The Annals of Probability
2011, Vol. 39, No. 3, 1097–1121
DOI: 10.1214/10-AOP567

THE RANK OF DILUTED RANDOM GRAPHS

BY CHARLES BORDENAVE, MARC LELARGE AND JUSTIN SALEZ

Université Toulouse III, INRIA and INRIA

We investigate the rank of the adjacency matrix of large diluted random graphs: for a sequence of graphs $(G_n)_{n\geq 0}$ converging locally to a Galton–Watson tree T (GWT), we provide an explicit formula for the asymptotic multiplicity of the eigenvalue 0 in terms of the degree generating function φ_* of T. In the first part, we show that the adjacency operator associated with T is always self-adjoint; we analyze the associated spectral measure at the root and characterize the distribution of its atomic mass at 0. In the second part, we establish a sufficient condition on φ_* for the expectation of this atomic mass to be precisely the normalized limit of the dimension of the kernel of the adjacency matrices of $(G_n)_{n\geq 0}$. Our proofs borrow ideas from analysis of algorithms, functional analysis, random matrix theory and statistical physics.

1. Introduction. In this paper we investigate asymptotical spectral properties of the adjacency matrix of large random graphs. To motivate our work, let us briefly mention its implications in the special case of Erdős–Rényi random graphs. Let $G_n = (V_n, E_n)$ be an Erdős–Rényi graph with connectivity $c > 0$ on the vertex set $V_n = \{1, \ldots, n\}$. In other words, we let each pair of distinct vertices ij belong to the edge-set E_n with probability c/n, independently of the other pairs. The adjacency matrix A_n of G_n is the $n \times n$ symmetric matrix defined by $(A_n)_{ij} = \mathbf{1}((ij) \in E_n)$. Let $\lambda_1(A_n) \geq \cdots \geq \lambda_n(A_n)$ denote the eigenvalues of A_n (with multiplicities) and

$$\mu_n = \frac{1}{n}\sum_{i=1}^{n} \delta_{\lambda_i(A_n)}$$

denote the spectral measure of A_n. Our main concern will be the rank of A_n

$$\operatorname{rank}(A_n) = n - \dim\ker(A_n) = n - n\mu_n(\{0\}).$$

THEOREM 1. (i) *There exists a deterministic symmetric measure μ such that, almost surely, for the weak convergence of probability measures,*

$$\lim_{n\to\infty} \mu_n = \mu.$$

(ii) *Let $0 < q < 1$ be the smallest solution to $q = \exp(-c\exp(-cq))$. Then almost surely,*

$$\lim_{n\to\infty} \mu_n(\{0\}) = \mu(\{0\}) = q + e^{-cq} + cqe^{-cq} - 1.$$

Received September 2009; revised March 2010.
MSC2010 subject classifications. Primary 05C80, 15A52; secondary 47A10.
Key words and phrases. Random graphs, adjacency matrix, random matrices, local weak convergence, Karp and Sipser algorithm.

In other words, almost surely,

$$\lim_{n\to\infty} \frac{\operatorname{rank}(A_n)}{n} = 2 - q - e^{-cq} - cqe^{-cq}. \tag{1}$$

Apart from an improvement of the convergence, part (i) is not new; the convergence in probability was first rigorously proved by Khorunzhy, Shcherbina and Vengerovsky [16] (for an alternative proof, see [10] [note that it only implies $\limsup_n \mu_n(\{0\}) \le \mu(\{0\})$]).

In the sparse case, that is, when the connectivity c grows with n like $a \log n$, the rank of A_n has been studied by Costello, Tao and Vu [13] and Costello and Vu [12]. Their results imply that for $a > 1$, with high probability $\dim\ker(A_n) = 0$ while for $0 < a < 1$, $\dim\ker(A_n)$ is of order of magnitude n^{1-a}. Our theorem answers one of their open questions in [12].

The formula (1) already appeared in a remarkable paper by Karp and Sipser [15] as the asymptotic size of the number of vertices left unmatched by a maximum matching of G_n. To be more precise, the function $G \mapsto \dim\ker(G)$ is easily checked to be invariant under "leaf removal," that is, if G' is the graph obtained from G by deleting a leaf and its unique neighbor, then $\dim\ker(G') = \dim\ker(G)$. Karp and Sipser [15] study the effect of iterating this leaf removal on the random graph G_n until only isolated vertices and a "core" with minimum degree at least 2 remain. They show that the asymptotic number of isolated vertices is approximately $(2 - q - e^{-cq} - cqe^{-cq})n$ as $n \to \infty$, and that the size of the core is $o(n)$ when $c \le e$. Thus, (1) follows by additivity of $G \mapsto \dim\ker(G)$ on disjoint components, as observed by Bauer and Golinelli [6]. However for $c > e$, the size of the core is not negligible and the same argument only leads to the following inequality:

$$\liminf_{n\to\infty} \frac{\dim\ker(A_n)}{n} \ge q + e^{-cq} + cqe^{-cq} - 1.$$

Bauer and Golinelli [6] conjecture that this lower bound should be the actual limit for all c, which is equivalent to saying that asymptotically the dimension of the kernel of the core is zero. The proof of this conjecture follows from our work (see Section 4).

Our results are not restricted to Erdős–Rényi graphs. They will in fact hold for any sequence $(G_n)_{n\ge1}$ of random graphs converging locally to a rooted Galton–Watson tree (GWT), provided the latter satisfies a certain degree condition. The precise definition of local convergence is recalled in Section 3. It was introduced by Benjamini and Schramm [7] and Aldous and Steele [3]. A rooted GWT (see [2]) is characterized by its *degree distribution* F_*, which can be any probability measure with finite mean on $\mathbb{N}$: the root Ø has offspring distribution F_* and all other genitors have offspring distribution F, where for all $k \ge 1$, $F(k-1) = kF_*(k)/\sum_\ell \ell F_*(\ell)$. In the case of Erdős–Rényi graphs with connectivity c, the limiting tree is simply a GWT with degree distribution $F_* = \operatorname{Poisson}(c)$.

The adjacency operator A of a GWT $[T = (V, E)]$ is a densely defined symmetric linear operator on the Hilbert space $\ell^2(V)$ defined for $\mathbf{i}, \mathbf{j}$ in V by

$$\langle Ae_{\mathbf{i}}, e_{\mathbf{j}} \rangle = \mathbf{1}(\mathbf{ij} \in E),$$

where for any $\mathbf{i} \in V$, $e_{\mathbf{i}}$ denotes the base function $\mathbf{j} \in V \mapsto \mathbf{1}(\mathbf{j} = \mathbf{i})$. As we will show, if F_* has a finite second moment, then A has almost surely a unique self-adjoint extension, which we also denote by A. Consequently, for any unitary vector $\psi \in \mathrm{Dom}(A)$, the spectral theorem guarantees the existence and uniqueness of a probability measure μ_ψ on $\mathbb{R}$, called the spectral measure associated with ψ, such that for any $k \geq 0$,

$$\langle A^k \psi, \psi \rangle = \int_{\mathbb{R}} x^k \, d\mu_\psi(x).$$

In particular, we may consider the spectral measure μ_T associated with the vector $e_\varnothing$, where Ø is the root of the rooted tree T. Our first main result is an explicit formula for $\mathbb{E}\mu_T(\{0\})$, the expected mass at zero of the spectral measure at the root Ø of a rooted GWT T.

THEOREM 2. *Let T be a GWT whose degree distribution F_* has a finite second moment, and let φ_* be the generating function of F_*. Then*, $\mathbb{E}\mu_T(\{0\}) = \max_{x \in [0,1]} M(x)$, *where*

$$M(x) = \varphi_*'(1) x \overline{x} + \varphi_*(1 - x) + \varphi_*(1 - \overline{x}) - 1 \qquad \textit{with } \overline{x} = \varphi_*'(1 - x)/\varphi_*'(1).$$

In the special case of regular trees, the measure μ_T can be explicitely computed and turns out to be absolutely continuous, so $\mu_T(\{0\}) = 0$. In contrast, one may construct GWTs with arbitrary large minimum degree and such that $\mathbb{E}\mu_T(\{0\}) > 0$. The following example is taken form [9] and is due to Picollelli and Molloy: set $d \geq 3$ and take $\varphi_*(x) = \frac{d}{1+d} x^d + \frac{1}{1+d} x^{d^3}$. Figure 1 gives a plot of M for the case $d = 3$, showing that $\mathbb{E}\mu_T(\{0\}) > 0$ in this case.

When F_* is a Poisson distribution with mean c, the corresponding quantity $\max_{x \in [0,1]} M(x)$ is precisely (1), and it already appeared in Zdeborová and Mézard [19], equation (38), as a "cavity method" prediction for the limiting fraction of unmatched vertices in a maximum matching.

To the best of our knowledge, the formula was unknown for general GWTs. However, Bauer and Gollineli [5] have computed explicitly the asymptotic rank of the uniform spanning tree on the complete graph of size n. Also Bhamidi, Evans and Sen [8] have recently analyzed the convergence of the spectrum of the adjacency matrix of growing random trees.

Our second main result (Theorem 13) states that for any sequence of random graphs $(G_n)_{n \geq 0}$ converging locally in distribution to a GWT, we have $\lim_n n^{-1} \operatorname{rank}(A_n) = 1 - \mathbb{E}\mu_T(\{0\})$, provided the first local extremum of the above function $x \mapsto M(x)$ is a global maximum on $[0, 1]$. We have left open the case where the global maximum of M is not the first local maximum (see Section 4).

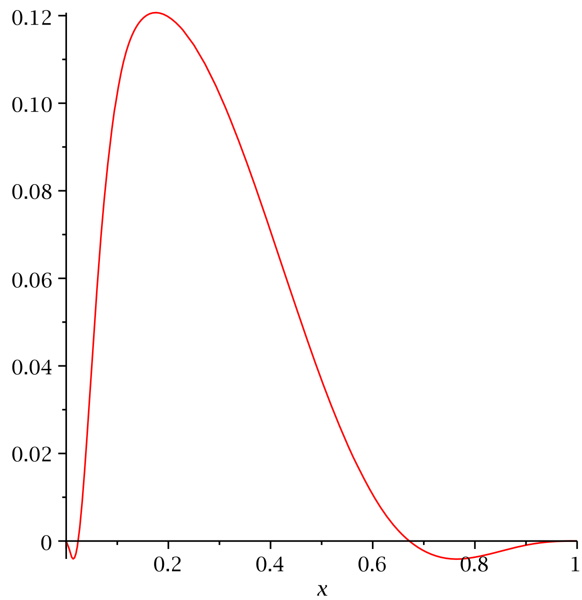

FIG. 1. *Plot of M for $\varphi_*(x) = \frac{d}{1+d}x^d + \frac{1}{1+d}x^{d^3}$, with $d = 3$.*

Our detailed analysis of the atomic mass at 0 of the limiting spectral measure μ remains only a small achievement for the global understanding of this measure. For example, for Erdős–Rényi graphs, the atomic part of μ is dense in $\mathbb{R}$, and nothing is known on the mass of atoms apart 0. There is also a conjecture about the absolutely continuous part μ_{ac} of the measure μ: we say that μ has *extended states* (resp., no extended state) at $E \in \mathbb{R}$ if the partition function $x \mapsto \mu_{ac}(-\infty, x)$ is differentiable at $x = E$ and its derivative is positive (resp., null). This notion was introduced in mathematical physics in the context of spectra of random Schrödinger operators; a recent treatment can be found in Aizenman, Sims and Warzel [1]. For Erdős–Rényi graphs, Bauer and Gollineli have conjectured that μ has no extended state at $E = 0$ when $0 < c \le e$, and has extended states at $E = 0$ when $c > e$. More generally, one may wonder whether $\mu_{ac} = 0$ when $0 < c \le e$. Finally, the existence of a singular continuous part in μ is apparently unknown.

The remainder of the paper is organized as follows: in Section 2, we analyze the adjacency operator of a GWT. In Section 2.3, we study $\mu_T(\{0\})$ and prove Theorem 2. In Section 3, we prove finally the convergence of the spectrum of finite graphs and the convergence of the rank. The proof of Theorem 1 is given in the Appendix.

2. Locally finite graphs and their adjacency operators. A rooted graph is the pair formed by a graph G with a distinguished vertex $\emptyset \in V$, called the root. There is a canonical way to define a distance on V: for each $u, v \in V$, the (graph)-distance is the minimal length of a path from u to v, if any, and ∞ otherwise. For a rooted graph G with root Ø and t an integer, we will denote by $(G)_t$ the rooted subgraph spanned by the vertices at distance at most t from the root. In all this section, we consider a locally finite rooted graph $G = (V, E)$ with root denoted by Ø.

2.1. *Adjacency operator.* Consider the Hilbert space

$$\ell^2(V) = \left\{\psi : V \to \mathbb{C}, \sum_{\mathbf{i}\in V} |\psi(\mathbf{i})|^2 < \infty\right\}$$

$$\text{with inner product } \langle\psi, \phi\rangle = \sum_{\mathbf{i}\in V} \psi(\mathbf{i})\overline{\phi(\mathbf{i})}.$$

Denote by $H_0 \subseteq \ell^2(V)$ the dense subspace of finitely supported functions, and by $(e_{\mathbf{i}})_{\mathbf{i}\in V}$ the canonical orthonormal basis of $\ell^2(V)$, that is, $e_{\mathbf{i}}$ is the coordinate function $\mathbf{j} \in V \mapsto \mathbf{1}(\mathbf{i} = \mathbf{j})$. By definition, the adjacency operator A of G is the densely-defined linear operator over $\ell^2(V)$ whose domain is H_0 and whose action on the basis vector $e_{\mathbf{i}}, \mathbf{i} \in V$, is

$$Ae_{\mathbf{i}} = \sum_{\mathbf{j}:\,\mathbf{ij}\in E} e_{\mathbf{j}}.$$

Note that $Ae_{\mathbf{i}} \in \ell^2(V)$ since G is locally finite. Moreover, for all $\mathbf{i}, \mathbf{j} \in V$,

$$\langle Ae_{\mathbf{i}}, e_{\mathbf{j}}\rangle = \mathbf{1}\{\mathbf{ij} \in E\} = \langle Ae_{\mathbf{j}}, e_{\mathbf{i}}\rangle.$$

Therefore, the operator A is symmetric, and we may now ask about the self-adjointness of its closure, which is again denoted by A. The answer of course depends upon G, but here is a simple sufficient condition that should suit all our needs in the present paper.

We define the boundary of a subset $S \subseteq V$ as $\partial S = \{\mathbf{ij} \in E : \mathbf{i} \in S, \mathbf{j} \notin S\}$, and the boundary degree $\Delta(\partial S)$ as the maximum number of boundary edges that are adjacent to the same vertex.

PROPOSITION 3. *For A to be self-adjoint, it is enough that V admits an exhausting sequence of finite subsets with bounded boundary degree*:

(A) *There exist finite subsets $S_1, S_2, \dots \subseteq V$ such that*

$$\bigcup_n S_n = V \quad \textit{and} \quad \sup_n \Delta(\partial S_n) < \infty.$$

PROOF. Denote by A^* the adjoint of A. By the basic criterion for self-adjointness (see, e.g., Reed and Simon [18], Theorem VIII.3), it is enough to show that 0 is the only vector $\psi \in \mathrm{Dom}(A^*)$ satisfying $A^*\psi = \pm i\psi$. Consider such a ψ (let us treat, say, the $+i$ case), and define the following flow along the oriented edges of G:

$$(\mathbf{i} \to \mathbf{j}) = \Im(\psi(\mathbf{i})\overline{\psi(\mathbf{j})}) = -(\mathbf{j} \to \mathbf{i}),$$

for all $\mathbf{ij} \in E$. The amount of flow created at vertex $\mathbf{i} \in V$ is then

$$\sum_{\mathbf{j}\,:\,\mathbf{ij}\in E} (\mathbf{i} \to \mathbf{j}) = \Im\left(\psi(\mathbf{i}) \sum_{\mathbf{j}\,:\,\mathbf{ij}\in E} \overline{\psi(\mathbf{j})}\right) = \Im\langle A(\psi(\mathbf{i})e_{\mathbf{i}}), \psi\rangle$$
$$= \Im\langle \psi(\mathbf{i})e_{\mathbf{i}}, A^*\psi\rangle = |\psi(\mathbf{i})|^2.$$

Now, by anti-symmetry of the flow, the total amount of flow created inside any finite subset $S \subseteq V$ must equal the total amount of flow escaping through the boundary ∂S

$$\sum_{\mathbf{i}\in S} |\psi(\mathbf{i})|^2 = \sum_{\mathbf{ij}\in\partial S} (\mathbf{i} \to \mathbf{j}).$$

Therefore, using $(\mathbf{i} \to \mathbf{j}) \le |\psi(\mathbf{i})||\psi(\mathbf{j})|$ and twice the Cauchy–Schwarz inequality, we find

$$\sum_{\mathbf{i}\in S} |\psi(\mathbf{i})|^2 \le \left(\sum_{\mathbf{i}\in\partial S^-} |\psi(\mathbf{i})|^2 \sum_{\mathbf{i}\in\partial S^-} \left(\sum_{\mathbf{j}\in S^c\cap N_{\mathbf{i}}} |\psi(\mathbf{j})|\right)^2\right)^{1/2}$$
$$\le \Delta(\partial S)\left(\sum_{\mathbf{i}\in\partial S^-} |\psi(\mathbf{i})|^2 \sum_{\mathbf{j}\in\partial S^+} |\psi(\mathbf{j})|^2\right)^{1/2},$$

where we have written $N_{\mathbf{i}}$ for the set of neighbors of $\mathbf{i}$, ∂S^- and ∂S^+ for the sets of vertices $\partial S \cap S$ and $\partial S \cap S^c$, respectively. Finally, take $S = S_n$, and let $n \to \infty$: the exhaustivity $\bigcup_n S_n = V$ ensures that the left-hand side tends to $\sum_{\mathbf{i}\in V} |\psi(\mathbf{i})|^2 = \|\psi\|^2$ and also that

$$\sum_{\mathbf{i}\in\partial S_n^-} |\psi(\mathbf{i})|^2 \xrightarrow[n\to\infty]{} 0 \quad \text{and} \quad \sum_{\mathbf{j}\in\partial S_n^+} |\psi(\mathbf{j})|^2 \xrightarrow[n\to\infty]{} 0.$$

Since $\sup_n \Delta(\partial S_n) < \infty$, the right-hand side vanishes, and we obtain the desired $\|\psi\| = 0$. $\square$

2.2. *Spectral measure*. We now assume that the adjacency operator A is self-adjoint. The spectral theorem then guarantees the validity of the *Borel functional calculus* on A: any measurable function $f : \mathbb{R} \to \mathbb{C}$ may now be rigorously applied to the operator A just as one would do with polynomials. Denoting by μ_G the spectral measure associated with the vector $e_\varnothing$, we may thus write

$$\langle f(A)e_\varnothing, e_\varnothing\rangle = \int_{\mathbb{R}} f(x)\, d\mu_G(x) \tag{2}$$

for any $f \in \mathcal{L}_{\mathbb{C}}(\mu_G)$. Taking $f(x) = x^n$ $(n \in \mathbb{N})$, we obtain in particular

$$\begin{aligned} \gamma_n &= \langle A^n e_\varnothing, e_\varnothing\rangle = \int x^n\, d\mu_G(x) \\ &= \#\{\text{paths of length } n \text{ from } \varnothing \text{ to } \varnothing \text{ in } G\}. \end{aligned} \tag{3}$$

Since $\|e_{\varnothing}\| = 1$, the spectral measure μ_T is a probability measure on $\mathbb{R}$. We will now study its Cauchy–Stieltjes transform. By definition, the Cauchy–Stieltjes transform of a probability measure μ on $\mathbb{R}$ is the holomorphic function m_μ defined on the upper complex half-plane $\mathbb{C}_+$ by

$$m_\mu : z \mapsto \int_{\mathbb{R}} \frac{d\mu(x)}{x - z}.$$

Note that m_μ belongs to the set $\mathcal{H}$ of holomorphic functions f on $\mathbb{C}_+$ satisfying

$$\forall z \in \mathbb{C}_+ \qquad \Im f(z) \geq 0 \quad \text{and} \quad |f(z)| \leq (\Im z)^{-1},$$

which is compact in the normed space of holomorphic functions on $\mathbb{C}^+$ (Montel's theorem).

Henceforth, we will assume that G is a rooted tree T. We write $\mathbf{j} \succ \mathbf{i}$ to mean that $\mathbf{i} \in V$ is an ancestor of $\mathbf{j} \in V$, and we let $T_{\mathbf{i}}$ be the subtree of T restricted to $\{\mathbf{j} \in V, \mathbf{j} \succeq \mathbf{i}\}$, rooted at $\mathbf{i}$. Its adjacency operator $A_{\mathbf{i}}$ is the projection of A on $\mathrm{Vect}(e_{\mathbf{j}}, \mathbf{j} \succeq \mathbf{i})$. Since it is also self-adjoint, we may consider its spectral measure $\mu_{T_{\mathbf{i}}}$ associated with the vector $e_{\mathbf{i}}$, and its Cauchy–Stieltjes transform $m_{T_{\mathbf{i}}}$. The recursive structure of trees implies a simple well-known recursion for the family $(m_{T_{\mathbf{i}}})_{\mathbf{i} \in V}$:

PROPOSITION 4. *The family $(m_{T_{\mathbf{i}}})_{\mathbf{i} \in V}$ is solution in $\mathcal{H}^V$ to the system of equations, for all $z \in \mathbb{C}_+$,*

$$f_{\mathbf{i}}(z) = -\Big(z + \sum_{\mathbf{j} \in D(\mathbf{i})} f_{\mathbf{j}}(z)\Big)^{-1}, \tag{4}$$

where $D(\mathbf{i}) = \{\mathbf{j} \succ \mathbf{i}, |\mathbf{j}| = |\mathbf{i}| + 1\}$ denotes the set of immediate children of $\mathbf{i}$.

PROOF. As we will see, the recursion follows from a classical operator version of the Schur complement formula (see, e.g., Proposition 2.1 in Klein [17] for a similar argument). We write the proof for completeness. Define the operator U on $\ell^2(V)$ by its matrix elements,

$$\langle U e_{\varnothing}, e_{\mathbf{i}} \rangle = \langle U e_{\mathbf{i}}, e_{\varnothing} \rangle = 1 = \langle A e_{\varnothing}, e_{\mathbf{i}} \rangle,$$

for all $\mathbf{i} \in D(\varnothing)$, and $\langle U e_{\mathbf{j}}, e_{\mathbf{k}} \rangle = 0$ otherwise. We then have the following decomposition:

$$A = U + \bigoplus_{\mathbf{i} \in D(\varnothing)} A_{\mathbf{i}},$$

where $A_{\mathbf{i}}$, is the projection of A on $V_{\mathbf{i}} = \mathrm{Vect}(e_{\mathbf{j}}, \mathbf{j} \succeq \mathbf{i})$. Since A and $\tilde{A} = \bigoplus_{\mathbf{i} \in D(\varnothing)} A_{\mathbf{i}}$ are self-adjoint operators, their respective resolvents

$$R : z \mapsto (A - zI)^{-1}, \qquad \tilde{R} : z \mapsto (\tilde{A} - zI)^{-1}$$

are well defined on $\mathbb{C}_+$, and the resolvent identity gives

$$R(z)U\tilde{R}(z) = R(z) - \tilde{R}(z). \tag{5}$$

In particular, for all $\mathbf{k} \in V$,

$$\langle R(z)U\tilde{R}(z)e_\varnothing, e_\mathbf{k}\rangle = \langle R(z)e_\varnothing, e_\mathbf{k}\rangle - \langle \tilde{R}(z)e_\varnothing, e_\mathbf{k}\rangle.$$

Now, using the definition of U, we may expand the left-hand side as

$$\left(\langle \tilde{R}(z)e_\varnothing, e_\mathbf{k}\rangle \sum_{\mathbf{i}\in D(\varnothing)} \langle R(z)e_\varnothing, e_\mathbf{i}\rangle\right) + \left(\langle R(z)e_\varnothing, e_\varnothing\rangle \sum_{\mathbf{i}\in D(\varnothing)} \langle \tilde{R}(z)e_\mathbf{i}, e_\mathbf{k}\rangle\right).$$

But $\tilde{R}(z)e_\varnothing = -z^{-1}e_\varnothing$ and each $V_\mathbf{i}, \mathbf{i} \in D(\varnothing)$, is stable for $\tilde{R}$. Therefore, in the special case where $\mathbf{k} = \varnothing$, the above equality simplifies into

$$-\frac{1}{z} \sum_{\mathbf{i}\in D(\varnothing)} \langle R(z)e_\varnothing, e_\mathbf{i}\rangle = \langle R(z)e_\varnothing, e_\varnothing\rangle + \frac{1}{z},$$

while for $\mathbf{k} \in D(\varnothing)$, it gives

$$\langle R(z)e_\varnothing, e_\varnothing\rangle\langle \tilde{R}(z)e_\mathbf{k}, e_\mathbf{k}\rangle = \langle R(z)e_\varnothing, e_\mathbf{k}\rangle.$$

Combining both, we finally obtain

$$\langle R(z)e_\varnothing, e_\varnothing\rangle = -\left(z + \sum_{\mathbf{i}\in D(\varnothing)} \langle \tilde{R}(z)e_\mathbf{i}, e_\mathbf{i}\rangle\right)^{-1},$$

which, by (2) with $f(x) = (x - z)^{-1}$, is precisely

$$m_{T_\varnothing}(z) = -\left(z + \sum_{\mathbf{i}\in D(\varnothing)} m_{T_\mathbf{i}}(z)\right)^{-1}.$$

□

When T is finite, the set of equations (4) uniquely determines the Cauchy–Stieltjes transforms $(m_{T_\mathbf{i}})_{\mathbf{i}\in V}$, which can be computed iteratively from the leaves up to the root. Under an extra condition on T, this extends to the infinite case. Recall that $(T)_n$ denote the truncation of T to the first n generations. In what follows, we will make the additional assumption

$$\limsup_{n\to\infty} |\partial(T)_n|^{1/n} < \infty. \tag{B}$$

PROPOSITION 5. *If T satisfies assumption (B), then $(m_{T_\mathbf{i}})_{\mathbf{i}\in V}$ is the unique solution in $\mathcal{H}^V$ to the system of equations (4), and for all $\mathbf{i} \in V$,*

$$m_{T_\mathbf{i}} = \lim_{n\to\infty} m_{(T_\mathbf{i})_n}, \tag{6}$$

in the sense of compact convergence on $\mathbb{C}_+$.

PROOF. If $(f_{\mathbf{i}})_{\mathbf{i}\in V} \in \mathcal{H}^V$ and $(g_{\mathbf{i}})_{\mathbf{i}\in V} \in \mathcal{H}^V$ are solutions to the system of equations (4), then we can write, for all $\mathbf{i} \in V$, $z \in \mathbb{C}_+$,

$$|f_{\mathbf{i}}(z) - g_{\mathbf{i}}(z)| = \left|\frac{\sum_{\mathbf{j}\in D(\mathbf{i})}(f_{\mathbf{j}}(z) - g_{\mathbf{j}}(z))}{(z + \sum_{\mathbf{j}\in D(\mathbf{i})} f_{\mathbf{j}}(z))(z + \sum_{\mathbf{j}\in D(\mathbf{i})} g_{\mathbf{j}}(z))}\right|$$
$$\leq \frac{1}{(\Im(z))^2}\sum_{\mathbf{j}\in D(\mathbf{i})}|f_{\mathbf{j}}(z) - g_{\mathbf{j}}(z)|.$$

Iterating this n times, and then using the uniform bound $|f_{\mathbf{j}}(z) - g_{\mathbf{j}}(z)| \leq 2 \times (\Im(z))^{-1}$, we obtain

$$|f_{\mathbf{i}}(z) - g_{\mathbf{i}}(z)| \leq \frac{1}{(\Im(z))^{2n}}\sum_{\mathbf{j}\in\partial(T_{\mathbf{i}})_n}|f_{\mathbf{j}}(z) - g_{\mathbf{j}}(z)| \leq \frac{2|\partial(T_{\mathbf{i}})_n|}{(\Im(z))^{2n+1}}.$$

Therefore, we see that under assumption (B),

$$\forall \mathbf{i} \in V \qquad |f_{\mathbf{i}}(z) - g_{\mathbf{i}}(z)| = 0$$

as soon as $\Im(z)$ is sufficiently large, hence for all $z \in \mathbb{C}_+$ by holomorphy. Finally, denote by M_n the denumerable vector of holomorphic functions $(m_{(T_{\mathbf{i}})_n})_{\mathbf{i}\in V} \in \mathcal{H}^V$. Since $\mathcal{H}$ is compact, the sequence $(M_n)_{n\geq 0}$ is relatively compact, and since each vector M_n satisfies the partial set of equations (4) corresponding to the truncated tree $(T)_n$, any limit point M_∞ must satisfy the global set of equations (4) corresponding to the full tree T, so M_∞ is nothing but $(m_{T_{\mathbf{i}}})_{\mathbf{i}\in V}$. Therefore, the sequence of vectors $(M_n)_{n\geq 0}$ converges to $M_\infty = (m_{T_{\mathbf{i}}})_{\mathbf{i}\in V}$, and this is exactly (6). $\square$

2.3. *Atomic mass at zero.* Our goal here is to characterize $\mu_T(\{0\})$, the atomic mass at zero of the spectral measure μ_T.

PROPOSITION 6. *If T satisfies assumption* (B), *then the family* $(\mu_{T_{\mathbf{i}}}(\{0\}))_{\mathbf{i}\in V}$ *is the largest solution in* $[0, 1]^V$ *to the system of equations*

$$x_{\mathbf{i}} = \left(1 + \sum_{\mathbf{j}\in D(\mathbf{i})}\left(\sum_{\mathbf{k}\in D(\mathbf{j})} x_{\mathbf{k}}\right)^{-1}\right)^{-1}, \tag{7}$$

with the conventions $1/0 = \infty$ *and* $1/\infty = 0$.

PROOF. Since T is acyclic, (3) ensures that the measures $\mu_{T_{\mathbf{i}}}$, $\mathbf{i} \in V$, are symmetric. Therefore, for all $t > 0$, $\mathbf{i} \in V$

$$m_{T_{\mathbf{i}}}(it) = \int_{\mathbb{R}}\frac{x}{x^2 + t^2}\,d\mu_{T_{\mathbf{i}}}(x) + i\int_{\mathbb{R}}\frac{t}{x^2 + t^2}\,d\mu_{T_{\mathbf{i}}}(x) = i\int_{\mathbb{R}}\frac{t}{x^2 + t^2}\,d\mu_{T_{\mathbf{i}}}(x).$$

Hence, if we define $h_{T_{\mathbf{i}}}(t) := -itm_{T_{\mathbf{i}}}(it) \in [0, 1]$, then by the dominated convergence theorem,

$$h_{T_{\mathbf{i}}}(t) = \int_{\mathbb{R}}\frac{t^2\,d\mu_{T_{\mathbf{i}}}(x)}{x^2 + t^2} \xrightarrow[t\to 0]{} \mu_{T_{\mathbf{i}}}(\{0\}).$$

But, iterating once equation (4), we get

$$h_{T_{\mathbf{i}}}(t) = \left(1 + \sum_{\mathbf{j} \in D(\mathbf{i})} \left(t^2 + \sum_{\mathbf{k} \in D(\mathbf{j})} h_{T_{\mathbf{k}}}(t)\right)^{-1}\right)^{-1}, \tag{8}$$

so that letting $t \to 0$ yields exactly that $(\mu_{T_{\mathbf{i}}}(\{0\}))_{\mathbf{i} \in V}$ must satisfy (7).

Again, when the rooted tree T is finite, this recursion characterizes the family $(\mu_{T_{\mathbf{i}}}(\{0\}))_{\mathbf{i} \in V}$, since it can be computed iteratively from the leaves up to the root. However, when T is infinite, (7) may admit several other solutions. Fortunately, among all of them, $(\mu_{T_{\mathbf{i}}}(\{0\}))_{\mathbf{i} \in V}$ is always the largest. To see why, consider any solution $(x_{\mathbf{i}})_{\mathbf{i} \in T} \in [0, 1]^V$. Fixing $t > 0$, let us show by induction that for all $n \in \mathbb{N}$,

$$\forall \mathbf{i} \in V \qquad x_{\mathbf{i}} \le h_{(T_{\mathbf{i}})_{2n}}(t) := -itm_{(T_{\mathbf{i}})_{2n}}(it). \tag{9}$$

This will conclude our proof since we may then let $n \to \infty$ to obtain $x_{\mathbf{i}} \le h_{T_{\mathbf{i}}}(t)$ by Proposition 5, and let finally $t \to 0$ to reach the desired $x_{\mathbf{i}} \le \mu_{T_{\mathbf{i}}}(\{0\})$. The base case $n = 0$ is trivial because the right-hand equals 1. Now, if (9) holds for some $n \in \mathbb{N}$, then for all $\mathbf{i} \in V$,

$$\begin{aligned} x_{\mathbf{i}} &= \left(1 + \sum_{\mathbf{j} \in D(\mathbf{i})} \left(\sum_{\mathbf{k} \in D(\mathbf{j})} x_{\mathbf{k}}\right)^{-1}\right)^{-1} \\ &\le \left(1 + \sum_{\mathbf{j} \in D(\mathbf{i})} \left(t^2 + \sum_{\mathbf{k} \in D(\mathbf{j})} h_{(T_{\mathbf{k}})_{2n}}(t)\right)^{-1}\right)^{-1} = h_{(T_{\mathbf{i}})_{2n+2}}(t), \end{aligned}$$

where the first equality follows from the fact that $(x_{\mathbf{i}})_{\mathbf{i} \in T}$ satisfies (7), the middle inequality from the induction hypothesis, and the last equality from (8) applied to $(T_{\mathbf{i}})_{2n+2}$. $\square$

2.4. *Galton–Watson trees*. We now apply the above results to Galton–Watson trees. Let F_* be a distribution on $\mathbb{N}$ with finite mean, and let T be a GWT with *degree* distribution F_*, that is, a random locally finite rooted tree obtained by a Galton–Watson branching process where the root has offspring distribution F_*, and all other genitors have offspring distribution F, where

$$\forall k \ge 1 \qquad F(k-1) = kF_*(k) \Big/ \sum_{\ell} \ell F_*(\ell). \tag{10}$$

In the rest of this paper, we will make the following second moment assumption on the distribution $F_* : \sum_k k^2 F_*(k) < \infty$, or equivalently $\sum_k kF(k) < \infty$. It is in fact a sufficient condition for all the previous results to hold almost surely.

PROPOSITION 7. *If F_* has a finite second moment, then T satisfies* (A) *and* (B) *with probability one. In particular, the adjacency operator A is almost surely self-adjoint, and the atomic mass at zero of the spectral measure at the root of T is characterized by the fixed-point equation* (7).

PROOF. Let N denote a generic random variable with law F. For (B), it is well known (and easy to check by a martingale argument) that the size of the nth generation in a GWT with offspring distribution F behaves like $\mathbb{E}^n N$ as $n \to \infty$, in the precise sense that almost surely, $n^{-1} \log|\partial(T)_n| \to \mathbb{E}N$, which is finite by assumption. As far as (A) is concerned now, if T is finite there is nothing to do. Now if T is infinite, we build an exhausting sequence of finite vertex subsets with uniformly bounded boundary degree as follows: the finite first moment assumption on F guarantees the existence of a large enough integer $\kappa \geq 1$ so that

$$\sum_{k \geq \kappa} kF(k) < 1. \tag{11}$$

For each vertex of T, color it in red if it has less than κ children and in blue otherwise. If the root Ø is red, set $S_1 = \{Ø\}$. Otherwise, the connected blue component containing the root is a GWT with average offspring $\sum_{k \geq \kappa} kF(k) < 1$, so it is almost-surely finite, and we define S_1 as the set of its vertices, together with their (red) external boundary vertices. Now for each external boundary vertex $\mathbf{i} \in \partial S_1^+$, we repeat the procedure on the subtree $T_{\mathbf{i}}$, and we define S_2 as the union of S_1 and all the resulting subsets. Iterating this procedure, we obtain an exhaustive sequence of subsets $S_1, S_2, \ldots \subseteq V$ whose boundary degree satisfies by construction $\Delta(\partial S_n) = \kappa$, which is exactly (A). $\square$

Owing to the recursive distributional nature of GWTs, the set of equations (7) defining $\mu_T(\{0\})$ takes the much nicer form of a Recursive distributional equation (RDE), which we now make explicit. We denote $\mathcal{P}(\mathbb{N})$ (resp., $\mathcal{P}([0, 1])$) the space of probability distributions on $\mathbb{N}$ ($[0, 1]$, resp.). Given $F, F' \in \mathcal{P}(\mathbb{N})$ and $\nu \in \mathcal{P}([0, 1])$, we denote by $\Theta_{F,F'}(\nu)$ the distribution of the $[0, 1]$-valued r.v.

$$Y = \frac{1}{1 + \sum_{i=1}^{N} (\sum_{j=1}^{N_i'} X_{ij})^{-1}}, \tag{12}$$

where $N \sim F$, $N_i' \sim F'$ and $X_{ij} \sim \nu$, all of them being independent. With this notation in hand, the previous result implies the following: if F^* has a finite second moment, then $\mu_T(\{0\})$ has distribution $\Theta_{F_*,F}(\nu_0^*)$, where F is given by (10) and ν_0^* is the largest solution to the RDE

$$\nu_0^* = \Theta_{F,F}(\nu_0^*). \tag{13}$$

The remainder of this section is dedicated to solving (13) when F_* has a finite second moment. We will assume that $F_*(0) + F_*(1) < 1$; otherwise $F = \delta_0$ and $\nu_0^* = \delta_1$ is clearly the only solution to (13). We let $\varphi_*(z) = \sum_{n \geq 0} F_*(n) z^n$ be the generating function of F_*. For any $x \in [0, 1]$, we set $\overline{x} = \varphi_*'(1 - x)/\varphi_*'(1)$, and we define

$$M(x) = \varphi_*'(1) x \overline{x} + \varphi_*(1 - x) + \varphi_*(1 - \overline{x}) - 1.$$

Observe that $M'(x) = \varphi_*''(1-x)(\overline{\overline{x}} - x)$, and therefore any $x \in [0,1]$ where M admits a local extremum must satisfy $x = \overline{\overline{x}}$. We will say that M admits a historical record at x if $x = \overline{\overline{x}}$ and $M(x) > M(y)$ for any $0 \le y < x$. Since $[0,1]$ is compact and M is analytic, there are only finitely many such records. In fact, they are in one-to-one correspondence with the solutions to the RDE (13).

THEOREM 8. *If $p_1 < \cdots < p_r$ are the locations of the historical records of M, then the RDE* (13) *admits exactly r solutions*; *moreover, these solutions can be stochastically ordered, say $\nu_1 < \cdots < \nu_r$, and for any $i \in \{1, \dots, r\}$*:

(i) $\nu_i(\{0\}^c) = p_i$;
(ii) $\Theta_{F_*,F}(\nu_i)$ *has mean* $M(p_i)$.

In particular, $\mathbb{E}[\mu_T(\{0\})] = \max_{x\in[0,1]} M(x)$.

It now remains to prove Theorem 8. The space $\mathcal{P}([0,1])$ is naturally equipped with:

- a natural topology, which is that of weak convergence,

$$\mu_n \xrightarrow[n\to\infty]{} \mu \quad \Longleftrightarrow \quad \int \varphi\, d\mu_n \xrightarrow[n\to\infty]{} \int \varphi\, d\mu$$

for any continuous function $\varphi : [0,1] \to \mathbb{R}$;

- a natural order, which is that of stochastic domination,

$$\mu_1 \le \mu_2 \quad \Longleftrightarrow \quad \int \varphi\, d\mu_1 \le \int \varphi\, d\mu_2$$

for any continuous, increasing function $\varphi : [0,1] \to \mathbb{R}$.

The proof is based on two lemmas, the first one being straightforward.

LEMMA 9. *For any $F, F' \in \mathcal{P}(\mathbb{N}) \setminus \{\delta_0\}$, $\Theta_{F,F'}$ is continuous and strictly increasing on $\mathcal{P}([0,1])$.*

LEMMA 10. *For any $\nu \in \mathcal{P}([0,1])$, letting $p = \nu(\{0\}^c)$, we have*:

(i) $\Theta_{F,F}(\nu)(\{0\}^c) = \overline{\overline{p}}$;
(ii) *if $\Theta_{F,F}(\nu) \le \nu$, then the mean of $\Theta_{F_*,F}(\nu)$ is at least $M(p)$*;
(iii) *if $\Theta_{F,F}(\nu) \ge \nu$, then the mean of $\Theta_{F_*,F}(\nu)$ is at most $M(p)$.*

In particular, if ν is a fixed point of $\Theta_{F,F}$, then $p = \overline{\overline{p}}$ and $\Theta_{F_*,F}(\nu)$ has mean $M(p)$.

PROOF. In (12) it is clear that $Y > 0$ if and only if for any $i \in \{1, \dots, N\}$, there exists $j \in \{1, \dots, N_i'\}$ such that $X_{ij} > 0$. Denoting by φ the generating function of F, this rewrites

$$\Theta_{F,F}(\nu)(\{0\}^c) = \varphi\bigl(1 - \varphi\bigl(1 - \nu(\{0\}^c)\bigr)\bigr).$$

But from (10) it follows that $\varphi(\cdot)=\varphi_*'(\cdot)/\varphi_*'(1)$, that is, $\varphi(1-x)=\overline{x}$, hence the first result.

Now let $X\sim\nu$, $N_*\sim F_*$, $N\sim F$, and let $S, S_1,\dots$ have the distribution of the sum of N i.i.d. copies of X, all these variables being independent. Then, $\Theta_{F_*,F}(\nu)$ has mean

$$
\begin{aligned}
\mathbb{E}\biggl[\frac{1}{1+\sum_{i=1}^{N_*}S_i^{-1}}\biggr] &= \mathbb{E}\biggl[\biggl(1-\frac{\sum_{i=1}^{N_*}S_i^{-1}}{1+\sum_{i=1}^{N_*}S_i^{-1}}\biggr)\mathbf{1}_{\{\forall i=1,\dots,N_*,S_i>0\}}\biggr]\\
&= \varphi_*(1-\overline{p})\\
&\quad -\varphi_*'(1)\mathbb{E}\biggl[\frac{S^{-1}}{S^{-1}+1+\sum_{i=1}^{N}S_i^{-1}}\mathbf{1}_{\{S>0,\forall i=1,\dots,\widehat{N}_*,S_i>0\}}\biggr]\\
&= \varphi_*(1-\overline{p})-\varphi_*'(1)\mathbb{E}\biggl[\frac{Y}{Y+S}\mathbf{1}_{\{S>0\}}\biggr],
\end{aligned}
$$

where the second and last lines follow from (10) and $Y\sim\Theta_{F,F}(\nu)$, respectively. Now, for any $s>0$, $x\mapsto\frac{x}{x+s}$ is increasing, and hence, depending on whether $\Theta_{F,F}(\nu)\geq\nu$ or $\Theta_{F,F}(\nu)\leq\nu$, $\Theta_{F_*,F}(\nu)$ has mean at most/least

$$
\begin{aligned}
&\varphi_*(1-\overline{p})-\varphi_*'(1)\mathbb{E}\biggl[\frac{X}{X+S}\mathbf{1}_{\{S>0\}}\biggr]\\
&\quad=\varphi_*(1-\overline{p})-p\varphi_*'(1)\mathbb{E}\biggl[\frac{1}{1+\widehat{N}}\mathbf{1}_{\{\widehat{N}\geq 1\}}\biggr]\qquad\text{with } \widehat{N}=\sum_{i=1}^{N}\mathbf{1}_{\{X_i>0\}}.
\end{aligned}\tag{14}
$$

But using the definition (10) and the well-known identity $(n+1)\binom{n}{d}=(d+1)\binom{n+1}{d+1}$, one can easily check that

$$
\begin{aligned}
&\varphi_*(1-\overline{p})-p\varphi_*'(1)\mathbb{E}\biggl[\frac{1}{1+\widehat{N}}\mathbf{1}_{\{\widehat{N}\geq 1\}}\biggr]\\
&\quad=\varphi_*(1-\overline{p})-p\varphi_*'(1)\sum_{n\geq 1}F(n)\sum_{d=1}^{n}\binom{n}{d}\frac{p^d(1-p)^{n-d}}{d+1}\\
&\quad=M(p). \qquad\square
\end{aligned}
$$

We now have all the ingredients we need to prove Theorem 8.

PROOF OF THEOREM 8. Let $p\in[0,1]$ such that $\overline{\overline{p}}=p$, and define $\nu_0=\text{Bernoulli}(p)$. From Lemma 10 we know that $\Theta_{F,F}(\nu_0)(\{0\}^c)=p$, and since Bernoulli$(p)$ is the largest element of $\mathcal{P}([0,1])$ putting mass p on $\{0\}^c$, we have $\Theta_{F,F}(\nu_0)\leq\nu_0$. Immediately, Lemma 9 guarantees that the limit

$$\nu_\infty=\lim_{k\to\infty}\searrow\Theta_{F,F}^k(\nu_0)$$

exists in $\mathcal{P}([0,1])$ and is a fixed point of $\Theta_{F,F}$. Moreover, by Fatou's lemma, the number $p_\infty = \nu_\infty(\{0\}^c)$ must satisfy $p_\infty \le p$. But then the mean of $\Theta_{F_*,F}(\nu_\infty)$ must be both:

- equal to $M(p_\infty)$ by Lemma 10 with $\nu = \nu_\infty$ and
- at least $M(p)$ since $\forall k \ge 0$, the mean of $\Theta_{F_*,F}(\Theta^k_{F,F}(\mu_0))$ is at least $M(p)$ [Lemma 10 with $\nu = \Theta^k_{F,F}(\nu_0)$].

We have just shown both $M(p) \le M(p_\infty)$ and $p_\infty \le p$. From this, we will now deduce the one-to-one correspondence between historical records of M and fixed points of $\Theta_{F,F}$. We treat each inclusion separately:

- If M admits a historical record at p, then clearly $p_\infty = p$, so ν_∞ is a fixed point satisfying $\nu_\infty(\{0\}^c) = p$.
- Conversely, considering a fixed point ν with $\nu(\{0\}^c) = p$, we want to deduce that M admits a historical record at p. We first claim that ν is the above defined limit ν_∞. Indeed, $\nu \le \text{Bernoulli}(p)$ implies $\nu \le \nu_\infty$ ($\Theta_{F,F}$ is increasing), and in particular $p \le p_\infty$. Therefore, $p = p_\infty$ and $M(p) = M(p_\infty)$. In other words, the two ordered distributions $\Theta_{F_*,F}(\nu) \le \Theta_{F_*,F}(\nu_\infty)$ share the same mean and hence are equal. This ensures $\nu = \nu_\infty$. Now, if $q < p$ is any historical record location, we know from part 1 that

  $$\lambda_\infty = \lim_{k\to\infty} \searrow \Theta^k_{F,F}(\text{Bernoulli}(q))$$

  is a fixed point of $\Theta_{F,F}$ satisfying $\lambda_\infty(\{0\}^c) = q$. But $q < p$, so $\text{Bernoulli}(q) < \text{Bernoulli}(p)$, hence $\lambda_\infty \le \nu_\infty$. Moreover, this limit inequality is strict because $\lambda_\infty(\{0\}^c) = q < p = \nu_\infty(\{0\}^c)$. Consequently, $\Theta_{F_*,F}(\lambda_\infty) < \Theta_{F_*,F}(\nu_\infty)$ and taking expectations, $M(q) < M(p)$. Thus, M admits a historical record at p. $\square$

3. Convergence of the spectral measure.

3.1. *Local convergence of rooted graphs.* In this paragraph, we briefly recall the framework of local convergence introduced by Benjamini and Schramm [7] and Aldous and Steele [3] (see also Aldous and Lyons [2]).

We recall that for integer t, $(G)_t$ is the rooted subgraph spanned by the vertices at distance at most t from the root. We consider the set $\mathcal{G}_*$ of all locally finite, connected rooted graphs, taken up to root-preserving isomorphism. With the terminology of combinatorics, $\mathcal{G}_*$ is the set of rooted unlabeled connected locally finite graphs. We define a metric on $\mathcal{G}_*$ by letting the distance between two rooted graphs G_1 and G_2 be $1/(1+T)$, where T is the supremum of those $t \ge 0$ such that there exists a root-preserving isomorphism from $(G_1)_t$ to $(G_2)_t$. Convergence with respect to this metric is called *local* convergence.

This makes $\mathcal{G}_*$ into a separable and complete metric space (see Section 2 in [2]). In particular, we can endow $\mathcal{G}_*$ with its Borel σ-algebra and speak about weak convergence of random elements in $\mathcal{G}_*$. Specifically, a sequence of probability distributions $\rho_1, \rho_2, \dots$ on $\mathcal{G}_*$ converges weakly to a probability distribution ρ, denoted by $\rho_n \Longrightarrow \rho$, if

$$\int_{\mathcal{G}^*} f\, d\rho_n \xrightarrow[n\to\infty]{} \int_{\mathcal{G}^*} f\, d\rho$$

for all bounded continuous function $f : \mathcal{G}_* \to \mathbb{R}$. This is called the *local weak* convergence.

Let us finally mention three important examples of random graph sequences that converge locally weakly to Galton–Watson trees. The Erdős–Rényi graphs with connectivity c on the vertex set $\{1, \dots, n\}$, rooted at $\varnothing = 1$ converges locally weakly to the GWT with degree distribution Poisson(c). The uniform k-regular ($k \geq 2$) graph on $\{1, \dots, n\}$, rooted at $\varnothing = 1$, converges weakly to the infinite k-regular tree. More generally, if F_* is a degree distribution on $\mathbb{N}$ with finite mean, the random graph-sequence with asymptotic degree distribution F_* converges to the GWT with degree distribution F_*. Note that in the above examples, the vertices are exchangeable and the choice $\varnothing = 1$ is arbitrary: equivalently, we could have chosen $\varnothing$ uniformly at random among all vertices, independently of the edge structure.

3.2. *Continuity of the spectral measure.* Since the elements of $\mathcal{G}^*$ have countably many vertices and are only considered up to isomorphism, we may without loss of generalities embed all vertices into the same, fixed generic vertex set V, say the set of finite words over integers: the root is represented by the empty-word $\varnothing$, and vertices at distance t from the root are represented by word of length t in the usual way. All adjacency operators can thus be viewed as acting on the same Hilbert space $\ell^2(V)$, their action being defined as zero on the orthogonal complement of the subspace spanned by their vertices. Note that this does not affect the spectral measure at the root μ_T.

If (G_n) is a converging sequence in $\mathcal{G}_*$, say to $G \in \mathcal{G}_*$, we may even relabel the vertices in a consistent way so that the root-preserving isomorphisms appearing in the definition of local convergence become identities: for every $t \in \mathbb{N}$, there exists $n_t \in \mathbb{N}$ such that

$$n \geq n_t \quad \Longrightarrow \quad (G_n)_t = (G)_t. \tag{15}$$

Fixing a word $\mathbf{i} \in V$, and setting t equal 1 plus the distance from $\mathbf{i}$ to the root above, we obtain that for all $n \geq n_t$, $\mathbf{i}$ is a vertex of G_n if and only if it is a vertex of G, and in that case its neighbors in G_n are exactly its neighbors in G. In other words, $A_n e_{\mathbf{i}} = A e_{\mathbf{i}}$. By linearity, it follows that any finitely supported vector $\psi : V \to \mathbb{C}$ must satisfy

$$A_n \psi \xrightarrow[n\to\infty]{\ell^2(V)} A\psi,$$

and since those ψ are dense in $\ell^2(V)$, Theorem VIII.25(a) in Reed and Simon [18] guarantees that $A_n \to A$ in the strong resolvent sense, provided of course that $A, A_1, \dots$ are self-adjoint. In particular, this implies the weak convergence of the corresponding spectral measures at the root and the compact convergence of their associated Cauchy–Stieltjes transforms,

$$m_{G_n} \xrightarrow[n\to\infty]{\mathcal{H}} m_G \quad \text{and} \quad \mu_{G_n} \xrightarrow[n\to\infty]{\mathcal{P}(\mathbb{R})} \mu_G.$$

Note that this last statement does not depend anymore on the way $G, G_1, \dots$ have been embedded. We have thus established the following continuity result:

PROPOSITION 11. *Let $G, G_1, G_2, \dots$ be elements of $\mathcal{G}^*$ whose adjacency operators are self-adjoint. Let $\mu_G, \mu_{G_1}, \dots$ denote the associated spectral measures at their root, and $m_G, m_{G_1}, \dots$ the corresponding Cauchy–Stieltjes transforms. If $G_n \xrightarrow[n\to\infty]{\mathcal{G}^*} G$, then*

$$m_{G_n} \xrightarrow[n\to\infty]{\mathcal{H}} m_G \quad \textit{and} \quad \mu_{G_n} \xrightarrow[n\to\infty]{\mathcal{P}(\mathbb{R})} \mu_G.$$

As a consequence, when $G, G_1, G_2, \dots$ are random elements of $\mathcal{G}^*$, the same implication holds with all convergences being replaced by their distributional versions. More precisely, if the law of G_n converges weakly to that of G, then

$$m_{G_n} \xrightarrow[n\to\infty]{\mathcal{P}(\mathcal{H})} m_G \quad \text{and} \quad \mu_{G_n} \xrightarrow[n\to\infty]{\mathcal{P}(\mathcal{P}(\mathbb{R}))} \mu_G.$$

3.3. *Connection with the empirical spectral measure of a finite graph.* In the case of a finite (nonrooted) graph $G_n = (V_n, E_n)$ on n vertices, the adjacency operator A_n is a particularly simple object: it is bounded and self-adjoint, and it has exactly n eigenvalues $\lambda_1(A_n) \geq \cdots \geq \lambda_n(A_n)$ (with multiplicities), all of them being real. Moreover, $\ell^2(V_n) \equiv \mathbb{C}^n$ admits an orthonormal basis of eigenvectors $(b_1, \dots, b_n)$, a priori different from the canonical orthonormal basis $(e_v)_{v\in V_n}$, such that

$$\forall x \in \mathbb{C}^n \qquad A_n x = \sum_{i=1}^n \lambda_i(A_n)\langle x, b_i\rangle b_i.$$

If (G_n, v) denotes the graph G_n when rooted at v, the spectral measure at the root is simply

$$\mu_{(G_n,v)} = \sum_{i=1}^n |\langle b_i, e_v\rangle|^2 \delta_{\lambda_i(A_n)}.$$

In fact $\mu_{(G_n,v)}$ can be interpreted as the local contribution of vertex v to the empirical spectral measure μ_n of G_n. Indeed, the above formula implies

$$\frac{1}{n}\sum_{v\in V_n} \mu_{(G_n,v)} = \frac{1}{n}\sum_{i=1}^n \delta_{\lambda_i(A_n)} = \mu_n. \tag{16}$$

Note that the left-hand side can be reinterpreted as the expectation of $\mu_{(G_n,\varnothing)}$ under a uniform choice of the root $\varnothing$. More generally, if G_n is a random graph on n vertices, we denote by $U(G_n)$ the random element of $\mathcal{G}^*$ obtained by rooting G_n at a uniformly chosen vertex, independently of the random edge-structure. Similarly, we define $U_2(G_n)$ as the random element $((G_n,\varnothing_1),(G_n,\varnothing_2))$ in $\mathcal{G}^*\times\mathcal{G}^*$, where $(\varnothing_1,\varnothing_2)$ is a uniformly chosen pair of vertices. Finally we let μ_n denote the (random) empirical spectral measure of the adjacency matrix of G_n. With this notation, we have the following corollary.

COROLLARY 12. *If $U(G_n)$ converges weakly to a rooted GWT T whose degree distribution F_* has a finite second moment, then*

$$\lim_{n\to\infty}\mathbb{E}\mu_n=\mathbb{E}\mu_T,$$

where μ_T denotes the local spectral measure at the root of T. If moreover $U_2(G_n)$ converges weakly to (T_1,T_2), two independent copies of T, then in probability,

$$\lim_{n\to\infty}\mu_n=\mathbb{E}\mu_T.$$

In the above-mentioned cases of Erdős–Rényi random graphs and random graphs with asymptotic degree distribution F_*, the assumption on $U_2(G_n)$ is easily checked. This corollary implies that the study of the limiting spectral measure of random tree-like graphs boils down to the study of the local spectral measure at the root of the limiting GWT. As we have seen, the latter is fully characterized by a simple RDE involving its Cauchy–Stieltjes transform. Note, however, that this result does not give the full statement of Theorem 1(i); the almost sure convergence will be considered later.

PROOF OF COROLLARY 12. By (16), we may write for any bounded continuous function $f:\mathbb{R}\to\mathbb{R}$,

$$\mathbb{E}\int_{\mathbb{R}} f\,d\mu_n=\frac{1}{n}\sum_{\varnothing\in V_n}\mathbb{E}\int_{\mathbb{R}} f\,d\mu_{(G_n,\varnothing)}\xrightarrow[n\to\infty]{}\mathbb{E}\int_{\mathbb{R}} f\,d\mu_T,$$

where the convergence follows from the weak convergence $U(G_n)\to T$ and the continuity result stated in Proposition 11. This is exactly saying that $\mathbb{E}\mu_n\to\mathbb{E}\mu_T$. If, moreover, $U_2(G_n)$ converges weakly to (T_1,T_2), then by the same argument,

$$\begin{aligned}\mathbb{E}\left(\int_{\mathbb{R}} f\,d\mu_n\right)^2 &= \frac{1}{n^2}\sum_{\varnothing_1\in V_n,\varnothing_2\in V_n}\mathbb{E}\left(\int_{\mathbb{R}} f\,d\mu_{(G_n,\varnothing_1)}\int_{\mathbb{R}} f\,d\mu_{(G_n,\varnothing_2)}\right)\\ &\xrightarrow[n\to\infty]{}\left(\mathbb{E}\int_{\mathbb{R}} f\,d\mu_T\right)^2,\end{aligned}$$

and therefore, the second moment method suffices to conclude that

$$\int_{\mathbb{R}} f\, d\mu_n \xrightarrow[n\to\infty]{P} \mathbb{E}\int_{\mathbb{R}} f\, d\mu_T,$$

which is exactly saying that $\mu_n \to \mathbb{E}\mu_T$ in probability. $\square$

3.4. *Main result*: *Convergence of the rank.* We are now in position to state the main result of this paper. We consider a sequence of finite random graphs $G_1, G_2, \ldots$ converging in distribution (once uniformly rooted) to a GWT whose degree distribution F_* has a finite second moment. As above, $\varphi_*(x) = \sum_k F_*(k)x^k$ denotes the generating function of F_*, and we consider the function

$$M : x \in [0,1] \mapsto \varphi_*'(1)x\overline{x} + \varphi_*(1-x) + \varphi_*(1-\overline{x}) - 1 \qquad \text{where } \overline{x} = \varphi_*'(1-x)/\varphi_*'(1).$$

Recall that $M'(x) = \varphi_*''(1-x)(\overline{\overline{x}} - x)$ so that $M(x)$ is a local extremum if and only if $\overline{\overline{x}} = x$.

THEOREM 13. *Assume that $U_2(G_n)$ converges weakly to (T_1, T_2), two independent copies of a GWT whose degree distribution F_* has a finite second moment. If the first local extremum of M is the global maximum, then in probability,*

$$\lim_{n\to\infty} \frac{1}{n}\operatorname{rank}(A_n) = 1 - \max_{x\in[0,1]} M(x).$$

Moreover, a simple sufficient condition for the assumption on M to hold is that φ_*'' is log-concave.

If the assumption $U_2(G_n) \to (T_1, T_2)$ is replaced by the weaker $U(G_n) \to T$, then we only have convergence of the expected rank.

The log-concavity of φ_*'' is a sufficient condition for the first local extremum of M to be a global maximum. Setting $h : x \mapsto \overline{\overline{x}} - x$, we find

$$\forall x \in (0,1) \qquad h''(x) = \frac{\varphi_*''(1-x)}{\varphi_*'(1)} \frac{\varphi_*''(1-\overline{x})}{\varphi_*'(1)} g(x)$$

with

$$g(x) = \frac{\varphi_*''(1-x)\varphi_*'''(1-\overline{x})}{\varphi_*'(1)\varphi_*''(1-\overline{x})} - \frac{\varphi_*'''(1-x)}{\varphi_*''(1-x)}.$$

Now, if φ_*'' is log-concave, then $x \mapsto \varphi_*'''(x)/\varphi_*''(x)$ is nonincreasing on $(0,1)$, and therefore, g is decreasing (as the difference of a decreasing function and a nondecreasing one). Consequently, h'' can vanish at most once on $(0,1)$, hence h' admits at most two zeros on $[0,1]$, and h at most three. The unique root x_c of $x = \overline{x}$ is always one of them, and if x is another one, then so is $\overline{x}$. Therefore, only two cases are possible:

- Either x_c is the only zero of h; then $h(0) > 0$ and $h(1) < 0$, so M is maximum at x_c,
- or h admits exactly three zeros $x_- < x_c < x_+$; in this case the decreasing function g has to vanish somewhere in $(0, 1)$, so h'' is positive and then negative on $(0, 1)$. Consequently, h is decreasing, then increasing, and then decreasing again. In other words, M is minimum at x_c and maximum at x_-, x_+.

In both cases, the first local extremum of M is its global maximum.

The remaining part of this section is devoted to the proof of Theorem 13. First, recall that $n^{-1}\operatorname{rank}(A_n) = 1 - \mu_n(\{0\})$. From Corollary 12, we have in probability,

$$\limsup_n \mu_n(\{0\}) \leq \mathbb{E}\mu_T(\{0\}).$$

In order to prove Theorem 13, it is thus sufficient to establish that

$$\liminf_n \mathbb{E}\mu_n(\{0\}) \geq \max_{x\in[0,1]} M(x). \tag{17}$$

To do so, we will use the Karp–Sipser leaf removal algorithm, which was introduced in [15] to efficiently build a matching (i.e., a subset of pairwise disjoint edges) on a finite graph.

For our purposes, the leaf removal algorithm on a locally finite graph $G = (V, E)$ can be described as an iterative procedure that constructs two nondecreasing sequences $(\mathcal{A}_t)_{t\geq 0}$ and $(\mathcal{B}_t)_{t\geq 0}$ of subsets of V as follows: we start with

$$\mathcal{A}_0 = \{v \in V : \deg(v; G) = 0\} \quad \text{and} \quad \mathcal{B}_0 = \varnothing.$$

Thus, $\mathcal{A}_0$ is simply the set of isolated vertices in G. Then, at each step $t \in \mathbb{N}$, we let G_t be the subgraph of G spanned by the vertex-set $V_t = V \setminus (\mathcal{A}_t \cup \mathcal{B}_t \cup P_t)$, where $P_0 = \varnothing$. We denote by

$$L_t = \{v \in V_t : \deg(v; G_t) = 1\}$$

the set of its leaves. We also introduce the set of vertices that are adjacent to those leaves,

$$W_t = \{v \in V_t \setminus L_t : \exists u \in L_t, (uv) \in E\}.$$

We add to P_t the set of pairs of adjacent vertices in L_t,

$$P_{t+1} = P_t \cup \{v \in L_t, \exists u \in L_t, (uv) \in E\}.$$

Then we set

$$\mathcal{A}_{t+1} = \mathcal{A}_t \cup \{u \in L_t : \exists v \in W_t, (uv) \in E\} \quad \text{and} \quad \mathcal{B}_{t+1} = \mathcal{B}_t \cup W_t.$$

In words, for any leaf u of G_t whose (unique) neighbor v is not a leaf, we add u to $\mathcal{A}_t$ and v to $\mathcal{B}_t$. Then to obtain G_{t+1}, all nodes in $\mathcal{A}_{t+1} \cup \mathcal{B}_{t+1} \cup P_{t+1}$ are removed from G (note that to obtain G_{t+1}, all leaves from G_t are removed with their adjacent vertices). If L_t becomes empty, we have $(\mathcal{A}_{t+1}, \mathcal{B}_{t+1}) = (\mathcal{A}_t, \mathcal{B}_t)$, and the algorithm stops. Finally, in the case where the graph G is finite, we define

$$\mathrm{LR}_t(G) = |\mathcal{A}_t(G)| - |\mathcal{B}_t(G)|. \tag{18}$$

Note that for any finite graph G, the sequence $(\mathrm{LR}_t(G))_{t\geq 0}$ is nondecreasing. Note also that the leaf removal algorithm is well defined for a (possibly infinite) locally finite graph, but the definition (18) makes sense only for finite graphs. The lemma below states a connection between these numbers and the rank of the adjacency matrix of G. It was first observed in [6], and a proof can be found in [14].

Although we will not need it here, let us make for completeness the following observation, which was the original reason why this algorithm was introduced for finite graphs: each time a vertex v is added to $\mathcal{B}_t$, one may arbitrarily associate it with one of its neighboring leaves $u_v \in \mathcal{A}_t$. Similarly, for every vertex v added to P_t, define u_v as its other neighboring leaf in P_t. The edge-set $\{(vu_v), v \in \mathcal{B}_t \cup P_t\}$ is then a matching of G, and it is contained in at least one maximum matching of G. Since the graph is finite, the algorithm stops at a finite time t^*. The subgraph of G spanned by the vertex-set $V \setminus (\mathcal{A}_{t^*} \cup \mathcal{B}_{t^*} \cup P_{t^*})$ is a graph with minimal degree at least 2 called the core of the graph.

LEMMA 14. *For any finite graph G with adjacency matrix A, and any $t \in \mathbb{N}$,*

$$\dim\ker(A) \geq |\mathrm{LR}_t(G)|.$$

PROOF. Let $u_1 \in L_0(G)$ be a leaf of G and v its unique neighboring vertex. Let $G' = G \setminus \{u_1, v\}$ and $A(G')$ the adjacency matrix of G', we have

$$\dim\ker A(G) = \dim\ker A(G')$$

(see [6]). Now, if $\{u_1, \dots, u_a\} \subset L_0(G)$, is the set of leaves adjacent to v, then $\{u_2, \dots, u_a\}$ are isolated vertices in G'. The vectors $e_{u_2}, \dots, e_{u_a}$ are thus eigenvectors of the kernel of A'. By orthogonal decomposition, we deduce that

$$\dim\ker A(G) = a - 1 + \dim\ker\bigl(A(G \setminus \{v, u_1, \dots, u_a\})\bigr).$$

By linearity, we obtain that for any integer t,

$$\begin{aligned}\dim\ker A(G) &= |\mathcal{A}_t(G)| - |\mathcal{B}_t(G)| + \dim\ker\bigl(A\bigl(G \setminus (\mathcal{A}_t \cup \mathcal{B}_t \cup P_t)\bigr)\bigr)\\ &\geq |\mathcal{A}_t(G)| - |\mathcal{B}_t(G)|.\end{aligned}$$ □

The lower bound (17) will now follow from the following proposition.

PROPOSITION 15. *Let T be a rooted GWT whose degree distribution F_* has a finite mean. Then*

$$\lim_{t\to\infty} \mathbb{P}\bigl(\varnothing \in \mathcal{A}_t(T)\bigr) - \mathbb{P}\bigl(\varnothing \in \mathcal{B}_t(T)\bigr) = M(x_0),$$

where $x_0 \in [0, 1]$ is the location of the first local extremum of M.

PROOF. The argument is close to that appearing in [15], Section 4. For any vertex $\mathbf{i} \neq \emptyset$, we run the leaf removal algorithm on $\tilde{T}_{\mathbf{i}}$ which is the tree $T_{\mathbf{i}}$ with an additional infinite path starting from $\mathbf{i}$. We first compute the corresponding probabilities $\alpha_t = \mathbb{P}(\mathbf{i} \in \mathcal{A}_t(\tilde{T}_{\mathbf{i}}))$ and $\beta_t = \mathbb{P}(\mathbf{i} \in \mathcal{B}_t(\tilde{T}_{\mathbf{i}}))$. For our purpose, adding the infinite path amounts to increase artificially the degree of the root by 1: to be a leaf in $\tilde{T}_{\mathbf{i}}$, the root needs to be isolated in $T_{\mathbf{i}}$. By construction, $\mathbf{i}$ is in $\mathcal{B}_t(\tilde{T}_{\mathbf{i}})$ if and only if one of its children $\mathbf{k}$ is in $\mathcal{A}_t(\tilde{T}_{\mathbf{k}})$. Hence if N denotes the number of children of $\mathbf{i}$, we have

$$\beta_t = \mathbb{E}[1 - (1-\alpha_t)^N] = 1 - \varphi(1-\alpha_t),$$

where φ is the generating function of N with distribution F given by (10). Similarly, $\mathbf{i}$ is in $\mathcal{A}_t(\tilde{T}_{\mathbf{i}})$ if and only if all its children $\mathbf{k}$ are in $\mathcal{B}_{t-1}(\tilde{T}_{\mathbf{k}})$, so that $\alpha_t = \varphi(\beta_{t-1})$. Hence for all $t \geq 1$, we have $\alpha_t = \varphi(1 - \varphi(1-\alpha_{t-1}))$ and $\alpha_0 = 0$. Since $x \mapsto \varphi(1-\varphi(1-x))$ is nondecreasing, α_t converges to α, the smallest fixed point of the equation $x = \varphi(1-\varphi(1-x))$, and β_t converges to $\beta = 1 - \varphi(1-\alpha)$. Note that $\varphi(x) = \varphi_*'(x)/\varphi_*'(1)$, where φ_* is the generating function of F_*. Hence, with the notation of Section 2.4, we have $\beta = 1 - \overline{\alpha}$, $\alpha = \overline{\overline{\alpha}}$. In particular, we get $x_0 = \alpha$.

We now compute $\mathbb{P}(\emptyset \in \mathcal{A}_t(T)) - \mathbb{P}(\emptyset \in \mathcal{B}_t(T))$. Recall that $D(\emptyset)$ is the set of neighbors of the root $\emptyset$. Here are all the possible cases:

- if $\forall \mathbf{i} \in D(\emptyset)$, $\mathbf{i} \in \mathcal{B}_{t-1}(\tilde{T}_{\mathbf{i}})$, then $\emptyset \in \mathcal{A}_t(T)$;
- if there exists $\mathbf{j} \in D(\emptyset) \setminus (\mathcal{B}_{t-1}(\tilde{T}_{\mathbf{j}}) \cup \mathcal{A}_t(\tilde{T}_{\mathbf{j}}))$ and $\forall \mathbf{i} \in D(\emptyset) \setminus \mathbf{j}$, $\mathbf{i} \in \mathcal{B}_{t-1}(\tilde{T}_{\mathbf{i}})$, then $\emptyset \in \mathcal{A}_t(T)$;
- if there exists $\mathbf{i} \neq \mathbf{j} \in D(\emptyset)$ such that $\mathbf{i} \in \mathcal{A}_t(\tilde{T}_{\mathbf{i}})$ and $\mathbf{j} \notin \mathcal{B}_{t-1}(\tilde{T}_{\mathbf{j}})$, then $\emptyset \in \mathcal{B}_t(T)$.

In all other cases, $\emptyset \notin \mathcal{A}_t(T) \cup \mathcal{B}_t(T)$. In summary, we have

$$\begin{aligned}
&\mathbb{P}\bigl(\emptyset \in \mathcal{A}_t(T)\bigr)\\
&\quad = \mathbb{P}\bigl(\forall \mathbf{i} \in D(\emptyset), \mathbf{i} \in \mathcal{B}_{t-1}(\tilde{T}_{\mathbf{i}})\bigr)\\
&\qquad + \mathbb{P}\bigl(\exists \mathbf{j} \in D(\emptyset) \setminus \bigl(\mathcal{B}_{t-1}(\tilde{T}_{\mathbf{j}}) \cup \mathcal{A}_t(\tilde{T}_{\mathbf{j}})\bigr), \forall \mathbf{i} \in D(\emptyset) \setminus \mathbf{j}, \mathbf{i} \in \mathcal{B}_{t-1}(\tilde{T}_{\mathbf{i}})\bigr)\\
&\quad = \varphi_*(\beta_{t-1}) + (1 - \beta_{t-1} - \alpha_t)\varphi_*'(\beta_{t-1}),\\
&\mathbb{P}\bigl(\emptyset \in \mathcal{B}_t(T)\bigr)\\
&\quad = \mathbb{P}\bigl(\exists \mathbf{i} \neq \mathbf{j} \in D(\emptyset), \mathbf{i} \in \mathcal{A}_t(\tilde{T}_{\mathbf{j}}), \mathbf{j} \notin \mathcal{B}_{t-1}(\tilde{T}_{\mathbf{j}})\bigr)\\
&\quad = \mathbb{P}\bigl(\exists \mathbf{i} \in D(\emptyset), \mathbf{i} \in \mathcal{A}_t(\tilde{T}_{\mathbf{i}})\bigr)\\
&\qquad - \mathbb{P}\bigl(\exists \mathbf{i} \in D(\emptyset), \mathbf{i} \in \mathcal{A}_t(\tilde{T}_{\mathbf{i}}), \forall \mathbf{j} \in D(\emptyset) \setminus \mathbf{i}, \mathbf{j} \in \mathcal{B}_{t-1}(\tilde{T}_{\mathbf{j}})\bigr)\\
&\quad = 1 - \varphi_*(1-\alpha_t) - \alpha_t \varphi_*'(\beta_{t-1}).
\end{aligned}$$

Hence,

$$\begin{aligned}\lim_{t\to\infty} \mathbb{P}\big(\emptyset \in \mathcal{A}_t(T)\big) - \mathbb{P}\big(\emptyset \in \mathcal{B}_t(T)\big) &= \varphi_*(\beta) + (1-\beta)\varphi_*'(\beta) + \varphi_*(1-\alpha) - 1 \\ &= M(\alpha) = M(x_0),\end{aligned}$$

where we have used the identities: $\beta = 1 - \overline{\alpha}$, $\varphi_*'(x)/\varphi_*'(1) = \overline{1-x}$ and $\overline{1-\beta} = \alpha$. □

PROOF OF THEOREM 13. As already pointed out, it is sufficient to prove (17). From Lemma 14, for any integer t,

$$\mathbb{E}\mu_n(\{0\}) \geq \frac{1}{n}\mathbb{E}\,\mathrm{LR}_t(G_n) = \mathbb{P}\big(\emptyset \in \mathcal{A}_t(G_n)\big) - \mathbb{P}\big(\emptyset \in \mathcal{B}_t(G_n)\big),$$

where $\emptyset$ is the uniformly drawn root of $U(G_n)$. Note that the events $\{\emptyset \in \mathcal{A}_t(G_n)\}$ and $\{\emptyset \in \mathcal{B}_t(G_n)\}$ belong to the σ-field generated by $(G_n, \emptyset)_{t+1}$. Thus the convergence of $U(G_n)$ implies that for any $t \in \mathbb{N}$,

$$\lim_{n\to\infty} \mathbb{P}\big(\emptyset \in \mathcal{A}_t(G_n)\big) - \mathbb{P}\big(\emptyset \in \mathcal{B}_t(G_n)\big) = \mathbb{P}\big(\emptyset \in \mathcal{A}_t(T)\big) - \mathbb{P}\big(\emptyset \in \mathcal{B}_t(T)\big),$$

where T is a rooted GWT with degree distribution F_* (this is a standard application of the objective method [3]). □

4. Conclusion. As explained in the Introduction, the condition on M in Theorem 13 is restrictive, and the convergence of the rank when this condition is not met (as in the example described in the Introduction) is left open. Without any condition on the function M, our work gives only the following bounds: assume that $U_2(G_n)$ converges weakly to (T_1, T_2), two independent copies of a GWT whose degree distribution F_* has a finite second moment, then in probability,

$$\begin{aligned}1 - \max_{x\in[0,1]} M(x) \leq \liminf_{n\to\infty} \frac{1}{n}\operatorname{rank}(A_n) &\leq \limsup_{n\to\infty} \frac{1}{n}\operatorname{rank}(A_n) \\ &\leq 1 - M(x_0),\end{aligned} \tag{19}$$

where x_0 is the first local extremum of M. For example, if the sequence of graphs converges weakly to a GWT with degree distribution F_* with $F_*(1) = 0$, that is, with no leaf, then $x_0 = 0$ and $M(0) = F_*(0)$ so that the upper bound in (19) is trivial.

Our proof for the upper bound on the rank of A_n relies on the analysis of the leaf removal algorithm on the graph G_n. As explained above, this algorithm when applied to a finite graph produces a matching and a subgraph of minimal degree 2 called the core. It turns out that the RDEs (12) and (13) also appear in the the analysis of the size of maximal matchings on graphs [11]. In particular, if the size of the core is $o(n)$, the leaf removal produces an (almost) maximal matching [with error $o(n)$], and the bounds in (19) match. If the size of the core is not negligible, but the bounds in (19) match (as, e.g., in the case where φ_*'' is log-concave), our result shows that the asymptotic size of the kernel of the core is zero. In [11], it is shown that this case corresponds to the situation where there is an (almost) perfect matching on the core of the graph. However, as soon as $M(x_0) \neq \max_{x\in[0,1]} M(x)$, for any maximal matching, there is a positive fraction of vertices in the core that are not matched [11]. In this latter case, the convergence of the rank is left open.

APPENDIX: PROOF OF THEOREM 1

In the case where F_* is the Poisson(c) distribution, we simply have

$$\forall x \in (0,1) \qquad \varphi(x) = \varphi_*(x) = \exp\bigl(c(x-1)\bigr),$$

whose second derivative is clearly log-concave. We may therefore apply Theorem 13 to the sequence of Erdős–Rényi graphs $(G_n)_{n\in\mathbb{N}}$. See Figure 2 for a plot of the corresponding function.

To complete the proof of Theorem 1, it only remains to improve the convergence in probability into an almost sure convergence. This is performed by a standard exploration procedure of the edges E_n of the graph G_n. For $1 \leq k \leq n$, we define the random variable in $\{0,1\}^k$,

$$X_k = (A_{ik})_{1\leq i\leq k}.$$

By construction, the variables $(X_k)_{1\leq k\leq n}$ are independent random variables. Note also that the upper half of the adjacency matrix A_n is precisely $(X_1,\dots,X_n)$ and we may safely write $A_n = A(X_1,\dots,X_n)$.

For $1 \leq i \leq n$, let $A_i(X_1,\dots,X_n)$ be the principal minor of A obtained by removing ith row and column. If $\lambda_1 \leq \cdots \leq \lambda_n$ and $\lambda_{1,i} \leq \cdots \leq \lambda_{n-1,i}$ denote the eigenvalues of $A(X_1,\dots,X_n)$ and $A_i(X_1,\dots,X_n)$, by the Cauchy interlacing theorem, for all $1 \leq j \leq n-1$,

$$\lambda_j \leq \lambda_{j,i} \leq \lambda_{j+1}.$$

In particular,

$$|\dim\ker A(X_1,\dots,X_n) - \dim\ker A_i(X_1,\dots,X_n)| \leq 1.$$

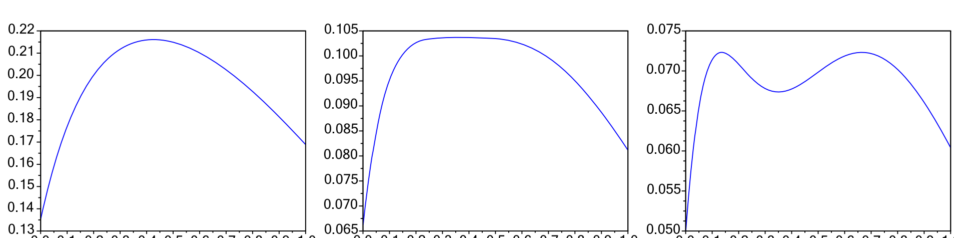

FIG. 2. *From left to right: plot of M for $c = 2$, $c = e$ and $c = 3$.*

We note that $A_i(X_1,\dots,X_n)$ does not depend on X_i. Therefore, for all $(x_j \in \{0,1\}^j)$, $1 \le j \le n$, $x_i' \in \{0,1\}^i$:

$$|\dim\ker A(x_1,\dots,x_{i-1},x_i,x_{i+1},\dots,x_n) \\ - \dim\ker A(x_1,\dots,x_{i-1},x_i',x_{i+1},\dots,x_n)| \le 2.$$

In other words, the function $(x_1,\dots,x_n) \mapsto \dim\ker A(x_1,\dots,x_n)$ is 2-Lipschitz for the Hamming distance. By a standard use of Azuma's martingale difference inequality we get

$$\mathbb{P}\bigl(|\dim\ker A(X_1,\dots,X_n) - \mathbb{E}\dim\ker A(X_1,\dots,X_n)| \ge t\bigr) \le 2\exp\left(\frac{-t^2}{8n}\right).$$

From the Borel–Cantelli lemma, we obtain that almost surely,

$$\lim_n \frac{\dim\ker A(X_1,\dots,X_n) - \mathbb{E}\dim\ker A(X_1,\dots,X_n)}{n} = 0.$$

Since we have already proved that $\mathbb{E}\dim\ker A(X_1,\dots,X_n)/n$ converges to $\max_{x\in[0,1]} M(x)$, we deduce that $\dim\ker A(X_1,\dots,X_n)/n$ convergences a.s. to $\max_{x\in[0,1]} M(x)$.

It remains to deal with the almost sure convergence in Theorem 1(i). We have already proved that μ_n converges in probability to μ. Henceforth $\mathbb{E}\mu_n$ converges to μ. It is thus sufficient to prove that almost surely, for all $t \in \mathbb{R}$, $\mu_n((-\infty,t]) - \mathbb{E}\mu_n((-\infty,t])$ converges to 0. The next lemma is a consequence of Lidskii's inequality. For a proof see Theorem 11.42 in [4].

LEMMA 16 (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 $\mu_B = \frac{1}{n}\sum_{i=1}^n \delta_{\lambda_i(B)}$. Then*

$$\sup_{t\in\mathbb{R}} |\mu_A((-\infty,t]) - \mu_B((-\infty,t])| \le \frac{1}{n}\operatorname{rank}(A-B).$$

Again, we view μ_n as a function of $(X_1,\dots,X_n)$, and write $\mu_n = \mu_{(X_1,\dots,X_n)}$. Note that for all $(x_j \in \{0,1\}^j)$, $1 \le j \le n$, $x_i' \in \{0,1\}^i$, $A(x_1,\dots,x_{i-1},x_i,x_{i+1},\dots,x_n) - A(x_1,\dots,x_{i-1},x_i',x_{i+1},\dots,x_n)$ has only the ith row possibly different from 0, and we get

$$\operatorname{rank}\bigl(A(x_1,\dots,x_{i-1},x_i,x_{i+1},\dots,x_n) - A(x_1,\dots,x_{i-1},x_i',x_{i+1},\dots,x_n)\bigr) \le 2.$$

Therefore from Lemma 16, for any real t,

$$\bigl|\mu_{(x_1,\dots,x_{i-1},x_i,x_{i+1},\dots,x_n)}((-\infty,t]) - \mu_{(x_1,\dots,x_{i-1},x_i',x_{i+1},\dots,x_n)}((-\infty,t])\bigr| \le \frac{1}{n}.$$

Again, Azuma's martingale difference inequality leads to

$$\mathbb{P}\bigl(\bigl|\mu_{(X_1,\dots,X_n)}((-\infty,t]) - \mathbb{E}\mu_{(X_1,\dots,X_n)}((-\infty,t])\bigr| \ge s\bigr) \le 2\exp\left(\frac{-ns^2}{2}\right).$$

We deduce similarly from the Borel–Cantelli lemma that $\mu_n((-\infty,t]) - \mathbb{E}\mu_n((-\infty,t])$ converges a.s. to 0 and the proof of Theorem 1 is complete.

REFERENCES

[1] AIZENMAN, M., SIMS, R. and WARZEL, S. (2006). Stability of the absolutely continuous spectrum of random Schrödinger operators on tree graphs. *Probab. Theory Related Fields* **136** 363–394. MR2257129

[2] ALDOUS, D. and LYONS, R. (2007). Processes on unimodular random networks. *Electron. J. Probab.* **12** 1454–1508. MR2354165

[3] ALDOUS, D. and STEELE, J. M. (2004). The objective method: Probabilistic combinatorial optimization and local weak convergence. In *Probability on Discrete Structures. Encyclopaedia Math. Sci.* **110** 1–72. Springer, Berlin. MR2023650

[4] BAI, Z. D. and SILVERSTEIN, J. W. (2006). *Spectral Analysis of Large Dimensional Random Matrices. Mathematics Monograph Series* **2**. Science Press, Beijing.

[5] BAUER, M. and GOLINELLI, O. (2000). On the kernel of tree incidence matrices. *J. Integer Seq.* **3** Art. 00.1.4, 1 HTML document (electronic). MR1750745

[6] BAUER, M. and GOLINELLI, O. (2001). Exactly solvable model with two conductor-insulator transitions driven by impurities. *Phys. Rev. Lett.* **86** 2621–2624.

[7] BENJAMINI, I. and SCHRAMM, O. (2001). Recurrence of distributional limits of finite planar graphs. *Electron. J. Probab.* **6** 13 pp. (electronic). MR1873300

[8] BHANIDI, S., EVANS, S. N. and SEN, A. (2009). Spectra of large random trees. Preprint.

[9] BOHMAN, T. and FRIEZE, A. (2010). Karp–Sipser on random graphs with a fixed degree sequence.

[10] BORDENAVE, C. and LELARGE, M. (2010). Resolvent of large random graphs. *Random Structures and Algorithms* **37** 332–352.

[11] BORDENAVE, C., LELARGE, M. and SALEZ, J. (2010). Matchings on infinite graphs. Preprint.

[12] COSTELLO, K. P. and VU, V. H. (2008). The rank of random graphs. *Random Structures Algorithms* **33** 269–285. MR2446482

[13] COSTELLO, K. P., TAO, T. and VU, V. (2006). Random symmetric matrices are almost surely nonsingular. *Duke Math. J.* **135** 395–413. MR2267289

[14] CVETKOVIĆ, D. M., DOOB, M. and SACHS, H. (1995). *Spectra of Graphs: Theory and Applications*, 3rd ed. Johann Ambrosius Barth, Heidelberg. MR1324340

[15] KARP, R. and SIPSER, M. (1981). Maximum matchings in sparse random graphs. In *Proc. of the Twenty-Second IEEE Annual Symposium on Foundations of Computer Science* 364–375. IEEE Computer Soc., Los Angeles, CA.

[16] KHORUNZHY, O., SHCHERBINA, M. and VENGEROVSKY, V. (2004). Eigenvalue distribution of large weighted random graphs. *J. Math. Phys.* **45** 1648–1672. MR2043849

[17] KLEIN, A. (1998). Extended states in the Anderson model on the Bethe lattice. *Adv. Math.* **133** 163–184. MR1492789

[18] REED, M. and SIMON, B. (1972). *Methods of Modern Mathematical Physics. I. Functional Analysis*. Academic Press, New York. MR0493419

[19] ZDEBOROVÁ, L. and MÉZARD, M. (2006). The number of matchings in random graphs. *J. Stat. Mech. Theory Exp.* 5 P05003, 24 pp. (electronic). MR2231662

C. BORDENAVE
CNRS UMR5219 AND INSTITUT
DE MATHÉMATIQUES DE TOULOUSE
UNIVERSITÉ TOULOUSE III
FRANCE
E-MAIL: charles.bordenave@math.univ-toulouse.fr

M. LELARGE
J. SALEZ
DÉPARTEMENT D'INFORMATIQUE, PROJET TREC
INRIA-ÉCOLE NORMALE SUPÉRIEURE
FRANCE
E-MAIL: marc.lelarge@ens.fr
justin.salez@ens.fr