



Universiteit
Leiden
The Netherlands

Spectral analysis of inhomogeneous network models

Malhotra, N.

Citation

Malhotra, N. (2025, November 20). *Spectral analysis of inhomogeneous network models*. Retrieved from <https://hdl.handle.net/1887/4283482>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/4283482>

Note: To cite this publication please use the final published version (if applicable).

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](https://arxiv.org/abs/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 $N\varepsilon_N \rightarrow \lambda \in (0, \infty)$. 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 $\lambda \rightarrow \infty$, 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 $\lambda \rightarrow \infty$, 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 $[N] = 1, \dots, N$, 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

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

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 $N\varepsilon_N \rightarrow \infty$, the average degree is unbounded, and when $N\varepsilon_N = O(1)$, 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 $N\varepsilon_N \rightarrow \lambda$. 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 $\lambda \rightarrow \infty$, 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 $N\varepsilon_N \rightarrow \infty$. 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 Bordeave 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 $\lambda \rightarrow \infty$) using this result. The second result, Theorem 2.3.9, provides an analytic character-

isation 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 $[N] = \{1, \dots, N\}$ 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 $\varepsilon_N \rightarrow 0$, $(w_i)_{i \geq 1}$ is a sequence of deterministic non-negative weights and $f : [0, \infty)^2 \rightarrow [0, \infty)$ 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

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

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

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

We are interested in studying the following object:

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

where $\lambda_1, \dots, \lambda_N$ 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:** $\varepsilon_N \rightarrow 0$ and $N\varepsilon_N \rightarrow \infty$. The connectivity regime with $N\varepsilon_N \gg C \log N$ falls in this regime.
- **Sparse Regime :** $\varepsilon_N \rightarrow 0$ and $N\varepsilon_N \rightarrow \lambda \in (0, \infty)$.

Dense regime

In literature, the dense regime is characterised by $\varepsilon_N \equiv \text{constant}$ 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 $\varepsilon_N \rightarrow 0$ and $N\varepsilon_N \rightarrow \infty$. 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

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

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

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

where H_k is a simple graph on k vertices with the edge set $E(H_k)$, $\mu_w^{\otimes k}(\cdot)$ is the k -fold product measure of $\mu_w(\cdot)$, and $\mathbf{w} = (w_1, \dots, w_k)$. 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

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

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:*

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

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 a unique analytic function $\mathcal{H}$ defined on $\mathbb{C}^+ \times [0, \infty)$ such that*

$$S_{\mu_f}(z) = \int_0^\infty \mathcal{H}(z, x) \mu_w(dx),$$

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

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

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 $r : [0, \infty) \rightarrow [0, \infty)$ 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 $\{\mathbb{G}_N\}_{N \geq 1}$ have a weak limit $\mathbb{G}$;
- For a uniformly chosen root $o_N \in \mathbb{G}_N$, the degree sequence of the rooted graph $(\deg(\mathbb{G}_N, o_N))_{N \geq 1}$ 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 $\mathbb{G}$;

then, there exists a unique probability measure μ_λ on $\mathbb{R}$ such that $\text{ESD}(\mathbf{A}_N) \implies \mu_\lambda$ weakly in probability as $N \rightarrow \infty$. 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 $\text{Poi}(\lambda)$ 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 $f : \mathbb{C}^+ \rightarrow \mathbb{C}^+$, 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

$$R_{\mathbf{A}_N}(z) = (\mathbf{A}_N - zI)^{-1}$$

for each $z \in \mathbb{C}^+$. The map $z \mapsto R_{\mathbf{A}_N}(z)(i, i)$ is in $\mathbb{H}$, and the Stieltjes transform of $\text{ESD}(\mathbf{A}_N)$ is given by $\text{tr } R_{\mathbf{A}_N}(z)$, where $\text{tr} = N^{-1} \text{Tr}$ 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 $S_{\mathbf{A}_N}(z)$ denote the Stieltjes transform of the empirical measure $\text{ESD}(\mathbf{A}_N)$. 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

$$\lim_{N \rightarrow \infty} S_{\mathbf{A}_N}(z) = \mathbb{E}X(z) \text{ in } L^1,$$

where $X(z)$ is such that:

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

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 $\text{Poi}(\lambda)$. 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 $f : [0, \infty)^2 \rightarrow [0, \infty)$ be a bounded, continuous function, with $|f| \leq C_f \in (0, \infty)$,

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

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

$$\mathbf{A}_N = \frac{1}{\sqrt{\lambda}} \mathbf{M}_N. \quad (2.8)$$

§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 $\text{ESD}(\mathbf{A}_N)$ (see (2.3)). In what follows, we will see that

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

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 $V_1, V_2, \dots, V_m$, for some $1 \leq m \leq k$. 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 $[k] = \{1, 2, \dots, k\}$ is said to be a *Special Symmetric partition* if it satisfies the following:

- All blocks of π are of even size.

2. Limiting spectra of inhomogeneous random graphs

- 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 $V_1, V_2, \dots, V_\ell$, such that there exist elements $\{a_{i_1}, a_{i_1+1}, \dots, a_{i_1+k_1}\} \in V_1$, $\{a_{i_2}, \dots, a_{i_2+k_2}\} \in V_2$, $\dots$, $\{a_{i_\ell}, \dots, a_{i_\ell+k_\ell}\} \in V_\ell$, with $a = a_{i_1} - 1$ and $b = a_{i_\ell+k_\ell} + 1$, such that $k_1, k_2, \dots, k_\ell$ 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 $[k] \rightarrow [k]$. 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.

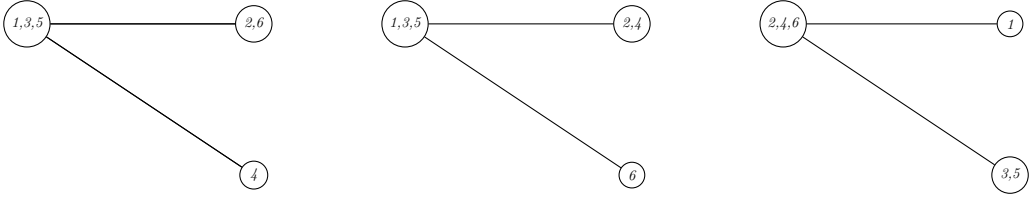
- *Initially consider the vertex set $V_{\gamma\pi} = [k]$ and perform a closed walk on $[k]$ as $1 \rightarrow 2 \rightarrow 3 \rightarrow \dots \rightarrow k \rightarrow 1$ and with each step of the walk, add an edge.*
- *Evaluate $\gamma\pi$, which will be of the form $\gamma\pi = \{V_1, V_2, \dots, V_m\}$ 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) $\pi_1 = \{\{1, 2, 5, 6\}, \{3, 4\}\},$ | (a) $\gamma\pi_1 = \{\{1, 3, 5\}, \{2, 6\}, \{4\}\},$ |
| (b) $\pi_2 = \{\{1, 2, 3, 4\}, \{5, 6\}\},$ | (b) $\gamma\pi_2 = \{\{1, 3, 5\}, \{2, 4\}, \{6\}\},$ |
| (c) $\pi_3 = \{\{1, 6\}, \{2, 3, 4, 5\}\}.$ | (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:



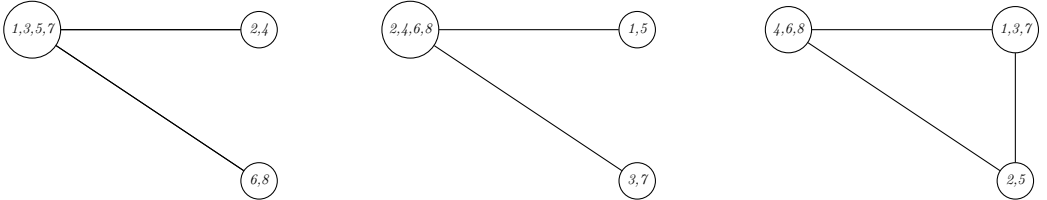
One can see that structurally the three graphs are the same. However, if we root them on V_1 , then the first two graphs are different from the third. Further, if we label the vertices as shown, all three graphs become distinct.

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

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

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

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

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 $\lambda \rightarrow \infty$,

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

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

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

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

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

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

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

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

$$S_{\mathbf{A}_N}(z) = \int_{\mathbb{R}} \frac{1}{x - z} \text{ESD}(\mathbf{A}_N)(dx) = \text{tr}(\mathbf{R}_{\mathbf{A}_N}(z)),$$

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

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

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

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

$$\mathcal{B} = \left\{ \phi : [0, \infty) \times [0, \infty) \rightarrow \mathbb{C} \text{ analytic} \left| \sup_{x, y \geq 0} \frac{|\phi(x, y)|}{\sqrt{1+y}} < \infty \right. \right\} \quad (2.11)$$

and take 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. 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

$$G_N(u, z) := \frac{1}{N} \sum_{i=1}^N e^{iur_i^N(z)} \quad (2.12)$$

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

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

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

$$d_f(y) = \int_0^\infty f(x, y) \mu_w(dx). \quad (2.13)$$

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

$$\begin{aligned} & \lim_{N \rightarrow \infty} \mathbb{E}[G_N(u, z)] \\ &= 1 - \sqrt{u} \int_0^\infty e^{-\lambda d_f(y)} \int_0^\infty \frac{J_1(2\sqrt{uv})}{\sqrt{v}} e^{ivz} e^{\lambda \phi^*(y, v/\lambda)} dv \mu_w(dy) \end{aligned} \quad (2.14)$$

and

$$\text{Var}[G_N(u, z)] \rightarrow 0.$$

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

$$\begin{aligned} & \phi^*(x, u) \\ &= F_z(\phi^*)(x, u) \\ &= d_f(x) - \int_0^\infty f(x, y) e^{-\lambda d_f(y)} \left(\sqrt{u} \int_0^\infty \frac{J_1(2\sqrt{uv})}{\sqrt{v}} e^{ivz} e^{\lambda \phi^*(y, v/\lambda)} dv \right) \mu_w(dy), \end{aligned} \quad (2.15)$$

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

$$J_1(x) = \frac{x}{2} \sum_{k=0}^{\infty} \frac{(-1)^k (x^2/4)^k}{k!(k+1)!}. \quad (2.16)$$

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 $z \in \mathbb{C}^+$,

$$S_{\mathbf{A}_N}(z) \rightarrow S_{\mu_\lambda}(z) \text{ in probability,}$$

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

$$S_{\mu_\lambda}(z) = i \int_0^\infty e^{-\lambda d_f(y)} \int_0^\infty e^{ivz} e^{\lambda \phi_z^*(y, v/\lambda)} dv \mu_w(dy), \quad z \in \mathbb{C}^+. \quad (2.17)$$

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

Corollary 2.3.11 (Stieltjes Transform as $\lambda \rightarrow \infty$).

For $\lambda \rightarrow \infty$, we have that

$$\lim_{\lambda \rightarrow \infty} S_{\mu_\lambda}(z) = S_{\mu_f}(z) \quad (2.18)$$

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

$$S_{\mu_f}(z) := \int_0^\infty \mathcal{H}(z, x) \mu_w(dx), \quad (2.19)$$

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

$$\phi^*(x, u) = 1 - \sqrt{u} \int_0^\infty e^{-\lambda} \left(\sqrt{u} \int_0^\infty \frac{J_1(2\sqrt{uv})}{\sqrt{v}} e^{ivz} e^{\lambda \phi^*(y, \frac{v}{\lambda})} dv \right) \mu_w(dy).$$

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

$$\widetilde{\phi}^*(u) = 1 - e^{-\lambda} \sqrt{u} \int_0^\infty \frac{J_1(2\sqrt{uv})}{\sqrt{v}} e^{ivz} e^{\lambda \widetilde{\phi}^*(v/\lambda)} dv. \quad (2.20)$$

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 $\widetilde{\phi}_z^*(u) = \mathbb{E}[e^{iuX(z)}]$ 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

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

where

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

§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

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

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

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

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 $m_\infty = o(\sqrt{m_1})$ and $N\varepsilon_N \rightarrow \lambda$. The connection probabilities will be given by

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

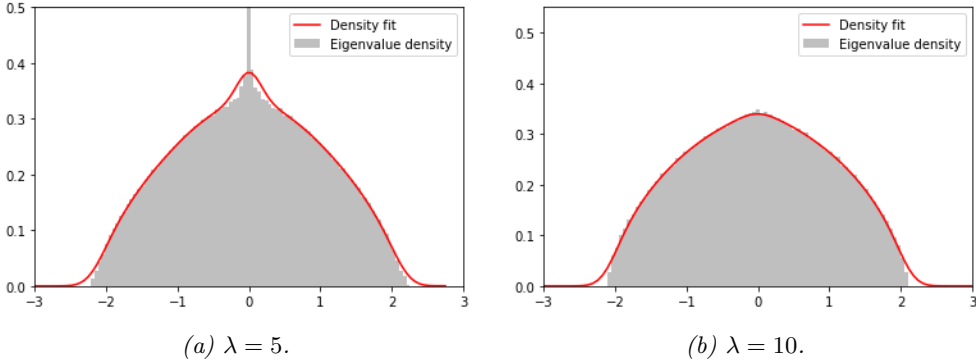


Figure 2.1: The homogeneous Erdős-Rényi Random Graph on 10,000 vertices.

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 be identified in the following way.

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

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

where $b_1(\sigma), \dots, b_{\#\sigma}(\sigma)$ 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 $\{\bar{1}, \dots, \bar{k}\}$, such that $\sigma \cup \bar{\sigma}$ is a non-crossing partition of $\{\bar{1}, 1, \dots, \bar{k}, k\}$. For example,

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

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

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

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

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

This also shows that

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

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 $w_i = d_i/\sqrt{m_1}$. Then,

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

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 ,

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

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 exists a coupling between $X \sim \text{Ber}(p)$ and $Y \sim \text{Ber}(q)$ 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 $\mathbf{M}_N^{\text{cl}}$ and $\mathbf{M}_N^{\text{grg}}$ 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 $N\varepsilon_N \rightarrow \lambda$ and $\mathbf{A}_N^{\text{cl}} = \lambda^{-1/2} \mathbf{M}_N^{\text{cl}}$ and $\mathbf{A}_N^{\text{grg}} = \lambda^{-1/2} \mathbf{M}_N^{\text{grg}}$. Then,

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

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

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

Therefore

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

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 $\text{ESD}(\mathbf{A}_N^{\text{cl}})$ converges weakly in probability to $\mu_{CL,\lambda}$ it follows that

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

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,

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

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 $\mathbf{A}_N^{\text{nr}} = \lambda^{-1/2} \mathbf{M}_N^{\text{nr}}$ where $\mathbf{M}_N^{\text{nr}}$ 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

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

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

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

Now, for $i \neq j$,

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

for some constant $C' > 0$. Therefore, for some new constant $C'_1 > 0$,

$$\mathbb{E} \left[d_L^3 (\text{ESD}(\mathbf{A}_N^{\text{grg}}), \text{ESD}(\mathbf{A}_N^{\text{nr}})) \right] \leq \frac{C'_1}{\lambda N} \frac{m_2^2}{m_1^2} \quad (2.21)$$

where $m_2 = \sum_{i=1}^N d_i^2$. Since W has compact support, we have that $\frac{m_2}{Nm_\infty} \rightarrow \mathbb{E}[W^2]$ and $\frac{m_1}{Nm_\infty} \rightarrow \mathbb{E}[W]$. 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

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

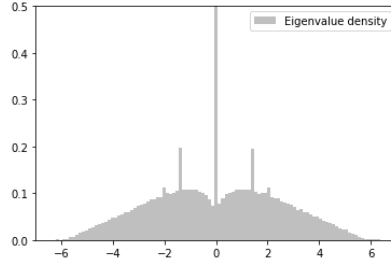


Figure 2.2: Spectral distributions for the Chung-Lu random graph, the generalised random graph, and the Norros-Reittu random graph on 10,000 vertices with $\{d_i\}_i$ uniformly generated integers in $[1, 5]$

Example 5: Inhomogeneous Random Graphs. Let $w_i = \frac{i}{N}$ and $f : [0, 1]^2 \rightarrow [0, 1]$ 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 $N\varepsilon_N \rightarrow \lambda$ 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.

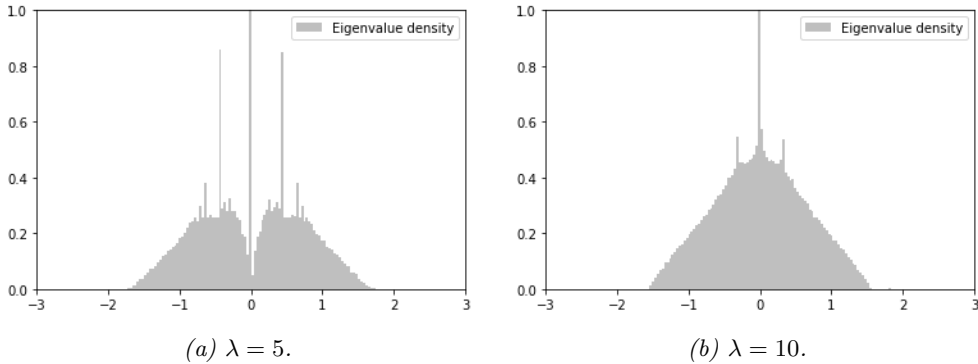


Figure 2.3: The Inhomogeneous Random Graph on 10,000 vertices, with the inhomogeneity function $f(x, y) = r_1(x)r_1(y) + r_2(x)r_2(y)$, where $r_1(x) = \frac{x}{1+x}$ and $r_2(x) = x$.

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 $W_N \xrightarrow{d} W$. This means that W_N has a distribution function $F_N(x)$ given by

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

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

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

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

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

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

$$\mathbb{E}[g(W_{N,1}, \dots, W_{N,k})] \rightarrow \mathbb{E}[g(W_1, \dots, W_k)]. \quad (2.22)$$

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 $\tilde{\mathbf{A}}_N$ be the matrix $\mathbf{A}_N$ with zero on the diagonal, and let d_L denote the Lévy distance. Then,

$$d_L(\text{ESD}(\tilde{\mathbf{A}}_N), \text{ESD}(\mathbf{A}_N)) \xrightarrow{\mathbb{P}} 0.$$

In particular, if $\text{ESD}(\mathbf{A}_N)$ converges weakly in probability to μ_λ , then so will $\text{ESD}(\tilde{\mathbf{A}}_N)$ 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

$$d_L^3(\text{ESD}(\tilde{\mathbf{A}}_N), \text{ESD}(\mathbf{A}_N)) \leq \frac{1}{N} \text{Tr}(D_N^2) = \frac{1}{N\lambda} \sum_{1 \leq i \leq N} a_{ii}^2.$$

Hence we have

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

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

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

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

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

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

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

$$\mathbb{E}[\text{tr}(\mathbf{A}_N^k)] = \frac{1}{N} \mathbb{E}[\text{Tr}(\mathbf{A}_N^k)] = \frac{1}{N \lambda^{k/2}} \sum_{1 \leq i_1, i_2, \dots, i_k \leq N} \mathbb{E}[a_{i_1 i_2} a_{i_2 i_3} \dots a_{i_k i_1}], \quad (2.24)$$

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 $\gamma = (1, 2, \dots, k)$ 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

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

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

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

Then,

$$\begin{aligned} \int x^k \bar{\mu}_{N,\lambda}(dx) = & \quad (2.26) \\ & \begin{cases} O(\lambda^{k/2} N^{-1}), & k \text{ odd} \\ \sum_{\pi \in SS(k)} \lambda^{|\gamma\pi| - 1 - k/2} \sum_{\mathbf{i} \in \mathcal{F}_{\gamma\pi}} \frac{1}{N^{|\gamma\pi|}} \prod_{(a,b) \in E_{\gamma\pi}} f(w_{i_a}, w_{i_b}) + O(\lambda^{k/2} N^{-1}), & k \text{ even} \end{cases} \end{aligned} \quad (2.27)$$

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

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

where $\mathbf{i} = (i_1, \dots, i_k)$. The term $a_{i_1 i_2} a_{i_2 i_3} \dots a_{i_k i_1}$ is associated with the closed walk $i_1 i_2 \dots i_k i_1$. 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

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

we can rewrite (2.24) as

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

Let π be a partition of $[k] := \{1, 2, \dots, k\}$ and $\gamma\pi = \{V_1, V_2, \dots, V_m\}$, 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

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

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

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

Note that since f is bounded, then the product is bounded. For a fixed k and a partition π of $[k]$, $|E_{\gamma\pi}| \leq k$. 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 $O(N^{-1})$. 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

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

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

$$\frac{1}{N\lambda^{k/2}} \mathbb{E}[\text{Tr}(\mathbf{A}_N^k)] \tag{2.29}$$

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

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 $G_{\gamma\pi}$.

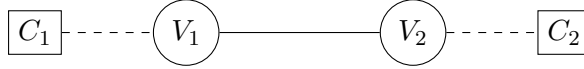


Figure 2.4: Graph associated to $\gamma\pi$ having blocks V_1 and V_2 with the edge between them traversed only once.

Thus, since the closed walk $1 \rightarrow 2, 2 \rightarrow 3, \dots, k \rightarrow 1$ 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 $\gamma\pi = \{V_1, \dots, V_l\}$, 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 $1 \rightarrow 2 \rightarrow 3 \rightarrow \dots \rightarrow k \rightarrow 1$, 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 $a - 1 \rightarrow a$ (or $a \rightarrow a + 1$) 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 $1 \rightarrow 2 \rightarrow \dots$ would imply that $a - 2 \in V_1$, which contradicts our claim. Now, if $t > 1$, the next edge $\{a - 1 \rightarrow a\}$ 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. $\square$

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

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. $\square$

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 $V = \{l_1, \dots, l_r\} \in \pi$ with $l_1 < l_2 < \dots < l_r$. 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. $\square$

Corollary 2.4.6 (Vanishing odd moments).

The odd moments vanish as $N \rightarrow \infty$.

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 $O(N^{-1})$. 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. $\square$

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.

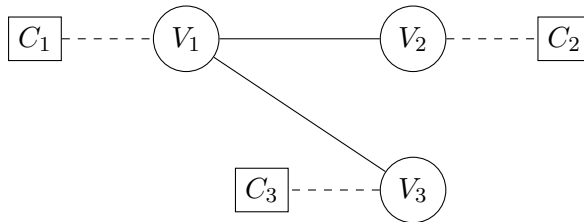


Figure 2.5: Graph associated to $\gamma\pi$ having blocks V_1 , V_2 and V_3 .

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 $\{k_1 \rightarrow k_1 + 1, \dots, k \rightarrow 1, 1 \rightarrow 2, \dots, k_1 - 1 \rightarrow k_1\}$. 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 $\{a_1 \rightarrow a_1 + 1, a_1 + 1 \rightarrow a_1 + 2, \dots, a_1 + e \rightarrow a_1 + e + 1\}$, 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 $\{a_2 \rightarrow a_2 + 1\}$. 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 . $\square$

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. $\square$

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.

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

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

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

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 $|V_{\gamma\pi}| = |\gamma\pi| \in \{2, 3, \dots, k/2 + 1\}$. In particular, for $\pi \in NC_2(k)$, $|\gamma\pi| = k/2 + 1$, and when π

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

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

§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

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

Proof. We shall proceed to compute the variance

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

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

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

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

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

Similarly, we have

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

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

$$1 \rightarrow 2 \rightarrow \dots k \rightarrow 1, k+1 \rightarrow k+2 \rightarrow \dots 2k \rightarrow k+1,$$

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

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

With this notation set up, one can see that

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

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

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

This completes the proof. $\square$

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

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

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):

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

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,

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

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

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

Thus,

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

Since $e^{-x} > 1 - x$, 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

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

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

$$S_{\mathbf{A}_N}(z) = \int_{\mathbb{R}} \frac{1}{x - z} \text{ESD}(\mathbf{A}_N)(dx) = \text{tr}(\mathbf{R}_{\mathbf{A}_N}(z)), \quad (2.36)$$

where tr denotes the normalised trace.

Lemma 2.5.1 (Resolvent Properties).

For any $z \in \mathbb{C}^+$, $1 \leq i, j \leq N$, the following properties are well-known for the resolvent $\mathbf{R}_{\mathbf{A}}$ of an $N \times N$ matrix $\mathbf{A}$.

- (i) **Analytic:** $z \mapsto \mathbf{R}_{\mathbf{A}}(z)(i, j)$ is an analytic function on $\mathbb{C}^+ \rightarrow \mathbb{C}^+$.
- (ii) **Bounded :** $\|\mathbf{R}_{\mathbf{A}}(z)\|_{\text{op}} \leq \Im(z)^{-1}$, where $\|\cdot\|_{\text{op}}$ denotes the operator norm.
- (iii) **Normal :** $\mathbf{R}_{\mathbf{A}}(z) \mathbf{R}_{\mathbf{A}}(z)^* = \mathbf{R}_{\mathbf{A}}(z)^* \mathbf{R}_{\mathbf{A}}(z)$.
- (iv) **Diagonals are bounded:** $|\mathbf{R}_{\mathbf{A}}(z)(i, j)| \leq \Im(z)^{-1}$.
- (v) **Trace bounded:** $|\text{tr}(\mathbf{R}_{\mathbf{A}}(z))| \leq \Im(z)^{-1}$. In particular,

$$|\text{tr}(\mathbf{R}_{\mathbf{A}}^p(z))| \leq \Im(z)^{-p}, \quad \text{for any } p \geq 1.$$

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

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

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 $\mathbf{R}_{\mathbf{A}}$ be the resolvent. Let $z \in \mathbb{C}^+$. Then for any fixed k , we have

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

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

$$\lim_{N \rightarrow \infty} S_{\mathbf{A}_N}(z) \rightarrow S_{\mu_\lambda}(z).$$

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 $\bar{S}_{\mathbf{A}_N}(z)$ denote the Stieltjes transform of $\bar{\mu}_{N,\lambda}$. Notice that $\bar{S}_{\mathbf{A}_N}(z) = \mathbb{E}[S_{\mathbf{A}_N}(z)]$. 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 $S_{\mu_N}(z)$ converges in probability to $S_\mu(z)$ for each $z \in \mathbb{C}^+$.

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

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

$$e^{iuz} = 1 - \sqrt{u} \int_0^\infty \frac{J_1(2\sqrt{uv})}{\sqrt{v}} e^{-ivz^{-1}} dv, \quad (2.37)$$

where $J_1(x)$ is the first-order Bessel function of the first kind given by (2.16). Note that for all $x \geq 0$, $|J_1(x)| \leq 1$ (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 $r_{jj}^N(z)$, the j^{th} diagonal entry of the $N \times N$ resolvent matrix, as our complex variable in $\mathbb{C}^+$. So we can get

$$e^{iur_{jj}^N(z)} = 1 - \sqrt{u} \int_0^\infty \frac{J_1(2\sqrt{uv})}{\sqrt{v}} e^{-iv(r_{jj}^N(z))^{-1}} dv. \quad (2.38)$$

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

$$S_{\mathbf{A}_N}(z) = \frac{\partial}{\partial u} \frac{1}{N} \sum_{j=1}^N e^{iur_{jj}^N(z)} \bigg|_{u=0}. \quad (2.39)$$

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 $R_{\mathbf{A}_N}(z)$. Let

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

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

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

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

$$\mathbb{E}[e^{r_{jj}^N}] = 1 - e^{-\lambda d_j} \sqrt{u} \int_0^\infty \frac{J_1(2\sqrt{uv})}{\sqrt{v}} e^{ivz} \mathbb{E} \left[e^{\lambda g_N(w_j, \frac{v}{\lambda}, z)} \right] dv + q_{N,\lambda}(u, z), \quad (2.42)$$

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

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 $z_1, z_2 \in \mathbb{C}^+$.

$$|e^{aiz_1} - e^{aiz_2}| \leq |a| |z_1 - z_2| \quad (2.43)$$

$$|e^a - e^b| \leq |a - b| e^{|a|+|b|} \quad (2.44)$$

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

$$e^{iur_{jj}^N} = 1 - \sqrt{u} \int_0^\infty \frac{J_1(2\sqrt{uv})}{\sqrt{v}} e^{ivz} \exp \left(iv \sum_{k,l \neq j} \tilde{r}_{kl}^{N-1} a_{kj} a_{lj} \right) dv. \quad (2.45)$$

Adding and subtracting the appropriate exponential to (2.45) yields

$$e^{iur_{jj}^N} = 1 - \sqrt{u} \int_0^\infty \frac{J_1(2\sqrt{uv})}{\sqrt{v}} e^{ivz} \exp \left(iv \sum_{k \neq j} \tilde{r}_{kk}^{N-1} a_{kj}^2 \right) dv + E_1, \quad (2.46)$$

where E_1 is an error term given by

$$E_1 = \sqrt{u} \int_0^\infty \frac{J_1(2\sqrt{uv})}{\sqrt{v}} e^{ivz} \left(\exp \left(iv \sum_{k,l \neq j} \tilde{r}_{kl}^{N-1} a_{kj} a_{lj} \right) - \exp \left(iv \sum_{k \neq j} \tilde{r}_{kk}^{N-1} a_{kj}^2 \right) \right) dv.$$

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 $|e^{ivz}| = |e^{i\zeta v} e^{-\eta v}| \leq e^{-\eta v}$. Thus,

$$\begin{aligned} |E_1| &= \left| \sqrt{u} \int_0^\infty \frac{J_1(2\sqrt{uv})}{\sqrt{v}} e^{ivz} \left(\exp \left(iv \sum_{k,l \neq j} \tilde{r}_{kl}^{N-1} a_{kj} a_{lj} \right) - \exp \left(iv \sum_{k \neq j} \tilde{r}_{kk}^{N-1} a_{kj}^2 \right) \right) dv \right| \\ &\leq \sqrt{u} \int_0^\infty \frac{ve^{-\eta v}}{\sqrt{v}} \sum_{k \leq N} \sum_{l \neq k} |\tilde{r}_{kl}^{N-1}| a_{kj} a_{lj} dv \\ &= \left(\sqrt{u} \int_0^\infty \sqrt{v} e^{-\eta v} dv \right) \sum_{k \leq N} \sum_{l \neq k} |\tilde{r}_{kl}^{N-1}| a_{kj} a_{lj} \end{aligned} \quad (2.47)$$

where in the last step, we use inequality (2.43) and the bound $|J_1(x)| \leq 1$ 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 $\tilde{r}_{kl}^{N-1}$. Thus, since f is bounded by a constant C_f , taking expectation on the summation gives us

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

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

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

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

$$\mathbb{E}[e^{iur_{jj}^N}] = 1 - \sqrt{u} \int_0^\infty \frac{J_1(2\sqrt{uv})}{\sqrt{v}} e^{ivz} \mathbb{E} \left[\exp \left(iv \sum_{k \neq j} \tilde{r}_{kk}^{N-1} a_{kj}^2 \right) \right] \, dv + q_{N,\lambda}(u, z), \quad (2.49)$$

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

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

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

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

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

$$|q'_k(N, \lambda)| \leq \frac{4C_f^2 \lambda^2}{N^2} = O \left(\frac{\lambda^2}{N^2} \right). \quad (2.51)$$

We can write

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

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 $|e^{i v \tilde{r}_{kk}^{N-1} / \lambda} - 1| \leq 2$, and again using (2.51), we have

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

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

$$\begin{aligned} |E_2| &\leq C_a \sum_{j=1}^N (2C_f \lambda)^{2j} N^j e^{-\frac{2j\lambda C_f}{N}} N^{-2j} \\ &= C_a \sum_{j=1}^N (2C_f \lambda)^{2j} N^{-j} e^{-\frac{2j\lambda C_f}{N}} = C_a \frac{4C_f \lambda^2 N^{-1} e^{-\frac{2\lambda C_f}{N}}}{1 - 4C_f \lambda^2 N^{-1} e^{-\frac{2\lambda C_f}{N}}}, \end{aligned}$$

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

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

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

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

where

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

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 $\mathbf{R}_{\mathbf{A}_N^t}(z)$, 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

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

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

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

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

Note that $\frac{d\mathbf{A}_N^t}{dt} = -J_N$, 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,

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

since the complex exponential $e^{ibr_{kk}^N(t)}$ 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

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

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

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

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

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

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

$$\begin{aligned} \mathbb{E}[|g_N(w_j, b, z) - \tilde{g}_{N-1}(w_j, b, z)|] &\leq \int_0^1 \frac{b}{N\eta} \mathbb{E} \left[\frac{C' \sqrt{N}}{\sqrt{\eta}} \left(\sum_{m=1}^N a_{mj}^2 \right)^{1/2} \right] dt \\ &\leq C' \int_0^1 \frac{b}{\sqrt{N}\eta^{3/2}} \left(\mathbb{E} \left[\sum_{m=1}^N a_{mj}^2 \right] \right)^{1/2} dt. \end{aligned}$$

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

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

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

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

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

$$\mathbb{E}[e^{iur_{jj}^N}] = 1 - e^{-\lambda d_j} \sqrt{u} \int_0^\infty \frac{J_1(2\sqrt{uv})}{\sqrt{v}} e^{ivz} \mathbb{E} \left[e^{\lambda g_N(w_j, \frac{v}{\lambda}, z)} \right] dv + q_{N,\lambda}(u, z),$$

which proves the proposition. $\square$

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

$$\begin{aligned} \mathbb{E}[g_N(x, u, z)] &= \frac{1}{N} \sum_{j=1}^N f(x, w_j) \\ &\quad - \frac{1}{N} \sum_{j=1}^N f(x, w_j) e^{-\lambda d_j} \sqrt{u} \int_0^\infty \frac{J_1(2\sqrt{uv})}{\sqrt{v}} e^{ivz} \mathbb{E} \left[e^{\lambda g_N(w_j, \frac{v}{\lambda}, z)} \right] dv \\ &\quad + q_{N,\lambda}(u, z). \end{aligned} \quad (2.59)$$

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

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

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

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

we have

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

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

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

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

Recall from (2.13) that we have

$$d_f(w_j) := \int f(x, w_j) \mu_w(dx).$$

Then, one simply gets

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

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

$$\begin{aligned} & \mathbb{E}[g_N(x, u, z)] \\ &= d_f(x) - \frac{1}{N} \sum_{j=1}^N f(x, w_j) e^{-\lambda d_f(w_j)} \left(\sqrt{u} \int_0^\infty \frac{J_1(2\sqrt{uv})}{\sqrt{v}} e^{ivz} \mathbb{E} \left[e^{\lambda g_N(w_j, \frac{v}{\lambda}, z)} \right] dv \right) \end{aligned} \quad (2.62)$$

$$+ \tilde{q}_{N,\lambda}(u, z),$$

where

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

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

$$\mathbf{I}_g(y) = f(x, y) e^{-\lambda d_f(y)} \left(\sqrt{u} \int_0^\infty \frac{J_1(2\sqrt{uv})}{\sqrt{v}} e^{ivz} \mathbb{E} \left[e^{\lambda g_N(y, \frac{v}{\lambda}, z)} \right] dv \right).$$

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,

$$\begin{aligned}
 & |\partial_y \mathbf{I}_g(y)| \\
 & \leq \left| \partial_y f(x, y) e^{-\lambda d_f(y)} \left(\sqrt{u} \int_0^\infty \frac{J_1(2\sqrt{uv})}{\sqrt{v}} e^{ivz} \mathbb{E} \left[e^{\lambda g_N(y, \frac{v}{\lambda}, z)} \right] dv \right) \right| \\
 & + \left| f(x, y) e^{-\lambda d_f(y)} \partial_y d_f(y) \left(\sqrt{u} \int_0^\infty \frac{J_1(2\sqrt{uv})}{\sqrt{v}} e^{ivz} \mathbb{E} \left[e^{\lambda g_N(y, \frac{v}{\lambda}, z)} \right] dv \right) \right| \\
 & + \left| f(x, y) e^{-\lambda d_f(y)} \left(\sqrt{u} \int_0^\infty \frac{J_1(2\sqrt{uv})}{\sqrt{v}} e^{ivz} \mathbb{E} \left[e^{\lambda g_N(y, \frac{v}{\lambda}, z)} \right] \partial_y g_N(y, v/\lambda, z) dv \right) \right|.
 \end{aligned} \tag{2.63}$$

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

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

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 $\partial_y f(x, y) \equiv \text{const.}$, we have $\Psi_y(x) \equiv \text{const.}$, which is integrable on $[0, \infty)$ since μ_w is a probability measure.

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

$$h(y, v) = \frac{J_1(2\sqrt{uv})}{\sqrt{v}} e^{ivz} \mathbb{E} \left[e^{\lambda g_N(y, v, z)} \right].$$

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

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

which again is bounded. Thus, $\partial_y h(y, v)$ exists, and is bounded above by $C_0 v^{-\frac{1}{2}} e^{-\eta v}$, 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. $\square$

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

$$\begin{aligned} & \mathbb{E}[g_N(x, u, z)] \\ &= d_f(x) \\ & - \int_0^\infty f(x, y) e^{-\lambda d_f(y)} \left(\sqrt{u} \int_0^\infty \frac{J_1(2\sqrt{uv})}{\sqrt{v}} e^{ivz} \mathbb{E} \left[e^{\lambda g_N(y, \frac{v}{\lambda}, z)} \right] dv \right) \mu_w(dy) \\ & + \tilde{q}_{N, \lambda}(u, z). \end{aligned} \tag{2.65}$$

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

$$\begin{aligned} & F_z(\phi)(x, u) \\ &= d_f(x) - \sqrt{u} \int_0^\infty f(x, y) e^{-\lambda d_f(y)} \left(\sqrt{u} \int_0^\infty \frac{J_1(2\sqrt{uv})}{\sqrt{v}} e^{ivz} e^{\lambda \phi(y, \frac{v}{\lambda}, z)} dv \right) \mu_w(dy). \end{aligned} \tag{2.66}$$

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 $\|\phi_1\|, \|\phi_2\| \leq C_f$. Then, using the norm we

defined in (2.11) and inequality 2.44, from (2.66) we get

$$\begin{aligned}
 & \|F_z(\phi_1) - F_z(\phi_2)\| \\
 & \leq \sup_{x,u \geq 0} \sqrt{\frac{1}{1+u}} \left| \int_0^\infty f(x,y) e^{-\lambda d_f(y)} \right. \\
 & \quad \times \left(\sqrt{u} \int_0^\infty \frac{J_1(2\sqrt{uv})}{\sqrt{v}} e^{ivz} \left(e^{\lambda \phi_1(y, \frac{v}{\lambda})} - e^{\lambda \phi_2(y, \frac{v}{\lambda})} \right) dv \right) \mu_w(dy) \Big| \\
 & \leq \sup_{u \geq 0} \sqrt{\frac{1}{1+u}} \int_0^\infty \int_0^\infty \frac{\lambda}{\sqrt{v}} e^{-\eta v} \left| \phi_1\left(y, \frac{v}{\lambda}\right) - \phi_2\left(y, \frac{v}{\lambda}\right) \right| \\
 & \quad \times e^{\lambda |\phi_1(y, \frac{v}{\lambda})| + \lambda |\phi_2(y, \frac{v}{\lambda})|} dv \mu_w(dy) \\
 & \leq \lambda \|\phi_1 - \phi_2\| \int_0^\infty \int_0^\infty \frac{\lambda}{\sqrt{v}} e^{-\eta v} \sup_{y, v \geq 0} \frac{\sqrt{1+v/\lambda}}{\sqrt{1+v/\lambda}} e^{\lambda |\phi_1(y, \frac{v}{\lambda})| + \lambda |\phi_2(y, \frac{v}{\lambda})|} dv \mu_w(dy) \\
 & \leq \lambda \|\phi_1 - \phi_2\| \int_0^\infty \int_0^\infty \frac{\sqrt{1+v/\lambda}}{\sqrt{v}} e^{-\eta v} \exp\left(\lambda \sqrt{1+v/\lambda} (\|\phi_1\| + \|\phi_2\|)\right) dv \mu_w(dy) \\
 & \leq \|\phi_1 - \phi_2\| \int_0^\infty \left(\frac{e^{-\eta v}}{\sqrt{v}} + \frac{e^{-\eta v}}{\sqrt{\lambda}} \right) e^{2C_f \sqrt{\lambda v}} dv \leq \frac{C_1}{\eta^{5/2}} \|\phi_1 - \phi_2\|,
 \end{aligned}$$

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

$$\int_0^\infty c_1 e^{-c_2 x + c_3 \sqrt{x}} dx$$

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 $F_z : B \rightarrow B$.

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

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

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 $\mathbb{E}[g_N(x, u, z)] \xrightarrow{N \rightarrow \infty} \phi^*(x, u)$. Further, we have

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

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

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

Let $1 \leq k \neq l \leq N$ and consider the covariance

$$A_{k,l} := \mathbb{E}[e^{iur_{kk}^N} e^{iur_{ll}^N}] - \mathbb{E}[e^{iur_{kk}^N}] \mathbb{E}[e^{iur_{ll}^N}]. \quad (2.67)$$

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

$$\begin{aligned} A_{k,l} &= \\ &- \mathbb{E}[T_j] - \mathbb{E}[T_k] \\ &+ u \int \int \frac{J_1(2\sqrt{uv_1})}{\sqrt{v_1}} \frac{J_1(2\sqrt{uv_2})}{\sqrt{v_2}} e^{i(v_1+v_2)z} \mathbb{E} \left[e^{iv_1 \sum_{l \neq j} \tilde{r}_{il}^{N-1} a_{jl}^2 + iv_2 \sum_{l \neq k} \tilde{r}_{il}^{N-1} a_{kl}^2} \right] dv_1 dv_2 \\ &+ \mathbb{E}[\tilde{T}_j] + \mathbb{E}[\tilde{T}_k] \\ &- u \int \int \frac{J_1(2\sqrt{uv_1})}{\sqrt{v_1}} \frac{J_1(2\sqrt{uv_2})}{\sqrt{v_2}} e^{i(v_1+v_2)z} \mathbb{E} \left[e^{\lambda g_N(w_j, \frac{v_1}{\lambda}, z) + \lambda g_N(w_k, \frac{v_2}{\lambda}, z)} \right] dv_1 dv_2, \end{aligned} \quad (2.68)$$

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 $C_I q_{N,\lambda}^2(u, z) + q_{N,\lambda}^2(u, z)$ 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

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

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

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

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

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

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

$$\mathbb{E}[|\delta_N(x, u, z)|^2] \leq C_1 C_f^2 \sup_{k,l} |A_{k,l}| = O\left(\frac{\lambda \sqrt{u}}{\eta^{5/2} \sqrt{N}}\right). \quad (2.71)$$

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

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

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 $\delta_N \rightarrow 0$ 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

$$\begin{aligned} & \mathbb{E}[g_N(x, u, z)] \\ & = d_f(x) - \int_0^\infty f(x, y) e^{-\lambda d_f(y)} \left(\sqrt{u} \int_0^\infty \frac{J_1(2\sqrt{uv})}{\sqrt{v}} e^{ivz} e^{\lambda \mathbb{E}[g_N(y, \frac{v}{\lambda}, z)]} dv \right) \mu_w(dy) \\ & + \tilde{q}_{N,\lambda}(u, z), \end{aligned} \quad (2.73)$$

where the error vanishes in the norm as

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

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

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

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

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

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,

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

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

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

As a quick remark, notice that

$$\|\phi^*\| \leq C_f, \quad (2.76)$$

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 $\lim_{N \rightarrow \infty} \mathbb{E}[g_N(x, u, z)]$ is an analytic function. Again from the identity theorem of complex analysis, since $\lim_{N \rightarrow \infty} \mathbb{E}[g_N]$ 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 $u \in [0, 1]$,

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

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

$$\mathbb{E} \left[e^{i u r_{jj}^N} \right] = 1 - e^{-\lambda d_j} \sqrt{u} \int_0^\infty \frac{J_1(2\sqrt{uv})}{\sqrt{v}} e^{ivz} \mathbb{E} \left[e^{\lambda g_N(w_j, \frac{v}{\lambda}, z)} \right] dv + q_{N,\lambda}(u, z),$$

and so,

$$\begin{aligned} \mathbb{E}[G_N(u, z)] &= \frac{1}{N} \sum_{j=1}^N \mathbb{E}[e^{i u r_{jj}^N}] \\ &= 1 - \frac{1}{N} \sum_{j=1}^N e^{-\lambda d_j} \sqrt{u} \int_0^\infty \frac{J_1(2\sqrt{uv})}{\sqrt{v}} e^{ivz} \mathbb{E} \left[e^{\lambda g_N(w_j, \frac{v}{\lambda}, z)} \right] dv + q_{N,\lambda}(u, z). \end{aligned} \quad (2.78)$$

Next, we see that the function

$$\tilde{\mathbf{I}}_g(y) = e^{-\lambda d_f(y)} \sqrt{u} \int_0^\infty \frac{J_1(2\sqrt{uv})}{\sqrt{v}} e^{ivz} e^{\lambda \phi^*(y, v/\lambda)} dv$$

is Lipschitz by using an argument similar to Lemma 2.5.6. Thus, we get

$$\begin{aligned} \mathbb{E}[G_N(u, z)] &= \\ 1 - \int_0^\infty e^{-\lambda d_f(y)} \left(\sqrt{u} \int_0^\infty \frac{J_1(2\sqrt{uv})}{\sqrt{v}} e^{ivz} \mathbb{E} \left[e^{\lambda g_N(y, \frac{v}{\lambda}, z)} \right] dv \right) \mu_w(dy) + \tilde{q}_{N,\lambda}(u, z). \end{aligned}$$

Since from Proposition 2.5.7 we have concentration for g_N , using inequality (2.44) we have that

$$\begin{aligned} \mathbb{E}[G_N(u, z)] &= \\ 1 - \int_0^\infty e^{-\lambda d_f(y)} \sqrt{u} \int_0^\infty \frac{J_1(2\sqrt{uv})}{\sqrt{v}} e^{ivz} e^{\lambda \mathbb{E}[g_N(y, \frac{v}{\lambda}, z)]} dv \mu_w(dy) + \tilde{q}_{N,\lambda}(u, z). \end{aligned}$$

Finally, taking the limit $N \rightarrow \infty$ gives us

$$\lim_{N \rightarrow \infty} \mathbb{E}[G_N(u, z)] = 1 - \int_0^\infty e^{-\lambda d_f(y)} \sqrt{u} \int_0^\infty \frac{J_1(2\sqrt{uv})}{\sqrt{v}} e^{ivz} e^{\lambda \phi^*(y, v/\lambda)} dv \mu_w(dy), \quad (2.79)$$

completing the proof of Theorem 2.3.9. $\square$

§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

$$\begin{aligned} \mathbb{E}[G_N(u, z)] &= \frac{1}{N} \mathbb{E} \sum_{j=1}^N e^{iur_{jj}^N} \\ &= 1 - \frac{1}{N} \sum_{j=1}^N e^{-\lambda d_j} \sqrt{u} \int_0^\infty \frac{J_1(2\sqrt{uv})}{\sqrt{v}} e^{ivz} \mathbb{E} \left[e^{\lambda g_N(w_j, \frac{v}{\lambda}, z)} \right] dv + q_{N, \lambda}(u, z). \end{aligned}$$

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 $\text{tr}(\mathbf{R}_{\mathbf{A}_N}(z))$ from the LHS of (2.78). On the other hand, we would first like to take $N \rightarrow \infty$ for the RHS to remove the error term. To interchange these operations, we have the following result.

Proposition 2.5.8.

Both the limits $\lim_{N \rightarrow \infty} \frac{\partial}{\partial u} \mathbb{E}[G_N(u, z)] \Big|_{u=0}$ and $\frac{\partial}{\partial u} \lim_{N \rightarrow \infty} \mathbb{E}[G_N(u, z)] \Big|_{u=0}$ exist and are equal.

Proof. We fix a $z \in \mathbb{C}^+$. Now, $\lim_{N \rightarrow \infty} \mathbb{E}[G_N(u, z)]$ 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

$$\begin{aligned} H_N(u, z) &= \frac{\mathbb{E}[G_N(u, z)] - \mathbb{E}[G_N(0, z)]}{u}, \\ H(u, z) &= \frac{G(u, z) - G(0, z)}{u}. \end{aligned}$$

Then,

$$\begin{aligned} \lim_{u \rightarrow 0} H_N(u, z) &= \frac{\partial}{\partial u} \mathbb{E}[G_N(u, z)] \Big|_{u=0}, \\ \lim_{u \rightarrow 0} H(u, z) &= \frac{\partial}{\partial u} G(u, z) \Big|_{u=0}. \end{aligned}$$

We would like to claim

$$\lim_{N \rightarrow \infty} \frac{\partial}{\partial u} \mathbb{E}[G_N(u, z)] \Big|_{u=0} = \frac{\partial}{\partial u} G(u, z) \Big|_{u=0}.$$

Thus, we want to interchange the order of limits. Note that

$$\lim_{N \rightarrow \infty} H_N(u, z) = H(u, z)$$

uniformly in $u \in (0, 1]$, and

$$\lim_{u \rightarrow 0} H_N(u, z) = \frac{\partial}{\partial u} \mathbb{E}[G_N(u, z)] \Big|_{u=0} = \mathbb{E}[\text{tr}(\mathbf{R}_{\mathbf{A}_N}(z))]$$

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 $\lim_{u \rightarrow 0} H(u, z)$ and $\lim_{N \rightarrow \infty} \mathbb{E}[\text{tr}(\mathbf{R}_{\mathbf{A}_N}(z))]$ exist and are equal. $\square$

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 $N \rightarrow \infty$ on the LHS of (2.78), and we do the reverse for the RHS of (2.78). Note that since $\lim_{N \rightarrow \infty} \mu_{N, \lambda} = \mu_\lambda$ in probability, $\mathbf{S}_{\mathbf{A}_N}(z) \rightarrow S_{\mu_\lambda}(z)$ and also $\bar{\mathbf{S}}_{\mathbf{A}_N}(z) \rightarrow S_{\mu_\lambda}(z)$ as $N \rightarrow \infty$ for all $z \in \mathbb{C}^+$. Thus, we then obtain using Proposition 2.5.8

$$\begin{aligned} & i S_{\mu_\lambda}(z) \\ &= i \lim_{N \rightarrow \infty} \bar{\mathbf{S}}_{\mathbf{A}_N}(z) \stackrel{(2.36)}{=} i \lim_{N \rightarrow \infty} \mathbb{E} \text{tr}(\mathbf{R}_{\mathbf{A}_N}(z)) \stackrel{(2.39)}{=} \lim_{N \rightarrow \infty} \frac{\partial}{\partial u} \mathbb{E}[G_N(u, z)] \Big|_{u=0} \\ &= \frac{\partial}{\partial u} \lim_{N \rightarrow \infty} \mathbb{E}[G_N(u, z)] \Big|_{u=0} \\ &\stackrel{(2.79)}{=} - \frac{\partial}{\partial u} \int_0^\infty e^{-\lambda d_f(y)} \sqrt{u} \int_0^\infty \frac{J_1(2\sqrt{uv})}{\sqrt{v}} e^{ivz} e^{\lambda \phi_z^*(y, \frac{v}{\lambda})} dv \mu_w(dy) \Big|_{u=0} \\ &= - \int_0^\infty e^{-\lambda d_f(y)} \frac{\partial}{\partial u} \sqrt{u} \int_0^\infty \frac{J_1(2\sqrt{uv})}{\sqrt{v}} e^{ivz} e^{\lambda \phi_z^*(y, \frac{v}{\lambda})} dv \mu_w(dy) \Big|_{u=0}. \quad (2.80) \end{aligned}$$

We now wish to evaluate the derivative on the RHS of (2.80). Let $K(u)$ denote

$$K(u) := \sqrt{u} \int_0^\infty \frac{J_1(2\sqrt{uv})}{\sqrt{v}} e^{ivz} e^{\lambda \phi_z^*(y, \frac{v}{\lambda})} dv. \quad (2.81)$$

Observe that

$$\sum_{k \geq 0} \int_0^\infty \frac{v^k}{k!(k+1)!} e^{-\eta v} dv = \sum_{k \geq 0} \frac{\Gamma(k+1)}{k!(k+1)! \eta^k} \leq e^{1/\eta} \quad (2.82)$$

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

$$\begin{aligned} K(u) &= \sqrt{u} \int_0^\infty \frac{J_1(2\sqrt{uv})}{\sqrt{v}} e^{ivz} e^{\lambda\phi_z^*(y, \frac{v}{\lambda})} dv \\ &= \sqrt{u} \int_0^\infty \frac{1}{\sqrt{v}} \sum_{k=0}^\infty \frac{(-1)^k (\sqrt{uv})^{2k+1}}{k!(k+1)!} e^{ivz} e^{\lambda\phi_z^*(y, \frac{v}{\lambda})} dv \\ &= \sum_{k=0}^\infty \frac{(-1)^k u^{k+1}}{k!(k+1)!} \int_0^\infty v^k e^{ivz} e^{\lambda\phi_z^*(y, \frac{v}{\lambda})} dv. \end{aligned}$$

Denote by $I_k(y)$ the integral

$$I_k(y) := \int v^k e^{ivz} e^{\lambda\phi_z^*(y, \frac{v}{\lambda})} dv.$$

Therefore,

$$\frac{K(u)}{u} = \sum_{k=0}^\infty \frac{(-1)^k u^k}{k!(k+1)!} I_k(y) = I_0(y) + \sum_{k \geq 1} \frac{(-1)^k u^k}{k!(k+1)!} I_k(y) =: I_0(y) + \sum_{k=1}^\infty a_k(u), \quad (2.83)$$

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

$$|I_k(y)| \leq \int_0^\infty v^k e^{-\eta v} e^{C_f \lambda} dv = \frac{e^{C_f \lambda}}{\eta^{k+1}} \Gamma(k+1).$$

Since $K(0) = 0$ and by (2.83) it follows that

$$\left. \frac{\partial}{\partial u} K(u) \right|_{u=0} = \lim_{u \rightarrow 0} \frac{K(u)}{u} = I_0(y) + \lim_{u \rightarrow 0} \sum_{k \geq 1} a_k(u), \quad (2.84)$$

Therefore we would like to evaluate $\lim_{u \rightarrow 0} \sum_{k \geq 1} a_k(u)$. Note that

$$|a_k(u)| \leq \frac{e^{C_f \lambda} \Gamma(k+1)}{\eta^{k+1} k!(k+1)!}$$

, as u is bounded by 1. Note that the series

$$\sum_{k \geq 1} \frac{\Gamma(k+1) e^{C_f \lambda}}{k!(k+1)! \eta^{k+1}} = \frac{e^{C_f \lambda}}{\eta^2} \sum_{k \geq 0} \frac{1}{\eta^k (k+2)!} \leq \frac{e^{C_f \lambda} e^{\frac{1}{\eta}}}{\eta^2}$$

converges, and consequently by the dominated convergence theorem, we have

$$\lim_{u \rightarrow 0} \sum_{k \geq 1} a_k(u) = \sum_{k \geq 1} \lim_{u \rightarrow 0} a_k(u) = 0.$$

Thus by (2.84) we have

$$\lim_{u \rightarrow 0} \frac{K(u)}{u} = I_0(y).$$

Therefore we get

$$i S_{\mu_\lambda}(z) = - \int_0^\infty e^{-\lambda d_f(y)} I_0(y) \mu_w(dy) - \int_0^\infty e^{-\lambda d_f(y)} \int_0^\infty e^{ivz} e^{\lambda \phi_z^*(y, \frac{v}{\lambda})} dv \mu_w(dy).$$

To conclude the argument, we use Lemma 2.5.3 with Theorem 2.3.7 to state that $S_{\mathbf{A}_N}(z)$ converges in probability to $S_{\mu_\lambda}(z)$ for each $z \in \mathbb{C}^+$. $\square$

We conclude with the proof of Corollary 2.3.11

Proof of Corollary 2.3.11. From Corollary 2.3.10, we have

$$S_{\mu_\lambda}(z) = i \int_0^\infty \int_0^\infty e^{ivz} e^{-\lambda d_f(y) + \lambda \phi^*(y, v/\lambda)} dv \mu_w(dy).$$

Recall that

$$\begin{aligned} \phi^*(x, u) &= d_f(x) - \int_0^\infty f(x, y) e^{-\lambda d_f(y)} \left(\sqrt{u} \int_0^\infty \frac{J_1(2\sqrt{uv})}{\sqrt{v}} e^{ivz} e^{\lambda \phi^*(y, \frac{v}{\lambda})} dv \right) \mu_w(dy) \end{aligned} \quad (2.85)$$

is the unique analytical solution of the fixed point equation as in (2.66). Expanding the Bessel function $J_1(x)$ in (2.85) using (2.16) gives

$$\begin{aligned} \phi^*(x, u) &= d_f(x) - \int_0^\infty f(x, y) e^{-\lambda d_f(y)} \left(\int_0^\infty \sum_{k \geq 0} \frac{(-1)^k u^{k+1} v^k}{k!(k+1)!} e^{ivz} e^{\lambda \phi^*(y, \frac{v}{\lambda})} dv \right) \mu_w(dy). \end{aligned} \quad (2.86)$$

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

$$\begin{aligned} &\sum_{k \geq 0} \int_0^\infty \left| \frac{(-1)^k u^{k+1} v^k}{k!(k+1)!} e^{ivz} e^{-\lambda d_f(y) + \lambda \phi^*(y, v/\lambda)} \right| dv \\ &\leq e^{C_f \lambda - \lambda d_f(y)} \sum_{k \geq 0} \frac{u^{k+1} \Gamma(k+1)}{k!(k+1)! \eta^{k+1}} \leq \frac{u}{\eta} e^{C_f \lambda - \lambda d_f(y)} e^{u/\eta}. \end{aligned}$$

Thus, by Fubini's Theorem, we can interchange the summand with the integral with respect to v , giving us

$$\phi^*(x, u) = d_f(x) - \int_0^\infty f(x, y) e^{-\lambda d_f(y)} \left(\sum_{k \geq 0} \frac{(-1)^k u^{k+1}}{k!(k+1)!} \int_0^\infty v^k e^{ivz} e^{\lambda \phi^*(y, v/\lambda)} dv \right) \mu_w(dy). \quad (2.87)$$

Now, denote by $\mathcal{H}^\lambda(z, y)$ the function

$$\mathcal{H}^\lambda(z, y) := i \int_0^\infty e^{ivz} e^{-\lambda d_f(y) + \lambda \phi^*(y, v/\lambda)} dv. \quad (2.88)$$

Then, by Corollary 2.3.10, we can see that $S_{\mu_\lambda}(z) = \int_0^\infty \mathcal{H}^\lambda(z, y) \mu_w(dy)$. From (2.87) we get that

$$\begin{aligned} \phi^*(x, u) &= d_f(x) - u \int_0^\infty f(x, y) \int_0^\infty e^{ivz} e^{-\lambda d_f(y) + \lambda \phi^*(y, v/\lambda)} dv \mu_w(dy) \\ &\quad - \int_0^\infty f(x, y) \sum_{k \geq 1} \int_0^\infty \frac{(-1)^k u^{k+1} v^k}{k!(k+1)!} e^{ivz} e^{-\lambda d_f(y) + \lambda \phi^*(y, v/\lambda)} dv \mu_w(dy), \end{aligned}$$

and so, we can write

$$\phi^*(x, u) = d_f(x) + iu \int_0^\infty f(x, y) \mathcal{H}^\lambda(z, y) \mu_w(dy) + T(x, u, \lambda, z) \quad (2.89)$$

where

$$\begin{aligned} T(x, u, \lambda, z) &:= - \int_0^\infty f(x, y) \sum_{k \geq 1} \int_0^\infty \frac{(-1)^k u^{k+1} v^k}{k!(k+1)!} e^{ivz} e^{-\lambda d_f(y) + \lambda \phi^*(y, v/\lambda)} dv \mu_w(dy). \end{aligned} \quad (2.90)$$

Substituting $u = v/\lambda$ for $v \in \mathbb{R}_+$ in (2.89) and multiplying throughout by λ , we have

$$-\lambda d_f(x) + \lambda \phi^*(x, v/\lambda) = iv \int_0^\infty f(x, y) \mathcal{H}^\lambda(z, y) \mu_w(dy) + \lambda T(x, v/\lambda, \lambda, z).$$

We begin by claiming the following:

Claim 2.5.9.

For any $x, u \geq 0$, we have

$$|e^{-\lambda d_f(x) + \lambda \phi^*(x, u)}| \leq 1. \quad (2.91)$$

Then, one can see that

$$\begin{aligned} |T(x, v/\lambda, \lambda, z)| &\leq \int_0^\infty f(x, y) \frac{v}{\lambda \eta} \left(\sum_{k \geq 1} \frac{v^k \Gamma(k+1)}{\eta^k \lambda^k k! (k+1)!} \right) \mu_w(dy) \\ &\leq \frac{v^2}{\lambda^2 \eta^2} e^{\frac{v}{\eta \lambda}} d_f(x), \end{aligned}$$

and so for each $v \in (0, \infty)$

$$\lim_{\lambda \rightarrow \infty} \lambda |T(x, v/\lambda, \lambda, z)| \rightarrow 0.$$

Thus, from (2.89), for any v we have

$$\lim_{\lambda \rightarrow \infty} (-\lambda d_f(x) + \lambda \phi^*(x, v/\lambda)) = iv \lim_{\lambda \rightarrow \infty} \int_0^\infty f(x, y) \mathcal{H}^\lambda(z, y) \mu_w(dy). \quad (2.92)$$

What remains now is to justify Claim 2.5.9, and taking the limit $\lambda \rightarrow \infty$ 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 $\phi^*(u) = \phi^*(x, v/\lambda)$. Then,

$$\widetilde{\phi^*}(u) = 1 - \sqrt{u} \int_0^\infty \frac{J_1(2\sqrt{uv})}{\sqrt{v}} e^{ivz} e^{-\lambda + \lambda \widetilde{\phi^*}(v/\lambda)} dv,$$

and from (2.92) we have $\lim_{\lambda \rightarrow \infty} (-\lambda + \lambda \widetilde{\phi^*}(v/\lambda)) = iv S_{\mu_f}(z)$. Moreover, from Corollary 2.3.10, we have

$$S_{\mu_\lambda}(z) = i \int_0^\infty e^{ivz} e^{-\lambda + \lambda \widetilde{\phi^*}(v/\lambda)} dv.$$

Since $f \equiv 1$, from (2.76) we have that $C_f = 1$ and $|\widetilde{\phi^*}| \leq 1$. Then, $|e^{-\lambda + \lambda \widetilde{\phi^*}}| \leq 1$, justifying Claim 2.5.9. Thus, the expression inside the integral is uniformly bounded by $e^{-\eta v}$. Using dominated convergence, we can pull the limit $\lambda \rightarrow \infty$ inside the integral to obtain

$$S_{\mu_f}(z) = i \int_0^\infty e^{ivz} e^{iv S_{\mu_f}(z)} dv = -\frac{1}{z + S_{\mu_f}(z)},$$

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 ,

$$\phi^*(x, u) = \lim_{N \rightarrow \infty} \frac{1}{N} \mathbb{E} \left[\sum_{i=1}^N f(x, w_i) e^{i u r_{ii}^N} \right].$$

Now, for any N , by trivially bounding the complex exponential $e^{iur_{ii}^N}$ by 1 for any i , we have that

$$\left| \frac{1}{N} \mathbb{E} \sum_{i=1}^N f(x, w_i) e^{iur_{ii}^N} \right| \leq \frac{1}{N} \sum_{i=1}^N |f(x, w_i)| = \frac{1}{N} \sum_{i=1}^N f(x, w_i).$$

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

$$\begin{aligned} -\frac{\lambda}{N} \sum_{i=1}^N f(x, w_i) + \lambda |\phi^*(x, u)| &\leq \lambda \sqrt{1+u} \frac{1}{\sqrt{1+u}} |\phi^*(x, u) - \mathbb{E}[g_N(x, u, z)]| \\ &\leq \lambda \sqrt{1+u} \|\phi^* - \mathbb{E}g_N\|_{\mathcal{B}}. \end{aligned} \quad (2.93)$$

Taking $N \rightarrow \infty$ on both sides in (2.93) yields that

$$-\lambda d_f(x) + \lambda |\phi^*(x, u)| \leq 0.$$

Using this, we conclude that

$$\left| e^{-\lambda d_f(x) + \lambda \phi^*(x, u)} \right| \leq e^{-\lambda d_f(x)} e^{\lambda |\phi^*(x, u)|} \leq 1 \quad (2.94)$$

for any x and u , proving Claim 2.5.9. Now, to evaluate $\lim_{\lambda \rightarrow \infty} S_{\mu_\lambda}(z)$, 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

$$S_{\mu_f}(z) = \lim_{\lambda \rightarrow \infty} S_{\mu_\lambda}(z) = i \int_0^\infty \int_0^\infty e^{ivz} \lim_{\lambda \rightarrow \infty} \left(e^{-\lambda d_f(y) + \lambda \phi^*(y, v/\lambda)} \right) dv \mu_w(dy).$$

and so, using (2.92), we get

$$S_{\mu_f}(z) = i \int_0^\infty \int_0^\infty e^{ivz} \lim_{\lambda \rightarrow \infty} e^{iv \int_0^\infty f(x, y) \mathcal{H}^\lambda(z, x) \mu_w(dx)} dv \mu_w(dy). \quad (2.95)$$

Recall from (2.88) that

$$\mathcal{H}^\lambda(z, y) = i \int_0^\infty e^{ivz} e^{-\lambda d_f(y) + \lambda \phi^*(y, v/\lambda)} dv.$$

Again using (2.94), we have that the integral is bounded in absolute value, and so, using DCT allows us to define

$$\mathcal{H}(z, y) := \lim_{\lambda \rightarrow \infty} \mathcal{H}^\lambda(z, y)$$

where $\int_0^\infty \mathcal{H}(z, y) \mu_w(dy) = S_{\mu_f}(z)$. 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 $\lambda \rightarrow \infty$ inside $\int_0^\infty f(x, y) \mathcal{H}^\lambda(z, x) \mu_w(dx)$. Thus, we obtain

$$\begin{aligned} S_{\mu_f}(z) &= i \int_0^\infty \int_0^\infty e^{ivz} e^{iv \int_0^\infty f(x, y) \mathcal{H}(z, x) \mu_w(dx)} dv \mu_w(dy) \\ &= - \int_0^\infty \frac{\mu_w(dy)}{z + \int_0^\infty f(x, y) \mathcal{H}(z, x) \mu_w(dx)}. \end{aligned}$$

The proof follows by observing that $\mathcal{H}(z, x)$ satisfies the analytic equation defined in (2.6). $\square$

§2.6 Appendix

Proposition 2.6.1 (Banach Space).

Let $X = [0, \infty)^2$ and consider the space $\mathcal{B}$ defined by

$$\mathcal{B} = \left\{ \phi : X \rightarrow \mathbb{C} \text{ analytic} \left| \sup_{x, y \geq 0} \frac{|\phi(x, y)|}{\sqrt{1+y}} < \infty \right. \right\}$$

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 $N_\epsilon \in \mathbb{N}$ such that for all $m, n > N_\epsilon$,

$$\|\phi_m - \phi_n\| < \epsilon.$$

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 $E = \bigcup_{m, n} E_{mn}$ 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

$$\psi(x, y) := \lim_{n \rightarrow \infty} \psi_n(x, y).$$

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 $\phi \in (\mathcal{B}, \|\cdot\|_{\mathcal{B}})$ $\square$

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 $f : A \times \Omega \rightarrow \mathbb{C}$ be such that for each $x \in A$, $\omega \mapsto f(x, \omega)$ is μ -integrable, and moreover for μ -a.e. ω , $x \mapsto f(x, \omega)$ is continuous. Consider the function $g : A \rightarrow \mathbb{C}$ defined by

$$g(x) = \int_{\Omega} f(x, \omega) \mu(d\omega).$$

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 μ -integrable function $h_x : \Omega \rightarrow \mathbb{C}$ 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

$$\partial_x g(x) = \int_{\Omega} \partial_x f(x, \omega) \mu(d\omega).$$