Electronic Journal of Probability

Vol. 5 (2000) Paper no. 17, pages 1–12.

Journal URL
http://www.math.washington.edu/~ejpecp/
Paper URL
http://www.math.washington.edu/~ejpecp/EjpVol5/paper17.abs.html

POISSON POINT PROCESS LIMITS IN SIZE-BIASED GALTON-WATSON TREES

Jochen Geiger
Universität Frankfurt
Fachbereich Mathematik, Postfach 11 19 32, D-60054 Frankfurt, Germany
geiger@mi.informatik.uni-frankfurt.de

Abstract Let $\widehat{T}_t$ be a critical binary continuous-time Galton-Watson tree size-biased according to the number of particles at time t. Decompose the population at t according to the particles' degree of relationship with a distinguished particle picked purely at random from those alive at t. Keeping track of the times when the different families grow out of the distinguished line of descent and the related family sizes at t, we represent this relationship structure as a point process in a time-size plane. We study limits of these point processes in the single- and some multitype case.

Keywords Galton-Watson process, random tree, point process, limit laws

AMS subject classification 60J80, 60G55

Research supported by Deutsche Forschungsgemeinschaft

Submitted to EJP on February 25, 2000. Accepted July 10, 2000.

1 Introduction

In this paper we introduce a point process representation of the genealogy of a random tree which we use to study asymptotics for a certain classical model of random trees. To explain the point process representation consider a branching population with one particle being distinguished. Say that at a fixed time t particles belong to the same family, if they have the same most recent common ancestor with the distinguished particle. This relationship structure can be represented by a point process on $(0,t] \times \{1,2,\ldots\}$ where a point has as its coordinates the minimal (backwards) time that family members have a common ancestor with the distinguished particle and the family size at t. Here we study limits of such point processes representing a critical single- or multitype Galton-Watson tree size-biased according to the number of particles at time t, where the distinguished particle is picked purely at random from the population at t.

In the single-type case we show that the Poisson point process Ψ_t representing the family structure of the size-biased Galton-Watson tree has a weak limit Ψ, say. The intensity of the Poisson point process Ψ reflects the asymptotic behavior of critical Galton-Watson branching processes and explains the Lévy-Khinchin representation of the limiting gamma distribution for the rescaled population size at t.

We then discuss the case where the particles may have different types and mutate at a fixed positive rate μ to a new type not present so far. Let Ψ_t^μ be the point process representing the subpopulation with the type of the distinguished particle at time t. Our main result states that the rescaled point process Ψ_t^μ converges in distribution as first $t \to \infty$ and then $\mu \to 0$, and identifies the limiting point process as a random portion $R_U\Psi$ of the weak limit Ψ from the single-type case. Here the relative size U is uniform on $(0,1)$, independent of Ψ, and R_U denotes the restriction to the time period $(0,U]$. This result and a scaling property of Ψ are applied to obtain an exponential limit law for the size of the distinguished type in equilibrium.

In Section 2 we first review the decomposition of the size-biased Galton-Watson tree along the distinguished line of descent of a particle chosen purely at random and give a precise description of its point process representation [formula (2.7)]. We then derive the limit laws for the single-type case. The multitype case is contained in Section 3.

Size-biased Galton-Watson trees arise as so-called Palm trees in branching particle systems [4, 13] and as limits of the pedigree of a randomly selected particle in a large population [11, 12]. They are closely related to Galton-Watson trees conditioned on non-extinction (see the remark at the end of Section 2) or on large total population size. The genealogy of critical Galton-Watson trees conditioned on non-extinction has been studied in [6, 8] and later in the super process context, see e.g. [15]. Galton-Watson trees conditioned on large total population size have been shown to converge to some limiting continuum tree [1]. For spinal decompositions and related constructions of random family trees, see e.g. [2, 9, 16].

2 The size-biased Galton-Watson tree

Consider a critical binary continuous-time Galton-Watson branching process $(Z_t)_{t\geq 0}$ starting with a single founding ancestor. In such a process particles have independent exponential lifetimes with rate one, say, and produce either two or no offspring with probability $p_0 = p_2 = 1/2$, the lifetimes and offspring numbers being independent (for background of Galton-Watson processes, see e.g. [3]). Let T be the random family tree obtained by having one edge for each particle produced with the length of an edge being the particle's lifetime. We think of T as a rooted planar tree with the distinguishable offspring of each particle ordered left and right. T is identified by its shape $[T]$ and the collection of the particles' lifetimes $L_i, 1 \leq i \leq n(T)$. Here $n(T)$ is the total progeny, i.e. the number of particles ever produced (including the root), and edges are labeled in some fixed deterministic manner, e.g. according to a depth-first search procedure (for a formal description of the space of trees we refer to [2, 17]). By the shape $[T]$ of a tree T we mean the tree obtained from T by setting all edges to constant length. See Figure 1 for an example of a realization of T.

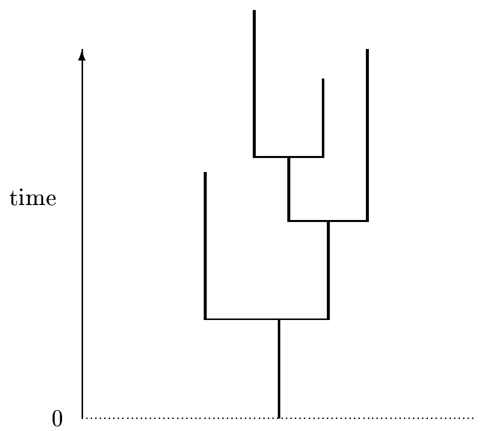

Figure 1: *A realization* $\mathbf{t}$ *of the Galton-Watson tree* T*. The tree* $\mathbf{t}$ *has total progeny* $n(\mathbf{t}) = 7$*. The left child of the founding ancestor at the bottom of the tree has no children, the right child has two children. The horizontal lines are merely to indicate connections.*

For a rooted binary tree $\mathbf{t}$ with edge lengths ℓ_i, $1 \le i \le n(\mathbf{t})$, let $T \in d\mathbf{t}$ be the event that T has shape $[\mathbf{t}]$ and edge lengths in $[\ell_i, \ell_i + d\ell_i)$. Since the offspring numbers and exponential lifetimes are assumed to be i.i.d. we have

$$\begin{aligned} \mathrm{P}(T \in d\mathbf{t}) &= \mathrm{P}([T] = [\mathbf{t}])\, \mathrm{P}(T \in d\mathbf{t} \,|\, [T] = [\mathbf{t}]) \\ &= 2^{-n(\mathbf{t})} \exp\Big(- \sum_{i=1}^{n(\mathbf{t})} \ell_i \Big)\, d\ell_1 \cdots d\ell_{n(\mathbf{t})}. \end{aligned} \tag{2.1}$$

Note that the distribution of T depends on $[T]$ only through $n(T)$, that is any shape with fixed total progeny is equally likely. By the assumed criticality, T is almost surely finite, which in particular means that the weights in (2.1) sum to 1.

The law of the *size-biased* critical binary continuous-time Galton-Watson tree $\widehat{T}_t$ is defined as

$$\mathrm{P}(\widehat{T}_t \in d\mathbf{t}) \;=\; z_t(\mathbf{t})\, \mathrm{P}(T \in d\mathbf{t}), \; t \ge 0, \tag{2.2}$$

where $z_t(\mathbf{t})$ is the number of edges at height t of the tree $\mathbf{t}$ (note that, by the assumed criticality, the normalizing constant $\mathrm{E}z_t(T)$ on the right-hand side of (2.2) equals 1, hence $\widehat{T}_t$ is really distributed according to a probability law). Relation (2.2) says that a tree with population size k at time t is chosen k times as likely as if sampling were according to Galton-Watson measure. In particular, if $\widehat{Z}_t = z_t(\widehat{T}_t)$ then $\widehat{Z}_t$ has the size-biased distribution of $Z_t = z_t(T)$, that is,

$$\mathrm{P}(\widehat{Z}_t = k) \;=\; k\, \mathrm{P}(Z_t = k), \; k \ge 0. \tag{2.3}$$

The size-biased Galton-Watson tree $\widehat{T}_t$ is a finite random tree of minimum height t. The tree is no longer symmetric about rehanging subtrees, which means that particles no longer evolve independently of each other nor homogeneously in time. However, as can be seen from the following explicit construction, the size-biased tree $\widehat{T}_t$ has a transparent structure if decomposed along the line of descent of a particle V, say, chosen purely at random from the population at time t. We construct a random binary continuous-time tree T_t^* and a random edge V^* of T_t^* at height t such that the pair (T_t^*, V^*) has distribution $\mathcal{L}(\widehat{T}_t, V)$. We start with the construction of the line of descent of the distinguished particle V^* alive at t. Take the points $T_1 < T_2 < \ldots$ of a homogeneous rate one Poisson process on $(0, t]$ to be the birth times of V^* and its ascendants (other than the founding ancestor) and let each particle in V^*'s line of descent (including V^*, excluding the founding ancestor) independently have probability $\frac{1}{2}$ of being the sibling to the left.

Otherwise the particle is positioned to the right. Having constructed V^*'s line of descent we have to determine the distribution of the descendant tree of V^* and of the descendant trees of the siblings of the particles in the distinguished line of descent: Let the siblings of of the particles in the distinguished line of descent found ordinary critical binary Galton-Watson trees and attach another independent such tree at the top (i.e. at height t) of the distinguished line of descent. Call the resulting tree T_t^*. See Figure 2 for an illustration of (T_t^*, V^*).

Note that particles not in the distinguished line of descent evolve as in an ordinary critical binary Galton-Watson branching process and so does the distinguished particle V^* after time t. We remark that V^* cannot be recovered from T_t^* unless $z_t(T_t^*) = 1$. In fact, the distinguished edge V^* is distributed uniformly on the edges of T_t^* at height t,

$$\mathrm{P}(V^* = e \,|\, T_t^* = \mathbf{t}) \;=\; z_t(\mathbf{t})^{-1}, \tag{2.4}$$

where e is any edge at height t of the finite tree $\mathbf{t}$ with minimum height t. The marginal distribution of the pair (T_t^*, V^*) on the space of trees is the size-biased distribution (2.2),

$$T_t^* \overset{d}{=} \widehat{T}_t. \tag{2.5}$$

[Indeed, let $\mathbf{t}$ have edge lengths $\ell_i, 1 \le i \le n(\mathbf{t})$, and assume for convenience that e's ascendants are labeled as e_1 = founding ancestor, $e_2, \dots, e_j = e$ and that $\ell_1, \dots, \ell_j$ are the corresponding lifetimes. Recall that the infinitesimal probability that a rate one Poisson process on $(0, t]$ has exactly $j-1$ points located in $[\sum_{i=1}^m \ell_i, \sum_{i=1}^m \ell_i + d\ell_m)$, $m = 1, \dots, j-1$, is $\exp(-t)d\ell_1 \cdots d\ell_{j-1}$. Also, note that each of the 2^{j-1} possible left/right arrangements of the particles in the distinguished line of descent is equally likely. Since the subtrees founded by particles off the distinguished line of descent have distribution (2.1) and the remaining lifetime of e after time t is $\sum_{i=1}^j \ell_i - t$, we see from the construction of (T_t^*, V^*) that

$$\begin{aligned}
\mathrm{P}(T_t^* \in d\mathbf{t}, V^* = e) \;&=\; \exp(-t)d\ell_1 \cdots d\ell_{j-1}\, 2^{-(j-1)} \\
&\qquad \cdot \tfrac{1}{2} \exp\Big(- \Big(\sum_{i=1}^{j} \ell_i - t \Big) \Big) d\ell_j \prod_{i=j+1}^{n(\mathbf{t})} \tfrac{1}{2} \exp(-\ell_i) d\ell_i \\
&=\; 2^{-n(\mathbf{t})} \exp\Big(- \sum_{i=1}^{n(\mathbf{t})} \ell_i \Big)\, d\ell_1 \cdots d\ell_{n(\mathbf{t})} \;=\; \mathrm{P}(T \in d\mathbf{t}).
\end{aligned} \tag{2.6}$$

Summing over all $z_t(\mathbf{t})$ edges of $\mathbf{t}$ at height t proves (2.5). Relation (2.4) follows from the fact that the right-hand side of (2.6) does not depend on e.]

Remark. The construction above is a special case of the construction of the conditioned family tree of a branching particle system described in [4] and is easily extended to general offspring distributions with finite mean. It is intimately related to the (infinite) size-biased discrete-time Galton-Watson tree constructed in [16].

To define the degree of relationship that a particle in $\widehat{T}_t$ alive at time t has with the distinguished particle V we introduce the notion of a particle's (genealogical) distance from V. A particle is said to have distance s from V, if the particle's most recent common ancestor with V died at time $t - s$. Write $\widehat{Z}_t(s)$, $0 < s \le t$, for the number of particles at t with distance s from V. We may decompose the population at some fixed time t as

$$\widehat{Z}_t \;=\; 1 + \sum_{0 < s \le t} \widehat{Z}_t(s).$$

Keeping track of the particles' distances from V we represent the relationship structure (or genealogy) of the population at t by the random finite set

$$\Psi_t = \{\, (s, \widehat{Z}_t(s)) : 0 < s \le t,\; \widehat{Z}_t(s) > 0 \}. \tag{2.7}$$

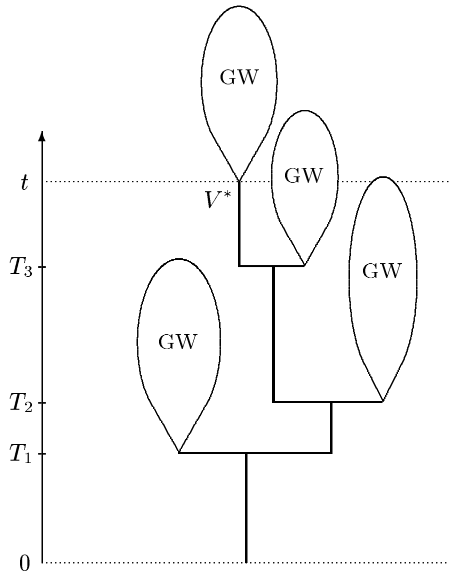

Figure 2: *Schematic representation of the random tree T_t^* and the random edge V^* of T_t^* at height t. The birth times $T_1 < T_2 < \ldots$ of V^* and its ascendants form a Poisson point process on $(0,t]$. The subtrees which grow to the sides of the distinguished line of descent and out at the top are independent copies of an ordinary critical binary Galton-Watson tree.*

We will identify the set Ψ_t and the simple point process $\sum_{(s,z)\,\in\Psi_t} \delta_{(s,z)}$, i.e. we do not distinguish between the random measure and its support. Note that we have replaced the complex structure of a random tree with an object that is much simpler to analyse and yet contains essential information on the relationship between particles.

The following statement is evident from the construction of (T_t^*, V^*) and the equality in law of $(\widehat{T}_t, V)$ and (T_t^*, V^*). We use notations $\mathbb{N} = \{1, 2, \ldots\}$ and $\mathbb{R}^+ = (0, \infty)$.

Lemma 2.1. *For any $t > 0$, the random set Ψ_t is a simple Poisson point process on $(0,t] \times \mathbb{N}$ with intensity measure*

$$\nu_t(ds \times \{k\}) = ds\, \mathrm{P}(Z_s = k). \tag{2.8}$$

Let R_s, $s > 0$, denote the restriction operator,

$$R_s A = A \cap (0,s] \times \mathbb{R}^+,\ \forall\, A \subset \mathbb{R}^+ \times \mathbb{R}^+. \tag{2.9}$$

Since the Ψ_t satisfy the consistency condition

$$\Psi_s \stackrel{d}{=} R_s \Psi_t,\ 0 < s \le t < \infty,$$

all Ψ_t, $t > 0$, can be constructed from one single Poisson point process Ψ_∞ on $\mathbb{R}^+ \times \mathbb{N}$ with intensity measure $\nu_\infty(ds \times \{k\}) = ds\, \mathrm{P}(Z_s = k)$.

A suitable rescaling of Ψ_t is to speed up time by t and assign mass t^{-1} to each particle. The asymptotic behavior of the rescaled point process

$$t^{-1}\Psi_t = \{(t^{-1}s, t^{-1}z) : (s,z) \in \Psi_t\}$$

is described by the following Proposition.

Proposition 2.2. *As $t \to \infty$,*

$$t^{-1}\Psi_t \xrightarrow{d} \Psi,$$

where Ψ *is a simple Poisson point process on* $(0,1]\times\mathbb{R}^+$ *with intensity*

$$\nu(ds\,dz)=\frac{4}{s^2}\exp\Big(-\frac{2z}{s}\Big)\,ds\,dz. \tag{2.10}$$

Here, $\xrightarrow{d}$ denotes convergence in distribution, which is just weak convergence of the joint distributions of $(t^{-1}\Psi_t(B_i),\ 1\le i\le n)$ for every finite family of relatively compact Borel subsets B_i of $(0,1]\times\mathbb{R}^+$.

Proof. Regard $t^{-1}\Psi_t$ as a point process on $(0,1]\times\mathbb{R}^+$ rather than on $(0,1]\times t^{-1}\mathbb{N}$. By the mapping theorem for Poisson point processes (see e.g. [14], p. 18), $t^{-1}\Psi_t$ is a simple Poisson point process with intensity measure $\tilde{\nu}_t=\nu_t(t\,\cdot)$. Hence, it is sufficient to show $\tilde{\nu}_t(B)\to\nu(B)$ for every relatively compact Borel subset B of $(0,1]\times\mathbb{R}^+$. The critical binary Galton-Watson process $(Z_s)_{s\ge 0}$ is a linear growth birth and death process whose distribution is explicitly known (see e.g. [7], p. 480). For any $s\ge 0$,

$$\mathrm{P}(Z_s>0) \;=\; \frac{2}{2+s}; \tag{2.11}$$

$$\mathrm{P}(Z_s\ge k\,|\,Z_s>0) \;=\; \Big(\frac{s}{2+s}\Big)^{k-1},\ k\ge 1. \tag{2.12}$$

Using Lemma 2.1 and relations (2.11) and (2.12), we obtain for any rectangle $C=[u_1,u_2]\times[z,\infty)\subset(0,1]\times\mathbb{R}^+$,

$$\begin{aligned}\lim_{t\to\infty}\tilde{\nu}_t(C) &= \lim_{t\to\infty}\int_{u_1t}^{u_2t} ds\,\mathrm{P}(Z_s\ge zt)\\ &= \lim_{t\to\infty}\int_{u_1}^{u_2} tds\,\mathrm{P}(Z_{st}>0)\,\mathrm{P}(Z_{st}\ge zt\,|\,Z_{st}>0)\\ &= \int_{u_1}^{u_2}\frac{2}{s}\exp\Big(-\frac{2z}{s}\Big)\,ds \;=\; \nu(C)\;<\infty.\end{aligned} \tag{2.13}$$

Since the rectangles C generate the σ-algebra of Borel subsets of $(0,1]\times\mathbb{R}^+$, we in particular have $\tilde{\nu}_t(B)\to\nu(B)$ for every relatively compact Borel subset B of $(0,1]\times\mathbb{R}^+$. This completes our proof of Proposition 2.2. $\square$

Remark. Note that to obtain (2.13) we have only used the asymptotic decay of the non-extinction probability $\mathrm{P}(Z_t>0)$ and the exponential limit law of Z_t conditioned on $Z_t>0$ following from (2.12). This asymptotic behavior holds for any critical Galton-Watson branching process with finite variance and is commonly referred to as Kolmogorov's asymptotic and Yaglom's exponential limit law (see e.g. Section I.9 in [3]).

The limiting point process Ψ from Proposition 2.2 satisfies the following scaling property which we state for further reference.

Lemma 2.3. *For any* $0<u\le 1$,

$$R_u\Psi\overset{d}{=}u\Psi. \tag{2.14}$$

Proof. The Poisson point process $R_u\Psi$ has intensity $1_{(0,u]}(s)\nu(ds\,dz)$ which coincides with $\nu(u^{-1}\cdot\,)$, the intensity of the Poisson point process $u\Psi$. $\square$

The following result describes the asymptotic behavior of $\widehat{Z}_t$ as $t \to \infty$.

Proposition 2.4. *The number of edges at height t in the size-biased Galton-Watson tree $\widehat{T}_t$ has a gamma limit law with shape parameter 2,*

$$\lim_{t\to\infty} \mathrm{P}(t^{-1}\widehat{Z}_t \le x) = \int_0^x 4y\, e^{-2y}\, dy, \;\; x \ge 0. \tag{2.15}$$

Proof. Let $g(A) = \sum_{(s,z)\in A} z$ for a countable set $A \subset \mathbb{R}^+ \times \mathbb{R}^+$. Note that $g(\Psi_t) = \widehat{Z}_t - 1$ is the total number of particles at time t other than the distinguished particle itself. We first show that $g(\Psi)$ has distribution as in (2.15). Write $p(A)$ for the projection of A onto its mass coordinate,

$$p(A) = \{z \mid \exists\, s \colon (s,z) \in A\}.$$

By the mapping theorem $p(\Psi)$ is a simple Poisson point process on $\mathbb{R}^+$ with intensity measure $\nu p = \nu((0,1] \times \cdot\,)$. Using a change of variables $u = s^{-1}$ we obtain

$$\begin{aligned} \nu p(dz) &= 4\int_0^1 \frac{ds}{s^2} \exp\left(-\frac{2z}{s}\right) dz \\ &= 4\int_1^\infty du \exp(-2zu)\, dz \;=\; 2\,\frac{e^{-2z}}{z}\, dz. \end{aligned} \tag{2.16}$$

It is well-known and easily verified that $\nu_\alpha(dz) = z^{-1}e^{-\alpha z}\, dz$ is the Lévy measure of the gamma process with scale parameter α (see e.g. [18], p. 80). Consequently, $g(\Psi) = \sum_{z\in p(\Psi)} z$ has the gamma distribution as in (2.15) with scale and shape parameter 2, respectively.

To establish the weak convergence

$$t^{-1}g(\Psi_t) \xrightarrow{d} g(\Psi) \;\text{ as } t \to \infty, \tag{2.17}$$

use (2.16) to deduce

$$\mathrm{E}g(\Psi) \;=\; \int_0^\infty z\,\nu p(dz) \;=\; 1$$

and note that Lemma 2.1 implies

$$\mathrm{E}g(\Psi_t) \;=\; \int_0^t ds\, \mathrm{E}Z_s \;=\; t.$$

Hence,

$$\lim_{t\to\infty} t^{-1}\mathrm{E}g(\Psi_t) = \mathrm{E}g(\Psi).$$

Assertion (2.17) now follows from Proposition 2.2 and the fact that $t^{-1}g(\Psi_t) = g(t^{-1}\Psi_t)$. This completes our proof of Proposition 2.4. □

Remarks. An alternate proof of Proposition 2.4 is as follows. Observe that $\widehat{Z}_t$ can also be obtained by size-biasing the law of Z_t conditioned on $Z_t > 0$. In view of Yaglom's exponential limit law (compare (2.12)) one simply has to justify that limiting and size-biasing procedures may be interchanged. (In general, a sufficient condition for the interchange is convergence of the means.) However, note that our proof of Proposition 2.4 explains the Lévy-Khinchin representation of the limiting gamma distribution.

Let T_t' be a critical binary Galton-Watson tree conditioned on non-extinction at time t. The genealogy of T_t' with the left-most particle at t being distinguished has a point process representation analogous to (2.7) (see [10]). This point process Ψ_t', say, is a Poisson point process with intensity $\nu_t' = \frac{1}{2}\nu_t$, i.e., the point process Ψ_t representing the size-biased critical Galton-Watson tree is the superposition of two independent copies of Ψ_t', the point process representing the conditioned Galton-Watson tree.

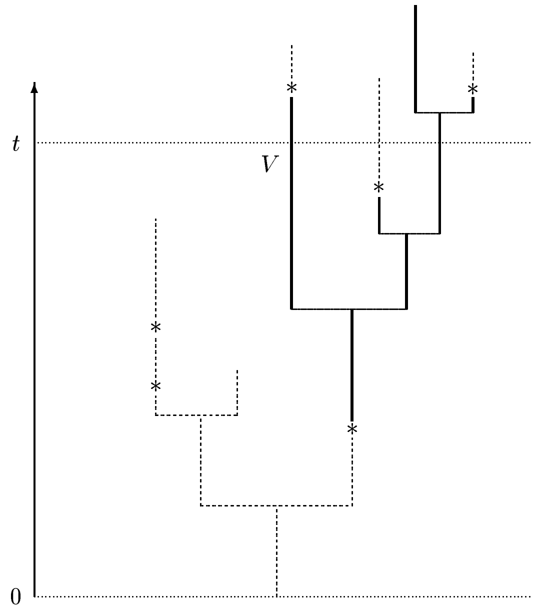

Figure 3: *A realization of the multitype size-biased Galton-Watson tree with the distinguished particle V. The asterisks on the edges indicate mutation events. The solid subtree is the component with the type of the distinguished particle V at t.*

3 Pruning the size-biased Galton-Watson tree

We now suppose that the particles may have different types. While alive particles independently mutate at rate $\mu > 0$, changing to a new type which was not present so far. (We may take the interval $(0, 1)$ for the set of possible types.) At birth particles inherit the type of their parent. Mutation and branching mechanisms are assumed independent. To indicate mutation events we put marks on the edges of the tree. If a particle mutates at time s, then a mark is put at height s on the edge corresponding to that particle. (The term "marked tree" is sometimes used for what we call a continuous-time Galton-Watson tree or, more general, for a tree-indexed Markov process. Here a mark is just meant to be some symbol.) An edge of length ℓ has a Poisson number of mutation marks with mean $\mu\ell$. The positions of the marks are independent and uniformly distributed on the edge. If we cut up the edges of the tree at mutation marks, then the tree falls to components of identical type (see Figure 3). For an account of pruned discrete-time Galton-Watson trees, see [2].

Our aim is to explore the genealogy of the subpopulation with the type of the distinguished particle V at t which is described by the component of the tree containing V (respectively, the piece of the edge V at height t, if the edge is cut up). As in the single type case we suppress some of the information on the relationship between particles and represent the *pruned* size-biased Galton-Watson tree with a distinguished particle by the random set

$$\Psi_t^\mu = \{(s, \widehat{Z}_t^\mu(s)) : 0 < s \le t,\ \widehat{Z}_t^\mu(s) > 0\}, \tag{3.1}$$

where $\widehat{Z}_t^\mu(s)$ is the number of particles at time t which at t have the same type as the distinguished particle V and have distance s from V. If we pass to the limit $t \to \infty$ in (3.1) we obtain a weak limiting point process Ψ_∞^μ, say. (Indeed, recall from Section 2 that the law of a subtree growing out of the distinguished line of descent at height $t - s$ does not depend on t. Hence, neither does the law of $\widehat{Z}_t^\mu(s)$, $0 < s \le t$.) We will refer to the point process Ψ_∞^μ as the equilibrium genealogy. Its law is described by the following lemma.

Lemma 3.1. *For any* $\mu > 0$,

$$\Psi^\mu_\infty \stackrel{d}{=} R_{T_\mu} \Phi_\mu,$$

where T_μ *is an exponential random variable with rate* μ, *independent of the simple Poisson point process* Φ_μ *on* $\mathbb{R}^+ \times \mathbb{N}$ *with intensity measure*

$$\rho_\mu(ds \times \{k\}) = ds\, \mathrm{P}(Z^\mu_s = k). \tag{3.2}$$

Here, $(Z^\mu_s)_{s \geq 0}$ *is a subcritical binary Galton-Watson branching process with* $p^\mu_2 = 1 - p^\mu_0 = \frac{1}{2}(1+\mu)^{-1}$ *and lifetime rates* $1 + \mu$.

Proof. By Lemma 2.1 independent critical binary Galton-Watson trees grow out of the distinguished line of descent at rate one. We claim that if the edges of a critical Galton-Watson tree are cut up at rate μ, then the component containing the root is a subcritical binary Galton-Watson tree with the parameters stated in the lemma. Indeed, note that the length of an edge in the pruned tree is the minimum of two independent exponential random variables X_1 and X_2 with rates 1 (original lifetime) and μ (mutation), respectively. The minimum of X_1 and X_2 has exponential distribution with rate $1+\mu$. For the offspring numbers note that a particle in the pruned tree has two children iff it had so before the pruning procedure and if the particle did not mutate. The first event has probability $p_2 = 1/2$, the second event independently has probability $\mathrm{P}(X_1 < X_2) = (1+\mu)^{-1}$. Finally, note that the particles in some pruned subtree have the same type as the distinguished particle V iff their distance from V is less than the time since the most recent mutation in the distinguished line of descent. This mutation event has occurred an independent exponential time T_μ ago. □

For $\mu > 0$ let $f_\mu : \mathbb{R}^+ \times \mathbb{R}^+ \to (0,1) \times \mathbb{R}^+$ be defined as

$$f_\mu(s,z) = (1 - e^{-\mu s},\ \mu z).$$

The following theorem describes the asymptotic behavior of the rescaled equilibrium genealogy $f_\mu(\Psi^\mu_\infty)$ as the mutation rate μ goes to zero.

Theorem 3.2. *As* $\mu \to 0$,

$$f_\mu(\Psi^\mu_\infty) \xrightarrow{d} R_U \Psi,$$

where R_u, $0 < u \leq 1$, *is the restriction operator introduced in (2.9),* Ψ *is the limiting Poisson point process from Proposition 2.2, and* U *has uniform distribution on* $(0,1)$, *independent of* Ψ.

Proof. In much the same way as in the proof of Proposition 2.2 we first show that

$$f_\mu(\Phi_\mu) \xrightarrow{d} \Psi \quad \text{as} \quad \mu \to 0. \tag{3.3}$$

By the mapping theorem $f_\mu(\Phi_\mu)$ is a simple Poisson point process on $(0,1] \times \mathbb{R}^+$ with intensity measure $\rho_\mu f_\mu = \rho_\mu(f^{-1}_\mu(\cdot))$, where f^{-1}_μ is the inverse of f_μ. To prove (3.3) it is sufficient to show $\rho_\mu f_\mu(B) \to \nu(B)$ for every relatively compact Borel subset B of $(0,1] \times \mathbb{R}^+$. For any $s \geq 0$ the distribution of Z^μ_s is given by (see e.g. [7], p. 480)

$$\begin{aligned} \mathrm{P}(Z^\mu_s > 0) &= \frac{2\mu e^{-\mu s}}{2\mu + 1 - e^{-\mu s}}; \\ \mathrm{P}(Z^\mu_s \geq k \mid Z^\mu_s > 0) &= \left(\frac{1 - e^{-\mu s}}{2\mu + 1 - e^{-\mu s}}\right)^{k-1}, \ k \geq 1. \end{aligned}$$

Hence, if we denote by h_μ^{-1} the inverse of $h_\mu(s) = 1 - e^{-\mu s}$, $s \geq 0$, and let $\lfloor \alpha \rfloor$ denote the integer part of $\alpha \in \mathbb{R}^+$, then

$$\mathrm{P}(Z^\mu_{h_\mu^{-1}(s)} > 0) \quad = \quad \frac{2\mu(1-s)}{2\mu + s}\,; \tag{3.4}$$

$$\mathrm{P}(Z^\mu_{h_\mu^{-1}(s)} \geq \mu^{-1}z \,|\, Z^\mu_{h_\mu^{-1}(s)} > 0) \quad = \quad \Big(\frac{s}{2\mu + s}\Big)^{\lfloor \mu^{-1}z \rfloor - 1}. \tag{3.5}$$

Using (3.4), (3.5) and the fact that $h'_\mu(s) = \mu\,(1 - h_\mu(s))$, we obtain for any rectangle $C = [u_1, u_2] \times [z, \infty) \subset \mathbb{R}^+ \times \mathbb{R}^+$

$$\begin{aligned}
\lim_{\mu \to 0} \rho_\mu f_\mu(C) \quad &= \quad \lim_{\mu\to 0} \int_{u_1}^{u_2} dh_\mu^{-1}(s)\,\mathrm{P}(Z^\mu_{h_\mu^{-1}(s)} \geq \mu^{-1}z) \\
&= \quad \lim_{\mu\to 0} \int_{u_1}^{u_2} \frac{ds}{h'_\mu(h_\mu^{-1}(s))}\ \mathrm{P}(Z^\mu_{h_\mu^{-1}(s)} > 0) \\
&\qquad\qquad \cdot\,\mathrm{P}(Z^\mu_{h_\mu^{-1}(s)} \geq \mu^{-1}z \,|\, Z^\mu_{h_\mu^{-1}(s)} > 0) \\
&= \quad \lim_{\mu\to 0} \int_{u_1}^{u_2} \frac{ds}{\mu(1-s)}\ \frac{2\mu\,(1-s)}{2\mu + s}\ \Big(1 - \frac{2\mu}{2\mu+s}\Big)^{\lfloor \mu^{-1}z \rfloor - 1} \\
&= \quad \int_{u_1}^{u_2} \frac{2ds}{s}\ \exp\Big(-\frac{2z}{s}\Big) \quad = \quad \nu(C) \;< \infty,
\end{aligned}$$

which establishes assertion (3.3). Now observe that

$$f_\mu \circ R_t = R_{h_\mu(t)} \circ f_\mu, \;\; \forall\, \mu, t > 0,$$

and that $h_\mu(T_\mu) \stackrel{d}{=} U$ since h_μ is the distribution function of T_μ. Hence, Lemma 3.1 implies

$$f_\mu(\Psi^\mu_\infty) \stackrel{d}{=} R_{h_\mu(T_\mu)} f_\mu(\Phi_\mu) \stackrel{d}{=} R_U f_\mu(\Phi_\mu),$$

where U is independent of Φ_μ. In view of (3.3) the claim of Theorem 3.2 now follows by passing to the limit $\mu \to 0$. $\square$

Let $\widehat{Z}^\mu_\infty = 1 + g(\Psi^\mu_\infty)$ denote the number of particles with the distinguished type in equilibrium. Our final result describes the asymptotic behavior of $\widehat{Z}^\mu_\infty$ as the mutation rate μ goes to zero.

Theorem 3.3. *The population size of the distinguished type in equilibrium has an exponential limit law,*

$$\lim_{\mu\to 0} \mathrm{P}(2\mu \widehat{Z}^\mu_\infty \geq x) \;=\; \exp(-x), \;\; x \geq 0.$$

Proof. Note that $\mu g(\Psi^\mu_\infty) \stackrel{a.s.}{=} g(f_\mu(\Psi^\mu_\infty))$ for any $\mu > 0$. Hence, Theorem 3.2 implies

$$\mu \widehat{Z}^\mu_\infty \stackrel{d}{\to} g(R_U \Psi) \;\; \text{as} \;\; \mu \to 0,$$

if we show

$$\lim_{\mu\to 0} \mu\,\mathrm{E} g(\Psi^\mu_\infty) = \mathrm{E} g(R_U \Psi). \tag{3.6}$$

By the scaling property of Ψ (Lemma 2.3) we have

$$g(R_U\Psi) \stackrel{d}{=} g(U\,\Psi) \stackrel{a.s.}{=} U g(\Psi),$$

that is $g(R_U\Psi)$ is an independent uniform contraction of a gamma distributed random variable with shape and scale parameter 2 (Proposition 2.4). Consequently, $g(R_U\Psi)$ has exponential distribution with rate two (see this e.g. by considering the interval covering the origin in a homogeneous Poisson process on $\mathbb{R}$). To verify (3.6) recall that $\mathrm{E}Z_s^\mu = e^{-\mu s}$ (see e.g. [7], p.457). Hence, Lemma 3.1 implies

$$\mathrm{E}g(\Psi_\infty^\mu) = \int_0^\infty \mathrm{P}(T_\mu \in dt) \int_0^t ds\, \mathrm{E}Z_s^\mu = (2\mu)^{-1} = \mu^{-1}\mathrm{E}g(R_U\Psi),$$

which completes our proof of Theorem 3.3. □

Remark. It is shown in [5] that the genealogy of the cluster at the origin in the multitype voter model is asymptotically described by the size-biased Galton-Watson tree. Combining this result with Theorem 3.3 explains a classical result by Sawyer [19] that the size of the type at the origin in the multitype voter model with mutation approaches an exponential limit distribution.

Acknowledgment. The author is grateful to an anonymous referee for helpful suggestions.

References

[1] Aldous, D. (1993) The continuum random tree III, *Ann. Probab.* **21**, 248-289.

[2] Aldous, D. and Pitman, J. (1998) Tree-valued Markov chains derived from Galton-Watson processes, *Ann. Inst. H. Poincaré* **34**, 637-686.

[3] Athreya, K. B. and Ney, P. (1972) *Branching Processes*, Springer, New York.

[4] Chauvin, B., Rouault, A. and Wakolbinger, A. (1991) Growing conditioned trees, *Stoch. Proc. Appl.* **39**, 117-130.

[5] Cox, J.T. and Geiger, J. (2000) The genealogy of a cluster in the multitype voter model, to appear in *Ann. Probab.*.

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

[7] Feller, W. (1968) *An Introduction to Probability Theory and Its Applications I*, 3rd ed., Wiley, New York.

[8] Fleischmann, K. and Siegmund-Schultze, R. (1977) The structure of reduced critical Galton-Watson processes, *Math. Nachr.* **79**, 233-241.

[9] Geiger, J. (1996) Size-biased and conditioned random splitting trees, *Stoch. Proc. Appl.* **65**, 187-207.

[10] Geiger, J. (1999) Elementary new proofs of classical limit theorems for Galton-Watson processes, *J. Appl. Prob.* **36**, 301–309.

[11] Jagers, P. and Nerman, O. (1984) The growth and composition of branching populations, *Adv. Appl. Prob.* **16**, 221–259.

[12] Joffe, A. and Waugh, W. A. O'N (1982) Exact distributions of kin numbers in Galton–Watson processes, *J. Appl. Prob.* **19**, 767–775.

[13] Kallenberg, O. (1977) Stability of critical cluster fields, *Math. Nachr.* **77**, 7–43.

[14] Kingman, J.F.C. (1993) *Poisson processes*, Clarendon Press, Oxford.

[15] Le Gall, J. F. (1991) Brownian excursions, trees and measure-valued branching processes, *Ann. Probab.* **19**, 1399-1439.

[16] LYONS, R., PEMANTLE, R. AND PERES, Y. (1995) Conceptual proofs of $L \log L$ criteria for mean behavior of branching processes, *Ann. Probab.* **23**, 1125-1138.

[17] NEVEU, J. (1986) Arbres et processus de Galton-Watson, *Ann. Inst. H. Poincaré* **22**, 199-207.

[18] ROGERS, L.C.G. AND WILLIAMS, D. (1994) *Diffusions, Markov Processes, and Martingales, Vol. 1: Foundations*, 2nd ed., Wiley, Chichester.

[19] SAWYER, S. (1979) A limit theorem for patch sizes in a selectively-neutral migration model, *J. Appl. Prob.* **16**, 482-495.