Universal techniques to analyze preferential attachment trees: Global and Local analysis

Shankar Bhamidi

August 19, 2007

Abstract

We use embeddings in continuous time Branching processes to derive asymptotics for various statistics associated with different models of preferential attachment. This powerful method allows us to deduce, with very little effort, under a common framework, not only local characteristics for a wide class of scale free trees, but also global characteristics such as the height of the tree, maximal degree, and the size and structure of the percolation component attached to the root. We exhibit our computations for a number of different graph models of preferential attachment. En-route we get exact results for a **large** number of preferential attachment models including not only the usual preferential attachment but also the preferential attachment with fitness as introduced by Barabasi et al ([6]) and the Competition Induced Preferential attachment of Berger et al ([5]) to name just two. While most of the techniques currently in vogue gain access to the asymptotic degree distribution of the tree, we show how the embedding techniques reveal significantly more information both on local and global characteristics of these trees. Again very soft arguments give us the asymptotic degree distribution and size of the maximal degree in some Preferential attachment network models (not just trees) formulated by Cooper and Frieze [11] and van der Hofstad et al [12]. In the process we find surprising connections between the degree distributions, Yule processes and α-stable subordinators. We end with a number of conjectures for the asymptotics for various statistics of these models including size of the maximal component in percolation on these trees.

Key words. Preferential attachment, branching processes, Fringe distribution, Poisson process, random trees, multi-type branching processes, percolation, age-dependent branching processes.

Contents

1 Introduction

In the last few years, coupled with the explosion in data on "real-world" networks, there has also been a phenomenal growth in the various network models proposed to explain the structure, growth and development of these networks. At a theoretical level, it is hoped that the study of these networks, as well as various dynamics on these models, including routing and ranking algorithms, models of congestion and flow and epidemic models would give us a clearer picture of the behavior of these processes on the real networks. In this context one model of network growth that has received a significant amount of attention is the preferential attachment model.

Preferential attachment models: At the simplest level the network grows as time evolves; new vertices are added into the network, and a new node joins the network by forming (a fixed number say one) links to existing nodes with some probability depending on some characteristic of these nodes, usually the degree or some function of the present degree of the existing nodes. See Section 2.1 for more details.

Our contribution: To a large extent many of the existing (rigorous) theoretical studies on these models focus solely on the degree distribution. Using very clever ideas, largely from Polya Urn theory coupled with concentration inequalities (see e.g. [5],[9]), researchers are able to prove the convergence of the degree distribution to some distribution (usually a power law) as the size of the network grows large. However these techniques don't seem powerful enough to handle more complicated functionals of the network wether they be local functionals such as the structure of the network about a typical vertex upto some distance k from the vertex, or more global functionals such as the diameter of the network, maximal degree, and degree of the root to name just a few. This is the ground we plan to break.

The essential fact we shall use is that many of the preferential attachment models considered in the literature can be embedded in continuous time branching processes $(\mathcal{F}_t)_{(t>0)}$ (see Section 6 for more details) via appropriate formulations of the rate of offspring birth of these processes; thus writing $\mathcal{G}_n$ as the network tree on n vertices and letting T_n denote the stopping time:

$$T_n = \inf\{t > 0 : |\mathcal{F}_t| = n\}$$

we have

$$\mathcal{F}_{T_n} =_d \mathcal{G}_n$$

While some studies have used quite clever formulations of continuous time Markov Chain ideas (see [5],[9]) to explore the degree distribution there does not seem to be any work done in exploring the connection between continuous time branching processes and preferential attachment trees. Our contention is that this embedding technique gives us *almost everything we would ever want to know it terms of asymptotics* for these networks, via a few very simple calculations. Though there have been beautiful proofs using the more abstract theory of age-dependent branching processes in some special cases interspersed in the literature([30] for the height) this are seems to be largely unexplored at least in the context of preferential attachment. Thus this study has two main aims:

(a) Bring to the attention of theoreticians the sheer power of this more abstract approach via explicit computations in a wide variety of models. Some of the following results (mainly regarding the degree distribution) are already known in the literature. However we perform far more intricate computations which have the potential of revealing much more information about the local structure of the network.

(b) Push the envelope and do more complicated analysis including global computations such as asymptotics for the maximal degree, degree of the root and structure of the percolation component of the root.

Other than the initial conceptual idea of embedding the attachment trees into continuous time branching process, none of the remaining computations are very hard. The basic point of this paper is that after the initial learning curve involved in understanding various properties of continuous time branching processes, almost all the subsequent calculations are almost trivially easy. In some sense much of the hard work has been done and asymptotics for at least the local structure of the network follows from the *stable age distribution theory* of general continuous time branching processes (see e.g. [19, 25, 2]). One of the basic point of this paper is that for many of the preferential attachment models, one can do explicit calculations using such results to get an enormous amount of asymptotic information about preferential attachment tree and network models. Global results such as height and size of the maximal degree, although requiring slightly more work can still be carried out in all the models we study. It is also not very clear to the author if there is possibly any other technique which gives the same amount of information, essentially because in almost all the cases, the *Malthusian growth* parameter corresponding to the particular model in question seems to play such a significant role. To cite just three examples,

(a) We find that in many of the models the percolation properties of the model depend on two quantities, the Malthusian growth of the original process and the Malthusian rate of growth of an associated pruned branching process. Both quantities do not seem to have an easy combinatorial interpretation and thus it is difficult to imagine getting information on the percolation properties of these trees other than through these continuous time branching process arguments. See e.g. Theorem 19 for more details.

(b) The maximal degree in a number of models turns out to be of order dictated by the Malthusian growth parameter, namely is of order n^γ where γ is the Malthusian rate of growth of the model. See e.g. Theorems 14, 18 and in full generality, Theorem 71 for more details.

(c) Finally the results on the fitness models also suggests that the maximal degree of the tree has asymptotics almost completely governed by the Malthusian rate of growth parameter. See Theorem 29 and 31 for more details

Finally we find that these methods are powerful enough to deal with the asymptotic degree distribution not just in the tree models but also in some models of growing networks. Also see Section 7 where we compare the techniques we use with the techniques currenly in vogue.

1.1 Related literature

Due the enormous amount of interest in these models and the explosive growth of research in this area it seems a fruitless task to provide a complete bibliography of all work, rigorous or otherwise. As we describe each of the models we analyze we shall try to give suitable references to the particular model. However here we shall mention some of the more wide-ranging works including those particularly pertinent to our work.

In the past few years along with the proposal of new preferential attachment schemes, there has also been a concerted effort by a number of mathematicians to both rigorously prove conjectured results on these models as well as discover further properties of these models. For the Barabasi-Albert preferential attachment model and in the more general case of the linear preferential attachment model we should mention the fundamental contribution of Bollobas and Riordan (see e.g. [8], [7] to name just two works in their far ranging effort).

In developing new schemes of preferential attachment as well as rigorously proving various aspects of other preferential attachment schemes, we have been heavily influenced by Christian Borgs, Jennifer Chayes and their co-workers. Rigorous results on the degree distribution for the preferential attachment scheme with fitness and the entire theoretical basis for the connection between the competition induced scheme and preferential attachment as wells as the computation of the asymptotic degree distribution of this scheme is due to them (see [9] and [5] respectively). We have mentioned only two of their far-reaching efforts to both derive rigorous results for the preferential attachment scheme as well as discovering how preferential attachment schemes underlie far more complicated attachment schemes.

In the recent past there has also been a parallel program to develop preferential attachment models in the spatial context. See [16] and the references therein. The techniques developed in this paper do not carry over to this setting however.

In terms of more general references, the studies closest to this methodology is the book [14] where branching random walk techniques are used to explore functionals such as the height for the Barabasi-Albert preferential attachment model, although the emphasis is still largely on recursions. As mentioned before [31] already uses embedding techniques to compute the limiting degree distribution and load for the Barabasi-Albert Preferential attachment tree and the Linear preferential attachment tree. Also see [23] for a nice historical account of generative power law models. Another good source of rigorous work is the forthcoming book [10].

Finally on the Statistical physics end of things see [22] for some examples of the models looked at and an exposition of the recursion method to derive asymptotics. Also [1],[26] and [13] provide beautiful surveys and extensive references. We define the various preferential attachment models in Section 2.1 and along the way, shall try to mention the works where the corresponding preferential attachment schemes were first defined.

1.2 Plan of the paper

Section 2.1 describes the basic construction of these models as well as the various cases we shall be dealing with. Due to a lack of space we shall not describe in overmuch detail the actual *applied motivations* for the creation of the various models but shall give representative references as to origin of each model model and rigorous computations if any related to the model. Section 3 delves into various aspects of random tree methodology and particular constructions of infinite trees which will be useful for stating our main local structural results. In Section 4 we describe the functionals of the trees we will be interested in. Section 5 contains all the results which we have obtained from this theory. Note that the results regarding the limiting degree distribution known, however the method of proof here is novel as compared to the usual techniques in vogue (Although see [31] where similar computations for the linear preferential attachment model have been done for the degree distribution). More importantly the degree distribution is just one among many functionals that we get asymptotics for. We devote a reasonable amount of space for the Standard preferential attachment model since this is what has been studied to the greatest degree. However similar results carry over almost trivially to most of the other modes. Section 6 contains the various results on branching process methodology that helps us in proving our main results and proofs of the actual results. Due to the large number of results, we have tried to show proofs illustrating the sort of computations that need to be performed and have skipped proofs which add nothing new to the discussion. Note that for one particular model (preferential attachment with fitness) we need some *multi-type* branching process methodology and thus deal with this in a separate section, Section 6.4. We finally conclude in Section 7 with a wide ranging discussion about the drawbacks of this method, how it compares with the current techniques being used, future work, conjectures and open problems.

2 Attachment models

2.1 Preferential attachment trees

We now outline the mathematical models of growing trees that we shall be interested in.

Basic Construction of Preferential Attachment trees:

The basic construction of these trees is as follows

- Start with one node at time one labeled 1 .
- At time n we have a tree on n vertices $[n] := \{1, 2, \ldots, n\}$. A new node $n+1$ is added to the tree by attaching one link from the new node to a pre-exisiting node. The existing node v is chosen from $[n]$ with probability proportional to some function $f(v, n)$. Call this function the **attractiveness** of node v at time n.

Depending on the choice of the function $f(v, n)$ we get different models. Since we will be mainly dealing with trees which we think of growing (directed) downwards away from the root, we shall let

$$D(v, n) = \texttt{out} - \texttt{deg}(v, n)$$

i.e. $D(v, n)$ is the out-degree of the vertex v at time n. Note that the actual degree of any vertex (other than the root) is just $D(v, n) + 1$.

Different Models

Albert-Barabasi-Simon-Yule Preferential attachment: This model is given by the attractiveness function $f(v, n) = D(v, n) + 1$. See Durrett [14] and the extensive references cited within for more details for it's extensive use and origin. For simplicity we shall refer to this model as the BA-model, However it has it's origins dating back to the beginning of the 20th century.

Linear Preferential Attachment model: This is modification of the previous model with $f(v, n) = [D(v, n) + 1] + a$ where $a \geq 0$. Again see Durrett [14] for origins etc.

Preferential attachment with additive fitness: Fix a probability measure ν on $\mathbb{R}^+$. We think of each node v being born with some fitness f_v chosen independently from ν and which remains equal to this value after birth. At time n an existing node v has attractiveness $f(v, n) = (D(v, n) + 1) + f_v$. See [15] where this model was first introduced.

Preferential Attachment with multiplicative fitness: Fix a probability measure ν on $\mathbb{R}^+$. We think of each node v being born with some fitness f_v chosen independently from ν and which remains constant and equal to this value after birth. At time n an existing node v has attractiveness $f(v, n) = f_v(D(v, n) + 1)$. See [6] and [9] for more details.

Competition Induced Preferential attachment model: Fix an integer A. Then this model is given by the following Markov scheme:

- There are two types of nodes, fertile and infertile nodes. Only fertile nodes are able to reproduce at the next step while infertile nodes remain dormant.
- At time $t = 1$ start with one fertile node.

- At each time t a new node is born which is designated as infertile. It is attached to an existing **fertile** node v with probability proportional to

$$\min(D(v,t)+1, A)$$

where as before $D(v,n)$ is the out-degree of the fertile node. Note this includes both fertile and infertile nodes.

- After attaching this newly infertile node if the number of infertile children attached to node v is greater than $A-1$ then choose one of these infertile children uniformly at random and make it fertile.

It turns out that this model is equivalent to an attachment model on the real line which starts with one node at the origin and where nodes arrive at uniform points on the unit interval and decide to attach themselves to pre-exsting nodes by minimizing some linear combination of hopcount to the origin and euclidean distance to the existing potential parent. See [5] for more details where the degree distribution is worked out. We shall see that the continuous time branching process approach results in much more information both locally and globally.

Sub-linear Preferential attachment: Here the attractiveness function is given by $f(v,n) = (D(v,n)+1)^{\alpha}$ where $0 < \alpha < 1$. See [14] and the references therein for known results.

Preferential Attachment with a cutoff: Fix some integer $A > 1$. Here the attractiveness is as in the Preferential Attachment scheme until the out-degree of a node reaches A and then remains constant. More precisely:

$$\begin{aligned} f(v,n) &= D(v,n)+1 \text{ if } D(v,n) \leq A \\ &= A+1 \text{ if } D(v,n) > A \end{aligned}$$

Note that in all of the above examples the attractiveness function can be written as

$$f(v,n) = f(D(v,n))$$

namely the attractiveness function at any time point is given by a (possibly random as in the fitness model) function of the degree at time n.

Definition 1 (Point process associated with an attractiveness function) *Given an attractiveness function as above the associated Point process on $\mathbb{R}^+$ is a Markov point process with the function f as it's rate function. More particularly $\mathcal{P}$ is a point process on $\mathbb{R}^+$ with*

$$\mathbb{P}(\mathcal{P}(t+dt) - \mathcal{P}(t) = 1 | \mathcal{P}(t) = k) = f(k).dt + o(dt)$$

with the initial condition $\mathcal{P}(0) = 0$.

Thus for example the point process associated with the random multiplicative fitness model is a point process which is constructed as follows:

(a) At time $t = 0$ choose a random variable X from the distribution ν.
(b) Conditional on X

$$\mathbb{P}(\mathcal{P}(t + dt) - \mathcal{P}(t) = 1 | \mathcal{P}(t) = k) = X \cdot (k + 1)$$

A crucial role in many of the computations is played by the standard Yule process which we define now:

Definition 2 *The* **Yule process** *with rate γ for some constant $\gamma > 0$ is the point process with the conditions:*
(a) $\mathcal{P}(0) = 0$
(b) $\mathcal{P}$ is a point process on $\mathbb{R}^+$ is a point process with

$$\mathbb{P}(\mathcal{P}(t + dt) - \mathcal{P}(t) = 1 | \mathcal{P}(t) = k) = \gamma \cdot (k + 1).dt + o(dt)$$

If $\gamma = 1$ then we shall call such a process the standard Yule process.

Remark: The Yule process shall appear as one of the central constructs in this paper. It's connection to preferential attachment models was observed quite early, see [3] for a nice account. The general practice is to assume the parametrization $\mathcal{P}(0) = 1$ and the rate function $f(k) = k$ instead of $k + 1$ but we find the above parametrization more convenient to state results.

Definition 3 *To deal with the Competition induced preferential attachment model we have to make a slight modification. Note that the first a node produces a fertile node only after producing $A - 1$ infertile nodes. The time to produce this fertile node has the same distribution as that of the amount of time it takes a Yule process to produce A nodes namely $\sum_1^A \frac{Y_i}{i}$ where $Y_i \sim \exp(1)$ distribution. After producing the first fertile node, each of the following nodes are fertile and are produced at rate A. Thus the associated point process of fertile nodes can be modeled by the following point process on $\mathbb{R}^+$.*
(a) Let $S_1 = \sum_1^A \frac{Y_i}{i}$
(b) $S_{i+1} = S_i + \frac{Y_{i+1}}{A}$.

Now consider $(S_1, S_2, \ldots,)$ as a point process on $\mathbb{R}^+$. This gives the times of birth of a fertile offspring of a node.

2.2 Preferential attachment networks

If each new node comes with more than one edge to connect to the network then instead of growing trees, one can also define a growing network. We analyze the particular case where the new node attaches to existing nodes with probability proportional to the present degree of the vertex, namely the analogue of the Albert-Barabasi-SImon-Yule model for networks. Note that various aspects of this model have been analyzed in fundamental papers see e.g. [8, 7]:

Construction 4 *(a) Let $(W_i)_{i \geq i}$ be iid positive integer valued random variables.*
(b) Start with a single node having W_1 self loops.
(c) Recursively construct a network as follows; once we have the network on n nodes, create a new node $(n + 1)$ with degree W_{n+1}. Let the out edges of this node be denoted by $1, 2, \ldots W_{n+1}$.

Edge 1 is attached to one of the n nodes with probability proportional to the degree. We think of node $n+1$ now having it's existence and starting out with degree 1. Recursively attached the other edges to nodes with probability proportional to the current degree, update the degree and iterate till all W_{n+1} edges have been attached.

This general setup is due to [11] and the particular formulation is from [12].

The reason as to why our methods easily carry over to this setup is because of the following fact which follows immediately from the above construction.

Lemma 5 *Let $\mathcal{N}_n$ denote the network constructed on n nodes. Then one construction of the above network is as follows:*
(a) Generate $(W_i)_{1\leq i\leq n}$. Let $S_n = \sum_1^n W_i$.
(b) Construct the tree on $\mathcal{G}_{S_n}$ nodes using the Preferential attachment scheme.
(c) Let $S_i = \sum_1^i W_i$ with $S_0 = 0$. Merge nodes $(S_{i-1}+1,\ldots,S_i)$ and call this node i.

Proof: Obvious.

Note that the above networks are allowed to have self-loops so that edges between the nodes $(S_{i-1}+1,\ldots,S_i)$ are retained. Such self-loops of course become extremely rare as the size of the network grows.

3 Methodology related to trees

Here we briefly describe various notations regarding both deterministic and random models of trees that we use in our study.

Continuous time Branching process: There are a number of different ways to define a Markov process with age dependent branching structure. We shall stick to the *Point process* model which is defined as follows:

Definition 6 *Fix a point process $\mathcal{P}$ on $\mathbb{R}^+$. Then a continuous time branching process $\mathcal{F}_t$, driven by $\mathcal{P}$ is defined as follows: We start with one individual at time zero. This individual gives birth to individuals at times of the point process $\mathcal{P}_0$ where $\mathcal{P}_0 =_d \mathcal{P}$. Each individual born has its own iid copy of the point process and gives birth to individuals at these times after its birth. The continuous time branching process associated with a particular preferential attachment model with attractiveness function f is a branching process as above where the driving point process is as given in Definition 1.*

Note that $\mathcal{F}_t$ contains the complete information of the population upto time t, namely not only does it contain information about the number of individuals born upto time t but also the entire genealogy of the population upto time t.

On most occasions we shall think of the point process in its equivalent formulation as a sequence of random variables $0 < S_1 \leq S_2 \leq \ldots$.

Sin trees: We restate the following definition of `sin`-trees from [2]. Call an infinite rooted tree (where each node has finite degree) a `sin`-tree (for single infinite path) if for each vertex w there exists a unique infinite path $(w_0 = w, w_1, w_2, \ldots)$ from w. The following definition encompasses a particular construction of a sin-tree.

Definition 7 *Fix $\alpha > 0$, a probability measure μ on $\mathbb{R}^+$ and a point process $\mathcal{P}$ on $\mathbb{R}^+$. Then we call the following construction an $(\alpha, \mu, \mathcal{P})$ random sin-tree (rooted at 0) :*

- *Let $X_0 \sim Exp(\alpha)$ and for $i \geq 1$, $X_i \sim \mu$. Consider the renewal process (started with a non uniform start) $S_n = \sum_0^n X_i$.*
- *Conditional on the sequence $(S_n)_{n\geq 0}$ let*
  1. *$\mathcal{F}_{X_0}$ be a continuous time branching process driven by $\mathcal{P}$ observed upto time X_0;*
  2. *for $n \geq 1$ let $\mathcal{F}_{S_n,S_{n-1}}$ be a continuous time branching process observed upto time S_n with the only difference being that the distribution of the points of birth of the first ancestor is $\mathcal{P}$ conditioned to have a birth X_n time units after the birth of the founding ancestor.*
  3. *Given $(S_n)_{n\geq 0}$ construct $\mathcal{F}_{X_0}, \mathcal{F}_{S_1,S_0}, \mathcal{F}_{S_2,S_3}, \ldots$ independent of each other.*
- *Finally construct the sin-tree via the following procedure; Let the infinite path be the integer lattice $0, 1, 2, \ldots$ with 0 designated as the root. Consider $\mathcal{F}_{X_0}$ to be rooted at 0 and for $n \geq 1$ consider $\mathcal{F}_{S_n,S_{n-1}}$ to be rooted at n.*

Denote such an infinite random tree by the notation $\mathcal{T}^{\mathtt{sin}}_{\alpha,\mu,\mathcal{P}}$. We shall call

$$\mathbf{f}_k(\mathcal{T}^{\mathtt{sin}}_{\alpha,\mu,\mathcal{P}}) = (\mathcal{F}_{X_0}, \mathcal{F}_{S_1,S_0}, \mathcal{F}_{S_2,S_3}, \ldots, \mathcal{F}_{S_k,S_{k-1}})$$

the `sin`*-tree $\mathcal{T}^{\mathtt{sin}}_{\alpha,\mu,\mathcal{P}}$ truncated at distance k on the infinite path.*

Although the definition seems slightly complicated, conceptually it is very simple: See Figure 1 for more details.

To make precise the idea of "convergence of the local neighborhood of rooted random tree about a random vertex " we restate Aldous's [2] definitions of fringe and extended fringe namely:

Definition 8 Fringe and extended fringe for finite trees: *Given a finite rooted tree $\mathcal{T}$ define the fringe subtree at a vertex v is defined as follows: Suppose v is at height h, let $(v = v_0, v_1, \ldots v_h = \rho)$ be the unique path from v to the root ρ. Let $b_0(\mathcal{T}, v)$ as the subtree consisting of all those vertices s such that the unique path from the root to s pass through v. This is called the* **fringe** *subtree rooted at v. For $1 \leq i \leq h$ let $b_i(\mathcal{T}, v)$ be the set of vertices x such that the path from the root to x hits v_i but does not hit v_{i-1}. The total collection $\mathbf{f}(\mathcal{T}, v) = (b_0(\mathcal{T}, v), b_1(\mathcal{T}, v), \ldots, b_h(\mathcal{T}, v))$ will be called the* **extended fringe** *of $\mathcal{T}$ rooted at v.*

See Figure 2 for details.

We come to probably the most important conceptual definition in this paper, namely the concept of convergence of trees, and extended fringes both in expectation and in empirical proportions.

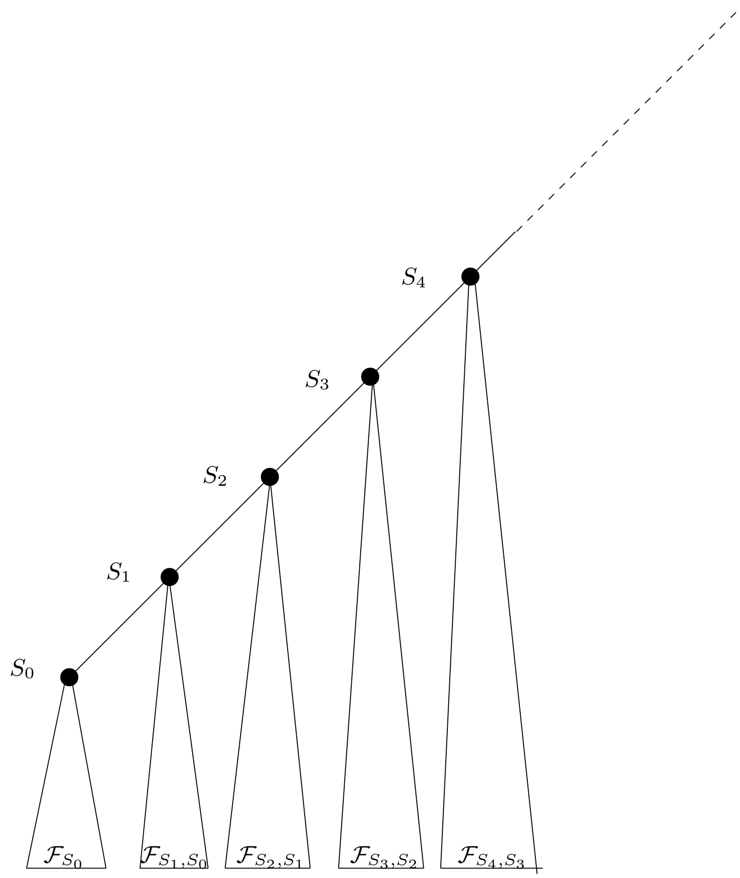


Figure 1: **The `sin`-tree construction**

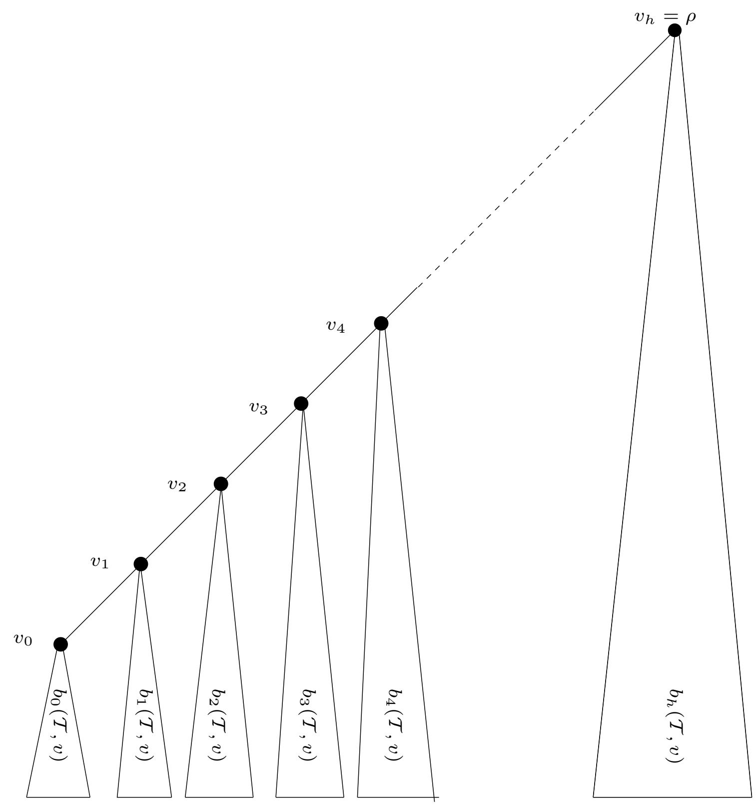


Figure 2: **The fringe and extended fringe construction**

For this we need some notation. Let $T_{\mathtt{tr}}$ be the set of finite rooted trees. Then note that $T_{\mathtt{tr}}^{\infty}$ corresponds to the space of `sin`-trees.

We will think of the extended fringe as a map from the tree space $T_{\mathtt{tr}}$ to the infinite product space $T_{\mathtt{tr}}^{\infty}$ via the map $\mathcal{T}$ to $\mathbf{f}(\mathcal{T}, v) = (b_0(\mathcal{T}, v), b_1(\mathcal{T}, v), \ldots, b_h(\mathcal{T}, v), \emptyset, \emptyset, \emptyset, \ldots)$. Call $\mathbf{f}_k(\mathcal{T}, v) = (b_0(\mathcal{T}, v), b_1(\mathcal{T}, v), \ldots, b_k(\mathcal{T}, v)$ be the extended fringe truncated at distance k on the path from v to the root.

Definition 9 (convergence of random trees) *Fix a random tree $\mathcal{T}_n$ on $|\mathcal{T}_n|$ vertices. We shall always use V_n to denote a vertex chosen uniformly at random among all the $|\mathcal{T}_n|$ nodes. We shall always assume that for our sequence of trees, the size explodes namely $|\mathcal{T}_n| \longrightarrow_d \infty$. (In most of the cases that we deal with $|\mathcal{T}_n| \uparrow \infty$ a.s.) Fix any $k \geq 1$ and trees $(t_1, t_2, \ldots t_k) \in \mathcal{T}_{\mathtt{tr}}^{k}$.*

We define two notions of convergence of the random trees $\mathcal{T}_n$ to the infinite sin tree $\mathcal{T}_{\alpha,\mu,\mathcal{P}}^{\mathtt{sin}}$:

(a) **Mean convergence:** *Say that a sequence of random trees $\mathcal{T}_n$ converges in mean fringe sense to the* `sin` *tree $\mathcal{T}_{\alpha,\mu,\mathcal{P}}^{\mathtt{sin}}$ if for any $k \geq 1$ and $(t_1, t_2, \ldots t_k)$*

$$\mathbb{P}(\mathbf{f}_k(\mathcal{T}_n, V_n) = (t_1, t_2, \ldots t_k)) \longrightarrow \mathbb{P}\left[\mathbf{f}_k(\mathcal{T}_{\alpha,\mu,\mathcal{P}}^{\mathtt{sin}}) = (t_1, \ldots t_k)\right]$$

as $n \to \infty$. We shall denote this notion of convergence as

$$\mathcal{T}_n \longrightarrow_{\mathbf{mn.-fr.}} \mathcal{T}_{\alpha,\mu,\mathcal{P}}^{\mathtt{sin}}$$

(b) **Convergence in probability:** *Say that a sequence of random trees $\mathcal{T}_n$ converges in probability-fringe sense to $\mathcal{T}^{\mathtt{sin}}_{\alpha,\mu,\mathcal{P}}$ if the empirical proportions converge in probability namely*

$$\frac{1}{|\mathcal{T}_n|}\sum_{v\in\mathcal{T}_n} 1_{\{\mathbf{f}_k(\mathcal{T}_n,v)=(t_1,\dots,t_k)\}} \longrightarrow_P \mathbb{P}\left[\mathbf{f}_k(\mathcal{T}^{\mathtt{sin}}_{\alpha,\mu,\mathcal{P}}] = (t_1,\dots t_k)\right)$$

as $n\to\infty$.

Denote this convergence by

$$\mathcal{T}_n \longrightarrow_{\mathbf{P-fr.}} \mathcal{T}^{\mathtt{sin}}_{\alpha,\mu,\mathcal{P}}$$

We record some simple facts of the above convergence.

Proposition 10 *(a) Convergence in probability-fringe sense implies convergence in mean-fringe sense.*
(b) Suppose $\mathcal{T}_n$ converges in mean-fringe sense to $\mathcal{T}^{\mathtt{sin}}_{\alpha,\mu,\mathcal{P}}$. Fix $k\geq 1$ and denote the number of nodes with degree k in $\mathcal{T}_n$ as $N_k(n)$. Then mean-fringe sense convergence implies

$$\mathbb{E}\left(\frac{N_k(n)}{|\mathcal{F}_n|}\right) \longrightarrow p_k$$

where $p_k = P(\deg(\rho)=k)$ where ρ is the root of $\mathcal{T}^{\mathtt{sin}}_{\alpha,\mu,\mathcal{P}}$.
(c) If the above convergence is in probability-fringe sense the this can be strengthened to

$$\frac{N_k(n)}{|\mathcal{F}_n|} \longrightarrow_P p_k$$

Proof: Obvious by the definition of mean-fringe sense and probability-fringe sense convergence.

Remark: Since we deal with trees where there is a natural sense of direction namely away from the root, we shall usually henceforth let $N_k(n)$ to be the number of nodes with out-degree k. Note that we only need to modify the above result so that p_k is the degree of the root in $\mathcal{F}_{X_0}$.

4 Functionals of interest

In this Section we briefly define the functionals we are interested in. Below we shall always use $\mathcal{T}$ to denote a tree. When we are interested in asymptotics we shall add a subscript n to denote a sequence of trees whose size is growing with n. We will be interested in the following functionals

Local Functionals: Let V_n denote a vertex picked uniformly at random from the tree. Then

- empirical degree distribution : Let $N_k(n)$ be the number of nodes with out-degree k in $\mathcal{T}_n$
- Fix an integer $k\geq 1$. Let $\mathcal{T}_k(V_n)$ denote the subtree consisting of all nodes within a distance k of V.

- Fringe subtree at V_n: Let $\mathcal{F}(V_n)$ denote the fringe subtree rooted at V_n.
- Load on V_n namely $Ld(V_n)$: This is defined as the number of vertices which have V_n in the shortest path from the root to the vertex, namely $Ld(V_n) = |\mathcal{F}(V_n)|$ where $\mathcal{F}(V_n)$ is the fringe subtree rooted at V_n.

Global Functionals

- Height of the tree $ht(\mathcal{T})$ or h_n.
- Degree of the root: $deg(\rho)$.
- Maximal degree of the tree: $\max(deg)$.
- Fix an integer k. Label the k^{th} vertex to be born by the integer k and let $S_n(k)$ denote the size of the subtree attached to vertex k at time n. We will occasionally abuse notation and call this the load on vertex k.

5 Results

Now we state the main results if this study.

A note about the organization of the paper: Since we wanted to make the paper "readable" and useful despite the number of different models analyzed we have taken the following editorial decision in the organization of the paper: When stating the results we shall deal with each individual model, state results of all the functionals analyzed for that model and then move on to the next model. When proving results however we take each functional say the height, prove the necessary result for that functional for all the models and then move onto the next functional. Since the method of proof for each functional is essentially the same across the models, we believe this makes the paper readable and concise.

5.1 Preferential attachment

Through out this section we shall let $\mathcal{T}_n$ denote the Preferential Attachment tree on n vertices. We shall also let V_n be a vertex in the tree chosen at random.

Theorem 11 (Detailed structure of the local neighborhood) *Let $\mathcal{T}_n$ denote the tree constructed via the preferential attachment method on n nodes. Then*

$$\mathcal{T}_n \longrightarrow_{\mathbf{P-fr.}} \mathcal{T}^{\mathtt{sin}}_{(2,Exp(1),\mathtt{Yule})}$$

i.e. the limiting sin tree has the same distribution as one of the random sin trees constructed in Section 3 with $\alpha = 2$, with the distribution on $\mathbb{R}^+$ identified as the $\exp(1)$ distribution and the associated point process is the rate one Yule process. Note that this convergence also takes place in the mean sense by Proposition 10.

Corollary 12 (Local properties) *Let* $L_k(n) = \sum_{v \in \mathcal{T}_n} 1\{|\mathbf{f}(\mathcal{T}_n, v)| = k\}$ *namely the number which have load* k *and let* $N_k(n) = \sum_{v \in \mathcal{T}_n} 1\{\mathtt{out}-\mathtt{deg}(v) = k\}$ *namely the number of nodes which have out-degree* k*. Then*

$$\frac{L_k(n)}{n} \longrightarrow_P \mathbb{P}(Z = k)$$

$$\frac{N_k(n)}{n} \longrightarrow_P \mathbb{P}(D = k)$$

The random variables Z *and* D*satisfy the following properties:*
(a) Let $g(s)$ *denote the generating function of* Z*. Then*

$$g(s) = 1 - \frac{1-s}{\sqrt{s}} \sinh^{-1}\left(\sqrt{\frac{s}{1-s}}\right)$$

(b) The random variable Z *has infinite expectation.*
(c) The degree distribution satisfies the following property:

$$P(D = k) = \frac{4}{(k+1)(k+2)(k+3)} \qquad \text{for } k = 0, 1, 2, \ldots$$

Before we state the next result we need some notation:

Let X be an exponential random variable with rate 2. Let $N_0(t), N_1(t), \ldots,$ be IID *standard* Yule processes independent of X.

Let T be another random variable independent of X_i with

$$P(T = k) = \frac{2}{k(k+1)} \qquad \text{for } k = 2, 3, \ldots$$

Finally define the random variables $Z_1 := N_0(X)$ and let $Z_2 := \sum_1^T N_i(X) + (T - 1)$. Note that because of the occurrence of the exponential random variable X, the two random variables are *not* independent.

Corollary 13 *For any node* v *let* $\mathtt{mother}(v)$ *denote the mother of the node* v*. For any two integers* $k \geq 0, l \geq 1$ *let* $X_{k,l}(n) = \sum_{v \in \mathcal{T}_n} 1\{\{\mathtt{out}-\mathtt{deg}(v) = k\} \cap \{\mathtt{out}-\mathtt{deg}(\mathtt{mother}(v)) = l\}\}$ *i.e the number of nodes whose degree is* k *and whose mother's degree is* l*.*

Then

$$\frac{1}{n} X_{k,l} \longrightarrow_P \mathbb{P}(Z_1 = k, Z_2 = l)$$

where Z_1 *and* Z_2 *are the random variables defined above.*

Remark: Note that Lemma 87 coupled with Corollary 13 implies that the asymptotic degree distribution of the mother of a uniformly chosen node has asymptotic degree distribution $q_k \sim 1/k^2$ namely with a degree two exponent.

Theorem 14 (Global properties) *(a) Let S_n^k denote the size of the subtree rooted at the k^{th} node to be born. Then*

$$\frac{S_n^k}{n} \longrightarrow_{a.s.} \beta(2, 2 \cdot (2k-1))$$

(b) Let $deg(\rho)$ denote the degree of the root vertex. Then

$$\frac{deg(\rho)}{\sqrt{n}} \longrightarrow_{a.s.} \frac{W}{\sqrt{\tilde{W}}}$$

Note that $2.\tilde{W} \sim \Gamma(2,1)$ distribution that is, it has the density xe^{-x} on $\mathbb{R}^+$.

Also note that W and $\tilde{W}$ are not independent but satisfy the following distributional identity: Let $\xi_1, \xi_2, \ldots$ be the points of birth of a rate one Yule process. Let $N(t)$ be the number of individuals of this Yule process born before time t. Let $\tilde{W}_1, \tilde{W}_2, \ldots$ independent of the ξ_i process, be independent and identically distributed as $\frac{1}{2}\Gamma(2,1)$ random variables. Then the limit random variables W and $\tilde{W}$ satisfy the following distributional identities:

$$\frac{N(t)}{e^t} \longrightarrow_{a.s.L^2} W$$

.

$$\tilde{W} = \sum_1^\infty e^{-2\xi_i} \tilde{W}_i$$

(c) Given any ϵ, $\exists$ a constant K_ε such that

$$\limsup_{n\to\infty} \mathbb{P}(\frac{\max(deg)}{\sqrt{n}} > K_\varepsilon) < \varepsilon$$

(d) The height of the tree, h_n satisfies the following asymptotics:

$$\frac{h_n}{\log n} \longrightarrow_P \frac{1}{2\gamma}$$

where γ is the unique positive root of the equation $\gamma e^{\gamma+1} = 1$

Note that (d) was proved in a very beautiful paper by Boris Pittel [30]. We generalize his technique to a much wider class of models.

It is very likely that our methods can be modified to prove this as well but we shall not explore this here.

Consider percolation on the tree namely retain each edge with probability p and delete otherwise. A component or percolation cluster is a connected tree of the resulting random forest. Denote by $\mathcal{C}_n(v)$ the size of the component attached to any vertex v and let $\max|\mathcal{C}_n|$ denote the size of the largest component. Then the next theorem says that the size of the percolation cluster at the root is of order $n^{\frac{1+p}{2}}$. Further it gives explicit structural limiting properties of this cluster including limiting degree distribution and asymptotics for the height of the cluster.

Theorem 15 (Percolation cluster connected to the root) *(a)Consider $\mathcal{C}_n(\rho)$ the percolation cluster at the root ρ. Then there exists a random variable $W_p > 0$ such that*

$$\frac{|\mathcal{C}_n(\rho)|}{n^{\frac{1+p}{2}}} \longrightarrow_{\mathcal{L}} W_p$$

as $n \to \infty$.
(b) Let $N_{\geq k}(n) = \sum_{v \in \mathcal{C}_n(\rho)} 1\{\mathtt{out} - \mathtt{deg}(\mathtt{v}) \geq k\}$, i.e. the number of nodes in the subtree corresponding to the percolation cluster about the root which have out degree greater than k. Then the empirical fraction

$$\frac{N_{\geq k}(n)}{|\mathcal{C}_n(\rho)|} \longrightarrow_P E\left[\prod_1^{\sum_1^k X_i} \left(\frac{i}{i+1+p}\right)\right]$$

as $n \to \infty$, where X_i are i.i.d. Geometric(p) random variables namely the have probability mass function

$$P(X = k) = p(1-p)^{k-1}$$

(c) There exists a constant $C_p > 0$ such that the height of the cluster

$$\frac{ht(C_n(\rho))}{\log n} \longrightarrow C_p$$

as $n \to \infty$.

Notation: To avoid repetition we shall let $N_k(n)$ be the number of nodes in the tree with out-degree k and $N_{\geq k}(n)$ the number of nodes with out degree greater than or equal to k.

5.2 Linear Preferential attachment

We first define the Malthusian rate of growth of the associated branching process in this setup:

Define

$$\gamma = \frac{a + 2 + \sqrt{a^2 + 4}}{2}$$

Note that γ satisfies the equation:

$$\frac{1}{\gamma - 1} + \frac{a}{\gamma} = 1$$

Define the following probability measure μ on $\mathbb{R}^+$ as the mixture of two exponential random variables with differing rates namely

$$\mu(*) = \frac{a}{\gamma}.\nu_{\exp(\gamma)}(*) + \frac{1}{\gamma - 1}.\nu_{\exp(\gamma-1)}(*)$$

Theorem 16 (Local properties) *For this model we have we convergence in probabilty fringe sense namely*

$$\mathcal{T}_n \longrightarrow_{\mathbf{P-fr.}} \mathcal{T}^{\mathtt{sin}}_{(\gamma, \mu, \mathcal{P}_{\mathbf{a}})}$$

as $n \to \infty$.

Corollary 17 (Limiting degree distribution) *Fix any $k \geq 0$. Then the empirical proportion of nodes with out degree k satisfies*

$$\frac{N_k(n)}{n} \longrightarrow_P \frac{\gamma}{(k+1+a)+\gamma} \prod_1^k \left(\frac{i+a}{i+a+\gamma} \right)$$

as $n \to \infty$. Note that the limiting distribution has a power law tail with exponent $\gamma + 1$.

Theorem 18 (Global properties) *(a) Let $deg(\rho)$ denote the degree of the root vertex. Then there exists a random variable $X > 0$ such that*

$$\frac{deg(\rho)}{n^{1/\gamma}} \longrightarrow_P X$$

(b) Given any ϵ, $\exists$ a constant K_ε such that

$$\limsup_{n\to\infty} \mathbb{P}(\frac{\max(deg)}{n^{1/\gamma}} > K_\varepsilon) < \varepsilon$$

Thus the maximal degree is of the same order as the degree of the root vertex. (c) There exists a constant $C_a > 0$ such that the height of the tree, h_n satisfies the following asymptotics:

$$\frac{h_n}{\log n} \longrightarrow_P C_a$$

Remark: Note that even is more is true namely Mori [24] proves that in both this and the Preferential attachment model (with $\gamma = 2$) there exists a random variable μ such that

$$\frac{\max(deg)}{n^{1/\gamma}} \longrightarrow_P \mu$$

It is probably true that our methods can be used to prove this stronger result here but we shall not explore the above in this study.

For simplicity we plot the graph of a versus the exponent γ^{-1} which seems to arise in many of the properties.

Theorem 19 (Percolation cluster about the root) *Fix $p > 0$. As before let $\mathcal{C}_n(\rho)$ denote the percolation component about the root. Define*

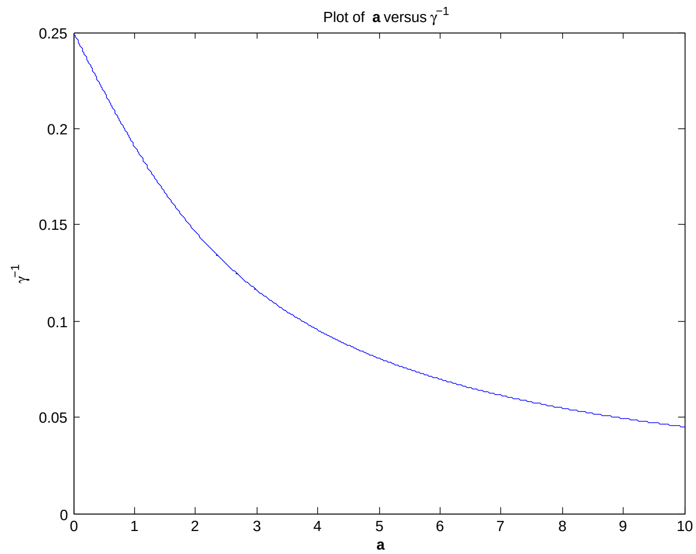
$$\gamma_p = \frac{1 + p(a+1) + \sqrt{(1+p(a+1))^2 - 4ap}}{2}$$

(a)There exists a random variable $W_{a,p} > 0$ such that

$$\frac{\mathcal{C}_n(\rho)}{n^{\frac{\gamma_p}{\gamma}}} \longrightarrow_P W_{a,p}$$

(b) Consider the number of nodes with out degree at least k in $\mathcal{C}_n(\rho)$. Then

$$\frac{N_{\geq k}(n)}{n} \longrightarrow_P E\left[\prod_1^{\sum_1^k X_i} \left(\frac{i+1+a}{i+1+a+\gamma_p} \right) \right]$$



as $n \to \infty$, where X_i are i.i.d. Geometric(p) random variables namely the have probability mass function

$$\mathbb{P}(X = k) = p(1-p)^{k-1}$$

(c) There exists a constant $C_{p,a} > 0$ such that the height of $\mathcal{C}_n(\rho)$ satisfies

$$\frac{ht(\mathcal{C}_n(\rho))}{\log n} \longrightarrow_P C_{p,a}$$

as $n \to \infty$.

5.3 Sub-linear Preferential attachment

Fix $0 < \alpha < 1$. Let $\theta = \theta(\alpha)$ be the unique solution to the equation

$$\sum_{1}^{\infty} \prod_{j=1}^{i} \left(\frac{j^{\alpha}}{\theta + j^{\alpha}} \right) = 1$$

Note that θ satisfies the following properties

Proposition 20 *(a)* $1 < \theta(\alpha) < 2$ *for all* $0 < \alpha < 1$
(b) $\theta(\alpha)$ *is a monotonically increasing function of* α.

We have the following structural results regarding these trees:

Theorem 21 (local properties) *Let* $\mathcal{T}_n$ *denote a random tree on* n *vertices, constructed via the sub-linear preferential attachment model. Then*

$$\frac{N_k(n)}{n} \longrightarrow_P \mathbb{P}(Z = k)$$

where Z *is an integer valued random variable with*

$$\mathbb{P}(Z \geq k) = \prod_{1}^{k} \frac{1}{1 + \theta/j^\alpha}$$

Note that the above implies that the tail of Z *satisfies*

$$\mathbb{P}(Z = k) \sim \frac{1}{k^\alpha} \exp(-\frac{\theta}{1-\alpha} k^{1-\alpha})$$

Note that in this setup as well there is convergence in the probability fringe sense to some `sin`-tree, however the sin measure does not have an easy expression so we do not state it explicitly.

Theorem 22 (Global properties) *(a) The random variable* $deg(\rho)$ *is* $O_P((\log n)^{\frac{1}{1-\alpha}})$. *More precisely*

$$\frac{deg(\rho)}{(\log n)^{\frac{1}{(1-\alpha)}}} \longrightarrow_P \left(\frac{1}{\theta(\alpha)}\right)^{\frac{1}{(1-\alpha)}}$$

where as before $\theta(\alpha)$ *is the Malthusian rate of growth.*

(b) Given any $\varepsilon > 0$ *there exists a constant* K_ε *such that*

$$\limsup_{n\to\infty} \mathbb{P}\left(\frac{\max(deg)}{(\log n)^{\frac{1}{1-\alpha}}} > K_\varepsilon\right) \leq \varepsilon$$

(c) There exists a constant $C_\alpha > 0$ *such that the height* h_n *satisfies the relation*

$$\frac{h_n}{\log n} \longrightarrow_P C_\alpha$$

as $n \to \infty$.

Remark: It might be a worthwhile task to see if one is able to explicitly derive the properties of the limit constants C_α that arise in the asymptotics for the height and their dependence on α. There is an explicit construction of the limit constant but in most of these models it does not seem amenable to any algebraic treatment.

5.4 Preferential attachment with a cutoff

Fix an integer $A > 1$. Denote by θ_A the unique positive root of the equation

$$\sum_1^{A-1} \prod_1^j \left(\frac{i}{i+\theta}\right) + \frac{A+\theta}{\theta} \prod_1^A \left(\frac{i}{i+\theta}\right) = 1$$

The solution θ_A satisfies the following properties

Proposition 23 *(a)* $1 < \theta_A < 2$ *for all* $A > 1$.
(b) θ_A *is a monotonically increasing function of* A

Theorem 24 (Local properties) *Let* θ_A *be as defined in Propn (23). Let* Π_A *be the associated point process on* $\mathbb{R}^+$ *and let* μ_A *be the measure on* $\mathbb{R}^+$ *defined by the density with respect to the Lebesgue measure as*

$$\mu_A(dx) = e^{-\theta_A x} \left[\sum_0^{A-1} (j+1) e^{-x} (1-e^{-x})^j + A(1-e^{-x})^A \right]$$

Then we have convergence in the probability fringe sense namely

$$\mathcal{T}_n \longrightarrow_{\mathbf{P-fr.}} \mathcal{T}^{\texttt{sin}}_{\theta_A, \mu_A, \mathcal{P}_A}$$

Corollary 25 (Limiting degree distribution) *For the number of nodes with out-degree* k *in the tree we have the following law of large numbers :*

$$\frac{N_k(n)}{n} \longrightarrow_P \frac{\theta_A}{\theta_A + k} \prod_1^{k-1} \left(\frac{i}{i+\theta_A}\right) \quad \textit{for } k \leq A$$

and

$$\frac{N_k(n)}{n} \longrightarrow_P \frac{\theta_A}{A+\theta_A} \left(\frac{A}{A+\theta_A}\right)^{k-1-A} \left[\prod_1^{k-1-A} \left(\frac{i}{i+\theta_A}\right) \right] \quad \textit{for } k > A$$

as $n \to \infty$

Theorem 26 (Global properties) *(a) For large* n *the root has asymptotic size* $\frac{A}{\theta_A} \log n$*. More precisely:*

$$\frac{deg(\rho)}{\log n} \to_P \frac{A}{\theta_A}$$

as $n \to \infty$.
(b) Given $\varepsilon > 0$ *we can choose a constant* K_ε *such that*

$$\limsup_{n \to \infty} \mathbb{P}\left(\frac{\max(deg)}{\log n} > K_\varepsilon\right) \leq \varepsilon$$

(c) For the height of the tree $\exists C_A > 0$ *such that*

$$\frac{h_n}{\log n} \longrightarrow_P C_A$$

5.5 Preferential attachment with fitness

In this Section we shall analyze only the "proper" case namely where the Malthusian parameter associated with the corresponding multi-tpe branching process actually exists. More precisely we make the following assumptions regarding the distribution ν on $\mathbb{R}^+$. We assume that the fitness assigned to each node is strictly positive namely:

$$\nu(0,\infty) = 1$$

Assumption 27 *(a) We assume in both the additive and multiplicative model that the probability measure ν is finitely supported namely there exists a constant $M > 0$ such that $\nu(M,\infty) = 0$. Note that this immediately implies that if $X \sim \nu$ then $E(X) < \infty$.*
(b) For the Multiplicative model we assume that the equation

$$\int_0^M \frac{s}{\alpha - s}\nu(ds) = 1 \tag{1}$$

has a solution for some $\alpha > M$. As a matter of notation call this root $\alpha_{\times,\nu}$, the "$\times$" symbol being used to remind us of it's connection to the multiplicative model. Define the new measure $\hat{\nu}_{\alpha_{\times,\nu}}$ via the equation:

$$\hat{\nu}_{\alpha_{\times,\nu}}(ds) = \frac{s}{\alpha - s}\nu(ds)$$

Note that by Equation (1), $\hat{\nu}_{\alpha_{\times,\nu}}$ is a probability measure.
(c) Assumptions 80 are satisfied for the fitness distribution and the associated multi-type branching process.

5.5.1 Additive model

This model was introduced in [15]. Compare this model to the Linear preferential attachment model. Note that in this model the attractiveness of a node at time n is $(D(v,n) + 1 + f_v)$ with f_v random whereas for the Linear Preferential attachment model we had attractiveness given by $(D(v,n) + 1 + a)$ where a was fixed and constant for all the nodes. Our results confirm the simulation based studies of [15] that at least the degree distribution looks similar to the Linear preferential model with parameter $a_\nu = \int_{\mathbb{R}^+} x \cdot \nu(dx)$ i.e. the mean of the fitness measure ν.

Throughout this section we let

$$\gamma_{+,\nu} = \frac{a_\nu + 2 + \sqrt{a_\nu^2 + 4}}{2}$$

Note that $\gamma_{+,\nu}$ is the unique positive solution of the equation:

$$\int_{\mathbb{R}^+} \left(\frac{1}{\gamma - 1} + \frac{r}{\gamma}\right)\nu(dr) = 1 \tag{2}$$

Define the measure $\hat{\nu}_{\gamma_{+,\nu}}$ by the equation

$$\hat{\nu}_{\gamma_{+,\nu}}(dr) = \left(\frac{1}{\gamma - 1} + \frac{r}{\gamma}\right)\nu(dr)$$

and note that Equation (2) implies that $\hat{\nu}_{\gamma_{+,\nu}}$ is a probability measure.

Theorem 28 *(a) Let $N_k(n)$ be the number of nodes with out- degree k in $\mathcal{T}_n$. Then*

$$\frac{N_k(n)}{n} \longrightarrow_P \int_0^M \left[\frac{\gamma_{+,\nu}}{k+1+r+\gamma_{+,\nu}} \prod_1^k \left(\frac{i+r}{i+r+\gamma_{+,\nu}}\right)\right] \nu(dr)$$

as $n \to \infty$.
(b) Fix any two sets $A, B \subset [0, M]$ and define

$$N_n(A,B) = \sum_{v \in \mathcal{T}_n} 1\{\texttt{fit}(v) \in A, \texttt{fit}(\texttt{mother}(v)) \in B\}$$

where $\texttt{fit}(v)$ denotes the fitness of a vertex v. Then the empirical proportions satisfy

$$\frac{N_n(A,B)}{n} \longrightarrow_P \mathbb{P}(Z_1 \in A, Z_2 \in B)$$

where Z_1 and Z_2 are independent, with $Z_1 \sim \nu$ and $Z_2 \sim \hat{\nu}_{\gamma_{+,\nu}}$.

Theorem 29 (Global properties) *(a) There exists a random variable $W_\nu > 0$ such that*

$$\frac{\deg(\rho)}{n^{\frac{1}{\gamma_{+,\nu}}}} \longrightarrow_{a.s.} W_\nu$$

as $n \to \infty$. Here as before $\gamma_{+,\nu}$ is the Malthusian rate of growth parameter.
(b) Given any $\varepsilon > 0$ there exists a constant K_ε dependent only on M and independent of ν such that:

$$\limsup_{n\to\infty} \mathbb{P}\left(\frac{\max \deg(\rho)}{n^{\frac{1}{\gamma_{+,\nu}}}} > K_\varepsilon\right) \le \varepsilon$$

(c) Consider the height h_n. Then there exists a constant $0 < C_\nu < \infty$ such that

$$\frac{h_n}{\log n} \longrightarrow_P C_\nu$$

Remark: It would be interesting to check if we find similar asymptotics for the size of the maximal degree if we did not assume finite support.

5.5.2 Multiplicative model

This model was introduced in [6] and the degree distribution rigorously analyzed in [9]. We first begin with asymptotic statements regarding the degree distribution:

Theorem 30 *(a) Let $N_k(n)$ be the number of nodes with out degree k. Then the empirical proportions satisfy*

$$\frac{N_k(n)}{n} \longrightarrow_P \int_0^M \left[\frac{\alpha_{\times,\nu}}{\alpha_{\times,\nu} + r\cdot(k+1)} \prod_1^k \left(\frac{r\cdot i}{\alpha_{\times,\nu} + r\cdot i}\right)\right] \nu(dr)$$

(b) As before for a vertex $v \in \mathcal{T}_n$, let $\texttt{mother}(v)$ be it's immediate predecessor in the tree. For $A, B \subset \mathbb{R}$ let

$$N_n(A,B) = \sum_{v \in \mathcal{T}_n} 1\{\texttt{fit}(v) \in A, \texttt{fit}(\texttt{mother}(v)) \in B\}$$

Then

$$\frac{N_n(A,B)}{n} \longrightarrow_P \mathbb{P}(Z_1 \in A, Z_2 \in B)$$

where Z_1 and Z_2 are independent with $Z_1 \sim \nu$ and $Z_2 \sim \hat{\nu}_{\alpha_{\times,\nu}}$.

Remark: We have analyzed completely, the "regular" setup where the Assumptions 80 are satisfied. Theorem 30(b) was originally proved in [9]. They also prove a similar result as Theorem 30(b) in the setup where the assumptions do not hold and a Malthusian growth parameter solving Equation (1) does *not* exist. Building on heuristics by [6], they identify a "Bose-Einstein" condensation phase where part of the mass escapes to M.

Theorem 31 (Global properties) *(a) The degree of the root node is of order $\Omega(n^{f_\rho/\alpha_{\times,\nu}})$. More precisely there exists a random variable $W_\nu > 0$ such that*

$$\frac{\deg(\rho)}{n^{\frac{f_\rho}{\alpha_{\times,\nu}}}} \longrightarrow_{a.s.} W_\nu$$

as $n \to \infty$. Here f_ρ is the fitness of the initial node and as before $\alpha_{\times,\nu}$ is the Malthusian rate of growth parameter.

(b)The correct asymptotic order of the maximal degree is $O(n^{\frac{m-o(1)}{\alpha_{\nu,\times}}})$. More precisely suppose the measure ν satisfies the following two properties:

1. *ν attaches no measure to the set (M, ∞), i.e. $\nu(M, \infty) = 0$.*
2. *Given any $\varepsilon > 0$, $\nu(M - \varepsilon, M] > 0$.*

Then given any $\varepsilon > 0$ there exists a node of degree $O(n^{\frac{M-\varepsilon}{\alpha_{\times,\nu}}})$ with high probability as $n \to \infty$ namely there exists a constant $K_\varepsilon^1 > 0$ such that

$$\liminf_{n\to\infty} \mathbb{P}\left(\frac{\max(\deg)}{n^{\frac{(M-\varepsilon)}{\alpha_{\times,\nu}}}} < K_\varepsilon^1\right) \le \varepsilon$$

Further there exists another constant K_ε^2 so that

$$\limsup_{n\to\infty} \mathbb{P}\left(\frac{\max(\deg)}{n^{\frac{M}{\alpha_{\times,\nu}}}} > K_\varepsilon^2\right) \le \varepsilon$$

(c) Consider the height h_n. Then there exists a constant $0 < C_\nu < \infty$ such that

$$\frac{h_n}{\log n} \longrightarrow_P C_\nu$$

As an example of the implications of the above result we work through a particular case of the fitness model and establish mathematical results first conjectured in Bianconi and Barabasi [6] where through heuristics and numerical simulation, they analyze the particular case when the fitness distribution is uniformly distributed on $[0, 1]$. We find the exact limiting degree distribution in this example. We go further and get lower bounds on the size of the maximal degree.

Corollary 32 *(a) Consider the Multiplicative fitness model with $\nu = U[0, 1]$. Using p_k to denote the asymptotic degree distribution we find that*

$$p_k \sim \frac{1}{k^{1+\alpha} \cdot \log k}$$

as $k \to \infty$. Here α is the Malthusian exponent $\alpha_{\times,\nu}$ computed for the uniform distribution and is the unique solution of the equation

$$1 - \frac{1}{\alpha} = e^{-\frac{2}{\alpha}}$$

(b) Given any ε with high probability for large n there exists a node of degree $n^{\frac{1-o(1)}{\alpha_{\times,\nu}}}$. More precisely there exists a constants $K_\varepsilon^1, K_\varepsilon^2 > 0$ such that

$$\limsup_{n\to\infty} \mathbb{P}\Big(\frac{\max(deg)}{n^{\frac{1-\varepsilon}{\alpha_{\times,\nu}}}} < K_\varepsilon^1\Big) \le \varepsilon$$

$$\limsup_{n\to\infty} \mathbb{P}\Big(\frac{\max(deg)}{n^{\frac{1}{\alpha_{\times,\nu}}}} > K_\varepsilon^2\Big) \le \varepsilon$$

5.6 Competition induced preferential attachment

Theorem 33 (Local Properties) *Fix an integer $A \ge 1$. Let $\alpha = 1$. Let $\mu(*)$ be the probability measure on $\mathbb{R}^+$ given by the density:*

$$\mu(dx) = A.\exp(-x)\,[1 - \exp(-x)]^{A-1}$$

Let $\mathcal{T}_n^{\texttt{fert}}$ denote the subtree of $\mathcal{T}_n$ consisting of only the fertile nodes. Then we have convergence in probability fringe sense namely

$$\mathcal{T}_n^{\texttt{fert}} \longrightarrow_{\mathbf{P-fr.}} \mathcal{T}_{\alpha,\mu,\mathcal{P}_A}^{\texttt{sin}}$$

Corollary 34 *Let $N_k(n)$ be the number of* **fertile** *vertices in $\mathcal{T}_n^{\texttt{fert}}$ with out-degree k. Then the limiting empirical proportions are qualitatively similar to the Preferential attachment with a cutoff model namely:*

$$\frac{N_k(n)}{|\mathcal{T}_n^{\texttt{fert}}|} \longrightarrow_P \frac{1}{1+k} \prod_1^{k-1} \left(\frac{i}{i+1}\right) \text{ for } k \le A$$

and

$$\frac{N_k(n)}{|\mathcal{T}_n^{\texttt{fert}}|} \longrightarrow_P \frac{1}{A+1} \left(\frac{A}{A+1}\right)^{k-1-A} \left[\prod_1^{k-1-A} \left(\frac{i}{i+1}\right)\right]$$

as $n \to \infty$

Theorem 35 (Global Properties) *(a) For large n the root has asymptotic size $A \log n$. More precisely:*

$$\frac{deg(\rho)}{\log n} \to_P A$$

as $n \to \infty$.
(b) Given $\varepsilon > 0$ we can choose a constant K_ε such that

$$\limsup_{n\to\infty} P(\frac{\max(deg)}{\log n} > K_\varepsilon) \leq \varepsilon$$

(c) For the height of the tree there exists a constant C_A depending on the parameter A such that

$$\frac{h_n}{\log n} \longrightarrow_P C_A$$

5.7 Preferential attachment Networks

Recall the construction, namely Construction 4 of these networks via the preferential attachment tree model. Recall that each node came with W_v edges, where $W_v \sim_{iid} \nu$ and these edges were attached to the nodes already present in a preferential method. Before we state the main results we need to setup some notation. Let us write the probability mass function of W as

$$P(W = j) = r_j$$

We shall consider three separate cases depending on the tail behavior of the random variable $W \sim \nu$.
(A) The random variables are a constant i.e. there is a constant integer $m \geq 2$ so that $r_m = 1$. Note that the $m = 1$ case is the usual preferential attachment model so we shall not consider it here.
(B) $\mathbb{E}(W) < \infty$ so that there exists $\alpha > 1$ so that

$$x^\alpha \mathbb{P}(W > x) \longrightarrow 0$$

as $x \to \infty$.
(C) Fix $0 < \alpha \leq 1$. Then there exists a constant C such that

$$x^\alpha \mathbb{P}(W > x) \longrightarrow C \tag{3}$$

as $x \to \infty$. Note that the last case belongs to the domain of attraction of the subordinator which has Levy measure:

$$\nu(dx) = \frac{C}{x^{1+\alpha}} \qquad x > 0$$

where C is the same constant as in Equation (3)

Let the associated Levy process be denoted by $(X_t)_{t\geq 0}$ with $X_0 = 0$. Let U be uniformly distributed in $[0,1]$ independent of $(X_t)_{t>0}$. Note that since the process X is increasing, $0 \leq \frac{X_U}{X_1} \leq 1$.

Let Z_α denote the positive random variable:

$$Z_\alpha = -\frac{1}{2}\left(\frac{1}{\alpha}\log(U) + \log(\frac{X_U}{X_1})\right) \tag{4}$$

Theorem 36 *Consider models (A) and (B). Let V_n be as before a randomly picked node from the network $\mathcal{N}_n$. Then*

$$\deg(V_n) \longrightarrow_L \sum_1^W N_i(T_2) \tag{5}$$

where $(N_i(.))$ are standard Yule processes and $T_2 \sim Exp(2)$ independent of N_i and $W \sim \nu$ independent of N_i and T_2. In particular

$$\mathbb{P}(\deg(V_n) = k) \longrightarrow \frac{2}{k(k+1)(k+2)} \sum_{j=1}^{k} r_j \cdot j(j+1) \tag{6}$$

as $n \to \infty$.

Write p_k for the asymptotic degree proportions. Then note that Theorem 36 implies the following results:

Corollary 37 *(A) In case of model A the asymptotic network degree distribution has a power law exponent of 3 i.e. $p_k \sim k^{-3}$ as $k \to \infty$*
(B) If $E(W^2) < \infty$ then $p_k \sim k^{-3}$
(C) If $r_j \sim j^{\alpha}$ for $2 \le \alpha < 3$ then $p_k \sim k^{\alpha}$

Thus the **smaller tail exponent** (when comparing usual linear preferential attachment and the distribution of W) always wins.

There is an even stronger result that can be proved namely:

Theorem 38 *Suppose the random variables W_i satisfy the property:*

$$\frac{\omega(n)}{n^2} \sum_1^n W_i^2 \longrightarrow_P 0$$

as $n \to \infty$ for some function $\omega(n) \uparrow \infty$ arbitrarily slowly.

Let $N_k(n)$ be the number of nodes with degree k. Then

$$\frac{N_k(n)}{n} \longrightarrow_P p_k$$

where p_k are the limiting degree distribution constants given by Theorem 36.

As a result we have the following Corollary which was also proved in [12] through completely different methods (they also get finite n error bounds, an aspect which we have not explored):

Corollary 39 *Suppose W has a $1 + \varepsilon$ moment for some $\varepsilon > 0$. Then*

$$\frac{N_k(n)}{n} \longrightarrow_P p_k$$

as $n \to \infty$.

See Section 6 for some details as to how to prove Theorem 38. The main point is that Theorem 36 with minor work implies Theorem 38. Thus the main work is in proving Theorem 36.

Theorem 40 *Consider Model (D) where we have $\mathbb{P}(W > x) \sim x^{-\alpha}$ for some $0 < \alpha \leq 1$. In this case we have*

$$\deg(V_n) \longrightarrow_L \sum_{1}^{W} N_i(Z_\alpha) \tag{7}$$

where as before $(N_i(.))$ are standard Yule processes and Z_α is as defined in Equation (4) independent of N_i and $W \sim \nu$ independent of N_i and Z_α.

Thus what this implies is that the expected number of nodes with degree k converges namely:

$$\mathbb{E}\left(\frac{N_k(n)}{n}\right) \longrightarrow \mathbb{P}(U = k)$$

where $U = \sum_1^W N_i(Z_\alpha)$.

Remark: Note that in this case we have not been able to prove that the "higher exponent always wins. The essential problem is that the concentration inequalities we use break down at this point and are not strong enough to show the concentration of the proportions of vertices with degree k about their mean value.

Global analysis: Now we come to a Theorem regarding the maximal degree distribution which probably requires more work using combinatorial techniques but which follows very easily from our techniques. Instead of dealing with completely general ν we shall look at special cases, depending on the extremal behavior of the random variables W_i. More precisely we shall look at 3 different cases depending on the tail behavior

(A) There exists a constant $m \geq 1$ so that $\mathbb{P}(W = m) = 1$ or has finite support namely there exists m such that $\mathbb{P}(W \leq m) = 1$.
(B) $W = \lfloor X \rfloor + 1$ where $X \sim Exp(1)$ and $\lfloor x \rfloor$ denotes the integer part of x.
(C) There exists an $\alpha > 2$ so that $\mathbb{P}(X > x) \sim \frac{1}{x^\alpha}$.
(D) There exists $\alpha < 1$ so that $\mathbb{P}(W > x) \sim \frac{1}{x^\alpha}$ where $\alpha < 1$

Case (A) is very simple and essentially follows from the already known asymptotics for the Preferential attachment model. The other models are slightly more tricky to analyze due to the unbounded support and the fact that as n becomes large, there are nodes which arrive with a large number of out edges, of order $O(\log n)$ for model (B) and of order $n^{\frac{1}{\alpha}}$ for model (C).

Theorem 41 *Consider the degree distribution of either (A) , (B) and (C). Then given any $\varepsilon > 0$ there exists constants $0 < K_\varepsilon^1 < K_\varepsilon^2 < \infty$ so that*

$$\limsup_{n\to\infty} \mathbb{P}\left(\frac{\max(\deg(\mathcal{N}_n))}{\sqrt{n}} \notin [K_\varepsilon^1, K_\varepsilon^2]\right) \leq \varepsilon$$

In the case where $\alpha < 1$ we get an easy upper bound. For this case write $p = \frac{1}{\alpha}$ and note that $p > 1$.

Theorem 42 *Given any $\varepsilon > 0$ there exists a constant $K_\varepsilon > 0$ such that*

$$\mathbb{P}\left(\max(\deg(\rho) > K_\varepsilon n^{\frac{p}{2}}\right) \geq 1 - \frac{\varepsilon}{2}$$

where ρ is the initial starting node.

Remark:

(1) The Case where $1 < \alpha < 2$ seems to be hard to analyze. See the Proof of Lemma 102 as to exactly where we need the condition $\alpha > 2$.

(2) Note that in the case $\alpha < 1$ extremal theory for the maximum of independent identically distributed random variables implies that the maximum out-degree $\max_{1 \leq i \leq n} W_i = O_P(n^p)$ where $p = \frac{1}{\alpha}$ whereas the above result implies that the degree of the root is $n^{\frac{p}{2}}$. Thus in this setup the size of the maximal degree and the degree of the root differ by several orders of magnitude, something we do not see in the other models.

6 Proofs

Due to sheer number of different models and results we hope to present and constraints on the physical size of the manuscript the organization of this section has proved to be a little tricky. We shall prove all the main results and whenever we feel that a particular result for some model can be proved via identical methods in some previously analyzed model we shall take an opportunity to skip the proof as it adds nothing new to the discussion. We shall also try to collect the common ingredients required in the proofs in the various models and then delve into each model separately.

In this section we shall analyze all models other than the random fitness models which require multitype branching process theory. In this Section whenever we say " all models in Section 2.1 " we mean all models excluding the above mentioned random fitness model.

All the proofs are based on the following obvious fact:

Proposition 43 *Consider any of the models in Section 2.1 which are given by attractiveness defined by $f(v,n) = f(Deg(v,n))$. Let $\mathcal{G}_n^f$ denote the random tree constructed on n vertices. Consider the associated Continuous time branching process as defined in Definition (6). For this branching process let $(N^f(t), \mathcal{F}^f(t))$ denote the size of the population and the entire population tree respectively at time t. Define the stopping time*

$$T_n = \inf\{t : N^f(t) = n\}$$

Then

$$\mathcal{F}^f(T_n) =_d \mathcal{G}_n^f$$

Proof: Obvious

■

6.1 Malthusian rate of growth

All the models in Section 2.1 are **regular** in the following sense: For the associated point process $\mathcal{P}$ of a preferential attachment model as in Definition 1, let the mean intensity measure $\mu(t)$ be defined as $\mu(t) = E(\mathcal{P}(t))$. Then for each of the above models there exists a constant $\theta > 0$ (depending on the distribution of the point process $\mathcal{P}$) such that

$$\theta \int_0^\infty e^{-\theta t}\mu(t)dt = 1 \tag{8}$$

This constant θ is called the Malthusian rate of growth for the particular branching process model and plays a crucial role in what is to follow. It can be explicitly computed in some particular cases. We start with a trivial Lemma exhibiting how it can be computed in the most general setup. See [31] where similiar such computations were carried out for the Linear preferential attachment model.

Lemma 44 *Given a point process $\mathcal{P}$ defined by the rate function f, the Malthusian rate of growth parameter θ is the unique positive solution of the equation*

$$\sum_0^\infty \prod_0^i \left(\frac{f(i)}{\theta + f(i)}\right) = 1$$

Proof: Uniqueness follows since the integral

$$\theta \int_0^\infty e^{-\theta t}\mu(t)dt = E(\mathcal{P}(X_\theta)) \tag{9}$$

where X_θ is an exponential random variable with rate θ independent of the point process $\mathcal{P}$. Thus the above function of θ is monotonically decreasing. Note that point process $\mathcal{P}$ can be constructed as follows: Let $(Y_i)_{i\geq 0}$ be independent exponential random variables with rate $f(i)$. Then the point process $\mathcal{P}$ on $\mathbb{R}^+$ is the collection of points S_i where

$$S_i = \sum_0^i Y_i$$

Now note that

$$\begin{aligned} E(\mathcal{P}(X_\theta)) &= E(\sum_{i=0}^\infty 1\{X_\theta > S_i\}) \\ &= \sum_{i=0}^\infty P(X_\theta > S_i) \\ &= \sum_{i=0}^\infty E(e^{-\theta S_i}) \end{aligned}$$

since for a rate θ exponential random variable $P(X_\theta > x) = e^{-\theta x}$.

Now use the fact that for an exponential rate $f(i)$ random variable

$$E(e^{-\theta Y_i}) = \frac{f(i)}{\theta + f(i)}$$

■

Although the above gives us a formula for computing θ in most cases we need to use special properties of the point process involved to get explicit values of θ. Below we compute explicitly the Malthusian rate of growth for some of the models in Section 2.1.

Proposition 45 *The Malthusian rate of growth is given by the following explicit values*
(a) Preferential attachment: $\theta = 2$.
(b) Linear Preferential attachment with parameter a*:* $\theta = \frac{a+2+\sqrt{a^2+4}}{2}$.
(c) Competition Induced Preferential attachment with parameter A*:* $\theta = 1$.

Proof:

Note that we have to solve the Equation (8). In the particular cases analyzed in this Lemma, the special properties of the Yule process help us to solve this particular equation.

For the preferential attachment model of part(a), the point process $\mathcal{P}$ is the Yule process. The following is known for the Yule process :

Lemma 46 *Consider the standard Yule process* $\mathcal{P}(t)$*. Then*
(a) The distribution at finite time is given by

$$\mathbb{P}(\mathcal{P}(t) = k) = e^{-t}(1 - e^{-t})^k \qquad k = 0, 1, 2, \dots$$

so that intensity measure is equal to

$$\mu(t) = e^t - 1$$

(b) $\frac{N(t)}{e^t}$ *is a martingale and*

$$\frac{N(t)}{e^t} \longrightarrow_{a.s.} W$$

where W *is a rate* 1 *exponential random variable.*

Substituting in Equation (8) gives part (a).

For the linear preferential attachment with parameter a the associated point process offspring distribution can be decomposed as $\mathcal{P}(t) = N(t) + Y(t)$ where $N(.)$ is a standard Yule process and $Y(t)$ is a Poisson process with rate a. Plugging this into Equation (8) implies we have to solve the equation

$$\frac{1}{\theta - 1} + \frac{a}{\theta} = 1$$

Solving the corresponding quadratic equation for the positive root gives us part (b) of the Lemma.

Finally in the case of the Competition Induced preferential attachment model, looking at the subtree of only fertile nodes we have a branching process driven by the point process of Definition 3. In this setup, solving the Equation (8) boils down to solving the equation

$$\frac{A+\theta}{\theta} \prod_{1}^{A} \left(\frac{i}{i+\theta} \right) = 1$$

The unique positive solution of this equation is $\theta = 1$.

■

In the other cases although we don't have explicit formulae for the Malthusian rate of growth, we prove the properties of this constant and how it depends on the model parameters namely:

Proof of Prop 20 and Prop 23 : Obvious from the probabilistic interpretation of the Malthusian rate of growth given by Equation (9). Alternatively this is obvious from Lemma 44

In the above setup, given the Malthusian rate of growth constant θ define the following measure:

Definition 47 *The* sin-*measure* ν *associated with a continuous time branching process is the measure on* $\mathbb{R}^+$ *with density*

$$\nu(dt) = e^{-\theta t}\mu(dt)$$

where μ *is the mean intensity measure of the offspring distribution* $\mathcal{P}$ *and* θ *is the Malthusian rate of growth parameter.*

6.2 Local computations

Now we state the essential results we use to prove our basic results. In their full generality they were proved by Jagers and Nerman ; see [19] and [25]. A very readable account is given in Aldous [2]. Before coming to the actual statement we need to make a very important technical assumption. In spirit, it implies that the Malthusian growth parameter gives the **true exponential rate of growth** of the Branching process.

Assumption 48 *For the Branching process driven by the point process* $\mathcal{P}$ *we assume that*
(a) There exists a Malthusian rate of growth parameter θ *i.e Equation (8) has an explicit positive solution.*
(b) The $x\log^+ x$ *condition is satisfied, that is, defining the random variable* $\hat{\mathcal{P}}(\theta) := \int_0^\infty e^{-\theta t}\mathcal{P}(dt)$ *where* θ *is the unique positive solution of Equation (8) and* $\log^+(x) = \{\max \log(x), 0\}$ *we assume:*

$$\mathbb{E}\left(\hat{\mathcal{P}}(\theta)\cdot\log^+(\hat{\mathcal{P}}(\theta))\right) < \infty \tag{10}$$

Theorem 49 ([19],[25]) *Consider a regular continuous time branching process driven by a point process* $\mathcal{P}$ *and Malthusian growth parameter* θ *that satisfies Assumption 48 Let* ν *be the associated* sin-*measure* ν *given by Definition 47. Let* $\mathcal{F}(t)$ *be the random tree at time* t*. Then*

$$\mathcal{F}(t) \longrightarrow_{\mathbf{P-fr.}} \mathcal{T}^{\mathtt{sin}}_{\theta,\nu,\mathcal{P}}$$

i.e the convergence takes place in the probability fringe sense. The same fact holds if we look at $\mathcal{F}^f(T_n)$ *where* T_n *are a sequence of increasing almost surely finite stopping times* T_n*, with* $T_n \uparrow \infty$ *a.s..*

An obvious corollary of this Theorem is the convergence of the degree distribution. Note that we do not need to use any concentration inequalities to prove this convergence.

6.2.1 Convergence of the empirical degree distribution:

Corollary 50 *As before fix any sequence of finite stopping times $T_n \uparrow \infty$ almost surely. Let $N_{\geq k}(n)$ be the number of nodes with out-degree greater than or equal to k in $\mathcal{F}(T_n)$. Then:*

$$\lim_{n\to\infty} \frac{N_{\geq k}}{n} \longrightarrow_P \prod_{0}^{k-1} \left(\frac{f(i)}{\theta + f(i)} \right) \qquad \textit{for } k = 1, 2, 3, \ldots$$

Proof: Theorem 49 gives us convergence in probability fringe sense and describes the limiting extended fringe distribution. By the definition of convergence in probability fringe sense (Definition 9) we have

$$\lim_{n\to\infty} \frac{N_{\geq k(n)}}{n} \longrightarrow_P \mathbb{P}(\mathcal{P}(X_\theta) \geq k)$$

where X_θ is an exponential rate θ random variable independent of the point process $\mathcal{P}$. Now note that

$$\begin{aligned} \mathbb{P}(\mathcal{P}(X_\theta) \geq k) &= \mathbb{P}(X_\theta > S_k) \\ &= \mathbb{E}(e^{-\theta S_k}) \end{aligned}$$

where S_j is the j^{th} point of the Point process $\mathcal{P}$. Now note that by the rates description of the point process $\mathcal{P}$ we have

$$S_j =_d \sum_{i=0}^{j-1} \frac{1}{f(i)} Y_i$$

where Y_i are distributed as iid rate one exponential random variables. Combining with the above equation gives us our result. ■

Corollary (50) proves all our degree asymptotic assertions in Section 5 using the appropriate attractiveness function f corresponding to the particular model of interest.

However before we fully use the above results we first need to verify that in our setup Assumption 48 is satisfied. We seperate the analysis into two cases, one where the offspring point process grows at an essentially Linear rate and another where it grows at an exponential rate.

Validity of the $x \log x$ assumption for our models:

Proposition 51 *Consider the Preferential attachment with a cutoff and the Competition induced preferential attachment models defined defined in Section 2.1. Then for both models, using $\mathcal{P}$ to denote the associated point process of the offspring distribution, we have:*

$$\mathbb{E}\left([\hat{\mathcal{P}}(\theta)]^2 \right) < \infty$$

where θ is the associated Malthusian rate of growth of parameter for the model.

Note that this is stronger than the condition required by Equation (10).

Proof:

Let θ be the Malthusian rate of growth in either case. Note that in both models, the associated point process is stochastically dominated by a Poisson process with rate A i.e.

$$\mathcal{P}_A(.) \leq_{st} Poi_A(.)$$

Thus it is enough to prove

$$\mathbb{E}([\widehat{P}oi(\theta_A)]^2) = \mathbb{E}(\left[\int_0^\infty e^{-\theta_A x} Poi_A(dx)\right]^2) < \infty$$

Now note that the by Fubini's theorem, for any counting measure $\xi(.)$ on $\mathbb{R}^+$, with $\xi(0) = 0$ we have

$$\int_0^\infty e^{-\theta x}\xi(dx) = \theta \int_0^\infty e^{-\theta x}\xi(0,x]dx = \mathbb{E}(\xi(0,T_\theta]|\xi)$$

where T_θ is an exponential random variable with rate θ independent of ξ.

Thus

$$\mathbb{E}\left([\int_0^\infty e^{-\theta x}\xi(dx)]^2 \middle|\, \xi \right) = \mathbb{E}\left(\xi(0,T_\theta^1]\cdot\xi(0,T_\theta^2]\middle|\,\xi\right)$$

where T_θ^i are iid rate θ exponential random variables independent of ξ.

Now using the lack of memory property of the Exponential random variable and the homogeneity of a rate A Poisson process we have:

$$\begin{aligned}
\mathbb{E}(Poi_A(0,T_\theta^1]\cdot Poi_A(0,T_\theta^2]) &= 2\cdot\mathbb{E}\left(Poi_A(0,T_\theta^1]\cdot Poi_A(0,T_\theta^2])1_{\{T_\theta^1<T_\theta^2\}}\right) \\
&= 2\mathbb{E}\left(Poi_A(0,T_\theta^1]\cdot\left(Poi_A(0,T_\theta^1] + Poi_A(T_\theta^1, T_\theta^1+T_\theta^2)\right)\right) \\
&= 2\cdot\mathbb{E}\left(Poi_A^2(T_\theta)\right) + 2\cdot[\mathbb{E}(Poi_A(T_\theta))]^2 \\
&= 2\left(\frac{2A^2}{\theta^2} + \frac{A}{\theta} + \frac{A^2}{\theta^2}\right) < \infty
\end{aligned}$$

where in the second line we have used the lack of memory property of the Exponential distribution.

■

Now we analyze the remaining three models. We first restate the crucial identity derived in the previous proof namely

Lemma 52 *For any point process ξ on $\mathbb{R}^+$, with $\xi(0) = 0$ we have*

$$\hat{\xi}(\theta) = \mathbb{E}(\xi(0,T_\theta]|\xi) \tag{11}$$

where T_θ is an exponential rate θ random variable independent of ξ.

In each of the models below we shall let ξ be the associated point process.

Proposition 53 *(a) Consider the Linear Preferential attachment model where the Malthusian growth parameter is $\theta = 2$. Then for any $1 < \beta < 2$ we have*

$$\mathbb{E}(\left[\hat{\xi}(\theta)\right]^{\beta}) < \infty$$

(b) Consider the Linear preferential attachment model with $a > 0$. Let γ_a be the Malthusian growth parameter. Then for any $1 < \beta \leq 2$ we have:

$$\mathbb{E}(\left[\hat{\xi}(\gamma_a)\right]^{\beta}) < \infty$$

(c) Consider the sub-linear preferential attachment model with parameter $\alpha < 1$. Let θ_α be the Malthusian growth parameter. Then for any $\beta > 1$ we have

$$\mathbb{E}(\left[\hat{\xi}(\theta_\alpha)\right]^{\beta}) < \infty$$

Proof:

Note that by Jensen's inequality applied to the function $f(x) = x^\beta$ for $\beta > 1$ coupled with Lemma 11 we have

$$\mathbb{E}([\hat{\xi}(\theta)]^{\beta}) \leq \mathbb{E}([\xi(0, T_\theta]]^{\beta}) \tag{12}$$

Thus we shall analyze the distribution of the random variable $\xi(0, T_\theta]$ occurring in the right side of Equation (12).

Thus in each case it is enough to prove that

$$\mathbb{E}([\xi(0, T_\theta]]^{\beta}) < \infty \tag{13}$$

Now note that Corollary 50 gives us the exact distribution of this random variable. More precisely the corollary implies that for any model, with θ denoting the Malthusian rate of growth and $f(i)$ denoting the attractiveness function f, we have:

$$\mathbb{P}(\xi(0, T_\theta] = k) = \frac{\theta}{\theta + f(k)} \prod_{0}^{k-1} \left(\frac{f(i)}{\theta + f(i)} \right) \tag{14}$$

For the usual preferential attachment model, using Equation (14)

$$P(\xi(0, T_\theta] = k) \sim \frac{1}{k^3}$$

This implies Equation (13) for the preferential attachment model.

For the linear preferential attachment model, identical computations imply that

$$P(\xi(0, T_{\theta_a}] = k) \sim \frac{1}{k^{\gamma_a + 1}}$$

Since for $a > 0$ we have $\gamma_a > 2$ we get part (b).

Finally to prove (c), note that Equation (14) implies that

$$P(\xi(0, T_{\theta_\alpha}] = k) \sim \frac{1}{k^\alpha} \exp(-\frac{\theta_\alpha}{1-\alpha} k^{1-\alpha})$$

This implies (c).

■

6.2.2 Computation of the sin-measures

Here we quickly show how to compute the `sin`- measures ν in whichever models have a nice analytic form for this measure.

Proposition 54 *(a) The sin measure for the usual preferential attachment model is the rate one exponential distribution, namely $\nu(dx) = e^{-x}dx$ on $\mathbb{R}^+$.*
(b) The `sin`*-measure for the Linear preferential attachment model is given by a mixture of a rate $(\gamma - 1)$ Exponential random variable and a rate γ Exponential random variable. Here $\gamma = \gamma(a)$ is the Malthusian rate of growth for this model.*
(c) For the Preferential attachment with cutoff at A, the `sin`*-measure is given by*

$$\nu(dx) = e^{-\theta_A x} \left[\sum_0^A (j+1) e^{-x} (1-e^{-x})^j + (A+1)(1-e^{-x})^A \right] dx$$

where θ_A is the Malthusian rate of growth parameter for the associate branching process.
(d) For the Competition induced preferential attachment model the `sin`*-measure is given by*

$$\nu(dx) = A e^{-x} [1 - e^{-x}]^{A-1} dx$$

Proof: (a) For the preferential attachment model note that $\theta = 2$. Since the offspring distribution is just the Yule process the mean intensity measure is given by $\mu(dt) = e^t dt$. Now use Definition 47 of the `sin`-measure to compute ν.
(b) Note that in this case the offspring distribution

$$\mathcal{P}(t) = N(t) + Poi_a(t)$$

where $N(.)$ is a rate one Yule process and Poi_a is a rate a Poisson process independent of $N(.)$. Thus the mean intensity measure satisfies the formulae:

$$\mu(dt) = (e^t + a) \cdot dt$$

Also note that the Malthusian rate of growth parameter satisfies the equation:

$$\frac{a}{\gamma} + \frac{1}{\gamma - 1} = 1$$

By the definition of the sin measure, we can write the sin measure in this case as:

$$\begin{aligned} \nu(dt) &= e^{-\gamma t} \cdot (e^t + a) dt \\ &= \frac{1}{\gamma - 1} \cdot \left[(\gamma - 1) e^{-(\gamma-1)t} \right] dt + \frac{a}{\gamma} \cdot \left[\gamma e^{-\gamma t} \right] dt \end{aligned}$$

This gives the result for the linear preferential attachment model.

(c) For the preferential attachment model with a cutoff note that the offspring distribution $\mathcal{P}(.)$ has the same rates as that of a standard Yule process till the number of offspring reaches A, and then remains constant $A+1$. A simple coupling argument with a Yule process then gives us

$$\mathbb{P}(\mathcal{P}(0,x] = k) = \mathbb{P}(N(0,x] = k) = e^{-x}(1-e^{-x})^k \qquad \forall x,\ \forall k < A \tag{15}$$

where $N(.)$ is a Yule process Thus

$$\mathbb{P}(\mathcal{P}(0,x] \geq A) = (1-e^{-x})^A \tag{16}$$

Note that the mean intensity measure satisfies the relation

$$\mu(dx) = \mathbb{P}(\mathcal{P}[x,x+dx) = 1) = \sum_0^\infty \mathbb{P}([x,x+dx) = 1|\mathcal{P}(0,x] = j) \cdot \mathbb{P}(\mathcal{P}(0,x] = j) \tag{17}$$

The rates of the Preferential attachment with a cutoff model are:

$$\begin{aligned}
\mathbb{P}([x,x+dx) = 1|\mathcal{P}(0,x] = j) &= (j+1)dx + o(dx) \qquad \text{for } j < A \\
&= A + 1dx + o(dx) \qquad \text{for } j \geq A
\end{aligned}$$

Combine the above with Equations (17), (15) and (16) to get the result.

(d) Recall the construction of the offspring distribution of the fertile nodes from Section 2.1 namely Definition 3. Once the first fertile node is produced, the process produces at a constant rate A. Equation (17) implies

$$\mu(dx) = A \cdot \mathbb{P}(S_1 \leq x)dx$$

Note that S_1 has the same distribution as the time of the A^{th} birth in a Yule process. Thus $\mathbb{P}(S_1 \leq x) = P(N(x) \geq A) = (1-e^x)^A$. This proves the result after noting that the Malthusian rate of growth in this case is $\theta_A = 1$ and using the definition of the `sin`-measure.

■

6.2.3 Proofs for the BA-Preferential Attachment model

In this section we provide the rest of the proofs for the special case of the usual preferential attachment model where we had found some extra results to illustrate the method of computations and had computed the limiting distribution of some more local functionals including the joint distribution of a node and it's mother, etc.

Theorem 11 follows from Theorem 49 and the explicit computation of the sin measure carried out in Proposition 54.

Proof of Corollary 12 :

Part (a): By Theorem 11 we have that the number of nodes with load k, namely $L_k(n)$ satisfies

$$\frac{L_k(n)}{n} \longrightarrow_P \mathbb{P}(|\mathcal{F}_0| = k)$$

while the number of nodes with degree k, namely $M_k(n)$ satisfies

$$\frac{N_k(n)}{n} \longrightarrow_P \mathbb{P}(\deg(\mathcal{F}_{0,\rho}) = k)$$

Where $\mathcal{F}_0$ is the fringe subtree (see Definition 7) in the construction of the limiting `sin-` tree corresponding to this model and $\mathcal{F}_{0,\rho}$ is used to denote the root of this tree. By previous computations, the fringe tree in this setup coressponds to:

$$\mathcal{F}_0 = \mathcal{F}(T_2)$$

where $\mathcal{F}(\cdot)$ is continuous time branching process with offspring distributions distributed as standard Yule processes and T_2 is an exponential random variable with rate 2 independent of $\mathcal{F}(\cdot)$. Thus we first need to find the distribution of $Z = |\mathcal{F}(T_2)|$. To do so, consider the following construction:

Construction 55 *Consider the following continuous time Markov process :*
(a) Start with one node at time $t = 0$.
(b) Each node lives for an exponential (rate 1) amount of time and then gives rise to 3 identical offspring and dies.
(c) Each of the offspring behave in the same manner namely each of them have their own independent exponential lifetimes before splitting into 3 new nodes and dying.

Let $N(t)$ denote the number of nodes *alive* at time t. Then it is easy to verify the following relations

$$|\mathcal{F}(t)| =_d \frac{N(t) + 1}{2}$$

Let $F(s, t) = \mathbb{E}(s^{N(t)})$. Then $N(t)$ is also a special time of Markov branching process (belonging to the class Bellman-Harris branching processes)Then the dynamics of the process allows us to simply compute (via solving a differential equation in t for a fixed s) F(s,t) as

$$F(s, t) = \frac{se^{-t}}{\sqrt{1 - (1 - e^{-2t})}}$$

See [4] Chapter III, Section 5 for a complete proof of this result. Simple algebraic manipulations using this now give us the generating function of Z.

Part (b): Differentiating this at 1 gives us that $\mathbb{E}(Z) = \infty$. Note that this also follows from [2], where Aldous proves that the expected size of a any limiting fringe tree is equal to infinity.

Part(c): This follows from the limiting degree computations carried out in Corollary 50 and are not repeated here.

■

Proof of Corollary 13: By Theorem 11 and the definition of convergence of fringe in probability sense and our knowledge about the limiting `sin`-tree, we know that writing:

$$X_{k,l}(n) = \sum_{v \in \mathcal{T}_n} 1\{\mathtt{out-deg}(v) = k, \mathtt{out-deg}(\mathtt{mother})(v) = l\}$$

we have

$$\frac{X_{k,l}}{n} \longrightarrow_P \mathbb{P}(Z_1 = k, Z_2 = l)$$

as $n \to \infty$, where (Z_1, Z_2) can be reconstructed as follows:

Let $X_1 \sim Exp(2)$ and $X_2 \sim Exp(1)$. Then $Z_1 \sim (\mathcal{P}_1(X_1) - 1)$ and $Z_2 \sim \mathcal{P}_2^{X_2}(X_1 + X_2)$ where $\mathcal{P}_i$ are independent with $\mathcal{P}_1(.)$ being a usual Yule process, and $\mathcal{P}_2^{X_2}(t)$ being a Yule process conditioned to have a point at the point X_2.

Now note that by the Markov property

$$\mathcal{P}^{X_2}(X_1 + X_2) =_d \sum_1^T \mathcal{P}_i(X_1)$$

where $T =_d \mathcal{P}^{X_2}(X_2)$ i.e the size of a Yule process at time X_2 conditioned to have a point at X_2.

To find explicit distributions of the quantities occurring in the previous paragraph, we collect the following Lemma for future reference. The proof follows from routine calculations using the exact distributional properties of the Yule process and are omitted.

Lemma 56 *Fix a constant $t > 0$. Let $\mathcal{P}^t(.)$ be a Yule process conditioned to have a birth at time t. Then*

(a) The distribution of $\mathcal{P}^t(t)$ is given by the probability mass function:

$$\mathbb{P}(\mathcal{P}^t(t) = k) = (k-1)e^{-2t}(1-e^{-t})^{k-2} \qquad \text{for } k = 2, 3, \ldots,$$

(b) For any $s > t$, the distribution of $\mathcal{P}^t(s)$ is given by the distributional identity

$$\mathcal{P}^t(s) =_d \sum_1^{\mathcal{P}^t(t)} Y_i(s-t)$$

where $Y_i(.)$ are independent Yule processes independent of $\mathcal{P}^t(t)$.

Thus integrating over the density of the conditioned point X_2, we find that the p.m.f of $\mathcal{P}^{X_1}(X_1)$ is given by:

$$\begin{aligned} P(\mathcal{P}^{X_1}(X_1) = k) &= \int_0^\infty P(N^t(t) = k)e^{-t}dt \\ &= \int_0^\infty (k-1)\left[e^{-2t}(1-e^{-t})^{k-2}\right]e^{-t}dt \\ &= \frac{2}{k(k+1)} \end{aligned}$$

Combining all these properties easily gives us Corollary 13.

■

6.3 Global computations

6.3.1 Time to reach size n

It is well known that the branching processes we consider grow at a truly exponential rate with the rate being the Malthusian rate of growth parameter. More precisely we state the following theorem from [19]:

Theorem 57 *Consider a branching process driven by a point process $\mathcal{P}$ satisfying Assumption 48. Then there exists a random variable $W > 0$ such that:*

$$\frac{N(t)}{e^{\theta t}} \longrightarrow_{a.s.,L^1} W$$

as $t \to \infty$, where θ is the Malthusian growth parameter associated with the process.

In particular this implies the following

Proposition 58 *Assume we are in the setup of Theorem 57. Consider the stopping times T_n defined as the first time the population grows to size n. Then*

$$\frac{n}{e^{\theta \cdot T_n}} \longrightarrow_{a.s.} W$$

In particular this implies the following relations:

$$\frac{n^{1/\theta}}{e^{T_n}} \longrightarrow_{a.s.} W^{1/\theta}, \qquad \frac{1}{\theta}\log n - T_n \longrightarrow_{a.s.} \frac{1}{\theta}\log W$$

and

$$\frac{T_n}{\frac{1}{\theta}\log n} \longrightarrow_{a.s.} 1 \tag{18}$$

Note that the we have crucially used the fact that $W > 0$. Thus Assumption 48 is absolutely essentially for us.

Recursive Distributional Equation for the Limiting random variable: For any given model (with N(t) denoting the size of the population at time t) let $(N_i(.))_{(1\leq i\leq\infty)}$ be *i.i.d.* copies of $N(.)$ and let $\xi_1, \xi_2, \ldots$ denote the times of birth of the offspring of the initial ancestor. Then note that

$$N(t) = 1 + \sum_{\xi_i \leq t} N_i(t - \xi_i)$$

Thus

$$\frac{N(t)}{e^{\theta t}} = e^{-\theta t} + \sum_{\xi_i \leq t} e^{-\theta \xi_i} \frac{N_i(t - \xi_i)}{e^{\theta(t-\xi_i)}}$$

Letting $t \to \infty$ we have

$$W =_d \sum_{1}^{\infty} e^{-\theta \xi_i} W_i \tag{19}$$

where W_i are i.i.d copies of W

Distribution of the limiting random variable: The explicit analytic form of the distribution of the limiting random variable W occurring in Theorem 57 is known in only a few cases. We collect what seems to be known, relevant to our context.

Lemma 59 *(a) Let $N(t)$ be the Yule process (where the offspring distribution of each node is a rate one Poisson process). Then*

$$\frac{N(t)}{e^t} \longrightarrow_{a.s.,L^2} W$$

where W has density e^{-x}, i.e is an Exponential rate 1 random variable.
(b) Let $N(t)$ be the branching process associated with the BA preferential attachment model (where the offspring distribution of each node is a standard Yule process). Then

$$\frac{N(t)}{e^{2t}} \longrightarrow_{a.s.L^2} W$$

where $2 \cdot W$ has a Gamma distribution namely $2 \cdot W$ has the density $x.e^{-x}$ on $\mathbb{R}^+$

Proof: (a) follows from Lemma 46.

(b) follows from Construction 55 and the explicit form of the probability generating function $F(s,t)$, by replacing s by $e^{-\lambda}$ and letting $t \to \infty$ for a fixed λ, to give us asymptotics for the Laplace transform of the limiting random variable .

We finally conclude this brief tutorial on continuous time branching processes with the following bound on the first moment. Although much more is known including bounds on the higher moments, see [18] for more details, the following shall suffice for our purposes:

Theorem 60 *Consider a continuous time branching process satisfying Assumptions 48. Let $N(t)$ be the size of the population at time t and let γ be the associated Malthusian rate of growth parameter. Then there exists a constant C depending on the distribution of the point process $\mathcal{P}$ of the offspring distribution, such that for any fixed time t,*

$$\mathbb{E}\left(N(t)\right) \leq Ce^{\gamma \cdot t}$$

6.3.2 Asymptotics for the root degree

Write $\mathcal{P}_0$ as the point process representing the times of birth of the offspring of the root. Then note that

$$deg(\rho) =_d \mathcal{P}_0(T_n) \tag{20}$$

where $deg(\rho)$ denotes the degree of the root ρ in $\mathcal{T}_n$. This is the fundamental equation which coupled with Proposition 58 allows us to derive all our asymptotics for the root degree.

Computations for different models:

(a) Preferential attachment: The offspring distribution $\mathcal{P}_0$ here is given by the Yule process. By Lemma 59(a) we have

$$\frac{deg(\rho)}{e^{T_n}} \longrightarrow_{a.s} W$$

where W is an exponential rate one random variable. This implies by Proposition 58 using the fact that in this case the Malthusian rate of growth is 2 that

$$\frac{deg(\rho)}{\sqrt{n}} \longrightarrow_{a.s.} \frac{W}{\sqrt{\tilde{W}}}$$

where $\tilde{W}$ is the random variable in Lemma 59(b). Equation (19) gives us the relation between W and $\tilde{W}$. This completes the proof for the BA preferential attachment model.

(b) Linear Preferential attachment: A similar argument works in this case as well, with the caveat that we are unable to find the explicit form of the limiting random variable.

(c) Sub-linear Preferential attachment: First note that in this case, the offspring distribution $\mathcal{P}(t)$ can be constructed as follows: Let Y_i be *iid* exponential rate 1 random variables. Define the partial sums:

$$S_n = \sum_{1}^{n} \frac{Y_i}{i^\alpha}$$

where $\alpha < 1$. Then

$$\mathcal{P}_0(t) = \sup\{n : S_n \leq t\}$$

We first start with a preliminary Lemma analyzing the growth rate of the above offspring distribution for large t.

Lemma 61 *(a) consider the partial sum sequence S_n as above. Then*

$$\frac{S_n}{n^{1-\alpha}} \longrightarrow_P 1$$

as$n \to \infty$ (b) The offspring distribution $\mathcal{P}$ satisfies:

$$\frac{\mathcal{P}(t)}{t^{\frac{1}{1-\alpha}}} \longrightarrow_P 1$$

Proof: Note that $E(S_n) = \sum_1^n \frac{1}{i^\alpha} \sim n^{1-\alpha}$. Also $Var(S_n) = \sum_1^n \frac{1}{i^{2\alpha}}$ so that $Var(S_n)/E(S_n) \to 0$ as $n \to \infty$. Now use Chebyshev's inequality.

To prove (b) we note that from (a) we have

$$\frac{(S_{\mathcal{P}(t)})^{\frac{1}{1-\alpha}}}{\mathcal{P}(t)} \longrightarrow_P 1$$

Also note that

$$\frac{(S_{\mathcal{P}(t)})^{\frac{1}{1-\alpha}}}{\mathcal{P}(t)} \leq \frac{t^{\frac{1}{1-\alpha}}}{\mathcal{P}(t)} \leq \frac{(S_{\mathcal{P}(t)+1})^{\frac{1}{1-\alpha}}}{\mathcal{P}(t)+1} \frac{\mathcal{P}(t)+1}{\mathcal{P}(t)}$$

Now let $t \to \infty$ and use (a) to prove (b).

■

To get asymptotics for the degree of the root note the following relations:

$$\deg(\rho) = \mathcal{P}_0(T_n)$$

where $\mathcal{P}_0$ is the offspring distribution of the root and T_n is the time for the branching process to reach size n.

Now note that Proposition 58 implies that

$$\frac{T_n}{\frac{1}{\theta}\log n} \longrightarrow_{a.s.} 1$$

where $\theta = \theta(\alpha)$ is the Malthusian rate of growth associated with the branching process with offspring distribution $\mathcal{P}$. Finally Lemma 61 implies that:

$$\frac{\mathcal{P}_0(T_n)}{T_n^{\frac{1}{1-\alpha}}} \longrightarrow_P 1$$

Combine to get the asymptotics for the root degree.

■

(d) Preferential attachment with a cutoff: Note that here we can construct the offspring point process as follows:

$$\mathcal{P}_0(t) = \min\{A, N(t)\} + Y_A(t - S_A) \tag{21}$$

where $N(.)$ is a Yule process independent of Y_A which is a rate A Poisson process independent of $N(.)$ (with the obvious convention that $Y_A(t) = 0$ for $t < \infty$) and S_A is the stopping time

$$S_A = \inf\{t : N(t) = A\}$$

By the law of large numbers for the Poisson process we have

$$\frac{Y_A(t)}{t} \longrightarrow_{a.s.} A$$

as $t \to \infty$ By Equation (21) we have similar asymptotics for the offspring distribution $\mathcal{P}_0$ namely:

$$\frac{\mathcal{P}_0(t)}{t} \longrightarrow_{a.s.} 1$$

This coupled with Proposition 58 and Equation (20) implies that:

$$\frac{deg(\rho)}{\frac{1}{\theta_A}\log n} \longrightarrow_{a.s.} A$$

■**(e) Competition induced preferential attachment:** Note that here $\theta_A = 1$ for all A Also note that from the description of the offspring distribution of fertile nodes, for any time t can be constructed as in the Cutoff model as:

$$\mathcal{P}_0(t) = \min\{A, N(t)\} + Y_A(t - S_A)$$

Also note that the actual total degree of the root at time t is $(A-1) + \mathcal{P}_0(t)$ since the birth of the first fertile node is preceded by the birth of $A - 1$ infertile nodes.

The rest of the analysis is now identical to Model (d) and we omit the details.

■

6.3.3 Asymptotics for the Maximal Degree:

The analysis of the maximal degree combines the following two ingredients:

(a) Fix any node v born say t units of time after the initial progenitor (namely the root). Then its degree at time T_n is equal to $\mathcal{P}_v(T_n - t)$ where $\mathcal{P}_v$ denotes the point process representing the offspring distribution for node v.
(b) Large deviations for the above quantity: Here we show that the degree of any node born a sizeable amount of time after the initial progenitor cannot have a large degree. Thus the size of the maximal degree has to be of the same order of magnitude as the size of the root degree.

The above two are enough to get probabilistic bounds on the size of the maximal degree. We exhibit the general mode of calculations via deriving explicit bounds in the case of the Preferential attachment with a cutoff at A and the BA-Preferential attachment model. Since the analysis for the maximal degree for all the other models follows similarly lines as the analysis of one of these two models, we shall omit explicit proofs for the other models.

Maximal degree analysis for the Cutoff model:

Here we analyze the first example, exhibiting the typical mode of analysis. First note that by Proposition 58 we know that given an $\varepsilon > 0$ we can choose a constant $B_\varepsilon > 0$ large enough so that

$$\limsup_{n\to\infty} \mathbb{P}(T_n > \frac{1}{\gamma_A}\log n + B_\varepsilon) \leq \frac{\varepsilon}{2} \tag{22}$$

where γ_A is the Malthusian rate of growth of this particular model and T_n denote the stop times for the population to reach size n. Let $\mathcal{F}(t)$ denote the population tree at time t. Write "max(t)" for the maximum degree of a node in $\mathcal{F}(t)$ and $t_n = \frac{1}{\gamma_A}\log n + B$. Then by Equation (22) it is enough to show that there exists a constant K_ε such that

$$\limsup_{n\to\infty} \mathbb{P}\left(\frac{\max(t_n)}{\log n} > K_\varepsilon\right) \leq \frac{\varepsilon}{2}$$

Now note that since the offspring distribution $\mathcal{P}$ in this case is stochastically dominated by a rate A Poisson point process, we have by large deviations for the Poisson distribution, for any fixed time $t > 0$ and any constant K:

$$\mathbb{P}\left(\mathcal{P}(t) > K \cdot At\right) \leq \exp(-c(K)t) \tag{23}$$

where $c(K)$ is the large deviations rate function for the Poisson distribution and $c(K) \to \infty$ as $K \to infty$. Thus we have

$$\mathbb{P}\left(\max t_n > K \cdot At_n\right) \leq \mathbb{P}(\mathcal{P}\left(t_n\right) > K \cdot At_n) \cdot \mathbb{E}(N(t_n))$$

The result now follows by combining Equation (23) with Theorem 60 on the size of the branching process by time t_n by choosing K large enough.

Maximal degree asymptotics for the BA-model:

We start with a "large deviations" lemma for the Yule process.

Lemma 62 *Let $\mathcal{P}(.)$ denote the standard Yule process started at time 0. Fix a constant K. Then*

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

Proof:

Note that $\mathcal{P}(t - t_0) \sim Geom(e^{t-t_0})$. Thus for any integer M

$$\mathbb{P}(\mathcal{P}(t - t_0) > M) = (1 - e^{-t-t_0})^M \le e^{-Me^{t-t_0}}$$

Now set $M = K \cdot e^t$.

■

The following Lemma collects the basic distributional facts we need to complete the proof:

Lemma 63 *(a) Consider the Branching process $N(t)$ driven by a standard Yule process offspring distribution as in the construction of the Linear preferential attachment model. Then for any fixed time t:*

$$N(t) \le_{st} Z_1(2t) + Z_2(2t)$$

where Z_1 and Z_2 are two independent standard Yule processes.
(b) Fix time $t > 0$ and t_0 with $t_0 + 1 < t$. Let the event $\mathcal{A}_{t,t_0,K}$ be:

> *There exists a node born in the interval $[t_0, t_0 + 1]$ such that the degree of the node by time t is greater than Ke^t.*

Then

$$\mathbb{P}\left(\mathcal{A}_{t,t_0,K}\right) \le e^{-(t_0+1)} + 2e^{-\left[Ke^{t_0} - 2(t_0+1) - \log 2(t_0+1)\right]}$$

Proof:

Part (a) follows via comparing the rates of the two Point processes $N(t)$ and $Z_1(2t) + Z_2(2t)$ (More precisely there is a coupling of the two process so that $N(t) \le Z_1(2t) + Z_2(2t)$ for all t).

Part (b): To prove this part, note that

$$\mathbb{P}(\mathcal{A}_{t,t_0,K}) \quad \le \quad \mathbb{P}\left(N[t_0, t_0 + 1] > 2(t_0 + 1)e^{2(t_0+1)}\right) + \mathbb{P}\left(\{N[t_0, t_0 + 1] \le 2(t_0 + 1)e^{2(t_0+1)}\} \cap \mathcal{A}_{t,t_0,K}\right)$$

where $N[a, b]$ are the number of births of the branching process $N(.)$ in the interval $[a, b]$.

Now note that by (a)

$$\mathbb{P}\left(N[t_0, t_0 + 1] > 2(t_0 + 1)e^{2(t_0+1)}\right) \le e^{-(t_0+1)}$$

Again by the union bound and Lemma 62 we have

$$\mathbb{P}\left(\{N[t_0, t_0 + 1] \le 2(t_0 + 1)e^{2(t_0+1)}\} \cap \mathcal{A}_{t,t_0,K}\right) \le 2(t_0 + 1)e^{2(t_0+1)}e^{-Ke^{t_0}}$$

Simple algebraic manipulations now give us the required result.

■

Now we are finally in a position to prove the following:

Proposition 64 *Given any $\varepsilon > 0$ there exists a constant L_ε such that*

$$\limsup_{n\to\infty} \mathbb{P}\left(\frac{\max(deg)}{\sqrt{n}} > L_\varepsilon\right) \leq \varepsilon$$

Proof: First choose $t_0 > 0$ large so that the infinite series

$$\sum_{i=0}^{\infty} e^{-\left[e^{t_0+i}-2(t_0+i)-\log 2(t_0+i)\right]} \leq \frac{\varepsilon}{3}$$

Then let $K = a \cdot t_0$ where $a > \log \frac{2}{\varepsilon} + 2t_0 + \log 2t_0$

Note that by Proposition 58 for any $\varepsilon > 0$ there exists B_ε so that

$$\limsup_{n\to\infty} \mathbb{P}(T_n > \frac{1}{2}\log n + B) \leq \frac{\varepsilon}{3}$$

Finally let $L_\varepsilon = Ke^B$. Note that

$$\mathbb{P}\left(\frac{\max(deg)}{\sqrt{n}} > L_\varepsilon\right) \quad \leq \quad \mathbb{P}(T_n > \frac{1}{2}\log n + B) + \mathbb{P}(\mathcal{A}_{\frac{1}{2}\log n + B, \leq t_0, K}) + \mathbb{P}(\mathcal{A}_{\frac{1}{2}\log n + B, > t_0, K})$$

where $\mathcal{A}_{\frac{1}{2}\log n + B, \leq t_0, K}$ is the event

> There is a node born **before** time t_0 whose degree by time $t = \frac{1}{2}\log n + B$ is greater that $K.e^t$

and $\mathcal{A}_{\frac{1}{2}\log n + B, > t_0, K}$ is the event

> There is a node born **after** time t_0 whose degree by time $t = \frac{1}{2}\log n + B$ is greater that $K.e^t$

By choice of t_0 and using Lemma 63 we have for all n

$$\mathbb{P}(\mathcal{A}_{\frac{1}{2}\log n + B, > t_0, K}) \leq \frac{\varepsilon}{4}$$

Finally note that by the union bound

$$\mathbb{P}(\mathcal{A}_{\frac{1}{2}\log n + B, \leq t_0, K}) \leq \mathbb{P}\left(N(t_0) > 2t_0 e^{2.t_0}\right) + 2t_0 e^{2.t_0} \mathbb{P}\left(\mathcal{P}(t) > Ke^t\right)$$

where for simplicity we gave written t for $\frac{1}{2}\log n + B$.

Note that by Lemma 63, by choice of t_0

$$\mathbb{P}\left(N(t_0) > 2t_0 \cdot e^{2t_0}\right) \leq e^{-t_0} \leq \frac{\varepsilon}{4}$$

Finally note that

$$2t_0 e^{2.t_0} \mathbb{P}(\mathcal{P}(t) > Ke^t) \leq 2t_0 e^{2t_0} e^{-K} \leq \frac{\varepsilon}{4}$$

by choice of K ■

Remark 1: The asymptotic probability bound for the maximal degree for the other models is similarly computed. The analysis is essentially split up into two general modes of analysis. In the first class of models the models are essentially similar to the cutoff model where the offspring distribution grows at the most super-linearly in time namely $\mathcal{P}(t) \sim t^\beta$ where $\beta \geq 1$. The proofs in this case tend to be relatively easy and follow from simple large deviation type inequalities. In the second class of models, the associated offspring distribution grows exponentially namely $\mathcal{P}(t) \sim e^{\lambda t}$ for some $\lambda > 0$. In this case we need more advanced "bracketing" techniques, according to the time when the node is born to complete the proof. This is exhibited in the analysis of the BA-preferential attachment model where we analyzed the behavior of nodes according to the interval $[t, t+1]$ when the node was born.

Remark 2: What the above computations also imply is that the maximal degree essentially occurs at a finite distance from the root ρ, namely in the local "neighborhood" of the root. More precisely we have for BA-Preferential attachment model and the Linear preferential attachment model, the following theorem:

Theorem 65 *Fix B and time t. Recall that $\mathcal{F}_t$ denotes the Branching process tree observed till time t.*

$$\limsup_{t\to\infty} \mathbb{P}\left(\textit{node with max degree in } \mathcal{F}_t \textit{ born at time} > B \textit{ after } \rho\right) \leq \varepsilon(B)$$

where $\varepsilon(B) \to 0$ as $B \to \infty$.

See Section 7 where we briefly mention how this theorem can presumably be used to prove distributional convergence of the maximal degree of $\mathcal{T}_n$, properly normalized.

6.3.4 Height

The analysis of the height in all of the above models follows almost immediately from a general result of Kingman [21]. This was essentially what was used in a beautiful paper of Pittel [30] to prove the result in the special case of the Linear attachment model. To paraphrase Kingman's result in our context, we first need some notation.

Consider a continuous time branching process as in Section 3, driven by a point process $\mathcal{P}(.)$, with mean intensity $\mu(t)$.

Define

$$\phi(\theta) = \theta \int_0^\infty e^{-\theta t} \mu(t) dt$$

Let B_k be the first time that an individual in the k^{th} generation (namely an individual at graph distance k from the root) is born. Then Kingman [21] proves the following

Theorem 66 *[21] Assume that there exists a θ_0 so that*

$$1 < \phi(\theta_0) < \infty$$

For $a > 0$ write

$$\mu(a) = \inf\{\phi(\theta) \cdot e^{\theta a}; \theta \geq \theta_0\}$$

and write

$$\gamma = \sup\{a : \mu(a) < 1\}$$

Then

$$\lim_{k\to\infty} \frac{B_k}{k} \longrightarrow_{a.s.} \gamma$$

To prove our results about heights note that, letting h_n be the height of the tree we have for the times of birth, the following inequality:

$$B_{h_n} \leq T_n \leq B_{h_n+1}$$

Dividing by h_n we have

$$\frac{B_{h_n}}{h_n} \leq \frac{T_n}{h_n} \leq \frac{B_{h_n+1}}{h_n}$$

Now letting $n \to \infty$ and using Kingman's result and Proposition 58 we have

$$\frac{h_n}{\frac{1}{\theta}\log n} \longrightarrow_{a.s.} \frac{1}{\gamma}$$

where θ is the Malthusian growth parameter corresponding to the particular model.

■

Thus we find for each model, the normalized random variables $h_n/\log n$ converges to a constant. Call this the **height limit constant** for the particular model. Although the above gives us a formula to compute the limit constants, other than the Preferential attachment model, it seems very hard to compute explicitly the limit constants in any of the other cases. It might be interesting to numerically solve for these constants. What is easily provable is that all the models have strict logarithmic growth namely:

Proposition 67 *For all the models we deal with, the limit constants associated with the height are strictly positive.*

Proof: Note that the limit constant in all the models is by construction equal to $\frac{\gamma}{\theta}$ where θ is the Malthusian rate of growth of the model and γ is as defined above. Thus to prove the Proposition, it is enough to show that for all our models, $\gamma > 0$. Note that by definition $\gamma := \sup\{\mu(a) < 1\}$ where

$$\mu(a) = \inf\{\beta \geq \theta : \phi(\beta)e^{\beta a}\}$$

Thus it is enough to show that

$$\mu(a) \to 0 \qquad \text{as} a \downarrow 0$$

This follows easily since for fixed a and θ,

$$\phi(\theta)e^{\theta a} = e^{\theta a} \cdot \mathbb{E}(\mathcal{P}(0, T_\theta])$$

where $\mathcal{P}$ is the point process of the offspring distribution and T_θ is independent of $\mathcal{P}$ and is a rate θ exponential, and thus mean $\frac{1}{\theta}$ exponential random variable. This implies in particular that $E(\mathcal{P}(0, T_\theta]) \to 0$ as $\theta \to \infty$. Thus $\mu(a) \to 0$ as $a \downarrow 0$.

■

6.3.5 Relative sizes of subtrees

Let v_k be the k^{th} node born into the population and let $L_k(n)$ be the load on k at time n, namely $L_n(k)$ is the size of the subtree attached to vertex k at time n. We are able to derive asymptotics for the special case of the Preferential attachment model, where the special structure of the offspring distribution and the explicit knowledge of the limiting random variable namely Lemma 59, help us to get a simple proof of the fact that in the case of the Preferential attachment model

$$\frac{L_n(k)}{n} \longrightarrow_L \beta(2, 2\cdot(2k-1))$$

To see this, consider the continuous time branching process construction of this model. Let $L_k(t)$ be the size of the subtree attached to vertex v_k at time t and note that $L_k(t) = 0$ when $t < T_k$. Thus it is enough to prove that

$$\frac{L_k(t)}{N(t)} \longrightarrow_{a.s.} B$$

where the random variable $B \sim \beta\left(2, 2\cdot(2k-1)\right)$

Also note that at the time of birth of v_k, namely T_k, the process grows at an infinitesimal rate of $2k-1$, whatever be the internal structure of the tree uptil this time. Suppose the $k-1$ preceding nodes have out degree $d_1, d_2, \ldots d_{k-1}$ with $d_k = 0$, This implies that at time T_k, vertex i reproduces at rate $d_i + 1$. We consider the following construction of the branching process after time T_k:

(a) Let $(\mathcal{F}^i(.))_{1\le i\le(2k-1)}$ be independent continuous time branching process driven by the Yule process.

(b) For $t > 0$ and $1 \le i \le k$, let $\mathcal{A}_i(t)$ as the size of the subtree attached to vertex i at time $t+T_k$. Then we can construct these processes as $\mathcal{A}_1(t) = \cup_1^{d_1+1}\mathcal{F}^i(t) - d_1$, $\mathcal{A}_2(t) = \cup_{d_1+2}^{d_2+3}\mathcal{F}^i(t) - d_2$, ... , $\mathcal{A}_k(t) = \mathcal{F}^{2k-1}(t)$. Note that by construction, for any $t > 0$, conditional on $\mathcal{F}(T_k)$, the size of the population at time $t + T_k$ satisfies the identity

$$N(T_k + t) = \sum_1^k |\mathcal{A}_k(t)| = \sum_1^{2k-1} |\mathcal{F}^i(t)| - (k-1)$$

Thus the proportion of the population which is a direct descendent of vertex k is given by

$$\frac{L_k(t+T_k)}{N(k+T_k)} = \frac{|\mathcal{F}^{2k-1}(t)|}{\sum_1^{2k-1} |\mathcal{F}^i(t)| - (k-1)} \tag{24}$$

$$= \frac{|\widetilde{\mathcal{F}}^{2k-1}(t)|}{\sum_1^{2k-1} |\widetilde{\mathcal{F}}^i(t)| - (k-1)/e^{2t}} \tag{25}$$

where we define $\widetilde{\mathcal{F}}^i(t) := \mathcal{F}^i(t)/e^{2t}$

Now note that by Lemma 59 we have

$$|\widetilde{\mathcal{F}}^i(t)| \longrightarrow_{a.s.} W_i/2$$

where W_i are iid and have a gamma distribution with density xe^{-x} on $\mathbb{R}^+$. This implies that

$$\frac{L_k(t)}{N(t)} \longrightarrow_{a.s.} \frac{W_{2k-1}}{\sum_1^{2k-1} W_i}$$

The result follows since by the properties of the Gamma distribution

$$\frac{W_{2k-1}}{\sum_1^{2k-1} W_i} \sim \beta(2, 2 \cdot (2k-1))$$

Remark: The above argument can be suitably modified to get convergence in many of the other preferential attachment models. However due to the lack of an explicit analytic knowledge of the distribution of the limiting random variable for the other models, we do not delve deeper into this aspect for the other models.

6.3.6 Asymptotics for the percolation component of the root

For proving properties regarding the percolation component of the root, we first consider the following coloring process associated with the original continuous time branching process.

Construction 68 *Fix any one of the models in Section 2.1. Denote the point process representing the offspring distribution as $\mathcal{P}(.) \equiv (S_1, S_2, \ldots)$. Let $\mathcal{T}_n$ denote the tree on n vertices, grown according to the particular attachment scheme of the model under consideration. Define the following coloring of the nodes of the branching process as:*
(a) Mark the root blue.
(b) Mark each new birth **blue** *with probability p or alternatively* **red** *with probability $(1-p)$.*

Denote the blue component attached to the root at time t by $\mathcal{F}^B(t)$. Then this subtree of $\mathcal{F}_t$ corresponds to the percolation component of $\mathcal{T}_{N(t)}$ where $\mathcal{T}_n$ is the tree on n vertices.

From the above construction we immediately have the following distributional fact about $\mathcal{F}^B(t)$

Proposition 69 *Fix any of the preferential attachment models for growing trees of Section 2.1. Let $\mathcal{P}(.) \equiv (S_1, S_2, \ldots)$ be the point process representing the offspring distribution and let $\mathcal{F}(.)$ be the branching process with offspring distribution $\mathcal{P}$. Let $X_1, X_2, \ldots$ be i.id. Geometric(p) random variables namely*

$$P(X_1 = k) = p(1-p)^{k-1} \qquad k = 1, 2, 3, \ldots$$

Consider the **pruned** *point process given by the points $\tilde{S}_i^p$ where*

$$\tilde{S}_i^p = S_{X_1 + \ldots X_k}$$

Then the blue component defined in construction 68, namely $\mathcal{F}^B(t)$, has the same distribution as branching process driven by the point process $\mathcal{P}^p$ with distribution

$$\mathcal{P}^p \equiv (\widetilde{S}_1^p, \widetilde{S}_2^p, \ldots)$$

Proof: Obvious.

■

Given any model of attachment, let γ_p denote the Malthusian parameter associated with the corresponding **pruned** branching process $(\mathcal{F}^B(t))_{t>0}$ Then note that by the above construction, since the percolation component of the root essentially behaves as a continuous time branching process, the whole theory of fringe and extended fringe distributions and `-sin`-trees carries over to this setup as well. In particular, the percolation component of the root also converges in *probability fringe* sense to an infinite `sin`-tree. We shall not analyze this `sin`-tree in it's entirety but show one can do analysis such as understand the limiting degree distribution of the percolation component of the root.

To analyze the local structure of the percolation component of the root, note that Corollary 50 implies the following:

Corollary 70 *Assume that the associated branching process $\mathcal{F}^B(.)$ satisfies the Assumptions 48. Let $N^p_{\geq k}(n)$ be the number of nodes in the percolation component of the root, $\mathcal{C}_n(\rho)$ which have out-degree greater than k. Then*

$$\frac{N_{\geq k}(n)}{n} \longrightarrow_P E(e^{-\gamma_p \tilde{S}^p_k})$$

as $n \to \infty$.

This fact is enough to prove the local structural assertions of Section 5 pertaining to the percolation cluster of the root namely Theorem 15(b) and Theorem 19(b) .

Finally to get asymptotics for the size of $\mathcal{C}_n(\rho)$ we have the following Theorem.

Theorem 71 *Let γ denote the Malthusian rate of growth of the original process and let γ_p denote the Malthusian rate of growth of the imbedded $\mathcal{F}^B(t)$ process. Assume that the associated branching process satisfies the Assumptions 48. Then there exists a random variable $W_p > 0$ such that*

$$\frac{|\mathcal{C}_n(\rho)|}{n^{\frac{\gamma_p}{\gamma}}} \longrightarrow_{a.s.} W_p$$

as $n \to \infty$.

Proof: Note that we have

$$\mathcal{C}_n(\rho) = \mathcal{F}^B(T_n)$$

where T_n are the stopping times as described in Proposition 58 and $\mathcal{F}^B(\cdot)$ is the pruned branching process whose construction was described above. Now note that by Theorem 57, we have

$$\frac{\mathcal{F}^B(t)}{e^{\gamma_p t}} \longrightarrow_{a.s.} X_p$$

where $X_p > 0$.

This in particular implies

$$\frac{|\mathcal{C}_n(\rho)|}{e^{\gamma_p T_n}} \longrightarrow_{a.s.} X_p$$

Also note that by Proposition 58

$$\frac{n^{\frac{\gamma_p}{\gamma}}}{e^{\gamma_p T_n}} \longrightarrow W^{\frac{\gamma_p}{\gamma}}$$

where W is the limiting random variable associated with the original process. Note that W and X_p are not independent. Combining and writing $W_p = \frac{X_p}{W^{\frac{\gamma_p}{\gamma}}}$, we get the result.

■

To use Theorem 71 we need to verify that the branching processes $\mathcal{F}^B(t)$ associated with the Preferential attachment model and the Linear Preferential attachment models satisfy the Assumptions 48. The calculations are very similar to the computations done in Proposition 53 and we don't repeat it here. Thus Theorem 71 implies Theorem 15(a) and 19(a).

Finally to prove statements regarding the heights of the percolation component of the root, we use Kingman's result [21]. The analysis is identical to the one carried out in Section 6.3.4 and we skip it here.

6.4 Preferential attachment with Fitness

To analyze the fitness models, we first need to introduce the concept of multi-type branching processes. The basic conceptual idea is the same as that of the usual branching processes introduced in Section 3. The minor variation is that each node is born with a particular *type* belonging to a type space. Although these processes can be constructed on abstract type spaces, all our processes are constructed with type space restricted to some bounded subset A of the interval $[0, M]$ for some fixed constant $M > 0$.

6.4.1 Multi-type branching processes:

We now delve into the properties of Multi-type branching processes. Most of the theory has been taken from [20] and paraphrased to suit our particular needs. The technical conditions used in Assumptions 73 are taken from [27]. The nice part of this abstract theory is that, in the setting of the fitness models, one can often do explicit computations, essentially because of the nice exact distributional properties of the Yule process.

Before describing multi-type branching processes we need some extra notation. Although necessary for mathematical completeness, on a first reading, this notation seems to obfuscate the essential idea, so we first describe the process in words.

> We have a continuous time branching process where each node v is born with a *type* $f_v \in \mathbb{R}^+$. This type determines the rate at which the node reproduces, and the corresponding types of children it has. Given a node's type it reproduces at times which are independent of the rest of the population and giving birth to nodes of type which depend only on the type of this parent node and independent of the rest of the population.

More precisely:

Notation: Let $\mathbb{N} = \{1, 2, 3, \ldots\}$. Identify the tree representing the genealogy of the population the with the labeling

$$\{0\} \cup \cup_{n=1}^{\infty} \mathbb{N}^n$$

with the understanding that 0 corresponds to the root, and $(i_1, i_2, \ldots i_k)$ is the i_k^{th} child of $(i_1, \ldots, i_{k-1})$. Let $(\mathcal{N}(\mathbb{R}^+ \times \mathbb{R}^+), \mathcal{M})$ denote the measure space of counting measures on $\mathbb{R}^+ \times \mathbb{R}^+$, with $\mathcal{M}$ denoting the sigma field making all projection maps $f_B : \mathcal{N}(\mathbb{R}^+ \times \mathbb{R}^+) \to \mathbb{N}$,

$$f_B(\mu) = \mu(B)$$

measurable for all $B \in \mathcal{B}(\mathbb{R}^+ \times \mathbb{R}^+)$. We shall think of the first co-ordinate as the type of the node born and the second co-ordinate representing the time of birth of the niode .

Thus all the point process measures $\mathbf{n}(.)$ in $\mathcal{N}(\mathbb{R}^+ \times \mathbb{R}^+)$we work with can be enumerated as

$$\mathbf{n}(.) \equiv \{[\sigma(1), \tau(1)], [\sigma(2), \tau(2)], [\sigma(3), \tau(3)], \ldots\}$$

with $\tau(1) < \tau(2) < \ldots$ representing the times of birth and $\sigma(i)$ representing the corresponding types of offspring.

Note that if we have counting measures such that say $(0, 0)$ is a limit point of points in the support of the counting measure or there exist more than one point with the same y co-ordinate, then obviously we cannot enumerate the points as described. For the type of counting measures we deal with, the above construction can always be done.

Finally we shall let Q denote the reproduction kernel, giving us the offspring time and type distribution conditional on the type of the parent. More precisely let $Q(.,.) : \mathbb{R}^+ \times \mathcal{M} \to [0, 1]$ satisfying the properties:

(a) For each fixed r, $Q(r, .)$ is a probability measure on $(\mathcal{N}(\mathbb{R}^+ \times \mathbb{R}^+), \mathcal{M})$
(b) For each fixed $B \in \mathcal{M}$, the function $Q(., B)$ is a measurable function on $(\mathbb{R}^+, \mathcal{B}(\mathbb{R}^+))$.

In our context the mean intensity measure μ of this kernel is extremely important and is defined below. Fix $r \in \mathbb{R}^+$ and let the random point process $\xi(.,.) \sim Q(r, .)$. Define $\mu(.,.) : \mathbb{R}^+ \times \mathcal{B}(\mathbb{R}^+ \times \mathbb{R}^+) \to \mathbb{R}^+$ as the measure:

$$\mu(r, dt \times ds) := \mathbb{E}_r(\xi(ds \times dt)) \tag{26}$$

where the subscript r in $\mathbb{E}_r$ is used to emphasize the fact that the distribution of ξ depends on r.

Construction 72 (Multi-type BP with type space $\mathbb{R}^+$:) *A multi-type branching process with reproduction kernel Q and initial probability distribution κ, is a Markov process constructed as follows:*

1. *Start with a single initial ancestor labeled* 0 *with type* $f_0 \in \mathbb{R}^+$, *chosen at random with distribution* κ.

2. *Given* f_0 *chose the reproduction point process of the root with distribution* $\xi_0 \sim Q(f_0, .)$. *Write* ξ_0 *as*

$$\xi_0 \equiv \{[\sigma(1), \tau(1)], [\sigma(2), \tau(2)], [\sigma(3), \tau(3)], \ldots\}$$

The interpretation being that that ξ_0 represents the reproduction process for the root, with the first child born at time $\tau(1)$ with type $\sigma(1)$, the second child born at time $\tau(2)$ with type $\sigma(2)$ and so on.

3. *The process is defined iteratively: Given the time of birth of a node $v \in \{0\} \cup \cup_{n=1}^{\infty} \mathbb{N}^n$ and it's type f_v conditional on the rest of the information upto time T_v, chose $\xi_v \sim Q(f_v, .)$. Write this reproduction process for node v as:*

$$\xi_v \equiv \{[\tau_v(1), \sigma_v(1)], [\tau_v(2), \sigma_v(2)], [\tau_v(3), \sigma_v(3)], \ldots\}$$

Then the times of birth of the offspring of node v are $T_v + \tau_v(1), T_v + \tau_v(2), \ldots$ with corresponding types $\sigma_v(1), \sigma_v(2), \ldots$.

To actually be able to construct such a Markov process without worrying about technicalities such as explosion etc, we need to make some technical assumptions on the mean intensity measure μ. These assumptions will involve the concepts of $\alpha-$ recurrence and irreducibility for general state Markov Chains. Since stating and developing these concepts in the general case requires a lot machinery we shall not state the full definitions but quickly mention what needs to be checked in our particular setup.

First we need some notation. Define the Laplace transform of the mean intensity measure μ (Equation (26)) as

$$\hat{\mu}(r, ds; \alpha) = \int_0^{\infty} e^{-\alpha t} \mu(r, ds \times dt) \tag{27}$$

For simplifying calculations later on in the proof, we note that a simple Fubini-type calculation implies that we have the more probabilistic interpretation of $\hat{\mu}$:

$$\hat{\mu}(r, ds, \alpha) = \mathbb{E}\left(\#\text{of offspring of type ds born in the interval } (0, T_\alpha] | \text{mother-type} = r\right) \tag{28}$$

where T_α is an exponential rate α variable independent of the offspring reproduction point process ξ.

Assumption 73 *(a) The process is Malthusian and supercritical, namely in the terminology developed in [28], the mean intensity kernel μ is is irreducible and there exists an $\alpha > 0$ such that the Laplace transform of the kernel μ, namely $\hat{\mu}(r, ds, \alpha)$ is $\alpha-$recurrent. This α is called the* **Malthusian rate of growth** *of the process. By general theory (see [33]), there exists a $\sigma-$ finite measure π on $\mathbb{R}^+$ and a $[\pi]$ a.e. finite and strictly positive function h (both unique upto multiplicative constants) such that*

$$\begin{aligned} \int_{\mathbb{R}^+} \pi(dr) \hat{\mu}(r, ds; \alpha) &= \pi(ds) \\ \int_{\mathbb{R}^+} h(s) \hat{\mu}(r, ds; \alpha) &= h(r) \end{aligned}$$

Also assume strong α-recurrence namely $h \in L^1[\pi]$ and:

$$0 < \beta = \int_{\mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{R}^+} t \cdot e^{-\alpha t} h(s) \mu(r, ds \times dt) \pi(dr) < \infty$$

Then π is a finite measure and we shall normalize h and π so that π is a probability measure and

$$\int_{\mathbb{R}^+} h d\pi = 1$$

(b) Assume that the reproduction kernel μ for the reproduction kernel Q satisfies the boundedness property

$$\sup_{s\in\mathbb{R}^+} \mu(s,\mathbb{R}^+\times[0,\varepsilon]) < \infty$$

Note that $\mu[s,\mathbb{R}^+\times[0,\varepsilon]]$ has the verbal meaning

> *"The expected number of offspring of all types born in the interval $[0,\varepsilon]$ starting from a mother of type s".*

(c) The **xlogx** *condition: Define the random variable*

$$\begin{aligned}\bar{\xi} &= \int_{\mathbb{R}^+\times\mathbb{R}^+} e^{-\alpha t} h(s)\xi(ds\times dt) \\ &= \int_0^\infty h(s)\mathbb{E}(\#\textit{of points of fitness } ds \textit{ in } [0,T_\alpha]\ |\xi)\end{aligned}$$

Here T_α is independent of ξ and has an exponential distribution with rate α ; α being the Malthusian rate of growth parameter and the expectation is taken over the randomness in T_α with ξ fixed.

Then we assume that

$$\mathbb{E}_\pi(\bar{\xi}\cdot\log^+(\bar{\xi})) < \infty$$

where π and h are as defined in part (b).

See [20] and the references cited within for more details. The essential point is that the above technical conditions are easily checkable in our setup.

6.4.2 Connection to the fitness model

Note that the input for the two fitness models was a probability measure ν with bounded support on $\mathbb{R}^+$. Similar to Proposition 43, we shall see that this branching process, sampled at times T_n, has the same distribution as our trees $\mathcal{T}_n$. This is the crucial link which helps us carry out our analysis. We state the exact specialization of the reproduction kernel Q and initial measure κ, for our two models. For both models, the initial probability measure, namely the fitness of the root is the same as ν, namely $\kappa = \nu$.

Additive fitness: Fix a fitness (or type) $s \in$ the support of ν. Then the random point process $\mathcal{P}$ on $\mathbb{R}^+\times\mathbb{R}^+$ with distribution $Q^+(s,\cdot)$ is constructed as follows: Let $\tau(1),\tau(2),\ldots$ be the points of a Markov point process $\mathcal{C}_s$ (constructed independently of the sequence σ) with transition rates:

$$\mathbb{P}\left(\mathcal{C}_s(t+dt)-\mathcal{C}_s(t)=1|\,\mathcal{C}_s(t)=k\right) = (k+1+s)\cdot dt + o(dt) \tag{29}$$

Then $\mathcal{P}$ is defined as:

$$\mathcal{P} := ((\sigma(1),\tau(1)),(\sigma(2),\tau(2)),\ldots)$$

We call Q^+ the *reproduction kernel* associated with the additive fitness model, and we call the associated multi-type branching process, the *additive fitness* branching process.

Multiplicative fitness: Fix a fitness (or type) $s \in$ the support of ν. Then the random point process $\mathcal{P}$ on $\mathbb{R}^+\times\mathbb{R}^+$ with distribution $Q^\times(s,\cdot)$ for the multiplicative fitness model is constructed as follows: Let $\tau(1), \tau(2), \ldots$ be the points of a Markov point process $\mathcal{C}_s$ (constructed independently of the sequence σ) with transition rates:

$$\mathbb{P}\left(\mathcal{C}_s(t+dt) - \mathcal{C}_s(t) = 1 \middle| \mathcal{C}_s(t) = k\right) = s(k+1) \cdot dt + o(dt) \tag{30}$$

Then $\mathcal{P}$ is defined as:

$$\mathcal{P} := ((\sigma(1), \tau(1)), (\sigma(2), \tau(2)), \ldots)$$

We call $Q^{\times}$ the *reproduction kernel* associated with the multiplicative fitness model, and we call the associated multi-type branching process, the *multiplicative fitness* branching process.

The connection between the fitness models for generating random trees and the multi-type branching processes with the respective kernels as above is given by the following trivial result:

Proposition 74 *Consider the additive (resp. multiplicative) fitness branching process denoted by $(N(t), \mathcal{F}(t))$ say. Assume that the fitness measure ν is supported on a bounded subset of $\mathbb{R}^+$, and satisfies the Assumptions of 73. Define T_n as the stopping times:*

$$T_n = \inf\{t : N(t) = n\}$$

Then we have the distributional identity

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

where $\mathcal{T}_n$ is the tree on n nodes constructed via the additive (resp. multiplicative) fitness scheme.

Proof: Obvious by our definition of the rates of the point processes.

■

We begin by collecting some simple facts about the mean intensity measure connected to the two models. First note that a simple Fubini-type computation implies that:

$$\mu(r, ds, [0,t]) = \mathbb{E}\left(\#\text{ of offspring with fitness ds born in time [0,t]} \middle| \text{mother-type} = r\right) \tag{31}$$

Lemma 75 *(a) For the additive fitness model:*

$$\mu(r, ds, [0,t]) = \nu(ds)(e^t - 1 + r \cdot t)$$

Thus the kernel $\hat{\mu}(.,.;\alpha)$ is given by the formula:

$$\hat{\mu}(r, ds, \alpha) = \nu(ds)\left(\frac{1}{\alpha - 1} + \frac{r}{\alpha}\right)$$

(b) For the multiplicative model we have:

$$\mu(r, ds, [0,t]) = \nu(ds)(e^{rt} - 1)$$

Thus the kernel $\hat{\mu}(.,.;\alpha)$ is given by the formula:

$$\hat{\mu}(r, ds; \alpha) = \nu(ds) \cdot \frac{r}{\alpha - r}$$

The kernel $\hat{\mu}$ for the multiplicative model is defined only for $\alpha > M$.

Proof: Note that for the additive model, for a mother of type r, the number of offspring in the interval [0,t] has the distribution

$$\mathcal{C}_r([0,t]) =_d N(t) + Poi_r(t)$$

where $N(.)$ is a standard Yule process independent of $Poi_r(.)$, which is a rate r poisson process. Also note that each offspring has type s with probability $\nu(ds)$. Now use Equation (31)

For the multiplicative model we have the same argument, except that in this case

$$\mathcal{C}_r([0,t]) =_d N(rt)$$

where $N(rt)$ is a standard Yule process. ■

The following Proposition describes the assumptions that need to be made on the fitness measure so that the general theory is applicable to our setting. We shall not give a complete proof since this would involve an unnecessarily long detail into the theory of general state Markov Chains. Let it suffice to say that the following is trivial to check via the methodology developed in [28].

Proposition 76 (Irreducibility) *Assume that the fitness measure ν is supported on some strictly positive and bounded subset $[0, M]$ of the real line. Then the kernel $\mu(r, ds)$ in both the additive and multiplicative models is irreducible with maximal irreducibility measure ν.*

Thus the irreducibility assumption of Assumption 73(a) is always satisfied.

To check the condition of α-recurrence and to explicitly compute the Malthusian rate of growth parameter, we have the following Proposition. We also show how to compute the eigen-measure π and the eigen-function h occurring in Assumption 73 for the two models. Note that these two quantities play a central role in describing local asymptotics, so it is very important to have explicit forms for these quantities.

Proposition 77 (α-recurrence) *Assume that ν is supported on some strictly positive and bounded subset $[0, M]$ of the real line. Also without loss of generality we shall assume that M is the essential supremum of the fitness measure ν.*

(a) For the additive fitness model note that the equation

$$g(\beta) = 1 \tag{32}$$

always has a unique positive solution, where $g(\beta)$ is the monotonically decreasing function on the interval $(1, \infty)$:

$$g(\beta) = \int_{\mathbb{R}^+} \nu(dr) \left(\frac{1}{\beta - 1} + \frac{r}{\beta} \right) \tag{33}$$

Denote this root of the equation by α. Then the kernel $\hat{\mu}(r, ds, \alpha)$ is α-recurrent. The eigen-measure π is given by $\pi = \nu$ and the eigen-function h is given by

$$h(r) = \left(\frac{1}{\alpha - 1} + \frac{r}{\alpha} \right)$$

(b) For the multiplicative fitness model assume that the equation

$$g(\beta) = 1 \tag{34}$$

where $g(.)$ is the monotonically decreasing function defined on (M, ∞)

$$g(\beta) = \int_{\mathbb{R}^+} \frac{r}{\beta - r} \cdot \nu(dr) \tag{35}$$

Denote the root of the equation by α. Then the kernel $\hat{\mu}(r, ds, \alpha)$ is α-recurrent. The eigen-measure π is given by $\pi = \nu$ and the eigen-function h is given by the formula

$$h(r) = \frac{r}{\alpha - r}$$

Remark: In either case the α is called the Malthusian rate of growth parameter for the model in question.

Proof: The explicit formulas for the eigen-measure and eigen-function follow from the defining and normalizing Equations of π and h in Assumption 73 where we use the formulas of $\hat{\mu}$ from Lemma 75. Thus we need to only prove α-recurrence. For this we first need some notation.

First define the collection of Borel sets getting positive measure under ν

$$\mathcal{B}^+ := \{A \subset [0, M] : \nu(A) > 0\}$$

For a fixed β, let us define the n^{th} iterate, $\hat{\mu}^{(n)}$ of the Laplace transform $\hat{\mu}(r, ds; \beta)$ recursively as:

$$\hat{\mu}^{(n)}(r, ds; \beta) = \int_0^M \hat{\mu}^{(n-1)}(r, dz; \beta)\hat{\mu}(z, ds; \beta)$$

Define the $R(\beta)$ kernel as:

$$R(r, \cdot; \beta) = \sum_1^\infty \hat{\mu}^{(n)}(r, \cdot, \beta)$$

Note that by the Fubini interpretation of the the kernel $\hat{\mu}$, namely Equation(28),for a fixed r $\hat{\mu}(r, \cdot; \beta)$ is monotonically decreasing in β. Thus define the *convergence parameter* α as the number:

$$\alpha := \inf\{\beta > 0 : R(r, A; \beta) < \infty \text{ for some r} \in [0, M],\ A \in \mathcal{B}^+\} \tag{36}$$

Now we come to the definition of α-recurrence (see [28],[33]) modified for our setup:

Proposition 78 *In our setup, the kernel $\hat{\mu}$ is called α-transient if*

$$R(r, A; \alpha) < \infty$$

for some $r \in [0, M]$ and $A \in \mathcal{B}^+$ and is called α-recurrent otherwise.

The proof of this fact is obvious from the definition of α-recurrence as stated in [28] and we shall not prove it.

To prove that solutions of Equations (33) and (35) is the same as α-recurrence as defined above, we need explicit formulas for the kernel $R(\cdot, \cdot; \beta)$ in our setup. Using the definitions of the various constructs it is not hard to prove

Lemma 79 *The n^{th} iterate of the kernel $\hat{\mu}$ for both models satisfy the relation:*

$$\hat{\mu}^{(n)}(r, ds, \beta) = [g(\beta)]^{n-1}\, \hat{\mu}(r, ds; \beta)$$

where g is the function defined by Equation (33) and (35) corresponding to the two cases.

Thus Lemma 79 proves the α-recurrence in our setup since

$$R(r, ds; \beta) = \hat{\mu}(r, ds; \beta) \sum_{1}^{\infty} [g(\beta)]^{n-1}$$

which converges for $g(\beta) < 1$ and diverges for $g(\beta) \geq 1$. The Malthusian rate of growth parameter defined in Proposition 77 is the same as the convergence parameter of the kernel $R(\cdot, \cdot; \beta)$.

■

Summarizing our findings and paraphrasing the general assumptions 73 in the special case of the fitness models implies we make the following assumptions on the fitness measure ν

Assumption 80 *(a)The measure ν has support on some finite bounded positive subset $[0, M]$, with M denoting the essential supremum of the measure ν.*
(b) For the multiplicative fitness model, we assume that there exists a positive solution of the equation:

$$\int_0^M \frac{s}{\beta - s} \nu(ds) = 1$$

For both models, let α denote the Malthusian rate of growth parameter as given by Eqns (33) and (35)
(c) Consider the random variable $\bar{\xi}$ constructed as follows:

1. *Choose a random fitness $X \sim \nu$.*

2. *Conditional on $X = s$, choose a random point process $\mathcal{C}_s$ with distribution as given by Equations (29) and (30) in the two cases. Let ξ denote the joint point process of births and fitnesses, where the times of birth are given by the times of $\mathcal{C}_s$ and each birth is independently given a fitness, distributed according to ν*

3. *Let T_α be independent exponentially distributed with rate α, where α is the Malthusian rate of growth parameter.*

4. *Finally define the random variable*

$$\bar{\xi} = \int_0^M h(s)\mathbb{E}\left(\#\ \text{of points of fitness } ds \text{ in } [0, T_\alpha] | \xi\right)$$

   where $h(\cdot)$ is the eigen-function.

Then we assume

$$\mathbb{E}(\bar{\xi} \log^+ \bar{\xi}) < \infty$$

6.4.3 Local Properties

We are now in a position to state the main result of [20], relevant for our local computations. However we need to recall some notation.

Definitions: Recall that $\mathcal{T}$ denoted the space of all finite rooted *unmarked* trees. Given a rooted tree $\mathcal{T}$ and a vertex $v \in \mathcal{T}$ recall the definition of the fringe subtree rooted at v $b_0(\mathcal{T}, v)$, namely Definition 8.

Fix a continuous time multi-type branching process $(\mathcal{F}(\cdot))_{t>0}$., Fix any time $t > 0$ and consider $\mathcal{F}(t)$ is a marked rooted tree (giving us all the information relevant for the population upto time t, such as times of birth etc) and consider the projection $g : \mathcal{F}(t) \to \mathcal{T}$, namely we retain only the lineage structure of the population up to time t and remove all marks from $\mathcal{F}(t)$.

Finally for any fixed time t, and any vertex $v \in \mathcal{F}(t)$, let f_v denote the fitness of the vertex v and let $f_{\mathtt{mom(v)}}$ denote the fitness of the mother of vertex of v.

Fix any two subsets $A_1, A_2 \subset [0, M]$ and a subset $T \subset \mathcal{T}$ and consider the empirical average:

$$X_t(A_1, A_2, T) = \frac{1}{|\mathcal{F}(t)|} \sum_{v \in \mathcal{F}(t)} 1\{f_v \in A_1, f_{\mathtt{mom(v)}} \in A_2, b_0(\mathcal{F}(t), v) \in T\}$$

Then we paraphrase the local results of [20] as follows:

Theorem 81 *Assume that the multi-type branching process satisfies all the Assumptions 73. Then the empirical averages satisfy*

$$X_t(A_1, A_2, T) \longrightarrow_P \mathbb{P}(f_1 \in A_1, f_2 \in A_2, \mathcal{F}^* \in T)$$

as $t \to \infty$.

Here the random variables f_1, f_2 and the random rooted tree $\mathcal{F}^*$ are constructed as follows:
(a) $f_1 \sim \pi$ where π is the eigen-measure constructed in Assumption 73.
(b) The conditional distribution of f_2 given f_1 is given by:

$$\mathbb{P}\left(f_2 \in ds | f_1 = r\right) = \pi(ds) \cdot \frac{\hat{\mu}(s, dr; \alpha)}{\pi(dr)}$$

(c) $\mathcal{F}^*$ is conditionally independent of f_2, given f_1 and has the distribution $g(\mathcal{F}(T_\alpha))$ where $\mathcal{F}(\cdot)$ is a multi-type branching process with the initial ancestor having fitness f_1; T_α is an exponential random variable with rate α independent of all the other random variables; and the process $\mathcal{F}(\cdot)$ and g is the projection of the marked random tree $\mathcal{F}(T_\alpha)$ into the unmarked space $\mathcal{T}$.

The same result holds if T_n are a sequence of almost surely finite stopping times with $T_n \uparrow \infty$.

This immediately gives us the necessary information to derive the local asymptotics in the two models via the embedding of the fitness models $\mathcal{T}_n$ into the associated multi-type branching processes as described in the beginning of Section 6.4.2. To see this, note that Lemma 75 gives us the formulaes for $\hat{\mu}$. Note that in both the models $\pi = \nu$ and $\hat{\mu}(s, dr; \alpha) = \nu(dr) \cdot h(s)$. Simple algebraic manipulations using the above theorem, now gives us Theorems 28 (b) and Theorems

30 (b) regarding the asymptotics for the empirical distribution of fitness of nodes and fitness of the mothers of the nodes.

To get asymptotics for the degree of nodes, note that by the above theorem, letting $X_k(n)$ denote the empirical distribution of nodes with degree k in $\mathcal{T}_n$, we have:

$$X_k(n) \longrightarrow_P \mathbb{P}\left(\deg(\rho_{\mathcal{F}^*}) = k\right)$$

where $\rho_{\mathcal{F}^*}$ denotes the root of the tree $\mathcal{F}^*$.

Now to compute the quantity on the right note by the joint distribution law specified in Theorem 81, we have

$$\mathbb{P}\left(deg(\rho_{\mathcal{F}^*}) = k\right) = \int_0^M \mathbb{P}\left(\deg(\rho_{\mathcal{F}_s(T_\alpha)}) = k\right) \nu(ds)$$

where $\mathcal{F}_s(T_\alpha)$ is the associated continuous time branching process where the initial ancestor has fitness s. Now note that conditional on the fitness, the degree of the root at time T_α has distribution $\mathcal{C}_s(T_\alpha)$ where $\mathcal{C}_s$ are the point process of offspring described by equations (29) and (30). Thus

$$\mathbb{P}\left(\deg(\rho_{\mathcal{F}_s(T_\alpha)}) = k\right) = \mathbb{P}(\mathcal{C}_s(T_\alpha) = k)$$

where T_α is independent of $\mathcal{C}_s$. Given the rates description of the point process $\mathcal{C}_s$, Equations (29) and (29), via computations similar Corollary 50, it is easy to compute that for the additive fitness model:

$$\mathbb{P}(\mathcal{C}_s(T_\alpha) = k) = \frac{\alpha}{k+1+s+\alpha} \prod_1^k \left(\frac{i+s}{i+s+\alpha}\right)$$

while for the multiplicative fitness model

$$\mathbb{P}(\mathcal{C}_s(T_\alpha) = k) = \frac{\alpha}{\alpha + s\cdot(k+1)} \prod_1^k \left(\frac{s\cdot i}{\alpha + s\cdot i}\right)$$

This computation gives us required result for asymptotics regarding the degree distribution, namely Theorem 28 (a) and Theorem 30 (a) respectively.

■

6.4.4 Global properties

We now state the main theorem required for performing global computations in our setup. It is essentially the same as Theorem 2 in [20], however reformulated to suit our needs.

Theorem 82 *Assume ν has bounded support and the corresponding (additive or multiplicative fitness) multi-type branching process model $(N(t), \mathcal{F}(t))$, driven by ν, satisfies all the Assumptions of 73. Then there exits an almost surely positive random variable W such that*

$$\frac{N(t)}{e^{\alpha t}} \longrightarrow_{L^1} W$$

as $t \to \infty$. Here α is the associated Malthusian rate of growth parameter computed in Lemma 77 for the two models.

In particular we have convergence in means as well namely there exits a constant C such that

$$\frac{E[N(t)]}{e^{\alpha t}} \longrightarrow C \tag{37}$$

as $t \to \infty$.

Let the stopping times $T_n := \inf\{t : N(t) = n\}$. Then analogous to Proposition 58 we have the following:

Proposition 83 *The stopping times T_n satisfy the following properties*

$$\frac{n}{e^{\alpha \cdot T_n}} \longrightarrow_P W$$

In particular this implies the following relations:

$$\frac{n^{1/\alpha}}{e^{T_n}} \longrightarrow_P W^{1/\alpha}, \qquad \frac{1}{\alpha}\log n - T_n \longrightarrow_P \frac{1}{\alpha}\log W$$

Thus

$$\frac{T_n}{\frac{1}{\alpha}\log n} \longrightarrow_P 1 \tag{38}$$

This is essentially enough to prove the various global properties.

Analysis of the root degree:

Write $\mathcal{P}_0(t)$ for the degree of the root at time t. Note that for both the additive and and multiplicative models the size of the degree grows as a Yule process. More precisely:

Lemma 84 *We have the following asymptotics for the root degree in the two fitness models:*
(a) For the additive model we have

$$\frac{\mathcal{P}_0(t)}{e^t} \longrightarrow_{a.s.} \tilde{W}$$

(b) For the multiplicative model we have

$$\frac{\mathcal{P}_0(t)}{e^{f_\rho \cdot t}} \longrightarrow_{a.s.} \tilde{W}$$

Here f_ρ is the fitness of the root in the multiplicative model and $\tilde{W}$ is an exponential rate one random variable.

Note that by definition of the stopping times and the tree $\mathcal{T}_n$, we have:

$$\rho(T_n) =_d \deg(\rho) \tag{39}$$

Thus combining Proposition 83, Lemma 84 and Equation (39) we have, for the additive model

$$\frac{\deg(\rho)}{n^{\frac{1}{\alpha}}} \longrightarrow_d \frac{\tilde{W}}{W^{\frac{1}{\alpha}}}$$

while for the multiplicative model we have

$$\frac{\deg(\rho)}{n^{\frac{f_\rho}{\alpha}}} \longrightarrow_d \frac{\tilde{W}}{W^{\frac{1}{\alpha}}}$$

■

Analysis of the maximal degree:

Here we analyze the two models separately.

Additive model: We start by noting that for a fixed B

$$\mathbb{P}\left(\frac{\max(deg)}{n^{\frac{1}{\gamma_{+,\nu}}}} > K\right) \leq \mathbb{P}\left(T_n > \frac{1}{\gamma_{+,\nu}}\log n + B\right) + \mathbb{P}\left(\max(t_n) > Kn^{\frac{1}{\gamma_{+,\nu}}}\right)$$

where for simplicity we have written $t_n = \frac{1}{\gamma_{+,\nu}}\log n + B$ and $\max(t_n)$ is the maximum offspring size (namely degree) of all nodes born before time t_n. By Proposition 83 given any $\varepsilon > 0$, we can choose $B_\varepsilon > 0$ a constant large such that ,

$$\limsup_{n\to\infty} \mathbb{P}\left(T_n > \frac{1}{\gamma_{+,\nu}}\log n + B\right) \leq \frac{\varepsilon}{2}$$

Thus it is enough to prove that given $\varepsilon > 0$ and the fixed constant B_ε, as above we can choose $K_\varepsilon > 0$ so that

$$\limsup_{n\to\infty} \mathbb{P}\left(\max(t_n) > K_\varepsilon n^{\frac{1}{\gamma_{+,\nu}}}\right) \leq \frac{\varepsilon}{2}$$

Since $e^{t_n} = Bn^{\frac{1}{\gamma_{+,\nu}}}$ it is enough to show the following

Proposition 85 *Let t_n be a fixed sequence $\uparrow \infty$. Then given any $\varepsilon > 0$ we can choose $K_\varepsilon < \infty$ so that*

$$\limsup_{n\to\infty} \mathbb{P}\left(\frac{\max(t_n)}{e^{t_n}} > K_\varepsilon\right) < \varepsilon$$

Proof: Dealing with nodes of different additive fitnesses turns out to be a nuisance. Via a simple coupling argument, note that it is enough to prove the above, for the case where all the nodes have additive fitness level M. Then note that we now have a linear preferential attachment model (see Section 2.1) with parameter M, namely the associated continuous time branching process has offspring distribution, the Markov point process $\mathcal{P}_M$ with rates

$$\mathbb{P}(\mathcal{P}_M(t+dt) - \mathcal{P}_M(t) = 1 | \mathcal{P}_M(t) = k) = (k+1+M)dt + o(dt) \tag{40}$$

For a fixed time $t_0 > 0$ note that

$$\mathbb{P}\left(\frac{\max(t_n)}{e^{t_n}} > K_\varepsilon\right) = \mathbb{P}\left(\mathcal{A}_{[0,t_0]}\right) + \mathbb{P}\left(\mathcal{A}_{(t_0,t_n]}\right)$$

where the events $\mathcal{A}_{[0,t_0]}$ is the event having the verbose description

> there exists a node born in the interval $[0, t_0]$ such that by time t_n, the number of offspring (direct descendants) it has is greater than Ke^{t_n}.

and $\mathcal{A}_{(t_0,t_n]}$ has an equivalent description for nodes born after time t_0. Thus it is enough to prove that the probabilities of the above two events are small for large n for K_ε chosen large enough.

The following simple Lemma plays a crucial role in what is to follow:

Lemma 86 *(a) Let $\mathcal{P}_M(.)$ be the offspring distribution point process in a linear preferential attachment model with parameter M as defined in Equation (40). Assume that the constant $K > M$. Then for any fixed times t and $0 \le t_0 < t$ we have*

$$\mathbb{P}(\mathcal{P}_M(t - t_0) > 2Ke^t) \le 2e^{-Ke^{t_0}}$$

where we have written $\mathcal{P}_M(t-t_0) = \mathcal{P}_M(0, t-t_0)$ for the number of offspring born in the interval $(0, t - t_0)$.

(b) Let $(N(t), \mathcal{F}(t))$ be the branching process with offspring distribution $\mathcal{P}_M$. Then there exists a constant C such that, for any fixed t and fixed constants A:

$$\mathbb{P}\left(N(t) > ACt^2e^{\gamma_M t}\right) \le \frac{1}{At^2}$$

where $\gamma_M = \frac{M+2+\sqrt{M^2+2}}{2}$ is the Malthusian rate of growth parameter associated with the process $N(t)$.

Proof: Part (a) follows from the fact that we can write

$$\mathcal{P}_M(.) = Y(.) + Poi_M(.)$$

where Y, Poi_M are independent rate one Yule and Rate M Poisson processes respectively. Now see the proof of Lemma 62.

(b) To prove (b) note that by Equation (37), there exists a constant C so that $E(N(t)) \le Ce^{\gamma_M t}$. Now use Markov's inequality.

■

To complete the proof choose K and t_0 integers, large so that the following conditions are satisfied:

(a) $\frac{C}{K}\sum_{m\ge t_0}\frac{1}{m^2} < \frac{\varepsilon}{10}$ where C is the constant in Lemma 86 (b) .
(b) $\sum_{m\ge t_0} Km^2e^{\gamma_M(m+1)}e^{-Ke^m} < \frac{\varepsilon}{10}$.
(c) $t_0^2e^{\gamma_M t_0}e^{-K} < \frac{\varepsilon}{10}$.

Then note that by the union bound

$$\mathbb{P}\left[\mathcal{A}_{[0,t_0]}\right] \le \mathbb{P}\left[N(t_0) > Kt_0^2e^{\gamma_M t_0}\right] + Kt_0^2e^{\gamma_M t_0}\cdot\mathbb{P}\left[\mathcal{P}(t_n) > Ke^{t_n}\right]$$

By choice of constants in condition (c) above and Lemma 86(a) we have

$$Kt_0^2e^{\gamma_M t_0}\cdot\mathbb{P}\left[\mathcal{P}(t_n) > Ke^{t_n}\right] < Kt_0^2e^{\gamma_M t_0}\cdot e^{-K} < \frac{\varepsilon}{10}$$

Again by Lemma 86(b) we have

$$\mathbb{P}\left[N(t_0) > Kt_0^2e^{\gamma_M t_0}\right] < \frac{C}{Kt_0^2} < \frac{\varepsilon}{10}$$

so that

$$\mathbb{P}\left[\mathcal{A}_{[0,t_0]}\right] \leq \frac{\varepsilon}{5}$$

To finish the proof, for any integer m with $m+1 \leq t_n$, write $\mathcal{A}_{[m,m+1]}$ as the event

> There exists a node born in the time interval $[m, m+1]$ so that by time t_n it has degree $> Ke^{t_n}$.

Note that

$$\mathbb{P}(\mathcal{A}_{(t_0,t_n]}) = \sum_{t_0+1}^{t_n} P(\mathcal{A}_{[m-1,m]})$$

Finally note that

$$\begin{aligned}\sum_{t_0+1}^{t_n} P(\mathcal{A}_{[m-1,m]}) &\leq \sum_{t_0+1}^{t_n} P(N[m-1,m] > K(m+1)^2 e^{\gamma_M(m+1)}) \\ &\quad + \sum_{t_0+1}^{t_n} K(m+1)^2 e^{\gamma_M(m+1)} P(\mathcal{P}(t_n - m) > Ke^{t_n}) \\ &\leq \frac{\varepsilon}{10} + \frac{\varepsilon}{10}\end{aligned}$$

by the conditions on the constants (a),(b) and (c).

■

Multiplicative model The proof of the upper bound is almost identical to the above proof for the additive model so we shall skip it. The lower bound states that with high probability, there exists a node with degree $> Ke^{\frac{M-\varepsilon}{\alpha_{\nu,\times}} \log n}$. To prove this first note that again by Propn 83 we can choose a constant B large enough so that

$$\mathbb{P}\left(T_n > \frac{1}{\alpha_{\nu,\times}} \log n - B\right) > 1 - \frac{\varepsilon}{4}$$

For simplicity, write $t_n = \frac{1}{\alpha_{\nu,\times}} \log n - B$. Then given any $\varepsilon > 0$, it is enough to prove that there exists a constant K such that by time t_n, there exists a node with degree $> Ke^{(M-\varepsilon)t_n}$ with probability $1 - \varepsilon/2$ for n large enough. Define

$$p_\varepsilon = \nu(M - \varepsilon, M] > 0$$

where ν is the fitness distribution. Since the offspring from the root come at the times of a Yule process of rate f_ρ, we can choose t_0 large so that

$$\mathbb{P}(\rho \text{ has an offspring in interval } [0, t_0] \text{ with fitness } > M - \varepsilon) > 1 - \frac{\varepsilon}{10}$$

Finally we can choose a constant K' so that for a Yule process $\mathcal{P}_{M-\varepsilon}$, growing at rate $M - \varepsilon$

$$\mathbb{P}\left(\mathcal{P}_{M-\varepsilon}(0, t_n - t_0) > K' e^{(M-\varepsilon)[t_n - t_0]}\right) > 1 - \frac{\varepsilon}{10}$$

The proof is concluded by writing $K^1_\varepsilon = \frac{K'}{e^{Mt_0}}$ and noting that the event that there exists a vertex with degree $Ke^{(M-\varepsilon)\cdot t_n}$ is contained in the event that a node of fitness level in $(M - \varepsilon, M]$ is born before time t_0 and by time t_n this node has more than $K' e^{(M-\varepsilon)[t_n - t_0]}$ offspring.

Analysis of the height: This is the same as for all the previous models namely, coupling Proposition 83 and Kingman's result, Theorem 66.

Proof of Corollary 32: Here we analyze the special case where the fitness distribution is the uniform measure on $[0,1]$. By Theorem 30, the Malthusian rate of growth parameter in this particular case is given by the unique positive solution of the equation

$$\int_0^1 \frac{s}{\alpha - s} ds = 1$$

i.e.

$$\int_0^1 \left[\frac{\alpha}{\alpha - s} - 1\right] ds = 1$$

Thus α is the unique solution of the equation

$$1 - \frac{1}{\alpha} = e^{-2/\alpha}$$

To show that this gives rise to a degree distribution with $p_k \sim \frac{1}{k^{1+\alpha} \log k}$, note that Theorem 30 implies that

$$f_k = \sum_{j \geq k} p_j = \int_0^1 \prod_1^k \left(\frac{x \cdot i}{\alpha + x \cdot i}\right) dx$$

It is thus enough to prove that $f_k \sim \frac{1}{k^\alpha \log k}$. Note that for each fixed i the factor $\left(\frac{x \cdot i}{\alpha + x \cdot i}\right)$ is an increasing function of x. Thus fix some $1 > \varepsilon > 0$, small. It is enough to prove that

$$f_k^\varepsilon = \int_\varepsilon^1 \prod_1^k \left(\frac{x \cdot i}{\alpha + x \cdot i}\right) dx \sim \frac{1}{k^{1+\alpha} \log k}$$

Let $i_\varepsilon = \inf\{j \geq 1 : j \cdot \varepsilon > \alpha\}$. Note that for any $x > \varepsilon$

$$\prod_{i_\varepsilon}^k \left(\frac{x \cdot i}{\alpha + x \cdot i}\right) = e^{-\sum_{j=i_\varepsilon}^k \log\left(1 + \frac{\alpha}{x \cdot j}\right)}$$

Now use the Taylor expansion for the log function to conclude that for large k

$$\prod_{i_\varepsilon}^k \left(\frac{x \cdot i}{\alpha + x \cdot i}\right) \sim \alpha \frac{\log k}{x}$$

so that

$$f_k^\varepsilon \sim \int_\varepsilon^1 e^{-\alpha \log(k)/x} dx = \int_1^{1/\varepsilon} \frac{e^{-y \cdot (\alpha \log k)}}{y^2} dy$$

Finally we get the result by noting that

$$\varepsilon^2 \int_1^{1/\varepsilon} e^{-y \cdot \alpha \log k} dy \leq \int_1^{1/\varepsilon} \frac{e^{-y \cdot (\alpha \log k)}}{y^2} dy \leq \int_1^{1/\varepsilon} e^{-y \cdot \alpha \log k} dy$$

and

$$\int_1^{1/\varepsilon} e^{-y \cdot \alpha \log k} dy \sim \frac{1}{\alpha \log k} e^{-\alpha \log k} = \frac{1}{\alpha k^\alpha \log k}$$

■

Proof for the Maximal degree: This follows immediately from Theorem 31

6.5 Preferential attachment networks

In this Section we shall provide a complete analysis for the Preferential attachment networks $\mathcal{N}_n$. As before let V_n denote a node picked uniformly at random from $\mathcal{N}_n$. We start with a basic Lemma, which helps us perform computations, once we have managed to establish the convergence of the distribution of $\deg(V_n)$. The basic plan of this Section will be as follows. We shall first assume the distributional convergence given by Equation (5) and (7) and derive the exact probability mass functions of the Limiting random variables as well as the power law exponents of these random variables. We shall then go on to the proof of the main results, namely Theorems 36 and 40.

6.5.1 Limiting probability mass functions

Lemma 87 *Let $(N_i(.))_{1\leq i\leq m}$ be independent Yule process, and let T_β be an exponential random variable independent of the sequence N_i. Define*

$$Z := \sum_{1}^{m} (N_i(T_\beta) + 1)$$

Note that $Z \geq m$. Then for any $k > m$

$$\mathbb{P}(Z = k) = \beta \frac{\Gamma(k) \cdot \Gamma(\beta + m)}{\Gamma(m) \cdot \Gamma(k + \beta + 1)}$$

Proof: Note that by Lemma 46 on the exact distributional properties of the Yule process, we have, for any $0 \leq s < 1$

$$\mathbb{E}\left(s^Z | T_\beta = t\right) = \left[\frac{se^{-t}}{(1 - s(1 - e^{-t}))}\right]^m$$

where $\frac{se^{-t}}{(1-s(1-e^{-t}))}$ is the probability generating function of random variable $X + 1$ where X is a geometric(e^{-t}) random variable. Writing $\phi_Z(s)$ for the probability generating function of Z we get

$$\phi_Z(s) = \beta \int_0^\infty e^{-\beta t} \left[\frac{se^{-t}}{(1 - s(1 - e^{-t}))}\right]^m dt$$

Now use the fact that

$$\mathbb{P}(Z = k) = \frac{1}{k!} \frac{d^k}{ds^k} \phi_Z(s)\bigg|_{s=0}$$

to get the result.

■

Proof of the limiting p.m.f, Equation 6:

Now note that Lemma 87 coupled with Equation (5) gives us the appropriate probability mass function of Equation (6).

Proof of Corollary 37:

From Equation (6), since we now have n explicit form of the probability mass function, simple algebraic derivations now imply Corollary 37.

6.5.2 Local Analysis: Proofs of Theorems 36 and 40

We first start by giving the basic idea of the proof. We will assume the simplest case, where each new node is born with a fixed number m of out edges, and tries to attach these edges recursively to the existing tree. Throughout this Section we shall be focussed on only the BA-preferential attachment model, so whenever we mention "continuous time" branching process, we mean the branching process associated with this model, namely with the Yule process as the offspring distribution. Now note that Theorem 11 has the following interpretation: Fix a large time $t > 0$ and pick a node uniformly at random from the tree $\mathcal{F}(t)$. Then

(a) The age of the node, say T^*, has approximately, an exponential distribution with rate 2.

(b) The degree of this node has distribution $\mathcal{P}(T^*)$, where $\mathcal{P}$ is a Yule process, independent of T^*.

Now suppose we pick a random node and the $m-1$ *consecutive* nodes born immediately after the random node chosen. Due to the exponential growth of the branching process, all these nodes have the same value of age but are essentially born at different places in the tree, so that if we write Z as the combined degree of these m nodes then the distribution of Z is approximately $\sum_{i=1}^{m} \mathcal{P}_i(T^*)$ where $\mathcal{P}_i$ are independent Yule processes, independent of the common age T^* as well. This gives us Theorem 36.

For the case where the distribution of out-edges have heavy tails, this heuristic fails and we have to resort to other techniques. Consider the construction of the network from te branching process where we first run the branching process and then coalesce consecutive nodes. Call the "age" of a node v in the network, the time of birth of the last node in $\mathcal{F}(t)$ which was used to construct the node v. Then we prove that the age of a randomly picked node in the network namely T^*, is no longer approximately exponentially distributed, but because of the presence of nodes with large out-degree, the age of this node has the distribution:

$$T^* =_d Z_\alpha$$

where Z_α has the distribution specified by Equation (4). Thus the distribution of the degree of the random node V_n is given approximately by

$$Z = \sum_{1}^{W} \mathcal{P}_i(T^*)$$

and this gives us Theorem 40. Our proof is essentially a formalization of these ideas.

We start with an initial Lemma which shall play a crucial role later.

Lemma 88 *Let $W_i > 0$ be i.i.d integer valued random variables and define the partial sums $S_i := \sum_1^i W_j$. Let N be uniformly distributed on $\{1, 2, \ldots, n\}$ independent of the sequence W_i. Then*
(a) If $E(W_i) < \infty$ then

$$\frac{S_N}{S_n} \longrightarrow_d U$$

as $n \to \infty$.
(b) If $x^\alpha \mathbb{P}(W > x) \to C$ for some $0 < \alpha \leq 1$ and some $C > 0$. Then

$$\frac{S_N}{S_n} \longrightarrow_d U^{\frac{1}{\alpha}} \cdot \frac{X_U}{X_1}$$

as$n \to \infty$, where $U \sim U(0,1)$ and $(X_t)_{t>0}$ is the associated Levy process with Levy measure C/x^α on $\mathbb{R}^+$.

Proof: We shall use the basic fact that

$$\frac{N}{n} \to_L U \tag{41}$$

Proof of (a) follows immediately from Equation (41) and the strong law of large number by writing

$$\frac{S_N}{S_n} = \frac{S_N/N}{S_n/n} \times \frac{N}{n}$$

Proof of (b) follows using Equation (41) and the fact that the process

$$\left(\frac{S_{[it]}}{n^{\frac{1}{\alpha}}}\right)_{0\le t\le 1} \longrightarrow_L (X_t)_{0\le t\le 1}$$

as $n \to \infty$.

■

We finally come to the main ingredients of the actual proof. Note that each node $v \in \{1,2,\ldots,n\}$ arrives with W_v edges where $W_v \sim_{iid} \nu$ and which are then connected to nodes already present in the network. We shall think of these edges as directed **away** from the node and directed towards the node they connect to, and shall say that node v has $\texttt{out}-\texttt{deg} = W_v$. Let V_n be a node picked uniformly at random from the network $\mathcal{N}_n$. Fix $m > 0$. Since we want to prove that

$$\deg(V_n) \longrightarrow_d \sum_1^W N_i(Z)$$

where Z is either T_2 or Z_α and $W \sim \nu$ independent of Z and the Yule processes N_i. Thus it is clearly enough to show for each fixed m, conditionally:

$$\mathcal{L}\left(\deg(V_n)|\texttt{out}-\texttt{deg}(V_n) = m\right) \longrightarrow_L \sum_1^m N_i(Z)$$

Thus fix m. All computations will be done conditional on $W_{V_n} = m$. We first start with an equivalent construction of the network $\mathcal{N}_n$ (conditional on the degree of the random node V_n, $W_{V_n} = m$).

Construction 89 (Network $\mathcal{N}_n$) *Consider the following construction:*

1. *Use the continuous time branching process construction to generate the BA-preferential attachment tree. Label the nodes in the order they are born with the root labeled as 1. Let the process be denoted as $(N(t), \mathcal{F}(t))_{t>0}$.*

2. *Fix $n > 0$. Then to generate $\mathcal{N}_n$ generate the following random varibles all of which are independent of each other and independent of the process $\mathcal{F}(t)$.*
   (a) Generate $N \sim Unif\{1,2,\ldots,n\}$.
   (b) Generate $W_1, W_2, W_{N-1}, W_{N+1}, \ldots W_n \sim_{iid} \nu$. Define $W_N := m$.
   (c) Define the partial sum sequence $S_i = \sum_1^i W_j$ with $S_0 = 0$.

3. *Define the sequence of stopping times* $T_{S_i} = \inf\{t > 0 : |\mathcal{F}(t)| = S_i\}$. *Consider the tree* $\mathcal{F}(T_{S_n})$. *For* $i > 0$ *coalesce the nodes* $S_{i-1}+1, \dots S_i$. *Call the resulting node, the* i^{th} **super-node** *and say that the* i^{th} *super-node has been fully born at time* T_{S_i}.

4. *Since we are interested in the degree of the* N^{th} *super-node call this* **ego**. *Note that by construction* **ego** *is made up of* m *normal nodes to be henceforth denoted as*

$$\mathbf{ego}(1), \mathbf{ego}(2), \dots, \mathbf{ego}(m)$$

labeled in the order of the time of birth, namely **ego** *(1) is the oldest node,* **ego** *(2) is the next oldest and so on.*

Lemma 90 *The resultant network constructed above, consisting of all the* **super-nodes** *is the same as* $\mathcal{N}_n$, *conditional on the event that the random node* N *node picked has* $\texttt{out} - \texttt{degree} = m$.

Proof: Obvious.

Some notation: Note that **ego** is made of the nodes $S_{N-1}+1, \dots S_{N-1}+m = S_N$ nodes to be born into the population. Also note that **ego** is **fully born** at time T_{S_N}. Write

$$L_n = T_{S_n} - T_{S_N} \tag{42}$$

for the **life span** of **ego** namely, the amount of time that elapses between the *full* birth of **ego** and the time for the entire network to be constructed namely T_{S_n}.

The following Lemma is essentially the heart of the whole argument. We will deal with slightly more complicated versions of this Lemma later but they tend to obfuscate the following Lemma in technical details and are essentially minor variants of the following Lemma.

Lemma 91 *As* $n \to \infty$ *we have*

$$L_n \longrightarrow_d Z$$

where Z *is a generic random variable, used to denote* T_2 *if* $E(W) < \infty$ *and* Z_α *otherwise.*

From here on end we shall use μ_Z to denote the distribution of Z.

Proof: The proof is almost obvious, and essentially combines the "exponential growth " of the branching process $\mathcal{F}(.)$ and the Law of large numbers result of Lemma 88. We want to show that

$$\mathbb{P}(L_n \geq x) \longrightarrow \mu_Z(x, \infty)$$

as $n \to \infty$.

Note that

$$\mathbb{P}(L_n > x) \quad = \quad \mathbb{P}\left(e^{2\cdot(T_{S_n} - T_{S_N})} > e^{2x}\right)$$

The above equation can be rewritten as

$$e^{2\cdot(T_{S_n} - T_{S_N})} = \frac{e^{2T_{S_n}}/S_n}{e^{T_{S_N}}/S_N} \times \frac{S_n}{S_N}$$

By the exponential rate of growth of the branching process, Theorem 57, there exists a single random variable $W > 0$ such that $e^{2T_{S_n}}/S_n \to W$ and $e^{T_{S_N}}/S_N \to W$ almost surely. The result then follows, since by Lemma 88 we have

$$e^{2\cdot(T_{S_n}-T_{S_N})} \longrightarrow_d \frac{1}{U}$$

if $E(W) < \infty$ and

$$e^{2\cdot(T_{S_n}-T_{S_N})} \longrightarrow_d \frac{X_1}{U^{\frac{1}{\alpha}} \cdot X_U}$$

where X_t is the associated Levy process.

■

We collect without proof some more basic properties of the branching process $\mathcal{F}(t)$.

Lemma 92 *(a) For the Markov process $\mathcal{F}(t)$, conditional on $\mathcal{F}(t)$, the rate of growth of $\mathcal{F}(t)$ depends only on the size of $|\mathcal{F}(t)|$, irrespective of the actual graph-theoretic structure of $\mathcal{F}(t)$ and is equal to $2|\mathcal{F}(t)| - 1$. More precisely*

$$\mathbb{P}\left(|\mathcal{F}(t,t+dt)| - |\mathcal{F}(t)| = 1|\, \mathcal{F}(t)\right) = (2 \cdot |\mathcal{F}(t)| - 1)dt + o(dt)$$

(b) Given any fixed leaf v in $\mathcal{F}(t)$, the probability that the next birth is a daughter of v is

$$1/(2 \cdot |\mathcal{F}(t)| - 1)$$

Remark: Thus Lemma was crucially used in Pittel's [30] beautiful paper, analyzing the height of the linear preferential attachment model.

Write the set G_n (G being a mnemonic for good) for the event

> none of the m nodes in $\mathcal{F}_{T_{S_n}}$ which make up the super node **ego** , share a mother-daughter relationship.

More precisely, there does not exist exist $1 \leq i < j \leq m$ such that $\mathbf{ego}(i)$ is a mother of $\mathbf{ego}(j)$. Write

$$A_k = \{\mathbf{ego}(k) \text{ is not a child of } \mathbf{ego}(i) \text{ for } i < k\}$$

so that we have the relation:

$$G_n = \cap_1^m A_i$$

By Lemma 92

$$\mathbb{P}(A_k|\mathcal{F}(T_{S_N}), \cap_1^{k-1} A_i) = \left(1 - \frac{k-1}{2 \cdot (S_{N-1} + k - 1) - 1}\right)$$

Taking expectations, we have the following fact:

Lemma 93 *For fixed n*

$$\mathbb{P}(G_n) = \mathbb{E}\left[\prod_1^m \left(1 - \frac{k-1}{2 \cdot (S_{N-1} + k - 1) - 1}\right)\right]$$

Since we have $S_N \to \infty$ a.s. this implies

$$\mathbb{P}(G_n) \to 1 \qquad \textit{as } n \to \infty$$

Thus Lemma 93 implies that with high probability:

(a) The nodes constituting the super-node **ego** are essentially born at different disjoint portions of the tree, and

(b) Till the last node **ego** (m) is born, none of the other nodes $\mathbf{ego}(1), \ldots, \mathbf{ego}(m-1)$ do not reproduce.

Thus by conditions (a) and (b), for a fixed time $t > 0$, write $\mathcal{A}(t)$ as the random forest consisting of the nodes $(\mathbf{ego}(i))_{1 \leq i \leq m}$ and its children observed at time $T_{S_N} + t$, namely t time units after the birth of all the nodes that make up **ego** . By the Markov property of the growing branching process $\mathcal{F}(\cdot)$, on the set G_n

$$\mathcal{A}(t) =_d \cup_1^m \mathcal{F}^i(t) \qquad \text{on } G_n \text{ conditional on } \mathcal{F}(T_{S_N}) \tag{43}$$

where $\mathcal{F}^i$ are independent branching process driven by the Yule point process and independent of $\mathcal{F}(T_{S_N})$ and with $\mathbf{ego}(i)$ as the root of $\mathcal{F}^i$.

In the above setup, for $t > 0$ write $\deg_i(t)$as the number of offspring of $\mathbf{ego}(i)$ at time t (after T_{S_N}). By the definition of the branching process and Equation (43), on the set G_n the random variable $\deg_i(t)$ has the same distribution as a Yule process upto time t. Note also that on the set G_n the degree of **ego** in $\mathcal{N}_n$, $\deg_{\mathcal{N}_n}(\mathbf{ego})$ is the same as

$$\deg_{\mathcal{N}_n}(\mathbf{ego}) = \sum_1^m \deg_i(L_n)$$

where L_n is the lifetime of the node **ego** , as defined in Equation (42).

To prove our result, it is enough to show that L_n and the process $\mathcal{A}(.)$ are "asymptotically independent". Define the space $\mathcal{H}_m$ as the space of all finite graphs with m distinguished vertices, labelled by the integers $\{1, 2, \ldots, m\}$. For a fixed time $t > 0$, think of $\mathcal{A}(t)$ as a random graph on $\mathcal{H}_m$. Then it is enough to prove, for any set $B \in \mathcal{H}_m$, and any fixed $t > 0$

$$\mathbb{P}\left(\mathcal{A}(t) \in B, L_N > t\right) \longrightarrow \mathbb{P}\left(\cup_1^m \mathcal{F}^i(t)\right) \cdot \mu_Z(t, \infty) \tag{44}$$

as $n \to \infty$, where Z is the generic random variable used to denote T_2 or Z_α in the two cases. This proves Theorems 36 and 40, since Equation (44) in particular implies

$$\deg_{\mathcal{N}_n}(\mathbf{ego}) = \sum_1^m \deg(i)(L_n) \longrightarrow_L \sum_1^m \deg(\mathcal{F}^i)(Z)$$

where Z is independent of $(\mathcal{F}^i)_{1 \leq i \leq m}$ and $\mathcal{F}^i$ are independent Yule processes.

Define $\texttt{For}_m \subset \mathcal{H}_m$, to be the subset of $\mathcal{H}_m$, consisting of graphs with m distinguished nodes where the graphs consist of m disjoint trees one tree corresponding to, and rooted at, each of the distinguished vertices. Note that Lemma 93 implies

$$\mathbb{P}(\mathcal{A}(t) \in \texttt{For}_m) \longrightarrow 1$$

as $n \to \infty$.

Fix any such forest say $\mathcal{E} \in \texttt{For}_m$. Let the number of nodes in the forest, $|\mathcal{E}| = k \geq m$. Then it is enough to prove that

$$\mathbb{P}(\mathcal{A}(t) = \mathcal{E}, L_n > t) \longrightarrow \mathbb{P}(\cup_1^m \mathcal{F}^i(t) = \mathcal{E}) \cdot \mu_Z(t, \infty)$$

By Lemma 93 we have

$$\mathbb{P}(\mathcal{A}(t) = \mathcal{E}, L_n > t) = \mathbb{P}(\{\mathcal{A}(t) = \mathcal{E}, L_n > t\} \cap G_n) + o(1)$$

as $n \to \infty$.

For $t > 0$, define the process $\tilde{\mathcal{F}}(t) := \mathcal{F}(T_{S_N} + t) \setminus \mathcal{A}(t)$, namely the process excluding the lineages of **ego**$(1), \ldots$ **ego**(m) seen t units of time after the birth of **ego** . By the Strong Markov property for the Markov process for the full branching process $\mathcal{F}(t)$, conditioning on $\mathcal{F}(T_{S_N})$, and using Equation (43), on the set G_n

$$\mathbb{P}\left(\{\mathcal{A}(t) = \mathcal{E}, L_n > t\} \middle| \mathcal{F}(T_{S_N})\right) = \mathbb{P}\left(\cup_1^m \mathcal{F}^i(t) = \mathcal{E}\right) \cdot \mathbb{P}\left(\tilde{\mathcal{F}}(t) < S_n - k | \mathcal{F}(T_{S_N})\right) \tag{45}$$

Now take expectations and note that by Lemma 92

$$L_n = T_{S_n} - T_{S_N} =_d \sum_{S_N}^{S_n} \frac{Y_i}{2i-1}$$

where $Y_i \sim_{iid} Exp(1)$ independent of all of the other variables. By Lemma 91

$$L_n := \sum_{S_N}^{S_n} \frac{Y_i}{2i-1} \longrightarrow_d Z$$

as $n \to \infty$. For any fixed k and m, this also implies that

$$L_n^k := \sum_{(S_N - m)}^{(S_n - k)} \frac{Y_i}{2i-1} \longrightarrow_d Z$$

Taking expectations in Equation (45) implies:

$$\mathbb{P}\left(\{\mathcal{A}(t) = \mathcal{E}, L_n > t\} \cap G_n\right) = \mathbb{P}\left(\cup_1^m \mathcal{F}^i(t) = \mathcal{E}\right) \cdot \mathbb{E}\left(C_n \cdot D_n\right)$$

where $C_n = 1\{G_n\}$ and $D_n = 1\{\tilde{\mathcal{F}}(t) < S_n - k\} = 1\{\tilde{T}_{S_n - k} > t\}$ and the random $T_{S_n - k}$ is the stopping time defined as

$$\tilde{T}_{S_n - k} := \inf\{s : \tilde{F}(s) = S_n - k\}$$

Note that on G_n, the distribution of $\tilde{T}_{S_n - k} =_d L_n^k$. Now use the fact that

$$C_n \to_P 1 \text{ and } D_n \to_L 1\{Z > t\}$$

to get that

$$\mathbb{P}\left(\{\mathcal{A}(t) = \mathcal{E}, L_n > t\} \cap G_n\right) \longrightarrow \mathbb{P}\left(\cup_1^m \mathcal{F}^i(t) = \mathcal{E}\right) \cdot \mu_Z(t, \infty)$$

■

Proof of Theorem 38

The proof of this theorem follows from a careful reformulation of the Azuma-Hoeffding inequality. The basic idea is very easy but to give a full rigorous proof takes some notation.

Note that the Theorem 36 can easily be modified, via Construction 89, where instead of dealing with a random sequence W_i, we deal with a deterministic sequence $(w_1, \ldots, w_n)$ satisfying some regularity conditions. Although we shall not fact, it is essentially obvious because we only use the Law of large numbers in Lemma 91, in the case where the random variables have finite mean.

Thus we first prove the following result which is a reformulation of Theorem 36 and deals with purely deterministic degree sequence.

Proposition 94 *Let* $(w_1, w_2, \ldots, w_n)$ *be a purely deterministic sequence of strictly positive integers satisfying*
(A) $\frac{1}{n}w_i \longrightarrow \mu$ *for some* $\mu > 0$.
(B) $\max_{1\leq i\leq n} \frac{w_i}{n} \to 0$ *as* $n \to \infty$.
(C)Let F_n *is the empirical distribution of the weight sequence* $(w_1, w_2, \ldots w_n)$*and* ν *be a distribution on integers. Then*

$$F_n \Rightarrow \nu$$

where the above denotes convergence in law.

Consider the previous construction of $\mathcal{N}_n$*, via running the Branching process first, and then coalescing nodes. Define* $L_k(n)$ *to be the number of degree* k *nodes in* $\mathcal{N}_n$.

Then

$$p_k^n = \frac{E(L_k(n))}{n} \longrightarrow p_k$$

as $n \to \infty$. *Here* p_k *is the asymptotic expected degree sequence when vertex* i *appears with* W_i *edges with* $W_i \sim_{iid} \nu$. *More precisely the sequence* p_k *denotes the probability mass function of the random variable* Z*, where* Z *has the distribution:*

$$Z =_d \sum_1^W \mathcal{P}_i(T_2)$$

with $W \sim \nu$

Proof: Go back to the proof of the Theorem and see that we have used only the strong law of large numbers and the condition (B).

■

To prove the main Theorem, the following is enough:

Proposition 95 *Consider a degree sequence* $(w_1, w_2, \ldots, w_n)$ *which satisfy the three stipulated conditions (A), (B) ,(C) of Proposition 94, and the added condition:*
(D) There exists a sequence $w(n) \uparrow \infty$ *such that*

$$\frac{w(n)}{n^2} \sum_1^n w_i^2 \to 0$$

Then for any $\delta > 0$,

$$\mathbb{P}\left(|\frac{N_k(n)}{n} - p_k| > \delta \right) \longrightarrow 0$$

as $n \to \infty$*, where* p_k *is the expected asymptotic degree sequence of Proposition 94.*

Proof: In the light of Proposition 94, since we have convergence for the expected values, note that it is enough to prove that:

$$\mathbb{P}\left(|\frac{L_k(n)}{n} - \mathbb{E}(\frac{L_k(n)}{n})| > \delta \right) \longrightarrow 0$$

as $n \to \infty$.

For the increasing sequence of graphs as $\mathcal{N}_m$ and consider the Martingale sequence

$$Y_i = \mathbb{E}(L_k(n)|\mathcal{N}_i) - \mathbb{E}(L_k(n)|\mathcal{N}_{i-1})$$

Note that $|Y_i| \leq 2w_i$

Now use Azume Hoeffding's inequality (see e.g. [32]) to conclude that:

$$\mathbb{P}\left(|\frac{L_k(n)}{n} - \mathbb{E}(\frac{L_k(n)}{n}| > \frac{\delta}{\sqrt{w(n)}}\right) \leq 2\exp\left(-\frac{\delta^2 n^2}{8w(n)\sum_1^n w_i^2}\right)$$

The result now follows from condition (D).

■

Proof of Corollary 39: The fact that W_i satisfies the conditions (A), (B) and (C) is obvious by the strong law of large numbers and the fact that W has a $1+\varepsilon$ moment. Condition (D) is also satisfied, since Theorem 1 in Section 22 of [17] easily implies that if W has a $1+\varepsilon$ moment then

$$\frac{w(n)}{n^2}\sum_1^n W_i^2 \longrightarrow_P 0$$

for $W_i \sim_{iid} \nu$ and the sequence $w_n = \log n$.

6.5.3 Global Analysis: Maximal degree

Here we shall provide complete proofs of Theorem 41 and Theorem 42. This essentially combines two of our previous ideas namely:

(a) The time required to observe the branching process before one can finish constructing the network, namely T_{S_n} is equal to $\frac{1}{2}\log n \overset{+}{-} O_P(1)$

(b) The maximal degree of the network essentially occurs at a finite distance from the root of the tree.

However making this precise requires quite a bit of work. They key ideas are as stated above. First note that for the bounded out-edge setup namely in model (A) where $\mathbb{P}(W \leq m) = 1$ for some integer m, the result is obvious, since by Construction 4, we have

$$\max(\deg(\mathcal{N}_n)) \leq_{st} m \cdot \max(\deg(\mathcal{T}_{m\cdot n}))$$

where $\mathcal{T}_{n\cdot m}$ is the Preferential attachment tree on $n \cdot m$ vertices. Further by Theorem 14, we have already proved that for large n, the maximal degree in $\mathcal{T}_{n\cdot m}$ is of order $n^{1/2}$.

Now consider the case where $W \sim \lfloor X \rfloor + 1$ where $X \sim Exp(1)$. We start with a simple Lemma regarding the extremal behavior of such a random variable:

Lemma 96 *Let W_i be iid and distributed as $\lfloor X \rfloor + 1$, where $X \sim Exp(1)$. Fix any $t > 0$. Then*

$$\mathbb{P}\left(\max_{1\leq i\leq e^t} W_i > 2t\right) \leq e^{-(t-1)}$$

Proof: This follows from the union bound and the fact that for an exponential random variable:

$$P(W_i > 2t) \leq P(X > 2t - 1) = e^{-(2t-1)}$$

Now we come to a stopping time Theorem which gives us asymptotic information on the amount of time we need to observe $\mathcal{F}(t)$ before we have enough data to construct $\mathcal{N}_n$.

Proposition 97 *Consider construction 4, where we first run a continuous time branching process $(N(t), \mathcal{F}(t))$ driven by a Yule offspring, then generate a sequence $W_i \sim_{iid} \nu$ having the above distribution and finally coalescing nodes $S_{i-1}+1, \ldots, S_i$ to form node i, where S_i was the partial sum sequence. Define the stopping time*

$$T_{S_n} = \inf\{t > 0 : N(t) = S_n\}$$

i.e. the random variable describing the amount of time needed to observe the branching process $\mathcal{F}(\cdot)$, to construct $\mathcal{N}_n$. Then there exists a random variable $0 < W < \infty$ such that

$$T_{S_n} - \frac{1}{2}\log n \longrightarrow_P \log W$$

Proof: Note that by Theorem 58,

$$\frac{S_n}{e^{2T_{S_n}}} \longrightarrow_{a.s} \tilde{W}$$

where $\{\tilde{W}\}$ has density $x \cdot e^{-x}$ on $\mathbb{R}^+$. Now note that by the Strong Law of Large numbers we have

$$\log S_n - \log n - \log \mu \longrightarrow_{a.s.} 0$$

where $\mu = E(X)$.

■.

Proof of the Lower bound in Theorem 41: This is easy since

$$\max(\deg(\mathcal{T}_n)) \leq_{st} \max(\deg(\mathcal{N}_n))$$

By Theorem 14(a), with high probability as $n \to \infty$, the root degree is of order $\Theta_P(\sqrt{n})$.

The upper bound requires more work. It involves the usual bracketing technique used in deriving the asymptotic bounds for the maximal degree in Theorem 14 and showing that the node with the maximum degree could not have been born at a time much after time 0. Here we have the added complication of making sure that we can ignore the effect of nodes born with a large out degree.

To simplify the proof, we shall define what is essentially a "moving average sequence" associated with the degree sequence. This moving average sequence is supposed to provide a caricature of the construction of the degree sequence, but is easier to analyze. Consider the following construction:

Construction 98 *Fix $B > 1$. Let $t_n = \frac{1}{2}\log n + B$*
(a) Consider $\mathcal{F}(t_n)$. Label each node in $\mathcal{F}(t_n)$ according to the order in which it was born, with the root labeled by the integer 1, the second node to be born is labeled by 2 and so on.
(b) Each node i born is also given a integer valued random variable W_i, independent with the distribution ν.

(c) For the time t_n defined above, let $\deg_i(t_n)$ be the degree of node i at time t_n. Consider the **moving average sequence** *defined as the sequence M_i with*

$$M_i = \sum_{j=i+1}^{i+W_i} \deg_j(t_n)$$

and $M_i = 0$, for $i > N(t_n) = |\mathcal{F}(t_n)|$.

Define the random variable

$$\mathcal{H}_n = \max_i M_i \tag{46}$$

We collect some properties of the above moving average sequence in the Lemma below. The Lemma basically says that to get asymptotics for the maximal degree in $\mathcal{N}_n$, it is enough to get asymptotics for the random variable $\mathcal{H}_n$.

Lemma 99 *Let $S_i = \sum_1^i W_j$ be the partial sum sequence and define the random variable M as,*

$$M := \sup\{j : S_j < |\mathcal{F}(t_n)|\}$$

Note that by using the construction 4, observing the branching process $\mathcal{F}(\cdot)$ up till time t_n, one can construct the BA-random preferential attachment network $\mathcal{N}_M$ on M nodes.

Then

$$\max(\deg(\mathcal{N}_M)) \leq_{st} \mathcal{H}_n$$

In particular,

$$\limsup_{n\to\infty} \mathbb{P}(S_n < t_n) \leq \frac{\varepsilon}{2}$$

implies, for any constant K

$$\limsup_{n\to\infty} \mathbb{P}\left(\max(\deg(\mathcal{N}_n)) \geq K \cdot \sqrt{n}\right) \leq \frac{\varepsilon}{2} + \limsup_{n\to\infty} \mathbb{P}\left(\mathcal{H}_n \geq K\sqrt{n}\right)$$

Proof: Observe that the subsequence of the moving average sequence:

$$(M_1, M_{S_1}, M_{S_3}, \ldots M_{S_{M-1}})$$

gives the degree sequence of $\mathcal{N}_M$.

■

Now note that by Prop 97 , given any $\varepsilon > 0$, we can choose a constant B large so that

$$\limsup_{n\to\infty} \mathbb{P}\left(N(\frac{1}{2}\log n + B) > S_n\right) > 1 - \frac{\varepsilon}{2} \tag{47}$$

Thus consider the previous construction for this B. Then it is obvious to complete the proof it is enough to prove:

Proposition 100 *Fix B so that Equation (47) is satisfied. Then given any $\varepsilon > 0$ we can find a constant $K_\varepsilon < \infty$ so that*

$$\limsup_{n\to\infty} \mathbb{P}\left(\mathcal{H}_n > K_\varepsilon\sqrt{n}\right) \leq \frac{\varepsilon}{2}$$

Proof:

We start with a basic Lemma which captures the behavior of both the number of nodes born in various intervals of time, as well as the relative size of the out-edges W_i of the node i.

Lemma 101 *For notational convenience, we shall without loss of generality assume that t_n is an integer. Let $m \leq t_n$ be any integer. Then*
(a) Define $N([m, m+1])$ as the number of births born in the time interval $[m, m+1]$ for the associated branching process $\mathcal{F}(\cdot)$. Then there exists a constant $C > 0$ such that:

$$\mathbb{P}\left(N(m, m+1) > e^{3(m+1)}\right) \leq C \cdot e^{-m}$$

(b)Define $J_m := \max\{W_j : j \text{ born in the interval } [m, m+1]\}$. Then

$$\mathbb{P}\left(J_m \geq 5 \cdot m\right) \leq 2Ce^{-m}$$

(c) Define the random variable $D_m := \max\{\deg_i(t_n) : \ i \text{ born in the time interval } [m, m+1]\}$. Then for any fixed constant $K > 0$,

$$\mathbb{P}\left(D_m > Ke^{t_n - \frac{m}{2}}\right) \leq Ce^{-m} + e^{3m} \cdot e^{-Ke^{\frac{m}{2}}}$$

(d) Finally define the random variable $\mathcal{H}_{n,m} := \max\{M_i : \ i \text{ is born in the interval } [m, m+1]\ \}$. Then for any $K > 0$

$$\mathbb{P}\left(\mathcal{H}_{n,m} > 5m \cdot Ke^{t_n - \frac{m}{2}}\right) \leq 3Ce^{-m} + e^{3m} \cdot e^{-Ke^{\frac{m}{2}}}$$

Proof: Part (a) follows from Theorem 60, since the number of individuals born before time m namely $N(m)$ has expectation $E(N(m)) \sim C \cdot e^{2m}$. Now use the Markov's inequality.

Part (b) follows from combining Part (a) with Lemma 96 on the extremal behavior of the distribution ν.

Part (c) follows from the union bound and Lemma 86 since

$$\mathbb{P}\left(D_m > Ke^{t_n - \frac{m}{2}}\right) \leq \mathbb{P}\left(N(m+1) > e^{m}\right) + \mathbb{P}\left(N(m+1) < e^{3m}, D_m > Ke^{t_n - \frac{m}{2}}\right)$$

Now note that for each of the nodes born in the interval $[m, m+1]$, the number of offspring this node has by time t_n is bounded stochastically by a Geometric random variable with parameter $e^{-(t_n - m)}$, (which we shall denote by $G(e^{-(t_n-m)})$), since the offspring distribution by the properties of the Yule process has a Geometric distribution. Thus by the union bound

$$\mathbb{P}\left(N(m+1) < e^{3m}, D_m > Ke^{t_n - \frac{m}{2}}\right) \leq e^{3m}\mathbb{P}\left(G(e^{-(t_n-m)}) > Ke^{t_n - \frac{m}{2}}\right)$$

Now use the large deviation properties of the Geometric distribution (see Lemma 62) to get the result.

Part (d) follows from combining (a), (b) and (c).

■.

To analyze the moving average sequence, define

$$\mathcal{H}_{n,\geq m} := \max\{M_i : \text{ i is born after time } m \}$$

and let $\mathcal{H}_{n,<m}$ have a similar interpretation for nodes born before time m. Note that for any choice of K and m,

$$\mathbb{P}\left(\mathcal{H}_n \geq Ke^{t_n}\right) \leq \mathbb{P}\left(\mathcal{H}_{n,\geq m} > Ke^{t_n}\right) + \mathbb{P}\left(\mathcal{H}_{n,<m} > Ke^{t_n}\right)$$

To complete the proof, fix a $K > 1$ and choose $m > 1$ large, so that the following two conditions are satisfied:

(a)

$$\sum_{j=m}^{\infty} 3Ce^{-m} + e^{-Ke^{m/2}+3m} \leq \frac{\varepsilon}{4}$$

and

(b)

$$5j \cdot e^{\frac{j}{2}} < 1 \qquad \text{for all } j \geq m$$

By Lemma 101 (d) we have for this choice of K and m

$$\mathbb{P}(\mathcal{H}_{n,\geq m} > Ke^{t_n}) \leq \frac{\varepsilon}{4}$$

Finally choose a $K_1 > K$ so that $e^{3m}e^{-K} < \frac{\varepsilon}{4}$. Then note that

$$\mathbb{P}(\mathcal{H}_{n,\leq m} \geq K_1 e^{t_n}) \leq \mathbb{P}(N(m) > e^{3m}) + e^{3m} P(\mathcal{P}(t_n) > K_1 e^{t_n})$$

where $\mathcal{P}(.)$ is a standard Yule process.

The argument of Lemma 101 and the choice of K and m gives that

$$\mathbb{P}(N(m) > e^{3m}) \leq Ce^{-m} \leq \frac{\varepsilon}{4}$$

Finally Lemma 86 gives us

$$\mathbb{P}(\mathcal{P}(t_n) > K_1 e^{t_n}) \leq e^{-K_1}$$

Combining and writing $\widehat{K}_\varepsilon = K_1 \cdot B$ we have

$$\mathbb{P}\left(\max(\deg(\mathcal{N}_n)) > \widehat{K}_\varepsilon \sqrt{n}\right) \leq \varepsilon$$

■

Model (C): The case where for some $\alpha > 2$ we have

$$\mathbb{P}(W > x) \sim \frac{1}{x^\alpha}$$

is very similar to the above analysis and we shall only describe the modifications that need to be made to Lemma 101. Note that due to the finiteness of the mean in this setup, Proposition 97 still holds in this setup. First we need some notation :

Since $\alpha > 2$ write $\frac{\alpha}{2} = 1 + \delta$ for some $\delta > 0$. Write $\beta := 1 + \delta/2$. As before we shall assume that t_n is a integer. Then we have the following Lemma

Lemma 102 *Let $m \leq t_n$ be any integer. Then*
(a) Define the random variable $N([m, m+1])$ as the number of births born in the time interval $[m, m+1]$ for the associated branching process. Then there exists a constant $C > 0$ such that

$$\mathbb{P}\left(N(m, m+1) > m^2 \cdot e^{2(m+1)}\right) \leq \frac{C}{m^2}$$

(b) Define $J_m := \max\{W_j : j$ born in the interval $[m, m+1]\}$. Then

$$\mathbb{P}\left(J_m \geq m^{\frac{4}{\alpha}} e^{\frac{2m}{\alpha}}\right) \leq 2\frac{C}{m^2}$$

(c) Define $D_m := \max\{\deg_i(t_n) :$ i born in the time interval $[m, m+1]\}$. Then for any fixed constant $K > 0$,

$$\mathbb{P}\left(D_m > Ke^{t_n - \frac{m}{\beta}}\right) \leq C\frac{1}{m^2} + m^2 e^{2m} \cdot e^{-Ke^{\frac{(\beta-1)\cdot m}{\beta}}}$$

(d) Finally define $\mathcal{H}_{n,m} := \max\{M_i :$ i is born in the interval $[m, m+1]$ $\}$. Then for any $K > 0$

$$\mathbb{P}\left(\mathcal{H}_{n,m} > m^{\frac{4}{\alpha}} \cdot Ke^{t_n - \frac{\delta \cdot m}{2\cdot(1+\delta)(1+\delta/2)}}\right) \leq 3C\frac{1}{m^2} + m^2 e^{2(m+1)} \cdot e^{-Ke^{\frac{(\beta-1)\cdot m}{\beta}}}$$

The rest of the analysis is similar and we shall skip the full discussion.

Proof of Theorem 42 : We first observe the counterpart of Prop 97 in this setup. Let ν be the distribution of W so that the tail behaves as

$$\nu[x, \infty) \sim \frac{1}{x^\alpha}$$

Then we have the following theorem for the amount of time we have to wait before the associated branching process $\mathcal{F}(\cdot)$ grows to size S_n.

Proposition 103 *Consider construction 4 where we first run a continuous time branching process $(N(t), \mathcal{F}(t))$ driven by a Yule offspring, then generate a sequence $W_i \sim_{iid} \nu$ having the above distribution and finally coalesce nodes $S_{i-1}+1, \ldots, S_i$ to form node i, where S_i is the partial sum sequence. Define*

$$T_{S_n} = \inf\{t > 0 : N(t) = S_n\}$$

for the time to observe the branching process $\mathcal{F}(\cdot)$, so as to construct $\mathcal{N}_n$. Then there exists a random variable $0 < W < \infty$ such that

$$T_{S_n} - \frac{p}{2}\log n \longrightarrow_L \log W$$

Proof: It is enough to observe

$$\frac{S_n}{e^{2S_n}} \longrightarrow_{a.s.} W$$

and

$$\frac{S_n}{n^p} \longrightarrow_L X_1$$

where $(X_t)_{t\geq 0}$ is a Levy process with Levy measure $\frac{1}{x^\alpha}$.

■

Define the random variable $\deg(\rho(\mathcal{F}_{T_{S_n}})) \leq \max(\mathcal{N}_n)$ for the degree of the root at time T_{S_n}. By Proposition 103, given $\varepsilon > 0$ we can find a $B > 0$ large, such that

$$\limsup_{n\to\infty} \mathbb{P}(T_{S_n} > \frac{p}{2}\log n - B) \geq 1 - \frac{\varepsilon}{2}$$

Write $t_n = \frac{p}{2}\log n - B$. Since the offspring distribution of the root is a rate one Yule process, the offspring at time t_n has the distribution $Geom(e^{-t_n})$. Thus we can find K such that

$$\limsup_{n\to\infty} \mathbb{P}(\deg(\rho(\mathcal{F}_{t_n})) \geq Ke^{t_n}) \geq 1 - \frac{\varepsilon}{2}$$

The result follows by noting that $e^{t_n} = \frac{1}{B}n^{\frac{p}{2}}$.

■

7 Conclusion

We end with a some general remarks, summarize our contribution and state some related conjectures:

(a) We went back to the very basics of embedding trees in continuous time branching processes to derive many results. The analysis of Urn schemes crucial to the conventional analysis of the degree distribution in many of the models, also use embedding into continuous time Markov processes to derive results for the degree distribution. However those methods are not pushed further to derive more explicit asymptotic information for local characteristics of the tree. We have used general branching process theory to derive much more detailed asymptotics for local and global aspects of the tree. The method does have the disadvantage of involving more mathematical abstraction such as point processes and renewal theory but seems to make up in the sheer amount of information it reveals.

(b) Much of the fundamental theory of continuous time branching process theory and the powerful theorems we have used, has been previously developed see e.g. [18], [19], [25], [2], [30]. The main contribution of this study has been to do explicit computations in a number of different models and show how this general theory actually helps us get explicit distributional asymptotics.

(c) Various special properties of the Yule process, such as the geometric distribution of finite time population sizes and the limiting random variable being exponentially distributed, allow us to do explicit computations in the BA-Preferential attachment model, where we were able to perform detailed asymptotics for various local statistics, including the empirical joint distribution of the degree of a node in the tree and the degree of it's mother. It is probably a little more difficult but presumably tractable to do similar computations in the more general setup of some of the other models.

(d) The work [5] analyzes the competition induced preferential model in some detail using clever continuous time Markov chains. Their analysis is equivalent to Corollary 34 in this paper. We see more information can be gleaned by using the abstract ideas of point process and continuous time branching processes. The work [5] analyzes some more preferential attachment schemes and it might be worth seeing if one can use the techniques exhibited in this paper to analyze those models as well. We have not pursued this line of thought in detail.

(e) Elimination of the use of concentration inequalities: The usual method of proof in these cases is via finding recurrence equations for means of random variables and concentration inequalities for random variables about their mean. See (i) for more details on the present techniques in use. However the general methods outlined in this paper, eliminate the need for such concentration inequalities, and result in one or two line arguments for statistics such as the asymptotic degree distribution.

(f)**Drawbacks:** It is fair to mention some of the drawbacks of the techniques outlined in this paper. They do not give rise to explicit finite n error bounds or rate of convergence results for quantities such as the degree distribution. See for example [12], where they analyze the Preferential attachment network models of Section 5.7 and establish finite n error bounds for the asymptotic degree distribution. The techniques used here also involve a higher technical overhead in terms of dealing with Point processes and renewal theory. However the exercise definitely seems worth the effort given the sheer amount of new information that results.

(g) It is the author's belief that the techniques used in this paper can be used to solve a wide variety of graph theoretic asymptotics for these random tree models. To state one concrete problem, note that we have proved, (see Theorem 71) that the size of the percolation component attached to the root is of order $n^{\gamma_p/\gamma}$ where γ is the Malthusian rate of growth of the original branching process and γ_p is the Malthusian rate of growth of the associated *pruned* branching process $\mathcal{F}^B(.)$. We believe something like the following is true:

Conjecture 104 *In most of the previous models there exists a random variable W_p such that if $\max(\mathcal{C}_n)$ denotes the largest percolation cluster then:*

$$\frac{|\max(\mathcal{C}_n)|}{n^{\frac{\gamma_p}{\gamma}}} \longrightarrow_P W_p$$

Another simple conjecture which should be provable easily is the "Thus the **smaller tail exponent** always wins !! " in Theorem 40. We state this as a Conjecture:

Conjecture 105 *Suppose $\mathbb{P}(W > j) \sim j^{-\alpha}$. Then the limiting degree distribution has the same power law exponent α.*

(h) We shall finally make some conjectures regarding the size of the maximal degree. Before we do so we need some notation. We shall split the class of models that we have analyzed into two classes:

Class I models: These are the models such as the BA preferential attachment model and the linear preferential attachment model where the point process $\mathcal{P}$ representing the offspring distribution of the associated branching process grows exponentially with time t, namely $\mathcal{P}(t) \sim e^{\lambda t}$.

Class II models: These are the models, including the attachment model with cutoff, the competition induced preferential attachment model and the sub-linear preferential attachment model where the process grows linearly or at most super linearly, namely $\mathcal{P}(t) \sim t^{\beta}$ for some $\beta \geq 1$.

Now note that for all the models, we have found probabilistic bounds of the form, given any $\varepsilon > 0$ there exists a constant K_ε such that

$$\limsup_{n\to\infty} \mathbb{P}\left(\frac{\max(deg)}{\omega_n} > K_\varepsilon\right) \leq \varepsilon$$

for some sequence of scaling constants $\omega_n \to \infty$. For e.g $\omega_n = \sqrt{n}$ for the BA-preferential attachment model. Now we believe the above techniques could be massaged to prove something stronger namely:

Conjecture 106 *(a) For the Class I models there exist random variables W (depending on the model) such that we have:*

$$\frac{\max(deg)}{\omega_n} \longrightarrow_d W$$

(b) For the Class II models we have:

$$\frac{\max(deg)}{\omega_n} \longrightarrow_P C$$

where C is a constant and depends on the model.

A starting point could be Theorem 65 which essentially states that to "observe" the size of the maximal degree, we need to only restrict our attention to nodes born within a large finite time say B after the root. Nodes born after this time have a very small probability of being the maximal degree node in $\mathcal{T}_n$.

(i) There are a number of Preferential attachment models formulated in the Spatial setting, see e.g. [16]. The methods exhibited in this study do not seem developed enough to work in the above settings but, it is the author's feeling that they might still be massaged to get something useful. Verifying the above statement seems a worthy topic of future research.

(j) Another model which is not amenable to the general *stable age distribution* theory is the superlinear preferential attachment model, where the attractiveness function is given by:

$$f(v,n) = (D(v,n)+1)^\beta$$

where $\beta > 1$. This model has been comprehensively analyzed in [29].

(k) **Comparison with existing techniques:** As remarked before, for most models, the main functional (and in many cases the only functional) to have been rigorously analyzed is the degree distribution. The standard way of analysis in may of the cases is a sfollows: Let $N_k(m)$ to be the number of degree k vertices in $\mathcal{T}_m$. Then to show that random variables $N_k(m)$ satisfy the property:

$$\frac{X_k(m)}{m} \longrightarrow_P p_k$$

for some probability mass function $(p_k)_{k\geq 1}$, the following two step procedure is followed:

1. Note that for all the tree models, the degree distribution vector $(N_1(m), N_2(m), \ldots)_{m\geq 1}$ is a Markov chain. This allows us to setup recursions for $X_k(m) = \mathbb{E}(N_k(m))$ via conditioning on what happens on step m, given the degree distribution at step $m-1$. It is then not very hard to prove via these recursions that (although there are some mathematical caveats involved, see e.g. Durrett [14] Chapter 4 more details)

$$\frac{\mathbb{E}\left(N_k(m)\right)}{m} \longrightarrow p_k \tag{48}$$

as $m \to \infty$.

2. Prove concentration for the random variables $N_k(m)$ via concentration inequalities such as the Azuma-Hoeffding inequality or through more subtle methods see e.g. [5]. These methods essentially state that for large m, $|N_k(m) - \mathbb{E}(N_k(m))| = o(m)$, which coupled with Equation (37) implies that
$$\frac{N_k(m)}{m} \longrightarrow_P p_k$$
as well. One of the contributions of this study has been to provide an alternate route to obtain such results. If one is able to formulate a related continuous branching process mechanism via an appropriate formulation of the rates which satisfies some easily checkable conditions, such as Malthusianity and the **xlogx** conditions, then we immediately get, via the *stable age distribution theory* of such processes, all of the above and in fact much more regarding the statistical properties of fringes and extended fringes. Not only this the above results in a lot of global information as well, which we hope Section 5 illustrates.

8 Acknowledgments

I would like to thank Elchanan Mossel and the organizers of the IPAM Random Networks conference in UCLA, May 2007 where I was first exposed to many of the preferential attachment models and was able to interact with many of the stalwarts in the field. I would also like to thank my advisor David Aldous for helpful advice and support and being always available. I would like to thank fellow graduate student Allan Sly for helpful discussions. I would like to thank Noam Berger and Yuval Peres for providing constant encouragement. We thank Steve Evans for providing helpful references regarding $\alpha-$recurrence theory, used in Section 6.4.1.

References

[1] Réka Albert and Albert-László Barabási. Statistical mechanics of complex networks. *Rev. Modern Phys.*, 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 J. Aldous. Stochastic models and descriptive statistics for phylogenetic trees, from Yule to today. *Statist. Sci.*, 16(1):23–34, 2001.

[4] K. B. Athreya and P. E. Ney. *Branching processes*. Dover Publications Inc., Mineola, NY, 2004. Reprint of the 1972 original [Springer, New York; MR0373040].

[5] N. Berger, C. Borgs, J. T. Chayes, R. M. D'Souza, and R. D. Kleinberg. Degree distribution of competition-induced preferential attachment graphs. *Combin. Probab. Comput.*, 14(5-6):697–721, 2005.

[6] Ginestra Bianconi and Albert-László Barabási. Bose-einstein condensation in complex networks. *Phys. Rev. Lett.*, 86(24):5632–5635, Jun 2001.

[7] Béla Bollobás and Oliver Riordan. Robustness and vulnerability of scale-free random graphs. *Internet Math.*, 1(1):1–35, 2003.

[8] Béla Bollobás and Oliver Riordan. The diameter of a scale-free random graph. *Combinatorica*, 24(1):5–34, 2004.

[9] Borgs Christian, Chayes Jennifer, Daskalakis C, and Roch Sebastien. An analysis of preferential attachment with fitness. *STOC*, 2007.

[10] Fan Chung and Linyuan Lu. *Complex graphs and networks*, volume 107 of *CBMS Regional Conference Series in Mathematics*. Published for the Conference Board of the Mathematical Sciences, Washington, DC, 2006.

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

[12] Maria Deijfen, Henri van den Esker, Remco van der Hofstad, and Gerard Hooghiemstra. A preferential attachment model with random initial degrees. 2007.

[13] S. N. Dorogovtsev and J. F. F. Mendes. *Evolution of networks*. Oxford University Press, Oxford, 2003. From biological nets to the Internet and WWW.

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

[15] G. Ergun and G. J. Rodgers. Growing random networks with fitness. *PHYSICA A*, 303:261, 2002.

[16] Abraham Flaxman, Alan Frieze, and Juan Vera. A geometric preferential attachment model of networks.

[17] B. V. Gnedenko and A. N. Kolmogorov. *Limit distributions for sums of independent random variables*. Addison-Wesley Publishing Company, Inc., Cambridge, Mass., 1954. Translated and annotated by K. L. Chung. With an Appendix by J. L. Doob.

[18] Peter Jagers. *Branching processes with biological applications*. Wiley-Interscience [John Wiley & Sons], London, 1975. Wiley Series in Probability and Mathematical Statistics—Applied Probability and Statistics.

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

[20] Peter Jagers and Olle Nerman. The asymptotic composition of supercritical multi-type branching populations. In *Séminaire de Probabilités, XXX*, volume 1626 of *Lecture Notes in Math.*, pages 40–54. Springer, Berlin, 1996.

[21] J. F. C. Kingman. The first birth problem for an age-dependent branching process. *Ann. Probability*, 3(5):790–801, 1975.

[22] P. L. Krapivsky and S. Redner. Rate Equation Approach for Growing Networks. In R. Pastor-Satorras, M. Rubi, and A. Diaz-Guilera, editors, *Statistical Mechanics of Complex Networks*, volume 625 of *Lecture Notes in Physics, Berlin Springer Verlag*, pages 3–22, 2003.

[23] Michael Mitzenmacher. A brief history of generative models for power law and lognormal distributions. *Internet Math.*, 1(2):226–251, 2004.

[24] Tamás F. Móri. The maximum degree of the Barabási-Albert random tree. *Combin. Probab. Comput.*, 14(3):339–348, 2005.

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

[26] M. E. J. Newman. The structure and function of complex networks. *SIAM Rev.*, 45(2):167–256 (electronic), 2003.

[27] S. Niemi and E. Nummelin. On nonsingular renewal kernels with an application to a semigroup of transition kernels. *Stochastic Process. Appl.*, 22(2):177–202, 1986.

[28] S. Niemi and E. Nummelin. On nonsingular renewal kernels with an application to a semigroup of transition kernels. *Stochastic Process. Appl.*, 22(2):177–202, 1986.

[29] Roberto Oliveira and Joel Spencer. Connectivity transitions in networks with super-linear preferential attachment. *Internet Math.*, 2(2):121–163, 2005.

[30] Boris Pittel. Note on the heights of random recursive trees and random m-ary search trees. *Random Structures Algorithms*, 5(2):337–347, 1994.

[31] Anna Rudas, Balint Toth, and Benedek Valko. Random trees and general branching processes. *Random Structures and Algorithms*, 2007.

[32] 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.

[33] Richard L. Tweedie. R-theory for Markov chains on a general state space. I. Solidarity properties and R-recurrent chains. *Ann. Probability*, 2:840–864, 1974.