Spectra of Random trees

Shankar Bhamidi, Steve Evans * and Arnab Sen

University of California
Department of Statistics
367 Evans Hall # 3860
Berkeley CA 94720-3860

March 9, 2008

Abstract

We analyze the spectral distribution of the adjacency matrix and the graph Laplacian for a wide variety of random trees. Using soft arguments which seem to be applicable in a wide variety of settings, we show that the empirical spectral distribution for a number of random tree models, converges to a constant (model dependent) distribution. We also analyze the kernel of the spectrum and prove asymptotic convergence to limit constants for the kernel of the spectrum. We then go on to analyze the joint distribution of the maximal eigenvalues of the adjacency matrix in the linear preferential attachment model (with parameter a). We first show that the for any fixed k, the maximal k degrees rescaled properly converge in distribution. Using results of Frieze [20] and Chung et al [14], this implies that the k largest eigenvalues rescaled by $n^{1/2\gamma_a}$ converge in distribution, where γ_a is the Malthusian rate of growth parameter for the associated continuous time branching process.

Key words.

MSC2000 subject classification. 60C05, 05C80, 90B15.

*NSFxxxx

1 Introduction

The theory of random matrices is one of the corner stones of modern probability, finding applications in such diverse fields as random partitions, high dimensional statistical analysis, nuclear physics, signal processing and wireless communication, quantum percolation and Free probability, and in the modern day world, in the operation of search engines such as Google. There seem to be mysterious connections between random matrix theory and various probabilistic constructs arising in classical statistical mechanics and integrable systems. The challenging mathematical problems that have arisen in this area have attracted many of the smartest physicists and mathematicians, thus giving rise to one of the most active areas of current mathematical research.

The starting point of many studies carried out in this field is a random (conjugate, in the complex case) symmetric $N \times N$ matrix A_N, either real or complex. The randomness in the matrix arises in some context dependent fashion, for e.g. one of the most popular models is the GUE model, where we assume that the entries are independent and identically distributed gaussian distributed random variables. Of interest are spectral statistics, such as spectral distribution, namely the empirical distribution of the spectra or the behavior of the maximal eigenvalues or the eigenvectors.

Of particular relevance in our context are the adjacency matrices of various models of random graphs. In the modern theory of graphs, matrices arise in many different guises, for e.g. as mentioned above, the adjacency matrix $A(G)$ of the graph or the Laplacian matrix (which is precisely defined below). In the recent past, with the availability of large-scale data on many real-world networks, there has been an explosion in the number of network models used to model real-world data. Other than the classical Erdos-Renyii random graph model, models such as the preferential attachment model have garnered significant amount of interest in the computer science and mathematical physics community. One way to explore the graph theoretic properties of any one of these network models is through their adjacency matrices. For example it is known (see e.g. [16]) that the Laplacian carries an enormous amount of information about the properties of the graph, ranging from the diffusive behavior of random walks on the graph, to the flow carrying properties of the graphs in question.

Our Work: In this study, we largely specialize to various models of random trees, although we make significant headway in the spectral analysis of one of the most popular modern models of network growth, the preferential attachment model. Random tree models arise in many varied contexts, ranging from the analysis of computer algorithms, to phylogenetics or the tree of life. Many of the various preferential attachment schemes are also largely random models of growing trees, see e.g. [8] for a survey of some of the more popular schemes. In the context of random matrix theory, this suggests the following question

> *What is the general math theory that could allow us to analyze various spectral statistics of interest applicable to a wide variety of models?*

In this study we develop such general theory and find some of the sharpest known results for the asymptotics of the maximal eigenvalues for some of these models. What is surprising is that in the analysis of the spectral distribution, the method of moments, one of the most popular work horses of random matrix theory fails, however we are still able to show convergence of the spectral distribution to a limiting non random distribution. At a mathematical level, many surprising connections appear between the theory of continued fractions and the analysis of the spectral distribution. Finally we are able to analyze the kernel of the spectrum (namely the number of zero eigenvalues, of crucial importance in areas such as quantum percolation see e.g. [6]) via its close relation to the graph theoretical constructs such as maximal matchings on trees. We also get lower bounds for the limit constants for the kernel of the spectrum, using an analysis of the performance of the Karp-Sipser algorithm on such trees. All in all, the study seems to be give rise to an extremely interesting blend of rigorous probability theory, including branching process theory and recursive distributional equations, algorithmic computer science, graph theory and classical mathematical constructs such as continued fractions.

Related work:

Random matrices seem to have arisen in many different, disparate areas of physics. For a classical references see [23]. As mentioned above our motivations are different from Dyson and Wigner's GUE theory. For a statistical physics view of how the study of such spectra gives information on the properties of disordered conductors, see [19] and the references therein. For a more recent work on the efficacy of the spectra of the adjacency matrix to differentiate between different graph topologies, especially in the phylogenetics setting, see e.g. [22]). There has been quite an extensive work on kernels of the spectra in the tree context, for e.g. see [6], [7] and the references therein.

The rigorous work most closely relevant to our setting are the series of contributions made by Fan Chung, L. Lu, Van Vu and co-workers see [15], [14] as well as by Frieze et al [20]. Quite closely related to our results for the empirical spectral distribution function are the recent works by Bordenave see [10]. They also carry out a much more detailed discussion of the issues involved in finding explicit properties of the limiting spectral measures. In the Euclidean context also see [11].

At a deterministic level, as mentioned above, there is a very large body of work exploring Interesting connections between the spectrum of the adjacency matrix and the actual topological properties of the trees or graphs in question. See [16] for an account of how the spectral profile determines an enormous amount of useful information about the graph.

1.1 Organization of the paper

In Section 2 we shall define the various graph theoretic definitions we shall use as well as the random tree models we shall pay particular attention to as well as the convergence of probability measures on tree space. Section 3 contains the results for various models. Section 4 contains the proofs of all the main results. We finally conclude in Section 5 we conclude with a wide ranging discussion on further problems and comments on how these results tie up with known results.

2 Notation and tree models

We shall use the following graph theoretical notions.

Matchings: A matching on a graph is a collection of edges which do not share any common node. If edges have weights then the weight of a matching is the sum of the weights of the edges in the matching. By default unless other specified, each edge has weight 1.

Stieltjes Transform: For any distribution function F on $\mathbb{R}$, define the Stieltjes transform as

$$s(z) = \int_{\mathbb{R}} \frac{1}{x - z} dF(x) \quad (z \in \mathbb{C}, \text{imag}(z) > 0)$$

Graph Laplacian:

Given a graph with adjacency matrix A, define the graph laplacian $\mathcal{L}(G)$ is defined as

$$\mathcal{L}(G) = I - D^{-1/2} A(G) D^{-1/2}$$

where D denotes the diagonal matrix with the diagonal entries being the degrees of the vertex, and I denoting the identity matrix.

2.1 Random tree models

There are an enormous number of random tree models, occurring both in the algorithmic community as well as in modeling real world networks community, see [2] and [8] for a description of some of the more popular models. Although our methods are quite general, to exhibit the typical method of computations we shall specialize to the following models (although we shall explicitly mention which results carry over more generally).

1. **Random Recursive tree:** This is the simplest model of constructing a rooted tree sequentially via the addition of a new node at each stage. Start with a single node (the root) at time 1 and (labeling nodes sequentially), after having constructed a tree $\mathcal{T}_n$ on $[n] := \{1, 2, \ldots, n\}$, the tree on $n + 1$ nodes is constructed via the addition of an edge from node $n + 1$ to a node chosen uniformly among the nodes $1, 2, \ldots, n$. See [26] for a survey of some of the properties of the random recursive tree.

2. **Linear Preferential Attachment model:** This is another model of sequential construction. As before we start with a single node (the root) at time 1. We shall think of the edges as directed away from the root (so that for all nodes other than the root, the degree d_v equals $D(v) + 1$, where $D(v)$ is the out-degree of the node v). After having constructed the tree on n vertices let $D(v, n)$ be the out degree of node $v \in [n]$ at time n. We construct a tree on $n + 1$ nodes via the addition of an edge between the new node $n+1$ and node in $[n]$ with probability proportional to $D(v, n) + 1 + a$ where a is a parameter of the process. There is an enormous amount of recent literature on this model. See [9], [18], [8] for relevant references.

3. **Uniform Random unordered labeled tree:** This is the uniform distribution on all rooted unordered trees on n labeled nodes, with one of the nodes selected as the root uniformly at random.

4. **Random Binary tree on n vertices:** There are various models of random binary trees. The one we shall consider is the following model of a randomly growing binary tree: At time 1 start with a root which has two leaves. Then at every stage a leaf is selected uniformly at random and "expanded" as two leaves.

2.2 Preferential attachment networks

Along with the random tree models, one of the most studied random network models, especially as a model for various real world networks has been the Preferential Attachment model (see e.g.

[1]). We shall consider only the simplest model although there are quite general extensions of the above model (see [17] where a general class of such models were introduced by Cooper and Frieze to to model the evolution of the web-graph). The model involves an integer valued parameter $m \geq 1$

(A) Start with a single node with m self loops.

(B) After creating the network with n nodes, a new node is "born", with m edges. It attaches each of these edges sequentially, with probability proportional to the present degree of the node in the network, the degrees of all the nodes are updated and the process is repeated till all the m edges have been added to the network and repeat the process. This gives the network on $n+1$ nodes.

We should mention a brief caveat that, once the first edge of the $n+1$ node is added to the network, then it is assumed to be part of the network so that subsequently the other edges could chose this node to add their edges. However, as the size of the network grows large this shall happen with vanishingly small property. The advantage of defining the model as above is because of the following connection between the $m = 1$ model, namely model 2 defined above and the $m \geq 1$ model.

Lemma 1 *Fix $n \geq 1$ and the parameter $m > 1$. Construct a random tree according to model (2) above on $n \cdot m$ nodes. Now merge nodes $1, 2, \ldots, m$ and calling this node 1, nodes $m+1, \ldots, 2m$ calling this node 2 and so on thus getting a network on n nodes. Then this random network has the same distribution as the model specified above, with parameter m.*

Proof: Obvious

2.3 Convergence on Tree space

We arrive at our first non-trivial definition, regarding the structure theory of infinite trees and a notion of convergence of large finite trees to infinite trees in some sense. See [2], [3] for the foundations of this theory as well as a wide variety of examples. We shall look at locally finite infinite trees with *one end* or *sin trees* (for single infinite path trees). This concept has been studied quite extensively in modern probability theory, see [2] for a vast array of examples.

Let $\mathbb{T}$ be the countable space of all finite rooted trees with the discrete topology. Consider the product space $\mathbb{T}^\infty = \{(t_0, t_1, t_2, \ldots) : t_i \in \mathbb{T}\}$ with the obvious induced topology. Note that any $\mathbf{t} = (t_0, t_2, \ldots) \in \mathbb{T}^\infty$ can be thought of as a tree with one end or a single infinite path via the identification of the roots of t_i with $i \in \mathbb{Z}^+$, and considering $\mathbb{Z}^+$ as the single infinite path. For $\mathbf{t} \in \mathbb{T}^\infty$, let us define the projection map onto the first k coordinates as f_k, namely $f_k(0, \mathbf{t}) = (t_0, t_2, \ldots t_{k-1})$, where the 0 in the parenthesis is used to emphasize the neighborhood "upto fixed distance k on the infinite path **from the root**".

For a finite rooted tree (with root ρ) we can define similar constructs. Fix a node v at distance h from the root and let $(v = v_0, v_1, \ldots, v_h = \rho)$ denote the path from the node to the root. For $k \leq h$ we write $f_k(v, t) = (t_0, t_1, \ldots, t_k)$ where t_0 is the subtree rooted at v_0 consisting of all nodes whose path from the root passes through v_0. Similarly, let t_i denote the subtree rooted at v_i consisting of all nodes for which the path from the root passes through v_i but not v_{i-1}. For all the models in this paper, the distance from the root to a typical vertex tends to infinity so it does not really matter how we define $f_k(v,t)$ for $k \geq h$, but for mathematical completeness define $f_k(v,t) = (t_0, t_1, \ldots, t_h, \emptyset, \ldots, \emptyset)$.

We are interested in a particularly strong notion of convergence of finite random trees to such `sin` trees, which is surprisingly satisfied in a wide array of probability models on finite trees.

Definition 2 *Let $\mathcal{T}_n = (\mathbf{V}_n, \mathbf{E}_n)$ be a sequence of random trees on n vertices with vertex labels $\mathbf{V}_n$ and edge set $\mathbf{E}_n$. Say that $\mathcal{T}_n$ converges in* **probability fringe sense** *to a* **sin** *random tree $\mathcal{T} \sim \mu$ if for any fixed $k \geq 1$ and any $\mathbf{t} = (t_0, t_2, \ldots, t_k) \in \mathbb{T}^k$, we have*

$$\frac{1}{n} \sum_{v \in V_n} \mathbb{1}\{f_k(v, \mathcal{T}_n) = \mathbf{t}\} \xrightarrow{P} \mathbb{P}_\mu \left(f_k(0, \mathcal{T}) = (t_0, \ldots t_{k-1})\right)$$

3 Results

Here we collect the basic results of our study.

3.1 Empirical Spectral distribution

We shall first prove the following general theorem and then shall describe what happens in the particular cases that we consider.

Theorem 3 *Let $\mathcal{T}_n$ denote a sequence of trees on n vertices. Suppose these trees converge in* **probability fringe sense** *to an infinite* `sin` *tree $\mathcal{T} \sim \mu$. Let F_n denote the empirical spectral distribution of either the adjacency matrix or the Graph Laplacian. Then there exists a probability distribution F on $\mathbb{R}$ such that*

$$F_n \stackrel{P}{\Rightarrow} F$$

Here convergence denotes convergence in the Levy-Prohrov metric on the space of measures on the real line.

Complementary to the above result, see [10] who uses similar such techniques to show convergence of the empirical distribution for some models of random networks.

Corollary 4 *Let $\mathcal{T}_n$ be any one of the four models of random trees described in Section 2.1 or the preferential attachment model of random networks. Then the empirical spectral distribution converges in probability to a (model dependent) probability distribution on $\mathbb{R}$.*

Remark: One of the unique properties of the proof of this theorem is that we are able to prove it without resorting to the usual work horse of random matrix theory, namely the method of moments. Typically, one method of showing results such as the above is to show that **for all** k,

$$\frac{1}{n} Tr(A^k) = \frac{1}{n} \sum_{1}^{n} \lambda_i^k = \int_{\mathbb{R}+} x^k dF_n(x) \xrightarrow{P} a_k$$

where the constants $(a_k)_{k \geq 1}$ uniquely determine a distribution. However in the case of the preferential attachment model, it is not hard to show that empirical spectral moments diverge except the second moment which converges to one.

Lemma 5 *Let A_n be the adjacency matrix of preferential attachment tree $\mathcal{T}_n$, then*

$$\frac{1}{n}\mathbb{E}Tr(A_n^4) \to \infty.$$

Proof. Let $d_n(k)$ be the degree of the k-th vertex in $\mathcal{T}_n$. We have, $\sum_{k=1}^n d_n(k) = 2(n-1)$ and $d_n(n) = 1$. At $(n+1)$-th step, $(n+1)$-th vertex is attached to the k-th vertex with probability $d_n(k)/2(n-1)$.

Claim 1. If $C_n = \sum_{k=1}^n d_n(k)^2$, then $(n-1)^{-1}\mathbb{E}(C_n)$ diverges to infinity.

Proof. Note that

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

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

Hence, we have the following recursive relation,

$$\mathbb{E}(C_{n+1}) - \mathbb{E}(C_n) = 1 + \mathbb{E}\sum_{k=1}^n (2d_n(k)+1)(d_n(k)/2(n-1)) = 2 + (n-1)^{-1}\mathbb{E}C_n.$$

Thus we get,

$$\mathbb{E}(C_{n+1}) - \mathbb{E}(C_n) \geq 2 \ \ \forall n \Rightarrow \liminf_n (n-1)^{-1}\mathbb{E}(C_n) \geq 2.$$

Again,

$$\liminf_n (\mathbb{E}(C_{n+1}) - \mathbb{E}(C_n)) \geq 4 \Rightarrow \liminf_n (n-1)^{-1}\mathbb{E}(C_n) \geq 4.$$

Proceeding similarly, we get that

$$\liminf_n (n-1)^{-1}\mathbb{E}(C_n) \geq 2k \text{ for every } k \geq 1.$$

Claim 2: $\mathbb{E}(Tr(A_{n+1}^4)) - \mathbb{E}(Tr(A_n^4)) \geq \mathbb{E}d_n(\mathcal{I}) = (2(n-1))^{-1}\mathbb{E}C_n$.

Proof. Note that

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

From claim 1 and 2, we can easily conclude that $n^{-1}\mathbb{E}(Tr(A_n^4)) \to \infty$.

Instead we take recourse to the Stieltjes transform approach. In passing we should remark that although typically this approach often requires hard analysis, most of our methods require very soft arguments.

3.2 Kernel of the spectrum

Theorem 6 *Let $(\mathcal{T}_n)_{n\geq 1}$ be the sequence of either the random recursive trees or the preferential attachment trees and let $A_n(\mathcal{T}_n)$ be the corresponding adjacency matrix. Let $\delta_n(\mathcal{T}_n)$ be the multiplicity of the zero eigenvalue for the adjacency matrix. Then*

$$\frac{\delta_n(\mathcal{T}_n)}{n} \xrightarrow{P} C(\mathcal{T})$$

for some (model dependent) constant $C(\mathcal{T})$.

3.3 Maximal Eigenvalues

Here we prove refined distributional asymptotics for the maximal eigenvalues of the adjacency matrix for the preferential attachment model. Recall that our model depended on a parameter a. Define

$$\gamma_a = a + 2 \tag{1}$$

Theorem 7 *(a) Fix any $k \geq 1$. Let $\mathcal{T}_n$ be the linear preferential attachment tree on n vertices. Let $\lambda_1 \geq \lambda_2 \geq \cdots \lambda_k$ be the k largest eigenvalues of the adjacency matrix of $\mathcal{T}_n^a$ and let $\Delta_1 \geq \Delta_2 \cdots \geq \Delta_k$ be the k largest degree. Then we have $\lambda_i = (1 + o(1))\sqrt{\Delta_i}$ for all $1 \leq i \leq k$ as $n \to \infty$.*

(b) There exist (non-degenerate) random variables $X_1 \geq X_2 \geq \cdots X_k > 0$ such that

$$\left(\frac{\Delta_1}{n^{1/\gamma_a}}, \frac{\Delta_2}{n^{1/\gamma_a}}, \ldots, \frac{\Delta_k}{n^{1/\gamma_a}}\right) \xrightarrow{d} (X_1, X_2, \ldots, X_k)$$

In particular, this implies via the results from [20], [14] that for the maximal eigenvalues we have

$$\left(\frac{\lambda_1}{n^{1/2\gamma_a}}, \frac{\lambda_2}{n^{1/2\gamma_a}}, \ldots, \frac{\lambda_k}{n^{1/2\gamma_a}}\right) \xrightarrow{d} (\sqrt{X_1}, \sqrt{X_2}, \ldots, \sqrt{X_k}) \tag{2}$$

Remark: Note that part(a) of the above theorem for the BA model (where $a = 0$) was proved by Frieze et al [20]. Here we give a separate proof (using continuous time branching process ideas but using similar conceptual techniques) and extend the proof to the general setup, and further find the exact scaling of the maximal eigenvalues.

Corollary 8 *For the Barabasi Albert tree, we have*

$$\left(\frac{\lambda_1}{n^{1/4}}, \frac{\lambda_2}{n^{1/4}}, \ldots, \frac{\lambda_k}{n^{1/4}}\right) \xrightarrow{d} (\sqrt{X_1}, \sqrt{X_2}, \ldots, \sqrt{X_k})$$

4 Proofs

We now delve into the proofs of the results. Before we start proving the main results, we first describe the continuous time branching process embedding that we use for some of our models and the corresponding limiting `sin` trees as defined in Section 2.1.

4.1 General theory

We first give a theorem regarding the probabilistic construction of the random tree models in question.

Proposition 9 *(a) Consider the Yule process $\mathcal{F}(t)$ which is a continuous time branching process where each individual lives forever and reproduces at unit rate. Define the stopping time $T_n = \inf\{t > 0 : |\mathcal{F}(t)| = n\}$ Then*

$$\mathcal{F}(T_n) \stackrel{d}{=} \mathcal{T}_n$$

where $\mathcal{T}_n$ is the random recursive tree on n vertices.

(b) Consider a continuous time branching process where now the offspring distribution of each individual has a Markov point process $N_a(\cdot)$ for his offspring distribution and where the point process is described by the rate equation

$$\mathbb{P}\left(N_a(t+dt) - N_a(t) = 1 | N_a(t) = k\right) = (k+1+a)dt + o(dt)$$

Let $T_n = \inf\{t > 0 : |\mathcal{F}(t)| = n\}$ Then

$$\mathcal{F}(T_n) \stackrel{d}{=} \mathcal{T}_n$$

where $\mathcal{T}_n$ is the Linear preferential attachment model on n vertices.

(c) Let $PGW^{(1)}$ be a (discrete time) branching process where each individual has Poisson number of children with mean 1. Note that $|PGW^{(1)}| < \infty$ a.s. Then

$$\mathcal{T}_n \stackrel{d}{=} PGW^{(1)} \Big| \{|PGW^{(1)}| = n\}$$

where $\mathcal{T}_n$ is the rooted uniform random unordered tree on n labeled vertices.

(d) Consider the random binary splitting process $\mathcal{F}(t)$ where we start with one individual at time zero and each individual lives for an exponential(1) amount of time and splits into two children. Let $T_n = \inf\{t > 0 : |\mathcal{F}(t)| = n\}$ Then

$$\mathcal{F}(T_n) \stackrel{d}{=} \mathcal{T}_n$$

where $\mathcal{T}_n$ is the random binary tree on n vertices.

Proof: The proofs are obvious, 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) follows from directly computing the probability of a rooted tree under the measure $\mathrm{PGW}^{(1)} \Big| \{|\mathrm{PGW}^{(1)}| = n\}$ and verifying using Cayley's theorem for the number of rooted, unlabeled vertices on n vertices that this corresponds to the uniform measure.

■

Remark: Part(c) is not the most useful construction in some cases (e.g. when checking the concentration of the maximal matching on the tree) so we give the following alternate construction, taken from [13]. Define the set $[n] := \{1, 2, \ldots, n\}$. Let $Y_0, Y_1, \ldots,$ be a sequence which is iid uniform distributed on the set $[n]$. Given $(Y_0, Y_1, \ldots)$, define the directed graph $\mathcal{T}(Y_0, Y_1, \ldots,)$ as

$$\mathcal{T}(Y_0, Y_1, \ldots) := \{(Y_{j-1}, Y_j) : Y_j \notin \{Y_0, \ldots, Y_{j-1}, j \geq 1\}\}$$

Think of $\mathcal{T}(Y_0, Y_1 \ldots,)$ as a tree labeled by $\{Y_0, Y_1, \ldots\}$ with root Y_0

Proposition 10 *The tree $\mathcal{T}(Y_0, Y_1, \ldots)$ has the same distribution as the rooted uniform unordered tree on n labelled vertices.*

Proof: See [13].

4.2 Empirical Spectral distribution

Here we shall give a proof of Theorem 3. The essential idea will be to show convergence in probability of the sequence of the (random) Stieltjes transforms $s_n(z)$ of the empirical distribution functions, to a limiting (non random) function $s(z)$. We first start with a few results on tightness and convergence of the Stieltjes transform.

Let the set

$$D = \{z \in \mathbb{C} : \mathcal{I}(z) > 0\}$$

Fix some $\alpha > 1$. We shall be interested in the set

$$D_\alpha = \{z \in \mathbb{C} : \mathcal{I}(z) > \alpha\}$$

Proposition 11 *Fix $\alpha > 1$. Let F_n be a sequence of random probability measures on $\mathbb{R}$ which are a.s. tight. Then $\exists$ a distribution function F such that*

$$d(F_n, F) \xrightarrow{P} 0$$

for any distance metric d which metrizes weak convergence of probability measures on $\mathbb{R}$, **if and only if** *there exists a function $s(\cdot)$ such that, for any subsequence F_{n_k}, there exists a further subsequence $F_{n'_k}$ and a null set N^* such that for all $w \notin N^*$*

$$s_{n'_k}(z) \longrightarrow s(z) \qquad \textit{for all } z \in D_\alpha \tag{3}$$

Proof: Standard results on weak convergence and uniqueness of Stieltjes transform (see Theorem 12.8 of Bai [5]) imply that F_n converges in probability to F if and only if Equation (3) holds, with D_α replaced by D, namely

$$s_{n'_k}(z) \longrightarrow s(z) \qquad \text{for all } z \in D \tag{4}$$

Thus we only need to show that equation (3) implies (4). This is obvious since Equation (3) implies that all sub-sequential limits agree with $s(z)$ on D_α. By Theorem 10.18 of [25], this implies that all subsequential limits agree on the whole of D, thus proving the uniqueness of subsequential limits.

■

The following Lemma essentially states that showing pointwise convergence for the Stieltjes transform is enough.

Lemma 12 *Fix $\alpha > 1$. Let F_n be a sequence of random probability measures which are almost surely tight. Suppose there exists a function $s(z)$ such that for each fixed $z \in D_\alpha$,*

$$s_n(z) \xrightarrow{P} s(z)$$

as $n \to \infty$. Then $\exists$ a distribution function F such $F_n \Rightarrow F$ in probability as $n \to \infty$.

Proof: First note that $\forall z \in D_\alpha$, $|s_n(z)| \leq 1/\alpha$. Fix a countable dense subset $D^* \subset D_\alpha$ (e.g. the set of all points in D_α which have rational coordinates). By the usual subsequence argument, given any subsequence F_{n_k} we can find a further subsequence such $F_{n'_k}$ such that there exists a null set N^* such that for all $\omega \notin N^*$,

$$s_{n'_k}(z) \longrightarrow s(z) \quad \text{for all } z \in D_\alpha$$

Now using Vitaly's theorem for the convergence of holomorphic functions (paraphrased from [5], Lemma 2.13) implies that F_n satisfy the conditions of Proposition 11, and this completes the proof.

Theorem 13 (Vitali's theorem) *Let $s_1, s_2, \ldots$ be analytic in D_α, a connected open set of $\mathbb{C}$, satisfying $|s_n(z)| \leq M$ for every n and z in D, and let $s_n(z)$ converges as $n \to \infty$ for each z in a subset of D^* having a limit point in D. Then there exists a function s, analytic in D, for which $s_n(z) \to s(z)$.*

■

We quickly prove the following trivial lemma regarding the tightness of all the distribution functions we analyze.

Lemma 14 *Consider a sequence of random trees $\mathcal{T}_n$ on n vertices, and let F_n denote the empirical spectral distribution of either the adjacency matrix or the Graph Laplacian. Then the sequence F_n are a.s. a tight sequence*

Proof: For the graph Laplacian, since all the eigenvalues are in the interval $[-1, 1]$, the proof is obvious. For the adjacency matrix, note that

$$\int_{\mathbb{R}} x^2 dF_n(x) = \frac{1}{n} \sum_1^n \lambda_i^2 = \frac{1}{n} Tr(A^2) = 1$$

This completes the proof.

■

Recall the definition of random `sin`-trees from Section 2.3. Associated with the `sin` tree we shall define a function $s(z)$. The remaining part of the proof essentially tries to show that the Stieltjes transform for the finite trees, $s_n(z)$ converge to the corresponding functions for the infinite tree.

We first start by collecting some basic properties of $s_n(z)$ for finite trees. For a finite rooted tree $\mathbf{t}$ with root v and adjacency matrix $A(\mathbf{t})$, write

$$s_n(z) = \frac{1}{n} \sum_1^n R_{vv}(\mathbf{t}, z)$$

Here $R_{vv}(\mathbf{t}, z) = (A(\mathbf{t}) - zI)^{-1}_{vv}$.

Let $c_1, c_2, \ldots, c_{n(v)}$ be the children of the root v. For any c_i let $\mathbf{t}_i$ be the subtree consisting of all descendants of c_i. Write

$$F(v, \mathbf{t}) = \sum_1^{n(v)} (A(\mathbf{t}_i) - zI)^{-1}_{c_i, c_i}$$

Now recall from Section 2.3, given a rooted tree $\mathcal{T}$ with root ρ, and a vertex $v \in \mathcal{T}$ at distance h from the root and path $(v_0 = v, v_1, \ldots, v_h = \rho)$, we had decomposed this tree as

$$\mathcal{T} = (\mathbf{t}_0(v_0), \mathbf{t}_1(v_0), \ldots, \mathbf{t}_h(v_0)) \tag{5}$$

where $\mathbf{t}_i$ was rooted at v_i. Call this the decomposition of $\mathcal{T}$ about the node v_0. For $i \geq 1$, write the tree $\mathbf{t}^{-i}(v_0)$, rooted at v_i as

$$\mathbf{t}^{-i}(v_0) = \mathcal{T} \setminus (\mathbf{t}_0(v_0), \ldots, \mathbf{t}_{i-1}(v_0))$$

Proposition 15 *Fix a rooted tree $\mathcal{T}$. (a) Fix any $z \in D_\alpha$. For all nodes v, we have*

$$\mathcal{I}(R_{vv}(\mathcal{T}, z)) \geq 0, \quad |R_{vv}(\mathcal{T}, z)| \leq 1/\alpha$$

(b) Fix any node v. Consider the decomposition of the tree $\mathcal{T}$ about the node v as given by Equation (5). Then we have the identity

$$R_{vv}(\mathcal{T}, z) = \frac{1}{-z - F(v, \mathbf{t}_0(v)) - R_{v_1 v_1}(\mathbf{t}^{-1}(v), z)} \tag{6}$$

Proof: To prove part(a) note that by spectral decomposition, we can write

$$(A - zI)^{-1} = U \text{diag}\left(\frac{1}{\lambda_i - z}\right) U^* = \sum_1^n \frac{1}{\lambda_i - z} u_i \overline{u_i}^t$$

for an orthonormal matrix $U = (u_1, u_2, \ldots, u_n)$, where λ_i are the eigenvalues of A. This implies that

$$R_{vv}(\mathcal{T}, z) = \sum_{j=1}^n \frac{1}{\lambda_j - z} |u_{vj}|^2 = \sum_{j=1}^n \frac{(\lambda_j - \bar{z})|u_{vj}|^2}{|\lambda_j - z|^2} \tag{7}$$

Since the λ_i are real, the first inequality implies

$$|R_{vv}(\mathcal{T}, z)| \leq \frac{1}{\alpha} \sum_{j=1}^n |u_{vj}|^2 = \frac{1}{\alpha}$$

since U is orthonormal. Since $\mathcal{I}(z) \geq 0$, the second equality in Equation (7) implies that $\mathcal{I}(R_{vv}(\mathcal{T}, z)) \geq 0$.

Part(b) follows from the general formula for computing the inverse of a matrix via partitioning it. In the linear algebra literature it is often called the Schur's theorem or Schur's formula.

■

In keeping with the previous notation, let

$$D_0 = \{z \in \mathbb{C} : \mathcal{I}(z) \geq 0\}$$

Fix $z = u + iv \in D_\alpha$ as before so that $v > \alpha$. This shall remain fixed through the remaining part of the argument. Also fix a sequence $x_0, x_1, \ldots \in D_0$. The following Lemma is obvious.

Lemma 16 *Consider the maps* $g_n(z,\cdot) : D_0 \mapsto \mathbb{C}$ *defined as*

$$g_n(z,y) = \frac{1}{-z - x_n - y}$$

Then g_n *maps* D_0 *into* D_0 *and further is a contraction map with contraction factor* $1/(\alpha)^2$*, namely* $\forall\ y, y' \in D_0$,

$$|g_n(z,y) - g(z,y')| \leq \frac{1}{\alpha^2}|y - y'|$$

In particular the sequence of maps $f_n = g_0 \circ g_1 \circ \cdots g_n$ *has a unique fixed point such that* $\forall\ y \in D_0$

$$\lim_{n\to\infty} f_n(z,0) = \lim_{n\to\infty} f_n(z,y) = \tilde{f}(z)$$

Now we come to the description of the limiting stieltjes transform for the various tree models. Recall the definition of $F(v, \mathbf{t})$ for a tree $\mathbf{t}$ rooted at v.

Fix a measure μ on the product tree space $\mathbb{T}^\infty$ and let $\mathcal{T} = (t_0, t_1, \ldots) \sim \mu$ where recall that we had identified the root of t_i with the integer i. Fix $z \in D_\alpha$ and define the (random) maps $g_n(z,\cdot) : D_0 \mapsto \mathbb{C}$

$$g_n(z,y) \ = \begin{cases} \frac{1}{-z-R_{00}(\mathbf{t}_0,z)-y} & \text{for } n = 0 \\ \\ \frac{1}{-z-F(n,\mathbf{t}_n)-y} & \text{for } n \geq 1 \end{cases} \tag{8}$$

Lemma 17 *Consider the (random) maps* $f_n(z,\cdot) : D_0 \mapsto \mathbb{C}$ *defined as* $f_n = g_0 \circ g_1 \circ \cdots \circ g_n$*. Then*

(A) g_n *maps* D_0 *into* D_0 *and are contractive with contraction factor* $1/\alpha$*, and thus* f_n *have a unique fixed point such that for any* $y \in D_0$

$$\lim_{n\to\infty} f_n(z,y) = \lim_{n\to\infty} f_n(z,0) = r(z) \quad a.s.$$

(B) Further for any $z \in D_\alpha$, $|r(z)| \leq 1/\alpha$ *a.s..*

Proof: (A) follows by using Lemma 16 where we let $x_0 = R_{00}(\mathbf{t}_0, z)$ and $x_n = F(n, \mathbf{t}_n)$, since using Proposition 15(a), we have that $x_n \in D_0$ for all n and hence the (random) function g_n satisfy the properties of Lemma 16 a.s.

To prove (B), note that repeated applications of the Proposition 15(b) implies that

$$f_n\left(z, R_{n+1,n+1}(\mathbf{t}, z)\right) = R_{00}(\mathbf{t}^{n+1}, 0)$$

where

$$\mathbf{t}^{n+1} = (\mathbf{t}_0, \mathbf{t}_1, \ldots, \mathbf{t}_{n+1})$$

By part (a) of Proposition 15, $|f_n\left(z, R_{n+1,n+1}(\mathbf{t}, z)\right)| \leq \frac{1}{\alpha}$ for all n. Further, since

$$|R_{n+1,n+1}(\mathbf{t}, z)| \leq \frac{1}{\alpha}$$

we have that

$$\lim_{n\to\infty} f_n\left(z, R_{n+1,n+1}(\mathbf{t}, z)\right) = r(z)$$

This finishes the proof.

■

Definition 18 *Given an infinite random* `sin` *tree* $\mathcal{T} \sim \mu$ *on* $\mathbb{T}^\infty$*, let the expected Stieltjes transform of the infinite sin tree be defined as*

$$s(z) = \mathbb{E}(r(z))$$

where $r(z)$ *is the random function constructed in Lemma 17. Note that this expectation exists since* $|r(z)| \leq 1/\alpha$

Proposition 19 *Fix some* $\alpha > 1$*. Let* $\mathcal{T}_n$ *be a sequence of random trees on* n *vertices, such that* $\mathcal{T}_n$ *converge in* **probability fringe sense** *to an infinite random sin tree* $\mathcal{T}$ *with expected Stieltjes transform as defined in Definition 18. Let* $s_n(z)$ *be the (random) Stieltjes transform of* $\mathcal{T}_n$*. Fix* $z \in D_\alpha$*. Then for each fixed* z

$$s_n(z) \xrightarrow{P} s(z)$$

as $n \to \infty$*.*

Note that this Proposition, along with the tightness lemma, Lemma 14 and the pointwise convergence lemma, Lemma 12 complete the proof of Theorem 3. Thus to complete the proof of Theorem 3, we just need to prove the above Proposition

Proof of Proposition 19:

We shall show that

$$\mathbb{E}(s_n(z)) \longrightarrow \mathbb{E}(s(z)) \tag{9}$$
$$\mathrm{Var}(s_n(z)) \longrightarrow 0 \tag{10}$$

To prove Equation (9) note that

$$\mathbb{E}(s_n(z)) = \mathbb{E}\left(\frac{1}{n}\sum_{v\in\mathcal{T}_n} R_{vv}(\mathcal{T}_n, z)\right) = \mathbb{E}(R_{V_n,V_n}(\mathcal{T}_n, z))$$

Here V_n is a node selected uniformly at random among all nodes. As in Section 2.3 let us decompose the finite tree $\mathcal{T}_n$ about the path from V_n to the root, namely $\mathcal{T}_n = (\mathbf{t}_0, \mathbf{t}_1, \ldots, \mathbf{t}_h)$ where $\mathbf{t}_i$ are trees rooted at v_i where $(v_0 = V_n, v_1, \ldots, v_h = \rho)$ is the path from the node V_n to the root.

Fix $z \in D_\alpha$ and define the sequence of random functions $g_k^{(n)}(z, \cdot) : D_0 \mapsto D_0$ as

$$g_k^{(n)}(z,y) = \begin{cases} \frac{1}{-z-R_{v_0,v_0}(\mathbf{t}_0,z)-y} & \text{for } k=0 \\ \frac{1}{-z-F(v_k,\mathbf{t}_k)-y} & \text{for } 1 \leq k \leq h \\ 0 & \text{for } k > h \end{cases}$$

By convention, write $v_k = \emptyset$ and $t_k = \emptyset$ for $k \geq h$ and $R_{v_k,v_k}(\mathbf{t}_k, z) = 0$. Then note that we can write, for any k,

$$R_{V_n,V_n}(\mathcal{T}_n, z) = g_0^{(n)} \circ g_1^{(n)} \cdots g_{k-1}^{(n)}(z, R_{v_k v_k}(\mathbf{t}_k, z)) \tag{11}$$

Define the function $f_k^{(n)}(z, \cdot) = g_0^{(n)} \circ g_1^{(n)} \cdots g_k^{(n)}(z, \cdot)$. Since all the functions $g_k^{(n)}$ satisfy the conditions of Lemma 17, they are contractions. Since $|R_{v_k v_k}(\mathbf{t}_k, z)| \leq 1/\alpha$, we have

$$|R_{V_n,V_n}(\mathcal{T}_n, z) - f_{k-1}^{(n)}(z, 0)| = |f_{k-1}^{(n)}(z, R_{v_k v_k}(\mathbf{t}_k, z)) - f_{k-1}^{(n)}(z, 0)| \leq \frac{1}{\alpha^{2k}} \tag{12}$$

Note that by probability fringe convergence of $\mathcal{T}_n$ to the limiting random tree $\mathcal{T} \sim \mu$, for any fixed $k \geq 0$, we have

$$|\mathbb{E}\left(f_{k-1}^n(z,0)\right) - \mathbb{E}\left(f_{k-1}(z,0)\right)| \longrightarrow 0 \tag{13}$$

as $n \to \infty$, where $f_{k-1}(z,\cdot)$ are the sequence of functions defined as compositions of g_i functions defined for the infinite tree as in Lemma 17.

Further by Lemma 17, we have

$$\mathbb{E}(f_k(z,0)) \longrightarrow s(z) \tag{14}$$

as $k \to \infty$.

Combining Equations (12), (13) and (14) we get that

$$\mathbb{E}(s_n(z)) = \mathbb{E}(R_{V_n,V_n}(\mathcal{T}_n, z)) \longrightarrow s(z)$$

as $n \to \infty$. This completes the proof for the Expectation.

To show that the variance tends to zero, note that it is enough to show that

$$\mathbb{E}\left[R_{V_n,V_n}(\mathcal{T}_n, z) \cdot R_{U_n,U_n}(\mathcal{T}_n, z)\right] \longrightarrow [s(z)]^2$$

Using limiting properties of the tree at a distance k from U_n and V_n, note that the following Lemma, telling us that neighborhoods of randomly selected nodes are essentially identically and asymptotically distributed as finite neighborhoods in the infinite `sin` tree, at least for large enough n, is more than enough to prove the variance result.

Lemma 20 *Recall the definition of $f_k(\mathcal{T}_n, v)$ from Section 2.3, giving us information upto distance k from the node v on the path to the root. Fix any $k \geq 1$ and 2 sets of trees $\mathbf{s} = (s_0, s_1, \ldots, s_k)$ and $\mathbf{t} = (t_0, t_1, \ldots, t_k)$. Then*

$$\mathbb{P}\left(f_k(\mathcal{T}_n, V_n) = \mathbf{t}, f_k(\mathcal{T}_n, U_n) = \mathbf{s}\right) \longrightarrow \mathbb{P}(f_k(\mathcal{T}, 0) = \mathbf{t}) \cdot \mathbb{P}(f_k(\mathcal{T}, 0) = \mathbf{s})$$

as $n \to \infty$

Proof: Note that the conditional probability

$$\mathbb{P}\left(f_k(\mathcal{T}_n, V_n) = \mathbf{t}, f_k(\mathcal{T}_n, U_n) = \mathbf{s} \middle| \mathcal{T}_n\right) = \left(\frac{1}{n}\sum_v \mathbb{1}\{f_k(v, \mathcal{T}_n) = \mathbf{t}\}\right) \cdot \left(\frac{1}{n}\sum_v \mathbb{1}\{f_k(v, \mathcal{T}_n) = \mathbf{s}\}\right)$$

Take expectations and use the definition of convergence in probability fringe sense to get the result.

■

This completes the proof of Theorem 3.

4.3 Size of maximum matching and kernel of the spectrum

Here we shall give two proofs of Theorem 6, each of which use fundamentally different techniques. While the first method uses a detailed analysis of the Karp-Sipser algorithm for trees, the second method uses a probabilistic reformulation of the cavity method developed in [3].

We first start with a simple lemma which draws a connection between Maximum matchings and the mass of the Spectral measure at 0.

Lemma 21 *Fix a tree $\mathcal{T}$ on n vertices and let A be the adjacency matrix of $\mathcal{T}$. Let $\delta(\mathcal{T})$ denote the number of zero eigenvalues of the adjacency matrix of $\mathcal{T}$. Then*

$$\delta(\mathcal{T}) = n - 2M(\mathcal{T})$$

where $M(\mathcal{T})$ is the number of edges in a maximum matching on $\mathcal{T}$.

Proof: Consider the characteristic polynomial $P(z) = |(z \cdot I - A)|$. Let $M_k(\mathcal{T})$ denote the set of k-matchings on the tree $\mathcal{T}$. Then note that by [12], Equation(5), we can write the characteristic polynomial as

$$P(z) = \sum_{k=0}^{M(\mathcal{T})} (-1)^k |M_k(\mathcal{T})| z^{n-2k}$$

This completes the proof. ■

The above Lemma demonstrates the importance of the understanding the number of edges in the Maximum matching on the tree. Both proofs essentially show that the expectation of the maximal matchings converge as $n \to \infty$. The following lemma shows that for the probabilistic tree models that we deal with, that is enough.

Lemma 22 *Suppose $(\mathcal{T}_n)$ is the sequence of one of the random trees defined in Section 2.1. Then*

$$\mathbb{P}(|M_n - \mathbb{E}(M_n)| > \lambda) \leq 2\exp(-\frac{\lambda^2}{4n})$$

Thus $\frac{M_n}{n} = \mathbb{E}\left(\frac{M_n}{n}\right) + o(1)$. In particular, if $\mathbb{E}\left(\frac{M_n}{n}\right) \to c_0$ for some constant c_0, then we have

$$\frac{M_n}{n} \xrightarrow{P} c_0$$

Proof: We shall prove the concentration result for the random recursive tree. The proof for the other models is similar. In particular the proof for the uniform random unordered tree on n labelled vertices uses the construction given in Proposition 10.

Now note that the random recursive tree $\mathcal{T}_n$ can be thought of as a function of $n-2$ independent random variables $U_3, U_4, \ldots, U_n$ where U_i is uniformly distributed on the set $1, 2, \ldots, (i-1)$. In particular we have that $M_n = f(U_1, U_2, \ldots, U_n)$. Now note that for any fixed $1 \leq i \leq n$, we have

$$|f(U_1, U_2, \ldots, U_{i-1}, U_i, U_{i+1}, \ldots, U_n) - f(U_1, \ldots U_{i-1} U_i', U_{i+1}, \ldots, U_n)| \leq 2$$

Thus by the Azuma-hoeffding inequality (see e.g. [27]) we have

$$\mathbb{P}(|M_n - \mathbb{E}(M_n)| > \lambda) \leq 2\exp(-\frac{\lambda^2}{4n})$$

This completes the proof.

■

Proof 1:

Our analysis on the size of maximum matchings of a tree relies heavily on an algorithm due to Karp and Sipser [], which is a simple randomized algorithm for finding an approximate maximum matching in a general sparse graph.

The algorithm can be described as follows:

Input $G = (V, E)$.

$M \leftarrow \emptyset$. While $E(G) \neq \emptyset$, do

Step I. If G has a vertex of degree 1, choose one (say x) randomly (or according to some deterministic order). Let $e = \{x, y\}$ be the unique edge in G incident on x.

Step II: If there is no vertex of degree 1, choose $e = \{x, y\} \in E$ randomly.

$G \leftarrow G \setminus \{x, y\}, M \leftarrow M \cup \{e\}$.

Output: M

We will now make very easy but immensely helpful observations about this algorithm when the input graph is a tree.

1. If the input G is a tree, we will never go into step II.
2. For tree Karp-Sipser algorithm is exact, that is, it will always produce one of the maximum matchings as an output. The proof of this fact can be done inductively once we note that any edge incident to a leaf of a tree can be extended to a maximum matching on the tree.

In what follows, given a graph $G = (V, E)$ and $V_1 \subseteq V$, $G \setminus V_1$ denotes the subgraph of G which is the restriction of G onto the vertex set $V \setminus V_1$.

Given a forest (union of trees) F, let $L(F) =$ the set of all leaves of F and $CL(F) =$ the set of all vertices of F which is connected to at least one leaf.

Given a tree T (may be infinite), we can partition the vertex set of T into $\{C_k : k \geq 1\}$ recursively in the following manner:

$$C_1 = L(T),\ C_2 = CL(T),\ T_1 = T \setminus C_1 \cup C_2.$$
$$\text{for } k > 1, \quad C_{2k-1} = L(T_{k-1}),\ C_{2k} = CL(T_{k-1}),\ T_k = T_{k-1} \setminus C_{2k-1} \cup C_{2k}.$$

From the facts noted about the Karp-Sipser algorithm for tree, the next lemma follows immediately.

Lemma 23 *The cardinality of the maximum matching of a tree T is given by*

$$M_n = \sum_{k=1}^{\infty} |C_{2k}(T)|. \tag{15}$$

Proposition 24 *Suppose $\mathcal{T}_n$ be a sequence of random rooted trees on n vertices which converge in probability fringe sense to a random sin tree $\mathcal{T}$ with law μ. Let M_n be the size of maximum matching on $\mathcal{T}_n$. Then*

$$\frac{\mathbb{E}M_n}{n} \rightarrow \sum_{k=1}^{\infty} \mathbb{P}_\mu(0 \in C_{2k}(\mathcal{T})).$$

Proof. For $v \in T$, let $\mathcal{N}_k(v)$ is the neighborhood of v in T which is within (graph) distance k from v. We now claim that $v \in C_k(T)$ can determined by only looking at $\mathcal{N}_k(v)$. The proof is obvious once we realize that we can alternatively define the sets C_k as follows

$$C_k(T) = \{v \in T : v \text{ is at exactly distance } k-1 \text{ away from } L(T)\}.$$

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

$$\begin{aligned} n^{-1} \sum_{k=m+1}^{\infty} \mathbb{E}|C_{2k}(\mathcal{T}_n)| &= \mathbb{P}(V \in \cup_{k=m+1}^{\infty} C_{2k}(\mathcal{T}_n)) \\ &\leq 1 - \mathbb{P}(\exists \text{ a leaf of } \ell \text{ within distance } 2m \text{ from } V) \\ &\leq 1 - \mathbb{P}(f_1(\mathcal{T}_n, V) \text{ contains a leaf within distance } 2m \text{ from } V) \\ &\overset{n\to\infty}{\to} 1 - \mathbb{P}_\mu(f_1(\mathcal{T}, 0) \text{ contains a leaf within distance } 2m \text{ from } 0) = o(m), \end{aligned}$$

where in the last step we have use the fact that $f_k(\mathcal{T}_n, V) \xrightarrow{d} f_k(\mathcal{T}, 0)$ on $\mathbb{T}^k$ for all $k \geq 1$ which is an easy consequence of convergence in fringe sense and Scheffe's theorem. Along the similar line, it can be shown that

$$\sum_{k=m+1}^{\infty} \mathbb{P}_\mu(0 \in C_{2k}(\mathcal{T})) = o(m).$$

On the other hand, note that since for any vertex $v \in \mathcal{T}_n, N_{2m}(v) \subseteq f_{2m}(\mathcal{T}_n, v)$, by continuous mapping theorem, we have for each $k \leq m$,

$$\mathbb{1}(V \in C_{2k}(\mathcal{T}_n) \xrightarrow{d} \mathbb{1}(0 \in C_{2k}(\mathcal{T})).$$

The rest of the proof is now immediate.

Remark 25 *Under the assumption of proposition 24, $n^{-1}\mathbb{E}I_n$ also converges where I_n is the size of a maximum independent set of $\mathcal{T}_n$. By Konig's theorem [], for a general bipartite graph, the size of a maximum matching is equal to the cardinality of a minimum vertex cover. On the other hand, complementation of a minimum vertex cover in any graph always yields a maximum independent set. Thus, we have, in our case, $I_n = n - M_n$.*

To get upper bounds on the asymptotic proportion of edges in the maximal matching we shall often use the following Lemma. Before describing the Lemma we shall need some notation. Fix any integer d and let $\mathcal{N}(d) \subset \mathcal{T}$ denote the set of vertices with degree greater than or equal to d. Let $\mathcal{I}(d) \subseteq \mathcal{N}(d)$ denote the subset of such vertices all of whose neighbors also have degree greater than or equal to d. Then we have the following Lemma

Lemma 26 *Fix and $d \geq 2$. The maximal matching $M(\mathcal{T})$ satisfies the upper bound*

$$M(\mathcal{T}) \leq n - \frac{|\mathcal{I}(d)|}{2} - \sum_{v \in \mathcal{N}(d)} (d(v) - 2)$$

Proof: ■

Proof 2:

We now give the second proof for the convergence of the normalized maximal matching. First note that the maximum matching on a tree need not be unique, namely there could be more than one subset of edges which could be maximum matchings on the tree. At a mathematical level, to avoid such complications we shall continuize the above combinatorial problem to deal with unique maximal *weight* matchings. Fix $0 < \varepsilon < 1$ Given a tree, for each edge e attach a random weight ξ_e (independent over edges) which are uniformly distributed in the interval $[1-\varepsilon, 1+\varepsilon]$. Recall the definition of matchings as defined in Section 2. The weight of any matching $S \subseteq \mathcal{T}$ is defined as $w(S) = \sum_{e \in S} w_e$. Let the random variable $M_n(\varepsilon)$ be the maximum weight matching.

We first start with a Lemma which shows why analyzing the asymptotics of $M_n(\varepsilon)$ gives us asymptotic information about the asymptotics of M_n.

Lemma 27 *Consider a sequence of finite random trees $\mathcal{T}_n$ on n vertices. Assume that for all $\varepsilon > 0$, we have*

$$\mathbb{E}\left(\frac{M_n(\varepsilon)}{n}\right) \longrightarrow c(\varepsilon)$$

for some limit constants $c(\varepsilon)$ as $n \to \infty$. Then there exists a constant c_0 such that

$$\mathbb{E}\left(\frac{M_n}{n}\right) \longrightarrow c_0$$

as $n \to \infty$.

Proof: Note that by the definition of M_n and $M_n(\varepsilon)$, it is easy to verify that

$$\left|\mathbb{E}\left(\frac{M_n}{n}\right) - \mathbb{E}\left(\frac{M_n(\varepsilon)}{n}\right)\right| \le \varepsilon \tag{16}$$

Fix any sequence $\varepsilon_n \to 0$. Since for all ε, the sequence of limit constants $0 < c(\varepsilon) < \frac{1}{2}$, we have by the Bolzano-Weirstrass theorem, that there exists a subsequence $\varepsilon_{n_k} \to 0$ and a constant c_0 such that $c(\varepsilon_{n_k}) \to c_0$. Now by Equation (16) we have

$$\mathbb{E}\left(\frac{M_n}{n}\right) \to c_0$$

as $n \to \infty$ and this completes the prove. ■.

Proof of Theorem 6: In lieu of Lemma 21 we see that it is enough to prove that the maximal matching $M_n(\mathcal{T}_n)$ satisfies the asymptotics:

$$\frac{M_n(\mathcal{T}_n)}{n} \xrightarrow{P} C_0(\mathcal{T})$$

Again using the concentration Lemma 22, it is enough to prove that the expectations converge. Finally by Lemma 27 it is enough to prove the following Proposition.

Proposition 28 *Fix any $0 < \varepsilon < 1$ and consider the Maximal weight matching $M_n(\varepsilon)$ as defined above. Then there exist constants $c(\varepsilon)$ such that*

$$\mathbb{E}\left(\frac{M_n(\varepsilon)}{n}\right) \overset{n\to\infty}{\longrightarrow} c(\varepsilon).$$

Proof: We shall prove the result for the random recursive tree. Essentially the same method applies for all the other random tree models. It is just more convenient to deal with a concrete case, for notational purposes. Thus through out the proof below, the random tree models we are dealing with are the random recursive trees. The proof is an elucidation of the Objective method as developed by Aldous and Steele [4]. The basic point is that for the tree models that we consider, the local neighborhood of a randomly chosen node asymptotically looks like the local neighborhood of a corresponding infinite tree, and this allows us to perform computations on the infinite tree.

The proof follows along conceptually similar directions as the proof for matchings on the uniform random tree given in [4]. To begin the proof, we first need some notation. We shall think of the tree as directed, with each edge directed from a node to it's parent. For any node v other than the root let $\grave{v}$ denote the edge from node v to its parent.

For any tree T let $M(T)$ denote the maximal partial matching and define $B(T)$ by the formulation $M(T) - B(T)$ is the maximal partial matching when we stipulate the additional restriction that the root does not belong to the matching. For any edge $e = (u, v)$ consider the two rooted trees (rooted at u and v respectively) formed by deleting edge e. Let $\mathcal{T}_n^{\texttt{small}}$ and $\mathcal{T}_n^{\texttt{big}}$ be the small subtree and the big subtree. Then note the following relations

(a) The maximal partial matching which does not contain edge $e = (u, v)$ is given by

$$M(\mathcal{T}_n^{\texttt{small}}(e)) + M(\mathcal{T}_n^{\texttt{big}}(e))$$

(b) The maximal partial matching that is constrained to contain edge e is given by

$$\xi_e + \left(M(\mathcal{T}_n^{\texttt{big}}(e)) - B(\mathcal{T}_n^{\texttt{big}}(e)) + M(\mathcal{T}_n^{\texttt{small}}(e)) - B(\mathcal{T}_n^{\texttt{small}}(e))\right)$$

Comparing we see that edge e is in the partial matching if and only if

$$\xi(e) > B(\mathcal{T}_n^{\texttt{big}}(e)) + B(\mathcal{T}_n^{\texttt{small}}(e))$$

Thus we can write the Maximal weight matching in the more useful form:

$$M_n(\mathcal{T}_n) = \sum_{e \in \mathcal{T}_n} \xi_e \mathbb{1}\{\xi_e > B(\mathcal{T}_n^{\texttt{big}}(e)) + B(\mathcal{T}_n^{\texttt{small}}(e))\}$$

Thus the expectation of M_n can be written as

$$\mathbb{E}(M_n) = (n-1)\mathbb{E}\left(\xi_{\mathbf{e}} \mathbb{1}\{\xi_{\mathbf{e}} > B(\mathcal{T}_n^{\texttt{big}}(\mathbf{e})) + B(\mathcal{T}_n^{\texttt{small}}(\mathbf{e}))\}\right)$$

where $\mathbf{e}$ is one of the $n-1$ edges chosen uniformly at random. Now choosing an edge and choosing a node at random are not the same, however in the case of the tree, it is easy to show via a simple coupling argument

$$\left|\mathbb{E}\left(\xi_{\mathbf{e}} \mathbb{1}\{\xi_{\mathbf{e}} > B(\mathcal{T}_n^{\texttt{big}}(\mathbf{e})) + B(\mathcal{T}_n^{\texttt{small}}(\mathbf{e}))\}\right) - \mathbb{E}\left(\xi_{(\mathbf{v},\mathbf{u})} \mathbb{1}\{\xi_{(\mathbf{v},\mathbf{u})} > B(\mathcal{T}_n^{\texttt{big}}((\mathbf{v},\mathbf{u}))) + B(\mathcal{T}_n^{\texttt{small}}((\mathbf{v},\mathbf{u})))\}\right)\right| \leq \frac{1}{n}$$

where $\mathbf{v}$ is a node chosen uniformly at random and $\mathbf{u}$ is it's mother. Thus it is enough to show that

$$\mathbb{E}\left(\xi_{(\mathbf{v},\mathbf{u})} \mathbb{1}\{\xi_{(\mathbf{v},\mathbf{u})} > B(\mathcal{T}_n^{\texttt{big}}((\mathbf{v},\mathbf{u}))) + B(\mathcal{T}_n^{\texttt{small}}((\mathbf{v},\mathbf{u})))\}\right) \longrightarrow C \tag{17}$$

Note that $\xi_{(\mathbf{v},\mathbf{u})}$ is independent of the random variables $\left(B(\mathcal{T}_n^{\texttt{big}}((\mathbf{v},\mathbf{u}))), B(\mathcal{T}_n^{\texttt{small}}((\mathbf{v},\mathbf{u})))\right)$. For simplicity of notation, we write $\mathbb{E}\left(\xi_{(\mathbf{v},\mathbf{u})} \mathbb{1}\{\xi_{(\mathbf{v},\mathbf{u})} > B(\mathcal{T}_n^{\texttt{big}}((\mathbf{v},\mathbf{u}))) + B(\mathcal{T}_n^{\texttt{small}}((\mathbf{v},\mathbf{u})))\}\right)$ as

$$\mathbb{E}\left(\xi \mathbb{1}\{\xi > B(\mathcal{T}_n^{\texttt{big}}(v)) + B(\mathcal{T}_n^{\texttt{small}}(v))\}\right)$$

Where v is a node selected uniformly at random from $\mathcal{T}_n$.

To prove our result it is thus enough to prove the following

Proposition 29 *There exist random variables (Y, Z) such that*

$$\left(B(\mathcal{T}_n^{\texttt{big}}(v)), B(\mathcal{T}_n^{\texttt{small}}(v))\right) \stackrel{d}{\Longrightarrow} (Y, Z)$$

as $n \to \infty$.

Proof: The proof of this proposition is slightly involved, however the conceptual idea is simple. We shall prove not only the distributional convergence as above, but shall also give an explicit description of the limit random variables. `xxx give brief outline of the proof`. Before we can proceed further we give an alternate description of the random recursive tree, which allows us to use continuous time branching process theory to glean detailed local asymptotics.

Lemma 30 *Consider the continuous time branching process $\mathcal{F}(t)$, where the population starts with one individual at time 0 who lives forever and produces at rate 1, and each of whose children also produce at rate 1 and live forever. For the above Markov process define*

$$T_n = \inf\{t : |\mathcal{F}(t) = n|\}$$

Then we have the identity

$$|\mathcal{F}(T_n)| \stackrel{d}{=} \mathcal{T}_n$$

where $\mathcal{T}_n$ is the random recursive tree.

Proof: Obvious.

■

The above process is called the Yule process and to fix notation, we shall call $\mathcal{F}(t)$, a Yule tree observed upto time t.

Now we shall define matchings on such trees and associated random variables called *bonuses.*

Definition 31 *Given a Yule tree $\mathcal{F}(t)$, where we attach to each edge, and independent random variable uniformly distributed on $[1-\varepsilon, 1+\varepsilon]$, let $M(t)$ be the random variable denoting the maximum weight matching on $\mathcal{F}(t)$. Further define the random variable $B(t) \geq 0$ by the formulation $M(t) - B(t)$ is the maximal weight matching under the additional constraint that we do not use the root.*

We give a simple distributional identity for the random variables $B(t)$

Lemma 32 *Let $U_1, U_2, \ldots$ be independent uniformly distributed on the interval $[0, t]$. Let $\xi_1, \xi_2, \ldots$ be uniform on $[1-\varepsilon, 1+\varepsilon]$. Conditional on U_i let $B^i(t-U_i)$ be conditionally independent with distribution $B(t-U_i)$. Let $N(t)$ be a Poisson random variable with mean t. Then*

$$B(t) \stackrel{d}{=} \max\left[0, \max_{1 \leq i \leq N(t)} \{\xi_i - B^i(t-U_i)\}\right]$$

Proof: We just condition on the times of birth of the children of the root upto time t and use the definitions of the maximal weight matchings and the bonus random variables.

■

We shall need the following property of the above random variables later.

Lemma 33 *There exists a constant $0 < c < 1$ such that for any t and ξ independent of $B(t)$ we have*

$$\mathbb{P}\left(B(t) > \xi\right) > c$$

Proof: Note that for any t, there exists c such that the probability at time t the root has a child v which is leaf, is greater than c. Then by the characterization given in Lemma 32, since conditional on the above event, we have $B_v(t - U_v) = 0$, thus

$$\mathbb{P}\left(B(t) > \xi\right) \geq c\mathbb{P}(\xi_v > \xi) = \frac{1}{2}c$$

■

We shall now describe the limiting random variables (Y, Z) that arise in Proposition 29. Let $X_0, X_1, \ldots,$ be independent $\exp(1)$ random variables, and let $S_i = \sum_0^i X_i$ be the partial sum sequence. It will be convenient for us to index the other variables with negative indices, although this gives rise to slightly clunky notation.

Conditional on the above sequence, let B_i for $i = -1, -2, \ldots$ be conditionally independent with $B_i \stackrel{d}{=} B(S_{-i})$. Let ξ_i, for $i = \{-1, -2, \ldots\}$ be independent $U[1-\varepsilon, 1+\varepsilon]$. Let the stopping time

$$T = \sup\{i \leq -1 : B_i \geq \xi_i\} \tag{18}$$

Note that by Lemma 33, $T < \infty$ a.s.

For $i = -1, -2, \ldots, -T$ define the sequence $(B_j^\infty)_{-T \leq j \leq 1}$ recursively as

$$B_j^\infty = \begin{cases} B_T & \text{j= T} \\ \max(\xi_j - B_{j-1}^\infty, B_j) & \text{otherwise} \end{cases} \tag{19}$$

By Equation (17), the following Proposition completes the proof.

Proposition 34 *We have the following explicit distributional convergence*

$$\left(B(\mathcal{T}_n^{\texttt{small}}(v)), B(\mathcal{T}_n^{\texttt{big}}(v))\right) \stackrel{d}{\Longrightarrow} (B_0, B_{-1}^\infty)$$

as $n \to \infty$ where B_0 is the bonus random variable defined as before and B_{-1}^∞ is given via the construction given by Equation (19).

Proof: As stated before, the crucial idea will be to define a limiting infinite, locally finite, rooted tree, whose neighborhood about the root gives us information about the neighborhood of a randomly chosen vertex in $\mathcal{T}_n$, for large n. The crucial connection in defining this infinite tree comes from the continuous time embedding given by Lemma 30.

Definition 35 *Let $\mathcal{T}^\infty$ be the following infinite locally fine rooted random tree*

(a) There is a single infinite path denoted as $\mathbb{Z}^+ = \{0, 1, 2, \ldots\}$. The tree is thought to be rooted at the origin 0.

(b) Let $X_0, X_1, \ldots$ be independent exponentially distributed with mean 1 *and let $(S_i)_{i \geq 0}$ be their partial sum sequence. Let $\mathcal{F}_0(\cdot), \mathcal{F}_1(\cdot), \ldots$ be independent Yule processes. We shall observe each for a finite amount of time as described in the next step.*

(c) Finally to finish the construction of the tree, think of the finite trees $\mathcal{F}_0(S_0)$ rooted at node 0, *$\mathcal{F}_1(S_1)$ rooted at node* 1 *and so. This gives an infinite tree rooted at the origin*

The point of defining the above infinite tree is the following theorem, which is essentially a special case of the stable age distribution theory for continuous time branching processes developed by Jagers and Nerman ([21], [24]).

`(xxx again note clunky notation of -i)`

Theorem 36 *Fix any k and set of finite rooted trees $t_0, t_1, t_2, \ldots t_k$. For the random tree $\mathcal{T}_n$ and a node V_n chosen u.a.r from $\mathcal{T}_n$, let $(v_0 = V_n, v_{-1}, \ldots, v_{-k})$ be the path from the V_n to the root of $\mathcal{T}_n$ truncated at distance k from V_n. Let $f_0(V_n)$ be the set of all nodes for which the path from the root passes through V_n, and for $i \geq 1$ let $f_{-i}(V_n)$ be the set of nodes for which the path from the root passes through v_i but not v_{i+1}. Think of $f_{-i}(V_n)$ as a tree rooted at v_i. Then we have*

$$\lim_{n\to\infty} \mathbb{P}\left(f_0(V_n) = t_0, \ldots, f_{-k}(V_n) = t_k\right) = \mathbb{P}\left(\mathcal{F}_0(S_0) = t_0, \ldots, \mathcal{F}_k(S_k) = t_k\right)$$

We shall now complete the proof by proving Proposition 34. To simplify notation we shall denote $f_{-i}(V_n) = \mathcal{T}_n^{-i}$ and let $\mathcal{F}_{-i} = \mathcal{F}(S_i)$. Recall that $B(\mathcal{T}_n^{\texttt{small}}(v)) = B(\mathcal{T}_n^0)$ denotes the bonus random variable for the tree First note that Theorem 36 immediately implies that for the bonus random variable $B(\mathcal{T}_n^0)$ we have

$$B(\mathcal{T}_n^0) \xrightarrow{d} B(\mathcal{F}_0) = B_0$$

as $n \to \infty$.

Also note that if we write $T_n = \sup\{-i : B(\mathcal{T}_n^{-i}) > \xi_{(-i,-i-1)}\}$ (where recall that $\xi_{(-i,-i-1)}$ denotes the edge weight corresponding to the edge (v_{-i}, v_{-i-1})) then the local convergence of the neighborhood gives us the following:

Lemma 37 *Consider the stopping time T_n defined above. Then we have*

$$T_n \xrightarrow{d} T$$

as $n \to \infty$, where T is the stopping defined in Equation (18).

We now turn to the more involved analysis of $B(\mathcal{T}_n^{\texttt{big}}(V_n))$. We however need some additional notation. In the context of the above decomposition of the local neighborhood about a node in terms of rooted trees on the path of the root to the node, we let the tree rooted at v_{-i} defined as

$$\mathcal{T}_n^{-i,\texttt{big}} := \mathcal{T}_n \setminus (\mathcal{T}_n^0, \mathcal{T}_n^{-1}, \ldots, \mathcal{T}_n^{-i+1})$$

In this notation, note that we are interested in the random variable $B(\mathcal{T}_n^{\texttt{big}}(V_n)) = B(\mathcal{T}_n^{-1,\texttt{big}})$. By the definitions of the bonuses we have the identity, for $T_n \leq j \leq -1$.

$$B(\mathcal{T}_n^j) = \begin{cases} B(\mathcal{T}_n^{T_n}) & \text{j= } T_n \\ \max\left(\xi_{j,j-1} - B(\mathcal{T}_n^{j-1,\texttt{big}}), B(\mathcal{T}_n^j)\right) & \text{otherwise} \end{cases} \tag{20}$$

Now compare with Equation (19) and use Lemma 37 to get the result.

■

4.4 Maximal Eigen values

Here we shall prove Theorem 7. Before proving the Theorem, we shall first prove some general properties of the branching process construction, namely Proposition 9(b), via which we embed the Preferential attachment model in a continuous time branching process. Part (a) of the following theorem (in a more general context) can be found in [8], so we shall not give the proof here. It essentially follows from the general theory of continuous time branching processes developed by Jagers and Nerman, specialized to our setup.

Theorem 38 *(a) Consider the continuous time branching process $\mathcal{F}(t)$, with offspring distribution $N(\cdot)$ as in Proposition 9(b). Then there exists a random variable $W > 0$, such that*

$$\frac{|\mathcal{F}(t)|}{e^{\gamma_a t}} \xrightarrow{a.s.} Z$$

as $t \to \infty$.

(b) Consider the offspring point process $N_a(t)$. Then $e^{-t} \cdot (N_a(t) + 1 + a)$ is an $\mathbb{L}^2$ martingale (in continuous time). In particular there exists a random variable $W_a > 0$ such that

$$\frac{N_a(t)}{e^t} \xrightarrow{a.s.} W_a$$

as $t \to \infty$.

Proof of (b): Note that the rate equations imply that

$$\mathbb{E}\left(dN_a(t)|\,\mathcal{F}(t)\right) = (N_a(t) + 1 + a)dt$$

where $\mathcal{F}(t)$ is the filtration generated by $N_a(\cdot)$. Thus writing $Y_a = e^{-t}\left(N_a(t) + 1 + a\right)$, this implies that

$$\mathbb{E}\left(dY_a(t)|\,\mathcal{F}(t)\right) = 0$$

The following is an obvious Corollary.

Corollary 39 *Consider the continuous time branching process $\mathcal{F}(t)$, with offspring distribution $N(\cdot)$ as in Proposition 9(b). Let $T_n = \inf\{t > 0 : |FF(t)| = n\}$. Then there exists a random variable $0 < Z < \infty$ such that*

$$T_n - \frac{1}{\gamma_a}\log n \xrightarrow{a.s.} Z$$

as $n \to \infty$

The following Lemma which describes the exact finite time distribution and also gives upper for the Yule process is well known.

Lemma 40 *(a) Suppose the parameter $a = 0$ so that $N_0(\cdot)$ is the offspring distribution process. Then for any fixed time, we have $N_0(t) \sim$ Geom(e^{-t}). In particular for any fixed $t > t_0$ we have*

$$\mathbb{P}(N_0(t - t_0) > Ke^t) \leq e^{-Ke^{t_0}}$$

(b) For $a > 0$, let $A = \lfloor a \rfloor + 1$. Then

$$\mathbb{P}(N_a(t - t_0) > Ke^t) \leq Ae^{-\frac{K}{A}e^{t_0}}$$

Proof: Part(a) is a well known property of the Yule process. To prove part (b), note that from the rate of growth of $N_a(\cdot)$, we can couple N_a with A independent Yule processes such that

$$N_a(\cdot) \leq_{st} \sum_1^A N_0^i(\cdot)$$

Now apply part(a) and the union bound.

■

Remark: We shall first prove Theorem 7 (b) regarding the asymptotics of the maximal degrees and then show the relation between the maximal degree and the maximum eigenvalues. The reason for this is as follows: the analysis of the maximal degrees allows us to conclude that the maximal degrees essentially occur within a "finite distance" from the root and that as we move further away from the root, the size of the degrees drops **very** quickly. Similar ideas will be used to understand the maximal eigenvalues.

Proof of Theorem 7(b):

Important Remark: For the next few paragraphs of the proof we shall assume that $a = 0$. The identical proof works for $a > 0$, but notation wise it is more convenient for us to think of $a = 0$. Note that for $a = 0$, $\gamma_a = 2$.

Fix $k \geq 1$ and recall that $\Delta_{\geq}\Delta_2 \cdot \geq \Delta_k$ denote the k maximal degree nodes. The essential idea will be to show that the maximal degrees all occur in a finite neighborhood about the root and that the degree of nodes in a finite neighborhood about the root, properly normalized converge in distribution. To start the argument fix $\varepsilon > 0$, which shall remain fixed through out the argument. Note that Corollary 39 implies that we can choose B_ε such that

$$\limsup_{n\to\infty} \mathbb{P}\left(|T_n - \frac{1}{2}\log n| > B_\varepsilon\right) \leq \varepsilon/2 \tag{21}$$

Fix this B_ε. In the construction of our Linear preferential attachment tree $\mathcal{T}_n$, we shall also attach with each node it's time of birth in the above construction. Fix $S > 0$ and write the event $A_{n,S}$ for the event

> There exists a node in $\mathcal{T}_n$ which was born after time S after the root, and which has degree greater than the root in $\mathcal{T}_n$.

Then we have the following Lemma

Lemma 41 *Given any $\varepsilon > 0$, there exists a constant $S_\varepsilon < \infty$ such that*

$$\limsup_{n\to\infty} \mathbb{P}(A_{n,S_\varepsilon}) \le \varepsilon$$

Proof: By Equation (21), note that it is enough to prove that we can find $S_\varepsilon < \infty$ such that

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

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

> For some time t in the interval $[\frac{1}{2}\log n - B_\varepsilon, \frac{1}{2}\log n + B_\varepsilon]$, in $\mathcal{F}(t)$ there exists a node, born after time S that has degree greater than the root.

Again, since the degrees of nodes (namely number of children of the nodes) are monotonically increasing functions of time t, it is enough to show that

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

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

> There exists a node born S units after the root such that degree of the node at time $(\frac{1}{2}\log n + B_\varepsilon)$ is greater than the degree of the root at time $(\frac{1}{2}\log n - B_\varepsilon)$.

For simplicity, let $t_-^n = \frac{1}{2}\log n - B_\varepsilon$ and $t_+^n = \frac{1}{2}\log n + B_\varepsilon$. Also for any fixed S, let

$$Z(S, t_+^n) = \max_{v:T_v > S} \deg(v, t_+^n)$$

Here T_v is the time of birth of node v and the $\deg(v, t_+^n)$ is the degree of the node at time t_+^n, namely the number of children node v has at time t_+^n. Let $X(\rho, t)$ denote the degree of the root at time t. Then note that the event $A''_{n,S} = \{Z(S, t_+^n) > X(\rho, t_-^n)\}$. Note that for any fixed constant K

$$\begin{aligned}\mathbb{P}([A''_{n,S}]^c) &\ge \mathbb{P}\left(X(\rho, t_-^n) > K, Z(S, t_+^n) < K\right)\\ &\ge \mathbb{P}(X(\rho, t_-^n) > K) + \mathbb{P}(Z(S, t_+^n) < K) - 1\end{aligned}$$

Thus it is enough to show that we can choose a sequence K_n and and then constants S_ε such that

$$\liminf_{n\to\infty} \mathbb{P}(X(\rho, t_-^n) > K_n) \ge 1 - \varepsilon/4 \tag{22}$$

$$\liminf_{n\to\infty} \mathbb{P}(Z(S_\varepsilon, t_+^n) < K_n) \ge 1 - \varepsilon/4 \tag{23}$$

To prove the first inequality, note that since the offspring distribution is a Yule process, $X(\rho, t_-^n) \sim Geom(\sqrt{n}/e^{B_\varepsilon})$, we can find a constant K_ε such that, defining $K_n = K_\varepsilon\sqrt{n}$, the first inequality is satisfied.

Now fix this sequence K_n. We need to show that we can choose a constant S_ε large enough such that for this sequence of K_n, the second inequality is satisfied. In this setup, a simple union bound does not work due to the exponential rate of growth of the whole branching process and we need to employing a *bracketing* technique.

For any fixed integer m, let $Z([m, m+1], t_+^n)$ be the maximum degree at time $t_{t_+^n}$ of all nodes born in the time interval $[m, m+1]$. Without loss of generality, assuming that S and t_+^n are integers, we have

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

Now note that by the union bound

$$\mathbb{P}(Z([m, m+1], t_+^n) > K_n) \leq \mathbb{E}(|\mathcal{F}(m+1)|)\mathbb{P}(N(t_+^n - S) > K_n) \tag{24}$$

where $|\mathcal{F}(m+1)|$ is the total number of nodes born by time $m+1$ and $N(t_+^n - S)$ has a Geometric distribution $Geom(e^{t_+^n - S})$ Recall that $\mathcal{F}(t)$ grows at rate e^{2t} and n particular (see [21]), there exists a constant C such that for all t, $\mathbb{E}(\mathcal{F}(t)) \leq Ce^{2t}$. Applying the large deviation Lemma for geometric distribution (Lemma 40(a)), we get

$$\mathbb{P}(Z(S, t_+^n) > K_n) \leq \sum_{S}^{t_+^n - 1} Ce^{2m} e^{-C' e^{m-1}}$$

where $C' = K_\varepsilon / e^{B_\varepsilon}$. Choosing S large enough, the second inequality now follows and the proof is completed.

Let us summarize in a Lemma what we have proved. We call all nodes born within time S after the initiating ancestor, the *S-neighborhood* of the root, denoted by $B(\rho, S)$. Let $\Delta^{S,n}$ denote the maximum degree in $\mathcal{T}_n$ among all nodes born within time S of the root. As before let Δ_1^n denote the maximum degree in $\mathcal{T}_n$

Lemma 42 *Given any ε, we can find $S_\varepsilon < \infty$ such that*

$$\liminf_{n \to \infty} \mathbb{P}(\Delta^{S_\varepsilon, n} = \Delta_1^n) \geq 1 - \varepsilon$$

A slightly more refined analysis shows that the above Lemma can be generalized for the k maximal degrees. We first need some convention. In the branching process construction, for any neighborhood $B(\rho, S)$ of the root, let $\Delta_1^{S,n} \geq \Delta_2^{S,n} \geq \cdots \geq \Delta_k^{S,n}$ be the k largest degrees in the $\mathcal{T}_n$ of the nodes of $B(\rho, S)$, with the convention that $\Delta_i^{S,n} = 0$ for $i \geq |B(\rho, S)|$ if $|B(\rho, S)| < k$. Let $\mathbf{\Delta}_k^{S,n} = (\Delta_1^{S,n}, \ldots, \Delta_k^{S,n})$; let the vector of k largest maximal degrees be denoted as $\mathbf{\Delta}_k^n = (\Delta_1, \ldots, \Delta_k)$. Then we have the following generalizations of the above Lemma

Lemma 43 *Given any $\varepsilon > 0$, we can find a constant $S_\varepsilon < \infty$ such that*

$$\liminf_{n \to \infty} \mathbb{P}(\mathbf{\Delta}_k^{S_\varepsilon, n} = \mathbf{\Delta}_k^n) \geq 1 - \varepsilon$$

A tightness Lemma: Now we come to the following routine tightness lemma which allows us to prove the existence of limits of random variables.

Lemma 44 *Let $X_n = (X_1^n, \ldots, X_n^k)$ be a sequence of $\mathbb{R}^k$ valued random variables. Suppose for each fixed $\varepsilon > 0$, there exists a sequence of $\mathbb{R}^k$ valued random variables $Y_n^{(\varepsilon)}$ such that*

$$\liminf_{n \to \infty} \mathbb{P}(X_n = Y_n^{(\varepsilon)}) \geq 1 - \varepsilon$$

Further assume that the collection of random variables $(Y_n^{(\varepsilon)})_{n\geq 1,\varepsilon>0}$ are tight. Finally assume that for each fixed $\varepsilon > 0$, there exists $Y^{(\varepsilon)}$ such that

$$Y_n^{(\varepsilon)} \xrightarrow{d} Y^{(\varepsilon)}$$

Then there exists a $\mathbb{R}^k$ valued random variable X_∞ such that

$$X_n \xrightarrow{d} X_\infty$$

as $n \to \infty$.

Proof: Fix any subsequential limit of Y_ε which exists and is finite since the sequence $(Y_{n,\varepsilon})_{n\geq 1,\varepsilon>0}$ is tight. This limiting random variable Y_0 will remain fixed for the rest of the argument. Then any subsequential limit of X_n has to coincide with Y_0 and this completes the proof.

■

We shall once again revert to statements regarding the model for general a. Recall the Malthusian rate of growth parameter γ_a. Applying Lemma 44 with $X_n = \mathbf{\Delta}_k^n/n^{1/\gamma_a}$ and $Y_n^\varepsilon = \mathbf{\Delta}_k^{n,S_\varepsilon}/n^{1/\gamma_a}$, to complete the proof that the maximal degrees properly normalized, converge to some distribution, it is thus enough to prove the following

Proposition 45 *Fix any $S > 0$ and consider the marked tree $B_n(\rho, S)$ where we mark each node $v \in B(\rho, S)$ with the real valued random variable* $\deg(v, \mathcal{T}_n)/n^{1/\gamma_a}$ *attached to it. Then*

$$B_n(\rho, S) \xrightarrow{d} B(\rho, S)$$

as $n \to \infty$. Here the convergence denotes convergence in the space of marked finite rooted trees. In particular

$$\frac{\mathbf{\Delta}_k^{n,S}}{n^{1/\gamma_a}} \xrightarrow{d} Y^S$$

as $n \to \infty$, for some $\mathbb{R}^k$ valued random variable Y^S.

Proof: Recall that we are going to think of $B_n(\rho, S)$ as a marked tree, where the topology of the tree consists of all those nodes in $\mathcal{T}_n$ which were born before time S in the continuous time embedding, and where we mark the edges from mother to daughter by the time after birth of the mother, that the daughter is born and each node is marked with $N_n(v) = \deg(v, n)/n^{1/\gamma_a}$. Thus we can write

$$B_n(\rho, S) = (\mathcal{F}(S), (N_n(v))_{v\in\mathcal{F}(S)})$$

where $\mathcal{F}(S)$ is the continuous time branching process observed till time S. Thus it is enough to prove the following:

Lemma 46 *Conditional on $\mathcal{F}(S)$, there exists a distribution $\mu(\mathcal{F}(S), \cdot)$ on $\mathbb{R}^{|\mathcal{F}(s)|}$ such that the vector*

$$(N_n(v))_{v\in\mathcal{F}(s)} \xrightarrow{d} \mu(\mathcal{F}(S), \cdot)$$

Proof: Conditional on $|\mathcal{F}(S)|$, we shall first analyze T_n, namely the amount of time needed by the continuous time branching process to grow to size n. Recall that unconditionally $\mathcal{F}(T_n) \stackrel{d}{=} \mathcal{T}_n$. The interesting fact about this model is that the total rate of growth of this at any time does not depend on the actual structure of the tree, rather depends only on the number of nodes k in the tree. To see why, suppose at any time t, there are k nodes and suppose the root has degree $d_\rho(t) = d_1(t)$ and let $d_i(t)$ denote the degree of vertex $i \geq 2$. Then note that by the rates of reproduction, we have that the rate of reproduction of the next individual into the population is

$$d_1(t) + 1 + a + \sum_{2}^{k}(d_i(t) + a) = 2k - 1 + ka$$

Thus conditional on $\mathcal{F}(S)$, the following Lemma in particular says that there exists a random variable $S_n^{|}\mathcal{F}(S)|$ which depends only on $|\mathcal{F}(S)|$ such that

$$T_n - \frac{1}{\gamma_a}\log n \xrightarrow{a.s.} W_k \tag{25}$$

as $n \to \infty$. Note that here $W_k = S + S_\infty^k$, where S_∞^k is the limiting random variable arising in the Lemma below.

Lemma 47 *Fix $a \geq 0$ and an integer $k \geq 1$. Let $(Y_i)_{i\geq 1}$ be iid, $Y_i \sim \exp(1)$. Consider the sequence*

$$S_n = \sum_{j=k}^{n}\left(\frac{Y_j}{2j-1+ja} - \frac{1}{2j-1+ja}\right)$$

Then there exists a random variable S_∞^k such that

$$S_n \xrightarrow{a.s.} S_\infty^k \quad \textit{and} \quad \mathbb{E}(S_\infty^k)^2 < \infty$$

Proof: Just observe that S_n^k is an $\mathbb{L}^2$ bounded martingale.

■

Now note that for any vertex $v \in \mathcal{F}(S)$, writing $P_v(t)$ for the offspring (namely degree of node v at time t), note that deg(v,n) $= P_v(T_n)$. By the asymptotics for the offspring distribution given by Theorem 38(b), we have

$$\frac{\deg(v,n)}{e^{T_n}} \xrightarrow{a.s.} W_v$$

Combining with Equation (25) on the asymptotics for the stop times T_n, we get

$$\frac{\deg(v,n)}{n^{1/\gamma_a}} \xrightarrow{a.s} \tilde{W}_v$$

for each $v \in \mathcal{F}(S)$ and this completes the proof of the Proposition.

■

We summarize the following fact about the random trees $\mathcal{F}(t_n)$ for some non random sequence t_n which follows from the above analysis. It shall be needed in the analysis of the maximal eigenvalues below.

Lemma 48 *Fix a sequence of constants $t_n \to \infty$ and consider $\mathcal{F}(t_n)$, where $\mathcal{F}(\cdot)$ denotes the the continuous time branching processes. Fix any two (integer) times $0 \leq t_0 < t_1 \leq t_n$. Let H denote* *the forest consisting of all edges such that both nodes that make the edge were born in the time interval* $[t_0, t_1]$*. Let* $\Delta_1(H)$ *denote the maximal degree of this graph. Then for any sequence* K_n

$$\mathbb{P}(\Delta_1(H) > K_n) \leq \sum_{t_0}^{t_1} Ce^{2m} e^{-K_n B e^{t_2 - m}}$$

for some universal constants C *and* B *(depending on the parameter* a *but independent of* n*).*

Proof: This is exactly what Equation (24) states.

■

Proof of Theorem 7(a):

Now we shall now derive asymptotics for the maximal eigenvalues. The proof follows along the same conceptual lines as [20]. However the technical tools are completely different in the sense that we shall use our continuous time embedding to get a completely different proof which generalizes very easily to the $a > 0$ setup.

The following Lemma collects the various linear algebra facts we need.

Lemma 49 *If* G *is a* d *star (namely a tree on* $d+1$ *vertices consisting of a single node connected to* d *leaves). Then* $\lambda_1(G) = \sqrt{d}$*. If* G *consists of* k *disjoint stars with the* i*'th star having* d_i *leaves with* $d_1 \geq d_2 \geq \cdots \geq d_k$*, then* $\lambda_i(G) = \sqrt{d_i}$*.*

Proof: The eigenvalues of a the adjacency matrix can be directly computed and this completes the proof.

Now the proof of Theorem 7(a) proceeds as follows: we shall construct subgraphs F_n of $\mathcal{T}_n$, which will be star forests such that

(A) $\lambda_i(F_n) = (1 + o(1))\sqrt{\Delta_i(\mathcal{T}_n)}$

(B) $\lambda_1(TT_n \setminus F_n) = o(1) \cdot n^{1/2\gamma_a}$

Note that Rayleigh's theorem states that for any graph $\mathcal{T}$

$$\lambda_i(\mathcal{T}) = \min_L \max_{x \in L, x \neq 0} \frac{\mathbf{x}^T A_{\mathcal{T}} \mathbf{x}}{\mathbf{x}^{\mathbf{T}} \mathbf{x}}$$

where L ranges over all $(n - i + 1)$ dimensional subspaces of $\mathbb{R}^n$. Thus for any partition F, $H = \mathcal{T} \setminus G$ of $\mathcal{T}$, we have

$$\min_L \max_{\mathbf{x} \in L, \mathbf{x} \neq 0} \frac{\mathbf{x}^T A_{\mathcal{T}} \mathbf{x}}{\mathbf{x}^{\mathbf{T}} \mathbf{x}} = \max_{\mathbf{x} \in L, \mathbf{x} \neq 0} \frac{\mathbf{x}^T A_F \mathbf{x}}{\mathbf{x}^{\mathbf{T}} \mathbf{x}} \pm O\left(\max_{x \neq 0} \frac{\mathbf{x}^T A_H \mathbf{x}}{\mathbf{x}^{\mathbf{T}} \mathbf{x}} \right)$$

Thus proving (A) and (B) is enough to prove the result. This the reasoning that allowed Fireze et al to conclude Theorem 7(a) for the case of the BA model. We shall now set about proving (A) and (B) for the more general linear preferential attachment model.

Consider the continuous time embedding of $\mathcal{T}_n$ in $\mathcal{F}(\cdot)$. Recall that we had $\mathcal{F}(T_n) = \mathcal{T}_n$ where $T_n = \inf\{t : |\mathcal{F}(t)| = n\}$. Also recall that we had refined asymptotics that said that $T_n \approx \frac{1}{\gamma_a} \log n$.

Call $t_n = \frac{1}{\gamma_a} \log n$. For ease of exposition we shall analyze the tree $\mathcal{F}(t_n)$ and shall find subsets $F_n \subset \mathcal{F}(t_n)$ of edges such that

(A) $\lambda_i(F_n) = (1 + o(1))\sqrt{\Delta_i(\mathcal{F}(t_n))}$

(B) $\lambda_1(\mathcal{F}(t_n) \setminus F_n) = o(1) \cdot \sqrt{\Delta_i(\mathcal{F}(t_n))}$

This will then complete the proof. Using Equation (21) and an analysis similar to the analysis of the maximal degree, this can be extended to the tree $\mathcal{F}(T_n) \stackrel{d}{=} \mathcal{T}_n$. Define the times t_0, t_1, t_2 as

$$t_0 = 0, t_1 = \frac{2}{3} t_n \quad t_2 = t_n - \log \log n \quad t_3 = t_n$$

For $i \geq 1$, let S_i denote the set of nodes born between times $[t_{i-1}, t_i]$.

Define F_n to be the star forest consisting of all edges where one end point is born before time t_1 and the other end point is born after time t_2. It is then easy to verify that with high probability

$$\Delta_i(F_n) = \Delta_i(\mathcal{F}(t_n)) - \frac{1}{\log n} O_P(\Delta_i(\mathcal{F}(t_n)))$$

Since $\Delta_i(\mathcal{F}(t_n)) = \Omega(e^{t_n})$, this implies that

$$\Delta_i(F_n) = (1 - o(1))\Delta_i(\mathcal{F}(t_n))$$

Since F_n is a star forest, we have by Lemma 49 that

$$\lambda_i(F_n) = \sqrt{\Delta_i(F_n)} = (1 + o(1))\Delta_i(\mathcal{F}(t_n))$$

Let $H_n = \mathcal{F}(t_n) \setminus F_n$ and let $H_{ij} = H_n(S_i S_j)$ be the set H_n restricted to (S_i, S_j) where the set (S_i, S_j) consists of those edges where one node is part of S_i and the other is part of S_j. Then it is easy to verify using Lemma 48 that

$$\lambda_1(H_{ij}) \leq \Delta_1(H_{ij}) = o(\Delta_k |\mathcal{F}(t_n)|)$$

This completes the proof.

5 Conclusion

Here we conclude with a wide ranging discussion and some open problems.

(a) Although we have not proved it explicitly here, using almost identical techniques to what we used to prove the convergence of the spectral distribution of the Adjacency matrix, we have the following result:

Theorem 50 *Suppose $\tilde{F}_n$ denotes the spectral distribution of the adjacency matrix of a family of random trees satisfying assumptions in Definition 2. Then there exists a probability distribution F on $\mathbb{R}$ such that*

$$\tilde{F}_n \stackrel{P}{\Rightarrow} F$$

Here convergence denotes convergence in the Levy-Prohrov metric on the space of measures on the real line.

(b) Although we have proved convergence of the distribution function as well as the maximum eigenvalues, we have not been able to derive any explicit properties of the limiting distributions. This seems to be another question worth pursuing. See [10] where there is a more complete discussion of the problems involved in solving the limiting recursive distributional equations. For some of the more classical models such as the uniform random tree on labelled vertices, there are close connections between the limiting Stieltjes transform and the classical theory of products of random matrices.

(c) **Analysis of the kernel for general random graphs:** Similar to Lemma 21, for more general graphs we have the following lemma. Note that a cycle cover of a graph (where we assume each edge is made of two directed edges) is the covering of the graph with disjoint directed cycles, where by disjoint we mean that the cycles do not share nodes. The weight of a directed cycle cover is the number of nodes present in the cycle cover.

Lemma 51 *Think of each edge of a graph as being composed of two directed edges. Then for a general graph the number of zero eigenvalues is the same as the number of nodes which do not belong to the maximal directed cycle cover of the graph.*

The following, which is probably already implicitly solved by Karp and Sipser seems like an important problem:

Conjecture 52 *For the Erdos-Renyii random graph $\mathcal{G}_n(n, \lambda/n)$, the maximal cover satisfies the following:*

For $\lambda < e$ we have

$$\mathbb{E}\left(\frac{|\mathcal{C}_n|}{n}\right) - \mathbb{E}\left(\frac{2M_n}{n}\right) \to 0$$

as $n \to \infty$. Further we have for the number of zero eigen values of $\mathcal{G}_n^\lambda$

$$\frac{\delta(G_n^\lambda)}{n} \longrightarrow_P a(\lambda)$$

as $n \to \infty$.

6 Acknowledgments

We would like to thank Charles Bordenave for many helpful discussions. S.B. would like to thank David Aldous for his constant encouragement and illuminating discussions regarding local weak convergence methodology.

References

[1] R. Albert and A.L. Barabási. Statistical mechanics of complex networks. *Reviews of Modern Physics*, 74(1):47–97, 2002.

[2] David Aldous. Asymptotic fringe distributions for general families of random trees. *Ann. Appl. Probab.*, 1(2):228–266, 1991.

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

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

[5] Z. Bai and J.W. Silverstein. *Spectral Analysis of Large Dimensional Random Matrices.* Science Press, 2006.

[6] M. Bauer and O. Golinelli. On the kernel of tree incidence matrices.

[7] M. Bauer and O. Golinelli. Random Incidence Matrices: Moments of the Spectral Density. *Journal of Statistical Physics*, 103(1):301–337, 2001.

[8] Shankar Bhamidi. *Universal Techniques for Preferential Attachment: Local and Global Analysis* . Preprint, 2007.

[9] Béla Bollobás and Oliver M. Riordan. Mathematical results on scale-free random graphs. In *Handbook of graphs and networks*, pages 1–34. Wiley-VCH, Weinheim, 2003.

[10] C. Bordenave and M. Lelarge. Resolvent of Large Random Graphs. 2007.

[11] Charles Bordenave. Eigenvalues of euclidean random matrices, 2006.

[12] Phillip Botti and Russell Merris. Almost all trees share a complete set of immanantal polynomials. *J. Graph Theory*, 17(4):467–476, 1993.

[13] M. Camarri and J. Pitman. Limit distributions and random trees derived from the birthday problem with unequal probabilities. *Electron. J. Probab*, 5(2):1–18, 2000.

[14] F. Chung, L. Lu, and V. Vu. Eigenvalues of Random Power law Graphs. *Annals of Combinatorics*, 7(1):21–33, 2003.

[15] F. Chung, L. Lu, and V. Vu. Spectra of random graphs with given expected degrees. *Proceedings of the National Academy of Sciences*, 100(11):6313–6318, 2003.

[16] F.R.K. Chung. *Spectral Graph Theory.* American Mathematical Society, 1997.

[17] C. Cooper and A. Frieze. A general model of web graphs. *Random Structures and Algorithms*, 22(3):311–335, 2003.

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

[19] S. N. Evangelou and E. N. Economou. Spectral density singularities, level statistics, and localization in a sparse random matrix ensemble. *Phys. Rev. Lett.*, 68(3):361–364, Jan 1992.

[20] T. Fenner, A. Flaxman, and A. Frieze. High Degree Vertices and Eigenvalues in the Preferential Attachment Graph, 2003.

[21] Peter Jagers and Olle Nerman. The growth and composition of branching populations. *Adv. in Appl. Probab.*, 16(2):221–259, 1984.

[22] Frederick A. Matsen and Steven N. Evans. Ubiquity of synonymity: almost all large binary trees are not uniquely identified by their spectra or their immanantal polynomials, 2005.

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

[24] Olle Nerman and Peter Jagers. The stable double infinite pedigree process of supercritical branching populations. *Z. Wahrsch. Verw. Gebiete*, 65(3):445–460, 1984.

[25] Walter Rudin. *Real and complex analysis*. McGraw-Hill Book Co., New York, third edition, 1987.

[26] Robert T. Smythe and Hosam M. Mahmoud. A survey of recursive trees. *Teor. Ĭmovīr. Mat. Stat.*, (51):1–29, 1994.

[27] J. Michael Steele. *Probability theory and combinatorial optimization*, volume 69 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1997.