Resolvent of Large Random Graphs

Charles Bordenave * and   Marc Lelarge[†]

June 22, 2009

Abstract

We analyze the convergence of the spectrum of large random graphs to the spectrum of a limit infinite graph. We apply these results to graphs converging locally to trees and derive a new formula for the Stieltjes transform of the spectral measure of such graphs. We illustrate our results on the uniform regular graphs, Erdös-Rényi graphs and graphs with a given degree sequence. We give examples of application for weighted graphs, bipartite graphs and the uniform spanning tree of n vertices.

MSC-class: 05C80, 15A52 (primary), 47A10 (secondary).

1   Introduction

Since the seminal work of Wigner [37], the spectral theory of large dimensional random matrix theory has become a very active field of research, see e.g. the monographs by Mehta [26], Hiai and Petz [21], or Bai and Silverstein [3], and for a review of applications in physics, see Guhr et al. [20]. It is worth noticing that the classical random matrix theory has left aside the dilute random matrices (i.e. when the number of non-zero entries on each row does not grow with the size of the matrix). In the physics literature, the analysis of dilute random matrices has been initiated by Rodgers and Bray [33]. In [7], Biroli and Monasson use heuristic arguments to anaylze the spectrum of the Laplacian of Erdös-Réyni random graphs and an explicit connection with their local approximation as trees is made by Semerjian and Cugliandolo in [35]. Also related is the recent cavity approach to the spectral density of sparse symmetric random matrices by Rogers et al. [34]. Rigorous mathematical treatments can be found in Bauer and Golinelli [4] and Khorunzhy, Scherbina and Vengerovsky [23] for Erdös-Réyni random graphs. In parallel, since McKay [25], similar questions have also appeared in graph theory and combinatorics, for a review, refer to Mohar and Weiss [29]. In this paper, we present a unified treatment of these issues, and prove under weak conditions the convergence of the empirical spectral distribution of adjacency and Laplacian matrices of large graphs.

---
*Institut de Mathématiques - Université de Toulouse & CNRS - France. Email: bordenave@math.univ-toulouse.fr

[†]INRIA-ENS - France. Email: marc.lelarge@ens.fr

Our main contribution is to connect this convergence to the local weak convergence of the sequence of graphs. There is a growing interest in the theory of convergence of graph sequences. The convergence of dense graphs is now well understood thanks to the work of Lovász and Szegedy [24] and a series of papers written with Borgs, Chayes Sós and Vesztergombi (see [11, 12] and references therein). They introduced several natural metrics for graphs and showed that they are equivalent. However these results are of no help in the case studied here of diluted graphs, i.e. when the number of edges scales as the number of vertices. Many new phenomena occur, and there are a host of plausible metrics to consider [9]. Our first main result (Theorem 1) shows that the local weak convergence implies the convergence of the spectral measure. Our second main result (Theorem 2) characterizes in term of Stieltjes transform the limit spectral measure of a large class of random graphs ensemble. The remainder of this paper is organized as follows: in the next section, we give our main results. In Section 3, we prove Theorem 1, in Section 4 we prove Theorem 2. Finally, in Section 5, we extend and apply our results to related graphs.

2 Main results

2.1 Convergence of the spectral measure of random graphs

Let G_n be a sequence of simple graphs with vertex set $[n] = \{1, \ldots, n\}$ and undirected edges set E_n. We denote by $A = A(G_n)$, the $n \times n$ adjacency matrix of G_n, in which $A_{ij} = 1$ if $(ij) \in E_n$ and $A_{ij} = 0$ otherwise. The Laplace matrix of G_n is $L(G_n) = D(G_n) - A(G_n)$, where $D = D(G_n)$ is the degree diagonal matrix in which $D_{ii} = \deg(G_n, i) := \sum_{j \in [n]} A_{ij}$ is the degree of i in G_n and $D_{ij} = 0$ for all $i \neq j$. The main object of this paper is to study the convergence of the empirical measures of the eigenvalues of A and L respectively when the sequence of graphs converges weakly as defined by Benjamini and Schramm [5] and Aldous and Steele [2] (a precise definition is given below). Note that the spectra of $A(G_n)$ or $L(G_n)$ do not depend on the labeling of the graph G_n. If we label the vertices of G_n differently, then the resulting matrix is unitarily equivalent to $A(G_n)$ and $L(G_n)$ and it is well-known that the spectra are unitarily invariant. For ease of notation, we define

$$\Delta_n = A(G_n) - \alpha D(G_n),$$

with $\alpha \in \{0, 1\}$ so that $\Delta_n = A(G_n)$ if $\alpha = 0$ and $\Delta_n = -L(G_n)$ if $\alpha = 1$. The empirical spectral measure of Δ_n is denoted by

$$\mu_{\Delta_n} = n^{-1} \sum_{i=1}^{n} \delta_{\lambda_i(\Delta_n)},$$

where $(\lambda_i(\Delta_n))_{1 \leq i \leq n}$ are the eigenvalues of Δ_n. We endow the set of measures on $\mathbb{R}$ with the usual weak convergence topology. This convergence is metrizable with the Lévy distance $L(\mu, \nu) = \inf\{h \geq 0 : \forall x \in \mathbb{R}, \mu((-\infty, x - h]) - h \leq \nu((-\infty, x]) \leq \mu((-\infty, x - h]) + h\}$.

We now define the local weak convergence introduced by Benjamini and Schramm [5] and Aldous and Steele [2]. For a graph G, we define the rooted graph (G, o) as the connected

component of G containing a distinguished vertex o of G, called the root. A homomorphism form a graph F to another graph G is an edge-preserving map form the vertex set of F to the vertex set of G. A bijective homomorphism is called an isomorphism. A rooted isomorphism of rooted graphs is an isomorphism of the graphs that takes the root of one to the root of the other. $[G, o]$ will denote the class of rooted graphs that are rooted-isomorphic to (G, o). Let $\mathcal{G}^*$ denote the set of rooted isomorphism classes of rooted connected locally finite graphs. Define a metric on $\mathcal{G}*$ by letting the distance between (G_1, o_1) and (G_2, o_2) be $1/(R+1)$, where R is the supremum of those $r \in \mathbb{N}$ such that there is some rooted isomorphism of the balls of radius r (for the graph-distance) around the roots of G_i. $\mathcal{G}^*$ is a separable and complete metric space [1]. For probability measures ρ, ρ_n on $\mathcal{G}^*$, we write $\rho_n \Rightarrow \rho$ when ρ_n converges weakly with respect to this metric.

Following [1], for a finite graph G, let $U(G)$ denote the distribution on $\mathcal{G}^*$ obtained by choosing a uniform random vertex of G as root. We also define $U_2(G)$ as the distribution on $\mathcal{G}^* \times \mathcal{G}^*$ of the pair of rooted graphs $((G, o_1), (G, o_2))$ where (o_1, o_2) is a uniform random pair of vertices of G. If $(G_n), n \in \mathbb{N}$, is a sequence of *deterministic* graphs with vertex set $[n]$ and ρ is a probability measure on $\mathcal{G}^*$, we say the random weak limit of G_n is ρ if $U(G_n) \Rightarrow \rho$. If $(G_n), n \in \mathbb{N}$, is a sequence of *random* graphs with vertex set $[n]$, we denote by $\mathbb{E}[\cdot] = \mathbb{E}_n[\cdot]$ the expectation with respect to the randomness of the graph G_n. The measure $\mathbb{E}[U(G_n)]$ is defined as $\mathbb{E}[U(G_n)](B) = \mathbb{E}[U(G_n)(B)]$ for any measurable event B on $\mathcal{G}^*$. Following Aldous and Steele [2], we will say that the random weak limit of G_n is ρ if $\mathbb{E}[U(G_n)] \Rightarrow \rho$. Note that the second definition generalizes the first one (take $\mathbb{E}_n = \delta_{G_n}$). In all cases, we denote by (G, o) a random rooted graph whose distribution of its equivalence class in $\mathcal{G}^*$ is ρ. Let $\deg(G_n, o)$ be the degree of the root under $U(G_n)$ and $\deg(G, o)$ be the degree of the root under ρ. They are random variables on $\mathbb{N}$ such that if the random weak limit of G_n is ρ then $\lim_{n \to \infty} \mathbb{E}[\deg(G_n, o) \leq k] = \rho(\{\deg(G, o) \leq k\})$.

We will make the following assumption for the whole paper:

A. The sequence of random variables $(\deg(G_n, o)), n \in \mathbb{N}$, is uniformly integrable.

Assumption (A) ensures that if the random weak limit of G_n is ρ, then the average degree of the root converges, namely $\lim_{n \to \infty} \mathbb{E}[\deg(G_n, o)] = \rho(\deg(G, o))$.

To prove our first main result, we will consider two assumptions, one, denoted by (D), for a given sequence of finite graphs and another, denoted by (R), for a sequence of random finite graphs.

D. As n goes to infinity, the random weak limit of G_n is ρ.

R. As n goes to infinity, $U_2(G_n) \Rightarrow \rho \otimes \rho$.

Of course, Assumption (R) implies (D). We are now ready to state our first main theorem:

Theorem 1 *(i) Let $G_n = ([n], E_n)$ be a sequence of graphs satisfying assumptions (D-A), then there exists a probability measure μ on $\mathbb{R}$ such that $\lim_{n \to \infty} \mu_{\Delta_n} = \mu$.*

3

(ii) Let $G_n = ([n], E_n)$ be a sequence of random graphs satisfying assumptions (R-A), then there exists a probability measure μ on $\mathbb{R}$ such that, $\lim_{n\to\infty} \mathbb{E}L(\mu_{\Delta_n}, \mu) = 0$.

In (ii), note that the stated convergence implies the weak convergence of the law of μ_{Δ_n} to δ_μ. Theorem 1 appeared under different settings, when the sequence of maximal degrees of the graphs G_n is bounded, see Colin de Verdière [13], Serre [36] and Elek [17].

2.2 Random graphs with trees as local weak limit

We now consider a sequence of random graphs $G_n, n \in \mathbb{N}$ which converges as n goes to infinity to a possibly infinite tree. In this case, we will be able to characterize the probability measure μ. Here, we restrict our attention to particular trees as limits but some cases outside the scope of this section are also analyzed in Section 5. A Galton-Watson Tree (GWT) with *offspring* distribution F is the random tree obtained by a standard Galton-Watson branching process with offspring distribution F. For example, the infinite k-ary tree is a GWT with offspring distribution δ_k, see Figure 1. A GWT with *degree* distribution F_* is a rooted random tree obtained by a Galton-Watson branching process where the root has offspring distribution F_* and all other genitors have offspring distribution F where for all $k \geq 1$, $F(k-1) = kF_*(k)/\sum_k kF_*(k)$ (we assume $\sum_k kF_*(k) < \infty$). For example the infinite k-regular tree is a GWT with degree distribution δ_k, see Figure 1. It is easy to check that a GWT with degree distribution F_* defines a unimodular probability measure on $\mathcal{G}^*$ (for a definition and properties of unimodular measures, refer to [1]). Note that if F_* has a finite second moment then F has a finite first moment.
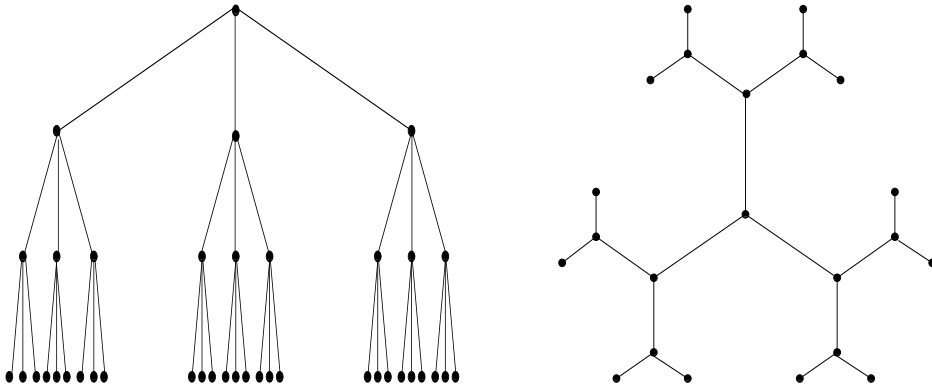

Figure 1: Left: representation of a 3-ary tree. Right: representation of a 3-regular tree.

Let $n \in \mathbb{N}$ and $G_n = ([n], E_n)$, be a random graph on the finite vertex set $[n]$ and edge set E_n. We assume that the following holds

RT. As n goes to infinity, $U_2(G_n)$ converges weakly to $\rho \otimes \rho$, where $\rho \in \mathcal{G}^*$ is the probability measure of GWT with degree distribution F_*.

4

Note that we have $\lim_{n\to\infty} \mathbb{E}[\deg(G_n, o)] = \sum_k k F_*(k) < \infty$ and

$$\lim_{n\to\infty} \mathbb{E}[\deg(G_n, v)| \, (o, v) \in E_n] = 1 + \sum_k k F(k) = 1 + \sum_k k(k-1)F_*(k)/\sum_k k F_*(k).$$

Under assumption (A), this last quantity might be infinite. To prove our next result, we need to strengthen it into

A'. Assumption (A) holds and $\sum_k k^2 F_*(k) < \infty$.

We mention three important classes of graphs which converge locally to a tree and which satisfy our assumptions.

Example 1 *Uniform regular graph.* The uniform k-regular graph on n vertices satisfies these assumptions with the infinite k-regular tree as local limit. It follows for example easily from Bollobás [8], see also the survey Wormald [38].

Example 2 *Erdös-Rényi graph.* Similarly, consider the Erdös-Rényi graph on n vertices where there is an edge between two vertices with probability p/n independently of everything else. This sequence of random graphs satisfies the assumptions with limiting tree the GWT with degree distribution $Poi(p)$.

Example 3 *Graphs with asymptotic given degree.* More generally, the usual random graph (called configuration model) with asymptotic degree distribution F_* satisfies this set of hypothesis provided that $\sum k(k-1)F_*(k) < \infty$ (e.g. see Chapter 3 in Durrett [15] and Molloy and Reed [30]).

Given a positive R and a finite rooted graph H, let $U(G_n)(H, R)$ and $GWT(H, R)$ be the probabilities that H is rooted isomorphic to the ball of radius R about the root of a graph chosen with distribution $U(G_n)$ and GWT respectively (recall that G_n is random and hence $U(G_n)$ is a random measure on $\mathcal{G}^*$). In the three examples above, it is easy to check that in probability $U(G_n)(H, R)$ converges to $GWT(H, R)$ and Assumption (RT) follows.

Recall that $\Delta_n = A(G_n) - \alpha D(G_n)$, with $\alpha \in \{0, 1\}$. We now introduce a standard tool of random matrix theory used to describe the empirical spectral measure μ_{Δ_n} (see Bai and Silverstein [3] for more details). Let $R_n(z) = (\Delta_n - zI_n)^{-1}$ be the resolvent of Δ_n. We denote $\mathbb{C}_+ = \{z \in \mathbb{C} : \Im z > 0\}$. Let $\mathcal{H}$ be the set of holomorphic functions f from $\mathbb{C}_+$ to $\mathbb{C}_+$ such that $|f(z)| \leq \frac{1}{\Im z}$. For all $i \in [n]$, the mapping $z \mapsto R_n(z)_{ii}$ is in $\mathcal{H}$ (see Section 3.1). We denote by $\mathcal{P}(\mathcal{H})$ the space of probability measures on $\mathcal{H}$. The Stieltjes transform of the empirical spectral distribution is given by:

$$m_n(z) \;\; = \;\; \int_{\mathbb{R}} \frac{1}{x - z} d\mu_{\Delta_n}(x) = \frac{1}{n} \mathrm{tr} R_n(z) = \frac{1}{n} \sum_{i=1}^n R_n(z)_{ii} \tag{1}$$

where $z \in \mathbb{C}_+$. Our main second result is the following.

Theorem 2 *Under assumptions (RT-A'),*

(i) There exists a unique probability measure $Q \in \mathcal{P}(\mathcal{H})$ such that for all $z \in \mathbb{C}_+$,

$$Y(z) \stackrel{d}{=} - \left(z + \alpha(N+1) + \sum_{i=1}^{N} Y_i(z) \right)^{-1}, \tag{2}$$

where N has distribution F and Y and Y_i are iid copies independent of N with law Q.

(ii) For all $z \in \mathbb{C}_+$, $m_n(z)$ converges as n tends to infinity in L^1 to $\mathbb{E}X(z)$, where for all $z \in \mathbb{C}_+$,

$$X(z) \stackrel{d}{=} - \left(z + \alpha N_* + \sum_{i=1}^{N_*} Y_i(z) \right)^{-1}, \tag{3}$$

where N_* has distribution F_* and Y_i are iid copies with law Q, independent of N_*.

Equation (2) is a Recursive Distributional Equation (RDE). In random matrix theory, the Stieltjes transform appears classically as a fixed point of a mapping on $\mathcal{H}$. For example, in the Wigner case [37] (i.e. the matrix $W_n = (A_{ij}/\sqrt{n})_{1 \le i,j \le n}$ where $A_{ij} = A_{ji}$ are iid copies of A with $var(A) = \sigma^2$), the Stieltjes tranform $m(z)$ of the limiting spectral measure satisfies for all $z \in \mathbb{C}_+$,

$$m(z) = - \left(z + \sigma^2 m(z) \right)^{-1}. \tag{4}$$

It is then easy to show that the limiting spectral measure is the semi circular law with radius 2σ.

We explain now why the situation is different in our case. Let $X_n(z) = R_n(z)_{oo} \in \mathcal{H}$ be the diagonal term of the resolvent matrix corresponding to the root of the graph G_n. From (1), we see that $m_n(z) = U(G_n)(X_n(z))$. In a first step, we will prove that the random variable $X_n(z)$ under $U(G_n)$ converges in distribution to $X(z)$ given by (3). Then we finish the proof of Theorem 2 with the following steps $\lim_n m_n(z) = \lim_n \mathbb{E}[U(G_n)(X_n(z))] = \mathbb{E}[X(z)]$. For the proof of the convergence of $X_n(z)$, we first consider the the case of uniformly bounded degrees in G, i.e. $\max_i \deg(G_n, i) \le \ell$. In this case, Mohar proved in [28] that it is possible to define the resolvent R of the possibly infinite graph (G, o) with law ρ. We then show that $X_n(z)$ converges weakly to $R(z)_{oo}$ under ρ. Thanks to the tree structure of (G, o), we can characterize the law of $R(z)_{oo}$ with the RDE (2). Recall that the offspring distribution of the root has distribution F_* whereas all other nodes have offspring distribution F which explains the formula (3) and (2) respectively. A similar approach was used in [10] for the spectrum of large random reversible Markov chains with heavy-tailed weights, where a more intricate tree structure appears.

Note that for all $z \in \mathbb{C}_+$, $m_n(z)$ and $X(z)$ are bounded by $\Im(z)^{-1}$, hence the convergence in Theorem 2(ii) of $m_n(z)$ to $\mathbb{E}X(z)$ holds in L^p for all $p \ge 1$. Under the restrictive assumption $\max_i \deg(G_n, i) \le \ell$, we are able to prove that on $\mathcal{H}$, X_n converges weakly to X but we do not know if this convergence holds in general. We also need Assumption (A') to prove the

uniqueness of the solution in (2) even though we know from Theorem 1 that the empirical spectral distribution converges under the weaker Assumption (A).

We end this section with two examples that appeared in the literature.

EXAMPLE 1   If G_n is the uniform k-regular graph on $[n]$, with $k \geq 2$, then G_n converges to the GWT with degree distribution δ_k. We consider the case $\alpha = 0$, looking for deterministic solutions of Y, we find: $Y(z) = -(z + (k-1)Y(z))^{-1}$, hence, in view of (4), Y is simply the Stieltjes transform of the semi-circular law with radius $2\sqrt{k-1}$. For $X(z)$ we obtain,

$$X(z) = -(z + kY(z))^{-1} = -\frac{2(k-1)}{(k-2)z + k\sqrt{z^2 - 4(k-1)}}. \tag{5}$$

In particular $\Im X(z) = \Im(z+kY(z))/|z+kY(z)|^2$. Using the formula $\mu[a,b] = \lim_{v \to 0^+} \frac{1}{\pi} \int_a^b \Im X(x + iv)dx$, valid for all continuity points $a < b$ of μ, we deduce easily that μ_{A_n} converges weakly to the probability measure $\mu(dx) = f(x)dx$ which has a density f on $[-2\sqrt{k-1}, 2\sqrt{k-1}]$ given by

$$f(x) = \frac{k}{2\pi} \frac{\sqrt{4(k-1) - x^2}}{k^2 - x^2},$$

and $f(x) = 0$ if $x \notin [-2\sqrt{k-1}, 2\sqrt{k-1}]$. This formula for the density of the spectral measure is due to McKay [25] and Kesten [22] in the context of simple random walks on groups. To the best of our knowledge, the proof of Theorem 2 is the first proof using the resolvent method of McKay's Theorem. It is interesting to notice that this measure and the semi-circle distribution are simply related by their Stieltjes transform see (5).

EXAMPLE 2   If G_n is a Erdös-Rényi graph on $[n]$, with parameter p/n then G_n converges to the GWT with degree distribution $Poi(p)$. In this case, F and F_* have the same distribution, thus for $\alpha = 0$, $X \overset{d}{=} Y$ has law Q. Theorem 2 improves a result of Khorunzhy, Shcherbina and Vengerovsky [23], Theorems 3 and 4, Equations (2.17), (2.24). Indeed, for all $z \in \mathbb{C}_+$, we use the formula $e^{iuw} = 1 - \sqrt{u} \int_0^\infty \frac{J_1(2\sqrt{ut})}{\sqrt{t}} e^{-itw^{-1}} dt$ valid for all $u \geq 0$ and $w \in \mathbb{C}_+$, and where $J_1(t) = \frac{t}{2} \sum_{k=0}^\infty \frac{(-t^2/4)^k}{k!(k+1)!}$ is the Bessel function of the first kind. Then with $f(u,z) = \mathbb{E}e^{uY(z)}$ and $\varphi(z) = \mathbb{E}z^N$ we obtain easily from (2) that for all $u \geq 0$, $z \in \mathbb{C}_+$,

$$f(u,z) = 1 - \sqrt{u} \int_0^\infty \frac{J_1(2\sqrt{ut})}{\sqrt{t}} e^{it(z+\alpha)} \varphi(e^{i\alpha t} f(t,z))dt.$$

If N is a Poison random variable with parameter p, then $\varphi(z) = \exp(p(z-1))$, and we obtain the results in [23].

3 Proof of Theorems 1

3.1 Random finite networks

It will be convenient to work with marked graphs that we call networks. First note that the space $\mathcal{H}$ of holomorphic functions f from $\mathbb{C}_+$ to $\mathbb{C}_+$ such that $|f(z)| \leq \frac{1}{\Im z}$ equipped with the topology induced by the uniform convergence on compact sets is a complete separable metrizable compact space (see Chapter 7 in [14]). We now define a network as a graph $G = (V, E)$ together with a complete separable metric space, in our case $\mathcal{H}$, and a map from V to $\mathcal{H}$. We use the following notation: G will denote a graph and $\overline{G}$ a network with underlying graph G. A rooted network $(\overline{G}, o)$ is the connected component of a network $\overline{G}$ of a distinguished vertex o of G, called the root. $[\overline{G}, o]$ will denote the class of rooted networks that are rooted-isomorphic to $(\overline{G}, o)$. Let $\overline{\mathcal{G}^*}$ denote the set of rooted isomorphism classes of rooted connected locally finite networks. As in [1], define a metric on $\overline{\mathcal{G}^*}$ by letting the distance between $(\overline{G}_1, o_1)$ and $(\overline{G}_2, o_2)$ be $1/(\alpha + 1)$, where α is the supremum of those $r \in \mathbb{N}$ such that there is some rooted isomorphism of the balls of (graph-distance) radius r around the roots of G_i such that each pair of corresponding marks has distance less than $1/r$. Note that the metric defined on $\mathcal{G}^*$ in Section 2.1 corresponds to the case of a constant mark attached to each vertex. It is now easy to extend the local weak convergence of graphs to networks. To fix notations, for a finite network $\overline{G}$, let $U(\overline{G})$ denote the distribution on $\overline{\mathcal{G}^*}$ obtained by choosing a uniform random vertex of G as root. If $(\overline{G}_n), n \in \mathbb{N}$, is a sequence of (possibly random) networks with vertex set $[n]$, we denote by $\mathbb{P}_n[\cdot]$ the expectation with respect to the randomness of the graph G_n. We will say that the random weak limit of $\overline{G}_n$ is $\overline{\rho}$ if $\mathbb{P}_n[U(\overline{G}_n)] \Rightarrow \overline{\rho}$.

In this paper, we consider the finite networks $\{G_n, (R_n(z)_{ii})_{i \in [n]}\}$, where we attach the mark $R_n(z)_{ii}$ to vertex i. We need to check that the map $z \mapsto R_n(z)_{ii}$ belongs to $\mathcal{H}$. First note that by standard linear algebra (see Lemma 7.2 in Appendix), we have

$$R_n(z)_{ii} = \frac{1}{(\Delta_n)_{ii} - z - \beta_i^t (\Delta_{n,i} - z I_{n-1})^{-1} \beta_i},$$

where β_i is the $((n-1) \times 1)$ ith column vector of Δ_n with the ith element removed and $\Delta_{n,i}$ is the matrix obtained form Δ_n with the ith row and column deleted (corresponding to the graph with vertex i deleted). An easy induction on n shows that $R_n(z)_{ii} \in \mathcal{H}$ for all $i \in [n]$.

3.2 Linear operators associated with a graph of bounded degree

We first recall some standard results that can be found in Mohar and Woess [29]. Let (G, o) be the connected component of a locally finite graph G containing the vertex o of G. There is no loss of generality in assuming that the vertex set of G is $\mathbb{N}$. Indeed if G is finite, we can extend G by adding isolated vertices. We assume that $\deg(G) = \sup\{\deg(G, u), u \in \mathbb{N}\} < \infty$. Let $A(G)$ be the adjacency matrix of G. We define the matrix $\Delta = \Delta(G) = A(G) - \alpha D(G)$, where $D(G)$ is the degree diagonal matrix and $\alpha \in \{0, 1\}$. Let $e_k = \{\delta_{ik} : i \in \mathbb{N}\}$ be the specified complete

orthonormal system of $L^2(\mathbb{N})$. Then Δ can be interpreted as a linear operator over $L^2(\mathbb{N})$, which is defined on the basis vector e_k as follows:

$$\langle \Delta e_k, e_j \rangle = \Delta_{kj}.$$

Since G is locally finite, Δe_k is an element of $L^2(\mathbb{N})$ and Δ can be extended by linearity to a dense subspace of $L^2(\mathbb{N})$, which is spanned by the basis vectors $\{e_k, k \in \mathbb{N}\}$. Denote this dense subspace H_0 and the corresponding operator Δ_0. The operator Δ_0 is symmetric on H_0 and thus closable (Section VIII.2 in [31]). We will denote the closure of Δ_0 by the same symbol Δ as the matrix. The operator Δ is by definition a closed symmetric transformation: the coordinates of $y = \Delta x$ are

$$y_i = \sum_j \Delta_{ij} x_j, \quad i \in \mathbb{N},$$

whenever these series converge.

The following lemma is proved in [28] for the case $\alpha = 0$ and the case $\alpha = 1$ follows by the same argument.

Lemma 2.1 *Assume that* $\deg(G) = \sup\{\deg(G, u), u \in V\} < \infty$, *then* Δ *is self-adjoint.*

3.3 Proof of Theorem 1 (i) - bounded degree

In this paragraph, we assume that there exists $\ell \in \mathbb{N}$ such that

$$\rho(\deg(G, o) \leq \ell) = 1. \tag{6}$$

Since $\mathcal{G}^*$ is a complete separable metric space, by the Skorokhod Representation Theorem (Theorem 7 in Appendix), we can assume that $U(G_n)$ and (G, o) are defined on a common probability space such that $U(G_n)$ converges almost surely to (G, o) in $\mathcal{G}^*$. As explained in previous section, we define the operators $\Delta = \Delta(G)$ and $\Delta_n = \Delta(G_n)$ on the Hilbert space $L^2(\mathbb{N})$. The convergence of $U(G_n)$ to (G, o) implies a convergence of the associated operators up to a re-indexing of $\mathbb{N}$ which preserves the root. To be more precise, let $(G, o)[r]$ denote the finite graph induced by the vertices at distance at most r from the root o in G. Then by definition there exists a.s. a sequence r_n tending to infinity and a random rooted isomorphism σ_n from $(G, o)[r_n]$ to $U(G_n)[r_n]$. We extend arbitrarily this isomorphism σ_n to all $\mathbb{N}$. For the basis vector $e_k \in L^2(\mathbb{N})$, we set $\sigma_n(e_k) = e_{\sigma_n(k)}$. For simplicity, we assume that e_0 corresponds to the root of the graph (so that $\sigma_n(0) = 0$) and we denote $e_0 = o$ to be consistent with the notation used for graphs. By extension, we define $\sigma_n(\phi)$ for each $\phi \in H_0$ the subspace of $L^2(\mathbb{N})$, which is spanned by the basis vectors $\{e_k, k \in \mathbb{N}\}$. Then the convergence of $U(G_n)$ to (G, o) implies that for all $\phi \in H_0$,

$$\sigma_n^{-1} \Delta_n \sigma_n \phi \to \Delta \phi \text{ a.s.} \tag{7}$$

9

By Theorem VIII.25(a) in [31], the convergence (7) and the fact that Δ is a self-adjoint operator (due to (6)) imply the convergence of $\sigma_n^{-1}\Delta_n\sigma_n \to \Delta$ in the strong resolvent sense:

$$\sigma_n^{-1}R_n(z)\sigma_n x - R(z)x \to 0, \text{ for any } x \in L^2(\mathbb{N}), \text{ and for all } z \in \mathbb{C}_+.$$

This last statement shows that the sequence of networks $U(\overline{G}_n)$ converges a.s. to $(\overline{G}, o)$ in $\overline{\mathcal{G}}^*$. In particular, we have

$$m_n(z) = \frac{1}{n}\sum_{i=1}^n \langle R_n(z)e_i, e_i \rangle = U(G_n)(\langle R_n(z)o, o\rangle) \to \int \langle R(z)o, o\rangle d\rho[G, o], \qquad (8)$$

by dominated convergence since $|\langle R_n(z)e_i, e_i\rangle| \leq (\Im z)^{-1}$.

Remark 1 *A trace operator* Tr *was defined in [1]. With their notation, we have*

$$\mathrm{Tr}(R(z)) = \int \langle R(z)o, o\rangle d\rho[G, o].$$

3.4 Proof of Theorem 1 (i) - general case

Let $\ell \in \mathbb{N}$, we define the graph $G_{n,\ell}$ on $[n]$ obtained from G_n by removing all edges adjacent to a vertex i, if $\deg(G_n, i) > \ell$. Therefore the matrix $\Delta(G_{n,\ell})$ denoted by $\Delta_{n,\ell}$ is equal to, for $i \neq j$

$$(\Delta_{n,\ell})_{ij} = \begin{cases} A(G_n)_{ij} & \text{if } \max\{\deg(G_n, i), \deg(G_n, j)\} \leq \ell \\ 0 & \text{otherwise} \end{cases}$$

and $(\Delta_{n,\ell})_{ii} = -\alpha\sum_{j\neq i}(\Delta_{n,\ell})_{ij}$. The empirical measure of the eigenvalues of $\Delta_{n,\ell}$ is denoted by $\mu_{\Delta_{n,\ell}}$. By the Rank Difference Inequality (Lemma 7.1 in Appendix),

$$L(\mu_{\Delta_n}, \mu_{\Delta_{n,\ell}}) \leq \frac{1}{n}\mathrm{rank}(\Delta_n - \Delta_{n,\ell}),$$

where L denotes the Lévy distance. The rank of $\Delta_n - \Delta_{n,\ell}$ is upper bounded by the number of rows different from 0, i.e. by the number of vertices with degree at least $\ell + 1$ or such that there exist a neighboring vertex with degree at least $\ell + 1$. By definition, each vertex with degree d is connected to d other vertices. It follows

$$\mathrm{rank}(\Delta_n - \Delta_{n,\ell}) \leq \sum_{i=1}^n (\deg(G_n, i) + 1)\mathbf{1}(\deg(G_n, i) > \ell),$$

and therefore:

$$L(\mu_{\Delta_n}, \mu_{\Delta_{n,\ell}}) \leq \int (\deg(G, o) + 1)\mathbf{1}(\deg(G, o) > \ell)dU(G_n)[G, o] =: p_{n,\ell}.$$

By assumptions (D-A), uniformly in ℓ,

$$p_{n,\ell} \to \int (\deg(G, o) + 1)\mathbf{1}(\deg(G, o) > \ell)d\rho[G, o],$$

10

where the right-hand term tends to 0 as ℓ tends to infinity. Fix $\epsilon > 0$, for ℓ sufficiently large and $n \geq N(\ell)$, we have

$$L(\mu_{\Delta_n}, \mu_{\Delta_{n,\ell}}) \leq \epsilon.$$

Hence we get, for $n \geq N(\ell)$, $q \in \mathbb{N}$,

$$L(\mu_{\Delta_{n+q}}, \mu_{\Delta_n}) \leq 2\epsilon + L(\mu_{\Delta_{n+q,\ell}}, \mu_{\Delta_{n,\ell}}).$$

By (8), if $m_{n,\ell}(z)$ denotes the Stieltjes transform of $\mu_{\Delta_{n,\ell}}$, we have for all $z \in \mathbb{C}_+$, $\lim_{n\to\infty} m_{n,\ell}(z) = m^{\ell}(z)$ for some $m^{\ell} \in \mathcal{H}$. Hence for the weak convergence $\lim_{n\to\infty} \mu_{\Delta_{n,\ell}} = \mu^{\ell}$ for the probability measure μ^{ℓ} whose Stieltjes transform is m^{ℓ}. It follows that the sequence μ_{Δ_n} is Cauchy and the proof of Theorem 1(i) is complete (recall that the set probability measures on $\mathbb{R}$ with the Lévy metric is complete).

3.5 Proof of Theorem 1 (ii) - bounded degree

We assume first the following

A". There exists $\ell \geq 1$ such that for all $n \in \mathbb{N}$ and $i \in [n]$, $\deg(G_n, i) \leq \ell$.

With this extra assumption, Δ_n and Δ are self adjoint operators and, as in Section 3.3, we deduce that $\mathbb{E}U(\overline{G}_n) \Rightarrow \overline{\rho}$. In particular, we have

$$\mathbb{E}[m_n(z)] \to \int \langle R(z)o, o \rangle d\overline{\rho}[G, o].$$

Hence, in order to prove Theorem 1 (ii) with the additional assumption (A"), it is sufficient to prove that, for all $z \in \mathbb{C}_+$,

$$\lim_{n\to\infty} \mathbb{E} |m_n(z) - \mathbb{E}[m_n(z)]|^2 = 0. \tag{9}$$

Take $z \in \mathbb{C}_+$ with $\Im z > 2\ell$. Notice, that by Vitali's convergence Theorem, it is sufficient to prove (9) for all z such that $\Im z > 2\ell$. By assumption (A"), we get $\left|(\Delta_n)_{ii}^k\right| \leq \ell^k$, and we may thus write for any integer $t \geq 1$,

$$
\begin{aligned}
R_n(z)_{ii}(z) &= -\sum_{k=0}^{t-1} \frac{(\Delta_n)_{ii}^k}{z^{k+1}} - \sum_{k=t}^{\infty} \frac{(\Delta_n)_{ii}^k}{z^{k+1}} \\
&=: X_i^{(n)}(t) + \epsilon_i^{(n)}(t).
\end{aligned}
$$

Note that $|\epsilon_i^{(n)}(t)| \leq \epsilon(t) := \sum_{k=t}^{\infty} \frac{\ell^k}{(2\ell)^{k+1}} \leq 2^{-t}$. We define $\overline{X}_i^{(n)}(t) = X_i^{(n)}(t) - \mathbb{E}\left[X_i^{(n)}(t)\right]$. Since $\left|X_i^{(n)}(t) - R_n(z)_{ii}\right| = \left|\epsilon_i^{(n)}(t)\right| \leq \epsilon(t)$, we have

$$|m_n(z) - \mathbb{E}[m_n(z)]| \leq \left|\frac{1}{n}\sum_{i=1}^{n} \overline{X}_i^{(n)}(t)\right| + 2\epsilon(t).$$

11

Therefore if for all $t \geq 1$, we manage to prove that in $L^2(\mathbb{P})$

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \overline{X}_i^{(n)}(t) = 0, \tag{10}$$

then the proof of (9) will be complete. We now prove (10):

$$\mathbb{E}\left(\frac{1}{n} \sum_{i=1}^{n} \overline{X}_i^{(n)}(t)\right)^2 = \frac{1}{n^2} \mathbb{E}\left(\sum_{i \neq j} \overline{X}_i^{(n)}(t)\overline{X}_j^{(n)}(t) + \sum_{i=1}^{n} \overline{X}_i^{(n)}(t)^2\right)$$

$$= \mathbb{E}\left(\overline{X}_{o_1}^{(n)}(p)\overline{X}_{o_2}^{(n)}(t)\right),$$

where (o_1, o_2) is a uniform pair of vertices. We then notice that $\overline{X}_i^{(n)}(t)$ is a measurable function of the ball of radius t and center i. Thus, by assumption (R), $\lim_n \mathbb{E}\overline{X}_{o_1}^{(n)}(t)\overline{X}_{o_2}^{(n)}(t) = \lim_n \mathbb{E}\overline{X}_{o_1}^{(n)}(t)\mathbb{E}\overline{X}_{o_2}^{(n)}(t) = 0$, and (10) follows. Hence we proved (9) under the assumption (A").

3.6 Proof of Theorem 1 (ii) - general case

We now relax assumption (A") by assumption (A). By the same argument as in Section 3.4, we get

$$\mathbb{E}L(\mu_{\Delta_n}, \mu_{\Delta_{n,\ell}}) \leq \mathbb{E}\int (\deg(G, o) + 1)\mathbf{1}(\deg(G, o) > \ell)dU(G_n)[G, o] =: \overline{p}_{n,\ell}.$$

By assumptions (R-A), uniformly in ℓ,

$$\overline{p}_{n,\ell} \to \mathbb{E}\int (\deg(G, o) + 1)\mathbf{1}(\deg(G, o) > \ell)d\rho[G, o],$$

where the right-hand term tends to 0 as ℓ tends to infinity. The end of the proof follows by the same argument, since we have now for $n \geq N(\ell)$, $q \geq 1$,

$$\mathbb{E}L(\mu_{\Delta_{n+q}}, \mu_{\Delta_n}) \leq 2\epsilon + \mathbb{E}L(\mu_{\Delta_{n+q,\ell}}, \mu_{\Delta_{n,\ell}}).$$

4 Proof of Theorem 2

4.1 Proof of Theorem 2 (i)

In this paragraph, we check the existence and the unicity of the solution of the RDE (2). Let $\Theta = \mathbb{N} \times \mathcal{H}^\infty$, where $\mathcal{H}^\infty$ is the usual infinite product space. We define a map $\psi : \Theta \to \mathcal{H}$ as follows

$$\psi(n, (h_i)_{i \in \mathbb{N}}) : \begin{array}{ccc} \mathbb{C}_+ & \to & \mathbb{C}_+ \\ z & \mapsto & -\left(z + \alpha(n+1) + \sum_{i=1}^{n} h_i(z)\right)^{-1}. \end{array}$$

Let Ψ be a map from $\mathcal{P}(\mathcal{H})$ to itself, where $\Psi(P)$ is the distribution of $\psi(N, (Y_i)_{i \in \mathbb{N}})$, where

12

(i) $(Y_i, i \geq 1)$ are independent with distribution P;

(ii) N has distribution F;

(iii) the families in (i) and (ii) are independent.

We say $Q \in \mathcal{P}(\mathcal{H})$ is a solution of the RDE (2) if $Q = \Psi(Q)$.

Lemma 2.2 *There exists a unique measure $Q \in \mathcal{P}(\mathcal{H})$ solution of the RDE (2).*

Proof. Let Ω be a bounded open set in the half plane $\{z \in \mathbb{C}_+ : \Im z \geq \sqrt{\mathbb{E}N} + 1\}$ (by assumption (A'), $\mathbb{E}N$ is finite). Let $\mathcal{P}(\mathcal{H})$ be the set of probability measures on $\mathcal{H}$. We define the distance on $\mathcal{P}(\mathcal{H})$

$$W(P, Q) = \inf \mathbb{E} \int_\Omega |X(z) - Y(z)| dz$$

where the infimum is over all possible coupling of the distributions P and Q where X has law P and Y has law Q. The fact that W is the distance follows from the fact that two holomorphic functions equal on a set containing a limit point are equal. The space $\mathcal{P}(\mathcal{H})$ equipped with the metric W gives a complete metric space.

Let X with law P, Y with law Q coupled so that $W(P, Q) = \mathbb{E} \int_\Omega |X(z) - Y(z)| dz$. We consider $(X_i, Y_i)_{i \in \mathbb{N}}$ iid copies of (X, Y) and independent of the variable N. By definition, we have the following

$$
\begin{aligned}
W(\Psi(P), \Psi(Q)) &\leq \mathbb{E} \int_\Omega |\psi(N, (X_i); z) - \psi(N, (Y_i); z)| dz \\
&\leq \mathbb{E} \int_\Omega \left| \frac{\sum_{i=1}^N X_i(z) - Y_i(z)}{\left(z + \alpha(N+1) + \sum_{i=1}^N X_i(z)\right)\left(z + \alpha(N+1) + \sum_{i=1}^N Y_i(z)\right)} \right| dz \\
&\leq \int_\Omega (\Im z)^{-2} \mathbb{E} \sum_{i=1}^N |X_i(z) - Y_i(z)| \, dz \\
&\leq \mathbb{E}N (\inf_{z \in \Omega} \Im z)^{-2} W(P, Q).
\end{aligned}
$$

Then since $\inf_{z \in \Omega} \Im z > \sqrt{\mathbb{E}N}$, Ψ is a contraction and from Banach fixed point Theorem, there exists a unique probability measure Q on $\mathcal{H}$ such that $\Psi(Q) = Q$. $\qquad \square$

4.2 Resolvent of a tree

In this paragraph, we prove the following proposition.

Proposition 2.1 *Let F_* be a distribution with finite mean, and $(T_n, 1)$ be a GWT rooted at 1 with degree distribution F_* stopped at generation n. Let $A(T_n)$ be the adjacency matrix of T_n,*

and $\Delta(T_n) = A(T_n) - \alpha D(T_n)$. Let $R^{(n,T)}(z) = (\Delta(T_n) - zI)^{-1}$ and $X^{(n,T)}(z) = R^{(n,T)}_{11}(z)$. For all $z \in \mathbb{C}_+$, as n goes to infinity $X^{(n,T)}(z)$ converges weakly to $X(z)$ defined by Equation (3).

We start with the following Lemma which explains where the RDE (2) comes from.

Lemma 2.3 *Let F be a distribution with finite mean, and $(T_n, 1)$ be a GWT rooted at 1 with offspring distribution F stopped at generation n. Let $A(T_n)$ be the adjacency matrix of T_n, and $\bar{\Delta}(T_n) = A(T_n) - \alpha(D(T_n) + V(T_n))$, where $V(T_n)_{11} = 1$ and $V(T_n)_{ij} = 0$ for all $(i,j) \neq (1,1)$. Let $R^{(n,T)}(z) = (\bar{\Delta}(T_n) - zI)^{-1}$ and $Y^{(n,T)}(z) = R^{(n,T)}_{11}(z)$. For all $z \in \mathbb{C}_+$, as n goes to infinity $Y^{(n,T)}(z)$ converges weakly to $Y(z)$ given by the RDE (2).*

Proof. For simplicity, we omit the superscript T and the variable z. We order the vertices of $(T_n, 1)$ according to a depth-first search in the tree. We denote by $N = D(T_n)_{11}$ the number of offsprings of the root and by $T_n^1, \cdots, T_n^N$ the subtrees of $T_n \backslash \{1\}$ ordered in the order of the depth-first search. With this ordering, we obtain a matrix $\bar{\Delta}(T_n)$ of the following shape:

$$
\begin{pmatrix}
-\alpha(N+1) & 1 & 0 & \cdots & 1 & 0 & \cdots & \cdots & 1 & 0 & \cdots \\
1 & & & & & & & & & & \\
0 & & \bar{\Delta}(T_n^1) & & & & & & & & \\
\vdots & & & & & & & & & & \\
1 & & & & & & & & & & \\
0 & & & & \bar{\Delta}(T_n^2) & & & & & & \\
\vdots & & & & & & & & & & \\
\vdots & & & & & & & & & & \\
1 & & & & & & & & & & \\
0 & & & & & & & & \bar{\Delta}(T_n^N) & & \\
\vdots & & & & & & & & & &
\end{pmatrix}.
$$

We then use a classical of Schur decomposition formula for $\bar{\Delta}(T_n) - zI$ (see Lemma 7.2 in Appendix)

$$
Y^{(n)} = -\left(z + \alpha(D(T_n)_{11} + 1) + \sum_{2 \leq i,j \leq n} \widetilde{R}^{(n-1)}_{ij} A(T_n)_{1i} A(T_n)_{1j} \right)^{-1}, \tag{11}
$$

where $\widetilde{R}^{(n-1)} = (\widetilde{\Delta}^{(n-1)} - zI)^{-1}$ with $\widetilde{\Delta}^{(n-1)}$ is the matrix obtained from $\bar{\Delta}(T_n)$ with the first row and column deleted. It follows that $\widetilde{R}^{(n-1)}$ is decomposable in the following diagonal block form

$$
\begin{pmatrix}
R(T_n^1) & & & \\
& R(T_n^2) & & \\
& & \cdots & \\
& & & R(T_n^N)
\end{pmatrix},
$$

14

where $R(T_n^i) = (\bar{\Delta}(T_n^i) - zI)^{-1}$. In particular, we get

$$\widetilde{R}_{ij}^{(n-1)} A(T_n)_{1i} A(T_n)_{1j} = 0 \quad \text{if} \quad i \neq j. \tag{12}$$

Indeed, if $A(T_n)_{1i} A(T_n)_{1j} = 1$ then both i and j are offsprings of 1 in separate subtrees so that $\widetilde{R}_{ij}^{(n-1)} = 0$.

Since T_n is a Galton Watson tree of depth n, the number of offsprings of the root, $N = D(T_n)_{11}$, has distribution F. Moreover the subtrees $T_n^1, \cdots, T_n^N$ are iid with common distribution T_{n-1}, and are independent of N. We now define

$$Y_i^{(n-1)} = \left((\bar{\Delta}(T_n^i) - zI)^{-1} \right)_{v_i, v_i} = \widetilde{R}_{v_i, v_i}^{(n-1)}.$$

It follows that $(Y_1^{(n-1)}, \cdots, Y_N^{(n-1)})$ are iid, independent of N, with the same common law than $Y^{(n-1)}$. From Equations (11), (12), we deduce

$$Y^{(n)} = -\left(z + \alpha(N+1) + \sum_{i=1}^{N} Y_i^{(n-1)} \right)^{-1},$$

In other words, with a slight abuse of notation, and identifying a random variable with its distribution, we have $Y^{(n)} = \Psi(Y^{(n-1)})$, where the mapping Ψ was defined in §4.1. The end of the proof follows directly from Lemma 2.2. □

Proof of Proposition 2.1. Again, we omit the superscript T and the variable z. As above, we use the decomposition formula:

$$X^{(n)} = -\left(z + \alpha D(T_n)_{11} + \sum_{2 \leq i,j \leq n} \widetilde{R}_{ij}^{(n-1)} A(T_n)_{1i} A(T_n)_{1j} \right)^{-1},$$

where $\widetilde{R}^{(n-1)} = (\widetilde{\Delta}^{(n-1)} - zI)^{-1}$ with $\widetilde{\Delta}^{(n-1)}$ is the matrix obtained from $\Delta(T_n)$ with the first row and column deleted. As above, since T_n is a tree $\widetilde{R}_{ij}^{(n-1)} A(T_n)_{1i} A(T_n)_{1j} = 0$ if $i \neq j$, so that we get

$$X^{(n)} = -\left(z + \alpha N_* + \sum_{i=1}^{N_*} Y_i^{(n-1)} \right)^{-1},$$

where N_* has distribution F_* and $Y_i^{(n-1)}$ are iid copies of $Y^{(n-1)}$, independent of N_*, defined in Lemma 2.3. Proposition 2.1 follows easily from Lemma 2.3. □

4.3 Proof of Theorem 2 (ii) - bounded degree

We first assume that Assumption (A") holds. Let T be a GWT with degree distribution F_* and T_n be the restriction of T to the set of vertices at distance at most n from the root (i.e. T_n

is stopped at generation n). Let $\Delta = \Delta(T)$ denote the operator associated to T, as in Section 3.2, Δ is self-adjoint and we may define for all $z \in \mathbb{C}_+$, $R(z) = (\Delta - zI)^{-1}$. Then by Theorem VIII.25(a) in [31], $\Delta(T_n)$ converges to Δ in the strong resolvent sense. Hence, by Proposition 2.1 we have

$$\langle R(z)o, o \rangle \overset{d}{=} X(z).$$

Now, as in §3.2, 3.5, $\Delta^{(n)}$ and Δ are self adjoint operators and $\mathbb{E}U(\overline{G}_n) \Rightarrow \overline{\rho}$. From Theorem 1, we have

$$m_n(z) \xrightarrow{L^1} \int \langle R(z)o, o \rangle d\overline{\rho}[G, o].$$

The proof of Theorem 2 (ii) is complete with the extra assumption (A").

4.4 Proof of Theorem 2 (ii) - general case

We now relax assumption (A") by assumption (A). From Theorem 1, it is sufficient to prove that $\lim_n \mathbb{E}m_n(z)$ converges to $\mathbb{E}X(z)$, where X is defined by Equation (3). By the same argument as in §3.4, it is sufficient to prove that, for the weak convergence on $\mathcal{H}$,

$$\lim_{\ell \to \infty} X^{(\ell)} = X$$

where $X^{(\ell)}$ is defined by Equation (3) with a degree distribution $F_*^{(\ell)}$ which converges weakly to F as ℓ goes to infinity. This continuity property is established in Lemma 7.3 (in Appendix). The proof of Theorem 2 is complete.

5 Applications and extensions

5.1 Weighted graphs

A weighted graph is a graph $G = (V, E)$ with attached weights on its edges. As in §2.1, we consider a sequence of graphs G_n on $[n]$. We define the symmetric matrix $W = (w_{ij})_{1 \le i,j \le n}$, where $(w_{ij})_{1 \le i \le j}$ is a sequence of iid real variables, independent of G_n and $w_{ij} = w_{ji}$. Let $\circ$ denote the Hadamard product (for all i, j, $(A \circ B)_{ij} = A_{ij}B_{ij}$) and let $T(G_n)$ be the diagonal matrix whose entry (i, i) is equal to $\sum_j w_{ij} A_{ij}^{(n)}$. We define $\mu_{\overline{\Delta}_n}$ as the spectral measure of the matrix $\overline{\Delta}_n = W \circ A(G_n) - \alpha T(G_n)$. If o denotes the uniformly picked root of G_n, we assume

B. The sequence of variables $\left(\sum_{k=1}^n A_{ok}^{(n)} |w_{ok}| \right), n \in \mathbb{N}$, is uniformly integrable.

An easy extension of Theorem 1 is

16

Theorem 3   *(i) Let $G_n = ([n], E_n)$ be a sequence of graphs satisfying assumptions (D-B). Then there exists a probability measure ν on $\mathbb{R}$ such that a.s. $\lim_{n\to\infty} \mu_{\overline{\Delta}_n} = \nu$.*

(ii) Let $G_n = ([n], E_n)$ be a sequence of random graphs satisfying Assumptions (R-B). Then there exists a probability measure ν on $\mathbb{R}$ such that, $\lim_{n\to\infty} \mathbb{E}L(\mu_{\overline{\Delta}_n}, \nu) = 0$.

The only difference with the proof of Theorem 1 appears in §3.4, for $\ell \in \mathbb{N}$, the matrix $\Delta_{n,\ell}$ is now equal to, for $i \neq j$

$$(\Delta_{n,\ell})_{ij} = \begin{cases} A(G_n)_{ij} w_{ij} & \text{if } \max\{\sum_{k=1}^n A_{ik}^{(n)} |w_{ik}|, \sum_{k=1}^n A_{jk}^{(n)} |w_{jk}|\} \leq \ell \\ 0 & \text{otherwise} \end{cases}$$

and $(\Delta_{n,\ell})_{ii} = -\alpha \sum_{j\neq i} (\Delta_{n,\ell})_{ij}$. The remainder is identical.

We may also state an analog of Theorem 2 for the case $\alpha = 0$, that is $\overline{\Delta}_n = W \circ A(G_n)$. We denote by s_n the Stieltjes transform of $\mu_{\overline{\Delta}_n}$. Assumption (A') is strengthen into

B'. Assumption (B) holds, $\sum_k k^2 F_*(k) < \infty$ and $\mathbb{E}[w_{12}^2] < \infty$.

The proof of the next result is a straightforward extension of the proof of Theorem 2.

Theorem 4 *Assume that assumptions (RT-B') hold and $\alpha = 0$ then*

(i) There exists a unique probability measure $P \in \mathcal{P}(\mathcal{H})$ such that for all $z \in \mathbb{C}_+$,

$$Y(z) \overset{d}{=} -\left(z + \sum_{i=1}^N |w_i|^2 Y_i(z) \right)^{-1},$$

where N has distribution F, w_i are iid copies with distribution w_{12}, Y and Y_i are iid copies with law P and the variables N, w_i, Y_i are independent.

(ii) For all $z \in \mathbb{C}_+$, $s_n(z)$ converges as n tends to infinity in L^1 to $\mathbb{E}X(z)$, where for all $z \in \mathbb{C}_+$,

$$X(z) \overset{d}{=} -\left(z + \sum_{i=1}^{N_*} |w_i|^2 Y_i(z) \right)^{-1},$$

where N_* has distribution F_*, w_i are iid copies with distribution w_{11}, Y_i are iid copies with law P and the variables N, w_i, Y_i are independent.

The case $\alpha = 1$ is more complicated: the diagonal term $T(G_n)$ introduces a dependence within the matrix which breaks the nice recursive structure of the RDE.

5.2 Bipartite graphs

In §2.2, we have considered a sequence of random graphs converging weakly to a GWT tree. Another important class of random graphs are the bipartite graphs. A graph $G = (V, E)$ is bipartite if there exists two disjoint subsets V^a, V^b, with $V^a \cup V = V$ such that all edges in E have an adjacent vertex in V^a and the other in V^b. In particular, if n and p are the cardinals of V^a and V^b, the adjacency matrix of a bipartite graph may be written as $\begin{pmatrix} 0 & M^* \\ M & 0 \end{pmatrix}$ for some $n \times p$ matrix M. The analysis of random bipartite graphs finds strong motivation in coding theory, see for example Richardson and Urbanke [32]. Note also that if M is an $n \times p$ matrix, the spectrum of M^*M may be obtained from the spectrum of $\begin{pmatrix} 0 & M^* \\ M & 0 \end{pmatrix}$. We may thus find another motivation in sparse statistical problem, see El Karoui [16].

The natural limit for random bipartite graphs is the following Bipartite Galton-Watson Tree (BGWT) with degree distribution (F_*, G_*) and scale $p \in (0, 1)$. The BGWT is obtained from a Galton-Watson branching process with alternated degree distribution. With probability p, the root has offspring distribution F_*, all odd generation genitors have an offspring distribution G, and all even generation genitors (apart from the root) have an offspring distribution F. With probability $1 - p$, the root has offspring distribution G_*, all odd generation genitors have an offspring distribution F, and all even generation genitors have an offspring distribution G.

We now consider a sequence (G_n) of random bipartite graphs satisfying assumptions $(R - A')$ with weak limit a BGWT with degree distribution (F_*, G_*) and scale $p \in (0, 1)$. The weak convergence of a natural ensemble of bipartite graphs toward a BGWT with degree distribution (F_*, G_*) and scale $p \in (0, 1)$ follows from [32], p being the proportion of vertices in V^a, F_* the asymptotic degree distribution of vertices in V^a and b_* the asymptotic degree distribution of vertices in V^a. As usual, we denote by μ_{Δ_n} the spectral measure of Δ_n, m_n is the Stieltjes transform of μ_{Δ_n}. We give without proof the following theorem which is a generalization of Theorem 2.

Theorem 5 *Under the foregoing assumptions,*

(i) *There exists a unique pair of probability measures $(R^a, R^b) \in \mathcal{P}(\mathcal{H}) \times \mathcal{P}(\mathcal{H})$ such that for all $z \in \mathbb{C}_+$,*

$$Y^a(z) \overset{d}{=} - \left(z + \alpha(N^a + 1) + \sum_{i=1}^{N^a} Y_i^b(z) \right)^{-1},$$

$$Y^b(z) \overset{d}{=} - \left(z + \alpha(N^b + 1) + \sum_{i=1}^{N^b} Y_i^a(z) \right)^{-1},$$

where N^a (resp. N^b) has distribution F (resp. G) and Y^a, Y_i^a (resp. Y^b, Y_i^b) are iid copies with law R^a (resp. R^b), and the variables N^b, Y_i^a, N^a, Y_i^b are independent.

(ii) For all $z \in \mathbb{C}_+$, $m_n(z)$ converges as n tends to infinity in L^1 to $p\mathbb{E}X^a(z) + (1-p)\mathbb{E}X^b(z)$ where for all $z \in \mathbb{C}_+$,

$$X^a(z) \overset{d}{=} -\left(z + \alpha N_*^a + \sum_{i=1}^{N_*^a} Y_i^b(z)\right)^{-1},$$

$$X^b(z) \overset{d}{=} -\left(z + \alpha N_*^b + \sum_{i=1}^{N_*^b} Y_i^a(z)\right)^{-1},$$

where N_*^a (resp. N_*^b) has distribution F_* (resp. G_*) and Y_i^a (resp. Y_i^b) are iid copies with law R^a (resp. R^b), independent of N_*^b (resp. N_*^a).

In the case $\alpha = 0$ and for bi-regular graphs (i.e. BGWT with degree distribution (δ_k, δ_l) and parameter p), the limiting spectral measure is already known and first derived by Godsil and Mohar [18], see also Mizuno and Sato [27] for an alternative proof.

5.3   Uniform random trees

The uniformly distributed tree on $[n]$ converges weakly to the Skeleton tree T_∞ which is defined as follows. Consider a sequence $T_0, T_1, \cdots$ of independent GWT with offspring distribution the Poisson distribution with intensity 1 and let $v_0, v_0, \cdots$ denote their roots. Then add all the edges (v_i, v_{i+1}) for $i \geq 0$. The distribution in $\mathcal{G}^*$ of the corresponding infinite tree is the Skeleton tree. See [2] for further properties and Grimmett [19] for the original proof of the weak convergence of the uniformly distributed tree on $[n]$ to the Skeleton tree T_∞.

Let $\mu_{A(G_n)}$ denote the spectral measure of the adjacency matrix of the random spanning tree T_n on $[n]$ drawn uniformly (for simplicity of the statement, we restrict ourselves to the case $\alpha = 0$). We denote by $m_n(z)$ the Stieltjes transform of $\mu_{A(G_n)}$.

As an application of Theorems 1, 2, we have the following:

Theorem 6 *Assume $\alpha = 0$.*

(i) There exists a unique probability measure $R \in \mathcal{P}(\mathcal{H})$ such that for all $z \in \mathbb{C}_+$,

$$X(z) \overset{d}{=} \left(W(z)^{-1} - X_1(z)\right)^{-1}, \tag{13}$$

where $W \overset{d}{=} (A(T_0) - zI)_{v_0 v_0}^{-1} \in \mathcal{H}$ is the resolvent taken at the root of a GWT with offspring distribution $Poi(1)$, X and X_1 have law R and the variables W and X_1 are independent.

(ii) For all $z \in \mathbb{C}_+$, $m_n(z)$ converges as n tends to infinity in L^1 to $\mathbb{E}X(z)$.

Sketch of Proof. The sequence T_n satisfy (R-A'), thus, from Theorem 1, it is sufficient to show that the resolvent operator $R = (A - zI)^{-1}$ of the Skeleton tree taken at the root satisfies the RDE (13). The distribution invariant structure of T_∞ implies that $X(z) \stackrel{d}{=} R(z)_{v_1 v_1}$. The number of offsprings of the root, v_1 of T_1 is a Poisson random variable with intensity 1, say N. We denote the offsprings of v_1 in T_1 by $v_{i_1}^1, \cdots, v_{i_N}^1$. From the Schur decomposition formula, we have

$$R(z)_{v_1 v_1} = - \left(z + \widetilde{R}(z)_{v_2 v_2} + \sum_{i=1}^{N} R^i(z)_{v_i^1 v_i^1} \right)^{-1},$$

where $\widetilde{R}(z)$ is the resolvent of the infinite tree obtained by removing T_1 and v_1 and $R^i(z)$ is the resolvent of the subtree of the descendants of v_i^1 in T_1. Now, by construction $\widetilde{R}(z)$ has the same distribution than $R(z)$, and $R^i(z)$ are independent copies, independent of $\widetilde{R}(z)$ of $W(z)$. Thus we obtain

$$X(z) \quad \stackrel{d}{=} \quad - \left(z + X_1(z) + \sum_{i=1}^{N} W_i(z) \right)^{-1}, \tag{14}$$

$$\stackrel{d}{=} \quad - \left(-W(z)^{-1} + X_1(z) \right)^{-1}, \tag{15}$$

where in (15), we have applied (2) for GWT with Poisson offspring distribution. The existence and unicity of the solution of (13) follows from (14) using the same proof as in Lemma 2.2. □

Appendix

For a proof of the next theorem, see e.g. Billingsley [6].

Theorem 7 (Skorokhod Representation Theorem) *Let μ_n, $n \in \mathbb{N}$, be a sequence of probability measures on a complete metric separable space S. Suppose that μ_n converges weakly to a probability measure μ on S. Then there exist random variables X_n, X defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that μ_n is the distribution of X_n, μ is the distribution of X, and for every $\omega \in \Omega$, $X_n(\omega)$ converges to $X(\omega)$.*

The next lemma is a consequence of Lidskii's inequality. For a proof see Theorem 11.42 in [3]

Lemma 7.1 (Rank difference inequality) *Let A, B be two $n \times n$ Hermitian matrices with empirical spectral measures $\mu_A = \frac{1}{n} \sum_{i=1}^{n} \delta_{\lambda_i(A)}$ and $\mu_B = \frac{1}{n} \sum_{i=1}^{n} \delta_{\lambda_i(A)}$. Then*

$$L(\mu_A, \mu_B) \leq \frac{1}{n} \mathrm{rank}(A - B).$$

The next lemma is a standard tool of random matrix theory.

Lemma 7.2 (Schur formula) *Let* $B = \begin{pmatrix} b_{11} & u^* \\ u & \tilde{B} \end{pmatrix}$, *denote a* $n \times n$ *hermitian invertible matrix,* u *being a vector of dimension* $n - 1$. *Then*

$$(B^{-1})_{11} = \left(b_{11} - u^* \tilde{B}^{-1} u \right)^{-1}.$$

Lemma 7.3 *Let* $(F_*^{(n)}), n \in \mathbb{N}$, *be a sequence of probability measures on* $\mathbb{N}$ *converging weakly to* F_* *such that* $\sup_n \sum k F_*^{(n)}(k) < \infty$. *Denote by* $X^{(n)}$ *and* X *the variable defined by (3) with degree distribution* $F_*^{(n)}$ *and* F_* *respectively. Then, for the weak convergence on* $\mathcal{H}$, $\lim_n X^{(n)} = X$.

Proof. Let F_* and F_*' be two probability measures on $\mathbb{N}$ with finite mean, and let $d_{TV}(F_*, F_*') = \sup_{A \subset \mathbb{N}} |\int_A F_*(dx) - \int_A F_*'(dx)| = 1/2 \sum_k |F_*(k) - F_*'(k)|$ be the total variation distance. Let N_*, N_*', N, N' denote variables with law F_*, F_*', F, F' respectively, and coupled so that $2\mathbb{P}(N_* \neq N_*') = d_{TV}(F_*, F'*)$ and $2\mathbb{P}(N \neq N') = d_{TV}(F, F')$ (the existence of these variables is guaranteed by the coupling inequality). We now reintroduce the distance defined in the proof of Lemma 2.2. Let Ω be a bounded open set in $\mathbb{C}_+$ with an empty intersection with the ball of center 0 and radius $\sqrt{\mathbb{E}N} + 1$. Let $\mathcal{P}(\mathcal{H})$ be the set of probability measures on $\mathcal{H}$. We define the distance on $\mathcal{P}(\mathcal{H})$

$$W(P, P') = \inf \mathbb{E} \int_\Omega |X(z) - X'(z)| dz$$

where the infimum is over all possible coupling of the distributions P and P' where X has law P and X' has law P'. With our assumptions, we may introduce the variables $X := X$ (with law P) and $X' := X'$ (with law P') defined by (3) with degree distribution F_* and F_*' respectively. The proof of the lemma will be complete if we prove that there exists C, not depending on F_* and F_*', such that

$$W(P, P') \leq C \max(d_{TV}(F_*, F_*'), d_{TV}(F, F')). \tag{16}$$

We denote by Y (with law Q) and Y' (with law Q') the variable defined by (2) with offspring distribution F and F', coupled so that $W(Q, Q') = \mathbb{E} \int_\Omega |Y(z) - Y'(z)| dz$. We consider $(Y_i, Y_i')_{i \in \mathbb{N}}$ iid copies of (Y, Y') and independent of the variable N_*. By definition, we have the following

$$
\begin{aligned}
W(P, P') &\leq \mathbb{E} \int_\Omega |X'(z) - X(z)| \, \mathbf{1}(N_* \neq N_*') dz \\
&\quad + \mathbb{E} \int_\Omega \left| \left(z + \alpha N_* + \sum_{i=1}^{N_*} Y_i(z) \right)^{-1} - \left(z + \alpha N_* + \sum_{i=1}^{N_*} Y_i'(z) \right)^{-1} \right| dz \\
&\leq d_{TV}(F_*, F_*') \int_\Omega (\Im z)^{-1} dz + \int_\Omega (\Im z)^{-2} \mathbb{E} \sum_{i=1}^{N_*} |Y_i(z) - Y_i'(z)| \, dz \\
&\leq d_{TV}(F_*, F_*') \int_\Omega (\Im z)^{-1} dz + \mathbb{E} N_* (\inf_{z \in \Omega} \Im z)^{-2} W(Q, Q'). \tag{17}
\end{aligned}
$$

We then argue as in the proof of Lemma 2.2. Since $\Psi(Q) = Q$ and $\Psi'(Q') = Q'$ (where Ψ' is defined as Ψ with the distribution F' instead of F), we get:

$$
\begin{aligned}
W(Q, Q') &= W(\Psi(Q), \Psi'(Q')) \\
&\leq \mathbb{E} \int_\Omega |\psi(N, (Y_i); z) - \psi(N', (Y_i'); z)| dz \\
&\leq d_{TV}(F, F') \int_\Omega (\Im z)^{-1} dz + \mathbb{E} \int_\Omega |\psi(N, (Y_i); z) - \psi(N, (Y_i'); z)| dz \\
&\leq d_{TV}(F, F') \int_\Omega (\Im z)^{-1} dz + \mathbb{E} N (\inf_{z \in \Omega} \Im z)^{-2} W(Q, Q').
\end{aligned}
$$

Then since $\inf_{z \in \Omega} \Im z > \sqrt{\mathbb{E} N}$, we deduce that

$$
W(Q, Q') \leq d_{TV}(F, F') \frac{\int_\Omega (\Im z)^{-1} dz}{1 - \mathbb{E} N (\inf_{z \in \Omega} \Im z)^{-2}}.
$$

This last inequality, together with (17), implies (16). $\qquad\square$

Acknowledgment

References

[1] D. Aldous and R. Lyons. Processes on unimodular random networks. *Electronic Journal of Probability*, 12:1454–1508, 2007.

[2] D. Aldous and J. M. Steele. The objective method: probabilistic combinatorial optimization and local weak convergence. In *Probability on discrete structures*, volume 110 of *Encyclopaedia Math. Sci.*, pages 1–72. Springer, Berlin, 2004.

[3] Z. D. Bai and J. W. Silverstein. *Spectral Analysis of Large Dimensional Random Matrices*. Mathematics Monograph Series 2. Science Press, Beijing, 2006.

[4] M. Bauer and O. Golinelli. Random incidence matrices: moments of the spectral density. *J. Statist. Phys.*, 103(1-2):301–337, 2001.

[5] I. Benjamini and O. Schramm. Recurrence of distributional limits of finite planar graphs. *Electron. J. Probab.*, 6:no. 23, 13 pp. (electronic), 2001.

[6] P. Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Inc., New York, second edition, 1999. A Wiley-Interscience Publication.

[7] G. Biroli and R. Monasson. A single defect approximation for localized states on random lattices. *Journal of Physics A: Mathematical and General*, 32(24):L255–L261, 1999.

[8] B. Bollobás. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European J. Combin.*, 1(4):311–316, 1980.

[9] B. Bollobas and O. Riordan. Sparse graphs: metrics and random models. *eprint arXiv: 0812.2656*, 2008.

[10] C. Bordenave, P. Caputo, and D. Chafai. Spectrum of large random reversible markov chains - heavy-tailed weights on the complete graph, 2009.

[11] C. Borgs, J. Chayes, L. Lovász, V. T. Sós, B. Szegedy, and K. Vesztergombi. Graph limits and parameter testing. In *STOC'06: Proceedings of the 38th Annual ACM Symposium on Theory of Computing*, pages 261–270. ACM, New York, 2006.

[12] C. Borgs, J. T. Chayes, L. Lovász, V. T. Sós, and K. Vesztergombi. Convergent sequences of dense graphs. I. Subgraph frequencies, metric properties and testing. *Adv. Math.*, 219(6):1801–1851, 2008.

[13] Y. Colin de Verdière. *Spectres de graphes*, volume 4 of *Cours Spécialisés [Specialized Courses]*. Société Mathématique de France, Paris, 1998.

[14] J. B. Conway. *Functions of one complex variable*. Springer-Verlag, New York, 1973. Graduate Texts in Mathematics, 11.

[15] R. Durrett. *Random graph dynamics*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, 2007.

[16] N. El Karoui. Spectrum estimation for large dimensional covariance matrices using random matrix theory. *Ann. Stat.*, page to appear.

[17] G. Elek. On limits of finite graphs. *Combinatorica*, 27(4):503–507, 2007.

[18] C. D. Godsil and B. Mohar. Walk generating functions and spectral measures of infinite graphs. In *Proceedings of the Victoria Conference on Combinatorial Matrix Analysis (Victoria, BC, 1987)*, volume 107, pages 191–206, 1988.

[19] G. R. Grimmett. Random labelled trees and their branching networks. *J. Austral. Math. Soc. Ser. A*, 30(2):229–237, 1980/81.

[20] T. Guhr, A. Müller-Groeling, and H. A. Weidenmüller. Random-matrix theories in quantum physics: common concepts. *Phys. Rep.*, 299(4-6):189–425, 1998.

[21] F. Hiai and D. Petz. *The semicircle law, free random variables and entropy*, volume 77 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2000.

[22] H. Kesten. Symmetric random walks on groups. *Trans. Amer. Math. Soc.*, 92:336–354, 1959.

[23] O. Khorunzhy, M. Shcherbina, and V. Vengerovsky. Eigenvalue distribution of large weighted random graphs. *J. Math. Phys.*, 45(4):1648–1672, 2004.

[24] L. Lovász and B. Szegedy. Limits of dense graph sequences. *J. Combin. Theory Ser. B*, 96(6):933–957, 2006.

[25] B. D. McKay. The expected eigenvalue distribution of a large regular graph. *Linear Algebra Appl.*, 40:203–216, 1981.

[26] M. L. Mehta. *Random matrices*, volume 142 of *Pure and Applied Mathematics (Amsterdam)*. Elsevier/Academic Press, Amsterdam, third edition, 2004.

[27] H. Mizuno and I. Sato. The semicircle law for semiregular bipartite graphs. *J. Combin. Theory Ser. A*, 101(2):174–190, 2003.

[28] B. Mohar. The spectrum of an infinite graph. *Linear Algebra Appl.*, 48:245–256, 1982.

[29] B. Mohar and W. Woess. A survey on spectra of infinite graphs. *Bull. London Math. Soc.*, 21(3):209–234, 1989.

[30] M. Molloy and B. Reed. A critical point for random graphs with a given degree sequence. In *Proceedings of the Sixth International Seminar on Random Graphs and Probabilistic Methods in Combinatorics and Computer Science, "Random Graphs '93" (Poznań, 1993)*, volume 6, pages 161–179, 1995.

[31] M. Reed and B. Simon. *Methods of modern mathematical physics. I. Functional analysis.* Academic Press, New York, 1972.

[32] T. Richardson and R. Urbanke. *Modern Coding Theory*. Cambridge University Press, 2008.

[33] G. J. Rodgers and A. J. Bray. Density of states of a sparse random matrix. *Phys. Rev. B (3)*, 37(7):3557–3562, 1988.

[34] T. Rogers, I. P. Castillo, R. Kühn, and K. Takeda. Cavity approach to the spectral density of sparse symmetric random matrices. *Physical Review E (Statistical, Nonlinear, and Soft Matter Physics)*, 78(3):031116, 2008.

[35] G. Semerjian and L. F. Cugliandolo. Sparse random matrices: the eigenvalue spectrum revisited. *Journal of Physics A: Mathematical and General*, 35(23):4837–4851, 2002.

[36] J.-P. Serre. Répartition asymptotique des valeurs propres de l'opérateur de Hecke T_p. *J. Amer. Math. Soc.*, 10(1):75–102, 1997.

[37] E. P. Wigner. *The collected works of Eugene Paul Wigner. Part A. The scientific papers. Vol. II. Nuclear physics.* Springer-Verlag, Berlin, 1996. Annotated by Herman Feshbach, Edited and with a preface by Arthur S. Wightman and Jagdish Mehra.

[38] N. C. Wormald. Models of random regular graphs. In *Surveys in combinatorics, 1999 (Canterbury)*, volume 267 of *London Math. Soc. Lecture Note Ser.*, pages 239–298. Cambridge Univ. Press, Cambridge, 1999.