

# 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 many 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 lower bounds 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, thereby significantly extending a result obtained by Aldous and Steele in the special case of uniform random trees.

We greatly generalize 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:* April 11, 2011.

*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 grants DMS-0405778 and DMS-0907630.

## CONTENTS

1. Introduction	3
2. Some representative random tree models	5
3. Probability fringe convergence of random trees	6
4. Statement of results	9
4.1. Convergence of the spectral distribution and atoms in the limiting spectral distribution	9
4.2. The proportion of zero eigenvalues and maximal matchings	11
4.3. Isospectrality	12
4.4. Largest eigenvalues and largest degrees	12
5. Convergence of spectral distributions	13
5.1. Interlacing inequalities and some of their consequences	13
5.2. Proof of Theorem 4.1(a)	15
5.3. Proof of Theorem 4.1(b)	16
5.4. Proof of Theorem 4.1(c)	16
6. Maximal matchings and the number of zero eigenvalues	17
6.1. Combinatorial preliminaries	17
6.2. Maximal weighted matchings: Proof of Theorem 4.7	20
6.3. Asymptotics of the number of zero eigenvalues	22
7. Largest eigenvalues: Proof of Theorem 4.10	25
8. An example: the random comb	29
8.1. Random walk on $\mathbb{Z}$ with random self-loops and the lamplighter group	35
References	37

## 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 interest 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 Remark 4.5, 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 a path of  $n$  points and 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 investigate the asymptotic moments of the spectral distribution for this ensemble and show that even in this simple case closed form expressions appear to be rather elusive, indicating that we should perhaps not expect simple characterizations of the limiting spectral distribution for more complex models.

Quite closely related to our results for the spectral distribution is the recent work [BL10a], 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  $\gamma$  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. Incidentally, the Karp-Sipser algorithm has been also used in a recent work [BL10b] to study the limiting proportion of

zero eigenvalues of random sparse graphs. 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 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 as 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  $\mathbb{T}^\infty$  and  $\mathbb{T}_*^\infty$  with the usual product topologies.

Consider a finite unlabeled rooted tree  $\mathbf{t} \in \mathbb{T}$  with root  $\rho$  and another distinguished vertex  $v$  that is at distance  $h$  from the root ( $v$  may coincide with  $\rho$ , 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  $\mathbb{T}_*^\infty$  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  $\Phi$  is a probability kernel from  $\mathbb{T}$  to  $\mathbb{T}_*^\infty$ .

**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  $\mathbb{T}^\infty$ -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  $\mathbb{T}_*^\infty$  converges weakly to the distribution of  $\mathcal{T}$  in the topology of weak convergence of probability measures on  $\mathbb{T}_*^\infty$ .

*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 the point process representing the offspring distribution of any individual is a pure birth point process started at 0 that can be described as follows: Whenever any individual has already given birth to  $k$  direct offspring, the individual produces a new offspring at rate  $k + 1 + a$ . In particular, at the time an individual is born, the individual generates new offspring at rate  $1 + a$ . Thus, the times that elapse between the birth of an individual and the successive births of the individual’s offspring, say  $(\beta_1, \beta_2, \dots)$ , may be written as  $\beta_i = \sum_{j=0}^{i-1} \eta_j$ , where the successive  $\eta_j$  are independent exponential random variables and  $\eta_j$  has rate  $j + 1 + a$ . Each individual in the population has its own independent and identically distributed copy of the above offspring point 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 linear preferential attachment tree on  $n$  vertices with parameter  $a > -1$ .*

(c) [Uniform random rooted unordered labeled tree] *Let  $Z_\infty$  be the complete family tree for a (discrete-time) Galton-Watson branching process with mean 1 Poisson offspring distribution. Note that  $Z_\infty$  is finite almost surely. The distribution of  $Z_\infty$  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  $\mathbb{T}^\infty$ . 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 preferential 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 processes 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 subject to the minor modifications that the point process describing the times at which the root individual gives birth is conditioned to have a birth at time  $X_i$  and the offspring born at this time and all its descendants are removed from the population. All other vertices give birth to according to the original offspring point process. 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  $Z_\infty$  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  $\rho_i$  to the root of  $\mathcal{F}_i(S_{i-1})$  and re-rooting the resulting tree at  $\rho_i$ . Then, the limiting **sin**-tree has the distribution of  $(\mathcal{U}_i)_{i=0}^\infty$ .



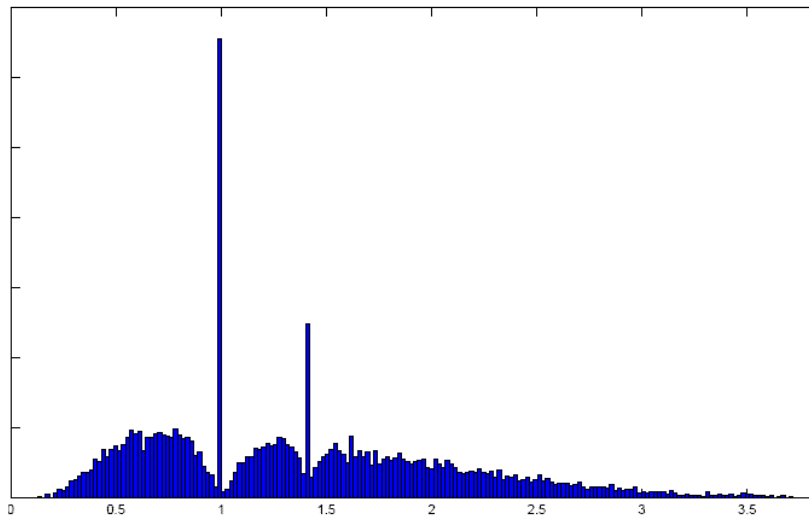


FIGURE 1. Empirical distribution for the positive eigenvalues of the random recursive tree with 200 vertices, averaged over 200 realizations.

#### 4. STATEMENT OF RESULTS

##### 4.1. Convergence of the spectral distribution and atoms in the limiting 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**  $\mathcal{T} = (\mathcal{T}^0, \mathcal{T}^1, \dots)$ . Let  $F_n$  denote the spectral distribution of the adjacency matrix of  $\mathcal{T}_n$ . Then the following are true.*

- (a) *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  $\mathbb{R}$ .*
- (b) *For any  $\gamma \in \mathbb{R}$ ,  $F_n(\{\gamma\})$  converges in distribution to a (model dependent) constant  $c_\gamma$  as  $n \rightarrow \infty$ . Moreover,  $F(\{\gamma\}) \geq c_\gamma$ .*
- (c) *Consider a forest  $\mathbf{u}$  composed of finitely many finite unlabeled rooted trees, and assume that some eigenvalue  $\gamma$  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  $c_\gamma = \lim_{n \rightarrow \infty} F_n(\{\gamma\})$  is strictly positive and hence  $\gamma$  is an atom of the limiting spectral distribution  $F$ .*

*Remark 4.2.* Simulations of the expected value the spectral distribution for various finite random trees are shown in Figure 1, Figure 2, and Figure 3. The large number of “spikes” in these figures is a reflection of parts (b) and (c) of Theorem 4.1 and the observation that the set of atoms of the limiting spectral distribution  $F$  is dense in the real line  $\mathbb{R}$  (resp. in the interval  $[-2\sqrt{2}, 2\sqrt{2}]$ ) for the random recursive tree, the linear preferential attachment tree and the uniform random tree (resp. for the random binary tree). To see these claims, first note that the adjacency matrix of

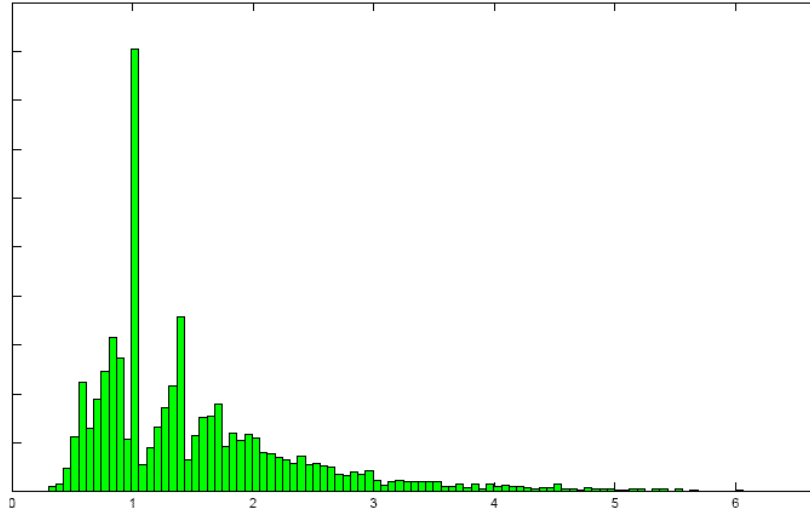


FIGURE 2. Empirical distribution for the positive eigenvalues of the preferential attachment tree ( $a = 0$ ) with 100 vertices, averaged over 200 realizations.

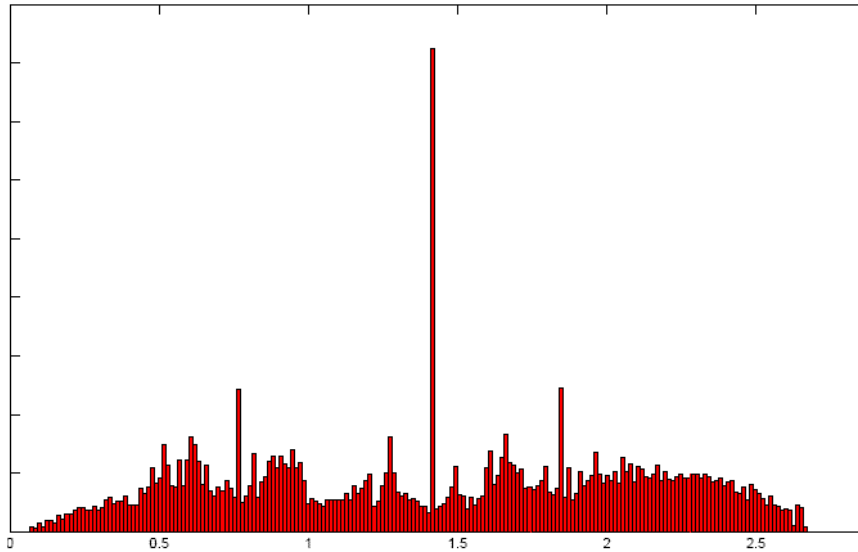


FIGURE 3. Empirical distribution for the positive eigenvalues of the random binary tree with 401 vertices, averaged over 100 realizations.

a forest  $\mathbf{u}$  has an eigenvalue  $\gamma$  with multiplicity greater than 1 if  $\gamma$  is an eigenvalue of more than one of the trees that make up  $\mathbf{u}$ . In particular, this condition holds if two or more of the trees that make up  $\mathbf{u}$  are equal to some common tree  $\mathbf{t}$ , and  $\gamma$  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, in the notation of Theorem 4.1, if  $\mathbf{u}$  is any forest of finite unlabeled rooted trees, then  $\mathbb{P}\{\mathcal{U} = \mathbf{u}\} > 0$ , and so any number  $\gamma$  that is the eigenvalue of the adjacency matrix of some 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. 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  $[-2\sqrt{2}, 2\sqrt{2}]$ , and because we can embed any binary tree into a complete binary tree of suitable height, we see that the limiting spectral measure in fact has this interval as its support.

*Remark 4.3.* 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.

*Remark 4.4.* 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 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 4.5.* The following result shows that the method of moments cannot be used for linear preferential attachment model with parameter  $a = 0$ . We omit the proof.

**Lemma 4.6.** *Let  $A_n$  be the adjacency matrix of the linear preferential attachment tree  $\mathcal{T}_n$  with  $a = 0$ . Then*

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

**4.2. The proportion of zero eigenvalues and maximal matchings.** Part (b) and (c) of Theorem 4.1 show that the limiting spectral distribution  $F$  will typically have many atoms. However, Theorem 4.1(c) provides a rather crude lower bounds on the mass of each atom. We obtain better lower bounds on the limiting proportion of zero eigenvalues in Subsection 6.3. 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 review 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 4.7.** *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  $\nu$  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.

**4.3. 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 4.8.** *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,*

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

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

**Theorem 4.9.** *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

$$(4.1) \quad \gamma_a := a + 2.$$

**Theorem 4.10.** *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  $\sigma$  and  $\tau$  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  $\sigma$  and  $\tau$ , 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.

We collect several simple facts in the following proposition.

**Proposition 5.1.** *In the above notation,*

- (a)  $d(P, Q) \leq (\frac{m}{n} - 1) \wedge 1$ .
- (b) *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  $B_k^\varepsilon$  is an  $n_k^\varepsilon \times n_k^\varepsilon$  principal sub-matrix of  $A_k$  with spectral distribution  $Q_k^\varepsilon$ . 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  $P_\infty = \lim_{\varepsilon \downarrow 0} Q_\infty^\varepsilon$ .*
- (c) *For each  $\gamma \in \mathbb{R}$ ,*

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

- (d) Let  $(A_k)_{k=1}^\infty, (B_k^\varepsilon)_{k=1}^\infty, m_k$  and  $n_k^\varepsilon$  be as in part (b). 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.* (a) By triangle inequality, it is enough to prove the assertion for  $n = m - 1$ . It follows immediately from the interlacing inequality (for example, see Lemma 3.3 in [Bai93]) that the Komogorov-Smirnov distance of  $P$  and  $Q$  (and hence the Lévy distance of  $P$  and  $Q$ ) is at most  $1/(m - 1)$ .

- (b) From Proposition 5.1(a),

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

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  $P_\infty$ .

Moreover,

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

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

(c) 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.

- (d) From part (c),

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

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 Proposition 5.1(b) finishes the proof.  $\square$

**Corollary 5.2.** *Consider a forest  $\mathbf{u}$  made up of finitely many finite unlabeled rooted trees, and assume that some eigenvalue  $\gamma$  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,  $\gamma$  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 5.1(c) 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(a).** 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  $N_n = \sum_{k=1}^K kR_{nk}$ . 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,

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

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  $\Xi_{nk}$  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  $\lambda_k$  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  $\Xi_k$ . The probability measure  $\lambda_k$  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

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

as  $n \rightarrow \infty$ .

An application of Proposition 5.1(b) completes the proof.

*Remark 5.3.* 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  $\ell$  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  $\ell$  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  $\ell$  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  $\Xi_k$  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  $\ell$ .

**5.3. Proof of Theorem 4.1(b).** The proof is almost identical to that of part (a) of the theorem in Subsection 5.2. Recall from that proof the constant  $K$ , the probabilities  $p_1, \dots, p_K$ , the probability distributions  $\lambda_k$ ,  $1 \leq k \leq K$ , on  $\mathbb{R}$ , and the random adjacency matrix  $B_n$  with distribution depending on  $K$  and  $n$ . Recall also that the probability measure  $\lambda_k$  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

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

as  $n \rightarrow \infty$ .

An application of Proposition 5.1(d) completes the proof.

**5.4. Proof of Theorem 4.1(c).** It follows from Corollary 5.2 that multiplicity of  $\gamma$  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  $\gamma$  (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(a),

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

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, Additional 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 for producing a maximal matching of a forest. It is a simplification of one due to Karp and Sipser [KS81] that is intended to find approximate maximal matchings of more general sparse graphs. The algorithm takes an initial forest and iteratively produces forests with the same set of vertices but smaller sets of edges while at the same time adding edges to a matching of the initial forest. We stress that a leaf of a forest is a vertex with degree one.

- Input a forest  $\mathbf{f}$  with vertices  $V(\mathbf{f})$  and edges  $E(\mathbf{f})$ .
- Initialize  $S \leftarrow \emptyset$ .
- While  $E(\mathbf{f}) \neq \emptyset$  do
  - \* Choose a leaf, say  $x$ , and let  $\{x, y\}$  be the unique edge in  $\mathbf{f}$  incident to  $x$ .
  - \* Set  $E(\mathbf{f}) \leftarrow \{e \in E(\mathbf{f}) : e \cap \{x, y\} = \emptyset\}$ , 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 forest, 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 no edge of the form  $\{y, w\}$  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 forest.  $\square$

Note that we are free to take any current leaf at each iteration of the “while” step of the algorithm. We start with some initial set of leaves and each iteration of the while step removes some leaves (by turning them into isolated vertices) as well as sometimes producing new leaves. We can therefore think of the leaves present after the completion of each while step as being labeled with the number of the step at which that vertex became a leaf, where the leaves in the initial forest are labeled with 0. We adopt the convention that in any iteration of the while step we take one of the current leaves with the lowest label.

Put  $i_0 = 0$  and define  $i_1, i_2, \dots$  inductively by setting  $i_{k+1}$  to be the number of iterations of the while step required until all of the leaves with labels at most  $i_k$  are turned into isolated vertices, where  $i_{k+1} = i_k$  if the forest after  $i_k$  iterations already consists of only isolated vertices. The numbers  $i_k$  are eventually constant and this final value is the cardinality of a maximal matching.

The iterations  $i_k + 1, \dots, i_{k+1}$  of the while step are of the following two types.

- Type I:** An iteration that removes all of the edges of the form  $\{y, z\}$ , where the vertex  $y$  is not a leaf with label at most  $i_k$  and there is a leaf  $x$  with label at most  $i_k$  such that  $\{y, x\}$  is an edge (so that  $y$  is at graph distance 1 from the leaves of the forest present after  $i_k$  iterations).
- Type II:** An iteration that removes an edge of the form  $\{y, z\}$  such that  $y$  and  $z$  are both leaves with label at most  $i_k$  (we say that  $\{y, z\}$  is an *isolated edge* in the forest present after  $i_k$  iterations).

Therefore, the cardinality of a maximal matching is the number of vertices that will be at graph distance 1 from the current leaves after  $i_k$  iterations of the while step for some  $k$  plus the number of edges in the initial forest that will eventually become isolated edges after  $i_k$  iterations of the while step for some  $k$ . We next introduce some notation to describe the sets of vertices and edges we have just characterized.

Write  $\mathbf{f}_k$ ,  $E_k(\mathbf{f})$ ,  $L_k(\mathbf{f})$ , and  $I_k(\mathbf{f})$ , respectively, for the forest, the set of edges, the set of leaves, and the set of isolated vertices after  $i_k$  iterations of the while step starting from the initial forest  $\mathbf{f}$ . Note that  $E_k(\mathbf{f})$  is obtained by removing all edges  $\{y, z\} \in E_{k-1}(\mathbf{f})$  such that there exists  $x \in L_{k-1}(\mathbf{f})$  with  $\{x, y\} \in E_{k-1}(\mathbf{f})$ . Equivalently,  $E_k(\mathbf{f})$  consists of exactly those edges  $\{u, v\} \in E_{k-1}(\mathbf{f})$  such that both vertices  $u$  and  $v$  are at graph distance at least 2 from  $L_{k-1}(\mathbf{f})$  in  $\mathbf{f}_{k-1}$ . This means that vertices that are distance 0 or 1 from  $L_{k-1}(\mathbf{f})$  in  $\mathbf{f}_{k-1}$  are isolated in  $\mathbf{f}_k$ , and vertices that are at graph distance 2 or greater from  $L_{k-1}(\mathbf{f})$  have degree in  $\mathbf{f}_k$  equal to the number of their neighbors in  $\mathbf{f}_{k-1}$  that are at graph distance 2 or greater from  $L_{k-1}(\mathbf{f})$ .

We further introduce new sets  $G_k(\mathbf{f})$ ,  $H_k(\mathbf{f})$  and  $J_k(\mathbf{f})$  as follows:

$$\begin{aligned} G_k(\mathbf{f}) &:= \{u \in L_k(\mathbf{f}) : \exists v \in L_k(\mathbf{f}) \text{ so that } \{u, v\} \in E_k(\mathbf{f})\}, \\ H_k(\mathbf{f}) &:= \{u \in V(\mathbf{f}) \setminus L_k(\mathbf{f}) : \exists v \in L_k(\mathbf{f}) \text{ so that } \{u, v\} \in E_k(\mathbf{f})\}, \\ J_k(\mathbf{f}) &:= (I_{k+1}(\mathbf{f}) \setminus I_k(\mathbf{f})) \setminus (G_k(\mathbf{f}) \cup H_k(\mathbf{f})). \end{aligned}$$

In words,  $G_k(\mathbf{f})$  is the set of leaves that are one of the two leaves of an isolated edge present after  $i_k$  iterations of the while step – these are the vertices that become isolated during iterations  $i_k + 1, \dots, i_{k+1}$  due to Type II steps,  $H_k(\mathbf{f})$  is the set of vertices that are graph distance 1 from the leaves after  $i_k$  iterations of the while step – these are the non-leaf vertices that become isolated during iterations  $i_k + 1, \dots, i_{k+1}$  due to Type I steps, and  $J_k(\mathbf{f})$  is the remaining set of vertices that become isolated during iterations  $i_k + 1, \dots, i_{k+1}$  (all due to Type I steps). Note that  $V(\mathbf{f})$  is the disjoint union of  $I_0(\mathbf{f})$  and  $G_k(\mathbf{f}), H_k(\mathbf{f}), J_k(\mathbf{f})$ ,  $k \geq 0$ , and so

$$\#V(\mathbf{f}) = \#I_0(\mathbf{f}) + \sum_{k=0}^{\infty} (\#G_k(\mathbf{f}) + \#H_k(\mathbf{f}) + \#J_k(\mathbf{f})).$$

Note that the forest  $\mathbf{f}_k$  can be obtained from the forest  $\mathbf{f}_{k-1}$  by deleting all the isolated edges of  $\mathbf{f}_{k-1}$  along with all the edges of  $\mathbf{f}_{k-1}$  that are incident to the vertices that are at a graph distance 1 from the leaves of  $\mathbf{f}_{k-1}$ . In particular,  $\mathbf{f}_k$  does not depend on the order in which we perform the leaf-removal operations (Type I and Type II) between iterations  $i_{k-1} + 1, \dots, i_k$  of the while step. Thus, by induction on  $k$ , it is easy to see that the sets of vertices  $G_k(\mathbf{f}), H_k(\mathbf{f})$  and  $J_k(\mathbf{f})$  are well defined in the sense that they do not depend on how we order the leaves of the initial forest  $\mathbf{f}_0$ .

Clearly, all the above objects can also be defined for an infinite forest  $\mathbf{f}$  such that every vertex is at a finite graph distance from a leaf, (that is, a vertex of degree one).

The discussion above leads immediately to the following result.

**Lemma 6.3.** *The cardinality of a maximal matching of a finite forest  $\mathbf{f}$  is*

$$M(\mathbf{f}) = \sum_{k=0}^{\infty} \#H_k(\mathbf{f}) + \frac{1}{2} \sum_{k=0}^{\infty} \#G_k(\mathbf{f}).$$

Consequently, the number of zero eigenvalues of the adjacency matrix of a finite tree  $\mathbf{t}$  is

$$\delta(\mathbf{t}) = \#V(\mathbf{t}) - 2M(\mathbf{t}) = \sum_{k=0}^{\infty} \#J_k(\mathbf{t}) - \sum_{k=0}^{\infty} \#H_k(\mathbf{t}).$$

*Example 6.4.* Consider the tree  $\mathbf{t}$  with vertices  $\{1, \dots, m\}$  and edges connecting successive integers. The cardinality of a maximal matching is obviously  $(m-1)/2$  when  $m$  is odd and  $m/2$  when  $m$  is even (so that  $\delta(\mathbf{t})$  is 1 when  $m$  is odd and 0 when  $m$  is even). There are four cases to consider in checking that this agrees with the formula of Lemma 6.3.

**Case I:**  $m$  is odd and  $(m-1)/2$  is odd ( $\Leftrightarrow m \equiv 3 \pmod{4}$ ).

Then,  $H_0(\mathbf{t}) = \{2, m-1\}$ ,  $H_1(\mathbf{t}) = \{4, m-3\}, \dots, H_{(m-3)/4}(\mathbf{t}) = \{(m+1)/2\}$ , all other  $H_k(\mathbf{t})$  are empty, and all  $G_k(\mathbf{t})$  are empty. The formula of Lemma 6.3 gives  $2 \times (m-3)/4 + 1 = (m-1)/2$ .

**Case II:**  $m$  is odd and  $(m-1)/2$  is even ( $\Leftrightarrow m \equiv 1 \pmod{4}$ ).

Then,  $H_0(\mathbf{t}) = \{2, m-1\}$ ,  $H_1(\mathbf{t}) = \{4, m-3\}$ ,  $\dots$ ,  $H_{(m-5)/4}(\mathbf{t}) = \{(m-1)/2, (m+3)/2\}$ , all other  $H_k(\mathbf{t})$  are empty, and all  $G_k(\mathbf{t})$  are empty. The formula of Lemma 6.3 gives  $2 \times ((m-5)/4 + 1) = (m-1)/2$ .

**Case III:**  $m$  is even and  $(m-2)/2$  is odd ( $\Leftrightarrow m \equiv 0 \pmod{4}$ ).

Then,  $H_0(\mathbf{t}) = \{2, m-1\}$ ,  $H_1(\mathbf{t}) = \{4, m-3\}$ ,  $\dots$ ,  $H_{(m-4)/4}(\mathbf{t}) = \{m/2, (m+2)/2\}$ , all other  $H_k(\mathbf{t})$  are empty, and all  $G_k(\mathbf{t})$  are empty. The formula of Lemma 6.3 gives  $2 \times ((m-4)/4 + 1) = m/2$ .

**Case IV:**  $m$  is even and  $(m-2)/2$  is even ( $\Leftrightarrow m \equiv 2 \pmod{4}$ ).

Then,  $H_0(\mathbf{t}) = \{2, m-1\}$ ,  $H_1(\mathbf{t}) = \{4, m-3\}$ ,  $\dots$ ,  $H_{(m-6)/4}(\mathbf{t}) = \{(m-2)/2, (m+4)/2\}$ , all other  $H_k(\mathbf{t})$  are empty,  $G_{(m-2)/4}(\mathbf{t}) = \{m/2, (m+2)/2\}$ , and all other  $G_k(\mathbf{t})$  are empty. The formula of Lemma 6.3 gives  $2 \times ((m-6)/4 + 1) + 1 = m/2$ .

**6.2. Maximal weighted matchings: Proof of Theorem 4.7.** 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  $\mathbb{P}\{\#\mathcal{T}^0 \leq K\}$  – 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  $p_k := \mathbb{P}\{\#\mathcal{T}^0 = k, \#\mathcal{T}^0 + \#\mathcal{T}^1 > K\}$ .

Moreover, the random probability measure  $\Xi_{nk}$  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

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

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 6.5 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  $\nu$ .

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  $\mu$  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  $\mathbb{P}\{\#\mathcal{T}^0 > K\}$  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  $\Xi_k$  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 6.5.** *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  $\Theta$ . Suppose that as  $i \rightarrow \infty$  the random variable  $L^i$  converges in distribution to  $\infty$  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  $i \rightarrow \infty$ .*

*Remark 6.6.* Let a sequence  $(\mathcal{T}_n)_{n=1}^\infty$  of random unlabeled rooted trees converge in the probability fringe sense to a random **sin**-tree  $\mathcal{T}$ . Let  $M_n$  be the number of vertices in  $\mathcal{T}_n$ . Consider the case when each edge-weight is identically one. 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,  $M_n^{-1}I(\mathcal{T}_n)$  also converges in distribution to a (model-dependent) constant  $\kappa_{\mathcal{T}} \geq 1/2$  as  $n \rightarrow \infty$ .

**6.3. Asymptotics of the number of zero eigenvalues.** If we combine Theorem 4.7 on the rescaled convergence of the total weight of a maximal weighted matching with Lemma 6.1 on the connection between the cardinality of a maximal matching and the number of zero eigenvalues of the adjacency matrix, then we get another proof of Theorem 4.1(b) on the convergence of  $F_n(\{\gamma\})$  in the special case when  $\gamma = 0$ . We now improve this result by using Lemma 6.3 to give a formula for the limit in terms of features of the limit **sin**-tree. We then show that how this formula may be used to get explicit lower bounds on the limit.

**Proposition 6.7.** *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  $F_n(\{0\})$  converges in distribution as  $n \rightarrow \infty$  to*

$$\sum_{k=0}^{\infty} \left( \mathbb{P}\{R \in J_k(\mathcal{T})\} - \mathbb{P}\{R \in H_k(\mathcal{T})\} \right).$$

*Proof.* In view of Theorem 4.1(b), it is enough to prove the convergence of  $F_n(\{0\})$  to the desired quantity in expectation. If  $V$  is a vertex chosen uniformly at random from  $\mathcal{T}_n$ , then, by Lemma 6.3 we can write  $\mathbb{E}[F_n(\{0\})] = M_n^{-1}\mathbb{E}[\delta(\mathcal{T}_n)]$  as

$$\begin{aligned} (6.1) \quad & \sum_{k=0}^{\infty} (M_n^{-1}\mathbb{E}[\#J_k(\mathcal{T}_n)] - M_n^{-1}\mathbb{E}[\#H_k(\mathcal{T}_n)]) \\ & = \sum_{k=0}^{\infty} (\mathbb{P}\{V \in J_k(\mathcal{T}_n)\} - \mathbb{P}\{V \in H_k(\mathcal{T}_n)\}). \end{aligned}$$

Given a tree  $\mathbf{t} \in \mathbb{T}$  with root  $\rho$  and a vertex  $v \in \mathbf{t}$ , write  $\mathcal{N}_k(v, \mathbf{t})$  for the subtree of  $\mathbf{t}$  induced by vertices that are at graph distance at most  $k$  from  $v$ . Note that whether or not a vertex  $v$  of  $\mathbf{t}$  belongs to the sets  $H_k(\mathbf{t})$  or  $J_k(\mathbf{t})$  can be determined by examining the neighborhood  $\mathcal{N}_{2k+4}(v, \mathbf{t})$ . Observe also that  $(t_0, t_1, \dots, t_h, *, *, \dots) \in \mathbb{T}_*^{\infty}$  is the decomposition of  $\mathbf{t}$  relative to  $\rho$  and  $v$ , then  $\mathcal{N}_k(v, \mathbf{t})$  can be reconstructed from  $(t_0, t_1, \dots, t_{k \wedge h})$ .

Recall that  $J_k(\mathcal{T}_n)$  and  $H_k(\mathcal{T}_n)$  are both subsets of  $I_{k+1}(\mathcal{T}_n) \setminus I_k(\mathcal{T}_n)$ , and so

$$|\mathbb{P}\{V \in J_k(\mathcal{T}_n)\} - \mathbb{P}\{V \in H_k(\mathcal{T}_n)\}| \leq \mathbb{P}\{V \in I_{k+1}(\mathcal{T}_n) \setminus I_k(\mathcal{T}_n)\}.$$

Moreover, for any nonnegative integer  $m$ ,

$$\begin{aligned} & \sum_{k=m}^{\infty} \mathbb{P}\{V \in I_{k+1}(\mathcal{T}_n) \setminus I_k(\mathcal{T}_n)\} \\ & = \mathbb{P}\{V \in V(\mathcal{T}_n) \setminus I_m(\mathcal{T}_n)\} \\ & \leq \mathbb{P}\{\text{the subtree of } \mathcal{T}_n \text{ below } V \text{ contains at least } 2m - 1 \text{ vertices}\} \\ & \rightarrow \mathbb{P}\{\#\mathcal{T}^0 \geq 2m - 1\} \end{aligned}$$

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 the analogous series involving the limiting **sin**-tree is also absolutely convergent.

Finally, it follows from the assumption of probability fringe convergence and our observations above about membership of  $H_k(\mathcal{T})$  and  $J_k(\mathcal{T})$  being locally determined that for each  $k \geq 1$ , the first  $k$  terms of the series (6.1) converge to the corresponding terms of the desired infinite series involving the limiting **sin**-tree.  $\square$

By construction, for any tree  $\mathbf{t}$ ,  $G_k(\mathbf{t}) \subseteq L_k(\mathbf{t})$  and  $L_k(\mathbf{t}) \setminus G_k(\mathbf{t}) \subseteq J_k(\mathbf{t})$ . Set  $K_k(\mathbf{t}) := J_k(\mathbf{t}) \setminus (L_k(\mathbf{t}) \setminus G_k(\mathbf{t}))$ . That is,  $K_k(\mathbf{t})$  consists of vertices that become isolated due to Type I steps during iterations  $i_k + 1, \dots, i_{k+1}$  of the Karp-Sipser algorithm but are of distance at least 2 from the leaves in the forest present after iteration  $i_k$ ; for example, if  $\mathbf{t}$  has vertices  $\{1, 2, 3, 4, 5\}$  and adjacent integers are joined by edges, then  $J_0(\mathbf{t}) = \{1, 3, 5\}$ ,  $L_0(\mathbf{t}) \setminus G_0(\mathbf{t}) = L_0(\mathbf{t}) = \{1, 5\}$ , and  $K_0(\mathbf{t}) = \{3\}$ . Note that for each  $v \in H_k(\mathbf{t})$  there exists  $u \in L_k(\mathbf{t}) \setminus G_k(\mathbf{t})$  such that  $\{v, u\} \in E_k(\mathbf{t})$ . Also, if  $v_1, v_2$  are distinct elements of  $H_k(\mathbf{t})$  and  $u_1, u_2 \in L_k(\mathbf{t}) \setminus G_k(\mathbf{t})$  are such that  $\{v_1, u_1\}, \{v_2, u_2\} \in E_k(\mathbf{t})$ , then  $u_1$  and  $u_2$  are also distinct. Consequently,  $\#L_k(\mathbf{t}) - \#G_k(\mathbf{t}) - \#H_k(\mathbf{t}) \geq 0$ . Applying this observation to  $\mathcal{T}_n$ , dividing by  $M_n$ , and taking the limit as  $n \rightarrow \infty$ , we deduce that the formula in Proposition 6.7 for the limit of  $F_n(\{0\})$  may be written as a sum over  $k$  of the sum of the two nonnegative terms  $\mathbb{P}\{R \in L_k(\mathcal{T})\} - \mathbb{P}\{R \in G_k(\mathcal{T})\} - \mathbb{P}\{R \in H_k(\mathcal{T})\}$  and  $\mathbb{P}\{R \in K_k(\mathcal{T})\}$ . We may give good lower bounds for the first few of these summands with relative ease.

We first find a lower bound on  $\mathbb{P}\{R \in L_0(\mathcal{T})\} - \mathbb{P}\{R \in G_0(\mathcal{T})\} - \mathbb{P}\{R \in H_0(\mathcal{T})\}$ . Note for any tree  $\mathbf{t}$  with 3 or more vertices that  $G_0(\mathbf{t}) = \emptyset$ . Observe also that

$$\begin{aligned} & \#L_0(\mathcal{T}_n) - \#H_0(\mathcal{T}_n) \\ &= \sum_{m=2}^{\infty} (m-1) \times \#\{u \in H_0(\mathcal{T}_n) : u \text{ is connected to exactly } m \text{ vertices in } L_0(\mathcal{T}_n)\} \\ &\geq \sum_{m=2}^{\infty} (m-1) \times \#\{u \in V(\mathcal{T}_n) : \text{the subtree below } u \text{ is an } m\text{-star}\}, \end{aligned}$$

where by a  $m$ -star we mean a unlabeled rooted tree with  $(m+1)$  vertices in which the root is connected to each of the other  $m$  vertices via an edge. Therefore,

$$\begin{aligned} \mathbb{P}\{R \in L_0(\mathcal{T})\} - \mathbb{P}\{R \in H_0(\mathcal{T})\} &= \lim_{n \rightarrow \infty} M_n^{-1} (\mathbb{E}[\#L_0(\mathcal{T}_n)] - \mathbb{E}[\#H_0(\mathcal{T}_n)]) \\ &\geq \sum_{m=2}^{\infty} (m-1) \times \mathbb{P}\{\mathcal{T}^0 \text{ is an } m\text{-star}\}. \end{aligned}$$

On the other hand, it is easy to check that  $\mathbb{P}\{R \in K_0(\mathcal{T})\} \geq \mathbb{P}\{\mathcal{T}^0 \in \mathbb{T}'' , \mathcal{T}^1 \in \mathbb{T}'\}$ , where  $\mathbb{T}' \subseteq \mathbb{T}$  is the set of finite unlabeled rooted trees for which the root has at least one child that has no children, and  $\mathbb{T}'' \subseteq \mathbb{T}$  is the set of finite unlabeled rooted trees for which the root has single child and that child in turn has at least one child that has no children.

As one might expect, finding good lower bounds on the terms  $\mathbb{P}\{R \in L_k(\mathcal{T})\} - \mathbb{P}\{R \in G_k(\mathcal{T})\} - \mathbb{P}\{R \in H_k(\mathcal{T})\}$  and  $\mathbb{P}\{R \in K_k(\mathcal{T})\}$  becomes increasingly difficult as  $k$  gets larger. However, we can still get crude lower bounds by computing the probability of appearance of special kinds of trees in the first few fringes in the

limiting **sin**-tree. For example,

$$\begin{aligned} & \mathbb{P}\{R \in L_k(\mathcal{T})\} - \mathbb{P}\{R \in G_k(\mathcal{T})\} - \mathbb{P}\{R \in H_k(\mathcal{T})\} \\ & \geq \mathbb{P}\{\mathcal{T}^0 \text{ is a complete binary tree of depth } (2k+1) \\ & \text{and } \mathcal{T}^i = \bullet \text{ for } 1 \leq i \leq 2k-2\}. \end{aligned}$$

where  $\bullet$  denotes the rooted tree with a single vertex. The proof follows along the same lines as the  $k=0$  case above. Furthermore,

$$\begin{aligned} & \mathbb{P}\{R \in K_k(\mathcal{T})\} \\ & \geq \mathbb{P}\{\mathcal{T}^0 \text{ is a path of length } (2k+2), \\ & \mathcal{T}^i = \bullet \text{ for } 1 \leq i \leq 2k, \\ & \text{and } \mathcal{T}^{2k+1} \text{ is a 1-star}\}. \end{aligned}$$

For the ensemble of linear preferential attachment trees with parameter  $a=0$ , it is well known (see, for example, [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  $n^{-1}\#L_0(\mathcal{T}_n)$  converges in distribution to  $2/3$ , and so  $\mathbb{P}\{R \in L_0(\mathcal{T})\} = 2/3$ . Hence,

$$\lim_{n \rightarrow \infty} F_n(\{0\}) \geq \mathbb{P}\{R \in L_0(\mathcal{T})\} - \mathbb{P}\{R \in H_0(\mathcal{T})\} \geq 2\mathbb{P}\{R \in L_0(\mathcal{T})\} - 1 = 1/3.$$

Now consider the ensemble of random recursive trees. Recall Construction 3.7(a). Let  $\xi_i, \xi'_i, i \geq 1$  and  $X$  be i.i.d. exponential random variables with rate 1. To get a lower bound on  $\lim_{n \rightarrow \infty} F_n(\{0\})$ , we may use the inequality  $\lim_{n \rightarrow \infty} F_n(\{0\}) \geq \sum_{m=2}^{\infty} (m-1) \times \mathbb{P}\{\mathcal{T}^0 \text{ is an } m\text{-star}\}$  where

$$\begin{aligned} \mathbb{P}\{\mathcal{T}^0 \text{ is an } m\text{-star}\} & \geq \mathbb{E} \left[ \mathbb{P} \left\{ \sum_{i=1}^m \xi_i \leq X, \sum_{i=1}^{m+1} \xi_i > X \mid X \right\} \prod_{i=1}^m \mathbb{P}\{\xi'_i > X \mid X\} \right] \\ & = \mathbb{E} \left[ \mathbb{P} \left\{ \sum_{i=1}^m \xi_i \leq X, \sum_{i=1}^{m+1} \xi_i > X \mid X \right\} e^{-mX} \right] \\ & = \mathbb{E} \left[ e^{-X} \frac{X^m}{m!} e^{-mX} \right] \\ & = \frac{1}{m!} \int_0^{\infty} x^m e^{-(m+2)x} dx \\ & = (m+2)^{-(m+1)}. \end{aligned}$$

For the uniform random trees, we can easily obtain lower bounds for various terms using the description of the fringes of the limiting **sin**-tree in terms of critical Poisson Galton-Watson trees. For example,

$$\begin{aligned} \mathbb{P}\{\mathcal{T}^0 \text{ is an } m\text{-star}\} & = \frac{e^{-1}}{m!} \times (e^{-1})^m = \frac{e^{-(m+1)}}{m!}, \\ \mathbb{P}\{\mathcal{T}^1 \in \mathbb{T}'\} & = 1 - \sum_{i=0}^{\infty} \frac{e^{-1}}{i!} \times (1 - e^{-1})^i = 1 - e^{-1}e^{1-e^{-1}}, \end{aligned}$$

and

$$\mathbb{P}\{\mathcal{T}^0 \in \mathbb{T}''\} = e^{-1} \times (1 - e^{-1}e^{1-e^{-1}}).$$



Therefore,

$$\begin{aligned} \lim_n F_n(\{0\}) &\geq \sum_{m=2}^{\infty} (m-1) \times \mathbb{P}\{\mathcal{T}^0 \text{ is a } m\text{-star}\} + \mathbb{P}\{\mathcal{T}^0 \in \mathbb{T}'', \mathcal{T}^1 \in \mathbb{T}'\} \\ &= e^{-1}(1 - (1 - e^{-1})e^{e^{-1}}) + e^{-1}(1 - e^{-1}e^{1-e^{-1}})^2. \end{aligned}$$

It's worth mentioning that [BG00] showed that the expected portion of zero eigenvalues of the empirical spectral distribution converges to  $2x_* - 1$  where  $x_*$  is the unique real root of the equation  $x = e^{-x}$ . But they do not prove any concentration result. In that sense, Proposition 6.7 in the special case of uniform random trees completes their result.

## 7. LARGEST EIGENVALUES: PROOF OF THEOREM 4.10

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 a geometric random variable with 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  $\ell$  is  $\ell$  – and the claimed distribution is well-known, see for example [Nor98], Chapter 2.

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  $\ell$  is  $\ell$ ). 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  $\ell$ . 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  $A^{\text{th}}$  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  $\tau$  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,  $\tau$  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  $S_\varepsilon$  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  $S_\varepsilon$  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  $B_\varepsilon$  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  $S_\varepsilon$  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  $S_\varepsilon$  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  $S_\varepsilon$  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  $\ell$ . 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 4.10.** 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  $\mathbf{X}_\infty$  such that  $\mathbf{X}_n$  converges in probability to  $\mathbf{X}_\infty$  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 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  $\mathbf{X}_\infty$ .  $\square$

## 8. AN EXAMPLE: THE RANDOM COMB

We have shown in previous sections that under quite general conditions the empirical spectral distributions of the adjacency matrices for many ensembles of random trees converge to a deterministic probability distribution as the number of vertices goes to infinity, and we have been able to deduce various properties of the limit. However, we have not identified the limit explicitly, except in highly regular, deterministic examples where the answer was already known. In this section we present an extremely simple ensemble of random trees, describe some of the ingredients that might go into identifying the limit, and conclude that even in this case ‘‘closed form’’ expressions for moments of the limit seem difficult to come by.

Consider the following construction of a random finite graph  $\mathcal{G}_n$  for a given positive integer  $n$ . The set of vertices of  $\mathcal{G}_n$  is  $\{1, \dots, 2n\}$ . Let  $\varepsilon_{n1}, \dots, \varepsilon_{nn}$  be

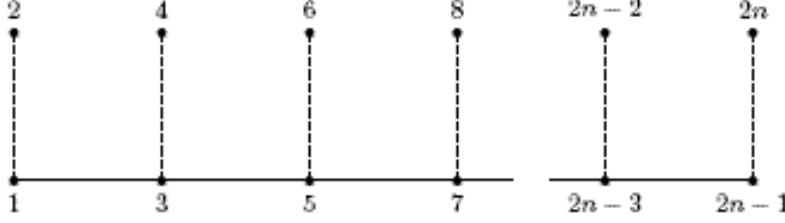


FIGURE 4. In the above graph  $\mathcal{G}_n$ , every horizontal edge always occurs whereas every vertical (dotted) edge occurs independently with probability  $1 - q$ . The comb tree  $\mathcal{T}_n$  is the unique connected component of  $\mathcal{G}_n$  that is not an isolated point.

independent, identically distributed random variables, with  $\mathbb{P}\{\varepsilon_{nk} = 0\} = \mathbb{P}\{\varepsilon_{nk} = 1\} = \frac{1}{2}$ . 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

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

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

Observe that, irrespective of the choice of the root, the sequence of random finite trees  $(\mathcal{T}_n)_{n=1}^\infty$  will not converge in the probability fringe sense because the cardinality of the subtree below a uniformly chosen point does not converge in distribution to a finite random variable. However, if we look at the empirical distribution of the subtrees spanned by the vertices within graph distance  $k$  of a vertex  $v$  as  $v$  ranges over the  $\mathcal{T}_n$ , then 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 further that these constants are the moments of a unique probability distribution. Thus, 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  $\delta_0$  is the unit point mass at 0. Since  $Z_n/n$  converges in probability to  $1/2$  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

$$H = \frac{1}{4}\delta_0 + \frac{3}{4}F,$$

where the probability measure  $F$  is the limit of the sequence  $F_n$ . In particular, to compute  $F$  it suffices to compute  $H$ .

Now define a random infinite graph  $\mathcal{G}$  with vertex set  $\mathbb{Z}$  as follows. Let  $\varepsilon_k, k \in \mathbb{Z}$ , be independent, identically distributed random variables, with  $\mathbb{P}\{\varepsilon_k = 0\} = \mathbb{P}\{\varepsilon_k = 1\} = \frac{1}{2}$ . 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  $m$ ,

$$\begin{aligned} & \int x^m H(dx) \\ &= \lim_{n \rightarrow \infty} \int x^m H_n(dx) \\ &= \lim_{n \rightarrow \infty} \mathbb{E} \left[ \int x^m H_n(dx) \right] \\ (8.1) \quad &= \lim_{n \rightarrow \infty} \mathbb{E} \left[ \frac{1}{2n} \text{tr } B_n^m \right] \\ &= \mathbb{E} \left[ \frac{1}{2} (B^m)_{11} + \frac{1}{2} (B^m)_{22} \right] \\ &= \frac{1}{2} \mathbb{E} \left[ \text{tr} \begin{pmatrix} (B^m)_{11} & (B^m)_{12} \\ (B^m)_{21} & (B^m)_{22} \end{pmatrix} \right]. \end{aligned}$$

Note that since  $B_{ij} = 0$  for  $|i - j| > 2$ , there is no problem defining  $B^m$ . Note also that these moments are zero when  $m$  is odd.

This observation suggests that we divide the matrix  $B$  into  $2 \times 2$  blocks with the rows (resp. columns) of the  $(i, j)$ <sup>th</sup> block indexed by  $\{2i + 1, 2i + 2\}$  (resp.  $\{2j + 1, 2j + 2\}$ ) and perform the matrix multiplications necessary to compute the powers of  $B$  blockwise. The resulting 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  $\Pi_x$  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  $2 \times 2$  zero matrix.

By analogy with the definition of  $\Pi_x$ , set

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

The following relations between these various matrices will be useful in computing the powers of the matrix  $B$ .

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

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  $r \geq 1$ ,

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

*Proof.* The first claim is obvious, because  $\Sigma^a$  either the  $2 \times 2$  identity matrix or  $\Sigma$  depending on whether  $a$  is even or odd.

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

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

where  $b_\ell$  is 0 or 1 depending on whether  $a_\ell$  is even or odd. This in turn may be rewritten as

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

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

The product is non-zero if and only if  $z_1 = z_2 = \cdots = z_r$ . This is equivalent to either  $b_1 = 0$  and  $b_2 = \cdots = b_r = 0$ , in which case the product is  $(\Pi_x)^r \Sigma^c = \Pi_x \Sigma^c$ , or  $b_1 = 1$  and  $b_2 = \cdots = b_r = 0$ , 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  $b_2 = \cdots = b_r = 0$  and  $b_1 + \cdots + b_{r+1}$  is even, in which case the trace is 1. This condition is equivalent to  $a_2, \dots, a_r$  and  $a_1 + \cdots + a_{r+1}$  all being even, and the result follows.  $\square$



Lemma 8.1 suggests the following construction. Consider the random (non-simple) graph  $\mathcal{Z}$  with vertex set  $\mathbb{Z}$ , where consecutive integers are connected by an edge and vertex  $k$  is connected to itself by a self-loop if and only if  $\epsilon_k = 1$ . Thus, with an appropriate re-labeling of vertices, the graph  $\mathcal{Z}$  is obtained from the infinite connected component of  $\mathcal{G}$  by replacing each “tooth” in that “comb” by a self-loop. As usual, a *path* in the graph  $\mathcal{Z}$  is a finite sequence of vertices (that is, a sequence of integers)  $(\mathbf{p}_0, \mathbf{p}_1, \dots, \mathbf{p}_\ell)$  such that there is an edge in  $\mathcal{Z}$  between  $\mathbf{p}_{i-1}$  and  $\mathbf{p}_i$  for each  $1 \leq i \leq \ell$ . We say that such a path connects the vertices  $x \in \mathbb{Z}$  and  $y \in \mathbb{Z}$  if  $\mathbf{p}_0 = x$  and  $\mathbf{p}_\ell = y$ , and we say that  $\ell$  is the *length* of the path. Define  $\mathbb{H}_\ell(x, y)$  to be the set of the paths in  $\mathcal{Z}$  of length  $\ell$  that connect  $x, y \in \mathbb{Z}$ . Then, (8.1) may be re-written as

$$(8.2) \quad \int x^m H(dx) = \frac{1}{2} \mathbb{E} \left[ \sum_{\mathbf{p} \in \mathbb{H}_m(0,0)} \text{tr}(\Xi_{\mathbf{p}_0, \mathbf{p}_1} \Xi_{\mathbf{p}_1, \mathbf{p}_2} \cdots \Xi_{\mathbf{p}_{m-1}, \mathbf{p}_m}) \right],$$

where

$$\Xi_{\mathbf{p}_{i-1}, \mathbf{p}_i} := \begin{cases} \Sigma, & \text{if } \mathbf{p}_i - \mathbf{p}_{i-1} = 0, \\ \Pi_x, & \text{if } \mathbf{p}_i - \mathbf{p}_{i-1} \in \{-1, +1\} \end{cases}$$

and we have used the observation that the sequences  $\{\epsilon_k\}_{k \in \mathbb{Z}}$  and  $\{\epsilon_{k+1}\}_{k \in \mathbb{Z}}$  have the same distribution.

Given a path  $(\mathbf{p}_0, \mathbf{p}_1, \dots, \mathbf{p}_\ell)$ , we say that the  $i^{\text{th}}$  step of the path is a *stutter* (resp. a *move*) if  $\mathbf{p}_i - \mathbf{p}_{i-1} = 0$  (resp.  $\mathbf{p}_i - \mathbf{p}_{i-1} \in \{-1, +1\}$ ). We can use the successive moves to split the stutters of a path into successive disjoint (possibly empty) runs: if the path  $\mathbf{p}$  has length  $\ell$  and  $\mathbf{s}(\mathbf{p})$  stutters, then there will be  $\ell - \mathbf{s}(\mathbf{p})$  moves and  $\ell - \mathbf{s}(\mathbf{p}) + 1$  runs of stutters. For  $1 \leq k \leq \ell - \mathbf{s}(\mathbf{p}) + 1$  write  $\mathbf{s}_k(\mathbf{p})$  for the number of stutters in the  $k^{\text{th}}$  such run, so that  $\mathbf{s}(\mathbf{p}) = \sum_{k=1}^{\ell - \mathbf{s}(\mathbf{p}) + 1} \mathbf{s}_k(\mathbf{p})$ .

For example, if  $\mathbf{p} = (0, 1, 2, 2, 2, 1, 0, 0) \in \mathbb{H}_7(0, 0)$ , then  $\mathbf{s}(\mathbf{p}) = 3$  and  $\mathbf{s}_1(\mathbf{p}) = 0$ ,  $\mathbf{s}_2(\mathbf{p}) = 0$ ,  $\mathbf{s}_3(\mathbf{p}) = 2$ ,  $\mathbf{s}_4(\mathbf{p}) = 0$ ,  $\mathbf{s}_5(\mathbf{p}) = 1$ .

For  $\ell \geq 1$  and  $0 \leq z < \ell$ , set

$$\mathbb{C}(\ell, z) := \left\{ \mathbf{p} \in \mathbb{H}_\ell(0, 0) : \mathbf{s}(\mathbf{p}) = z, \mathbf{s}_1(\mathbf{p}) + \mathbf{s}_{\ell - z + 1}(\mathbf{p}) \in 2\mathbb{Z}, \right. \\ \left. \mathbf{s}_k(\mathbf{p}) \in 2\mathbb{Z}, 1 < k < \ell - z + 1 \right\}$$

and put

$$\mathbb{C}(\ell, \ell) := \left\{ \mathbf{p} \in \mathbb{H}_\ell(0, 0) : \mathbf{s}(\mathbf{p}) = \ell, \mathbf{s}(\mathbf{p}) \in 2\mathbb{Z} \right\}.$$

Because of Lemma 8.1, we may rewrite (8.2) in terms of the expected sizes of the sets  $\mathbb{C}(m, z)$  as follows

$$\int x^m H(dx) = \frac{1}{2} \left[ 2\mathbb{E}[\#\mathbb{C}(m, m)] + \sum_{z=0}^{m-1} \mathbb{E}[\#\mathbb{C}(m, z)] \right].$$

Observe that  $\mathbb{E}[\#\mathbb{C}(m, m)] = \mathbb{P}\{\epsilon_0 = 1\} = \frac{1}{2}$  when  $m$  is even. For a positive integer  $m$  and  $0 \leq z < m$  we can write the set  $\mathbb{C}(m, z)$  as the disjoint union of the sets  $\mathbb{B}(m, z, j)$ ,  $0 \leq j \leq \frac{z}{2}$ , where

$$\mathbb{B}(m, z, j) := \left\{ \mathbf{p} \in \mathbb{H}_m(0, 0) : \mathbf{s}(\mathbf{p}) = m, \mathbf{s}_1(\mathbf{p}) + \mathbf{s}_{m - z + 1}(\mathbf{p}) = 2j, \right. \\ \left. \mathbf{s}_k(\mathbf{p}) \in 2\mathbb{Z}, 1 < k < m - z + 1 \right\}.$$

We can express the cardinalities of the sets  $\mathbb{B}(m, z, j)$  as follows.

**Case I.**  $j = 0$ . We have

$$\begin{aligned} \#\mathbf{B}(m, z, 0) &= \sum_{x, y = \pm 1} \#\{\mathbf{p} \in \mathbf{H}_{m-2}(x, y) : \mathbf{s}(\mathbf{p}) = z, \mathbf{s}_k(\mathbf{p}) \in 2\mathbb{Z}, 1 \leq k \leq m - z - 1\} \\ &= \sum_{x, y = \pm 1} \#\left\{\mathbf{p} \in \mathbf{H}_{m-z/2-2}(x, y) : \mathbf{s}(\mathbf{p}) = \frac{z}{2}\right\}. \end{aligned}$$

The first equality follows by considering the first and the last steps of a path in  $\mathbf{H}_m(0, 0)$ . The second inequality is then obtained by realizing that each path  $\mathbf{p}$  in the set

$$\{\mathbf{p} \in \mathbf{H}_{m-2}(x, y) : \mathbf{s}(\mathbf{p}) = z, \mathbf{s}_k(\mathbf{p}) \in 2\mathbb{Z}, 1 \leq k \leq m - z - 1\}$$

may be transformed to a path in the set

$$\left\{\mathbf{p} \in \mathbf{H}_{m-z/2-2}(x, y) : \mathbf{s}(\mathbf{p}) = \frac{z}{2}\right\}$$

by removing for  $1 \leq k \leq m - z - 1$  every second stutter in the  $k^{\text{th}}$  run of stutters of size  $\mathbf{s}_k(\mathbf{p}) \in 2\mathbb{Z}$ , and that this transformation is a bijection.

Because the distribution of the sequence  $\{\epsilon_k\}_{k \in \mathbb{Z}}$  is invariant under the group of transformations of the index set  $\mathbb{Z}$  generated by the transformations  $k \mapsto k \pm 1$  and  $k \mapsto -k$ , the distribution of the random graph  $\mathcal{Z}$  inherits similar invariance and we have

$$(8.3) \quad \begin{aligned} \mathbb{E}[\#\mathbf{B}(m, z, 0)] &= 2\mathbb{E}\left[\#\left\{\mathbf{p} \in \mathbf{H}_{m-z/2-2}(0, 0) : \mathbf{s}(\mathbf{p}) = \frac{z}{2}\right\}\right] \\ &\quad + 2\mathbb{E}\left[\#\left\{\mathbf{p} \in \mathbf{H}_{m-z/2-2}(0, 2) : \mathbf{s}(\mathbf{p}) = \frac{z}{2}\right\}\right]. \end{aligned}$$

**Case II.**  $0 < j \leq \frac{z}{2}$ . By reasoning similar to the above, we deduce that

$$\begin{aligned} \#\mathbf{B}(m, z, j) &= \sum_{x, y = \pm 1} \#\{\mathbf{p} \in \mathbf{H}_{m-2j-2}(x, y) : \mathbf{s}(\mathbf{p}) = z - 2j, \\ &\quad \mathbf{s}_k(\mathbf{p}) \in 2\mathbb{Z}, 1 \leq k \leq m - z - 1\} \mathbf{1}_{\epsilon_0=1} \\ &= \sum_{x, y = \pm 1} \#\left\{\mathbf{p} \in \mathbf{H}_{m-2j-2-(z-2j)/2}(x, y) : \mathbf{s}(\mathbf{p}) = \frac{(z-2j)}{2}\right\} \mathbf{1}_{\epsilon_0=1}. \end{aligned}$$

The first equality is obtained by considering both the first and the last steps of a path in  $\mathbf{H}_m(0, 0)$  that are moves (that is, are not stutters). Taking expectation and again using the symmetry properties of the distribution of  $\mathcal{Z}$ , we have

$$(8.4) \quad \begin{aligned} \mathbb{E}[\#\mathbf{B}(m, z, j)] &= \mathbb{E}\left[\#\left\{\mathbf{p} \in \mathbf{H}_{m-z/2-j-2}(0, 0) : \mathbf{s}(\mathbf{p}) = \frac{z}{2} - j\right\} \mid \epsilon_1 = 1\right] \\ &\quad + \mathbb{E}\left[\#\left\{\mathbf{p} \in \mathbf{H}_{m-\frac{z}{2}-j-2}(0, 2) : \mathbf{s}(\mathbf{p}) = \frac{z}{2} - j\right\} \mid \epsilon_1 = 1\right]. \end{aligned}$$

Combining these identities, we obtain for  $m$  even that

$$\begin{aligned}
(8.5) \quad & \int x^m H(dx) \\
&= \frac{1}{2} \left( 1 + 2 \sum_{z=0}^{m-1} \mathbb{E} \left[ \# \left\{ \mathbf{p} \in \mathbb{H}_{m-z/2-2}(0,0) : \mathbf{s}(\mathbf{p}) = \frac{z}{2} \right\} \right] \right. \\
&+ 2 \sum_{z=0}^{m-1} \mathbb{E} \left[ \# \left\{ \mathbf{p} \in \mathbb{H}_{m-z/2-2}(0,2) : \mathbf{s}(\mathbf{p}) = \frac{z}{2} \right\} \right] \\
&+ \sum_{z=0}^{m-1} \sum_{1 \leq j \leq z/2} \mathbb{E} \left[ \# \left\{ \mathbf{p} \in \mathbb{H}_{m-z/2-j-2}(0,0) : \mathbf{s}(\mathbf{p}) = \frac{z}{2} - j \right\} \mid \epsilon_1 = 1 \right] \\
&\left. + \sum_{z=0}^{m-1} \sum_{1 \leq j \leq z/2} \mathbb{E} \left[ \# \left\{ \mathbf{p} \in \mathbb{H}_{m-z/2-j-2}(0,2) : \mathbf{s}(\mathbf{p}) = \frac{z}{2} - j \right\} \mid \epsilon_1 = 1 \right] \right).
\end{aligned}$$

It is clear (8.5) from that moments of the limit of the empirical spectral distribution of the adjacency matrix of the random comb model can be expressed in terms of quantities involving the random walk on the random graph  $\mathcal{Z}$ . However, since computing these quantities appears to involve taking an expectation with respect to the random graph  $\mathcal{Z}$  (or at least with respect to a neighborhood of 0 that grows with the order  $m$  of the moment), it is not clear how much of an advance this observation is over the formula (8.1). We show in the next subsection that this computation for a ‘random walk in random environment’ can be reformulated as a computation for a random walk on a deterministic graph, thereby removing the explicit expectation over the random environment.

**8.1. Random walk on  $\mathbb{Z}$  with random self-loops and the lamplighter group.** The lamplighter group is the wreath product of the group  $\mathbb{Z}/2\mathbb{Z}$  of integers modulo 2 and the group of integers  $\mathbb{Z}$ ; that is, it is the semi-direct product of the infinite product group  $(\mathbb{Z}/2\mathbb{Z})^{\mathbb{Z}}$  and  $\mathbb{Z}$ , where  $\mathbb{Z}$  is thought of as a group of automorphisms on  $(\mathbb{Z}/2\mathbb{Z})^{\mathbb{Z}}$  by associating  $k \in \mathbb{Z}$  with the automorphism of  $(\mathbb{Z}/2\mathbb{Z})^{\mathbb{Z}}$  that sends the sequence  $(c_i)_{i \in \mathbb{Z}}$  to the sequence  $(c_{i-k})_{i \in \mathbb{Z}}$ .

More concretely, the lamplighter group may be taken to be  $\{0,1\}^{\mathbb{Z}} \times \mathbb{Z}$  as a set. Elements can be thought of as consisting of two components: a doubly infinite sequence of street lamps that are each either ‘off’ (0) or ‘on’ (1) and an integer giving the location of the lamplighter. The group operation is defined by

$$((c_i)_{i \in \mathbb{Z}}, k) \cdot ((d_j)_{j \in \mathbb{Z}}, \ell) = ((c_h + d_{h-k})_{h \in \mathbb{Z}}, k + \ell),$$

where the additions in  $\{0,1\}$  are, of course, performed modulo 2. The identity element is  $\mathbf{id} := (\mathbf{0}, 0)$ , where  $\mathbf{0} \in \{0,1\}^{\mathbb{Z}}$  is the sequence consisting of all zeros.

Write  $\mathbf{e}_n \in \{0,1\}^{\mathbb{Z}}$  for the sequence that has 1 in the  $n^{\text{th}}$  coordinate and 0 elsewhere. Set  $a := (\mathbf{e}_0, 0)$  and  $t := (\mathbf{0}, 1)$ . Observe that

$$((c_i)_{i \in \mathbb{Z}}, k) \cdot a = ((c'_i)_{i \in \mathbb{Z}}, k),$$

where  $c'_i = c_i$  except for  $i = k$ , for which  $c'_k = c_k + 1 \pmod{2}$ , and

$$((c_i)_{i \in \mathbb{Z}}, k) \cdot t = ((c_i)_{i \in \mathbb{Z}}, k + 1).$$

Thus, right multiplication by the element  $a$  corresponds to the lamplighter flipping the state of the lamp at his current location and staying where he is, while right multiplication by the element  $t$  corresponds to the lamplighter leaving the configuration of lamps unchanged and taking one step to the right.

The elements  $a$  and  $t$  generate the lamplighter group and we consider the Cayley graph corresponding to the symmetric set of generators  $\{t, t^{-1}, a\}$ . This is an infinite 3-regular graph.

Given  $\gamma \in \mathbb{R}$ , define  $A_\gamma : (\{0, 1\}^{\mathbb{Z}} \times \mathbb{Z})^2 \rightarrow \mathbb{R}$  by

$$A_\gamma((\mathbf{u}, x), (\mathbf{v}, y)) = \begin{cases} 1, & \text{if } (\mathbf{u}, x)^{-1}(\mathbf{v}, y) \in \{t, t^{-1}\}, \\ \gamma, & \text{if } (\mathbf{u}, x)^{-1}(\mathbf{v}, y) = a, \\ 0, & \text{otherwise.} \end{cases}$$

That is,  $A_\gamma$  is the adjacency matrix of the Cayley graph of the lamplighter group modified so that edges corresponding to the generator  $a = a^{-1}$  have “weight”  $\gamma$  rather than 1.

**Lemma 8.2.** *For any  $x \in \mathbb{Z}$  and non-negative integers  $0 \leq k \leq m$*

$$\mathbb{E}\left[\#\{\mathbf{p} \in \mathbf{H}_m(0, x) : \mathbf{s}(\mathbf{p}) = k\}\right]$$

*is the coefficient of  $\alpha^k$  in*

$$\sum_{\ell=0}^m \binom{m}{\ell} (\alpha/2)^{m-\ell} A_{\alpha/2}^\ell(\mathbf{id}, (\mathbf{0}, x))$$

*and*

$$\mathbb{E}\left[\#\{\mathbf{p} \in \mathbf{H}_m(0, x) : \mathbf{s}(\mathbf{p}) = k\} \mid \epsilon_1 = 1\right]$$

*is the coefficient of  $\alpha^k$*

$$\sum_{\ell=0}^m \binom{m}{\ell} (\alpha/2)^{m-\ell} \left[ A_{\alpha/2}^\ell(\mathbf{id}, (\mathbf{0}, x)) + A_{\alpha/2}^\ell(\mathbf{id}, (\mathbf{e}_1, x)) \right].$$

*Proof.* We will only establish the first claim. The proof of the second claim is similar and we leave it to the reader.

For  $\alpha \in \mathbb{R}$  define a random doubly infinite tridiagonal matrix  $T_\alpha$  with rows and columns indexed by  $\mathbb{Z}$  by setting  $T_\alpha(i, i+1) = T_\alpha(i-1, i) = 1$  and  $T_\alpha(i, i) = \alpha \epsilon_i$  for all  $i \in \mathbb{Z}$ . Note that  $T_\alpha$  is obtained from the adjacency matrix of the random graph  $\mathcal{Z}$  by replacing each entry that corresponds to a self-loop (that is, a one on the diagonal) by  $\alpha$ . It is clear that

$$\mathbb{E}\left[\#\{\mathbf{p} \in \mathbf{H}_m(0, x) : \mathbf{s}(\mathbf{p}) = k\}\right]$$

is the coefficient of  $\alpha^k$  in  $\mathbb{E}[T_\alpha^m(0, x)]$ .

Set  $\overline{T}_\alpha = T_\alpha - \frac{\alpha}{2}I$ , where  $I$  is the identity matrix. Then,

$$\mathbb{E}[T_\alpha^m(0, x)] = \sum_{\ell=0}^m \binom{m}{\ell} (\alpha/2)^{m-\ell} \mathbb{E}[\overline{T}_\alpha^\ell(0, x)].$$

We claim that

$$\mathbb{E}[\overline{T}_\alpha^\ell(0, x)] = A_{\alpha/2}^\ell(\mathbf{id}, (\mathbf{0}, x)).$$

To see this, note first that if we define a graph  $\overline{\mathcal{Z}}$  with vertex set  $\mathbb{Z}$  by placing edges between successive integers and self-loops at each integer, then  $\overline{T}_\alpha$  is the adjacency

matrix of  $\overline{\mathcal{Z}}$  except that the diagonal entries (which are all ones) have been replaced by i.i.d. symmetric random variables with values in the set  $\{\pm \frac{\alpha}{2}\}$ . Therefore, we can expand  $\overline{T}_\alpha^\ell(0, x)$  out as a sum of products of entries of  $\overline{T}_\alpha(0, x)$ , with each product in the sum corresponding to a path of length  $\ell$  from 0 to  $x$  in the graph  $\overline{\mathcal{Z}}$ . In order that the expectation of the product corresponding to such a path  $\mathbf{p}$  is non-zero, it is necessary that any self-loop in  $\overline{\mathcal{Z}}$  is traversed an even number of times, in which case the expectation is  $(\frac{\alpha}{2})^{\mathbf{s}(\mathbf{p})}$ , where, as above,  $\mathbf{s}(\mathbf{p})$  is the number of stutters in the path. Such paths can be associated bijectively with paths in the Cayley graph of the lamplighter group where the lamplighter goes from position 0 to position  $x$  in  $\ell$  steps in such a way that if every lamp is “off” at the beginning of his journey, then every lamp is also “off” at the end of his journey and a total of  $\mathbf{s}(\mathbf{p})$  flips of lamp states have been made. We can expand  $A_{\alpha/2}^\ell(\mathbf{id}, (\mathbf{0}, x))$  as a sum of products with each product corresponding to one of the latter collection of paths in the Cayley graph, and the value of such a product for the path in the Cayley graph associated with the path  $\mathbf{p}$  in  $\overline{\mathcal{Z}}$  is  $(\frac{\alpha}{2})^{\mathbf{s}(\mathbf{p})}$ .  $\square$

In view of (8.5) and Lemma 8.2, the moments of the probability measure  $H$  can be recovered from powers of the matrices  $A_\gamma$ . Computing the  $\ell^{\text{th}}$  power of the matrix  $A_\gamma$  is equivalent to computing the  $\ell$ -fold convolution of the measure  $\mu$  on the lamplighter group that assigns mass  $\gamma$  to the generator  $a$  and mass 1 to each of the generators  $t$  and  $t^{-1}$ . One might therefore hope to use the representation theory of the lamplighter group (see, for example, [BW05, LNW08, ST08]) to compute the relevant quantities. For example, Corollary 3.5 of [ST08] gives a formula for the entries of  $A_\gamma^\ell$  when  $\gamma = 1$ . The treatment in [ST08] is for the finite lamplighter group in which the lamps are indexed by the group  $\mathbb{Z}/n\mathbb{Z}$  instead of  $\mathbb{Z}$ , but when  $n$  is large enough, say  $n > 2\ell$ , the random walker cannot feel this finiteness within  $\ell$  steps and so the formula still applies. Unfortunately, the formula involves a sum over all subsets of  $\mathbb{Z}/n\mathbb{Z}$  and we are effectively led back to performing a computation for each possible realization of the random graph  $\mathcal{Z}$  and taking the expectation over all such realizations!

**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)
- [Bai93] Z. D. Bai, *Convergence rate of expected spectral distributions of large random matrices. I. Wigner matrices*, Ann. Probab. **21** (1993), no. 2, 625–648. MR 1217559 (95a:60039)
- [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)
- [BL10a] C. Bordenave and M. Lelarge, *Resolvent of large random graphs*, Random Structures & Algorithms **37** (2010), no. 3, 332–352.
- [BL10b] Charles Bordenave and Marc Lelarge, *The rank of diluted random graphs*, Proceedings of the Twenty-First Annual ACM-SIAM Symposium on Discrete Algorithms (Philadelphia, PA, USA), SODA '10, Society for Industrial and Applied Mathematics, 2010, pp. 1389–1402.
- [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)
- [BW05] Laurent Bartholdi and Wolfgang Woess, *Spectral computations on lamplighter groups and Diestel-Leader graphs*, J. Fourier Anal. Appl. **11** (2005), no. 2, 175–202. MR 2131635 (2006e:20052)
- [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.
- [LNW08] Franz Lehner, Markus Neuhauser, and Wolfgang Woess, *On the spectrum of lamplighter groups and percolation clusters*, Math. Ann. **342** (2008), no. 1, 69–89. MR 2415315 (2009d:60329)
- [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)
- [Nor98] J. R. Norris, *Markov chains*, Cambridge Series in Statistical and Probabilistic Mathematics, vol. 2, Cambridge University Press, Cambridge, 1998, Reprint of 1997 original. MR 1600720 (99c:60144)
- [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. Imovir. Mat. Stat. (1994), no. 51, 1–29. MR MR1445048 (97k:60027)
- [ST08] Fabio Scarabotti and Filippo Tolli, *Harmonic analysis of finite lamplighter random walks*, J. Dyn. Control Syst. **14** (2008), no. 2, 251–282. MR 2390215 (2009e:43009)

- [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.