Wireless random-access networks and spectra of random graphs
Sfragara, M.

Author: Sfragara, M.
Title: Wireless random-access networks and spectra of random-graphs
Issue Date: 2020-10-28

PART II

SPECTRA OF INHOMOGENEOUS ERDŐS-RÉNYI RANDOM GRAPHS

CHAPTER 5

Spectral distribution of the adjacency and the Laplacian matrix

This chapter is based on:
A. Chakrabarty, R.S. Hazra, F. den Hollander, M. Sfragara. *Spectra of adjacency and Laplacian matrices of Inhomogeneous Erdős-Rényi random graphs.* Random Matrices: Theory and Applications, 2020.

Abstract

We consider inhomogeneous Erdős-Rényi random graphs G_N on N vertices in the non-sparse non-dense regime. The edge between the pair of vertices $\{i, j\}$ is retained with probability $\varepsilon_N\, f(\frac{i}{N}, \frac{j}{N})$, $1 \leq i \neq j \leq N$, independently of other edges, where $f\colon\ [0,1]^2 \to [0,\infty)$ is a continuous function such that $f(x,y) = f(y,x)$ for all $(x,y) \in [0,1]^2$. We study the empirical distribution of both the adjacency matrix A_N and the Laplacian matrix Δ_N associated with G_N, in the limit as $N \to \infty$ when $\lim_{N\to\infty} \varepsilon_N = 0$ and $\lim_{N\to\infty} N\varepsilon_N = \infty$. In particular, we show that the empirical spectral distributions of A_N and Δ_N, after appropriate scaling and centering, converge to deterministic limits weakly in probability. For the special case where $f(x,y) = r(x)r(y)$ with $r\colon\ [0,1] \to [0,\infty)$ a continuous function, we give an explicit characterization of the limiting distributions. Furthermore, we apply our results to constrained random graphs, Chung-Lu random graphs and social networks.

§5.1 Introduction and main results

In Section 5.1.1 we define the mathematical model. In Section 5.1.2 we state the existence of the limiting spectral distributions for the adjancency and Laplacian matrices after suitable scaling. In Section 5.1.3 we identify those limiting spectral distributions under the assumption that the connection probabilities have a multiplicative structure. In Section 5.1.4 we generalize our results to graphs where the connection probabilities are randomized. In Section 5.1.5 we anticipate some of the applications that we will discuss later and give an outline of the remainder of the chapter.

§5.1.1 Setting

We refer to Section 1.2.3 for a general introduction to spectra of Erdős-Rényi random graphs. We focus on *inhomogeneous Erdős-Rényi random graphs* and consider the non-dense non-sparse regime, where the degrees of the vertices diverge sublinearly with the size of the graph.

Let $f\colon\ [0,1]^2 \to [0,\infty)$ be a continuous function, satisfying

$$f(x,y) = f(y,x) \qquad \forall\,(x,y) \in [0,1]^2. \tag{5.1}$$

A sequence of positive real numbers $(\varepsilon_N\colon\ N \geq 1)$ is fixed that satisfies

$$\lim_{N\to\infty} \varepsilon_N = 0, \qquad \lim_{N\to\infty} N\varepsilon_N = \infty. \tag{5.2}$$

Consider the random graph G_N on the set of vertices $\{1,\ldots,N\}$ where, for each (i,j) with $1 \leq i < j \leq N$, an edge is present between vertices i and j with probability

$$\varepsilon_N f(\tfrac{i}{N},\tfrac{j}{N}), \tag{5.3}$$

independently of other pairs of vertices. In particular, G_N is an undirected graph with no self loops and no multiple edges. Boundedness of f ensures that $\varepsilon_N f(\frac{i}{N},\frac{j}{N}) \leq 1$ for all $1 \leq i < j \leq N$ when N is large enough. If $f \equiv c$ with c a constant, then G_N is the Erdős-Rényi graph with edge retention probability $\varepsilon_N c$. For general f, G_N can be thought of as an inhomogeneous version of the Erdős-Rényi graph.

We next define our two main objects of interest. We refer to Section 1.2.2 for more details.

Definition 5.1.1 (Adjacency and Laplacian matrices).
The *adjacency matrix* of G_N is denoted by A_N and defined as in (1.15). Clearly, A_N is a symmetric random matrix whose diagonal entries are zero and whose upper triangular entries are independent Bernoulli random variables, i.e.,

$$A_N(i,j) \triangleq \mathrm{BER}\big(\varepsilon_N f\big(\tfrac{i}{N},\tfrac{j}{N}\big)\big), \qquad 1 \leq i \neq j \leq N. \tag{5.4}$$

The *Laplacian matrix* of G_N is denoted Δ_N and defined as in (1.17).

§5.1.2 Existence of the limiting spectral distribution

We recall the definition of *empirical spectral distribution (ESD)* in (1.19). It is the probability measure that puts mass $1/N$ at every eigenvalue, respecting its algebraic multiplicity.

Our first theorem states the existence of the limiting spectral distribution of A_N after suitable scaling.

Theorem 5.1.2 (Existence of the limiting spectral distribution of A_N). *There exists a symmetric probability measure μ on $\mathbb{R}$ such that*

$$\lim_{N\to\infty} \mathrm{ESD}\big((N\varepsilon_N)^{-1/2} A_N\big) = \mu \quad \textit{weakly in probability,} \tag{5.5}$$

and μ is compactly supported. Furthermore, if

$$\min_{0\le x,y\le 1} f(x,y) > 0, \tag{5.6}$$

then μ is absolutely continuous with respect to Lebesgue measure.

Our second theorem is the analogue of Theorem 5.1.2 with A_N replaced by Δ_N.

Theorem 5.1.3 (Existence of the limiting spectral distribution of Δ_N). *There exists a symmetric probability measure ν on $\mathbb{R}$ such that*

$$\lim_{N\to\infty} \mathrm{ESD}\big((N\varepsilon_N)^{-1/2}(\Delta_N - D_N)\big) = \nu \quad \textit{weakly in probability,} \tag{5.7}$$

where

$$D_N = \mathrm{Diag}\left(\mathbb{E}\big[\Delta_N(1,1)\big], \dots, \mathbb{E}\big[\Delta_N(N,N)\big]\right). \tag{5.8}$$

Furthermore, if

$$f \not\equiv 0, \tag{5.9}$$

then the support of ν is unbounded.

The ESD of a random matrix is a random probability measure. Note that μ and ν are both deterministic, i.e., a law of large numbers is in force.

Theorems 5.1.2–5.1.3 are existential, in the sense that explicit descriptions of μ and ν are missing. We have some control on the Stieltjes transform of μ. In the proof of Theorem 5.1.2 (in Lemma 5.2.3) we will see that the ESD of $(N\varepsilon_N)^{-1/2}A_N$ has the same limit as the ESD of

$$\bar{A}_N(i,j) = \sqrt{\frac{1}{N} f\left(\frac{i}{N}, \frac{j}{N}\right)}\, G_{i\wedge j, i\vee j} \tag{5.10}$$

with $(G_{i,j} : 1 \le i \le j)$ a family of i.i.d. standard Gaussian random variables. Such random matrices are known in the literature as Wigner matrices with a variance profile (see, for example, [107], [129], [164], [185]). The limiting spectral distribution of $\bar{A}_N$ matches with the one of certain symmetric random matrices with dependent entries (see [135] for details). It turns out that, by using the combinatorics of non-crossing partitions, we can derive a recursive equation for the Stieltjes transform of μ, i.e.,

$$G_\mu(z) = \int_{\mathbb{R}} \frac{1}{z-x}\,\mu(dx), \quad z \in \mathbb{C}\setminus\mathbb{R}. \tag{5.11}$$

It turns out that

$$G_\mu(z) = \int_0^1 \mathcal{H}(z,x)\,dx, \tag{5.12}$$

where $\mathcal{H}(z,x)$, $x \in [0,1]$, is the unique analytic solution of the integral equation

$$z\mathcal{H}(z,x) = 1 + \mathcal{H}(z,x)\int_0^1 \mathcal{H}(z,y)f(x,y)\,dy, \qquad x \in [0,1]. \tag{5.13}$$

The form of $\mathcal{H}(z,x)$ can also be expressed in terms of non-crossing partitions and the function $f(x,y)$ (see [134, Section 4.1] for details). We mention that the above measure is similar to the limiting measure in [196, Theorem 3.4]. There it is shown that a graphon sequence W_N can be associated with a Wigner matrix with a variance profile $(s_{i,j} : 1 \le i,j \le N)$. If the sequence of graphons W_N converges in the cut norm to W with $W(x,y) = f(x,y)$, then the limiting measure matches with μ.

The description of ν through its Stieltjes transform is hard to obtain, although, just like before, the ESD of $(N\varepsilon_N)^{-1/2}(\Delta_N - D_N)$ turns out to be the same as that of

$$\tilde{\Delta}_N = \bar{A}_N + Y_N, \tag{5.14}$$

where Y_N is a diagonal matrix of order N defined by

$$Y_N(i,i) = Z_i\sqrt{\frac{1}{N}\sum_{1\le j\le N,\, j\ne i} f\left(\frac{i}{N},\frac{j}{N}\right)}, \qquad 1 \le i \le N, \tag{5.15}$$

where $(Z_i\colon\, i \ge 1)$ is a family of i.i.d. standard normal random variables, independent of $(G_{i,j}\colon\, 1 \le i \le j)$. Suppose that Y_N is a deterministic diagonal matrix, embedded in $L^\infty[0,1]$ (as a step function). For the case where this function converges to a function h in the $\|\cdot\|_\infty$ norm, the limiting spectral distribution of $\bar{A}_N + Y_N$ was studied in [185] (see also [186, Theorem 22.7.2]). In our case, due to the presence in Y_N of Gaussian random variables (which have unbounded support) and the fact that the spectral norm of Y_N tends to infinity as $N \to \infty$, the existing results cannot be applied. One of the major contributions of our paper is to overcome this hurdle. Also, our proofs ensure that ν has a finite moment generating function (see (5.123) below) and unbounded support.

§5.1.3 Identification of the limiting spectral distribution

Our next theorem identifies μ and ν under the additional assumption that f has a *multiplicative structure*, i.e.,

$$f(x,y) = r(x)r(y), \qquad (x,y) \in [0,1]^2, \tag{5.16}$$

for some continuous function $r\colon [0,1] \to [0,\infty)$. The statement is based on the theory of (possibly unbounded) self-adjoint operators affiliated with a W^*-probability space. Recall Section 1.2.4 for an introduction to free probability theory. A few extra relevant definitions are given below. For details the reader is referred to [108, Section 5.2.3].

Definition 5.1.4 (Operators affiliated with a W^*-probability space).
A C^*-algebra $\mathcal{A} \subset B(\mathcal{H})$, with $\mathcal{H}$ a Hilbert space, is a W^*-algebra when $\mathcal{A}$ is closed under the weak operator topology. If, in addition, τ is a state such that there exists a unit vector $\xi \in \mathcal{H}$ satisfying

$$\tau(a) = \langle a\xi, \xi\rangle \qquad \forall\, a \in \mathcal{H}, \tag{5.17}$$

then $(\mathcal{A}, \tau)$ is a W^*-probability space. In that case a densely defined self-adjoint (possibly unbounded) operator T on $\mathcal{H}$ is said to be *affiliated with* $\mathcal{A}$ if $h(T) \in \mathcal{A}$ for any bounded measurable function h defined on the spectrum of T, where $h(T)$ is defined by the spectral theorem. Finally, for an affiliated operator T, its *law* $\mathcal{L}(T)$ is the unique probability measure on $\mathbb{R}$ satisfying

$$\tau(h(T)) = \int_{\mathbb{R}} h(x)(\mathcal{L}(T))(dx) \tag{5.18}$$

for every bounded measurable $h\colon \mathbb{R} \to \mathbb{R}$.

The distribution of a single self-adjoint operator is defined above. For two or more self-adjoint operators $T_1, \dots, T_n$, a description of their *joint distribution* is a specification of

$$\tau(h_1(T_{i_1}) \cdots h_k(T_{i_k})), \tag{5.19}$$

for all $k \geq 1$, all $i_1, \dots, i_k \in \{1, \dots, n\}$, and all bounded measurable functions $h_1, \dots, h_k$ from $\mathbb{R}$ to itself. Once the above is specified, it is immediate to see that $\mathcal{L}(p(T_1, \dots, T_k))$ can be calculated for any polynomial p in k variables such that $p(T_1, \dots, T_k)$ is self-adjoint.

Definition 5.1.5 (Free independence of operators).
Let $(\mathcal{A}, \tau)$ be a W^*-probability space and $a_1, a_2 \in \mathcal{A}$. Then a_1 and a_2 are *freely independent* if

$$\tau(p_1(a_{i_1}) \cdots p_n(a_{i_n})) = 0, \tag{5.20}$$

for all $n \geq 1$, all $i_1, \dots, i_n \in \{1, 2\}$ with $i_j \neq i_{j+1}$, $j = 1, \dots, n-1$, and all polynomials $p_1, \dots, p_n$ in one variable satisfying

$$\tau(p_j(a_{i_j})) = 0, \qquad j = 1, \dots, n. \tag{5.21}$$

For (possibly unbounded) operators $a_1, \dots, a_k$ and $b_1, \dots, b_m$ affiliated with $\mathcal{A}$, the collections $(a_1, \dots, a_k)$ and $(b_1, \dots, b_m)$ are *freely independent* if and only if

$$p(h_1(a_1), \dots, h_k(a_k)) \quad \text{and} \quad q(g_1(b_1), \dots, g_m(b_m)), \tag{5.22}$$

are freely independent for all bounded measurable $h_1, \dots, h_k$ and $g_1, \dots, g_m$, and all polynomials p and q in k and m non-commutative variables, respectively. It is immediate that the two operators in the above display are bounded, and hence belong to $\mathcal{A}$.

We are now in a position to state our next theorem.

Theorem 5.1.6 (Identification of the limiting spectral distribution). *If f is as in* (5.16)*, then*

$$\mu = \mathcal{L}\Big(r^{1/2}(T_u)T_s r^{1/2}(T_u)\Big), \tag{5.23}$$

and

$$\nu = \mathcal{L}\Big(r^{1/2}(T_u)T_s r^{1/2}(T_u) + \alpha r^{1/4}(T_u)T_g r^{1/4}(T_u)\Big), \tag{5.24}$$

where

$$\alpha = \left(\int_0^1 r(x)\,dx\right)^{1/2}. \tag{5.25}$$

Here, T_g and T_u are commuting self-adjoint operators affiliated with a W^*-probability space $(\mathcal{A}, \tau)$ such that, for bounded measurable functions h_1, h_2 from $\mathbb{R}$ to itself,

$$\tau(h_1(T_g)h_2(T_u)) = \left(\int_{\mathbb{R}} h_1(x)\phi(x)\,dx\right)\left(\int_0^1 h_2(u)\,du\right), \tag{5.26}$$

with ϕ the standard normal density. Furthermore, T_s has a standard semicircle distribution and is freely independent of (T_g, T_u).

The right-hand side of (5.23) is the same as the free multiplicative convolution of the standard semicircle law and the law of $r(U)$, where U is a standard uniform random variable.

The fact that T_g and T_u commute, together with (5.26), specifies their joint distribution. In fact, they are standard normal and standard uniform, respectively, independently of each other in the *classical sense*. Free independence of T_s and (T_g, T_u), plus the fact that the former follows the standard semicircle law, specifies the joint distribution of T_s, T_g, T_u.

In order to admit the unbounded operator T_g, a W^*-probability space is needed. If all the operators would have been bounded, then a C^*-probability space would have sufficed.

§5.1.4 Randomization

Theorem 5.1.2 can be generalized to the situation where the function f is random. Such a randomization helps us to address the applications listed in Section 5.4. Suppose that $(\varepsilon_N \colon N \geq 1)$ is a sequence of positive numbers satisfying (5.2). Suppose further that, for every $N \geq 1$, $(R_{Ni} \colon 1 \leq i \leq N)$ is a collection of non-negative random variables (defined on the same probability space) such that there is a deterministic $C < \infty$ for which

$$\sup_{N \geq 1} \max_{1 \leq i \leq N} R_{Ni} \leq C \quad \text{almost surely.} \tag{5.27}$$

In addition, suppose that there is a probability measure μ_r on $\mathbb{R}$ such that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \delta_{R_{Ni}} = \mu_r \quad \text{weakly almost surely.} \tag{5.28}$$

The non-negativity of R_{Ni} and (5.27) ensure that μ_r is concentrated on $[0, C]$. Furthermore, the first line of (5.2) ensures that the additional assumption

$$\sup_{N\geq 1} \varepsilon_N \leq \frac{1}{C^2} \tag{5.29}$$

entails no loss of generality.

For fixed N and conditional on $(R_{N1}, \ldots, R_{NN})$, the random graph G_N is constructed as before, except that there is an edge between i and j with probability $\varepsilon_N R_{Ni} R_{Nj}$, which is at most 1 by (5.29) for all $1 \leq i < j \leq N$. In other words, G_N has two levels of randomness: one in the choice of $(R_{N1}, \ldots, R_{NN})$ and one in the choice of the set of edges. Once again, A_N is the adjacency matrix of G_N. The following is a *randomized* version of Theorem 5.1.2.

Theorem 5.1.7 (Limiting spectral distribution of A_N).
Under the assumptions (5.2) *and* (5.27)–(5.28),

$$\lim_{N\to\infty} \mathrm{ESD}\big((N\varepsilon_N)^{-1/2} A_N\big) = \mu_r \boxtimes \mu_s \quad \textit{weakly in probability,} \tag{5.30}$$

where μ_s is the standard semicircle law.

§5.1.5 Applications and outline

As we will see in Section 5.4, our results can be applied in various ways. A first application consists in *constrained random graphs*. Given a sequence of positive integers, among the probability distributions for which the sequence of average degrees matches the given sequence, called the soft configuration model, the one that maximizes the entropy is the canonical Gibbs measure. It is known that, under a sparsity condition, the connection probabilities arising out of the canonical Gibbs measure asymptotically have a multiplicativestructure (see [187]). We show that our results on the adjacency matrix can be easily extended to cover such situations. Another important application consists in *Chung-Lu type random graph*, which are used to model *sociability patterns in networks*. We show how to use the rescaled empirical spectral distribution and free probability to statistically recover the underlying sociability distribution.

Outline of the chapter. The remainder of this chapter is organized as follows. In Section 5.2 a number of technical lemmas are proved. These serve as preparation for the proofs of our main theorems, which are given in Section 5.3. In Section 5.4, the above applications are discussed, organized into three propositions. Appendix E collects a few basic facts that are needed along the way.

§5.2 Preparatory approximations

The proofs of our main theorems rely on several preparatory approximations, which we organize in Lemmas 5.2.1–5.2.4 and 5.2.6 below. Along the way we need several basic results, which we collect in Appendix E.

§5.2.1 Centering

The first approximation is that the mean of each off-diagonal entry of A_N and Δ_N can be subtracted, with negligible perturbation in the respective empirical spectral distributions.

Lemma 5.2.1 (Centering).
Let A_N^0 and Δ_N^0 be $N \times N$ matrices defined by

$$\begin{aligned} A_N^0(i,j) &= (N\varepsilon_N)^{-1/2}\big(A_N(i,j) - \mathbb{E}[A_N(i,j)]\big), && (5.31)\\ \Delta_N^0(i,j) &= (N\varepsilon_N)^{-1/2}\big(\Delta_N(i,j) - \mathbb{E}[\Delta_N(i,j)]\big), && (5.32) \end{aligned}$$

for all $1 \le i,j \le N$. Then

$$\begin{aligned} &\lim_{N\to\infty} L\big(\mathrm{ESD}(A_N^0), \mathrm{ESD}((N\varepsilon_N)^{-1/2}A_N)\big) = 0 \quad \textit{in probability,}\\ &\lim_{N\to\infty} L\big(\mathrm{ESD}(\Delta_N^0), \mathrm{ESD}((N\varepsilon_N)^{-1/2}(\Delta_N - D_N))\big) = 0 \quad \textit{in probability,} \end{aligned} \qquad (5.33)$$

where $L(\eta_1,\eta_2)$ denotes the Lévy distance between the probability measures η_1 and η_2, and D_N is the diagonal matrix defined in (5.8).

Proof. An appeal to Lemma E.1 shows that

$$\begin{aligned} &L^3\big(\mathrm{ESD}(A_N^0), \mathrm{ESD}((N\varepsilon_N)^{-1/2}A_N)\big)\\ &\le \frac{1}{N^2\varepsilon_N}\sum_{i,j=1}^{N} \mathbb{E}^2[A_N(i,j)]\\ &= \frac{1}{N^2\varepsilon_N}\sum_{i\neq j} \varepsilon_N^2 f^2\left(\frac{i}{N},\frac{j}{N}\right)\\ &= \varepsilon_N \int_{[0,1]^2} f^2(x,y)\,dx\,dy\,[1+o(1)], \qquad N\to\infty. \end{aligned} \qquad (5.34)$$

The first claim follows by recalling that $\varepsilon_N \to 0$. The proof the second claim is verbatim the same. ☐

§5.2.2 Gaussianisation

One of the crucial steps in studying the scaling properties of ESD is to replace each entry by a Gaussian random variable.

Lemma 5.2.2 (Gaussianisation).
Let $(G_{i,j}\colon\ 1 \le i \le j)$ be a family of i.i.d. standard Gaussian random variables. Define $N \times N$ matrices A_N^g and Δ_N^g by

$$A_N^g(i,j) = \begin{cases} \sqrt{\frac{1}{N} f\big(\frac{i}{N},\frac{j}{N}\big)\big(1-\varepsilon_N f\big(\frac{i}{N},\frac{j}{N}\big)\big)}\, G_{i\wedge j, i\vee j}, & i \ne j, \\ 0, & i = j, \end{cases} \tag{5.35}$$

$$\Delta_N^g(i,j) = \begin{cases} A_N^g(i,j), & i \ne j, \\ -\sum_{k=1,k\ne i}^N A_N^g(i,k), & i = j. \end{cases} \tag{5.36}$$

Fix $z \in \mathbb{C}\setminus\mathbb{R}$ and a three times continuously differentiable function $h\colon \mathbb{R}\to\mathbb{R}$ such that

$$\max_{0\le j\le 3} \sup_{x\in\mathbb{R}} |h^{(j)}(x)| < \infty. \tag{5.37}$$

For an $N \times N$ real symmetric matrix M, define

$$H_N(M) = \frac{1}{N}\operatorname{Tr}\big((M - zI_N)^{-1}\big), \tag{5.38}$$

where I_N is the identity matrix of order N. Then

$$\lim_{N\to\infty} \mathbb{E}\big[h\big(\Re H_N(A_N^g)\big) - h\big(\Re H_N(A_N^0)\big)\big] = 0, \tag{5.39}$$

$$\lim_{N\to\infty} \mathbb{E}\big[h\big(\Im H_N(A_N^g)\big) - h\big(\Im H_N(A_N^0)\big)\big] = 0, \tag{5.40}$$

and

$$\lim_{N\to\infty} \mathbb{E}\big[h\big(\Re H_N(\Delta_N^g)\big) - h\big(\Re H_N(\Delta_N^0)\big)\big] = 0, \tag{5.41}$$

$$\lim_{N\to\infty} \mathbb{E}\big[h\big(\Im H_N(\Delta_N^g)\big) - h\big(\Im H_N(\Delta_N^0)\big)\big] = 0, \tag{5.42}$$

where $\Re$ and $\Im$ denote the real and the imaginary part of a complex number, respectively.

Proof. We only prove (5.41). The proofs of the other claims are similar. We use ideas from [136]. Let $z = u + iv \in \mathbb{C}^+$ and $n = N(N-1)/2$. Define $\phi\colon\ \mathbb{R}^n \to \mathbb{C}$ as

$$\phi(x) = H_N(\Delta(x)) \tag{5.43}$$

where $\Delta(x)$ is the $N \times N$ symmetric Laplacian matrix given by

$$\Delta(x)(i,j) = \begin{cases} -\sum_{k=1,k\ne i}^N x_{i,k}, & i = j, \\ x_{i\wedge j, i\vee j}, & i \ne j. \end{cases} \tag{5.44}$$

Note that $\partial\Delta(x)/\partial x_{ij}$ is the $N\times N$ matrix that has -1 at the i-th and j-th diagonal and 1 at (i,j)-th and (j,i)-th entry. The following identities were derived in [136, Section 2]:

$$\begin{aligned}
\frac{\partial \phi}{\partial x_{i,j}} &= -N^{-1}\operatorname{Tr}\left(\frac{\partial \Delta}{\partial x_{i,j}}K^2\right),\\
\frac{\partial^2 \phi}{\partial x_{i,j}^2} &= 2N^{-1}\operatorname{Tr}\left(\frac{\partial \Delta}{\partial x_{i,j}}K\frac{\partial \Delta}{\partial x_{i,j}}K^2\right),\\
\frac{\partial^3 \phi}{\partial x_{i,j}^3} &= -6N^{-1}\operatorname{Tr}\left(\frac{\partial \Delta}{\partial x_{i,j}}K\frac{\partial \Delta}{\partial x_{i,j}}K\frac{\partial \Delta}{\partial x_{i,j}}K^2\right),
\end{aligned} \tag{5.45}$$

where $K(x) = (\Delta(x) - zI)^{-1}$. Now using these identities we get

$$\left\|\frac{\partial \phi}{\partial x_{ij}}\right\|_\infty \le \frac{4}{|\Im z|^2}\frac{1}{N}, \qquad \left\|\frac{\partial^2 \phi}{\partial x_{ij}^2}\right\|_\infty \le \frac{8}{|\Im z|^3}\frac{1}{N}, \qquad \left\|\frac{\partial^3 \phi}{\partial x_{ij}^3}\right\|_\infty \le \frac{48}{|\Im z|^4}\frac{1}{N}. \tag{5.46}$$

If we define

$$\lambda_2(\phi) = \sup\left\{\left\|\frac{\partial \phi}{\partial x_{i,j}}\right\|_\infty^2, \ \left\|\frac{\partial^2 \phi}{\partial x_{i,j}^2}\right\|_\infty\right\}, \tag{5.47}$$

$$\lambda_3(\phi) = \sup\left\{\left\|\frac{\partial \phi}{\partial x_{i,j}}\right\|_\infty^3, \left\|\frac{\partial^2 \phi}{\partial x_{i,j}^2}\right\|_\infty^2, \left\|\frac{\partial^3 \phi}{\partial x_{i,j}^3}\right\|_\infty\right\}, \tag{5.48}$$

then there exist constants C_2 and C_3 depending on $\Im z$ such that $\lambda_2(\phi) \le C_2 N^{-1}$ and $\lambda_3(\phi) \le C_3 N^{-1}$. Hence, using $\lambda_r(\Re\phi) \le \lambda_r(\phi)$ and

$$U = \Re\big(H_N(\Delta_N^0)\big), \qquad V = \Im\big(H_N(\Delta_N^g)\big), \tag{5.49}$$

we have from [136, Theorem 1.1]

$$\begin{aligned}
&\big|\mathbb{E}[h(U)] - \mathbb{E}[h(V)]\big|\\
&\le C_1(h)\lambda_2(\phi)\sum_{1\le i\ne j\le N}\big(\mathbb{E}[A_N^0(i,j)^2; |A_N^0(i,j)| > K]\\
&\qquad\qquad + \mathbb{E}[A_N^g(i,j)^2; |A_N^g(i,j)| > K]\big)\\
&\quad + C_2(h)\frac{\lambda_3(\phi)}{(N\varepsilon_N)^{3/2}}\sum_{i\ne j}\big(\mathbb{E}[A_N^0(i,j); |A_N^0(i,j)| > K]\\
&\qquad\qquad + \mathbb{E}[A_N^g(i,j)^3; |A_N^g(i,j)| > K]\big).
\end{aligned} \tag{5.50}$$

Using the fact that $\varepsilon_N \to 0$, we have that $\mathbb{E}[A_N^0(i,j)^4] = \mathrm{O}(N^{-2}\varepsilon_N^{-1})$. Also

$$\mathbb{P}(|A_N^0(i,j)| > K) \le \mathrm{O}(N^{-1}). \tag{5.51}$$

So, by the Cauchy-Schwartz inequality and the above bounds, we have

$$\mathbb{E}[A_N^0(i,j)^2; |A_N^0(i,j)| > K] \le \mathrm{O}\big(\varepsilon_N^{-1/2}N^{-3/2}\big). \tag{5.52}$$

Since $N\varepsilon_N \to \infty$, we have

$$\lambda_2(\phi)\sum_{1\le i\ne j\le N}\mathbb{E}[A_N^0(i,j)^2; |A_N^0(i,j)| > K] \le CN^{-1/2}\varepsilon_N^{-1/2}, \tag{5.53}$$

which tends to 0 as $N \to \infty$. Similarly, we have

$$\lambda_3(\phi) \sum_{i \neq j} \mathbb{E}[A_N^0(i,j)^3; |A_N^0(i,j)| > K] \leq \frac{C}{N^{5/2} \varepsilon_N^{3/2}} N^2 \varepsilon_N, \tag{5.54}$$

which also tends to 0 as $N \to \infty$. Using Gaussian tail bounds, we can also show that the other two terms in (5.50) tend to 0 as $N \to \infty$, which settles (5.41). In order to prove (5.42), a similar computation can be done for the imaginary part in (5.49). The proofs of (5.39) and (5.40) are analogous (and, in fact, closer to the argument in [136]). □

§5.2.3 Leading order variance

Next, we show that another minor tweak to the entries of A_N^g and Δ_N^g results in a negligible perturbation.

Lemma 5.2.3 (Leading order variance).
Define an $N \times N$ matrix A_N by

$$\bar{A}_N(i,j) \quad = \quad \sqrt{\frac{1}{N} f\left(\frac{i}{N}, \frac{j}{N}\right)}\, G_{i \wedge j, i \vee j}, \qquad 1 \leq i, j \leq N, \tag{5.55}$$

and let

$$\bar{\Delta}_N = \bar{A}_N - X_N, \tag{5.56}$$

where X_N is a diagonal matrix of order N defined by

$$X_N(i,i) = \sum_{k=1, k \neq i}^{N} \bar{A}_N(i,k), \qquad 1 \leq i \leq N. \tag{5.57}$$

Then

$$\lim_{N \to \infty} L\big(\mathrm{ESD}(A_N^g), \mathrm{ESD}(\bar{A}_N)\big) \quad = \quad 0 \quad \textit{in probability}, \tag{5.58}$$

$$\lim_{N \to \infty} L\big(\mathrm{ESD}(\Delta_N^g), \mathrm{ESD}(\bar{\Delta}_N)\big) \quad = \quad 0 \quad \textit{in probability.} \tag{5.59}$$

Proof. To prove (5.59), yet another application of Lemma E.1 implies that

$$
\begin{aligned}
&\mathbb{E}\big[L^3\big(\mathrm{ESD}(\Delta_N^g), \mathrm{ESD}(\bar{\Delta}_N)\big)\big] \\
&\le \frac{1}{N}\mathbb{E}\big[\operatorname{Tr}\big((\Delta_N^g - \bar{\Delta}_N)^2\big)\big] \\
&= \frac{1}{N}\sum_{1\le i\ne j\le N} \operatorname{Var}(\bar{A}_N(i,j) - A_N^g(i,j)) \\
&\quad + \frac{1}{N}\sum_{i=1}^N \operatorname{Var}\Big(\sum_{j=1, j\ne i}^N (\bar{A}_N(i,j) - A_N^g(i,j))\Big) + \frac{1}{N^2}\sum_{i=1}^N f\Big(\frac{i}{N}, \frac{i}{N}\Big) \qquad (5.60) \\
&= \frac{4}{N^2}\sum_{1\le i<j\le N} f\Big(\frac{i}{N}, \frac{j}{N}\Big)\Big(1 - \sqrt{1 - \varepsilon_N f\Big(\frac{i}{N}, \frac{j}{N}\Big)}\Big)^2 \\
&\quad + \frac{1}{N^2}\sum_{i=1}^N f\Big(\frac{i}{N}, \frac{i}{N}\Big),
\end{aligned}
$$

which tends to 0 as $N \to \infty$ because f is bounded. Thus, (5.59) follows. The proof of (5.58) is similar. □

§5.2.4 Decoupling

The (diagonal) entries of X_N are nothing but the row sums of $\bar{A}_N$. However, the correlation between an entry of $\bar{A}_N$ and that of X_N is small. The following decoupling lemma shows that it does not hurt when the entries of X_N are replaced by a mean-zero Gaussian random variable of the same variance that is independent of $\bar{A}_N$.

Lemma 5.2.4 (Decoupling).
Let $(Z_i \colon i \ge 1)$ be a family of i.i.d. standard normal random variables, independent of $(G_{i,j} \colon 1 \le i \le j)$. Define a diagonal matrix Y_N of order N by

$$
Y_N(i,i) = Z_i \sqrt{\frac{1}{N}\sum_{j=1, j\ne i}^N f\Big(\frac{i}{N}, \frac{j}{N}\Big)}, \qquad 1 \le i \le N, \tag{5.61}
$$

and let

$$
\tilde{\Delta}_N = \bar{A}_N + Y_N. \tag{5.62}
$$

Then, for every $k \in \mathbb{N}$,

$$
\lim_{N\to\infty} \frac{1}{N}\mathbb{E}\big[\operatorname{Tr}\big((\tilde{\Delta}_N)^{2k} - (\bar{\Delta}_N)^{2k}\big)\big] = 0, \tag{5.63}
$$

and

$$
\lim_{N\to\infty} \frac{1}{N^2}\mathbb{E}\big[\operatorname{Tr}^2\big((\tilde{\Delta}_N)^k\big) - \operatorname{Tr}^2\big((\bar{\Delta}_N)^k\big)\big] = 0. \tag{5.64}
$$

Proof. Without loss of generality we may assume that $f \le 1$. For $N \ge 1$, define the $N \times N$ matrices $\bar{M}_N$ and $\tilde{M}_N$ by

$$\bar{M}_N(i,j) = \begin{cases} N^{-1/2} G_{i\wedge j, i\vee j}, & i \neq j, \\ N^{-1/2} G_{i,i} - \sum_{k=1,\, k\neq i}^{N} \bar{M}_N(i,k), & i = j, \end{cases} \tag{5.65}$$

and

$$\tilde{M}_N(i,j) = \begin{cases} \bar{M}_N(i,j), & i \neq j, \\ N^{-1/2} G_{i,i} + Z_i \sqrt{\frac{N-1}{N}}, & i = j. \end{cases} \tag{5.66}$$

Note that, in the special case where f is identically 1, $\bar{M}_N$ and $\tilde{M}_N$ are identical to $\bar{\Delta}_N$ and $\tilde{\Delta}_N$, respectively. For $k \in \mathbb{N}$ and Π a partition of $\{1, \dots, 2k\}$, let

$$\Psi(\Pi, N) = \Big\{ i \in \{1, \dots, N\}^{2k} \colon\ i_u = i_v \iff u, v \text{ belong to the same block of } \Pi \Big\}. \tag{5.67}$$

For fixed Π and N, an immediate application of Wick's formula shows that, for all $i, j \in \Psi(\Pi, N)$,

$$\mathbb{E}\Bigg[\prod_{u=1}^{2k} \bar{M}_N(i_u, i_{u+1}) \Bigg] = \mathbb{E}\Bigg[\prod_{u=1}^{2k} \bar{M}_N(j_u, j_{u+1}) \Bigg], \tag{5.68}$$

with the convention that $i_{2k+1} \equiv i_1$, and

$$\mathbb{E}\Bigg[\prod_{u=1}^{2k} \tilde{M}_N(i_u, i_{u+1}) \Bigg] = \mathbb{E}\Bigg[\prod_{u=1}^{2k} \tilde{M}_N(j_u, j_{u+1}) \Bigg], \tag{5.69}$$

Therefore, for any $i \in \Psi(\Pi, N)$, we can unambiguously define

$$\psi(\Pi, N) = \mathbb{E}\Bigg[\prod_{u=1}^{2k} \bar{M}_N(i_u, i_{u+1}) \Bigg] - \mathbb{E}\Bigg[\prod_{u=1}^{2k} \tilde{M}_N(i_u, i_{u+1}) \Bigg]. \tag{5.70}$$

As shown in [128, Lemma 4.12], for a fixed Π,

$$\lim_{N\to\infty} N^{-1} |\psi(\Pi, N)| |\Psi(\Pi, N)| = 0. \tag{5.71}$$

An immediate observation is that, for all $1 \le i, j, i', j' \le N$,

$$\mathrm{Cov}\big(\tilde{M}_N(i,j), \tilde{M}_N(i',j')\big) = 0 \quad \text{if} \quad (i \wedge j, i \vee j) \neq (i' \wedge j', i' \vee j'), \tag{5.72}$$

and likewise for $\tilde{\Delta}_N$. Furthermore,

$$\mathrm{Var}\big(\tilde{M}_N(i,j)\big) = \mathrm{Var}\big(\bar{M}_N(i,j)\big), \qquad 1 \le i, j \le N, \tag{5.73}$$

and likewise for $\tilde{\Delta}_N$ and $\bar{M}_N$. For $N \ge 1$ and $1 \le i, j, i', j' \le N$, define

$$\eta_N(i,j,i',j') = \begin{cases} \frac{\mathrm{Cov}\left(\bar{\Delta}_N(i,j), \bar{\Delta}_N(i',j')\right)}{\mathrm{Cov}\left(\bar{M}_N(i,j), \bar{M}_N(i',j')\right)}, & \text{if the denominator is non-zero,} \\ 0, & \text{otherwise.} \end{cases} \tag{5.74}$$

CHAPTER 5

It is easy to check that the assumption $f \le 1$ ensures that $|\eta_N(i,j,i',j')| \le 1$. Therefore, for all N and $1 \le i,j,i',j' \le N$,

$$
\begin{aligned}
\mathrm{Cov}\big(\bar{\Delta}_N(i,j), \bar{\Delta}_N(i',j')\big) &= \eta_N(i,j,i',j')\mathrm{Cov}\big(\bar{M}_N(i,j), \bar{M}_N(i',j')\big),\\
\mathrm{Cov}\big(\tilde{\Delta}_N(i,j), \tilde{\Delta}_N(i',j')\big) &= \eta_N(i,j,i',j')\mathrm{Cov}\big(\tilde{M}_N(i,j), \tilde{M}_N(i',j')\big).
\end{aligned}
$$

For fixed Π, N and $i \in \Psi(\Pi, N)$, by an appeal to Wick's formula the above implies that there exists a $\xi(i,N) \in [-1,1]$ such that

$$
\mathbb{E}\Bigg[\prod_{u=1}^{2k} \bar{\Delta}_N(i_u, i_{u+1})\Bigg] - \mathbb{E}\Bigg[\prod_{u=1}^{2k} \tilde{\Delta}_N(i_u, i_{u+1})\Bigg] = \xi(i,N)\psi(\Pi,N), \tag{5.75}
$$

and therefore, by (5.71),

$$
\begin{aligned}
&\sum_{i\in\Psi(\Pi,N)} \Bigg|\mathbb{E}\Bigg[\prod_{u=1}^{2k} \bar{\Delta}_N(i_u, i_{u+1})\Bigg] - \mathbb{E}\Bigg[\prod_{u=1}^{2k} \tilde{\Delta}_N(i_u, i_{u+1})\Bigg]\Bigg|\\
&\quad = \sum_{i\in\Psi(\Pi,N)} |\xi(i,N)||\psi(\Pi,N)| \le |\psi(\Pi,N)||\Psi(\Pi,N)| = o(N), \qquad N \to \infty.
\end{aligned} \tag{5.76}
$$

Since this holds for every partition Π of $\{1,\dots,2k\}$, (5.63) follows. The proof of (5.64) follows along similar lines. $\square$

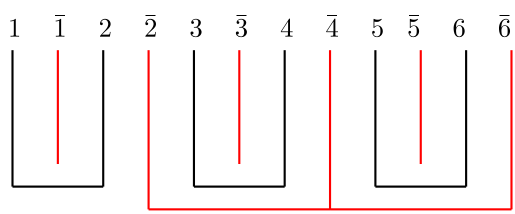
§5.2.5 Combinatorics from free probability

The final preparation is a general result from random matrix theory. To state this, the following notions from the theory of free probability are borrowed. We refer to Section 1.2.4 for an introduction to free probability theory and to [186] for more details.

Definition 5.2.5 (Kreweras complement).
For an even positive integer k, $NC_2(k)$ is the set of non-crossing pair partitions of $\{1,\dots,k\}$. For $\sigma \in NC_2(k)$, its *Kreweras complement* $K(\sigma)$ is the maximal non-crossing partition $\bar{\sigma}$ of $\{\bar{1},\dots,\bar{k}\}$, such that $\sigma \cup \bar{\sigma}$ is a non-crossing partition of $\{1,\bar{1},\dots,k,\bar{k}\}$. For example,

$$
\begin{gathered}
K(\{(1,4),(2,3)\}) = \{(1,3),(2),(4)\},\\
K(\{(1,2),(3,4),(5,6)\}) = \{(1),(2,4,6),(3),(5)\}.
\end{gathered} \tag{5.77}
$$

The second example is illustrated as



For $\sigma \in NC_2(k)$ and $N \geq 1$, define

$$S(\sigma, N) = \big\{ i \in \{1, \ldots, N\}^k \colon\ i_u = i_v \iff u, v \text{ belong to the same block of } K(\sigma) \big\} \tag{5.78}$$

and

$$C(k, N) = \{1, \ldots, N\}^k \setminus \Bigg(\bigcup_{\sigma \in NC_2(k)} S(\sigma, N) \Bigg). \tag{5.79}$$

In other words, $S(\sigma, N)$ is the same as $\Psi(K(\sigma), N)$ defined in (5.67).

Lemma 5.2.6 (Trace of product of random matrices).
Suppose that, for each $N \geq 1$, $W_{N,1}, \ldots, W_{N,k}$ are $N \times N$ real (and possibly asymmetric) random matrices, where k is a positive even number. Suppose further that, for each $u = 1, \ldots, k$,

$$\max_{1 \leq i,j \leq N} \mathbb{E}\big[W_{N,u}(i,j)^k \big] = O(N^{-k/2}) \tag{5.80}$$

and

$$\lim_{N \to \infty} \mathbb{E}\Bigg[\Bigg(\frac{1}{N} \sum_{i \in C(k,N)} P_i \Bigg)^2 \Bigg] = 0, \tag{5.81}$$

and that, for every $\sigma \in NC_2(k)$, there exists a deterministic and finite $\beta(\sigma)$ such that

$$\lim_{N \to \infty} \mathbb{E}\Bigg(\frac{1}{N} \sum_{i \in S(\sigma,N)} P_i \Bigg) \quad = \quad \beta(\sigma), \tag{5.82}$$

$$\lim_{N \to \infty} \mathbb{E}\Bigg[\Bigg(\frac{1}{N} \sum_{i \in S(\sigma,N)} P_i \Bigg)^2 \Bigg] \quad = \quad \beta(\sigma)^2, \tag{5.83}$$

where

$$P_i = W_{N,1}(i_1, i_2) \cdots W_{N,k-1}(i_{k-1}, i_k) W_{N,k}(i_k, i_1), \qquad i \in \{1, \ldots, N\}^k. \tag{5.84}$$

Furthermore, let $V_1, V_2, \ldots$ be i.i.d. random variables drawn from some distribution with all moments finite, independent of $(W_{N,j} \colon\ N \geq 1,\ 1 \leq j \leq k)$, and let

$$U_N = \mathrm{Diag}(V_1, \ldots, V_N), \qquad N \geq 1. \tag{5.85}$$

Then, for all choices of $n_1, \ldots, n_k \geq 0$,

$$\lim_{N \to \infty} \frac{1}{N} \mathrm{Tr}\left(U_N^{n_1} W_{N,1} \cdots U_N^{n_k} W_{N,k} \right) = c \quad \text{ in } L^2 \tag{5.86}$$

for some deterministic $c \in \mathbb{R}$ (depending on $k, n_1, \ldots, n_k$).

Proof. The fact that the sets $S(\sigma, N)$ are disjoint for different $\sigma \in NC_2(k)$ allows us to write

$$\mathrm{Tr}\left(U_N^{n_1} W_{N,1} \cdots U_N^{n_k} W_{N,k} \right) = \sum_{\sigma \in NC_2(k)} \sum_{i \in S(\sigma,N)} \tilde{P}_i + \sum_{i \in C(k,N)} \tilde{P}_i, \tag{5.87}$$

Chapter 5

where

$$\tilde{P}_i = \prod_{j=1}^{k} \big(V_{i_j}^{n_j} W_{N,j}(i_j, i_{j+1})\big), \qquad i \in \{1, \ldots, N\}^k. \tag{5.88}$$

In order to show that the second sum in the right-hand side is negligible after scaling by N, the independence of $(V_1, V_2, \ldots)$ and $(W_{N,j}\colon\ N \geq 1,\ 1 \leq j \leq k)$, together with the fact that the common distribution of the former has finite moments, implies that

$$\mathbb{E}\left[\left(\frac{1}{N}\sum_{i\in C(k,N)} \tilde{P}_i\right)^2\right] \leq KN^{-2} \sum_{i,j\in C(k,N)} \mathbb{E}[P_i P_j],$$

where K is a finite constant. Assumption (5.81) shows that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{i\in C(k,N)} \tilde{P}_i = 0 \quad \text{in } L^2. \tag{5.89}$$

In order to complete the proof, it suffices to show that for every $\sigma \in NC_2(k)$ there exists a $\theta(\sigma) \in \mathbb{R}$ with

$$\lim_{N\to\infty} \frac{1}{N} \sum_{i\in S(\sigma,N)} \tilde{P}_i = \theta(\sigma) \qquad \text{in } L^2. \tag{5.90}$$

To that end, fix $\sigma \in NC_2(k)$ and note that, for $i \in S(\sigma, N)$,

$$\mathbb{E}[\tilde{P}_i] = \mathbb{E}[P_i]\mathbb{E}\left[\prod_{j=1}^{k} V_{i_j}^{n_j}\right] = \mathbb{E}[P_i] \prod_{u\in K(\sigma)} \mathbb{E}\left[V_1^{\sum_{j\in u} n_j}\right], \tag{5.91}$$

the product in the last line being taken over every block u of $K(\sigma)$. Putting

$$\theta(\sigma) = \beta(\sigma) \prod_{u\in K(\sigma)} \mathbb{E}\left[V_1^{\sum_{j\in u} n_j}\right], \tag{5.92}$$

we see that (5.82) gives

$$\lim_{N\to\infty} \mathbb{E}\left[\frac{1}{N}\sum_{i\in S(\sigma,N)} \tilde{P}_i\right] = \theta(\sigma). \tag{5.93}$$

Let us call $i, j \in \mathbb{N}^k$ “disjoint” if no coordinate of i matches any coordinate of j, i.e.,

$$\min_{1\leq u,v\leq k} |i_u - j_v| \geq 1. \tag{5.94}$$

Since $K(\sigma)$ has exactly $\frac{1}{2}k + 1$ blocks, (5.80) implies that

$$\lim_{N\to\infty} N^{-2} \sum_{\substack{i,j\in S(\sigma,N)\\ i,j \text{ not disjoint}}} \mathbb{E}[\tilde{P}_i \tilde{P}_j] = 0. \tag{5.95}$$

If $i, j \in S(\sigma, N)$ are disjoint, then it is immediate that

$$\mathbb{E}[\tilde{P}_i \tilde{P}_j] = \Big(\prod_{u \in K(\sigma)} \mathbb{E}\big[V_1^{\sum_{j \in u} n_j}\big] \Big)^2 \mathbb{E}[P_i P_j]. \tag{5.96}$$

The above two displays, in conjunction with (5.83), show that

$$\lim_{N \to \infty} \mathbb{E}\Bigg[\Big(\frac{1}{N} \sum_{i \in S(\sigma, N)} \tilde{P}_i \Big)^2 \Bigg] = \theta(\sigma)^2. \tag{5.97}$$

This, along with (5.93), establishes (5.90), from which the proof follows. □

§5.3 Proofs of the main results

In this section we prove the theorems in Section 5.1. In Section 5.3.1 we prove Theorems 5.1.2–5.1.3 on the existence of the limiting spectral distributions of A_N and Δ_N. In Section 5.3.2 we identify those distributions by proving Theorem 5.1.6. In Section 5.3.3 we prove Theorem 5.1.7.

§5.3.1 Proof: existence

Proof of Theorem 5.1.2. From [129, Theorem 2.1] we know that, as $N \to \infty$,

$$\lim_{N \to \infty} \mathrm{ESD}(\bar{A}_N) = \mu \quad \text{weakly in probability,} \tag{5.98}$$

for a compactly supported symmetric probability measure μ. Lemma 5.2.3 immediately tells us that

$$\lim_{N \to \infty} \mathrm{ESD}(A_N^g) = \mu \quad \text{weakly in probability,} \tag{5.99}$$

and hence for h and H_N as in Lemma 5.2.2,

$$\lim_{N \to \infty} \mathbb{E}\big[h\big(\Re H_N(A_N^g)\big)\big] = h\Big(\Re \int_{\mathbb{R}} \frac{1}{x - z}\, \mu(dx)\Big). \tag{5.100}$$

The claim in (5.39) shows that A_N^g can be replaced by A_N^0 in the above display. Since the right-hand side is deterministic and the above holds for any h satisfying the hypothesis of Lemma 5.2.2, it follows that

$$\lim_{N \to \infty} \Re H_N(A_N^0) = \Re \int_{\mathbb{R}} \frac{1}{x - z}\, \mu(dx) \quad \text{in probability.} \tag{5.101}$$

A similar argument works for the imaginary part, which shows that

$$\lim_{N \to \infty} \mathrm{ESD}(A_N^0) = \mu \quad \text{weakly in probability.} \tag{5.102}$$

Lemma 5.2.1 completes the proof of (5.5).

Finally, if f is bounded away from 0, then the combination of [129, Lemma 3.1] and [120, Corollary 2] implies that μ is absolutely continuous with respect to the Lebesgue measure (see also [131]). □

A close inspection of the proof reveals that it suffices to assume that f is bounded and Riemann integrable instead of continuous. In other words, if f is symmetric and bounded, and its set of discontinuities has Lebesgue measure zero, then the result holds. However, continuity will be used later in (5.107) in the proof of Theorem 5.1.3. Furthermore, if $\varepsilon_N = 1$ for all N, then

$$\lim_{N\to\infty} \mathrm{ESD}\big(N^{-1/2}(A_N - \mathbb{E}(A_N))\big) = \mu_{\sqrt{f(1-f)}} \quad \text{weakly in probability,} \tag{5.103}$$

where the right-hand side is the probability measure obtained after replacing f with $\sqrt{f(1-f)}$ in [129, Theorem 2.1].

Proof of Theorem 5.1.3. The proof comes in three steps.

1 (Riemann approximation). For $N \geq 1$, define the $N \times N$ diagonal matrix Q_N by

$$Q_N(i,i) = F(i/N) Z_i, \qquad 1 \leq i \leq N, \tag{5.104}$$

where

$$F(x) = \left(\int_0^1 f(x,y)\, dy \right)^{1/2}, \qquad x \in [0,1], \tag{5.105}$$

and $(Z_i\colon\ i \geq 1)$ is as in Lemma 5.2.4. Lemma E.2 implies that

$$\left| \left(\frac{1}{N} \mathrm{Tr}\left((\tilde{\Delta}_N)^k\right)\right)^{1/k} - \left(\frac{1}{N} \mathrm{Tr}\left((\bar{A}_N + Q_N)^k\right)\right)^{1/k} \right| \leq \left(\frac{1}{N} \mathrm{Tr}\left((Y_N - Q_N)^k\right)\right)^{1/k}. \tag{5.106}$$

Since, f being continuous,

$$\begin{aligned} &\mathbb{E}\big[N^{-2}\, \mathrm{Tr}^2\left((Y_N - Q_N)^k\right)\big] \\ &= O(1) \sup_{x\in[0,1]} \left[F(x) - \left(\frac{1}{N} \sum_{j=1, j\neq [Nx]/N}^{N} f\left(x, \frac{j}{N}\right)\right)^{1/2} \right]^{2k}, \qquad N \to \infty, \end{aligned} \tag{5.107}$$

and it tends to 0 as $N \to \infty$, we get that, for every even k,

$$\left(\frac{1}{N} \mathrm{Tr}\left((\tilde{\Delta}_N)^k\right)\right)^{1/k} - \left(\frac{1}{N} \mathrm{Tr}\left((\bar{A}_N + Q_N)^k\right)\right)^{1/k} \tag{5.108}$$

tends to 0 in L^{2k} as $N \to \infty$.

Our next step is to show that, for every even integer k,

$$\lim_{N\to\infty} \frac{1}{N} \mathrm{Tr}\left((\bar{A}_N + Q_N)^k\right) = \gamma_k \quad \text{in } L^2 \tag{5.109}$$

for some $\gamma_k \in \mathbb{R}$. The above will follow once we show that, for all $m \geq 1$ and $n_1, \ldots, n_m \geq 0$,

$$\lim_{N\to\infty} \frac{1}{N} \mathrm{Tr}\left(Q_N^{n_1} \bar{A}_N \cdots Q_N^{n_m} \bar{A}_N\right) = \theta \quad \text{in } L^2 \tag{5.110}$$

for some $\theta \in \mathbb{R}$ (depending on $m, n_1, \ldots, n_m$). To that end, define the diagonal matrices U_N and B_N by

$$U_N(i,i) = Z_i \quad \text{and} \quad B_N(i,i) = F(i/N), \tag{5.111}$$

for $i = 1, \ldots, N$. Observe that

$$Q_N = B_N U_N = U_N B_N, \tag{5.112}$$

and hence the left-hand side of (5.110) is the same as

$$\frac{1}{N} \operatorname{Tr}\left(U_N^{n_1} W_{N,1} \cdots U_N^{n_m} W_{N,m}\right), \tag{5.113}$$

where

$$W_{N,j} = B_N^{n_j} \bar{A}_N, \qquad j = 1, \ldots, m. \tag{5.114}$$

In order to apply Lemma 5.2.6 we need to verify its hypotheses.

2 (Verification of the hypotheses). Our next claim is that $W_{N,1}, \ldots, W_{N,m}$ satisfy (5.80)–(5.83). To that end, observe that for $N \geq 1$ and $j = 1, \ldots, m$,

$$W_{N,j}(u,v) = F^{n_j}\left(\frac{u}{N}\right) f^{1/2}\left(\frac{u}{N}, \frac{v}{N}\right) N^{-1/2} G_{u \wedge v,\, u \vee v}, \qquad 1 \leq u, v \leq N. \tag{5.115}$$

Let

$$H_j(x,y) = F^{n_j}(x) f^{1/2}(x,y), \qquad (x,y) \in [0,1]^2. \tag{5.116}$$

Fix a partition Π of $\{1, \ldots, m\}$. Recall the notation $\Psi(\Pi, N)$ introduced in the proof of Lemma 5.2.4. Clearly, for every $i \in \Psi(\Pi, N)$,

$$\mathbb{E}\left[\prod_{j=1}^m W_{N,j}(i_j, i_{j+1})\right] = N^{-m/2} \psi(\Pi) \left(\prod_{j=1}^m H_j\left(\frac{i_j}{N}, \frac{i_{j+1}}{N}\right)\right), \tag{5.117}$$

where

$$\psi(\Pi) = \mathbb{E}\left[\prod_{j=1}^m G_{i_j \wedge i_{j+1}, i_j \vee i_{j+1}}\right], \tag{5.118}$$

which does not depend on $i \in \Psi(\Pi, N)$. The standard arguments leading to a proof via the method of moments of the Wigner semicircle law show that

$$\begin{aligned} &\lim_{N \to \infty} N^{-m/2+1}\, \psi(\Pi)\, |\Psi(\Pi, N)| \\ &\quad = \begin{cases} 1, & \text{if } m \text{ is even, and } \Pi = K(\sigma) \text{ for some } \sigma \in NC_2(m), \\ 0, & \text{otherwise.} \end{cases} \end{aligned} \tag{5.119}$$

Assume for the moment that m is even, and let $\sigma \in NC_2(m)$. It is known that $K(\sigma)$ has $m/2 + 1$ blocks. Define a function $\mathcal{L}_\sigma : \{1, \ldots, m\} \to \{1, \ldots, \frac{1}{2}m + 1\}$ such that

$\mathcal{L}_\sigma(j) = \mathcal{L}_\sigma(k)$ if and only if j, k are in the same block of $K(\sigma)$. It follows that for $\Pi = K(\sigma)$,

$$\begin{aligned}
&\lim_{N\to\infty} \frac{1}{N} \sum_{i\in\Psi(\Pi,N)} \mathbb{E}\Big[\prod_{j=1}^{m} W_{N,j}(i_j, i_{j+1})\Big] \\
&= \int_{[0,1]^{(m/2)+1}} \prod_{(u,v)\in\sigma, u<v} H_u\big(x_{\mathcal{L}_\sigma(u)}, x_{\mathcal{L}_\sigma(v)}\big)\, dx_1 \cdots dx_{(m/2)+1}.
\end{aligned} \tag{5.120}$$

This shows that hypothesis (5.82) holds. The hypotheses (5.81) and (5.83) follow similarly by an analogue of the standard arguments, while (5.80) is trivial.

Thus, $W_{N,1}, \ldots, W_{N,m}$ and U_N satisfy the hypotheses of Lemma 5.2.6. The claim of that lemma shows that the random variable in (5.113) converges in L^2 to a finite deterministic constant as $N \to \infty$, i.e., (5.110) holds. This in turn proves (5.109), which in conjunction with (5.108) shows that

$$\lim_{N\to\infty} \frac{1}{N} \operatorname{Tr}\big((\tilde{\Delta}_N)^k\big) = \gamma_k \quad \text{in } L^2. \tag{5.121}$$

Lemma 5.2.4 asserts that

$$\lim_{N\to\infty} \frac{1}{N} \operatorname{Tr}\big((\bar{\Delta}_N)^k\big) = \gamma_k \quad \text{in } L^2, \tag{5.122}$$

and hence also in probability.

3 (Uniqueness of the limiting measure). Equation (5.109) ensures that there exists a symmetric probability measure on $\mathbb{R}$ whose k-th moment is γ_k for every even integer k. Our next claim is that such a measure is unique, i.e., $(\gamma_k\colon\ k \geq 1)$ determines the measure. It is not obvious how to check Carleman's condition, and therefore we argue as follows. It suffices to exhibit a probability measure ν whose odd moments are zero and whose k-th moment is γ_k for even k such that

$$\int_{\mathbb{R}} e^{tx} \nu(dx) < \infty \qquad \forall\, t \in \mathbb{R}. \tag{5.123}$$

To do so we bring in the notion of a non-commutative probability space (NCP), which is defined in Appendix E. For $K > 0$ and $N \geq 1$, define

$$U_{NK} = \operatorname{Diag}\big(Z_1 \mathbf{1}(|Z_1| \leq K), \ldots, Z_1 \mathbf{1}(|Z_N| \leq K)\big), \tag{5.124}$$

and

$$Q_{NK} = B_N U_{NK}. \tag{5.125}$$

The arguments leading to (5.110) can be easily tweaked to show that, for fixed $K > 0$ and a fixed polynomial p in two non-commuting variables,

$$\lim_{N\to\infty} \frac{1}{N} \mathbb{E}\big[\operatorname{Tr}\big(p(\bar{A}_N, Q_{NK})\big)\big] \tag{5.126}$$

exists. Lemma E.4 implies that there exist self-adjoint elements q and a in a tracial NCP $(\mathcal{A}, \phi)$ such that the above limit equals $\phi\left[p\left(a, q\right)\right]$ for every polynomial p in two non-commuting variables. Hence

$$\lim_{N\to\infty} \mathrm{EESD}\big[p(\bar{A}_N, Q_{NK})\big] = \mathcal{L}(p\left(a, q\right)) \quad \text{in distibution,} \tag{5.127}$$

for any symmetric polynomial p, where EESD denotes the expectation of ESD. Theorem 5.1.2 implies that the limiting spectral distribution of $\bar{A}_N$, which is $\mathcal{L}(a)$ by (5.127), is compactly supported, and hence a is a bounded element. The spectrum of q is clearly a subset of $[-K, K]$. The second claim in Lemma E.4 allows us to assume that $(\mathcal{A}, \phi)$ is a W^*-probability space.

Let

$$\nu_K = \mathcal{L}(a + q). \tag{5.128}$$

If C is a finite constant such that

$$-C\mathbf{1} \leq a \leq C\mathbf{1}, \tag{5.129}$$

then clearly

$$a + q \leq C\mathbf{1} + q. \tag{5.130}$$

Applying the method of moments to Q_{NK}, we find by an appeal to (5.127) that the law of q is the same as the law of

$$F(V)Z_1\mathbf{1}(|Z_1| \leq K),$$

where V is standard uniform independently of Z_1, and F is as in (5.105). Under the assumption that $f \leq 1$, which represents no loss of generality,

$$\int_{\mathbb{R}} e^{tx}\big(\mathcal{L}(q)\big)(dx) \leq e^{t^2/2}, \qquad t \in \mathbb{R}. \tag{5.131}$$

By [118, Corollary 3.3] applied to (5.130), it follows that

$$\int_{\mathbb{R}} e^{tx}\nu_K(dx) \leq \int_{\mathbb{R}} e^{tx}\big(\mathcal{L}(C\mathbf{1} + q)\big)(dx) \leq \exp\left(\tfrac{1}{2}t^2 + tC\right), \qquad t > 0. \tag{5.132}$$

Lemma E.1 applied to $\bar{A}_N + Q_{NK_1}$ and $\bar{A}_N + Q_{NK_1}$ shows that

$$\sup_{N\geq 1} L\big(\mathrm{EESD}[\bar{A}_N + Q_{NK_1}], \mathrm{EESD}[\bar{A}_N + Q_{NK_2}]\big) \tag{5.133}$$

is small for large K_1 and K_2. Thus, $(\nu_K\colon\ K > 0)$ is Cauchy in the Lévy metric, and hence there exists a probability measure ν such that

$$\lim_{K\to\infty} \nu_K = \nu. \tag{5.134}$$

This, along with (5.132), establishes that

$$\int_{\mathbb{R}} e^{tx}\nu(dx) \leq \exp\left(\tfrac{1}{2}t^2 + tC\right), \qquad t > 0, \tag{5.135}$$

and

$$\lim_{K\to\infty}\int_{\mathbb{R}} x^k \nu_K(dx) = \int_{\mathbb{R}} x^k \nu(dx), \qquad k \geq 1. \tag{5.136}$$

Clearly,

$$\int_{\mathbb{R}} x^k \nu_K(dx) = \lim_{N\to\infty} N^{-1}\mathbb{E}\big[\operatorname{Tr}\big((\bar{A}_N + Q_{NK})^k\big)\big]. \tag{5.137}$$

Therefore, by keeping track of the limit in (5.126), we can show (with some effort) that

$$\lim_{K\to\infty}\int_{\mathbb{R}} x^k \nu_K(dx) = \begin{cases}\gamma_k, & k \text{ even},\\ 0, & k \text{ odd}.\end{cases} \tag{5.138}$$

Thus, ν has the desired moments. By extending (5.135) to the case $t < 0$, we see that (5.123) follows. Thus, ν is the only symmetric probability measure whose even moments are (γ_k).

Equation (5.122) and the claim proved above show that

$$\lim_{N\to\infty} \mathrm{ESD}(\bar{\Delta}_N) = \nu \quad \text{weakly in probability}. \tag{5.139}$$

Hence Lemmas 5.2.1–5.2.3 imply that

$$\lim_{N\to\infty} \mathrm{ESD}\big((N\varepsilon_N)^{-1/2}(\Delta_N - D_N)\big) = \nu \quad \text{weakly in probability}. \tag{5.140}$$

as in the proof of Theorem 5.1.2.

It remains to show that if f is not identically zero, then the support of ν is unbounded. To that end, recall that (5.109), together with the fact that ν is the only symmetric probability measure whose even moments are (γ_k), establish that

$$\lim_{N\to\infty} \mathrm{ESD}(\bar{A}_N + Q_N) = \nu \quad \text{weakly in probability}, \tag{5.141}$$

where $\bar{A}_N$ and Q_N are as in (5.55) and (5.104), respectively. Fix $0 < p < 1/2$, and for any $N \times N$ real symmetric matrix Σ, enumerate its eigenvalues in descending order by $\lambda_1(\Sigma), \dots, \lambda_N(\Sigma)$. Weyl's inequality (see [189, Equation (1.54)]) implies that

$$\lambda_{2\lceil Np\rceil-1}(Q_N) \leq \lambda_{\lceil Np\rceil}(\bar{A}_N + Q_N) + \lambda_{\lceil Np\rceil}(-\bar{A}_N), \tag{5.142}$$

where $\lceil x\rceil$ denotes the smallest integer larger than or equal to x. Therefore

$$\begin{aligned}\limsup_{N\to\infty} \lambda_{\lceil Np\rceil}(\bar{A}_N + Q_N) &\geq \limsup_{N\to\infty} \lambda_{2\lceil Np\rceil-1}(Q_N) - \liminf_{N\to\infty} \lambda_{\lceil Np\rceil}(-\bar{A}_N)\\ &\geq \limsup_{N\to\infty} \lambda_{2\lceil Np\rceil-1}(Q_N) - C,\end{aligned} \tag{5.143}$$

where C is as in (5.129). Letting $p \to 0$ and appealing to Lemma E.6, we find that

$$\begin{aligned}\sup(\mathrm{Supp}(\nu)) &= \lim_{p\to 0}\limsup_{N\to\infty} \lambda_{\lceil Np\rceil}(\bar{A}_N + Q_N)\\ &\geq \lim_{p\to 0}\limsup_{N\to\infty} \lambda_{2\lceil Np\rceil-1}(Q_N) - C = \infty,\end{aligned} \tag{5.144}$$

where the last line uses the fact that, as $N \to \infty$, $\mathrm{ESD}(Q_N)$ converges weakly in probability to the distribution of $F(V)Z_1$, the support of which is unbounded because f is not identically zero. $\square$

§5.3.2 Proof: identification

Proof of Theorem 5.1.6. Let $(G_{i,j}\colon\ 1 \le i \le j)$ and $(Z_i\colon\ i \ge 1)$ be as in Lemma 5.2.4. For $N \ge 1$, define the $N \times N$ matrices

$$\begin{aligned} G_N(i,j) &= N^{-1/2} G_{i\wedge j, i\vee j}, \qquad 1 \le i,j \le N, && (5.145)\\ R_N &= \mathrm{Diag}\left(\sqrt{r(1/N)}, \ldots, \sqrt{r(1)}\right), && (5.146)\\ U_N &= \mathrm{Diag}(Z_1, \ldots, Z_N). && (5.147) \end{aligned}$$

The notation U_N is exactly as in the proof of Theorem 5.1.3. Let $\bar{A}_N$ and Q_N be as in (5.55) and (5.104), respectively. Observe that, under the assumption (5.16),

$$\bar{A}_N = R_N G_N R_N, \tag{5.148}$$

and

$$Q_N = \alpha R_N^{1/2} U_N R_N^{1/2}, \tag{5.149}$$

where α is as defined in the statement of Theorem 5.1.6. Proceeding as in the proofs of Theorems 5.1.2–5.1.3, wee see that it suffices to show that

$$\lim_{N\to\infty} \mathrm{ESD}\left(R_N G_N R_N\right) = \mathcal{L}\big(r^{1/2}(T_u) T_s T^{1/2}(T_u)\big) \quad \text{weakly in probability} \tag{5.150}$$

and

$$\begin{aligned} &\lim_{N\to\infty} \mathrm{ESD}\big(R_N G_N R_N + \alpha R_N^{1/2} U_N R_N^{1/2}\big)\\ &\quad = \mathcal{L}\big(r^{1/2}(T_u) T_s T^{1/2}(T_u) + \alpha r^{1/4}(T_u) T_g r^{1/4}(T_u)\big) \quad \text{weakly in probability,} \end{aligned} \tag{5.151}$$

where T_s, T_g, T_u are as in the statement. Define U_{NK} to be the "truncated" version of U_N, for a fixed $K > 0$, as in the proof of Theorem 5.1.3. Both (5.150) and (5.151) will follow once we show that

$$\lim_{N\to\infty} \frac{1}{N} \mathrm{Tr}\left(p\big(R_N^{1/2}, U_{NK}, G_N\big)\right) = \tau\big(p(T_r, T_g', T_s)\big) \quad \text{in probability,} \tag{5.152}$$

where $T_r = r^{1/4}(T_u)$ and $T_g' = T_g \mathbb{1}_{\{|T_g| \le K\}}$, for any symmetric polynomial p in three non-commuting variables. It is a well known fact that, for all $k \ge 1$,

$$\lim_{N\to\infty} \frac{1}{N} \mathrm{Tr}(G_N^k) = \tau(T_s^k) \quad \text{in probability.} \tag{5.153}$$

Since R_N and U_{NK} are diagonal matrices, they commute. This, in conjunction with the strong law of large numbers, implies that, for any $k \ge 1$, $m_1, \ldots, m_k$ and $n_1, \ldots, n_k \ge 0$,

$$\begin{aligned} &\lim_{N\to\infty} \frac{1}{N} \mathrm{Tr}\left(R_N^{m_1} U_{NK}^{n_1} \cdots R_N^{m_k} U_{NK}^{n_k}\right)\\ &\quad = \int_0^1 r^{(m_1+\ldots+m_k)/4}(u)\, du \int_{-K}^{K} (2\pi)^{-1/2} x^{n_1+\ldots+n_k} e^{-x^2/2}\, dx \quad \text{almost surely} \end{aligned} \tag{5.154}$$

The above, in conjunction with (5.26) and the fact that T_g and T_r commute, implies that

$$\lim_{N\to\infty}\frac{1}{N}\operatorname{Tr}\left(p\big(R_N^{1/2},U_{NK}\big)\right)=\tau\big(p(T_r,T_g')\big)\quad\text{almost surely}\tag{5.155}$$

for any polynomial p in two variables.

Thus, all that remains to show is the asymptotic free independence of T_s and (T_r, T_g'), which is precisely the claim of Lemma E.5, i.e., (5.153) and (5.155) imply (5.152). Applying (5.152) to $p(x,y,z)=x^2zx^2$ and $p(x,y,z)=x^2zx^2+\alpha xyx$, we get the truncated versions of (5.150) and (5.151), respectively. Yet another application of Lemma E.1 allows us to let $K\to\infty$, obtaining (5.150) and (5.151). This completes the proof of (5.23) and (5.24). □

§5.3.3 Proof: randomization

Proof of Theorem 5.1.7. As before, Lemma 5.2.1 and (5.2) imply that the mean of the entries of A_N can be subtracted at the cost of a negligible perturbation of the ESD. The inequalities (5.2) and (5.27) ensure that the Gaussianization as in Lemma 5.2.2 goes through by conditioning on $R_{N1},\dots,R_{NN}$. That is, if $(G_{ij}\colon 1\le i\le j)$ is a collection of i.i.d. standard normal random variables that are independent of $(R_{Ni}\colon 1\le i\le N, N\ge 1)$, W_N^g is an $N\times N$ matrix defined by

$$W_N^g(i,j)=G_{i\wedge j,i\vee j},1\le i,j\le N,\tag{5.156}$$

and

$$\Theta_N=\operatorname{Diag}\left(\sqrt{R_{N1}},\dots,\sqrt{R_{NN}}\right),\tag{5.157}$$

then the ESD of $A_N/\sqrt{N\varepsilon_N}$ is close to that of $\Theta_N W_N^g\Theta_N/\sqrt{N}$.

The assumptions (5.27) and (5.28) imply that, for $k\ge 1$,

$$\lim_{N\to\infty}\frac{1}{N}\operatorname{Tr}\left(\Theta_N^{2k}\right)=\int_{\mathbb{R}}x^k\mu_r(dx)\quad\text{almost surely.}\tag{5.158}$$

Finally, Lemma E.5 together with (5.27) shows the asymptotic free independence of Θ_N and W_N^g, that is,

$$\lim_{N\to\infty}\operatorname{ESD}\big(N^{-1/2}\Theta_N W_n^g\Theta_N\big)=\mu_r\boxtimes\mu_s\quad\text{weakly in probability.}\tag{5.159}$$

This completes the proof. □

§5.4 Applications

In this section we discuss three applications, explained in Sections 5.4.1–5.4.3.

§5.4.1 Constrained random graphs

Let $\mathcal{S}_N$ be the set of all simple graphs on N vertices. Suppose that we fix the degrees of the vertices, namely, vertex i has degree k_i^*. Here, $k^*=(k_i^*\colon 1\le i\le N)$ is a sequence of positive integers of which we only require that they are graphical, i.e., there is at least one simple graph with these degrees. The so-called *canonical ensemble* P_N is the unique probability distribution on S_N with the following two properties.

(I) The *average degree* of vertex i, defined by $\sum_{G\in\mathcal{S}_N} k_i(G)P_N(G)$, equals k_i^* for all $1 \le i \le N$.

(II) The *entropy* of P_N, defined by $-\sum_{G\in\mathcal{S}_N} P_N(G)\log P_N(G)$, is maximal.

The name canonical ensemble comes from Gibbs theory in equilibrium statistical physics. The probability distribution P_N describes a random graph of which we have no prior information other than the average degrees, and is called the *soft configuration model.* It is known that, because of property (II), P_N takes the form (see [169])

$$P_N(G) = \frac{1}{Z_N(\theta^*)} \exp\Bigg[-\sum_{i=1}^{N} \theta_i^* k_i(G)\Bigg], \qquad G \in \mathcal{S}_N, \tag{5.160}$$

where $\theta^* = (\theta_i^*\colon 1 \le i \le N)$ is a sequence of real-valued Lagrange multipliers that must be chosen in such a way that property (I) is satisfied. The normalization constant $Z_N(\theta^*)$, which depends on θ^*, is called the partition function in Gibbs theory.

The gradients of the constraints in property (I) are linearly independent vectors and the matching of property (I) uniquely fixes θ^*. It turns out that

$$P_N(G) = \prod_{1\le i<j\le N} (p_{ij}^*)^{A_N[G](i,j)}\,(1-p_{ij}^*)^{1-A_N[G](i,j)}, \qquad G \in \mathcal{S}_N, \tag{5.161}$$

where $A_N[G]$ is the adjacency matrix of G, and p_{ij}^* represent a reparameterisation of the Lagrange multipliers, namely,

$$p_{ij}^* = \frac{x_i^* x_j^*}{1 + x_i^* x_j^*}, \qquad 1 \le i \ne j \le N, \tag{5.162}$$

with $x_i^* = e^{-\theta_i^*}$ (see [187] for more details). Thus, we see that P_N is nothing other than an inhomogeneous Erdős-Rényi random graph where the probability that vertices i and j are connected by an edge equals p_{ij}^*. In order to match property (I), these probabilities must satisfy

$$k_i^* = \sum_{j=1, j\ne i}^{N} p_{ij}^*, \qquad 1 \le i \le N, \tag{5.163}$$

which constitutes a set of N equations for the N unknowns $x_1^*, \dots, x_N^*$.

In order to state the next result, we need to make some assumptions on the sequence $(k_{Ni}^*\colon 1 \le i \le N)$. For the sake of notational simplification, the dependence on N will be suppressed from the notation.

Proposition 5.4.1 (Theorem 5.1.7 for constrained random graphs). *Let $(k_i^*\colon 1 \le i \le N)$ be a graphical sequence of positive integers. Define*

$$m_N = \max_{1\le \ell\le N} k_\ell^*. \tag{5.164}$$

Assume that

$$\lim_{N\to\infty} m_N = \infty, \qquad \lim_{N\to\infty} m_N/\sqrt{N} = 0, \tag{5.165}$$

and

$$\lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} \delta_{k_i^*/m_N} = \mu_r \qquad \textit{weakly}, \tag{5.166}$$

for some probability measure μ_r. Let x_i^* and p_{ij}^* be determined by (5.162) *and* (5.163). *Let A_N be the adjacency matrix of an inhomogeneous Erdős-Rényi random graph on N vertices, with p_{ij}^* the probability of an edge being present between vertices i and j for $1 \le i \ne j \le N$. Then*

$$\lim_{N\to\infty} \mathrm{ESD}\big((N\varepsilon_N)^{-1/2} A_N\big) = \mu_r \boxtimes \mu_s \qquad \textit{weakly in probability.} \tag{5.167}$$

Proof. Abbreviate

$$\sigma_N = \sum_{\ell=1}^{N} k_\ell^*. \tag{5.168}$$

In [187] it is shown that

$$\max_{1\le \ell\le N} x_\ell^* = o(1), \qquad N\to\infty, \tag{5.169}$$

in which case (5.162) and (5.163) give

$$x_i^* = \frac{k_i^*}{\sqrt{\sigma_N}}\,[1+o(1)] \quad \text{ and } \quad p_{ij}^* = \frac{k_i^* k_j^*}{\sqrt{\sigma_N}}\,[1+o(1)], \qquad N\to\infty, \tag{5.170}$$

with the error term uniform in $1 \le i \ne j \le N$. Pick

$$\varepsilon_N = \frac{m_N^2}{\sigma_N}. \tag{5.171}$$

It follows from (5.165) that

$$\lim_{N\to\infty} \varepsilon_N = 0, \qquad \lim_{N\to\infty} N\varepsilon_N = \infty. \tag{5.172}$$

As in the proof of Theorem 5.1.7, Lemmas 5.2.1–5.2.2 imply that the upper triangular entries of A_N can be replaced by independent mean-zero normal random variables. In other words, if $(G_{ij}\colon\ 1 \le i \le j)$ are i.i.d. standard normal, and A_N^g is the random matrix defined by

$$A_N^g(i,j) = \sqrt{p_{ij}^*}\, G_{i\wedge j, i\vee j}, \qquad 1 \le i,j \le N, \tag{5.173}$$

with $p_{ii}^* = 0$ for all $1 \le i \le N$, then $\mathrm{ESD}\big((N\varepsilon_N)^{-1/2} A_N\big)$ and $\mathrm{ESD}\big((N\varepsilon_N)^{-1/2} A_N^g\big)$ are asymptotically close. The second part of (5.170) implies that

$$\sqrt{p_{ij}^*} = \sqrt{\varepsilon_N \frac{k_i^* k_j^*}{m_N^2}\,[1+o(1)]}, \qquad N\to\infty, \tag{5.174}$$

uniformly in $1 \le i \ne j \le N$, and hence

$$\sum_{i,j=1}^{N} \left(\sqrt{p_{ij}^*} - \sqrt{\varepsilon_N \frac{k_i^* k_j^*}{m_N^2}} \right)^2 = o(N^2 \varepsilon_N), \qquad N \to \infty. \tag{5.175}$$

In other words, if $\tilde{A}_N$ is defined by

$$\tilde{A}_N(i,j) = \sqrt{\frac{k_i^* k_j^*}{m_N^2}}\, G_{i \wedge j, i \vee j}, \qquad 1 \le i,j \le N, \tag{5.176}$$

then

$$\lim_{N\to\infty} \frac{1}{N} \mathbb{E}\big[\operatorname{Tr}\big((N\varepsilon_N)^{-1/2} A_N^g - N^{-1/2}\tilde{A}_N\big)^2\big] = 0. \tag{5.177}$$

Lemma E.1 implies that

$$\lim_{N\to\infty} L\big(\mathrm{ESD}\big((N\varepsilon_N)^{-1/2} A_N^g\big), \mathrm{ESD}\big(N^{-1/2}\tilde{A}_N\big)\big) = 0 \quad \text{in probability.} \tag{5.178}$$

Finally, by an appeal to Lemma E.5, (5.166) implies that

$$\lim_{N\to\infty} \mathrm{ESD}\big(N^{-1/2}\tilde{A}_N\big) = \mu_r \boxtimes \mu_s \quad \text{weakly in probability,} \tag{5.179}$$

where μ_s is the standard semicircle law. Hence

$$\lim_{N\to\infty} \mathrm{ESD}\big((N\varepsilon_N)^{-1/2} A_N\big) = \mu_r \boxtimes \mu_s \quad \text{weakly in probability,} \tag{5.180}$$

and this completes the proof. $\square$

Remark 5.4.2 (Example).
We look at a concrete example of a graphical sequence $(k_i^*\colon 1 \le i \le N)$ satisfying (5.165)–(5.166). For $N \ge 1$, let

$$k_i^* = \lfloor i^{1/3} \rfloor, \qquad 1 \le i \le N, \tag{5.181}$$

where $\lfloor x \rfloor$ denotes the greatest integer smaller than or equal to x. Then [165, Theorem 7.12] implies that $(k_i^*\colon 1 \le i \le N)$ is graphical for N large enough. Since $m_N = \lfloor N^{1/3} \rfloor$, it is immediate that (5.165) holds and that

$$\lim_{N\to\infty} \left(\frac{1}{N} \sum_{i=1}^{N} \delta_{k_i^*/m_N} \right)(\cdot) = \mathbb{P}\big(U^{1/3} \in \cdot\,\big) \quad \text{weakly,} \tag{5.182}$$

with U a standard uniform random variable.

§5.4.2 Chung-Lu graphs

The following random graph introduced by [140] is similar to the one discussed in Section 5.4.1. For $N \ge 1$, let $(d_{Ni}\colon 1 \le i \le N)$ be a sequence of positive real numbers. Abbreviate

$$m_N = \max_{1 \le i \le N} d_{Ni}, \qquad \sigma_N = \sum_{i=1}^{N} d_{Ni}. \tag{5.183}$$

Assume that

$$\lim_{N\to\infty} \frac{m_N^2}{\sigma_N} = 0, \qquad \lim_{N\to\infty} N\frac{m_N^2}{\sigma_N} = \infty, \tag{5.184}$$

and

$$\lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} \delta_{d_{Ni}/m_N} = \mu_r \quad \text{weakly} \tag{5.185}$$

for some measure μ_r on $\mathbb{R}$. Consider an inhomogeneous Erdős-Rényi graph on N vertices where an edge exists between i and j, $1 \le i \neq j \le N$, with probability $d_{Ni}d_{Nj}/\sigma_N$. Such graph is called a Chung-Lu graph. If A_N denotes its adjacency matrix, then the following result follows from Theorem 5.1.7.

Proposition 5.4.3 (Theorem 5.1.7 for Chung-Lu graphs).
Under the hypotheses mentioned above,

$$\lim_{N\to\infty} \mathrm{ESD}\big((N\varepsilon_N)^{-1/2}A_N\big) = \mu_r \boxtimes \mu_s \quad \textit{weakly in probability,} \tag{5.186}$$

where

$$\varepsilon_N = \frac{m_N^2}{\sigma_N} \tag{5.187}$$

and μ_s is the standard semicircle law.

§5.4.3 Social networks

Consider a community consisting of N individuals. Data is available on whether the i-th individual and the j-th individual are acquainted, for every pair (i,j) with $1 \le i,j \le N$. Based on this data, the *sociability pattern* of the community has to be inferred statistically. Examples arise in social networks and collaboration networks.

The above situation can be modeled in several ways, one being the following. Denote by ρ the *sociability distribution* of the community, which is a compactly supported probability measure on $[0,\infty)$. Let $(R_i)_{1\le i\le N}$ be i.i.d. random variables drawn from ρ. Think of R_i as the *sociability index* of the i-th individual. Fix $\varepsilon_N > 0$ such that $\varepsilon_N m^2 \le 1$, where m is the supremum of the support of ρ, so that

$$0 \le \varepsilon_N R_i R_j \le 1, \qquad 1 \le i \neq j \le N. \tag{5.188}$$

Suppose that, conditional on $(R_i)_{1\le i\le N}$, the i-th and the j-th individual are acquainted with probability $\varepsilon_N R_i R_j$. In other words, the graph in which the vertices are individuals and the edges are mutual acquaintances is an inhomogeneous Erdős-Rényi random graph G_N with random connection parameters that are controlled by ρ. The data that is available is the adjacency matrix A_N of this graph. The goal is to draw information about ρ from this data. This statistical inference problem boils down to estimating ρ from A_N. Without loss of generality we assume that ρ is standardized, i.e.,

$$\int_0^\infty x\rho(dx) = 1. \tag{5.189}$$

Proposition 5.4.4 (Theorem 5.1.7 for social networks).
Under the assumptions $N^{-1} \ll \varepsilon_N \ll 1$ and (5.189),

$$\lim_{N\to\infty} \mathrm{ESD}\left(\sqrt{\frac{N}{\mathrm{Tr}(A_N^2)}}\, A_N\right) = \rho \boxtimes \mu_s \quad \textit{weakly in probability}, \tag{5.190}$$

where μ_s is the standard semicircle law.

Proof. It is immediate that

$$\lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} \delta_{R_i} = \rho \quad \text{weakly almost surely.} \tag{5.191}$$

Theorem 5.1.7 implies that if $N^{-1} \ll \varepsilon_N \ll 1$, then

$$\lim_{N\to\infty} \mathrm{ESD}\big((N\varepsilon_N)^{-1/2} A_N\big) = \rho \boxtimes \mu_s \quad \text{weakly in probability.} \tag{5.192}$$

Since $A_N(i,j)$ is either 0 or 1,

$$\mathbb{E}\big[\,\mathrm{Tr}(A_N^2)\big] = \sum_{i,j=1}^{N} \mathbb{E}[A_N(i,j)] = \sum_{1\le i\ne j\le N} \varepsilon_N \mathbb{E}[R_i R_j] = \varepsilon_N N(N-1), \tag{5.193}$$

where the last equality follows from (5.189). Consequently,

$$\lim_{N\to\infty} \frac{1}{N^2\varepsilon_N}\mathbb{E}\big[\,\mathrm{Tr}(A_N^2)\big] = 1. \tag{5.194}$$

The fact that the variance equals the sum of the expectation of the conditional variance and the variance of the conditional expectation, implies that

$$\begin{aligned}
\mathrm{Var}\big(\,\mathrm{Tr}(A_N^2)\big) &= \mathrm{Var}\left(2\sum_{1\le i<j\le N} A_N(i,j)\right)\\
&= 4E\left(\sum_{1\le i<j\le N} \varepsilon_N R_i R_j(1-\varepsilon_N R_i R_j)\right) + 4\mathrm{Var}\left(\sum_{1\le i<j\le N} \varepsilon_N R_i R_j\right)\\
&= O(N^2\varepsilon_N) + 4\varepsilon_N^2 \sum_{1\le i<j\le N}\ \sum_{1\le k<l\le N} \mathrm{Cov}(R_iR_j, R_kR_l)\\
&= O(N^3\varepsilon_N^2), \qquad N\to\infty,
\end{aligned} \tag{5.195}$$

where the last line follows from the observation that if i,j,k,l are distinct, then $\mathrm{Cov}(R_iR_j, R_kR_l)$ vanishes. Hence,

$$\lim_{N\to\infty} \mathrm{Var}\left(\frac{1}{N^2\varepsilon_N}\,\mathrm{Tr}(A_N^2)\right) = 0. \tag{5.196}$$

The above in conjunction with (5.194) shows that

$$\lim_{N\to\infty} \frac{1}{N^2\varepsilon_N}\,\mathrm{Tr}(A_N^2) = 1 \quad \text{in probability.} \tag{5.197}$$

This, together with (5.192), completes the proof. □

Thus, $\rho \boxtimes \mu_s$ can in principle be statistically estimated from A_N. Subsequently, ρ can be computed because the moments of $\rho \boxtimes \mu_s$ are functions of the moments of ρ, as shown below. We know from [183, Equation (14.5)] that, for $n \geq 1$,

$$\int_{\mathbb{R}} x^{2n} \, \rho \boxtimes \mu_s(dx) = \sum_{\sigma \in NC_2(2n)} \prod_{j=1}^{n+1} \int_{\mathbb{R}} x^{l_j(\sigma)} \rho(dx), \tag{5.198}$$

where $l_1(\sigma), \ldots, l_{n+1}(\sigma)$ are block sizes of $K(\sigma)$, the Kreweras complement of σ. With the help of the above, the n-th moment of ρ can be written in terms of the $2n$-th moment of $\rho \boxtimes \mu_s$, and the first $n-1$ moments of ρ. Therefore, the moments of ρ can be recursively computed from those of $\rho \boxtimes \mu_s$. Since ρ is compactly supported, it can be computed from its moments.

§E Appendix: basic facts

The following is [114, Corollary A.41], and is also a corollary of the Hoffman-Wielandt inequality.

Lemma E.1 (Lévy distance between empirical spectral distributions).
If L denotes the Lévy distance between two probability measures, then for $N \times N$ symmetric matrices A and B,

$$L^3\big(\mathrm{ESD}(A), \mathrm{ESD}(B)\big) \leq \frac{1}{N} \operatorname{Tr}\big((A-B)^2\big). \tag{5.199}$$

The following is a consequence of the Minkowski and k-Hoffman-Wielandt inequalities. The latter can be found in Exercise 1.3.6 of [189].

Lemma E.2 (Difference of traces).
For real symmetric matrices A and B of the same order, and an even positive integer k,

$$\big| \operatorname{Tr}^{1/k}(A^k) - \operatorname{Tr}^{1/k}(B^k) \big| \leq \operatorname{Tr}^{1/k}\big((A-B)^k\big). \tag{5.200}$$

Definition E.3 (Non-commutative probability space).
A *non-commutative probability space (NCP)* $(\mathcal{A}, \phi)$ is a unital $*$-algebra $\mathcal{A}$ equipped with a linear functional $\phi \colon \mathcal{A} \to \mathbb{C}$ that is unital, i.e.,

$$\phi(\mathbf{1}) = 1, \tag{5.201}$$

and positive, i.e.,

$$\phi(a^* a) \geq 0 \qquad \forall\, a \in \mathcal{A}. \tag{5.202}$$

An NCP $(\mathcal{A}, \phi)$ is *tracial* if

$$\phi(ab) = \phi(ba), \qquad a, b \in \mathcal{A}. \tag{5.203}$$

Lemma E.4 (Limit of polynomials in an NCP).
Suppose that, for every $n \in \mathbb{N}$, $(\mathcal{A}_n, \phi_n)$ is a tracial NCP, and there exist self-adjoint $a_{n1}, \ldots, a_{nk} \in \mathcal{A}_n$ such that, for every polynomial p in k non-commuting variables,

$$\lim_{n \to \infty} \phi_n\big(p(a_{n1}, \ldots, a_{nk})\big) = \alpha_p \in \mathbb{C}. \tag{5.204}$$

Then there exists a tracial NCP $(\mathcal{A}_\infty, \phi_\infty)$ *and self-adjoint* $a_{\infty 1}, \ldots, a_{\infty k} \in \mathcal{A}_\infty$ *such that, for every polynomial* p *in* k *non-commuting variables,*

$$\phi_\infty\big(p(a_{\infty 1}, \ldots, a_{\infty k})\big) = \alpha_p. \tag{5.205}$$

Furthermore, if

$$\sup_{1 \le i \le k,\, j \ge 1} \Big(\phi_\infty\big(a_{\infty i}^{2j}\big)\Big)^{1/2j} < \infty, \tag{5.206}$$

then $(\mathcal{A}_\infty, \phi_\infty)$ *can be embedded into a* W^**-probability space.*

Proof. Let

$$\mathcal{A}_\infty = \mathbb{C}[X_1, \ldots, X_k], \tag{5.207}$$

the set of all polynomials in k non-commuting variables. For a monomial

$$p = \alpha X_{i_1} \ldots X_{i_m}, \tag{5.208}$$

define

$$p^* = \overline{\alpha} X_{i_m} \ldots X_{i_1}. \tag{5.209}$$

This defines the $*$-operation on the whole of $\mathcal{A}$. Let

$$\phi_\infty(p) = \alpha_p \qquad \forall\, p \in \mathcal{A}_\infty. \tag{5.210}$$

It is immediate from (5.204) that ϕ_∞ is positive and unital, i.e., $(\mathcal{A}_\infty, \phi_\infty)$ is an NCP. The desired conclusions are ensured by defining

$$a_{\infty 1} = X_1,\ \ldots,\ a_{\infty k} = X_k. \tag{5.211}$$

Finally, (5.206) implies that $a_{\infty,1}, \ldots, a_{\infty,k}$ are bounded. Hence, by going from polynomials to continuous functions with the help of the Bolzano-Weierstrass theorem, we can embed $(\mathcal{A}_\infty, \phi_\infty)$ into a W^*-probability space. □

The next lemma follows from [182, Theorem 4.20] (which is due to Voiculescu) and the discussion immediately following it.

Lemma E.5 (Polynomials and independence in an NCP).
Suppose that W_N *is an* $N \times N$ *scaled standard Gaussian Wigner matrix, i.e., a symmetric matrix whose upper triangular entries are i.i.d. normal with mean zero and variance* $1/N$*. Let* D_N^1 *and* D_N^2 *be (possibly random)* $N \times N$ *symmetric matrices such that there exists a deterministic* C *satisfying*

$$\sup_{N \ge 1, i=1,2} \|D_N^i\| \le C < \infty, \tag{5.212}$$

where $\|\cdot\|$ *denotes the usual matrix norm (which for a symmetric matrix is the same as the largest absolute value of its eigenvalues). Furthermore, assume that there is a* W^**-probability space* $(\mathcal{A}, \tau)$ *in which there are self-adjoint elements* d_1 *and* d_2 *such that, for any polynomial* p *in two variables, it*

$$\lim_{N \to \infty} \frac{1}{N} \operatorname{Tr}\big(p(D_N^1, D_N^2)\big) = \tau\big(p(d_1, d_2)\big) \quad \textit{almost surely.} \tag{5.213}$$

Finally, suppose that (D_N^1, D_N^2) is independent of W_N. Then there exists a self-adjoint element s in $\mathcal{A}$ (possibly after expansion) that has the standard semicircle distribution and is freely independent of (d_1, d_2), and is such that

$$\lim_{N\to\infty} \frac{1}{N} \operatorname{Tr}\big(p(W_N, D_N^1, D_N^2)\big) = \tau\big(p(s, d_1, d_2)\big) \quad \textit{almost surely} \tag{5.214}$$

for any polynomial p in three variables.

Lemma E.6 (Support of the limiting measure of random variables). *Suppose that for all $n \geq 1$, $Z_{n1} \geq \ldots \geq Z_{nn}$ are random variables such that*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} \delta_{Z_{nj}} = \mu \qquad \textit{weakly in probability,} \tag{5.215}$$

for some probability measure μ on $\mathbb{R}$, where δ_x is the probability measure that puts mass 1 at x. Then,

$$\lim_{p\to 0} \limsup_{n\to\infty} Z_{n\,\lceil np\rceil} = \sup(\operatorname{Supp}(\mu)) \quad \textit{almost surely,} \tag{5.216}$$

where $\lceil x\rceil$ denotes the smallest integer larger than or equal to x.

Proof. Our first claim is that if $x \in \mathbb{R}$ and $0 < p < 1$ are such that

$$\mu((-\infty, x)) < 1 - p, \tag{5.217}$$

then

$$\limsup_{n\to\infty} Z_{n\,\lceil np\rceil} \geq x \quad \text{almost surely.} \tag{5.218}$$

To see why, fix p, x as above and $\varepsilon > 0$ such that $\mu(\{x-\varepsilon\}) = 0$, and note that the hypothesis implies that

$$\lim_{n\to\infty} \frac{1}{n} \big|\{1 \leq j \leq n : Z_{nj} \leq x - \varepsilon\}\big| = \mu((-\infty, x-\varepsilon]) \quad \text{in probability.} \tag{5.219}$$

Therefore,

$$\begin{aligned} &P\left(\limsup_{n\to\infty} Z_{n\,\lceil np\rceil} \leq x - \varepsilon\right) \\ &\leq P\left(\frac{1}{n}\big|\{1 \leq j \leq n : Z_{nj} \leq x - \varepsilon\}\big| \geq 1 - \frac{1}{n}\lceil np\rceil \text{ for large } n\right) \\ &\leq \limsup_{n\to\infty} P\left(\frac{1}{n}\big|\{1 \leq j \leq n : Z_{nj} \leq x - \varepsilon\}\big| \geq 1 - \frac{1}{n}\lceil np\rceil\right) = 0, \end{aligned} \tag{5.220}$$

where the last step follows from (5.219) and the observation that

$$\lim_{n\to\infty} 1 - \frac{1}{n}\lceil np\rceil = 1 - p > \mu((-\infty, x)) \geq \mu((-\infty, x-\varepsilon]). \tag{5.221}$$

Since $\varepsilon > 0$ can be chosen to be arbitrarily small such that $\mu(\{x-\varepsilon\}) = 0$, (5.218) follows.

It is immediate to see that $\limsup_{n\to\infty} Z_{n\,\lceil np\rceil}$ is monotone in p, and hence the almost sure limit exists as $p \to 0$. Furthermore,

$$\lim_{p\to 0}\limsup_{n\to\infty} Z_{n\,\lceil np\rceil} \le \alpha \quad \text{almost surely,} \tag{5.222}$$

where

$$\alpha = \sup(\mathrm{Supp}(\mu)). \tag{5.223}$$

To complete the proof, choose x_k such that $x_k \to \alpha$ and $x_k < \alpha$. Since α is the right end point of the support of μ, it follows that

$$\mu((-\infty, x_k)) < 1. \tag{5.224}$$

Choosing

$$0 < p_k < [1 - \mu((-\infty, x_k))] \wedge \frac{1}{k}, \qquad k \ge 1, \tag{5.225}$$

we see that (5.218) implies

$$\limsup_{n\to\infty} Z_{n\,\lceil np_k\rceil} \ge x_k \quad \text{almost surely.} \tag{5.226}$$

Therefore, since $x_k \to \alpha$,

$$\liminf_{k\to\infty}\limsup_{n\to\infty} Z_{n\,\lceil np_k\rceil} \ge \alpha \quad \text{almost surely.} \tag{5.227}$$

Since $p_k \to 0$, the left-hand side above equals that of (5.222). $\square$