

FLUCTUATION BOUNDS FOR CONTINUOUS TIME BRANCHING PROCESSES AND EVOLUTION OF GROWING TREES WITH A CHANGE POINT

BY SAYAN BANERJEE^{1,a}, SHANKAR BHAMIDI^{1,b} AND IAIN CARMICHAEL^{2,c}

¹*Department of Statistics and Operations Research, UNC Chapel Hill*

^a*sayan@email.unc.edu,*

^b*bhamidi@email.unc.edu*

²*Department of Statistics, UC Berkeley*

^c*iaian@berkeley.edu*

We consider dynamic random trees constructed using an attachment function $f : \mathbb{N} \rightarrow \mathbb{R}_+$ where, at each step of the evolution, a new vertex attaches to an existing vertex v in the current tree with probability proportional to $f(\text{degree}(v))$. We explore the effect of a change point in the system; the dynamics are initially driven by a function f until the tree reaches size $\tau(n) \in (0, n)$, at which point the attachment function switches to another function, g , until the tree reaches size n . Two change point time scales are considered, namely the *standard model* where $\tau(n) = \gamma n$, and the *quick big bang model* where $\tau(n) = n^\gamma$, for some $0 < \gamma < 1$. In the former case, we obtain deterministic approximations for the evolution of the empirical degree distribution (EDF) in sup-norm and use these to devise a provably consistent non-parametric estimator for the change point γ . In the latter case, we show that the effect of pre-change point dynamics asymptotically vanishes in the EDF, although this effect persists in functionals such as the maximal degree. Our proofs rely on embedding the discrete time tree dynamics in an associated (time) inhomogeneous continuous time branching process (CTBP). In the course of proving the above results, we develop novel mathematical techniques to analyze both homogeneous and inhomogeneous CTBPs and obtain rates of convergence for functionals of such processes, which are of independent interest.

1. Introduction. Driven by the explosion in the amount of data on various real world networks, the last few years have seen the emergence of many new mathematical network models. Motivations behind these models are diverse [1, 15, 25, 38, 39, 47] including **(a)** extracting unexpected patterns in the network (e.g. community detection); **(b)** understanding properties of dynamics (e.g. the spread of epidemics); **(c)** understanding mechanistic reasons for the emergence of empirically observed properties of real world systems. An important niche is the setting of networks that evolve over time. In the context of probabilistic combinatorics, these models have been studied for decades in the vast field of *recursive trees*, e.g. see [12, 24, 26, 35] and the references therein.

To fix ideas, consider the general random tree model called *non-uniform random recursive trees* [45]. Fix $n \geq 1$ and an *attachment function* $f : \{0, 1, 2, \dots\} \rightarrow (0, \infty)$. A sequence of random trees $\{\mathcal{T}_k : 1 \leq k \leq n\}$ is grown as follows ($\mathcal{T}_k$ has k vertices labelled by the integers $[k] := \{1, \dots, k\}$). For $k = 1$, $\mathcal{T}_1$ has one vertex, which we call the “root.” For fixed $k \geq 2$, $\mathcal{T}_k$ is constructed conditional on $\mathcal{T}_{k-1}$ as follows. A new vertex, k , is born into the system and attaches to a previously existing vertex $v \in [k-1]$ with probability proportional

MSC2020 subject classifications: Primary 60C05; secondary 05C80.

Keywords and phrases: continuous time branching processes, temporal networks, change point detection, random networks, stable age distribution theory, Malthusian rate of growth, inhomogeneous branching processes.

to $f(\deg(v))$, where $\deg(v)$ denotes the number of children of v (which is one less than its graph degree in $\mathcal{T}_{k-1}$). Thus,

$$\mathbb{P}(k \text{ attaches to } v \in [k-1] \mid \mathcal{T}_{k-1}) := \frac{f(\deg(v))}{\sum_{u=1}^{k-1} f(\deg(u))}.$$

The vertex that k selects is called the “parent” and the edge is directed from the parent to the new “child” vertex. The case of $f(\cdot) \equiv 1$ corresponds to the famous class of *random recursive trees* [44]. The specific case of “linear preferential attachment” when f is affine was considered in [10] to provide a generative story for heavy tailed degree distributions of real networks.

Next, consider the non-uniform random recursive tree model with a *change point*. Here, the random tree is grown according to one rule till some (possibly random) time called the change point, after which the dynamics switch to another rule. In detail, let $1 \leq \tau < n$ and $f_0, f_1 : \{0, 1, 2, \dots\} \rightarrow (0, \infty)$ be two attachment functions. For $1 \leq k \leq \tau$ the process evolves according to the *initializer* function f_0 i.e. node k attaches to pre-existing vertex $v \in [k-1]$ with probability proportional to $f_0(\deg(v))$. After the change point for $k \in [\tau+1, n]$ the process evolves according to f_1 . We denote this sequence of random trees by $\{\mathcal{T}_k^\theta : 1 \leq k \leq n\}$, where $\theta = (f_0, f_1, \tau)$. While the focus of this paper is on one change point, the methodology allows one to derive analogous results for multiple change points.

1.1. *Informal description of our aims and results.* This paper has the following two major aims for the models described above:

- (a) Asymptotics in the large network limit as well as corresponding functionals have been derived for a host of random tree models [2, 13, 28]. One major driver of research has been proving convergence of the empirical distribution of these functionals to model dependent constants. Establishing (even suboptimal) rates of convergence for these models has been non-trivial other than for models related to urn models e.g. see the seminal work of Janson [32]. The aim of this paper is to develop robust methodology for proving such error bounds for general models. These error bounds play a key role in understanding robustness properties of network source detection problems, see e.g. [9].
- (b) We aim to understand the effect of change points on structural properties of the network. Analogous to classical change point detection, we start by considering models with a change point at time $\tau = \lfloor \gamma n \rfloor$ for $0 < \gamma < 1$ (referred to as the *standard model* in the sequel). Using techniques involving embedding into inhomogeneous continuous time branching processes, we approximate the empirical degree distribution (EDF) as it evolves in time by deterministic curves derived from the attachment functions f_0 and f_1 . Using this, we devise a non-parametric estimator for γ . Counter-intuitively, we find that irrespective of the value of γ , structural properties of the network such as the tail of the degree distribution are **only** determined by model parameters **before** the change point. Motivated by this we consider other time scales of the change point, namely, when the change happens at time $\tau = \lfloor n^\gamma \rfloor$ for $0 < \gamma < 1$ (the *quick big bang* model) to understand the extent of this long range dependence phenomenon. In this case, we show that the effect of the pre-change point dynamics asymptotically vanishes in the EDF. However, for specific examples of attachment functions f_0 and f_1 , we conclude that this change point has a drastic effect on asymptotics for the maximal degree.

2. Preliminaries. We use $\leq_{\text{st}}$ for stochastic domination between two real valued probability measures. For $J \geq 1$, let $[J] := \{1, 2, \dots, J\}$. A random variable Y with rate λ exponential distribution is denoted by $Y \sim \exp(\lambda)$. Write $\mathbb{Z}$ for the set of integers, $\mathbb{R}$ for

the real line, $\mathbb{N}$ for the set of natural numbers and let $\mathbb{Z}_+ := \{0, 1, 2, \dots\}$, $\mathbb{R}_+ := (0, \infty)$. Write $\xrightarrow{\text{a.s.}}, \xrightarrow{P}, \xrightarrow{d}$ for convergence almost surely, in probability and in distribution respectively. For non-negative function g and another function f both defined on $\mathbb{N}$, we write $f(n) = O(g(n))$ when $|f(n)|/g(n)$ is uniformly bounded, and $f(n) = o(g(n))$ when $\lim_{n \rightarrow \infty} f(n)/g(n) = 0$. Furthermore, write $f(n) = \Theta(g(n))$ if $f(n) = O(g(n))$ and $g(n) = O(f(n))$. A sequence of events $(A_n)_{n \geq 1}$ occurs *with high probability* (whp) when $\mathbb{P}(A_n) \rightarrow 1$. For some σ -field $\mathcal{F}$, an integrable random variable X and non-random constant C , when we write $\mathbb{E}(X|\mathcal{F}) \leq C$, this will formally mean that the bound holds with probability one. For a sequence of increasing rooted trees $\{\mathcal{T}_n : n \geq 1\}$ (random or deterministic), we assume that edges are directed from parent to child (with the root as the original progenitor). For exposition purposes we will write *degree* for *out-degree* i.e. the number of children of a vertex. This should not be confused with the *total degree* or *graph degree*, which is the sum of incoming and outgoing edges (and thus the graph degree of a vertex is always one more than the out-degree in our tree networks). For $n \geq 1$ and $k \geq 0$, let $D_n(k)$ be the number of vertices in $\mathcal{T}_n$ with degree k ; thus $D_n(0)$ counts the number of leaves in $\mathcal{T}_n$.

2.1. Organization of the paper. The rest of Section 2 defines key objects required to state our main results. Section 3 contains the main results. In Section 4 we discuss the relevance of this work and related literature. The remaining sections are devoted to the proofs of the main results.

2.2. Branching processes. Fix an attachment function f . For $i \geq 0$ let $E_i \sim \exp(f(i))$, $i \geq 0$ be a sequence of independent exponential random variables. Define for $i \geq 1$, $L_i := \sum_{j=0}^{i-1} E_j$. Let ξ_f be the point process on $\mathbb{R}_+$:

$$(2.1) \quad \xi_f := (L_1, L_2, \dots).$$

Abusing notation, write for $t \geq 0$,

$$(2.2) \quad \xi_f[0, t] := \#\{i : L_i \leq t\}, \quad \mu_f[0, t] := \mathbb{E}(\xi_f[0, t]).$$

Here we view μ_f as a measure on $(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$. We will also need variants of the above objects: for fixed $k \geq 0$, let $\xi_f^{(k)}$ denote the k -shifted version of the point process ξ_f where the first inter-arrival time is E_k i.e. define the sequence, $L_i^{(k)} = E_k + E_{k+1} + \dots + E_{k+i-1}$, $i \geq 1$ and then let

$$(2.3) \quad \xi_f^{(k)} := (L_1^{(k)}, L_2^{(k)}, \dots), \quad \mu_f^{(k)}[0, t] := \mathbb{E}(\xi_f^{(k)}[0, t]).$$

As above, $\xi_f^{(k)}[0, t] := \#\{i : L_i^{(k)} \leq t\}$. We abbreviate $\xi_f[0, t]$ as $\xi_f(t)$ and similarly $\mu_f(t)$, $\xi_f^{(k)}(t)$, $\mu_f^{(k)}(t)$. Define the Ulam-Harris set $\mathcal{J} := \cup_{d=0}^{\infty} \mathbb{N}^d$ where $\mathbb{N} = \{1, 2, \dots\}$ and $\mathbb{N}^0 = \{\emptyset\}$ for the root of the tree.

DEFINITION 2.1 (Continuous time Branching process (CTBP) [4, 29]). *Fix an attachment function $f : \mathbb{Z}_+ \rightarrow \mathbb{R}_+$. A continuous time branching process driven by f , written as $\{\text{BP}_f(t) : t \geq 0\}$, is a $\mathcal{J}$ -valued process, started with one individual $\emptyset$ (the root) at time $t = 0$, such that every individual $x \in \mathcal{J}$ born into the system gives birth to offspring $\{(x, i) : i \in \mathbb{N}\}$ with birth times given by an independent copy of the point process ξ_f defined in (2.1). For $t \geq 0$, $\text{BP}_f(t)$ denotes the set of individuals alive at time t and $Z_f(t) := |\text{BP}_f(t)|$ denotes the size of this set. For $x \in \mathcal{J}$, let σ_x denote the birth time of x .*

In analogy with the original tree model, we will often refer to individuals in the branching process as vertices and the number of children of an individual in the population at

time t as its *degree* at time t . In our construction, by the strict positivity assumption on the attachment function, individuals continue to reproduce forever. Let

$$(2.4) \quad m_f(t) := \mathbb{E}(Z_f(t)), \quad t \geq 0,$$

If $\limsup_{k \rightarrow \infty} f(k)/k < \infty$, it can be shown [29, Chapter 3] that for all $t > 0$, $m_f(t) < \infty$, and that $m_f(\cdot)$ is strictly increasing with $m_f(t) \uparrow \infty$ as $t \uparrow \infty$. In the sequel, to simplify notation we will suppress dependence on f and write $\text{BP}(\cdot)$, $m(\cdot)$ etc. The connection between CTBP and the discrete random tree models is given by the following result and is the starting point of the Athreya-Karlin embedding [3].

LEMMA 2.2. *Fix an attachment function f and consider the sequence of random trees $\{\mathcal{T}_m : 2 \leq m \leq n\}$ constructed using attachment function f . Consider the continuous time construction in Definition 2.1 and define for $m \geq 1$ the stopping times $T_m := \inf\{t \geq 0 : |\text{BP}_f(t)| = m\}$. Then viewed as a sequence of growing random labelled rooted trees we have, $\{\text{BP}_f(T_m) : 2 \leq m \leq n\} \stackrel{d}{=} \{\mathcal{T}_m : 2 \leq m \leq n\}$.*

2.3. *Continuous embedding of model with single change point.* The continuous time embedding of the tree model without change point has a natural extension to the model with a single change point $\tau \in \mathbb{N}$. Individuals in the population reproduce according to independent copies of the point process ξ_{f_0} up till the time $t(\tau)$ when the total population size is τ . After this time, individuals continue to reproduce independently as follows. An individual of degree ℓ at time $t(\tau)$ reproduces according to the point process $\xi_{f_1}^{(\ell)}$. New individuals born into the system after time $t(\tau)$ reproduce according to ξ_{f_1} . Recalling the notation for driving parameters $\theta = (f_0, f_1, \tau)$ and denoting this *inhomogeneous* branching process by $\text{BP}_\theta(\cdot)$, the same approach used to prove Lemma 2.2 shows that $\{\text{BP}_\theta(T_m^\theta) : 2 \leq m \leq n\} \stackrel{d}{=} \{\mathcal{T}_m^\theta : 2 \leq m \leq n\}$, where $T_m^\theta := \inf\{t \geq 0 : |\text{BP}_\theta(t)| = m\}$, $m \geq 1$. Note that $t(\tau) = T_\tau^\theta$. We will refer to this random time $t(\tau)$ as the change point for the branching process $\text{BP}_\theta(\cdot)$. When θ is clear from context, we will often drop the superscript (or subscript) θ from associated quantities for notational convenience.

REMARK 2.3. There exists a common probability space $(\Omega^*, \mathcal{F}^*, \mathbb{P}^*)$ on which the process $\{\text{BP}_\theta(T_m^\theta) : 2 \leq m \leq n\}$, and hence $\{\mathcal{T}_m^\theta : 2 \leq m \leq n\}$, can be constructed for all n . $(\Omega^*, \mathcal{F}^*, \mathbb{P}^*)$ can be taken to be a probability space on which the countable i.i.d. collection $\{\xi_{f_0, x} : x \in \mathcal{J}\}$ and i.i.d. collection $\{\xi_{f_1, x} : x \in \mathcal{J}\}$ of point processes are defined. Individual x with $\sigma_x < t(\tau)$ uses $\xi_{f_0, x}$ to reproduce until the stopping time $t(\tau)$ when the total population size hits τ . If x has k children at time $t(\tau)$, it uses the k -shifted version $\xi_{f_1, x}^{(k)}$ of the point process $\xi_{f_1, x}$ for future reproduction. For $x \in \mathcal{J}$ with $\sigma_x \geq t(\tau)$, x has all its reproduction according to $\xi_{f_1, x}$. Future references to convergence in probability and almost sure convergence for the associated branching processes and trees with change point will all be implicitly assumed to take place on $(\Omega^*, \mathcal{F}^*, \mathbb{P}^*)$.

2.4. *Assumptions on attachment functions.* Here we set up assumptions as well as constructions needed to state the main results. We mainly follow [29, 30, 37, 43].

ASSUMPTION 2.4. (i) *Every attachment function f is strictly positive and can grow at most linearly,*

$$\sup_{k \geq 0} f(k)/(k+1) = C < \infty.$$

(ii) Consider the following function $\hat{\rho} : (0, \infty) \rightarrow (0, \infty]$ defined via,

$$(2.5) \quad \hat{\rho}(\lambda) := \sum_{k=1}^{\infty} \prod_{i=0}^{k-1} \frac{f(i)}{\lambda + f(i)}.$$

Define $\underline{\lambda} := \inf\{\lambda > 0 : \hat{\rho}(\lambda) < \infty\}$. We assume,

$$(2.6) \quad \lim_{\lambda \downarrow \underline{\lambda}} \hat{\rho}(\lambda) > 1.$$

Using (ii) of the above Assumption, let $\lambda^* := \lambda^*(f)$ be the unique λ such that

$$(2.7) \quad \hat{\rho}(\lambda^*) = 1.$$

The *Malthusian rate of growth* parameter, λ^* , is intricately tied to the growth rate of a continuous time branching process. In fact, recalling that $Z_f(t)$ denotes the size of the branching process at time t , under Assumption 2.4, $e^{-\lambda^* t} Z_f(t)$ converges in probability as $t \rightarrow \infty$ to a finite random variable [37, Theorem 3.1] (see also Lemma 6.6 below). While not obvious, λ^* plays an essential role in all law of large number results about the **discrete time** tree model with attachment function f (see, for example, (3.1) below).

REMARK 2.5. The requirement (2.6) is a standard assumption in branching process literature that implies almost sure convergence of a broad collection of branching process statistics of ratio-type [37, Theorem 6.3]. It is algebraic in nature and can be checked for a given attachment function using the explicit form of $\hat{\rho}(\cdot)$ given in (2.5). In particular, (2.6) is satisfied for $f(\cdot) \equiv 1$ (easy to check that $\hat{\rho}(\lambda) = \lambda^{-1}, \lambda > 0$), $f(k) = k + 1 + \beta, k \geq 0$, for any $\beta > 0$ [43, Section 4.2] and $f(k) = (k + 1)^\alpha, k \geq 0$, for any $\alpha \in (0, 1)$ [34, Lemma 10]. See [8, Lemma 7.8] for additional checkable conditions for (2.6) to hold.

3. Main Results.

3.1. *Convergence rates for model without change point.* Consider a continuous time branching process with attachment function f and Malthusian rate λ^* . For $k \geq 0, t \geq 0$, let $D(k, t)$ denote the number of vertices in $\text{BP}_f(t)$ with degree k and abbreviate $Z_f(t)$ to $Z(t)$. Let $\lambda^* = \lambda^*(f)$ be as in (2.7). Define the probability mass function $\mathbf{p}(f) := \{p_k : k \geq 0\}$ via,

$$(3.1) \quad p_k = p_k(f) := \int_0^\infty \lambda^* e^{-\lambda^* t} \mathbb{P}(\xi_f(t) = k) dt = \frac{\lambda^*}{\lambda^* + f(k)} \prod_{j=0}^{k-1} \frac{f(j)}{\lambda^* + f(j)}, \quad k \geq 0.$$

For $k = 0$, $\prod_{j=0}^{k-1}$ is taken to be 1. The last equality above follows from standard calculations involving exponential distributions (see, for example, the proof of Theorem 2 (a) in [43]). Following the seminal work of [29, 30, 37, 43] for each $k \geq 0$, $D(k, t)/Z(t) \rightarrow p_k$ in probability as $t \rightarrow \infty$. However to get consistent change point estimators we need to strengthen this convergence to a sup-norm convergence on a time interval whose size goes to infinity with growing t as well as obtain a quantitative rate for this convergence. Such results have been obtained for very specific attachment functions via functional central limit theorems but do not extend to the setting of general attachment functions; see e.g. [32]; specific to linear attachment see [16, 42, 47]. The following assumptions on the attachment function will play a crucial role in this section.

ASSUMPTION 3.1. *There exists $C^* \geq 0$ such that $\lim_{k \rightarrow \infty} f(k)/k = C^*$.*

ASSUMPTION 3.2. *$\text{Var}\left(\int_0^\infty e^{-\lambda^* t} \xi_f(dt)\right) < \infty$.*

REMARK 3.3. Assumption 3.2 might at first sight seem opaque. Here we give three conceptually easier sufficient conditions that cover a wide array of functions. Throughout we assume Assumption 2.4. Assumption 3.2 holds if any of the following three conditions hold.

- (a) **Diverging attachment functions:** $f(k) \rightarrow \infty$ as $k \rightarrow \infty$. See [43, Proof of Lemma 1].
- (b) **Finite variance of the degree distribution:** $\sum_{k=0}^{\infty} k^2 p_k(f) < \infty$. To see this note

$$\begin{aligned} \mathbb{E} \left[\left(\int_0^{\infty} e^{-\lambda^* t} \xi_f(dt) \right)^2 \right] &= \mathbb{E} \left[\left(\int_0^{\infty} \lambda^* e^{-\lambda^* t} \xi_f(t) dt \right)^2 \right] \leq \mathbb{E} \left(\int_0^{\infty} \lambda^* e^{-\lambda^* t} \xi_f^2(t) dt \right) \\ &= \int_0^{\infty} \lambda^* e^{-\lambda^* t} \sum_{k=1}^{\infty} k^2 \mathbb{P}(\xi_f(t) = k) dt = \sum_{k=1}^{\infty} k^2 \left(\int_0^{\infty} \lambda^* e^{-\lambda^* t} \mathbb{P}(\xi_f(t) = k) dt \right) = \sum_{k=1}^{\infty} k^2 p_k(f) < \infty \end{aligned}$$

where the last equality follows from (3.1). For a given f , the finiteness of the above sum can possibly be checked using the explicit formula for $p_k(f)$ given in (3.1).

- (c) **Lower boundedness and asymptotic linearity:** If $\inf_{k \geq 0} f(k) > 0$ and $\lim_{k \rightarrow \infty} f(k)/k = C^* \geq 0$ (Assumption 3.1).

This assertion was largely proven in [9] which we now explain. In Lemma 9.1 we show that under these assumptions, $\lim_{k \rightarrow \infty} f(k)/k < \lambda^*$ and hence $\hat{\rho}(\lim_{k \rightarrow \infty} f(k)/k) > 1$. Then [9, Proposition 5.7] shows that in this case $\int_0^{\infty} e^{-\lambda^* t} \xi_f(dt)$ has finite **exponential** moments and thus, in particular, Assumption 3.2 on finiteness of the second moment holds.

Fix a sequence of growing trees $\{\mathcal{T}_m : m \geq 2\}$ and recall that for any $N \geq 2$ and $k \geq 0$, $D_N(k)$ denotes the number of vertices in $\mathcal{T}_N$ with degree k . The following theorem establishes convergence of the empirical degree distribution to its limit in a certain ‘uniform’ sense and furnishes a rate for this convergence.

THEOREM 3.4. *Consider a continuous time branching process with attachment function f that satisfies Assumptions 2.4, 3.1 and 3.2. Let $\mathbf{p}(f)$ be the limiting degree distribution as in (3.1). There exist $\omega^* \in (0, 1)$, $\epsilon^{**} \in (0, 1)$, such that for any $\epsilon \leq \epsilon^{**}$,*

$$n^{\omega^*} \sum_{k=0}^{\infty} 2^{-k} \left(\sup_{t \in [0, 2\epsilon \log n / \lambda^*]} \left| \frac{D(k, \frac{1-\epsilon}{\lambda^*} \log n + t)}{Z(\frac{1-\epsilon}{\lambda^*} \log n + t)} - p_k \right| \right) \xrightarrow{P} 0.$$

Thus for a sequence of non-uniform recursive trees $\{\mathcal{T}_m : m \geq 2\}$ grown using attachment function f ,

$$n^{\omega^*} \sum_{k=0}^{\infty} 2^{-k} \sup_{n^{1-\epsilon} \leq N \leq n^{1+\epsilon}} |D_N(k)/N - p_k| \xrightarrow{P} 0.$$

The analysis of branching processes in continuous time starts via scoring individuals existing at any fixed time t via so called characteristics, measuring individuals (and their offspring) in various phases of their life, weighting existing individuals using these scores and then deriving asymptotics as $t \rightarrow \infty$. Such characteristics can in principle depend on the entire set of descendants (not just immediate offspring) of an individual, including ones that are born at future times. We refer the interested reader to [30, 31] for further discussion on the importance of such characteristics and [2] for describing the importance of such results in the context of local weak convergence of large discrete random structures. An important technical contribution of this paper is the next result, Theorem 3.5, regarding rates of convergence for normalized counts associated with general characteristics.

We introduce some notation related to functionals of branching processes, closely following [30, 37]. Recall that the individuals in the population are indexed by $\mathcal{J} = \cup_{d=0}^{\infty} \mathbb{N}^d$ and for $x \in \mathcal{J}$, σ_x denotes the birth time of x . Let $\{\xi_{f,x}, x \in \mathcal{J}\}$ be i.i.d. copies of the point process ξ_f (see (2.1)), where each $\xi_{f,x}$ is defined on some probability space $(\Omega_x, \mathcal{A}_x, \mathbb{P}_x)$. $\xi_{f,x}$ encodes the times of birth of children of x . The underlying probability space for the branching process (without a change point) is taken to be $(\Omega, \mathcal{A}, \mathbb{P}) = \prod_{x \in \mathcal{J}} (\Omega_x, \mathcal{A}_x, \mathbb{P}_x)$. Elements of Ω are denoted by $\omega = \{\omega_x : x \in \mathcal{J}\}$. For each $x \in \mathcal{J}$, define the shift operator $S_x : \Omega \rightarrow \Omega$ which maps $\{\omega_y : y \in \mathcal{J}\}$ to $\{\omega_{xy} : y \in \mathcal{J}\}$. Thus, the shift operator S_x maps $\emptyset$ and its descendants to x and its descendants. A characteristic $\phi : \mathbb{R} \times \Omega \rightarrow \mathbb{R}_+$ is a $\mathcal{B}(\mathbb{R}) \times \mathcal{A}$ -measurable, separable, non-negative random process. We assume $\phi(t, \omega) = 0$ for every $t < 0, \omega \in \Omega$. Later in (3.3), we will make further assumptions on the stochastic process $\{\phi(t, \omega), t \in \mathbb{R}\}$.

Informally, for each $t \geq 0$, $\phi(t)$ can be thought of as a ‘score’ assigned to the root at time t , namely when the root is of **age** t . For each $x \in \mathcal{J}$, the characteristic corresponding to x , naturally obtained from ϕ , is defined by $\phi_x(t, \omega) := \phi(t, S_x(\omega)), t \geq 0$. Thus, $\phi_x(t)$ can be thought of as the score given to x based on x and its descendants when x is of age t . We suppress the dependence of ϕ, ϕ_x on ω and write $\phi(t), \phi_x(t)$ for $\phi(t, \omega)$ and $\phi_x(t, \omega)$ respectively.

For any characteristic ϕ , define $Z_f^\phi(t) := \sum_{x \in \mathcal{J}} \phi_x(t - \sigma_x) = \sum_{x \in \text{BP}_f(t)} \phi_x(t - \sigma_x)$. This can be thought of as the sum of ϕ -scores, or aggregate ϕ -score, of all individuals in $\text{BP}_f(t)$. In particular, the age of individual x in $\text{BP}_f(t)$ is $t - \sigma_x$, and hence its contribution to the aggregate ϕ -score is $\phi_x(t - \sigma_x)$. Write $m_f^\phi(t) = \mathbb{E}(Z_f^\phi(t))$ and $M_f^\phi(t) = \mathbb{E}(e^{-\lambda^* t} Z_f^\phi(t))$. Note the characteristics $\phi(t) = \mathbb{1}\{t \geq 0\}$ and $\phi(t) = \mathbb{1}\{\xi(t) = k\}, k \geq 0$, count the total number of vertices and number of vertices of degree k at time t respectively. For these two specific characteristics we write the associated scores as $Z_f(t)$ and $Z_f^{(k)}(t)$ respectively; analogously we write $m_f(t), m_f^{(k)}(t)$ and $M_f(t), M_f^{(k)}(t)$. It is easy to check that for a general (integrable) characteristic ϕ , $M_f^\phi(t)$ satisfies the renewal equation

$$(3.2) \quad M_f^\phi(t) = e^{-\lambda^* t} \mathbb{E}(\phi(t)) + \int_0^t M_f^\phi(t-s) e^{-\lambda^* s} \mu_f(ds).$$

Write $M_f^\phi(\infty) = \lim_{t \rightarrow \infty} M_f^\phi(t)$ when the limit exists. Following [37], for $t \geq 0$, let $\mathcal{J}(t) = \{x = (x', i) : \sigma_{x'} \leq t \text{ and } t < \sigma_x < \infty\}$ denote the set of individuals born after time t to parents who were born at or before time t . Write $W_t := \sum_{x \in \mathcal{J}(t)} e^{-\lambda^* \sigma_x}$. By Corollary 2.5 of [37], W_t converges almost surely to a finite random variable W_∞ as $t \rightarrow \infty$. By Theorem 3.1 of [37], $e^{-\lambda^* t} Z_f^\phi(t) \xrightarrow{P} W_\infty M_f^\phi(\infty)$ for any $\phi \in \mathcal{C}$.

For this article, we are interested in the following class of characteristics (where once again recall $\emptyset$ denotes the root of the tree):

$$(3.3) \quad \mathcal{C} := \{\phi \text{ with càdlàg paths} : \exists \text{ a non-random } b_\phi > 0 \text{ such that } \phi(t) \leq b_\phi(\xi_{f,\emptyset}(t) + 1) \text{ for all } t \geq 0\}.$$

THEOREM 3.5. *Consider a continuous time branching process with attachment function f that satisfies Assumptions 2.4 and 3.2. There exist positive constants C_1, C_2 such that for any $b_\phi > 1$ and any characteristic $\phi \in \mathcal{C}$ satisfying $\phi(t) \leq b_\phi(\xi_{f,\emptyset}(t) + 1)$ for all $t \geq 0$,*

$$\mathbb{E} \left| e^{-\lambda^* t} Z_f^\phi(t) - W_\infty M_f^\phi(\infty) \right| \leq C_1 b_\phi e^{-C_2 t}, \quad t \geq 0.$$

REMARK 3.6. The constants ω^* in Theorem 3.4 and C_1, C_2 in Theorem 3.5 are explicitly computable from our proof techniques. However, they depend on the Malthusian rate and $\underline{\lambda}$ (see (2.6)) and thus we have not tried to derive an explicit form of these objects.

3.2. Sup-norm convergence of degree distribution for the standard model. We start by studying the model under the following assumption which we refer to as the “standard” model owing to the analogous assumptions for change point methodology in time series:

ASSUMPTION 3.7. *There exist $0 < \gamma < 1$ such that the change point is $\tau = \lfloor n\gamma \rfloor$.*

To simplify notation we will drop $\lfloor \cdot \rfloor$. Recall the sequence of random trees $\{\mathcal{T}_m^\theta : 2 \leq m \leq n\}$. For any $0 < t \leq 1$ and $k \geq 0$, write $D_n(k, \mathcal{T}_{nt}^\theta)$ for the number of vertices with degree k when the tree is of size nt . Fix initializer attachment function f_0 and let $\lambda_0^* = \lambda^*(f_0)$ be as in (2.7). Define the probability mass function $\{p_k^0 : k \geq 0\}$ via (3.1) with (λ_0^*, f_0) in place of (λ^*, f) . As before write f_1 for the attachment function after change point.

Recall the continuous time embedding of $\{\mathcal{T}_m^\theta : 2 \leq m \leq n\}$ into an inhomogeneous branching process $\text{BP}_\theta(\cdot)$ as described in Section 2.3. At the change point of $\text{BP}_\theta(\cdot)$, different individuals have different degrees, and their offspring process after the change point need to be quantified in terms of their degree at the change point. We now introduce some key quantities required in this quantification. Recall $m_{f_1}(\cdot)$ from (2.4). For fixed $k \geq 0$, recall the functions $\mu_{f_1}^{(k)}[0, \cdot]$ from (2.3) and define, for $t \geq 0$, $m_{f_1}^{(k)}(t) := \mathbb{E} \left(\sum_{x \in \text{BP}_{f_1}(t)} \mathbb{1} \{ \xi_{f_1, x}(t - \sigma_x) = k \} \right)$, which denotes the expected number of individuals with k children in $\text{BP}_{f_1}(t)$. It can be checked (using the continuity estimates obtained in Lemmas 7.2 and 7.9) that for any $k \geq 0$, $t \geq 0$, $m_{f_1}^{(k)}(t) = \int_0^t \mathbb{P}(\xi_{f_1}(u) = k) m_{f_1}(t - du)$.

For $\ell, k \geq 0$, define

(3.4)

$$\lambda_\ell(t) = 1 + \int_0^t m_{f_1}(t - s) \mu_{f_1}^{(\ell)}(ds), \quad \lambda_\ell^{(k)}(t) = \mathbb{P}(\xi_{f_1}^{(\ell)}(t) = k - \ell) + \int_0^t m_{f_1}^{(k)}(t - s) \mu_{f_1}^{(\ell)}(ds).$$

Given that an individual is of degree ℓ at the change point, $\lambda_\ell(t)$ (respectively, $\lambda_\ell^{(k)}(t)$) denotes the expected number of descendants (respectively, the expected number of descendants having degree k), including possibly itself, t time units after the change point. Let $\mathcal{P}$ denote the collection of all probability measures on $\mathbb{Z}_+$. For each $a > 0$, consider the functional $\Phi_a : \mathcal{P} \rightarrow \mathcal{P}$ given by

$$(3.5) \quad \Phi_a(\mathbf{p}) = \left(\sum_{\ell=0}^{\infty} p_\ell \lambda_\ell^{(k)}(a) / \sum_{\ell=0}^{\infty} p_\ell \lambda_\ell(a) \right)_{k \geq 0}$$

where $\mathbf{p} = (p_0, p_1, \dots) \in \mathcal{P}$. Write $(\Phi_a(\mathbf{p}))_k$ for the k -th co-ordinate of the above map. Let $\mathbf{p}^i = \mathbf{p}(f_i) := (p_0^i, p_1^i, \dots)$ for $i = 0, 1$ denote the limiting degree distribution for a non-uniform random recursive tree grown with attachment function f_i (i.e. without any change point). Informally, $\Phi_a(\mathbf{p}^0)$ shows how the degree distribution in the continuous time embedding evolves in a units of time after the change point. Corollary 8.2 shows that for each $t > \gamma$, there is a unique $0 < a_t < \infty$ such that

$$(3.6) \quad \sum_{k=0}^{\infty} p_k^0 \left[\int_0^{a_t} m_{f_1}(a_t - s) \mu_{f_1}^{(k)}(ds) \right] = (t - \gamma) / \gamma.$$

Recall the continuous time embedding of $\{\mathcal{T}_m^\theta : 2 \leq m \leq n\}$ in BP_θ described in Section 2.3. Conceptually here, for $t > \gamma$, a_t denotes (in the large n limit) the time required in the continuous time embedding for the process starting at $\mathcal{T}_{n\gamma}$ (i.e. at the change point) to reach size nt . Set $a_t = 0$ for $t \leq \gamma$.

Suppose f_0, f_1 satisfy Assumption 2.4. The following theorem shows that the empirical degree distribution of the (discrete) standard model can be approximated uniformly on compact time intervals after the change point by a deterministic curve, obtained using the continuous time embedding.

THEOREM 3.8. *For each fixed $k \geq 0$ and $s \in [\gamma, 1]$, $\sup_{t \in [\gamma, s]} |D_n(k, \mathcal{T}_{nt}^\theta) / nt - (\Phi_{a_t}(\mathbf{p}^0))_k| \xrightarrow{P} 0$.*

There is a probabilistic way to view the limit. Write α for a_1 .

CONSTRUCTION 3.9 (X_{BC}). *Generate $D \sim \{p_k^0 : k \geq 0\}$. Conditional on $D = k$, generate point process $\xi_{f_1}^{(k)}$ and let $\mathfrak{C} = \xi_{f_1}^{(k)}[0, \alpha]$. Now set $X_{BC} = D + \mathfrak{C}$.*

CONSTRUCTION 3.10 (X_{AC} , Age). (a) *Generate $D \sim \{p_k^0 : k \geq 0\}$. Conditional on $D = k$, generate Age supported on the interval $[0, \alpha]$ with distribution*

$$\mathbb{P}(\text{Age} > u) := \int_0^{\alpha-u} m_{f_1}(\alpha - u - s) d\mu_{f_1}^{(k)}(ds) / \int_0^\alpha m_{f_1}(\alpha - s) d\mu_{f_1}^{(k)}(ds), \quad 0 \leq u \leq \alpha.$$

(b) *Conditional on D and Age, let $X_{AC} = \xi_{f_1}[0, \text{Age}]$, with ξ_{f_1} as in (2.2).*

Conceptually in the above notation, ‘BC’ stands for ‘before change’ and ‘AC’ stands for ‘after change’. Thus (in the large n limit), X_{BC} denotes the final degree (when the tree is of size n) of an individual which had degree D at the change point. X_{AC} denotes the final degree of an individual born $\alpha - \text{Age}$ time units after the change point. Now, slightly abusing notation, let $\theta = (f_0, f_1, \gamma)$. Let D_θ be the integer valued random variable defined as follows: with probability γ , $D_\theta = X_{BC}$ and with probability $1 - \gamma$, $D_\theta = X_{AC}$. The following is a restatement of the convergence result implied by Theorem 3.8 for time $t = 1$.

THEOREM 3.11 (Standard model). *Fix $k \geq 0$ and let $D_n(k)$ denote the number of vertices with degree k in the tree $\mathcal{T}_n^\theta$. Under Assumption 2.4 on the attachment functions f_0, f_1 and Assumption 3.7 on the change point γ , we have that $D_n(k) / n \xrightarrow{P} \mathbb{P}(D_\theta = k)$.*

Write $\mathbf{p}(\theta)$ for the pmf of D_θ . The next result, albeit intuitively reasonable, is non-trivial to prove in the generality of the models considered in the paper.

COROLLARY 3.12. *Assume that $\mathbf{p}^0 \neq \mathbf{p}^1$. Then for any $0 < \gamma < 1$ one has $\mathbf{p}^0 \neq \mathbf{p}(\theta)$. Thus the change point always changes the degree distribution.*

For the following corollary, we say that a random variable X has an exponential tail if there exist positive constants C_1, C_2 such that $\mathbb{P}(X > x) \leq C_1 \exp\{-C_2 x\}$ for all $x \geq 0$. We say X has a power law tail with exponent $\kappa > 0$ if there exist positive constants C_1, C_2 such that $C_1 x^{-\kappa} \leq \mathbb{P}(X > x) \leq C_2 x^{-\kappa}$ for all $x \geq 1$.

COROLLARY 3.13 (Initializer wins under the standard model). *The initializer function f_0 determines the tail behavior of D_θ in the sense that*

- (i) *If in the model without change point using f_0 , the degree distribution has an exponential tail then so does the model with change point irrespective of $\gamma > 0$ and f_1 .*
- (ii) *If in the model without change point using f_0 , the degree distribution has a power law tail with exponent $\kappa > 0$ then so does model with change point irrespective of $\gamma > 0$ and f_1 .*

COROLLARY 3.14 (Maximum degree under the standard model). *Suppose the initializer is linear with $f_0(i) = i + 1 + \alpha$ for $i \geq 0$. For $k \geq 1$, let $M_n(k)$ be the size of the k -th maximal degree. If f_1 satisfies Assumption 2.4 then $M_n(k) / n^{1/(\alpha+2)}$ is a tight collection of random variables bounded away from zero as $n \rightarrow \infty$.*

REMARK 3.15. Corollary 3.14 shows the initializer determines the behavior of the maximal degree in the case of a linear initializer under the standard model. In the absence of a change point, for each fixed $k \geq 1$, $M_n(k)/n^{1/(\alpha+2)} \xrightarrow{d} X_k(\alpha)$ for a non-degenerate random variable $X_k(\alpha)$ with $\mathbb{P}(X_k(\alpha) > 0) = 1$ [36]. Thus the above result shows that irrespective of the second attachment function f_1 , the maximal degree asymptotics for linear preferential attachment remain unaffected. Proof of the above result follows via analogous arguments as [14, Proof of Theorem 2.2] and thus is not provided in this paper.

3.3. *The quick big bang model.* Now consider the case where the change point scales like $o(n)$ i.e. happens “early” in the evolution of the process. We call this version of the process “quick big bang” to fix the idea that the change happens way back in the origin of the process (akin to the “big bang”), but despite this change close to the origin of the process (relative to the entire time scale), the effect of this can be felt and observed all the way till the present via carefully chosen functionals. Let $\{p_k^1 : k \geq 0\}$ be the probability mass function as in (3.1), but using the function f_1 in place of f to obtain λ^* in (2.7) and in (3.1). For $\alpha > 0$ and any non-negative measure μ , let $\hat{\mu}(\alpha) := \int_0^\infty \alpha e^{-\alpha t} \mu(t) dt$. We work under the following assumption.

ASSUMPTION 3.16. $\mathbb{E}(\hat{\xi}_f(\lambda^*) |\log(\hat{\xi}_f(\lambda^*))|) < \infty$.

REMARK 3.17. Assumption 3.16 is, in some sense, the ‘minimal assumption’ required to ensure non-degeneracy of the random variable $W_\infty := \limsup_{t \rightarrow \infty} e^{-\lambda^* t} Z_f(t)$ [37, Proposition 1.1]. In particular, $W_\infty > 0$ almost surely if Assumption 3.16 is satisfied and $W_\infty = 0$ almost surely if Assumption 3.16 fails.

Recall that in the previous section, one of the messages was that the initializer function f_0 determined various macroscopic properties of the degree distribution for the standard model.

THEOREM 3.18 (Initializer loses under the quick big bang). *Suppose $\tau_1 = n^\gamma$ for fixed $0 < \gamma < 1$. If f_0 satisfies Assumption 2.4 and f_1 satisfies Assumptions 2.4, 3.1 and 3.16, the limiting degree distribution **does not** feel the effect of the change point or the initializer attachment function f_0 in the sense that for any fixed $k \geq 0$, $D_n(k)/n \xrightarrow{P} p_k^1$ as $n \rightarrow \infty$.*

REMARK 3.19. The form $\tau_1 := n^\gamma$ was assumed for simplicity. We believe the proof techniques are robust enough to handle any $\tau_1 = \omega_n$, where $\omega_n = o(n)$ and $\omega_n \uparrow \infty$. We defer this to future work.

The next result implies the maximal degree *does* feel the effect of the change point. Instead of proving a general result we consider the following special cases. Let $M_n(1)$ denote the maximal degree in $\mathcal{T}_n^\theta$. Fix two deterministic positive sequences $\{a_n\}_{n \geq 1}, \{b_n\}_{n \geq 1}$ with $a_n, b_n \uparrow \infty$ and $a_n/b_n \rightarrow 0$ as $n \rightarrow \infty$. For a sequence of non-negative random variables $\{M_n\}_{n \geq 1}$, say that $a_n \ll M_n \ll b_n$ with high probability as $n \rightarrow \infty$ if $M_n/a_n \xrightarrow{P} \infty$ and $M_n/b_n \xrightarrow{P} 0$ as $n \rightarrow \infty$.

THEOREM 3.20 (Maximal degree under quick big bang). *Assume $\tau_1 = n^\gamma$ and consider:*

- (a) **Uniform $\rightsquigarrow$ Linear:** *Suppose $f_0(\cdot) \equiv 1$ whilst $f_1(k) = k + 1 + \alpha$ for fixed $\alpha > 0$. Then for any sequence $\omega_n \uparrow \infty$, with high probability as $n \rightarrow \infty$, $n^{\frac{1-\gamma}{2+\alpha}} \log n / \omega_n \ll M_n(1) \ll n^{\frac{1-\gamma}{2+\alpha}} (\log n)^2$.*

- (b) **Linear** $\rightsquigarrow$ **Uniform**: Suppose $f_0(k) = k + 1 + \alpha$ for fixed $\alpha > 0$ whilst $f_1(\cdot) \equiv 1$. Then for any sequence $\omega_n \uparrow \infty$, with high probability as $n \rightarrow \infty$, $n^{\frac{\gamma}{2+\alpha}} \log n / \omega_n \ll M_n(1) \ll n^{\frac{\gamma}{2+\alpha}} (\log n)^2$.
- (c) **Linear** $\rightsquigarrow$ **Linear**: Suppose $f_0(k) = k + 1 + \alpha$ whilst $f_1(k) = k + 1 + \beta$ where $\alpha \neq \beta$. Then $M_n(1) / n^{\eta(\alpha, \beta)}$ is tight and bounded away from zero where

$$\eta(\alpha, \beta) := (\gamma(2 + \beta) + (1 - \gamma)(2 + \alpha)) / (2 + \alpha)(2 + \beta).$$

REMARK 3.21. Writing $\tilde{M}_n := M_n(1) / n^{\eta(\alpha, \beta)}$, in (c) by bounded away from zero we mean $\{1/\tilde{M}_n : n \geq 1\}$ is tight. This result shows that while the initializer does not affect the limiting degree distribution in the quick big bang model (Theorem 3.18), it can influence the maximal degree. It is instructive to compare the above results to the setting without change point. For the uniform $f \equiv 1$ model, it is known [23, 46] that the maximal degree scales like $\log n$ whilst for the linear preferential attachment, the maximal degree scales like $n^{1/(\alpha+2)}$ [36]. Thus, for example, in the ‘Linear $\rightsquigarrow$ Uniform’ case, Theorem 3.18 implies that the limiting degree distribution in this case is the same as that of the *uniform random recursive tree* (URRT) namely Geometric with parameter 1/2; however Theorem 3.20 (b) implies that the maximal degree scales polynomially in n and **not** like $\log n$ as in the URRT.

3.4. *Change point detection.* In the context of the standard model, now consider the issue of change point detection from an observation of the network. Consider any two sequences $h_n \rightarrow \infty, b_n \rightarrow \infty$ satisfying $\log h_n / \log n \rightarrow 0, \log b_n / \log n \rightarrow 0$ as $n \rightarrow \infty$. Define:

$$\hat{T}_n = \inf \left\{ t \geq \frac{1}{h_n} : \sum_{k=0}^{\infty} 2^{-k} \left| \frac{D_n(k, \mathcal{T}_{[nt]}^{\theta})}{nt} - \frac{D_n(k, \mathcal{T}_{[n/h_n]}^{\theta})}{n/h_n} \right| > \frac{1}{b_n} \right\}.$$

The following theorem establishes the consistency of the above estimator.

THEOREM 3.22. Assume that $\mathbf{p}^0 \neq \mathbf{p}^1$. Suppose f_0 satisfies Assumptions 2.4, 3.1 and 3.2, and f_1 satisfies Assumptions 2.4 and 3.16. Then $\hat{T}_n \xrightarrow{P} \gamma$.

4. Discussion. (i) **Random recursive trees:** Random recursive trees have now been studied for decades [22, 24, 27, 35, 44]. For specific examples such as the uniform attachment or the linear attachment model with $f(i) := i + 1$, one can use the seminal work of Janson [32] via a so-called “super ball” argument to obtain functional central limit theorems for the degree distribution. Obtaining quantitative error bounds let alone weak convergence results in the general setting considered in this paper is much more involved. Regarding proof techniques, we proceed via embedding the discrete time models into continuous time branching processes and then using martingale/renewal theory arguments for the corresponding continuous time objects; this approach goes back all the way to [3]. Limit results for the corresponding CTBPs in the setting of interest for this paper were developed in [29, 30, 37]. One contribution of this work is to derive quantitative versions for this convergence, a topic less explored, but required to answer questions regarding statistical estimation of the change point. In the context of growing random trees (without change point) with either uniform or linear attachment functions, understanding the effect of the initial seed graph and in particular constructing algorithms to estimate the root (the so-called “Adam problem”) has inspired intense activity over the last decade. See for example [18, 19, 21] for more details.

(ii) **General change point:** Change point detection, especially in the context of univariate time series, has matured into a vast field, see [17, 20]. Even in this context, consistent

estimation especially in the setting of multiple change points is non-trivial and requires specific assumptions see e.g. [48]; in the context of econometric time series see [5–7]; for applications in the biological sciences see [41, 49]. The only pre-existing work on change point in the context of growing networks formulated in this paper that we are aware of was carried out in [14] where one assumed linear attachment functions. Regarding the estimator proposed in this paper in Theorem 3.22, we do not believe the estimator is “optimal” in terms of rates of convergence. These issues are deferred to future work.

5. Overview of the proofs. The rest of the paper proves the main results. Section 6 lays out some preliminary constructions and estimates used subsequently in the paper. Section 7 deals with the continuous time version of the change point model analyzed for a fixed time a after the change point. Theorem 7.1 proved in this section estimates, for a general characteristic $\phi \in \mathcal{C}$, the L^1 -error in approximating the aggregate ϕ -score at time a of all individuals born after the change point with a weighted linear combination of the degree counts at the change point. This estimate, apart from directly yielding a law of large numbers (see second part of Theorem 7.1), turns out to be crucial in most subsequent proofs.

The estimates derived in Section 7 are then used in Section 8 to analyze the standard model and prove the main theorems in this setting (Theorems 3.8 and 3.11) as well as Corollary 3.13. Corollary 3.12 follows directly from Lemma 11.3 and requires an in-depth analysis of the fluid limits derived in Theorem 3.8 and is postponed to Section 11.

Section 9 contains proofs of the quick big bang model. We note here that all the estimates obtained in Sections 7 and 8 to analyze the model for a fixed time a after the change point explicitly exhibit the dependence on a . This turns out to be crucial in Section 9 where we take $a = \eta_0 \log n$ and the estimates above still hold if η_0 is sufficiently small. Roughly speaking, we partition the interval $[T_{n'}, T_n]$ into finitely many subintervals of size at most $\eta_0 \log n$ and ‘bootstrap’ the estimates obtained in Sections 7 and 8 to prove Theorem 3.18.

In Section 10, we prove Theorems 3.4 and 3.5. We conclude in Section 11 with the proof of Theorem 3.22 on the change point detection estimator.

6. Initial constructions. This section is devoted to some preliminary constructions and estimates that will then be repeatedly used in the proofs. The first set of lemmas deal with properties of *linear* preferential attachment and an important class of offspring processes associated to it.

DEFINITION 6.1 (Rate ν Affine κ PA model). *Fix $\nu > 0, \kappa \geq 0$. A branching process with attachment function $f(i) = \nu(i + 1) + \kappa, i \geq 0$, will be called a linear PA branching process with rate ν and affine parameter κ . Denote this as $\{\text{PA}_{\nu, \kappa}(t) : t \geq 0\}$.*

DEFINITION 6.2 (Rate ν Yule process). *The offspring process $\xi_{\nu, 0}(t)$ associated with a $\text{PA}_{\nu, 0}(\cdot)$ process is called a rate ν Yule process. Thus, the rate of birth of new individuals in a Yule process is proportional to the size of the current population. We write $\{Y_\nu(t) : t \geq 0\}$ for this process.*

LEMMA 6.3 ([40, Section 2.5]). *Fix $t > 0$ and rate $\nu > 0$. Then $Y_\nu(t)$ has a Geometric distribution with parameter $p = e^{-\nu t}$. Precisely, $\mathbb{P}(Y_\nu(t) = k) = e^{-\nu t}(1 - e^{-\nu t})^{k-1}, k \geq 1$. The process $\{Y_\nu(t) \exp(-\nu t) : t \geq 0\}$ is an $\mathbb{L}^2$ bounded martingale and thus $\exists W > 0$ such that $Y_\nu(t) \exp(-\nu t) \xrightarrow{\text{a.s.}} W$. Further $W \sim \exp(1)$.*

Next we derive moment bounds for the attachment point processes for linear preferential attachment.

LEMMA 6.4. *Fix $\nu > 0, \kappa \geq 0$. Let $\xi_{\nu, \kappa}(t)$ be the offspring process associated with a $\text{PA}_{\nu, \kappa}(\cdot)$ process. Then with respect to the natural filtration the following two processes are martingales which give the proceeding two expectations:*

$$M_1(t) := e^{-\nu t} \xi_{\nu, \kappa}(t) - \frac{\nu + \kappa}{\nu} (e^{\nu t} - 1), \quad t \geq 0$$

$$M_2(t) := e^{-2\nu t} \xi_{\nu, \kappa}(t)^2 - \int_0^t (2\kappa + 3\nu) \xi_{\nu, \kappa}(s) e^{-2\nu s} ds - \frac{\nu + \kappa}{2\nu} (1 - e^{-2\nu t}), \quad t \geq 0.$$

$$\mathbb{E} \xi_{\nu, \kappa}(t) = \frac{\nu + \kappa}{\nu} (e^{\nu t} - 1), \text{ and } \mathbb{E} (\xi_{\nu, \kappa}(t))^2 = \frac{(2\kappa + 3\nu)(\nu + \kappa)}{2\nu^2} (e^{\nu t} - 1)^2 + \frac{\nu + \kappa}{2\nu} (e^{2\nu t} - 1).$$

PROOF. We sketch the proof. Let $\mathcal{F}(t)$ be the natural filtration corresponding to the continuous time branching process with attachment function f . Note that $\xi_{\nu, \kappa}(t) \leadsto \xi_{\nu, \kappa}(t) + 1$ at rate $\nu(\xi_{\nu, \kappa}(t) + 1) + \kappa$. This can be used to check $\mathbb{E}[dM_1(t)|\mathcal{F}(t)] = 0$ showing $M_1(t)$ is a martingale. Similarly, $\xi_{\nu, \kappa}(t)^2 \leadsto \xi_{\nu, \kappa}(t)^2 + 2\xi_{\nu, \kappa}(t) + 1$ at rate $\nu(\xi_{\nu, \kappa}(t) + 1) + \kappa$. This expression can similarly be used to check $M_2(t)$ is a martingale. The first expectation claimed in the lemma follows immediately by setting the expectation of $M_1(t)$ equal to zero. The second expectation follows by computing the expectation of $M_2(t)$ and then using the expectation of $\xi_{\nu, \kappa}(t)$. ■

We now derive expressions for moments of the process $\text{PA}_{\nu, \kappa}$. To simplify notation, when possible we will suppress dependence on ν, κ and write the above as $\text{PA}(\cdot)$. Note the proof of Proposition 6.5 is similar to the proof of Lemma 6.4 and is therefore omitted.

PROPOSITION 6.5. *Fix $\nu > 0, \kappa \geq 0$. With respect to the natural filtration the following processes are martingales:*

$$M_1(t) := e^{-(2\nu + \kappa)t} (|\text{PA}_{\nu, \kappa}(t)| - 1) - \frac{\nu + \kappa}{2\nu + \kappa} (1 - e^{-(2\nu + \kappa)t}), \quad t \geq 0$$

$$M_2(t) := (|\text{PA}_{\nu, \kappa}(t)| - 1)^2 - \int_0^t ((4\nu + 2\kappa)(|\text{PA}_{\nu, \kappa}(s)| - 1)^2 + (4\nu + 3\kappa)(|\text{PA}_{\nu, \kappa}(s)| - 1) + (\nu + \kappa)) ds, \quad t \geq 0.$$

In particular, for any fixed $a > 0$, $\exists C$ (dependent on ν and κ) such that for $0 \leq t \leq a$

$$(6.1) \quad \mathbb{E}(|\text{PA}_{\nu, \kappa}(t)|) - 1 \leq C e^{(2\nu + \kappa)a} t; \quad \mathbb{E}(|\text{PA}_{\nu, \kappa}(t)| - 1)^2 \leq C e^{(4\nu + 2\kappa)a} t.$$

Recall the class of characteristics $\mathcal{C}$ defined in (3.3), the Malthusian rate of growth λ^* and the mean measure of the offspring process μ_f . Let $m^* := \int_{\mathbb{R}_+} u e^{-\lambda^* u} \mu_f(du)$. For any fixed characteristic $\chi \in \mathcal{C}$ and any $\alpha > 0$, define,

$$\hat{\chi}(\alpha) := \int_0^\infty \alpha e^{-\alpha t} \chi(t) dt, \quad \hat{\mu}_f(\alpha) := \int_0^\infty \alpha e^{-\alpha t} \mu_f(t) dt.$$

It can be checked that for any $\alpha > 0$, recalling $\hat{\rho}$ from Assumption 2.4 (ii), $\hat{\rho}(\alpha) = \hat{\mu}_f(\alpha) = \int_0^\infty e^{-\alpha t} \mu_f(dt)$. Recall the definitions of $Z_f^\chi(t)$, $M_f^\chi(t)$, $Z_f(t)$ and $M_f(t) = \mathbb{E}(e^{-\lambda^* t} Z_f(t))$ from Section 3.1. The first part of the following lemma is a consequence of [37, Theorem 6.3] and the second part follows from [37, Theorem 5.4 and Corollary 3.3].

LEMMA 6.6. (i) *Under Assumption 2.4 (ii), for any characteristic $\chi \in \mathcal{C}$, $Z_f^\chi(t)/Z_f(t) \xrightarrow{a.s.} \mathbb{E}(\hat{\chi}(\lambda^*))$.*
(ii) *Under Assumptions 2.4 and 3.16, there exists a strictly positive random variable W_∞ with $\mathbb{E}(W_\infty) = 1$ such that for characteristics $\chi \in \mathcal{C}$, $e^{-\lambda^* t} Z_f^\chi(t) \xrightarrow{a.s., \mathbb{L}^1} W_\infty \mathbb{E}(\hat{\chi}(\lambda^*)) / \lambda^* m^*$.*

7. Change point model run for fixed time a after change point: point-wise convergence for general characteristics. In this section we consider growing the tree (in continuous time) using attachment function f_0 till it reaches size γn , and using the second attachment function f_1 for a constant time a after the change point i.e. f_1 is used for $t \in [T_{\gamma n}, T_{\gamma n} + a]$. We will also assume throughout this section that f_0, f_1 satisfy Assumption 2.4. We count vertices *born after the change point* according to a general characteristic $\phi \in \mathcal{C}$ (defined in (3.3)) and prove a law of large numbers for this aggregate ϕ -score at time a as $n \rightarrow \infty$ (see Theorem 7.1). This will be a key tool in the rest of the paper. For notational convenience we consider the change point as time $t = 0$ (i.e. $t = s$ corresponds to actual time $T_{\gamma n} + s$ for any $s \in [0, a]$). Recall the continuous time embedding of the model in an inhomogeneous branching process $\text{BP}_\theta(\cdot)$ described in Section 2.3. For $t \geq 0$, we will write $\text{BP}_n(t) := \text{BP}_\theta(t + T_{\gamma n})$ to denote the branching process at time t (i.e. time t after the change point).

7.1. Notation. Let λ_i^* be the Malthusian parameter for the branching process with attachment function $f_i, i = 0, 1$. For $k \geq 0, i = 0, 1$, recall $\xi_{f_i}^{(k)}(\cdot)$ and $\mu_{f_i}^{(k)}(\cdot)$ from (2.3) with f_i in place of f . For $0 \leq s \leq t$, let $\xi_{f_i}^{(k)}[s, t] := \xi_{f_i}^{(k)}(t) - \xi_{f_i}^{(k)}(s)$ and $\mu_{f_i}^{(k)}[s, t] := \mu_{f_i}^{(k)}(t) - \mu_{f_i}^{(k)}(s)$. For the branching process (without change point) with attachment function f_1 , and for any characteristic ϕ , recall $Z_{f_1}^\phi(t), m_{f_1}^\phi(t), Z_{f_1}$ from Section 3.1 (just after (3.3)). Let $v_{f_1}^\phi(t) = \text{Var}(Z_{f_1}^\phi(t))$. Recall the class of characteristics defined in (3.3). For $\phi \in \mathcal{C}$, an easy computation implies there exists non-random $b_\phi > 0$ such that $Z_{f_1}^\phi(t) \leq 2b_\phi Z_{f_1}(t)$ for every $t \geq 0$. Moreover, by Assumption 2.4(i) on f_1 , $Z_{f_1}(\cdot)$ is stochastically dominated by $|\text{PA}_{2C,0}(\cdot)|$ (see Definition 6.1), in the sense that there exists a coupling $(Z_{f_1}(\cdot), |\text{PA}_{2C,0}(\cdot)|)$ satisfying $Z_{f_1}(t) \leq |\text{PA}_{2C,0}(t)|$ for all $t \geq 0$. Combining these observations and using (6.1), we obtain for any $\phi \in \mathcal{C}$,

$$(7.1) \quad \sup_{t \in [0, a]} m_{f_1}^\phi(t) \leq 2b_\phi \mathbb{E}(Z_{f_1}(a)) \leq C_1 e^{C_2 a}, \quad \sup_{t \in [0, a]} v_{f_1}^\phi(t) \leq 4b_\phi^2 \mathbb{E}(Z_{f_1}^2(a)) \leq C_1 e^{C_2 a}$$

where C_1, C_2 depend on ϕ but not on a . For $\phi \in \mathcal{C}$, let $Z_n^\phi(a)$ denote the aggregate ϕ -score at time a (see Section 3, just after (3.3)) of all individuals born after the change point, namely

$$Z_n^\phi(a) := \sum_{x \in \text{BP}_n(a) \setminus \text{BP}_n(0)} \phi_x(a - \sigma_x).$$

For $k \geq 0, 0 \leq t \leq a$, let $\lambda_k^\phi(t) = \int_0^t m_{f_1}^\phi(t-s) \mu_{f_1}^{(k)}(ds)$ denote the expected aggregate ϕ -score at time t of all descendants that are born in $(0, t]$ to a vertex with degree k at time 0. For $k \geq 0$, let $D_n(k, 0)$ denote the number of vertices with degree k at the change point time 0. Let $\mathcal{F}_n(0)$ denote the σ -field containing the information on the entire branching process till time $T_{\gamma n}$, the change point.

The following is the main result proved in this section.

THEOREM 7.1. *Suppose f_0, f_1 satisfy Assumption 2.4. Fix any $\phi \in \mathcal{C}$. There exist deterministic positive constants $C, C' < \infty$ such that for every $a > 0$ and $n \geq 2$,*

$$\mathbb{E} \left[\left| Z_n^\phi(a) - \sum_{k=0}^{\infty} D_n(k, 0) \lambda_k^\phi(a) \right| \middle| \mathcal{F}_n(0) \right] \leq C e^{C' a} \sqrt{n}.$$

Furthermore, as $n \rightarrow \infty$, $Z_n^\phi(a)/n \xrightarrow{\text{P}} \gamma \sum_{k=0}^{\infty} p_k^0 \lambda_k^\phi(a)$.

7.2. Definitions. In this section we define constructs for the branching process $\text{BP}_n(\cdot)$ which will be used in the proof of Theorem 7.1 (and are motivated by the proof outline in Section 7.3). For notational simplicity, since a is fixed in this section, we will write Z_n^ϕ for $Z_n^\phi(a)$ and λ_k^ϕ for $\lambda_k^\phi(a)$. **For the rest of this section, unless specified otherwise, we always work conditional on $\mathcal{F}_n(0)$** so that expectation operations such as $\mathbb{E}(\cdot)$ and $\text{Var}(\cdot)$ for the associated branching process statistics post change point in the ensuing results mean $\mathbb{E}(\cdot|\mathcal{F}_n(0))$ and $\text{Var}(\cdot|\mathcal{F}_n(0))$. Divide the interval $[0, a] := \cup_{i=0}^{n^\delta-1} [ian^{-\delta}, (i+1)a n^{-\delta}]$ into subintervals of size $an^{-\delta}$. We will eventually take limits as $\delta \rightarrow \infty$.

(i) **System at change point:** Define the filtration $\{\mathcal{F}_n(t) : t \geq 0\} := \{\sigma(\text{BP}_n(t)) : t \geq 0\}$ (information till t time units after change point). For fixed $k \geq 0$, we write $\mathcal{D}_n(k, t)$ for the set of vertices with degree k at time t and let $D_n(k, t) := |\mathcal{D}_n(k, t)|$. The initial set $\mathcal{D}_n(k, 0)$ which arose from the pre-change point dynamics will play a special role. Label the vertices in $\mathcal{D}_n(k, 0)$ in the order they were born into $\text{BP}_n(0)$ as $\mathcal{D}_n(k, 0) := \{v_1^{(k)}, v_2^{(k)}, \dots, v_{D_n(k, 0)}^{(k)}\}$. Let $\mathcal{D}_n(0) := \cup_{k \geq 0} \mathcal{D}_n(k, 0)$.

(ii) **Descendants in small intervals:** For $0 \leq i \leq n^\delta - 1$, $k \geq 0$ and vertex $v_j^{(k)} \in \mathcal{D}_n(k, 0)$, let $\mathcal{V}_n^{(k)}(i, j)$ denote the set of children born in the interval $[ian^{-\delta}, (i+1)an^{-\delta}]$ to $v_j^{(k)}$. Let $N_n^{(k)}(i, j) := |\mathcal{V}_n^{(k)}(i, j)|$ be the number of such vertices. Write $N_n^{(k)}(i) := \sum_{j=1}^{D_n(k, 0)} N_n^{(k)}(i, j)$ for the total number of children of vertices that were of degree k at the change point, born in the time interval $[ian^{-\delta}, (i+1)an^{-\delta}]$.

(iii) **Good and bad vertices:** Call a vertex in $\mathcal{V}_n^{(k)}(i, j)$ a *good* vertex if it does **not** give birth to any children by time $(i+1)an^{-\delta}$. Let $\tilde{\mathcal{V}}_n^{(k)}(i, j) \subseteq \mathcal{V}_n^{(k)}(i, j)$ denote the set of good children of $v_j^{(k)}$ born in the interval $[ian^{-\delta}, (i+1)an^{-\delta}]$. Let $\tilde{N}_n^{(k)}(i, j) := |\tilde{\mathcal{V}}_n^{(k)}(i, j)|$ be the number of such vertices. Let $\tilde{N}_n^{(k)}(i) := \sum_{j=1}^{D_n(k, 0)} \tilde{N}_n^{(k)}(i, j)$ be the total number of *good* children of vertices which originally had degree k at the change point born in the interval $[ian^{-\delta}, (i+1)an^{-\delta}]$. Let $\mathcal{B}_n^{(k)}(i, j) := \mathcal{V}_n^{(k)}(i, j) \setminus \tilde{\mathcal{V}}_n^{(k)}(i, j)$ be the collection of bad children namely those in $\mathcal{V}_n^{(k)}(i, j)$ who **have** reproduced by time $(i+1)an^{-\delta}$. Let $B_n^{(k)}(i, j) = |\mathcal{B}_n^{(k)}(i, j)|$. Let $\mathcal{R}_n^{(k)}(i, j)$ be the set of descendants of vertices in $\mathcal{B}_n^{(k)}(i, j)$ (excluding the parent vertices in $\mathcal{B}_n^{(k)}(i, j)$) born in the time interval $[ian^{-\delta}, (i+1)an^{-\delta}]$ and let $R_n^{(k)}(i, j) := |\mathcal{R}_n^{(k)}(i, j)|$.

(iv) **Vertices counted by a characteristic:** For $0 \leq i \leq n^\delta - 1$, $k \geq 0$, $1 \leq j \leq D_n(k, 0)$ and $x \in \mathcal{V}_n^{(k)}(i, j)$, let $Z_n^{(k), \phi}(i, j, x)$ be the aggregate ϕ -score (defined in Section 3, just after (3.3)) of x and its descendants at time a . More precisely, denoting the set of x and its descendants at time a by $\mathcal{Q}_n^{(k)}(i, j, x)$, $Z_n^{(k), \phi}(i, j, x) := \sum_{y \in \mathcal{Q}_n^{(k)}(i, j, x)} \phi_y(a - \sigma_y)$. Write $Z_n^{(k), \phi} = \sum_{j=1}^{D_n(k, 0)} \sum_{i=0}^{n^\delta-1} \sum_{x \in \mathcal{V}_n^{(k)}(i, j)} Z_n^{(k), \phi}(i, j, x)$ for the aggregate ϕ -score at time a of all vertices of degree k at the change point along with their descendants at time a . Thus, $Z_n^\phi = \sum_{k=0}^\infty Z_n^{(k), \phi}$. For $k \geq 0$, let $\tilde{Z}_n^{(k), \phi} = \sum_{j=1}^{D_n(k, 0)} \sum_{i=0}^{n^\delta-1} \sum_{x \in \tilde{\mathcal{V}}_n^{(k)}(i, j)} Z_n^{(k), \phi}(i, j, x)$ denote the aggregate ϕ -score at time a of all *good* vertices born in $[0, a]$ which are descendants of vertices of degree k at the change point.

(v) **Technical conditioning tool:** For $0 \leq s < t \leq \infty$, let $\mathcal{G}[s, t]$ be the σ -field generated by the biographies of all individuals born in $[s, t]$ over the same time interval. Formally,

$$\mathcal{G}[s, t] := \sigma(\{s \leq \sigma_x \leq t - u\} \cap \{\xi_{x, f_1}(u) = k\}, x \in \mathcal{I} \setminus \mathcal{D}_n(0), u \in [0, t - s], k \in \mathbb{Z}_+\).$$

Moreover, let $\mathcal{G}_0$ denote the σ -field generated by the entire biographies of the individuals at time 0, namely,

$$\mathcal{G}_0 := \sigma(\{\xi_{x, f_1}(u) = k\}, x \in \mathcal{D}_n(0), u \in [0, \infty), k \in \mathbb{Z}_+).$$

The following σ -field will play a crucial role in controlling fluctuations of aggregate ϕ -scores of good vertices

$$\mathcal{G}_n := \sigma\left(\mathcal{G}_0 \bigcup_{0 \leq i \leq n^\delta - 1} \mathcal{G}[ian^{-\delta}, (i+1)an^{-\delta}]\right).$$

7.3. Proof of Theorem 7.1. We first give an outline of the proof. We discretize the time interval $[0, a]$ into small subintervals of the form $\{[ian^{-\delta}, ((i+1)a)n^{-\delta}]\}_{0 \leq i \leq n^\delta - 1}$, for $\delta > 0$. For an individual born in the interval $[ian^{-\delta}, ((i+1)a)n^{-\delta}]$, the final aggregate ϕ -score of its descendants at time a (counting the parent) is estimated by the expected aggregate ϕ -score of a degree zero parent and its descendants tracked till time $a - ((i+1)a)n^{-\delta}$, which equals $m_{f_1}^\phi\left(a - \frac{(i+1)a}{n^\delta}\right)$. For this approximation to be valid, we need to show that the total number of bad vertices defined above is small in an appropriate sense. In fact, a number of lemmas below are ‘continuity estimates’ that validate this discrete approximation. These lemmas are very general and are also used in subsequent sections. As the expected number of children born in the time interval $[ian^{-\delta}, ((i+1)a)n^{-\delta}]$ to a vertex having degree k at time zero equals $\mu_{f_1}^{(k)}\left[\frac{ia}{n^\delta}, \frac{(i+1)a}{n^\delta}\right]$, and there are $D_n(k, 0)$ degree k vertices at time zero, Z_n^ϕ is thus estimated by the Riemann sum $\sum_{k=0}^\infty D_n(k, 0) \sum_{i=0}^{n^\delta - 1} m_{f_1}^\phi\left(a - \frac{(i+1)a}{n^\delta}\right) \mu_{f_1}^{(k)}\left[\frac{ia}{n^\delta}, \frac{(i+1)a}{n^\delta}\right]$. This Riemann sum can then be shown to be close to $\sum_{k=0}^\infty D_n(k, 0) \lambda_k^\phi$.

We fix a characteristic $\phi \in \mathcal{C}$ throughout the proof. The main tools used are Lemmas 7.10, 7.11 below. We will need a number of supporting results which we now embark upon. For the rest of this section we write $C_1, C_2, C_3, C_4, C, C', c, a_0$ for generic non-random constants which are independent of a, n, δ, k , whose values might change between lines and inequalities. We start with a technical lemma controlling the number of children a vertex with degree k at time 0 can produce within a fixed interval.

LEMMA 7.2. *For any $k \geq 0$ and any interval $[b, b + \eta] \subseteq [0, a]$,*

$$\mathbb{E}\left[\xi_{f_1}^{(k)}[b, b + \eta]\right] \leq C_1 e^{C_2 a} (k + 1)\eta, \quad \mathbb{E}\left[\xi_{f_1}^{(k)}[b, b + \eta]^2\right] \leq C_3 e^{C_4 a} \{(k + 1)^2 \eta^2 + (k + 1)\eta\}.$$

PROOF. By Assumption 2.4(i), the process $\{U(t) := \xi_{f_1}^{(k)}(t/C) : t \geq 0\}$ is stochastically dominated by the offspring process $\{P_k(t) : t \geq 0\}$ of a rate 1 affine k PA model (see Definition 6.1), namely a point process constructed using attachment function $f^{(k)}(i) = k + 1 + i$ for $i \geq 0$ with initial condition $P_k(0) = 0$. From the first moment computed in Lemma 6.4 (with $\nu = 1$ and $\kappa = k$) we obtain $\mathbb{E}(P_k(t)) = (1 + k)(e^t - 1)$. We show how to use the first moment of $P_k(\cdot)$ to obtain the first assertion in the lemma. The second assertion follows from the same argument using the second moment of $P_k(\cdot)$ which is also obtained from Lemma 6.4. Conditioning on $\xi_{f_1}^{(k)}(b)$ and using the Markov property we get,

$$(7.2) \quad \mathbb{E}\xi_{f_1}^{(k)}[b, b + \eta] = \sum_{d=0}^\infty \mathbb{P}\left(\xi_{f_1}^{(k)}(b) = d\right) \mathbb{E}\xi_{f_1}^{(k+d)}(\eta)$$

Now for any fixed $k \geq 0$ and $t \leq a$, using domination by the corresponding linear PA process, we get

$$(7.3) \quad \mathbb{E}[\xi_{f_1}^{(k)}(t)] \leq \mathbb{E}(P_k(tC)) = e^{tC}(1 + k)(1 - e^{-tC}) \leq e^{Ca}C(1 + k)t.$$

Using this bound twice in (7.2) gives,

$$\mathbb{E}\xi_{f_1}^{(k)}[b, b + \eta] \leq C e^{Ca} \eta \sum_{d=0}^\infty \mathbb{P}\left(\xi_{f_1}^{(k)}(b) = d\right) (1 + k + d) = C e^{Ca} \eta (1 + k + \mathbb{E}(\xi_{f_1}^{(k)}(b))) \leq C' e^{C'' a} (k + 1)\eta$$

where C', C'' are constants that do not depend on k, a . This completes the proof. $\blacksquare$

Recall from Section 7.2 (ii) the random variable $N_n^{(k)}(i, j) \stackrel{d}{=} \xi_{f_1}^{(k)}[i a n^{-\delta}, (i+1) a n^{-\delta}]$. Using Lemma 7.2 now gives the following result.

COROLLARY 7.3. *For all $1 \leq j \leq D_n(k, 0)$, $\mathbb{E}(N_n^{(k)}(i, j)) \leq C_1 e^{C_2 a} (k+1) n^{-\delta}$ and $\mathbb{E}[N_n^{(k)}(i, j)^2] \leq C_3 e^{C_4 a} \{(k+1)^2 n^{-2\delta} + (k+1) n^{-\delta}\}$.*

The next Lemma bounds the number of “bad” vertices and their descendants born within small intervals.

LEMMA 7.4. *For any k, i, j , $\mathbb{E}(R_n^{(k)}(i, j)) \leq C_1 e^{C_2 a} (k+1) n^{-2\delta}$.*

PROOF. For every child $u \in \mathcal{V}_n^{(k)}(i, j)$, write $\text{BP}(\cdot; u)$ for the branching process lineage emanating from u . Conditional on $\mathcal{V}_n^{(k)}(i, j)$, using Assumption 2.4(i) on f_1 , generate a collection of independent rate C affine 0 linear PA branching processes (see Definition 6.1) $\{Y_\ell : 1 \leq \ell \leq |\mathcal{V}_n^{(k)}(i, j)|\}$ such that $|\text{BP}(\cdot; u)| \leq |Y_\ell(\cdot)|$. Now note that $X_\ell(t) := |Y_\ell(t)| - 1$ is the number of descendants of the root for this branching process by time t . Using this construction we have the trivial inequality $R_n^{(k)}(i, j) \leq \sum_{\ell=1}^{N_n^{(k)}(i, j)} X_\ell[0, a n^{-\delta}]$. This implies

$$\mathbb{E}(R_n^{(k)}(i, j)) \leq \mathbb{E}(N_n^{(k)}(i, j)) \mathbb{E}\left(X_1[0, a n^{-\delta}]\right).$$

The lemma follows from this bound upon using Corollary 7.3 for moments of $N_n^{(k)}(i, j)$ and (6.1) for moments of $X_1[0, a n^{-\delta}]$. $\blacksquare$

The next lemma bounds fluctuations of aggregate ϕ -scores of *good* descendants of ancestors who were of degree k at the change point.

LEMMA 7.5. *For any $k \geq 0$, $\text{Var}\left(\tilde{Z}_n^{(k), \phi}\right) \leq C e^{C' a} \{(k+1)^2 n^{-\delta} + (k+1)\} D_n(k, 0)$.*

PROOF. By construction we have

$$(7.4) \quad \text{Var}\left(\tilde{Z}_n^{(k), \phi}\right) = \text{Var}\left(\sum_{j=1}^{D_n(k, 0)} \sum_{i=0}^{n^\delta-1} \sum_{x \in \tilde{\mathcal{V}}_n^{(k)}(i, j)} Z_n^{(k), \phi}(i, j, x)\right)$$

$$(7.5) \quad = D_n(k, 0) \text{Var}\left(\sum_{i=0}^{n^\delta-1} \sum_{x \in \tilde{\mathcal{V}}_n^{(k)}(i, 1)} Z_n^{(k), \phi}(i, 1, x)\right).$$

We analyze the variance term on the right by first conditioning on $\mathcal{G}_n$. Note that,

$$(7.6) \quad \mathbb{E}\left[\text{Var}\left(\sum_{i=0}^{n^\delta-1} \sum_{x \in \tilde{\mathcal{V}}_n^{(k)}(i, 1)} Z_n^{(k), \phi}(i, 1, x) \middle| \mathcal{G}_n\right)\right] = \mathbb{E}\left[\sum_{i=0}^{n^\delta-1} \tilde{N}_n^{(k)}(i, 1) v_{f_1}^\phi\left(a - (i+1) a n^{-\delta}\right)\right] \\ \leq C_1 e^{C_2 a} (k+1) n^{-\delta} n^\delta \mathbb{E}(Z_{f_1}^2(a)) \leq C e^{C' a} (k+1).$$

The first equality comes from noting that $\tilde{\mathcal{V}}_n^{(k)}(i, 1)$ is $\mathcal{G}_n$ measurable, the collection $\{Z_n^{(k), \phi}(i, 1, x) \mid x \in \tilde{\mathcal{V}}_n^{(k)}(i, 1), 1 \leq i \leq n^\delta - 1\}$ are conditionally independent given $\mathcal{G}_n$ and further, conditionally on $\mathcal{G}_n$, for each $0 \leq i \leq n^\delta - 1$ and $x \in \tilde{\mathcal{V}}_n^{(k)}(i, 1)$, $Z_n^{(k), \phi}(i, 1, x)$ is distributed as $Z_{f_1}^\phi(a - (i+1) a n^{-\delta})$, since x has no children by time $(i+1) a n^{-\delta}$. The second

inequality follows by using Corollary 7.3 for $N_n^{(k)}(i, 1)$ and (7.1). Similarly

$$(7.7) \quad \text{Var} \left(\mathbb{E} \left[\sum_{i=0}^{n^\delta-1} \sum_{x \in \tilde{\mathcal{V}}_n^{(k)}(i,1)} Z_n^{(k),\phi}(i, 1, x) \middle| \mathcal{G}_n \right] \right) = \text{Var} \left(\sum_{i=0}^{n^\delta-1} \tilde{N}_n^{(k)}(i, 1) m_{f_1}^\phi \left(a - (i+1) a n^{-\delta} \right) \right) \\ \leq 4c^2 \left(\mathbb{E}(Z_{f_1}(a)) \right)^2 \sum_{i=0}^{n^\delta-1} \mathbb{E} \left[\left(\tilde{N}_n^{(k)}(i, 1) \right)^2 \right] \leq C e^{C'a} \left((k+1)^2 n^{-\delta} + (k+1) \right).$$

Here we use Corollary 7.3 in the second inequality. Using (7.6) and (7.7) to bound the variance term in the right of (7.5) completes the proof. $\blacksquare$

The next lemma provides tight bounds on expectations of aggregate ϕ -scores of descendants of good vertices. Recall $\mu_{f_1}^{(k)}$ denotes the mean measure for the offspring process of a vertex which had degree k at the change point.

LEMMA 7.6. *For any $k \geq 0$,*

$$\varepsilon_n := \left| \mathbb{E} \left[\tilde{Z}_n^{(k),\phi} \right] - D_n(k, 0) \sum_{i=0}^{n^\delta-1} m_{f_1}^\phi \left(a - (i+1) a n^{-\delta} \right) \mu_{f_1}^{(k)} \left[i a n^{-\delta}, (i+1) a n^{-\delta} \right] \right| \\ \leq C e^{C'a} (k+1) D_n(k, 0) n^{-\delta}.$$

PROOF. First note,

$$\mathbb{E} \left[\tilde{Z}_n^{(k),\phi} \right] = \sum_{i=0}^{n^\delta-1} \sum_{j=1}^{D_n(k,0)} \mathbb{E} \left[\sum_{x \in \tilde{\mathcal{V}}_n^{(k)}(i,j)} Z_n^{(k),\phi}(i, j, x) \right] \\ = \sum_{i=0}^{n^\delta-1} D_n(k, 0) \mathbb{E} \left[\mathbb{E} \left[\sum_{x \in \tilde{\mathcal{V}}_n^{(k)}(i,1)} Z_n^{(k),\phi}(i, 1, x) \middle| \mathcal{G}_n \right] \right] \\ = D_n(k, 0) \sum_{i=0}^{n^\delta-1} m_{f_1}^\phi \left(a - (i+1) a n^{-\delta} \right) \mathbb{E} \left[\tilde{N}_n^{(k)}(i, 1) \right].$$

Here the third equality follows from the fact that $\tilde{\mathcal{V}}_n^{(k)}(i, 1)$ is $\mathcal{G}_n$ measurable and for fixed i and, conditional on $\mathcal{G}_n$, for each $x \in \tilde{\mathcal{V}}_n^{(k)}(i, 1)$, $Z_n^{(k),\phi}(i, j, x) \stackrel{d}{=} Z_{f_1}^\phi \left(a - (i+1) a / n^\delta \right)$. Applying equation (7.1), the error term ε_n in the statement of the lemma can be bounded as,

$$(7.8) \quad \varepsilon_n \leq 2c D_n(k, 0) m_{f_1}^\phi(a) \sum_{i=0}^{n^\delta-1} \mathbb{E} \left[N_n^{(k)}(i, 1) - \tilde{N}_n^{(k)}(i, 1) \right].$$

Next using that the total number of descendants of bad vertices in an interval bounds the number of bad vertices in this interval since each bad vertex has at least one child, we get using Lemma 7.4,

$$0 \leq \mathbb{E} \left[N_n^{(k)}(i, 1) - \tilde{N}_n^{(k)}(i, 1) \right] = \mathbb{E}[B_n^{(k)}(i, 1)] \leq \mathbb{E}[R_n^{(k)}(i, j)] \leq C_1 e^{C_2 a} (k+1) n^{-2\delta}.$$

Using this and (7.1) in (7.8) completes the proof. $\blacksquare$

LEMMA 7.7. *There exists a positive constant $a_0 < \infty$ independent of n, δ such that for $k \geq 0$ and $a \leq \frac{\delta}{a_0} \log n$,*

$$\mathbb{E} \left[Z_n^{(k),\phi} - \tilde{Z}_n^{(k),\phi} \right] \leq C e^{C'a} n^{-\delta} (k+1) D_n(k, 0).$$

PROOF.

$$\begin{aligned}
 (7.9) \quad \mathbb{E} \left[Z_n^{(k),\phi} - \tilde{Z}_n^{(k),\phi} \right] &\leq \mathbb{E} \left[\sum_{j=1}^{D_n(k,0)} \sum_{i=0}^{n^\delta-1} \sum_{x \in \mathcal{V}_n^{(k)}(i,j)} Z^{(k),\phi}(i,j,x) \mathbb{1}_{\{B_x\}} \right] \\
 &= D_n(k,0) \sum_{i=0}^{n^\delta-1} \mathbb{E} \left[\sum_{x \in \mathcal{V}_n^{(k)}(i,1)} Z^{(k),\phi}(i,1,x) \mathbb{1}_{\{B_x\}} \right],
 \end{aligned}$$

where B_x is the event that the vertex x is bad namely has one or more descendants in the interval that it was born. Now, recalling that $\phi \in \mathcal{C}$, note that for a fixed i , conditional on the number of births $N_n^{(k)}(i,1)$, we have

$$(7.10) \quad \sum_{x \in \mathcal{V}_n^{(k)}(i,1)} Z^{(k),\phi}(i,1,x) \mathbb{1}_{\{B_x\}} \leq_{\text{st}} \sum_{l=1}^{N_n^{(k)}(i,1)} 2b_\phi |PA^{(l)}[0,a]| \mathbb{1}_{\{\tilde{B}_l\}},$$

where $\{PA^{(l)} : l \geq 1\}$ is a collection of linear PA branching processes with parameters $\nu = C$ and $\kappa = 0$ (independent of $N_n^{(k)}(i,1)$) and $\tilde{B}_l := \{|PA^{(l)}[0,a/n^\delta]| \geq 2\}$, namely the root of $PA^{(l)}$ has at least one child by time a/n^δ (here C can be taken to be the same constant appearing in Assumption 2.4(i)). Using this in (7.9) implies,

$$(7.11) \quad \mathbb{E} \left[Z_n^{(k),\phi} - \tilde{Z}_n^{(k),\phi} \right] \leq 2c D_n(k,0) \sum_{i=1}^{n^\delta-1} \mathbb{E}(N_n^{(k)}(i,1)) \mathbb{E}(|PA^{(1)}[0,a]| \mathbb{1}_{\{\tilde{B}_1\}}).$$

Conditioning on the number of births $Y(a/n^\delta)$ of the root of $PA^{(1)}$ in $[0,a/n^\delta]$ and by the Markov property,

$$\mathbb{E}(|PA^{(1)}[0,a]| \mathbb{1}_{\{\tilde{B}_1\}}) \leq \sum_{j=1}^{\infty} \mathbb{P}\left(Y(an^{-\delta}) = j\right) \mathbb{E}(PA^{(1),j}[0,a]),$$

where $PA^{(1),j}$ is a modified linear PA process with $\nu = C, \kappa = 0$ with the modification that the offspring process of the root of $PA^{(1),j}$ is constructed using attachment function $f(i) := C(j+i+1)$ for $i \geq 0$. Comparing rates, it is easy to see that for each $j \geq 1$, $PA^{(1),j}[0,a] \leq_{\text{st}} U_j(a)$, where $U_j(a)$ is constructed by first running a linear PA process $PA_{\nu,\kappa}$ with $\nu = C$ and $\kappa = Cj$ and then setting $U_j(a) = |PA_{\nu,\kappa}[0,a]|$. By Lemma 6.3 for $Y(a/n^\delta)$ and Proposition 6.5 for $\mathbb{E}(U_j(a))$, we get $a_0 > 0$ such that whenever $a \leq \frac{\delta}{a_0} \log n$,

$$(7.12) \quad \mathbb{E}(|PA^{(1)}[0,a]| \mathbb{1}_{\{\tilde{B}_1\}}) \leq \sum_{j=1}^{\infty} \left(Can^{-\delta}\right)^j e^{a(2C+Cj)} \leq Ce^{C'a} n^{-\delta}.$$

In (7.11), using this bound and using Corollary 7.3 for $\mathbb{E}(N_n^{(k)}(i,1))$ completes the proof. ■

LEMMA 7.8. *For any $k \geq 0$, whenever $a \leq \frac{\delta}{a_0} \log n$,*

$$\begin{aligned}
 \omega_n &:= \mathbb{E} \left| Z_n^\phi - \sum_{k=0}^{\infty} D_n(k,0) \sum_{i=0}^{n^\delta-1} m_{f_1}^\phi \left(a - (i+1)an^{-\delta} \right) \mu_{f_1}^{(k)} \left[ian^{-\delta}, (i+1)an^{-\delta} \right] \right| \\
 &\leq Ce^{C'a} \left(n^{1-\delta} + \sqrt{n} + n^{-\delta/2} \left(\sum_{k=1}^{\infty} (k+1)^2 D_n(k,0) \right)^{1/2} \right).
 \end{aligned}$$

PROOF. We can write $\omega_n := \omega_n^{(1)} + \omega_n^{(2)} + \omega_n^{(3)}$ where $\omega_n^{(1)} := Z_n^\phi - \tilde{Z}_n^\phi$, $\omega_n^{(2)} := \tilde{Z}_n^\phi - \mathbb{E}(\tilde{Z}_n^\phi)$ and

$$\omega_n^{(3)} := \mathbb{E}(\tilde{Z}_n^\phi) - \sum_{k=0}^{\infty} D_n(k, 0) \sum_{i=0}^{n^\delta-1} m_{f_1}^\phi \left(a - (i+1)an^{-\delta} \right) \mu_{f_1}^{(k)} \left[ian^{-\delta}, (i+1)an^{-\delta} \right].$$

Now fix $\varepsilon > 0$. Using Lemma 7.7 we get,

$$(7.13) \quad \mathbb{E}(|\omega_n^{(1)}|) \leq Ce^{C'a} n^{-\delta} \sum_{k=0}^{\infty} (k+1) D_n(k, 0) \leq 2\gamma Ce^{C'a} n^{1-\delta},$$

since $\sum_{k=1}^{\infty} (k+1) D_n(k, 0) = 2\gamma n - 1$ for tree $\mathcal{T}_{n\gamma}$. Next using Lemma 7.5 and Jensen's inequality,

$$(7.14) \quad \begin{aligned} \mathbb{E}(|\omega_n^{(2)}|) &\leq Ce^{C'a} \left(\sum_{k=1}^{\infty} \left((k+1)^2 n^{-\delta} + (k+1) \right) D_n(k, 0) \right)^{1/2} \\ &\leq Ce^{C'a} \left(n^{-\delta/2} \left(\sum_{k=1}^{\infty} (k+1)^2 D_n(k, 0) \right)^{1/2} + \sqrt{n} \right). \end{aligned}$$

Finally using Lemma 7.6 gives,

$$(7.15) \quad |\omega_n^{(3)}| \leq Ce^{C'a} \sum_{k=0}^{\infty} (k+1) D_n(k, 0) n^{-\delta} \leq Ce^{C'a} n^{1-\delta}.$$

Combining (7.13), (7.14) and (7.15) completes the proof. $\blacksquare$

The next lemma establishes Lipschitz continuity of $m_{f_1}^\phi(t)$ in t for any $\phi \in \mathcal{C}$.

LEMMA 7.9. *For any $k \geq 0$ and any $\eta \in [0, 1]$, $\sup_{t \in [0, a]} |m_{f_1}^\phi(t + \eta) - m_{f_1}^\phi(t)| \leq Ce^{C'a}\eta$.*

PROOF. Let $\bar{\tau}_1$ be the time of the first birth for the branching process with attachment function f_1 . For any $t \in [0, a]$ and $\eta \in [0, 1]$, using the Markov property at time η , we obtain

$$(7.16) \quad \begin{aligned} m_{f_1}^\phi(t + \eta) &= \mathbb{E} \left[Z_{f_1}^\phi(t + \eta) \right] = \mathbb{E} \left[Z_{f_1}^\phi(t + \eta) \mathbb{1}(\bar{\tau}_1 > \eta) \right] + \mathbb{E} \left[Z_{f_1}^\phi(t + \eta) \mathbb{1}(\bar{\tau}_1 \leq \eta) \right] \\ &= \mathbb{E} \left[Z_{f_1}^\phi(t) \right] \mathbb{E} \left[\mathbb{1}(\bar{\tau}_1 > \eta) \right] + \mathbb{E} \left[Z_{f_1}^\phi(t + \eta) \mathbb{1}(\bar{\tau}_1 \leq \eta) \right] = m_{f_1}^\phi(t) (1 - \mathbb{P}(\bar{\tau}_1 \leq \eta)) + \mathbb{E} \left[Z_{f_1}^\phi(t + \eta) \mathbb{1}(\bar{\tau}_1 \leq \eta) \right]. \end{aligned}$$

Using the strong Markov property at $\bar{\tau}_1$, we can write the second term above as $\mathbb{E} \left[Z_{f_1}^\phi(t + \eta) \mathbb{1}(\bar{\tau}_1 \leq \eta) \right] = \mathbb{E} \left[\mathbb{E} \left(Z_{f_1}^\phi(t + \eta) \mid \mathcal{F}_{\bar{\tau}_1} \right) \mathbb{1}(\bar{\tau}_1 \leq \eta) \right]$, where $\mathcal{F}_{\bar{\tau}_1}$ denotes the associated stopped sigma field. Note that at time $\bar{\tau}_1$, there are two vertices, one with degree one and the other with degree zero. Thus, conditional on $\mathcal{F}_{\bar{\tau}_1}$, for $i = 1, 2$, if $U_i(t)$ is distributed as the size of the linear PA process PA_{v, κ_i} with $v = C$ and $\kappa_i = C(i-1)$ at time t (where C is the same constant appearing in Assumption 2.4(i)), we have

$$\mathbb{E} \left(Z_{f_1}^\phi(t + \eta) \mid \mathcal{F}_{\bar{\tau}_1} \right) \leq 2c \mathbb{E} \left(Z_{f_1}(t + \eta) \mid \mathcal{F}_{\bar{\tau}_1} \right) \leq 2c \mathbb{E}(U_1(a+1) + U_2(a+1)) \leq Ce^{C'a}$$

for constants C, C' not depending on η, a, t , where we used Proposition 6.5 to get the last inequality. Using this bound and (7.1) in (7.16), we obtain

$$|m_{f_1}^\phi(t + \eta) - m_{f_1}^\phi(t)| = \left| -m_{f_1}^\phi(t) \mathbb{P}(\bar{\tau}_1 \leq \eta) + Ce^{C'a} \mathbb{P}(\bar{\tau}_1 \leq \eta) \right| \leq 2Ce^{C'a} \mathbb{P}(\bar{\tau}_1 \leq \eta) \leq C'' e^{C'a} \eta$$

for a constant C'' not depending on η, a, t , where the last equality comes from the fact $\bar{\tau}_1 \sim \text{Exp}(f_1(0))$. $\blacksquare$

LEMMA 7.10. Recall $\lambda_k^\phi = \int_0^a m_{f_1}^\phi(a-s)\mu_{f_1}^{(k)}(ds)$. For any $k \geq 0$, whenever $a \leq \frac{\delta}{a_0} \log n$,

$$\mathbb{E} \left| Z_n^\phi - \sum_{k=0}^{\infty} D_n(k, 0) \lambda_k^\phi \right| \leq C e^{C'a} \left(n^{1-\delta} + \sqrt{n} + n^{-\delta/2} \left(\sum_{k=1}^{\infty} (k+1)^2 D_n(k, 0) \right)^{1/2} \right).$$

PROOF. By Lemma 7.8 it is enough to show, for positive constants C, C' not depending on a, n, δ ,

$$\bar{\omega}_n^* := \left| \sum_{k=0}^{\infty} D_n(k, 0) \lambda_k^\phi - \sum_{k=0}^{\infty} D_n(k, 0) \sum_{i=0}^{n^\delta-1} m_{f_1}^\phi \left(a - \frac{(i+1)a}{n^\delta} \right) \mu_{f_1}^{(k)} \left[\frac{ia}{n^\delta}, \frac{(i+1)a}{n^\delta} \right] \right| \leq C e^{C'a} n^{1-\delta}. \quad (7.17)$$

Using Lemma 7.9,

$$\begin{aligned} \bar{\omega}_n^* &\leq \sum_{k=0}^{\infty} D_n(k, 0) \int_0^a \sum_{i=0}^{n^\delta-1} \left| m_{f_1}^\phi(a-s) - m_{f_1}^\phi \left(a - \frac{(i+1)a}{n^\delta} \right) \right| \mathbb{1} \left(s \in \left[\frac{ia}{n^\delta}, \frac{(i+1)a}{n^\delta} \right] \right) \mu_{f_1}^{(k)}(ds) \\ &\leq C e^{C'a} n^{-\delta} \sum_{k=0}^{\infty} D_n(k, 0) \int_0^a \sum_{i=0}^{n^\delta-1} \mathbb{1} \left(s \in \left[\frac{ia}{n^\delta}, \frac{(i+1)a}{n^\delta} \right] \right) \mu_{f_1}^{(k)}(ds) = C e^{C'a} n^{-\delta} \sum_{k=0}^{\infty} D_n(k, 0) \mu_{f_1}^{(k)}[0, a] \\ &\leq (C e^{C'a})^2 a n^{-\delta} \sum_{k=0}^{\infty} (k+1) D_n(k, 0) = (C e^{C'a})^2 a n^{-\delta} (2\gamma n - 1), \end{aligned}$$

where the last inequality uses Lemma 7.2 and the last equality uses $\sum_{k=0}^{\infty} (k+1) D_n(k, 0) = 2\gamma n - 1$. ■

LEMMA 7.11. Let $\phi \in \mathcal{C}$. As $n \rightarrow \infty$, $n^{-1} \sum_{k=1}^{\infty} D_n(k, 0) \lambda_k^\phi \xrightarrow{a.s.} \gamma \sum_{k=1}^{\infty} p_k^0 \lambda_k^\phi$.

PROOF. Let χ be the characteristic $\chi(t) = \sum_{k=0}^{\infty} \lambda_k^\phi \mathbb{1} \{ \xi_{f_0}(t) = k \}$. Note by equation (7.1) and Lemma 7.2 that $\lambda_k^\phi \leq C e^{C'a} (k+1)$ and thus $\chi \in \mathcal{C}$. Now apply Lemma 6.6 (i). ■

Completing the proof of Theorem 7.1: By letting $\delta \rightarrow \infty$ and keeping $n \geq 2$ fixed in Lemma 7.10, the first claim follows. Lemma 7.11 then gives the second claim.

8. Proofs: Sup-norm convergence of degree distribution for the standard model.

We will assume throughout this section that f_0, f_1 satisfy Assumption 2.4.

8.1. *Proof of Theorems 3.8 and 3.11.* Here we prove convergence results for the empirical degree distribution post change-point. As before, time starts at the change point, i.e. $t = 0$ represents the time $T_{\gamma n}$. We focus on the characteristic $\phi(t) = \mathbb{1} \{ \xi_{f_1}(t) = k \}$ for fixed $k \geq 0$ and denote the corresponding $Z_{f_1}^\phi$ and $m_{f_1}^\phi$ by $Z_{f_1}^{(k)}$ and $m_{f_1}^{(k)}$. $\text{BP}_n(t)$ denotes the branching process at time t (i.e. t time units after the change point).

8.1.1. *Notation.* We will use the following notation for fixed $t \geq 0$ in this section.

- (i) Recall $n\gamma$ is the number of vertices born **before** the change point. Let $Z_{AC,n}(t) :=$ number of vertices at time t who were born **after** the change point. $Z_n(t) := n\gamma + Z_{AC,n}(t)$ denotes the total number of vertices in the system at time t .
- (ii) Let $\mathcal{D}_n^{BC}(k, t)$ be the set of vertices with degree k at time t who were born **before** the change point $T_{\gamma n}$. Let $D_n^{BC}(k, t) = |\mathcal{D}_n^{BC}(k, t)|$. Similarly, let $\mathcal{D}_n^{AC}(k, t)$ be the set of vertices with degree k at time t who were born **after** the change point. Let $D_n^{AC}(k, t) = |\mathcal{D}_n^{AC}(k, t)|$. Let $D_n(k, t) = D_n^{BC}(k, t) + D_n^{AC}(k, t)$ be the total number of vertices with degree k at time t .

- (iii) Let $\lambda_\ell^{AC}(t) = \int_0^t m_{f_1}(t-s)\mu_{f_1}^{(\ell)}(ds)$ and $\lambda_\ell^{AC,(k)}(t) = \int_0^t m_{f_1}^{(k)}(t-s)\mu_{f_1}^{(\ell)}(ds)$. Let $\lambda_\ell(t) = 1 + \lambda_\ell^{AC}(t)$ and $\lambda_\ell^{(k)}(t) = \mathbb{P}\left(\xi_{f_1}^{(\ell)}(t) = k - \ell\right) + \lambda_\ell^{AC,(k)}(t)$.
- (iv) Let $q_k(t) := \mathbb{P}\left(\xi_{f_1}^{(k)}(t) > 1\right)$.

The following is the main theorem proved in this section. As will be seen below, Theorems 3.8 and 3.11 are consequences of this theorem.

THEOREM 8.1. *For any $k \geq 0$, $a > 0$, $\epsilon > 0$, as $n \rightarrow \infty$,*

$$\mathbb{P}\left(\sup_{t \in [0, a]} \left| D_n(k, t) - n \sum_{\ell=0}^{\infty} \gamma p_\ell^0 \lambda_\ell^{(k)}(t) \right| > \epsilon n\right) \rightarrow 0, \quad \mathbb{P}\left(\sup_{t \in [0, a]} \left| Z_n(t) - n \sum_{\ell=0}^{\infty} \gamma p_\ell^0 \lambda_\ell(t) \right| > \epsilon n\right) \rightarrow 0.$$

Assuming the above result for the time being, we now describe how Theorem 8.1 (coupled with a technical continuity result, Lemma 8.4) is enough to prove Theorems 3.8 and 3.11. Recall for $m \geq 1$, $T_m = \inf\{t \geq 0 : |\text{BP}_n(t)| = m\}$.

COROLLARY 8.2. *Let $G(t) := \sum_{\ell=0}^{\infty} p_\ell^0 \lambda_\ell^{AC}(t)$, $t \geq 0$. For any $s \in [\gamma, 1]$, let a_s be the unique solution to $G(a_s) = (s - \gamma)/\gamma$. Then for any $s \in [\gamma, 1]$, $\sup_{t \in [\gamma, s]} |T_{[tn]} - a_t| \xrightarrow{P} 0$ as $n \rightarrow \infty$.*

PROOF. As f_1 is a strictly positive function, it is easy to see that $G(t)$ is strictly increasing in t and $G(\gamma) = 0$. By Lemma 8.4 proved below, G (hence G^{-1}) is continuous. Moreover since $m_{f_1}(t) \geq 1$ and $\lambda_\ell^{AC}(t) \geq \mu_{f_1}^{(\ell)}(t) \uparrow \infty$ we see $G(t) \rightarrow \infty$ as $t \rightarrow \infty$. Therefore $G(a_s) = \frac{s-\gamma}{\gamma}$ has a unique solution for $s \in [\gamma, 1]$.

Next fix $s \in [\gamma, 1]$ and let a_s be as above. For any $\eta > 0$, choosing $\epsilon = \frac{G(a_s + \eta) - G(a_s)}{2\gamma}$, the second assertion in Theorem 8.1 readily implies $\mathbb{P}(Z_n(a_s + \eta) > sn + 1) \rightarrow 1$. Similarly, it follows that $\mathbb{P}(Z_n(a_s - \eta) < sn - 1) \rightarrow 1$. Therefore, $T_{[sn]} \xrightarrow{P} a_s$. From this, Theorem 8.1, and the definition of G , $\lambda^{(\ell)}$, we conclude that

$\frac{1}{n} \sup_{t \in [0, T_{[sn]}]} |Z_n(t) - \gamma n(1 + G(t))| \xrightarrow{P} 0$ which implies $\sup_{t \in [\gamma, s]} \left| \frac{t-\gamma}{\gamma} - G(T_{[tn]}) \right| \xrightarrow{P} 0$. By continuity of G^{-1} , this implies $\sup_{t \in [\gamma, s]} \left| G^{-1}\left(\frac{t-\gamma}{\gamma}\right) - T_{[tn]} \right| \xrightarrow{P} 0$ which proves the corollary. ■

PROOF OF THEOREM 3.8. Fix $s \in [\gamma, 1]$. It follows from Lemma 8.4 and Corollary 8.6 proved below that $t \mapsto \Phi_t(\mathbf{p}^0)$ is continuous and hence, from Corollary 8.2 for each fixed $k \geq 0$,

$$(8.1) \quad \sup_{t \in [\gamma, s]} \left| (\Phi_{T_{[tn]}}(\mathbf{p}^0))_k - (\Phi_{a_t}(\mathbf{p}^0))_k \right| \xrightarrow{P} 0.$$

It is easy to see that

$$(8.2) \quad \sup_{t \in [\gamma, s]} \left| \frac{D_n(k, T_{[tn]})}{tn} - (\Phi_{T_{[tn]}}(\mathbf{p}^0))_k \right| \leq \frac{1}{\gamma n} \left(\sup_{t \in [0, T_{[sn]}]} \left| D_n(k, t) - n \sum_{\ell=0}^{\infty} \gamma p_\ell^0 \lambda_\ell^{(k)}(t) \right| + \sup_{t \in [0, T_{[sn]}]} \left| Z_n(t) - n \sum_{\ell=0}^{\infty} \gamma p_\ell^0 \lambda_\ell(t) \right| \right) \xrightarrow{P} 0.$$

The theorem follows from (8.1) and (8.2). ■

PROOF OF THEOREM 3.11. Follows immediately from Theorem 3.8. ■

Proof of Theorem 8.1: The rest of this Section is devoted to the proof of this result. We start with a brief outline of the proof. We start by partitioning the interval $[0, a]$ into subintervals $[t_j, t_{j+1}]_{1 \leq j \leq an^{\bar{\theta}} - 1}$ and showing by means of some continuity estimates that $D_n(k, t)$ and $Z_n(t)$ do not vary too much as t varies within each such subinterval (see Lemmas 8.8 and 8.10). We then use Theorem 7.1 (for vertices born post-change point) and a variance computation (8.17) (for vertices born pre-change point) to show that $|D_n(k, t) - \sum_{\ell=0}^{\infty} D_n(\ell, 0) \lambda_{\ell}^{(k)}(t)|$ and $|Z_n(t) - \sum_{\ell=0}^{\infty} D_n(\ell, 0) \lambda_{\ell}(t)|$ are small for each $t = t_j$. This, combined with the continuity estimates, implies that the above quantities are small uniformly for all $t \in [0, a]$ for appropriately chosen partitions. Finally, a law of large numbers type argument along with continuity estimates is used to show $|\frac{1}{n} \sum_{\ell=0}^{\infty} D_n(\ell, 0) \lambda_{\ell}^{(k)}(t) - \gamma \sum_{\ell=0}^{\infty} p_{\ell}^0 \lambda_{\ell}^{(k)}(t)|$ and $|\frac{1}{n} \sum_{\ell=0}^{\infty} D_n(\ell, 0) \lambda_{\ell}(t) - \gamma \sum_{\ell=0}^{\infty} p_{\ell}^0 \lambda_{\ell}(t)|$ are uniformly small in t , which proves Theorem 8.1.

For the remaining portion of this section C, C', C'', n_0 will denote generic positive constants not depending on n, a, k, ℓ, t whose values might change from line to line and between inequalities.

LEMMA 8.3. $q_k(t) \leq C(k+1)t$ where C is the constant appearing in Assumption 2.4(i) on f_1 .

PROOF. Let $\bar{\tau}_1^k$ be the time of the first birth to a vertex started with degree k . Note $\bar{\tau}_1^k \sim \text{Exp}(f_1(k))$. Thus $\mathbb{P}(\bar{\tau}_1^k < t) = 1 - e^{-f_1(k)t} \leq f_1(k)t \leq C(k+1)t$. The final inequality comes from Assumption 2.4(i) on f_1 . ■

LEMMA 8.4. For any $\ell, k \geq 0$ and $t, t+s \leq a$,

$$|\lambda_{\ell}(t+s) - \lambda_{\ell}(t)| \leq Ce^{C'a}(\ell+1)s, \quad |\lambda_{\ell}^{AC, (k)}(t+s) - \lambda_{\ell}^{AC, (k)}(t)| \leq Ce^{C'a}(\ell+1)s.$$

PROOF. We will only prove the first inequality. The second one follows similarly.

$$\begin{aligned} |\lambda_{\ell}(t+s) - \lambda_{\ell}(t)| &\leq \int_0^t |m_{f_1}(t+s-x) - m_{f_1}(t-x)| \mu_{f_1}^{(\ell)}(dx) + \int_t^{t+s} m_{f_1}(t+s-x) \mu_{f_1}^{(\ell)}(dx) \\ &\leq Ce^{C'a}s \mathbb{E} \left[\xi_{f_1}^{(\ell)}[0, t] \right] + Ce^{C'a} m_{f_1}(t+s) \mathbb{E} \left[\xi_{f_1}^{(\ell)}[t, t+s] \right] \leq Ce^{2C'a} a(\ell+1)s + Ce^{2C'a}(\ell+1)s \end{aligned}$$

where the second inequality uses Lemma 7.9 and the third inequality uses Lemma 7.2 and (7.1). ■

LEMMA 8.5. For $k \geq \ell$ and $t, t+s \leq a$, $\left| \mathbb{P} \left(\xi_{f_1}^{(\ell)}(t+s) = k-\ell \right) - \mathbb{P} \left(\xi_{f_1}^{(\ell)}(t) = k-\ell \right) \right| \leq Ce^{C'a}(k+1)s$.

PROOF. We prove this inequality in two steps. By repeated applications of the Markov property, Markov's inequality and Lemma 7.2,

$$\begin{aligned} &\mathbb{P} \left(\xi_{f_1}^{(\ell)}(t+s) = k-\ell \right) \\ &= \sum_{d=0}^{k-\ell-1} \mathbb{P} \left(\xi_{f_1}^{(\ell)}(t) = d \right) \mathbb{P} \left(\xi_{f_1}^{(d+\ell)}(s) = k-\ell-d \right) + \mathbb{P} \left(\xi_{f_1}^{(\ell)}(t) = k-\ell \right) \mathbb{P} \left(\xi_{f_1}^{(k)}(s) = 0 \right) \\ &\leq \sum_{d=0}^{k-\ell-1} \mathbb{P} \left(\xi_{f_1}^{(\ell)}(t) = d \right) \mathbb{E} \xi_{f_1}^{(d+\ell)}(s) + \mathbb{P} \left(\xi_{f_1}^{(\ell)}(t) = k-\ell \right) \end{aligned}$$

$$\begin{aligned}
&\leq Ce^{C'a}s \left(\mathbb{E} \left(\xi_{f_1}^{(\ell)}(t) \right) + \ell + 1 \right) + \mathbb{P} \left(\xi_{f_1}^{(\ell)}(t) = k - \ell \right) \\
&\leq C'' e^{2C'a}(\ell + 1)s + \mathbb{P} \left(\xi_{f_1}^{(\ell)}(t) = k - \ell \right).
\end{aligned}$$

We now show the opposite inequality.

$$\mathbb{P} \left(\xi_{f_1}^{(\ell)}(t+s) = k - \ell \right) \geq \mathbb{P} \left(\xi_{f_1}^{(\ell)}(t) = k - \ell \right) \mathbb{P} \left(\xi_{f_1}^{(k)}(s) = 0 \right) = \mathbb{P} \left(\xi_{f_1}^{(\ell)}(t) = k - \ell \right) \left(1 - \mathbb{P} \left(\xi_{f_1}^{(k)}(s) \geq 1 \right) \right)$$

Thus

$$\begin{aligned}
\mathbb{P} \left(\xi_{f_1}^{(\ell)}(t+s) = k - \ell \right) - \mathbb{P} \left(\xi_{f_1}^{(\ell)}(t) = k - \ell \right) &\geq -\mathbb{P} \left(\xi_{f_1}^{(\ell)}(t) = k - \ell \right) \mathbb{P} \left(\xi_{f_1}^{(k)}(s) \geq 1 \right) \\
&\geq -\mathbb{E} \xi_{f_1}^{(k)}(s) \geq -Ce^{C'a}(k+1)s.
\end{aligned}$$

The second inequality uses Markov's inequality and the last inequality comes from Lemma 7.2. $\blacksquare$

An immediate consequence of Lemmas 8.4 and 8.5 is

COROLLARY 8.6. *For any $k, \ell > 0$ and $t, t+s < a$, $|\lambda_{\ell}^{(k)}(t+s) - \lambda_{\ell}^{(k)}(t)| \leq Ce^{C'a}(k+\ell+2)s$.*

COROLLARY 8.7. *For any k and $t, t+s < a$, $\sum_{\ell=0}^{\infty} D_n(\ell, 0) |\lambda_{\ell}^{(k)}(t+s) - \lambda_{\ell}^{(k)}(t)| \leq Ce^{C'a}(k+3)sn$.*

PROOF. By the above Corollary 8.6 (with k fixed)

$$\sum_{\ell=0}^{\infty} D_n(\ell, 0) |\lambda_{\ell}^{(k)}(t) - \lambda_{\ell}^{(k)}(t+s)| \leq Ce^{C'a}s \sum_{\ell=0}^{\infty} (k+\ell+2) D_n(\ell, 0) \leq Ce^{C'a}(k+3)s\gamma n$$

since $\sum_{\ell=0}^{\infty} D_n(\ell, 0) = \gamma n$ and $\sum_{\ell=0}^{\infty} \ell D_n(\ell, 0) = \gamma n - 1$. $\blacksquare$

For the rest of this section, unless specified otherwise, we always work conditional on $\mathcal{F}_n(0)$ so that expectation operations such as $\mathbb{P}(\cdot)$, $\mathbb{E}(\cdot)$ and $\text{Var}(\cdot)$ in the ensuing results mean $\mathbb{P}(\cdot|\mathcal{F}_n(0))$, $\mathbb{E}(\cdot|\mathcal{F}_n(0))$ and $\text{Var}(\cdot|\mathcal{F}_n(0))$ respectively.

We will use Theorem 7.1 crucially in what follows for two significant characteristics. Taking $\phi(t) = \mathbb{1}_{\{t \geq 0\}}$ in Theorem 7.1, there exist deterministic positive constants $C, C' < \infty$ independent of a, n such that for every $n \geq 2$,

$$(8.3) \quad \sup_{t \in [0, a]} \mathbb{E} \left| Z_{AC, n}(t) - \sum_{k=0}^{\infty} D_n(k, 0) \lambda_k^{AC}(t) \right| < Ce^{C'a} \sqrt{n}.$$

Taking any $k \geq 0$ and setting $\phi(t) = \mathbb{1}_{\{\xi_{f_1}(t) = k\}}$ in Theorem 7.1, there exist deterministic positive constants $C, C' < \infty$ independent of a, n, k such that for every $n \geq 2$,

$$(8.4) \quad \sup_{t \in [0, a]} \mathbb{E} \left| D_n^{AC}(k, t) - \sum_{\ell=0}^{\infty} D_n(\ell, 0) \lambda_{\ell}^{AC, (k)}(t) \right| < Ce^{C'a} \sqrt{n}.$$

Take any $\tilde{\theta} \in (0, 1/2)$. Take $\omega \in (0, 1)$ such that $\omega > \max(1 - \tilde{\theta}, \frac{1}{2} + \tilde{\theta})$. Now let $\{t_i\}_{i=0}^{n^{\tilde{\theta}}-1}$ be an equispaced partition of $[0, a]$ of mesh $an^{-\tilde{\theta}}$.

LEMMA 8.8. *Let $\{t_j\}, \tilde{\theta}$ and ω be as above. Fix $\epsilon \in (0, 1)$ and k . Then we have*

$$\sum_{j=0}^{n^{\tilde{\theta}}-1} \mathbb{P} \left(\sup_{t \in [t_j, t_{j+1}]} |D_n(k, t) - D_n(k, t_j)| > \epsilon n^{\omega} \right) \leq Ce^{C'a} \epsilon^{-2} n^{-(\omega - \tilde{\theta} - \frac{1}{2})}.$$

PROOF. Condition on $\mathcal{F}_n(t_j)$. Fix j and consider $t \in [t_j, t_{j+1}]$. We clearly have the following lower bound on $D_n(k, t)$:

$$D_n(k, t) \geq D_n(k, t_j) - Y_1$$

where Y_1 is the number of degree k vertices at time t_j which have given birth by time t_{j+1} . Note that $Y_1 \stackrel{d}{=} \text{Bin}\left(D_n(k, t_j), q_k(an^{-\tilde{\theta}})\right)$. We also have the following upper bound on $D_n(k, t)$:

$$(8.5) \quad D_n(k, t) \leq (Z_{AC,n}(t_{j+1}) - Z_{AC,n}(t_j)) + Y_2 + D_n(k, t_j)$$

where Y_2 denotes the number of vertices existing at time t_j of degree strictly less than k which have given birth by time t_{j+1} . Note that $Y_2 \stackrel{d}{=} \sum_{\ell=0}^{k-1} \text{Bin}\left(D_n(\ell, t_j), q_\ell(an^{-\tilde{\theta}})\right)$. To see this upper bound, note that the degree k vertices at time t originate from vertices either existing at time t_j or new vertices born in the time interval $[t_j, t]$. The latter is bounded by $Z_{AC,n}(t_{j+1}) - Z_{AC,n}(t_j)$, namely, the total number of new births in the time interval $[t_j, t_{j+1}]$. The former is bounded by the sum of the number of vertices which are of degree k at time t_j and have not given birth by time t (which is bounded by $D_n(k, t_j)$) and the number of vertices of lower degree at time t_j which have grown to degree k at time t (which is bounded by Y_2). These two bounds give the following

$$|D_n(k, t) - D_n(k, t_j)| \leq (Z_{AC,n}(t_{j+1}) - Z_{AC,n}(t_j)) + Y_1 + Y_2.$$

Note the right hand side does not depend on t . We now have for all $0 \leq j \leq n^{\tilde{\theta}} - 1$ and $t \in [t_j, t_{j+1}]$.

$$\begin{aligned} & \sup_{j \leq n^{\tilde{\theta}} - 1} \mathbb{P} \left(\sup_{t \in [t_j, t_{j+1}]} |D_n(k, t) - D_n(k, t_j)| > \epsilon n^\omega \right) \\ & \leq \sup_{j \leq n^{\tilde{\theta}} - 1} \left[\mathbb{P} \left(\sum_{\ell=0}^k \text{Bin}\left(D_n(\ell, t_j), q_\ell(an^{-\tilde{\theta}})\right) > \epsilon n^\omega / 2 \right) + \mathbb{P}(|Z_{AC,n}(t_{j+1}) - Z_{AC,n}(t_j)| > \epsilon n^\omega / 2) \right] \\ & \leq C e^{C'a} \epsilon^{-2} n^{\frac{1}{2} - \tilde{\theta} - \omega} + C e^{C'a} \epsilon^{-1} n^{\frac{1}{2} - \omega} \end{aligned}$$

where the second inequality comes from Lemmas 8.9 and 8.10 which are proved below. The result now follows after taking the sum of these terms. $\blacksquare$

LEMMA 8.9. *Let $\{t_j\}, \tilde{\theta}$ and ω be as above and let $\epsilon \in (0, 1)$. Then there exist constants C'', n_0 such that for all $n \geq n_0$ and all $a \leq C'' \log n$,*

$$\sup_{j \leq n^{\tilde{\theta}}} \mathbb{P} \left(\sum_{\ell=0}^k \text{Bin}\left(D_n(\ell, t_j), q_\ell(an^{-\tilde{\theta}})\right) > \epsilon n^\omega / 2 \right) \leq C e^{C'a} \epsilon^{-2} n^{\frac{1}{2} - \tilde{\theta} - \omega}.$$

PROOF. Let $A_j = \left\{ Z_n(t_j) < (\gamma + \epsilon/8) n^{\tilde{\theta} + \omega} \right\}$. Note $\sum_{\ell=0}^{\infty} (\ell + 1) D_n(\ell, t_j) = 2Z_n(t_j) - 1$, so on the event A_j ,

$$(8.6) \quad \sum_{\ell=0}^{\infty} (\ell + 1) D_n(\ell, t_j) < 2(\gamma + \epsilon/8) n^{\tilde{\theta} + \omega}.$$

Applying Chebyshev's inequality, on the event A_j , we have

$$\mathbb{P} \left(\sum_{\ell=0}^k \text{Bin}\left(D_n(\ell, t_j), q_\ell(an^{-\tilde{\theta}})\right) > \frac{\epsilon}{2} n^\omega \middle| \mathcal{F}_n(t_j) \right)$$

$$\begin{aligned}
&\leq \frac{4}{\epsilon^2 n^{2\omega}} \sum_{\ell=0}^k \text{Var} \left(\text{Bin} \left(D_n(\ell, t_j), q_\ell \left(a n^{-\tilde{\theta}} \right) \right) \middle| \mathcal{F}_n(t_j) \right) \\
&\leq \frac{4}{\epsilon^2 n^{2\omega}} \sum_{\ell=0}^k D_n(\ell, t_j) q_\ell \left(a n^{-\tilde{\theta}} \right) \left(1 - q_\ell \left(a n^{-\tilde{\theta}} \right) \right) \leq \frac{4}{\epsilon^2 n^{2\omega}} \frac{C a}{n^{\tilde{\theta}}} \sum_{\ell=0}^k D_n(\ell, t_j) (\ell + 1) \\
(8.7) \quad &\leq \frac{4}{\epsilon^2 n^{2\omega}} \frac{C a}{n^{\tilde{\theta}}} \left[2 \left(\gamma + \frac{\epsilon}{8} \right) n^{\tilde{\theta} + \omega} \right] \leq \frac{C' a}{\epsilon^2 n^{\omega}}
\end{aligned}$$

for C' not depending on j , where the first inequality is from Chebyshev's inequality the third inequality is a consequence of Lemma 8.3, and the fourth inequality follows from the definition of A_j . We now have

$$(8.8) \quad \mathbb{P} \left(\sum_{\ell=0}^k \text{Bin} \left(D_n(\ell, t_j), q_\ell \left(a n^{-\tilde{\theta}} \right) \right) > \epsilon n^{\omega} / 2 \right) \leq \frac{C' a}{\epsilon^2 n^{\omega}} + \mathbb{P} \left(Z_n(t_j) \geq \left(\gamma + \epsilon/8 \right) n^{\tilde{\theta} + \omega} \right).$$

Now, we control the second term above. By Lemma 7.2 (and the fact the integral is over a bounded interval) $\lambda_\ell(a) \leq C e^{C'a} (\ell + 1)$. As $\tilde{\theta} + \omega > 1$, we can clearly choose C'', n_0 such that for all $n \geq n_0$ and all $a \leq C'' \log n$, $\frac{\epsilon}{16} n^{\tilde{\theta} + \omega} > (1 + \gamma) C e^{C'a} n$. For such n, a ,

$$\sum_{\ell=0}^{\infty} D_n(\ell, 0) \lambda_\ell(t_j) \leq C e^{C'a} \sum_{\ell=0}^{\infty} (\ell + 1) D_n(\ell, 0) = C e^{C'a} (2\gamma n - 1) < \epsilon n^{\tilde{\theta} + \omega} / 16.$$

Consequently,

$$\begin{aligned}
(8.9) \quad &\mathbb{P} \left(Z_n(t_j) \geq \left(\gamma + \frac{\epsilon}{8} \right) n^{\tilde{\theta} + \omega} \right) \leq \mathbb{P} \left(Z_n(t_j) - \gamma n \geq \left(\gamma + \frac{\epsilon}{8} \right) n^{\tilde{\theta} + \omega} - \gamma n \right) \leq \mathbb{P} \left(Z_n(t_j) - \gamma n > \frac{\epsilon}{8} n^{\tilde{\theta} + \omega} \right) \\
&= \mathbb{P} \left(Z_{AC,n}(t_j) > \frac{\epsilon}{8} n^{\tilde{\theta} + \omega} \right) \leq \mathbb{P} \left(\left| Z_{AC,n}(t_j) - \sum_{\ell=0}^{\infty} D_n(\ell, 0) \lambda_\ell(t_j) \right| > \frac{\epsilon}{16} n^{\tilde{\theta} + \omega} \right) \\
&\leq \frac{16}{\epsilon} \frac{1}{n^{\tilde{\theta} + \omega}} \mathbb{E} \left| Z_{AC,n}(t_j) - \sum_{\ell=0}^{\infty} D_n(\ell, 0) \lambda_\ell(t_j) \right| \leq \frac{16}{\epsilon} C e^{C'a} \frac{1}{n^{\tilde{\theta} + \omega - \frac{1}{2}}}
\end{aligned}$$

for C, C' not depending on j , where the last inequality comes from (8.3), (8.7) and (8.9) and the fact that $\tilde{\theta} < 1/2$. The result now follows. $\blacksquare$

LEMMA 8.10. *Let $\{t_j\}$, $\tilde{\theta}$ and ω be as above and let $\epsilon > 0$. Then*

$$\sup_{j \leq n^{\tilde{\theta}-1}} \mathbb{P} \left(\left| Z_{AC,n}(t_{j+1}) - Z_{AC,n}(t_j) \right| > \frac{\epsilon}{2} n^{\omega} \right) \leq C e^{C'a} \epsilon^{-1} n^{\frac{1}{2} - \omega}.$$

PROOF. Applying the triangle inequality,

$$\begin{aligned}
\left| Z_{AC,n}(t_{j+1}) - Z_{AC,n}(t_j) \right| &\leq \left| Z_{AC,n}(t_{j+1}) - \sum_{\ell=0}^{\infty} D_n(\ell, 0) \lambda_\ell(t_{j+1}) \right| + \left| Z_{AC,n}(t_j) - \sum_{\ell=0}^{\infty} D_n(\ell, 0) \lambda_\ell(t_j) \right| \\
&\quad + \sum_{\ell=0}^{\infty} D_n(\ell, 0) \left| \lambda_\ell(t_{j+1}) - \lambda_\ell(t_j) \right|.
\end{aligned}$$

Note by Lemma 8.4 and the fact that $t_{j+1} - t_j = a n^{-\tilde{\theta}}$

$$(8.10) \quad \sum_{\ell=0}^{\infty} D_n(\ell, 0) \left| \lambda_\ell(t_{j+1}) - \lambda_\ell(t_j) \right| \leq C e^{C'a} \frac{a}{n^{\tilde{\theta}}} \sum_{\ell=0}^{\infty} D_n(\ell, 0) (\ell + 1) = C e^{C'a} \frac{a}{n^{\tilde{\theta}}} (2\gamma n - 1) \leq C'' a e^{C'a} n^{1 - \tilde{\theta}}.$$

From equation (8.3) we get $\sup_{j \leq n^{\tilde{\theta}-1}} \mathbb{E} |Z_n(t_j) - \sum_{\ell=0}^{\infty} D_n(\ell, 0) \lambda_\ell(t_j)| \leq C e^{C'a} \sqrt{n}$. Putting this all together, using (8.10), the fact that $\omega > (1 - \tilde{\theta})$ and Markov's inequality we get for large enough n

$$\begin{aligned} \mathbb{P} \left(|Z_{AC,n}(t_{j+1}) - Z_{AC,n}(t_j)| > \frac{\epsilon}{2} n^\omega \right) &= \mathbb{P} \left(|Z_n(t_{j+1}) - Z_n(t_j)| > \frac{\epsilon}{2} n^\omega \right) \\ &\leq \mathbb{P} \left(\left| Z_n(t_j) - \sum_{\ell=0}^{\infty} D_n(\ell, 0) \lambda_\ell(t_j) \right| + \left| Z_n(t_{j+1}) - \sum_{\ell=0}^{\infty} D_n(\ell, 0) \lambda_\ell(t_{j+1}) \right| > \frac{\epsilon}{4} n^\omega \right) \\ &\leq \frac{2}{\epsilon} n^{-\omega} \left(\mathbb{E} \left| Z_n(t_j) - \sum_{\ell=0}^{\infty} D_n(\ell, 0) \lambda_\ell(t_j) \right| + \mathbb{E} \left| Z_n(t_{j+1}) - \sum_{\ell=0}^{\infty} D_n(\ell, 0) \lambda_\ell(t_{j+1}) \right| \right) \leq 2C e^{C'a} \epsilon^{-1} n^{\frac{1}{2}-\omega} \end{aligned}$$

for C, C' not depending on j , which proves the lemma. $\blacksquare$

LEMMA 8.11. *There exist positive constants C, C' such that for each k and $\epsilon \in (0, 1)$,*

$$\begin{aligned} \mathbb{P} \left(\sup_{t \in [0, a]} \left| D_n(k, t) - \sum_{\ell=0}^{\infty} D_n(\ell, 0) \lambda_\ell^{(k)}(t) \right| > \epsilon(k+1) n^\omega \right) &\leq C e^{C'a} \epsilon^{-2} n^{\tilde{\theta} + \frac{1}{2} - \omega} \\ \mathbb{P} \left(\sup_{t \in [0, a]} \left| Z_n(t) - \sum_{\ell=0}^{\infty} D_n(\ell, 0) \lambda_\ell(t) \right| > \epsilon n^\omega \right) &\leq C e^{C'a} \epsilon^{-2} n^{\tilde{\theta} + \frac{1}{2} - \omega}. \end{aligned}$$

PROOF. Fix k and $\epsilon \in (0, 1)$. Note that

$$\begin{aligned} &\mathbb{P} \left(\sup_{t \in [0, a]} \left| D_n(k, t) - \sum_{\ell=0}^{\infty} D_n(\ell, 0) \lambda_\ell^{(k)}(t) \right| > \epsilon n^\omega \right) \\ &\leq \sum_{j=0}^{n^{\tilde{\theta}}-1} \mathbb{P} \left(\sup_{t \in [t_j, t_{j+1}]} \left| D_n(k, t) - \sum_{\ell=0}^{\infty} D_n(\ell, 0) \lambda_\ell^{(k)}(t) \right| > \epsilon n^\omega \right) \\ &\leq \sum_{j=0}^{n^{\tilde{\theta}}-1} \left[\mathbb{P} \left(\sup_{t \in [t_j, t_{j+1}]} |D_n(k, t) - D_n(k, t_j)| > \frac{\epsilon}{3} n^\omega \right) + \mathbb{P} \left(\left| D_n(k, t_j) - \sum_{\ell=0}^{\infty} D_n(\ell, 0) \lambda_\ell^{(k)}(t_j) \right| > \frac{\epsilon}{3} n^\omega \right) \right. \\ &\quad \left. + \mathbb{P} \left(\sup_{t \in [t_j, t_{j+1}]} \sum_{\ell=0}^{\infty} D_n(\ell, 0) \left| \lambda_\ell^{(k)}(t) - \lambda_\ell^{(k)}(t_j) \right| > \frac{\epsilon}{3} n^\omega \right) \right]. \end{aligned} \tag{8.11}$$

By Lemma 8.8,

$$(8.12) \quad \sum_{j=0}^{n^{\tilde{\theta}}-1} \mathbb{P} \left(\sup_{t \in [t_j, t_{j+1}]} |D_n(k, t) - D_n(k, t_j)| > \frac{\epsilon}{3} n^\omega \right) \leq C e^{C'a} \epsilon^{-2} n^{\tilde{\theta} + \frac{1}{2} - \omega}.$$

By Corollary 8.7, $\sup_{j \leq n^{\tilde{\theta}-1}} \sup_{t \in [t_j, t_{j+1}]} \sum_{\ell=0}^{\infty} D_n(\ell, 0) \left| \lambda_\ell^{(k)}(t) - \lambda_\ell^{(k)}(t_j) \right| \leq C e^{C'a} (k + \gamma + 2) n^{1-\tilde{\theta}}$ and hence, as $\omega > 1 - \tilde{\theta}$, there exists n_0 not depending on k such that for all $n \geq n_0$,

$$(8.13) \quad \sum_{j=0}^{n^{\tilde{\theta}}-1} \mathbb{P} \left(\sup_{t \in [t_j, t_{j+1}]} \sum_{\ell=0}^{\infty} D_n(\ell, 0) \left| \lambda_\ell^{(k)}(t) - \lambda_\ell^{(k)}(t_j) \right| > \frac{\epsilon(k+1)}{3} n^\omega \right) = 0.$$

Finally we control the second term appearing in the sum (8.11). It is sufficient to show

$$(8.14) \quad \sup_{j \leq n^{\bar{\theta}}} \mathbb{P} \left(\left| D_n(k, t_j) - \sum_{\ell=0}^{\infty} D_n(\ell, 0) \lambda_{\ell}^{(k)}(t_j) \right| > (\epsilon/3) n^{\omega} \right) \leq C e^{C'a} \epsilon^{-2} n^{\frac{1}{2}-\omega}.$$

By the triangle inequality and definitions of $D_n(k, t)$, and $\lambda_{\ell}^{(k)}(t)$, we see that for each fixed j, k ,

$$(8.15) \quad \begin{aligned} \left| D_n(k, t_j) - \sum_{\ell=0}^{\infty} D_n(\ell, 0) \lambda_{\ell}^{(k)}(t_j) \right| &\leq \left| D_n^{BC}(k, t_j) - \sum_{\ell=0}^k D_n(\ell, 0) \mathbb{P} \left(\xi_{f_1}^{(\ell)}(t_j) = k - \ell \right) \right| \\ &\quad + \left| D_n^{AC}(k, t_j) - \sum_{\ell=0}^{\infty} D_n(\ell, 0) \lambda_{\ell}^{AC, (k)}(t_j) \right|. \end{aligned}$$

By (8.4) and Markov's inequality,

$$(8.16) \quad \sup_{j \leq n^{\bar{\theta}}} \mathbb{P} \left(\left| D_n^{AC}(k, t_j) - \sum_{\ell=0}^{\infty} D_n(\ell, 0) \lambda_{\ell}^{AC, (k)}(t_j) \right| > \frac{\epsilon}{6} n^{\omega} \right) \leq 6C e^{C'a} \epsilon^{-1} n^{\frac{1}{2}-\omega}.$$

We now control the first term appearing in the bound in equation (8.15) by showing

$$(8.17) \quad \sup_{t \in [0, a]} \mathbb{E} \left[\left(D_n^{BC}(k, t) - \sum_{\ell=0}^k D_n(\ell, 0) \mathbb{P} \left(\xi_{f_1}^{(\ell)}(t) = k - \ell \right) \right)^2 \right] \leq Cn.$$

Fix k and $t \in [0, a]$. Define a collection of mutually independent random variables $\{\xi_{f_1, m}^{(\ell)}(t) \mid 1 \leq m \leq D_n(\ell, 0), 0 \leq \ell \leq k\}$ where $\xi_{f_1, m}^{(\ell)}(t) \sim \xi_{f_1}^{(\ell)}(t)$. Note that

$$D_n^{BC}(k, t) \stackrel{d}{=} \sum_{\ell=0}^k \sum_{m=1}^{D_n(\ell, 0)} \mathbb{1} \left(\xi_{f_1, m}^{(\ell)}(t) = k - \ell \right),$$

i.e. a vertex that was born before the change point and was of degree ℓ at the change point has to add $k - \ell$ new births to reach degree k at time t . Therefore,

$$\begin{aligned} &\mathbb{E} \left[\left(D_n^{BC}(k, t) - \sum_{\ell=0}^k D_n(\ell, 0) \mathbb{P} \left(\xi_{f_1}^{(\ell)}(t) = k - \ell \right) \right)^2 \right] \\ &= \mathbb{E} \left[\left(\sum_{\ell=0}^k \sum_{m=1}^{D_n(\ell, 0)} \mathbb{1} \left(\xi_{f_1, m}^{(\ell)}(t) = k - \ell \right) - \sum_{\ell=0}^k D_n(\ell, 0) \mathbb{P} \left(\xi_{f_1}^{(\ell)}(t) = k - \ell \right) \right)^2 \right] \\ &= \mathbb{E} \left[\left\{ \sum_{\ell=0}^k \sum_{m=1}^{D_n(\ell, 0)} \left(\mathbb{1} \left(\xi_{f_1, m}^{(\ell)}(t) = k - \ell \right) - \mathbb{P} \left(\xi_{f_1}^{(\ell)}(t) = k - \ell \right) \right) \right\}^2 \right]. \end{aligned}$$

Note that

$$\sum_{\ell=0}^k \sum_{m=1}^{D_n(\ell, 0)} \left(\mathbb{1} \left(\xi_{f_1, m}^{(\ell)}(t) = k - \ell \right) - \mathbb{P} \left(\xi_{f_1}^{(\ell)}(t) = k - \ell \right) \right) \stackrel{d}{=} \sum_{\ell=0}^k \sum_{m=1}^{D_n(\ell, 0)} Y_{\ell, m}$$

Where the random variables $\{Y_{\ell, m} \mid 1 \leq m \leq D_n(\ell, 0), 0 \leq \ell \leq k\}$ are mutually independent, supported on $[-1, 1]$ and $\mathbb{E} Y_{\ell, m} = 0$. Thus,

$$\mathbb{E} \left[\left(\sum_{\ell=0}^k \sum_{m=1}^{D_n(\ell, 0)} Y_{\ell, m} \right)^2 \right] = \sum_{\ell=0}^k \sum_{m=1}^{D_n(\ell, 0)} \mathbb{E} \left[Y_{\ell, m}^2 \right] \leq C \sum_{\ell=0}^k D_n(\ell, 0) = C\gamma n$$

which proves (8.17). Using (8.17) and Chebychev's inequality, we get

$$(8.18) \quad \sup_{j \leq n^{\hat{\theta}}} \mathbb{P} \left(\left| D_n^{BC}(k, t_j) - \sum_{\ell=0}^k D_n(\ell, 0) \mathbb{P} \left(\xi_{f_1}^{(\ell)}(t_j) = k - \ell \right) \right| > (\epsilon/6)n^{\omega} \right) \leq C\epsilon^{-2}n^{1-2\omega}.$$

Using (8.16) and (8.18) in (8.15), we obtain (8.14). The first assertion in the lemma follows by using (8.12), (8.13) and (8.14) in (8.11). The second assertion follows similarly upon noting that $Z_{AC,n}(t)$ is increasing in t and using (8.3), Lemma 8.10 and the first bound in Lemma 8.4. $\blacksquare$

Now, we proceed towards removing the conditioning on $\mathcal{F}_n(0)$ to complete the proof of Theorem 8.1. We need the following Corollary to Lemma 7.11.

COROLLARY 8.12. *Fix $k \geq 0$, $\epsilon > 0$ and let $s_1, \dots, s_m \in [0, a]$ be m fixed time points. Then, almost surely, there exists $n_0 \geq 1$ such that for all $n \geq n_0$,*

$$\begin{aligned} \sup_{1 \leq j \leq m} \left| \frac{1}{n} \sum_{\ell=0}^{\infty} D_n(\ell, 0) \lambda_{\ell}^{(k)}(s_j) - \gamma \sum_{\ell=0}^{\infty} p_{\ell}^0 \lambda_{\ell}^{(k)}(s_j) \right| &\leq \epsilon, \\ \sup_{1 \leq j \leq m} \left| \frac{1}{n} \sum_{\ell=0}^{\infty} D_n(\ell, 0) \lambda_{\ell}(s_j) - \gamma \sum_{\ell=0}^{\infty} p_{\ell}^0 \lambda_{\ell}(s_j) \right| &\leq \epsilon. \end{aligned}$$

PROOF. Follows from Lemma 7.11 and the union bound. $\blacksquare$

LEMMA 8.13. *Let $\{p_k(f) : k \geq 0\}$ as in (3.1) be the asymptotic degree distribution using attachment function f satisfying Assumption 2.4. Then $\sum_{k=0}^{\infty} k p_k(f) = 1$.*

PROOF. Recall that $p_k(f) = t_{k-1} - t_k$ where $t_k := \prod_{i=0}^k \frac{f(i)}{\lambda^* + f(i)}$ and λ^* is the Malthusian parameter for the corresponding preferential attachment branching process. Therefore, $\sum_{k=1}^{\infty} k p_k(f) = \sum_{k=0}^{\infty} k(t_{k-1} - t_k) = \sum_{k=0}^{\infty} t_k$. By the definition of λ^* and t_k we see $\sum_{k=1}^{\infty} t_k = 1$, proving the lemma. $\blacksquare$

LEMMA 8.14. *For any $k \geq 0$,*

$$\sup_{t \in [0, a]} \left| \frac{1}{n} \sum_{\ell=0}^{\infty} D_n(\ell, 0) \lambda_{\ell}^{(k)}(t) - \gamma \sum_{\ell=0}^{\infty} p_{\ell}^0 \lambda_{\ell}^{(k)}(t) \right| \xrightarrow{a.s.} 0, \quad \sup_{t \in [0, a]} \left| \frac{1}{n} \sum_{\ell=0}^{\infty} D_n(\ell, 0) \lambda_{\ell}(t) - \gamma \sum_{\ell=0}^{\infty} p_{\ell}^0 \lambda_{\ell}(t) \right| \xrightarrow{a.s.} 0.$$

PROOF. Fix $\epsilon > 0$. Let $0 = s_1 < s_2 < \dots < s_m = a$ be a partition such that $|s_{j+1} - s_j| \leq \epsilon$. By Corollary 8.7,

$$\sup_{1 \leq j \leq m} \sup_{t \in [s_j, s_{j+1}]} \left| \frac{1}{n} \sum_{\ell=0}^{\infty} D_n(\ell, 0) \lambda_{\ell}^{(k)}(t) - \frac{1}{n} \sum_{\ell=0}^{\infty} D_n(\ell, 0) \lambda_{\ell}^{(k)}(s_j) \right| \leq C e^{C'a} (k+3) \epsilon.$$

Similarly, using Corollary 8.6,

$$\begin{aligned} \sup_{1 \leq j \leq m-1} \sup_{t \in [s_j, s_{j+1}]} \left| \gamma \sum_{\ell=0}^{\infty} p_{\ell}^0 \lambda_{\ell}^{(k)}(t) - \gamma \sum_{\ell=0}^{\infty} p_{\ell}^0 \lambda_{\ell}^{(k)}(s_j) \right| &\leq \sup_{1 \leq j \leq m-1} \sup_{t \in [s_j, s_{j+1}]} \gamma \sum_{\ell=0}^{\infty} p_{\ell}^0 \left| \lambda_{\ell}^{(k)}(t) - \lambda_{\ell}^{(k)}(s_j) \right| \\ &\leq C e^{C'a} \epsilon \gamma \sum_{\ell=0}^{\infty} p_{\ell}^0 (k + \ell + 2) = C e^{C'a} \gamma (k+3) \epsilon. \end{aligned}$$

By Corollary 8.12, almost surely, there exists $n_0 \geq 1$ such that for all $n \geq n_0$,

$$\sup_{1 \leq j \leq m} \left| \frac{1}{n} \sum_{\ell=0}^{\infty} D_n(\ell, 0) \lambda_{\ell}^{(k)}(s_j) - \gamma \sum_{\ell=0}^{\infty} p_{\ell}^0 \lambda_{\ell}^{(k)}(s_j) \right| \leq \epsilon.$$

From the above, we now have that for $n \geq n_0$,

$$\begin{aligned} & \sup_{t \in [0, a]} \left| \frac{1}{n} \sum_{\ell=0}^{\infty} D_n(\ell, 0) \lambda_{\ell}^{(k)}(t) - \gamma \sum_{\ell=0}^{\infty} p_{\ell}^0 \lambda_{\ell}^{(k)}(t) \right| \\ & \leq \sup_{1 \leq j \leq m-1} \sup_{t \in [s_j, s_{j+1}]} \left| \frac{1}{n} \sum_{\ell=0}^{\infty} D_n(\ell, 0) \lambda_{\ell}^{(k)}(t) - \frac{1}{n} \sum_{\ell=0}^{\infty} D_n(\ell, 0) \lambda_{\ell}^{(k)}(s_j) \right| \\ & \quad + \sup_{1 \leq j \leq m-1} \sup_{t \in [s_j, s_{j+1}]} \left| \gamma \sum_{\ell=0}^{\infty} p_{\ell}^0 \lambda_{\ell}^{(k)}(t) - \gamma \sum_{\ell=0}^{\infty} p_{\ell}^0 \lambda_{\ell}^{(k)}(s_j) \right| \\ & \quad + \sup_{1 \leq j \leq m} \left| \frac{1}{n} \sum_{\ell=0}^{\infty} D_n(\ell, 0) \lambda_{\ell}^{(k)}(s_j) - \gamma \sum_{\ell=0}^{\infty} p_{\ell}^0 \lambda_{\ell}^{(k)}(s_j) \right| \leq C e^{C'a} (k+3) \epsilon \end{aligned}$$

which proves the first assertion of the lemma. The second assertion follows similarly using Corollary 8.12 and the first bound in Lemma 8.4. $\blacksquare$

PROOF OF THEOREM 8.1. The theorem follows from Lemmas 8.11 and 8.14. $\blacksquare$

PROOF OF COROLLARY 3.13. The essential message of this Corollary 3.13 is that the tail of the distribution prescribed by the initializer function always wins. Recall that the limit random variable D_{θ} is a mixture of the distributions of X_{BC} and X_{AC} . $\blacksquare$

LEMMA 8.15. *The random variable X_{AC} always has an exponential tail.*

Proof: By construction, note that $X_{AC} \leq_{st} \xi_{f_1}[0, \alpha]$. Assumption 2.4 on the attachment functions implies that there exists $\bar{\kappa} > 0$ such that $\max(f_0(i), f_1(i)) \leq \bar{\kappa}(i+1)$ for all i . In particular $\xi_{f_1}[0, \alpha] \leq_{st} Y_{\bar{\kappa}}[0, \alpha]$ where $Y_{\bar{\kappa}}(\cdot)$ is a rate $\bar{\kappa}$ Yule process (Definition 6.2). Using Lemma 6.3 completes the proof. $\blacksquare$

Thus it is enough to consider X_{BC} and show that this random variable has the same tail behavior as the random variable $D \sim \{p_k^0 : k \geq 1\}$. Once again by construction, $X_{BC} \leq_{st} D + \sum_{i=1}^D Y_{\bar{\kappa}, i}[0, \alpha]$, where $\{Y_{\bar{\kappa}, i}(\cdot) : i \geq 1\}$ is an infinite collection of independent Yule processes (independent of D) having the same distribution as $Y_{\bar{\kappa}}(\cdot)$. Let $\mu := \mathbb{E}(Y_{\bar{\kappa}, i}[0, \alpha])$. Note $\mu > 1$. Conditioning on the value of D we see that for $x \geq 1$, $\mathbb{P}(X_{BC} > x) \leq \mathcal{E}$ where $\mathcal{E} = \sum_{j=1}^{x/2\mu} \mathbb{P}(D = j) \mathbb{P}(\sum_{i=1}^j Y_{\bar{\kappa}, i}[0, \alpha] > x - j) + \mathbb{P}(D > x/2\mu)$. Further for $x \geq 1$,

$$(8.19) \quad \mathcal{E} \leq \mathbb{P}\left(\sum_{i=1}^{x/2\mu} Y_{\bar{\kappa}, i}[0, \alpha] > x \left(1 - \frac{1}{2\mu}\right)\right) + \mathbb{P}(D > x/2\mu).$$

Standard large deviation bounds for the law of $Y_{\bar{\kappa}, i}$ implies that there exists constants C_1, C_2 such that for all x , $\mathbb{P}\left(\sum_{i=1}^{x/2\mu} Y_{\bar{\kappa}, i}[0, \alpha] > x \left(1 - \frac{1}{2\mu}\right)\right) \leq C_1 \exp(-C_2 x)$. Thus in the setting of Corollary 3.13(i), assuming D has exponential tails, one finds using (8.19) that there exist finite constants C'_1, C'_2 such that $\mathbb{P}(X_{BC} > x) \leq C'_1 \exp(-C'_2 x)$. This completes the proof of Corollary 3.13(i). A similar argument, along with the obvious inequality $\mathbb{P}(D > x) \leq \mathbb{P}(X_{BC} > x)$, verifies Corollary 3.13(ii). $\blacksquare$

9. Proofs: Quick Big bang.

9.1. *Proof of Theorem 3.18.* Throughout this section we assume that f_0 satisfies Assumption 2.4 and f_1 satisfies Assumptions 2.4, 3.1 and 3.16. For notational convenience, instead of considering the change point at n^γ and evolving the tree till size n , we will consider the problem of the change point being at n and evolving the tree till size $n^{1+\lambda_1^*\theta}$ for some $\theta > 0$ (where λ_1^* is the Malthusian rate corresponding to f_1). For this section, $t = 0$ represents time T_n (the first time the total population size of the associated continuous time branching process has n vertices). It is easy to see that Theorem 3.18 is equivalent to Theorem 9.14 proved below.

We first give a proof outline. We again use the embedding of the discrete time network model into the associated continuous time branching process. Recall the notation from Section 8. From Lemma 8.11, for $k \geq 0$, there exists $\eta_0 > 0$ such that for $\eta \leq \eta_0$,

$$(9.1) \quad \frac{1}{n} \sup_{t \in [0, \eta \log n]} \left| D_n(k, t) - \sum_{\ell=0}^{\infty} D_n(\ell, 0) \lambda_\ell^{(k)}(t) \right| \xrightarrow{P} 0, \text{ as } n \rightarrow \infty.$$

Similarly, using Lemma 8.11, we obtain $\eta_0 > 0$ such that for all $\eta \leq \eta_0$,

$$(9.2) \quad \frac{1}{n} \sup_{t \in [0, \eta \log n]} \left| Z_n(t) - \sum_{\ell=0}^{\infty} D_n(\ell, 0) \lambda_\ell(t) \right| \xrightarrow{P} 0, \text{ as } n \rightarrow \infty.$$

(9.1) and (9.2) immediately imply for any $\eta \leq \eta_0$,

$$(9.3) \quad \begin{aligned} \frac{1}{n^{1+\eta\lambda_1^*}} D_n(k, \eta \log n) - \frac{1}{n^{1+\eta\lambda_1^*}} \sum_{\ell=0}^{\infty} D_n(\ell, 0) \lambda_\ell^{(k)}(\eta \log n) &\xrightarrow{P} 0, \\ \frac{1}{n^{1+\eta\lambda_1^*}} Z_n(\eta \log n) - \frac{1}{n^{1+\eta\lambda_1^*}} \sum_{\ell=0}^{\infty} D_n(\ell, 0) \lambda_\ell(\eta \log n) &\xrightarrow{P} 0 \end{aligned}$$

as $n \rightarrow \infty$. Thus, before the total population has grown too big, i.e. is of size $n^{1+\eta\lambda_1^*}$ for some $\eta \leq \eta_0$, one can approximate the empirical degree distribution and rescaled total population size by the normalized sums appearing in (9.3). For each $\ell \geq 0$, $n^{-1} D_n(\ell, 0)$, converges to the classical limit degree distribution of the system without change point i.e. $p_l^0 = p_l(f_0)$ as in (3.1). Thus, in lieu of (9.3), one needs to understand how the quantities $n^{-\eta\lambda_1^*} \lambda_\ell^{(k)}(\eta \log n)$ and $n^{-\eta\lambda_1^*} \lambda_\ell(\eta \log n)$ behave for large n . Lemmas 9.1 to 9.7 use techniques from renewal theory to quantify rates of convergence and characterize properties of the limits of these quantities in this general setup. This can be used to prove an analogue of Theorem 3.18 for the branching process in the regime where the approximation (9.3) is valid i.e. for $\eta \leq \eta_0$. To extend this proof to the general case, we develop a ‘bootstrapping procedure’ laid out in Lemma 9.11 where we use results from Section 7 and the lemmas proved in this section to show that for each $j \geq 0$, the ‘quick big bang’ phenomenon holds when the population is of size $n^{1+\eta\lambda_1^*}$ for some $\eta \leq (j+1)\eta_0$ if it holds for all $\eta \leq j\eta_0$. The rest of the section translates these results to the network model in discrete time.

Define for each $\ell \geq 0$ and $\beta > 0$, the β -Laplace transform of the measure $\mu_{f_1}^{(\ell)}$ given by

$$w_\ell(\beta) := \int_0^\infty e^{-\beta s} \mu_{f_1}^{(\ell)}(ds).$$

We will simply write w_ℓ for $w_\ell(\lambda_1^*)$. We need the following technical lemmas. Recall from Assumption 2.4 (ii) that there exists $\beta_1 \in (0, \lambda_1^*)$ such that $\hat{\rho}(\beta_1) < \infty$. Recall C^* from Assumption 3.1 applied to f_1 .

LEMMA 9.1. $\beta_1 \geq C^*$.

PROOF. If $C^* = 0$, there is nothing to prove. So we assume $C^* > 0$. For any $\epsilon \in (0, C^*)$, by Assumption 3.1, there exists $j_0 \geq 1$ such that for all $j \geq j_0$, $f_1(j) \geq (C^* - \epsilon)j$. Finiteness of $\hat{\rho}(\beta_1)$ implies that

$$(9.4) \quad \sum_{k=1}^{\infty} \prod_{i=0}^{k-1} \frac{f_1(i+j_0)}{\beta_1 + f_1(i+j_0)} < \infty.$$

For any $k \geq 1$, noting that $x \mapsto \frac{x}{\beta_1+x}$ is a strictly increasing function and, $\log(1+x) \leq x$ for any $x \geq 0$, and $\sum_{j=j_1}^{j_2} \frac{1}{j} \leq \int_{j_1-1}^{j_2} \frac{dx}{x}$ for any $j_2 \geq j_1 \geq 1$,

$$\begin{aligned} \log \left[\prod_{i=0}^{k-1} \frac{f_1(i+j_0)}{\beta_1 + f_1(i+j_0)} \right] &\geq \log \left[\prod_{i=0}^{k-1} \frac{\frac{\beta_1}{C^*-\epsilon} + i + j_0}{\frac{\beta_1}{C^*-\epsilon} + i + j_0} \right] = - \sum_{i=0}^{k-1} \log \left[1 + \frac{\beta_1}{(C^*-\epsilon)(i+j_0)} \right] \\ &\geq - \frac{\beta_1}{C^*-\epsilon} \sum_{i=0}^{k-1} \frac{1}{i+j_0} \geq - \frac{\beta_1}{C^*-\epsilon} \int_{j_0-1}^{j_0+k-1} \frac{dx}{x} = - \frac{\beta_1}{C^*-\epsilon} \log \left(\frac{j_0+k-1}{j_0-1} \right) \end{aligned}$$

and thus

$$\prod_{i=0}^{k-1} \frac{f_1(i+j_0)}{\beta_1 + f_1(i+j_0)} \geq \left(\frac{j_0-1}{j_0+k-1} \right)^{\frac{\beta_1}{C^*-\epsilon}}.$$

Thus, (9.4) holds only if $\beta_1 > C^* - \epsilon$. As $\epsilon > 0$ is arbitrary, this proves the lemma. $\blacksquare$

REMARK 9.2. Lemma 9.1 shows that if f satisfies Assumptions 2.4 and 3.1, then $\lambda^* > C^*$. In addition, if f satisfies $\inf_{i \geq 0} f(i) > 0$, then [9, Proposition 5.7], implies $\mathbb{E} \left(\exp \left\{ \delta \int_0^\infty e^{-\lambda^* t} \xi_f(dt) \right\} \right) < \infty$ for some $\delta > 0$ and, in particular, Assumption 3.2.

LEMMA 9.3. For any $\beta \in (\beta_1, \lambda_1^*]$, there exists a constant $C(\beta) > 0$ such that $w_\ell(\beta) \leq C(\beta)(\ell+1)$ for $\ell \geq 0$.

PROOF. Fix any $\beta \in (\beta_1, \lambda_1^*]$ and $\ell \geq 0$. Since $\int_0^\infty e^{-\beta s} \mu_{f_1}(ds) = \sum_{k=1}^\infty \prod_{i=0}^{k-1} \frac{f_1(i)}{\beta + f_1(i)}$, the sum on the right hand side is finite. Note that

$$w_\ell(\beta) = \int_0^\infty e^{-\beta s} \mu_{f_1}^{(\ell)}(ds) = \sum_{k=1}^\infty \prod_{i=\ell}^{\ell+k-1} \frac{f_1(i)}{\beta + f_1(i)} = \frac{\sum_{k=1}^\infty \prod_{i=0}^{\ell+k-1} \frac{f_1(i)}{\beta + f_1(i)}}{\prod_{i=0}^{\ell-1} \frac{f_1(i)}{\beta + f_1(i)}} < \infty.$$

Choose and fix $\epsilon > 0$ such that $C^* + 2\epsilon < \beta$ (which is possible by Lemma 9.1). By Assumption 3.1, there exists $j_0 \geq 1$ such that for all $j \geq j_0$, $f_1(j) \leq (C^* + \epsilon)j$. For any $\ell \geq j_0$, using the facts that $x \mapsto \frac{x}{\beta+x}$ is a strictly increasing function and, $\log(1+x) \geq \frac{x}{1+x}$ for any $x \geq 0$, and $\sum_{j=j_1}^{j_2} \frac{1}{j} \geq \int_{j_1}^{j_2+1} \frac{dx}{x}$ for any $j_2 \geq j_1 \geq 1$, we obtain for any $\ell \geq j_0$,

$$\begin{aligned} \log \left[\prod_{i=\ell}^{2\ell-1} \frac{f_1(i)}{\beta + f_1(i)} \right] &\leq \log \left[\prod_{i=\ell}^{2\ell-1} \frac{i}{\frac{\beta}{C^*+\epsilon} + i} \right] = - \sum_{i=\ell}^{2\ell-1} \log \left[1 + \frac{\beta}{(C^*+\epsilon)i} \right] \\ &\leq - \sum_{i=\ell}^{2\ell-1} \frac{\frac{\beta}{(C^*+\epsilon)i}}{1 + \frac{\beta}{(C^*+\epsilon)i}} \leq - \frac{\frac{\beta}{C^*+\epsilon}}{1 + \frac{\beta}{(C^*+\epsilon)\ell}} \sum_{i=\ell}^{2\ell-1} \frac{1}{i} \leq - \frac{\frac{\beta}{C^*+\epsilon}}{1 + \frac{\beta}{(C^*+\epsilon)\ell}} \int_\ell^{2\ell} \frac{dx}{x} = - \frac{\frac{\beta}{C^*+\epsilon}}{1 + \frac{\beta}{(C^*+\epsilon)\ell}} \log 2. \end{aligned}$$

Take $\ell_1 \geq j_0$ such that $\frac{\frac{\beta}{C^*+\epsilon}}{1 + \frac{\beta}{(C^*+\epsilon)\ell_1}} \geq \frac{\beta}{C^*+2\epsilon}$. From the above calculation, for all $\ell \geq \ell_1$,

$$\prod_{i=\ell}^{2\ell-1} \frac{f_1(i)}{\beta + f_1(i)} \leq 2^{-\frac{\beta}{C^*+2\epsilon}}. \text{ Using this bound iteratively, we obtain for any } j \geq 1, \prod_{i=\ell}^{2^j \ell - 1} \frac{f_1(i)}{\beta + f_1(i)} \leq$$

$2^{-\frac{\beta j}{C^*+2\epsilon}}$. Thus, for all $\ell \geq \ell_1$,

$$\begin{aligned} w_\ell(\beta) &= \sum_{k=1}^{\infty} \prod_{i=\ell}^{\ell+k-1} \frac{f_1(i)}{\beta + f_1(i)} \leq \ell + \sum_{j=0}^{\infty} \sum_{k=2^j\ell}^{2^{j+1}\ell-1} \prod_{i=\ell}^{\ell+k-1} \frac{f_1(i)}{\beta + f_1(i)} \leq \ell + \sum_{j=0}^{\infty} 2^j \ell \prod_{i=\ell}^{2^j\ell-1} \frac{f_1(i)}{\beta + f_1(i)} \\ &= \ell \left[1 + \sum_{j=0}^{\infty} 2^{\left(1-\frac{\beta}{C^*+2\epsilon}\right)j} \right] = \left(\frac{2-2^{\left(1-\frac{\beta}{C^*+2\epsilon}\right)}}{1-2^{\left(1-\frac{\beta}{C^*+2\epsilon}\right)}} \right) \ell \end{aligned}$$

where the sum converges as $C^* + 2\epsilon < \beta$. This proves the lemma. $\blacksquare$

Recall the class of characteristics $\mathcal{C}$ defined in (3.3). For given $\phi \in \mathcal{C}$ and initial values $\{\lambda_\ell^\phi(0) \in [0, 1] : \ell \geq 0\}$, define for each $\ell \geq 0$,

$$(9.5) \quad \lambda_\ell^\phi(t) = \lambda_\ell^\phi(0) + \int_0^t m_{f_1}^\phi(t-s) \mu_{f_1}^{(\ell)}(ds).$$

Note that this definition generalizes the expected aggregate ϕ -score of offsprings of a degree ℓ parent defined in Section 7 (see just before Theorem 7.1) in that we allow for a general initial value $\lambda_\ell^\phi(0) \in [0, 1]$. Hence, we keep the same notation. Two special instances of $\lambda_\ell^\phi(\cdot)$ that we have already used extensively are given by taking $\phi(t) = \mathbb{1}\{t \geq 0\}$, $t \geq 0$, $\lambda_\ell^\phi(0) = 1, \ell \geq 0$, which we denoted by $\lambda_\ell(\cdot)$, and $\phi(t) = \mathbb{1}\{\xi(t) = k\}$, $t \geq 0$, $\lambda_\ell^\phi(0) = \mathbb{P}(\xi_{f_1}^{(\ell)}(t) = k - \ell)$ for $\ell \geq 0, k \geq 0$, denoted by $\lambda_\ell^{(k)}(\cdot)$ (see (3.4)).

LEMMA 9.4. *Let $\phi \in \mathcal{C}$ such that $\lim_{t \rightarrow \infty} e^{-\lambda_1^* t} m_{f_1}^\phi(t) = c_\phi$. Recall $\lambda_\ell^\phi(\cdot)$ defined in (9.5). There is a constant $C > 0$ for which the following holds: for any $\epsilon > 0$, there exists $t(\epsilon) > 0$ such that for any $\ell \geq 0$,*

$$\sup_{t \geq t(\epsilon)} \left| e^{-\lambda_1^* t} \lambda_\ell^\phi(t) - w_\ell c_\phi \right| \leq C\epsilon(\ell + 1).$$

PROOF. In this proof, C, C', C'' will denote generic positive constants not depending on t, ℓ whose values might change from line to line. From (9.5) and the definition of w_ℓ , we have for any $t \geq 0$,

$$(9.6) \quad e^{-\lambda_1^* t} \lambda_\ell^\phi(t) - w_\ell c_\phi = \lambda_\ell^\phi(0) e^{-\lambda_1^* t} - c_\phi \int_t^\infty e^{-\lambda_1^* s} \mu_{f_1}^{(\ell)}(ds) + \int_0^t \left(e^{-\lambda_1^* (t-s)} m_{f_1}^\phi(t-s) - c_\phi \right) e^{-\lambda_1^* s} \mu_{f_1}^{(\ell)}(ds).$$

Choose any $\epsilon > 0$. Fix any $\vartheta > 0$ such that $\lambda_1^* - \vartheta > \beta_1$. As $\lim_{t \rightarrow \infty} e^{-\lambda_1^* t} m_{f_1}^\phi(t) = c_\phi$ and $\sup_{t < \infty} e^{-\lambda_1^* t} m_{f_1}^\phi(t) < \infty$ (which holds because the limit as $t \rightarrow \infty$ exists and as $\phi \in \mathcal{C}$, therefore for each $a > 0$, $\sup_{t \in [0, a]} m_{f_1}^\phi(t) \leq C \sup_{t \in [0, a]} m_{f_1}(t) < \infty$ by virtue of (7.1)), there exists $t_0 > 0$ such that for all $t \geq t_0$, $\left| e^{-\lambda_1^* t} m_{f_1}^\phi(t) - c_\phi \right| \leq \epsilon$ and $e^{-\vartheta t} \left(\sup_{z < \infty} e^{-\lambda_1^* z} m_{f_1}^\phi(z) + c_\phi \right) \leq \epsilon$. Thus, for any $t \geq 2t_0$,

$$\sup_{s \leq t} e^{-\vartheta s} \left| e^{-\lambda_1^* (t-s)} m_{f_1}^\phi(t-s) - c_\phi \right| \leq \epsilon.$$

Thus, applying Lemma 9.3 with $\beta = \lambda_1^* - \vartheta$, we conclude that for any $t \geq 2t_0$,

$$\begin{aligned} \int_0^t \left| e^{-\lambda_1^* (t-s)} m_{f_1}^\phi(t-s) - c_\phi \right| e^{-\lambda_1^* s} \mu_{f_1}^{(\ell)}(ds) &= \int_0^t e^{-\vartheta s} \left| e^{-\lambda_1^* (t-s)} m_{f_1}^\phi(t-s) - c_\phi \right| e^{-(\lambda_1^* - \vartheta)s} \mu_{f_1}^{(\ell)}(ds) \\ &\leq \epsilon w_\ell(\lambda_1^* - \vartheta) \leq C\epsilon(\ell + 1). \end{aligned}$$

Moreover, as $\int_0^\infty e^{-(\lambda_1^* - \theta)s} \mu_{f_1}^{(\ell)}(ds) \leq C(\ell + 1)$, for $t \geq 0$, $c_\phi \int_t^\infty e^{-\lambda_1^* s} \mu_{f_1}^{(\ell)}(ds) \leq C'(\ell + 1)e^{-\theta t}$. Using these in (9.6) and recalling $\lambda_\ell(0) \in [0, 1]$ for each ℓ , we obtain for $t \geq 2t_0$,

$$\left| e^{-\lambda_1^* t} \lambda_\ell^\phi(t) - w_\ell c_\phi \right| \leq e^{-\lambda_1^* t} + C'(\ell + 1)e^{-\theta t} + C\epsilon(\ell + 1).$$

Thus, there exists $t_1 \geq 2t_0$ such that for all $\ell \geq 0$ and all $t \geq t_1$, $\left| e^{-\lambda_1^* t} \lambda_\ell^\phi(t) - w_\ell c_\phi \right| \leq C''\epsilon(\ell + 1)$. $\blacksquare$

LEMMA 9.5. *Let $\phi \in \mathcal{C}$ such that $\lim_{t \rightarrow \infty} e^{-\lambda_1^* t} m_{f_1}^\phi(t) = c_\phi$. Fix any $\eta > 0, a \in \mathbb{R}$. Then as $n \rightarrow \infty$,*

$$n^{-(1+\eta\lambda_1^*)} \sum_{\ell=0}^{\infty} D_n(\ell, 0) \lambda_\ell^\phi(\eta \log n + a) \xrightarrow{P} c_\phi e^{\lambda_1^* a} \sum_{\ell=0}^{\infty} p_\ell^0 w_\ell.$$

PROOF. In this proof, once again C, C', C'' will denote generic positive constants not depending on n, t, ℓ whose values might change from line to line. Note that

$$(9.7) \quad \left| n^{-(1+\eta\lambda_1^*)} \sum_{\ell=0}^{\infty} D_n(\ell, 0) \lambda_\ell^\phi(\eta \log n + a) - c_\phi e^{\lambda_1^* a} \sum_{\ell=0}^{\infty} p_\ell^0 w_\ell \right| \\ \leq n^{-1} \sum_{\ell=0}^{\infty} D_n(\ell, 0) \left| \lambda_\ell^\phi(\eta \log n + a) n^{-\eta\lambda_1^*} - w_\ell c_\phi e^{\lambda_1^* a} \right| + c_\phi e^{\lambda_1^* a} \left| \sum_{\ell=0}^{\infty} n^{-1} D_n(\ell, 0) w_\ell - \sum_{\ell=0}^{\infty} p_\ell^0 w_\ell \right|.$$

To show that the second term goes to zero in probability, consider the characteristic $\chi(t) = \sum_{\ell=0}^{\infty} w_\ell \mathbb{1}_{\{\xi_{f_1}(t) = \ell\}}$. By Lemma 9.3, $w_\ell \leq C(\ell + 1)$ and hence, $\chi \in \mathcal{C}$. Thus, by Lemma 6.6 (i),

$$(9.8) \quad \left| \sum_{\ell=0}^{\infty} n^{-1} D_n(\ell, 0) w_\ell - \sum_{\ell=0}^{\infty} p_\ell^0 w_\ell \right| \xrightarrow{P} 0 \quad \text{as } n \rightarrow \infty.$$

To show that the first term in the bound (9.7) goes to zero in probability, take any $\epsilon > 0$. Recalling

$\sum_{\ell=0}^{\infty} D_n(\ell, 0) = n$ and $\sum_{\ell=0}^{\infty} (\ell + 1) D_n(\ell, 0) = 2n - 1$, and taking $t = \eta \log n + a$ for any $n \geq e^{(t(\epsilon) - a)/\eta}$ in Lemma 9.4, we obtain

$$\sum_{\ell=0}^{\infty} n^{-1} D_n(\ell, 0) \left| \lambda_\ell^\phi(\eta \log n + a) n^{-\eta\lambda_1^*} - w_\ell c_\phi e^{\lambda_1^* a} \right| \leq n^{-1} C'' e^{\lambda_1^* a} \epsilon \sum_{\ell=0}^{\infty} (\ell + 1) D_n(\ell, 0) \leq 2C'' e^{\lambda_1^* a} \epsilon.$$

As $\epsilon > 0$ is arbitrary, the first term in (9.7) converges to zero as $n \rightarrow \infty$ and completes the proof. $\blacksquare$

Define $m_1^* := \int_0^\infty u e^{-\lambda_1^* u} \mu_{f_1}(du)$. For $\ell \geq 0, k \geq 0$, recall $\lambda_\ell(\cdot)$ and $\lambda_\ell^{(k)}(\cdot)$ from (3.4).

COROLLARY 9.6. *Fix any $\eta > 0$ and $k \geq 0$. Then as $n \rightarrow \infty$*

$$n^{-(1+\eta\lambda_1^*)} \sum_{\ell=0}^{\infty} D_n(\ell, 0) \lambda_\ell(\eta \log n) \xrightarrow{P} \sum_{\ell=0}^{\infty} p_\ell^0 w_\ell / \lambda_1^* m_1^*,$$

$$n^{-(1+\eta\lambda_1^*)} \sum_{\ell=0}^{\infty} D_n(\ell, 0) \lambda_\ell^{(k)}(\eta \log n) \xrightarrow{P} p_k^1 \sum_{\ell=0}^{\infty} p_\ell^0 w_\ell / \lambda_1^* m_1^*.$$

PROOF. Follows from Lemma 9.5 upon using the explicit formulas

$$\lambda_\ell(t) = 1 + \int_0^t m_{f_1}(t-s) \mu_{f_1}^{(\ell)}(ds), \quad \lambda_\ell^{(k)}(t) = \mathbb{P}\left(\xi_{f_1}^{(\ell)}(t) = k - \ell\right) + \int_0^t m_{f_1}^{(k)}(t-s) \mu_{f_1}^{(\ell)}(ds), \quad t \geq 0,$$

and observing by Lemma 6.6 (ii)

$$(9.9) \quad \lim_{t \rightarrow \infty} e^{-\lambda_1^* t} m_{f_1}(t) = (\lambda_1^* m_1^*)^{-1}, \quad \lim_{t \rightarrow \infty} e^{-\lambda_1^* t} m_{f_1}^{(k)}(t) = p_k^1 (\lambda_1^* m_1^*)^{-1}.$$

■

LEMMA 9.7. *There exists $\eta_0 > 0$ such that for any $\eta \leq \eta_0$, the following limits hold as $n \rightarrow \infty$:*

- (i) $n^{-(1+\eta\lambda_1^*)} Z_n(\eta \log n) \xrightarrow{P} \sum_{\ell=0}^{\infty} p_\ell^0 w_\ell / \lambda_1^* m_1^*,$
- (ii) *For any $k \geq 0$, $n^{-(1+\eta\lambda_1^*)} D_n(k, \eta \log n) \xrightarrow{P} p_k^1 \sum_{\ell=0}^{\infty} p_\ell^0 w_\ell / \lambda_1^* m_1^*.$*

PROOF. (i) and (ii) follow from (9.2) and (9.1) respectively along with Corollary 9.6. ■

COROLLARY 9.8. $\sum_{\ell=0}^{\infty} p_\ell^1 w_\ell = \lambda_1^* m_1^*.$

PROOF. Note that Lemma 9.7 (i) holds in the special case where f_0 is taken to be f_1 (the model without change point). In this case, $p_\ell^0 = p_\ell^1$ for all $\ell \geq 0$. By Lemma 6.6 (ii), $Z_n(\eta_0 \log n) e^{-\lambda_1^* (T_n + \eta_0 \log n)} \xrightarrow{a.s.} W_\infty / \lambda_1^* m_1^*$. Moreover, as $Z(T_n) = n$, therefore, applying Lemma 6.6 (ii) again, $n^{-1} e^{\lambda_1^* T_n} = e^{\lambda_1^* T_n} / Z(T_n) \xrightarrow{a.s.} \lambda_1^* m_1^* / W_\infty$. Using these observations, we obtain

$$n^{-(1+\eta_0\lambda_1^*)} Z_n(\eta_0 \log n) = n^{-1} e^{\lambda_1^* T_n} Z_n(\eta_0 \log n) e^{-\lambda_1^* (T_n + \eta_0 \log n)} \xrightarrow{a.s.} 1.$$

Comparing this with Lemma 9.7 (i) with $f_0 = f_1$ gives the result. ■

Recall that for any $k \geq 0$, $\xi_{f_1}^{(k)}(\cdot)$ is the point process denoting the distribution of birth times of children of a vertex which is of degree k at time zero. The following lemma gives an estimate on the second moment of $\xi_{f_1}^{(k)}(t)$ under Assumption 3.1.

LEMMA 9.9. *There exists $C > 0$ and $\beta' < \lambda_1^*$ such that for any $k \geq 0, t \geq 0$, $\mathbb{E}\left(\xi_{f_1}^{(k)}(t)\right)^2 \leq C(k+1)^2 e^{2\beta' t}.$*

PROOF. By Assumption 3.1 and Lemma 9.1, for any $\beta' \in (\beta_1, \lambda_1^*)$, there exists $\ell_0 \geq 0$ such that for all $\ell \geq \ell_0$, $f_1(\ell) \leq \beta' \ell$. Let $m = \max_{\ell \leq \ell_0} f_1(\ell)$. It is clear that $\xi_{f_1}^{(k)}(\cdot)$ is stochastically dominated by the offspring process of a continuous time branching process with linear attachment function $f^*(\ell) = \beta' \ell + 1 + (m + \beta' k)$, $\ell \geq 0$. Applying the second moment obtained in Lemma 6.4 (with $\nu = \beta'$ and $\kappa = 1 + m + \beta'(k-1)$) the lemma follows. ■

For $\eta > 0, j \geq 0$, let $D_n(k, j, \eta)$ denote the number of vertices of degree k at time $(j+1)\eta \log n$ that were born before time $j\eta \log n$ (including possibly the ones at time zero).

LEMMA 9.10. *For any $\eta > 0, j \geq 0$, as $n \rightarrow \infty$, $\sum_{k=0}^{\infty} (k+1) D_n(k, j, \eta) / \left(Z_n(j\eta \log n) n^{\lambda_1^* \eta} \right) \xrightarrow{P} 0.$*

PROOF. We will condition on $\mathcal{F}_n(j\eta \log n)$ throughout the proof. For $1 \leq m \leq D_n(\ell, j\eta \log n)$, denote by $\xi_{f_1, m}^{(\ell)}(t)$ the degree at time $t + j\eta \log n$ of the m -th vertex of degree ℓ at time $j\eta \log n$. Observe that

$$\begin{aligned} \sum_{k=0}^{\infty} (k+1) D_n(k, j, \eta) &= \sum_{k=0}^{\infty} (k+1) \sum_{\ell=0}^k \sum_{m=1}^{D_n(\ell, j\eta \log n)} \mathbb{1}_{\{\xi_{f_1, m}^{(\ell)}(\eta \log n) = k - \ell\}} \\ &= \sum_{\ell=0}^{\infty} \sum_{m=1}^{D_n(\ell, j\eta \log n)} \sum_{k=\ell}^{\infty} (k+1) \mathbb{1}_{\{\xi_{f_1, m}^{(\ell)}(\eta \log n) = k - \ell\}} = \sum_{\ell=0}^{\infty} \sum_{m=1}^{D_n(\ell, j\eta \log n)} (\ell + 1 + \xi_{f_1, m}^{(\ell)}(\eta \log n)) \\ &= \sum_{\ell=0}^{\infty} (\ell + 1) D_n(\ell, j\eta \log n) + \sum_{\ell=0}^{\infty} \sum_{m=1}^{D_n(\ell, j\eta \log n)} \xi_{f_1, m}^{(\ell)}(\eta \log n) \\ &= 2Z_n(j\eta \log n) - 1 + \sum_{\ell=0}^{\infty} \sum_{m=1}^{D_n(\ell, j\eta \log n)} \xi_{f_1, m}^{(\ell)}(\eta \log n). \end{aligned}$$

Thus, it suffices to show that as $n \rightarrow \infty$,

$$(9.10) \quad \frac{1}{Z_n(j\eta \log n)} \sum_{\ell=0}^{\infty} \sum_{m=1}^{D_n(\ell, j\eta \log n)} \frac{1}{n^{\lambda_1^* \eta}} \xi_{f_1, m}^{(\ell)}(\eta \log n) \xrightarrow{P} 0.$$

Note that using Lemma 9.9,

$$\begin{aligned} &\text{Var} \left(\frac{1}{Z_n(j\eta \log n)} \sum_{\ell=0}^{\infty} \sum_{m=1}^{D_n(\ell, j\eta \log n)} \frac{1}{n^{\lambda_1^* \eta}} \xi_{f_1, m}^{(\ell)}(\eta \log n) \right) \\ &\leq \frac{1}{Z^2(j\eta \log n) n^{2\lambda_1^* \eta}} \sum_{\ell=0}^{\infty} \sum_{m=1}^{D_n(\ell, j\eta \log n)} \mathbb{E} \left(\xi_{f_1, m}^{(\ell)}(\eta \log n) \right)^2 \leq \frac{C n^{2\beta' \eta}}{Z^2(j\eta \log n) n^{2\lambda_1^* \eta}} \sum_{\ell=0}^{\infty} (\ell+1)^2 D_n(\ell, j\eta \log n). \end{aligned}$$

Denoting the maximum degree at time $j\eta \log n$ of the branching process by $D^{\max}$, note that $D^{\max} + 1 \leq Z_n(j\eta \log n)$ and hence,

$$\sum_{\ell=0}^{\infty} (\ell+1)^2 D_n(\ell, j\eta \log n) \leq (D^{\max} + 1) \sum_{\ell=0}^{\infty} (\ell+1) D_n(\ell, j\eta \log n) \leq Z_n(j\eta \log n) (2Z_n(j\eta \log n) - 1).$$

Using this in the above variance bound, we get

$$\text{Var} \left(\frac{1}{Z_n(j\eta \log n)} \sum_{\ell=0}^{\infty} \sum_{m=1}^{D_n(\ell, j\eta \log n)} \frac{1}{n^{\lambda_1^* \eta}} \xi_{f_1, m}^{(\ell)}(\eta \log n) \right) \leq \frac{2C n^{2\beta' \eta} Z^2(j\eta \log n)}{Z^2(j\eta \log n) n^{2\lambda_1^* \eta}} = \frac{2C}{n^{2(\lambda_1^* - \beta') \eta}} \rightarrow 0$$

as $n \rightarrow \infty$ and hence,

$$(9.11) \quad \begin{aligned} &\frac{1}{Z_n(j\eta \log n)} \sum_{\ell=0}^{\infty} \sum_{m=1}^{D_n(\ell, j\eta \log n)} \frac{1}{n^{\lambda_1^* \eta}} \xi_{f_1, m}^{(\ell)}(\eta \log n) \\ &\quad - \frac{1}{Z_n(j\eta \log n)} \sum_{\ell=0}^{\infty} \sum_{m=1}^{D_n(\ell, j\eta \log n)} \frac{1}{n^{\lambda_1^* \eta}} \mathbb{E} \left(\xi_{f_1, m}^{(\ell)}(\eta \log n) \right) \xrightarrow{P} 0. \end{aligned}$$

By Lemma 9.3, we obtain $\beta \in (\lambda_1^* - 1, \lambda_1^*)$ such that $w_{\ell}(\beta) = \int_0^{\infty} e^{-\beta s} \mu_{f_1}^{(\ell)}(ds) \leq C(\beta)(\ell + 1)$.

This implies for any m, ℓ , $\mathbb{E} \left(\xi_{f_1, m}^{(\ell)}(\eta \log n) \right) \leq C(\beta) n^{\beta \eta} (\ell + 1)$ and consequently,

$$(9.12) \quad \frac{1}{Z_n(j\eta \log n)} \sum_{\ell=0}^{\infty} \sum_{m=1}^{D_n(\ell, j\eta \log n)} \frac{1}{n^{\lambda_1^* \eta}} \mathbb{E} \left(\xi_{f_1, m}^{(\ell)}(\eta \log n) \right)$$

$$\leq \frac{1}{n^{(\lambda_1^* - \beta)\eta}} \frac{C(\beta)}{Z_n(j\eta \log n)} \sum_{\ell=0}^{\infty} (\ell+1) D_n(\ell, j\eta \log n) \leq \frac{2C(\beta)}{n^{(\lambda_1^* - \beta)\eta}} \rightarrow 0$$

as $n \rightarrow \infty$. From (9.11) and (9.12), the proof of (9.10), and hence the lemma, is complete. $\blacksquare$

LEMMA 9.11. *Let $\phi \in \mathcal{C}$ such that $\lim_{t \rightarrow \infty} e^{-\lambda_1^* t} m^\phi(t) = c_\phi$. Fix any $\eta_0 \in (0, 1/(2C'))$, where C' is the constant appearing in Theorem 7.1. Then for any $j \geq 0$, $\eta \in (0, \eta_0]$ and $a \in \mathbb{R}$, as $n \rightarrow \infty$:*

$$(9.13) \quad \frac{1}{n^{1+(j\eta_0+\eta)\lambda_1^*}} \sum_{\ell=0}^{\infty} D_n(\ell, j\eta_0 \log n) \lambda_\ell^\phi(\eta \log n + a) \xrightarrow{P} c_\phi e^{\lambda_1^* a} \sum_{\ell=0}^{\infty} p_\ell^0 w_\ell.$$

PROOF. We will proceed by induction on $j \geq 0$. Suppose for some $j' \geq 0$, (9.13) holds for all $0 \leq j \leq j'$, $\eta \in (0, \eta_0]$ and $a \in \mathbb{R}$. Taking $\phi(t) = \mathbb{1}\{t \geq 0\}$ and $\eta = \eta_0$ and recalling $\lim_{t \rightarrow \infty} e^{-\lambda_1^* t} m_{f_1}(t) = \frac{1}{\lambda_1^* m_1^*}$, we obtain for any $0 \leq j \leq j'$ and any $a \in \mathbb{R}$,

$$(9.14) \quad \frac{1}{n^{1+(j+1)\eta_0\lambda_1^*}} Z_n((j+1)\eta_0 \log n + a) \xrightarrow{P} \frac{1}{\lambda_1^* m_1^*} e^{\lambda_1^* a} \sum_{\ell=0}^{\infty} p_\ell^0 w_\ell.$$

Fix any $\phi \in \mathcal{C}$. Note that for any $\eta \leq \eta_0$,

$$(9.15) \quad \left| \frac{1}{n^{1+((j'+1)\eta_0+\eta)\lambda_1^*}} \sum_{\ell=0}^{\infty} D_n(\ell, (j'+1)\eta_0 \log n) \lambda_\ell^\phi(\eta \log n + a) - c_\phi e^{\lambda_1^* a} \sum_{\ell=0}^{\infty} p_\ell^0 w_\ell \right|$$

$$\leq \sum_{\ell=0}^{\infty} \frac{D_n(\ell, (j'+1)\eta_0 \log n)}{n^{1+(j'+1)\eta_0\lambda_1^*}} \left| \frac{\lambda_\ell^\phi(\eta \log n + a)}{n^{\eta\lambda_1^*}} - c_\phi e^{\lambda_1^* a} w_\ell \right|$$

$$+ c_\phi e^{\lambda_1^* a} \left| \sum_{\ell=0}^{\infty} \frac{D_n(\ell, (j'+1)\eta_0 \log n)}{n^{1+(j'+1)\eta_0\lambda_1^*}} w_\ell - \sum_{\ell=0}^{\infty} p_\ell^0 w_\ell \right|.$$

For any $\epsilon > 0$, by Lemma 9.4, there exists $n_0 \geq 1$ and $C'' > 0$ such that for all $n \geq n_0$, $\ell \geq 0$,

$$\left| \frac{\lambda_\ell^\phi(\eta \log n + a)}{n^{\eta\lambda_1^*}} - c_\phi e^{\lambda_1^* a} w_\ell \right| \leq C'' e^{\lambda_1^* a} \epsilon (\ell + 1)$$

and hence,

$$\sum_{\ell=0}^{\infty} \frac{D_n(\ell, (j'+1)\eta_0 \log n)}{n^{1+(j'+1)\eta_0\lambda_1^*}} \left| \frac{\lambda_\ell^\phi(\eta \log n + a)}{n^{\eta\lambda_1^*}} - c_\phi e^{\lambda_1^* a} w_\ell \right|$$

$$\leq C'' e^{\lambda_1^* a} \epsilon \sum_{\ell=0}^{\infty} \frac{(\ell+1) D_n(\ell, (j'+1)\eta_0 \log n)}{n^{1+(j'+1)\eta_0\lambda_1^*}} \leq 2C'' e^{\lambda_1^* a} \epsilon \frac{Z_n((j'+1)\eta_0 \log n)}{n^{1+(j'+1)\eta_0\lambda_1^*}}.$$

Therefore, using (9.14) with $j = j'$, and as $\epsilon > 0$ is arbitrary, the first term in the bound (9.15) converges to zero in probability. To estimate the second term in (9.15), consider the characteristic $\chi(t) = \sum_{\ell=0}^{\infty} w_\ell \mathbb{1}\{\xi_{f_1}(t) = \ell\}$ and note that by Lemma 9.3, $\chi \in \mathcal{C}$. Recall Z_n^χ from Section 7 (see Notation (iv)) with $\mathcal{F}_n(0)$ replaced by $\mathcal{F}_n(j'\eta_0 \log n)$ (that is, time starting at $T_n + j'\eta_0 \log n$) and take $a = \eta_0 \log n$. As Z_n^χ denotes the aggregate χ -score of all vertices born in the interval $[j'\eta_0 \log n, (j'+1)\eta_0 \log n]$,

(9.16)

$$\begin{aligned} \frac{1}{n^{1+(j'+1)\eta_0\lambda_1^*}} \left| \sum_{\ell=0}^{\infty} D_n(\ell, (j'+1)\eta_0 \log n) w_\ell - Z_n^\chi \right| &\leq \frac{C(\lambda_1^*)}{n^{1+(j'+1)\eta_0\lambda_1^*}} \sum_{\ell=0}^{\infty} (\ell+1) D_n(\ell, j', \eta_0) \\ &= \frac{Z_n(j' \eta_0 \log n)}{n^{1+j' \eta_0 \lambda_1^*}} \frac{C(\lambda_1^*)}{Z_n(j' \eta_0 \log n) n^{\eta_0 \lambda_1^*}} \sum_{\ell=0}^{\infty} (\ell+1) D_n(\ell, j', \eta_0) \xrightarrow{P} 0 \end{aligned}$$

as $n \rightarrow \infty$, which follows from Lemma 9.10 and by (9.14) with $j = j' - 1$ for $j' \geq 1$ and the trivial observation that $\frac{Z_n(j' \eta_0 \log n)}{n^{1+j' \eta_0 \lambda_1^*}} \xrightarrow{P} 0$ when $j' = 0$. Here $C(\lambda_1^*)$ is the constant appearing in Lemma 9.3. Recall η_0 is chosen such that $\frac{C e^{C' \eta_0 \log n}}{\sqrt{n}} \rightarrow 0$ as $n \rightarrow \infty$, where C, C' are the constants appearing in Theorem 7.1. Thus, recalling $\lambda_\ell^\chi(t) = \int_0^t m_{f_1}^\chi(t-s) \mu_{f_1}^{(\ell)}(ds)$, by Theorem 7.1 and (9.14),

(9.17)

$$\begin{aligned} \frac{1}{n^{1+(j'+1)\eta_0\lambda_1^*}} \left| Z_n^\chi - \sum_{\ell=0}^{\infty} D_n(\ell, j' \eta_0 \log n) \lambda_\ell^\chi(\eta_0 \log n) \right| &\leq \frac{C e^{C' \eta_0 \log n}}{n^{1+(j'+1)\eta_0\lambda_1^*}} \sqrt{Z_n(j' \eta_0 \log n)} \\ &\leq \frac{C e^{C' \eta_0 \log n}}{\sqrt{n}} \sqrt{\frac{Z_n(j' \eta_0 \log n)}{n^{1+(j'+1)\eta_0\lambda_1^*}}} \xrightarrow{P} 0. \end{aligned}$$

By (9.16) and (9.17), we obtain

$$\begin{aligned} (9.18) \quad &\left| \sum_{\ell=0}^{\infty} \frac{D_n(\ell, (j'+1)\eta_0 \log n)}{n^{1+(j'+1)\eta_0\lambda_1^*}} w_\ell - \sum_{\ell=0}^{\infty} \frac{D_n(\ell, j' \eta_0 \log n)}{n^{1+(j'+1)\eta_0\lambda_1^*}} \lambda_\ell^\chi(\eta_0 \log n) \right| \\ &\leq \frac{1}{n^{1+(j'+1)\eta_0\lambda_1^*}} \left| \sum_{\ell=0}^{\infty} D_n(\ell, (j'+1)\eta_0 \log n) w_\ell - Z_n^\chi \right| \\ &\quad + \frac{1}{n^{1+(j'+1)\eta_0\lambda_1^*}} \left| Z_n^\chi - \sum_{\ell=0}^{\infty} D_n(\ell, j' \eta_0 \log n) \lambda_\ell^\chi(\eta_0 \log n) \right| \xrightarrow{P} 0. \end{aligned}$$

Next, we will show that

$$(9.19) \quad e^{-\lambda_1^* t} m_{f_1}^\chi(t) \rightarrow 1 \text{ as } t \rightarrow \infty.$$

To see this, first note that it follows from Assumption 2.4 (ii) that there exists $\beta < \lambda_1^*$ such that $\mathbb{E}(\xi_{f_1}(t)) \leq C e^{\beta t}$. Moreover, $w_\ell \leq C(\ell+1)$ for all $\ell \geq 0$. These observations imply

$$\begin{aligned} \sum_{k=0}^{\infty} \sup_{t \in [k, k+1]} \left[e^{-\lambda_1^* t} \mathbb{E}(\chi(t)) \right] &\leq C \sum_{k=0}^{\infty} \sup_{t \in [k, k+1]} \left[e^{-\lambda_1^* t} \sum_{\ell=0}^{\infty} (\ell+1) \mathbb{P}(\xi_{f_1}(t) = \ell) \right] \\ &= C \sum_{k=0}^{\infty} \sup_{t \in [k, k+1]} \left[e^{-\lambda_1^* t} \mathbb{E}(\xi_{f_1}(t) + 1) \right] \leq C' \sum_{k=0}^{\infty} \sup_{t \in [k, k+1]} \left[e^{-\lambda_1^* t} e^{\beta t} \right] \leq C' e^{\beta} \sum_{k=0}^{\infty} e^{-(\lambda_1^* - \beta)k} < \infty \end{aligned}$$

where $C, C' > 0$ are constants. Thus, by Proposition 2.2 of [37] and Corollary 9.8, it follows that

$$\lim_{t \rightarrow \infty} e^{-\lambda_1^* t} m_{f_1}^\chi(t) = \frac{1}{\lambda_1^* m_1^*} \sum_{\ell=0}^{\infty} w_\ell \lambda_1^* \int_0^\infty e^{-\lambda_1^* s} \mathbb{P}(\xi_{f_1}(s) = \ell) ds = \frac{1}{\lambda_1^* m_1^*} \sum_{\ell=0}^{\infty} w_\ell p_\ell^1 = 1.$$

Using this, the definition of λ_ℓ^χ , the fact that $\chi \in \mathcal{C}$ and the induction hypothesis, we obtain

$$(9.20) \quad \frac{1}{n^{1+(j'+1)\eta_0\lambda_1^*}} \sum_{\ell=0}^{\infty} D_n(\ell, j' \eta_0 \log n) \lambda_\ell^\chi(\eta_0 \log n) \xrightarrow{P} \sum_{\ell=0}^{\infty} p_\ell^0 w_\ell \text{ as } n \rightarrow \infty.$$

From (9.18) and (9.20), the second term in (9.15) goes to 0 in probability as $n \rightarrow \infty$ which shows

$$\left| \frac{1}{n^{1+((j'+1)\eta_0+\eta)\lambda_1^*}} \sum_{\ell=0}^{\infty} D_n(\ell, (j'+1)\eta_0 \log n) \lambda_{\ell}^{\phi} (\eta \log n + a) - c_{\phi} e^{\lambda_1^* a} \sum_{\ell=0}^{\infty} p_{\ell}^0 w_{\ell} \right| \xrightarrow{P} 0$$

establishing (9.13) for all $j \leq j' + 1$. (9.13) holds for $j = 0$ by Lemma 9.5. Thus, the lemma is proved. $\blacksquare$

LEMMA 9.12. *For any $k \geq 0, \theta > 0$ and $a \in \mathbb{R}$, as $n \rightarrow \infty$:*

$$n^{-(1+\theta\lambda_1^*)} Z_n(\theta \log n + a) \xrightarrow{P} e^{\lambda_1^* a} \sum_{\ell=0}^{\infty} p_{\ell}^0 w_{\ell} / \lambda_1^* m_1^*, \quad \frac{D_n(k, \theta \log n + a)}{Z_n(\theta \log n + a)} \xrightarrow{P} p_k^1.$$

PROOF. Note for any $\eta_0 > 0$ we can write $\theta = j\eta_0 + \eta$ for some $j \geq 0$. The first assertion follows by the argument used to derive (9.14). To prove the second assertion, fix any $k \geq 0$. Take $\eta_0 > 0$ in Lemma 9.11 small enough so that $C e^{C'\eta_0 \log n} \epsilon^{-2} n^{-(\omega-\tilde{\theta}-\frac{1}{2})} \rightarrow 0$, where $C, C', \omega, \tilde{\theta}$ are as in Lemma 8.11. Recall that the bound obtained in Lemma 8.11 conditionally on $\mathcal{F}_n(0)$ was in terms of deterministic constants and n , the total number of vertices at time 0. Replacing $\mathcal{F}_n(0)$ by $\mathcal{F}_n(j\eta_0 \log n)$ and time starting from $T_n + j\eta_0 \log n$, Lemma 8.11 (with n replaced by $Z_n(j\eta_0 \log n)$, the total number of vertices at time $j\eta_0 \log n$) implies,

$$\frac{1}{Z_n(j\eta_0 \log n)} D_n(k, \theta \log n + a) - \frac{1}{Z_n(j\eta_0 \log n)} \sum_{\ell=0}^{\infty} D_n(\ell, j\eta_0 \log n) \lambda_{\ell}^{(k)} (\eta \log n + a) \xrightarrow{P} 0, \text{ as } n \rightarrow \infty.$$

From Lemma 9.11 (taking $\phi(t) = \mathbb{1}\{t \geq 0\}$), $Z_n(j\eta_0 \log n) / Z_n(\theta \log n + a) \xrightarrow{P} 0$ if $\eta > 0$, and $Z_n(j\eta_0 \log n) / Z_n(\theta \log n + a) \xrightarrow{P} e^{-\lambda_1^* a}$ if $\eta = 0$ and thus, multiplying both sides of the above by

$Z_n(j\eta_0 \log n) / Z_n(\theta \log n + a)$, we obtain

$$(9.21) \quad \frac{D_n(k, \theta \log n + a)}{Z_n(\theta \log n + a)} - \frac{1}{Z_n(\theta \log n + a)} \sum_{\ell=0}^{\infty} D_n(\ell, j\eta_0 \log n) \lambda_{\ell}^{(k)} (\eta \log n + a) \xrightarrow{P} 0, \text{ as } n \rightarrow \infty.$$

Taking $\phi(t) = \mathbb{1}\{\xi_{f_1}(t) = k\}$, we see that $\lambda_{\ell}^{\phi} = \lambda_{\ell}^{(k)}$ for each $\ell \geq 0$. Moreover, recall from (9.9) $\lim_{t \rightarrow \infty} e^{-\lambda_1^* t} m_{f_1}^{(k)}(t) = p_k^1 / \lambda_1^* m_1^*$. Thus, from Lemma 9.11,

$$(9.22) \quad \frac{1}{n^{1+\theta\lambda_1^*}} \sum_{\ell=0}^{\infty} D_n(\ell, j\eta_0 \log n) \lambda_{\ell}^{(k)} (\eta \log n + a) \xrightarrow{P} \frac{p_k^1}{\lambda_1^* m_1^*} e^{\lambda_1^* a} \sum_{\ell=0}^{\infty} p_{\ell}^0 w_{\ell}.$$

Using (9.22) and the first assertion of the lemma in (9.21), the second assertion follows. $\blacksquare$

Let $a_0 := \frac{1}{\lambda_1^*} \log \left(\frac{\lambda_1^* m_1^*}{\sum_{\ell=0}^{\infty} p_{\ell}^0 w_{\ell}} \right)$ and $T_n^{\theta} := T_{n^{1+\lambda_1^* \theta}}$ be the first time the branching process has $n^{1+\lambda_1^* \theta}$ vertices.

LEMMA 9.13. $T_n^{\theta} - \theta \log n \xrightarrow{P} a_0$.

PROOF. Follows immediately from the first assertion of Lemma 9.12. $\blacksquare$

THEOREM 9.14. *For any $k \geq 0, \theta > 0$, as $n \rightarrow \infty$, $n^{-(1+\lambda_1^* \theta)} D_n(k, T_n^{\theta}) \xrightarrow{P} p_k^1$.*

PROOF. In the proof, we will abbreviate $z^* = \frac{1}{\lambda_1^* m_1^*} \sum_{\ell=0}^{\infty} p_{\ell}^0 w_{\ell}$. Fix any $k \geq 0$, $\theta > 0$. Take any $\epsilon \in (0, 1)$. By the same argument as in the proof of Lemma 8.8,

(9.23)

$$\sup_{t \leq 2\epsilon} |D_n(k, \theta \log n + a_0 - \epsilon + t) - D_n(k, \theta \log n + a_0 - \epsilon)| \leq (Z_n(\theta \log n + a_0 + \epsilon) - Z_n(\theta \log n + a_0 - \epsilon)) + Y_n.$$

where, conditionally on $\mathcal{F}_n(\theta \log n + a_0 - \epsilon)$, Y_n has the same distribution as the random variable $\sum_{\ell=0}^k \text{Bin}(D_n(\ell, \theta \log n + a_0 - \epsilon), q_{\ell}(2\epsilon))$. Observe that by the first assertion in Lemma 9.12, for small enough ϵ ,

$$(9.24) \quad n^{-(1+\lambda_1^*\theta)} (Z_n(\theta \log n + a_0 + \epsilon) - Z_n(\theta \log n + a_0 - \epsilon)) \xrightarrow{P} e^{\lambda_1^*\epsilon} - e^{-\lambda_1^*\epsilon} \leq 4\lambda_1^*\epsilon.$$

Note that for any $C > 0$,

$$(9.25) \quad \mathbb{P}(Y_n > C\sqrt{\epsilon}n^{1+\lambda_1^*\theta}) \leq \mathbb{P}(Y_n > C\sqrt{\epsilon}n^{1+\lambda_1^*\theta}, Z_n(\theta \log n + a_0 - \epsilon) \leq \epsilon^{-1/2}n^{1+\lambda_1^*\theta}) \\ + \mathbb{P}(Z_n(\theta \log n + a_0 - \epsilon) > \epsilon^{-1/2}n^{1+\lambda_1^*\theta}).$$

For ϵ sufficiently small, by the first assertion of Lemma 9.12, as $n \rightarrow \infty$,

$$(9.26) \quad \mathbb{P}(Z_n(\theta \log n + a_0 - \epsilon) > \epsilon^{-1/2}n^{1+\lambda_1^*\theta}) \rightarrow 0.$$

Let $\mathcal{H}_n := \mathcal{F}_n(\theta \log n + a_0 - \epsilon)$. Using Lemma 8.3,

$$\mathbb{E}(Y_n | \mathcal{H}_n) = \sum_{\ell=0}^k D_n(\ell, \theta \log n + a_0 - \epsilon) q_{\ell}(2\epsilon) \leq C'\epsilon \sum_{\ell=0}^k (\ell+1) D_n(\ell, \theta \log n + a_0 - \epsilon) \\ \leq 2C'\epsilon Z_n(\theta \log n + a_0 - \epsilon).$$

Thus, choosing $C > 4C'$, using Chebychev's inequality, conditionally on $\mathcal{H}_n$ on the event $\{Z_n(\theta \log n + a_0 - \epsilon) \leq \epsilon^{-1/2}n^{1+\lambda_1^*\theta}\}$,

$$(9.27) \quad \mathbb{P}(Y_n > C\sqrt{\epsilon}n^{1+\lambda_1^*\theta} | \mathcal{H}_n) \leq \mathbb{P}\left(Y_n - \mathbb{E}(Y_n | \mathcal{H}_n) > \frac{C}{2}\sqrt{\epsilon}n^{1+\lambda_1^*\theta} | \mathcal{H}_n\right) \\ \leq \frac{4\text{Var}(Y_n | \mathcal{H}_n)}{C^2\epsilon n^{2(1+\lambda_1^*\theta)}} = \frac{4\sum_{\ell=0}^k D_n(\ell, \theta \log n + a_0 - \epsilon) q_{\ell}(2\epsilon) (1 - q_{\ell}(2\epsilon))}{C^2\epsilon n^{2(1+\lambda_1^*\theta)}} \\ \leq \frac{4C'\epsilon \sum_{\ell=0}^k (\ell+1) D_n(\ell, \theta \log n + a_0 - \epsilon)}{C^2\epsilon n^{2(1+\lambda_1^*\theta)}} \leq \frac{8C'Z_n(\theta \log n + a_0 - \epsilon)}{C^2n^{2(1+\lambda_1^*\theta)}} \leq \frac{8C'}{C^2\sqrt{\epsilon}n^{1+\lambda_1^*\theta}} \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

Using (9.26) and (9.27) in (9.25), we conclude

$$(9.28) \quad \mathbb{P}(Y_n > C\sqrt{\epsilon}n^{1+\lambda_1^*\theta}) \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

Using (9.24), (9.28) and (9.23), we conclude that there exist $C_0 > 0, \epsilon_0 > 0$ such that for all $\epsilon \in (0, \epsilon_0)$,

(9.29)

$$\mathbb{P}\left(\sup_{t \leq 2\epsilon} |D_n(k, \theta \log n + a_0 - \epsilon + t) - D_n(k, \theta \log n + a_0 - \epsilon)| > C_0\sqrt{\epsilon}n^{1+\lambda_1^*\theta}\right) \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

From (9.29) and Lemma 9.13, as $n \rightarrow \infty$,

$$(9.30) \quad \mathbb{P}\left(|D_n(k, T_n^{\theta}) - D_n(k, \theta \log n + a_0 - \epsilon)| > C_0\sqrt{\epsilon}n^{1+\lambda_1^*\theta}\right) \leq \mathbb{P}\left(|T_n^{\theta} - \theta \log n - a_0| > 2\epsilon\right) \\ + \mathbb{P}\left(\sup_{t \leq 2\epsilon} |D_n(k, \theta \log n + a_0 - \epsilon + t) - D_n(k, \theta \log n + a_0 - \epsilon)| > C_0\sqrt{\epsilon}n^{1+\lambda_1^*\theta}\right) \rightarrow 0.$$

For any $\epsilon > 0$,

$$(9.31) \quad \mathbb{P} \left(\left| \frac{D_n(k, T_n^\theta)}{n^{1+\lambda_1^* \theta}} - p_k^1 \right| > 2C_0 \sqrt{\epsilon} \right) \leq \mathbb{P} \left(\left| \frac{D_n(k, T_n^\theta)}{n^{1+\lambda_1^* \theta}} - \frac{D_n(k, \theta \log n + a_0 - \epsilon)}{n^{1+\lambda_1^* \theta}} \right| > C_0 \sqrt{\epsilon} \right) \\ + \mathbb{P} \left(\left| \frac{D_n(k, \theta \log n + a_0 - \epsilon)}{n^{1+\lambda_1^* \theta}} - p_k^1 \right| > C_0 \sqrt{\epsilon} \right).$$

By Lemma 9.12,

$$\frac{D_n(k, \theta \log n + a_0 - \epsilon)}{n^{1+\lambda_1^* \theta}} = \frac{D_n(k, \theta \log n + a_0 - \epsilon)}{Z_n(\theta \log n + a_0 - \epsilon)} \frac{Z_n(\theta \log n + a_0 - \epsilon)}{n^{1+\lambda_1^* \theta}} \xrightarrow{P} p_k^1 e^{-\lambda_1^* \epsilon},$$

and therefore, there is an $\epsilon_1 \leq \epsilon_0$ such that for all $\epsilon \in (0, \epsilon_1)$,

$$(9.32) \quad \left| D_n(k, \theta \log n + a_0 - \epsilon) n^{-(1+\lambda_1^* \theta)} - p_k^1 \right| \xrightarrow{P} p_k^1 (1 - e^{-\lambda_1^* \epsilon}) \leq p_k^1 \lambda_1^* \epsilon < C_0 \sqrt{\epsilon}.$$

For $\epsilon \in (0, \epsilon_1)$, using (9.30) and (9.32) in (9.31), we conclude $\mathbb{P} \left(\left| \frac{D_n(k, T_n^\theta)}{n^{1+\lambda_1^* \theta}} - p_k^1 \right| > 2C_0 \sqrt{\epsilon} \right) \rightarrow 0$ as $n \rightarrow \infty$ proving the theorem. $\blacksquare$

9.2. Proof of Theorem 3.20. We prove (a) of the theorem; (b) and (c) follow via straightforward modifications of these arguments. For (a), construct the continuous time branching process $\text{BP}_\theta(\cdot)$ with change point as in Section 2.3 with $\tau = n^\gamma$. To ease notation later in the section, write $\text{BP}_n(\cdot) := \text{BP}_\theta(\cdot)$. Thus $\text{BP}_n(T_{n^\gamma})$ is a random tree obtained by running a continuous time branching process with attachment function $f_0 \equiv 1$ till it reaches size n^γ after which all vertices switch to reproducing using attachment function f_1 as in (a) of the Theorem. We are interested in the random tree $\mathcal{T}_n = \text{BP}_n(T_n)$, where as before for any m , $T_m := \inf\{t \geq 0 : |\text{BP}_n(t)| = m\}$.

PROPOSITION 9.15. *For the process $\text{BP}_n(\cdot)$ as constructed above:*

- (a) *The stopping time T_{n^γ} satisfies, $T_{n^\gamma} - \gamma \log n \xrightarrow{\text{a.s.}} \tilde{W}$, where $\tilde{W} = -\log W$ and $W \sim \exp(1)$.*
- (b) *Let $\omega_n \rightarrow \infty$ arbitrarily slowly. Then there exists a constant $C > 0$ independent of ω_n such that*

$$\mathbb{P} \left(\sup_{t \geq 0} |n^{-\gamma} e^{-(2+\alpha)t} |\text{BP}_n(t + T_{n^\gamma})| - 1| > \omega_n n^{-\gamma/2} \right) \leq C/\omega_n^2.$$

In particular whp as $n \rightarrow \infty$, $|(T_n - T_{n^\gamma}) - (1 - \gamma) \log n / (2 + \alpha)| \leq \omega_n n^{-\gamma/2}$.

PROOF. Part (a) follows from Lemma 6.3 upon noting that T_{n^γ} has the same distribution as the hitting time of n^γ by a Yule process with rate 1. To prove (b), recall that for $t > T_{n^\gamma}$, all individuals switch to offspring dynamics modulated by f_1 . For the rest of the proof, we proceed conditional on $\text{BP}_n(T_{n^\gamma})$. Using Proposition 6.5, the following two processes are martingales

$$M_1(t) := (e^{-(2+\alpha)t} |\text{BP}_n(t + T_{n^\gamma})| - n^\gamma) + (1 - e^{-(2+\alpha)t}) / (2 + \alpha), \quad t \geq 0,$$

$$M_2(t) := e^{-2(2+\alpha)t} |\text{BP}_n(t + T_{n^\gamma})|^2 - \int_0^t \alpha e^{-2(2+\alpha)s} |\text{BP}_n(s + T_{n^\gamma})| ds - e^{-2(2+\alpha)t} / 2(2+\alpha), \quad t \geq 0,$$

Using these expressions, it can be deduced that $\sup_{t \geq 0} \mathbb{E}(M_1^2(t)) \leq Cn^\gamma$ for some constant $C > 0$. Doob's $\mathbb{L}^2$ -maximal inequality then proves the first assertion of Proposition 9.15 (b) which then results in the second assertion in (b). $\blacksquare$

We now construct two approximating processes BP_n^+ and BP_n^- for $\mathcal{T}_n^\theta$ (grown completely in continuous time). Fix constant $B > 0$ and sequence $\omega_n \rightarrow \infty$ such that $\omega_n = o(\log n) \uparrow \infty$. For the rest of this Section let $t_n^\pm := \frac{(1-\gamma)}{(2+\alpha)} \log n \pm \frac{\omega_n}{n^{\gamma/2}}$. Define the process $\{\text{BP}_n^+(t) : 0 \leq t \leq \gamma \log n + B + t_n^+\}$ as follows: **(a)** Run a continuous time branching process driven by $f_0(\cdot) \equiv 1$ for time $\gamma \log n + B$; **(b)** After this time, every vertex switches dynamics so that it reproduces at rate equal to the number of children $+1 + \alpha$. Run this process for **an additional** time t_n^+ . Write $\tilde{\mathcal{T}}_n^+(B, \omega_n) = \text{BP}_n^+(\gamma \log n + B + t_n^+)$ for the random rooted tree at the end of this process. Analogously define $\{\text{BP}_n^-(t) : 0 \leq t \leq \gamma \log n - B + t_n^-\}$ and $\tilde{\mathcal{T}}_n^-(B, \omega_n) := \text{BP}_n^-(\gamma \log n - B + t_n^-)$ where in the above construction we wait till time $\gamma \log n - B$ before switching dynamics and run the new dynamics for additional time t_n^- .

By Proposition 9.15, given any $\varepsilon > 0$ we can choose a constant $B = B(\varepsilon)$ for which we can produce a coupling between $\mathcal{T}_n$ and $\tilde{\mathcal{T}}_n^+(B, \omega_n)$ such that for all large n , with probability at least $1 - \varepsilon$, $\mathcal{T}_n \subseteq \tilde{\mathcal{T}}_n^+(B, \omega_n)$ where we see the object on the left as a subtree of the object on the right with the same root. A similar assertion holds with $\tilde{\mathcal{T}}_n^-(B, \omega_n) \subseteq \mathcal{T}_n$. Using these couplings, the following proposition completes the proof of part (a) of Theorem 3.20 with part (a) of the proposition proving the lower bound while part (b) proving the upper bound. In the following, we will denote the root of the respective trees by ρ^* .

PROPOSITION 9.16. *Fix $B > 0$ and $\omega_n = o(\log n) \uparrow \infty$.*

- (a) *Consider the degree of the root $D_n^-(\rho^*)$ in $\tilde{\mathcal{T}}_n^-(B, \omega_n)$. Then $D_n^-(\rho^*) \geq \frac{\gamma}{4} n^{(1-\gamma)/(2+\alpha)} \log n$ whp as $n \rightarrow \infty$.*
- (b) *Consider the maximal degree $M_n^+(1)$ in $\tilde{\mathcal{T}}_n^+(B, \omega_n)$. Then $\exists$ constant $C > 0$ such that whp as $n \rightarrow \infty$, $M_n^+(1) \ll C n^{(1-\gamma)/(2+\alpha)} (\log n)^2$.*

Proof: We start with (a). Each individual in the original branching process driven by $f_0(\cdot) \equiv 1$ before time $\gamma \log n - B$ reproduces according to a rate one Poisson process. In particular standard bounds for a Poisson random variable imply that the degree of the root in the branching process at time $\gamma \log n - B$, denoted by $\deg_n(\rho^*, \gamma \log n - B)$, satisfies

$$(9.33) \quad \deg_n(\rho^*, \gamma \log n - B) \geq \frac{3}{4} \gamma \log n \text{ whp as } n \rightarrow \infty.$$

Now let $\{Y_{(i)}(\cdot) : i \geq 1\}$ be a collection of independent rate one Yule processes. Comparing rates for the evolution of the degree of the root after $\gamma \log n - B$ we get that

$$(9.34) \quad D_n^-(\rho^*) \geq_{\text{st}} \sum_{i=1}^{\deg_n(\rho^*, \gamma \log n - B)} Y_{(i)}(t_n^-).$$

Using (9.33), (9.34), Lemma 6.3 and standard lower tail bounds for the Geometric distribution [33, Theorem 3.1] finishes the proof.

Let us now prove (b). Recall that after the change point, dynamics are modulated by $f_1(\cdot) := \cdot + 1 + \alpha$. Let A denote the smallest integer $\geq \alpha + 1$. Let ξ_{f_1} be point process associated with f_1 as in (2.1). Comparing rates we see that $\xi_{f_1}(\cdot) \leq_{\text{st}} \sum_{i=1}^{A+2} Y_{(i)}(\cdot)$, where as before $\{Y_{(i)}(\cdot) : i \geq 1\}$ is a collection of independent rate one Yule processes. For every vertex $v \in \tilde{\mathcal{T}}_n^+(B, \omega_n)$ write $\deg_n(v)$ for the final degree of the vertex at time $\gamma \log n + B + t_n^+$ when we have finished constructing the process $\text{BP}_n^+(\cdot)$. As below Theorem 3.4, for any $v \in \text{BP}_n^+$, let σ_v denote the time of birth of vertex v into the system. We split the proof of (b) into two cases (loosely corresponding to the maximal degree of vertices after and before change point respectively):

(b1) Maximal degree for vertices born after $\gamma \log n + B$: Define the following collection of vertices

$$\mathbb{L}_n = \left\{ v \in \tilde{\mathcal{T}}_n^+(B, \omega_n) : \sigma_v \in [\gamma \log n + B, \gamma \log n + B + t_n^+], \deg_n(v) > C(A+2)n^{\frac{1-\gamma}{2+\alpha}} (\log n)^2 \right\},$$

where A is above is the smallest integer $\geq \alpha + 1$ and C is an appropriate constant chosen later in the proof. We now show that we can choose C such that $\mathbb{E}(|\mathbb{A}_n|) \rightarrow 0$, as $n \rightarrow \infty$. This would then imply

$$(9.35) \quad \mathbb{P}(\exists v \in \tilde{\mathcal{T}}_n^+(B, \omega_n), \sigma_v \geq \gamma \log n + B, \deg_n(v) > C(A+2)n^{\frac{1-\gamma}{2+\alpha}}(\log n)^2) \rightarrow 0.$$

For the rest of the proof let $k'_n := C(A+2)n^{\frac{1-\gamma}{2+\alpha}}(\log n)^2$ and $k_n = k'_n/(A+2) = Cn^{\frac{1-\gamma}{2+\alpha}}(\log n)^2$. Fix $s \geq 0$ and consider a vertex born at some time $s + \gamma \log n + B \in [\gamma \log n + B, \gamma \log n + B + t_n^+]$. Thus this vertex has time $t_n^+ - s$ to evolve its degree. Using the bound on ξ_{f_1} namely the offspring process of each new vertex born at $t > \gamma \log n + B$ by a sum of Yule process above, by Lemma 6.3 the probability that such a vertex has degree greater than $(A+2)k_n$ by time t_n^+ is bounded by $\mathbb{P}(\text{geom}(e^{-(t_n^+ - s)}) \geq (A+2)k_n) \leq e^{-k_n e^{t_n^+ - s}}$. Next note that for any $t \geq \gamma \log n + B$, new vertices are produced at rate $(2+\alpha)|\text{BP}_n^+(t)| - 1$. As in the proof of Proposition 9.15, the process $M(s) := e^{-(2+\alpha)s}|\text{BP}_n^+(s + \gamma \log n + B)| + (2+\alpha)^{-1}e^{-(2+\alpha)s}$, $s \geq 0$ is a martingale. Noting $\mathbb{E}|\text{BP}_n^+(\gamma \log n + B)| = e^B n^\gamma$ we get that $\mathbb{E}|\text{BP}_n^+(s + \gamma \log n + B)| \leq C' n^\gamma e^{(2+\alpha)s}$ for $0 \leq s \leq t_n^+$ where C' is a constant depending only on B, α . Thus,

$$\mathbb{E}(|\mathbb{L}_n|) \leq C'' n^\gamma \int_0^{t_n^+} e^{-k_n e^{-(t_n^+ - s)}} e^{(2+\alpha)s} ds,$$

where C'' depends only on B, α . The following completes the proof of (9.35).

LEMMA 9.17. $I_n := n^\gamma \int_0^{t_n^+} e^{-C(\log n)^2 n^{\frac{1-\gamma}{2+\alpha}} e^{-(t_n^+ - s)}} e^{(2+\alpha)s} ds \rightarrow 0$ for sufficiently large C as $n \rightarrow \infty$.

PROOF. Writing $a := \frac{1-\gamma}{2+\alpha}$ and $b := 2 + \alpha$, algebraic manipulations result in:

$$I_n \leq n^\gamma (\log n)^{-2b} e^{\frac{b w_n}{n^{\gamma/2}}} \Gamma\left(b, C(\log n)^2 e^{-\frac{w_n}{n^{\gamma/2}}}\right) := \mathcal{E}_n.$$

where $\Gamma(b, z) = \int_z^\infty e^{-t} t^{b-1} dt$ is the upper incomplete Gamma function. It is known that $\Gamma(b, z) = \Omega(z^{b-1} e^{-z})$ as $z \rightarrow \infty$. Thus $\mathcal{E}_n \sim n^\gamma C \log n e^{-\frac{w_n}{n^{\gamma/2}}} (\log n)^{-2} e^{-\frac{w_n}{n^{\gamma/2}}} \rightarrow 0$. ■

(b2) Maximal degree for vertices born before $\log n + B$:

To simplify notation let $\Delta_n := \gamma \log n + B$, $Y_n := \gamma \log n + B + t_n^+$. For fixed vertex v born into $\text{BP}_n^+(\cdot)$ and for time $t \leq Y_n$, let $\deg(v, t)$ denote the degree of this vertex v in $\text{BP}_n^+(t)$ with the convention that $\deg(v, t) := 0$ for $t < \sigma_v$. Write $\deg_n(v) := \deg(v, Y_n)$ for the final degree of v in $\tilde{\mathcal{T}}_n^+(B, \omega_n)$. Fix $C > 0$ and let $\mathbb{B}_n$ be the set of vertices born before $\gamma \log n + B$ whose final degree is too large i.e. $\mathbb{B}_n := \{v \in \tilde{\mathcal{T}}_n^+(B, \omega_n) : \sigma_v \leq \gamma \log n + B, \deg_n(v) > Cn^{\frac{1-\gamma}{2+\alpha}}(\log n)^2\}$, where as before, $\deg_n(v) := \deg(v, Y_n)$ is the degree of vertex v in the final tree $\tilde{\mathcal{T}}_n^+(B, \omega_n)$.

PROPOSITION 9.18. We can choose $C < \infty$ such that $\mathbb{P}(|\mathbb{B}_n| \geq 1) \rightarrow 0$ as $n \rightarrow \infty$.

PROOF. Consider the tree $\text{BP}_n^+(\Delta_n)$. Let $M_n(\Delta_n) := \max_{v \in \text{BP}_n^+(\Delta_n)} \deg(v, \Delta_n)$ be the maximal degree of vertices in $\text{BP}_n^+(\Delta_n)$ at time Δ_n . Let $\ell_n := 10e \log n$ and fix a sequence $\omega_n \uparrow \infty$. By the union bound,

$$\begin{aligned} \mathbb{P}(|\mathbb{B}_n| \geq 1) &\leq \mathbb{P}(|\mathbb{B}_n| \geq 1, |\text{BP}_n^+(\Delta_n)| < \omega_n n^\gamma, M_n(\Delta_n) \leq \ell_n) \\ &\quad + \mathbb{P}(|\text{BP}_n^+(\Delta_n)| \geq \omega_n n^\gamma) + \mathbb{P}(M_n(\Delta_n) > \ell_n). \end{aligned}$$

Lemmas 9.19 and 9.20 bound the three terms on the right and complete the proof of the Proposition. ■

LEMMA 9.19. *Let $\omega_n = \log n$. We can choose constant $C < \infty$ such that as $n \rightarrow \infty$,*

$$\mathbb{P}(|\mathbb{B}_n| \geq 1, |\text{BP}_n^+(\Delta_n)| < \omega_n n^\gamma, M_n(\Delta_n) \leq \ell_n) \rightarrow 0.$$

PROOF. Let $\mathbb{G}_n = \{|\text{BP}_n^+(\Delta_n)| < \omega_n n^\gamma, M_n(\Delta_n) \leq \ell_n\}$. It is sufficient to show we can choose constant C such that $\mathbb{P}(|\mathbb{B}_n| \geq 1 | \mathbb{G}_n) \rightarrow 0$. Conditional on $\mathbb{G}_n$, we will construct a stochastic process that bounds the growth of the maximal degree of the vertices in $\text{BP}_n^+(\Delta_n)$ for times $t \geq \Delta_n$. Let $\{X_i(\cdot) : 1 \leq i \leq n^\gamma \omega_n\}$ be a collection of i.i.d. stochastic processes with distribution $X(\cdot) = \sum_{j=1}^{\ell_n + A + 2} Y_j(\cdot)$, where $\{Y_j(\cdot) : j \geq 1\}$ is collected of i.i.d. rate one Yule processes. Recall that $t_n^+ = \frac{1-\gamma}{2+\alpha} \log n + \frac{\omega_n}{n^{\gamma/2}}$. Let $\mathcal{M}_n := \max_{1 \leq i \leq \omega_n n^\gamma} X_i(t_n^+)$.

On the event $\mathbb{G}_n$, the number of vertices $|\text{BP}_n^+(\Delta_n)| \leq \omega_n n^\gamma$ and further the maximal degree of any vertex at time Δ_n is $\leq \ell_n$. Thus on $\mathbb{G}_n$, for any $v \in \text{BP}_n^+(\Delta_n)$, comparing rates for the point process representing the evolution of degrees for $t > \Delta_n$, we see that $\deg(v, \cdot) \leq_{\text{st}} X(\cdot)$ with X as above. The time translation makes the precise formulation clunky but in brief, on the set $\mathbb{G}_n$, for any $v \in \text{BP}_n^+(\Delta_n)$, we can construct $\{(\deg(v, \Delta_n + s), X(s)) : 0 \leq s \leq t_n^+\}$ on a common probability space so that for all $0 \leq s \leq t_n^+$, $\deg(v, \Delta_n + s) \leq X(s)$. Thus on the event $\mathbb{G}_n$, the maximal degree at time Υ_n of vertices born before time Δ_n satisfies $\max_{v \in \text{BP}_n^+(\Delta_n)} \deg(v, \Upsilon_n) \leq_{\text{st}} \mathcal{M}_n$. The rest of the proof analyzes $\mathcal{M}_n$. The union bound gives,

$$(9.36) \quad \mathbb{P}(|\mathbb{B}_n| \geq 1 | \mathbb{G}_n) \leq \mathbb{P}(\mathcal{M}_n \geq C n^{\frac{1-\gamma}{2+\alpha}} (\log n)^2) \leq \omega_n n^\gamma \mathbb{P}(X(t_n^+) \geq C n^{\frac{1-\gamma}{2+\alpha}} (\log n)^2).$$

By Lemma 6.3 for any $t \geq 0$ and $\lambda > 0$, with $m = \ell_n + A + 2$,

$$\mathbb{P}(X(t) > \lambda) \leq m \mathbb{P}(\text{geom}(e^{-t}) > (\lambda/m)) \leq m \exp[-(\lambda/m)e^{-t}].$$

Plugging in $t = t_n^+$, $\lambda = C n^{\frac{1-\gamma}{2+\alpha}} (\log n)^2$ we get that the last term in (9.36) can be bounded by $K \omega_n n^\gamma n^{-C} \log n$ which goes to zero for sufficiently large C . ■

LEMMA 9.20. *For C large enough as $n \rightarrow \infty$, $\mathbb{P}(|\text{BP}_n^+(\Delta_n)| \geq \omega_n n^\gamma) \rightarrow 0$, and $\mathbb{P}(M_n(\Delta_n) > \ell_n) \rightarrow 0$.*

PROOF. The second assertion follows from standard bounds for the maximal degree of the random recursive tree [23]. We omit the proof. We prove the first assertion. The size of the tree grows according to a rate one Yule process. Thus by Lemma 6.3, $|\text{BP}_n(\Delta_n)| \sim \text{geom}(e^{-(\gamma \log n + B)})$. Thus

$$\mathbb{P}(|\text{BP}_n^+(\Delta_n)| \geq \omega_n n^\gamma) \leq \exp[-\omega_n n^\gamma e^{-\gamma \log n - B}] \rightarrow 0, \quad \text{as } n \rightarrow \infty. \quad \blacksquare$$

10. Proofs: Convergence rates for model without change point. This section is dedicated to proving Theorem 3.4 and Theorem 3.5. We need the following lemma which quantifies the rate of convergence of solutions of renewal equations to their limit as time goes to infinity.

LEMMA 10.1. *Consider a continuous time branching process with attachment function f that satisfies Assumption 2.4. Fix $\beta \in (0, \lambda^*)$. There exist positive constants C_1, C_2 such that the following holds: if h solves the renewal equation*

$$h(t) = e^{-\lambda^* t} \phi(t) + \int_0^t h(t-s) e^{-\lambda^* s} \mu_f(ds)$$

with any ϕ satisfying $|\phi(s)| \leq C_\phi e^{\beta s}$ for all $s \geq 0$, for some $C_\phi > 0$, then $h(\infty) := \lim_{t \rightarrow \infty} h(t)$ exists and we have, for all $t \geq 0$, $|h(\infty) - h(t)| \leq C_1 C_\phi e^{-C_2 t}$.

PROOF. In the proof, C, C' will denote generic positive constants, not depending on C_ϕ or the choice of ϕ , whose values might change from line to line. We will use estimates about quantitative rates of convergence for renewal measures derived in [11] in the setting of the point process with i.i.d. inter-arrival times having distribution $e^{-\lambda^* s} \mu_f(ds)$. By Assumption 2.4 (ii), it is clear that the measure $e^{-\lambda^* s} \mu_f(ds)$ satisfies $\int_0^\infty e^{\beta' s} e^{-\lambda^* s} \mu_f(ds) < \infty$ for some $\beta' > 0$ and thus, Assumption 1 of [11] is satisfied. Moreover, for any Borel set A in $[0, 1]$, denoting by E the first time the root reproduces (which has an exponential distribution with rate $f(0)$), note that

$$\mu_f(A) \geq \mathbb{E}(\mathbb{1}\{E \in A\}) = \int_A f(0) e^{-f(0)x} dx \geq f(0) e^{-f(0)} \int_A dx$$

and consequently, the distribution of the inter-arrival time is *spread out* in the sense of Assumption 2 of [11] taking $c = 1/2, L = 1/2$ and $\tilde{\eta} = f(0) e^{-(\lambda^* + f(0))}$. Thus, Corollary 1 of [11] holds for the point process under consideration. For any $x \geq 0$, denote by U^x the renewal measure corresponding to the associated point process with time started at x . The stationary version of this point process corresponds to a random starting time whose law is $\mu^*(ds) = m^{*-1} s e^{-\lambda^* s} \mu_f(ds)$ (called the *stationary delay distribution*), where $m^* = \int_0^\infty u e^{-\lambda^* u} \mu_f(du)$. From translation invariance, it follows that the renewal measure associated to this stationary version is given by $U^*(ds) = m^{*-1} ds$. By Corollary 1 of [11], there exist constants $C, C' > 0$ and $\beta'' < \beta'$ such that for any Borel set $D \subset (0, \infty)$ and any $x, t \geq 0$,

$$|U^x(D+t) - U^0(D+t)| \leq C e^{\beta'' x} e^{-C' t} (U^0((0, \sup D)) + 1).$$

Integration both sides of the above relation over x with respect to the stationary delay distribution $\mu^*(dx)$ and using Fubini's theorem and the fact that $\int_0^\infty e^{\beta' s} e^{-\lambda^* s} \mu_f(ds) < \infty$, we obtain

$$|U^*(D+t) - U^0(D+t)| \leq C e^{-C' t} (U^0((0, \sup D)) + 1).$$

This, in turn, implies that for any $t \geq 0$, if $U_{M,t}^*$ and $U_{M,t}^0$ denote the measures defined by $U_{M,t}^*(D) = U^*(D+t)$ and $U_{M,t}^0(D) = U^0(D+t)$ for any Borel set $D \subset [0, M]$, then using the fact that $\lim_{t \rightarrow \infty} t^{-1} U^0([0, t]) = \frac{1}{m^*}$ (which follows from the elementary renewal theorem),

$$(10.1) \quad \|U_{M,t}^* - U_{M,t}^0\|_{TV} \leq C M e^{-C' t}.$$

From standard results in renewal theory, $h(t) = \int_0^t e^{-\lambda^* (t-s)} \phi(t-s) U^0(ds), t \geq 0$, and $h(\infty) := \lim_{t \rightarrow \infty} h(t)$ exists with $h(\infty) = \int_0^\infty e^{-\lambda^* s} \phi(s) U^*(ds)$. Thus, for $t \geq 0$,

$$(10.2) \quad |h(\infty) - h(t)| = \left| \int_0^\infty e^{-\lambda^* s} \phi(s) U^*(ds) - \int_0^t e^{-\lambda^* (t-s)} \phi(t-s) U^0(ds) \right| \\ \leq \left| \int_0^t e^{-\lambda^* s} \phi(s) U^*(ds) - \int_0^t e^{-\lambda^* (t-s)} \phi(t-s) U^0(ds) \right| + \int_t^\infty e^{-\lambda^* s} \phi(s) U^*(ds).$$

As $|\phi(s)| \leq C_\phi e^{\beta s}$ for all s ,

$$(10.3) \quad \int_t^\infty e^{-\lambda^* s} \phi(s) U^*(ds) \leq C_\phi m^{*-1} \int_t^\infty e^{-(\lambda^* - \beta)s} ds = \frac{C_\phi}{m^* (\lambda^* - \beta)} e^{-(\lambda^* - \beta)t}.$$

To estimate the first term in the bound (10.2), note that for $t \geq 0$,

$$(10.4) \quad \left| \int_0^t e^{-\lambda^* s} \phi(s) U^*(ds) - \int_0^t e^{-\lambda^* (t-s)} \phi(t-s) U^0(ds) \right|$$

$$\begin{aligned}
&\leq \int_0^{t/2} e^{-\lambda^*(t-s)} \phi(t-s) U^*(ds) + \int_0^{t/2} e^{-\lambda^*(t-s)} \phi(t-s) U^0(ds) \\
&\quad + \left| \int_{t/2}^t e^{-\lambda^*(t-s)} \phi(t-s) U^*(ds) - \int_{t/2}^t e^{-\lambda^*(t-s)} \phi(t-s) U^0(ds) \right| \\
&\leq C_\phi e^{-(\lambda^*-\beta)t/2} U^*([0, t/2]) + C_\phi e^{-(\lambda^*-\beta)t/2} U^0([0, t/2]) + C_\phi \|U_{t/2, t/2}^* - U_{t/2, t/2}^0\|_{TV} \leq C_1' C_\phi e^{-C_2' t}
\end{aligned}$$

for constants $C_1', C_2' > 0$ not depending on ϕ , where we used (10.1) and the observations that $U^*([0, t/2]) = t/2 m^*$ and $\lim_{t \rightarrow \infty} t^{-1} U^0([0, t/2]) = 1/(2m^*)$. The lemma follows using (10.3) and (10.4) in (10.2). $\blacksquare$

PROOF OF THEOREM 3.5. We bound $\left| e^{-\lambda^* t} Z_f^\phi(t) - W_\infty M_f^\phi(\infty) \right|$ using the same techniques as in the proof of Theorem 3.1 of [37]. For each term appearing in the bound, we show that they are small in a suitable sense using renewal theoretic methods and variance computations.

In the proof, $C, C', C'', C_1, C_2, \beta', \beta$ denote generic positive constants depending neither on b_ϕ nor the choice of ϕ . Following [37], write $x = (x', i)$ when x is the i -th child of x' and define for any $t, c \geq 0$,

$$\mathcal{J}(t) = \{x = (x', i) : \sigma_{x'} \leq t \text{ and } t < \sigma_x < \infty\}, \quad \mathcal{J}(t, c) = \{x = (x', i) : \sigma_{x'} \leq t \text{ and } t + c < \sigma_x < \infty\}.$$

Let $\bar{T}_t$ denote the number of vertices born by time t and let $\mathcal{A}_n$ be the filtration generated by the entire biographies of the first n vertices (see [37] for detailed definitions). Define $\mathcal{F}_t = \mathcal{A}_{\bar{T}_t}$. For any $s > 0$, write $\phi = \phi_s + \phi'_s$ where $\phi_s(u) = \phi(u) \mathbb{1}\{u < s\}$ and $\phi'_s(u) = \phi(u) \mathbb{1}\{u \geq s\}$. Note that

$$\begin{aligned}
(10.5) \quad \mathbb{E} \left| e^{-\lambda^* t} Z_f^\phi(t) - W_\infty M_f^\phi(\infty) \right| &\leq \mathbb{E} \left| e^{-\lambda^* t} (Z_f^\phi(t) - Z_f^{\phi_s}(t)) \right| + \mathbb{E} \left| e^{-\lambda^* t} Z_f^{\phi_s}(t) - W_\infty M_f^{\phi_s}(\infty) \right| \\
&\quad + \mathbb{E} \left(\left| M_f^{\phi_s}(\infty) - M_f^\phi(\infty) \right| W_\infty \right).
\end{aligned}$$

Recall that, by (2.6) appearing in Assumption 2.4 (ii), $\underline{\lambda} < \lambda^*$ and hence, there exists $\beta' \in (\underline{\lambda}, \lambda^*)$ such that

$$(10.6) \quad e^{-\beta' t} \mathbb{E}(\xi_f(t)) = \mathbb{E}(\xi_f(t)) \int_t^\infty \beta' e^{-\beta' u} du \leq \int_0^\infty \beta' e^{-\beta' u} \mathbb{E}(\xi_f(u)) du = \hat{\rho}(\beta') < \infty.$$

Using this, the third term in the bound (10.5) can be bounded as

$$\begin{aligned}
(10.7) \quad \mathbb{E} \left(\left| M_f^{\phi_s}(\infty) - M_f^\phi(\infty) \right| W_\infty \right) &= M_f^{\phi'_s}(\infty) = \frac{1}{m^*} \int_s^\infty e^{-\lambda^* u} \mathbb{E}(\phi(u)) du \\
&\leq \frac{b_\phi}{m^*} \int_s^\infty e^{-\lambda^* u} \mathbb{E}(\xi_f(u) + 1) du \leq C b_\phi e^{-(\lambda^* - \beta')s}.
\end{aligned}$$

The first term in the bound (10.5) can be bounded as

$$(10.8) \quad \mathbb{E} \left| e^{-\lambda^* t} (Z_f^\phi(t) - Z_f^{\phi_s}(t)) \right| = \mathbb{E} \left(e^{-\lambda^* t} Z_f^{\phi'_s}(t) \right) \leq \left| M_f^{\phi'_s}(t) - M_f^{\phi'_s}(\infty) \right| + M_f^{\phi'_s}(\infty).$$

By the fact that $M_f^{\phi'_s}(t)$ satisfies the renewal equation (3.2) (with ϕ'_s in place of ϕ) and Lemma 10.1, for $t \geq 0$, $\left| M_f^{\phi'_s}(t) - M_f^{\phi'_s}(\infty) \right| \leq C_1 b_\phi e^{-C_2 t}$. Using this estimate and (10.7) in (10.8), we obtain

$$(10.9) \quad \mathbb{E} \left| e^{-\lambda^* t} (Z_f^\phi(t) - Z_f^{\phi_s}(t)) \right| \leq C_1 b_\phi e^{-C_2 t} + C b_\phi e^{-(\lambda^* - \beta')s}.$$

Using (10.7) and (10.9) in (10.5), for any $t, s \geq 0$,

$$(10.10) \quad \mathbb{E} \left| e^{-\lambda^* t} Z_f^\phi(t) - W_\infty M_f^\phi(\infty) \right| \leq \mathbb{E} \left| e^{-\lambda^* t} Z_f^{\phi_s}(t) - W_\infty M_f^{\phi_s}(\infty) \right| + C_1 b_\phi e^{-C_2 t} + 2C b_\phi e^{-(\lambda^* - \beta') s}.$$

Now, we estimate the first term in the above bound. Observe that as $\phi_s(u) = 0$ for all $u \geq s$, every individual that contributes to $Z_f^{\phi_s}(t+s)$ must be born after time t . Therefore, $Z_f^{\phi_s}(t+s) = \sum_{x \in \mathcal{J}(t)} Z_{f,x}^{\phi_s}(t+s-\sigma_x)$ where for any vertex x and any $u \geq 0$, $Z_{f,x}^{\phi_s}(u)$ denotes the aggregate ϕ -score at time $\sigma_x + u$ treating the vertex x as the root. For $t, c \geq 0$ such that $s \geq c$, write

$$X(t, s, c) = \sum_{x \in \mathcal{J}(t) \setminus \mathcal{J}(t, c)} e^{-\lambda^* \sigma_x} \left(e^{-\lambda^* (t+s-\sigma_x)} Z_{f,x}^{\phi_s}(t+s-\sigma_x) - M_f^{\phi_s}(t+s-\sigma_x) \right).$$

and write $W_t = \sum_{x \in \mathcal{J}(t)} e^{-\lambda^* \sigma_x}$, $W_{t,c} = \sum_{x \in \mathcal{J}(t, c)} e^{-\lambda^* \sigma_x}$. Following equation (3.36) in [37], we obtain

$$\begin{aligned} & \left| e^{-\lambda^* (t+s)} Z_f^{\phi_s}(t+s) - W_\infty M_f^{\phi_s}(\infty) \right| \\ & \leq |X(t, s, c)| + \sum_{x \in \mathcal{J}(t) \setminus \mathcal{J}(t, c)} e^{-\lambda^* \sigma_x} \left| M_f^{\phi_s}(t+s-\sigma_x) - M_f^{\phi_s}(\infty) \right| \\ & \quad + \left| \sum_{x \in \mathcal{J}(t, c)} e^{-\lambda^* \sigma_x} \left(e^{-\lambda^* (t+s-\sigma_x)} Z_{f,x}^{\phi_s}(t+s-\sigma_x) - M_f^{\phi_s}(\infty) \right) \right| + M_f^{\phi_s}(\infty) |W_t - W_\infty|. \end{aligned}$$

(10.11)

Note that

$$(10.12) \quad \text{Var}(X(t, s, c) | \mathcal{F}_t) = \sum_{x \in \mathcal{J}(t) \setminus \mathcal{J}(t, c)} e^{-2\lambda^* \sigma_x} V_f^{\phi_s}(t+s-\sigma_x)$$

where $V_f^{\phi_s}(t) = \text{Var} \left(e^{-\lambda^* t} Z_f^{\phi_s}(t) \right)$. Recall $m_f^{\phi_s}(t) = \mathbb{E} \left(Z_f^{\phi_s}(t) \right)$ and $v_f^{\phi_s}(t) = \text{Var} \left(Z_f^{\phi_s}(t) \right)$. From Theorem 3.2 of [30], $v_f^{\phi_s}(t) = h \star U(t)$, where $h(t) = \text{Var} \left(\phi_s(t) + \int_0^t m_f^{\phi_s}(t-u) \xi_f(du) \right)$ and $U(\cdot) = \sum_{\ell=0}^\infty \mu_f^{\star \ell}(\cdot)$ denotes the renewal measure. As $\phi_s(t) \leq b_\phi(\xi_f(t) + 1)$ for all t , using Assumption 3.2,

$$\begin{aligned} & e^{-2\lambda^* t} \mathbb{E}(\phi_s(t))^2 \leq (b_\phi)^2 \mathbb{E} \left(e^{-\lambda^* t} (1 + \xi_f(t)) \right)^2 \\ (10.13) \quad & \leq 2(b_\phi)^2 \mathbb{E} \left(e^{-2\lambda^* t} + \lambda^{*2} \left(\int_t^\infty e^{-\lambda^* u} \xi_f(u) du \right)^2 \right) \leq C(b_\phi)^2. \end{aligned}$$

As $\mathbb{E}(\xi_f(t) + 1) \leq C e^{\beta' t}$ by (10.6), therefore $\mathbb{E}(\phi_s(t)) \leq b_\phi \mathbb{E}(\xi_f(t) + 1) \leq b_\phi C e^{\beta' t}$. Hence, by the fact that $M_f^{\phi_s}(t)$ satisfies the renewal equation (3.2) and Lemma 10.1, for $t \geq 0$,

$$(10.14) \quad \left| M_f^{\phi_s}(t) - M_f^{\phi_s}(\infty) \right| \leq C_1 b_\phi e^{-C_2 t}.$$

Moreover,

$$(10.15) \quad M_f^{\phi_s}(\infty) = (m^*)^{-1} \int_0^\infty e^{-\lambda^* u} \mathbb{E}(\phi_s(u)) du \leq (m^*)^{-1} b_\phi \int_0^\infty \mathbb{E} \left(e^{-\lambda^* u} (1 + \xi_f(u)) \right) du \leq C b_\phi.$$

Using (10.14) and (10.15), we obtain for all $t \geq 0$,

$$(10.16) \quad M_f^{\phi_s}(t) \leq C' b_\phi.$$

From (10.13) and (10.16), we conclude for all $t \geq 0$,

$$\begin{aligned} e^{-2\lambda^* t} h(t) &= \text{Var} \left(e^{-\lambda^* t} \phi_s(t) + \int_0^t e^{-\lambda^* (t-u)} m_f^{\phi_s}(t-u) e^{-\lambda^* u} \xi_f(du) \right) \\ &\leq 2e^{-2\lambda^* t} \mathbb{E}(\phi_s(t))^2 + 2 \mathbb{E} \left(\int_0^t M_f^{\phi_s}(t-u) e^{-\lambda^* u} \xi_f(du) \right)^2 \\ &\leq 2C(b_\phi)^2 + 2(Cb_\phi)^2 \mathbb{E} \left(\int_0^\infty e^{-\lambda^* u} \xi_f(du) \right)^2 \leq C'(b_\phi)^2. \end{aligned}$$

Thus, for all $t \geq 0$,

$$\begin{aligned} (10.17) \quad V_f^{\phi_s}(t) &= \int_0^\infty e^{-2\lambda^* (t-u)} h(t-u) e^{-2\lambda^* u} U(du) \\ &\leq C'(b_\phi)^2 \int_0^\infty e^{-2\lambda^* u} U(du) = C'(b_\phi)^2 \sum_{\ell=0}^\infty \hat{\mu}_f(2\lambda^*)^\ell = \frac{C'(b_\phi)^2}{1 - \hat{\mu}_f(2\lambda^*)} = C''(b_\phi)^2. \end{aligned}$$

Using this bound in (10.12), we obtain

$$\mathbb{E}(\text{Var}(X(t, s, c) | \mathcal{F}_t)) \leq C''(b_\phi)^2 \mathbb{E} \left(\sum_{x \in \mathcal{J}(t) \setminus \mathcal{J}(t, c)} e^{-2\lambda^* \sigma_x} \right) \leq C''(b_\phi)^2 e^{-\lambda^* t} \mathbb{E}(W_t) = C''(b_\phi)^2 e^{-\lambda^* t}.$$

Moreover, $\mathbb{E}(X(t, s, c) | \mathcal{F}_t) = 0$. Thus, we obtain

$$(10.18) \quad \mathbb{E}|X(t, s, c)| \leq \sqrt{\mathbb{E}(X(t, s, c))^2} = \sqrt{\text{Var}(X(t, s, c))} \leq \sqrt{C''} b_\phi e^{-\lambda^* t/2}.$$

Using (10.14),

$$(10.19) \quad \mathbb{E} \left(\sum_{x \in \mathcal{J}(t) \setminus \mathcal{J}(t, c)} e^{-\lambda^* \sigma_x} \left| M_f^{\phi_s}(t+s-\sigma_x) - M_f^{\phi_s}(\infty) \right| \right) \leq C_1 b_\phi e^{-C_2(s-c)} \mathbb{E}(W_t) = C_1 b_\phi e^{-C_2(s-c)}.$$

To estimate the third term in the bound (10.11), observe that upon conditioning on $\mathcal{F}_t$ and noting that $\sup_{t < \infty} M_f^{\phi_s}(t) \leq C' b_\phi$,

$$\begin{aligned} (10.20) \quad &\mathbb{E} \left(\left| \sum_{x \in \mathcal{J}(t, c)} e^{-\lambda^* \sigma_x} \left(e^{-\lambda^* (t+s-\sigma_x)} Z_{f, x}^{\phi_s}(t+s-\sigma_x) - M_f^{\phi_s}(\infty) \right) \right| \right) \\ &\leq \mathbb{E} \left(\sum_{x \in \mathcal{J}(t, c)} e^{-\lambda^* \sigma_x} \left(M_f^{\phi_s}(t+s-\sigma_x) + M_f^{\phi_s}(\infty) \right) \right) \leq C' b_\phi \mathbb{E}(W_{t, c}). \end{aligned}$$

Consider the characteristic $\phi^c(v) = e^{\lambda^* v} \left(\int_{v+c}^\infty e^{-\lambda^* u} \xi_f(du) \right)$, $v \geq 0$. Then $W_{t, c} = e^{-\lambda^* t} Z_f^{\phi^c}(t)$. Note that

$$\begin{aligned} \mathbb{E}(\phi^c(t)) &= e^{\lambda^* t} \mathbb{E} \left(\int_{t+c}^\infty e^{-\lambda^* u} \xi_f(du) \right) = e^{\lambda^* t} \mathbb{E} \left(\int_{t+c}^\infty \lambda^* e^{-\lambda^* v} (\xi_f(v) - \xi_f(t+c)) dv \right) \\ &\leq e^{\lambda^* t} \mathbb{E} \left(\int_{t+c}^\infty \lambda^* e^{-\lambda^* v} \xi_f(v) dv \right) \leq C e^{\lambda^* t} \left(\int_{t+c}^\infty \lambda^* e^{-\lambda^* v} e^{\beta' v} dv \right) \leq \frac{C \lambda^* e^{\lambda^* t}}{\lambda^* - \beta'} e^{-(\lambda^* - \beta')t} = \frac{C \lambda^* e^{\beta' t}}{\lambda^* - \beta'}. \end{aligned}$$

Hence, by Lemma 10.1,

$$(10.21) \quad \left| M_f^{\phi^c}(t) - M_f^{\phi^c}(\infty) \right| \leq C_1 e^{-C_2 t}.$$

Moreover, by Lemma 3.5 of [37], $M_f^{\phi^c}(\infty) = \int_c^\infty (1 - \mu_{f,\lambda^*}(u)) du / \int_0^\infty (1 - \mu_{f,\lambda^*}(u)) du$ where $\mu_{f,\lambda^*}(u) = \int_0^u e^{-\lambda^* v} \mu_f(dv)$. Now, for any $u \geq 0$,

$$\begin{aligned} 1 - \mu_{f,\lambda^*}(u) &= \int_u^\infty e^{-\lambda^* v} \mu_f(dv) \leq \int_u^\infty \lambda^* e^{-\lambda^* v} \mu_f(v) dv \\ &\leq C \int_u^\infty \lambda^* e^{-\lambda^* v} e^{\beta' v} dv = \frac{C\lambda^*}{\lambda^* - \beta'} e^{-(\lambda^* - \beta')u} \end{aligned}$$

and hence,

$$\int_c^\infty (1 - \mu_{f,\lambda^*}(u)) du \leq \int_c^\infty \frac{C\lambda^*}{\lambda^* - \beta'} e^{-(\lambda^* - \beta')u} du = \frac{C\lambda^*}{(\lambda^* - \beta')^2} e^{-(\lambda^* - \beta')c}.$$

This bound implies that there exists $C > 0$ such that for all $c > 0$,

$$(10.22) \quad M_f^{\phi^c}(\infty) \leq C e^{-(\lambda^* - \beta')c}.$$

Combining (10.21) and (10.22), we have $\mathbb{E}(W_{t,c}) = M_f^{\phi^c}(t) \leq C_1 e^{-C_2 t} + C e^{-(\lambda^* - \beta')c}$. Using this in (10.20),

$$(10.23) \quad \mathbb{E} \left(\left| \sum_{x \in \mathcal{J}(t,c)} e^{-\lambda^* \sigma_x} \left(e^{-\lambda^* (t+s-\sigma_x)} Z_{f,x}^{\phi^s}(t+s-\sigma_x) - M_f^{\phi^s}(\infty) \right) \right| \right) \leq C' b_\phi \left(e^{-C_2 t} + e^{-(\lambda^* - \beta')c} \right).$$

To estimate the last term in the bound (10.11), observe that for any $t \geq 0$, $W_\infty = \sum_{x \in \mathcal{J}(t)} e^{-\lambda^* \sigma_x} W_\infty^x$, where W_∞^x corresponds to W_∞ treating vertex x as the root (and hence are i.i.d. and have the same distribution as W_∞). Moreover, by Theorem 4.1 of [30], $\text{Var}(W_\infty) < \infty$. Using these observations,

$$\begin{aligned} \mathbb{E}(W_t - W_\infty)^2 &= \mathbb{E} \left(\sum_{x \in \mathcal{J}(t)} e^{-\lambda^* \sigma_x} (1 - W_\infty^x) \right)^2 = \text{Var}(W_\infty) \mathbb{E} \left(\sum_{x \in \mathcal{J}(t)} e^{-2\lambda^* \sigma_x} \right) \\ &\leq \text{Var}(W_\infty) e^{-\lambda^* t} \mathbb{E}(W_t) = \text{Var}(W_\infty) e^{-\lambda^* t}. \end{aligned}$$

Together with the fact that $\sup_{t < \infty} M_f^{\phi^s}(t) \leq C' b_\phi$, this implies that for $t \geq 0$,

$$(10.24) \quad \mathbb{E} \left| M_f^{\phi^s}(\infty) | W_t - W_\infty \right| \leq \sqrt{\mathbb{E} \left(M_f^{\phi^s}(\infty) | W_t - W_\infty \right)^2} \leq C' b_\phi e^{-\lambda^* t/2}.$$

Using (10.18), (10.19), (10.23) and (10.24) and the bound (10.11), we obtain $D, D_1, D_2, D_3 > 0$ not depending on b_ϕ, t, s, c such that

$$(10.25) \quad \mathbb{E} \left(\left| e^{-\lambda^* (t+s)} Z_f^{\phi^s}(t+s) - W_\infty M_f^{\phi^s}(\infty) \right| \right) \leq D b_\phi (e^{-D_1 t} + e^{-D_2 c} + e^{-D_3 (s-c)}).$$

On taking $t - s$ in place of t in (10.25), we obtain for any $s, t, c \geq 0$ such that $t \geq s \geq c$,

$$(10.26) \quad \mathbb{E} \left(\left| e^{-\lambda^* t} Z_f^{\phi^s}(t) - W_\infty M_f^{\phi^s}(\infty) \right| \right) \leq D b_\phi (e^{-D_1 (t-s)} + e^{-D_2 c} + e^{-D_3 (s-c)}).$$

Using (10.26) in (10.10), we obtain for any $s, t, c \geq 0$ such that $t \geq s \geq c$,

$$\mathbb{E} \left| e^{-\lambda^* t} Z_f^{\phi^s}(t) - W_\infty M_f^{\phi^s}(\infty) \right| \leq D b_\phi (e^{-D_1 (t-s)} + e^{-D_2 c} + e^{-D_3 (s-c)}) + C_1 b_\phi e^{-C_2 t} + 2C b_\phi e^{-(\lambda^* - \beta')s}.$$

The theorem now follows by taking $s = t/2$ and $c = t/4$. ■

Recall $\lambda_\ell, \lambda_\ell^{(k)}$ for $k, \ell \geq 0$ from (3.4), with f_1 replaced by f (this section considers the model without change point). The following lemma uses the exponential convergence rate established in Theorem 3.5 along with some continuity estimates to furnish a quantitative sup-norm bound on appropriate statistics on suitably chosen intervals.

LEMMA 10.2. *Consider a continuous time branching process with attachment function f that satisfies Assumptions 2.4, 3.1 and 3.2. There exist $\omega_1 \in (0, 1), \epsilon^* \in (0, 1)$ and positive constants C, ω_2 such that for all $\epsilon \leq \epsilon^*$ and all $T \in [\frac{1-\epsilon}{\lambda^*} \log n, \frac{1+\epsilon}{\lambda^*} \log n]$,*

$$\mathbb{E} \left(n^{\omega_1} \sup_{t \in [0, 2\epsilon \log n / \lambda^*]} \left| e^{-\lambda^* T} \sum_{\ell=0}^{\infty} \lambda_{\ell}(t) D(\ell, T) - \frac{1}{\lambda^* m^*} \sum_{\ell=0}^{\infty} \lambda_{\ell}(t) p_{\ell} W_{\infty} \right| \right) \leq C n^{-\omega_2}$$

and for any $k \geq 0$,

$$\mathbb{E} \left(n^{\omega_1} \sup_{t \in [0, 2\epsilon \log n / \lambda^*]} \left| e^{-\lambda^* T} \sum_{\ell=0}^{\infty} \lambda_{\ell}^{(k)}(t) D(\ell, T) - \frac{1}{\lambda^* m^*} \sum_{\ell=0}^{\infty} \lambda_{\ell}^{(k)}(t) p_{\ell} W_{\infty} \right| \right) \leq C(k+1) n^{-\omega_2}.$$

PROOF. For any t , consider the characteristic $\phi(s) = \sum_{\ell=0}^{\infty} \lambda_{\ell}(t) \mathbb{1}\{\xi_f(s) = \ell\}$. Then $Z_f^{\phi}(s) = \sum_{\ell=0}^{\infty} \lambda_{\ell}(t) D(\ell, s)$. By Lemma 6.6 (ii), $\lim_{t \rightarrow \infty} e^{-\lambda^* t} m_f(t) = \frac{1}{\lambda^* m^*}$. Moreover, as Assumption 3.1 holds, by Lemma 9.3, there exists a constant $C > 0$ such that for each $\ell \geq 0$, $w_{\ell} \leq C(\ell+1)$. Thus, there exists a constant $C' > 0$ such that for any $\ell \geq 0$,

$$\sup_{t \geq 0} e^{-\lambda^* t} \lambda_{\ell}(t) \leq 1 + w_{\ell} \left(\sup_{t \geq 0} e^{-\lambda^* t} m_f(t) \right) \leq C'(\ell+1).$$

Hence, the hypotheses of Theorem 3.5 hold with $b_{\phi} = C' e^{\lambda^* t}$. Consequently, for any $\epsilon \in (0, 1)$, any $t \in [0, 2\epsilon \log n / \lambda^*]$ and any $T \in [\frac{1-\epsilon}{\lambda^*} \log n, \frac{1+\epsilon}{\lambda^*} \log n]$,

$$\begin{aligned} \mathbb{E} \left(\left| e^{-\lambda^* T} \sum_{\ell=0}^{\infty} \lambda_{\ell}(t) D(\ell, T) - \frac{1}{\lambda^* m^*} \sum_{\ell=0}^{\infty} \lambda_{\ell}(t) p_{\ell} W_{\infty} \right| \right) \\ \leq C_1 C e^{\lambda^* t} e^{-\frac{C_2(1-\epsilon)}{\lambda^*} \log n} \leq C_1 C e^{2\epsilon \log n} e^{-\frac{C_2(1-\epsilon)}{\lambda^*} \log n}. \end{aligned}$$

Therefore, choosing ϵ^* small enough, there exists $\theta_1 > 0$ such that for any $\epsilon \leq \epsilon^*$, any $t \in [0, 2\epsilon \log n / \lambda^*]$ and any $T \in [\frac{1-\epsilon}{\lambda^*} \log n, \frac{1+\epsilon}{\lambda^*} \log n]$,

$$(10.27) \quad \mathbb{E} \left(\left| e^{-\lambda^* T} \sum_{\ell=0}^{\infty} \lambda_{\ell}(t) D(\ell, T) - \frac{1}{\lambda^* m^*} \sum_{\ell=0}^{\infty} \lambda_{\ell}(t) p_{\ell} W_{\infty} \right| \right) \leq n^{-\theta_1}.$$

Take any $\theta_2 \in (0, \theta_1)$ and a partition of $[0, 2\epsilon \log n / \lambda^*]$ into $t_0 < t_1 < \dots < t_{\lfloor (2\epsilon \log n / \lambda^*) n^{\theta_2} \rfloor + 1}$ of mesh $n^{-\theta_2}$. By Lemma 8.4, for any j and any $t \in [t_j, t_{j+1}]$, there exist constants $C, C' > 0$ independent of ϵ, n such that

$$\begin{aligned} (10.28) \quad & \left| e^{-\lambda^* T} \sum_{\ell=0}^{\infty} \lambda_{\ell}(t) D(\ell, T) - \frac{1}{\lambda^* m^*} \sum_{\ell=0}^{\infty} \lambda_{\ell}(t) p_{\ell} W_{\infty} \right| \\ & - \left| e^{-\lambda^* T} \sum_{\ell=0}^{\infty} \lambda_{\ell}(t_j) D(\ell, T) - \frac{1}{\lambda^* m^*} \sum_{\ell=0}^{\infty} \lambda_{\ell}(t_j) p_{\ell} W_{\infty} \right| \\ & \leq e^{-\lambda^* T} \sum_{\ell=0}^{\infty} |\lambda_{\ell}(t) - \lambda_{\ell}(t_j)| D(\ell, T) + \frac{1}{\lambda^* m^*} \sum_{\ell=0}^{\infty} |\lambda_{\ell}(t) - \lambda_{\ell}(t_j)| p_{\ell} W_{\infty} \\ & \leq \frac{C n^{C'\epsilon}}{n^{1-\epsilon+\theta_2}} \sum_{\ell=0}^{\infty} (\ell+1) D(\ell, T) + \frac{C n^{C'\epsilon}}{n^{\theta_2}} \sum_{\ell=0}^{\infty} (\ell+1) p_{\ell} W_{\infty} \leq \frac{2C}{n^{1-(1+C')\epsilon+\theta_2}} Z(T) + \frac{2C}{n^{\theta_2-C'\epsilon}} W_{\infty}. \end{aligned}$$

Using (10.27), (10.28) and the union bound, we obtain for any $\omega' > 0$,

$$\begin{aligned}
& \mathbb{E} \left(n^{\omega'} \sup_{t \in [0, 2\epsilon \log n / \lambda^*]} \left| e^{-\lambda^* T} \sum_{\ell=0}^{\infty} \lambda_{\ell}(t) D(\ell, T) - \frac{1}{\lambda^* m^*} \sum_{\ell=0}^{\infty} \lambda_{\ell}(t) p_{\ell} W_{\infty} \right| \right) \\
& \leq \mathbb{E} \left(n^{\omega'} \sup_{1 \leq j \leq \lfloor (2\epsilon \log n / \lambda^*) n^{\theta_2} \rfloor + 1} \left| e^{-\lambda^* T} \sum_{\ell=0}^{\infty} \lambda_{\ell}(t_j) D(\ell, T) - \frac{1}{\lambda^* m^*} \sum_{\ell=0}^{\infty} \lambda_{\ell}(t_j) p_{\ell} W_{\infty} \right| \right) \\
& \quad + \mathbb{E} \left(\frac{2C n^{\omega'}}{n^{1-(1+C')\epsilon + \theta_2}} Z(T) + \frac{2C n^{\omega'}}{n^{\theta_2 - C'\epsilon}} W_{\infty} \right) \\
& \leq n^{\omega'} \sum_{j=0}^{\lfloor (2\epsilon \log n / \lambda^*) n^{\theta_2} \rfloor + 1} \mathbb{E} \left(\left| e^{-\lambda^* T} \sum_{\ell=0}^{\infty} \lambda_{\ell}(t_j) D(\ell, T) - \frac{1}{\lambda^* m^*} \sum_{\ell=0}^{\infty} \lambda_{\ell}(t_j) p_{\ell} W_{\infty} \right| \right) \\
& \quad + n^{\omega'} \mathbb{E} \left(\frac{2C}{n^{1-(1+C')\epsilon + \theta_2}} Z(T) + \frac{2C}{n^{\theta_2 - C'\epsilon}} W_{\infty} \right) \leq \frac{C'' \epsilon \log n}{n^{\theta_1 - \theta_2 - \omega'}} + \frac{C''}{n^{\theta_2 - (2+C')\epsilon - \omega'}} + \frac{C''}{n^{\theta_2 - C'\epsilon - \omega'}}
\end{aligned}$$

for some constant $C'' > 0$. Taking $\epsilon^* < \theta_2 / (2 + C')$ and any $\omega' < \min\{\theta_1 - \theta_2, \theta_2 - (2 + C')\epsilon^*, 1\}$, this proves the first assertion in the lemma. The second assertion follows similarly upon noting that $\lambda_{\ell}^{(k)} \leq \lambda_{\ell}$ for each $k \geq 0$ (and thus the constant C in the expectation bound can be chosen uniformly over k) and using Corollary 8.6 in place of Lemma 8.4 (which accounts for the $(k + 1)$ in the bound). $\blacksquare$

PROOF OF THEOREM 3.4. Take $\epsilon^{**} \leq \epsilon^*$ (where ϵ^* is as in Lemma 10.2) and any $\epsilon \leq \epsilon^{**}$. We abbreviate

$$\begin{aligned}
\mathcal{S}_n &:= \sup_{t \in [0, 2\epsilon \log n / \lambda^*]} \left| \sum_{\ell=0}^{\infty} \lambda_{\ell}(t) D\left(\ell, \frac{1-\epsilon}{\lambda^*} \log n\right) - \frac{n^{1-\epsilon}}{\lambda^* m^*} \sum_{\ell=0}^{\infty} \lambda_{\ell}(t) p_{\ell} W_{\infty} \right|, \\
\mathcal{S}_n^{(k)} &:= \sup_{t \in [0, 2\epsilon \log n / \lambda^*]} \left| \sum_{\ell=0}^{\infty} \lambda_{\ell}^{(k)}(t) D\left(\ell, \frac{1-\epsilon}{\lambda^*} \log n\right) - \frac{n^{1-\epsilon}}{\lambda^* m^*} \sum_{\ell=0}^{\infty} \lambda_{\ell}^{(k)}(t) p_{\ell} W_{\infty} \right|.
\end{aligned}$$

Observe that for any $k \geq 0$, using the fact that $\lambda_{\ell}(\cdot)$ is an increasing function and $\lambda_{\ell}(0) = 1$ for each $\ell \geq 0$,

$$\begin{aligned}
& \sup_{t \in [0, 2\epsilon \log n / \lambda^*]} \left| \frac{\sum_{\ell=0}^{\infty} \lambda_{\ell}^{(k)}(t) D\left(\ell, \frac{1-\epsilon}{\lambda^*} \log n\right)}{\sum_{\ell=0}^{\infty} \lambda_{\ell}(t) D\left(\ell, \frac{1-\epsilon}{\lambda^*} \log n\right)} - \frac{\sum_{\ell=0}^{\infty} \lambda_{\ell}^{(k)}(t) p_{\ell}}{\sum_{\ell=0}^{\infty} \lambda_{\ell}(t) p_{\ell}} \right| \\
& \leq \frac{\mathcal{S}_n^{(k)}}{\sum_{\ell=0}^{\infty} \lambda_{\ell}(t) D\left(\ell, \frac{1-\epsilon}{\lambda^*} \log n\right)} + \frac{\mathcal{S}_n \left(\sum_{\ell=0}^{\infty} \lambda_{\ell}^{(k)}(t) p_{\ell} W_{\infty} \right)}{\left(\sum_{\ell=0}^{\infty} \lambda_{\ell}(t) p_{\ell} W_{\infty} \right) \left(\sum_{\ell=0}^{\infty} \lambda_{\ell}(t) D\left(\ell, \frac{1-\epsilon}{\lambda^*} \log n\right) \right)} \\
& \leq \frac{\mathcal{S}_n^{(k)}}{\sum_{\ell=0}^{\infty} \lambda_{\ell}(0) D\left(\ell, \frac{1-\epsilon}{\lambda^*} \log n\right)} + \frac{\mathcal{S}_n}{\left(\sum_{\ell=0}^{\infty} \lambda_{\ell}(0) D\left(\ell, \frac{1-\epsilon}{\lambda^*} \log n\right) \right)} = \frac{\mathcal{S}_n^{(k)}}{Z\left(\frac{1-\epsilon}{\lambda^*} \log n\right)} + \frac{\mathcal{S}_n}{Z\left(\frac{1-\epsilon}{\lambda^*} \log n\right)}.
\end{aligned}$$

Recalling ω_1 from Lemma 10.2,

$$\begin{aligned}
& n^{\omega_1} \sum_{k=0}^{\infty} 2^{-k} \left(\sup_{t \in [0, 2\epsilon \log n / \lambda^*]} \left| \frac{\sum_{\ell=0}^{\infty} \lambda_{\ell}^{(k)}(t) D\left(\ell, \frac{1-\epsilon}{\lambda^*} \log n\right)}{\sum_{\ell=0}^{\infty} \lambda_{\ell}(t) D\left(\ell, \frac{1-\epsilon}{\lambda^*} \log n\right)} - \frac{\sum_{\ell=0}^{\infty} \lambda_{\ell}^{(k)}(t) p_{\ell}}{\sum_{\ell=0}^{\infty} \lambda_{\ell}(t) p_{\ell}} \right| \right) \\
& \leq \frac{n^{1-\epsilon}}{Z\left(\frac{1-\epsilon}{\lambda^*} \log n\right)} \sum_{k=0}^{\infty} 2^{-k} \left(\frac{\mathcal{S}_n^{(k)}}{n^{1-\epsilon-\omega_1}} + \frac{\mathcal{S}_n}{n^{1-\epsilon-\omega_1}} \right).
\end{aligned}$$

Using Lemma 10.2, for any $\eta > 0$,

$$\begin{aligned} \mathbb{P} \left(\sum_{k=0}^{\infty} 2^{-k} \left(\frac{\mathcal{S}_n^{(k)} + \mathcal{S}_n}{n^{1-\epsilon-\omega_1}} \right) > \eta \right) &\leq \eta^{-1} \sum_{k=0}^{\infty} 2^{-k} \frac{1}{n^{1-\epsilon-\omega_1}} \mathbb{E} \left(\mathcal{S}_n^{(k)} + \mathcal{S}_n \right) \\ &\leq \eta^{-1} \sum_{k=0}^{\infty} 2^{-k} (k+2) C n^{-\omega_2} \leq C' \eta^{-1} n^{-\omega_2} \end{aligned}$$

for positive constants C, C' . Moreover, $\frac{n^{1-\epsilon}}{Z \left(\frac{1-\epsilon}{\lambda^*} \log n \right)} \xrightarrow{P} \frac{\lambda^* m^*}{W_{\infty}}$ as $n \rightarrow \infty$ by Lemma 6.6. Combining these,

$$(10.29) \quad n^{\omega_1} \sum_{k=0}^{\infty} 2^{-k} \left(\sup_{t \in [0, 2\epsilon \log n / \lambda^*]} \left| \frac{\sum_{\ell=0}^{\infty} \lambda_{\ell}^{(k)}(t) D \left(\ell, \frac{1-\epsilon}{\lambda^*} \log n \right)}{\sum_{\ell=0}^{\infty} \lambda_{\ell}(t) D \left(\ell, \frac{1-\epsilon}{\lambda^*} \log n \right)} - \frac{\sum_{\ell=0}^{\infty} \lambda_{\ell}^{(k)}(t) p_{\ell}}{\sum_{\ell=0}^{\infty} \lambda_{\ell}(t) p_{\ell}} \right| \right) \xrightarrow{P} 0.$$

Moreover, it is straightforward to check that

$$\begin{aligned} (10.30) \quad &\sup_{t \in [0, 2\epsilon \log n / \lambda^*]} \left| \frac{D \left(k, \frac{1-\epsilon}{\lambda^*} \log n + t \right)}{Z \left(\frac{1-\epsilon}{\lambda^*} \log n + t \right)} - \frac{\sum_{\ell=0}^{\infty} \lambda_{\ell}^{(k)}(t) D \left(\ell, \frac{1-\epsilon}{\lambda^*} \log n \right)}{\sum_{\ell=0}^{\infty} \lambda_{\ell}(t) D \left(\ell, \frac{1-\epsilon}{\lambda^*} \log n \right)} \right| \\ &\leq \frac{1}{Z \left(\frac{1-\epsilon}{\lambda^*} \log n \right)} \sup_{t \in [0, 2\epsilon \log n / \lambda^*]} \left| D \left(k, \frac{1-\epsilon}{\lambda^*} \log n + t \right) - \sum_{\ell=0}^{\infty} \lambda_{\ell}^{(k)}(t) D \left(\ell, \frac{1-\epsilon}{\lambda^*} \log n \right) \right| \\ &\quad + \frac{1}{Z \left(\frac{1-\epsilon}{\lambda^*} \log n \right)} \sup_{t \in [0, 2\epsilon \log n / \lambda^*]} \left| Z \left(\frac{1-\epsilon}{\lambda^*} \log n + t \right) - \sum_{\ell=0}^{\infty} \lambda_{\ell}(t) D \left(\ell, \frac{1-\epsilon}{\lambda^*} \log n \right) \right|. \end{aligned}$$

Abbreviate

$$\begin{aligned} \hat{\mathcal{S}}_n^{(k)} &:= \sup_{t \in [0, 2\epsilon \log n / \lambda^*]} \left| D \left(k, \frac{1-\epsilon}{\lambda^*} \log n + t \right) - \sum_{\ell=0}^{\infty} \lambda_{\ell}^{(k)}(t) D \left(\ell, \frac{1-\epsilon}{\lambda^*} \log n \right) \right|, \\ \hat{\mathcal{S}}_n &:= \sup_{t \in [0, 2\epsilon \log n / \lambda^*]} \left| Z \left(\frac{1-\epsilon}{\lambda^*} \log n + t \right) - \sum_{\ell=0}^{\infty} \lambda_{\ell}(t) D \left(\ell, \frac{1-\epsilon}{\lambda^*} \log n \right) \right|. \end{aligned}$$

By conditioning on $\mathcal{F}_n \left(\frac{1-\epsilon}{\lambda^*} \log n \right)$ and applying Lemma 8.11, we obtain $\omega'_1 \in (0, 1), \omega'_2 > 0$ not depending on ϵ such that for any $\eta > 0$,

$$\begin{aligned} (10.31) \quad &\mathbb{P} \left(\sum_{k=0}^{\infty} 2^{-k} \left(\frac{\hat{\mathcal{S}}_n^{(k)}}{Z \left(\frac{1-\epsilon}{\lambda^*} \log n \right)^{1-\omega'_1}} \right) > \eta \mid \mathcal{F}_n \left(\frac{1-\epsilon}{\lambda^*} \log n \right) \right) \\ &= \mathbb{P} \left(\sum_{k=0}^{\infty} 2^{-k} \left(\frac{\hat{\mathcal{S}}_n^{(k)}}{Z \left(\frac{1-\epsilon}{\lambda^*} \log n \right)^{1-\omega'_1}} \right) > \sum_{k=0}^{\infty} \left(\frac{3}{2} \right)^{-k} \frac{\eta}{3} \mid \mathcal{F}_n \left(\frac{1-\epsilon}{\lambda^*} \log n \right) \right) \\ &\leq \sum_{k=0}^{\infty} \mathbb{P} \left(\frac{\hat{\mathcal{S}}_n^{(k)}}{Z \left(\frac{1-\epsilon}{\lambda^*} \log n \right)^{1-\omega'_1}} > \left(\frac{4}{3} \right)^k \frac{\eta}{3} \mid \mathcal{F}_n \left(\frac{1-\epsilon}{\lambda^*} \log n \right) \right) \\ &\leq C e^{C' 2\epsilon \log n / \lambda^*} \eta^{-2} Z \left(\frac{1-\epsilon}{\lambda^*} \log n \right)^{-\omega'_2} \sum_{k=0}^{\infty} (k+1)^2 \left(\frac{3}{4} \right)^{2k} = C' n^{2C' \epsilon / \lambda^*} \eta^{-2} Z \left(\frac{1-\epsilon}{\lambda^*} \log n \right)^{-\omega'_2} \end{aligned}$$

for positive constants C, C' . As $\frac{n^{1-\epsilon}}{Z(\frac{1-\epsilon}{\lambda^*} \log n)} \xrightarrow{P} \frac{\lambda^* m^*}{W_\infty}$, the bound above converges to zero almost surely if ϵ^{**} is chosen sufficiently small and $\epsilon \leq \epsilon^{**}$. Similarly,

$$(10.32) \quad \mathbb{P} \left(\sum_{k=0}^{\infty} 2^{-k} \left(\frac{\hat{\mathcal{J}}_n}{Z(\frac{1-\epsilon}{\lambda^*} \log n)^{1-\omega'_1}} \right) > \epsilon \mid \mathcal{F}_n \left(\frac{1-\epsilon}{\lambda^*} \log n \right) \right) \leq C' n^{2C'\epsilon/\lambda^*} \epsilon^{-2} Z \left(\frac{1-\epsilon}{\lambda^*} \log n \right)^{-\omega_2}.$$

Using (10.30), (10.31), (10.32) and recalling that $\frac{n^{1-\epsilon}}{Z(\frac{1-\epsilon}{\lambda^*} \log n)} \xrightarrow{P} \frac{\lambda^* m^*}{W_\infty}$ as $n \rightarrow \infty$, we conclude

$$(10.33) \quad n^{(1-\epsilon)\omega'_1} \sum_{k=0}^{\infty} 2^{-k} \left(\sup_{t \in [0, 2\epsilon \log n / \lambda^*]} \left| \frac{D(k, \frac{1-\epsilon}{\lambda^*} \log n + t)}{Z(\frac{1-\epsilon}{\lambda^*} \log n + t)} - \frac{\sum_{\ell=0}^{\infty} \lambda_\ell^{(k)}(t) D(\ell, \frac{1-\epsilon}{\lambda^*} \log n)}{\sum_{\ell=0}^{\infty} \lambda_\ell(t) D(\ell, \frac{1-\epsilon}{\lambda^*} \log n)} \right| \right) \xrightarrow{P} 0.$$

Choosing $\omega^* = \min\{\omega_1, (1-\epsilon)\omega'_1\}$, we conclude from (10.29) and (10.33) that

$$(10.34) \quad n^{\omega^*} \sum_{k=0}^{\infty} 2^{-k} \left(\sup_{t \in [0, 2\epsilon \log n / \lambda^*]} \left| \frac{D(k, \frac{1-\epsilon}{\lambda^*} \log n + t)}{Z(\frac{1-\epsilon}{\lambda^*} \log n + t)} - \frac{\sum_{\ell=0}^{\infty} \lambda_\ell^{(k)}(t) p_\ell}{\sum_{\ell=0}^{\infty} \lambda_\ell(t) p_\ell} \right| \right) \xrightarrow{P} 0.$$

Finally, we claim that for each $k \geq 0, t \geq 0$,

$$(10.35) \quad \sum_{\ell=0}^{\infty} \lambda_\ell^{(k)}(t) p_\ell / \sum_{\ell=0}^{\infty} \lambda_\ell(t) p_\ell = p_k.$$

To see this, observe that the following limits hold as $n \rightarrow \infty$: $\frac{Z(\frac{1-\epsilon}{\lambda^*} \log n + t)}{n^{1-\epsilon}} \xrightarrow{P} \frac{e^{\lambda^* t} W_\infty}{\lambda^* m^*}$, and $\frac{D(k, \frac{1-\epsilon}{\lambda^*} \log n + t)}{n^{1-\epsilon}} \xrightarrow{P} \frac{p_k e^{\lambda^* t} W_\infty}{\lambda^* m^*}$. Thus $\frac{D(k, \frac{1-\epsilon}{\lambda^*} \log n + t)}{Z(\frac{1-\epsilon}{\lambda^*} \log n + t)} \xrightarrow{P} p_k$. But from (10.34),

$$\frac{D(k, \frac{1-\epsilon}{\lambda^*} \log n + t)}{Z(\frac{1-\epsilon}{\lambda^*} \log n + t)} \xrightarrow{P} \frac{\sum_{\ell=0}^{\infty} \lambda_\ell^{(k)}(t) p_\ell}{\sum_{\ell=0}^{\infty} \lambda_\ell(t) p_\ell}.$$

(10.35) follows from the above two observations. The theorem now follows from (10.34) and (10.35). $\blacksquare$

11. Proofs: Change point detection. Throughout this section, we assume that f_0 satisfies Assumptions 2.4, 3.1 and 3.2, and f_1 satisfies Assumptions 2.4 and 3.16. Recall $\lambda_\ell, \lambda_\ell^{(k)}$ for $k, \ell \geq 0$ defined in (3.4) and the functional $\Phi_a : \mathcal{P} \rightarrow \mathcal{P}$ defined for each $a > 0$ in (3.5).

LEMMA 11.1. $\lim_{a \rightarrow \infty} \Phi_a(\mathbf{p}) = \mathbf{p}^1$ (where the limit is taken in the coordinate-wise sense).

PROOF. As f_1 satisfies Assumptions 2.4 and 3.16, for each $k \geq 0$, by Lemma 6.6 (ii), $\lim_{t \rightarrow \infty} e^{-\lambda_1^* t} m_{f_1}(t) = (\lambda_1^* m_1^*)^{-1}$ and $\lim_{t \rightarrow \infty} e^{-\lambda_1^* t} m_{f_1}^{(k)}(t) = p_k^1 / (\lambda_1^* m_1^*)$ and consequently,

$$(11.1) \quad \lim_{t \rightarrow \infty} e^{-\lambda_1^* t} \lambda_\ell(t) = w_\ell / (\lambda_1^* m_1^*), \quad \lim_{t \rightarrow \infty} e^{-\lambda_1^* t} \lambda_\ell^{(k)}(t) = p_k^1 w_\ell / (\lambda_1^* m_1^*).$$

Moreover, it is easy to see from (3.4) that for any $\ell, k \geq 0$, $e^{-\lambda_1^* t} \lambda_\ell(t) \leq 1 + \left(\sup_{u \geq 0} e^{-\lambda_1^* u} m_{f_1}(u) \right) w_\ell$ and $e^{-\lambda_1^* t} \lambda_\ell^{(k)}(t) \leq 1 + \left(\sup_{u \geq 0} e^{-\lambda_1^* u} m_{f_1}(u) \right) w_\ell$ for all $t \geq 0$ and this bound is finite. By this observation, we can apply the dominated convergence theorem and (11.1) in the formula of $\Phi_a(\mathbf{p})$ to obtain the lemma. $\blacksquare$

LEMMA 11.2. *For any $s, t \geq 0$ and any $j, k \geq 0$,*

$$\sum_{\ell=0}^{\infty} \lambda_j^{(\ell)}(t) \lambda_{\ell}(s) = \lambda_j(s+t), \quad \sum_{\ell=0}^{\infty} \lambda_j^{(\ell)}(t) \lambda_{\ell}^{(k)}(s) = \lambda_j^{(k)}(s+t).$$

Consequently, for any $\mathbf{p} \in \mathcal{P}$, we have $\Phi_s(\Phi_t(\mathbf{p})) = \Phi_{s+t}(\mathbf{p})$.

PROOF. We will only prove the first assertion. The second one follows similarly. Denote by $\text{BP}^{(j)}(\cdot)$ the continuous time branching process with attachment function $i \mapsto f_1(i+j)$ and denote by $D_n^{(j)}(\ell, t)$ the corresponding number of vertices of degree ℓ at time t (excluding the root). Then

$$\begin{aligned} \mathbb{E}\left(|\text{BP}^{(j)}(t+s)| \mid \mathcal{F}_n(t)\right) &= \sum_{\ell=j}^{\infty} \mathbb{1}\left\{\xi_{f_1}^{(j)}(t) = \ell - j\right\} \left(1 + \int_0^s m_{f_1}(s-v) \mu_{f_1}^{(\ell)}(dv)\right) \\ &\quad + \sum_{\ell=0}^{\infty} D_n^{(j)}(\ell, t) \left(1 + \int_0^s m_{f_1}(s-v) \mu_{f_1}^{(\ell)}(dv)\right) \end{aligned}$$

where the first term denotes the expected number of vertices born to the root (counting the root itself) in the time interval $[t, t+s]$ and the second term denotes the expected number of vertices born in the time interval $[t, t+s]$ to those vertices born in the time interval $(0, t]$, both expectations conditional on $\mathcal{F}_n(t)$. Taking expectation on both sides of the above expression and noting that $\lambda_j(t+s) = \mathbb{E}(|\text{BP}^{(j)}(t+s)|)$ and $\mathbb{E}\left(D_n^{(j)}(\ell, t)\right) = \int_0^t m_{f_1}^{(\ell)}(t-u) \mu_{f_1}^{(j)}(du)$, we obtain

$$\begin{aligned} \lambda_j(t+s) &= \sum_{\ell=0}^{\infty} \left(\mathbb{P}\left(\xi_{f_1}^{(j)}(t) = \ell - j\right) + \int_0^t m_{f_1}^{(\ell)}(t-u) \mu_{f_1}^{(j)}(du) \right) \left(1 + \int_0^s m_{f_1}(s-v) \mu_{f_1}^{(\ell)}(dv)\right) \\ &= \sum_{\ell=0}^{\infty} \lambda_j^{(\ell)}(t) \lambda_{\ell}(s). \end{aligned}$$

To prove the semigroup property, note that for each $k \geq 0$,

$$\begin{aligned} (\Phi_s(\Phi_t(\mathbf{p})))_k &= \left(\frac{\sum_{\ell=0}^{\infty} (\Phi_t(\mathbf{p}))_{\ell} \lambda_{\ell}^{(k)}(s)}{\sum_{\ell=0}^{\infty} (\Phi_t(\mathbf{p}))_{\ell} \lambda_{\ell}(s)} \right) = \left(\frac{\sum_{\ell=0}^{\infty} \left(\sum_{j=0}^{\infty} p_j \lambda_j^{(\ell)}(t) \right) \lambda_{\ell}^{(k)}(s)}{\sum_{\ell=0}^{\infty} \left(\sum_{j=0}^{\infty} p_j \lambda_j^{(\ell)}(t) \right) \lambda_{\ell}(s)} \right) \\ &= \frac{\sum_{j=0}^{\infty} p_j \left(\sum_{\ell=0}^{\infty} \lambda_j^{(\ell)}(t) \lambda_{\ell}^{(k)}(s) \right)}{\sum_{j=0}^{\infty} p_j \left(\sum_{\ell=0}^{\infty} \lambda_j^{(\ell)}(t) \lambda_{\ell}(s) \right)} = \frac{\sum_{j=0}^{\infty} p_j \lambda_j^{(k)}(s+t)}{\sum_{j=0}^{\infty} p_j \lambda_j(s+t)} = (\Phi_{s+t}(\mathbf{p}))_k. \end{aligned}$$

■

LEMMA 11.3. *For any $a > 0$ and any $\mathbf{p} \in \mathcal{P}$ such that $\mathbf{p} \neq \mathbf{p}^1$, we have $\Phi_a(\mathbf{p}) \neq \mathbf{p}$.*

PROOF. Suppose there exists $a > 0$ and $\mathbf{p} \neq \mathbf{p}^1$ such that $\Phi_a(\mathbf{p}) = \mathbf{p}$. Then by Lemma 11.2, for any $n \geq 1$, $\Phi_{na}(\mathbf{p}) = \mathbf{p}$. Letting $n \rightarrow \infty$ and using Lemma 11.1, we obtain $\mathbf{p}^1 = \mathbf{p}$ which gives a contradiction. ■

PROOF OF THEOREM 3.22. Recall ω^* , ϵ^{**} from Theorem 3.4 applied to the branching process with attachment function f_0 and fix any $\epsilon \leq \epsilon^{**}$. Let λ_0^* denote the associated

Malthusian rate. Take any $n_0 \geq 1$ such that $h_n \geq 1/\gamma$ for all $n \geq n_0$. Observe that for any $\eta > 0$ and any $n \geq n_0$,

$$\begin{aligned} & \mathbb{P} \left(n^{\omega^*} \sum_{k=0}^{\infty} 2^{-k} \sup_{1/h_n \leq t \leq \gamma} \left| \frac{D(k, T_{\lfloor nt \rfloor})}{\lfloor nt \rfloor} - p_k^0 \right| > \eta \right) \\ & \leq \mathbb{P} \left(n^{\omega^*} \sum_{k=0}^{\infty} 2^{-k} \left(\sup_{t \in [0, 2\epsilon \log n / \lambda_0^*]} \left| \frac{D\left(\ell, \frac{1-\epsilon}{\lambda_0^*} \log n + t\right)}{Z\left(\frac{1-\epsilon}{\lambda_0^*} \log n + t\right)} - p_k^0 \right| > \eta \right) \right) \\ & \quad + \mathbb{P} \left(T_{\lfloor n/h_n \rfloor} < \frac{1-\epsilon}{\lambda_0^*} \log n \right) + \mathbb{P} \left(T_{\lfloor n\gamma \rfloor} > \frac{1+\epsilon}{\lambda_0^*} \log n \right). \end{aligned}$$

The first term in the above bound converges to zero by Theorem 3.4. Further,

$$\mathbb{P} \left(T_{\lfloor n/h_n \rfloor} < \frac{1-\epsilon}{\lambda_0^*} \log n \right) \rightarrow 0$$

because $\lambda_0^* T_{\lfloor n/h_n \rfloor} / \log(n/h_n) \xrightarrow{P} 1$ as $n \rightarrow \infty$ by Lemma 6.6 (ii) and by assumption, $\log h_n / \log n \rightarrow 0$. Similarly, $\mathbb{P} \left(T_{\lfloor n\gamma \rfloor} > \frac{1+\epsilon}{\lambda_0^*} \log n \right) \rightarrow 0$ because $\lambda_0^* T_{\lfloor n\gamma \rfloor} / \log(n\gamma) \xrightarrow{P} 1$ as $n \rightarrow \infty$. Thus, we conclude

$$(11.2) \quad n^{\omega^*} \sum_{k=0}^{\infty} 2^{-k} \sup_{1/h_n \leq t \leq \gamma} \left| \frac{D(k, T_{\lfloor nt \rfloor})}{\lfloor nt \rfloor} - p_k^0 \right| \xrightarrow{P} 0$$

as $n \rightarrow \infty$ which, along with the fact that $\omega^* \in (0, 1)$, implies

$$n^{\omega^*} \sum_{k=0}^{\infty} 2^{-k} \sup_{1/h_n \leq t \leq \gamma} \left| \frac{D(k, T_{\lfloor nt \rfloor})}{nt} - \frac{D(k, T_{\lfloor n/h_n \rfloor})}{n/h_n} \right| \xrightarrow{P} 0.$$

As $\frac{\log b_n}{\log n} \rightarrow 0$ as $n \rightarrow \infty$, the above implies $b_n \sum_{k=0}^{\infty} 2^{-k} \sup_{1/h_n \leq t \leq \gamma} \left| \frac{D(k, T_{\lfloor nt \rfloor})}{nt} - \frac{D(k, T_{\lfloor n/h_n \rfloor})}{n/h_n} \right| \xrightarrow{P} 0$. From this observation and the definition of $\hat{T}_n$, we conclude that

$$(11.3) \quad \mathbb{P}(\hat{T}_n \geq \gamma) \rightarrow 1 \text{ as } n \rightarrow \infty.$$

Moreover, by Theorem 3.8, for any $t > \gamma$ and any $k \geq 0$, $\left| \frac{D(k, T_{\lfloor tn \rfloor})}{tn} - (\Phi_{a_t}(\mathbf{p}^0))_k \right| \xrightarrow{P} 0$ and hence, by (11.2) and the dominated convergence theorem, as $n \rightarrow \infty$,

$$\sum_{k=0}^{\infty} 2^{-k} \left| \frac{D(k, T_{\lfloor nt \rfloor})}{nt} - \frac{D(k, T_{\lfloor n/h_n \rfloor})}{n/h_n} \right| \xrightarrow{P} \sum_{k=0}^{\infty} 2^{-k} |(\Phi_{a_t}(\mathbf{p}^0))_k - p_k^0|.$$

As $a_t > 0$ for each $t > \gamma$ and $\mathbf{p}^0 \neq \mathbf{p}^1$, by Lemma 11.3, $\Phi_{a_t}(\mathbf{p}^0) \neq \mathbf{p}^0$ and hence, the limit above is strictly positive. From the definition of $\hat{T}_n$ and the above, we conclude that for each $t > \gamma$,

$$(11.4) \quad \mathbb{P}(\hat{T}_n \leq t) \rightarrow 1 \text{ as } n \rightarrow \infty.$$

The theorem follows from (11.3) and (11.4). ■

Acknowledgements. SBh and IC were partially supported by NSF grants DMS-1613072, DMS-1606839 and ARO grant W911NF-17-1-0010. SBh is partially supported by NSF DMS-2113662. SBa is partially supported by the NSF CAREER award DMS-2141621. SBa and SBh were also supported in part by the NSF RTG grant DMS-2134107. We thank three anonymous referees and an associate editor for many suggestions that lead to a significant improvement in the original submission.

References.

- [1] R. Albert and A.-L. Barabási, *Statistical mechanics of complex networks*, Rev. Modern Phys. **74** (2002), no. 1, 47–97.
- [2] D. Aldous, *Asymptotic fringe distributions for general families of random trees*, Ann. Appl. Probab. **1** (1991), no. 2, 228–266.
- [3] K. B. Athreya and S. Karlin, *Embedding of urn schemes into continuous time Markov branching processes and related limit theorems*, The Annals of Mathematical Statistics **39** (1968), no. 6, 1801–1817.
- [4] K. B. Athreya and P. E. Ney, *Branching processes*, Springer-Verlag, New York-Heidelberg, 1972. Die Grundlehren der mathematischen Wissenschaften, Band 196.
- [5] J. Bai, *Estimating multiple breaks one at a time*, Econometric Theory **13** (1997), no. 3, 315–352.
- [6] J. Bai and P. Perron, *Estimating and testing linear models with multiple structural changes*, Econometrica **66** (1998), no. 1, 47–78.
- [7] J. Bai and P. Perron, *Computation and analysis of multiple structural change models*, Journal of Applied Econometrics **18** (2003), no. 1, 1–22.
- [8] S. Banerjee and S. Bhamidi, *Persistence of hubs in growing random networks*, Probability Theory and Related Fields **180** (2021), no. 3, 891–953.
- [9] S. Banerjee and S. Bhamidi, *Root finding algorithms and persistence of jordan centrality in growing random trees*, The Annals of Applied Probability **32** (2022), no. 3, 2180–2210.
- [10] A. L. Barabási and R. Albert, *Emergence of scaling in random networks*, Science **286** (1999), no. 5439, 509–512.
- [11] J.-B. Bardet, A. Christen, and J. Fontbona, *Quantitative exponential bounds for the renewal theorem with spread-out distributions*, Markov Process. Related Fields **23** (2017), no. 1, 67–86.
- [12] F. Bergeron, P. Flajolet, and B. Salvy, *Varieties of increasing trees*, CAAP '92 (Rennes, 1992), 1992, pp. 24–48.
- [13] S. Bhamidi, *Universal techniques to analyze preferential attachment trees: Global and local analysis*, In preparation. Version August (2007).
- [14] S. Bhamidi, J. Jin, and A. Nobel, *Change point detection in network models: preferential attachment and long range dependence*, The Annals of Applied Probability **28** (2018), no. 1, 35–78.
- [15] B. Bollobás, *Random graphs*, Second, Cambridge Studies in Advanced Mathematics, vol. 73, Cambridge University Press, Cambridge, 2001.
- [16] B. Bollobás, O. Riordan, J. Spencer, and G. Tusnády, *The degree sequence of a scale-free random graph process*, Random Structures & Algorithms **18** (May 2001), no. 3, 279–290.
- [17] B. E. Brodsky and B. S. Darkhovsky, *Nonparametric methods in change-point problems*, Mathematics and its Applications, vol. 243, Kluwer Academic Publishers Group, Dordrecht, 1993.
- [18] S. Bubeck, L. Devroye, and G. Lugosi, *Finding Adam in random growing trees*, Random Structures & Algorithms **50** (2017), no. 2, 158–172.
- [19] S. Bubeck, E. Mossel, and M. Z. Rácz, *On the influence of the seed graph in the preferential attachment model*, IEEE Transactions on Network Science and Engineering **2** (2015), no. 1, 30–39.
- [20] M. Csörgő and L. Horváth, *Limit theorems in change-point analysis*, Wiley Series in Probability and Statistics, John Wiley & Sons, Ltd., Chichester, 1997. With a foreword by David Kendall.
- [21] N. Curien, T. Duquesne, I. Kortchemski, and I. Manolescu, *Scaling limits and influence of the seed graph in preferential attachment trees*, Journal de l'Ecole polytechnique-Mathématiques **2** (2015), 1–34.
- [22] L. Devroye, *Branching processes and their applications in the analysis of tree structures and tree algorithms*, Probabilistic methods for algorithmic discrete mathematics, 1998, pp. 249–314.
- [23] L. Devroye and J. Lu, *The strong convergence of maximal degrees in uniform random recursive trees and dags*, Random Structures & Algorithms **7** (1995), no. 1, 1–14.
- [24] M. Drmota, *Random trees: An interplay between combinatorics and probability*, SpringerWienNewYork, Vienna, 2009.
- [25] R. Durrett, *Random graph dynamics*, Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, Cambridge, 2007.
- [26] P. Flajolet and R. Sedgewick, *Analytic combinatorics*, Cambridge University Press, Cambridge, 2009.
- [27] C. Goldschmidt and J. B. Martin, *Random recursive trees and the Bolthausen-Sznitman coalescent*, Electron. J. Probab. **10** (2005), no. 21, 718–745.
- [28] C. Holmgren, S. Janson, et al., *Fringe trees, Crump-Mode-Jagers branching processes and m-ary search trees*, Probability Surveys **14** (2017), 53–154.
- [29] P. Jagers, *Branching processes with biological applications*, Wiley-Interscience [John Wiley & Sons], London-New York-Sydney, 1975. Wiley Series in Probability and Mathematical Statistics—Applied Probability and Statistics.
- [30] P. Jagers and O. Nerman, *The growth and composition of branching populations*, Advances in Applied Probability **16** (1984), no. 2, 221–259.

- [31] P. Jagers and O. Nerman, *Limit theorems for sums determined by branching and other exponentially growing processes*, Stochastic Process. Appl. **17** (1984), no. 1, 47–71.
- [32] S. Janson, *Functional limit theorems for multitype branching processes and generalized Pólya urns*, Stochastic Processes and their Applications **110** (2004), no. 2, 177–245.
- [33] S. Janson, *Tail bounds for sums of geometric and exponential variables*, Statistics & Probability Letters **135** (2018), 1–6.
- [34] V. Jog and P.-L. Loh, *Analysis of centrality in sublinear preferential attachment trees via the Crump-Mode-Jagers branching process*, IEEE Transactions on Network Science and Engineering **4** (2016), no. 1, 1–12.
- [35] H. M. Mahmoud, *Pólya urn models*, Texts in Statistical Science Series, CRC Press, Boca Raton, FL, [2009] ©2009.
- [36] T. F. Móri, *Degree distribution nearby the origin of a preferential attachment graph*, Electronic Communications in Probability **12** (2007), 276–282.
- [37] O. Nerman, *On the convergence of supercritical general (CMJ) branching processes*, Probability Theory and Related Fields **57** (1981), no. 3, 365–395.
- [38] M. E. J. Newman, *Networks: An introduction*, Oxford University Press, Oxford, 2010.
- [39] M. E. Newman, *The structure and function of complex networks*, SIAM review **45** (2003), no. 2, 167–256.
- [40] J. R. Norris, *Markov chains*, Cambridge Series in Statistical and Probabilistic Mathematics, vol. 2, Cambridge University Press, Cambridge, 1998. Reprint of 1997 original.
- [41] A. B. Olshen, E. Venkatraman, R. Lucito, and M. Wigler, *Circular binary segmentation for the analysis of array-based DNA copy number data*, Biostatistics **5** (2004), no. 4, 557–572.
- [42] S. I. Resnick and G. Samorodnitsky, *Asymptotic normality of degree counts in a preferential attachment model*, Advances in Applied Probability **48** (2016), no. A, 283–299.
- [43] A. Rudas, B. Tóth, and B. Valkó, *Random trees and general branching processes*, Random Structures & Algorithms **31** (2007), no. 2, 186–202.
- [44] R. T. Smythe and H. M. Mahmoud, *A survey of recursive trees*, Theory of Probability and Mathematical Statistics **51** (1995), 1–28.
- [45] J. Szymański, *On a nonuniform random recursive tree*, Random graphs '85 (Poznań, 1985), 1987, pp. 297–306.
- [46] J. Szymański, *On the maximum degree and the height of a random recursive tree*, Random graphs '87 (Poznań, 1987), 1990, pp. 313–324.
- [47] R. van der Hofstad, *Random graphs and complex networks. Vol. 1*, Cambridge Series in Statistical and Probabilistic Mathematics, [43], Cambridge University Press, Cambridge, 2017.
- [48] Y.-C. Yao, *Estimating the number of change-points via Schwarz' criterion*, Statistics & Probability Letters **6** (1988), no. 3, 181–189.
- [49] N. R. Zhang and D. O. Siegmund, *A modified Bayes information criterion with applications to the analysis of comparative genomic hybridization data*, Biometrics **63** (2007), no. 1, 22–32.