possible to generate random variable from $\{T_{W^m}, T_g\}$ that is free from T_s. Analogously, one can do the same for $\{T_W, T_g\}$.

Consider a self-adjoint polynomial Q_m of $\{T_{W^m}, T_g, T_s\}$ and let the law of this polynomial be given by ν_m. Similarly, let Q be the same self-adjoint polynomial of $\{T_W, T_g, T_s\}$ and ν be its law. Then using $pT_{W^m}p = pT_W p$ and (4.36) and [Bercovici and Voiculescu, 1993, Corollary 4.5 and Theorem 3.9] we have that $d_\infty(\nu_m, \nu) \leq \varepsilon$ for all m large enough. Here d_∞ is the Kolmogorov distance. Picking $Q(x, y, z) = x^{1/2}yx^{1/2} + cx^{1/4}zx^{1/4}$ for some constant $c = \sqrt{\mathbf{E}[W]}$, completes the proof. $\square$

§4.6 Appendix

In this section we collect some technical lemmas that are used in the proofs of our main results.

§4.6.1 Technical lemmas

For bounding the d_L distance between the ESDs of two matrices, we will need the following inequality, due to Hoffman and Wielandt (see Bai and Silverstein [2010, Corollary A.41]).

Proposition 4.6.1 (Hoffman-Wielandt inequality).
Let $\mathbf{A}$ and $\mathbf{B}$ be two $N \times N$ normal matrices and let $\mathrm{ESD}(\mathbf{A})$ and $\mathrm{ESD}(\mathbf{B})$ be their ESDs, respectively. Then,

$$d_L\left(\mathrm{ESD}(\mathbf{A}), \mathrm{ESD}(\mathbf{B})\right)^3 \leq \frac{1}{N} \operatorname{Tr}\left[(\mathbf{A}-\mathbf{B})(\mathbf{A}-\mathbf{B})^*\right]. \tag{4.37}$$

Here $\mathbf{A}^*$ denotes the conjugate transpose of $\mathbf{A}$. Moreover, if $\mathbf{A}$ and $\mathbf{B}$ are two Hermitian matrices of size $N \times N$, then

$$\sum_{i=1}^{N} (\lambda_i(\mathbf{A}) - \lambda_i(\mathbf{B}))^2 \leq \operatorname{Tr}[(\mathbf{A}-\mathbf{B})^2]. \tag{4.38}$$

The next two straightforward lemmas control the tail of the product of two Pareto random variables and the expectation of a truncated Pareto.

Lemma 4.6.2.
Let X and Y be two independent Pareto r.v.'s with parameters β_1 and β_2 respectively, with $\beta_1 \leq \beta_2$. There exist constants $c_1 = c_1(\beta_1, \beta_2) > 0$ and $c_2 = c_2(\beta_1) > 0$ such that

$$\mathbf{P}(XY > t) = \begin{cases} c_1 t^{-\beta_1} & \text{if } \beta_1 < \beta_2 \\ c_2 t^{-\beta_1} \log t & \text{if } \beta_1 = \beta_2. \end{cases}$$

Lemma 4.6.3.
Let X be a Pareto random variable with law $\mathbf{P}$ and parameter $\beta > 1$. For any $m > 0$ it holds

$$\mathbf{E}\left[X \mathbb{1}_{X \geq m}\right] = \frac{\beta}{(\beta - 1)} m^{1-\beta}.$$

We state one final auxiliary lemma related to the approximation of sums by integrals.

Lemma 4.6.4.
Let $\beta \in (0, 1]$. Then there exists a constant $c_1 = c_1(\beta) > 0$ such that

$$\frac{1}{N} \sum_{i \neq j \in \mathbf{V}_N} \frac{1}{\|i - j\|^\beta} \sim c_1 \max\{N^{1-\beta}, \log N\}. \tag{4.39}$$

If instead $\beta > 1$, there exists a constant $c_2 > 0$ such that

$$\frac{1}{N} \sum_{i \neq j \in \mathbf{V}_N} \frac{1}{\|i - j\|^\beta} \sim c_2 \,.$$

We end this section by quoting, for the reader's convenience, the following lemma from Chakrabarty et al. [2016, Fact 4.3].

Lemma 4.6.5.
Let (Σ, d) be a complete metric space, and let $(\Omega, \mathcal{A}, P)$ be a probability space. Suppose that $\left(X_{mn} : (m, n) \in \{1, 2, \ldots, \infty\}^2 \backslash \{\infty, \infty\}\right)$ is a family of random elements in Σ, that is, measurable maps from Ω to Σ, the latter being equipped with the Borel σ-field induced by d. Assume that

(1) for all fixed $1 \leq m < \infty$

$$\lim_{n \to \infty} d\left(X_{mn}, X_{m\infty}\right) = 0 \text{ in } P\text{-probability.}$$

(2) For all $\varepsilon > 0$,

$$\lim_{m \to \infty} \limsup_{n \to \infty} P\left(d\left(X_{mn}, X_{\infty n}\right) > \varepsilon\right) = 0.$$

Then, there exists a random element $X_{\infty\infty}$ of Σ such that

$$\lim_{m\to\infty} d\left(X_{m\infty}, X_{\infty\infty}\right) = 0 \text{ in } P\text{-probability} \tag{4.40}$$

and

$$\lim_{n\to\infty} d\left(X_{\infty n}, X_{\infty\infty}\right) = 0 \text{ in } P\text{-probability.}$$

Furthermore, if $X_{m\infty}$ is deterministic for all m, then so is $X_{\infty\infty}$, and (4.40) simplifies to

$$\lim_{m\to\infty} d\left(X_{m\infty}, X_{\infty\infty}\right) = 0. \tag{4.41}$$

Lemma 4.6.6 (Fact A.4 Chakrabarty et al. [2021b]).
Suppose that W_N is an $N \times N$ scaled standard Gaussian Wigner matrix, i.e., a symmetric matrix whose upper triangular entries are i.i.d. normal with mean zero and variance $1/N$. Let D_N^1 and D_N^2 be (possibly random) $N \times N$ symmetric matrices such that there exists a deterministic C satisfying

$$\sup_{N\geq 1, i=1,2} \left\|D_N^i\right\| \leq C < \infty$$

where $\|\cdot\|$ denotes the usual matrix norm (which is same as the largest singular value for a symmetric matrix). Furthermore, assume that there is a W^*-probability space $(\mathcal{A}, \varphi)$ in which there are self-adjoint elements d_1 and d_2 such that, for any polynomial p in two variables, it

$$\lim_{N\to\infty} \frac{1}{N} \operatorname{Tr}\left(p\left(D_N^1, D_N^2\right)\right) = \varphi\left(p\left(d_1, d_2\right)\right) \text{ a.s.}$$

Finally, suppose that $\left(D_N^1, D_N^2\right)$ is independent of W_N. Then there exists a self-adjoint element s in $\mathcal{A}$ (possibly after expansion) that has the standard semicircle distribution and is freely independent of (d_1, d_2), and is such that

$$\lim_{N\to\infty} \frac{1}{N} \operatorname{Tr}\left(p\left(W_N, D_N^1, D_N^2\right)\right) = \varphi\left(p\left(s, d_1, d_2\right)\right) \text{ a.s.}$$

for any polynomial p in three variables.

Lemma 4.6.7 (Wick's formula).
Let $(X_1, X_2, \ldots, X_n)$ be a real Gaussian vector, then, and $\mathcal{P}_2(k)$ the set of pair partitions of $[k]$. Then, for any $1 \leq k \leq n$,

$$\mathbb{E}[X_{i_1} \cdots X_{i_k}] = \sum_{\pi\in\mathcal{P}_2(k)} \prod_{(r,s)\in\pi} \mathbb{E}[X_{i_r} X_{i_s}]\,. \tag{4.42}$$

We borrow the following definition from Avena et al. [2023, Definition 2.3].

Definition 4.6.8 (Graph associated to a partition).
For a fixed $k \geq 1$, let γ denote the cyclic permutation $(1, 2, \dots, k)$. For a partition π, we define $G_{\gamma\pi} = (\mathbf{V}_{\gamma\pi}, E_{\gamma\pi})$ as a rooted, labelled directed graph associated with any partition π of $[k]$, constructed as follows.

- *Initially consider the vertex set $\mathbf{V}_{\gamma\pi} = [k]$ and perform a closed walk on $[k]$ as $1 \to 2 \to 3 \to \cdots \to k \to 1$ and with each step of the walk, add an edge.*
- *Evaluate $\gamma\pi$, which will be of the form $\gamma\pi = \{V_1, V_2, \dots, V_m\}$ for some $m \geq 1$ where $\{V_i\}_{1 \leq i \leq m}$ are disjoint blocks. Then, collapse vertices in $\mathbf{V}_{\gamma\pi}$ to a single vertex if they belong to the same block in $\gamma\pi$, and collapse the corresponding edges. Thus, $\mathbf{V}_{\gamma\pi} = \{V_1, \dots, V_m\}$.*
- *Finally root and label the graph as follows.*
  - Root: *we always assume that the first element of the closed walk (in this case '1') is in V_1, and we fix the block V_1 as the root.*
  - Label: *each vertex V_i gets labelled with the elements belonging to the corresponding block in $\gamma\pi$.*

Chapter Four

CHAPTER 5

Discussion and future directions

In this short chapter, we show some simulations of spectral distributions of random graph models discussed in the previous chapters. We focus on the cases not covered by our main results and compare them with the previous simulations. This leaves us with many open directions for the future.

§5.1 Introduction

This thesis establishes new results in the study of spectral analysis of inhomogeneous random graph models, providing further insight into the area and opening the door to multiple directions for further research. We discuss some of the open directions in this chapter.

Chapter 2 extends results of the homogeneous Erdős-Rényi random graph to the inhomogeneous setting, providing a characterisation of the limiting spectral measure of the adjacency matrix. While the limit is not explicitly known, we provide a combinatorial expression for the moments and an analytic description of the Stieltjes transform, which complements the random graph description that one can obtain from Bordenave and Lelarge [2010], since the model has a local weak limit that is a multi-type branching process (see van der Hofstad [2024]). However, the results are restricted to the setting where vertex weights (w_i) are deterministic, and a natural extension would be to consider random weights with f an almost surely continuous connectivity function. As seen in Remark 2.3.13, our proof techniques require that W (which is the random variable such that $w_{o_N} \xrightarrow{d} W$) is compactly supported. What also remains unknown is the rate of convergence of the measure μ_λ to μ_f, even in the homogeneous setting where $\mu_f = \mu_{sc}$. These questions naturally arise due to the works of Bai and Silverstein [2010], Augeri [2025], Jung and Lee [2018], Tran et al. [2013], however, we believe that the fixed-point equation as in Theorem 2.3.9 needs further analysis to describe the rate of convergence of Stieltjes transforms. Furthermore, the extension of results from Bordenave et al. [2011], Coste and Salez [2021], Salez [2020] remains an open question in the inhomogeneous setting.

Spectral properties of kernel-based random graphs (as introduced in Jorritsma et al. [2023]), and in particular of the scale-free percolation model, are a relatively untouched topic. Chapters 3 and 4 now provide a foundation for this topic. We consider random Pareto weights on the vertices, with tail exponent $\tau - 1$, $\tau > 1$. The spectral properties of the adjacency matrix are described in Chapter 3 for kernel-based random graphs with the kernel structure

$$\kappa(x, y) = (x \vee y)(x \wedge y)^\sigma$$

where $\sigma < \tau - 2$. One extension would be to consider a far more general kernel. The above multiplicative structure simplifies calculations significantly. We also restrict ourselves to $\tau > 2$, where the weights have finite mean. This is crucial in the truncation step, since for a truncation at $m > 1$, the error rate is $m^{2-\tau}$. We believe that this is a technical assumption. Analogously, we do not consider $\sigma \geq \tau - 1$. Due to the rank one nature of the kernel when $\sigma = 1$, we can characterise the limiting measure using tools from free probability.

Consequently, we observe an interesting tail asymptotic, where the measure has a power-law tail with exponent $2(\tau - 1)$. When $\sigma \neq 1$, this becomes more challenging, and we believe that the tail may not have a power-law decay but rather a more complicated behaviour. A more interesting direction is the case when $\alpha > 1$, with $\tau > 2$. This yields a sparse random graph, for which the existence of a limiting measure is guaranteed by Bordenave and Lelarge [2010]. However, since the local weak limit of the random graph is not locally tree-like, there is no description of the measure. This will require a novel approach, and the spectrum of the centred and non-centred adjacency matrices will differ.

The Laplacian matrix of the scale-free percolation model is analysed in Chapter 4. The existence of the limiting measure is achieved by computing the moments, which is far more challenging than computing the moments of the adjacency matrix. We believe that this will be the primary challenge when attempting to extend the results to the kernel as described in Chapter 3. We also restrict ourselves to $\tau > 3$, and an extension to $\tau > 2$ will require better bounds in the Gaussianisation step, as well as ensuring that the decoupling of the diagonal holds. Decoupling is an essential step for the moment method, without which the approach becomes highly complicated.

Outline of the chapter

The first part of the chapter is devoted to the homogeneous Erdős-Rényi random graph $\mathrm{ER}_N(p)$ with $p = \lambda/N$. We simulate the spectrum of the adjacency matrix for increasing λ to illustrate that, for a λ such that $1 < \lambda < \log N$, μ_λ starts taking the shape of μ_{sc} (with possible atoms). We then simulate the spectrum of the Laplacian matrix, moving from the sparse to the dense case, and show why centring becomes essential as the graph becomes more dense.

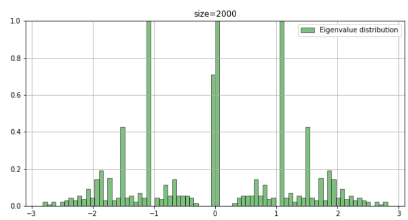
The second part of the chapter showcases simulations for the scale-free percolation model. We simulate the spectra of the adjacency matrix for a combination of α and τ, to analyse the cases where $\alpha(\tau-1) > 1$ and $\alpha(\tau-1) < 1$. We also simulate the spectrum of the long-range percolation model with increasing α, to illustrate the sparse setting. We compare the resolvent matrices of the long-range percolation model, GOE model, and $\mathrm{ER}_N(\lambda/N)$ with $\lambda > 1$. We conclude with the centred Laplacian matrix of the scale-free percolation model for varying τ, namely the infinite mean regime, the infinite variance regime, and the finite variance regime.

§5.2 Erdős-Rényi Random graph

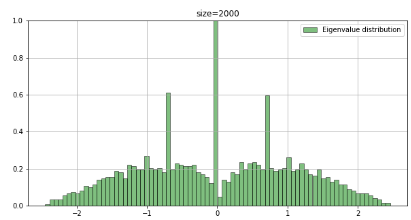
Consider the homogeneous Erdős-Rényi random graph $\mathrm{ER}_N(p)$ on N vertices, with $p = \lambda/N$ for some $\lambda \in (0, \infty)$. If $\mathbf{A}_{\mathbb{G}_N}$ is the adjacency matrix of this graph, then define $\mathbf{A}_N = \lambda^{-1/2}\mathbf{A}_{\mathbb{G}_N}$ as the scaled adjacency matrix. This falls under the setting of Chapter 2 as a special case. In particular, Theorem 2.3.7 (and also results from Jung and Lee [2018], Bordenave and Lelarge [2010], Tran et al. [2013]) tells us that there exists a unique limiting measure μ_λ such that $\lim_{N\to\infty} \mathrm{ESD}(\mathbf{A}_N) = \mu_\lambda$ in probability, and $\mu_\lambda \implies \mu_{sc}$ as $\lambda \to \infty$. Further, from Bordenave and Lelarge [2010], if $\mathbf{\Delta}_N$ is the scaled Laplacian matrix of this graph, then there exists a unique limiting measure ν_λ such that $\lim_{N\to\infty} \mathrm{ESD}(\mathbf{\Delta}_N) = \nu_\lambda$ in probability. It follows from Khorunzhy et al. [2004] that $\nu_\lambda \implies \mu_{sc} \boxplus \mu_g$, where μ_g is the Gaussian law.

§5.2.1 Adjacency matrix

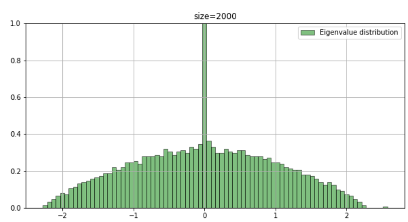
Consider the scaled adjacency matrix $\mathbf{A}_N$ of this graph. In Chapter 2, we see that in the limit $N \to \infty$, the ESD of $\mathbf{A}_N$ and that of the centred matrix $\mathbf{A}_N - \mathbb{E}[\mathbf{A}_N]$ are close in probability, and so we can study the non-centred matrix directly.



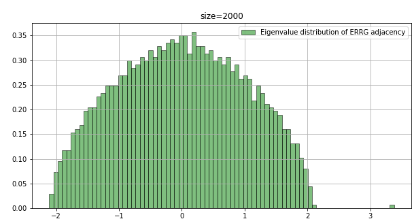
(a) $\lambda = 0.9$.



(b) $\lambda = 1.9$.



(c) $\lambda = 3.5$.



(d) $\lambda = 9$.

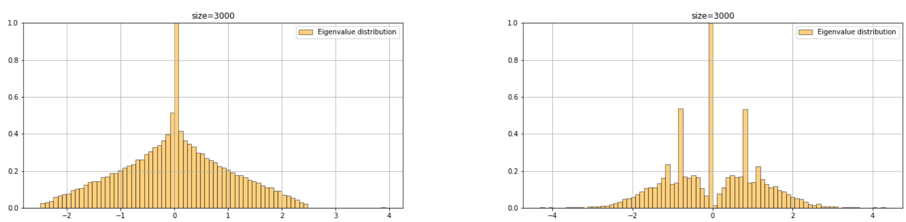
Figure 5.1: Eigenvalue distributions of the adjacency of $\mathrm{ER}_N(\lambda/N)$ *with* $N = 2000$.

In Figure 5.1, we see the eigenvalue distributions of this matrix with $N = 2000$ for varying values of λ. For $\lambda < 1$, we observe "spikes", indicating that the measure has many atoms (in line with Salez [2020]). For $\lambda > 1$, we observe a continuous part, indicating the presence of a density (in line with Arras and Bordenave [2023]). When $\lambda > \log N$ ($\lambda = 9$), we observe a distribution that resembles the semicircle law, with an outlier that is the largest eigenvalue, which is of the order $\sqrt{\lambda}$ (see Erdős et al. [2013]).

The interesting case is when λ is "large", but smaller than $\log N$. We observe at $\lambda = 3.5$ for $N = 2000$ that there is a spike at the eigenvalue 0, indicating the presence of an atom. However, the remaining distribution begins to take the shape of a semicircle distribution. This indicates that the rate of convergence in λ is relatively fast. While we were not able to prove this, we believe that the metric defined by Stieltjes transforms as in Augeri [2025] can aid in determining this rate of convergence. Through moments, we heuristically see a possible candidate for the convergence rate. The $2k-$th moment of μ_λ is

$$\int_{\mathbb{R}} x^{2k} \mu_\lambda(\mathrm{d}\, x) = C_k + \mathrm{Err}(\lambda^{-1}) = \int_{\mathbb{R}} x^{2k} \mu_{sc}(\mathrm{d}\, x) + \mathrm{Err}(\lambda^{-1}) \,,$$

where $\mathrm{Err}(\lambda^{-1})$ is an error term with leading order λ^{-1} and C_k is the $k-$th Catalan number. We leave the optimal rate of convergence as an open problem.



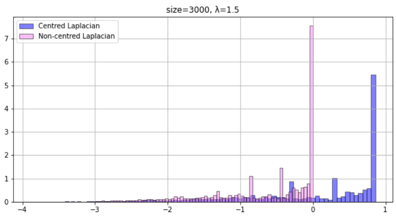
(a) Degrees $d_i \sim \mathrm{Unif}[1, 10]$.

(b) Degrees $d_i \sim \mathrm{Pareto}(\tau - 1)$*, with* $\tau = 3.5$.

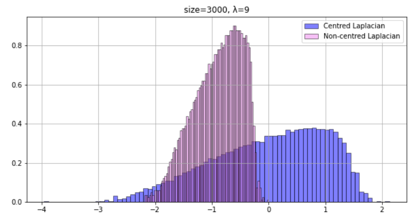
Figure 5.2: Spectral distributions of adjacency matrices of IER models, with edge probability $p_{ij} = \frac{d_i d_j}{m_1 + d_i d_j} \wedge 1$*, where* $(d_i)_{i=1}^N$ *is a given degree sequence and* $m_1 = \sum_{i=1}^N d_i$*.* $N = 3000$*.*

§5.2.2 Laplacian matrix

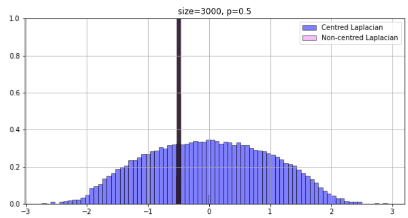
From Bordenave and Lelarge [2010], we have the existence of ν_λ for the ESD of the graph Laplacian when the graph is sparse. We see in Figure 5.3 that the spectra of the centred and non-centred Laplacian differ significantly, in particular when the sparsity parameter increases. For dense graphs with a fixed p, the spectrum of the Laplacian is a Dirac mass at p (see Bryc et al. [2006]), which is what we observe in Figure 5.3c. It is only meaningful to study the spectrum of the centred Laplacian in the dense setting. Understanding the ESD and identifying the limiting measure in the general inhomogeneous setting is still an open problem. Also, it is unclear whether for any $\lambda > 0$, the limiting measure always has an absolutely continuous spectrum. It would be interesting to derive the behaviour of the atoms for $\lambda < 1$.



(a) $p = \lambda/N$, $\lambda = 1.5$.



(b) $p = \lambda/N$, $\lambda = 9$.



(c) $p = 0.5$, with the non-centred Laplacian scaled by N instead of $\sqrt{Np(1-p)}$.

Figure 5.3: Spectral distributions of the Laplacian matrices of ERRG, with $N = 3000$.

§5.3 Scale-Free percolation

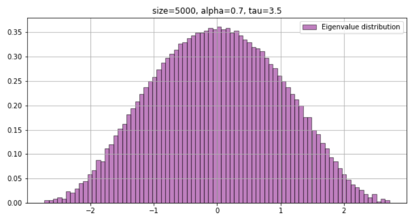
Let us consider the model from Chapter 4, which is a special case of the model from Chapter 3. We take the discrete torus on N vertices and an i.i.d. sequence of Pareto weights $(W_i)_{i=1}^N$. Conditionally on the weights, we add edges independently with probability

$$p_{ij} = \frac{W_i W_j}{\|i-j\|^\alpha} \wedge 1 ,$$

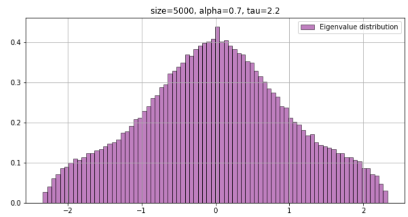
where $\alpha > 0$ is a parameter of choice and $\|\cdot\|$ is the torus distance. In the dense case, we scale the adjacency and Laplacian matrices with the scaling factor $c_N \sim N^{1-\alpha}$ for $\alpha < 1$. In the sparse case, when $\alpha > 1$, we scale by a constant scaling $\zeta(\alpha)$, which is the Riemann-Zeta function evaluated at α.

§5.3.1 Adjacency matrix

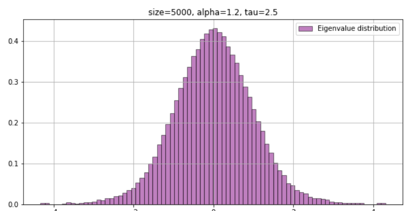
The degrees of vertices in the model are heavy-tailed with parameter $\gamma := \alpha(\tau - 1)$ (see Deijfen et al. [2013], Cipriani and Salvi [2024]). We simulate the eigenvalue distribution of the scaled adjacency matrix $\mathbf{A}_N$ for the regimes $\gamma < 1$ and $\gamma > 1$. For $\gamma > 1$, we have two sub-regimes, namely when $\alpha < 1$ and $\alpha > 1$, and similarly for $\gamma < 1$, giving us a total of 4 regimes, as in Figure 5.4. While we have theoretical results for Figure 5.4a, wherein we also see that the centred and non-centred adjacency matrices are spectrally close, we believe extension to the setting simulated in Figure 5.4b should be possible with some modifications to deal with infinite-mean weights, though the spectrum may differ in the centred and non-centred cases. The eigenvalue distributions in Figures 5.4a and 5.4c look similar, where the parameter $\gamma > 1$. Similarly, the eigenvalue distribution in Figures 5.4b and 5.4d have a similar shape, where $\gamma < 1$. This indicates that γ possibly plays a role in the limiting spectrum, though we do not see this in Chapters 3 and 4. We believe that the limiting measures exist in all regimes after appropriate scaling and may be random in certain cases.



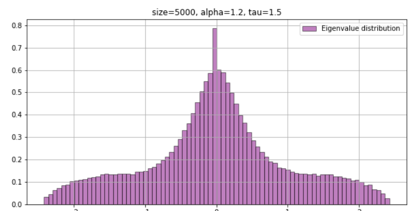
(a) $\alpha = 0.7$, $\tau = 3.5$.



(b) $\alpha = 0.7$, $\tau = 2.2$.

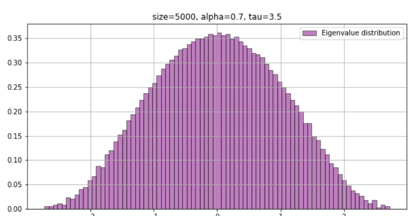


(c) $\alpha = 1.2$, $\tau = 2.5$.

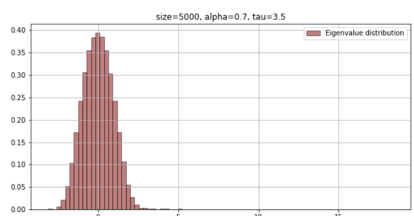


(d) $\alpha = 1.2$, $\tau = 1.5$.

Figure 5.4: Spectral distributions of the centred adjacency matrices of scale-free percolation, with $N = 5000$.



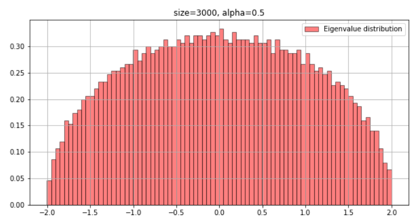
(a) Centred adjacency.



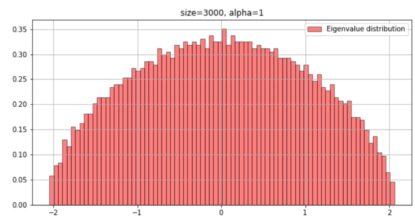
(b) Non-centred adjacency.

Figure 5.5: Spectral distributions of the centred and non-centred adjacency matrices of scale-free percolation, with $N = 5000$, $\alpha = 0.7$, $\tau = 3.5$.

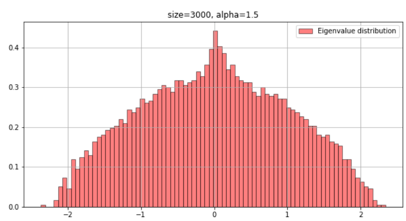
For the long-range percolation model (that is, for $W_i \equiv 1$) in Figure 5.6, we observe the semicircle law when $\alpha < 1$. At $\alpha = 1$, the shape is still semicircle-like, though we observe some concentration towards the centre. For $\alpha \in (1, 2)$, we still observe the presence of a density, with possible atoms at 0, and $\alpha = 2$, this density begins to break down. For $\alpha > 2$, where the model behaves similarly to bond percolation, the spectrum starts to break down. Such transitions in forms of percolative behaviour in different regimes have already been observed in long-range percolation theory (Berger [2002]). It would be interesting to see this behaviour in the spectrum also.



(a) $\alpha = 0.5$.



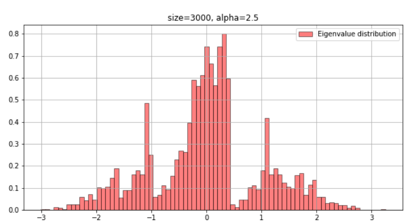
(b) $\alpha = 1$.



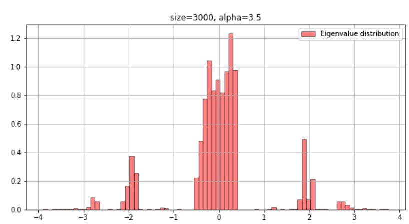
(c) $\alpha = 1.5$.

(d) $\alpha = 2$.

Figure 5.6: Spectral distributions of the centred adjacency matrices of long-range percolation, with $N = 3000$.



(e) $\alpha = 2.5$.



(f) $\alpha = 3.5$.

Figure 5.6: (continued)

§5.3.2 Resolvent Matrix

Recall that for a random matrix $\mathbf{A}_N$, one can define the resolvent as $\mathrm{R}_{\mathbf{A}_N}(z) = (\mathbf{A}_N - z\,\mathrm{I})^{-1}$. For some models, there is a concentration on the diagonal of the resolvent matrix, which makes computation easier. For example, let $\mathbf{A}_N$ be the GOE, with $\mathbf{A}_N(i,j) = \mathbf{A}_N(j,i) \stackrel{d}{=} N^{-1/2}\mathrm{N}(0,1)$. With the following heuristic, we can see how concentration on the diagonal of the resolvent occurs:

With Schur's complement formula from Bordenave [2019], we have

$$r_{ii} = -\frac{1}{z + \sum_{j,k\neq i} \tilde{r}_{jk}\mathbf{A}_N(i,j)\mathbf{A}_N(i,k)},$$

where $\tilde{r}_{ij} := \mathrm{R}_{\mathbf{A}_N^{(i)}}(z) = (\mathbf{A}_N^{(i)} - z\,\mathrm{I})^{-1}$, and $\mathbf{A}_N^{(i)}$ is $\mathbf{A}_N$ with the $i-$th row and column deleted. We briefly recall the heuristics from Chapter 2.1. Taking expectation, we get

$$\begin{aligned}
\mathbb{E}[r_{ii}] &= -\mathbb{E}\left[\frac{1}{z + \sum_{j,k\neq i} \tilde{r}_{jk}\mathbf{A}_N(i,j)\mathbf{A}_N(i,k)}\right] \\
&\approx -\frac{1}{z + \mathbb{E}\left[\sum_{j,k\neq i} \tilde{r}_{jk}\mathbf{A}_N(i,j)\mathbf{A}_N(i,k)\right]} \\
&\approx -\frac{1}{z + \mathbb{E}\left[\sum_{j\neq i} r_{jj}\mathbf{A}_N(i,j)^2\right]} = -\frac{1}{z + \mathrm{tr}(\mathrm{R}_{\mathbf{A}_N}(z))},
\end{aligned}$$

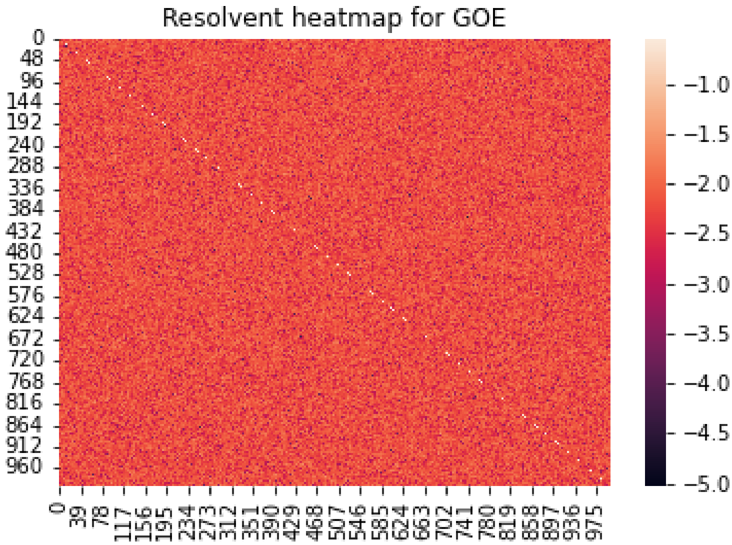
and so for N large, the diagonal terms are in some sense "replaced" by the Stieltjes transform of μ_{sc}, with the off-diagonal terms vanishing as $N \to \infty$.

This concentration may not happen in other models. Notably, in the sparse case of the long-range percolation model, we see that there seems to be significant mass at the off-diagonal terms.

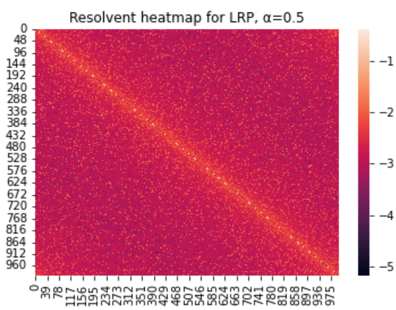
This suggests that understanding the local convergence for these models is a significant challenge, as most methods require a critical understanding of the resolvent matrix, which roughly concentrates around the diagonal for the classical Gaussian models (Anderson et al. [2010], Bordenave [2019]).

§5.3.3 Laplacian matrix

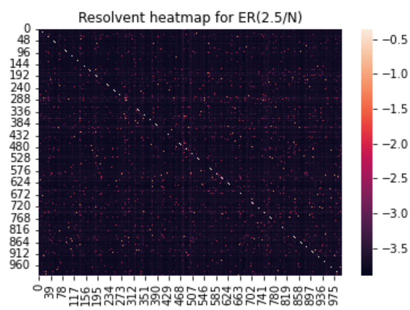
For the scaled Laplacian matrix of the scale-free percolation model, we have theoretical results for the existence of a limiting distribution when the weights have finite variance, as in Figure 5.8a. We observe that as τ decreases, that is, the weights become more heavy-tailed, the mass at 0 for the measure increases as well, and when we have infinite mean weights, as in Figure 5.8c, there is an



(a) GOE.



(b) LRP, with $\alpha = 0.5$.

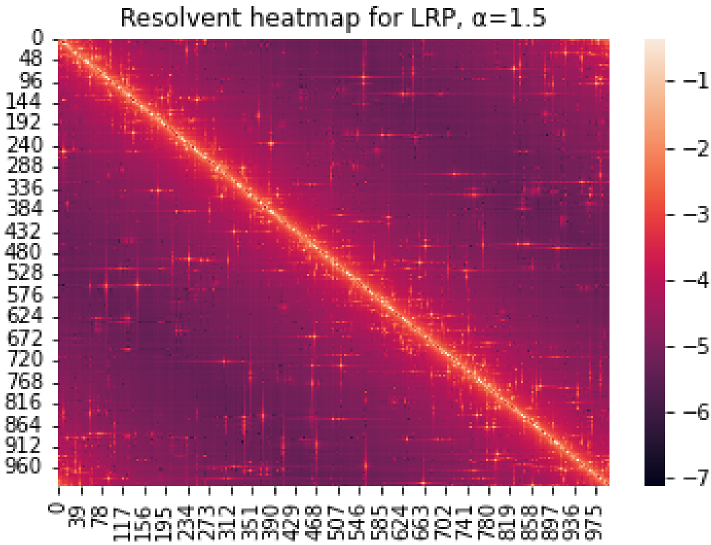


(c) $\mathrm{ER}_N(2.5/N)$.

Figure 5.7: Logarithmic resolvent heat maps for centred adjacencies of LRP, GOE, *and ERRG models, with* $N = 1000$. *We take* $z = 1 + 2i$, *evaluate the resolvent, and compute the absolute values of the entries. We add* N^{-2} *to each entry and compute the logarithm of the value and plot a heat map.*

indication of an atom present at 0. We expect the results to be true under the assumption of finite mean for the weights. We remark that the Gaussianisation and decoupling steps may fail when we have infinite variance for the weights, and so, a new approach has to be taken to tackle the problem. We leave the case of finite mean and infinite variance open.

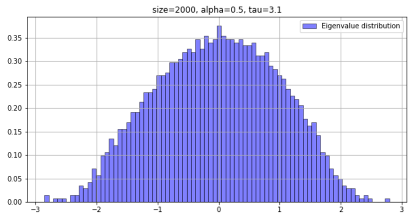
In Figure 5.9, we simulate the eigenvalue distributions of the centred Laplacian matrix of the LRP and SFP models, when $\alpha > 1$. We observe that for the LRP, the spectrum breaks down when $\alpha > 2$, as in Figure 5.9b, whereas



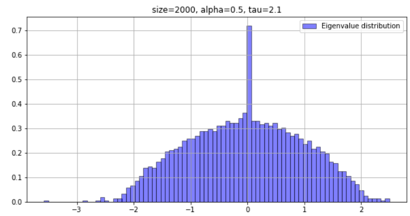
(d) LRP, with $\alpha = 1.5$.

Figure 5.7: (continued)

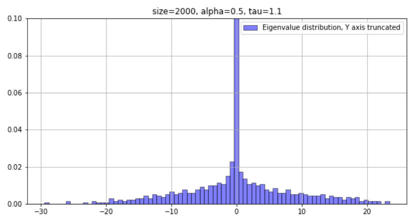
we observe a density-like shape in Figure 5.9a. For the SFP models, we keep $\alpha = 1.5$ fixed, and observe that the distribution skews less when the weights become more heavy-tailed and the graph becomes denser, as in Figure 5.9d.



(a) $\tau = 3.1$.

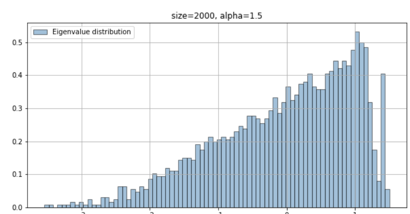


(b) $\tau = 2.1$.

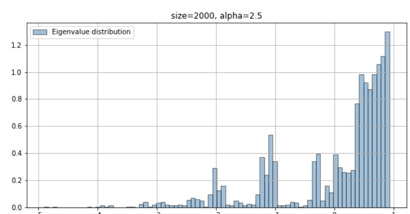


(c) $\tau = 1.1$.

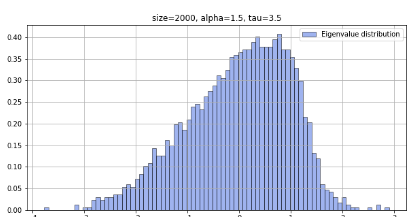
Figure 5.8: Spectral distributions of the centred Laplacian matrix of scale-free percolation, with $N = 2000$, $\alpha = 0.5$.



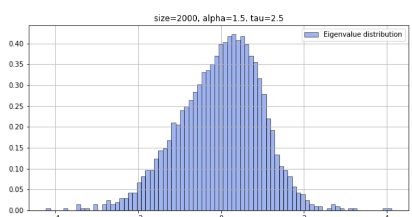
(a) LRP, $\alpha = 1.5$.



(b) LRP, $\alpha = 2.5$.



(c) SFP, $\alpha = 1.5$, $\tau = 3.5$.



(d) SFP, $\alpha = 1.5$, $\tau = 2.5$.

Figure 5.9: Spectral distributions of the centred Laplacian matrix of long-range and scale-free percolation, with $N = 2000$.

Bibliography

E. Abbe. Community detection and stochastic block models: recent developments. *Journal of Machine Learning Research (JMLR)*, 18(1):Paper No. 177, 86, 2017.

E. Abbe, E. Boix-Adsera, P. Ralli, and C. Sandon. Graph powering and spectral robustness. *SIAM Journal on Mathematics of Data Science*, 2(1):132–157, 2020.

M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover, New York, ninth dover printing, tenth gpo printing edition, 1964.

O. Ajanki, L. Erdős, and T. Krüger. *Quadratic vector equations on complex upper half-plane*, volume 261. Memoirs of American Mathematical Society, 2019.

P. Akara-pipattana and O. Evnin. Random matrices with row constraints and eigenvalue distributions of graph laplacians. *Journal of Physics. A. Mathematical and Theoretical*, 56(29):Paper No. 295001, 28, 2023.

D. Aldous. Brownian excursions, critical random graphs and the multiplicative coalescent. *The Annals of Probability*, 25(2):812–854, 1997.

D. Aldous and R. Lyons. Processes on unimodular random networks. *Electronic Journal of Probability*, 12:no. 54, 1454–1508, 2007.

N. Anantharaman, M. Ingremeau, M. Sabri, and B. Winn. Absolutely continuous spectrum for quantum trees. *Communications in Mathematical Physics*, 383:537–594, 2021.

G. W. Anderson, A. Guionnet, and O. Zeitouni. *An introduction to random matrices*, volume 118. Cambridge University Press, 2010.

O. Arizmendi and V. Pérez-Abreu. The S-transform of symmetric probability measures with unbounded supports. *Proceedings of the American Mathematical Society*, 137(9):3057–3066, 2009.

A. Arras and C. Bordenave. Existence of absolutely continuous spectrum for galton–watson random trees. *Communications in Mathematical Physics*, 403 (1):495–527, 2023.

F. Augeri. Large deviations of the empirical spectral measure of supercritical sparse Wigner matrices. *Advances in Mathematics*, 466:Paper No. 110156, 53, 2025.

L. Avena, R. S. Hazra, and N. Malhotra. Limiting spectra of inhomogeneous random graphs. *arXiv preprint arXiv:2312.02805*, 2023.

L. Avena, F. Capannoli, R. S. Hazra, and M. Quattropani. Meeting, coalescence and consensus time on random directed graphs. *The Annals of Applied Probability*, 34(5):4940–4997, 2024.

Z. Bai and J. W. Silverstein. *Spectral analysis of large dimensional random matrices*, volume 20. Springer, 2010.

Z. D. Bai. Methodologies in spectral analysis of large dimensional random matrices, a review. *Statistica Sinica*, 9(3):611–677, 1999.

M. Banna and T. Mai. Berry-Esseen bounds for the multivariate $\mathcal{B}$-free CLT and operator-valued matrices. *Transactions of the American Mathematical Society*, 376(6):3761–3818, 2023.

V. Bansaye and M. Salvi. Branching processes and homogenization for epidemics on spatial random graphs. *Electronic Journal of Probability*, 29:1–37, 2024.

A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999.

M. Bauer and O. Golinelli. Random incidence matrices: moments of the spectral density. *Journal of Statistical Physics*, 103(1-2):301–337, 2001.

F. Benaych-Georges, A. Guionnet, and C. Male. Central limit theorems for linear statistics of heavy tailed random matrices. *Communications in Mathematical Physics*, 329(2):641–686, 2014.

F. Benaych-Georges, C. Bordenave, and A. Knowles. Largest eigenvalues of sparse inhomogeneous Erdős–Rényi graphs. *The Annals of Probability*, 47(3): 1653–1676, 2019.

F. Benaych-Georges, C. Bordenave, and A. Knowles. Spectral radii of sparse random matrices. *Annales de l'Institut Henri Poincaré Probabilités et Statistiques*, 56(3):2141–2161, 2020.

I. Benjamini and O. Schramm. Recurrence of distributional limits of finite planar graphs. *Electronic Journal of Probability*, 6:no. 23, 13, 2001.

H. Bercovici and D. Voiculescu. Free convolution of measures with unbounded support. *Indiana University Mathematics Journal*, 42(3):733–773, 1993.

N. Berger. Transience, recurrence and critical behavior for long-range percolation. *Communications in Mathematical Physics*, 226(3):531–558, 2002.

G. Bet, K. Bogerd, and V. Jacquier. First-order asymptotics for the structure of the inhomogeneous random graph. *arXiv preprint arXiv:2306.06396*, 2023.

S. Bhamidi, R. Van Der Hofstad, and J. S. van Leeuwaarden. Scaling limits for critical inhomogeneous random graphs with finite third moments. *Electronic Journal of Probability*, 15:1682–1702, 2010.

S. Bhamidi, S. N. Evans, and A. Sen. Spectra of large random trees. *Journal of Theoretical Probability*, 25(3):613–654, 2012.

F. M. Bianchi, D. Grattarola, and C. Alippi. Spectral clustering with graph neural networks for graph pooling. In *International conference on machine learning*, pages 874–883. PMLR, 2020.

P. Biane. On the free convolution with a semi-circular distribution. *Indiana University Mathematics Journal*, 46(3):705–718, 1997.

P. Billingsley. *Probability and measure.* Wiley Series in Probability and Statistics. John Wiley & Sons, Inc., Hoboken, NJ, anniversary edition, 2012. ISBN 978-1-118-12237-2.

B. Bollobás, S. Janson, and O. Riordan. The phase transition in inhomogeneous random graphs. *Random Structures & Algorithms*, 31(1):3–122, 2007.

C. Bordenave. Lecture notes on random matrix theory. *IMPA*, 2019.

C. Bordenave and M. Lelarge. Resolvent of large random graphs. *Random Structures & Algorithms*, 37(3):332–352, 2010.

C. Bordenave, M. Lelarge, and J. Salez. The rank of diluted random graphs. *The Annals of Probability*, 39(3):1097–1121, 2011.

C. Bordenave, P. Caputo, and D. Chafaï. Spectrum of Markov generators on sparse random graphs. *Communications on Pure and Applied Mathematics*, 67 (4):621–669, 2014.

C. Bordenave, M. Lelarge, and L. Massoulié. Non-backtracking spectrum of random graphs: community detection and non-regular Ramanujan graphs. In *2015 IEEE 56th Annual Symposium on Foundations of Computer Science*, pages 1347–1357. IEEE, 2015.

C. Bordenave, A. Sen, and B. Virág. Mean quantum percolation. *Journal of the European Mathematical Society*, 19(12):3679–3707, 2017.

A. Bose, K. Saha, A. Sen, and P. Sen. Random matrices with independent entries: beyond non-crossing partitions. *Random Matrices. Theory and Applications*, 11(2):Paper No. 2250021, 42, 2022.

N. Broutin, T. Duquesne, and M. Wang. Limits of multiplicative inhomogeneous random graphs and lévy trees: Limit theorems. *Probability Theory and Related Fields*, 181:865–973, 2021.

W. Bryc, A. Dembo, and T. Jiang. Spectral measure of large random Hankel, Markov and Toeplitz matrices. *The Annals of Probabality*, 34(1):1–38, 2006.

R. Carmona and J. Lacroix. *Spectral theory of random Schrödinger operators.* Springer Science & Business Media, 2012.

A. Chakrabarty and R. S. Hazra. Remarks on absolute continuity in the context of free probability and random matrices. *Proceedings of the American Mathematical Society*, 144(3):1335–1341, 2016.

A. Chakrabarty, R. S. Hazra, and D. Sarkar. Limiting spectral distribution for Wigner matrices with dependent entries. *Acta Phys. Polon. B*, 46(9):1637–1652, 2015. ISSN 0587-4254.

A. Chakrabarty, R. S. Hazra, and D. Sarkar. From random matrices to long range dependence. *Random Matrices. Theory and Applications*, 5(2):1650008, 52, 2016.

A. Chakrabarty, S. Chakraborty, and R. S. Hazra. A note on the folklore of free independence. *Statistics and Applications*, 19(1):187–198, 2021a.

A. Chakrabarty, R. S. Hazra, F. den Hollander, and M. Sfragara. Spectra of adjacency and Laplacian matrices of inhomogeneous Erdős-Rényi random graphs. *Random Matrices. Theory and Applications*, 10(1):Paper No. 2150009, 34, 2021b.

A. Chakrabarty, R. S. Hazra, F. den Hollander, and M. Sfragara. Large deviation principle for the maximal eigenvalue of inhomogeneous Erdős-Rényi random graphs. *Journal of Theoretical Probability*, 35(4):2413–2441, 2022.

C. Champion, B. Mélanie, B. Rémy, L. Jean-Michel, and R. Laurent. Robust spectral clustering using Lasso regularization. *arXiv preprint arXiv:2004.03845*, 2020.

A. Chatterjee and R. S. Hazra. Spectral properties for the Laplacian of a generalized Wigner matrix. *Random Matrices. Theory and Applications*, 11 (3):Paper No. 2250026, 66, 2022.

S. Chatterjee. A simple invariance theorem. *arXiv preprint math/0508213*, 2005.

U. Chatterjee, K. Huang, R. Karmakar, B. R. V. Kumar, G. Lugosi, N. Malhotra, A. Mandal, and M. A. Tarafdar. Detecting weighted hidden cliques. *arXiv preprint arxiv:2506.21543*, 2025.

D. Cheliotis and M. Louvaris. The limit of the operator norm for random matrices with a variance profile. *arXiv preprint arXiv:2404.13795*, 2024.

F. Chung and L. Lu. The average distances in random graphs with given expected degrees. *Proceedings of the National Academy of Sciences*, 99(25): 15879–15882, 2002.

F. Chung, L. Lu, and V. Vu. Eigenvalues of random power law graphs. *Annals of Combinatorics*, 7(1):21–33, 2003.

F. R. Chung. *Spectral graph theory*, volume 92 of *CBMS Regional Conference Series in Mathematics*. American Mathematical Society, 1997. ISBN 0-8218-0315-8.

A. Cipriani and M. Salvi. Scale-free percolation mixing time. *Stochastic Processes and their Applications*, 167:104236, 2024.

A. Cipriani, R. S. Hazra, N. Malhotra, and M. Salvi. The spectrum of dense kernel-based random graphs. *arXiv preprint arXiv:2502.09415*, 2025.

S. Coste and J. Salez. Emergence of extended states at zero in the spectrum of sparse random graphs. *The Annals of Probability*, 49(4):2012–2030, 2021.

J. Dalmau and M. Salvi. Scale-free percolation in continuous space: quenched degree and clustering coefficient. *Journal of Applied Probability*, 58(1):106–127, 2021.

N. M. M. De Abreu. Old and new results on algebraic connectivity of graphs. *Linear algebra and its applications*, 423(1):53–73, 2007.

M. Deijfen, R. van der Hofstad, and G. Hooghiemstra. Scale-free percolation. *Annales de l'Institut Henri Poincaré Probabilités et Statistiques*, 49(3):817–838, 2013.

P. Deprez and M. V. Wüthrich. Scale-free percolation in continuum space. *Communications in Mathematics and Statistics*, 7(3):269–308, 2019.

P. Deprez, R. S. Hazra, and M. V. Wüthrich. Inhomogeneous long-range percolation for real-life network modeling. *Risks*, 3(1):1–23, 2015.

L. Devroye and N. Fraiman. Connectivity of inhomogeneous random graphs. *Random Structures & Algorithms*, 45(3):408–420, 2014.

X. Ding and T. Jiang. Spectral distributions of adjacency and Laplacian matrices of random graphs. *The Annals of Applied Probability*, 20(6):2086–2117, 2010.

P. Dionigi, D. Garlaschelli, R. S. Hazra, F. den Hollander, and M. Mandjes. Central limit theorem for the principal eigenvalue and eigenvector of Chung–Lu random graphs. *Journal of Physics; Complexity*, 4(1), 2023.

R. Ducatez, A. Guionnet, and J. Husson. Large deviation principle for the largest eigenvalue of random matrices with a variance profile. *arXiv preprint arXiv:2403.05413*, 2024.

I. Dumitriu and Y. Zhu. Sparse general Wigner-type matrices: local law and eigenvector delocalization. *Journal of Mathematical Physics*, 60(2):023301, 16, 2019.

F. J. Dyson. A brownian-motion model for the eigenvalues of a random matrix. *Journal of Mathematical Physics*, 3(6):1191–1198, 1962.

L. Erdős and H.-T. Yau. *A dynamical approach to random matrix theory*, volume 28 of *Courant Lecture Notes in Mathematics*. Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, RI, 2017. ISBN 978-1-4704-3648-3.

P. Erdős and A. Rényi. On random graphs. *Publicationes mathematicae*, 6 (26):290–297, 1959.

L. Erdős, A. Knowles, H.-T. Yau, and J. Yin. Spectral statistics of erdős-rényi graphs i: Local semicircle law. *The Annals of Probability*, 41(3B):2279–2375, 2013.

I. Gallagher, A. Jones, A. Bertiger, C. E. Priebe, and P. Rubin-Delanchy. Spectral embedding of weighted graphs. *Journal of the American Statistical Association*, 119(547):1923–1932, 2024.

L. Geisinger. Convergence of the density of states and delocalization of eigenvectors on random regular graphs. *Journal of Spectral Theory*, 5(4):783–827, 2015.

E. N. Gilbert. Random graphs. *Annals of Mathematical Statistics*, 30(4): 1141–1144, 1959.

P. Gracar and A. Grauer. The contact process on scale-free geometric random graphs. *Stochastic Processes and their Applications*, 173:104360, 2024.

P. Gracar, L. Lüchtrath, and P. Mörters. Percolation phase transition in weight-dependent random connection models. *Advances in Applied Probability*, 53(4):1090–1114, 2021.

W. H. Haemers. Hoffman's ratio bound. *arXiv preprint arXiv:2102.05529*, 2021.

N. Hao and M. Heydenreich. Graph distances in scale-free percolation: the logarithmic case. *Journal of Applied Probability*, 60(1):295–313, 2023.

R. S. Hazra and N. Malhotra. Spectral properties of the Laplacian of scale-free percolation models. *arXiv preprint arXiv:2504.17552*, 2025.

R. S. Hazra and K. Maulik. Free subexponentiality. *The Annals of Probability*, 41(2):961–988, 2013.

M. Heydenreich, T. Hulshof, and J. Jorritsma. Structures in supercritical scale-free percolation. *The Annals of Applied Probability*, pages 2569–2604, 2017.

J. Huang and B. Landon. Spectral statistics of sparse Erdős–Eényi graph Laplacians. *Annales de l'Institut Henri Poincaré Probabilités et Statistiques*, 56(1):120–154, 2020.

H. J. Hupkes, M. Jukić, P. Stehlík, and V. Švígler. Propagation reversal for bistable differential equations on trees. *SIAM Journal on Applied Dynamical Systems*, 22(3):1906–1944, 2023.

J. Husson. Large deviations for the largest eigenvalue of matrices with variance profiles. *Electronic Journal of Probability*, 27:Paper No. 74, 44, 2022.

S. Janson, D. E. Knuth, T. Łuczak, and B. Pittel. The birth of the giant component. *Random Structures & Algorithms*, 4(3):233–358, 1993.

T. Jiang. Empirical distributions of Laplacian matrices of large dilute random graphs. *Random Matrices. Theory and Applications*, 1(03):1250004, 2012.

I. M. Johnstone. On the distribution of the largest eigenvalue in principal components analysis. *Annals of Statistics*, 29(2):295–327, 2001.

J. Jorritsma, J. Komjáthy, and D. Mitsche. Cluster-size decay in supercritical kernel-based spatial random graphs. *arXiv preprint arXiv:2303.00724*, 2023.

J. Jorritsma, J. Komjáthy, and D. Mitsche. Large deviations of the giant in supercritical kernel-based spatial random graphs. *arXiv preprint arXiv:2404.02984*, 2024.

P. Jung and J. Lee. Delocalization and limiting spectral distribution of Erdős-Rényi graphs with constant expected degree. *Electronic Communications in Probability*, 23:Paper No. 92, 13, 2018.

O. Khorunzhy, M. Shcherbina, and V. Vengerovsky. Eigenvalue distribution of large weighted random graphs. *Journal of Mathematical Physics*, 45(4): 1648–1672, 2004.

B. Kołodziejek and K. Szpojankowski. A phase transition for tails of the free multiplicative convolution powers. *Advances in Mathematics*, 403:Paper No. 108398, 50, 2022.

J. Komjáthy and B. Lodewijks. Explosion in weighted hyperbolic random graphs and geometric inhomogeneous random graphs. *Stochastic Processes and their Applications*, 130(3):1309–1367, 2020.

J. Komjáthy, J. Lapinskas, J. Lengler, and U. Schaller. Four universal growth regimes in degree-dependent first passage percolation on spatial random graphs I. *arXiv preprint arXiv:2309.11840*, 2023.

R. J. LeVeque. *Finite Difference Methods for Ordinary and Partial Differential Equations: Steady-State and Time-Dependent Problems*. Society for Industrial and Applied Mathematics, 2007. ISBN 978-0-89871-629-0.

G. Lin. Recent developments on the moment problem. *Journal of Statistical Distributions and Applications*, 4(1):1–17, 2017.

L. Lovász and B. Szegedy. Limits of dense graph sequences. *Journal of Combinatorial Theory, Series B*, 96(6):933–957, 2006.

C. Male. The limiting distributions of large heavy wigner and arbitrary random matrices. *Journal of Functional Analysis*, 272(1):1–46, 2017.

J. A. Mingo and R. Speicher. *Free probability and random matrices*, volume 35 of *Fields Institute Monographs.* Springer, New York; Fields Institute for Research in Mathematical Sciences, Toronto, ON, 2017. ISBN 978-1-4939-6941-8; 978-1-4939-6942-5.

H. L. Montgomery. The pair correlation of zeros of the zeta function. *Analytic number theory, Proceedings of Symposia in Pure Mathematics*, 24:181–193, 1973.

C. M. Newman and L. S. Schulman. One dimensional $1/|j-i|^\alpha$ percolation models: The existence of a transition for $s \leq 2$. *Communications in Mathematical Physics*, 104(4):547–571, 1986.

A. Nica and R. Speicher. *Lectures on the combinatorics of free probability*, volume 13 of *London Mathematical Society Lecture Note Series.* Cambridge University Press, Cambridge, 2006. ISBN 978-0-521-85852-6; 0-521-85852-6.

R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani. Epidemic processes in complex networks. *Reviews of Modern Physics*, 87(3):925–979, 2015.

M. Pernici. Noncrossing partition flow and random matrix models. *arXiv preprint arXiv:2106.02655*, 2021.

S. I. Resnick. *Extreme values, regular variation, and point processes*, volume 4. Springer Science & Business Media, 2008.

W. Rudin. *Principles of mathematical analysis.* International Series in Pure and Applied Mathematics. McGraw-Hill Book Co., New York-Auckland-Düsseldorf, third edition, 1976.

W. Rudin. *Functional Analysis.* International Series in Pure and Applied Mathematics. McGraw-Hill, Inc., New York, second edition, 1991. ISBN 0-07-054236-8.

O. Ryan. On the limit distributions of random matrices with independent or free entries. *Communications in Mathematical Physics*, 193:595–626, 1998.

J. Salez. *Some implications of local weak convergence for sparse random graphs.* Thesis, Université Pierre et Marie Curie—Paris VI. Ecole Normale Supérieure de Paris—ENS Paris, 2011.

J. Salez. Spectral atoms of unimodular random trees. *Journal of the European Mathematical Society*, 22(2):345–363, 2020.

L. S. Schulman. Long range percolation in one dimension. *Journal of Physics A: Mathematical and General*, 16(17):L639, 1983.

A. Sen and B. Virág. Absolute continuity of the limiting eigenvalue distribution of the random Toeplitz matrix. *Electronic Communications in Probability*, 16: 706–711, 2011.

R. Speicher. Free probability theory. In G. Akemann, J. Baik, and P. D. Francesco, editors, *The Oxford Handbook of Random Matrix Theory*, chapter 23, pages 465–491. Oxford University Press, 2011.

R. Speicher. *Lecture Notes on "Random Matrices"*. EMS Series of Lectures in Mathematics. EMS Press, Berlin, 2024. ISBN 978-3-98547-075-4; 978-3-98547-575-9.

D. Spielman. Spectral graph theory. *Combinatorial scientific computing*, 18 (18), 2012.

T. Tao. *Topics in Random Matrix Theory*, volume 132 of *Graduate studies in mathematics*. American Mathematical Society, 2012.

C. A. Tracy and H. Widom. Level-spacing distributions and the Airy kernel. *Communications in Mathematical Physics*, 159(1):151–174, 1994.

L. V. Tran, V. H. Vu, and K. Wang. Sparse random graphs: Eigenvalues and eigenvectors. *Random Structures & Algorithms*, 42(1):110–134, 2013.

J. A. Tropp. *An Introduction to Matrix Concentration Inequalities*. Foundations and Trends in Machine Learning, 2015.

R. van der Hofstad. Critical behavior in inhomogeneous random graphs. *Random Structures & Algorithms*, 42(4):480–508, 2013.

R. van der Hofstad. *Random graphs and complex networks. Vol. I*, volume [43] of *Cambridge Series in Statistical and Probabilistic Mathematics, [43]*. Cambridge University Press, Cambridge, 2017. ISBN 978-1-107-17287-6.

R. van der Hofstad. *Random graphs and complex networks- Vol. II*, volume 2. Cambridge Series in Statistical and Probabilistic Mathematics, 2024. ISBN 9781107174009.

R. van der Hofstad and J. Komjáthy. Explosion and distances in scale-free percolation. *arXiv preprint arXiv:1706.02597*, 2017.

R. van der Hofstad, P. Van Der Hoorn, and N. Maitra. Local limits of spatial inhomogeneous random graphs. *Advances in Applied Probability*, 55(3):793–840, 2023.

R. van der Hofstad, P. van der Hoorn, C. Kerriou, N. Maitra, and P. Mörters. Condensation in scale-free geometric graphs with excess edges. *arXiv preprint arXiv:2405.20425*, 2024.

D. Voiculescu. Limit laws for Random Matrices and Free Products. *Inventiones Mathematicae*, 104(1):201–220, 1991.

D.-V. Voiculescu. Symmetries of some reduced free product C*-algebras. In *Operator Algebras and Their Connections with Topology and Ergodic Theory*, volume 1132 of *Lecture Notes in Mathematics*, pages 556–588. Springer, 1985.

D. J. Watts and S. H. Strogatz. Collective dynamics of small-world networks. *Nature*, 393(6684):440, 1998.

E. P. Wigner. Characteristic vectors of bordered matrices with infinite dimensions. *Annals of Mathematics*, 62(3):548–564, 1955.

E. P. Wigner. On the distribution of the roots of certain symmetric matrices. *Annals of Mathematics*, 67(2):325–327, 1958.

J. Wishart. The generalised product moment distribution in samples from a normal multivariate population. *Biometrika*, 20A(1/2):32–52, 1928.

X. Zhu and Y. Zhu. Central limit theorems for linear spectral statistics of inhomogeneous random graphs with graphon limits. *arXiv preprint arXiv:2412.19352*, 2024.

Y. Zhu. A graphon approach to limiting spectral distributions of Wigner-type matrices. *Random Structures & Algorithms*, 56(1):251–279, 2020.

List of publications

- L. Avena, R. S. Hazra, **N. Malhotra**.
  Limiting spectra of inhomogeneous random graphs.
  arXiv preprint arXiv:2312.02805, **2023**
- A. Cipriani, R. S. Hazra, **N. Malhotra**, M. Salvi.
  The spectrum of dense kernel-based random graphs.
  arXiv preprint arXiv:2502.09415, **2025**.
- R. S. Hazra, **N. Malhotra**.
  Spectral properties of the Laplacian of scale-free percolation models.
  arXiv preprint arXiv:2504.17552, **2025**.
- U. Chatterjee, K. Huang, R. Karmakar, B. R. V. Kumar, G. Lugosi, **N. Malhotra**, A. Mandal, M. A. Tarafdar.
  Detecting weighted hidden cliques.
  arXiv preprint arxiv:2506.21543, **2025**.

Summary

This thesis comprises five chapters, three of which contain the core mathematical content. We study the limiting spectral distributions of random graph models with vertex inhomogeneity. In particular, we focus on the adjacency matrix of the inhomogeneous Erdős–Rényi random graph in the sparse regime, as well as the adjacency and Laplacian matrices for random graph models that incorporate spatial structure.

Random graph models provide a mathematical framework for understanding complex networks observed in fields such as physics, biology, computer science, and the social sciences. The classical Erdős–Rényi model, in which edges are added independently with equal probability, serves as a foundational model that continues to yield deep insights. Spectral graph theory plays a key role in this context, connecting the eigenvalues and eigenvectors of the adjacency and Laplacian matrices of graphs to structural and geometric properties of graphs. For instance, the Perron–Frobenius theorem ensures a unique largest eigenvalue for the adjacency matrix of a connected graph, with a corresponding positive eigenvector. More broadly, the spectrum gives information about the graph connectivity, subgraph counts, the chromatic number, and other topological features. Laplacian eigenvalues are central in the study of diffusion, mixing times of random walks, and spectral clustering algorithms. Notably, the Kirchhoff Matrix–Tree Theorem relates the determinant of the combinatorial Laplacian to the count of the spanning trees of the graph. These connections make spectral analysis a powerful tool for studying the geometry of complex networks. Chapter 1 provides a detailed introduction to spectral graph theory, random graphs, and random matrices.

The spectra of the adjacency and Laplacian matrices are well understood in the dense Erdős-Rényi random graph model. In the sparse case, three main analytical techniques are used: (i) characterising the limiting spectrum via local weak limits such as Galton–Watson trees; (ii) using combinatorial methods and special symmetric partitions to compute the moments of the limiting spectral measure; (iii) deriving the Stieltjes transform of the limiting measure using a fixed-point equation in an appropriate Banach space. In Chapter 2, we extend the Erdős–Rényi random graph model by incorporating deterministic vertex weights to introduce inhomogeneity, where now edges are added independently with a probability proportional to a function of the vertex weights. We study this model in the sparse setting, where the connectivity function is bounded. We analyse the empirical spectral distribution of the adjacency matrix using the moment method and the Stieltjes transform, and describe the limiting distribution through homomorphism densities, symmetric partitions, and a fixed-point equation.

Real-world networks often exhibit spatial structure in addition to vertex inhomogeneity. In Chapter 3, we consider a kernel-based random graph model on a discrete torus, where the vertices are equipped with random weights that follow a power-law distribution, and connection probabilities between two vertices depend directly on a function of the two weights and that is inversely proportional to the torus distance between the two vertices. Using the method of moments, we study the adjacency matrix of this model and show the existence and uniqueness of a limiting spectral measure. We further analyse the measure through its prelimit to show that it is absolutely continuous and non-degenerate. We characterise the Stieltjes transform of this measure through a fixed-point equation. When the kernel is rank-one, that is, it has a product structure, we identify the limiting measure explicitly as a free multiplicative convolution between the semicircle law and the Pareto law using tools from free probability.

In Chapter 4, we focus on the centred Laplacian matrix of the rank-one model, which is known as the Scale-Free percolation model. Using the method of moments, we show the existence of a unique limiting spectral measure. We further identify the measure in terms of the spectral distribution of some non-commutative unbounded operators, again using techniques from free probability theory.

In Chapter 5, we present simulations and provide a brief discussion of examples that fall outside the restrictions assumed in the previous chapters.

Samenvatting

Dit proefschrift bestaat uit vijf hoofdstukken, waarvan er drie de kern van de wiskundige inhoud bevatten. We bestuderen limieten van spectrale verdelingen van toevallige graafmodellen met puntinhomogeniteiten. In het bijzonder richten we ons op de nabuur-matrix van de inhomogene Erdős–Rényi graaf in het ijle regime, en op de nabuur- en Laplace-matrices voor modellen die een ruimtelijke structuur bevatten.

Toevallige graafmodellen bieden een wiskundig kader voor het begrijpen van complexe netwerken in vakgebieden zoals natuurkunde, biologie, informatica en sociale wetenschappen. Het klassieke Erdős–Rényi-model, waarin lijnen onafhankelijk met gelijke waarschijnlijkheid worden toegevoegd, dient als een fundamenteel model dat diepgaande inzichten blijft opleveren. Spectrale grafentheorie speelt een sleutelrol in deze context, door de eigenwaarden en eigenvectoren van de nabuur- en Laplace-matrices van graafmodellen te verbinden met structurele en geometrische eigenschappen van graafmodellen. De stelling van Perron-Frobenius garandeert bijvoorbeeld een unieke grootste eigenwaarde voor de nabuur-matrix van een verbonden graaf, met een bijbehorende positieve eigenvector. Breder gezien geeft het spectrum informatie over de connectiviteit van de graaf, het aantal subgrafen van een bepaald type, het chromatische getal en andere topologische kenmerken. Laplace-eigenwaarden spelen een centrale rol in de studie van diffusie, mengtijden van toevallige wandelingen en spectrale clusteringalgoritmen. Met name de matrix-boomstelling van Kirchhoff relateert de determinant van de combinatorische Laplace-matrix aan het aantal opspannende bomen van de graaf. Deze verbindingen maken spectrale analyse een krachtig hulpmiddel voor het bestuderen van de geometrie van complexe netwerken. Hoofdstuk 1 biedt een gedetailleerde inleiding tot spectrale grafentheorie, toevallige grafen en toevallige matrices.

De spectra van de nabuur- en Laplace-matrices worden goed begrepen in het Erdős-Rényi-model in het dichte regime. In het ijle regime worden drie belangrijke analytische technieken gebruikt: (i) karakterisering van het limietspectrum via lokale zwakke limieten zoals Galton-Watson-bomen; (ii) het gebruik van combinatorische methoden en speciale symmetrische partities om de momenten van de limietspectraalmaat te berekenen; (iii) het afleiden van de Stieltjestransformatie van de limietmaat met behulp van een vaste-puntvergelijking in een

geschikte Banach-ruimte. In hoofdstuk 2 breiden we het Erdős-Rényi-model uit door deterministische puntgewichten te integreren om inhomogeniteit te introduceren, waarbij lijnen nu onafhankelijk worden toegevoegd met een waarschijnlijkheid evenredig met een functie van de puntgewichten. We bestuderen dit model in het ijle regime, waar de connectiviteitsfunctie begrensd is. We analyseren de empirische spectrale verdeling van de nabuur-matrix met behulp van de momentenmethode en de Stieltjestransformatie, en beschrijven de limietverdeling met behulp van homomorfismedichtheden, symmetrische partities en een vaste-puntvergelijking.

Netwerken vertonen vaak een ruimtelijke structuur naast puntinhomogeniteit. In hoofdstuk 3 beschouwen we een kernelgebaseerd toevallig graafmodel op een discrete torus, waarbij de punten zijn voorzien van toevallige gewichten die een machtswetverdeling volgen, en de verbindingskansen tussen twee punten afhangen van een functie van de twee gewichten die omgekeerd evenredig is met de torusafstand tussen de twee punten. Met behulp van de momentenmethode bestuderen we de nabuur-matrix van dit model en tonen we het bestaan en de uniciteit van een spectrale limietmaat aan. We analyseren de maat via zijn prelimiet om aan te tonen dat deze absoluut-continu en niet-ontaard is. We karakteriseren de Stieltjestransformatie van deze maat via een vaste-puntvergelijking. Wanneer de kernel rang-1 is, dat wil zeggen, een productstructuur heeft, identificeren we de limietmaat expliciet als een vrije multiplicatieve convolutie tussen de halvecirkelwet en de Paretowet, met behulp van hulpmiddelen uit de vrije kansrekening.

In hoofdstuk 4 richten we ons op de gecentreerde Laplace-matrix van het rang-1-model, ook wel bekend als het schaalvrije percolatiemodel. Met behulp van de momentenmethode tonen we het bestaan van een unieke spectrale limietmaat aan. We identificeren de maat verder aan de hand van de spectrale verdeling van enkele niet-commutatieve onbegrensde operatoren, wederom met behulp van technieken uit de vrije kansrekening.

In hoofdstuk 5 presenteren we simulaties en geven we een korte bespreking van voorbeelden die buiten de veronderstelde beperkingen van de voorgaande hoofdstukken vallen.

Acknowledgements

This thesis is a conclusion of a four-year journey, which comprises numerous people and experiences that shaped not just the content of this thesis itself, but also of my life.

I begin by thanking my supervisors; Rajat, Luca, and Frank. I am deeply grateful to Rajat for our blackboard-filled discussions, for borrowing a significant portion of my stress, and for effortlessly switching between the roles of a supervisor and a collaborator as required. This thesis was a result of the invaluable support, encouragement, and caffeine that he provided, creating a sufficiently informal space for me to present my ideas. I sincerely appreciate Luca for the several recaps that ensured that the problem was ingrained in me. While Florence was more than just a bike ride away, voice memos bridged the gap in a way I could not have imagined. Last, but not least, I express my gratitude to my promoter, Frank, for his enthusiasm in closely following my projects and for his guidance and mentorship wherever I needed it. I learned a lot about the mathematics community and about being a mathematician from him.

I am grateful to my reading committee for reviewing my thesis and providing feedback. I thank my collaborators, Alessandra and Michele, for our discussions and my visits to London, which went smoothly despite a few technical glitches and pink comments.

I thank the NETWORKS program for funding my research, and the group members who made the training weeks fun. I also thank the Leiden Mathematical Institute for supporting me in this journey, and the POD group of the institute for the nice atmosphere they created.

To my friends and colleagues at the institute, I am grateful for the vibrant workplace they created through memorable lunches and colloquia, untimed coffee and tea breaks, and post-lunch NYT games. I am immensely thankful to my two paranymphs, Federico and Jolien, and to all the wonderful friends I made in Leiden. The photography walks, weekly swims, boxing after intense warm-ups, bike rides despite my complaints about the wind, and the experiments of building up their spice tolerances are all unforgettable memories that made Leiden a home away from home.

I would like to thank Neeraja, who developed my interest in the topic of probability and random graphs during my master's and who introduced me to the probability community in India, where I first met Rajat.

As certain long-range network models suggest, some bonds can be exceptionally strong even when they are halfway across the world. I express my deepest gratitude to my parents, Ajeeta, the Viswamitras, and Raghav for their unwavering love, encouragement, guidance, and belief in me. Our annual end-of-year time together was a welcome pause and provided a much-needed breather. I cannot thank Kausthub, Abhijit, and Nikhil enough, for being a constant and reliable support system from the very start of my mathematical journey in India all the way through to my PhD. Their friendship ensured that even the most challenging days were a breeze.

Lastly, for every step of the journey until this moment, and for celebrating every high and supporting me through every low, I am indebted to Aishwarya. My journey to becoming a better mathematician is strongly correlated with my journey to becoming a better person, and her unconditional love and patience have been a foundation for both.

Curriculum Vitae

Nandan Malhotra was born on September 14th 1998, in New Delhi, India. He grew up in New Delhi, where he attended the Mother's International School and obtained his high school graduation certificate in 2016.

He then moved to Mohali (Punjab) to join a dual-degree BS-MS programme at the Indian Institute of Science Education and Research (IISER), Mohali. He was awarded the DST-INSPIRE Fellowship by the Government of India. He obtained a Bachelor of Science and Master of Science in Mathematics in 2021. During his time at IISER Mohali, he undertook summer internships at the Indian Institute of Science (IISc) Bangalore and the École Normale Supérieure (ENS) Rennes. In his final year, he wrote his master's thesis entitled "Preferential Attachment Trees with fitness", under the supervision of dr. N. Sahasrabudhe (IISER Mohali) and dr. K. Avrachenkov (INRIA Sophia Antipolis).

In September 2021, he moved to Leiden to start his PhD as a member of the probability group at the Mathematical Institute of Leiden University, within the framework of the NETWORKS consortium. His research was funded in part by an NWO grant and further under a Marie Skłodowska-Curie grant. He worked under the supervision of prof. dr. W.T.F. den Hollander (Leiden University), dr. R.S. Hazra (Leiden University) and dr. L. Avena (University of Florence).

During his PhD, he served as a teaching assistant for various courses in Leiden. He attended workshops, conferences and summer schools, and presented his research on several occasions in Brazil, Germany, India, Italy, and The Netherlands. He collaborated on research projects and open problems with researchers in India, Italy, The Netherlands, the UK, and the USA. He was an active member of the mathematical community in these four years as an organiser of the PhD Colloquium and the probability seminar at the Mathematical Institute of Leiden University. He also served as a board member of the Young KWG, a section founded under the Royal Dutch Mathematical Society (KWG) to foster a community of young mathematicians in the early stages of their careers.