Spectral analysis of inhomogeneous network models

Malhotra, N.

CHAPTER 3

Adjacency spectra of kernel-based random graphs

This chapter is based on:
A. Cipriani, R.S. Hazra, N. Malhotra, M. Salvi. *Spectrum of dense kernel-based random graphs.* [*arxiv:2502:09415*], 2025.

Abstract

Kernel-based random graphs (KBRGs) are a broad class of random graph models that account for inhomogeneity among vertices. We consider KBRGs on a discrete $d-$dimensional torus $\mathbf{V}_N$ of size N^d. Conditionally on an i.i.d. sequence of Pareto weights $(W_i)_{i\in\mathbf{V}_N}$ with tail exponent $\tau-1>0$, we connect any two points i and j on the torus with probability

$$p_{ij}=\frac{\kappa_\sigma(W_i,W_j)}{\|i-j\|^\alpha}\wedge 1$$

for some parameter $\alpha>0$ and $\kappa_\sigma(u,v)=(u\vee v)(u\wedge v)^\sigma$ for some $\sigma\in(0,\tau-1)$. We focus on the adjacency operator of this random graph and study its empirical spectral distribution. For $\alpha<d$ and $\tau>2$, we show that a non-trivial limiting distribution exists as $N\to\infty$ and that the corresponding measure $\mu_{\sigma,\tau}$ is absolutely continuous with respect to the Lebesgue measure. $\mu_{\sigma,\tau}$ is given by an operator-valued semicircle law, whose Stieltjes transform is characterised by a fixed point equation in an appropriate Banach space. We analyse the moments of $\mu_{\sigma,\tau}$ and prove that the second moment is finite even when the weights have infinite variance. In the case $\sigma=1$, corresponding to the so-called scale-free percolation random graph, we can explicitly describe the limiting measure and study its tail.

§3.1 Introduction

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

$$\mathbb{P}\left(i \leftrightarrow j \mid W_i, W_j\right) = \kappa(W_i, W_j)\|i-j\|^{-\alpha} \wedge 1\,, \tag{3.1}$$

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

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

In the above $\sigma_{\text{pa}} = \alpha(\tau-1)/d - 1$, where $\tau - 1$ is the exponent of the tail distribution of the weights, such that the kernel κ_{pa} mimics the form that appears in preferential attachment models [Jorritsma et al., 2023], while the trivial kernel κ_{triv} corresponds to the classical long-range percolation model [Schulman, 1983, Newman and Schulman, 1986]. The kernel κ_{prod} yields a model which is substantially equivalent to scale-free percolation, introduced in Deijfen et al. [2013], which has connection probabilities of the form

$$1 - \exp\left(-W_i W_j \|i-j\|^{-\alpha}\right).$$

Various percolation properties for kernel-based spatial random graphs are known on $\mathbb{Z}^d$ and beyond (Deprez et al. [2015], Hao and Heydenreich [2023], van der Hofstad and Komjáthy [2017], Gracar et al. [2021], Jorritsma et al. [2024], see also Deprez and Wüthrich [2019], Dalmau and Salvi [2021] for a version of the same in the continuum) as well as the behaviour of interacting particle systems on them [Berger, 2002, Heydenreich et al., 2017, Komjáthy and Lodewijks, 2020, Cipriani and Salvi, 2024, Gracar and Grauer, 2024, Bansaye and Salvi, 2024, Komjáthy et al., 2023]. In contrast, their spectral properties, to the best of the authors' knowledge, have received less attention.

As a branch of random matrix theory, the study of the spectrum of random graphs has wide applications ranging from the study of random Schrödinger operators [Carmona and Lacroix, 2012, Geisinger, 2015] and quantum chaos in physics, to the analysis of community structures [Bordenave et al., 2015] and diffusion processes in network science, to the problems of spectral clustering [Champion et al., 2020] and graph embeddings [Gallagher et al., 2024] in data science. Many challenges remain unsolved in this area, even for the simplest models. As a prominent example, for bond percolation on $\mathbb{Z}^2$ it is known that the expected spectral measure has a continuous component if and only if $p > p_c$, but this result has not yet been established in higher dimensions [Bordenave et al., 2017]. In this chapter, we begin the study of spectral properties of spatial inhomogeneous random graphs, which in turn have been proposed as models for several real-world networks (see e.g. Dalmau and Salvi [2021]).

We will work with KBRGs in the typical setting where the weights (W_i) have support in $[1, \infty)$ and the kernel κ is an increasing function of the weights. Let us recall that in this case the vertices of KBRG random graphs on $\mathbb{Z}^d$ have almost surely infinite degree as soon as $\alpha < d$. Thus, as it happens in many percolation problems, the regime $\alpha > d$ would be the most appealing (and the toughest to tackle). In the present work we will focus instead on the dense case $\alpha < d$. We consider the discrete torus with N^d vertices equipped with the torus distance $\|\cdot\|$. The weights are sampled independently from a Pareto distribution with parameter $\tau - 1$ with $\tau > 2$. Conditionally on the weights, vertices i and j are connected independently from other pairs with probability given by (3.1) with a kernel of the form $\kappa_\sigma(w, v) := (w \vee v)(w \wedge v)^\sigma$. It is worth noting a difference between our connection probability and that studied recently in Jorritsma et al. [2023], van der Hofstad et al. [2023], where the connection probabilities are given by

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

The two forms can be made equivalent through a simple modification of the weights and an appropriate choice of α.

We call $\mathbb{G}_N$ the random graph obtained with this procedure and study the empirical spectral distribution of its adjacency matrix, appropriately scaled. Note that when $\alpha = 0$ we recover the (inhomogeneous) Erdős–Rényi random graph (modulo a tweak inserting a suitable tuning parameter ε_N) . In recent years, there has been significant research on inhomogeneous Erdős–Rényi random graphs, which can be equivalently modelled by Wigner matrices with a variance profile. The limiting spectral distribution of the adjacency matrix of such graphs has been studied in Chakrabarty et al. [2021b], Zhu [2020], Bose et al. [2022], while local eigenvalue statistics have been analysed in Dumitriu and Zhu [2019], Ajanki et al. [2019]. Zhu and Zhu [2024] studies the fluctuations of the linear eigenvalue statistics for a wide range of such inhomogeneous graphs. Additionally, various properties of the largest eigenvalue have been investigated in Cheliotis and Louvaris [2024], Husson [2022], Chakrabarty et al. [2022], Ducatez et al. [2024]. One of the most significant properties of the limiting spectral measure for random graphs is its absolute continuity with respect to the Lebesgue measure, which is closely tied to the concept of mean quantum percolation [Bordenave et al., 2017, Anantharaman et al., 2021, Arras and Bordenave, 2023]. Quantum percolation investigates whether the limiting measure has a non-trivial absolutely continuous spectrum. Recently, it was shown in Arras and Bordenave [2023] that the adjacency operator of a supercritical Poisson Galton-Watson tree has a non-trivial absolutely continuous part when the average degree is sufficiently large. Additionally, Bordenave et al. [2017] demonstrated that supercritical bond percolation on $\mathbb{Z}^d$ has a non-trivial absolutely continuous part for $d = 2$. These results motivate similar questions for KBRGs.

Our contributions: Results and proofs

Here below we showcase our main results and the novelties of our proofs Recall that we work in the regime $\alpha < d$ and $\tau > 2$. We also restrict to values of σ in $(0, \tau - 1)$.

(a) In Theorem 3.2.1 we show that, after scaling the adjacency matrix of $\mathbb{G}_N$ by $c_0 N^{(d-\alpha)/2}$, the empirical spectral distribution converges weakly in probability to a deterministic measure $\mu_{\sigma,\tau}$. The classical approach to proving the convergence of the empirical distribution is generally through either the method of moments or the Stieltjes transform. However, the limiting measure is expected to be heavy-tailed (see Figure 3.3) and so it is not determined by its moments. As a consequence, we cannot directly apply the method of moments. To overcome this issue, we pass through a truncation argument where we impose a maximal value to the weights, reducing the problem to well-behaved measures. To simplify the method of moments, we further reduce the model by substituting the adjacency matrix of $\mathbb{G}_N$ with a Gaussian matrix whose entries are centred and have roughly the same variance as before. This is made possible by a classical result of Chatterjee [2005]. Once we have shifted our attention to this simpler Gaussianised matrix with bounded weights, we can use the classical method of moments using finding its moments is made possible by a combinatorial argument on partitions and their graphical representation. Finally we remove the truncation effect.

(b) In Theorem 3.2.2 we investigate the graph corresponding to κ_{prod}, that is, when $\sigma = 1$. In this case we can explicitly identify $\mu_{1,\tau}$ as the free multiplicative convolution of the semicircle law and the measure of the weight distribution. In the $\sigma = 1$ case the moment expression derived in Theorem 3.2.1

simplifies, so the challenge is to recover the limiting measure from those moments. This is made possible thanks to the extension of the free multiplicative convolution to measures with unbounded support by Arizmendi and Pérez-Abreu [2009]. Furthermore, we show that $\mu_{1,\tau}$ has power-law tails with exponent $2(\tau-1)$. This is based on a Breiman-type argument for free multiplicative convolutions [Kołodziejek and Szpojankowski, 2022].

(c) In Theorem 3.2.3 we explicitly derive the second moment of $\mu_{\sigma,\tau}$ and prove that it is finite and non-degenerate. The proof is based on the ideas of Chakrabarty et al. [2016, Theorem 2.2]. This result is noteworthy because our weight distribution may exhibit infinite variance in the chosen range of parameters. To show that the second moment is finite, we need to establish the uniform integrability of a sequence of measures converging to the limiting measure. This is achieved through an extension of Skorohod's representation theorem for measures that converge weakly in probability.

(d) In Theorem 3.2.4 we prove that $\mu_{\sigma,\tau}$ is absolutely continuous. What makes the result possible is that we are able to split the original matrix as a free sum of a standard Wigner matrix and another Wigner matrix with a carefully chosen variance profile (yielding, as a by-product, another characterisation of the limit measure $\mu_{\sigma,\tau}$). Once this is established, the result is a consequence of Biane [1997].

(e) In Theorem 3.2.5 we provide an analytical description of $\mu_{\sigma,\tau}$ when $\tau > 3$ and $\sigma < \tau - 2$. Removing the truncation in the method of moments proof of Theorem 3.2.1 does not yield an explicit characterisation of the limiting measure. On the other hand, certain moment recursions for the truncated Gaussian matrix that appear in the proof can be used to derive properties of $\mu_{\sigma,\tau}$ through the Stieltjes transform. When the weights are bounded, the limiting measure corresponds to the operator-valued semicircle law (Speicher [2011]). Its transform can be expressed in terms of functions solving an analytic recursive equation (see Avena et al. [2023], Zhu [2020] for similar results in other random graph ensembles). In our case, when the weights are heavy-tailed, this is no longer possible. We achieve instead the convergence of the analytic recursive equation by constructing a suitable Banach space and demonstrating that it forms a contractive mapping.

Outline of the article.
In Section 3.2 we will define the model and state precisely the main results. In Section 3.3 we will give some auxiliary results which will be used to prove the main theorems in the rest of the article. More precisely, in Section 3.4 we will prove the existence of the limiting ESD, and in Section 3.5 we will give estimates

on its tail behaviour. In Section 3.6 we will prove the non-degeneracy of the limiting measure and in Section 3.7 we will show its absolute continuity. Finally, Section 3.8 is devoted to describing the Stieltjes transform of the limiting ESD.

§3.2 Set-up and main results

§3.2.1 Random graph models

To introduce our models, we use $a \wedge b$ to denote the minimum of two real numbers a and b, and $a \vee b$ to denote their maximum.

(a) **Vertex set:** the vertex set is $\mathbf{V}_N := \{1, 2, \ldots, N\}^d$. The vertex set is equipped with torus the distance $\|i - j\|$, where

$$\|i - j\| = \sum_{\ell=1}^{d} |i_\ell - j_\ell| \wedge (N - |i_\ell - j_\ell|).$$

(b) **Weights:** the weights $(W_i)_{i \in \mathbf{V}_N}$ are i.i.d. random variables sampled from a Pareto distribution W (whose law we denote by $\mathbf{P}$) with parameter $\tau - 1$, where $\tau > 1$. That is,

$$\mathbf{P}(W > t) = t^{-(\tau-1)} \mathbf{1}_{\{t \geq 1\}} + \mathbf{1}_{\{t < 1\}}. \tag{3.2}$$

(c) **Kernel:** the kernel function $\kappa_\sigma : [0, \infty) \times [0, \infty) \to [0, \infty)$ determines how the weights interact. In this article, we focus on kernel functions of the form

$$\kappa_\sigma(w, v) := (w \vee v)(w \wedge v)^\sigma, \tag{3.3}$$

where $\sigma \geq 0$.

(d) **Long-range parameter:** $\alpha > 0$ tunes the influence of the distance between vertices on their connection probability.

(e) **Connectivity function:** conditional on the weights, each pair of distinct vertices i and j is connected independently with probability $P^W(i \leftrightarrow j)$ given by

$$P^W(i \leftrightarrow j) := \mathbb{P}(i \leftrightarrow j \mid W_i, W_j) = \frac{\kappa_\sigma(W_i, W_j)}{\|i - j\|^\alpha} \wedge 1. \tag{3.4}$$

We will be using the short-hand notation $p_{ij} := \mathbb{P}(i \leftrightarrow j \mid W_i, W_j)$ for convenience. Note that the graph does not have self-loops (see Remark 3.4.1).

The associated graph is connected, as nearest neighbours with respect to the torus distance are always linked.

§3.2.2 Spectrum of a random graph

Let us denote the random graph generated by our choice of edge probabilities by $\mathbb{G}_N$. Let $\mathbb{A}_{\mathbb{G}_N}$ denote the adjacency matrix (operator) associated with this random graph, defined as

$$\mathbb{A}_{\mathbb{G}_N}(i,j) = \begin{cases} 1 & \text{if } i \leftrightarrow j, \\ 0 & \text{otherwise.} \end{cases}$$

Since the graph is finite, the adjacency matrix is always self-adjoint and has real eigenvalues. For $\alpha < d$, the eigenvalues require a scaling, which turns out to be independent of the kernel in our setup. Here we assume $\sigma \in (0, \tau - 1)$ and $\tau > 2$, ensuring that the vertex weights $(W_i)_{i \in \mathbf{V}_N}$ have finite mean. We define the scaling factor as

$$c_N = \frac{1}{N^d} \sum_{i \neq j \in \mathbf{V}_N} \frac{1}{\|i-j\|^\alpha} \sim c_0 N^{d-\alpha}, \tag{3.5}$$

where c_0 is a constant depending on α and d, and for two functions $f(\cdot)$ and $g(\cdot)$ we use $f(t) \sim g(t)$ to indicate that their quotient $f(t)/g(t)$ tends to one as t tends to infinity. The scaled adjacency matrix is then defined as

$$\mathbf{A}_N := \frac{\mathbb{A}_{\mathbb{G}_N}}{\sqrt{c_N}}. \tag{3.6}$$

The empirical measure that assigns a mass of $1/N^d$ to each eigenvalue of the $N^d \times N^d$ random matrix $\mathbf{A}_N$ is called the Empirical Spectral Distribution (ESD) of $\mathbf{A}_N$, denoted as

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

where $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_{N^d}$ are the eigenvalues of $\mathbf{A}_N$.

§3.2.3 Main results

We are now ready to state the main result of this article. Let μ_W denote the law of W. Here onwards, let $\mathbb{P} = \mathbf{P} \otimes P^W$ represent the joint law of the weights and the edge variables. Note that $\mathbb{P}$ depends on N, but we omit this dependence for simplicity. Let $\mathbb{E}, \mathbf{E}$, and E^W denote the expectation with respect to $\mathbb{P}, \mathbf{P}$, and P^W respectively. Furthermore, if $(\mu_N)_{N \geq 0}$ is a sequence of probability measures, we write $\lim_{N \to \infty} \mu_N = \mu_0$ to denote that μ_0 is the weak limit of the measures μ_N. Since the empirical spectral distribution is a random probability

Chapter Three

measure, we require the notion of convergence in probability in the context of weak convergence.

The Lévy-Prokhorov distance $d_L : \mathcal{P}(\mathbb{R})^2 \to [0, +\infty)$ between two probability measures μ and ν on $\mathbb{R}$ is defined as

$$d_L(\mu, \nu) := \inf \left\{ \varepsilon > 0 \mid \mu(A) \leq \nu(A^\varepsilon) + \varepsilon \text{ and } \nu(A) \leq \mu(A^\varepsilon) + \varepsilon \quad \forall A \in \mathcal{B}(\mathbb{R}) \right\},$$

where $\mathcal{B}(\mathbb{R})$ denotes the Borel σ-algebra on $\mathbb{R}$, and A^ε is the ε-neighbourhood of A. For a sequence of random probability measures $(\mu_N)_{N \geq 0}$, we say that

$$\lim_{N\to\infty} \mu_N = \mu_0 \text{ in } \mathbb{P}\text{-probability}$$

if, for every $\varepsilon > 0$,

$$\lim_{N\to\infty} \mathbb{P}(d_L(\mu_N, \mu_0) > \varepsilon) = 0.$$

The first result states the existence of the limiting spectral distribution of the scaled adjacency matrix.

Theorem 3.2.1 (Limiting spectral distribution).
Consider the random graph $\mathbb{G}_N$ on $\mathbf{V}_N$ with connection probabilities given by (3.4) *with parameters $\tau > 2$, $0 < \alpha < d$ and $\sigma \in (0, \tau - 1)$. Let $\mathrm{ESD}(\mathbf{A}_N)$ be the empirical spectral distribution of $\mathbf{A}_N$ defined in* (3.6)*. Then there exists a deterministic measure $\mu_{\sigma,\tau}$ on $\mathbb{R}$ such that*

$$\lim_{N\to\infty} \mathrm{ESD}(\mathbf{A}_N) = \mu_{\sigma,\tau} \qquad \textit{in } \mathbb{P}\textit{–probability}.$$

The remaining results focus of the properties of the limiting measure. First we note that when we set $\sigma = 1$ we can explicitly identify the limiting measure in terms of free multiplicative convolution. We refer the reader to Anderson et al. [2010, Section 5.2.3] for an exposition on free multiplicative and additive convolutions.

For two probability measures μ and ν the free multiplicative convolution $\mu \boxtimes \nu$ of the two measures is defined as the law of the product ab of free, random, non-commutative operators a and b, with laws μ and ν respectively. The free multiplicative convolution for two non-negatively supported measures was introduced in Bercovici and Voiculescu [1993]. Note that the semicircle law is not non-negatively supported and hence we use the extended definition of Arizmendi and Pérez-Abreu [2009] for the multiplicative convolution.

Theorem 3.2.2 (Limiting ESD for $\boldsymbol{\sigma = 1}$).
Consider the KBRG for $\sigma = 1$, while α, τ are as in the assumptions of Theorem 3.2.1. The the limiting spectral distribution $\mu_{1,\tau}$ is given by

$$\mu_{1,\tau} = \mu_{sc} \boxtimes \mu_W \,,$$

where μ_{sc} is the semicircle law

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

and $\boxtimes$ is the free multiplicative convolution of the two measures. Moreover, the limiting measure $\mu_{1,\tau}$ has a power-law tail, that is,

$$\mu_{1,\tau}(x,\infty) \sim \frac{1}{2}\big(m_1(\mu_W)\big)^{\tau-1}x^{-2(\tau-1)} \quad \text{as } x\to\infty,$$

where $m_1(\nu)$ denotes the first moment of the probability measure ν.

In the general case, it is hard to explicitly identify the limiting measure, so we present some characterisations of it. Since we do not impose that $\tau > 3$ and consequently the weights can have infinite variance, it is not immediate if the second moment of the limiting measure is non-degenerate and finite. We prove this in the following result.

Theorem 3.2.3 (Non-degeneracy of the limiting measure).
Under the assumptions of Theorem 3.2.1, the second moment of the limiting measure $\mu_{\sigma,\tau}$ is given by

$$\int_{\mathbb{R}} x^2\mu_{\sigma,\tau}(\mathrm{d}\, x) = (\tau-1)^2\int_1^\infty\int_1^\infty \frac{1}{(x\wedge y)^{\tau-\sigma}(x\vee y)^{\tau-1}}\,\mathrm{d}\, x\,\mathrm{d}\, y \in (0,\infty).$$

Moreover, for $p\in\mathbb{N}$ and $p<(\tau-1)/(\sigma\vee 1)$, we have $\int_{\mathbb{R}}|x|^{2p}\,\mu_{\sigma,\tau}(\mathrm{d}\ x)<\infty$.

We state the following result as an independent theorem as the absolute continuity of the KBRG model deserves to be treated separately.

Theorem 3.2.4 (Absolute continuity).
Let $\tau > 2$ and $\sigma\in(0,\tau-1)$, then $\mu_{\sigma,\tau}$ is symmetric and absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}$.

We conclude the main results by providing an analytic description of the limiting measure in terms of its Stieltjes transform when we slightly restrict our parameters. Recall that, for $z\in\mathbb{C}^+$, where $\mathbb{C}^+$ denotes the upper half-plane of the complex plane, the Stieltjes transform of a measure μ on $\mathbb{R}$ is given by

$$\mathrm{S}_\mu(z) = \int_{\mathbb{R}}\frac{1}{x-z}\mu(\mathrm{d}x)\,. \tag{3.7}$$

Theorem 3.2.5 (Stieltjes transform).
Let $0<\alpha<d$, $\tau>3$ and $\sigma<\tau-2$. Then there exists a unique analytic function a^* on $\mathbb{C}^+\times[1,\infty)$ such that

$$\mathrm{S}_{\mu_{\sigma,\tau}}(z) = \int_1^\infty a^*(z,x)\mu_W(\mathrm{d}x),$$

where we recall that μ_W is the law of the random variable W.

Chapter Three

The function a^* in the above theorem turns out to be a fixed point of a contraction mapping on an appropriate Banach space. The equation above shares similarities with the quadratic vector equations introduced and studied in Ajanki et al. [2019], although in our setting the measures have unbounded support. The properties and the proof of Theorem 3.2.5 are discussed in Section 3.8.

Remark 3.2.6 (Higher dimensions).
While we have presented our results for $0 < \alpha < d$, our proofs are worked out in the $d = 1$ setup. This is in order to avoid notational complications that would especially affect the clarity of Theorem 3.2.1. The limiting spectral distribution and its properties remain unchanged for $d > 1$.

§3.2.4 Examples, simulations and discussion

Firstly, in Figure 3.1 we plot the eigenvalue distribution of the adjacency matrix of two realisations of kernel-based graphs with different parameters, indicated at the top of the image. Secondly, in Figure 3.2 we sample 10 realisations of scale-

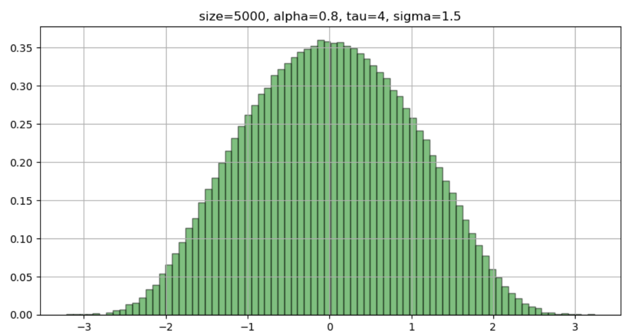

Figure 3.1: Eigenvalue distribution a KBRG realisation.

free percolation adjacency matrices of size 4000×4000 with $\sigma = 1$ and plot their eigenvalues (in green). We superpose on them the eigenvalues of the product $P_N G_N P_N$ of a GUE matrix G_N with a diagonal matrix P_N with i.i.d. entries distributed as $\sqrt{Pareto(\tau)}$ (in blue). Note that by Nica and Speicher [2006, Remark 14.2], Chakrabarty et al. [2021a, Remark 4.3], the a.s. limiting ESD of $P_N G_N P_N$ is $\mu_{sc} \boxtimes \mu_W$. All matrices are centred and rescaled by the sample second moment. Thirdly, to elucidate the tail behaviour of the limiting ESD when $\sigma = 1$ (Theorem 3.2.2) we draw in Figure 3.3 the empirical survival function of the eigenvalues of a matrix of size 7000×7000 in $x \geq 1.5$.

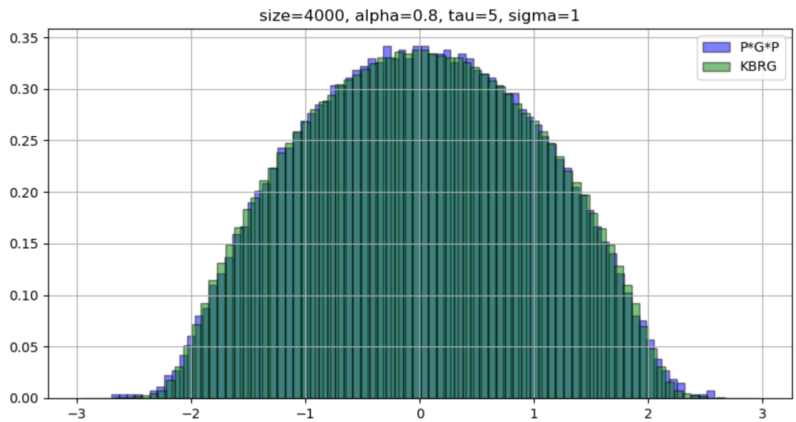

Figure 3.2: KBRG eigenvalue distribution and $P_N G_N P_N$ distribution.

Finally, we provide in Figure 3.4 a simulation of the eigenvalues of the Gaussian matrix $\tilde{\mathbf{A}}_{N,m,g}$ (see (3.24)) when $\alpha = 0$ and $N = 6000$. We compare this picture with the right-hand side of Figure 3.1, which has a small α. We conjecture that the atom appearing in the latter is due to high connectivity of the kernel-based realisation (if $\alpha = 0$, for all i, j we have that p_{ij} is identically one in (3.4)), whilst in the Gaussian setup this trivialization does not arise.

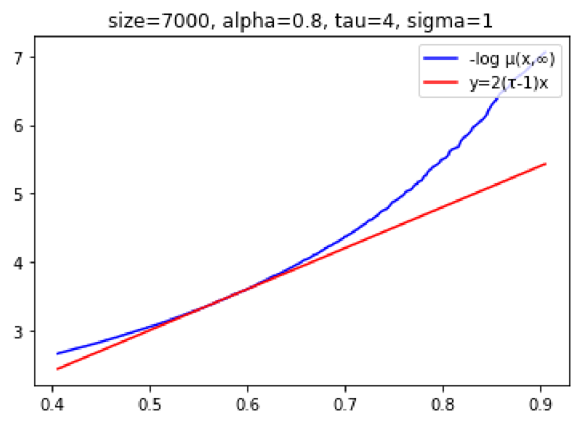

Figure 3.3: Negative of the log-empirical survival function and tails of Theorem 3.2.2 for $x \geq 1.5$.

Remark 3.2.7 (Sparse case).
We expect the case $\alpha > d$ to be very different due to the sparse nature of the graph. There has been a significant development in the area of spectral prop-

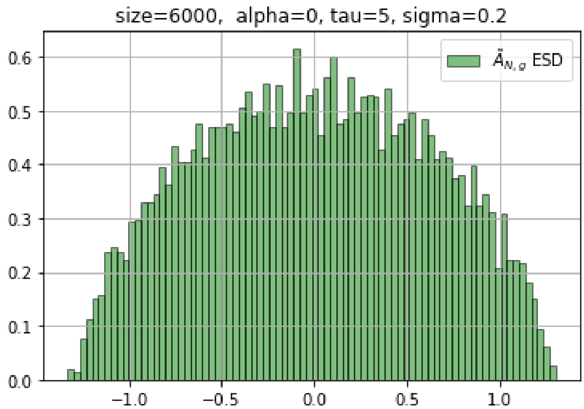

Figure 3.4: ESD for $\tilde{\mathbf{A}}_{N,m,g}$.

erties of sparse random graphs using the techniques of local weak convergence [Bordenave and Lelarge, 2010, Bordenave et al., 2017, 2011]. However, it is not immediately clear whether these techniques can be employed in our framework in order to determine the properties of the limiting measure: the underlying random graph generated in our model will not be tree-like to begin with. We plan to address this case in a future work.

§3.3 Notation and preliminary lemmas

In this section, we fix some notation and collect some technical lemmas that will be used in the proofs of our main results.

§3.3.1 Notation

We will use the Landau notation o_N, O_N indicating in the subscript the variable under which we take the asymptotic (typically this variable will grow to infinity unless otherwise specified). Universal positive constants are denoted as $c, c_1, \ldots,$ and their value may change with each occurrence. For an $N \times N$ matrix $A = (a_{ij})_{i,j=1}^N$ we use $\mathrm{Tr}(A) := \sum_{i=1}^N a_{ii}$ for the trace and $\mathrm{tr}(A) := N^{-1}\,\mathrm{Tr}(A)$ for the normalised trace. When $n \in \mathbb{N}$ we write $[n] := \{1, 2, \ldots, n\}$. We denote the cardinality of a set A as $\#A$, and, with a slight abuse of notation, $\#\sigma$ also denotes the number of cycles in a permutation σ.

§3.3.2 Technical lemmas

The following proposition, known as the *Hoffman-Wielandt inequality*, follows from Bai and Silverstein [2010, Corollary A.41].

Proposition 3.3.1 (Hoffman-Wielandt inequality).
Let $\mathbf{A}$ *and* $\mathbf{B}$ *be two* $N \times N$ *normal matrices and let* $\mathrm{ESD}(\mathbf{A})$ *and* $\mathrm{ESD}(\mathbf{B})$ *be their ESDs, respectively. Then,*

$$d_L\left(\mathrm{ESD}(\mathbf{A}), \mathrm{ESD}(\mathbf{B})\right)^3 \leq \frac{1}{N} \operatorname{Tr}\left[(\mathbf{A}-\mathbf{B})(\mathbf{A}-\mathbf{B})^*\right]. \tag{3.8}$$

Here $\mathbf{A}^*$ *denotes the conjugate transpose of* $\mathbf{A}$*. Moreover, if* $\mathbf{A}$ *and* $\mathbf{B}$ *are two Hermitian matrices of size* $N \times N$*, then*

$$\sum_{i=1}^{N} (\lambda_i(\mathbf{A}) - \lambda_i(\mathbf{B}))^2 \leq \operatorname{Tr}[(\mathbf{A}-\mathbf{B})^2]. \tag{3.9}$$

The next two straightforward lemmas control the tail of the product of two Pareto random variables and the expectation of a truncated Pareto.

Lemma 3.3.2.
Let X *and* Y *be two independent Pareto r.v.'s with parameters* β_1 *and* β_2 *respectively, with* $\beta_1 \leq \beta_2$*. There exist constants* $c_1 = c_1(\beta_1, \beta_2) > 0$ *and* $c_2 = c_2(\beta_1) > 0$ *such that*

$$\mathbf{P}(XY > t) = \begin{cases} c_1 t^{-\beta_1} & \text{if } \beta_1 < \beta_2 \\ c_2 t^{-\beta_1} \log t & \text{if } \beta_1 = \beta_2. \end{cases}$$

Lemma 3.3.3.
Let X *be a Pareto random variable with law* $\mathbf{P}$ *and parameter* $\beta > 1$*. For any* $m > 0$ *it holds*

$$\mathbf{E}\left[X \mathbb{1}_{X \geq m}\right] = \frac{\beta}{(\beta - 1)} m^{1-\beta}.$$

We state one final auxiliary lemma related to the approximation of sums by integrals.

Lemma 3.3.4.
Let $\beta \in (0, 1]$*. Then there exists a constant* $c_1 = c_1(\beta) > 0$ *such that*

$$\frac{1}{N} \sum_{i \neq j \in \mathbf{V}_N} \frac{1}{\|i-j\|^\beta} \sim c_1 \max\{N^{1-\beta}, \log N\}. \tag{3.10}$$

If instead $\beta > 1$*, there exists a constant* $c_2 > 0$ *such that*

$$\frac{1}{N} \sum_{i \neq j \in \mathbf{V}_N} \frac{1}{\|i-j\|^\beta} \sim c_2\,.$$

We end this section by quoting, for the reader's convenience, the following lemma from Chakrabarty et al. [2016, Fact 4.3].

Lemma 3.3.5.
Let (Σ, d) be a complete metric space, and let $(\Omega, \mathcal{A}, P)$ be a probability space. Suppose that $\big(X_{mn} : (m,n) \in \{1,2,\dots,\infty\}^2 \backslash \{\infty,\infty\}\big)$ is a family of random elements in Σ, that is, measurable maps from Ω to Σ, the latter being equipped with the Borel σ-field induced by d. Assume that

(1) for all fixed $1 \le m < \infty$

$$\lim_{n\to\infty} d\left(X_{mn}, X_{m\infty}\right) = 0 \text{ in } P\text{-probability.}$$

(2) For all $\varepsilon > 0$,

$$\lim_{m\to\infty} \limsup_{n\to\infty} P\left(d\left(X_{mn}, X_{\infty n}\right) > \varepsilon\right) = 0.$$

Then, there exists a random element $X_{\infty\infty}$ of Σ such that

$$\lim_{m\to\infty} d\left(X_{m\infty}, X_{\infty\infty}\right) = 0 \text{ in } P\text{-probability} \tag{3.11}$$

and

$$\lim_{n\to\infty} d\left(X_{\infty n}, X_{\infty\infty}\right) = 0 \text{ in } P\text{-probability.}$$

Furthermore, if $X_{m\infty}$ is deterministic for all m, then so is $X_{\infty\infty}$, and (3.11) simplifies to

$$\lim_{m\to\infty} d\left(X_{m\infty}, X_{\infty\infty}\right) = 0. \tag{3.12}$$

§3.4 Existence and Uniqueness

The proof of Theorem 3.2.1 is split into several parts and we will now briefly sketch them.

(1) **Truncation**: The first part of the proof is a truncation argument on the unbounded weights $(W_i)_{i\in\mathbf{V}_N}$. We construct a new sequence $(W_i^m)_{i\in\mathbf{V}_N}$ that is obtained by truncating the original weights at a value $m > 1$. We construct another scaled adjacency matrix $\mathbf{A}_{N,m}$, with entries $\mathbf{A}_{N,m}(i,j)$ distributed as Bernoulli random variables with parameter p_{ij}^m given by (3.4) with the weights substituted by the truncated ones. We then show (see Lemma 3.4.2) that the empirical measure $\mathrm{ESD}(\mathbf{A}_N)$ is well approximated by $\mathrm{ESD}(\mathbf{A}_{N,m})$, that is, their Lévy distance vanishes in probability in the limit $m \to \infty$.

(2) **Gaussianisation**: In the second part, we aim to Gaussianise $\mathbf{A}_{N,m}$ using the ideas of Chatterjee [2005]. We begin with the construction of a centred matrix $\overline{\mathbf{A}}_{N,m}$, that is obtained by subtracting out the expectation from each entry of $\mathbf{A}_{N,m}$. We then Gaussianise $\overline{\mathbf{A}}_{N,m}$, that is, we pass to another matrix $\mathbf{A}_{N,g}$ with each entry $\mathbf{A}_{N,g}(i,j)$ being a normal random variable with mean 0 and the same variance $p_{ij}^m(1-p_{ij}^m)$ as the corresponding entry of $\overline{\mathbf{A}}_{N,m}$. Lastly, we tweak the variances of $\mathbf{A}_{N,g}$ to obtain a Gaussian random matrix $\tilde{\mathbf{A}}_{N,m,g}$ with entries $\tilde{\mathbf{A}}_{N,m,g}(i,j)$ having mean 0 and variance equal to r_{ij}^m, the "unbounded version" of p_{ij}^m (see (3.13)). Thanks to (3.8), we can show (Lemma 3.4.3, Lemma 3.4.4 and Lemma 3.4.6) that in this whole process we did not lose too much: the Lévy distance between the empirical measures $\mathrm{ESD}(\mathbf{A}_{N,m})$ and $\mathrm{ESD}(\tilde{\mathbf{A}}_{N,m,g})$ is small in probability. We remark here that the order of the errors in Lemmas 3.4.3 and 3.4.6 is $N^{-\alpha}$, and these steps fail for $\alpha = 0$.

(3) **Identification of the limit**: We then proceed to analyse the limit of the measure $\mathrm{ESD}(\tilde{\mathbf{A}}_{N,m,g})$ as N goes to infinity. We use Wick's formula to compute its expected moments and use a concentration argument to show the existence of a unique limiting measure

$$\mu_{\sigma,\tau,m} := \lim_{N\to\infty} \mathrm{ESD}(\tilde{\mathbf{A}}_{N,m,g})$$

using Proposition 3.4.9. We conclude the proof of Theorem 3.2.1 by letting the truncation m go to infinity: using Lemma 3.3.5 we can show that there is a unique limiting measure $\mu_{\sigma,\tau}$ such that $\mu_{\sigma,\tau} := \lim_{m\to\infty} \mu_{\sigma,\tau,m}$. In the case $\sigma = 1$ calculations become explicit.

Remark 3.4.1 (Self-loops).
We can use Proposition 3.3.1 to show that having self-loops in the model will not affect the limiting spectral distribution. Let $\mathbf{A}_N$ be the scaled adjacency matrix of the model as defined in (3.6)*. Now, consider*

$$D_N = c_N^{-1/2}\,\mathrm{Diag}(1,\ldots,1)$$

to be the $N\times N$ diagonal matrix with all diagonal entries "1", scaled by a factor of $\sqrt{c_N}$, and $\mathbf{A}_{N,SL} = \mathbf{A}_N + D_N$. If we extend the definition of p_{ij} for the case $i=j$ as $p_{ii}=1$, then $\mathbf{A}_{N,SL}$ will be the scaled adjacency of the random graph with self-loops. Using (3.8)*, we get*

$$d_L^3(\mu_{\mathbf{A}_N},\mu_{\mathbf{A}_{N,SL}}) \le \frac{1}{N}\,\mathrm{Tr}[(\mathbf{A}_N-\mathbf{A}_{N,SL})^2] = \frac{1}{N}\,\mathrm{Tr}[D_N^2] = \frac{N}{Nc_N} = \mathrm{O}(c_N^{-1}).$$

§3.4.1 Truncation

Now we show that for our analysis the weights can be truncated. More precisely, let $m > 1$ be a truncation threshold and define $W_i^m = W_i \mathbf{1}_{W_i \le m}$ for any $i \in \mathbf{V}_N$. For all $N \in \mathbb{N}$, we define a new random graph with vertex set $\mathbf{V}_N$ and connection probability as follows: conditional on the weights $(W_i^m)_{i\in\mathbf{V}_N}$ we connect $i, j \in \mathbf{V}_N$ with probability

$$p_{ij}^m = r_{ij}^m \wedge 1 \qquad \text{with} \qquad r_{ij}^m = \frac{(W_i^m \vee W_j^m)(W_i^m \wedge W_j^m)^\sigma}{\|i-j\|^\alpha} \qquad i \neq j \in \mathbf{V}_N\,. \tag{3.13}$$

Let $\mathbf{A}_{N,m}$ be the corresponding adjacency matrix scaled by $\sqrt{c_N}$ and let its ESD be denotes by $\mathrm{ESD}(\mathbf{A}_{N,m})$.

It will be useful later to have the two following easy bounds (following from Lemma 3.3.4):

$$\sum_{i\neq j\in\mathbf{V}_N} r_{ij}^m \le m^{1+\sigma} N c_N\,, \qquad\qquad \sum_{i\neq j\in\mathbf{V}_N} (r_{ij}^m)^t \le c\, m^{2+2\sigma} \max\{N^{1-t\alpha}, \log N\}\,, \tag{3.14}$$

for some constant $c > 0$ and $t > 1$ a real number. The second bound is not optimal, since for some $t > 1$ such that $t\alpha > 1$, the upper bound will just be a constant depending on t and α. However, for our computations, this bound suffices.

Lemma 3.4.2 (Truncation).
For every $\delta > 0$ one has

$$\limsup_{m\to\infty} \lim_{N\to\infty} \mathbb{P}\left(d_L(\mathrm{ESD}(\mathbf{A}_N), \mathrm{ESD}(\mathbf{A}_{N,m})) > \delta\right) = 0\,.$$

Proof. By (3.8) we have that

$$\begin{aligned}
&\mathbb{E}\left[d_L^3\left(\mathrm{ESD}(\mathbf{A}_N),\, \mathrm{ESD}(\mathbf{A}_{N,m})\right)\right] \\
&\le \frac{1}{Nc_N}\mathbb{E}\left[\mathrm{Tr}\left((\mathbf{A}_N - \mathbf{A}_{N,m})^2\right)\right] \\
&= \frac{1}{Nc_N}\sum_{i\neq j\in\mathbf{V}_N} \mathbb{E}\left[(\mathbf{A}_N(i,j) - \mathbf{A}_{N,m}(i,j))^2 \mathbf{1}_{\mathbf{A}_N(i,j)\neq \mathbf{A}_{N,m}(i,j)}\right] \\
&\le \frac{1}{Nc_N}\sum_{i\neq j\in\mathbf{V}_N} \mathbb{P}\left(\mathbf{A}_N(i,j) \neq \mathbf{A}_{N,m}(i,j)\right).
\end{aligned} \tag{3.15}$$

For fixed i, j we will analyse $\mathbb{P}\left(\mathbf{A}_N(i,j) \neq \mathbf{A}_{N,m}(i,j)\right)$ as follows. We notice that $\mathbf{A}_N(i,j) \neq \mathbf{A}_{N,m}(i,j)$ can occur only if one between W_i and W_j exceeds m. Calling

$$A = \{W_i \ge m > W_j\} \quad \text{and} \quad B = \{W_i \ge W_j \ge m\} \tag{3.16}$$

we have, by symmetry of W_i and W_j, that $\mathbb{P}\left(\mathbf{A}_N(i,j) \neq \mathbf{A}_{N,m}(i,j)\right)$ equals

$$2\mathbb{P}\left(\{\mathbf{A}_N(i,j) \neq \mathbf{A}_{N,m}(i,j)\} \cap A\right) + 2\mathbb{P}\left(\{\mathbf{A}_N(i,j) \neq \mathbf{A}_{N,m}(i,j)\} \cap B\right).$$

Notice that on the events A and B the variable $\mathbf{A}_{N,m}(i,j)$ is always 0. So we can bound

$$\begin{aligned}
&\mathbb{P}\left(\{\mathbf{A}_N(i,j) \neq \mathbf{A}_{N,m}(i,j)\} \cap A\right) \\
&= \mathbb{P}\left(\{\mathbf{A}_N(i,j) = 1\} \cap A\right) \\
&\leq \mathbf{E}\Big[\frac{\kappa_\sigma(W_i, W_j)}{\|i-j\|^\alpha}\mathbf{1}_A\Big] \leq \frac{\mathbf{E}[W_i \mathbf{1}_{W_i \geq m}]\mathbf{E}[W_j^\sigma]}{\|i-j\|^\alpha} \leq c\frac{m^{2-\tau}}{\|i-j\|^\alpha}
\end{aligned}$$

for some constant $c > 0$, where we have used Lemma 3.3.3 and the fact that $\mathbb{E}[W_j^\sigma] < \infty$. Analogously we can bound the second summand by

$$\begin{aligned}
&\mathbb{P}\left(\{\mathbf{A}_N(i,j) \neq \mathbf{A}_{N,m}(i,j)\} \cap B\right) \\
&\leq \mathbf{E}\left[\frac{W_i W_j^\sigma}{\|i-j\|^\alpha}\mathbf{1}_B\right] \leq \frac{\mathbf{E}[W_i \mathbf{1}_{W_i \geq m}]\mathbf{E}[W_j^\sigma]}{\|i-j\|^\alpha} \\
&\leq c\frac{m^{2-\tau}}{\|i-j\|^\alpha}.
\end{aligned}$$

Plugging these estimates back into (3.15) we obtain

$$\mathbb{E}\left[d_L^3\left(\mathrm{ESD}(\mathbf{A}_N), \mathrm{ESD}(\mathbf{A}_{N,m})\right)\right] \leq \frac{4c}{N c_N}\sum_{i \neq j \in \mathbf{V}_N}\frac{m^{2-\tau}}{\|i-j\|^\alpha} = 4cm^{2-\tau}.$$

We can then conclude by applying Markov's inequality:

$$\begin{aligned}
&\limsup_{m\to\infty}\lim_{N\to\infty}\mathbb{P}\left(d_L\left(\mathrm{ESD}(\mathbf{A}_N), \mathrm{ESD}(\mathbf{A}_{N,m})\right) > \delta\right) \\
&\leq \limsup_{m\to\infty}\lim_{N\to\infty}\frac{\mathbb{E}\left[d_L^3\left(\mathrm{ESD}(\mathbf{A}_N), \mathrm{ESD}(\mathbf{A}_{N,m})\right)\right]}{\delta^3} \\
&= 0
\end{aligned}$$

since $\tau > 2$. □

§3.4.2 Centring

Let $1 < m \leq \infty$ and $\overline{\mathbf{A}}_{N,m}$ be the centred and rescaled truncated adjacency matrix, i.e. the matrix defined as

$$\overline{\mathbf{A}}_{N,m}(i,j) = \mathbf{A}_{N,m}(i,j) - E^W[\mathbf{A}_{N,m}(i,j)], \quad i \neq j \in \mathbf{V}_N. \tag{3.17}$$

Note that here $m = \infty$ corresponds to the matrix with non-truncated weights. The following lemma says that the centring does not affect the limiting spectral distribution.

Chapter Three

Lemma 3.4.3 (Centring).
For any $m \in (1, \infty]$, under the conditions in Theorem 3.2.1, we have, for all $\delta > 0$,

$$\lim_{N\to\infty} \mathbb{P}\left(d_L\left(\mathrm{ESD}(\mathbf{A}_{N,m}), \mathrm{ESD}(\overline{\mathbf{A}}_{N,m})\right) > \delta\right) = 0\,,$$

where $\mathrm{ESD}(\overline{\mathbf{A}}_{N,m})$ is the empirical spectral distribution of $\overline{\mathbf{A}}_{N,m}$.

Proof. By (3.8) we have

$$\begin{aligned}
\mathbb{E}\left[d_L^3\left(\mathrm{ESD}(\mathbf{A}_{N,m}), \mathrm{ESD}(\overline{\mathbf{A}}_{N,m})\right)\right] &\le \frac{1}{N}\mathbb{E}\left[\mathrm{Tr}(E^W[\mathbf{A}_{N,m}]^2)\right] \\
&= \frac{1}{Nc_N}\sum_{i\neq j\in\mathbf{V}_N} \mathbf{E}[p_{ij}^m]^2 \\
&\le \frac{1}{Nc_N}\sum_{i\neq j\in\mathbf{V}_N} \frac{\mathbf{E}\left[(W_i\vee W_j)(W_i\wedge W_j)^\sigma\right]^2}{\|i-j\|^{2\alpha}} \\
&\le \frac{c}{Nc_N}\max\{N^{1-2\alpha}, \log N\}. \qquad (3.18)
\end{aligned}$$

Here c is some constant as for $\tau > 2$ and $\sigma < \tau - 1$ we have

$$\mathbf{E}\left[(W_i\vee W_j)(W_i\wedge W_j)^\sigma\right] = 2\mathbf{E}\left[W_i W_j^\sigma \mathbf{1}_{W_i > W_j}\right] \le 2\mathbf{E}[W_i]\mathbf{E}[W_j^\sigma] < \infty.$$

In the last inequality we used Lemma 3.3.4. The result follows by applying Markov's inequality. $\square$

§3.4.3 Gaussianisation

Let $\{G_{i,j}, 1 \le i \le j\}$ be a family of i.i.d. standard Gaussian random variables, independent of the weights and the graph. Define a symmetric $N \times N$ matrix $\mathbf{A}_{N,m,g}$ by

$$\mathbf{A}_{N,m,g}(i,j) = \begin{cases} \frac{\sqrt{p_{ij}^m(1-p_{ij}^m)}}{\sqrt{c_N}} G_{i\wedge j, i\vee j} & \text{for } 1 \le i \neq j \le N \\ 0 & \text{for } i = j. \end{cases} \qquad (3.19)$$

Notice that the entries of $\mathbf{A}_{N,m,g}$ have the same mean and variance of the corresponding entries of $\overline{\mathbf{A}}_{N,m}$. Consider a three-times continuously differentiable function $h : \mathbb{R} \to \mathbb{R}$ such that

$$\max_{0\le k\le 3}\sup_{x\in\mathbb{R}}\left|h^{(k)}(x)\right| < \infty$$

where $h^{(k)}$ denotes the k-th derivative. For an $N \times N$ real symmetric matrix $\mathbf{M}_N$ define the *resolvent* of $\mathbf{M}_N$ as

$$R_{M_N}(z) = (\mathbf{M}_N - z\,\mathrm{I}_N)^{-1}, \qquad z \in \mathbb{C}^+,$$

where I_N is the $N \times N$ identity matrix. In particular, if $\mu := \mu_{\mathbf{M}_N}$ is the ESD of $\mathbf{M}_N$, the relation between the Stieltjes transform $\mathrm{S}_{\mathbf{M}_N}$ of $\mu_{\mathbf{M}_N}$ and resolvent can be expressed as

$$H(\mathbf{M}_N) := \mathrm{S}_{\mathbf{M}_N}(z) = \operatorname{tr}(R_{M_N}(z)),\ z \in \mathbb{C}^+ \tag{3.20}$$

[Bai and Silverstein, 2010, Section 1.3.2]. The next result shows that the real and imaginary parts of the Stieltjes transform of $\mu_{\overline{\mathbf{A}}_{N,m}}$ are close to those of $\mu_{\mathbf{A}_{N,m,g}}$. Since one knows that the convergence of the ESD is equivalent to showing the convergence of the corresponding Stieltjes transform, one can shift the problem to the Gaussianised setup and work with the matrix $\mathbf{A}_{N,m,g}$.

Lemma 3.4.4 (Gaussianisation).
Consider the matrix $\overline{\mathbf{A}}_{N,m}$ defined in Subsection 3.4.1 and the matrix $\mathbf{A}_{N,m,g}$ defined in (3.19)*. For any three-times continuously differentiable function $h : \mathbb{R} \to \mathbb{R}$ such that*

$$\max_{0\le k\le 3} \sup_{x\in\mathbb{R}} \left| h^{(k)}(x) \right| < \infty$$

we have

$$\lim_{N\to\infty} \left| \mathbb{E}\left[h\left(\Re H\left(\mathbf{A}_{N,m,g}\right)\right)\right] - \mathbb{E}\left[h\left(\Re H\left(\overline{\mathbf{A}}_{N,m}\right)\right)\right] \right| = 0,$$
$$\lim_{N\to\infty} \left| \mathbb{E}\left[h\left(\Im H\left(\mathbf{A}_{N,m,g}\right)\right)\right] - \mathbb{E}\left[h\left(\Im H\left(\overline{\mathbf{A}}_{N,m}\right)\right)\right] \right| = 0\,,$$

where $\Re$ and $\Im$ denote the real and imaginary parts respectively and $h^{(k)}$ denotes the k-th derivative of h.

To prove the above lemma, we will need the following result from Chatterjee [2005].

Theorem 3.4.5 (Chatterjee [2005, Theorem 1.1]).
Let $\mathbf{X} = (X_1, \dots, X_n)$ and $\mathbf{Y} = (Y_1, \dots, Y_n)$ be two vectors of independent random variables with finite second moments, taking values in some open interval I and satisfying, for each $i, \mathbb{E}X_i = \mathbb{E}Y_i$ and $\mathbb{E}X_i^2 = \mathbb{E}Y_i^2$. Let $f : I^n \to \mathbb{R}$ be three-times differentiable in each argument. If we set $U = f(\mathbf{X})$ and $V = f(\mathbf{Y})$, then for any thrice differentiable $h : \mathbb{R} \to \mathbb{R}$ and any $K > 0$,

$$|\mathbb{E}h(U) - \mathbb{E}h(V)| \le C_1(h)\lambda_2(f) \sum_{i=1}^{n} \left[\mathbb{E}\left[X_i^2 \mathbf{1}_{|X_i|>K}\right] + \mathbb{E}\left[Y_i^2 \mathbf{1}_{|Y_i|>K}\right]\right]$$
$$+ C_2(h)\lambda_3(f) \sum_{i=1}^{n} \left[\mathbb{E}\left[|X_i|^3 \mathbf{1}_{|X_i|\le K}\right] + \mathbb{E}\left[|Y_i|^3 \mathbf{1}_{|Y_i|\le K}\right]\right]$$

where $C_1(h) = \|h'\|_\infty + \|h''\|_\infty, C_2(h) = \frac{1}{6}\|h'\|_\infty + \frac{1}{2}\|h''\|_\infty + \frac{1}{6}\|h'''\|_\infty$ and

$$\lambda_s(f) := \sup\left\{|\partial_i^q f(x)|^{\frac{s}{q}} : 1 \le i \le n, 1 \le q \le s, x \in I^n\right\},$$

where ∂_i^q denotes q-fold differentiation with respect to the i-th coordinate.

Proof of Lemma 3.4.4. We prove this for the real part of the Stieltjes transform. The bounds for the imaginary part remain the same. We fix a complex number $z \in \mathbb{C}^+$, given by $z = \Re(z) + i\eta$ with $\eta > 0$.

Let $n = N(N-1)/2$ and $\mathbf{x} = (x_{ij})_{1\le i<j\le N} \in \mathbb{R}^n$. Define $\mathrm{R}(\mathbf{x})$ to be the matrix-valued differentiable function given by

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

where $\mathbf{M}_N(\cdot)$ is the matrix-valued differentiable function that maps a vector in $\mathbb{R}^n$ to the space of $N \times N$ Hermitian matrices, given by

$$\mathbf{M}_N(\mathbf{x})_{ij} = \begin{cases} c_N^{-1/2} x_{ij} & \text{if } i < j, \\ c_N^{-1/2} x_{ji} & \text{if } i > j, \\ 0 & \text{if } i = j. \end{cases}$$

Since $\mathbf{M}_N$ is symmetric, it has all real eigenvalues. The function $H(\mathbf{M}_N(\mathbf{x}))$ admits partial derivatives of all orders. In particular, we denote for any $\mathbf{u} \in \{(i,j)\}_{1\le j<i\le n}$ the partial derivative as $\partial H/\partial x_{\mathbf{u}}$. For any $\mathbf{u} \in \{(i,j)\}_{1\le j<i\le n}$, using the identity $(\mathbf{M}_N(\mathbf{x}) - z\,\mathrm{I})\mathrm{R}(\mathbf{x}) = \mathrm{I}_N$ we have

$$\frac{\partial \mathrm{R}(\mathbf{x})}{\partial x_{\mathbf{u}}} = -\mathrm{R}(\mathbf{x})(\partial_{\mathbf{u}}\mathbf{M}_N)\mathrm{R}(\mathbf{x}).$$

By iterative application of derivatives, three identities were derived in Chatterjee [2005]:

$$\begin{aligned}
\frac{\partial H}{\partial x_{\mathbf{u}}} &= -\frac{1}{N}\,\mathrm{Tr}\left(\frac{\partial \mathbf{M}_N(\mathbf{x})}{\partial x_{\mathbf{u}}}\mathrm{R}(\mathbf{x})^2\right), \\
\frac{\partial^2 H}{\partial x_{\mathbf{u}}^2} &= \frac{2}{N}\,\mathrm{Tr}\left(\frac{\partial \mathbf{M}_N(\mathbf{x})}{\partial x_{\mathbf{u}}}\mathrm{R}(\mathbf{x})\frac{\partial \mathbf{M}_N(\mathbf{x})}{\partial x_{\mathbf{u}}}\mathrm{R}(\mathbf{x})^2\right), \\
\frac{\partial^3 H}{\partial x_{\mathbf{u}}^3} &= -\frac{6}{N}\,\mathrm{Tr}\left(\frac{\partial \mathbf{M}_N(\mathbf{x})}{\partial x_{\mathbf{u}}}\mathrm{R}(\mathbf{x})\frac{\partial \mathbf{M}_N(\mathbf{x})}{\partial x_{\mathbf{u}}}\mathrm{R}(\mathbf{x})\frac{\partial \mathbf{M}_N(\mathbf{x})}{\partial x_{\mathbf{u}}}\mathrm{R}(\mathbf{x})^2\right).
\end{aligned}$$

Note that $\partial_{ij}\mathbf{M}_N(\mathbf{x})$ is a matrix with $c_N^{-1/2}$ at the $(i,j)^{\text{th}}$ and $(j,i)^{\text{th}}$ entry, and 0 everywhere else. Using the bounds on Hilbert-Schmidt norms and following the exact argument regarding the bounds in equations (4), (5) and (6) in Chatterjee [2005] we get that

$$\left\|\frac{\partial H}{\partial x_{\mathbf{u}}}\right\|_\infty \le \frac{2}{\eta N\sqrt{c_N}}, \quad \left\|\frac{\partial^2 H}{\partial x_{\mathbf{u}}^2}\right\|_\infty \le \frac{4}{\eta^3 N c_N}, \quad \left\|\frac{\partial^3 H}{\partial x_{\mathbf{u}}^3}\right\|_\infty \le \frac{12}{\eta^4 N c_N^{3/2}}.$$

Hence

$$\lambda_2(H) \leq 4 \max\left\{\frac{1}{\eta^4}, \frac{1}{\eta^3}\right\} \frac{1}{Nc_N}$$

and

$$\lambda_3(H) \leq 12 \max\left\{\frac{1}{\eta^6}, \frac{1}{\eta^{9/2}}, \frac{1}{\eta^4}\right\} \frac{1}{Nc_N^{3/2}}.$$

Conditional on the weights $(W_i)_{i\geq 1}$, consider the following sequence of independent random variables. Let $\mathbf{X}_b = (X^b_{ij})_{1\leq i<j\leq N}$ be a vector with $X^b_{ij} \sim \mathrm{Ber}(p^m_{ij}) - p^m_{ij}$. Similarly, take another vector $\mathbf{X}_g = (X^g_{ij})_{1\leq i<j\leq N}$ with $X^g_{ij} \sim \mathcal{N}\left(0, p^m_{ij}(1-p^m_{ij})\right)$. Then,

$$\overline{\mathbf{A}}_{N,m} = \mathbf{M}_N(\mathbf{X}_b) \ \text{ and } \ \mathbf{A}_{N,g} = \mathbf{M}_N(\mathbf{X}_g)$$

in law. We have that

$$\begin{aligned}
&\left|\mathbb{E}\left[h\left(\Re H_z\left(\mathbf{A}_{N,m,g}\right)\right) - h\left(\Re H_z\left(\overline{\mathbf{A}}_{N,m}\right)\right)\right]\right| \\
&= \left|\mathbf{E}\left[E^W\left[h\left(\Re H_z\left(\mathbf{A}_{N,m,g}\right)\right) - h\left(\Re H_z\left(\overline{\mathbf{A}}_{N,m}\right)\right)\right]\right]\right|.
\end{aligned}$$

Conditionally on the weights, the sequences $\mathbf{X}_g$ and $\mathbf{X}_b$ form two vectors of independent random variables, with $E^W[X^b_{ij}] = E^W[X^g_{ij}]$ and $E^W[(X^b_{ij})^2] = E^W[(X^g_{ij})^2]$. Then, using Theorem 3.4.5 on the conditional expectation

$$E^W[h\left(\Re H_z\left(\mathbf{A}_{N,m,g}\right)\right) - h\left(\Re H_z\left(\overline{\mathbf{A}}_{N,m}\right)\right)],$$

we have that

$$\begin{aligned}
&\left|\mathbf{E}\left[E^W[h\left(\Re H_z\left(\mathbf{A}_{N,m,g}\right)\right) - h\left(\Re H_z\left(\overline{\mathbf{A}}_{N,m}\right)\right)]\right]\right| \\
&\leq C_1(h)\lambda_2(H) \sum_{1\leq i<j\leq N} \mathbb{E}[(X^b_{ij})^2 \mathbf{1}_{|X^b_{ij}|>K_N}] + \mathbb{E}[(X^g_{ij})^2 \mathbf{1}_{|X^g_{ij}|>K_N}] \qquad (3.21)\\
&+ C_2(h)\lambda_3(H) \sum_{1\leq i<j\leq N} \mathbb{E}[(X^b_{ij})^3 \mathbf{1}_{|X^b_{ij}|\leq K_N}] + \mathbb{E}[(X^g_{ij})^3 \mathbf{1}_{|X^g_{ij}|\leq K_N}], \qquad (3.22)
\end{aligned}$$

where K_N is a (possibly) $N-$dependent truncation and where we have used that $|\partial^p_{\mathbf{u}}\Re H| = |\Re\partial^p_{\mathbf{u}} H| \leq |\partial^p_{\mathbf{u}} H|$. Now using the fact that $r/p > 0$ we have $|\partial^p_{\mathbf{u}}\Re H|^{\frac{r}{p}} \leq |\partial^p_{\mathbf{u}} H|^{\frac{r}{p}}$, and therefore

$$\lambda_r(\Re H) \leq \lambda_r(H).$$

We begin by evaluating (3.21). To compute the Bernoulli term, notice that X^b_{ij} are uniformly bounded by 1, so, for any $K_N > 1$, we automatically have that

$$\sum_{1\leq i<j\leq N} \mathbb{E}[(X^b_{ij})^2 \mathbf{1}_{|X^b_{ij}|>K_N}] = 0.$$

For the Gaussian term, we apply the Cauchy-Schwarz inequality (with respect to $\mathbb{E}$). Using also the trivial bound $p_{ij}^m \le r_{ij}^m$ and Markov's inequality, we obtain

$$\begin{aligned}\sum_{1\le i<j\le N} \mathbb{E}[(X_{ij}^g)^2 \mathbf{1}_{|X_{ij}^g|>K_N}] &\le \sum_{1\le i<j\le N} \mathbb{E}[(X_{ij}^g)^4]^{1/2}\mathbb{P}(|X_{ij}^g|>K_N)^{1/2} \\ &\le 3 \sum_{1\le i<j\le N} \mathbf{E}[(r_{ij}^m)^2]^{1/2} \frac{\mathbb{E}[(X_{ij}^g)^2]^{1/2}}{K_N} \le 3 \sum_{1\le i<j\le N} \mathbf{E}[(r_{ij}^m)^2]^{1/2} \frac{\mathbf{E}[r_{ij}^m]^{1/2}}{K_N} \\ &\overset{(3.14)}{=} \mathrm{O}_N(N\cdot K_N^{-1}\max\{N^{1-3\alpha/2}, \log N\}).\end{aligned}$$

We thus conclude that (3.21) is of order

$$(3.21) = \mathrm{O}_N(c_N^{-1} K_N^{-1} \max\{N^{1-3\alpha/2}, \log N\}).$$

For (3.22), we use that for any random variable X we have the bound

$$\mathbb{E}[|X|^3 \mathbf{1}_{|X|\le K}] \le K\mathbb{E}[X^2]\,.$$

Hence we can bound

$$\begin{aligned}&\sum_{1\le i<j\le N} \mathbb{E}[(X_{ij}^b)^3 \mathbf{1}_{|X_{ij}^b|\le K_N} + (X_{ij}^g)^3 \mathbf{1}_{|X_{ij}^g|\le K_N}] \\ &\le K_N \sum_{1\le i<j\le N} \mathbb{E}[(X_{ij}^b)^2 + (X_{ij}^g)^2] \\ &\le 2K_N \sum_{1\le i<j\le N} \mathbf{E}[r_{ij}^m] \overset{(3.14)}{=} \mathrm{O}_N(K_N N c_N)\,.\end{aligned}$$

This yields that (3.22) is of order $\mathrm{O}_N(K_N c_N^{-1/2})$. Choosing $K_N = \mathrm{O}_N 1$ gives us that

$$\left|\mathbb{E}\left[h\left(\Re H\left(\mathbf{A}_{N,m,g}\right)\right)\right] - \mathbb{E}\left[h\left(\Re H\left(\overline{\mathbf{A}}_{N,m}\right)\right)\right]\right| = o_N(1)\,. \tag{3.23}$$

A similar argument holds for the imaginary part $\Im(H)$ and this completes the proof. $\square$

Simplification of the variance structure

To conclude Gaussianisation, we would like to construct a final matrix $\tilde{\mathbf{A}}_{N,m,g}$ with a simpler variance structure than that of $\mathbf{A}_{N,m,g}$. We let its entries be

$$\tilde{\mathbf{A}}_{N,m,g}(i,j) = \frac{\sqrt{r_{ij}^m}}{\sqrt{c_N}} G_{i\wedge j, i\vee j} \quad 1 \le i,j \le N \tag{3.24}$$

where r_{ij}^m is as in (3.13) and the $\{G_{i,j} : i \ge j\}$ are the i.i.d. collection of Gaussian variables used in (3.19). We need to prove that the ESD of this matrix gives asymptotically a good approximation of the ESD of $\mathbf{A}_{N,m,g}$.

Lemma 3.4.6 (Simplification of variance).
For any $\delta > 0$

$$\lim_{N\to\infty} \mathbb{P}\left(d_L(\mathrm{ESD}(\mathbf{A}_{N,m,g}), \mathrm{ESD}(\tilde{\mathbf{A}}_{N,m,g})) > \delta\right) = 0\,.$$

Proof. Construct a matrix $L_{N,g}$ with entries

$$L_{N,g}(i,j) = \begin{cases} \frac{\sqrt{p_{ij}^m}}{\sqrt{c_N}} G_{i\wedge j, i\vee j} & 1 \le i \ne j \le N \\ 0 & 1 \le i = j \le N \end{cases}$$

where $p_{ij}^m = r_{ij}^m \wedge 1$. By (3.8), we have that

$$\begin{aligned} \mathbb{E}[d_L^3(\mathrm{ESD}(\mathbf{A}_{N,m,g}), \mathrm{ESD}(L_{N,g}))] &\le \frac{1}{Nc_N} \sum_{i\ne j\in \mathbf{V}_N} \mathbb{E}\left[G_{i,j}^2 p_{ij}^m \left(\sqrt{1-p_{ij}^m} - 1\right)^2\right] \\ &\le \frac{1}{Nc_N} \sum_{i\ne j\in \mathbf{V}_N} \mathbf{E}[p_{ij}^m |(1-p_{ij}^m) - 1|] \\ &\le \frac{1}{Nc_N} \sum_{i\ne j\in \mathbf{V}_N} \mathbf{E}[(r_{ij}^m)^2] \overset{(3.14)}{=} o_N(1). \end{aligned}$$

For $i \ne j \in \mathbf{V}_N$ define the events $\mathcal{A}_{ij} = \{r_{ij}^m \le 1\}$. Construct yet another matrix $\tilde{L}_{N,g}$ as

$$\tilde{L}_{N,g}(i,j) = L_{N,g}(i,j)\mathbf{1}_{\mathcal{A}_{ij}} + \frac{X_{ij}}{\sqrt{c_N}} \mathbf{1}_{\mathcal{A}_{ij}^c}$$

where, conditional on the weights, $X_{ij} \sim \mathcal{N}\left(0, r_{ij}^m\right)$ are mutually independent and independent of the $\{G_{i,j}\}_{i>j}$. It is easy to see that $\tilde{L}_{N,g} = \tilde{\mathbf{A}}_{N,m,g}$ in distribution. So, comparing $L_{N,g}$ with $\tilde{L}_{N,g}$, using (3.8) we get

$$\begin{aligned} &\mathbb{E}[d_L^3(\mathrm{ESD}(\tilde{L}_{N,g}), \mathrm{ESD}(L_{N,g}))] \le \frac{1}{N} \sum_{i\ne j\in \mathbf{V}_N} \mathbb{E}[(L_{N,g}(i,j) - \tilde{L}_{N,g}(i,j))^2] \\ &= \frac{1}{N} \sum_{i\ne j\in \mathbf{V}_N}^{N} \mathbb{E}[(L_{N,g}(i,j) - \tilde{L}_{N,g}(i,j))^2 \mathbf{1}_{\mathcal{A}_{ij}^c}] \\ &= \frac{1}{N} \sum_{i\ne j\in \mathbf{V}_N}^{N} \mathbb{E}\left[\left(\frac{\sqrt{p_{ij}^m}}{\sqrt{c_N}} G_{i\wedge j, i\vee j} - \frac{X_{ij}}{\sqrt{c_N}}\right)^2 \mathbf{1}_{\mathcal{A}_{ij}^c}\right]. \end{aligned}$$

Using that the $G_{i,j}$ are centred and independent of the weights, and the Cauchy-Schwarz inequality, we can develop the square to obtain a further upper bound of the form

$$
\begin{aligned}
&\frac{1}{Nc_N}\sum_{i\neq j\in \mathbf{V}_N}^{N} \mathbf{E}[G^2_{i\wedge j, i\vee j}\mathbf{1}_{\mathcal{A}^c_{ij}}] + \mathbf{E}[X^2_{ij}\mathbf{1}_{\mathcal{A}^c_{ij}}] \\
&\le \frac{1}{Nc_N}\sum_{i\neq j\in \mathbf{V}_N}^{N} \mathbf{P}(\mathcal{A}^c_{ij}) + \mathbf{E}[X^4_{ij}]^{1/2}\mathbf{P}(\mathcal{A}^c_{ij})^{1/2} \\
&\le \frac{1}{Nc_N}\sum_{i\neq j\in \mathbf{V}_N}^{N} \mathbf{P}(\mathcal{A}^c_{ij}) + \frac{3\mathbf{E}[(W_i^m \vee W_j^m)^2(W_i^m \wedge W_j^m)^{2\sigma}]^{1/2}}{\|i-j\|^\alpha}\mathbf{P}(\mathcal{A}^c_{ij})^{1/2} \\
&= o_N(1)
\end{aligned}
$$

since

$$
\mathbf{P}(\mathcal{A}^c_{ij}) \le \mathbf{P}\left(W_i W_j^\sigma \ge \|i-j\|^\alpha\right) \le \frac{c}{\|i-j\|^{\alpha\left((\tau-1)\wedge\frac{\tau-1}{\sigma}\right)}}.
$$

Using the triangle inequality, we get

$$
\mathbb{E}[d_L^3(\mathrm{ESD}(\mathbf{A}_{N,m,g}), \mathrm{ESD}(\tilde{\mathbf{A}}_{N,m,g}))] = o_N(1)\,.
$$

We conclude the proof using Markov's inequality. □

§3.4.4 Moment method

Preliminary results: combinatorial setup

We will recall here the combinatorics features of partitions we need in the chapter, and refer the reader for a detailed exposition to Nica and Speicher [2006, Chapter 9].

For $k \ge 1$, denote by $\mathcal{P}(2k)$ the set of partitions of $[2k]$, and by $NC(2k) :=NC([2k])$ the set of non-crossing partitions of $\{1, 2, \dots, 2k\}$. When we write a partition, we order its blocks in such a way that the first block always contains 1, and the $(i+1)$th block contains the smallest element not belonging to any of the previous i blocks.

In what follows, we shall use Wick's formula. Let $(X_1, \dots, X_n)$ be a real Gaussian vector, then

$$
\mathbb{E}[X_{i_1}\cdots X_{i_k}] = \sum_{\pi\in\mathcal{P}_2(2k)}\prod_{(r,s)\in\pi}\mathbb{E}[X_{i_r}X_{i_s}], \tag{3.25}
$$

where $\mathcal{P}_2(2k)$ denotes the pair partitions of $[2k]$.

Any partition $\pi \in \mathcal{P}(k)$ can be realised as a *permutation* of $[k]$, that is, a bijective mapping $[k] \to [k]$. Let S_k denote the set of permutations on k

elements. Let $\gamma = (1, 2, \ldots, k) \in S_k$ be the shift by 1 modulo k. We will be interested in the composition of two permutations γ and π, denoted by $\gamma\pi$, which will be seen below as a partition.

As an example, consider $\pi = \{\{1, 2\}, \{3, 4\}\}$ and $\gamma = (1, 2, 3, 4)$. To compute $\gamma\pi$, we read π as $(1, 2)(3, 4)$, and compute $\gamma\pi = (1, 3)(2)(4)$. We finally read $\gamma\pi$ as $\{\{1, 3\}, \{2\}, \{4\}\}$. We now define a graph associated to a partition, borrowing the definition from Avena et al. [2023, Definition 2.3].

Definition 3.4.7 (Graph associated to a partition).
For a fixed $k \geq 1$, let γ denote the cyclic permutation $(1, 2, \ldots, k)$. For a partition π, we define $G_{\gamma\pi} = (\mathbf{V}_{\gamma\pi}, E_{\gamma\pi})$ as a rooted, labelled directed graph associated with any partition π of $[k]$, constructed as follows.

- *Initially consider the vertex set $\mathbf{V}_{\gamma\pi} = [k]$ and perform a closed walk on $[k]$ as $1 \to 2 \to 3 \to \cdots \to k \to 1$ and with each step of the walk, add an edge.*

- *Evaluate $\gamma\pi$, which will be of the form $\gamma\pi = \{V_1, V_2, \ldots, V_m\}$ for some $m \geq 1$ where $\{V_i\}_{1 \leq i \leq m}$ are disjoint blocks. Then, collapse vertices in $\mathbf{V}_{\gamma\pi}$ to a single vertex if they belong to the same block in $\gamma\pi$, and collapse the corresponding edges. Thus, $\mathbf{V}_{\gamma\pi} = \{V_1, \ldots, V_m\}$.*

- *Finally root and label the graph as follows.*
  - Root: *we always assume that the first element of the closed walk (in this case '1') is in V_1, and we fix the block V_1 as the root.*
  - Label: *each vertex V_i gets labelled with the elements belonging to the corresponding block in $\gamma\pi$.*

For the partitions $\pi = \{\{1, 2\}, \{3, 4\}\}$, $\gamma\pi = \{\{1, 3\}, \{2\}, \{4\}\}$, Figure 3.5 illustrates this procedure.

The following lemma is an exercise in Nica and Speicher [2006, Exercise 22.15] and explains also why non-crossing pair partitions will have the dominant role in the computations that follow. We will denote as $NC_2(2k)$ the set of non-crossing pair partitions of $[2k]$. For a partition π we let $\#\pi$ the number of its blocks.

Lemma 3.4.8.
Given $\pi \in \mathcal{P}_2(2k)$, one has $\#\gamma\pi \leq k + 1$ and the equality holds if and only $\pi \in NC_2(2k)$. If $\pi \in NC_2(2k)$, the graph $G_{\gamma\pi}$ is a rooted tree.

Finally, given $\pi \in NC_2(2k)$, we define the map $\mathcal{T} = \mathcal{T}_\pi : [2k] \to [k + 1]$ as follows. By Lemma 3.4.8, we know that $\#\gamma\pi = k + 1$ and let $\gamma\pi = \{V_1, V_2, \ldots, V_{k+1}\}$. Define

$$\mathcal{T}_\pi(i) = j \quad \text{if} \quad i \in V_j. \tag{3.26}$$

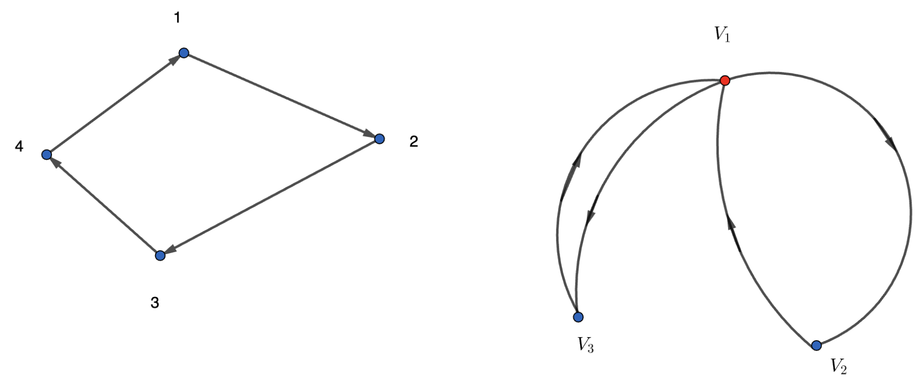

Figure 3.5: Left: closed walk on $[4]$. *Right: graph associated to* $\gamma\pi = \{\{1,3\},\{2\},\{4\}\}$. *The root is in red.*

Moment characterisation

We are now ready to give the proofs on Gaussianisation leading to the main result of this subsection, the proof of Theorem 3.2.1.

Proposition 3.4.9.
Let $\tilde{\mathbf{A}}_{N,m,g}$ *be defined as in* (3.24). *Let* $\mathrm{ESD}(\tilde{\mathbf{A}}_{N,m,g})$ *be its empirical spectral distribution. Then, for* $k \in \mathbb{N}$, *one has*

$$\lim_{N\to\infty} \mathbb{E}\left[\int_{\mathbb{R}} x^{2k}\, \mathrm{ESD}(\tilde{\mathbf{A}}_{N,m,g})(\mathrm{d}\, x)\right] = M_{2k} \tag{3.27}$$

and odd moments are zero. Moreover,

$$\lim_{N\to\infty} \mathrm{Var}\left(\int_{\mathbb{R}} x^{2k}\, \mathrm{ESD}(\tilde{\mathbf{A}}_{N,m,g})(\mathrm{d}\, x)\right) = 0, \tag{3.28}$$

where

$$M_{2k} = \sum_{\pi\in NC_2(2k)} \mathbf{E}\left[\prod_{(u,v)\in E(G_{\gamma\pi})} \kappa_\sigma(W_u^m, W_v^m)\right] < \infty, \tag{3.29}$$

where κ_σ *is as in* (3.3) *and* $E(G_{\gamma\pi})$ *is the edge set of the tree* $G_{\gamma\pi}$. *Moreover, there exists a unique compactly supported symmetric and deterministic measure* $\mu_{\sigma,\tau,m}$ *characterised by the moment sequence* $\{M_{2k}\}_{k\in\mathbb{N}}$ *such that*

$$\lim_{N\to\infty} \mathrm{ESD}(\tilde{\mathbf{A}}_{N,m,g}) = \mu_{\sigma,\tau,m} \quad \textit{in } \mathbb{P}\textit{-probability.} \tag{3.30}$$

Proof. Let $\{G_{i,j} : 1 \le i < j \le N\}$ be a sequence of standard independent centred Gaussian random variables as in (3.24) which is also independent of $(W_i)_{i\in[N]}$. Let $\mathcal{G}$ be the matrix

$$\mathcal{G}(i,j) = \begin{cases} \|i-j\|^{-\alpha/2} G_{i\wedge j, i\vee j} & i \neq j \\ 0 & i = j \end{cases} \tag{3.31}$$

Observe that

$$\tilde{\mathbf{A}}_{N,m,g} \stackrel{d}{=} \Upsilon_{\sigma,m} \circ \mathcal{G},$$

where $\Upsilon_{\sigma,m}$ is the matrix with elements

$$\Upsilon_{\sigma,m}(i,j) = \sqrt{\frac{\kappa_\sigma(W_i^m, W_j^m)}{c_N}}$$

and $\circ$ denotes the Hadamard product. Using Wick's formula (3.25) we have

$$\begin{aligned}
\mathbb{E}\left[\operatorname{tr}\left(\tilde{\mathbf{A}}_{N,m,g}^{2k}\right)\right] &= \frac{1}{N c_N^k} \sum_{1\le i_1,\dots,i_{2k}\le N}{}^{'} \mathbb{E}\left[\prod_{\ell=1}^{2k} \Upsilon_{\sigma,m}(i_\ell, i_{\ell+1}) \prod_{\ell=1}^{2k} \mathcal{G}(i_\ell, i_{\ell+1})\right] \\
&= \frac{1}{N c_N^k} \sum_{1\le i_1,\dots,i_{2k}\le N}{}^{'} \mathbb{E}\left[\prod_{\ell=1}^{2k} \kappa_\sigma^{1/2}(W_{i_\ell}^m, W_{i_{\ell+1}}^m)\right] \\
&\qquad \times \sum_{\pi\in\mathcal{P}_2(2k)} \prod_{(r,s)\in\pi} \mathbb{E}\left[\mathcal{G}(i_r, i_{r+1})\mathcal{G}(i_s, i_{s+1})\right] \\
&= \frac{1}{N c_N^k} \sum_{1\le i_1,\dots,i_{2k}\le N}{}^{'} \mathbb{E}\left[\prod_{\ell=1}^{2k} \kappa_\sigma^{1/2}(W_{i_\ell}^m, W_{i_{\ell+1}}^m)\right] \\
&\qquad \times \sum_{\pi\in\mathcal{P}_2(2k)} \prod_{(r,s)\in\pi} \frac{1}{\|i_r - i_{r+1}\|^\alpha} \mathbf{1}_{\{i_r, i_{r+1}\} = \{i_s, i_{s+1}\}},
\end{aligned} \tag{3.32}$$

where we set $i_{2k+1} = i_1$ to ease notation, and $(r,s) \in \pi$ means $\pi(r) = s$ and $\pi(s) = r$. Here the $\sum'$ indicates the sum over all the indices $(i_1, \dots, i_{2k})$ such that $i_\ell \neq i_{\ell+1}$ for $\ell \in [2k]$. The condition $\{i_r, i_{r+1}\} = \{i_s, i_{s+1}\}$ is satisfied in two cases:

C1) $i_r = i_{s+1}$ and $i_s = i_{r+1}$, that is, $i_r = i_{\gamma\pi(r)}$ and $i_s = i_{\gamma\pi(s)}$, or

C2) $i_r = i_s$ and $i_{r+1} = i_{s+1}$, that is, $i_r = i_{\pi(r)}$ and $i_{r+1} = i_{\pi(r)+1}$.

As we are going to show, the limit of (3.32) will be supported on permutations $\pi \in NC_2(2k)$ and such that Case 1) is true for all $(r,s) \in \pi$. To prove this, let us define

$$\operatorname{Cat}_{\pi,k} = \{\mathbf{i} = (i_1, \dots, i_{2k}) \in [N]^{2k} : i_r \neq i_{r+1}, \, i_r = i_{\gamma\pi(r)} \; \forall r \in [2k]\}.$$

When the condition $i_r = i_{\gamma\pi(r)}$ holds for all r, we see that $\mathbf{i}$ is constant on the blocks of $\gamma\pi$. We construct a graph $G(\mathbf{i})$ associated to $\mathbf{i} \in \mathrm{Cat}_{\pi,k}$ by performing a closed walk $i_1 \to i_2 \to \ldots i_{2k} \to i_1$, and then collapsing elements i_r, i_s into the same vertex if r, s belong to the same block in $\gamma\pi$. We then collapse multiple edges. After this, we see that $G(\mathbf{i}) = G_{\gamma\pi}$. Thus, when we sum over $\mathbf{i} \in \mathrm{Cat}_{\pi,k}$, the count is over $\#\gamma\pi$ many indices.

We split the summation in (3.32) into two parts: a first sum over the non-crossing pairings and $\mathbf{i} \in \mathrm{Cat}_{\pi,k}$, and a second part with all the other terms, that we call $\mathcal{R}_1$. Since we take $\mathbf{i} \in \mathrm{Cat}_{\pi,k}$, $\mathbf{i}$ is constant on the blocks of $\gamma\pi$. Using this property, we obtain

$$
\begin{aligned}
&\mathbb{E}\left[\mathrm{tr}\left(\tilde{\mathbf{A}}^{2k}_{N,m,g}\right)\right] \\
&= \sum_{\pi \in NC_2(2k)} \frac{1}{N c_N^k} \sum_{\mathbf{i} \in \mathrm{Cat}_{\pi,k}} \mathbb{E}\left[\prod_{j=1}^{2k} \kappa_\sigma^{1/2}(W^m_{i_j}, W^m_{i_{j+1}})\right] \prod_{(r,s)\in\pi} \frac{1}{\|i_r - i_{r+1}\|^\alpha} + \mathcal{R}_1 \\
&= \sum_{\pi \in NC_2(2k)} \frac{1}{N c_N^k} \sum_{\mathbf{i} \in \mathrm{Cat}_{\pi,k}} \mathbb{E}\left[\prod_{(u,v)\in E(G_{\gamma\pi})} \kappa_\sigma(W^m_u, W^m_v)\right] \prod_{(r,s)\in\pi} \frac{1}{\|i_r - i_{r+1}\|^\alpha} + \mathcal{R}_1
\end{aligned}
$$

where in the last line we have used that $\mathbf{i}$ is constant on the blocks of $\gamma\pi$. Since the inner expectation no longer depends on $\mathbf{i}$, we get that

$$
\begin{aligned}
&\mathbb{E}\left[\mathrm{tr}\left(\tilde{\mathbf{A}}^{2k}_{N,m,g}\right)\right] \\
&= \sum_{\pi \in NC_2(2k)} \mathbb{E}\left[\prod_{(u,v)\in E(G_{\gamma\pi})} \kappa_\sigma(W^m_u, W^m_v)\right] \frac{1}{N c_N^k} \sum_{\mathbf{i} \in \mathrm{Cat}_{\pi,k}} \prod_{(r,s)\in\pi} \frac{1}{\|i_r - i_{r+1}\|^\alpha} \\
&+ \mathcal{R}_1.
\end{aligned}
$$

Now we make the following two claims which will finish the proof.

Claim 3.4.10.
The following hold.

a) For any $\pi \in NC_2(2k)$,

$$
\lim_{N\to\infty} \frac{1}{N c_N^k} \sum_{\mathbf{i} \in \mathrm{Cat}_{\pi,k}} \prod_{(r,s)\in\pi} \frac{1}{\|i_r - i_{r+1}\|^\alpha} = 1.
$$

b) We have that $\lim_{N\to\infty} \mathcal{R}_1 = 0$.

With the above claim, whose proof is deferred to page 131, we have that (3.27) holds. Moreover, the odd moments are identically 0, since there are no non-crossing pair partitions for tuples of the form $\{1, 2, \ldots, 2k+1\}, k \in \mathbb{N}$. We now need to now show that (3.28) holds.

We introduce some new notation to prove (3.28). Let $\mathbf{j} = (j_1, \ldots, j_{2k})$. Let $P(\mathbf{i})$ denote the expectation

$$P(\mathbf{i}) \overset{(3.31)}{:=} \mathbb{E}\left[\prod_{\ell=1}^{2k} \kappa_\sigma^{1/2}(W_{i_\ell}^m, W_{i_{\ell+1}}^m)\mathcal{G}(i_\ell, i_{\ell+1})\right],$$

and $P(\mathbf{i}, \mathbf{j})$ be

$$P(\mathbf{i}, \mathbf{j}) := \mathbb{E}\left[\prod_{\ell=1}^{2k} \kappa_\sigma^{1/2}(W_{i_\ell}^m, W_{i_{\ell+1}}^m)\mathcal{G}(i_\ell, i_{\ell+1}) \prod_{p=1}^{2k} \kappa_\sigma^{1/2}(W_{i_p}^m, W_{i_{p+1}}^m)\mathcal{G}(i_p, i_{p+1})\right]$$

(with the usual cyclic convention that $2k+1$ equals 1 for subscripts of indices). We can then see that

$$\mathrm{Var}\left(\int_{\mathbb{R}} x^{2k}\, \mathrm{ESD}(\tilde{\mathbf{A}}_{N,m,g})(\mathrm{d}\, x)\right) = \frac{1}{N^2 c_N^{2k}} \sum_{\mathbf{i},\mathbf{j}:[2k]\to[N]} \left[P(\mathbf{i}, \mathbf{j}) - P(\mathbf{i})P(\mathbf{j})\right]. \tag{3.33}$$

Note that if the terms involving $\mathbf{i}$ and $\mathbf{j}$ are completely different, that is, if the product of the terms $\mathcal{G}(i_1, i_2) \cdots \mathcal{G}(i_{2k}, i_1)$ is independent of $\mathcal{G}(j_1, j_2) \cdots \mathcal{G}(j_{2k}, j_1)$, then $P(\mathbf{i}, \mathbf{j}) = P(\mathbf{i})P(\mathbf{j})$, and (3.33) becomes identically 0. Hence, we have

$$\mathrm{Var}\left(\int_{\mathbb{R}} x^{2k} \mu_{\tilde{\mathbf{A}}_{N,m,g}}(\mathrm{d}\, x)\right) = \frac{1}{N^2 c_N^{2k}} \sum_{\mathbf{i},\mathbf{j}:[2k]\to[N]}^{(\geq 1)} P(\mathbf{i}, \mathbf{j}), \tag{3.34}$$

where $\sum^{(\geq 1)}$ is over $\mathbf{i}, \mathbf{j}$ such that there is *at least* one matching of the form $\tilde{\mathbf{A}}_{N,m,g}(i_r, i_{r+1}) = \tilde{\mathbf{A}}_{N,m,g}(j_s, j_{s+1})$ for some $1 \leq r,\ s \leq 2k-1$. If there is only one entry of $\mathbf{i}$, say i_1, equal to only one entry of $\mathbf{j}$, say j_1, then we still have

$$E^W\left[\prod_{\ell=1}^{2k} \mathcal{G}(i_\ell, i_{\ell+1})\mathcal{G}(j_\ell, j_{\ell+1})\right] = 0$$

since all entries $\mathcal{G}(i_\ell, i_{\ell+1})$ are independent (even if $i_1 = j_1$) and centred. All the more, $P(\mathbf{i}, \mathbf{j}) = 0$, so let us pass to having two equal indices, that is, a matching.

Let us consider the case when there is *exactly one* matching. Since both indices in $\mathbf{i}$ and $\mathbf{j}$ can be reordered without affecting the variance,without loss of generality we can assume that the matching is $(i_1, i_2) = (j_1, j_2)$, and the rest of the indices of $\mathbf{i}$ are different from the ones in $\mathbf{j}$. One now has $\mathbf{i}' = (i_3, \ldots, i_{2k})$

and $\mathbf{j}' = (j_3, \ldots, j_{2k})$ with $2k-2$ indices each, and so we can construct partitions π, π' for each of them independently.

For the ease of notation, let

$$a_{i,j} := \kappa_\sigma^{1/2}(W_i^m, W_j^m)\mathcal{G}(i, j)$$

and let $\sum^{(1)}$ be the sum over $\mathbf{i}, \mathbf{j}$ such that there is exactly one matching between $\mathbf{i}$ and $\mathbf{j}$. Using Wick's formula in the second equality, we have

$$\begin{aligned}
&\frac{1}{N^2 c_N^{2k}} \sum_{\mathbf{i},\mathbf{j}:[2k]\to[N]}^{(1)} P(\mathbf{i},\mathbf{j}) \\
&= \frac{1}{N^2 c_N^{2k}} \sum_{\mathbf{i},\mathbf{j}:[2k]\to[N]}^{(1)} \mathbf{E}\left[E^W\left[\prod_{\ell=1}^{2k} a_{i_\ell, i_{\ell+1}} a_{j_\ell, j_{\ell+1}}\right]\right] \\
&= \frac{1}{N^2 c_N^{2k}} \sum_{\mathbf{i},\mathbf{j}:[2k]\to[N]} \mathbf{E}\Bigg[E^W[a_{i_1,i_2}^2] \sum_{\pi,\pi'\in\mathcal{P}_2(\{3,\ldots,2k\})} \prod_{(r,s)\in\pi} E^W[a_{i_r, i_{r+1}} a_{i_s, i_{s+1}}] \\
&\qquad \times \prod_{(r',s')\in\pi'} E^W[a_{j_{r'}, j_{r'+1}} a_{j_{s'}, j_{s'+1}}]\Bigg]. \qquad (3.35)
\end{aligned}$$

Following the idea of the proof for (3.27), we assume Claim 3.4.10 to be true to obtain the optimal order. We will consider $\mathbf{i}', \mathbf{j}' \in \mathrm{Cat}_{\pi,k-1}$, and notice that

$$E^W[a_{\ell,\ell'}^2] \le \frac{m^{1+\sigma}}{\|\ell - \ell'\|^\alpha}. \qquad (3.36)$$

Interchanging summands, we obtain

$$\begin{aligned}
(3.35) &= \frac{1}{N^2 c_N^{2k}} \mathbf{E}\Bigg[\sum_{\pi,\pi'\in\mathcal{P}_2(\{3,\ldots,2k\})} \sum_{\substack{\mathbf{i}',\mathbf{j}'\in\mathrm{Cat}_{\pi,k-1},\\ i_1\ne i_2\in[N]}} E^W\left[a_{i_1,i_2}^2\right] \prod_{(r,s)\in\pi} E^W\left[a_{i_r i_{\gamma\pi(r)}}^2\right] \\
&\qquad \times \prod_{(r',s')\in\pi'} E^W\left[a_{j_{r'} j_{\gamma\pi(r')}}^2\right]\Bigg] + \mathcal{R}_1' \\
&\overset{(3.36)}{\le} \frac{1}{N^2 c_N^{2k}} \sum_{\pi,\pi'\in\mathcal{P}_2(\{3,\ldots,2k\})} \sum_{\substack{\mathbf{i}',\mathbf{j}'\in\mathrm{Cat}_{\pi,k-1},\\ i_1\ne i_2\in[N]}} \frac{m^{1+\sigma}}{\|i_1 - i_2\|^\alpha} \prod_{(r,s)\in\pi} \frac{m^{1+\sigma}}{\|i_r - i_{\gamma\pi(r)}\|^\alpha} \\
&\qquad \times \prod_{(r',s')\in\pi'} \frac{m^{1+\sigma}}{\|j_{r'} - j_{\gamma\pi(r')}\|^\alpha} + \mathcal{R}_1', \qquad (3.37)
\end{aligned}$$

where $\mathcal{R}_1'$ is an error term such that $\lim_{N\to\infty} \mathcal{R}_1' = 0$, which follows from Claim 3.4.10. The contributing terms of the right-hand side of (3.37) can be upper-bounded by

$$\frac{1}{N^2 c_N^{2k}} \sum_{\pi,\pi'\in\mathcal{P}_2(\{3,\ldots,2k\})} \sum_{\substack{\mathbf{i}:\mathbf{i}'\in\mathrm{Cat}_{\pi,k-1},\\ i_1\neq i_2}} \frac{m^{1+\sigma}}{\|i_1-i_2\|^\alpha} \prod_{(r,s)\in\pi} \frac{m^{1+\sigma}}{\|i_r - i_{\gamma\pi(r)}\|^\alpha}$$
$$\times \sum_{\mathbf{j}'\in\mathrm{Cat}_{\pi',k-1}} \prod_{(r',s')\in\pi'} \frac{m^{1+\sigma}}{\|j_{r'} - j_{\gamma\pi(r')}\|^\alpha}$$
$$= \frac{1}{N^2 c_N^{2k}} \sum_{\pi,\pi'\in\mathcal{P}_2(\{3,\ldots,2k\})} \sum_{\substack{\mathbf{i}:\mathbf{i}'\in\mathrm{Cat}_{\pi,k-1},\\ i_1\neq i_2}} \frac{m^{1+\sigma}}{\|i_1-i_2\|^\alpha} \prod_{(r,s)\in\pi} \frac{m^{1+\sigma}}{\|i_r - i_{\gamma\pi(r)}\|^\alpha} O_N(N c_k^{k-1}).$$

Analogously, the sum over $\mathbf{i}$ conditioned on $\mathbf{i}' \in \mathrm{Cat}_{\pi,k-1}$ will be at most of order $N c_N^k$. Since the sum over partitions is finite and independent of N, we obtain

$$\frac{1}{N^2 c_N^{2k}} \sum_{\mathbf{i},\mathbf{j}:[2k]\to[N]}^{(1)} P(\mathbf{i},\mathbf{j}) = \mathrm{O}_N(c_N^{-1}).$$

More generally, if one has t pairings of the form $(i_1, i_2) = (j_1, j_2), \ldots, (i_{t-1}, i_t) = (j_{t-1}, j_t)$, one can use the same argument and instead obtain a faster error of the order of c_N^{-t+1}, simply due to the set $(j_{t+1}, j_2, \ldots, j_{2k})$ now having only $2k - t$ independent indices from $\mathbf{i}$. Thus, we conclude

$$\mathrm{Var}\left(\int_{\mathbb{R}} x^{2k} \mu_{\tilde{\mathbf{A}}_{N,m,g}}(\mathrm{d}\, x)\right) = \mathrm{O}_N(c_N^{-1}). \tag{3.38}$$

This proves (3.28).

To conclude, one can see that

$$M_{2k} \le (m^{1+\sigma})^k C_k, \tag{3.39}$$

where C_k is the k^{th} Catalan number. Since $\sum_{k\ge1} C_k^{-1/2k} = \infty$, so Carleman's condition implies that $\{M_{2k}\}_{k\ge1}$ uniquely determine the limiting measure. Therefore we can find C, $R > 0$ such that for all $k \ge 1$ we have $M_{2k} \le CR^{2k}$. In turn, it is a straightforward exercise to show that this implies that $\mu_{\tau,\sigma,m}$ is compactly supported, and since it has odd moments equal to zero it is symmetric. To conclude the proof of Proposition 3.4.9 we use for example Tao [2012, pg. 134].

Proof of Claim 3.4.10. We first show a). Fix $\pi \in NC_2(2k)$. Recall that $\mathbf{i} \in \mathrm{Cat}_{\pi,k}$ is constant on the blocks of $\gamma\pi$. Therefore the number of free indices over which we can construct $\mathbf{i}$ is $\#\gamma\pi = k+1$ (Lemma 3.4.8).

For any $\pi \in NC_2(2k)$, there exists *at least* one block of the form $(r, r+1) \in \pi$, where $1 \leq r \leq 2k$, and $2k+1$ is identified with "1". Then, $\{r+1\} \in \gamma\pi$ is a singleton, and consequently, i_{r+1} is a free index under $\gamma\pi$, that is, under the summation over indices $i_1, \ldots, i_{2k}$, i_{r+1} runs from 1 to N independent of other indices. Moreover, as $\mathbf{i} \in \mathrm{Cat}_{\pi,k}$, we have $i_r = i_{r+2}$. If we remove the block $(r, r+1)$ from π, we obtain $\pi' \in NC_2(2k-2)$ as a new partition on $\{1, 2, \ldots, r-1, r+2, \ldots 2k\}$. Let $\mathbf{i}'$ be the tuple $(i_1, i_2, \ldots, i_{r-1}, i_{r+2}, \ldots, i_{2k})$. We then have $\mathbf{i}' \in \mathrm{Cat}_{\pi',k-1}$. So, we can write

$$\frac{1}{Nc_N^k} \sum_{\mathbf{i} \in \mathrm{Cat}_{\pi,k}} \prod_{(r,s) \in \pi} \frac{1}{\|i_r - i_s\|^\alpha}$$
$$= \frac{1}{Nc_N^k} \sum_{\mathbf{i}' \in \mathrm{Cat}_{\pi',k-1}} \left(\prod_{(r,s) \in \pi'} \frac{1}{\|i_r - i_s\|^\alpha} \right) \left(\sum_{i_{r+1}=1}^{N} \frac{1}{\|i_{r+1} - i_{r+2}\|^\alpha} \right). \tag{3.40}$$

We now proceed inductively. For $k = 1$ the result is given by (3.5). Assume now that we have shown, for some $k-1 \geq 0$ and any $\pi' \in NC_2(2(k-1))$, that

$$\lim_{N\to\infty} \frac{1}{Nc_N^{k-1}} \sum_{\mathbf{i}' \in \mathrm{Cat}_{\pi',k-1}} \prod_{(r,s) \in \pi'} \frac{1}{\|i_r - i_s\|^\alpha} = 1. \tag{3.41}$$

We need to show the same statement holds for k, which is precisely Claim 3.4.10a). Now, we have that

$$(3.40) = \frac{1}{Nc_N^{k-1}} \sum_{\mathbf{i}' \in \mathrm{Cat}_{\pi',k-1}} \left(\prod_{(r,s) \in \pi'} \frac{1}{\|i_r - i_s\|^\alpha} \right) \left(\frac{1}{c_N} \sum_{i_{r+1}=1}^{N} \frac{1}{\|i_{r+2} - i_{r+1}\|^\alpha} \right). \tag{3.42}$$

Taking the limit $N \to \infty$, we have that the second factor in brackets above by (3.5), and then the remaining expression equals 1 by the induction hypothesis (3.41). This proves a).

To show b), we now analyse $\mathcal{R}_1$ explicitly. We have to deal with two cases:

b.1) $\pi \in \mathcal{P}_2(2k)$ and $\mathbf{i} \notin \mathrm{Cat}_{\pi,k}$.

b.2) $\pi \in \mathcal{P}_2(2k) \setminus NC_2(k)$ and $\mathbf{i} \in \mathrm{Cat}_{\pi,k}$.

Note that for both cases the following factor involving the weights will not play any role:

$$\mathbb{E}\left[\prod_{j=1}^{2k} \kappa_\sigma^{1/2}(W_{i_j}^m, W_{i_{j+1}}^m) \right] \leq m^{k(1+\sigma)}.$$

We first deal with Case b.2). From Lemma 3.4.8 we have $\#\gamma\pi \leq k$ and hence

$$\sum_{\pi\in\mathcal{P}_2(2k)\setminus NC_2(2k)} \frac{1}{Nc_N^k} \sum_{\mathbf{i}\in\mathrm{Cat}_{\pi,k}} \mathbb{E}\left[\prod_{j=1}^{2k} \kappa_\sigma^{1/2}(W^m_{i_j}, W^m_{i_{j+1}})\right] \prod_{(r,s)\in\pi} \frac{1}{\|i_r - i_{r+1}\|^\alpha}$$
$$\leq m^{k(1+\sigma)} \sum_{\pi\in\mathcal{P}_2(2k)\setminus NC_2(2k)} \frac{1}{Nc_N^k} \sum_{i_1\in[N]} \sum_{i_2,\ldots,i_k\in[N]} \frac{1}{\|i_2\|^\alpha \ldots \|i_k\|^\alpha}, \tag{3.43}$$

where (3.43) follows from $\mathbf{i}$ being constant on the cycles of $\gamma\pi$. Thus, we get that the terms involved in Case b.2) give a contribution of the order

$$(3.43) \leq cm^{k(1+\sigma)} \sum_{\pi\in\mathcal{P}_2(2k)\setminus NC_2(2k)} \frac{1}{N^{1+k(1-\alpha)}} N^{1+(k-1)(1-\alpha)} = \mathrm{O}_N \frac{1}{N^{1-\alpha}} = o_N(1)\,. \tag{3.44}$$

We now show that the contribution from b.1) is also negligible. Begin by fixing a partition π. For any tuple $\mathbf{i}$, we construct a corresponding graph $G(\mathbf{i})$ (recall that when $\mathbf{i} \in \mathrm{Cat}_{\pi,k}$ we ended up with $G(\mathbf{i}) = G_{\gamma\pi}$). For $\mathbf{i} \notin \mathrm{Cat}_{\pi,k}$, $G(\mathbf{i})$ is constructed by a closed walk $i_1 \to i_2 \to \ldots i_{2k} \to i_1$, thereby adding the edges $(i_p, i_{p+1})_{p=1}^{2k}$ with $i_{2k+1} = i_1$. We then collapse indices i_r, i_s into the same vertex when $\{i_r, i_{r+1}\} = \{i_s, i_{s+1}\}$, which can be justified by (3.32). We then proceed by collapsing the multiple edges and looking at the skeleton graph $G(\mathbf{i})$, with vertex set $V(\mathbf{i})$. Hence, we see that

$$\begin{aligned}
&\sum_{\pi\in\mathcal{P}_2(2k)} \frac{1}{Nc_N^k} \sum_{\mathbf{i}:[2k]\to[N]}{}' \mathbb{E}\left[\prod_{j=1}^{2k} \kappa_\sigma^{1/2}(W^m_{i_j}, W^m_{i_{j+1}})\right] \prod_{(r,s)\in\pi} \frac{1}{\|i_r - i_{r+1}\|^\alpha} \\
&\leq m^{k(1+\sigma)} \sum_{\pi\in\mathcal{P}_2(2k)} \frac{1}{Nc_N^k} N^{1+(\#V(\mathbf{i})-1)(1-\alpha)} \\
&\leq \mathrm{O}N^{(\#V(\mathbf{i})-k-1)(1-\alpha)}.
\end{aligned} \tag{3.45}$$

since $m > 1$ is fixed and the sum over the set $\mathcal{P}_2(2k)$ is finite. We see that the only non-trivial contribution comes when $\#V(\mathbf{i}) = k+1$, which signifies that $G(\mathbf{i})$ is a tree. Now we claim that for any $\pi \in \mathcal{P}_2(2k)$ and $\mathbf{i} \notin \mathrm{Cat}_{\pi,k}$ we have $\#V(\mathbf{i}) < k+1$.

When $\mathbf{i} \notin \mathrm{Cat}_{\pi,k}$, it implies that there exists at least one $(r,s) \in \pi$, such that $i_r = i_s$ and $i_{r+1} = i_{s+1}$. Let us begin by assuming that there exists *exactly one* such pair. Observe that due to the restrictions in $\sum'$, no pair-wise indices are same, hence s can neither be $r+1$, nor $r-1$. Now consider the reduced partition $\pi' = \pi \setminus (r, s)$. Observe that $\pi' \in \mathcal{P}_2(2k)(\{1, \dots, r-1, r+1, \dots, s-1, s+1, \dots, 2k\})$. Note that now $\mathbf{i}' \in \mathrm{Cat}_{\pi', k-1}$, so its contribution to (3.37) is of the order of $N^{1+(k-1)(1-\alpha)}$, which comes from the tree $G(\mathbf{i}')$ on k vertices, and where $\mathbf{i}'$ are the $(2k-2)$ indices which are obtained by removal of (i_r, i_{r+1}). So, all we are left to show is that due to Case 2), i_r and i_s will not give rise to a new vertex in $G(\mathbf{i})$.

Now, there exists an $r < e < s-1$ such that $(e, s-1) \in \pi$. Due to Case 2), we have that $i_r = i_s$ contribute to the same vertex in $G(\mathbf{i})$. Also $i_e = i_s$ and $i_{e+1} = i_{s-1}$ due to Case 1). This implies that $i_r = i_s = i_e$, where i_e is already a contributing index in $G(\mathbf{i}')$. This implies that $G(\mathbf{i})$ is a tree on at most k vertices, and hence $\#V(\mathbf{i}) \le k$. This shows that the contribution in (3.45) goes to 0.

The case for which there is more than one pair breaking the constraint in $\mathrm{Cat}_{\pi, k}$ leads to an even smaller order. When none of the pairs satisfy the constraint then $i_r = i_{\pi(r)}$ for all r and hence i is constant on the blocks of π. So $\#V(\mathbf{i}) \le k$ and again the contribution in (3.45) goes to 0, thus proving the claim. □

We wish to highlight that Proposition 3.4.9 is in fact more general, and works beyond the kernels κ_σ defined in (3.3).

Remark 3.4.11.
The statement of Proposition 3.4.9 holds when we replace the entries of $\tilde{\mathbf{A}}_{N,m,g}$ in (3.24) by

$$\sqrt{\frac{\kappa(W_i, W_j)}{c_N \,\|i-j\|^\alpha}}\, G_{i\wedge j, i\vee j} \qquad 1 \le i, j \le N$$

for any function $\kappa : [1, \infty)^2 \to [0, \infty)$ which is symmetric and such that, for all $k \in \mathbb{N}$,

$$\mathbb{E}\left[\prod_{j=1}^{2k} \sqrt{\kappa(X_j, X_{j+1})}\right] < \infty \tag{3.46}$$

where $X_1, \dots, X_{2k}$ are i.i.d. random variables in $[1, \infty)$.

In our case the kernels $\kappa(x, y) := \kappa_\sigma(x, y)\mathbf{1}_{x,y\le m}$ satisfy (3.46).

Proof of Theorem 3.2.1. To prove the final result, we shall use Lemma 3.3.5 with the complete metric space $\Sigma = \mathcal{P}(\mathbb{R})$ and metric d_L. Recall also the definition of $\tilde{\mathbf{A}}_{N,m,g}$ resp. $\overline{\mathbf{A}}_{N,m}$ of (3.24) resp. (3.17). In Proposition 3.4.9 we have shown that there exists a (deterministic) measure $\mu_{\sigma,\tau,m}$ such that, for every $m > 0$,

$$\lim_{N\to\infty} \mathrm{ESD}(\tilde{\mathbf{A}}_{N,m,g}) = \mu_{\sigma,\tau,m} \text{ in } \mathbb{P}\text{–probability.}$$

Hence for any h satisfying the assumptions of Lemma 3.4.4 and H as in (3.20) it follows that

$$\lim_{N\to\infty} \mathbb{E}\left[h\left(\Re H\left(\tilde{\mathbf{A}}_{N,m,g}\right)\right)\right] = h\left(\Re\, \mathrm{S}_{\mu_{\sigma,\tau,m}}(z)\right).$$

and thus, by means of Lemma 3.4.4 and Lemma 3.4.6,

$$\lim_{N\to\infty} \mathbb{E}\left[h\left(\Re H\left(\overline{\mathbf{A}}_{N,m}\right)\right)\right] = h\left(\Re\, \mathrm{S}_{\mu_{\sigma,\tau,m}}(z)\right).$$

Since the above holds true for any h satisfying the assumptions of Lemma 3.4.4 and $\mu_{\sigma,\tau,m}$ is deterministic, it follows that

$$\lim_{N\to\infty} \Re H\left(\overline{\mathbf{A}}_{N,m}\right) = \Re\, \mathrm{S}_{\mu_{\sigma,\tau,m}}(z) \text{ in } \mathbb{P}\text{–probability.}$$

A similar argument for the imaginary part shows that

$$\lim_{N\to\infty} \Im H\left(\overline{\mathbf{A}}_{N,m}\right) = \Im\, \mathrm{S}_{\mu_{\sigma,\tau,m}}(z) \text{ in } \mathbb{P}\text{–probability.}$$

Combining the real and imaginary parts, we have, for any $z \in \mathbb{C}^+$,

$$\lim_{N\to\infty} \mathrm{S}_{\mathrm{ESD}(\overline{\mathbf{A}}_{N,m})}(z) = \mathrm{S}_{\mu_{\sigma,\tau,m}}(z) \quad \text{in } \mathbb{P}\text{–probability.}$$

Since the convergence of the Stieltjes transform characterises weak convergence, we have

$$\lim_{N\to\infty} \mathrm{ESD}(\overline{\mathbf{A}}_{N,m}) = \mu_{\sigma,\tau,m} \quad \text{in } \mathbb{P}\text{–probability.}$$

From Lemma 3.4.6 and Lemma 3.4.3, it also follows that, for every $\delta > 0$ and $m > 0$,

$$\limsup_{N\to\infty} \mathbb{P}(d_L(\mu_{\mathbf{A}_{N,m}}, \mu_{\sigma,\tau,m}) > \delta) = 0.$$

This shows condition (1) of Lemma 3.3.5. Condition (2) follows from Lemma 3.4.2 where we have proved that

$$\limsup_{m\to\infty} \lim_{N\to\infty} \mathbb{P}\left(d_L(\mu_{\mathbf{A}_{N,m}}, \mu_{\mathbf{A}_N}) > \delta\right) = 0.$$

Thus, it follows from Lemma 3.3.5 that there exists a deterministic measure $\mu_{\sigma,\tau}$ such that

$$\lim_{m\to\infty} d_L(\mu_{\sigma,\tau,m}, \mu_{\sigma,\tau}) = 0, \tag{3.47}$$

and hence using the triangle inequality the result follows.

§3.5 Scale-Free Percolation: A special case

Proof of Theorem 3.2.2. Step 1: identification. We are now dealing with the special case of $\sigma = 1$. We go back to the moments of $\mu_{\sigma,\tau,m}$. Let $\gamma\pi = (V_1, \ldots, V_{k+1})$ and let $\ell_i = \#V_i$ (with a slight abuse of notation, we are viewing here V_i as a set rather than a cycle). Since $\sigma = 1$, $\kappa_\sigma(W_u^m, W_v^m) = W_u^m W_v^m$. It follows that

$$\begin{aligned} M_{2k} &= \sum_{\pi \in NC_2(2k)} \mathbb{E}\left[\prod_{(u,v)\in E(G_{\gamma\pi})} W_u^m W_v^m\right] \\ &= \sum_{\pi \in NC_2(2k)} \prod_{i=1}^{k+1} \mathbb{E}[(W_1^m)^{\ell_i}] \\ &= \int_{\mathbb{R}} x^{2k} \mu_{sc} \boxtimes \mu_{W,m}(\mathrm{d}\, x). \end{aligned}$$

The last equality follows from the combinatorial expression of the moments of the free multiplicative convolution of the semicircle element with an element whose law is given by $\mu_{W,m}$ (see Nica and Speicher [2006, Theorem 14.4]). Consider the map $x \mapsto x^2$ from $\mathbb{R} \to [0, \infty)$ and let μ^2 be the push-forward of a probability measure μ under this mapping, so that μ_{sc} is pushed forward to μ_{sc}^2. Then by Bercovici and Voiculescu [1993, Corollary 6.7] it follows that

$$\lim_{m\to\infty} \mu_{W,m} \boxtimes \mu_{sc}^2 \boxtimes \mu_{W,m} = \mu_W \boxtimes \mu_{sc}^2 \boxtimes \mu_W .$$

A consequence of Arizmendi and Pérez-Abreu [2009, Lemma 8] is that

$$\mu_{W,m} \boxtimes \mu_{sc}^2 \boxtimes \mu_{W,m} = (\mu_{sc} \boxtimes \mu_{W,m})^2$$

and

$$\mu_W \boxtimes \mu_{sc}^2 \boxtimes \mu_W = (\mu_{sc} \boxtimes \mu_W)^2. \tag{3.48}$$

Thus

$$\lim_{m\to\infty} (\mu_{sc} \boxtimes \mu_{W,m})^2 = (\mu_{sc} \boxtimes \mu_W)^2.$$

Observe that $\mu_{sc} \boxtimes \mu_{W,m}$ and $\mu_{sc} \boxtimes \mu_W$ are symmetric around the origin [Arizmendi and Pérez-Abreu, 2009, Theorem 7], hence we have that

$$\lim_{m\to\infty} d_L(\mu_{sc} \boxtimes \mu_W, \mu_{sc} \boxtimes \mu_{W,m}) = \lim_{m\to\infty} d_L(\mu_{sc} \boxtimes \mu_W,\, \mu_{1,\tau,m},) = 0.$$

Theorem 3.2.1 then implies that the ESD$(\mathbf{A}_N)$ converges to $\mu_{sc} \boxtimes \mu_W$ weakly in probability.

Step 2: tail asymptotics. In the following we use the recent results of Kołodziejek and Szpojankowski [2022, Lemma 7.2] from which we also borrow the notation. The free probability analogue of the classical Breiman's lemma is as follows: let μ, ν be probability measures and

$$\mu(x,\infty) \sim x^{-\beta} L(x) \tag{3.49}$$

with $L(\cdot)$ a slowly varying function [Kołodziejek and Szpojankowski, 2022, Definition 1.1]. Assume furthermore that the $\lfloor \beta+1 \rfloor$-th moment of ν exists:

$$m_{\lfloor \beta+1 \rfloor}(\nu) < \infty.$$

Then

$$\mu \boxtimes \nu(x,\infty) \sim m_1^{\beta}(\nu)\mu(x,\infty)$$

with $m_1(\nu)$ the first moment of ν.

Since $\mu_W \boxtimes \mu_{sc}$ is a symmetric measure we have, using Kołodziejek and Szpojankowski [2022, equation (7.3)] and (3.48),

$$\mu_W \boxtimes \mu_{sc}(x,\infty) = \frac{1}{2}(\mu_W \boxtimes \mu_{sc})^2(x^2,\infty) = \frac{1}{2}\mu_W \boxtimes \mu_{sc}^2 \boxtimes \mu_W(x^2,\infty). \tag{3.50}$$

By the commutativity and associativity of the free multiplicative convolution [Nica and Speicher, 2006, Remark 14.2] we have $\mu_W \boxtimes \mu_{sc}^2 \boxtimes \mu_W = \mu_{sc}^2 \boxtimes \mu_W \boxtimes \mu_W$. Let $\nu_W \coloneqq \mu_W \boxtimes \mu_W$. Then a consequence of Kołodziejek and Szpojankowski [2022, Theorem 1.3(iv)] is that

$$\nu_W(x,\infty) \sim (m_1(\mu_W))^{\tau-1}\, \mu_W(x,\infty). \tag{3.51}$$

Therefore ν_W satisfies (3.49) with $\beta \coloneqq \tau - 1$, and clearly $m_{\lfloor \tau \rfloor}(\mu_{sc}^2) < \infty$. Thus, applying Kołodziejek and Szpojankowski [2022, Lemma 7.2],

$$\begin{aligned}
(\mu_{sc} \boxtimes \nu_W)(x,\infty) &\overset{(3.50)}{=} \frac{1}{2}\mu_W \boxtimes \mu_{sc}^2 \boxtimes \mu_W(x^2,\infty) \\
&\sim \frac{1}{2}\left(m_1(\mu_{sc}^2)\right)^{\tau-1} \nu_W(x^2,\infty) \\
&\overset{(3.51)}{\sim} \frac{1}{2}\left(m_1(\mu_{sc}^2)\right)^{\tau-1} (m_1(\mu_W))^{\tau-1}\, \mu_W(x^2,\infty) \\
&\sim \frac{1}{2}\left(m_1(\mu_{sc}^2)\right)^{\tau-1} (m_1(\mu_W))^{\tau-1}\, x^{-2(\tau-1)}.
\end{aligned}$$

We can conclude noting that $m_1(\mu_W)$ is finite since $\tau > 2$ and $m_1(\mu_{sc}^2) = m_2(\mu_{sc}) = 1$ [Arizmendi and Pérez-Abreu, 2009, Proposition 5 a)]. $\square$

Chapter Three

§3.6 Non-degeneracy of the limiting measure

The proof of Theorem 3.2.3 follows the arguments in Chakrabarty et al. [2016, Theorem 2.2]. A key observation is that the limiting measure $\mu_{\sigma,\tau}$ does not depend on the parameter α. This will allow us to deal with an easier model, formally corresponding to the case $\alpha = 0$, that does not feel the influence of the torus' geometry. The lack of geometry also allows us to work on a unique probability space. More precisely, let $(G_{i,j})_{i,j\geq 1}$ be an i.i.d. sequence of $\mathcal{N}(0,1)$ random variables, and let $(W_i)_{i\geq 1}$ be an i.i.d. sequence of Pareto-distributed random variables with parameter $\tau - 1$. Assume they are defined on the same probability space $(\Omega, \mathcal{F}, \mathbf{P})$. Define the $N \times N$ matrix

$$B_{N,m} = N^{-1/2}\sqrt{\kappa_\sigma(W_i^m, W_j^m)}\, G_{i\wedge j, i\vee j}\,.$$

Let $B_{N,\infty}$ denote the matrix with non-truncated weights. The following result can be proven exactly as in Proposition 3.4.9.

Proposition 3.6.1.
Let $\mathrm{ESD}(B_{N,m})$ *be the empirical spectral distribution of* $B_{N,m}$. *Then for all* $m \geq 1$,

$$\lim_{N\to\infty} \mathrm{ESD}(B_{N,m}) = \mu_{\sigma,\tau,m} \qquad \text{in } \mathbf{P}\text{-probability.}$$

Moreover,

$$\lim_{N\to\infty} \mathrm{ESD}(B_{N,\infty}) = \mu_{\sigma,\tau} \qquad \text{in } \mathbf{P}\text{-probability.}$$

We use this result to prove Theorem 3.2.3. Recall that, for a distribution function F, the generalised inverse is given by

$$F^{\leftarrow}(y) := \inf\{x \in \mathbb{R} : F(x) \geq y\}, \quad 0 < y < 1.$$

Proof of Theorem 3.2.3. From Proposition 3.6.1, it follows that there exists a subsequence $(N_k)_{k\geq 1}$ such that $\mu_{N_k,m}$ converges weakly almost surely to $\mu_{\sigma,\tau,m}$; that is,

$$\lim_{k\to\infty} d_L(\mathrm{ESD}(B_{N_k,m}), \mu_{\sigma,\tau,m}) = 0 \quad \mathbf{P}\text{-almost surely.} \tag{3.52}$$

For a $n \times n$ matrix A, let us denote by $\lambda_1(A) \leq \lambda_2(A) \leq \cdots \leq \lambda_n(A)$ its eigenvalues. For fixed integers $1 \leq k < \infty$, $1 < m < \infty$, define the following random variables on the probability space $(\Omega \times (0,1), \mathcal{F} \otimes \mathcal{B}(0,1), \mathsf{P} = \mathbf{P} \times \mathrm{Leb})$:

$$Z_{k,m}(\omega, x) = \lambda_{\lceil N_k x\rceil}\big(B_{N_k,m}(\omega)\big), \quad \omega \in \Omega,\ x \in (0,1),$$

and

$$Z_{k,\infty}(\omega, x) := \lambda_{\lceil N_k x \rceil}\big(B_{N_k,\infty}(\omega)\big), \quad \omega \in \Omega,\ x \in (0,1).$$

Let F_m be the distribution function of $\mu_{\sigma,\tau,m}$ (we suppress the dependence on σ and τ in F_m for ease of notation), and define

$$Z_{\infty,m}(\omega, x) := F_m^{\leftarrow}(x), \quad \omega \in \Omega,\ x \in (0,1).$$

Now consider $L^2(\Omega \times (0,1))$ with the P measure. This is a complete metric space, with $d(X,Y) = \mathsf{E}[(X-Y)^2]$. Our aim is to use Lemma 3.3.5 applied to the sequence of random variables $Z_{k,m}$. We proceed therefore to check assumptions (1) and (2) of the lemma. These will directly follow if we prove that

$$\lim_{k\to\infty} \mathsf{E}\left[(Z_{k,m} - Z_{\infty,m})^2\right] = 0 \tag{3.53}$$

and

$$\lim_{m\to\infty}\lim_{k\to\infty} \mathsf{E}\left[(Z_{k,m} - Z_{k,\infty})^2\right] = 0. \tag{3.54}$$

We start by (3.53). First of all we show that

$$\lim_{k\to\infty} Z_{k,m} = Z_{\infty,m} \quad \mathsf{P}\text{-almost surely.} \tag{3.55}$$

Define

$$\begin{aligned} A &:= A' \times (0,\,1) \\ &:= \left\{\omega \in \Omega : \lim_{k\to\infty} d_L(\mathrm{ESD}(B_{N_k,m}), \mu_{\sigma,\tau,m}) = 0,\ \forall m > 1\right\} \times (0,1)\,. \end{aligned}$$

Observe that $\mathsf{P}(A) = 1$ due to (3.52) and $\mathrm{Leb}(0,1) = 1$. To prove (3.55), it suffices to show that, for all $\omega \in A'$,

$$\lim_{k\to\infty} Z_{k,m}(\omega,x) = Z_{\infty,m}(\omega,x), \quad x \in (0,1). \tag{3.56}$$

Let $F_{k,m}(\omega,\cdot)$ be the distribution function of $\mathrm{ESD}(B_{N_k,m}(\omega))$. On A, we have $F_{k,m}(\omega,x) \to F_m(x)$ for all x at which F_m is continuous. Note that

$$Z_{k,m}(\omega,x) = F_{k,m}^{\leftarrow}(\omega,x).$$

It then follows from Resnick [2008, Proposition 0.1] that for all $x \in (0,1)$

$$\lim_{k\to\infty} F_{k,m}^{\leftarrow}(x) = F_m^{\leftarrow}(x).$$

Thus, we have proved (3.55).

Chapter Three

Next, we show that for all $m \geq 1$,

$$\{Z_{k,m}^2 : 1 \leq k < \infty\} \text{ is uniformly integrable.} \tag{3.57}$$

It suffices to show that $\sup_{k\geq 1} \mathsf{E}[Z_{k,m}^4] < \infty$. Since $\lceil N_k x \rceil$ is constant on intervals of length $1/N_k$, it easily follows that

$$\begin{aligned}\lim_{k\to\infty} \mathsf{E}[Z_{k,m}^4] &= \lim_{k\to\infty} \frac{1}{N_k} \mathbf{E}\left[\sum_{i=1}^{N_k} \lambda_i(B_{N_k,m})^4\right] \\ &= \lim_{k\to\infty} \frac{1}{N_k} \mathbf{E}\,\mathrm{Tr}(B_{N_k,m}^4) = \int_{\mathbb{R}} x^4\, \mu_{\sigma,\tau,m}(\mathrm{d}\,x) < \infty\end{aligned}$$

using (3.27) and (3.29), hence (3.57) is proven. Using this and (3.55), we obtain (3.53).

We move to (3.54). To prove this note that

$$\begin{aligned}\mathsf{E}\left[(Z_{k,m} - Z_{k,\infty})^2\right] &= \frac{1}{N_k}\mathbf{E}\left[\sum_{j=1}^{N_k} \big(\lambda_j(B_{N_k,m}) - \lambda_j(B_{N_k,\infty})\big)^2\right] \\ &\overset{(3.9)}{\leq} \frac{1}{N_k}\mathbf{E}\left[\mathrm{Tr}\left((B_{N_k,m} - B_{N_k,\infty})^2\right)\right] \\ &= \frac{1}{N_k}\mathbf{E}\left[\sum_{i,j=1}^{N_k} \big(B_{N_k,m}(i,j) - B_{N_k,\infty}(i,j)\big)^2\right].\end{aligned}$$

Reasoning as in the proof of Lemma 3.4.2, it follows that

$$\begin{aligned}&\frac{1}{N_k}\mathbf{E}\left[\sum_{i,j=1}^{N_k} (B_{N_k,m}(i,j) - B_{N_k,\infty}(i,j))^2\right] \\ &= \frac{1}{N_k^2}\sum_{i,j=1}^{N_k} \mathbf{E}\left[\left(\sqrt{\kappa_\sigma(W_i^m, W_j^m)} - \sqrt{\kappa_\sigma(W_i, W_j)}\right)^2\right] \\ &\leq \frac{2}{N_k^2}\sum_{i,j=1}^{N_k} \mathbf{E}\left[\kappa_\sigma(W_i, W_j)\mathbf{1}_{W_j<m<W_i}\right] \\ &+ \frac{2}{N_k^2}\sum_{i,j=1}^{N_k} \mathbf{E}\left[\kappa_\sigma(W_i, W_j)\mathbf{1}_{W_i\geq W_j>m}\right].\end{aligned}$$

We can use similar bounds as for Lemma 3.4.2, which yield that both summands have order at most $m^{2-\tau}$. Hence (3.54) follows, since $\tau > 2$.

Since we have now checked assumptions (1) and (2) of Lemma 3.3.5, it follows that there exists $Z_\infty \in L^2(\Omega \times (0,1))$ such that

$$\lim_{m\to\infty} \mathsf{E}\left[(Z_{\infty,m} - Z_\infty)^2\right] = 0.$$

Let U be a uniform random variable on $(0,1)$. Then $F_m^{\leftarrow}(U)$ has the same distribution as $\mu_{\sigma,\tau,m}$. Since $\mu_{\sigma,\tau,m}$ converges weakly to $\mu_{\sigma,\tau}$ by (3.47), Z_∞ has law $\mu_{\sigma,\tau}$. Hence

$$\lim_{m\to\infty} \mathsf{E}[Z_{\infty,m}^2] = \lim_{m\to\infty} \int_{\mathbb{R}} x^2\, \mu_{\sigma,\tau,m}(\mathrm{d}\,x) = \int_{\mathbb{R}} x^2\, \mu_{\sigma,\tau}(\mathrm{d}\,x),$$

and

$$\lim_{m\to\infty} \int_{\mathbb{R}} x^2\, \mu_{\sigma,\tau,m}(\mathrm{d}\,x) = (\tau-1)^2 \int_1^\infty \int_1^\infty \frac{1}{(x\wedge y)^{\tau-\sigma}(x\vee y)^{\tau-1}}\,\mathrm{d}\,x\,\mathrm{d}\,y$$

which can be easily obtained from (3.29) with $k=1$. This completes the proof of the first part.

Since $\lim_{m\to\infty} \mu_{\sigma,\tau,m} = \mu_{\sigma,\tau}$ weakly, we apply Fatou's lemma to obtain

$$\int x^{2p}\, \mu_{\sigma,\tau}(\mathrm{d}\,x) \le \liminf_{m\to\infty} \int x^{2p}\, \mu_{\sigma,\tau,m}(\mathrm{d}\,x) = \lim_{m\to\infty} M_{2p},$$

where, recalling (3.29),

$$M_{2p} = \sum_{\pi\in NC_2(2p)} \mathbf{E}\left[\prod_{(u,v)\in E(G_{\gamma\pi})} \kappa_\sigma(W_u^m, W_v^m)\right].$$

For $\sigma > 0$, we observe that $(x\wedge y)^\sigma (x\vee y) \le (xy)^{\sigma\vee 1}$. Thus,

$$M_{2p} \le \sum_{\pi\in NC_2(2p)} \prod_{i=1}^{p+1} \mathbb{E}\left[(W_i^m)^{(\sigma\vee 1)\# V_i}\right], \tag{3.58}$$

where $\{V_1,\dots,V_{p+1}\}$ are the blocks of $\gamma\pi$. Due to Lemma 3.4.8, it follows that $\max_{1\le i\le p+1} \# V_i \le p$, typically achieved by partitions π such that

$$\gamma\pi = \{(1,3,\dots,2p-1),(2),(4),\dots,(2p)\}.$$

This shows that the maximum moment bound required for the right-hand side of (3.58) to remain finite is $\mathbb{E}[(W_i)^{p(\sigma\vee 1)}]$. Since W_i has a tail index of $\tau-1$, if $p(\sigma\vee 1) < \tau - 1$, then $\mathbb{E}[(W_i)^{p(\sigma\vee 1)}] < \infty$. Therefore, M_{2p} is uniformly bounded in m, completing the proof of the theorem. □

§3.7 Absolute continuity and symmetry

We begin by showing absolute continuity. We shall use the following fact from Chakrabarty and Hazra [2016, Fact 2.1], which follows from Nica and Speicher [2006, Proposition 22.32].

Lemma 3.7.1.
Assume that, for each N, A_N is a $N \times N$ Gaussian Wigner matrix scaled by $\sqrt{N}$, that is, $(A_N(i,j) : 1 \le i \le j \le N)$ are i.i.d. normal random variables with mean zero and variance $1/N$, and $A_N(j,i) = A_N(i,j)$. Suppose that B_N is a $N \times N$ random matrix, such that for all $k \ge 1$

$$\lim_{N\to\infty} \frac{1}{N} \mathrm{Tr}\left(B_N^k\right) = \int_{\mathbb{R}} x^k \mu(\mathrm{d}\, x)$$

in probability, for some compactly supported (deterministic) probability measure μ. Furthermore, let the families $(A_N : N \ge 1)$ and $(B_N : N \ge 1)$ be independent. Then for all $k \ge 1$

$$\lim_{N\to\infty} \frac{1}{N} \mathrm{E}_{\mathcal{F}} \, \mathrm{Tr}\left[(A_N + B_N)^k\right] = \int_{\mathbb{R}} x^k \mu \boxplus \mu_{sc}(\mathrm{d}\, x)$$

in probability, where $\mathcal{F} := \sigma(B_N : N \ge 1)$ and $\mathrm{E}_{\mathcal{F}}$ denotes the conditional expectation with respect to $\mathcal{F}$.

Proof of Theorem 3.2.4. We consider the truncated weights $(W_i^m)_{i\ge1}$. Let Γ_m be an $N \times N$ matrix with entries given by

$$\Gamma_m(i,j) = \sqrt{\kappa_\sigma(W_i^m, W_j^m)}.$$

Given $\delta \in (0,1)$, define the function $g_{\delta,m}$ such that

$$g_{\delta,m}(W_i^m, W_j^m)^2 = \left(\sqrt{\kappa_\sigma(W_i^m, W_j^m)} - \delta\right)^2 + 2\delta\left(\sqrt{\kappa_\sigma(W_i^m, W_j^m)} - \delta\right).$$

As a consequence

$$g_{\delta,m}(W_i^m, W_j^m)^2 + \delta^2 = \kappa_\sigma(W_i^m, W_j^m)\,. \tag{3.59}$$

Define the matrix $\Gamma_{g_{\delta,m}}(i,j) = g_{\delta,m}(W_i^m, W_j^m)$. Let $\{G_{i,j}\}_{1\le i,j\le N}$ be i.i.d. standard Gaussian random variables, independent of the sequence $(W_i)_{i\ge1}$. Denote by $\mathfrak{G}_N$ the matrix with entries

$$\mathfrak{G}_N(i,j) = \frac{1}{\sqrt{N}} G_{i\wedge j, i\vee j}\,.$$

Define

$$\mathbf{B}^{(1)}_{N,m} = \Gamma_m \circ \mathfrak{G}_N\,.$$

Similarly, define

$$\mathbf{B}^{(2)}_{N,m} = \Gamma_{g_{\delta,m}} \circ \mathfrak{G}_N\,.$$

Lastly, consider a sequence of i.i.d. standard Gaussian random variables $(G'_{i,j})_{1\le i,j\le N}$, independent of the sigma field $\mathcal{F}$ generated by $(W_i)_{i\ge 1}, (G_{i,j})_{i,j\ge 1}$. Define a matrix $\mathbf{B}^{(3)}_{N,m}$ with entries

$$\mathbf{B}^{(3)}_{N,m}(i,j) = \frac{1}{\sqrt{N}} G'_{i\wedge j, i\vee j}\,.$$

We claim that, conditionally on $(W_i)_{i\in[N]}$,

$$\mathbf{B}^{(1)}_{N,m} \overset{d}{=} \mathbf{B}^{(2)}_{N,m} + \delta \mathbf{B}^{(3)}_{N,m}. \tag{3.60}$$

Indeed, conditionally on $(W_i)_{i\in[N]}$, the entries of $\mathbf{B}^{(1)}_{N,m}$, $\mathbf{B}^{(2)}_{N,m}$, and $\mathbf{B}^{(3)}_{N,m}$ are normally distributed. Thus, it is sufficient to compare the mean and variance of the entries. All the variables in question have mean zero and the variances match, too, due to (3.59). Following Proposition 3.6.1, there exists a measure $\mu_{g_{\delta,m}}$ such that

$$\lim_{N\to\infty} \frac{1}{N} \operatorname{Tr}\left((\mathbf{B}^{(2)}_{N,m})^k \right) = \int_{\mathbb{R}} x^k \, \mu_{g_{\delta,m}}(\mathrm{d}\, x)$$

in probability. In particular, we recall the expression for the even moments of $\mu_{g_{\delta,m}}$ given in (3.29):

$$M_{2k} = \sum_{\pi\in NC_2(2k)} \mathbb{E}\left[\prod_{(u,v)\in E(G_{\gamma\pi})} g^2_{\delta,m}(W^m_u, W^m_v) \right].$$

Since $g^2_{\delta,m}(W^m_u, W^m_v) \le \kappa_\sigma(W^m_u, W^m_v)$, it follows that $\mu_{g_{\delta,m}}$ is uniquely determined by its moments, and is also compactly supported (Corollary 3.4.11). This verifies the first condition of Lemma 3.7.1. Since $\mathbf{B}^{(3)}_{N,m}$ is a standard Wigner matrix, it follows from Lemma 3.7.1 that

$$\lim_{N\to\infty} \frac{1}{N} \mathbb{E}_{\mathcal{F}} \left[\operatorname{Tr}\left((\mathbf{B}^{(2)}_{N,m} + \delta \mathbf{B}^{(3)}_{N,m})^k \right) \right] = \int_{\mathbb{R}} x^k \, (\mu_{g_{\delta,m}} \boxplus \mu_{sc,\delta})(\mathrm{d}\, x),$$

where $\mu_{sc,\delta}$ is the semicircular law with variance δ^2 and density

$$\mu_{sc,\delta}(\mathrm{d}\, x) = \frac{1}{2\pi\delta} \sqrt{4 - \left(\frac{x}{\delta}\right)^2}\, \mathbf{1}_{|x|\le 2\delta}\, \mathrm{d}\, x, \quad x\in\mathbb{R}.$$

Since both $\mu_{g_{\delta,m}}$ and $\mu_{sc,\delta}$ are compactly supported, so is $\mu_{g_{\delta,m}} \boxplus \mu_{sc,\delta}$, and thus the measure is completely determined by its moments.

From Proposition 3.4.9 we have

$$\lim_{N\to\infty} \mathbb{E}\left[\frac{1}{N}\mathbb{E}_{\mathcal{F}}[\mathrm{Tr}(\mathbf{B}_{N,m}^{(1)})^k]\right] = \int_{\mathbb{R}} x^k \, \mu_{\sigma,\tau,m}(\mathrm{d}\, x)$$

and

$$\lim_{N\to\infty} \mathrm{Var}\left(\frac{1}{N}\mathbb{E}_{\mathcal{F}}[\mathrm{Tr}((\mathbf{B}_{N,m}^{(1)})^k)]\right) \le \lim_{N\to\infty} \mathrm{Var}\left(\frac{1}{N}\,\mathrm{Tr}((\mathbf{B}_{N,m}^{(1)})^k)\right) = 0.$$

Thus,

$$\lim_{N\to\infty} \frac{1}{N}\mathbb{E}_{\mathcal{F}}\left[\mathrm{Tr}(\mathbf{B}_{N,m}^{(1)})^k\right] = \int_{\mathbb{R}} x^k \, \mu_{\sigma,\tau,m}(\mathrm{d}\, x)$$

in probability. Since the measures are uniquely determined by their moments, this shows that

$$\mu_{\sigma,\tau,m} = \mu_{g_{\delta,m}} \boxplus \mu_{sc,\delta}. \tag{3.61}$$

We show that there exists μ_{g_δ} such that

$$\lim_{m\to\infty} d_L(\mu_{g_{\delta,m}}, \mu_{g_\delta}) = 0. \tag{3.62}$$

If we can prove this, using Bercovici and Voiculescu [1993, Proposition 4.13] it will follow that

$$\lim_{m\to\infty} d_L(\mu_{g_{\delta,m}} \boxplus \mu_{sc,\delta}, \mu_{g_\delta} \boxplus \mu_{sc,\delta}) \le \lim_{m\to\infty} d_L(\mu_{g_{\delta,m}}, \mu_{g_\delta}) = 0. \tag{3.63}$$

To show (3.62), we employ Lemma 3.3.5. Note that, from Remark 3.4.11, we get that for any fixed $m \ge 1$ one has

$$\lim_{N\to\infty} d_L\big(\mu_{\mathbf{B}_{N,m}^{(2)}}, \mu_{g_{\delta,m}}\big) = 0 \quad \text{ in } \mathbb{P}\text{-probability}$$

where $\mu_{\mathbf{B}_{N,m}^{(2)}}$ is the empirical spectral distribution of $\mathbf{B}_{N,m}^{(2)}$.

This establishes condition (1) of Lemma 3.3.5. To complete the proof, we need to verify condition (2), namely,

$$\lim_{m\to\infty} \limsup_{N\to\infty} \mathbb{P}\big(d_L(\mathrm{ESD}(\mathbf{B}_{N,m}^{(2)}), \mathrm{ESD}(\mathbf{B}_N^{(2)})) > \varepsilon\big) = 0. \tag{3.64}$$

Here $\mathbf{B}_N^{(2)}$ is defined as $\mathbf{B}_{N,\infty}^{(2)}$ with $m = \infty$. From Proposition 3.3.1 we see that

$$\begin{aligned} d_L\left(\mathrm{ESD}(\mathbf{B}_{N,m}^{(2)}), \mathrm{ESD}(\mathbf{B}_N^{(2)})\right)^3 &\le \frac{1}{N}\,\mathrm{Tr}\left((\mathbf{B}_{N,m}^{(2)} - \mathbf{B}_N)^2\right) \\ &= \frac{1}{N^2}\sum_{i,j=1}^{N} \left(\Gamma_{g_{\delta,m}}(i,j) - \Gamma_{g_{\delta,\infty}}(i,j)\right)^2 G_{i\wedge j, i\vee j}^2 \,. \end{aligned}$$

Hence we have

$$\mathbb{E}\left[d_L\big(\operatorname{ESD}(\mathbf{B}^{(2)}_{N,m}),\operatorname{ESD}(\mathbf{B}^{(2)}_{N})\big)^3\right] \le \frac{1}{N^2}\sum_{i\neq j=1}^{N}\mathbf{E}[\big(\Gamma_{g_{\delta,m}}(i,j)-\Gamma_{g_{\delta,\infty}}(i,j)\big)^2]$$
$$\le \frac{2}{N^2}\sum_{i\neq j=1}^{N}\mathbf{E}\left[g_{\delta,\infty}(W_i,\,W_j)^2\left(\mathbf{1}_{W_j<m<W_i}+\mathbf{1}_{W_i>W_j>m}\right)\right]$$
$$\le \frac{2}{N^2}\sum_{i\neq j=1}^{N}\mathbf{E}\left[\kappa_\sigma(W_i,W_j)\left(\mathbf{1}_{W_j<m<W_i}+\mathbf{1}_{W_i>W_j>m}\right)\right].$$

Just as in the proof of (3.54), it follows that the last term is bounded by $Cm^{2-\tau}$. Thus, using Markov's inequality, condition (2) of Lemma 3.3.5 holds, too. In conclusion, we can show that there exists μ_{g_δ} such that

$$\lim_{m\to\infty} d_L(\mu_{g_{\delta,m}}\boxplus\mu_{sc,\delta},\mu_{\sigma,\tau}) \overset{(3.61)}{=} \lim_{m\to\infty} d_L(\mu_{\sigma,\tau,m},\mu_{\sigma,\tau}) \overset{(3.47)}{=} 0$$
$$\overset{(3.63)}{=} \lim_{m\to\infty} d_L(\mu_{g_{\delta,m}}\boxplus\mu_{sc,\delta},\mu_{g_\delta}\boxplus\mu_{sc,\delta}).$$

Therefore it must be that $\mu_{\sigma,\tau}=\mu_{g_\delta}\boxplus\mu_{sc,\delta}$. The right-hand side is absolutely continuous, as shown by Biane [1997, Corollary 2].

Finally, to show symmetry, we see that $\mu_{\sigma,\tau}$ does not give weight to singletons by absolute continuity. Therefore, in light of the weak convergence stated in (3.47),

$$\mu_{\sigma,\,\tau}(-\infty,\,-x) = \lim_{m\to\infty}\mu_{\sigma,\,\tau,\,m}(-\infty,\,-x)$$
$$= \lim_{m\to\infty}\mu_{\sigma,\,\tau,\,m}(x,\,+\infty) = \mu_{\sigma,\,\tau}(x,\,+\infty)$$

for all $x\ge 0$. This completes the proof. □

§3.8 Stieltjes transform of the limiting measure

To prove Theorem 3.2.5, we first identify the Stieltjes transform for the measure $\mu_{\sigma,\tau,m}$. We then proceed to take the limit $m\to\infty$, which requires a functional analytic approach. Throughout this section, we fix $z\in\mathbb{C}^+$, given as $z=\xi+i\eta$ with $\eta>0$. If μ is a probability measure having all its moments $\{m_k\}_{k\ge1}$, it follows from the definition of Stieltjes transform (3.7) that, for any $z\in\mathbb{C}^+$,

$$\mathrm{S}_\mu(z) = -\sum_{k\ge0}\frac{m_k}{z^{k+1}}, \tag{3.65}$$

where the Laurent series on the right-hand side of (3.65) converges for $|z|>R>0$, with $\operatorname{supp}(\mu)=[-R,R]$.

§3.8.1 Stieltjes transform for truncated weights

To derive a characterisation of the limiting measure $\mu_{\sigma,\tau}$, we need to first study the truncated version $\mu_{\sigma,\tau,m}$. We borrow ideas from the proof of Chakrabarty et al. [2015, Theorem 4.1]. The main result of this subsection will be Proposition 3.8.1, which requires a few technical lemmas to prove. The results in this subsection hold for the regime $\tau > 2$ and $\sigma < \tau - 1$, as before.

We have that the (even) moments for the measure $\mu_{\sigma,\tau,m}$ are given by (3.29). Using these, we derive a representation of $\mathrm{S}_{\mu_{\sigma,\tau,m}}(z)$.

Proposition 3.8.1.
For $\tau > 2$ and $\sigma \in (0, \tau - 1)$ there exists a function $a(z, x) = a_m(z, x)$ defined on $\mathbb{C}^+ \times [1, \infty)$ such that

$$\mathrm{S}_{\mu_{\sigma,\tau,m}}(z) = \int_1^\infty a(z, x)\mu_{W,m}(\mathrm{d}\, x)\,,$$

where $\mu_{W,m}$ is the law of the truncated weights (W_i^m). Moreover, $a(z, x)$ satisfies the following recursive equation:

$$a(z, x)\left(z + \int_1^\infty a(z, y)\kappa_\sigma(x, y)\mu_{W,m}(\mathrm{d}\, y)\right) = -1\,. \tag{3.66}$$

Before tackling the proof of the proposition, we lay the ground with two auxiliary results. For any $k \geq 1$ and $\pi \in NC_2(2k)$, recall the map $\mathcal{T}_\pi$ of (3.26), where $\gamma\pi = \{V_1, \ldots, V_{k+1}\}$. Consider the mapping $L_\pi : [1, \infty)^{k+1} \to \mathbb{R}$ defined as

$$L_\pi(\mathbf{x}) = \kappa_\sigma^{1/2}(x_{\mathcal{T}_\pi(1)}, x_{\mathcal{T}_\pi(2)})\kappa_\sigma^{1/2}(x_{\mathcal{T}_\pi(2)}, x_{\mathcal{T}_\pi(3)}) \ldots \kappa_\sigma^{1/2}(x_{\mathcal{T}_\pi(2k)}, x_{\mathcal{T}_\pi(1)}) \tag{3.67}$$

and the function $H_\pi : \mathbb{R} \to \mathbb{R}^+$ given as

$$H_\pi(y) = \int_{[1,\infty)^k} L_\pi(y,\, x_2, \ldots,\, x_{k+1})\mu_{W,m}^{\otimes k}(\mathrm{d}\, \mathbf{x}')\,, \tag{3.68}$$

where we are integrating over $\mathbf{x}' = (x_2, \ldots, x_{k+1}) \in [1, \infty)^k$.

Lemma 3.8.2.
Let $\{M_{2k}\}_{k\geq 1}$ be as in (3.29). Then

$$M_{2k} = \sum_{\pi \in NC_2(2k)} \int_1^\infty H_\pi(y)\mu_{W,m}(\mathrm{d}\, y).$$

Proof of Lemma 3.8.2. We begin by evaluating the integral on the right-hand side. We have

$$\begin{aligned}
&\int_1^\infty H_\pi(y)\mu_{W,m}(\mathrm{d}\,y) \\
&= \int_1^\infty \int_{[1,\infty)^k} L_\pi(y, x_2, \ldots, x_{k+1})\mu_{W,m}^{\otimes k}(\mathrm{d}\,\mathbf{x}')\mu_{W,m}(\mathrm{d}\,y) \\
&= \int_{[1,\infty)^{k+1}} \kappa_\sigma^{1/2}(x_{\mathcal{T}_\pi(1)}, x_{\mathcal{T}_\pi(2)}) \cdots \kappa_\sigma^{1/2}(x_{\mathcal{T}_\pi(2k)}, x_{\mathcal{T}_\pi(1)})\mu_{W,m}^{\otimes k+1}(\mathrm{d}\,\mathbf{x}).
\end{aligned}$$

We know that, for $\pi \in NC_2(2k)$, $\#\gamma\pi = k+1$ and so the graph $G_{\gamma\pi}$ has $k+1$ vertices. Furthermore, when we perform a closed walk of the form $1 \to 2 \to \ldots \to 2k \to 1$ on the (unoriented) graph $G_{\gamma\pi}$, we traverse each edge *exactly* twice. In particular, the product $\kappa_\sigma^{1/2}(x_{\mathcal{T}_\pi(1)}, x_{\mathcal{T}_\pi(2)}) \cdots \kappa_\sigma^{1/2}(x_{\mathcal{T}_\pi(2k)}, x_{\mathcal{T}_\pi(1)})$ has $2k$ terms with k matchings, and so

$$\kappa_\sigma^{1/2}(x_{\mathcal{T}_\pi(1)}, x_{\mathcal{T}_\pi(2)}) \cdots \kappa_\sigma^{1/2}(x_{\mathcal{T}_\pi(2k)}, x_{\mathcal{T}_\pi(1)}) = \prod_{(u,v)\in E(G_{\gamma\pi})} \kappa_\sigma(x_u, x_v)\,.$$

We then have that

$$\begin{aligned}
\int_1^\infty H_\pi(y)\mu_{W,m}(\mathrm{d}\,y) &= \int_{[1,\infty)^{k+1}} \prod_{(u,v)\in E(G_{\gamma\pi})} \kappa_\sigma(x_u, x_v)\mu_{W,m}^{\otimes k+1}(\mathrm{d}\,\mathbf{x}) \\
&= \mathbb{E}\left[\prod_{(u,v)\in E(G_{\gamma\pi})} \kappa_\sigma(W_u^m, W_v^m)\right],
\end{aligned}$$

which concludes the proof. $\square$

We show now some properties of H_π that will help us in the upcoming computations.

Lemma 3.8.3.
Let $k \geq 1$ and let H_π be as defined in (3.68)*. Let $\pi \in NC_2(2k)$. Then,*

(1) If $\pi = (1, 2k) \cup \pi_1$, where π_1 is a non-crossing pair partition of $\{2, \ldots, 2k-1\}$, then,

$$H_\pi(y) = \int_1^\infty H_{\pi_1}(x)\kappa_\sigma(x, y)\mu_{W,m}(\mathrm{d}\,x). \tag{3.69}$$

(2) If $\pi = \pi_1 \cup \pi_2$, then $H_\pi(\cdot) = H_{\pi_1}(\cdot)H_{\pi_2}(\cdot)$.

Proof of Lemma 3.8.3. We first prove property (1). Let $\pi = (1, 2k) \cup \pi_1$. Then, $\gamma\pi = \{(1), V_2, \dots, V_{k+1}\}$. We know that $2 \in V_2$ and then $\gamma\pi(2k) = 2 \in V_2$. Now, fix $x_1 = y$. Then

$$
\begin{aligned}
&H_\pi(y) \\
&= \int_{[1,\infty)^k} L_\pi(y, x_2, \dots, x_{k+1}) \mu_{W,m}^{\otimes k}(\mathrm{d}\,\mathbf{x}') \\
&= \int_{[1,\infty)^k} \kappa_\sigma^{1/2}(y, x_2)\kappa_\sigma^{1/2}(x_2, x_{\mathcal{T}_\pi(3)}) \dots \kappa_\sigma^{1/2}(x_{\mathcal{T}_\pi(2k-1)}, x_2)\kappa_\sigma^{1/2}(x_2, y)\mu_{W,m}^{\otimes k}(\mathrm{d}\,\mathbf{x}') \\
&= \int_1^\infty \kappa_\sigma(y, x_2) \\
&\qquad \times \int_{[1,\infty)^{k-1}} \kappa_\sigma^{1/2}(x_2, x_{\mathcal{T}_\pi(3)}) \dots \kappa_\sigma^{1/2}(x_{\mathcal{T}_\pi(2k-1)}, x_2)\mu_{W,m}^{\otimes k-1}(\mathrm{d}\,\mathbf{x}'')\mu_{W,m}(\mathrm{d}\,x_2) \\
&= \int_1^\infty \kappa_\sigma(y, x_2) H_{\pi_1}(x_2)\mu_{W,m}(\mathrm{d}\,x_2),
\end{aligned}
$$

which is what we desired.

For property (2), let $\pi = \pi_1 \cup \pi_2$, with $\pi_1 \in NC_2(\{1, 2, \dots, 2r\})$ and $\pi_2 \in NC_2(\{2r+1, \dots, 2k\})$ and let us consider the function $H_\pi(y)$ with $y = x_1 = x_{\mathcal{T}_\pi(1)}$. Then,

$$
H_\pi(y) = \int_{[1,\infty)^k} \kappa_\sigma^{1/2}(y, x_{\mathcal{T}_\pi(2)}) \dots \kappa_\sigma^{1/2}(x_{\mathcal{T}_\pi(2r)}, x_{\mathcal{T}_\pi(2r+1)}) \dots \kappa_\sigma^{1/2}(x_{\mathcal{T}_\pi(2k)}, y)\mu_{W,m}^{\otimes k}(\mathrm{d}\,\mathbf{x}').
$$

We now claim that this integral can be split up into two integrals. First, consider the element $x_{\mathcal{T}_\pi(1)}$. Since we assume that '1' maps to $V_1 \in \gamma\pi$, all elements of V_1 are mapped to y. To understand where other elements are mapped, we will state a claim and see its consequences to this proof, and then prove it on 149.

Claim 3.8.4.
Under $\gamma\pi$, the elements $\{2, \dots, 2r\}$ are mapped to the blocks

$$
V_1 \cup \{V_2, \dots, V_{r'}\} \subset \gamma\pi\,,
$$

and the elements $\{2r+1, \dots, 2k\}$ are mapped to the blocks

$$
V_1 \cup \{V_{r'+1}, \dots, V_{k+1}\} \subset \gamma\pi\,,
$$

where $r' < k+1$ is some index. In particular $\gamma\pi(2r+1) \in V_1$.

From this claim we have that

$$
\begin{aligned}
&H_\pi(y)\\
&= \int_{[1,\infty)^k} \kappa_\sigma^{1/2}(y, x_{\mathcal{T}_\pi(2)}) \dots \kappa_\sigma^{1/2}(x_{\mathcal{T}_\pi(2r)}, x_{\mathcal{T}_\pi(2r+1)}) \dots \kappa_\sigma^{1/2}(x_{\mathcal{T}_\pi(2k)}, y) \mu_{W,m}^{\otimes k}(\mathrm{d}\,\mathbf{x}')\\
&= \int_{[1,\infty)^k} \kappa_\sigma^{1/2}(y, x_{\mathcal{T}_\pi(2)}) \dots \kappa_\sigma^{1/2}(x_{\mathcal{T}_\pi(2r)}, y) \dots \kappa_\sigma^{1/2}(x_{\mathcal{T}_\pi(2k)}, y) \mu_{W,m}^{\otimes k}(\mathrm{d}\,\mathbf{x}')\\
&= \int_{[1,\infty)^{r'}} \kappa_\sigma^{1/2}(y, x_{\mathcal{T}_\pi(2)}) \dots \kappa_\sigma^{1/2}(x_{\mathcal{T}_\pi(2r)}, y) \mu_{W,m}^{\otimes r'}(\mathrm{d}\,\mathbf{x}^{(r')})\\
&\times \int_{[1,\infty)^{k-r'}} \kappa_\sigma^{1/2}(y, x_{\mathcal{T}_\pi(2r+2)}) \dots \kappa_\sigma^{1/2}(x_{\mathcal{T}_\pi(2k)}, y) \mu_{W,m}^{\otimes (k-r')}(\mathrm{d}\,\mathbf{x}^{(k-r')})\\
&= H_{\pi_1}(y) H_{\pi_2}(y).
\end{aligned}
$$

This concludes the proof. $\square$

Proof of Claim 3.8.4. Let γ_1 resp. γ_2 be the shift by one on $[2r]$ resp. $\{2r+1, \dots, 2k\}$. To prove this claim, it suffices to analyse the special indices $\{1, 2r, 2r+1, 2k\}$, since γ_1 and γ_2 are cyclic permutations on $[2r]$ and $\{2r+1, \dots, 2k\}$, respectively. We will be using the fact that all elements in a block of $\gamma\pi$ must be either all odd or all even [Avena et al., 2023, Property 1], and that any pairing in π must have one element odd and the other even [Avena et al., 2023, Property 2].

(a) We already have $1 \in V_1$. Now, let $(o_1, 2k) \in \pi_2$, for some o_1 such that $o_1 \geq 2r+1$. Then, o_1 must be odd. Now, $o_1 + 1$ is even, and cannot belong to V_1. Thus $\gamma\pi(2k) = o_1 + 1 \in \{V_{r'+1}, \dots, V_{2k}\}$. This takes care of the index $2k$.

(b) Let us continue with $(o_2, 2r) \in \pi_1$ for some o_2. We know that o_2 must be odd. Thus, $\gamma\pi(2r) = o_2 + 1 \in \{V_2, \dots, V_{r'}\} =: \gamma_1\pi_1 \setminus V_1$. This resolves the case of $2r$.

(c) Lastly, by construction, $\gamma\pi(o_2) = 2r+1$, which brings us to the last special element. Since o_2 and $2r+1$ belong to the same block in $\gamma\pi$, it suffices to show that this block is V_1, that is, the block to which element 1 belongs. Now, if $(1, o_2 - 1) \in \pi_1$, we are done, since $\gamma\pi(1) = o_2$. Suppose not, and let $(1, e_1) \in \pi_1$ for some even integer e_1. Similarly as before, if now $(e_1 + 1, o_2 - 1) \in \pi_1$, we are done. Since π_1 and π_2 act on the first $2r$ elements and the remaining $2k - 2r$ elements respectively, then, by the non-crossing nature, there is a sequence of even integers $\{e_i\}_{i=1}^t$ such that $(1, e_1), (e_1 + 1, e_2), \dots, (e_t + 1, o_2 - 1) \in \pi_1$. Computing $\gamma\pi$ recursively gives us that $\gamma\pi(1) = o_2$, and so $\gamma\pi(2r+1) \in V_1$.

This proves the claim. $\square$

We are now ready to prove Proposition 3.8.1.

Proof of Proposition 3.8.1. We now derive the Stieltjes transform of the measure $\mu_{\sigma,\tau,m}$. Using (3.65) and Proposition 3.4.9, we have that

$$\mathrm{S}_{\mu_{\sigma,\tau,m}}(z) = -\sum_{k\geq 0} \frac{M_{2k}}{z^{2k+1}}.$$

Using Lemma 3.8.2 we substitute the expression for M_{2k}. We have

$$\begin{aligned} \mathrm{S}_{\mu_{\sigma,\tau,m}}(z) &= -\sum_{k\geq 0} \frac{1}{z^{2k+1}} \int_1^\infty \sum_{\pi\in NC_2(2k)} H_\pi(x)\mu_{W,m}(\mathrm{d}\,x) \\ &= -\int_1^\infty \sum_{k\geq 0} \sum_{\pi\in NC_2(2k)} \frac{H_\pi(x)}{z^{2k+1}} \mu_{W,m}(\mathrm{d}\,x), \end{aligned} \tag{3.70}$$

where we could interchange the integral and the sum by Fubini's theorem. Now, we define the function $a(z,x)$ as

$$a(z,x) := -\sum_{k\geq 0} \sum_{\pi\in NC_2(2k)} \frac{H_\pi(x)}{z^{2k+1}}. \tag{3.71}$$

Then using (3.70) we have

$$\mathrm{S}_{\mu_{\sigma,\tau,m}}(z) = \int_1^\infty a(z,x)\mu_{W,m}(\mathrm{d}\,x).$$

We now state some properties of the function $a(z,x)$. Firstly, for any $z\in\mathbb{C}^+$ the map $x\mapsto a(z,x)$ is in $L^\infty([1,\infty),\mu_{W,m})$ as H_π is bounded . Secondly, for any $x\in[1,\infty)$, the map $z\mapsto a(z,x)$ is analytic in $\mathbb{C}$, which follows from the Laurent series expansion. Finally we see that $a(z,x)$ lies in $\mathbb{C}^+$, for any $z\in\mathbb{C}^+$ and $x>1$. Indeed, for any $\Im(z)>0$, the expansion on the right-hand side of (3.71) will always have a non-trivial imaginary part. Thus, since $a(\cdot,\cdot)$ is analytic, it will either lie completely in $\mathbb{C}^-$ or $\mathbb{C}^+$, since it can never take values in $\mathbb{R}$. However, $\mathrm{S}_{\mu_{W,m}}(z)\in\mathbb{C}^+$, and thus, $a(z,x)\in\mathbb{C}^+$ for any $z\in\mathbb{C}^+$ and $x>1$.

To write down a functional recursion for $a(\cdot,\cdot)$ it is convenient to use the notion of words. Any partition π can be associated to a *word* w, with any elements in $i,j\in[2k]$ being associated with the same letter in w if i,j are in the same block of π. For example, $\pi=\{\{1,2\},\{3,4\}\}$ can be written as $w=aabb$. In particular, any partition $\pi\in NC_2(2k)$ can be associated to a word w of the form $w = aw_1aw_2$, where w_1, w_2 are words that can be empty. For any word w associated to a partition π, let $H_\pi = H_w$. Furthermore, for $w \in NC_2(2k)$ we mean a word w whose associated partition π is in $NC_2(2k)$. Then we have, using Lemma 3.8.3 in the third equality,

$$\begin{aligned} a(z,x) &= -\sum_{k\geq 0}\sum_{w\in NC_2(2k)} \frac{H_w(x)}{z^{2k+1}} \\ &= -\frac{1}{z} - \sum_{k\geq 1}\sum_{\substack{w\in NC_2(2k)\\ w=aw_1aw_2}} \frac{H_{aw_1aw_2}(x)}{z^{2k+1}} \\ &= -\frac{1}{z} - \sum_{k\geq 1}\sum_{\substack{w\in NC_2(2k)\\ w=aw_1aw_2}} \frac{H_{aw_1a}(x)H_{w_2}(x)}{z^{2k+1}} \\ &= -\frac{1}{z} - \frac{1}{z}\sum_{k\geq 1}\sum_{\ell=1}^{k}\sum_{w_1\in NC_2(2\ell-2)} \frac{H_{aw_1a}(x)}{z^{2\ell-2+1}} \sum_{w_2\in NC_2(2k-2\ell)} \frac{H_{w_2}(x)}{z^{2k-2\ell+1}}. \end{aligned} \tag{3.72}$$

One can see that the word aw_1a has as corresponding partition $(1,2\ell)\cup\pi_1$, with $\pi_1 \in NC_2(2\ell-2)$. Using (3.69) from Lemma 3.8.3, we have

$$\begin{aligned} &a(z,x) \\ &= -\frac{1}{z} - \frac{1}{z}\sum_{k\geq 1}\sum_{\ell=1}^{k}\sum_{w_1\in NC_2(2\ell-2)} \frac{1}{z^{2\ell-1}}\int_1^\infty H_{w_1}(y)\kappa_\sigma(x,y)\mu_{W,m}(\mathrm{d}\,y) \\ &\qquad\qquad \times \sum_{w_2\in NC_2(2k-2\ell)} \frac{H_{w_2}(x)}{z^{2k-2\ell+1}} \\ &= -\frac{1}{z} - \frac{1}{z}\sum_{k\geq 1}\sum_{\ell=1}^{k}\sum_{\pi_2\in NC_2(2k-2\ell)} \frac{H_{\pi_2}(x)}{z^{2k-2\ell+1}} \\ &\qquad\qquad \times \sum_{\pi_1\in NC_2(2\ell-2)} \frac{1}{z^{2\ell-1}}\int_1^\infty H_{\pi_1}(y)\kappa_\sigma(x,y)\mu_{W,m}(\mathrm{d}\,y) \\ &= -\frac{1}{z} - \frac{a(z,x)}{z}\int_1^\infty a(z,y)\kappa_\sigma(x,y)\mu_{W,m}(\mathrm{d}\,y). \end{aligned}$$

Thus, we have (3.66), which completes the proof of Proposition 3.8.1. □

Remark 3.8.5.
Equation (3.66) gives an analytic description of a in terms of the recursive equation. Now, for any $z \in \mathbb{C}^+$, we have that

$$z = i\int_0^\infty \mathrm{e}^{-itz^{-1}}\,\mathrm{d}\,t. \tag{3.73}$$

Since $a(z,x) \in \mathbb{C}^+$ for any fixed $x \in [1,\infty)$, applying (3.73) *to $a(z,x)$ and using* (3.66) *gives us that*

$$a(z,x) = i\int_0^\infty \mathrm{e}^{itz} \exp\left\{it\int_1^\infty a(z,y)\kappa_\sigma(x,y)\mu_{W,m}(\mathrm{d}\,y)\right\}\mathrm{d}\,t. \tag{3.74}$$

An immediate consequence of (3.74) *is that $a(z,x)$ is uniformly bounded in x and m. Indeed, if we take $z = \xi + i\eta$ with $\eta > 0$, we have that*

$$\begin{aligned} |a(z,x)| &\le \int_0^\infty \mathrm{e}^{-\eta t}\left|\exp\left\{it\int_1^\infty a(z,y)\kappa_\sigma(x,y)\mu_{W,m}(\mathrm{d}\,y)\right\}\right|\mathrm{d}\,t \\ &\le \int_0^\infty \mathrm{e}^{-\eta t}\,\mathrm{d}\,t = \frac{1}{\eta}. \end{aligned} \tag{3.75}$$

The bound in the second line holds since $a(z,x) \in \mathbb{C}^+$, and so

$$\int_1^\infty a(z,y)\kappa_\sigma(x,y)\mu_{W,m}(\mathrm{d}\,y) \in \mathbb{C}^+$$

as $\kappa_\sigma \ge 1$.

§3.8.2 Limiting Stieltjes transform

We now set up the framework required to prove Theorem 3.2.5. For the remainder of this section, denote $a_z(x) := a(z,x)$, which implicitly depends on m. We wish to extend Proposition 3.8.1 to the measure $\mu_{\sigma,\tau}$ by passing to the limit $m \to \infty$. We have a natural candidate for the function a^* in Theorem 3.2.5, which should be the limit of $a(\cdot,\cdot)$ as m tends to infinity. We now formalise this idea through a series of lemmas.

Since our goal now is to show Theorem 3.2.5 we are going to work for the remainder of this section with the following parameters:

(a) $\tau > 3$,

(b) $\sigma < \tau - 2$, and

(c) a parameter β such that $2 \vee 1 + \sigma < \beta < \tau - 1$.

Let $\overline{\mathbb{C}}^+ = \mathbb{C}^+ \cup \mathbb{R}$ be the closure of $\mathbb{C}^+$, and let ν be the measure defined as

$$\nu(\mathrm{d}\,x) = x^{-\beta}\,\mathrm{d}\,x. \tag{3.76}$$

Consider the space $L^1([1,\infty),\nu)$ of all functions $f : [1,\infty) \to \overline{\mathbb{C}}^+$ that are L^1−integrable with respect to ν.

Definition 3.8.6.
Let $\mathcal{B}$ denote the Banach space $\mathcal{B} := (L^1([1,\infty),\nu), \|\cdot\|_1)$, where the norm $\|\cdot\|_1$ is the L^1 norm with respect to ν as in (3.76), which is defined for $f \in L^1([1,\infty),\nu)$ as

$$\|f\|_1 := \int_1^\infty |f(x)| x^{-\beta} \,\mathrm{d}\, x. \tag{3.77}$$

Recall that $\mu_{W,m}$ denotes the law of the truncated weights $(W_x^m)_x$, given as

$$\mu_{W,m}(\cdot) = c_m^{-1} \mu_w(\cdot) \mathbb{1}_{\{\cdot \le m\}},$$

where $c_m = 1 - m^{-(\tau-1)}$ is a normalizing constant converging to 1 as m tends to infinity, and μ_W is the Pareto law defined in (3.2). For $z \in \mathbb{C}^+$, let T_z denote the map

$$T_z f(\cdot) = i \int_0^\infty \mathrm{e}^{itz} \exp\left\{ it \int_1^\infty f(y) \kappa_\sigma(\cdot, y) \mu_W(\mathrm{d}\, y) \right\} \mathrm{d}\, t. \tag{3.78}$$

Then, we have the following result.

Lemma 3.8.7.
There exists a constant $\tilde{c} = \tilde{c}(\tau, \sigma, \beta)$ such that, for all $z \in \mathbb{C}^+$ with $\Im(z)^2 = \eta^2 > \tilde{c}$ $T_z : \mathcal{B} \to \mathcal{B}$ is a contraction mapping, with a contraction constant $\tilde{c}\eta^{-2}$.

Proof of Lemma 3.8.7. We first need to show that, for any $f \in \mathcal{B}$, one has $T_z f \in \mathcal{B}$. Indeed, for $x \ge 1$ it holds that

$$\left| T_z f(x) \right| \le \int_0^\infty \mathrm{e}^{-\eta t} \left| \exp\left\{ it \int_1^\infty f(y) \kappa_\sigma(x, y) \mu_W(\mathrm{d}\, y) \right\} \right| \mathrm{d}\, t \le \frac{1}{\eta},$$

where the last inequality holds as $f(y) \in \overline{\mathbb{C}^+}$ for any $y \ge 1$, and thus the second complex exponential is bounded by 1. Since $|T_z f(\cdot)|$ is uniformly bounded, it is L^1–integrable with respect to ν, and so $T_z(\mathcal{B}) \subseteq \mathcal{B}$.

Now, we wish to show T_z is a contraction. Let us take $f_1, f_2 \in \mathcal{B}$. Recall that for any $z_1, z_2 \in \overline{\mathbb{C}^+}$ and $t > 0$, we have

$$|\mathrm{e}^{itz_1} - \mathrm{e}^{itz_2}| \le t|z_1 - z_2|. \tag{3.79}$$

Then, for any $x \in [1, \infty)$ we have that

$$\begin{aligned}
&|T_z f_1(x) - T_z f_2(x)| \\
&= \left| i \int_0^\infty \mathrm{e}^{itz} \left(\mathrm{e}^{it \int_1^\infty f_1(y)\kappa_\sigma(x,y)\mu_W(\mathrm{d}\, y)} - \mathrm{e}^{it \int_1^\infty f_2(y)\kappa_\sigma(x,y)\mu_W(\mathrm{d}\, y)} \right) \mathrm{d}\, t \right| \\
&\le \int_0^\infty \mathrm{e}^{-\eta t} \left| \mathrm{e}^{it \int_1^\infty f_1(y)\kappa_\sigma(x,y)\mu_W(\mathrm{d}\, y)} - \mathrm{e}^{it \int_1^\infty f_2(y)\kappa_\sigma(x,y)\mu_W(\mathrm{d}\, y)} \right| \mathrm{d}\, t \\
&\le \int_0^\infty \mathrm{e}^{-\eta t} t \left| \int_1^\infty (f_1(y) - f_2(y))\, \kappa_\sigma(x,y) \mu_W(\mathrm{d}\, y) \right| \mathrm{d}\, t,
\end{aligned} \tag{3.80}$$

Chapter Three

where in (3.80) we use (3.79). Now, evaluating the integral over t in (3.80), we obtain

$$\begin{aligned}&|T_z f_1(x) - T_z f_2(x)|\\ &\leq \frac{(\tau-1)}{\eta^2}\int_1^\infty |f_1(y)-f_2(y)|\kappa_\sigma(x,y)y^{-\tau}\,\mathrm{d}\,y,\end{aligned} \tag{3.81}$$

where we explicitly write down the Pareto law $\mu_W(\mathrm{d}\,y) := (\tau-1)y^{-\tau}\,\mathrm{d}\,y$. Recall that $\kappa_\sigma(x,y) = (x\wedge y)(x\vee y)^\sigma$. Thus, (3.81) becomes

$$\begin{aligned}&|T_z f_1(x) - T_z f_2(x)|\\ &\leq \frac{\tau-1}{\eta^2}\Big(\int_1^x |f_1(y)-f_2(y)|xy^{\sigma-\tau}\,\mathrm{d}\,y + \int_x^\infty |f_1(y)-f_2(y)|x^\sigma y^{1-\tau}\,\mathrm{d}\,y\Big).\end{aligned}$$

Integrating with respect to ν gives us

$$\begin{aligned}&\|T_z f_1 - T_z f_2\|_1\\ &\leq \frac{\tau-1}{\eta^2}\int_1^\infty\left(x\int_1^x |f_1(y)-f_2(y)|y^{\sigma-\tau}\,\mathrm{d}\,y\right)x^{-\beta}\,\mathrm{d}\,x\\ &\qquad + \frac{\tau-1}{\eta^2}\int_1^\infty\left(x^\sigma\int_x^\infty |f_1(y)-f_2(y)|y^{1-\tau}\,\mathrm{d}\,y\right)x^{-\beta}\,\mathrm{d}\,x\\ &= \frac{\tau-1}{\eta^2}\left(\int_1^\infty |f_1(y)-f_2(y)|y^{\sigma-\tau}\int_y^\infty x^{1-\beta}\,\mathrm{d}\,x\,\mathrm{d}\,y\right.\\ &\qquad\qquad \left. + \int_1^\infty |f_1(y)-f_2(y)|y^{1-\tau}\int_1^y x^{\sigma-\beta}\,\mathrm{d}\,x\,\mathrm{d}\,y\right).\end{aligned} \tag{3.82}$$

Using $\beta > 2$, the first integral in (3.82) can be bounded by

$$\begin{aligned}&\int_1^\infty |f_1(y)-f_2(y)|y^{\sigma-\tau}\int_y^\infty x^{1-\beta}\,\mathrm{d}\,x\,\mathrm{d}\,y\\ &= c_1\int_1^\infty |f_1(y)-f_2(y)|y^{-\beta}y^{2+\sigma-\tau}\,\mathrm{d}\,y \leq c_1\|f_1-f_2\|_1\,,\end{aligned} \tag{3.83}$$

since $y^{2+\sigma-\tau} \leq 1$ and $c_1 = 1/(\beta-2)$. Similarly, the second integral in (3.82) gives us

$$\begin{aligned}\int_1^\infty |f_1(y)-f_2(y)|y^{1-\tau}\int_1^y x^{\sigma-\beta}\,\mathrm{d}\,x\,\mathrm{d}\,y &\leq c_2\int_1^\infty |f_1(y)-f_2(y)|y^{1-\tau}\,\mathrm{d}\,y\\ &\leq c_2\|f_1-f_2\|_1,\end{aligned} \tag{3.84}$$

with $c_2 = 1/(\beta-1-\sigma)$, where for the last line we have used $1-\tau < -\beta$. Combining (3.83) and (3.84) in (3.82) gives us that

$$\|T_z f_1 - T_z f_2\|_1 \leq \frac{\tilde{c}}{\eta^2}\|f_1-f_2\|_1\,, \tag{3.85}$$

where $\tilde{c}$ is a constant depending on τ, σ and β. Thus, taking $\eta > 0$ to be sufficiently large such that $\eta > \sqrt{\tilde{c}}$ gives us that T_z is a contraction mapping on $\mathcal{B}$, hence proving the result. □

The following corollary is immediate from the Banach fixed-point theorem for contraction mappings.

Corollary 3.8.8.
Let $T_z : \mathcal{B} \to \mathcal{B}$ be the contraction map given in (3.78). Then, there exists a unique analytic function $a_z^* \in \mathcal{B}$ such that $T_z(a_z^*) = a_z^*$.

We know from (3.74) that

$$a_z(x) = i\int_0^\infty \mathrm{e}^{itz} \exp\left\{it\int_1^\infty c_m^{-1} a_z(y)\kappa_\sigma(x,y)\mathbb{1}_{\{y\le m\}}\mu_W(\mathrm{d}\, y)\right\}\mathrm{d}\, t\,. \tag{3.86}$$

Define $\tilde{a}_z$ as

$$\tilde{a}_z(x) = i\int_0^\infty \mathrm{e}^{itz} \exp\left\{it\int_1^\infty c_m^{-1} a_z(y)\kappa_\sigma(x,y)\mu_W(\mathrm{d}\, y)\right\}\mathrm{d}\, t\,. \tag{3.87}$$

Then, $\tilde{a}_z = T_z(c_m^{-1}a_z)$. We now have the following lemma.

Lemma 3.8.9.
Let a_z and $\tilde{a}_z$ be as in (3.86) and (3.87), respectively. Then,

$$\|a_z - \tilde{a}_z\|_1 \le \frac{C(m)}{\eta^3}\,,$$

where $C(m)$ is a constant depending on m such that $\lim_{m\to\infty} C(m) = 0$.

Proof of Lemma 3.8.9. Since $a_z \in \mathcal{B}$, we again use (3.79) to get

$$\begin{aligned}|a_z(x) - \tilde{a}_z(x)| &\le \int_0^\infty \mathrm{e}^{-\eta t} t\left|\int_m^\infty c_m^{-1} a_z(y)\kappa_\sigma(x,y)\mu_W(\mathrm{d}\, y)\right|\mathrm{d}\, t\\ &\le \frac{\tau-1}{c_m\eta^2}\int_m^\infty |a_z(y)|\kappa_\sigma(x,y)y^{-\tau}\,\mathrm{d}\, y\,,\end{aligned} \tag{3.88}$$

where we evaluate the integral over t to get the factor of η^{-2} in (3.88). Recall that $c_m = 1 - m^{-(\tau-1)}$. Using (3.75), we have that

$$|a_z(x) - \tilde{a}_z(x)| \le \frac{\tau-1}{c_m\eta^3}\int_m^\infty \kappa_\sigma(x,y)y^{-\tau}\,\mathrm{d}\, y. \tag{3.89}$$

Since $\kappa_\sigma(x,y) \le (xy)^{1\vee\sigma}$, we have

$$|a_z(x) - \tilde{a}_z(x)| \le \frac{\tau-1}{c_m\eta^3}x^{1\vee\sigma}\int_m^\infty y^{(1\vee\sigma)-\tau}\,\mathrm{d}\, y = \frac{(\tau-1)m^{(1\vee\sigma)-(\tau-1)}}{c_m((\tau-1)-(1\vee\sigma))\eta^3}x^{1\vee\sigma}, \tag{3.90}$$

where we use the fact that $\tau > \max(2, 1+\sigma)$, and so the integral evaluated in (3.90) is finite. Define

$$c(m) := \frac{(\tau-1)c_m^{-1} m^{(1\vee\sigma)-(\tau-1)}}{(\tau-1)-(1\vee\sigma)}.$$

Since c_m tends to one, and $m^{(1\vee\sigma)-(\tau-1)}$ tends to zero we have $c(m) = o_m(1)$. Now, integrating both sides of (3.90) against $x^{-\beta}\,\mathrm{d}\,x$ gives us

$$\|a_z - \tilde{a}_z\|_1 \leq \frac{c(m)}{\eta^3}\int_1^\infty x^{1\vee\sigma-\beta}\,\mathrm{d}\,x = \frac{C(m)}{\eta^3}\,, \tag{3.91}$$

since $\beta > 2 \vee 1 + \sigma$, and where $C(m) = o_m(1)$, completing the proof. □

We are now at the penultimate step, where we have the necessary tools to show the convergence of a_z to a_z^* in the space $\mathcal{B}$.

Lemma 3.8.10.
Let a_z^* be the unique fixed point of the contraction map T_z defined in (3.78). *Then, we have that*

$$\lim_{m\to\infty} \|a_z - a_z^*\|_1 = 0\,. \tag{3.92}$$

Proof of Lemma 3.8.10. We have, using Lemma 3.8.9 and the fact that T_z is a contraction, that

$$\begin{aligned}
\|a_z - a_z^*\|_1 &\leq \|a_z - \tilde{a}_z\|_1 + \|\tilde{a}_z - a_z^*\|_1 \\
&\leq C(m)\eta^{-3} + \|T_z(c_m^{-1}a_z) - T_z(a_z^*)\|_1 \\
&\leq C(m)\eta^{-3} + \tilde{c}\eta^{-2}\|c_m^{-1}a_z - a_z^*\|_1 \\
&\leq C(m)\eta^{-3} + \tilde{c}\eta^{-2}c_m^{-1}\|a_z - a_z^*\|_1 + \tilde{c}\eta^{-2}\|a_z^*\|_1|c_m^{-1} - 1|\,.
\end{aligned}$$

Thus, choosing $\eta > 0$ such that $0 < 1 - \tilde{c}c_m^{-1}\eta^{-2} < 1$, we have that

$$\|a_z - a_z^*\|_1 \leq \frac{1}{1 - C_\tau c_m^{-1}\eta^{-2}}\left(C(m)\eta^{-3} + C_\tau\eta^{-2}\|a_z^*\|_1|c_m^{-1} - 1|\right)\,. \tag{3.93}$$

Now, as $m \to \infty$, we have that $C(m) \to 0$, and $c_m \to 1$. Since $\|a_z^*\| < \infty$, we have that the right-hand side of (3.93) goes to 0 as $m \to \infty$. Thus, $\|a_z - a_z^*\|_1 \to 0$ as $m \to \infty$ for z in an appropriate domain $D_\eta \subset \mathbb{C}^+$. However, in the complex variable z, the domains of a_z and a_z^* are $\mathbb{C}^+$. Since the convergence holds for an open set of this domain (that is, in $D_\eta \subset \mathbb{C}^+$), by the identity theorem of complex analysis, the convergence holds everywhere in $\mathbb{C}^+$, that is, for each $z \in \mathbb{C}^+$. □

We now proceed towards a proof of Theorem 3.2.5, and to achieve this we wish to take the limit $m \to \infty$ to characterise $\mathrm{S}_{\mu_{\sigma,\tau}}$. We know that since $\lim_{m\to\infty} \mu_{\sigma,\tau,m} = \mu_{\sigma,\tau}$, then for each $z \in \mathbb{C}^+$, $\lim_{m\to\infty} \mathrm{S}_{\mu_{\sigma,\tau,m}}(z) = \mathrm{S}_{\mu_{\sigma,\tau}}(z)$.

Proof of Theorem 3.2.5. Let a_z^* be the unique fixed point of the contraction mapping T_z as in Corollary 3.8.8, and let $\mathrm{S}_{\mu_{\sigma,\tau}}(z)$ be the Stieltjes transform of $\mu_{\sigma,\tau}$ for any $z \in \mathbb{C}^+$. We wish to show that

$$\mathrm{S}_{\mu_{\sigma,\tau}}(z) = \int_1^\infty a_z^*(x)\mu_W(\mathrm{d}\,x).$$

We have that

$$\begin{aligned}
&\left|\int_1^\infty a_z(x)\mu_{W,m}(\mathrm{d}\,x) - \int_1^\infty a_z^*(x)\mu_W(\mathrm{d}\,x)\right| \\
&\le \left|\int_1^\infty a_z(x)\mu_{W,m}(\mathrm{d}\,x) - \int_1^\infty a_z^*(x)\mu_{W,m}(\mathrm{d}\,x)\right| \\
&\quad + \left|\int_1^\infty a_z^*(x)\mu_{W,m}(\mathrm{d}\,x) - \int_1^\infty a_z^*(x)\mu_W(\mathrm{d}\,x)\right|. \qquad (3.94)
\end{aligned}$$

The first term in (3.94) can be evaluated as

$$\begin{aligned}
&\left|\int_1^\infty a_z(x)\mu_{W,m}(\mathrm{d}\,x) - \int_1^\infty a_z^*(x)\mu_{W,m}(\mathrm{d}\,x)\right| \\
&\le (\tau-1)c_m^{-1}\int_1^m |a_z(x) - a_z^*(x)|x^{-\tau}\,\mathrm{d}\,x \\
&\le (\tau-1)c_m^{-1}\int_1^\infty |a_z(x) - a_z^*(x)|x^{-\beta}x^{\beta-\tau}\,\mathrm{d}\,x \\
&\le (\tau-1)c_m^{-1}\|a_z - a_z^*\|_1 = o_m(1), \qquad (3.95)
\end{aligned}$$

as $x^{\beta-\tau} \le 1$, and $\|a_z - a_z^*\|_1 = o_m(1)$ from Lemma 3.8.10. The second term of (3.94) can be evaluated as

$$\begin{aligned}
&\left|\int_1^\infty a_z^*(x)\mu_{W,m}(\mathrm{d}\,x) - \int_1^\infty a_z^*(x)\mu_W(\mathrm{d}\,x)\right| \\
&\le c_m^{-1}\left|\int_1^m a_z^*(x)\mu_W(\mathrm{d}\,x) - \int_1^\infty a_z^*(x)\mu_W(\mathrm{d}\,x)\right| + \left|\int_1^\infty a_z^*(x)\mu_W(\mathrm{d}\,x)\right| |c_m^{-1} - 1| \\
&\le \frac{(\tau-1)}{c_m\eta}\int_m^\infty x^{-\tau}\,\mathrm{d}\,x + \frac{|c_m^{-1}-1|}{\eta} = \frac{(\tau-1)m^{1-\tau}}{c_m\eta} + \frac{|c_m^{-1}-1|}{\eta} = o_m(1), \qquad (3.96)
\end{aligned}$$

since $|a_z^*| \le \eta^{-1}$. Combining (3.95) and (3.96) completes the proof of the theorem. $\square$

Chapter Three