

SPECTRA OF LARGE RANDOM TREES

SHANKAR BHAMIDI, STEVEN N. EVANS, AND ARNAB SEN

ABSTRACT. We analyze the eigenvalues of the adjacency matrices of a wide variety of random trees. Using general, broadly applicable arguments based on the interlacing inequalities for the eigenvalues of a principal submatrix of a Hermitian matrix and a suitable notion of local weak convergence for an ensemble of random trees that we call probability fringe convergence, we show that the empirical spectral distributions for each of a number of random tree models converge to a deterministic (model dependent) limit as the number of vertices goes to infinity.

Moreover, the masses assigned by the empirical spectral distributions to individual points also converge in distribution to constants. We conclude for ensembles such as the linear preferential attachment models, random recursive trees, and the uniform random trees that the limiting spectral distribution has a set of atoms that is dense in the real line. We obtain precise asymptotics on the mass assigned to zero by the empirical spectral measures via the connection between the number of zero eigenvalues of the adjacency matrix of a tree and the cardinality of a maximal matching on the tree. In particular, we employ a simplified version of an algorithm due to Karp and Sipser to construct maximal matchings and understand their properties. Moreover, we show that the total weight of a weighted matching is asymptotically equivalent to a constant multiple of the number of vertices when the edge weights are independent, identically distributed, non-negative random variables with finite expected value.

We greatly extend a celebrated result obtained by Schwenk for the uniform random trees by showing that if any ensemble converges in the probability fringe sense and a very mild further condition holds, then, with probability converging to one, the spectrum of a realization is shared by at least one other (non-isomorphic) tree.

For the the linear preferential attachment model with parameter $a > -1$, we show that for any fixed k the k largest eigenvalues jointly converge in distribution to a non-trivial limit when rescaled by $n^{1/2\gamma_a}$, where $\gamma_a = a + 2$ is the Malthusian rate of growth parameter for an associated continuous time branching process.

Date: March 17, 2009.

1991 *Mathematics Subject Classification.* 60C05, 05C80, 90B15.

Key words and phrases. eigenvalue, random matrix, random graph, adjacency matrix, graph Laplacian, interlacing, preferential attachment, recursive random tree, Yule tree, local weak convergence, probability fringe convergence, maximal matching, Karp-Sipser algorithm, branching process, isospectral, exchange property.

SB research supported in part by NSF Grant DMS-0704159, PIMS, and NSERC Canada.

SNE supported in part by NSF grant DMS-0405778.

CONTENTS

1. Introduction	3
2. Some representative random tree models	5
3. Probability fringe convergence of random trees	6
4. Statement of results	8
4.1. Convergence of the spectral distribution	8
4.2. Atoms in the spectral distribution	11
4.3. The proportion of zero eigenvalues and maximal matchings	12
4.4. Isospectrality	13
4.5. Largest eigenvalues and largest degrees	13
5. Convergence of spectral distributions	14
5.1. Interlacing inequalities and some of their consequences	14
5.2. Proof of Theorem 4.1	16
5.3. Proof of Theorem 4.5	17
5.4. Proof of Theorem 4.6	18
6. Maximal matchings and the number of zero eigenvalues	18
6.1. Combinatorial preliminaries	18
6.2. Asymptotics of the number of zero eigenvalues	19
6.3. Maximal weighted matchings: Proof of Theorem 4.9	21
7. Largest eigenvalues: Proof of Theorem 4.12	22
8. An example	27
References	33

1. INTRODUCTION

The study of large random matrices and their eigenvalues is one of the primary themes of current research in probability. It finds applications in such diverse fields as number theory, random partitions, free probability and operator algebras, high-dimensional statistical analysis, nuclear physics, signal processing, wireless communication, quantum percolation, and the operation of search engines. Some recent book length expositions are [Dei00, BI01, Meh04, TV04, Gui09, For09].

The objects of study in this field are usually random real symmetric or complex Hermitian matrices. For example, one of the most popular models is the *Gaussian unitary ensemble (GUE)*, where the matrices are Hermitian, the entries above the diagonal are independent, identically distributed, standard complex Gaussian random variables, and the entries on the diagonal are independent, identically distributed, standard real Gaussian random variables. Much is now known about the asymptotic behavior of objects such as the empirical distribution of the eigenvalues and the behavior of the maximal eigenvalue.

Here we investigate random matrices with substantially greater structure and complexity than the GUE, namely the adjacency matrices of random graphs, although our methods are also applicable to the closely related Laplacian matrices. The recent availability of large amounts of data has led to an explosion in the number of models used to model real-world networks, and dynamically grown models such as various preferential attachment schemes have attracted significant interest from the computer science and mathematical physics community. It is known (see, for example, the monographs [Chu97, GR01, Big93, CRS97, CDS95, CDGT88]) that a surprising diversity of features of a graph are determined, at least in part, by the behavior of the eigenvalues of its adjacency and Laplacian matrices.

We concentrate on the adjacency matrices of various ensembles of random trees. Random trees arise in numerous contexts, ranging from the analysis of database and search algorithms in computer science to models of phylogenies (that is, evolutionary family trees) in biology. Moreover, many of the preferential attachment schemes for networks are also largely random models of growing trees (see, for example, [Bha07] for a survey of some of the more popular schemes). We note that, although trees are rather simple graphs, the analysis of their eigenvalues is still rather challenging, and even in the case of highly symmetric deterministic trees explicit formulae for spectra have only been found recently [Roj08, RR07a, Roj07, RR07b, Roj06a, Roj06b, RS05].

We introduce a general technique based on the concept of *probability fringe convergence* for showing that the *spectral distributions* (that is, empirical distributions of the eigenvalues) of the adjacency matrices of an ensemble of random trees converge in the topology of weak convergence of probability measures on the line to a deterministic limit as the number of vertices goes to infinity, and we show how this technique applies in several natural examples.

The notion of probability fringe convergence is a type of *local weak convergence* for random graphs that involves the convergence in distribution of suitably defined neighborhoods of a vertex picked uniformly from a random graph as the size of the graph goes to infinity. Surveys of this general methodology are [Ald91a, AS04]. Such convergence results for random trees where the limit is described in terms of a continuous-time branching processes go back to [JN84, NJ84]. The first (to

our knowledge) use of such techniques in various general models of preferential attachment is [RTV07]. Such notions are further explored in [Bha07].

The key algebraic element of our proof of convergence of the spectral distributions is the set of interlacing inequalities between the eigenvalues of a Hermitian matrix and the eigenvalues of one of its principal sub-matrices. The interlacing inequalities allow us to break a large tree up into a forest of smaller trees by deleting a small proportion of edges and conclude that the spectral distribution of the tree is close to that of the forest which, in turn, is a convex combination of the spectral distributions of its component sub-trees. If the decomposition into a forest is done appropriately, then the resulting sub-trees are “locally defined” in a sense that allows probability fringe convergence to be brought to bear to show that the spectral distribution of the forest converges.

We note that interlacing has found other applications in algebraic graph theory [BN04, Fio99, Hae95].

Another interesting technical aspect of our work is that the *method of moments*, one of the most commonly used tools in random matrix theory, fails for some natural ensembles because, as we observe in Subsection 4.1, expected values of moments of the spectral distribution go to infinity.

While our method for showing that the spectral distribution converges is quite general, it does not provide any sort of characterization of the limiting distribution. In Section 8 we look at an extremely simple random tree that is obtained by taking the tree consisting of an arc of n points and connecting an independently connecting an edge to each point with equal probability, so that the resulting tree resembles a comb with some of its teeth missing. Our probability fringe convergence methodology does not apply immediately to this ensemble of random trees, but a straightforward modification of it does. We compute the asymptotic moments of the spectral distribution for this ensemble and show that even in this simple case the answer is rather intricate, indicating that we should perhaps not expect simple characterizations of the limit for more complex models.

Quite closely related to our results for the spectral distribution is the recent work [BL07], where similar local weak convergence techniques are combined with Stieltjes transform methods to prove various limiting results for families of random graphs.

We extend our results on the convergence of the spectral distribution in two different directions.

First, we show for any $\gamma \in \mathbb{R}$ that the proportion of eigenvalues that have the value γ converges to a constant under the assumption of probability fringe convergence. Moreover, we give a simple sufficient condition for the limit to be positive and apply this condition to show for several models that the limiting spectral distribution has a set of atoms that is dense in \mathbb{R} . We pay particular attention to the proportion of zero eigenvalues, a quantity of importance in areas such as quantum percolation [BG00, BG01]. It is possible to obtain much more exact information on the limiting proportion because of the connection between the number of zero eigenvalues of the adjacency matrix of a tree and the cardinality of a maximal matching. In particular, we use a simplified version of the Karp-Sipser algorithm [KS81] to construct maximal matchings. We also use our methods to obtain the asymptotic behavior of the total weight of a maximal weighted matching when the edge weights are given by independent, identically distributed, non-negative random variables.

Second, we obtain results on the joint convergence in distribution of the suitably normalized k largest eigenvalues for the preferential attachment tree. These results extend and sharpen those in [CLV03b, CLV03a, FFF05, FFF03], where it was shown that the k largest eigenvalues are asymptotically equivalent to the square roots of the k largest out-degrees. The weak convergence of the suitably rescaled maximum out-degree was obtained in [Mór05] using martingale methods. However it is not clear how to extend this technique to describe the asymptotics for the k largest out-degrees for $k \geq 2$. We prove our more general results using an approach that is essentially completely different.

2. SOME REPRESENTATIVE RANDOM TREE MODELS

An enormous number of random tree models have been developed by computer scientists working on the analysis of algorithms and the mathematical modeling of real world networks: see [Ald91a, Bha07] for a description of some of the more popular models. Although our methods apply quite generally, it will be useful to have the following models in mind when it comes to checking how the hypotheses of our results may be verified in particular instances.

Random recursive tree: This is the simplest model of constructing a rooted tree sequentially via the addition of a new vertex at each stage. Start with a single vertex (the root) at time 1. Label the vertex added at stage n by n , so the tree \mathcal{T}_n that has been constructed by stage n has vertex set $[n] := \{1, 2, \dots, n\}$. Construct the tree at stage $n+1$ by adding an edge from vertex $n+1$ to a vertex chosen uniformly among the vertices $1, 2, \dots, n$. We refer the reader to [SM94] for a survey of some of the properties of the random recursive tree.

Linear preferential attachment tree: This is another sequential construction. As before, start with a single vertex (the root) at time 1. Suppose the tree on n vertices labeled by $[n]$ has been constructed. Think of the edges as directed away from the root and let $D(v, n)$ be the out-degree of vertex $v \in [n]$ at time n (that is, $D(v, n)$ is the number of children of vertex v at time n). Construct a tree on $n+1$ vertices via the addition of an edge between the new vertex $n+1$ and the vertex v in $[n]$ with probability proportional to $D(v, n) + 1 + a$, where $a > -1$ is a parameter of the process. There is an enormous amount of recent literature on this model. We refer the reader to [BR03, Dur07, Bha07] for relevant references.

Uniform random rooted unordered labeled tree: By Cayley's theorem, there are n^{n-1} rooted trees on n labeled vertices (we think of trees as abstract graphs and so we don't consider a particular embedding of a tree in the plane when it comes to deciding whether two trees are "the same" or "different" – this is the import of the adjective "unordered"). Choose one of these trees uniformly at random. Since we are interested only in the structure of the tree, the labeling will be irrelevant.

Random binary tree: There are various models of random rooted binary trees. The one we shall consider is the following sequential construction. Start at time 1 with the three vertex tree consisting of a root and two leaves. At each stage, choose a leaf uniformly and attach two new leaves to it by two new edges.

3. PROBABILITY FRINGE CONVERGENCE OF RANDOM TREES

The key to understanding the asymptotic properties of the spectra of random trees such as those introduced in Section 2 is that they converge “locally” to appropriate locally finite infinite trees. We define the relevant notion of local convergence in this section, and then show how it applies to the models of Section 2.

We first need to say be precise about what we mean by the term *finite rooted tree*. So far, we have talked about trees as particular types of graphs. That is, we have thought of a tree as being described by a finite set of vertices and a finite set of edges that are unordered pairs of vertices. A rooted tree has then been defined as such an object with a particular distinguished vertex that we call the root. This point of view is useful for describing constructions of random trees. However, we will often wish to consider two trees as being the same if they are *isomorphic* in the usual graph-theoretic sense: that is, if they have the same *shape* and only differ by a labeling of the vertices. A tree in this latter sense is thus an isomorphism class of trees thought of as graphs. When we wish to distinguish these two notions we will use standard terminology and speak of *labeled* and *unlabeled* trees, respectively. Continuing in this vein, we take two rooted trees (thought of graphs) to be the same if there is a graph-theoretic isomorphism from one to the other that preserves the root, and we call the corresponding equivalence classes *unlabeled rooted trees*. Even more generally, we may consider unlabeled trees with several distinguished vertices.

Let \mathbb{T} be the countable space of all finite unlabeled rooted trees. Set $\mathbb{T}_* = \mathbb{T} \sqcup \{*\}$, where $*$ is an adjoined point. Equip \mathbb{T} and \mathbb{T}_* with the respective discrete topologies, and equip the Cartesian products \mathbb{T}^∞ and \mathbb{T}_*^∞ with the usual product topologies.

Consider a finite unlabeled rooted tree $\mathbf{t} \in \mathbb{T}$ with root ρ and another distinguished vertex v that is at distance h from the root (v may coincide with ρ , in which case $h = 0$). Let $(v = v_0, v_1, \dots, v_h = \rho)$ denote the unique path from the vertex v to the root. Write t_0 for the subtree rooted at $v_0 = v$ that consists of all vertices for which the path to the root passes through v_0 , and for $1 \leq k \leq h$, write t_k for the subtree rooted at v_k that consists of all vertices for which the path from the root passes through v_k but not through v_{k-1} . Write $\Phi(\mathbf{t}, \cdot)$ for the probability distribution on \mathbb{T}_*^∞ that places mass $(\#\mathbf{t})^{-1}$ at each of the sequences $(t_0, t_1, \dots, t_h, *, *, \dots) \in \mathbb{T}_*^\infty$ as v ranges over the $\#\mathbf{t}$ vertices of \mathbf{t} . It is clear that Φ is a probability kernel from \mathbb{T} to \mathbb{T}_*^∞ .

Definition 3.1. Let $(\mathcal{T}_n)_{n=1}^\infty$ be a sequence of random finite unlabeled rooted trees, and suppose that \mathcal{T} is a \mathbb{T}^∞ -valued random variable. The sequence $(\mathcal{T}_n)_{n=1}^\infty$ converges in the *probability fringe sense* to \mathcal{T} if the sequence $\Phi(\mathcal{T}_n, \cdot)$ of random probability measures on \mathbb{T}_*^∞ converges weakly to the distribution of \mathcal{T} in the topology of weak convergence of probability measures on \mathbb{T}_*^∞ .

Remark 3.2. The definition requires that the empirical distribution of the sub-trees below the various vertices of \mathcal{T}_n converges. However, it demands much more than this: for each $k \geq 1$, the joint empirical distribution of the sub-tree below a vertex and the sub-trees below each of its k most recent ancestors must also converge.

Remark 3.3. Note that any sequence $(t_0, t_1, \dots) \in \mathbb{T}^\infty$ may be thought of as a locally finite unlabeled rooted tree with one end (that is, with a single semi-infinite path) via the identification of the roots of t_k , $k \in \mathbb{Z}^+$, as the successive vertices on

the unique semi-infinite path from the root. We call such trees **sin-trees** (for single infinite path trees).

Remark 3.4. The terminology “probability fringe convergence” is not standard. In the literature, the convergence of the local structure around a **uniformly chosen** vertex of \mathcal{T}_n to the structure around the root for some limiting random **sin-tree** is an instance of what has been termed “local weak convergence” by Aldous, see [AS04]. Our definition is somewhat stronger.

A powerful technique for establishing probability fringe convergence of an ensemble of random trees is to first show that each member of the ensemble can be constructed as the family tree of a suitable stopped continuous-time branching process. (For us, a continuous-time branching process is the sort of object considered in [Jag89]: individuals give birth to a possibly random number of offspring at the arrival times of a point process up to a possibly infinite death time, and those offspring go on to behave as independent copies of their parent.) The next result describes such embeddings for the ensembles of Section 2.

Proposition 3.5. (a) [Random recursive tree] *Consider a continuous time branching process that starts with a single progenitor, individuals live forever, and individuals produce a single offspring at each arrival time of a unit rate Poisson process (this process is sometimes called the Yule process, but the usage of that terminology is not completely consistent in the literature). Write $\mathcal{F}(t) \in \mathbb{T}$ for the corresponding family tree at time $t \geq 0$. Set $T_n := \inf\{t > 0 : \#\mathcal{F}(t) = n\}$. Then $\mathcal{F}(T_n)$ has the same distribution as \mathcal{T}_n , where \mathcal{T}_n is the random recursive tree on n vertices.*

(b) [Linear preferential attachment tree] *Consider a continuous-time branching process that starts with a single progenitor, individuals live forever, and individuals produce a single offspring at each arrival time of a pure birth process with birth rate $k + 1 + a$ in state k . We stress that the state of the pure birth process represents the **number of birth events** for the given individual and, in particular, the pure birth process starts in state 0 for each individual. Write $\mathcal{F}(t) \in \mathbb{T}$ for the corresponding family tree at time $t \geq 0$. Set $T_n := \inf\{t > 0 : \#\mathcal{F}(t) = n\}$. Then, $\mathcal{F}(T_n)$ has the same distribution as \mathcal{T}_n , where \mathcal{T}_n is the linear preferential attachment tree on n vertices with parameter $a > -1$.*

(c) [Uniform random rooted unordered labeled tree] *Let Z_∞ be the complete family tree for a (discrete-time) Galton-Watson branching process with mean 1 Poisson offspring distribution. Note that Z_∞ is finite almost surely. The distribution of Z_∞ conditioned on $\#Z_\infty = n$ is the same as that of \mathcal{T}_n , where \mathcal{T}_n is the object obtained by taking the uniform random rooted unordered tree on n labeled vertices and removing the labeling.*

(d) [Random binary tree] *Consider a continuous-time branching process that starts with a single progenitor, individuals live until a rate 1 exponential time, at which time they produce two offspring (we will refer to this process as the random binary splitting process). Write $\mathcal{F}(t) \in \mathbb{T}$ for the corresponding family tree at time $t \geq 0$. Set $T_n := \inf\{t > 0 : \#\mathcal{F}(t) = n\}$. Then, $\mathcal{F}(T_n)$ has the same distribution as \mathcal{T}_n , where \mathcal{T}_n is the random binary tree on n vertices.*

Proof. Parts (a), (b) and (d) follow from the comparison of the rates of the production of the offspring and the corresponding growth dynamics of the associated tree

\mathcal{T}_n . Part (c) is well-known and follows from randomly ordering the offspring of each individual to obtain an ordered (that is, planar) tree, computing the conditional probability distribution of the resulting rooted ordered tree, randomly labeling the vertices of the rooted ordered tree, and verifying that the randomly labeled tree is uniformly distributed using Cayley's theorem for the number of rooted labeled trees on n vertices (see, for example, [Ald91b]). \square

We now describe briefly the limiting **sin**-trees for the models considered above. Recall that a **sin**-tree can be thought of as an element of \mathbb{T}^∞ . The following proposition follows from well-known results, and we give the appropriate references for each specific construction.

Proposition 3.6. *Each of the four ensembles of Section 2. converges in the probability fringe sense, (as defined in Definition 3.1). The limiting random **sin**-tree for each model is described explicitly in Construction 3.7.*

Construction 3.7. (a) [*Random recursive tree*: [JN84, NJ84, Ald91a]] Let $\mathcal{F}_i(\cdot)$ be independent rate one Yule processes. Let X_0, X_1, \dots be independent rate 1 exponential random variables and put $S_i = \sum_{j=0}^i X_j$. Then, the limiting **sin**-tree has the distribution of $(\mathcal{F}_i(S_i))_{i=0}^\infty$.

(b) [*Linear referential attachment*: [NJ84, JN84, Bha07]] Let $(X_i)_{i=0}^\infty$ be independent exponential random variables, where X_0 has rate $2+a$ and each X_i , $i > 0$ has rate $1+a$. Let $(\mathcal{F}_i)_{i=0}^\infty$ be continuous time branching process that are conditionally independent given $(X_i)_{i=0}^\infty$, with the conditional distribution of \mathcal{F}_i being that in part (b) of Proposition 3.5 with the minor modification that the point process describing the birth times for the root is conditioned to have a birth at time $\sum_{j=0}^i X_j$. Then, the limiting **sin**-tree has the distribution of $(\mathcal{F}_i(\sum_{j=0}^i X_j))_{i=0}^\infty$.

(c) [*Uniform random tree*: [Gri81]] The limiting **sin**-tree has the distribution of an infinite sequence of independent copies of the critical Poisson Galton-Watson tree Z_∞ of part (c) of Proposition 3.5.

(d) [*Random binary tree*: [Ald91a]] Let $(\mathcal{F}_i)_{i=0}^\infty$ be independent random binary splitting processes as in part (d) of Proposition 3.5. Let $(X_i)_{i=0}^\infty$ be independent rate 1 exponential random variables and set $S_i = \sum_{j=0}^i X_j$. Define \mathbb{T} -valued random variables $(\mathcal{U}_i)_{i=0}^\infty$ as follows. Put $\mathcal{U}_0 = \mathcal{F}_0(S_0)$. For $i \geq 1$, \mathcal{U}_i is constructed by attaching a new vertex ρ_i to the root of $\mathcal{F}_i(S_{i-1})$ and re-rooting the resulting tree at ρ_i . Then, the limiting **sin**-tree has the distribution of $(\mathcal{U}_i)_{i=0}^\infty$.

4. STATEMENT OF RESULTS

4.1. Convergence of the spectral distribution.

Theorem 4.1. *Suppose that $(\mathcal{T}_n)_{n=1}^\infty$ is a sequence of random finite unlabeled rooted trees that converges in the probability fringe sense to a **sin**-tree \mathcal{T} . Let F_n denote the spectral distribution of the adjacency matrix of \mathcal{T}_n . Then there exists a (model dependent) deterministic probability distribution F such that F_n converges in distribution to F in the topology of weak convergence of probability measures on \mathbb{R} .*

Simulations of the expected value the spectral distribution for various finite random trees are shown in Figure 1, Figure 2, and Figure 3.

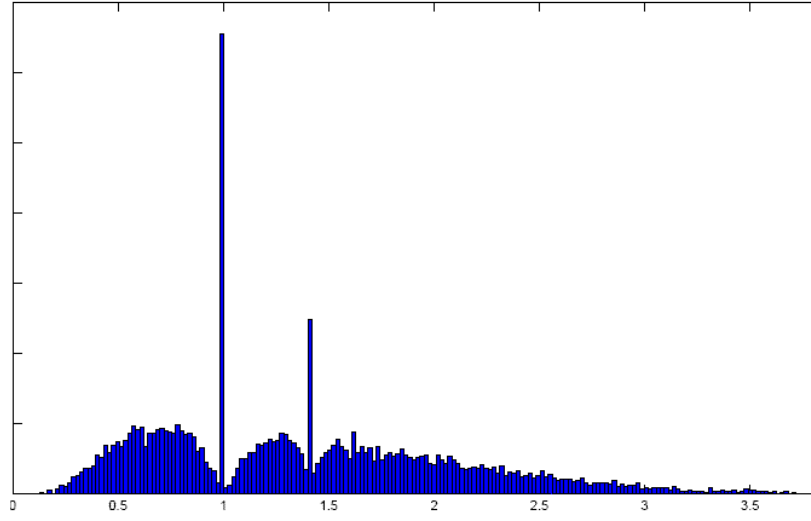


FIGURE 1. Empirical distribution for the positive eigenvalues of the random recursive tree with 200 vertices, averaged over 200 realizations.

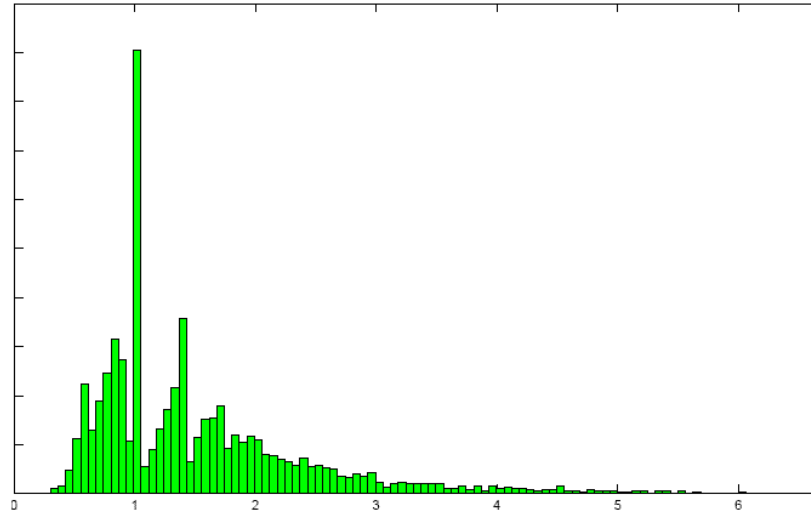


FIGURE 2. Empirical distribution for the positive eigenvalues of the preferential attachment tree with 100 vertices, averaged over 200 realizations.

Remark 4.2. Recall that the *graph Laplacian* of a tree \mathbf{t} with adjacency matrix A is the matrix $A - D$, where D is the diagonal matrix recording the degrees of the vertices of \mathbf{t} (we caution the reader that some authors refer to the negative of this matrix as the Laplacian). The methods we use to establish Theorem 4.1 can also be used to show that if the sequence $(\mathcal{T}_n)_{n=1}^\infty$ converges in the probability fringe

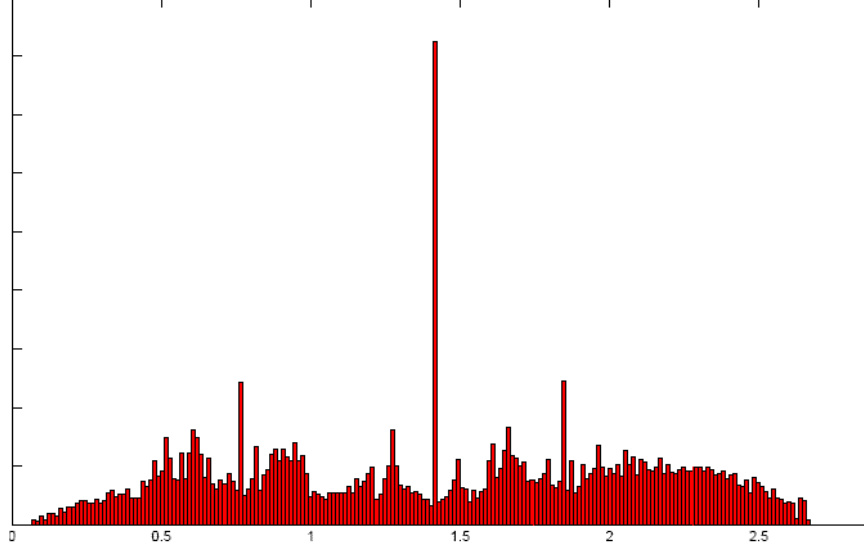


FIGURE 3. Empirical distribution for the positive eigenvalues of the random binary tree with 401 vertices, averaged over 100 realizations.

sense, then the spectral distribution of the Laplacian matrix of \mathcal{T}_n converges in distribution to a deterministic probability distribution on \mathbb{R} .

Remark 4.3. One of the interesting aspects of our proof of Theorem 4.1 is that it does not employ the usual work-horse of random matrix theory, namely the method of moments. A standard method for establishing convergence of the spectral distribution F_n of an $n \times n$ random matrix A_n to a deterministic probability distribution F as $n \rightarrow \infty$ is to show for all $k \geq 1$ that

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[\frac{1}{n} \text{tr}(A^k) \right] \lim_{n \rightarrow \infty} = \mathbb{E} \left[\int_{\mathbb{R}} x^k dF_n(x) \right] = a_k,$$

where the constants $(a_k)_{k=1}^{\infty}$ are the moments of a unique distribution F , and then show that

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[\frac{1}{n^2} \text{tr}(A^k)^2 \right] = a_k^2.$$

The following result shows that this strategy cannot be used for linear preferential attachment model.

Lemma 4.4. *Let A_n be the adjacency matrix of the linear preferential attachment tree \mathcal{T}_n with $a = 0$. Then*

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[\frac{1}{n} \text{tr}(A_n^4) \right] = \infty.$$

Proof. Write $d_n(k)$ for the degree of the k -th vertex in \mathcal{T}_n . Note that $\sum_{k=1}^n d_n(k) = 2(n-1)$ and $d_n(n) = 1$. By construction, at the $(n+1)$ st step the $(n+1)$ st vertex is attached to the k -th vertex with probability $d_n(k)/2(n-1)$. Set $C_n := \sum_{k=1}^n d_n(k)^2$.

We first show that

$$(4.1) \quad \lim_{n \rightarrow \infty} \frac{1}{n-1} \mathbb{E}[C_n] = \infty.$$

Note that

$$\mathbb{E}[C_{n+1}] - \mathbb{E}[C_n] = \mathbb{E}[d_{n+1}(n+1)^2] + \mathbb{E}[(1 + d_n(\mathcal{I}))^2 - d_n(\mathcal{I})^2],$$

where \mathcal{I} is the (random) vertex amongst the vertices $\{1, 2, \dots, n\}$ to which the vertex $(n+1)$ is attached. Hence,

$$\mathbb{E}[C_{n+1}] - \mathbb{E}[C_n] = 1 + \mathbb{E} \left[\sum_{k=1}^n (2d_n(k) + 1) \frac{d_n(k)}{2(n-1)} \right] = 2 + \frac{1}{n-1} \mathbb{E}[C_n].$$

Thus,

$$\mathbb{E}[C_{n+1}] - \mathbb{E}[C_n] \geq 2 \quad \forall n,$$

and hence

$$\liminf_{n \rightarrow \infty} \frac{1}{n-1} \mathbb{E}[C_n] \geq 2.$$

Consequently,

$$\liminf_{n \rightarrow \infty} \mathbb{E}[C_{n+1}] - \mathbb{E}[C_n] \geq 4,$$

and so

$$\liminf_{n \rightarrow \infty} \frac{1}{n-1} \mathbb{E}[C_n] \geq 4.$$

Proceeding in this fashion, we establish that

$$\liminf_{n \rightarrow \infty} \frac{1}{n-1} \mathbb{E}[C_n] \geq 2k$$

for every $k \geq 1$, and hence (4.1) holds.

Note also that

$$\text{tr}(A_{n+1}^4) - \text{tr}(A_n^4) \geq A_{n+1}^4(n+1, n+1) = A_{n+1}^2(\mathcal{I}, \mathcal{I}) \geq A_n^2(\mathcal{I}, \mathcal{I}) = d_n(\mathcal{I}),$$

and so

$$(4.2) \quad \mathbb{E}[\text{tr}(A_{n+1}^4)] - \mathbb{E}[\text{tr}(A_n^4)] \geq \mathbb{E}[d_n(\mathcal{I})] = \frac{1}{2(n-1)} \mathbb{E}[C_n].$$

The result follows immediately from combining (4.1) and (4.2). \square

4.2. Atoms in the spectral distribution.

Theorem 4.5. *Let $(\mathcal{T}_n)_{n=1}^\infty$ a sequence of random finite unlabeled rooted trees that converges in the probability fringe sense to a **sin**-tree. For any $\gamma \in \mathbb{R}$, $F_n(\{\gamma\})$ converges in distribution to a (model dependent) constant c_γ as $n \rightarrow \infty$. If F is the limiting spectral distribution guaranteed by Theorem 4.1, then $F(\{\gamma\}) \geq c_\gamma$.*

Theorem 4.6. *Suppose that $(\mathcal{T}_n)_{n=1}^\infty$ is a sequence of random finite unlabeled rooted trees that converges in the probability fringe sense to a **sin**-tree $\mathcal{T} = (\mathcal{T}^0, \mathcal{T}^1, \dots)$. Consider a forest \mathbf{u} composed of finitely many finite unlabeled rooted trees, and assume that some eigenvalue γ of the adjacency matrix of \mathbf{u} has multiplicity $L > 1$. Write \mathcal{U} for the random forest obtained by deleting the root of \mathcal{T}^0 from \mathcal{T}^0 , and suppose that $\mathbb{P}\{\mathcal{U} = \mathbf{u}\} > 0$. Then, the constant c_γ of Theorem 4.5 is strictly positive and hence γ is an atom of the limiting spectral distribution F of Theorem 4.1.*

Remark 4.7. The condition that the adjacency matrix of the forest \mathbf{u} has an eigenvalue γ with multiplicity greater than 1 certainly holds if γ is an eigenvalue of more than one of the trees that make up \mathbf{u} . In particular, it holds if two or more of the trees that make up \mathbf{u} are equal to some common tree \mathbf{t} , and γ is an eigenvalue of \mathbf{t} . It is clear for the random recursive tree, the linear preferential attachment tree, and the uniform random tree, that if \mathbf{u} is any forest of finite unlabeled rooted trees, then $\mathbb{P}\{\mathcal{U} = \mathbf{u}\} > 0$, and so any number γ that is the eigenvalue of the adjacency matrix of a finite tree will be an atom of the limiting spectral distribution F for these models. From Theorem 7 of [RR07a], the eigenvalues of the adjacency matrix of the rooted tree in which every non-leaf vertex has out-degree d and each leaf is distance $k - 1$ from the root are

$$2\sqrt{d} \cos\left(\frac{\pi\ell}{j+1}\right), \quad j = 1, \dots, k, \ell = 1, \dots, j,$$

with given multiplicities. It follows that the atoms of the limiting spectral distribution F are dense in \mathbb{R} for the random recursive tree, the linear preferential attachment tree, and the uniform random tree. A similar argument shows that the limiting spectral distribution for the random binary tree has a set of atoms that is dense in the interval $[-2\sqrt{2}, 2\sqrt{2}]$. Because we can embed any binary tree into a complete binary tree of suitable height, we see that the limiting spectral measure has this interval as its support.

Remark 4.8. In light of the previous remark, it is natural to inquire whether the limiting spectral distribution F is purely discrete or whether it also has a continuous component. Our methods do not suffice to resolve this question.

4.3. The proportion of zero eigenvalues and maximal matchings. Theorem 4.5 and Theorem 4.6 show that the limiting spectral distribution F will typically have many atoms. However, Theorem 4.6 provides a rather crude lower bounds on the mass of each atom. We give an expression for the limiting proportion of zero eigenvalues in Subsection 6.2 and use this formula to obtain better lower bounds. The key tool we use is the intimate connection we recall in Subsection 6.1 between the number of zero eigenvalues of the adjacency matrix of a tree and *maximal matchings* on the tree – a notion that we now recall briefly.

Suppose that G is a graph with vertex set V and edge set E and for each edge $e \in E$ there is a corresponding *weight* $w(e)$. Recall that a *matching* of G is a subset of $S \subseteq E$ such that no two edges in S share a common vertex. A matching S^* is *maximal* for the system of weights $\{w(e) : e \in E\}$ if $\sum_{e \in S^*} w(e) \geq \sum_{e \in S} w(e)$ for any other matching S . There may be several maximal matchings but the total weight $\sum_{e \in S^*} w(e)$ is, of course, the same for all of them. When no weights are mentioned explicitly, they are assumed to be all 1, and the total weight of a maximal matching in this case is just the maximal possible cardinality of a matching.

Although we only need the case when all the weights are 1 to investigate the proportion of zero eigenvalues, our methods establish the following more general result without much further effort.

Theorem 4.9. *Consider a sequence $(\mathcal{T}_n)_{n=1}^\infty$ of random trees that converge in the probability fringe sense to a random **sin-tree** $\mathcal{T} = (\mathcal{T}^0, \mathcal{T}^1, \dots)$. Write M_n for the number of vertices of \mathcal{T}_n and $M(\mathcal{T}_n)$ for the total weight of a maximal matching on \mathcal{T}_n when the associated system of edge weights is a collection of independent and*

identically distributed \mathbb{R}_+ -valued random variables $X_n(e)$ with a common distribution ν that has finite expected value. Then, $M_n^{-1}M(\mathcal{T}_n)$ converges in distribution to a (model dependent) constant $c_{\mathcal{T},\nu}$ as $n \rightarrow \infty$.

Using their *objective method*, Aldous and Steele [AS04] show that $M_n^{-1}\mathbb{E}[M(\mathcal{T}_n)]$ converges in the case of the ensemble of uniform random trees. Moreover, they characterize the limit in terms of the fixed points of certain distributional identities.

4.4. Isospectrality. A result of Schwenk [Sch73] states that the probability the adjacency matrix of a realization of the uniform random tree has the same spectrum as some other (non-isomorphic) tree converges to one as the number of vertices goes to infinity. Schwenk's method was developed further in [BM93]. The key idea is to first establish that a certain pair of non-isomorphic finite rooted trees \mathbf{t}_1 and \mathbf{t}_2 with the same number of vertices have the following *exchange property*: If \mathbf{t}' is any finite rooted tree with \mathbf{t}_1 as a subtree, then replacing \mathbf{t}_1 by \mathbf{t}_2 produces a tree \mathbf{t}'' with the same adjacency matrix spectrum as that of \mathbf{t}' . If one can then show that a given sequence $(\mathcal{T}_n)_{n=1}^\infty$ is such that $\mathbb{P}\{\mathbf{t}_1 \text{ is a subtree of } \mathcal{T}_n\} \rightarrow 1$ as $n \rightarrow \infty$, then $\mathbb{P}\{\mathcal{T}_n \text{ shares its spectrum with another tree}\} \rightarrow 1$ as $n \rightarrow \infty$. Pairs of trees with exchange property are exhibited in [Sch73, BM93]. Pairs of binary trees (that is, every non-leaf vertex has out-degree 2) with the exchange property are found in [ME06]. The following result is sufficiently obvious that we will not provide a proof. It applies to all four of the models in Section 2, with the pair $\mathbf{t}_1, \mathbf{t}_2$ being, for example, the binary trees in [ME06].

Proposition 4.10. *Consider a sequence $(\mathcal{T}_n)_{n=1}^\infty$ of random finite unlabeled rooted trees that converges in the probability fringe sense to a **sin-tree** $\mathcal{T} = (\mathcal{T}^0, \mathcal{T}^1, \dots)$. Suppose for some pair $\mathbf{t}_1, \mathbf{t}_2 \in \mathbb{T}$ with the exchange property that $\mathbb{P}\{\mathcal{T}^0 = \mathbf{t}_1\} > 0$. Then,*

$$\lim_{n \rightarrow \infty} \mathbb{P}\{\mathcal{T}_n \text{ shares its spectrum with another tree}\} = 1.$$

4.5. Largest eigenvalues and largest degrees. The following result is proved in [FFF05, FFF03, CLV03a] in the case $a = 0$. The proof extends readily to general $a > -1$.

Theorem 4.11. *Let $(\mathcal{T}_n)_{n=1}^\infty$ be the ensemble of linear preferential attachment trees. Fix any $k \geq 1$. Write $\lambda_{n,1} \geq \lambda_{n,2} \geq \dots \geq \lambda_{n,k}$ for the k largest eigenvalues of the adjacency matrix of \mathcal{T}_n and denote by $\Delta_{n,1} \geq \Delta_{n,2} \geq \dots \geq \Delta_{n,k}$ the k largest out-degrees of \mathcal{T}_n . Then, $\lambda_{n,i}/\sqrt{\Delta_{n,i}}$ converges in distribution to 1 as $n \rightarrow \infty$ for $1 \leq i \leq k$.*

We complement this result by establishing the following theorem. Recall that the linear preferential attachment model depends on a parameter $a > -1$. Define the corresponding *Malthusian parameter* by

$$(4.3) \quad \gamma_a := a + 2.$$

Theorem 4.12. *There exist random variables $X_1 \geq X_2 \geq \dots \geq X_k > 0$ that such that*

$$\left(\frac{\Delta_{n,1}}{n^{1/\gamma_a}}, \frac{\Delta_{n,2}}{n^{1/\gamma_a}}, \dots, \frac{\Delta_{n,k}}{n^{1/\gamma_a}} \right)$$

converges in distribution to (X_1, X_2, \dots, X_k) as $n \rightarrow \infty$. Hence,

$$\left(\frac{\lambda_{n,1}}{n^{1/2\gamma_a}}, \frac{\lambda_{n,2}}{n^{1/2\gamma_a}}, \dots, \frac{\lambda_{n,k}}{n^{1/2\gamma_a}} \right)$$

converges in distribution to $(\sqrt{X_1}, \sqrt{X_2}, \dots, \sqrt{X_k})$ as $n \rightarrow \infty$.

5. CONVERGENCE OF SPECTRAL DISTRIBUTIONS

5.1. Interlacing inequalities and some of their consequences. Suppose that A is an $m \times m$ Hermitian matrix and B is an $n \times n$ principal sub-matrix of A for $1 \leq n \leq m$ (that is, B is formed by deleting $m - n$ rows and columns of A with the same indices).

Write $\mu_1 \leq \dots \leq \mu_m$ for the eigenvalues of A and $\nu_1 \leq \dots \leq \nu_n$ for the eigenvalues of B . The interlacing theorem (see, for example, [HJ90]) gives that $\mu_k \leq \nu_k \leq \mu_{k+m-n}$ for $1 \leq k \leq n$.

Write $P := \frac{1}{m} \sum_{i=1}^m \delta_{\mu_i}$ for the spectral distribution of A and $Q := \frac{1}{n} \sum_{i=1}^n \delta_{\nu_i}$ for the spectral distribution of B .

We wish to compare P and Q . To this end, we recall that the Lévy distance between two probability measures σ and τ on \mathbb{R} is given by

$$d(\sigma, \tau) := \inf\{\varepsilon > 0 : S(x - \varepsilon) - \varepsilon < T(x) < S(x + \varepsilon) + \varepsilon, \forall x \in \mathbb{R}\},$$

where S and T are the cumulative distribution functions of σ and τ , respectively – see, for example, [Zol01]. The Lévy distance is a metric that metrizes weak convergence of probability measures on \mathbb{R} , and the space of probability measures on \mathbb{R} is complete with respect to this metric. Note also that if $\sigma \preceq \tau$, where \preceq denotes stochastic domination (that is, $S(x) \geq T(x)$ for all $x \in \mathbb{R}$), and σ' and τ' are two other probability measures with $\sigma \preceq \sigma' \preceq \tau$ and $\sigma \preceq \tau' \preceq \tau$, then $d(\sigma', \tau') \leq d(\sigma, \tau)$. Finally, observe that if X and Y are two random variables on the same probability space that have marginal distributions σ and τ , respectively, then $d(\sigma, \tau) \leq \mathbb{P}\{X \neq Y\}$.

We return to the comparison of P and Q . Define two new probability measures by $P_* = \frac{1}{n} \sum_{i=1}^n \delta_{\mu_i}$ and $P^* = \frac{1}{n} \sum_{i=1}^n \delta_{\mu_{i+m-n}}$. It is clear that $P_* \preceq P \preceq P^*$ and, by the interlacing theorem, that $P_* \preceq Q \preceq P^*$. Thus,

$$d(P, Q) \leq d(P_*, P^*).$$

Suppose that on some probability space we have a random variable I that is uniformly distributed on $\{1, \dots, n\}$. Define a pair of random variables (Z_*, Z^*) by

$$(Z_*, Z^*) = \begin{cases} (\mu_I, \mu_I), & \text{on the event } \{I \geq m - n + 1\}, \\ (\mu_I, \mu_{n+I}), & \text{on the event } \{I \leq m - n\}. \end{cases}$$

Note that Z_* has marginal distribution P_* and Z^* has marginal distribution P^* . By construction, $\mathbb{P}\{Z_* \neq Z^*\} \leq [(m - n) \wedge n]/n$. We have therefore established the following result.

Proposition 5.1. *In the above notation, $d(P, Q) \leq (\frac{m}{n} - 1) \wedge 1$.*

Corollary 5.2. *Consider a sequence $(A_k)_{k=1}^\infty$ of Hermitian matrices, with $(A_k)_{k=1}^\infty$ being $m_k \times m_k$ and having spectral distribution P_k . For each $\varepsilon > 0$, let $(B_k^\varepsilon)_{k=1}^\infty$ be such that B_k^ε is an $n_k^\varepsilon \times n_k^\varepsilon$ principal sub-matrix of A_k with spectral distribution Q_k^ε . Suppose for every $\varepsilon > 0$ that $Q_\infty^\varepsilon = \lim_{k \rightarrow \infty} Q_k^\varepsilon$ exists and $\limsup_{k \rightarrow \infty} m_k/n_k^\varepsilon \leq 1 + \varepsilon$. Then, $P_\infty = \lim_{k \rightarrow \infty} P_k$ exists and is given by $P_\infty = \lim_{\varepsilon \downarrow 0} Q_\infty^\varepsilon$.*

Proof. From Proposition 5.1,

$$\begin{aligned} \limsup_{k, \ell \rightarrow \infty} d(P_k, P_\ell) &\leq \limsup_{k \rightarrow \infty} d(P_k, Q_k^\varepsilon) \\ &\quad + \limsup_{k, \ell \rightarrow \infty} d(Q_k^\varepsilon, Q_\ell^\varepsilon) + \limsup_{\ell \rightarrow \infty} d(Q_\ell^\varepsilon, P_\ell) \\ &\leq 2\varepsilon \end{aligned}$$

for each $\varepsilon > 0$. The sequence $(P_k)_{k=1}^\infty$ is thus Cauchy in the Lévy metric, and hence it converges weakly to a limit P_∞ .

Moreover,

$$d(P_\infty, Q_\infty^\varepsilon) = \lim_{k \rightarrow \infty} d(P_k, Q_k^\varepsilon) \leq \varepsilon,$$

and so $P_\infty = \lim_{\varepsilon \downarrow 0} Q_\infty^\varepsilon$. \square

Proposition 5.3. *In the notation at the beginning of this subsection,*

$$\begin{aligned} &|\#\{1 \leq k \leq m : \mu_k = \gamma\} - \#\{1 \leq k \leq n : \nu_k = \gamma\}| \\ &= |mP(\{\gamma\}) - nQ(\{\gamma\})| \\ &\leq (m - n) \end{aligned}$$

for all $\gamma \in \mathbb{R}$.

Proof. Suppose that $p = \#\{1 \leq k \leq m : \mu_k = \gamma\}$, with $\mu_{a+1} = \dots \mu_{a+p} = \gamma$, and $q = \#\{1 \leq k \leq n : \nu_k = \gamma\}$, with $\nu_{b+1} = \dots \nu_{b+q} = \gamma$. It follows from the interlacing inequalities that $\nu_{a+1} \leq \mu_{a+1}$, provided $a+1 \leq n$, and $\nu_{a+p-(m-n)} \leq \mu_{a+p}$ provided $a+p-(m-n) \geq 1$. Hence, $q \geq p - (m - n)$. Similarly, $\nu_{b+1} \leq \mu_{b+1+(m-n)}$ and $\mu_{b+q} \leq \nu_{b+q}$, so that $p \geq q - (m - n)$. Thus, $|p - q| \leq (m - n)$, as required. \square

Corollary 5.4. *Consider a sequence $(A_k)_{k=1}^\infty$ of Hermitian matrices, with $(A_k)_{k=1}^\infty$ being $m_k \times m_k$ and having spectral distribution P_k . For each $\varepsilon > 0$, let $(B_k^\varepsilon)_{k=1}^\infty$ be such that B_k^ε is an $n_k^\varepsilon \times n_k^\varepsilon$ principal sub-matrix of A_k with spectral distribution Q_k^ε . Suppose for some fixed $\gamma \in \mathbb{R}$ that for every $\varepsilon > 0$ the limit $\lim_{k \rightarrow \infty} Q_k^\varepsilon(\{\gamma\})$ exists and $\limsup_{k \rightarrow \infty} m_k/n_k^\varepsilon \leq 1 + \varepsilon$. Then, $\lim_{k \rightarrow \infty} P_k(\{\gamma\})$ exists and is given by $\lim_{\varepsilon \downarrow 0} \lim_{k \rightarrow \infty} Q_k^\varepsilon(\{\gamma\})$.*

Proof. From Proposition 5.3,

$$|m_k P_k(\{\gamma\}) - n_k^\varepsilon Q_k^\varepsilon(\{\gamma\})| \leq (m_k - n_k^\varepsilon),$$

and so

$$|P_k(\{\gamma\}) - Q_k^\varepsilon(\{\gamma\})| \leq \left(1 - \frac{n_k^\varepsilon}{m_k}\right) + \left(\frac{m_k}{n_k^\varepsilon} - 1\right).$$

An argument using completeness similar to that in the proof of Corollary 5.2 finishes the proof. \square

Corollary 5.5. *Consider a forest \mathbf{u} made up of finitely many finite unlabeled rooted trees, and assume that some eigenvalue γ of the adjacency matrix of \mathbf{u} has multiplicity L . Suppose that A is the adjacency matrix of a finite unlabeled rooted tree \mathbf{t} with m vertices, and suppose that there are K vertices v of \mathbf{t} such that the forest formed by deleting v from the subtree below v produces the forest \mathbf{u} . Then, γ is an eigenvalue of the matrix A with multiplicity at least $KL - m + (m - K) = K(L - 1)$.*

Proof. The proof follows immediately by applying Proposition 5.3 to the matrix B that is the adjacency matrix of the graph obtained by deleting the K designated vertices from \mathbf{t} . The matrix B is block diagonal, and some of its blocks can be collected into K identical larger blocks that each form a copy of the adjacency matrix of the forest \mathbf{u} . It remains to observe that the set of eigenvalues of a block diagonal matrix is the union (including multiplicities) of the sets of eigenvalues of the respective blocks. \square

5.2. Proof of Theorem 4.1. Suppose that the random tree \mathcal{T}_n has M_n vertices and adjacency matrix A_n .

Fix a positive integer K . The construction of several objects in the proof will depend on K , but our notation will not record this.

Denote by W_n the set of vertices v of \mathcal{T}_n such that the sub-tree below v (including v) contains at most K vertices. Put $N_n := \#W_n$. In the notation of Section 3, $N_n/M_n = \Phi(\mathcal{T}_n, \{(t_0, t_1, \dots) : \#t_0 \leq K\})$.

In order to avoid conflicting notation, write the limit **sin**-tree \mathcal{T} as $(\mathcal{T}^0, \mathcal{T}^1, \dots)$. By the assumption of probability fringe convergence, N_n/M_n converges in distribution to the constant $\mathbb{P}\{\#\mathcal{T}^0 \leq K\}$. The latter constant can be made arbitrarily close to 1 by choosing K sufficiently large.

Denote by \mathcal{U}_n the subgraph of \mathcal{T}_n induced by the set of vertices W_n . That is, the graph \mathcal{U}_n has vertex set W_n and two vertices in \mathcal{U}_n are connected by an edge if they are connected by an edge in \mathcal{T}_n . The graph \mathcal{U}_n is a forest.

Write X_{nk} , $1 \leq k \leq K$, for the set of vertices v of \mathcal{T}_n with the following two properties:

- the subtree below v contains k vertices,
- if w is first vertex (other than v) on the path to the root from v , then w is on the path to the root for more than K vertices (that is, the subtree below w contains more than K vertices).

The set of roots of the trees in the forest \mathcal{U}_n is the disjoint union $\bigcup_{k=1}^K X_{nk}$. Put $R_{nk} := \#X_{nk}$, so that $N_n = \sum_{k=1}^K k R_{nk}$. It follows from the assumption of probability fringe convergence that $R_{nk}/M_n = \Phi(\mathcal{T}_n, \{(t_0, t_1, \dots) : \#t_0 = k, \#t_0 + \#t_1 > K\})$ converges in distribution to the constant $p_k := \mathbb{P}\{\#\mathcal{T}^0 = k, \#\mathcal{T}^0 + \#\mathcal{T}^1 > K\}$. Of course, the value of p_k depends on K and may be 0. However,

$$\begin{aligned} \sum_{k=1}^K k p_k &= \lim_{n \rightarrow \infty} \sum_{k=1}^K k \frac{R_{nk}}{M_n} \\ &= \lim_{n \rightarrow \infty} \frac{N_n}{M_n} \\ &= \mathbb{P}\{\#\mathcal{T}^0 \leq K\}. \end{aligned}$$

Moreover, if we write

$$\Xi_{nk} := \frac{M_n}{R_{nk}} \Phi(\mathcal{T}_n, \cdot \cap \{(t_0, t_1, \dots) : \#t_0 = k, \#t_0 + \#t_1 > K\})$$

for the empirical distribution of the subtrees rooted at the vertices in X_{nk} (with some suitable convention when $R_{nk} = 0$), then Ξ_{nk} is concentrated on the finite set of trees with k vertices and $\Xi_{nk}(\{\mathbf{t}\})$ converges in distribution when $p_k > 0$ to the constant

$$\Xi_k(\{\mathbf{t}\}) := \mathbb{P}\{\mathcal{T}^0 = \mathbf{t} \mid \#\mathcal{T}^0 = k, \#\mathcal{T}^0 + \#\mathcal{T}^1 > K\}$$

for each such tree.

Denote by λ_k the distribution of an eigenvalue picked independently and uniformly at random from the k eigenvalues (counting possible multiplicities) of the $k \times k$ adjacency matrix of a k -vertex random tree with distribution Ξ_k . The probability measure λ_k is concentrated on the finite set of real numbers that are the possible eigenvalues of some tree with k vertices.

Write B_n for the adjacency matrix of the forest \mathcal{U}_n . This is a block diagonal matrix with R_{nk} many $k \times k$ blocks for $1 \leq k \leq K$. Recall that the set of eigenvalues of a block diagonal Hermitian matrix is the union of the eigenvalues of the blocks (including multiplicities). Thus, the spectral distribution of B_n converges in distribution to the deterministic probability measure

$$\frac{\sum_{k=1}^K kp_k \lambda_k}{\sum_{k=1}^K kp_k}$$

as $n \rightarrow \infty$.

An application of Corollary 5.2 completes the proof.

Remark 5.6. It is instructive to consider what the various objects that appeared in the proof look like in a simple example. Suppose that \mathcal{T}_n is the deterministic tree with $2^{n+1} - 1$ vertices in which every non-leaf vertex has out-degree 2 and each leaf is distance n from the root. We say that \mathcal{T}_n is a complete binary tree of height n . It is clear that \mathcal{T}_n converges in the probability fringe sense to a random **sin**-tree $(\mathcal{T}^0, \mathcal{T}^1, \dots)$, where \mathcal{T}^0 is a complete binary tree of height H with $\mathbb{P}\{H = h\} = 2^{-h}$, $h = 0, 1, \dots$, and \mathcal{T}^i consists of a root connected by an edge to the root of a complete binary tree of height $H + i - 1$ for $i \geq 1$.

If $2^{n+1} - 1 \geq K$ and ℓ is the unique integer such that $2^{\ell+1} - 1 \leq K < 2^{\ell+2} - 1$, then W_n is the set of vertices of \mathcal{T}_n that are within distance at most ℓ of the leaves. Thus, $N_n = 2^{n-\ell}(2^{\ell+1} - 1)$. Moreover, the set X_{nk} is empty unless $k = 2^{\ell+1}$, in which case X_{nk} is the set of vertices of \mathcal{T}_n that are at distance exactly ℓ from the leaves and $R_{nk} = 2^{n-\ell}$.

The sub-probability distribution $(p_k)_{k=1}^K$ assigns mass $2^{-\ell}$ to $2^{\ell+1} - 1$ and 0 elsewhere, while the probability measure Ξ_k is the point mass at the complete binary tree of height h when k is of the form $2^{h+1} - 1$. The spectral distribution of B_n converges to the spectral distribution of the complete binary tree of height ℓ .

5.3. Proof of Theorem 4.5. The proof is almost identical to that of Theorem 4.1 in Subsection 5.2. Recall from that proof the constant K , the probabilities p_1, \dots, p_K , the probability distributions λ_k , $1 \leq k \leq K$, on \mathbb{R} , and the random adjacency matrix B_n with distribution depending on K and n . Recall also that the probability measure λ_k is concentrated on the finite set of real numbers that are the possible eigenvalues of some tree with k vertices.

It follows from the argument in Subsection 5.2 that the mass assigned by the spectral distribution of B_n to $\gamma \in \mathbb{R}$ converges in distribution to the deterministic probability measure

$$\frac{\sum_{k=1}^K kp_k \lambda_k(\{\gamma\})}{\sum_{k=1}^K kp_k}$$

as $n \rightarrow \infty$.

An application of Corollary 5.4 completes the proof.

5.4. Proof of Theorem 4.6. It follows from Corollary 5.5 that multiplicity of γ as an eigenvalue of the adjacency matrix of \mathcal{T}_n is at least $(L-1)$ times the number of vertices v of \mathcal{T}_n such that the forest formed by deleting v from the subtree below v produces the forest \mathbf{u} . By the assumption of probability fringe convergence, the proportion of eigenvalues of the adjacency matrix of \mathcal{T}_n that have the value γ (that is, $F_n(\{\gamma\})$) satisfies

$$\mathbb{P}\{F_n(\{\gamma\}) > (L-1)\mathbb{P}\{\mathcal{U} = \mathbf{u}\} - \varepsilon\} \rightarrow 1$$

as $n \rightarrow \infty$ for any $\varepsilon > 0$. Moreover, because F_n converges weakly to F in distribution by Theorem 4.1,

$$\mathbb{P}\{F(\{\gamma\}) > F_n(\{\gamma\}) - \varepsilon\} \rightarrow 1$$

as $n \rightarrow \infty$ for any $\varepsilon > 0$. Combining these observations establishes that

$$F(\{\gamma\}) \geq (L-1)\mathbb{P}\{\mathcal{U} = \mathbf{u}\} > 0,$$

as required.

6. MAXIMAL MATCHINGS AND THE NUMBER OF ZERO EIGENVALUES

6.1. Combinatorial preliminaries. The following lemma is standard, but we include the proof for completeness.

Lemma 6.1. *Consider a tree \mathbf{t} with n vertices and adjacency matrix A . Let $\delta(\mathbf{t})$ denote the number of zero eigenvalues A . Then*

$$\delta(\mathbf{t}) = n - 2M(\mathbf{t})$$

where $M(\mathbf{t})$ is the cardinality of a maximal matching of \mathbf{t} .

Proof. It follows from the usual expansion of the determinant that the characteristic polynomial of the adjacency matrix of \mathbf{t} is given by

$$\det(zI - A) = \sum_{k=0}^{\lfloor n/2 \rfloor} (-1)^k N_k(\mathbf{t}) z^{n-2k},$$

where $N_k(\mathbf{t})$ is the number of matchings of \mathbf{t} that contain k edges (see, for example, Addition Result 7b of [Big93]), and the result follows immediately. \square

Our analysis of the cardinality of a maximal matching for a tree relies on the following “greedy” algorithm. It is a simplification of one due to Karp and Sipser [KS81] that is intended to find approximate matchings of more general sparse graphs.

- Input a rooted tree \mathbf{t} with vertices $V(\mathbf{t})$ and edges $E(\mathbf{t})$.
- Initialize $S \leftarrow \emptyset$.
- While $E(\mathbf{t}) \neq \emptyset$ do
 - * Choose a leaf, say x , and let $\{x, y\}$ be the unique edge in \mathbf{t} incident to x .
 - * Set $V(\mathbf{t}) \leftarrow V(\mathbf{t}) \setminus \{x, y\}$, $E(\mathbf{t}) \leftarrow \{e \in E(\mathbf{t}) : e \cap \{x, y\} = \emptyset\}$, and $S \leftarrow S \cup \{\{x, y\}\}$.
- Output the matching S .

Lemma 6.2. *The algorithm produces a maximal matching as its output.*

Proof. Let x be any leaf of the tree, and write $\{x, y\}$ for the unique incident edge. Note that every maximal matching either contains the edge $\{x, y\}$ or an edge of the form $\{y, z\}$ for some vertex $z \neq x$, because otherwise $\{x, y\}$ could be added to a putative maximal matching that contains neither edge to produce a matching with a larger cardinality. Also, note that replacing any edge of the form $\{y, z\}$ with $z \neq x$ that appears in some matching by the edge $\{x, y\}$ results in a collection of edges that is also a matching and has the same cardinality. It follows that the edge $\{x, y\}$ must belong to at least one maximal matching.

The result now follows by induction on the number of edges in the tree. \square

In what follows, given a graph G with vertex set V and edge set E , and a set of vertices $V' \subseteq V$, let $G \setminus V'$ denote the subgraph of G with vertex set $V \setminus V'$ and edge set $\{\{x, y\} \in E : x, y \in V \setminus V'\}$. In the usual graph theory terminology, $G \setminus V'$ is the subgraph of G induced by $V \setminus V'$.

Given a tree \mathbf{t} (which may be infinite), partition the vertex set of \mathbf{t} into disjoint sets $\{C_k(\mathbf{t}) : k \geq 1\}$ in the following recursive manner:

$$\begin{aligned} C_1(\mathbf{t}) &:= L(\mathbf{t}) := \text{leaves of } \mathbf{t}, \\ \mathbf{t}_1 &:= \mathbf{t} \setminus C_1, \\ \text{for } k > 1, \quad C_k(\mathbf{t}) &= L(\mathbf{t}_{k-1}), \\ \mathbf{t}_k &:= \mathbf{t}_{k-1} \setminus C_k(\mathbf{t}). \end{aligned}$$

This definition makes sense since a tree remains connected (and hence a tree) after deleting its leaf set.

It is easy to argue that $C_k(\mathbf{t})$ admits the following alternative description

$$(6.1) \quad C_k(\mathbf{t}) = \{v \in \mathbf{t} : v \text{ is graph distance } k-1 \text{ from } L(\mathbf{t})\}.$$

The following result is immediate from Lemma 6.2.

Lemma 6.3. *The cardinality of the maximum matching of a tree \mathbf{t} is given by*

$$M(\mathbf{t}) = \sum_{k=1}^{\infty} \#C_{2k}(\mathbf{t}).$$

6.2. Asymptotics of the number of zero eigenvalues.

Proposition 6.4. *Consider a sequence $(\mathcal{T}_n)_{n=1}^{\infty}$ of random unlabeled rooted trees, where \mathcal{T}_n has M_n vertices. Suppose that $(\mathcal{T}_n)_{n=1}^{\infty}$ converges in the probability fringe sense to a random **sin-tree** $\mathcal{T} = (\mathcal{T}^0, \mathcal{T}^1, \dots)$ and write R for the root of \mathcal{T}^0 . Then, $M_n^{-1}\mathbb{E}[M(\mathcal{T}_n)]$ converges to $\sum_{k=1}^{\infty} \mathbb{P}\{R \in C_{2k}(\mathcal{T})\}$ as $n \rightarrow \infty$, and hence $F_n(\{0\})$ converges in distribution to $\sum_{k=1}^{\infty} (-1)^{k-1} \mathbb{P}\{R \in C_k(\mathcal{T})\}$ as $n \rightarrow \infty$.*

Proof. The claim regarding the convergence of $F_n(\{0\})$ follows from Theorem 4.5, Lemma 6.1, and Lemma 6.3 once we have established the claim regarding convergence of $M_n^{-1}\mathbb{E}[M(\mathcal{T}_n)]$.

Given a tree $\mathbf{t} \in \mathbb{T}$ with root ρ and a vertex $v \in \mathbf{t}$, write $\mathcal{N}_k(v)$ for the subtree of \mathbf{t} induced by vertices that are at graph distance at most k from v . It follows from (6.1) that whether or not a vertex v of \mathbf{t} belongs to the set $C_k(\mathbf{t})$ can be determined by examining the neighborhood $\mathcal{N}_k(v)$. Observe also that $(t_0, t_1, \dots, t_h, *, *, \dots) \in \mathbb{T}_*^{\infty}$ is the decomposition of \mathbf{t} relative to ρ and v , then $\mathcal{N}_k(v)$ can be reconstructed from $(t_0, t_1, \dots, t_{k \wedge h})$.

Note that if V is a vertex chosen uniformly at random from \mathcal{T}_n then $M_n^{-1}\mathbb{E}[\#C_k(\mathcal{T}_n)] = \mathbb{P}\{V \in C_k(\mathcal{T}_n)\}$. Hence,

$$\begin{aligned} M_n^{-1} \sum_{k=m+1}^{\infty} \mathbb{E}[\#C_{2k}(\mathcal{T}_n)] &= \mathbb{P}\left\{V \in \bigcup_{k=m+1}^{\infty} C_{2k}(\mathcal{T}_n)\right\} \\ &\leq 1 - \mathbb{P}\{\exists \text{ a leaf of } \mathcal{T}_n \text{ within distance } 2m \text{ of } V\} \\ &\leq 1 - \mathbb{P}\{\exists \text{ a leaf of } \mathcal{T}_n \text{ below } V \text{ and within distance } 2m \text{ of } V\} \\ &\rightarrow 1 - \mathbb{P}\{\exists \text{ a leaf of } \mathcal{T}^0 \text{ within distance } 2m \text{ of } R\} \end{aligned}$$

as $n \rightarrow \infty$, by the assumption of probability fringe convergence. The last term clearly converges to 0 as $m \rightarrow \infty$. A similar argument shows that

$$\lim_{m \rightarrow \infty} \sum_{k=m+1}^{\infty} \mathbb{P}\{R \in C_{2k}(\mathcal{T})\} = 0.$$

Finally, it follows from the assumption of probability fringe convergence and our observations above about membership of $C_{2k}(\mathcal{T})$ being locally determined, that $\mathbb{P}\{V \in C_{2k}(\mathcal{T}_n)\}$ converges to $\mathbb{P}\{R \in C_{2k}(\mathcal{T})\}$ as $n \rightarrow \infty$ for each $k \leq m$. \square

Remark 6.5. It follows from Lemma 6.1 that $M_n^{-1}M(\mathcal{T}_n)$ converges in distribution to $\sum_{k=1}^{\infty} \mathbb{P}\{R \in C_{2k}(\mathcal{T})\}$ as $n \rightarrow \infty$ under the assumptions of Proposition 6.4.

Remark 6.6. Write $I(\mathcal{T}_n)$ for the cardinality of a maximal independent set for \mathcal{T}_n . By König's theorem [BM76], for a general bipartite graph the cardinality of a maximal matching is equal to the cardinality of a minimal vertex cover. On the other hand, complementation of a minimal vertex cover in any graph always yields a maximal independent set. Thus, $I(\mathcal{T}_n) = M_n - M(\mathcal{T}_n)$ in our case. Consequently, $M_n^{-1}I(\mathcal{T}_n)$ converges in distribution to $\sum_{k=1}^{\infty} \mathbb{P}\{R \in C_{2k-1}(\mathcal{T})\}$ as $n \rightarrow \infty$.

Although Proposition 6.4 gives an exact expression for the limiting proportion of zero eigenvalues, the quantities in the alternating sum are difficult to compute. In order to show that the limit is positive, it is convenient to note that from Lemma 6.1 and Lemma 6.3 we get the lower bound

$$\begin{aligned} \delta(\mathbf{t}) &= \sum_{k=1}^{\infty} (-1)^{k-1} \#C_k(\mathbf{t}) \\ &\geq 2\#C_1(\mathbf{t}) + 2\#C_3(\mathbf{t}) - \sum_{k=1}^{\infty} \#C_k(\mathbf{t}) \\ &= 2\#C_1(\mathbf{t}) + 2\#C_3(\mathbf{t}) - n \end{aligned}$$

on the number of zero eigenvalues of the tree \mathbf{t} with n vertices.

For the ensemble of linear preferential attachment trees with $a = 0$, it is quite well known, see [Dur07], that the proportion of vertices with degree d converges in distribution to $p_d = 4/d(d+1)(d+2)$. Specializing to $d = 1$, we see that $n^{-1}\#C_1(\mathcal{T}_n)$ converges in distribution to $2/3$. Thus

$$\liminf_{n \rightarrow \infty} \mathbb{P}\{F_n(\{0\}) \leq 1/3 - \varepsilon\} \rightarrow 0$$

as $n \rightarrow \infty$ for any $\varepsilon > 0$. In particular, $F(\{0\}) \geq 1/3$ for the ensemble of linear preferential attachment trees with $a = 0$.

On the other hand, for the ensemble of random recursive trees, it is known, see e.g. [Ald91a] that $p_d = (1/2)^d$, and hence $n^{-1}\#C_1(\mathcal{T}_n)$ converges to $1/2$ in distribution as $n \rightarrow \infty$. So, unlike the preferential attachment ensemble, we cannot ignore the more complicated term $2\#C_3(\mathcal{T}_n)$ here. However, from the description of the limiting random **sin**-tree for this model, it is clear that $n^{-1}\#C_3(T)$ converges to a strictly positive constant as $n \rightarrow \infty$.

6.3. Maximal weighted matchings: Proof of Theorem 4.9. We will use the same construction as we used in the proof of Theorem 4.1 in Subsection 5.2.

Recall that for a fixed positive integer K this construction produced for each n a set of vertices W_n of \mathcal{T}_n with cardinality N_n such that N_n/M_n , where M_n is the number of vertices of \mathcal{T}_n , converged in distribution to $\mathbb{P}\{\#T^0 \leq K\}$ – a constant that can be made arbitrarily close to 1 by choosing K sufficiently large.

The subgraph of \mathcal{T}_n induced by W_n was the forest \mathcal{U}_n rooted at the points $\bigcup_{k=1}^K X_{nk}$ and $\#X_{nk}/M_n$ converged in distribution to the constant $p_k := \mathbb{P}\{\#T^0 = k, \#T^0 + \#T^1 > K\}$.

Moreover, the random probability measure Ξ_{nk} given by the empirical distribution of the subtrees rooted at the vertices in X_{nk} was concentrated on the finite set of trees with k vertices and $\Xi_{nk}(\{\mathbf{t}\})$ converged in distribution when $p_k > 0$ to the constant

$$\Xi_k(\{\mathbf{t}\}) := \mathbb{P}\{T^0 = \mathbf{t} \mid \#T^0 = k, \#T^0 + \#T^1 > K\}$$

for each such tree \mathbf{t} .

Write $M(\mathcal{T}_n)$ (respectively, $M(\mathcal{U}_n)$) for the total weight of a maximal matching on \mathcal{T}_n (respectively, \mathcal{U}_n) for the independent, identically distributed edge weights $X_n(e)$, where e ranges over the edges of \mathcal{T}_n .

Note that a maximal matching on \mathcal{U}_n is obtained by separately constructing maximal matchings on each component subtree of \mathcal{U}_n . It follows from Lemma 6.7 below that $M_n^{-1}M(\mathcal{U}_n)$ converges in distribution to

$$\sum_{k=1}^K p_k \sum_{\mathbf{t}: \#\mathbf{t}=k} \Xi_k(\{\mathbf{t}\})\mu(\mathbf{t}),$$

where $\mu(\mathbf{t})$ is the expected value of the total weight of a maximal matching on \mathbf{t} when the weights of the edges are independent and identically distributed with common distribution ν .

Observe that any matching on \mathcal{U}_n is also a matching on \mathcal{T}_n and that the restriction of any matching on \mathcal{T}_n to \mathcal{U}_n is a matching on \mathcal{U}_n . Thus,

$$M(\mathcal{U}_n) \leq M(\mathcal{T}_n) \leq M(\mathcal{U}_n) + \sum_{e \in E(\mathcal{T}_n) \setminus E(\mathcal{U}_n)} X_n(e),$$

where $E(\mathcal{T}_n)$ (respectively, $E(\mathcal{U}_n)$) is the set of edges of \mathcal{T}_n (respectively, \mathcal{U}_n).

There is an element of $E(\mathcal{T}_n) \setminus E(\mathcal{U}_n)$ for each vertex of \mathcal{T}_n other than the root that is not a vertex of \mathcal{U}_n and one for each root of a subtree in the forest \mathcal{U}_n . Thus, writing μ for the common expected value of the edge weights,

$$\mathbb{E} \left[M_n^{-1} \sum_{e \in E(\mathcal{T}_n) \setminus E(\mathcal{U}_n)} X_n(e) \mid \mathcal{T}_n \right] = M_n^{-1} \left[(M_n - N_n - 1)_+ + \sum_{k=1}^K \#X_{nk} \right] \mu.$$

Note from above that $1 - M_n^{-1}N_n$ converges in distribution to the constant $\mathbb{P}\{\#T^0 > K\}$ and $M_n^{-1} \sum_{k=1}^K \#X_{nk}$ converges in distribution to the constant

$\sum_{k=1}^K p_k = \mathbb{P}\{\#\mathcal{T}^0 \leq K, \#\mathcal{T}^0 + \#\mathcal{T}^1 > K\}$ as $n \rightarrow \infty$. Both of these constants converge to 0 as $K \rightarrow \infty$. It follows that

$$\lim_{K \rightarrow \infty} \lim_{n \rightarrow \infty} \mathbb{P} \left\{ M_n^{-1} \sum_{e \in E(\mathcal{T}_n) \setminus E(\mathcal{U}_n)} X_n(e) > \varepsilon \right\} = 0$$

for all $\varepsilon > 0$.

Therefore, $M_n^{-1}M(\mathcal{T}_n)$ converges in distribution as $n \rightarrow \infty$ to the constant

$$\lim_{K \rightarrow \infty} \sum_{k=1}^K p_k \sum_{\mathbf{t}: \#\mathbf{t}=k} \Xi_k(\{\mathbf{t}\})\mu(\mathbf{t}),$$

where we stress that p_k and Ξ_k depend on K , even though this is not indicated by our notation.

The following lemma, which we used above, is a straightforward consequence of the strong law of large numbers.

Lemma 6.7. *For $i = 1, 2, \dots$ let L^i be a positive integer-valued random variable and $\theta_1^i, \dots, \theta_{L^i}^i$ be random variables taking values in a finite set Θ . Suppose that as $i \rightarrow \infty$ the random variable L^i converges in distribution to ∞ and for each $\theta \in \Theta$ the random variable*

$$\frac{\#\{1 \leq j \leq L^i : \theta_j^i = \theta\}}{L^i}$$

converges in distribution to a constant $\pi(\theta)$. Let $\xi_1^i, \dots, \xi_{L^i}^i$ be \mathbb{R}_+ -valued random variables that are conditionally independent given $\theta_1^i, \dots, \theta_{L^i}^i$, and such that

$$\mathbb{P}\{\xi_j^i \in A \mid \theta_1^i, \dots, \theta_{L^i}^i\} = \Pi(\theta_j^i; A)$$

for some collection of Borel probability measures $(\Pi(\theta; \cdot))_{\theta \in \Theta}$. Suppose that

$$v(\theta) := \int_{\mathbb{R}_+} x \Pi(\theta; dx) < \infty$$

for all $\theta \in \Theta$. Then,

$$\frac{\sum_{j=1}^{L^i} \xi_j^i}{L^i}$$

converges in distribution to

$$\sum_{\theta \in \Theta} \pi(\theta) v(\theta)$$

as $i \rightarrow \infty$.

7. LARGEST EIGENVALUES: PROOF OF THEOREM 4.12

We first recall from Proposition 3.5(b) how \mathcal{T}_n , the linear preferential attachment tree on n vertices with parameter $a > -1$, can be constructed from a particular continuous-time branching process.

Denote by $N_a = (N_a(t))_{t \geq 0}$ a pure birth process that starts with a single progenitor and when there have been k births a new birth occurs at rate $k + 1 + a$. Recall that $\mathcal{F}(t) \in \mathbb{T}$ is the family tree at time $t \geq 0$ of the continuous-time branching process in which the birth process of each individual is a copy of N_a . Then, \mathcal{T}_n has the same distribution as $\mathcal{F}(T_n)$, where $T_n := \inf\{t > 0 : \#\mathcal{F}(t) = n\}$.

We now record some useful facts about the birth process N_a . Recall the Malthusian rate of growth parameter $\gamma_a := a + 2$.

Lemma 7.1. (a) For any fixed time $t \geq 0$, the random variable $\mathbb{P}\{N_0(t) = k\} = (1 - e^{-t})^k e^{-t}$, $k = 0, 1, \dots$. That is, $N_0(t)$ is distributed as the number of failures until the first success in a sequence of independent Bernoulli trials with common success probability e^{-t} .

(b) For $a > -1$, set $A := \lceil a + 1 \rceil$. Then,

$$\mathbb{P}\{N_a(t) > Ke^t\} \leq Ae^{-\frac{K}{A}}$$

for all $K > 0$ and $t \geq 0$.

Proof. For part(a), note that $N_0 + 1$ is a Yule process – the birth rate in state ℓ is ℓ – and the claimed distribution is well-known.

To prove part (b), suppose that $M = (M(t))_{t \geq 0}$ is a Yule process started in state A (that is, M is pure birth process and the birth rate in state ℓ is ℓ). Then, $(M(t) - A)_{t \geq 0}$ is a pure birth process that starts in state 0 and has birth rate $\ell + A \geq \ell + 1 + a$ in state ℓ . It is therefore possible to couple M and N_a in such a way that $N_a(t) \leq M(t) - A$ for all $t \geq 0$. Observe that M has the same distribution as $\sum_{i=1}^A (N_0^i + 1)$, and so $M - A$ has the same distribution as $\sum_{i=1}^A N_0^i$. We could prove (b) using the fact that $M(t)$ is distributed as the number of trials before the A^{th} success in a sequence of independent Bernoulli trials with common success probability e^{-t} , but it is more straightforward to use a simple union bound.

Observe that from part(a) and the inequality $1 - x \leq \exp(-x)$ that, for any $C \geq 0$,

$$\begin{aligned} \mathbb{P}\{N_0(t) > Ce^t\} &= (1 - e^{-t})^{\lfloor Ce^t \rfloor + 1} \\ &\leq (\exp(-e^{-t}))^{\lfloor Ce^t \rfloor + 1} \\ &= \exp(-e^{-t}(\lfloor Ce^t \rfloor + 1)) \\ &\leq e^{-C}, \end{aligned}$$

and hence,

$$\begin{aligned} \mathbb{P}\{N_a(t) > Ke^t\} &\leq \sum_{i=1}^A \mathbb{P}\{N_0^i(t) > \frac{K}{A}e^t\} \\ &\leq Ae^{-\frac{K}{A}}. \end{aligned}$$

□

Theorem 7.2. (a) There exists a random variable $Z_a > 0$ such that

$$\lim_{t \rightarrow \infty} \frac{\#\mathcal{F}(t)}{e^{\gamma_a t}} = Z_a \quad \text{almost surely.}$$

(b) There exists a constant C such that $\mathbb{E}[\#\mathcal{F}(t)] \leq Ce^{\gamma_a t}$.

(c) For the random variable Z_a of part (a),

$$\lim_{n \rightarrow \infty} T_n - \frac{1}{\gamma_a} \log n = -\log Z_a \quad \text{almost surely.}$$

(d) There exists a random variable $W_a > 0$ such that

$$\lim_{t \rightarrow \infty} \frac{N_a(t)}{e^t} = W_a \quad \text{almost surely.}$$

Proof. Parts (a) and (b) (in a more general context) can be found in [Bha07], so we shall not give the proof here. They essentially follow from the general theory of continuous time branching processes developed by Jagers and Nerman.

Part (c) follows immediately from part (a) and the relationship $\#\mathcal{F}(T_n) = n$.

Turning to part (d), note that

$$N_a(t) - \int_0^t (N_a(s) + 1 + a) dt, \quad t \geq 0,$$

is a local martingale with bounded variation. Stochastic calculus shows that the process $(e^{-t} \cdot (N_a(t) + 1 + a))_{t \geq 0}$ is also a local martingale with bounded variation. The fact that the latter process is bounded in \mathbb{L}^2 and hence, in particular, a true martingale follows from Lemma 7.1(b).

It follows from the martingale convergence theorem that $e^{-t}N_a(t)$ converges almost surely and in \mathbb{L}^2 to a random variable W_a .

It remains to show that W_a is strictly positive almost surely. Consider first the case $a \geq 0$. From a comparison of branching rates similar to that in the proof of Lemma 7.1, it is possible to couple N_a and N_0 so that $N_0(t) \leq N_a(t)$ for all $t \geq 0$. Note that W_0 has an exponential distribution with mean 1, and so W_a is certainly almost surely positive.

Consider now the case $-1 < a < 0$. Let \tilde{N}_a be N_a started in the initial state 1 rather than 0, and put $\hat{N}_a = \tilde{N}_a - 1$. Another comparison of branching rates shows that it is possible to couple \hat{N}_a and N_0 so that $N_0(t) \leq \hat{N}_a(t)$ for all $t \geq 0$. Thus, $\lim_{t \rightarrow \infty} e^{-t}N_a(t)$ is stochastically greater than the strictly positive random variable $e^{-\tau}W_0$, where the random variable τ is independent of W_0 and has the same distribution as the time taken for N_a to go from 0 to 1 (that is, τ has an exponential distribution with rate $1 + a$). \square

Fix $k \geq 1$. Recall that $\Delta_{n,1} \geq \Delta_{n,2} \geq \dots \geq \Delta_{n,k}$ are the k largest out-degrees in \mathcal{T}_n (the out-degree of a vertex is its number of children). We will show that the vertices with these out-degrees occur in a *finite* neighborhood about the root and that the out-degrees of vertices in a finite neighborhood about the root converge in distribution when properly normalized.

Lemma 7.3. *In the branching process construction of the linear preferential attachment tree \mathcal{T}_n , let $\Delta_{n,1}^S$ denote the maximum out-degree in \mathcal{T}_n among all vertices born before time S . Given any $\varepsilon > 0$, there exists a finite constant S_ε such that*

$$\liminf_{n \rightarrow \infty} \mathbb{P}\{\Delta_{n,1}^{S_\varepsilon} = \Delta_{n,1}\} \geq 1 - \varepsilon.$$

Proof. Fix $\varepsilon > 0$. Write $A_{n,S}$ for the event

There exists a vertex in \mathcal{T}_n which was born after time S and has out-degree greater than the root.

The claim of the lemma may be rephrased as a statement that there is a finite constant S_ε such that

$$\limsup_{n \rightarrow \infty} \mathbb{P}(A_{n,S_\varepsilon}) \leq \varepsilon.$$

Note from Theorem 7.2(c) that there is a constant B_ε such that

$$\limsup_{n \rightarrow \infty} \mathbb{P}\left\{\left|T_n - \frac{1}{\gamma_a} \log n\right| > B_\varepsilon\right\} \leq \varepsilon/2.$$

Set $t_-^n := \frac{1}{\gamma_a} \log n - B_\varepsilon$ and $t_+^n := \frac{1}{\gamma_a} \log n + B_\varepsilon$. It is enough to prove that there exists a finite constant S_ε such that

$$\limsup_{n \rightarrow \infty} \mathbb{P}(A'_{n, S_\varepsilon}) \leq \varepsilon/2,$$

where $A'_{n, S}$ is the event

There exists a vertex born after time S that has out-degree greater than the root for some time t in the interval $[t_-^n, t_+^n]$.

Furthermore, since the out-degrees of vertices increase with time, it is enough to show that

$$\limsup_{n \rightarrow \infty} \mathbb{P}(A''_{n, S_\varepsilon}) \leq \varepsilon/2,$$

where $A''_{n, S}$ is the event

There exists a vertex born after time S such that the out-degree of the vertex at time t_+^n is greater than the out-degree of the root at time t_-^n .

For $t \geq 0$ and a time interval $I \subseteq [0, t]$ denote by $Z(I, t)$ the maximum out-degree at time t of all vertices born in the time interval I . Let $\zeta_\rho(t)$ denote the out-degree of the root at time t . Note that

$$A''_{n, S} = \{Z([S, t_+^n], t_+^n) > \zeta_\rho(t_-^n)\}.$$

Observe also, that for any constant K ,

$$\begin{aligned} \mathbb{P}(A''_{n, S}) &\leq \mathbb{P}\{\zeta_\rho(t_-^n) \leq K \text{ or } Z([S, t_+^n], t_+^n) > K\} \\ &\leq \mathbb{P}\{\zeta_\rho(t_-^n) \leq K\} + \mathbb{P}\{Z([S_\varepsilon, t_+^n], t_+^n) > K\}. \end{aligned}$$

It thus suffices to show that there is a sequence K_n and a constant S_ε such that

$$(7.1) \quad \limsup_{n \rightarrow \infty} \mathbb{P}\{\zeta_\rho(t_-^n) \leq K_n\} \leq \varepsilon/4$$

and

$$(7.2) \quad \limsup_{n \rightarrow \infty} \mathbb{P}\{Z([S, t_+^n], t_+^n) > K_n\} \leq \varepsilon/4.$$

It follows from Theorem 7.2(d) that the inequality (7.1) holds with $K_n = K_\varepsilon n^{1/\gamma_a}$ for a suitable constant $K_\varepsilon > 0$.

Turning to the inequality (7.2), assume without loss of generality that S and t_+^n are integers. In that case,

$$Z(S, t_+^n) = \max_{S \leq m \leq t_+^n - 1} Z([m, m+1], t_+^n).$$

Note that, by the union bound,

$$\mathbb{P}\{Z([m, m+1], t_+^n) > K_n\} \leq \mathbb{E}[\#\mathcal{F}(m+1)] \mathbb{P}\{N_a(t_+^n - m) > K_n\}.$$

Applying Theorem 7.2(b) and Lemma 7.1(b) gives

$$\mathbb{P}\{Z(S, t_+^n) > K_n\} \leq \sum_{m=S}^{t_+^n - 1} C e^{\gamma_a(m+1)} A e^{-C' e^m},$$

where $C' = K_\varepsilon / (A e^{B_\varepsilon})$. The inequality (7.2) follows upon choosing $S = S_\varepsilon$ large enough. \square

A slightly more detailed analysis shows that Lemma 7.3 can be generalized to the k maximal out-degrees for any fixed k . Let $\Delta_{n,1}^S \geq \Delta_{n,2}^S \geq \dots \geq \Delta_{n,k}^S$ be the k largest out-degrees in \mathcal{T}_n from among the vertices that are born before time S , with the convention that $\Delta_i^{S,n} = 0$ for $i \geq \#\mathcal{F}(S)$ when $\#\mathcal{F}(S) < k$. We leave the proof of the following result to the reader.

Lemma 7.4. *For any $\varepsilon > 0$ there exists a finite constant S_ε such that*

$$\liminf_{n \rightarrow \infty} \mathbb{P}\{(\Delta_{n,1}^{S_\varepsilon}, \dots, \Delta_{n,k}^{S_\varepsilon}) = (\Delta_{n,1}, \dots, \Delta_{n,k})\} \geq 1 - \varepsilon.$$

Proposition 7.5. *Fix $S > 0$ and consider the marked tree $\mathcal{T}_n^\#$ constructed by marking each vertex v of the tree $\mathcal{F}(S)$ with $n^{-1/\gamma_a} D(v, n) \in \mathbb{R}_+$, where $D(v, n)$ is the out-degree of v in \mathcal{T}_n . Then, $\mathcal{T}_n^\#$ converges almost surely as $n \rightarrow \infty$ to the tree $\mathcal{F}(S)$ equipped with a set of marks that are strictly positive almost surely.*

Proof. For any vertex $v \in \mathcal{F}(S)$, write $\zeta_v(t)$ for the out-degree (that is, number of offspring) of v at time t , so that $\zeta_v(T_n) = D(v, n)$. Note that $\zeta_v(S)$ can be computed by only looking at $\mathcal{F}(S)$ (recall that our trees are rooted). Conditional on $\mathcal{F}(S)$, the processes $\hat{\zeta}_v := (\zeta_v(S+t) - \zeta_v(S))_{t \geq 0}$, $v \in \mathcal{F}(S)$, are conditionally independent. Note that the conditional distribution of $\hat{\zeta}_v$ is that of a pure birth process that starts in state 0 and has birth rate $\zeta_v(S) + \ell + 1 + a$ in state ℓ . It follows from Theorem 7.2(d) that $e^{-t} \hat{\zeta}_v(t)$ converges almost surely as $t \rightarrow \infty$ to a random variable that has a conditional distribution which is that of the strictly positive random variable $W_{a+\zeta_v(S)}$. Hence, by Theorem 7.2(c),

$$\lim_{n \rightarrow \infty} e^{-(T_n-S)} \hat{\zeta}_v(T_n - S) = \lim_{n \rightarrow \infty} e^S Z_a n^{-1/\gamma_a} \hat{\zeta}_v(T_n - S)$$

exists almost surely and the limit is strictly positive almost surely.

The result follows because $n^{-1/\gamma_a} D(v, n) = n^{-1/\gamma_a} (\zeta_v(S) + \hat{\zeta}_v(T_n - S))$. \square

Corollary 7.6. *The random vector $n^{-1/\gamma_a} (\Delta_{n,1}^S, \dots, \Delta_{n,k}^S)$ converges almost surely to a random vector $(Y_1^S, Y_2^S, \dots, Y_k^S)$ as $n \rightarrow \infty$, where $Y_1^S \geq Y_2^S \geq \dots \geq Y_k^S > 0$ almost surely.*

Completion of the proof of Theorem 4.12. Given Corollary 7.6 and Lemma 7.4, the proof is completed by applying the following elementary result with $X_{n,i} = n^{-1/\gamma_a} \Delta_{n,i}$ and $Y_{n,i}^\varepsilon = n^{-1/\gamma_a} \Delta_{n,i}^{S_\varepsilon}$.

Lemma 7.7. *Let $(\mathbf{X}_n)_{n=1}^\infty = ((X_{n,1}, \dots, X_{n,k}))_{n=1}^\infty$ be a sequence of \mathbb{R}^k -valued random variables. Suppose for each fixed $\varepsilon > 0$ that there exists a sequence of \mathbb{R}^k -valued random variables $(\mathbf{Y}_n^{(\varepsilon)})_{n=1}^\infty = ((Y_{n,1}^\varepsilon, \dots, Y_{n,k}^\varepsilon))_{n=1}^\infty$ on the same probability space such that*

$$\liminf_{n \rightarrow \infty} \mathbb{P}\{\mathbf{X}_n = \mathbf{Y}_n^{(\varepsilon)}\} \geq 1 - \varepsilon.$$

Suppose further that for each $\varepsilon > 0$ there exists a random vector $\mathbf{Y}_\infty^{(\varepsilon)}$ such that $\mathbf{Y}_n^{(\varepsilon)}$ converges in probability to $\mathbf{Y}_\infty^{(\varepsilon)}$ as $n \rightarrow \infty$. Then, there exists an \mathbb{R}^k -valued random variable \mathbf{X}_∞ such that \mathbf{X}_n converges in probability to \mathbf{X}_∞ as $n \rightarrow \infty$.

Proof. Convergence in probability for the space of \mathbb{R}^k -valued random variables is metrized by the metric r , where $r(\mathbf{X}, \mathbf{Y}) := \mathbb{E}[|\mathbf{X} - \mathbf{Y}| \wedge 1]$ and where $|\cdot|$ denotes the

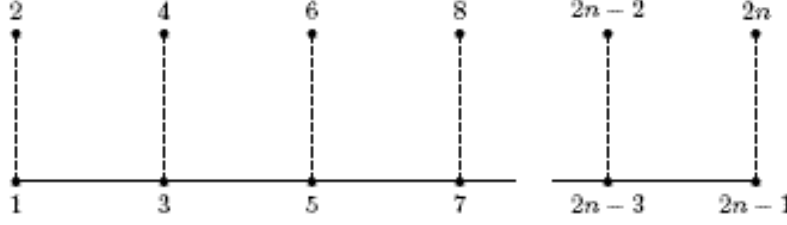


FIGURE 4. In the above graph \mathcal{G}_n , every horizontal edge always occurs whereas every vertical (dotted) edge occurs independently with probability $1 - q$. The comb tree \mathcal{T}_n is the unique connected component of \mathcal{G}_n that is not an isolated point.

Euclidean norm on \mathbb{R}^k . Moreover, this metric space is complete. By assumption,

$$\begin{aligned}
 \limsup_{m,n \rightarrow \infty} r(\mathbf{X}_m, \mathbf{X}_n) &\leq \limsup_{m \rightarrow \infty} r(\mathbf{X}_m, \mathbf{Y}_m^{(\varepsilon)}) \\
 &\quad + \limsup_{m,n \rightarrow \infty} r(\mathbf{Y}_m^{(\varepsilon)}, \mathbf{Y}_n^{(\varepsilon)}) + \limsup_{n \rightarrow \infty} r(\mathbf{Y}_n^{(\varepsilon)}, \mathbf{X}_n) \\
 &\leq 2 \limsup_{n \rightarrow \infty} \mathbb{P}\{\mathbf{Y}_n^\varepsilon \neq \mathbf{X}_n\} \\
 &\leq 2\varepsilon.
 \end{aligned}$$

The sequence $(\mathbf{X}_n)_{n=1}^\infty$ is thus Cauchy in the metric r , and hence it converges in probability to a limit \mathbf{X}_∞ . \square

8. AN EXAMPLE

Consider the following construction of a random finite graph \mathcal{G}_n for a given positive integer n . The vertices of \mathcal{G}_n are $\{1, \dots, 2n\}$. Let $\varepsilon_{n1}, \dots, \varepsilon_{nn}$ be independent, identically distributed random variables, with $\mathbb{P}\{\varepsilon_{nk} = 0\} = q$ and $\mathbb{P}\{\varepsilon_{nk} = 1\} = 1 - q$, where $0 < q < 1$. There is always an edge between the vertices $2k - 1$ and $2k + 1$ for $1 \leq k \leq n - 1$, and there is an edge between the vertices $2k - 1$ and $2k$ for $1 \leq k \leq n$ if and only if $\varepsilon_{nk} = 1$. There are no other edges.

The graph \mathcal{G}_n consists of a large connected component \mathcal{T}_n with vertices

$$\{1, 3, \dots, 2n - 1\} \cup \{2k : 1 \leq k \leq n, \varepsilon_{nk} = 1\}$$

plus the (possibly empty) set of isolated points $\{2k : 1 \leq k \leq n, \varepsilon_{nk} = 0\}$. The graph \mathcal{T}_n is a tree. It resembles a comb with some of the teeth missing, see Figure 4.

No matter how we chose a root, the sequence of random finite trees $(\mathcal{T}_n)_{n=1}^\infty$ would not converge in the probability fringe sense because the cardinality of the subtree below a uniformly chosen point would not converge in distribution to a finite random variable. However, if we look at the empirical distribution of the tree within graph distance k of a vertex v as v ranges over the \mathcal{T}_n , then it is not hard to see that this random measure converges to a deterministic limit for every k . This observation and the fact that the vertices of \mathcal{T}_n have degree at most 3 shows that the moments of the spectral distribution of \mathcal{T}_n converge as $n \rightarrow \infty$ to finite constants, and these constants are the moments of a distribution that is determined

by its moments. Therefore, the spectral distribution of \mathcal{T}_n converges in distribution to a deterministic limit as $n \rightarrow \infty$, and in order to compute the moments of that limit, it suffices to compute the limits of the expectations of the moments of the spectral distribution of \mathcal{T}_n .

Write $Z_n = \#\{1 \leq k \leq n : \varepsilon_{nk} = 0\}$. By permuting indices, it is possible to re-write the adjacency matrix of \mathcal{G}_n in block form, where the upper-left block has dimensions $(2n - Z_n) \times (2n - Z_n)$ and is the adjacency matrix of \mathcal{T}_n , while the lower-right block is the zero matrix of dimensions $Z_n \times Z_n$. Therefore, if we write F_n (respectively, H_n) for the empirical distribution of the eigenvalues of the adjacency matrix of \mathcal{T}_n (respectively, \mathcal{G}_n), then

$$H_n = \frac{Z_n}{2n} \delta_0 + \left(1 - \frac{Z_n}{2n}\right) F_n,$$

where δ_0 is the unit point mass at 0. Since Z_n/n converges in probability to q as $n \rightarrow \infty$, the limiting behavior of F_n is determined by that of H_n and *vice-versa*: H_n converges in probability to a non-random probability measure H and

$$H = \frac{q}{2} \delta_0 + \left(1 - \frac{q}{2}\right) F,$$

where the probability measure F is the limit of the sequence F_n .

Define a random infinite graph \mathcal{G} with vertex set \mathbb{Z} as follows. Let ε_k , $k \in \mathbb{Z}$, be independent, identically distributed random variables, with $\mathbb{P}\{\varepsilon_k = 0\} = q$ and $\mathbb{P}\{\varepsilon_k = 1\} = 1 - q$. There is an edge between the vertices $2k - 1$ and $2k + 1$ for all $k \in \mathbb{Z}$, and there is an edge between the vertices $2k - 1$ and $2k$ for $k \in \mathbb{Z}$ if and only if $\varepsilon_k = 1$. There are no other edges.

Let B_n (resp. B) denote the adjacency matrix of \mathcal{G}_n (resp. \mathcal{G}). For each non-negative integer m ,

$$\begin{aligned} & \int x^m H(dx) \\ &= \lim_{n \rightarrow \infty} \int x^m H_n(dx) \\ &= \lim_{n \rightarrow \infty} \mathbb{E} \left[\int x^m H_n(dx) \right] \\ &= \lim_{n \rightarrow \infty} \mathbb{E} \left[\frac{1}{2n} \text{tr } B_n^m \right] \\ &= \mathbb{E} \left[\frac{1}{2} (B^m)_{11} + \frac{1}{2} (B^m)_{22} \right] \end{aligned}$$

(because $B_{ij} = 0$ for $|i - j| > 2$, there is no problem defining B^m).

Divide the matrix B into 2×2 blocks with the rows (resp. columns) of the $(i, j)^{\text{th}}$ block indexed by $\{2i + 1, 2i + 2\}$ (resp. $\{2j + 1, 2j + 2\}$). This block form matrix is block tridiagonal. The entries in the diagonals above and below the main diagonal are always the matrix

$$\Pi_x := \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

We use this notation because we can think of Π_x as the matrix for the orthogonal projection onto the x -axis in a two-dimensional (x, y) coordinate system. The entry

in the (k, k) diagonal block is the matrix

$$\begin{pmatrix} 0 & \varepsilon_{k+1} \\ \varepsilon_{k+1} & 0 \end{pmatrix}.$$

If $\varepsilon_{k+1} = 1$, then this is the matrix

$$\Sigma := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

that permutes the two coordinates. Otherwise, it is the 2×2 zero matrix. By analogy with the definition of Π_x , set

$$\Pi_y := \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

We have the following relations between these matrices

$$\begin{aligned} \Pi_x^2 &= \Pi_x \\ \Pi_y^2 &= \Pi_y \\ \Pi_x \Pi_y &= \Pi_y \Pi_x = 0 \\ \Sigma^2 &= I \\ \Sigma \Pi_x &= \Pi_y \Sigma \\ \Sigma \Pi_y &= \Pi_x \Sigma \\ \text{tr } \Pi_x &= \text{tr } \Pi_y = 1 \\ \text{tr } \Sigma \Pi_x &= \text{tr } \Pi_x \Sigma = 0 \\ \text{tr } \Sigma \Pi_y &= \text{tr } \Pi_y \Sigma = 0. \end{aligned}$$

A consequence of these relations is the following.

Lemma 8.1. *For $a \geq 0$,*

$$\text{tr } \Sigma^a = \begin{cases} 2, & \text{if } a \text{ is even,} \\ 0, & \text{if } a \text{ is odd.} \end{cases}$$

For any $r \geq 1$,

$$\text{tr } (\Sigma^{a_1} \Pi_x \Sigma^{a_2} \Pi_x \dots \Sigma^{a_r} \Pi_x \Sigma^{a_{r+1}}) = \begin{cases} 1, & \text{if } a_1 + a_{r+1}, a_2, \dots, a_r \text{ are all even,} \\ 0, & \text{otherwise.} \end{cases}$$

Proof. The first claim is obvious, because Σ^a either the 2×2 identity matrix or Σ depending on whether a is even or odd.

For the second claim, first observe that the product in question may be rewritten as

$$\Sigma^{b_1} \Pi_x \Sigma^{b_2} \Pi_x \dots \Sigma^{b_r} \Pi_x \Sigma^{b_{r+1}},$$

where b_ℓ is 0 or 1 depending on whether a_ℓ is even or odd. This in turn may be rewritten as

$$\Pi_{z_1} \Pi_{z_2} \dots \Pi_{z_r} \Sigma^c,$$

where z_ℓ , $1 \leq \ell \leq r$, is x or y depending on whether $b_1 + \dots + b_\ell$ is even or odd, and c is 0 or 1 depending on whether $b_1 + \dots + b_{r+1}$ is even or odd.

The product is non-zero if and only if $z_1 = z_2 = \dots = z_r$. This is equivalent to either $b_1 = 0$ and $b_2 = \dots = b_r = 0$, in which case the product is $(\Pi_x)^r \Sigma^c = \Pi_x \Sigma^c$, or $b_1 = 1$ and $b_2 = \dots = b_r = 0$, in which case the product is $(\Pi_y)^r \Sigma^c = \Pi_y \Sigma^c$.

Furthermore, even if the product is non-zero, and hence of the form $\Pi_x \Sigma^c$ or $\Pi_y \Sigma^c$, the trace is zero if $c = 1$. Otherwise, the trace is 1.

Thus, the trace is zero unless $b_2 = \dots = b_r = 0$ and $b_1 + \dots + b_{r+1}$ is even, in which case the trace is 1. This condition is equivalent to a_2, \dots, a_r and $a_1 + \dots + a_{r+1}$ all being even, and the result follows. \square

In order to carry out the combinatorial calculations involved in computing the expected value of $(B^m)_{11} + (B^m)_{22}$, it will be helpful to recast them as computations involving a certain killed random walk on a random subinterval of the integers.

Let U_1, U_2, \dots be independent, identically distributed random variables that are independent of the random infinite matrix B and have common distribution

$$\mathbb{P}\{U_k = -1\} = \mathbb{P}\{U_k = 0\} = \mathbb{P}\{U_k = +1\} = \frac{1}{3}.$$

Set $W_m = U_1 + \dots + U_m$ (with $W_0 = 0$). Note that

$$\begin{aligned} & \mathbb{E} \left[\begin{pmatrix} (B^m)_{11} & (B^m)_{12} \\ (B^m)_{21} & (B^m)_{22} \end{pmatrix} \right] \\ &= 3^m \mathbb{E} [\Xi_1 \Xi_2 \dots \Xi_m \mathbf{1}\{W_m = 0 \text{ \& } \alpha < W_\ell < \beta, 1 \leq \ell \leq m\}], \end{aligned}$$

where

$$\begin{aligned} \alpha &:= \max\{k \leq 0 : \varepsilon_{k+1} = 0\}, \\ \beta &:= \min\{k \geq 0 : \varepsilon_{k+1} = 0\} \end{aligned}$$

and

$$\Xi_\ell := \begin{cases} \Pi_x, & \text{if } U_\ell \in \{-1, +1\}, \\ \Sigma, & \text{if } U_\ell = 0. \end{cases}$$

Lemma 8.2. *The random variables α and β satisfy*

$$\mathbb{P}\{\beta - \alpha = N\} = \begin{cases} q, & \text{if } N = 0, \\ 0, & \text{if } N = 1, \\ (N-1)(1-q)^{N-1}q^2, & \text{if } N \geq 2, \end{cases}$$

and

$$\mathbb{P}\{0 - \alpha = j \mid \beta - \alpha = N\} = \frac{1}{N-1}, \quad N \geq 2, \quad 1 \leq j \leq N-1.$$

Proof. It is clear that

$$\mathbb{P}\{\beta - \alpha = 0\} = \mathbb{P}\{\varepsilon_1 = 0\} = q$$

and $\mathbb{P}\{\beta - \alpha = 1\} = 0$.

For $N \geq 2$ and $1 \leq j \leq N-1$,

$$\begin{aligned} & \mathbb{P}\{\beta - \alpha = N, 0 - \alpha = j\} \\ &= \mathbb{P}\{\varepsilon_{-j+1} = 0, \varepsilon_{-j+2} = \dots = \varepsilon_{-j+N} = 1, \varepsilon_{-j+N+1} = 0\} \\ &= (1-q)^{N-1}q^2 \\ &= (N-1)(1-q)^{N-1}q^2 \times \frac{1}{N-1}, \end{aligned}$$

and the remainder of the result follows. \square

Consider the moment generating function $\varphi(t) := \sum_{m=0}^{\infty} \frac{t^m}{m!} \int x^m H(dx)$. From the above, $\varphi(t) = \varphi(-t)$ for all $t \in \mathbb{R}$ and

$$\exp(-3t)\varphi(t) = \frac{1}{2} \mathbb{E}[\text{tr}(\Xi_1 \Xi_2 \cdots \Xi_M) \mathbf{1}\{W_M = 0 \text{ \& } \alpha < W_\ell < \beta, 1 \leq \ell \leq M\}]$$

for $t \geq 0$, where M is a Poisson distributed random variable with expected value $3t$ that is independent of the random infinite matrix B and the increments U_1, U_2, \dots

Put $\mathcal{I}' := \{1 \leq \ell \leq M : U_\ell = \pm 1\}$ and $\mathcal{I}'' := \{1 \leq \ell \leq M : U_\ell = 0\}$. Then $M' := \#\mathcal{I}'$ and $M'' := \#\mathcal{I}''$ are independent Poisson distributed random variables with expected values $2t$ and t , respectively. Write $\mathcal{I}' = \{\ell_1, \ell_2, \dots, \ell_{M'}\}$, with $\ell_1 < \ell_2 < \dots < \ell_{M'}$. The set \mathcal{I}'' is the disjoint union of sets $\mathcal{I}_1'', \mathcal{I}_2'', \dots, \mathcal{I}_{M'+1}''$, where

$$\mathcal{I}_1'' := \{1 \leq \ell \leq M : \ell < \ell_1\},$$

$$\mathcal{I}_j'' := \{1 \leq \ell \leq M : \ell_{j-1} < \ell < \ell_j\}, \quad 1 \leq j \leq M',$$

and

$$\mathcal{I}_{M'+1}'' := \{1 \leq \ell \leq M : \ell_{M'} < \ell\}.$$

Conditional on $M' = m'$, the random variables $\#\mathcal{I}_1'', \#\mathcal{I}_2'', \dots, \#\mathcal{I}_{m'+1}''$ are independent and identically distributed with a common distribution that is Poisson with expected value $t/(m' + 1)$.

In view of Lemma 8.1, the following well-known result will be useful.

Lemma 8.3. *The probability that a Poisson random variable with expected value λ takes an even value is $\frac{1}{2}[1 + \exp(-2\lambda)]$.*

Proof. The probability in question is

$$\sum_{k=0}^{\infty} \exp(-\lambda) \frac{\lambda^{2k}}{(2k)!} = \exp(-\lambda) \frac{1}{2} [\exp(\lambda) + \exp(-\lambda)] = \frac{1}{2} [1 + \exp(-2\lambda)].$$

□

The conditional distribution of the sequence $0 = W_0, W_{\ell_1}, \dots, W_{\ell_{m'}}$ given $M' = m'$ is that of a symmetric simple random walk with m' steps, and we need the following result.

Lemma 8.4. *Fix an integer $N \geq 2$. Suppose that S_0, S_1, S_2, \dots is a symmetric simple random walk. Then*

$$\sum_{j=1}^{N-1} \mathbb{P}\{S_n = j \text{ \& } 0 < S_\ell < N, 1 \leq \ell \leq n \mid S_0 = j\} = \sum_{k=1}^{N-1} \left(\cos \frac{k\pi}{N} \right)^n$$

for $n = 0, 1, 2, \dots$. This quantity is zero when n is odd.

Proof. Consider the transition matrix P of a symmetric simple random walk killed on exiting $\{1, 2, \dots, N-1\}$. That is, P is the $(N-1) \times (N-1)$ symmetric matrix with rows and columns both indexed by $\{1, 2, \dots, N-1\}$ and entries

$$P_{ij} = \begin{cases} \frac{1}{2}, & \text{if } |i - j| = 1, \\ 0, & \text{otherwise.} \end{cases}$$

It is a straightforward (and well-known) calculation that the eigenvalues of P are $\cos \frac{k\pi}{N}$, $1 \leq k \leq N-1$, and the corresponding normalized eigenvectors are $\sqrt{\frac{2}{N}}(\sin \frac{k\pi}{N} j)_{j=1}^{N-1}$. Hence,

$$(P^n)_{ij} = \sum_{k=1}^{N-1} \left(\cos \frac{k\pi}{N} \right)^n \frac{2}{N} \left(\sin \frac{k\pi}{N} i \right) \left(\sin \frac{k\pi}{N} j \right).$$

Thus,

$$\begin{aligned} \sum_{j=1}^{N-1} (P^n)_{jj} &= \sum_{k=1}^{N-1} \left(\cos \frac{k\pi}{N} \right)^n \frac{2}{N} \sum_{j=1}^{N-1} \left(\sin \frac{k\pi}{N} j \right)^2 \\ &= \sum_{k=1}^{N-1} \left(\cos \frac{k\pi}{N} \right)^n, \end{aligned}$$

as claimed. \square

Observe from Lemma 8.1 that for $t \geq 0$ the moment generating function φ of H satisfies

$$\begin{aligned} \exp(-3t)\varphi(t) &= \frac{1}{2} \mathbb{E} [\text{tr} (\Xi_1 \Xi_2 \cdots \Xi_M) \mathbf{1}\{W_M = 0 \text{ \& } \alpha < W_\ell < \beta, 1 \leq \ell \leq M\}] \\ &= \frac{1}{2} \mathbb{E} \left[\text{tr} \left(\Sigma^{\# \mathcal{I}_1''} \Pi_x \Sigma^{\# \mathcal{I}_2''} \cdots \Sigma^{\# \mathcal{I}_{M'}''} \Pi_x \Sigma^{\# \mathcal{I}_{M'+1}''} \right) \right. \\ &\quad \left. \times \mathbf{1}\{W_M = 0 \text{ \& } \alpha < W_\ell < \beta, 1 \leq \ell \leq M\} \right] \\ &= \frac{1}{2} \left(2 \times \mathbb{P}\{M' = 0, M'' \text{ is even, \& } \alpha < 0 < \beta\} \right. \\ &\quad \left. + 1 \times \mathbb{P}\{M' > 0, \# \mathcal{I}_1'' + \# \mathcal{I}_{M'+1}'', \# \mathcal{I}_2'', \dots, \# \mathcal{I}_{M'}'', \text{ are all even,} \right. \\ &\quad \left. S_{M'} = 0, \text{ \& } \alpha < S_k < \beta, 1 \leq k \leq M'\} \right) \end{aligned}$$

where $S_k = W_{\ell_k}$ for $1 \leq k \leq M'$.

Conditioning on M' taking the even value $2h$ and $\beta - \alpha$ taking the value N , we hence, from Lemmas 8.3 and 8.4,

$$\begin{aligned} \exp(-3t)\varphi(t) &= \exp(-2t) \frac{1}{2} [1 + \exp(-2t)] (1 - q) \\ &\quad + \frac{1}{2} \sum_{h=1}^{\infty} \exp(-2t) \frac{(2t)^{2h}}{2h!} \\ &\quad \times \left\{ \frac{1}{2} \left[1 + \exp \left(-\frac{4t}{2h+1} \right) \right] \right\} \left\{ \frac{1}{2} \left[1 + \exp \left(-\frac{2t}{2h+1} \right) \right] \right\}^{2h-1} \\ &\quad \times \sum_{N=2}^{\infty} (N-1)(1-q)^{N-1} q^2 \frac{1}{N-1} \sum_{k=1}^{N-1} \left(\cos \frac{k\pi}{N} \right)^{2h} \end{aligned}$$

for $t \geq 0$. Thus,

$$\begin{aligned} \varphi(t) &= (1-q) \cosh t \\ &+ \frac{1}{2} \sum_{h=1}^{\infty} \frac{(2t)^{2h}}{2h!} \\ &\quad \times \cosh \left(-\frac{2t}{2h+1} \right) \left\{ \cosh \left(-\frac{t}{2h+1} \right) \right\}^{2h-1} \\ &\quad \times \sum_{N=2}^{\infty} (1-q)^{N-1} q^2 \sum_{k=1}^{N-1} \left(\cos \frac{k\pi}{N} \right)^{2h}. \end{aligned}$$

for $t \geq 0$. The right-hand side thought of as a function of $t \in \mathbb{R}$ is an even function, and it follows that the right-hand side is the correct expression for $\varphi(t)$ for all $t \in \mathbb{R}$.

Finally, if we write ψ for the moment generating function of F , then

$$\psi(t) = \frac{\varphi(t) - \frac{q}{2}}{1 - \frac{q}{2}}.$$

Acknowledgments. We thank Charles Bordenave for helpful discussions. S.B. thanks David Aldous for his constant encouragement and illuminating discussions regarding local weak convergence methodology. S.N.E. conducted part of this research while he was a Visiting Fellow at the Mathematical Sciences Institute of the Australian National University.

REFERENCES

- [Ald91a] David Aldous, *Asymptotic fringe distributions for general families of random trees*, Ann. Appl. Probab. **1** (1991), no. 2, 228–266. MR MR1102319 (92j:60009)
- [Ald91b] David Aldous, *The continuum random tree I*, Ann. Probab. **19** (1991), 1–28.
- [AS04] David Aldous and J. Michael Steele, *The objective method: probabilistic combinatorial optimization and local weak convergence*, Probability on discrete structures, Encyclopaedia Math. Sci., vol. 110, Springer, Berlin, 2004, pp. 1–72. MR MR2023650 (2005e:60018)
- [BG00] M. Bauer and O. Golinelli, *On the kernel of tree incidence matrices*, J. Integer Seq. **3** (2000), no. 1, Article 00.1.4, 1 HTML document (electronic). MR MR1750745 (2001b:05138)
- [BG01] ———, *Random incidence matrices: Moments of the spectral density*, Journal of Statistical Physics **103** (2001), no. 1, 301–337.
- [Bha07] Shankar Bhamidi, *Universal techniques for preferential attachment: Local and global analysis*, 2007, Preprint available at <http://www.stat.berkeley.edu/users/shanky/preferent.pdf>.
- [BI01] Pavel Bleher and Alexander Its (eds.), *Random matrix models and their applications*, Cambridge University Press, 2001.
- [Big93] Norman Biggs, *Algebraic graph theory*, second ed., Cambridge Mathematical Library, Cambridge University Press, Cambridge, 1993. MR MR1271140 (95h:05105)
- [BL07] C. Bordenave and M. Lelarge, *Resolvent of large random graphs*, 2007, [arXiv:0801.0155v2](https://arxiv.org/abs/0801.0155v2).
- [BM76] J. A. Bondy and U. S. R. Murty, *Graph theory with applications*, American Elsevier Publishing Co., Inc., New York, 1976. MR MR0411988 (54 #117)
- [BM93] Phillip Botti and Russell Merris, *Almost all trees share a complete set of immanantal polynomials*, J. Graph Theory **17** (1993), no. 4, 467–476. MR MR1231010 (94g:05053)
- [BN04] Béla Bollobás and Vladimir Nikiforov, *Graphs and Hermitian matrices: eigenvalue interlacing*, Discrete Math. **289** (2004), no. 1-3, 119–127. MR MR2106034 (2005j:05058)

- [BR03] Béla Bollobás and Oliver M. Riordan, *Mathematical results on scale-free random graphs*, Handbook of graphs and networks, Wiley-VCH, Weinheim, 2003, pp. 1–34. MR MR2016117 (2004j:05108)
- [CDGT88] Dragoš M. Cvetković, Michael Doob, Ivan Gutman, and Aleksandar Torgašev, *Recent results in the theory of graph spectra*, Annals of Discrete Mathematics, vol. 36, North-Holland Publishing Co., Amsterdam, 1988. MR MR926481 (89d:05130)
- [CDS95] Dragoš M. Cvetković, Michael Doob, and Horst Sachs, *Spectra of graphs*, third ed., Johann Ambrosius Barth, Heidelberg, 1995, Theory and applications. MR MR1324340 (96b:05108)
- [Chu97] Fan R. K. Chung, *Spectral graph theory*, CBMS Regional Conference Series in Mathematics, vol. 92, Published for the Conference Board of the Mathematical Sciences, Washington, DC, 1997. MR MR1421568 (97k:58183)
- [CLV03a] F. Chung, L. Lu, and V. Vu, *Eigenvalues of random power law graphs*, Annals of Combinatorics **7** (2003), no. 1, 21–33.
- [CLV03b] ———, *Spectra of random graphs with given expected degrees*, Proceedings of the National Academy of Sciences **100** (2003), no. 11, 6313–6318.
- [CRS97] D. Cvetković, P. Rowlinson, and S. Simić, *Eigenspaces of graphs*, Encyclopedia of Mathematics and its Applications, vol. 66, Cambridge University Press, Cambridge, 1997. MR MR1440854 (98f:05111)
- [Dei00] Percy Deift, *Orthogonal polynomials and random matrices: A Riemann-Hilbert approach*, Courant Lecture Notes, American Mathematical Society, 2000.
- [Dur07] Rick Durrett, *Random graph dynamics*, Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, Cambridge, 2007. MR MR2271734
- [FFF03] Abraham Flaxman, Alan Frieze, and Trevor Fenner, *High degree vertices and eigenvalues in the preferential attachment graph*, Approximation, randomization, and combinatorial optimization, Lecture Notes in Comput. Sci., vol. 2764, Springer, Berlin, 2003, pp. 264–274. MR MR2080798 (2005e:05124)
- [FFF05] ———, *High degree vertices and eigenvalues in the preferential attachment graph*, Internet Math. **2** (2005), no. 1, 1–19. MR MR2166274 (2006e:05157)
- [Fio99] M. A. Fiol, *Eigenvalue interlacing and weight parameters of graphs*, Linear Algebra Appl. **290** (1999), no. 1-3, 275–301. MR MR1673001 (2000c:05100)
- [For09] Peter Forrester, *Log-gases and random matrices*, 2009, Preprint of book available at <http://www.ms.unimelb.edu.au/matpjf/matpjf.html>.
- [GR01] Chris Godsil and Gordon Royle, *Algebraic graph theory*, Graduate Texts in Mathematics, vol. 207, Springer-Verlag, New York, 2001. MR MR1829620 (2002f:05002)
- [Gri81] G. R. Grimmett, *Random labelled trees and their branching networks*, J. Austral. Math. Soc. Ser. A **30** (1980/81), no. 2, 229–237. MR MR607933 (82g:05042)
- [Gui09] Alice Guionnet, *Large random matrices: Lectures on macroscopic asymptotics: École d'été de Probabilités de Saint-Flour XXXVI 2006*, Lecture Notes in Mathematics, Springer, 2009.
- [Hae95] Willem H. Haemers, *Interlacing eigenvalues and graphs*, Linear Algebra Appl. **226/228** (1995), 593–616. MR MR1344588 (96e:05110)
- [HJ90] Roger A. Horn and Charles R. Johnson, *Matrix analysis*, Cambridge University Press, Cambridge, 1990, Corrected reprint of the 1985 original. MR MR1084815 (91i:15001)
- [Jag89] Peter Jagers, *General branching processes as Markov fields*, Stochastic Process. Appl. **32** (1989), no. 2, 183–212. MR MR1014449 (91d:60208)
- [JN84] Peter Jagers and Olle Nerman, *The growth and composition of branching populations*, Adv. in Appl. Probab. **16** (1984), no. 2, 221–259. MR MR742953 (86j:60193)
- [KS81] R. Karp and M. Sipser, *Maximum matchings in sparse random graphs*, 22nd Annual Symposium on Foundations of Computer Science (1981), 364–375.
- [ME06] Frederick A. Matsen and Steven N. Evans, *Ubiquity of synonymy: almost all large binary trees are not uniquely identified by their spectra or their immanantal polynomials*, U.C. Berkeley Department of Statistics Technical Report No. 698, 2006.
- [Meh04] Madan Lal Mehta, *Random matrices*, Third ed., Pure and Applied Mathematics, vol. 142, Academic Press, 2004.
- [Mór05] Tamás F. Móri, *The maximum degree of the Barabási-Albert random tree*, Combin. Probab. Comput. **14** (2005), no. 3, 339–348. MR MR2138118 (2006a:60014)

- [NJ84] Olle Nerman and Peter Jagers, *The stable double infinite pedigree process of supercritical branching populations*, Z. Wahrsch. Verw. Gebiete **65** (1984), no. 3, 445–460. MR MR731231 (85e:60091)
- [Roj06a] Oscar Rojo, *On the spectra of certain rooted trees*, Linear Algebra Appl. **414** (2006), no. 1, 218–243. MR MR2209241 (2006m:05156)
- [Roj06b] ———, *The spectra of some trees and bounds for the largest eigenvalue of any tree*, Linear Algebra Appl. **414** (2006), no. 1, 199–217. MR MR2209240 (2007j:05145)
- [Roj07] ———, *The spectra of a graph obtained from copies of a generalized Bethe tree*, Linear Algebra Appl. **420** (2007), no. 2-3, 490–507. MR MR2278225 (2007g:05116)
- [Roj08] ———, *Spectra of weighted generalized Bethe trees joined at the root*, Linear Algebra Appl. **428** (2008), no. 11-12, 2961–2979. MR MR2416602
- [RR07a] Oscar Rojo and María Robbiano, *An explicit formula for eigenvalues of Bethe trees and upper bounds on the largest eigenvalue of any tree*, Linear Algebra Appl. **427** (2007), no. 1, 138–150. MR MR2353161 (2008g:05131)
- [RR07b] ———, *On the spectra of some weighted rooted trees and applications*, Linear Algebra Appl. **420** (2007), no. 2-3, 310–328. MR MR2278210 (2007i:05118)
- [RS05] Oscar Rojo and Ricardo Soto, *The spectra of the adjacency matrix and Laplacian matrix for some balanced trees*, Linear Algebra Appl. **403** (2005), 97–117. MR MR2140275 (2006b:05081)
- [RTV07] Anna Rudas, Bálint Tóth, and Benedek Valkó, *Random trees and general branching processes*, Random Structures Algorithms **31** (2007), no. 2, 186–202. MR MR2343718
- [Sch73] Allen J. Schwenk, *Almost all trees are cospectral*, New directions in the theory of graphs (Proc. Third Ann Arbor Conf., Univ. Michigan, Ann Arbor, Mich., 1971), Academic Press, New York, 1973, pp. 275–307. MR MR0384582 (52 #5456)
- [SM94] Robert T. Smythe and Hosam M. Mahmoud, *A survey of recursive trees*, Teor. ĭmovĭr. Mat. Stat. (1994), no. 51, 1–29. MR MR1445048 (97k:60027)
- [TV04] M. Tulino and S. Verdu, *Random matrix theory and wireless communications*, Foundations and Trends in Communications and Information Theory, Now Publishers, 2004.
- [Zol01] V.M. Zolotarev, *Lévy metric*, Encyclopaedia of Mathematics (Michiel Hazewinkel, ed.), Kluwer Academic Publishers, 2001.

E-mail address: shankar@math.ubc.ca

DEPARTMENT OF MATHEMATICS, THE UNIVERSITY OF BRITISH COLUMBIA, ROOM 121, 1984 MATHEMATICS ROAD, VANCOUVER, B.C., CANADA V6T 1Z2

E-mail address: evans@stat.Berkeley.EDU

E-mail address: arnab@stat.Berkeley.EDU

DEPARTMENT OF STATISTICS #3860, UNIVERSITY OF CALIFORNIA AT BERKELEY, 367 EVANS HALL, BERKELEY, CA 94720-3860, U.S.A.