The genealogy of Galton-Watson trees

Samuel G.G. Johnston*

Abstract

Take a continuous-time Galton-Watson tree and pick k distinct particles uniformly from those alive at a time T. What does their genealogical tree look like? The case $k = 2$ has been studied by several authors, and the near-critical asymptotics for general k appear in Harris, Johnston and Roberts (2018) [9]. Here we give the full picture.

1 Introduction

Let L be a random variable taking values in $\{0, 1, 2, \ldots\}$. Consider a continuous-time Galton-Watson tree starting with one initial particle, branching at rate 1, and with offspring distributed like L.

Let N_t be the number of particles alive at time t, and write $f(s) := \mathbb{E}[s^L]$ and $F_t(s) := \mathbb{E}[s^{N_t}]$ for the generating functions associated with the process. Let $T > 0$, and on the event $\{N_T \geq k\}$ pick k distinct particles $U_1, \ldots, U_k$ uniformly from those alive at time T. For each earlier time $t \in [0, T]$, define the equivalence relation $\sim_t$ on $\{1, \ldots, k\}$ by

$$i \sim_t j \iff U_i \text{ and } U_j \text{ share a common ancestor alive at time } t.$$

We let $\pi_t^{k,L,T}$ denote the random partition of $\{1, \ldots, k\}$ corresponding to this equivalence relation. The process $(\pi_t^{k,L,T})_{t \in [0,T]}$, defined on the event $\{N_T \geq k\}$, is a right-continuous partition-valued stochastic process characterising the entire genealogical tree of $U_1, \ldots, U_k$.

Our goal is to describe the law of $(\pi_t^{k,L,T})_{t \in [0,T]}$ conditioned on the event $\{N_T \geq k\}$, with a view towards the asymptotic regime $T \to \infty$. We find that as $T \to \infty$, there are marked differences in the qualitative behaviour of $(\pi_t^{k,L,T})_{t \in [0,T]}$ depending on the mean number of offspring

$$m := f'(1).$$

---

*University of Graz, Austria. E-mail: sggjohnston@gmail.com

Before we state our results in full generality in Section 3, we give an impression of the structure we expect to encounter by exploring the special case $k = 2$, which features as the focus of a chapter in the recent book [3], and on which the majority of the related literature concentrates.

2 The case $k = 2$

The case $k = 2$ amounts to choosing two particles uniformly from those alive at a time T from a tree with offspring distributed like L, and studying the time $\tau^{L,T}$ in $[0, T]$ at which they last shared a common ancestor. In terms of the partition process $(\pi_t^{2,L,T})_{t \in [0,T]}$, $\tau^{L,T}$ is the time at which the single block $\{\{1, 2\}\}$ splits into the pair of singletons $\{\{1\}, \{2\}\}$.

The following characterisation of the law of $\tau^{L,T}$ (which we will generalise later) was first given by Lambert [14].

Lemma 2.1 (Lambert [14], Corollary 1). On $\{N_T \geq 2\}$, pick two distinct particles uniformly from those alive at time T. Let $\tau^{L,T} \in [0, T]$ be the time at which they last shared a common ancestor. Then

$$\mathbb{P}\left(\tau^{L,T} \in [t, T], N_T \geq 2\right) = \int_0^1 (1 - s) \frac{F''_{T-t}(s)}{F'_{T-t}(s)} F'_T(s) ds, \tag{2.1}$$

where $F_t(s) = \mathbb{E}[s^{N_t}]$.

Although Lambert's result gives a powerful implicit characterisation of the distribution of $\tau^{L,T}$, it is difficult to infer qualitative properties of this random variable directly from (2.1). When $T \to \infty$ however, it is possible to gain a more intuitive insight. Unsurprisingly, different qualitative behaviours arise depending on whether the underlying Galton-Watson tree is supercritical, critical, or subcritical. These cases correspond to $m > 1$, $m = 1$, and $m < 1$ respectively (where $m = f'(1)$).

In the remainder of this section we will exploit classical limit theory of Galton-Watson trees in conjunction with (2.1) to show that conditioned on $\{N_T \geq 2\}$, we have the following limiting behaviour in $\tau^{L,T}$ as $T \to \infty$:

- When the tree is supercritical, $\tau^{L,T}$ remains near the beginning of the interval $[0, T]$. That is, we have the convergence in distribution

$$\tau^{L,T} \xrightarrow{D} \bar{\tau}^L$$

  as $T \to \infty$, where $\bar{\tau}^L$ is a $[0, \infty)$-valued random variable depending on the law of L.

- When the tree is critical, $\tau^{L,T}$ grows linearly in T. That is, we have the convergence in distribution

$$\tau^{L,T}/T \xrightarrow{D} \bar{\tau}^{\text{crit}}$$

  as $T \to \infty$, where $\bar{\tau}^{\text{crit}}$ is a $[0, 1]$-valued random variable universal in all critical offspring distributions with finite variance.

- When the tree is subcritical, $\tau^{L,T}$ remains near the end of the interval $[0, T]$. That is, we have the convergence in distribution

$$T - \tau^{L,T} \xrightarrow{D} \bar{\nu}^L$$

  as $T \to \infty$, where $\bar{\nu}^L$ is a $[0, \infty)$-valued random variable depending on the law of L.

In all three cases we are able to obtain integral formulas for the law of the limit variables.

2.1 The case $k = 2$, supercritical

First we consider the supercritical case $m > 1$. Bühler [5] first observed that when two particles are chosen uniformly from a supercritical tree at a large time, their most recent common ancestor was a member of one of the first generations in tree. More recently, Athreya [1] showed that conditioned on $\{N_T \geq 2\}$, $\tau^{L,T}$ converges in distribution to a $[0, \infty)$-valued random variable $\bar{\tau}^L$ as $T \to \infty$.

Without too much concern at this stage for technical details, we now outline how it is possible to use limit theory for supercritical trees in conjunction with Lambert's formula (2.1) to obtain a formula for the law of the limit variable $\bar{\tau}^L$.

When the tree is supercritical and the *Kesten-Stigum condition* $\mathbb{E}[L \log_+ L] < \infty$ holds, the non-negative and unit-mean martingale $W_t := N_t e^{-(m-1)t}$ converges to a well behaved limit W_∞ [12]. This suggests that at a large time T, the population size is of order $e^{(m-1)T}$, and it would be useful to study the scaling $s = e^{(-ve^{-(m-1)t})}$ in the generating function $F_t(s)$. Indeed, if we let $\varphi(v) := \mathbb{E}[e^{-vW_\infty}]$ denote the Laplace transform of the martingale limit W_∞, in Lemma 6.1 we will show that

$$\lim_{T \to \infty} e^{-k(m-1)T} F_{T-t}^{(k)}(e^{-ve^{-(m-1)T}}) = (-1)^k e^{-k(m-1)t} \varphi^{(k)}(ve^{-(m-1)t}), \quad k \geq 0, \qquad (2.2)$$

where $F_t^{(k)}(s) := \frac{\partial^k}{\partial s^k} F_t(s)$. Assuming for now we can take the limit inside the integral, with the change of variable $s = e^{-ve^{-(m-1)T}}$ in (2.1), using (2.2) in the final equality below we obtain

$$\mathbb{P}(\bar{\tau}^L > t, \text{survival}) := \lim_{T \to \infty} \mathbb{P}\left(\tau^{L,T} \in [t, T], N_T \geq 2\right)$$

$$= \lim_{T \to \infty} \int_0^1 (1 - s) \frac{F_{T-t}''(s)}{F_{T-t}'(s)} F_T'(s) ds$$

$$= \int_0^\infty ve^{-(m-1)t} \frac{\varphi''(ve^{-(m-1)t})}{\varphi'(ve^{-(m-1)t})} \varphi'(v) dv. \qquad (2.3)$$

The formula (2.3) appears to be new, and corresponds to the special case $k = 2$ of our main result for supercritical trees, Theorem 3.5.

2.2 The case $k = 2$, critical

We now move onto the critical case $m = 1$, which has received a lot of attention from different authors [2, 6, 9, 17, 18]. Under the second moment assumption $f''(1) < \infty$, Zubkov [18] found that conditioned on $\{N_T \geq 2\}$, $\tau^{L,T}/T$ converges in distribution to a $[0, 1]$-valued random variable $\bar{\tau}^{\text{Crit}}$ as $T \to \infty$.

Like in the supercritical case, it is possible to use limit theory for critical trees in conjunction with (2.1) to obtain the law of $\bar{\tau}^{\text{Crit}}$. Namely, the Kolmogorov-Yaglom exponential limit law [4, III.7] states that for critical trees with finite variance

$$\lim_{T \to \infty} T\mathbb{P}(N_T > 0) = \frac{1}{c}, \quad \lim_{T \to \infty} \mathbb{P}\left(\frac{N_T}{cT} > x \,\middle|\, N_T > 0\right) = e^{-x}, \qquad (2.4)$$

where $c := f''(1)/2$. In Lemma 6.2, we use the exponential limit law (2.4) to show that for $k \geq 1$ and $a \in (0, 1]$

$$\lim_{T \to \infty} T^{-k+1} F_{aT}^{(k)}\left(e^{-\frac{\theta}{cT}}\right) = (ac)^{k-1} \frac{k!}{(1 + a\theta)^{k+1}}. \qquad (2.5)$$

For $u \in [0, 1]$, set $t = uT$ and take the change of variable $s = \exp\left(-\frac{\theta}{cT}\right)$ in (2.1). Assuming

we can take the limit inside the integral, using (2.5) in the third equality below we obtain

$$\mathbb{P}(\bar{\tau}^{\mathsf{Crit}} \in [u,1]) := \lim_{T \to \infty} \mathbb{P}\left(\tau^{L,T}/T \in [u,1] \mid N_T \geq 2\right)$$

$$= \lim_{T \to \infty} \frac{1}{\mathbb{P}(N_T \geq 2)} \int_0^1 (1-s) \frac{F''_{(1-u)T}(s)}{F'_{(1-u)T}(s)} F'_T(s) ds$$

$$= \int_0^\infty \frac{2(1-u)}{(1+(1-u)\theta)} \frac{\theta}{(1+\theta)^2} d\theta$$

$$= \frac{2(1-u)}{u^2} \left(\log\left(\frac{1}{1-u}\right) - u \right). \tag{2.6}$$

Various formulas for the law of $\bar{\tau}^{\mathsf{Crit}}$ have appeared in the literature. Durrett [6] gave (2.6) in terms of a power series, Athreya [2] gave an expression in terms of sums of exponential random variables, and O'Connell [17] (and more recently, Harris, Johnston and Roberts [9]) obtained (2.6) as written in the more general near-critical setting. We refer the reader to [9, Section 3] for further discussion.

2.3 The case $k = 2$, subcritical

Finally, we look at the subcritical case $m < 1$. On the overwhelmingly rare event that a subcritical tree manages to survive until a large time T, the law of the number of particles alive conditioned on survival converges to a quasi-stationary limit [4, Section III.7]. By this, we mean that there exist non-negative numbers $\{c_j : j \geq 1\}$ satisfying $\sum_{j \geq 1} c_j = 1$ such that

$$\lim_{T \to \infty} \mathbb{P}(N_T = j | N_T > 0) = c_j. \tag{2.7}$$

Lambert showed in [14] (and also Athreya in [2]) that conditioned on $\{N_T \geq 2\}$, the difference $\upsilon^{L,T} := T - \tau^{L,T}$ converges in distribution to a $[0,\infty)$-valued random variable $\bar{\upsilon}^L$ as $T \to \infty$. Lambert also gave an implicit formula for the distribution of the limit variable $\bar{\upsilon}^L$, which Le [15] inverted to obtain

$$\mathbb{P}(\bar{\upsilon}^L < t) = \frac{1}{1-c_1} \int_0^1 (1-s) \frac{F''_t(s)}{F'_t(s)} C'(s) ds, \tag{2.8}$$

where $C(s) := \sum_{j \geq 1} c_j s^j$ is the generating function of the quasi-stationary limit. By replacing t with $T - t$ in (2.1), and using the fact (due to (2.7)) that

$$\lim_{T \to \infty} \frac{F'_T(s)}{\mathbb{P}(N_T \geq 2)} = \lim_{T \to \infty} \frac{\mathbb{P}(N_T \geq 1)}{\mathbb{P}(N_T \geq 2)} \mathbb{E}[N_T s^{N_T - 1} | N_T \geq 1] = \frac{1}{1-c_1} C'(s), \tag{2.9}$$

it is straightforward to sketch a proof of (2.8).

3 Main results

3.1 Overview of results

Let us now give a brief overview of our main results, which will be stated formally in the sequel. Our results for general k run analagously to Section 2 – first we provide integral formulas for the law of $(\pi_t^{k,L,T})_{t \in [0,T]}$ for fixed (finite) T, then we study the $T \to \infty$ asymptotics of these integrals in the supercritical, critical, and subcritical cases.

The fixed-T results, Theorem 3.1, Theorem 3.2, and Theorem 3.4, characterise the law of $(\pi_t^{k,L,T})_{t \in [0,T]}$ in three different ways, first in terms of its finite dimensional distributions, second in terms of its random splitting times, and third as a mixture of

Markov processes. In all cases, explicit formulas are obtained, each in the form of an integral equation involving various generating functions associated with the process.

For instance, the case $n = 1$ of Theorem 3.1 gives the one-dimensional distributions of $(\pi_t^{k,L,T})_{t \in [0,T]}$. Namely, for any partiton γ of $\{1, \ldots, k\}$,

$$\mathbb{P}(\pi_t^{k,L,T} = \gamma, N_T \geq k) = \int_0^1 \frac{(1-s)^{k-1}}{(k-1)!} F_t^{|\gamma|} (F_{T-t}(s)) \prod_{\Gamma \in \gamma} F_{T-t}^{|\Gamma|}(s) ds,$$

where F_t^j denotes the jth derivative of $F_t(s)$ with respect to s.

We are able to understand the combinatorial nature of the products in these integral formulas by relating them to the *Faà di Bruno formula* [10], which states that for k-times differentiable f and g,

$$(f \circ g)^k = \sum_{\gamma \in \Pi^k} \left(f^{|\gamma|} \circ g \right) \prod_{\Gamma \in \gamma} g^{|\Gamma|}, \tag{3.1}$$

where Π^k is the set of partitions of $\{1, \ldots, k\}$, $|\gamma|$ is the number of blocks of a partition γ and $|\Gamma|$ are the block sizes, and h^j denotes the jth derivative of h. It transpires that there is a class of Markov processes whose finite dimensional distributions may be given in terms of a generalisation of the Faà di Bruno for the semigroup $(F_t(s))_{t \geq 0}$. Theorem 3.4 states that $(\pi_t^{k,L,T})_{t \in [0,T]}$ is a random mixture of these processes.

We then send the picking time $T \to \infty$, and study the asymptotic behaviour of the process $(\pi_t^{k,L,T})_{t \in [0,T]}$. As in the case $k = 2$ discussed in Section 2, we will see analogous differences in the asymptotic behaviour of $(\pi_t^{k,L,T})_{t \in [0,T]}$ depending on the mean $m := f'(1)$ of the offspring distribution. Under certain conditions, and in each case conditioned on $\{N_T \geq k\}$, we have the following as $T \to \infty$:

- In the supercritical case $m > 1$, Theorem 3.5 states that we have the distributional convergence

$$(\pi_t^{k,L,T})_{t \in [0,T]} \to (\bar{\pi}_t^{k,L})_{t \in [0,\infty)}, \tag{3.2}$$

  for a limit process $(\bar{\pi}_t^{k,L})_{t \in [0,\infty)}$ depending on the law of L. We will characterise the law of the limit process $(\bar{\pi}_t^{k,L})_{t \in [0,\infty)}$ in terms of integral formulas involving the Laplace transform of the martingale limit.

- In the critical case $m = 1$, Theorem 3.6 states that there exists a universal stochastic process $(\bar{\pi}_t^{k,\text{crit}})_{t \in [0,1]}$ such that

$$(\pi_{tT}^{k,L,T})_{t \in [0,1]} \to (\bar{\pi}_t^{k,\text{crit}})_{t \in [0,1]},$$

  for every critical offspring distribution with finite variance. This result is not new, and was covered in detail (and in greater generality) in Harris, Johnston and Roberts [9].

- In the subcritical case $m < 1$, conditioned on survival until a large time T, the common ancestors of a sample of k particles chosen at T existed near the end of the time interval $[0, T]$. Here it makes more sense to consider $(\rho_t^{k,L,T})_{t \in [0,T]}$ – the right-continuous modification of $(\pi_{T-t}^{k,L,T})_{t \in [0,T]}$. Our subcritical result, Theorem 3.7, states that

$$(\rho_t^{k,L,T})_{t \in [0,T]} \to (\bar{\rho}_t^{k,L})_{t \in [0,\infty)}, \tag{3.3}$$

  for a limit process $(\bar{\rho}_t^{k,L})_{t \in [0,\infty)}$ depending on the law of L. We will characterise the law of the limit process $(\bar{\rho}_t^{k,L})_{t \in [0,\infty)}$ in terms of integral formula involving the generating function of the quasi-stationary limit.

Finally, we consider the relationships between the processes $(\pi_t^{k,L,T})_{t\in[0,T]}$ for different values of k. Theorem 3.8 states that conditioned on the event $\{N_T \geq k+j\}$, the process obtained by projecting $(\pi_t^{k+j,L,T})_{t\in[0,T]}$ onto $\{1,\ldots,k\}$ has the same law as $(\pi_t^{k,L,T})_{t\in[0,T]}$. Corollary 3.9 states that the limiting processes $(\bar{\pi}_t^{k,L})_{t\geq0}$ and $(\bar{\pi}_t^{k,\mathsf{crit}})_{t\in[0,1]}$ appearing in Theorem 3.5 and Theorem 3.6 also satisfy a projectivity property.

3.2 Definitions

Before stating the main results in full, we need to introduce some more notation and definitions. We start by giving a brief formal description of the continuous time Galton-Watson tree. Let L be a $\{0,1,2,\ldots\}$-valued random variable and let $f(s) := \mathbb{E}[s^L]$ be its generating function. Under the probability measure $\mathbb{P}$, we start at time 0 with one particle which we call $\varnothing$. The particle $\varnothing$ lives for a unit-mean and exponentially distributed length of time $\tau_\varnothing$ until it dies, and is replaced by a random number of offspring with labels $1,2,\ldots,L_\varnothing$, where $L_\varnothing$ is distributed like L and is independent of $\tau_\varnothing$. These offspring then independently repeat this behaviour. That is, for each u born at some time, u lives a length of time τ_u distributed like $\tau_\varnothing$ and at death is replaced by offspring with labels $u1, u2, \ldots, uL_u$, where L_u is distributed like L. Here, τ_u and L_u are independent of each other and of the past. We write $\mathcal{N}_t$ for the set of particles alive at time t, $N_t = |\mathcal{N}_t|$ for the number alive at t, and let $F_t(s) := \mathbb{E}[s^{N_t}]$. We remark that $F_t(s)$ enjoys the semigroup property $F_{t_1} \circ F_{t_2} = F_{t_1+t_2}$.

For $u \neq v$, we write $u < v$ if u is an ancestor of v (or equivalently, v is a descendent of u) and $u \leq v$ if $u < v$ or $u = v$. Throughout we will use the terminology *ancestor* and *descendent* weakly, so that u is both an ancestor and a descendent of itself.

A *partition* γ of a non-empty set A is a collection of disjoint non-empty subsets of A, or *blocks*, whose union is A. We write $|\gamma|$ for the number of blocks in γ, and for a block $\Gamma \in \gamma$, we write $|\Gamma|$ for the number of elements in Γ. We write Π^A for the collection of partitions of A, and write $\Pi^k := \Pi^{\{1,\ldots,k\}}$. If B is a non-empty subset of A, and α is a partition of A, we write α^B (or $\alpha|^B$ when there are other superscripts present) for the projection of α onto B:

$$\alpha^B := \{A' \cap B \text{ non-empty} : A' \in \alpha\}.$$

When projecting a partition α onto the set $\{1,\ldots,k\}$, we will write α^k (or $\alpha|^k$) in place of $\alpha^{\{1,\ldots,k\}}$.

For partitions α, β, we say α can break into β, written $\alpha \prec \beta$ (or $\beta \succ \alpha$), if each block of α is a union of blocks in β. For example, $\{\{1,2,4\},\{3\}\} \prec \{\{1\},\{2,4\},\{3\}\}$. An *n-chain* (or just *chain*) of partitions is a sequence of partitions $\boldsymbol{\gamma} = (\gamma_1,\ldots,\gamma_n)$ with the property that $\gamma_i \prec \gamma_{i+1}$ for every i. Let Π_n^A denote the set of n-chains of partitions of A, and for $k \geq 1$ write $\Pi_n^k := \Pi_n^{\{1,\ldots,k\}}$. Using the conventions $\gamma_0 = \{\{1,\ldots,k\}\}$ and $\gamma_{n+1} = \{\{1\},\{2\},\ldots,\{k\}\}$, for each $0 \leq i \leq n$, every block $\Gamma \in \gamma_i$ is the union of $b_i(\Gamma) \geq 1$ blocks of γ_{i+1}. Ordering blocks by their least element, we call the doubly indexed array $(b_i(\Gamma) : 0 \leq i \leq n, \Gamma \in \gamma_i)$ the *fragmentation numbers* associated with the chain $\boldsymbol{\gamma}$, and $(b_i(\Gamma) : \Gamma \in \gamma_i)$ the *fragmentation numbers at the level i*.

We will also use the terminology *chain* for sequence of partitions satisfying $\gamma_i \succ \gamma_{i+1}$. Adopting the convention $\gamma_0 = \{\{1\},\ldots,\{k\}\}$ and $\gamma_{n+1} = \{\{1,\ldots,k\}\}$, for each $0 \leq i \leq n$, each block $\Gamma \in \gamma_i$ is the union of $m_i(\Gamma)$ blocks of γ_{i-1}. We call the array $(m_i(\Gamma) : i \leq n, \Gamma \in \gamma_i)$ the *merger numbers*.

A *mesh* $(t_i)_{i\leq n}$ of a time interval $[0,T]$ is a collection of times $0 < t_1 < \ldots < t_n < T$. Given a mesh $(t_i)_{i\leq n}$, we set $\Delta t_i := t_{i+1} - t_i$ (employing the convention $t_0 = 0, t_{n+1} = T$).

Whenever h is a function, h^j or $h^j(s)$ will refer to the jth-derivative of the function (and $h(s)^j$ for the jth exponent). In particular, we will write $F_t^j(s)$ for the jth-derivative of $F_t(s)$ with respect to s.

Recall that under $\mathbb{P}$ we have a continuous-time Galton-Watson tree starting with one initial particle, branching at rate 1 and with offspring numbers distributed like L. Additionally under $\mathbb{P}$, and on the event $\{N_T \geq k\}$, pick k distinct particles $U_1, \ldots, U_k$ uniformly from those alive at time T. For each time $t \in [0, T]$, we define the equivalence relation $i \sim_t j$ if and only if U_i and U_j share a common ancestor alive at time t. We let $\pi_t^{k,L,T}$ denote the random partition of $\{1, \ldots, k\}$ corresponding to this equivalence relation. The resulting process $(\pi_t^{k,L,T})_{t \in [0,T]}$, defined on the event $\{N_T \geq k\}$, is a right-continuous partition-valued stochastic process satisfying

$$\pi_0^{k,L,T} = \Big\{\{1, 2, \ldots, k\}\Big\}, \quad \pi_T^{k,L,T} = \Big\{\{1\}, \{2\}, \ldots, \{k\}\Big\}.$$

Furthermore, $(\pi_t^{k,L,T})_{t \in [0,T]}$ is a *fragmentation process*, in the sense that blocks break as time passes:

$$t_1 < t_2 \implies \pi_{t_1}^{k,L,T} \prec \pi_{t_2}^{k,L,T}.$$

Let $(\rho_t^{k,L,T})_{t \in [0,T]}$ be the right-continuous modification of $(\pi_{T-t}^{k,L,T})_{t \in [0,T]}$. Then the stochastic process $(\rho_t^{k,L,T})_{t \in [0,T]}$ is a *coalescent process*, in the sense that blocks merge together as time passes:

$$t_1 < t_2 \implies \rho_{t_1}^{k,L,T} \succ \rho_{t_2}^{k,L,T}.$$

We call the discontinuities $\tau_1 < \ldots < \tau_n$ of $(\pi_t^{k,L,T})_{t \in [0,T]}$ *split times*, since these times correspond to a block splitting into two or more blocks.

The event $\{(\pi_t^{k,L,T})_{t \in [0,T]} \text{ is binary}\}$ refers to the event that at every split time, a block splits into exactly two blocks. Note that

$$\{(\pi_t^{k,L,T})_{t \in [0,T]} \text{ is binary}\} = \{(\pi_t^{k,L,T})_{t \in [0,T]} \text{ has } k-1 \text{ split times}\}.$$

When the offspring generating function is of the form $f(s) = \alpha + \gamma s + \beta s^2$, no particle in the tree has more than two offspring upon death, and we call the underlying Galton-Watson tree a *birth-death process*. It follows from that when the underlying tree is a birth-death process

$$\mathbb{P}\left((\pi_t^{k,L,T})_{t \in [0,T]} \text{ is binary} \mid N_T \geq k\right) = 1.$$

Whenever the tree is not a birth-death process, there is a positive probability that before time T some particle in the underlying Galton-Watson tree is replaced by three or more offspring upon death, and hence

$$\mathbb{P}\left((\pi_t^{k,L,T})_{t \in [0,T]} \text{ is binary} \mid N_T \geq k\right) < 1$$

for every $k \geq 3$.

3.3 Hypotheses

We need to ensure that there actually *are* at least k particles alive at time T with positive probability, and that we can choose uniformly from them. To be more precise, we must ensure that both $\mathbb{P}(N_T \geq k) > 0$, and $\mathbb{P}(N_T < \infty) = 1$. The inequality $\mathbb{P}(N_T \geq k) > 0$ is guaranteed to hold by virtue of our first hypothesis, which states that

$$f''(1) > 0. \tag{3.4}$$

In addition to (3.4), we insist that the following non-explosion hypothesis holds:

$$\int_{1-\epsilon}^{1} \frac{ds}{|f(s) - s|} = \infty, \qquad \forall \epsilon \in (0, 1). \tag{3.5}$$

This condition (3.5) is equivalent to our second requirement that $\mathbb{P}(N_t < \infty) = 1$ for t, and holds whenever $f'(1) < \infty$ [8, Chapter II, Theorem 9.1]. We emphasize that both hypotheses (3.4) and (3.5) are in force in the remainder of this paper.

We are now ready to state our main results, which we split into three sections. The results in Section 3.4 concern fixed and finite T. The results in Section 3.5 concern the asymptotic regime in which T is sent to ∞. The results in Section 3.6 concern the projectivity of the partition processes.

3.4 Fixed-T results

The results in this section describe the law of the process $(\pi_t^{k,L,T})_{t \in [0,T]}$ for fixed times T in terms of the generating functions $F_t(s) = \mathbb{E}[s^{N_t}]$ and $f(s) = \mathbb{E}[s^L]$.

Our first fixed-T result, Theorem 3.1, is a generalisation of Lambert's equation (2.1), giving the finite dimensional distributions of the stochastic process $(\pi_t^{k,L,T})_{t \in [0,T]}$.

Theorem 3.1. For any mesh $(t_i)_{i \leq n}$, and any chain of partitions $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_n)$ of $\{1, \ldots, k\}$,

$$\mathbb{P}(\pi_{t_1}^{k,L,T} = \gamma_1, \ldots, \pi_{t_n}^{k,L,T} = \gamma_n, \ N_T \geq k) = \int_0^1 \frac{(1-s)^{k-1}}{(k-1)!} \prod_{i=0}^{n} \prod_{\Gamma \in \gamma_i} F_{\Delta t_i}^{b_i(\Gamma)} \left(F_{T-t_{i+1}}(s) \right) ds, \tag{3.6}$$

where $\Delta t_i = t_{i+1} - t_i$.

Our next result, Theorem 3.2, characterises the law of $(\pi_t^{k,L,T})_{t \in [0,T]}$ in terms of its split times. To this end, let $\boldsymbol{\eta} = (\eta_0, \ldots, \eta_n)$ be a chain of partitions such that $\eta_0 = \{\{1, \ldots, k\}\}$, $\eta_n = \{\{1\}, \ldots, \{k\}\}$. We say $\boldsymbol{\eta}$ is *maximal* if η_i is obtained from η_{i-1} by breaking precisely one block of η_{i-1} into $q_i \geq 2$ blocks in η_i.

Roughly speaking, given a maximal chain $\boldsymbol{\eta} = (\eta_0, \ldots, \eta_n)$, the following theorem gives the joint density of the n times that the process 'jumps' from the value η_{i-1} to η_i, characterising the joint law of the split times of $(\pi_t^{k,L,T})_{t \in [0,T]}$.

Theorem 3.2. Let $\boldsymbol{\eta} = (\eta_0, \ldots, \eta_n)$ be a maximal chain, and let $a_1 < b_1 < a_2 < b_2 < \ldots a_n < b_n$. Then

$$\mathbb{P}(\pi_{a_i}^{k,L,T} = \eta_{i-1}, \pi_{b_i}^{k,L,T} = \eta_i \ \forall \ i = 1, \ldots, n, \ N_T \geq k) \tag{3.7}$$

$$= \int_{a_1}^{b_1} \ldots \int_{a_n}^{b_n} du_1 \ldots du_n \int_0^1 \frac{(1-s)^{k-1}}{(k-1)!} F_T'(s) \prod_{i=1}^{n} f^{q_i}(F_{T-u_i}(s)) F_{T-u_i}'(s)^{q_i-1} ds, \tag{3.8}$$

where $q_i = 1 + |\eta_i| - |\eta_{i-1}|$.

The following lemma is a generalisation of the Faà di Bruno formula (3.1), shedding light on the products occuring in the integral in (3.6).

Lemma 3.3. Let $g_0, \ldots, g_n$ be k-times differentiable. Then

$$(g_0 \circ g_1 \circ \ldots \circ g_n)^k = \sum_{\boldsymbol{\gamma} \in \Pi_n^k} \prod_{i=0}^{n} \prod_{\Gamma \in \gamma_i} g_i^{b_i(\Gamma)} \circ g_{i+1} \circ \ldots \circ g_n. \tag{3.9}$$

In particular, for any semigroup $(F_t)_{t \geq 0}$ of k-times differentiable functions, and any mesh $(t_i)_{i \leq n}$ of $[0, T]$, we have

$$F_T^k(s) = \sum_{\boldsymbol{\gamma} \in \Pi_n^k} \prod_{i=0}^{n} \prod_{\Gamma \in \gamma_i} F_{\Delta t_i}^{b_i(\Gamma)} \left(F_{T-t_{i+1}}(s) \right). \tag{3.10}$$

The identity (3.10) is used to interpret $(\pi_t^{k,L,T})_{t \in [0,T]}$ as a mixture of Markov processes as follows.

Recall that if B is a subset of A and α is a partition of A, α^B is the projection of α onto B. We say a Π^k-valued (time-inhomogeneous) Markov fragmentation process $(\tilde{\pi}_t)_{t \in [0,T]}$ has the *independent blocks property* if given $\{\tilde{\pi}_{t_0} = \gamma\}$, the stochastic processes $\{(\tilde{\pi}_t^\Gamma)_{t \in [t_0,T]} : \Gamma \in \gamma\}$ are conditionally independent. Any such process under a law $\mathbb{P}$ is characterised by the quantities

$$\mathbb{P}(\pi_{t_2}^\Gamma = \delta | \pi_{t_1} = \gamma),$$

where $t_1 < t_2 < T$, γ is a partition of $\{1, \ldots, k\}$, Γ is a block of γ, and δ is a partition of the block Γ.

In Section 5.3 we show using (3.10) that there exists a Markov process $(\tilde{\pi}_t)_{t \in [0,T]}$ starting from $\tilde{\pi}_0 = \{\{1, \ldots, k\}\}$ under a probability law $\mathbb{R}_s^{k,L,T}$ with the independent blocks property and transition density

$$\mathbb{R}_s^{k,L,T}(\tilde{\pi}_{t_2}^\Gamma = \delta | \tilde{\pi}_{t_1} = \gamma) := \frac{F_{t_2-t_1}^{|\delta|}(F_{T-t_2}(s)) \prod_{\Delta \in \delta} F_{T-t_2}^{|\Delta|}(s)}{F_{T-t_1}^{|\Gamma|}(s)}, \quad t_2 \geq t_1. \tag{3.11}$$

Our final fixed-T result, Theorem 3.4, states that the process $(\pi_t^{k,L,T})_{t \in [0,T]}$ can be constructed as a random mixture $m^{k,L,T}$ of processes with laws given by $\{\mathbb{R}_s^{k,L,T} : s \in [0,1]\}$, where the mixture measure is given by

$$m^{k,L,T}(ds) := \frac{(1-s)^{k-1} F_T^k(s)}{(k-1)! \mathbb{P}(N_T \geq k)}. \tag{3.12}$$

Theorem 3.4. The conditional law of $(\pi_t^{k,L,T})_{t \in [0,T]}$ on the event $\{N_T \geq k\}$ is given by

$$\mathbb{P}(\pi_{t_1}^{k,L,T} = \gamma_1, \ldots, \pi_{t_n}^{k,L,T} = \gamma_n \mid N_T \geq k) = \int_0^1 m^{k,L,T}(ds) \mathbb{R}_s^{k,L,T}(\tilde{\pi}_{t_1} = \gamma_1, \ldots, \tilde{\pi}_{t_n} = \gamma_n).$$

In particular, $m^{k,L,T}(ds)$ is a probability measure and $(\pi_t^{k,L,T})_{t \in [0,T]}$ is a mixture of Markov processes with the independent blocks property.

3.5 Asymptotic-T results

We now move on to results concerning the asymptotic behaviour of $(\pi_t^{k,L,T})_{t \in [0,T]}$ as $T \to \infty$. Below we say a collection of partition-valued stochastic processes $\{(\pi_t^T)_{t \geq 0} : T > 0\}$ converge in distribution to a stochastic process $(\bar{\pi}_t)_{t \geq 0}$ as $T \to \infty$ if the finite dimensional distributions converge:

$$\lim_{T \to \infty} \mathbb{P}(\pi_{t_1}^T = \gamma_1, \ldots, \pi_{t_n}^T = \gamma_n) = \mathbb{P}(\bar{\pi}_{t_1} = \gamma_1, \ldots, \bar{\pi}_{t_n} = \gamma_n).$$

First we will consider the supercritical case $m > 1$. Given the stochastic process $(\pi_t^{k,L,T})_{t \in [0,T]}$ defined on $[0,T]$, we define its extension $(\pi_t^{k,L,T})_{t \geq 0}$ to all of $[0,\infty)$ by setting

$$\pi_t^{k,L,T} = \{\{1\}, \ldots, \{k\}\}$$

whenever $t > T$.

Theorem 3.5. Let $m > 1$ and $\mathbb{E}[L \log_+ L] < \infty$. Then as $T \to \infty$, conditioned on $\{N_T \geq k\}$ the process $(\pi_t^{k,L,T})_{t \geq 0}$ converges in distribution to a stochastic process

$(\bar{\pi}_t^{k,L})_{t \geq 0}$ with finite dimensional distributions given by

$$
\mathbb{P}(\bar{\pi}_{t_1}^{k,L} = \gamma_1, \ldots, \bar{\pi}_{t_n}^{k,L} = \gamma_n)
$$
$$
= \frac{(-1)^k e^{-k(m-1)t_n}}{1 - \varphi(\infty)} \int_0^\infty \frac{v^{k-1}}{(k-1)!} \prod_{i=0}^{n-1} \prod_{\Gamma \in \gamma_i} F_{\Delta t_i}^{b_i(\Gamma)} \left(\varphi(e^{-(m-1)t_{i+1}}v) \right) \prod_{\Gamma \in \gamma_n} \varphi^{|\Gamma|}(e^{-(m-1)t_n}v) dv,
$$
(3.13)

where $\varphi(v) := \mathbb{E}[e^{-vW_\infty}]$ is the Laplace transform of the martingale limit $W_\infty := \lim_{T \to \infty} N_T e^{-(m-1)T}$.

Next we consider the critical case $m = 1$.

Theorem 3.6. There exists a universal stochastic process $(\bar{\pi}_t^{k,\mathsf{crit}})_{t \in [0,1]}$ such that for any tree with $m = 1$ and $f''(1) < \infty$, the process $(\pi_{Tt}^{k,L,T})_{t \in [0,1]}$ conditioned on $\{N_T \geq k\}$ converges in distribution to $(\bar{\pi}_t^{k,\mathsf{crit}})_{t \in [0,1]}$ as $T \to \infty$. Moreover, the finite dimensional distributions of $(\bar{\pi}_t^{k,\mathsf{crit}})_{t \in [0,1]}$ are given by

$$
\mathbb{P}(\bar{\pi}_{t_1}^{k,\mathsf{crit}} = \gamma_1, \ldots, \bar{\pi}_{t_n}^{k,\mathsf{crit}} = \gamma_n)
$$
$$
= \prod_{i=0}^n \prod_{\Gamma \in \gamma_i} b_i(\Gamma)! \int_0^\infty \frac{\theta^{k-1}}{(k-1)!} \prod_{i=0}^n (\Delta t_i)^{|\gamma_{i+1}| - |\gamma_i|} \left(\frac{1 + (1 - t_{i+1})\theta}{1 + (1 - t_i)\theta} \right)^{|\gamma_{i+1}|} d\theta.
$$
(3.14)

Theorem 3.6 has already appeared (albeit from a different perspective) in Harris, Johnston and Roberts [9], who show that the genealogical tree corresponding to $(\bar{\pi}_t^{k,\mathsf{crit}})_{t \in [0,1]}$ is binary, and the $k - 1$ split times of $(\bar{\pi}_t^{k,\mathsf{crit}})_{t \in [0,1]}$ have joint probability density function

$$
P(u_1, \ldots, u_{k-1}) = k \int_0^\infty \frac{\theta^{k-1}}{(1+\theta)^2} \prod_{i=1}^{k-1} \frac{1}{(1 + \theta(1 - u_i))^2} d\theta.
$$
(3.15)

It is possible to derive (3.15) from our formula (3.14) by letting $t_{2i-1} = u_i, t_{2i} = u_i + h_i$ for $i = 1, \ldots, k-1$, and sending every $h_i \downarrow 0$. The resulting discrepancy by a factor $\frac{k!(k-1)!}{2^k}$ is a matter of counting tree topologies: there are $\frac{k!(k-1)!}{2^k}$ ranked binary trees with k labelled leaves and $(k-1)$ ranked internal nodes [16].

Finally, we look at the subcritical case $m < 1$. Recall that $(\rho_t^{k,L,T})_{t \in [0,T]}$ is the right-continuous modification of $(\pi_{T-t}^{k,L,T})_{t \in [0,T]}$. We define the extension $(\rho_t^{k,L,T})_{t \geq 0}$ to all of $[0, \infty)$ by setting

$$
\rho_t^{k,L,T} := \{\{1, \ldots, k\}\}
$$

whenever $t > T$.

Theorem 3.7. Let $m < 1$ and $\mathbb{E}[L \log_+ L] < \infty$. Then as $T \to \infty$, conditioned on $\{N_T \geq k\}$, the process $(\rho_t^{k,L,T})_{t \geq 0}$ converges in distribution to a stochastic process $(\bar{\rho}_t^{k,L})_{t \geq 0}$ with finite dimensional distributions given by

$$
\mathbb{P}(\bar{\rho}_{t_1}^{k,L} = \gamma_1, \ldots, \bar{\rho}_{t_n}^{k,L} = \gamma_n)
$$
$$
= \frac{e^{-(m-1)t_n}}{1 - \sum_{j=1}^{k-1} c_j} \int_0^1 \frac{(1-s)^{k-1}}{(k-1)!} C^{|\gamma_n|}(F_{t_n}(s)) \prod_{i=1}^n \prod_{\Gamma \in \gamma_i} F_{\Delta t_{j-1}}^{m_j(\Gamma)}(F_{t_{j-1}}(s)) ds,
$$
(3.16)

where $c_j := \lim_{T \to \infty} \mathbb{P}(N_T = j | N_T > 0)$ and $C(s) := \sum_{j \geq 1} c_j s^j$.

3.6 Projectivity

Let k and j be positive integers. Recall that for $\gamma \in \Pi^{k+j}$, $\gamma|^k \in \Pi^k$ is the projection of γ onto $\{1, \ldots, k\}$. Writing $(\pi_t^k)_{t \in [0,T]} := (\pi_t^{k,L,T})_{t \in [0,T]}$ for the remainder of this section, it is natural to expect from the definition of $(\pi_t^k)_{t \in [0,T]}$ that the projected process $\left(\pi_t^{k+j}\big|^k\right)_{t \in [0,T]}$ is closely related to $(\pi_t^k)_{t \in [0,T]}$ – a property we call *projectivity*. The following generalisation of Theorem 3.1 clarifies this connection.

Theorem 3.8. On the event $\{N_T \geq k + j\}$, the processes $\left(\pi_t^{k+j}\big|^k\right)_{t \in [0,T]}$ and $(\pi_t^k)_{t \in [0,T]}$ are identical in law, and have finite dimensional distributions given by

$$
\mathbb{P}(\pi_{t_1}^k = \gamma_1, \ldots, \pi_{t_n}^k = \gamma_n, \ N_T \geq k + j)
$$
$$
= \int_0^1 \frac{(1-s)^{k+j-1}}{(k+j-1)!} \frac{\partial^j}{\partial s^j} \left(\prod_{i=0}^{n} \prod_{\Gamma \in \gamma_i} F_{\Delta t_i}^{b_i(\Gamma)} \left(F_{T-t_{i+1}}(s) \right) \right) ds. \tag{3.17}
$$

Theorem 3.8 has two immediate corollaries. The first of these, Corollary 3.9, states that when the underlying Galton-Watson tree is either supercritical or critical, the discrepancy in the conditioning disappears in the limit.

Corollary 3.9. Let $(\bar{\pi}_t^{k,L})_{t \geq 0}$ be defined as in Theorem 3.5, and let $(\pi_t^{k,\mathsf{crit}})_{t \in [0,1]}$ be defined as in Theorem 3.6. Then the processes $\left(\bar{\pi}_t^{k+j,L}\big|^k\right)_{t \geq 0}$ and $(\bar{\pi}_t^{k,L})_{t \geq 0}$ are identical in law, and the processes $\left(\bar{\pi}_t^{k+j,\mathsf{crit}}\big|^k\right)_{t \in [0,1]}$ and $(\bar{\pi}_t^{k,\mathsf{crit}})_{t \in [0,1]}$ are identical in law.

Proof. In both the supercritical and critical cases, the proof is a consequence of Theorem 3.8 and the fact that

$$
\lim_{T \to \infty} \mathbb{P}(N_T \geq k + j \mid N_T \geq k) = 1. \tag{3.18}
$$

See [4, Chapter III] for details on (3.18). $\qquad\square$

We emphasise that no exact analogue of Corollary 3.9 holds in the subcritical case, since due to (2.7) we have

$$
\lim_{T \to \infty} \mathbb{P}(N_T \geq k + j \mid N_T \geq k) = \frac{\sum_{i \geq k+j} c_j}{\sum_{i \geq k} c_j} < 1.
$$

The second corollary of Theorem 3.8 gives the finite dimensional distributions of $(\pi_t^k)_{t \in [0,T]}$ on the event $\{N_T = k + j\}$.

Corollary 3.10. For any $j \geq 0$,

$$
\mathbb{P}(\pi_{t_1}^k = \gamma_1, \ldots, \pi_{t_n}^k = \gamma_n, \ N_T = k + j) = \frac{1}{(k+j)!} \frac{\partial^j}{\partial s^j} \left\{ \prod_{i=0}^{n} \prod_{\Gamma \in \gamma_i} F_{\Delta t_i}^{b_i(\Gamma)} (F_{T-t_{i+1}}(s)) \right\} \Bigg|_{s=0}.
$$

Proof. Let $S(k,j)$ denote the integral in (3.17). Then

$$
\mathbb{P}(\pi_{t_1}^k = \gamma_1, \ldots, \pi_{t_n}^k = \gamma_n, \ N_T = k + j) = S(k,j) - S(k, j+1).
$$

Now integrate by parts. $\qquad\square$

3.7 Further discussion of related literature

In [7], Grosjean and Huillet examined the genealogical structure of discrete-time Galton-Watson trees, providing the following extension of Lambert's equation (2.1). On the event $\{N_T \geq k\}$, let $\tau^{k,L,T}$ be the time at which k distinct particles chosen uniformly at time T last shared a common ancestor. Then [7, Proposition 2.2] states that

$$\mathbb{P}(\tau^{k,L,T} > t, N_T \geq k) = \int_0^1 \frac{(1-s)^{k-1}}{(k-1)!} \frac{F_T'(s)}{F_{T-t}'(s)} F_{T-t}^k(s) ds. \tag{3.19}$$

Note by the semigroup property $F_{t_1}(F_{t_2}(s)) = F_{t_1+t_2}(s)$ that $F_t'(F_{T-t}(s)) = \frac{F_T'(s)}{F_{T-t}'(s)}$, and also that by definition $\{\tau^{k,L,T} > t\} = \{\pi_t^{k,L,T} = \{\{1, \ldots, k\}\}\}$. Combining these two facts, we see that (3.19) corresponds to the special case $n = 1$, $\gamma_1 = \{\{1, \ldots, k\}\}$ of Theorem 3.1.

In [15], Le studied the coalescent structure of continuous-time Galton-Watson trees starting with $x \geq 1$ individuals. (We remark that when $x > 1$, the random initial partition $\pi_0^{k,L,T}$ of $(\pi_t^{k,L,T})_{t \in [0,T]}$ may have more than one block.) Le gave an implicit representation for the split times $\tau_1 < \ldots < \tau_{k-1}$ of $(\pi_t^{k,L,T})_{t \in [0,T]}$ on the joint event

$$A := \left\{ \pi_0^{k,L,T} = \{\{1, \ldots, k\}\}, \ (\pi_t^{k,L,T})_{t \in [0,T]} \text{ is binary} \right\}.$$

Namely, in a tree starting with $x \geq 1$ individuals, [15, Theorem 4.2] states that

$$\mathbb{E}_x[N_T^{(k)} s^{N_T - k}, \ \tau_i \in dt_i \ \forall \ i, A, N_T \geq k]$$
$$= \frac{k!(k-1)!}{2^{k-1}} x F_T'(s) F_T(s)^{x-1} \prod_{i=1}^{k-1} F_{T-t_i}'(s) f''(F_{T-t_i}(s)) dt_i, \tag{3.20}$$

for any $0 < t_1 < \ldots < t_{k-1} < T$, where $n^{(k)} := n(n-1) \ldots (n-k+1)$.

We can relate (3.20) to a special case of Theorem 3.2 by calling upon an inversion formula we prove below, Lemma 4.6, which states that when N is a $\{0, 1, \ldots\}$-valued random variable and X is a non-negative random variable on some probability space, then

$$\int_0^1 \frac{(1-s)^{k-1} \mathbb{E}[N^{(k)} s^{N-k} X]}{(k-1)!} ds = \mathbb{E}[X \mathbb{1}_{N \geq k}]. \tag{3.21}$$

Without too much concern for technicalities surrounding whether $X := \mathbb{1}\{\tau_i \in dt_i \ \forall \ i, A\}$ constitutes a well-defined random variable, by applying (3.21) to Le's formula (3.20) we obtain

$$\mathbb{P}_x(\tau_i \in dt_i \ \forall \ i, A, N_T \geq k)$$
$$= \frac{k!}{2^{k-1}} \int_0^1 (1-s)^{k-1} x F_T'(s) F_T(s)^{x-1} \prod_{i=1}^{k-1} F_{T-t_i}'(s) f''(F_{T-t_i}(s)) ds. \tag{3.22}$$

Setting $x = 1$ in (3.22), and accounting for a topological factor of $\frac{k!(k-1)!}{2^{k-1}}$ (counting the number of binary trees with k labelled leaves and $k-1$ ranked internal nodes [16]), (3.22) corresponds the special case of Theorem 3.2 obtained by setting

$$(q_1, \ldots, q_n) = (\underbrace{2, 2, \ldots, 2}_{k-1 \text{ times}}).$$

Finally, let us discuss Harris, Johnston and Roberts [9], who looked at the genealogical structure of Galton-Watson trees in two cases. First they considered the genealogy of

birth-death processes for fixed times T, and thereafter they studied the genealogy of trees under the near-critical scaling limit

$$f'(1) = 1 + \mu/T + o(1/T), \quad f''(1) = \sigma^2 + o(1), \quad T \to \infty. \tag{3.23}$$

In the birth-death case where $f(s) = \alpha + \beta s^2$, the process $(\pi_t^{k,L,T})_{t\in[0,T]}$ is binary, and [9, Proposition 20] states that conditioned on $\{N_T \geq k\}$, the joint density of the $k-1$ split times of $(\pi_t^{k,L,T})_{t\in[0,T]}$ is given by

$$P(t_1, \ldots, t_{k-1})$$
$$= \frac{k!(\beta e^{(\beta-\alpha)T} - \alpha)^k (\beta - \alpha)^{2k-1}}{(e^{(\beta-\alpha)T} - 1)^{k-1} e^{(\beta-\alpha)T}} \int_0^1 (1-s)^{k-1} \prod_{j=0}^{k-1} \frac{e^{(\beta-\alpha)(T-t_j)}}{(\beta(1-s)e^{(\beta-\alpha)(T-t_j)} + \beta s - \alpha)^2} ds,$$
$$\tag{3.24}$$

whenever $\alpha \neq \beta$ (with a similar formula holding when $\alpha = \beta$). It is possible to obtain (3.24) from Theorem 3.2 of the present paper by setting $n = k-1$, $(c_1, \ldots, c_n) = (2, \ldots, 2)$, and using the fact that

$$F_t(s) = \frac{\alpha(1-s)e^{(\beta-\alpha)t} + \beta s - \alpha}{\beta(1-s)e^{(\beta-\alpha)t} + \beta s - \alpha}, \tag{3.25}$$

where (3.25) can be derived using Kolmogorov's forward equation (see for instance, [4, Chapter III, Section 5]).

As for the near-critical scaling limit (3.23), [9, Theorem 3] states that conditioned on $\{N_T \geq k\}$, the process $(\pi_{Tt}^{k,L,T})_{t\in[0,1]}$ converges in distribution to a binary process $(\pi_t^{k,\mathsf{crit},\mu})_{t\in[0,1]}$. (We have translated this result into our notation – the process we study in Theorem 3.6 of the present paper corresponds to the special case $\mu = 0$.) Moreover, according to [9, Section 2.3], the $k-1$ split times of the process $(\pi_t^{k,\mathsf{crit},\mu})_{t\in[0,1]}$ have probability density function

$$f_k(t_1, \ldots, t_{k-1}) = \begin{cases} k(r\mu)^{k-1}(1 - e^{-r\mu}) \displaystyle\int_0^\infty \theta^{k-1} \prod_{i=0}^{k-1} \frac{e^{r\mu(1-t_i)}}{(1 + \theta(e^{r\mu(1-t_i)} - 1))^2} \, d\theta & \text{if } \mu > 0 \\[2em] k \displaystyle\int_0^\infty \theta^{k-1} \prod_{i=0}^{k-1} \frac{1}{(1 + \theta(1-t_i))^2} \, d\theta & \text{if } \mu = 0 \\[2em] k(-1)^k(r\mu)^{k-1}(1 - e^{-r\mu}) \displaystyle\int_0^\infty \theta^{k-1} \prod_{i=0}^{k-1} \frac{e^{r\mu(1-t_i)}}{(1 - \theta(e^{r\mu(1-t_i)} - 1))^2} \, d\theta & \text{if } \mu < 0, \end{cases}$$

with the convention $t_0 = 0$. Harris et al also look at the topology of $(\pi_t^{k,\mathsf{crit},\mu})_{t\in[0,1]}$, showing the tree drawn out is topologically equivalent to *Kingman's coalescent* [13].

The results in [9] are obtained using *multiple spines*, a collection of stochastic processes

$$\left((\xi_t^1)_{t\geq 0}, \ldots, (\xi_t^k)_{t\geq 0}\right)$$

that 'flow' through the tree forward in time, and introduced a change of measure $\mathbb{Q}^{k,T}$ on these spines that biases their behaviour such a way that the spines $(\xi_T^1, \ldots, \xi_T^k)$ represent a uniform sample of k distinct particles at time T.

When the underlying Galton-Watson tree is a birth-death process (respectively, a near-critical tree), the subtree traced out by the spines under this change of measure $\mathbb{Q}^{k,T}$ is binary (respectively, asymptotically binary), and hence the law of this subtree has a tractable expression in terms of its $k-1$ split times. A side effect of the change of measure $\mathbb{Q}^{k,T}$ is that it distorts the law of the underlying Galton-Watson tree, and the

main challenge in [9] was the inversion of this change of measure using a variant of Campbell's formula.

We were inspired in part by [9], and there is some contentual overlap in Section 4 of the present paper and [9, Section 4], which we now discuss. The methodology used in [9] is reliant on the fact that the tree traced out by the spines is binary (or equivalently $(\pi_t^{k,L,T})_{t\in[0,T]}$ is binary), and hence has a law easily expressed in terms of the split times. Though we also use multiple spines in working towards our main result, Theorem 3.1, our method differs in order to encompass non-binary spine trees. Here the spines are indexed by the natural numbers, and the changes of measure are based on a more general class of Radon-Nikodym derivatives associated with partitions α of finite subsets A of $\mathbb{N}$. The resulting changes of measure $\mathbb{Q}^{\alpha,T}$ encourages the spines in the set A to flow through the tree in such away that:

- Spines ξ_T^i and ξ_T^j are following the same particle at time T $\iff$ i and j are in the same block of α.
- The carriers of spines at time T represent a uniform sample of $|\alpha|$ distinct particles at time T.

Moreover, in inverting the changes of measure, rather than using Campbell's formula as in [9], we use a more concise (but arguably less natural) inversion formula – Lemma 4.6 – based around the beta integral. In summary, though our approach in Section 4 to proving Theorem 3.1 is perhaps less intuitive, it works in the more general non-binary setting, and in fewer pages.

Having proved Theorem 3.1, our methods for deriving the remainder of our fixed-T results, built on generalisations of Faà di Bruno's formula, are altogether different from those used in [9]. Our asymptotic-T results are consequences of Theorem 3.1, and though the critical case has already appeared in [9], the results in the supercritical and subcritical cases are new.

3.8 Organisation of the paper

The rest of the paper is structured as follows. In Section 4 we introduce multiple spines and a collection of changes of measure, ultimately leading to a proof of Theorem 3.1 under a moment assumption. In Section 5, we lift this moment assumption, proving the fixed-T results of Section 3.4 in full generality, and we also prove the projectivity result Theorem 3.8. In Section 6, we give proofs of the asymptotic results of Section 3.5.

4 Spines partitions and changes of measure

In this section we introduce spines, our tool for calculating the distributions of genealogical trees associated with uniformly chosen particles. For each $n \in \mathbb{N}$, we associate a line of descent $(\xi_t^n)_{t\geq 0}$ that flows through a continuous-time Galton-Watson tree forward in time, choosing uniformly a branch to follow next at branching points. We call this line of descent $(\xi_t^n)_{t\geq 0}$ the n-spine. The idea of this section is to create a change of measure under which the first k spines $(\xi_t^1, \ldots, \xi_t^k)$ flow through a tree forward in time in such a way that the subtree they trace out is equal in law to $(\pi_t^{k,L,T})_{t\in[0,T]}$.

In order to avoid confusion about different measures, throughout Section 4 we write $\mathbb{P}[\cdot]$ rather than $\mathbb{E}[\cdot]$ for the expectation operator associated with $\mathbb{P}$. Moreover, through this section we will assume that $\mathbb{P}[N_t^k] < \infty$ for all t. In Section 5.2 we show this assumption can be lifted, and the results we derive continue to hold.

Section 4 of the present paper initially runs in parallel to [9, Section 4], in some cases generalising the results that hold there for a specific partition to any partition of a finite subset of $\mathbb{N}$. More specifically, Lemma 4.2 of the present paper is lifted directly from [9,

Lemma 13], and the derivation of the following equation, (4.7), replicates the derivation of [9, Equation (10)]. We also mention that the special cases $\alpha = \{\{1\}, \{2\}, \ldots, \{k\}\}$ of both Lemma 4.1 and Lemma 4.3 of the present paper appear as [9, Lemma 6] and [9, Lemma 14] respectively.

4.1 Spines indexed by $\mathbb{N}$

Suppose under a measure $\mathbb{P}$ we have continuous-time Galton-Watson tree with offspring distribution L. Recall we write $\mathcal{N}_t$ for the set of particles alive at time t, and $N_t = |\mathcal{N}_t|$. For technical reasons, we append a cemetery state Δ to the statespace, and write $\hat{\mathcal{N}}_t = \mathcal{N}_t \cup \Delta$.

Additionally under $\mathbb{P}$, for each $n \in \mathbb{N}$, there is a right-continuous stochastic process $(\xi_t^n)_{t \geq 0}$ called the n-*spine* defined as follows.

- At each time $t \geq 0$, the n-spine takes values in $\hat{\mathcal{N}}_t$ – that is $\xi_t^n \in \hat{\mathcal{N}}_t$. If $u \in \mathcal{N}_t$ and $\xi_t^n = u$, we say that the n-spine is *following* u, and that u is *carrying* the n-spine.
- If a particle carrying the n-spine just before time t dies at time t and is replaced by $p \geq 1$ particles $v_1, \ldots, v_p$, then the n-spine chooses uniformly among the p offspring a particle to follow next. If the particle carrying the n-spine dies at a time t and is replaced by no offspring, we send the n-spine to the cemetery state Δ for the remainder of time. That is, $\xi_r^n = \Delta$ for all $r \geq t$.
- The n-spines don't affect the behaviour of the particles they are following. That is, if a particle u is carrying the n-spine at time t, then this particle still branches at rate 1 and has offspring distributed like L.
- The set of n-spines $\{(\xi_t^n)_{t \geq 0} : n \in \mathbb{N}\}$ are independent of one another – that is, if a particle carrying some spines dies and is replaced by p offspring, each of these spines chooses uniformly an offspring to follow next, independently of the others.

So in essence, under $\mathbb{P}$ the n-spines are simply a set of labels that flow forward in time through a continuous-time Galton-Watson tree without affecting the law of the underlying tree.
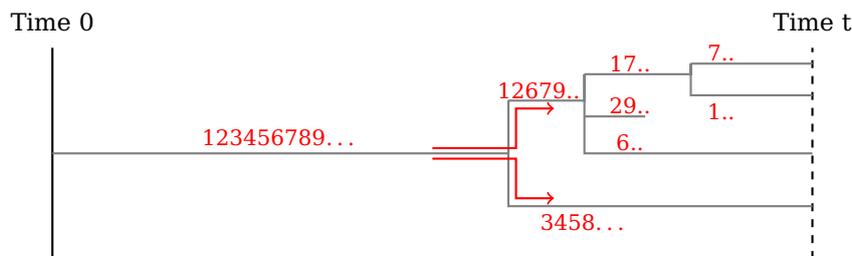

Figure 1: The spines flow through the tree forward in time.

4.2 The spine partition change of measure $\mathbb{Q}^{\alpha, T}$

For any set S and $k \geq 0$, let $S^{(k)}$ be the set of distinct k-tuples from S, and for $n \geq 0$, write

$$n^{(k)} = \begin{cases} n(n-1)(n-2) \ldots (n-k+1) & \text{if } n \geq k \\ 0 & \text{otherwise.} \end{cases}$$

Note that $|S^{(k)}| = |S|^{(k)}$. Let $\mathcal{F}_t^{\varnothing}$ be the σ-algebra containing all the information about the underlying continuous-time Galton-Watson tree up until time t, but without any knowledge of which particles the spines are following. For a subset A of $\mathbb{N}$, we call the

set of processes $\{(\xi_t^A)_{t \geq 0} : a \in A\}$ the A-spines. Let $(\mathcal{F}_t^A)_{t \geq 0}$ be the filtration containing all the information about the underlying tree and the carriers of A-spines until time t:

$$\mathcal{F}_t^A = \sigma\Big(\mathcal{F}_t^\varnothing; (\xi_s^a)_{s \in [0,t]}, a \in A\Big).$$

We note that our notation is consistent with taking A to be the empty set $\varnothing$. Furthermore, if $B \subset A$, then $\mathcal{F}_t^B \subset \mathcal{F}_t^A$ for each t, and in particular, $\mathcal{F}_t^\varnothing \subset \mathcal{F}_t^A$.

We now examine the probabilities, conditional on $\mathcal{F}_T^\varnothing$-knowledge, that a given spine is following a given particle in $\mathcal{N}_T$. For a particle in $u \in \mathcal{N}_T$, let $Q(u)$ be the product of offspring sizes of ancestors of u:

$$Q(u) = \prod_{v < u} L_v.$$

Note that for the a-spine to be following particle $u \in \mathcal{N}_T$, for each strict ancestor v of u, the a-spine must have chosen the 'correct' offspring of the L_v offspring of v to continue following. Hence

$$\mathbb{P}(\xi_T^a = u | \mathcal{F}_T^\varnothing) = Q(u)^{-1}. \tag{4.1}$$

Since the spines behave independently of one another, the probability that the A-spines are following a list $(u_a : a \in A)$ of (possibly non-distinct) members of $\mathcal{N}_t$ is

$$\mathbb{P}\big(\cap_{a \in A} \{\xi_t^a = u_a\}\big|\mathcal{F}_t^\varnothing\big) = \prod_{a \in A} Q(u_a)^{-1}. \tag{4.2}$$

We emphasise that since in general, the quantities $Q(u)^{-1}$ in (4.1) vary for different $u \in \mathcal{N}_T$, under $\mathbb{P}$ the spines are more likely to be following some particles than others.

Define the $(\mathcal{F}_t^A)_{t \geq 0}$-adapted, $(\Pi^A \cup \{\Delta\})$-valued process $(\theta_t^A)_{t \geq 0}$ as follows. If there is an element a of A such that $\xi_t^a = \Delta$, set $\theta_t^A = \Delta$. Otherwise, (that is, if every spine in A is following a living particle at time t) let θ_t^A be the partition of A defined by the equivalence relation

$$a \sim_t b \iff \text{Spines } a \text{ and } b \text{ are following the same particle at time } t, \text{ i.e. } \xi_t^a = \xi_t^b \in \mathcal{N}_t.$$

For the remainder of Section 4, fix a finite subset A of $\mathbb{N}$ and fix a partition α of A into k blocks. Define

$$\hat{\zeta}_{\alpha,t} = \mathbb{1}\{\theta_t^A = \alpha\} \prod_{a \in A} Q(\xi_t^a). \tag{4.3}$$

The following lemma gives the $\mathbb{P}$-conditional expectation of $\hat{\zeta}_{\alpha,t}$ given $\mathcal{F}_t^\varnothing$.

Lemma 4.1. For any $t \geq 0$, $\mathbb{P}[\hat{\zeta}_{\alpha,t}|\mathcal{F}_t^\varnothing] = N_t^{(k)}$.

Proof. Note that if $\alpha = \{A_1, \ldots, A_k\}$, we can decompose the event $\{\theta_t^A = \alpha\}$ into the disjoint union

$$\{\theta_t^A = \alpha\} = \bigcup_{(u_1, \ldots, u_k) \in \mathcal{N}_T^{(k)}} \cap_{i=1}^k \cap_{a \in A_i} \{\xi_t^a = u_i\}.$$

It follows that

$$\mathbb{P}[\hat{\zeta}_{\alpha,t}|\mathcal{F}_t^\varnothing] = \mathbb{P}\bigg[\mathbb{1}\{\theta_t^A = \alpha\} \prod_{a \in A} Q(\xi_t^a)\bigg|\mathcal{F}_t^\varnothing\bigg]$$

$$= \mathbb{P}\bigg[\sum_{u \in \mathcal{N}_t^{(k)}} \prod_{i=1}^k \prod_{a \in A_i} \mathbb{1}\{\xi_t^a = u_i\} Q(\xi_t^a)\bigg|\mathcal{F}_t^\varnothing\bigg]$$

$$= \sum_{u \in \mathcal{N}_t^{(k)}} \mathbb{P}\bigg[\prod_{i=1}^k \prod_{a \in A_i} \mathbb{1}\{\xi_t^a = u_i\} Q(u_i)\bigg|\mathcal{F}_t^\varnothing\bigg]$$

since $\mathcal{N}_t \in \mathcal{F}_t^\varnothing$. Now since the spines are independent,

$$\mathbb{P}\left[\prod_{i=1}^k \prod_{a \in A_i} \mathbb{1}\{\xi_t^a = u_i\}Q(u_i)\Big|\mathcal{F}_t^\varnothing\right] = \prod_{i=1}^k \prod_{a \in A_i} \mathbb{P}\left[\mathbb{1}\{\xi_t^a = u_i\}Q(u_i)\Big|\mathcal{F}_t^\varnothing\right].$$

Now, $Q(u_i) \in \mathcal{F}_t^\varnothing$, so for every a, i,

$$\mathbb{P}\left[\mathbb{1}\{\xi_t^a = u_i\}Q(u_i)\Big|\mathcal{F}_t^\varnothing\right] = Q(u_i)\mathbb{P}(\xi_t^a = u_i|\mathcal{F}_t^\varnothing) = 1$$

by (4.1), and hence

$$\mathbb{P}\left[\prod_{i=1}^k \prod_{a \in A_i} \mathbb{1}\{\xi_t^a = u_i\}Q(u_i)\Big|\mathcal{F}_t^\varnothing\right] = \prod_{i=1}^k \prod_{a \in A_i} 1 = 1. \tag{4.4}$$

Finally,

$$\mathbb{P}[\hat{\zeta}_{\alpha,t}|\mathcal{F}_t^\varnothing] = \sum_{u \in \mathcal{N}_t^{(k)}} 1 = |\mathcal{N}_t^{(k)}| = N_t^{(k)}. \qquad \square$$

Now let T be a fixed time. By the previous lemma $\mathbb{P}[\hat{\zeta}_{\alpha,T}] = \mathbb{P}[\mathbb{P}[\hat{\zeta}_{\alpha,T}|\mathcal{F}_T^\varnothing]] = \mathbb{P}[N_T^{(k)}]$, so the random variable

$$\zeta_{\alpha,T} = \frac{\hat{\zeta}_{\alpha,t}}{\mathbb{P}[N_T^{(k)}]}$$

has unit mean, and we can define a new probability measure $\mathbb{Q}^{\alpha,T}$ on $\mathcal{F}_T^A$ by setting

$$\frac{d\mathbb{Q}^{\alpha,T}}{d\mathbb{P}}\bigg|_{\mathcal{F}_T^A} = \zeta_{\alpha,T}. \tag{4.5}$$

Moreover, by Lemma 4.1 we have

$$\frac{d\mathbb{Q}^{\alpha,T}}{d\mathbb{P}}\bigg|_{\mathcal{F}_T^\varnothing} = \mathbb{P}[\zeta_{\alpha,T}|\mathcal{F}_T^\varnothing] = \frac{N_T^{(k)}}{\mathbb{P}[N_T^{(k)}]} =: Z_{k,T}. \tag{4.6}$$

4.3 Uniformity properties of $\mathbb{Q}^{\alpha,T}$

The goal of this section is to prove that under $\mathbb{Q}^{\alpha,T}$, conditional on $\mathcal{F}_T^\varnothing$, the particles the A-spines are following at the time T are equally likely to be any k-tuple alive. We then exploit this property to relate the spine process $(\theta_t^A)_{t \in [0,T]}$ to the process $(\pi_t^{k,L,T})_{t \in [0,T]}$ associated with choosing k particles uniformly from those alive at time T.

The following result, giving a relationship between projections and changes of measure, is lifted directly from Harris, Johnston and Roberts [9, Lemma 13].

Lemma 4.2. Suppose that $\mathbb{Q}$ and $\mathbb{P}$ are probability measures on the σ-algebra $\tilde{\mathcal{F}}$, and that $\mathcal{F}$ is a sub-σ-algebra of $\tilde{\mathcal{F}}$. If

$$\frac{d\mathbb{Q}}{d\mathbb{P}}\bigg|_{\tilde{\mathcal{F}}} = \zeta \quad \text{and} \quad \frac{d\mathbb{Q}}{d\mathbb{P}}\bigg|_{\mathcal{F}} = Z,$$

then for any non-negative $\tilde{\mathcal{F}}$-measurable X,

$$Z\mathbb{Q}[X|\mathcal{F}] = \mathbb{P}[\zeta X|\mathcal{F}] \quad \mathbb{P}\text{-almost surely.}$$

Proof. For any $S \in \mathcal{F}$,

$$\mathbb{P}[Z\mathbb{Q}[X|\mathcal{F}]S] = \mathbb{Q}[\mathbb{Q}[X|\mathcal{F}]S] = \mathbb{Q}[XS] = \mathbb{P}[\zeta XS].$$

Since $Z\mathbb{Q}[X|\mathcal{F}]$ is $\mathcal{F}$-measurable, it therefore satisfies the definition of conditional expectation of ζX with respect to $\mathcal{F}$ under $\mathbb{P}$. $\qquad \square$

Working in parallel with the derivation of [9, Equation (10)], by applying Lemma 4.2 using (4.5) and (4.6), with $\tilde{\mathcal{F}} = \mathcal{F}_T^A$, and $\mathcal{F} = \mathcal{F}_T^\varnothing$, we find that for any non-negative $\mathcal{F}_T^A$-measurable random variable X, on the event $\{Z_{k,T} > 0\}$,

$$\mathbb{Q}^{\alpha,T}[X|\mathcal{F}_T^\varnothing] = \frac{1}{Z_{k,T}} \mathbb{P}[X\zeta_{\alpha,T}|\mathcal{F}_T^\varnothing]. \tag{4.7}$$

Note $\zeta_{\alpha,T}$ is supported on $\{\theta_T^A = \alpha\}$, and hence $\mathbb{Q}^{\alpha,T}(\theta_T^A = \alpha) = 1$. In particular, since α partitions A into k blocks, under $\mathbb{Q}^{\alpha,T}$ there must be at least k distinct particles alive at time T for the spines to follow, and hence

$$\mathbb{Q}^{\alpha,T}(N_T \geq k) = 1.$$

In summary, $\mathbb{Q}^{\alpha,T}$-almost surely the A-spines at time T are distributed across k different particles in $\mathcal{N}_T$ and induce the partition α of A at time T. The following lemma tells us that given knowledge of the tree but not the spines, under $\mathbb{Q}^{\alpha,T}$ the k-tuple of carriers of A-spines are equally likely to be any k-tuple alive.

Lemma 4.3. The $\mathbb{Q}^{\alpha,T}$-conditional probability given $\mathcal{F}_T^\varnothing$ that the A-spines are following a particular k-tuple $(u_1, \ldots, u_k) \in \mathcal{N}_T^{(k)}$ equals $1/N_T^{(k)}$. That is

$$\mathbb{Q}^{\alpha,T}\left(\cap_{i=1}^k \cap_{a \in A_i} \{\xi_t^a = u_i\} \Big| \mathcal{F}_T^\varnothing \right) = \frac{1}{N_T^{(k)}}.$$

Proof. Note that if $N_T \geq k$ then $Z_{k,T} > 0$. Then by (4.7), for any $u \in \mathcal{N}_T^{(k)}$,

$$\mathbb{Q}^{\alpha,T}\left(\cap_{i=1}^k \cap_{a \in A_i} \{\xi_t^a = u_i\} \Big| \mathcal{F}_T^\varnothing \right) = \frac{1}{Z_{k,T}} \mathbb{P}\left[\zeta_{\alpha,T} \mathbb{1}\{\cap_{i=1}^k \cap_{a \in A_i} \{\xi_t^a = u_i\}\} \Big| \mathcal{F}_T^\varnothing \right]$$

$$= \frac{\mathbb{P}[N_T^{(k)}]}{N_T^{(k)}} \frac{1}{\mathbb{P}[N_T^{(k)}]} \mathbb{P}\left[\prod_{i=1}^k \prod_{a \in A_i} \mathbb{1}\{\xi_t^a = u_i\} Q(u_i) \Big| \mathcal{F}_t^\varnothing \right]$$

$$= \frac{1}{N_T^{(k)}},$$

where the third equality follows from (4.4). □

Given $u = (u_1, \ldots, u_k) \in \mathcal{N}_T^{(k)}$, for each $t \in [0,T]$, let $\pi(u)_t$ be the partition of $\{1, \ldots, k\}$ defined by setting

i and j are in the same block of $\pi(u)_t \iff u_i$ and u_j share a common time-t ancestor.

Let $(\pi_t^{k,L,T})_{t \in [0,T]}$ be the partition process associated with picking k distinct particles $U_1, \ldots, U_k$ uniformly from those alive at time T (as defined in the introduction). It follows from the definition of $(\pi_t^{k,L,T})_{t \in [0,T]}$ that

$$\mathbb{1}_{N_T \geq k} \mathbb{P}(\pi_{t_1}^{k,L,T} = \gamma_1, \ldots, \pi_{t_n}^{k,L,T} = \gamma_n | \mathcal{F}_T^\varnothing)$$

$$= \frac{\mathbb{1}_{N_T \geq k}}{N_T^{(k)}} \sum_{u \in \mathcal{N}_T^{(k)}} \mathbb{1}\left\{ \pi(u)_{t_1} = \gamma_1, \ldots, \pi(u)_{t_n} = \gamma_n \right\}. \tag{4.8}$$

We use the notation $\mathbb{Q}^{k,T}$ for the change of measure $\mathbb{Q}^{\{\{1\},\ldots,\{k\}\},T}$ associated with partitioning $\{1, \ldots, k\}$ into singletons, and similarly we write $\theta_t^k := \theta_t^{\{1,\ldots,k\}}$. The following corollary is the main idea of this section, relating the genealogical process $(\pi_t^{k,L,T})_{t \in [0,T]}$ to the spine process $(\theta_t^k)_{t \in [0,T]}$.

Corollary 4.4. We have the identity

$$\mathbb{1}_{N_T \geq k} \mathbb{P}(\pi_{t_1}^{k,L,T} = \gamma_1, \ldots, \pi_{t_n}^{k,L,T} = \gamma_n | \mathcal{F}_T^{\varnothing}) = \mathbb{1}_{N_T \geq k} \mathbb{Q}^{k,T} \left(\theta_{t_1}^k = \gamma_1, \ldots, \theta_{t_n}^k = \gamma_n | \mathcal{F}_T^{\varnothing} \right).$$

Proof. Let $u = (u_1, \ldots, u_k) \in \mathcal{N}_T^{(k)}$. Note that on the event $\{(\xi_T^1, \ldots, \xi_T^k) = (u_1, \ldots, u_k)\}$ (written $\{\xi_T^{(k)} = u\}$ below) the processes $(\theta_t^k)_{t \in [0,T]}$ and $(\pi(u)_t)_{t \in [0,T]}$ are identical. Using this fact in the second equality below, we have

$$\mathbb{1}_{N_T \geq k} \mathbb{Q}^{k,T} \left(\theta_{t_1}^k = \gamma_1, \ldots, \theta_{t_n}^k = \gamma_n | \mathcal{F}_T^{\varnothing} \right)$$

$$= \mathbb{1}_{N_T \geq k} \mathbb{Q}^{k,T} \left[\sum_{u \in \mathcal{N}_T^{(k)}} \mathbb{1}\{\xi_T^{(k)} = u\} \mathbb{1} \left\{ \theta_{t_1}^k = \gamma_1, \ldots, \theta_{t_n}^k = \gamma_n \right\} \Big| \mathcal{F}_T^{\varnothing} \right]$$

$$= \mathbb{1}_{N_T \geq k} \mathbb{Q}^{k,T} \left[\sum_{u \in \mathcal{N}_T^{(k)}} \mathbb{1}\{\xi_T^{(k)} = u\} \mathbb{1} \left\{ \pi(u)_{t_1} = \gamma_1, \ldots, \pi(u)_{t_n} = \gamma_n \right\} \Big| \mathcal{F}_T^{\varnothing} \right].$$

Using the property that $\mathcal{N}_T$ is $\mathcal{F}_T^{\varnothing}$ is measurable in the first equality below, and Lemma 4.3 in the second, we yield

$$= \mathbb{1}_{N_T \geq k} \sum_{u \in \mathcal{N}_T^{(k)}} \mathbb{1} \left\{ \pi(u)_{t_1} = \gamma_1, \ldots, \pi(u)_{t_n} = \gamma_n \right\} \mathbb{Q}^{k,T} \left[\mathbb{1}\{\xi_T^{(k)} = u\} \Big| \mathcal{F}_T^{\varnothing} \right]$$

$$= \frac{\mathbb{1}_{N_T \geq k}}{N_T^{(k)}} \sum_{u \in \mathcal{N}_T^{(k)}} \mathbb{1} \left\{ \pi(u)_{t_1} = \gamma_1, \ldots, \pi(u)_{t_n} = \gamma_n \right\}.$$

By (4.8), this proves the result. $\qquad \square$

4.4 The joint law of N_T and $(\theta_t^A)_{t \in [0,T]}$ under $\mathbb{Q}^{\alpha,T}$

The following theorem gives the joint law of the process $(\theta_t^A)_{t \in [0,T]}$ and N_T.

Theorem 4.5. Let $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_n)$ be a chain of partitions of A such that $\gamma_n \prec \alpha$. Then

$$\mathbb{Q}^{\alpha,T}(s^{N_T}, \theta_{t_1}^A = \gamma_1, \ldots, \theta_{t_n}^A = \gamma_n) = \frac{s^k}{\mathbb{P}[N_T^{(k)}]} \prod_{i=0}^{n} \prod_{\Gamma \in \gamma_i} F_{\Delta t_i}^{b_i(\Gamma)}(F_{T-t_{i+1}}(s)), \qquad (4.9)$$

where $(b_i(\Gamma) : i = 0, 1, \ldots, n)$ are the fragmentation numbers associated with the partition sequence $(\gamma_0, \gamma_1, \ldots, \gamma_n, \gamma_{n+1})$ obtained by setting $\gamma_0 := \{A\}$ and $\gamma_{n+1} := \alpha$.

Proof. We proceed by induction. The case $n = 0$ follows immediately from (4.6), since

$$\mathbb{Q}^{\alpha,T}(s^{N_T}) = \mathbb{P} \left[\frac{N_T^{(k)}}{\mathbb{P}[N_T^{(k)}]} s^{N_T} \right] = \frac{s^k}{\mathbb{P}[N_T^{(k)}]} F_T^k(s).$$

Now we consider the general case $n \geq 0$. Using the definition (4.5) in the first equality below and the tower property in the second,

$$\mathbb{Q}^{\alpha,T}(s^{N_T}, \theta_{t_1}^A = \gamma_1, \ldots, \theta_{t_n}^A = \gamma_n)$$

$$= \frac{1}{\mathbb{P}[N_T^{(k)}]} \mathbb{P} \left[\mathbb{1}\{\theta_T^A = \alpha\} \prod_{a \in A} Q(\xi_T^a) \, \mathbb{1}\{\theta_{t_1}^A = \gamma_1, \ldots, \theta_{t_n}^A = \gamma_n\} s^{N_T} \right]$$

$$= \frac{1}{\mathbb{P}[N_T^{(k)}]} \mathbb{P} \left[\mathbb{P} \left[\mathbb{1}\{\theta_T^A = \alpha\} \prod_{a \in A} Q(\xi_T^a) \, \mathbb{1}\{\theta_{t_1}^A = \gamma_1, \ldots, \theta_{t_n}^A = \gamma_n\} s^{N_T} \Big| \mathcal{F}_{t_n}^A \right] \right]$$

$$= \frac{1}{\mathbb{P}[N_T^{(k)}]} \mathbb{P} \left[H(\boldsymbol{\gamma}, s)_{t_n} \right],$$

where

$$H(\boldsymbol{\gamma}, s)_{t_n} = \left(\prod_{a \in A} Q(\xi_{t_n}^a)\right) \mathbb{1}\{\theta_{t_1}^A = \gamma_1, \ldots, \theta_{t_n}^A = \gamma_n\} \mathbb{P}\left[\prod_{a \in A} \frac{Q(\xi_T^a)}{Q(\xi_{t_n}^a)} \mathbb{1}\{\theta_T^A = \alpha\} s^{N_T} \Big| \mathcal{F}_{t_n}^A\right].$$

Note that on the event $\{\theta_{t_n}^A = \gamma_n\}$, the set $\{\xi_{t_n}^a : a \in \Gamma\}$ is a singleton for each block Γ of γ_n. In particular, on $\{\theta_{t_n}^A = \gamma_n\}$, we can decompose $\mathcal{N}_T$ into the disjoint union

$$\mathcal{N}_T = \left(\bigcup_{\Gamma \in \gamma_n} \mathcal{N}_{t_n, T}^\Gamma\right) \cup \{u \in \mathcal{N}_T : \nexists a \in A : \xi_{t_n}^a \leq u\},$$

where $\mathcal{N}_{t_n, T}^\Gamma$ is the set of time-T particles descended from the single element of $\{\xi_{t_n}^a : a \in \Gamma\}$ of $\mathcal{N}_{t_n}$. Let $N_{t_n, T}^\Gamma$ be the size of $\mathcal{N}_{t_n, T}^\Gamma$ and let $\hat{N}_{t_n, T}$ be the size of the set $\{u \in \mathcal{N}_T : \nexists a \in A : \xi_{t_n}^a \leq u\}$. Then

$$H(\boldsymbol{\gamma}, s)_{t_n} = \left(\prod_{a \in A} Q(\xi_{t_n}^a)\right) \mathbb{1}\{\theta_{t_1}^A = \gamma_1, \ldots, \theta_{t_n}^A = \gamma_n\} \mathbb{P}\left[\prod_{a \in A} \frac{Q(\xi_T^a)}{Q(\xi_{t_n}^a)} \mathbb{1}\{\theta_T^A = \alpha\} s^{N_T} \Big| \mathcal{F}_{t_n}^A\right]$$

$$= \hat{\zeta}_{\gamma_n, t_n} \mathbb{1}\{\theta_{t_1}^A = \gamma_1, \ldots, \theta_{t_{n-1}}^A = \gamma_{n-1}\}$$

$$\times \mathbb{P}\left[s^{\hat{N}_{t_n, T}} \prod_{\Gamma \in \gamma_n} \prod_{a \in A} \frac{Q(\xi_T^a)}{Q(\xi_{t_n}^a)} \mathbb{1}\{\theta_T^\Gamma = \alpha^\Gamma\} s^{N_{t_n, T}^\Gamma} \Big| \mathcal{F}_{t_n}^A\right] \tag{4.10}$$

Now on the event $\{\theta_{t_n}^A = \gamma_n\}$, since the spines in different blocks of Γ are following different particles from t_n onwards, the random variables

$$\left\{\prod_{a \in A} \frac{Q(\xi_T^a)}{Q(\xi_{t_n}^a)} \mathbb{1}\{\theta_T^\Gamma = \alpha^\Gamma\} s^{N_{t_n, T}^\Gamma} : \Gamma \in \gamma_n\right\}$$

are conditionally independent of each other, of $\hat{N}_{t_n, T}$, and of $\mathcal{F}_{t_n}^A$, and are distributed like copies of $\hat{\zeta}_{\alpha^\Gamma, T-t_n} s^{N_{T-t_n}}$. In particular, using Lemma 4.1 in the second equality below we have

$$\mathbb{P}\left[\prod_{\Gamma \in \gamma_n} \prod_{a \in A} \frac{Q(\xi_T^a)}{Q(\xi_{t_n}^a)} \mathbb{1}\{\theta_T^\Gamma = \alpha^\Gamma\} s^{N_{t_n, T}^\Gamma} \Big| \mathcal{F}_{t_n}^A\right] = \prod_{\Gamma \in \gamma_n} \mathbb{P}\left[\hat{\zeta}_{\alpha^\Gamma, T-t_n} s^{N_{T-t_n}}\right]$$

$$= \prod_{\Gamma \in \gamma_n} \mathbb{P}\left[N_{T-t_n}^{(|\alpha^\Gamma|)} s^{N_{T-t_n}}\right]$$

$$= \prod_{\Gamma \in \gamma_n} s^{|\alpha^\Gamma|} F_{T-t_n}^{|\alpha^\Gamma|}(s). \tag{4.11}$$

Moreover, since on the event $\{\theta_{t_n}^A = \gamma_n\}$ there are $N_{t_n} - |\gamma_n|$ particles not carrying an A-spine at time t_n,

$$\mathbb{P}[s^{\hat{N}_{t_n, T}} | \mathcal{F}_{t_n}^A] = F_{T-t_n}(s)^{N_{t_n} - |\gamma_n|}. \tag{4.12}$$

Finally, writing $\tilde{s} := F_{T-t_n}(s)$, and then using (4.5) in the first equality below and the inductive hypothesis in the second, we have

$$\mathbb{P}\left[\hat{\zeta}_{\gamma_n, t_n} \mathbb{1}\{\theta_{t_1}^A = \gamma_1, \ldots, \theta_{t_{n-1}}^A = \gamma_{n-1}\} \tilde{s}^{N_t - |\gamma_n|}\right]$$

$$= \mathbb{P}[N_{t_n}^{(|\gamma_n|)}] \mathbb{Q}^{\gamma_n, t_n}\left[\mathbb{1}\{\theta_{t_1}^A = \gamma_1, \ldots, \theta_{t_{n-1}}^A = \gamma_{n-1}\} \tilde{s}^{N_t - |\gamma_n|}\right]$$

$$= \prod_{i=0}^{n-1} \prod_{\Gamma \in \gamma_i} F_{\Delta t_i}^{b_i(\Gamma)}(F_{t_n - t_{i+1}}(\tilde{s})). \tag{4.13}$$

Combining (4.11), (4.12) and (4.13), by taking expectations of (4.10) we obtain

$$\mathbb{P}[H(\boldsymbol{\gamma}, s)_{t_n}] = s^{\sum_{\Gamma \in \gamma_n} |\alpha^\Gamma|} \prod_{i=0}^{n-1} \prod_{\Gamma \in \gamma_i} F_{\Delta t_i}^{b_i(\Gamma)}(F_{t_n - t_{i+1}}(\tilde{s})) \prod_{\Gamma \in \gamma_n} F_{T-t_n}^{|\alpha^\Gamma|}(s).$$

Note that $|\alpha^\Gamma| = b_n(\Gamma)$, and in particular, $\sum_{\Gamma \in \gamma_n} |\alpha^\Gamma| = k$. Finally, by using the semigroup property $F_{t_n - t_{i+1}}(\tilde{s}) = F_{T-t_{i+1}}(s)$, (4.9) follows. $\qquad\square$

4.5 Inverting the change of measure

Lemma 4.6. Under a probability measure $\mathbb{P}$, let N be a $\{0, 1, \ldots\}$-valued random variable and let X be a $[0, \infty)$-valued random variable. Then

$$\int_0^1 \frac{(1-s)^{k-1}}{(k-1)!} \mathbb{P}[X \, N^{(k)} s^{N-k}] ds = \mathbb{P}[X \mathbb{1}_{N \geq k}]. \tag{4.14}$$

Proof. Recall from the definition of the beta function that

$$\frac{(x-1)!(y-1)!}{(x+y-1)!} = \int_0^1 s^{x-1}(1-s)^{y-1} ds.$$

It follows that for $n \geq k$ we have the identity

$$\frac{1}{n^{(k)}} = \frac{(n-k)!}{n!} = \frac{1}{(k-1)!} \int_0^1 (1-s)^{k-1} s^{n-k} ds. \tag{4.15}$$

By interchanging the order of expectation and integration we may write

$$\int_0^1 \frac{(1-s)^{k-1}}{(k-1)!} \mathbb{P}[X \, N^{(k)} s^{N-k}] ds = \mathbb{P}\left[X \mathbb{1}_{N \geq k} N^{(k)} \frac{1}{(k-1)!} \int_0^1 (1-s)^{k-1} s^{N-k} ds \right]$$
$$= \mathbb{P}[X \mathbb{1}_{N \geq k}]. \qquad\square$$

We are now ready to wrap things together to prove Theorem 3.1 under the assumption $\mathbb{P}[N_t^{(k)}] < \infty$ for all t. Namely for any partition chain γ_i and mesh $(t_i)_{i \leq n}$ we now show that

$$\mathbb{P}(\pi_{t_1}^{k,L,T} = \gamma_1, \ldots, \pi_{t_n}^{k,L,T} = \gamma_n, \, N_T \geq k) = \int_0^1 \frac{(1-s)^{k-1}}{(k-1)!} \prod_{i=0}^{n} \prod_{\Gamma \in \gamma_i} F_{\Delta t_i}^{b_i(\Gamma)} \left(F_{T-t_{i+1}}(s) \right) ds. \tag{4.16}$$

Proof of Theorem 3.1 under k^{th}-moment assumption. By Lemma 4.6

$$\mathbb{P}(\pi_{t_1}^{k,L,T} = \gamma_1, \ldots, \pi_{t_n}^{k,L,T} = \gamma_n, \, N_T \geq k)$$
$$= \int_0^1 \frac{(1-s)^{k-1}}{(k-1)!} \mathbb{P}\left[\mathbb{1}\left\{ \pi_{t_1}^{k,L,T} = \gamma_1, \ldots, \pi_{t_n}^{k,L,T} = \gamma_n \right\} N_T^{(k)} s^{N_T - k} \right] ds. \tag{4.17}$$

Recall that $\mathbb{Q}^{k,T}$ is the change of measure associated with the partition $\{\{1\}, \ldots, \{k\}\}$. Under the assumption $\mathbb{P}[N_T^{(k)}] < \infty$, using (4.6) in the first equality below and Corollary 4.4 in the second, we have

$$\mathbb{P}\left[\mathbb{1}\left\{ \pi_{t_1}^{k,L,T} = \gamma_1, \ldots, \pi_{t_n}^{k,L,T} = \gamma_n \right\} N_T^{(k)} s^{N_T - k} \right]$$
$$= \mathbb{P}[N_T^{(k)}] \mathbb{Q}^{k,T} \left[s^{N_T - k} \mathbb{P}\left(\pi_{t_1}^{k,L,T} = \gamma_1, \ldots, \pi_{t_n}^{k,L,T} = \gamma_n \Big| \mathcal{F}_T^\varnothing \right) \right]$$
$$= \mathbb{P}[N_T^{(k)}] \mathbb{Q}^{k,T} \left[s^{N_T - k} \mathbb{1}\left\{ \theta_{t_1}^k = \gamma_1, \ldots, \theta_{t_n}^k = \gamma_n \right\} \right]. \tag{4.18}$$

Finally, by Theorem 4.5

$$\mathbb{P}[N_T^{(k)}]\mathbb{Q}^{k,T}\left[s^{N_T-k}\mathbb{1}\left\{\theta_{t_1}^k=\gamma_1,\ldots,\theta_{t_n}^k=\gamma_n\right\}\right]=\prod_{i=0}^{n}\prod_{\Gamma\in\gamma_i}F_{\Delta t_i}^{b_i(\Gamma)}\left(F_{T-t_{i+1}}(s)\right). \quad (4.19)$$

Combining (4.17), (4.18) and (4.19) yields Theorem 3.1 under the assumption $\mathbb{P}[N_t^{(k)}]<\infty$ for all t. $\qquad\square$

5 Proofs of fixed-T results

In the sequel, we will return to writing $\mathbb{E}$ for the expectation associated with the probability measure $\mathbb{P}$. In Section 5.1 we prove a generalisation of the Faà di Bruno formula for semigroups $(F_t)_{t\geq 0}$, which will be used in Section 5.2 to lift the $\mathbb{E}[N_t^{(k)}]<\infty$ assumption from Theorem 3.1.

5.1 Generalisations of the Faà di Bruno formula

In this section we prove several generalisations of the Faà di Bruno formula

$$(f\circ g)^k=\sum_{\gamma\in\Pi^k}\left(f^{|\gamma|}\circ g\right)\prod_{\Gamma\in\gamma}g^{|\Gamma|}. \quad (5.1)$$

Recall that for $\gamma\in\Pi^{k+j}$, $\gamma|^k\in\Pi^k$ is its projection onto $\{1,\ldots,k\}$. Given a chain $\boldsymbol{\gamma}=(\gamma_1,\ldots,\gamma_n)\in\Pi_n^{k+j}$, let $\boldsymbol{\gamma}|^k\in\Pi_n^k$ be the chain defined by projecting γ_i onto $\{1,\ldots,k\}$ for each $i=1,\ldots,n$. That is, $(\boldsymbol{\gamma}|^k)_i:=\gamma_i|^k$ for each i.

Finally, we recall that h^j or $h^j(s)$ will refer to the jth derivative of a function j. (We write $h(s)^j$ for the jth exponent).

Lemma 5.1. For each chain $\boldsymbol{\gamma}=(\gamma_1,\ldots,\gamma_n)\in\Pi_n^k$,

$$\left(\prod_{i=0}^{n}\prod_{\Gamma\in\gamma_i}g_i^{b_i(\Gamma)}\circ g_{i+1}\circ\ldots\circ g_n\right)^j=\sum_{\boldsymbol{\eta}\in\Pi_n^{k+j}:\boldsymbol{\eta}|^k=\boldsymbol{\gamma}}\prod_{i=0}^{n}\prod_{H\in\eta_i}g_i^{b_i(H)}\circ g_{i+1}\circ\ldots\circ g_n. \quad (5.2)$$

Proof. First we prove the case $j=1$. By the Leibniz rule,

$$\left(\prod_{i=0}^{n}\prod_{\Gamma\in\gamma_i}g_i^{b_i(\Gamma)}\circ g_{i+1}\circ\ldots\circ g_n\right)^1$$

$$=\sum_{0\leq i\leq n,\Gamma\in\gamma_i}g_i^{b_i(\Gamma)+1}\circ g_{i+1}\circ\ldots\circ g_n\prod_{l=i+1}^{n}g_l^1\circ g_{l+1}\circ\ldots\circ g_n$$

$$\times\prod_{0\leq p\leq n,\tilde{\Gamma}\in\gamma_p:(p,\tilde{\Gamma})\neq(i,\Gamma)}g_p^{b_p(\tilde{\Gamma})}\circ g_{p+1}\circ\ldots\circ g_n. \quad (5.3)$$

We now define a bijection

$$\eta:\{(i,\Gamma):0\leq i\leq n,\Gamma\in\gamma_i\}\to\{\boldsymbol{\zeta}\in\Pi_n^{k+1}:\boldsymbol{\zeta}|^k=\boldsymbol{\gamma}\}$$

as follows. Let $\eta_j:=\eta(i,\Gamma)_j$ be the jth partition in the chain $\eta(i,\Gamma)$. Then

- for each $j\leq i$, η_j is formed from γ_j by joining $\{k+1\}$ to the block containing Γ.
- for each $j\geq i+1$, η_j is formed from γ_j by adding the singleton set $\{k+1\}$ to γ_j.

The map $(i, \Gamma) \mapsto \eta(i, \Gamma)$ is clearly injective. Furthermore to see that it is surjective, for any chain $\boldsymbol{\zeta} \in \Pi_n^{k+1}$ such that $\boldsymbol{\zeta}|^k = \boldsymbol{\gamma}$, let i be the largest $0 \leq j \leq n$ such that $k+1$ is contained in a non-singleton block $\Gamma \cup \{k+1\}$ of ζ_j. Then $\boldsymbol{\zeta} = \eta(i, \Gamma)$.

Now consider the fragmentation numbers of the partition $\boldsymbol{\eta} = \eta(i, \Gamma)$. The fragmentation numbers of $\boldsymbol{\eta}$ and $\boldsymbol{\gamma}$ are identical at the levels $p = 0, 1, \ldots, i-1$. The fragmentation numbers $(b_i(H) : H \in \eta_i)$ at the level i of $\eta = \eta(i, \Gamma)$ are given by replacing the entry $b_i(\Gamma)$ with $b_i(\Gamma) + 1$. Finally, for $p = i+1, \ldots, n$, the fragmentation numbers $(b_p(H) : H \in \eta_p)$ of $\boldsymbol{\eta}$ at the level p are obtained by adding an entry 1 to the end of the array $(b_p(\Gamma) : \Gamma \in \gamma_p)$.

It follows that

$$
\prod_{j=0}^{n} \prod_{H \in \eta(i,\Gamma)_j} g_j^{b_j(H)} \circ g_{j+1} \circ \ldots \circ g_n
$$

$$
= g_i^{b_i(\Gamma)+1} \circ g_{i+1} \circ \ldots \circ g_n \prod_{l=i+1}^{n} g_l^1 \circ g_{l+1} \circ \ldots \circ g_n
$$

$$
\times \prod_{0 \leq p \leq n, \tilde{\Gamma} \in \gamma_p : (p,\tilde{\Gamma}) \neq (i,\Gamma)} g_p^{b_p(\tilde{\Gamma})} \circ g_{p+1} \circ \ldots \circ g_n. \tag{5.4}
$$

Using (5.4) in the first equality below, and the fact that η is a bijection in the second, we have

$$
\sum_{0 \leq i \leq n, \Gamma \in \gamma_i} g_i^{b_i(\Gamma)+1} \circ g_{i+1} \circ \ldots \circ g_n \prod_{l=i+1}^{n} g_l^1 \circ g_{l+1} \circ \ldots \circ g_n
$$

$$
\times \prod_{0 \leq p \leq n, \tilde{\Gamma} \in \gamma_p : (p,\tilde{\Gamma}) \neq (i,\Gamma)} g_p^{b_p(\tilde{\Gamma})} \circ g_{p+1} \circ \ldots \circ g_n
$$

$$
= \sum_{0 \leq i \leq n, \Gamma \in \gamma_i} \prod_{j=0}^{n} \prod_{H \in \eta(i,\Gamma)_j} g_j^{b_j(H)} \circ g_{j+1} \circ \ldots \circ g_n
$$

$$
= \sum_{\boldsymbol{\zeta} \in \Pi_n^{k+1} : \boldsymbol{\zeta}|^k = \boldsymbol{\gamma}} \prod_{j=0}^{n} \prod_{Z \in \zeta_j} g_j^{b_j(Z)} \circ g_{j+1} \circ \ldots \circ g_n,
$$

proving the result for $j = 1$.

Now we prove the general case $j \geq 1$ by induction. Suppose the result holds for all $j' \leq j$. Then by the inductive hypothesis we have

$$
\left(\prod_{i=0}^{n} \prod_{\Gamma \in \gamma_i} g_i^{b_i(\Gamma)} \circ g_{i+1} \circ \ldots \circ g_n \right)^{j+1} = \sum_{\boldsymbol{\eta} \in \Pi_n^{k+j} : \boldsymbol{\eta}|^k = \boldsymbol{\gamma}} \left(\prod_{i=0}^{n} \prod_{H \in \eta_i} g_i^{b_i(H)} \circ g_{i+1} \circ \ldots \circ g_n \right)^1 . \tag{5.5}
$$

Using the case $j = 1$ for each term in the sum on the right-hand-side of (5.5), we obtain

$$
\left(\prod_{i=0}^{n} \prod_{\Gamma \in \gamma_i} g_i^{b_i(\Gamma)} \circ g_{i+1} \circ \ldots \circ g_n \right)^{j+1}
$$

$$
= \sum_{\boldsymbol{\eta} \in \Pi_n^{k+j} : \boldsymbol{\eta}|^k = \boldsymbol{\gamma}} \sum_{\boldsymbol{\theta} \in \Pi_n^{k+j+1} : \boldsymbol{\theta}|^{k+j} = \boldsymbol{\eta}} \prod_{i=0}^{n} \prod_{\Theta \in \theta_i} g_i^{b_i(\Theta)} \circ g_{i+1} \circ \ldots \circ g_n. \tag{5.6}
$$

Noting we have the disjoint union

$$
\cup_{\boldsymbol{\eta} \in \Pi_n^{k+j} : \boldsymbol{\eta}|^k = \boldsymbol{\gamma}} \{ \boldsymbol{\theta} \in \Pi_n^{k+j+1} : \boldsymbol{\theta}|^{k+j} = \boldsymbol{\eta} \} = \{ \boldsymbol{\theta} \in \Pi_n^{k+j+1} : \boldsymbol{\theta}|^k = \boldsymbol{\gamma} \}, \tag{5.7}
$$

it follows from using (5.7) in (5.6) that

$$
\left(\prod_{i=0}^{n} \prod_{\Gamma \in \gamma_i} g_i^{b_i(\Gamma)} \circ g_{i+1} \circ \ldots \circ g_n \right)^{j+1} = \sum_{\boldsymbol{\theta} \in \Pi_n^{k+j+1} : \boldsymbol{\theta}|^k = \boldsymbol{\gamma}} \prod_{i=0}^{n} \prod_{\Theta \in \theta_i} g_i^{b_i(\Theta)} \circ g_{i+1} \circ \ldots \circ g_n.
$$

This proves the result holds for $j + 1$, and hence by induction, (5.2) holds for every $j \geq 1$. $\qquad\square$

Lemma 5.2. Let $g_0, \ldots, g_n$ be k-times differentiable. Then

$$
(g_0 \circ g_1 \circ \ldots \circ g_n)^k = \sum_{\boldsymbol{\gamma} \in \Pi_n^k} \prod_{i=0}^{n} \prod_{\Gamma \in \gamma_i} g_i^{b_i(\Gamma)} \circ g_{i+1} \circ \ldots \circ g_n. \tag{5.8}
$$

Furthermore, for any semigroup $(F_t)_{t \geq 0}$ of k-times differentiable functions, and any mesh $(t_i)_{i \leq n}$ of $[0, T]$, we have

$$
F_T^k = \sum_{\boldsymbol{\gamma} \in \Pi_n^k} \prod_{i=0}^{n} \prod_{\Gamma \in \gamma_i} F_{\Delta t_i}^{b_i(\Gamma)} \circ F_{T - t_{i+1}}. \tag{5.9}
$$

In particular, for every $s \in [0, 1]$ and every chain $\boldsymbol{\gamma}$ in Π_n^k

$$
F_T^k(s) \geq \prod_{i=0}^{n} \prod_{\Gamma \in \gamma_i} F_{\Delta t_i}^{b_i(\Gamma)} \left(F_{T - t_{i+1}}(s) \right) \geq 0. \tag{5.10}
$$

Proof. To see that (5.8) is true for $k = 1$, we note that Π_n^1 contains the single chain $\boldsymbol{\gamma}^* = (\gamma_1, \ldots, \gamma_n)$ where $\gamma_i = \{\{1\}\}$ for each i, the breakage number of the single block $\{1\}$ in each partition is simply 1, and the result holds by comparison with the chain rule. For general k, by using the case $k = 1$ to obtain the first equality below and (5.2) to obtain the second, we have

$$
\begin{aligned}
(g_0 \circ \ldots \circ g_n)^k &= \left(\prod_{i=0}^{n} \prod_{\Gamma \in \gamma_i^*} g_i^{b_i(\Gamma)} \circ g_{i+1} \circ \ldots \circ g_n \right)^{k-1} \\
&= \sum_{\boldsymbol{\eta} \in \Pi_n^k : \boldsymbol{\eta}|^1 = \boldsymbol{\gamma}^*} \prod_{i=0}^{n} \prod_{H \in \eta_i} g_i^{b_i(H)} \circ g_{i+1} \circ \ldots \circ g_n \\
&= \sum_{\boldsymbol{\eta} \in \Pi_n^k} \prod_{i=0}^{n} \prod_{H \in \eta_i} g_i^{b_i(H)} \circ g_{i+1} \circ \ldots \circ g_n,
\end{aligned}
$$

where the final equality holds from the fact that *every* chain $\boldsymbol{\eta}$ in Π_n^k satisfies $\boldsymbol{\eta}|^1 = \boldsymbol{\gamma}^*$.

Now we prove (5.9). If $0 \leq t_1 \leq \ldots \leq t_n \leq T$ is a mesh of $[0, T]$, and $(F_t)_{t \geq 0}$ is a semigroup, then we obtain (5.9) by setting $g_i := F_{\Delta t_i}$, and noting

$$
g_{i+1} \circ \ldots \circ g_n = F_{\Delta t_{i+1}} \circ \ldots \circ F_{\Delta t_n} = F_{\Delta t_{i+1} + \ldots + \Delta t_n} = F_{T - t_{i+1}}.
$$

Finally, to prove (5.10), note that for every $s \in [0, 1]$, $F_t(s) \in [0, 1]$, and for every $j \geq 1$, $F_t^j(s) \geq 0$. It follows that for all $j \geq 1$, and $s \in [0, 1]$, $F_{t_1}^j(F_{t_2}(s)) \geq 0$ for all $t_1, t_2 \geq 0$. This shows that the summands in (5.9) are non-negative, and (5.10) follows. $\qquad\square$

5.2 Proof of Theorem 3.1 without moment assumption

In this section we prove that we can relax the assumption that $\mathbb{E}[N_t^{(k)}] < \infty$ for each t and Theorem 3.1 continues to hold. In order to use a coupling argument in lifting this condition, first we require the following result.

Given a Galton-Watson tree, we call $(N_t)_{t \geq 0}$ the *process* associated with the tree, where N_t is the number of particles in the tree at time t.

Lemma 5.3. Let $\bar{N}_t = |\bigcup_{s \in [0,t]} \mathcal{N}_s|$ be the number of particles who have ever lived up until time t. Then, provided the non-explosion hypothesis (3.5) holds, $\mathbb{P}(\bar{N}_t < \infty) = 1$.

Proof. Suppose we have a continuous-time Galton-Watson process $N := (N_t)_{t \geq 0}$ with offspring generating function $f(s) = \mathbb{E}[s^L] = p_0 + p_1 s + s^2 g(s)$ satisfying the non-explosion hypothesis. Couple N with another process $M := (M_t)_{t \geq 0}$ with generating function $f^*(s) = (p_0 + p_1 + g(s))s^2$ as follows. Every time an particle in the process N has 0 or 1 children, the corresponding particle in the process M has 2 children. Writing $\bar{M}_t$ for the number who have ever lived until t in the M-process, clearly $\mathbb{P}(\bar{N}_t \leq \bar{M}_t) = 1$, and it is straightforward to verify that $f^*(s)$ also satisfies the non-explosion hypothesis, and hence M_t is almost surely finite.

Consider in the process M that every particle is replaced by at least two particles upon death, and hence there were at most $\frac{1}{2} M_t$ parents of particles alive at time t. A similar argument says that there can have been at most $\frac{1}{4} M_t$ grandparents, and so forth. It follows that the we can bound above the number who have ever lived: $\bar{M}_t \leq \sum_{i \geq 0} 2^{-i} M_t = 2 M_t$.

Since $2 M_t \geq \bar{M}_t \geq \bar{N}_t$, the latter quantity is almost surely finite. $\qquad\square$

The following lemma, a variant of the dominated convergence theorem, will be used in the proofs of Theorem 3.1, Theorem 3.5 and Theorem 3.7.

Lemma 5.4. Let $g, (g_n)$, and $h, (h_n)$ be measurable functions on a probability space $(\Omega, \mathcal{A}, \mu)$, with $|g_n| \leq h_n$ for all n, and such that $g_n \to g$, $h_n \to h$, and $\mu h_n \to \mu h$. Then $\mu g_n \to \mu g$.

Proof. See [11, Theorem 1.21]. $\qquad\square$

We are now ready to prove that Theorem 3.1 holds for every offspring distribution with generating function satisfying hypotheses (3.4) and (3.5).

Proof of Theorem 3.1. Our proof idea as follows. To calculate the distribution of the process $(\pi_t^{k,L,T})_{t \in [0,T]}$, first calculate $(\pi_{n,t}^{k,L,T})_{t \in [0,T]}$ from a tree where offspring sizes are bounded by n. Any such tree clearly satisfies $\mathbb{E}[L^k] < \infty$, and therefore ([4, Section III.6]) $\mathbb{E}[N_t^{(k)}] < \infty$ for every t and the formula (4.16) applies. Then we send $n \to \infty$, showing the formula (4.16) converge suitably.

Let L be a random variable, and let Tree_T be the continuous-time Galton-Watson tree run until time T and $(N_t)_{t \in [0,T]}$ be the corresponding process for the number of particles alive. We couple the tree Tree_T with a tree with bounded branching as follows. Let $\mathsf{Tree}_{n,T}$ be the tree with offspring distribution $L\mathbb{1}_{L \leq n}$ taken by replacing any birth of size greater than n Tree_T with a birth of size zero, and let $(N_{n,t})_{t \in [0,T]}$ be the associated process for the number of particles alive.

By Lemma 5.3, $\mathbb{P}(\bar{N}_t < \infty) = 1$, and thus $\mathbb{P}(\bar{N}_t \leq n) \uparrow 1$ as $n \to \infty$. Note that $\{\bar{N}_t \leq n\}$ ensures $\{\mathsf{Tree}_{n,T} = \mathsf{Tree}_T\}$, since if at most n particles have ever lived, no particle ever had more than n offspring. It follows that $\mathbb{P}(\mathsf{Tree}_{n,T} = \mathsf{Tree}_T) \uparrow 1$ as $n \to \infty$. In particular, $N_{n,t} \to N_t$ almost surely and hence $\mathbb{P}(N_{n,T} \geq k) \to \mathbb{P}(N_T \geq k)$.

If we pick k particles from $\text{Tree}_{n,T}$ and call the partition process $(\pi_{n,t}^{k,L,T})_{t\in[0,T]}$, it follows that $(\pi_{n,t}^{k,L,T})_{t\in[0,T]}$ converges in distribution to $(\pi_t^{k,L,T})_{t\in[0,T]}$ as $n\to\infty$, since the partition processes correspond to subtrees of the trees $\text{Tree}_{n,T}$ and Tree_T respectively.

It remains to check that for a process $(N_{n,t})_{t\geq 0}$ with offspring distributed like $L\mathbb{1}_{L\leq n}$ and generating function $F_{n,t}(s)$, that as $n\uparrow\infty$,

$$\int_0^1 \frac{(1-s)^{k-1}}{(k-1)!\mathbb{P}(N_{n,T}\geq k)}\prod_{i=0}^n\prod_{\Gamma\in\gamma_i}F_{n,\Delta t_i}^{b_i(\Gamma)}\Big(F_{n,T-t_{i+1}}(s)\Big)ds \tag{5.11}$$

$$\to \int_0^1 \frac{(1-s)^{k-1}}{(k-1)!\mathbb{P}(N_T\geq k)}\prod_{i=0}^n\prod_{\Gamma\in\gamma_i}F_{\Delta t_i}^{b_i(\Gamma)}\Big(F_{T-t_{i+1}}(s)\Big)ds. \tag{5.12}$$

First, let us establish that for all (j,t,s), that as $n\to\infty$ $F_{n,t}^j(s)\to F_t^j(s)$. Now $N_{n,t}\uparrow N_t$ almost surely. If $s=1$, $F_{n,t}^j(1)=\mathbb{E}[N_{n,t}^{(j)}]\uparrow\mathbb{E}[N_t^{(j)}]=F_t^j(1)$ by the monotone convergence theorem. If $s<1$, then the function $n\mapsto n^{(j)}s^{n-j}$ is bounded for $n\in\{0,1,2,\ldots\}$, and hence $F_{n,t}^j(s)=\mathbb{E}[N_{n,t}^{(j)}s^{N_{n,t}-j}]\to\mathbb{E}[N_t^{(j)}s^{N_t-j}]=F_t^j(s)$ by the bounded convergence theorem.

To see the convergence of (5.11) to (5.12), by (5.10) we have the domination relation

$$H(s):=\frac{(1-s)^{k-1}F_T^k(s)}{(k-1)!\mathbb{P}(N_T\geq k)}\geq\frac{(1-s)^{k-1}\prod_{i=0}^n\prod_{\Gamma\in\gamma_i}F_{\Delta t_i}^{b_i(\Gamma)}\Big(F_{T-t_{i+1}}(s)\Big)}{(k-1)!\mathbb{P}(N_T\geq k)}=:G(s)\geq 0. \tag{5.13}$$

Similarly,

$$H_n(s):=\frac{(1-s)^{k-1}F_{n,T}^k(s)}{(k-1)!\mathbb{P}(N_{n,T}\geq k)} \tag{5.14}$$

$$\geq G_n(s):=\frac{(1-s)^{k-1}\prod_{i=0}^n\prod_{\Gamma\in\gamma_i}F_{n,\Delta t_i}^{b_i(\Gamma)}\Big(F_{n,T-t_{i+1}}(s)\Big)}{(k-1)!\mathbb{P}(N_{n,T}\geq k)}\geq 0. \tag{5.15}$$

By our assumption $\mathbb{P}(N_T\geq k)>0$, and hence by setting $X=1, N=N_T$ in Lemma 4.6, for every n, $\int_0^1 H_n(s)ds=1=\int_0^1 H(s)ds$. Trivially, $\int_0^1 H_n(s)ds\to\int_0^1 H(s)ds$. So $G_n(s)\to G(s)$ pointwise, $H_n(s)\to H(s)$ pointwise, $\int_0^1 H_n(s)ds\to\int_0^1 H(s)ds$, and $H_n(s)\geq G_n(s)\geq 0$. It follows by Lemma 5.4 that $\int_0^1 G_n(s)ds\to\int_0^1 G(s)ds$. $\qquad\square$

5.3 Proof of Theorem 3.4

In this section we will prove Theorem 3.4, which states that $(\pi_t^{k,L,T})_{t\in[0,T]}$ has a representation in terms of a mixture of Markov processes.

Proof of Theorem 3.4. In light of Theorem 3.1 we may write

$$\mathbb{P}(\pi_{t_1}^{k,L,T}=\gamma_1,\ldots,\pi_{t_n}^{k,L,T}=\gamma_n\mid N_T\geq k)=\int_0^1 m^{k,L,T}(ds)R_{t_1,\ldots,t_n}(ds),$$

where

$$R_{t_1,\ldots,t_n}^{k,L,T}(s,\boldsymbol{\gamma}):=\frac{\prod_{i=0}^n\prod_{\Gamma\in\gamma_i}F_{\Delta t_i}^{b_i(\Gamma)}\big(F_{T-t_{i+1}}(s)\big)}{F_T^k(s)}, \tag{5.16}$$

and

$$m^{k,L,T}(ds):=\frac{(1-s)^{k-1}F_T^k(s)}{(k-1)!\mathbb{P}(N_T\geq k)}ds. \tag{5.17}$$

First of all, we remark that setting $X = 1$ and $N = N_T$ in Lemma 4.6 establishes that $m^{k,L,T}(ds)$ is a probability measure on $[0,1]$. Furthermore, $R^{k,L,T}_{t_1,\ldots,t_n}(s,\cdot)$ is also a probability measure on Π^k_n. To see this, note that by (5.10) that $R^{k,L,T}_{t_1,\ldots,t_n}(s,\boldsymbol{\gamma}) \geq 0$ for all $\boldsymbol{\gamma} \in \Pi^k_n$ and using (5.9) it follows that

$$\sum_{\boldsymbol{\gamma} \in \Pi^k_n} R^{k,L,T}_{t_1,\ldots,t_n}(s,\boldsymbol{\gamma}) = 1.$$

It remains to show that $R^{k,L,T}_{t_1,\ldots,t_n}(s,\cdot)$ are the finite dimensional distributions of a Markov process $(\tilde{\pi}_t)_{t\in[0,T]}$ satisfying the independent blocks property and with transition density given by (3.11).

Now suppose $\boldsymbol{\gamma} := (\gamma_1,\ldots,\gamma_n)$ and $\boldsymbol{\gamma}' := (\gamma_1,\ldots,\gamma_n,\gamma_{n+1},\ldots,\gamma_{n+m})$ are partition chains, and $t_1 < \ldots < t_{n+m}$ are times, then

$$\frac{R^{k,L,T}_{t_1,\ldots,t_{n+m}}(s,\boldsymbol{\gamma}')}{R^{k,L,T}_{t_1,\ldots,t_n}(s,\boldsymbol{\gamma})} = \frac{\prod_{i=n}^m \prod_{\Gamma \in \gamma_i} F^{b_i(\Gamma)}_{\Delta t_i} F_{T-t_{i+1}}(s)}{\prod_{\Gamma \in \gamma_n} F^{|\Gamma|}_{T-t_n}(s)}$$

$$= \prod_{\Gamma \in \gamma_n} \frac{\prod_{i=n}^m \prod_{\Gamma' \in \gamma_i^\Gamma} F^{b_i(\Gamma')}_{\Delta t_i} F_{T-t_{i+1}}(s)}{F^{|\Gamma|}_{T-t_n}(s)}. \tag{5.18}$$

Since this expression only depends on $(\gamma_1,\ldots,\gamma_n)$ through the partition γ_n, $R^{k,L,T}_{t_1,\ldots,t_n}(s,\cdot)$ are the finite dimensional distributions of a Markov process. Additionally, since this expression factorises over $\Gamma \in \gamma_n$, this Markov process satisfies the independent blocks property.

Finally, we need to show the transition density of $(\tilde{\pi}_t)_{t\in[0,T]}$ is given by (3.11). Given $\{\tilde{\pi}_t = \gamma\}$ and a block $\Gamma \in \gamma$, we want to obtain the conditional probability that $\{\pi^\Gamma_{t'} = \delta\}$ for a later time $t < t' < T$. To this end note that by setting $m = n+1$ in (5.18), and letting δ be a partition of $\Gamma \in \gamma_n$, we see that this conditional probability is equal to

$$\frac{F^{|\delta|}_{t_{n+1}-t_n}(F_{T-t_{n+1}}(s)) \prod_{\Delta \in \delta} F^{|\Delta|}_{T-t_{n+1}}(s)}{F^{|\Gamma|}_{T-t_n}(s)},$$

establishing (3.11). $\qquad\square$

5.4 Proof of the split time representation, Theorem 3.2

Proof of Theorem 3.2. Let $u_1 < \ldots < u_n$ be times in $[0,T]$ and let $h_1,\ldots,h_n$ be small positive reals. First we consider the small-(h_i) asymptotics of the quantity

$$\mathbb{R}^{k,L,T}_s(\tilde{\pi}_{u_i} = \eta_{i-1}, \tilde{\pi}_{u_i+h_i} = \eta_i, \quad \forall i = 1,\ldots,n).$$

For $i = 1,\ldots,n$, let $\boldsymbol{\gamma} = (\gamma_1,\ldots,\gamma_{2n})$ be given by $\gamma_{2i-1} = \eta_{i-1}$ and $\gamma_{2i} = \eta_i$, and let $t_{2i-1} = u_i$ and $t_{2i} = u_i + h_i$. Then

$$\mathbb{R}^{k,L,T}_s(\tilde{\pi}_{u_i} = \eta_{i-1}, \tilde{\pi}_{u_i+h_i} = \eta_i, \quad \forall i = 1,\ldots,n) = R^{k,L,T}_{t_1,\ldots,t_{2n}}(s,\boldsymbol{\gamma}).$$

Letting $u_0 = 0, h_0 = 0, u_{n+1} = T$, using (5.16), we have

$$R^{k,L,T}_{t_1,\ldots,t_{2n}}(s,\boldsymbol{\gamma}) = \frac{1}{F^k_T(s)} \prod_{i=0}^n F'_{u_{i+1}-(u_i+h_i)}(F_{T-u_{i+1}}(s))^{|\eta_i|}$$

$$\times \prod_{i=1}^n F^{q_i}_{h_i}(F_{T-(u_i+h_i)}(s)) F'_{h_i}(F_{T-(u_i+h_i)}(s))^{|\eta_i|-1}. \tag{5.19}$$

It is straightforward to verify by the definition of $F_t(s)$ that for small h,

$$F_h^j(s) = hf^j(s) + o(h), j \geq 2, \quad F_h'(s) = 1 + o(1). \tag{5.20}$$

Using (5.20) we obtain

$$\lim_{h_i \downarrow 0} \frac{1}{h_1 \ldots h_n} R_{t_1, \ldots, t_{2n}}^{k, L, T}(s, \boldsymbol{\gamma}) = \frac{\prod_{i=0}^n F_{u_{i+1} - u_i}'(F_{T - u_{i+1}}(s))^{|\eta_i|} \prod_{i=1}^n f^{q_i}(F_{T - u_i}(s))}{F_T^k(s)}.$$

Finally, note by the semigroup identity that $F_t'(F_{T-t}(s)) = F_T'(s)/F_{T-t}'(s)$, and that $|\eta_i| - |\eta_{i-1}| = q_i - 1$, and hence

$$\prod_{i=0}^n F_{\Delta u_i}'(F_{T - u_{i+1}}(s))^{|\eta_i|} = F_T'(s) \prod_{i=1}^n F_{T - u_i}'(s)^{q_i - 1},$$

and hence

$$\lim_{h_i \downarrow 0} \frac{1}{h_1 \ldots h_n} R_{t_1, \ldots, t_{2n}}^{k, L, T}(s, \boldsymbol{\gamma}) = \frac{F_T'(s) \prod_{i=1}^n f^{q_i}(F_{T - u_i}(s)) F_{T - u_i}'(s)^{q_i - 1}}{F_T^k(s)}.$$

It follows that for any collection of time intervals $[a_1, b_1], \ldots, [a_n, b_n]$ such that $b_i < a_{i+1}$,

$$\mathbb{R}_s^{k, L, T}(\tilde{\pi}_{a_i} = \eta_{i-1}, \tilde{\pi}_{b_i} = \eta_i \ \forall \ i = 1, \ldots, n)$$
$$= \int_{a_1}^{b_1} \ldots \int_{a_n}^{b_n} du_1 \ldots du_n \frac{F_T'(s) \prod_{i=1}^n f^{q_i}(F_{T - u_i}(s)) F_{T - u_i}'(s)^{q_i - 1}}{F_T^k(s)}. \tag{5.21}$$

Now by Theorem 3.4, conditional on the event $\{N_T \geq k\}$, the process $(\pi_t^{k, L, T})_{t \in [0, T]}$ has law $\int_0^1 m^{k, L, T}(ds) \mathbb{R}_s^{k, L, T}$. Using this fact in conjunction with (5.21) (and Fubini's theorem in the final equality) we have

$$\mathbb{P}(\pi_{a_i}^{k, L, T} = \eta_{i-1}, \pi_{b_i}^{k, L, T} = \eta_i \ \forall \ i = 1, \ldots, n \mid N_T \geq k)$$
$$= \int_0^1 m^{k, L, T}(ds) \int_{a_1}^{b_1} \ldots \int_{a_n}^{b_n} du_1 \ldots du_n \frac{F_T'(s) \prod_{i=1}^n f^{q_i}(F_{T - u_i}(s)) F_{T - u_i}'(s)^{q_i - 1}}{F_T^k(s)}$$
$$= \int_{a_1}^{b_1} \ldots \int_{a_n}^{b_n} du_1 \ldots du_n \int_0^1 \frac{(1 - s)^{k-1}}{(k - 1)! \mathbb{P}(N_T \geq k)} F_T'(s) \prod_{i=1}^n f^{q_i}(F_{T - u_i}(s)) F_{T - u_i}'(s)^{q_i - 1} ds. \quad \square$$

5.5 Proof of the projective extension, Theorem 3.8

In this section we will use the abbreviation $(\pi_t^k)_{t \in [0, T]} := (\pi_t^{k, L, T})_{t \in [0, T]}$.

Proof of Theorem 3.8. If on the event $\{N_T \geq k + j\}$, $k + j$ distinct particles $U_1, \ldots, U_{k+j}$ are chosen uniformly from those alive at time T, then the first k of them $U_1, \ldots, U_k$ represent a uniformly chosen sample of k-distinct particles. It follows by definition that on the event $\{N_T \geq k + j\}$ that $\left(\pi_t^{k+j}\big|^k\right)_{t \in [0, T]}$ and $(\pi_t^k)_{t \in [0, T]}$ are identical in law.

It remains to compute the law of the stochastic process $\left(\pi_t^{k+j}\big|^k\right)_{t \in [0, T]}$. Note we can write

$$\mathbb{P}(\pi_{t_1}^{k+j}\big|^k = \gamma_1, \ldots, \pi_{t_n}^{k+j}\big|^k = \gamma_n, \ N_T \geq k + j)$$
$$= \sum_{\boldsymbol{\eta} \in \Pi_n^{k+j} : \boldsymbol{\eta}|^k = \boldsymbol{\gamma}} \mathbb{P}(\pi_{t_1}^{k+j} = \eta_1, \ldots, \pi_{t_n}^{k+j} = \eta_n, \ N_T \geq k + j)$$
$$= \sum_{\boldsymbol{\eta} \in \Pi_n^{k+j} : \boldsymbol{\eta}|^k = \boldsymbol{\gamma}} \int_0^1 \frac{(1 - s)^{k+j-1}}{(k + j - 1)!} \prod_{i=0}^n \prod_{H \in \eta_i} F_{\Delta t_i}^{b_i(H)} \left(F_{T - t_{i+1}}(s)\right) ds$$
$$= \int_0^1 \frac{(1 - s)^{k+j-1}}{(k + j - 1)!} \sum_{\boldsymbol{\eta} \in \Pi_n^{k+j} : \boldsymbol{\eta}|^k = \boldsymbol{\gamma}} \prod_{i=0}^n \prod_{H \in \eta_i} F_{\Delta t_i}^{b_i(H)} \left(F_{T - t_{i+1}}(s)\right) ds.$$

Setting $g_i := F_{\Delta t_i}$ in Lemma 5.2, we have

$$\sum_{\boldsymbol{\eta} \in \Pi_n^{k+j} : \boldsymbol{\eta}|^k = \boldsymbol{\gamma}} \prod_{i=0}^{n} \prod_{H \in \eta_i} F_{\Delta t_i}^{b_i(H)}\left(F_{T-t_{i+1}}(s)\right) ds = \frac{\partial^j}{\partial s^j}\left(\prod_{i=0}^{n} \prod_{\Gamma \in \gamma_i} F_{\Delta t_i}^{b_i(\Gamma)}\left(F_{T-t_{i+1}}(s)\right)\right),$$

and the result follows. $\qquad\square$

6 Proofs of asymptotic-T results

6.1 Supercritical asymptotics

When $m = \mathbb{E}[L] > 1$ and the Kesten-Stigum condition $\mathbb{E}[L \log_+ L] < \infty$ holds, the unit-mean and non-negative martingale $W_t := N_t e^{-(m-1)t}$ converges to almost-surely to a non-degenerate limit $W := W_\infty$, with the properties that $\mathbb{E}[W] = 1$ and

$$\{W > 0\} = \{N_t > 0 \,\forall\, t\} \text{ almost surely.} \tag{6.1}$$

Defining $\varphi : [0, \infty) \to [0, 1]$ by $\varphi(v) := \mathbb{E}[e^{-vW}]$, we note that by Fubini's theorem,

$$\mathbb{E}[W^k e^{-vW}] = (-1)^k \varphi^k(v). \tag{6.2}$$

The following lemma gives us the scaling limit of the generating function derivatives under the change of variable $s = e^{-ve^{-(m-1)T}}$.

Lemma 6.1. For every $v > 0$, $t \geq 0$, and non-negative integer k,

$$\lim_{T \to \infty} e^{-k(m-1)T} F_{T-t}^k(e^{-ve^{-(m-1)T}}) = (-1)^k e^{-k(m-1)t} \varphi^k(ve^{-(m-1)t}).$$

Proof. Write

$$e^{-k(m-1)T} F_{T-t}^k(e^{-ve^{-(m-1)T}}) = e^{-k(m-1)T} \mathbb{E}\left[N_{T-t}^{(k)} e^{-ve^{-(m-1)T}N_{T-t}}\right]$$
$$= e^{-k(m-1)t} \mathbb{E}[h_{v,t}^T(W_{T-t})],$$

where $h_{v,t}^T(x) = \left(\prod_{i=0}^{k-1}(x - ie^{-(m-1)(T-t)})\right) \exp\left(-ve^{-(m-1)t}(x - ke^{-(m-1)(T-t)})\right)$. Note that as $T \to \infty$, $h_{v,t}^T(x)$ converges uniformly on $[0, \infty)$ to

$$h_{v,t}(x) := x^k e^{-ve^{-(m-1)t}x}.$$

Since $W_{T-t} \to W$ almost surely as $T \to \infty$, it follows that $h_{v,t}^T(W_{T-t})$ converges almost surely to $h_{v,t}(W)$. Since $h_{v,T}^T(x)$ are bounded as T varies, it follows by the bounded convergence theorem that

$$\mathbb{E}[h_{v,t}^T(W_{T-t})] \to \mathbb{E}[h_{v,t}(W)] = (-1)^k \varphi^k(ve^{-(m-1)t}). \qquad\square$$

We are now ready to prove our main result for supercritical trees.

Proof of Theorem 3.5. By Theorem 3.1,

$$\mathbb{P}(\pi_{t_1}^{k,L,T} = \gamma_1, \ldots, \pi_{t_n}^{k,L,T} = \gamma_n \mid N_T \geq k) \tag{6.3}$$

$$= \frac{1}{\mathbb{P}(N_T \geq k)} \int_0^1 \frac{(1-s)^{k-1}}{(k-1)!} \prod_{i=0}^{n} \prod_{\Gamma \in \gamma_i} F_{\Delta t_i}^{b_i(\Gamma)}\left(F_{T-t_{i+1}}(s)\right) ds. \tag{6.4}$$

For a fixed mesh $(t_i)_{i \leq n}$, the only interval $[t_i, t_{i+1}]$ that grows with T is the final one $[t_n, T]$, and hence it is convenient to write

$$= \frac{1}{\mathbb{P}(N_T \geq k)} \int_0^1 \frac{(1-s)^{k-1}}{(k-1)!} \prod_{i=0}^{n-1} \prod_{\Gamma \in \gamma_i} F_{\Delta t_i}^{b_i(\Gamma)}\left(F_{T-t_{i+1}}(s)\right) \prod_{\Gamma \in \gamma_n} F_{T-t_n}^{|\Gamma|}(s) ds, \tag{6.5}$$

where we note $b_n(\Gamma) = |\Gamma|$ by definition. Applying the change of variable $s = e^{-ve^{-(m-1)T}}$, we can write

$$\mathbb{P}(\pi_{t_1}^{k,L,T} = \gamma_1, \ldots, \pi_{t_n}^{k,L,T} = \gamma_n \mid N_T \geq k) = \int_0^\infty G^T(v)dv, \tag{6.6}$$

where

$$G^T(v) := e^{-(m-1)T} e^{-e^{-(m-1)T}v} \frac{(1 - e^{-e^{-(m-1)T}v})^{k-1}}{(k-1)!\mathbb{P}(N_T \geq k)} \prod_{i=0}^{n-1} \prod_{\Gamma \in \gamma_i} F_{\Delta t_i}^{b_i(\Gamma)}\left(F_{T-t_{i+1}}(e^{-e^{-(m-1)T}v})\right)$$
$$\times \prod_{\Gamma \in \gamma_n} F_{T-t_n}^{|\Gamma|}(e^{-e^{-(m-1)T}v}).$$

Note by (6.1) that $\mathbb{P}(N_T \geq k) \to 1 - \varphi(\infty)$, and by using Lemma 6.1 it can then be seen that as $T \to \infty$, $G^T(v)$ converges pointwise to

$$G(v) := \frac{e^{-k(m-1)t_n}v^{k-1}}{(k-1)!(1-\varphi(\infty))} \prod_{i=0}^{n-1} \prod_{\Gamma \in \gamma_i} F_{\Delta t_i}^{b_i(\Gamma)}\left(\varphi(e^{-(m-1)t_i}v)\right) \prod_{\Gamma \in \gamma_n} (-1)^{|\Gamma|}\varphi^{|\Gamma|}(e^{-(m-1)t_n}v).$$

It remains to establish that $\int_0^\infty G^T(v)dv \to \int_0^\infty G(v)dv$, which we prove using the dominated convergence theorem, Lemma 5.4. To this end, let

$$H^T(v) := e^{-(m-1)T} e^{-ve^{-(m-1)T}} \frac{(1 - e^{-e^{-(m-1)T}v})^{k-1}}{(k-1)!\mathbb{P}(N_T \geq k)} F_T^k(e^{-e^{-(m-1)T}v}).$$

Setting $s = e^{-e^{-(m-1)T}v}$ in (5.10), we see that $0 \leq G^T(v) \leq H^T(v)$. Furthermore, $H^T(v)$ converges pointwise to

$$H(v) := \frac{(-1)^k v^{k-1} \varphi^k(v)}{(k-1)!(1-\varphi(\infty))}. \tag{6.7}$$

Finally, we note that for each T, by changing variable $s = e^{-e^{-(m-1)T}v}$ and then setting $X = 1$ in Lemma 4.6, it can be seen that $H^T(v)dv$ is a probability measure. We now show that $H(v)dv$ is also a probability measure. Using (6.2) in the first equality below and Fubini's theorem in the second,

$$\int_0^\infty H(v)dv = \frac{1}{(k-1)!(1-\varphi(\infty))} \int_0^\infty v^{k-1}\mathbb{E}[W^k e^{-vW}]dv$$
$$= \frac{1}{(k-1)!(1-\varphi(\infty))} \mathbb{E}\left[p(W)\right],$$

where by the definition of the gamma integral, $p(w) := \mathbb{1}_{w>0} \int_0^\infty (vw)^k e^{-vw} \frac{dv}{v} = (k-1)!\mathbb{1}_{w>0}$. Using $\mathbb{P}(W > 0) = 1 - \varphi(\infty)$, we obtain

$$\int_0^\infty H(v)dv = \frac{1}{(k-1)!(1-\varphi(\infty))} \mathbb{E}\left[(k-1)!\mathbb{1}_{W>0}\right] = 1. \tag{6.8}$$

So $H^T(v) \geq G^T(v) \geq 0$, $H(v) \geq G(v)$, $H^T(v)$ converge pointwise to $H(v)$ and we trivially have $\int_0^\infty H^T(v)dv = 1 \to \int_0^1 H(v)dv = 1$. It follows by Lemma 5.4 that $\int_0^\infty G^T(v)dv \to \int_0^\infty G(v)dv$. $\qquad\square$

6.2 Critical asymptotics

Recall from Section 2 the Kolmogorov-Yaglom exponential limit law [4, III.7], which states that when $f'(1) = 1$

$$\lim_{T\to\infty} T\mathbb{P}(N_T > 0) = \frac{1}{c}, \quad \lim_{T\to\infty} \mathbb{P}\left(\frac{N_T}{cT} > x \,\middle|\, N_T > 0\right) = e^{-x}, \tag{6.9}$$

where $c := f''(1)/2$.

The following lemma gives us the scaled asymptotics of $F_t(s)$ when $f'(1) = 1$ and $c = f''(1)/2 < \infty$.

Lemma 6.2. For any $\theta \in [0, \infty)$, and $a > 0$, $b \geq 0$, $j \geq 1$,

$$\lim_{T\to\infty} (cT)^{-(j-1)} F_{aT}^j \left(F_{bT}(\exp(-\theta/cT))\right) = a^{j-1} j! \left(\frac{1 + \theta b}{1 + \theta(a+b)}\right)^{j+1}.$$

Proof. Throughout this proof, Z will refer to a standard exponential random variable, and we note

$$\mathbb{E}[Z^j e^{-\phi Z}] = \frac{j!}{(1+\phi)^{j+1}}.$$

By (6.9), conditioned on $\{N_{bT} > 0\}$, the random variable N_{bT}/cT converges in distribution to bZ as $T \to \infty$. Now since $x \mapsto e^{-\theta x}$ is bounded for $x \geq 0$, it follows that for any $b > 0$ we have

$$\lim_{T\to\infty} T\left(1 - F_{bT}(\exp(-\theta/cT))\right) = \lim_{T\to\infty} T(1 - \mathbb{E}[\exp(-\theta N_{bT}/cT)])$$

$$= \lim_{T\to\infty} T\mathbb{P}(N_{bT} > 0)\Big(1 - \mathbb{E}[\exp(-\theta N_{bT}/cT)|N_{bT} > 0]\Big)$$

$$= \frac{1}{cb}(1 - \mathbb{E}[e^{-\theta b Z}]),$$

where Z is a standard exponential, and hence

$$\lim_{T\to\infty} T(1 - F_{bT}(\exp(-\theta/cT))) = \frac{1}{c}\frac{\theta}{1 + \theta b}. \tag{6.10}$$

Note that (6.10) is also true for $b = 0$, since $F_0(s) = s$. It follows that for any $b \geq 0$,

$$\lim_{T\to\infty} F_{bT}(\exp(-\theta/cT))^{cT} = \exp\left(-\frac{\theta}{1 + \theta b}\right).$$

Moreover, conditional on $\{N_{aT} > 0\}$,

$$\frac{N_{aT}^{(j)}}{(caT)^j} F_{bT}(\exp(-\theta/cT))^{N_{aT}-j} = \frac{N_{aT}^{(j)}}{(caT)^j} F_{bT}(\exp(-\theta/cT))^{caT \times (N_{aT}/caT - j/caT)}$$

converges in distribution to $Z^j \exp\left(-\frac{\theta}{1+\theta b} aZ\right)$. It follows that

$$\lim_{T\to\infty} (cT)^{-(j-1)} F_{aT}^j \left(F_{bT}(\exp(-\theta/cT))\right)$$

$$= a^{j-1} \lim_{T\to\infty} (caT)\mathbb{P}(N_{aT} > 0) \lim_{T\to\infty} \mathbb{E}\left[\frac{N_{aT}^{(j)}}{(caT)^j} F_{bT}(\exp(-\theta/cT))^{N_{aT}-j} \,\middle|\, N_{aT} > 0\right]$$

$$= a^{j-1} \mathbb{E}\left[Z^j \exp\left(-\frac{\theta}{1+\theta b} aZ\right)\right]$$

$$= a^{j-1} j! \left(\frac{1 + \theta b}{1 + \theta(a+b)}\right)^{j+1},$$

as required. □

We now prove our main result for critical trees.

Proof of Theorem 3.6. Let $f'(1) = 1, c = f''(1)/2 < \infty$. Then by (3.6), for elements $t_1 < \ldots < t_n$ of $[0,1]$,

$$\mathbb{P}(\pi_{Tt_1}^{k,L,T} = \gamma_1, \ldots, \pi_{Tt_n}^{k,L,T} = \gamma_n | N_T \geq k)$$
$$= \frac{1}{\mathbb{P}(N_T \geq k)} \int_0^1 \frac{(1-s)^{k-1}}{(k-1)!} \prod_{i=0}^n \prod_{\Gamma \in \gamma_i} F_{T\Delta t_i}^{b_i(\Gamma)} \left(F_{T(1-t_{i+1})}(s) \right) ds.$$

Taking the change of variable $s = e^{-\theta/cT}$, we may write

$$\mathbb{P}(\pi_{Tt_1}^{k,L,T} = \gamma_1, \ldots, \pi_{Tt_n}^{k,L,T} = \gamma_n | N_T \geq k) = \int_0^\infty G^T(\theta) d\theta,$$

where

$$G^T(\theta) := \frac{e^{-\theta/cT}}{cT\mathbb{P}(N_T \geq k)} \frac{(cT)^{k-1}(1 - e^{-\theta/cT})^{k-1}}{(k-1)!} \frac{1}{(cT)^{k-1}}$$
$$\times \prod_{i=0}^n \prod_{\Gamma \in \gamma_i} F_{T(\Delta t_i)}^{b_i(\Gamma)} \left(F_{T(1-t_{i+1})}(e^{-\theta/cT}) \right).$$

Now note that given any chain of partitions $(\gamma_1, \ldots, \gamma_n)$, we have by the definition of the fragmentation numbers $b_i(\Gamma)$,

$$\sum_{\Gamma \in \gamma_i} (b_i(\Gamma) - 1) = |\gamma_{i+1}| - |\gamma_i|,$$

and in particular, $\sum_{i=0}^n \sum_{\Gamma \in \gamma_i} b_i(\Gamma) = k - 1$. Using these facts in conjunction with Lemma 6.2 we have for every θ,

$$G^T(\theta) \to G(\theta) := \prod_{i=0}^n \prod_{\Gamma \in \gamma_i} b_i(\Gamma)! \frac{\theta^{k-1}}{(k-1)!} \prod_{i=0}^n (\Delta t_i)^{|\gamma_{i+1}| - |\gamma_i|} \left(\frac{1 + (1-t_{i+1})\theta}{1 + (1-t_i)\theta} \right)^{|\gamma_{i+1}|}.$$

It remains to show that $\int_0^\infty G^T(\theta) d\theta \to \int_0^\infty G(\theta) d\theta$. To this end, define

$$H^T(\theta) := \frac{e^{-\theta/cT}}{cT\mathbb{P}(N_T \geq k)} \frac{(cT)^{k-1}(1 - e^{-\theta/cT})^{k-1}}{(k-1)!} \frac{1}{(cT)^{k-1}} F_T^k(e^{-\theta/cT}).$$

Setting $s = e^{-\theta/cT}$ in (5.10) we see that $H^T(\theta) \geq G^T(\theta) \geq 0$. We also note that by Lemma 6.2 that $H^T(\theta)$ converges pointwise to

$$H(\theta) := \frac{k\theta^{k-1}}{(1+\theta)^{k+1}}.$$

Furthermore, by taking the change of variable $s = e^{-\theta/cT}$ and using Lemma 4.6, it can be seen each $H^T(\theta)d\theta$ is a probability measure on $[0, \infty)$. It is also straightforward to verify that $H(\theta)d\theta$ is also a probability measure on $[0, \infty)$.

In particular, we trivially have $\int_0^\infty H^T(\theta)d\theta \to \int_0^\infty H(\theta)d\theta$, and it follows from Lemma 5.4 that $\int_0^\infty G^T(\theta)d\theta \to \int_0^\infty G(\theta)d\theta$. $\qquad \square$

6.3 Subcritical asymptotics

Finally, we prove our main result in the subcritical case. This proof is more straightforward than the supercritical and critical cases since on survival until a large time T, there are only a constant order of particles alive at time T, and as a result, no scaling is needed in the generating functions.

Recall that when $m < 1$, there exist non-negative numbers $\{c_j : j \geq 1\}$ satisfying $\sum_{j \geq 1} c_j = 1$ such that

$$\lim_{T \to \infty} \mathbb{P}(N_T = j | N_T > 0) = c_j. \tag{6.11}$$

We set $C(s) := \sum_{j \geq 1} c_j s^j$, and note that according to [4, I.11] that $C'(1) < \infty \iff \mathbb{E}[L \log_+ L] < \infty$.

Lemma 6.3. Suppose $f'(1) < 1$ and $\mathbb{E}[L \log_+ L] < \infty$. Then for all $l \geq 1$,

$$\lim_{T \to \infty} \frac{F_{T-t}^l(F_t(s))}{\mathbb{P}(N_T \geq k)} = \frac{e^{-(m-1)t}}{1 - \sum_{j=1}^{k-1} c_j} C^l(F_t(s)). \tag{6.12}$$

Proof. First note that

$$\lim_{T \to \infty} e^{-(m-1)T} \mathbb{P}(N_T > 0) = \lim_{T \to \infty} e^{-(m-1)T} \frac{\mathbb{E}[N_T]}{\mathbb{E}[N_T | N_T > 0]} = \frac{1}{C'(1)}.$$

Now write

$$\lim_{T \to \infty} \frac{F_{T-t}^l(F_t(s))}{\mathbb{P}(N_T \geq k)} = \lim_{T \to \infty} \frac{\mathbb{P}(N_T > 0)}{\mathbb{P}(N_T \geq k)} \lim_{T \to \infty} \frac{\mathbb{P}(N_{T-t} > 0)}{\mathbb{P}(N_T > 0)} \lim_{T \to \infty} \frac{F_{T-t}^l(F_t(s))}{\mathbb{P}(N_{T-t} > 0)},$$

and use the definition (6.11). $\qquad\square$

We now prove our main result for subcritical trees.

Proof of Theorem 3.7. By replacing t_i with $T - t_{n-i+1}$ in Theorem 3.1, we have the following formula for the finite dimensional distributions of $(\rho_t^{k,L,T})_{t \in [0,T]}$:

$$\mathbb{P}(\rho_{t_1}^{k,L,T} = \gamma_1, \ldots, \rho_{t_n}^{k,L,T} = \gamma_n | N_T \geq k)$$
$$= \frac{1}{\mathbb{P}(N_T \geq k)} \int_0^1 \frac{(1-s)^{k-1}}{(k-1)!} \prod_{i=1}^{n+1} \prod_{\Gamma \in \gamma_i} F_{\Delta t_{i-1}}^{m_i(\Gamma)}(F_{t_{i-1}}(s)) ds, \tag{6.13}$$

where $\gamma_1 \succ \ldots \succ \gamma_n$ is a chain with merger numbers $(m_i(\Gamma))$.

Note that $m_{n+1}(\{1, \ldots, k\}) = |\gamma_n|$. Now for fixed $(t_i)_{i \leq n}$, as we send $T \to \infty$ the only time interval $[t_i, t_{i+1}]$ in (6.13) that grows with T is $[t_n, T]$. For this reason it is useful to write

$$= \int_0^1 \frac{(1-s)^{k-1}}{(k-1)!} \prod_{i=1}^{n} \prod_{\Gamma \in \gamma_i} F_{\Delta t_{i-1}}^{m_i(\Gamma)}(F_{t_{i-1}}(s)) \frac{F_{T-t_n}^{|\gamma_n|}(F_{t_n}(s))}{\mathbb{P}(N_T \geq k)} ds =: \int_0^1 G^T(s) ds.$$

By (6.12), for each $s \in [0, 1]$,

$$G^T(s) \to G(s) := \frac{e^{-(m-1)t_n}}{1 - \sum_{j=1}^{k-1} c_j} \frac{(1-s)^{k-1}}{(k-1)!} \prod_{i=1}^{n} \prod_{\Gamma \in \gamma_i} F_{\Delta t_{i-1}}^{m_i(\Gamma)}(F_{t_{i-1}}(s)) C^{|\gamma_n|}(F_{t_n}(s)).$$

It remains to establish that $\int_0^1 G^T(s)ds \to \int_0^1 G(s)ds$. To this end, define

$$H^T(s) := \frac{(1-s)^{k-1}F_T^k(s)}{(k-1)!\mathbb{P}(N_T \geq k)}$$

By replacing t_i with $T - t_{n+1-i}$ and γ_i with γ_{n+1-i} (and noting $m_i(\Gamma) = b_{n-i}(\Gamma)$) in (5.10), we obtain

$$H^T(s) \geq G^T(s) \geq 0.$$

Furthermore, by applying (6.12) with $t = 0$ and $l = k$, we see that $H^T(s)$ converges pointwise to

$$H(s) := \frac{(1-s)^{k-1}}{(k-1)!(1 - \sum_{j=1}^{k-1} c_j)} C^k(s).$$

Finally, by Theorem 3.4, $H^T(s)ds = m^{k,L,T}(ds)$ is a probability measure on $[0,1]$, and furthermore, by Lemma 4.6, so is $H(s)ds$. We trivially have $\int_0^1 H^T(s)ds \to \int_0^1 H(s)ds$, and hence by Lemma 5.4, $\int_0^1 G^T(s)ds \to \int_0^1 G(s)ds$. $\qquad\square$

References

[1] Athreya, K. B.: Coalescence in the recent past in rapidly growing populations. *Stoch. Proc. Appl.* **122**, (2012), 3757–3766. MR-2965924

[2] Athreya, K. B.: Coalescence in critical and subcritical Galton-Watson branching processes. *J. Appl. Prob.* **49**, (2012), 627–638. MR-3012088

[3] Athreya, K. B.: Coalescence in branching processes, in *Branching Processes and Their Applications*. **219**, Lecture Notes in Statistics, Springer, (2016). MR-3587978

[4] Athreya, K. B. and Ney, P. E.: *Branching Processes*, Springer-Verlag, New York, (1972). MR-0373040

[5] Bühler, W.: Generations and degree of relationship in supercritical Markov branching processes. *Z. Wahrscheinlichkeitstheorie verw. Geb.* **18**, (1971), 141–152. MR-0295444

[6] Durrett, R.: The genealogy of critical branching processes. *Stoch. Proc. Appl.* **8**(1), (1978), 101–116. MR-0511879

[7] Grosjean, N. and Huillet, T.: On the genealogy and coalescence times of Bienayme-Galton-Watson branching processes, *Stoch. Models* **34**(1), (2018). MR-3776534

[8] Harris, T. E.: *The Theory of Branching Processes*, Berlin, Springer, (1963). MR-0163361

[9] Harris, S. C., Johnston, S. G. G. and Roberts, M.: The coalescent structure of continuous-time Galton-Watson trees, arXiv:1703.00299

[10] Johnson, W. P.: The curious history of Faà Di Bruno's formula. *Am. Math. Mon.* **109**(3), (2002), 217–234. MR-1903577

[11] Kallenberg, O.: Foundations of Modern Probability. Springer-Verlag, New York, (1997). MR-1464694

[12] Kesten, H. and Stigum, B. P.: A limit theorem for multidimensional Galton-Watson processes. *Ann. Math. Statistics* **37**(5), (1966), 1211–1223. MR-0198552

[13] Kingman, J. F. C.: The coalescent. *Stoch. Proc. Appl.* **13**(3), (1982), 235–248. MR-0671034

[14] Lambert, A.: Coalescence times for the branching process. *Adv. Appl. Prob.* **35**(4), (2003), 1071–1089. MR-2014270

[15] Le, V.: Coalescence times for the Bienaymé-Galton-Watson process. *J. Appl. Probab.* **51**, (2014), 209-218. MR-3189452

[16] Mutagh, F.: Counting dendrograms: A survey. *Discrete Appl. Math.* **7**(2), (1984), 191–199. MR-0727923

[17] O'Connell, N.: The genealogy of branching processes and the age of our most recent common ancestor. *Adv. Appl. Probab.* **27**(2), (1995), 418–442. MR-1334822

[18] Zubkov, A. M.: Limiting distributions of the distance to the closest common ancestor. *Theory Prob. Appl.* **20**(3), (1976), 602–612. MR-0397915

Electronic Journal of Probability
Electronic Communications in Probability

Advantages of publishing in EJP-ECP

- Very high standards

- Free for authors, free for readers

- Quick publication (no backlog)

- Secure publication (LOCKSS[1])

- Easy interface (EJMS[2])

Economical model of EJP-ECP

- Non profit, sponsored by IMS[3], BS[4] , ProjectEuclid[5]

- Purely electronic

Help keep the journal free and vigorous

- Donate to the IMS open access fund[6] (click here to donate!)

- Submit your best articles to EJP-ECP

- Choose EJP-ECP over for-profit journals

---

[1]LOCKSS: Lots of Copies Keep Stuff Safe http://www.lockss.org/
[2]EJMS: Electronic Journal Management System http://www.vtex.lt/en/ejms.html
[3]IMS: Institute of Mathematical Statistics http://www.imstat.org/
[4]BS: Bernoulli Society http://www.bernoulli-society.org/
[5]Project Euclid: https://projecteuclid.org/
[6]IMS Open Access Fund: http://www.imstat.org/publications/open.htm