Growth of Galton-Watson trees: immigration and lifetimes

Xiao'ou Cao* Matthias Winkel†

April 20, 2010

Abstract

We study certain consistent families $(F_\lambda)_{\lambda\geq 0}$ of Galton-Watson forests with lifetimes as edge lengths and/or immigrants as progenitors of the trees in F_λ. Specifically, consistency here refers to the property that for each $\mu \leq \lambda$, the forest F_μ has the same distribution as the subforest of F_λ spanned by the black leaves in a Bernoulli leaf colouring, where each leaf of F_λ is coloured in black independently with probability μ/λ. The case of exponentially distributed lifetimes and no immigration was studied by Duquesne and Winkel and related to the genealogy of Markovian continuous-state branching processes. We characterise here such families in the framework of arbitrary lifetime distributions and immigration according to a renewal process, related to Sagitov's (non-Markovian) generalisation of continuous-state branching renewal processes, and similar processes with immigration.

AMS 2000 subject classifications: 60J80.

Keywords: Galton-Watson process, continuous-state branching process, random tree, immigration, age-dependent branching, geometric infinite divisibility, backbone decomposition

1 Introduction

Galton-Watson branching processes are a classical model for the evolution of population sizes, see e.g. [1, 22]. More specifically, there is an interest in the underlying genealogical trees. In the most basic model, there is a single *progenitor* that produces i children with probability $q(i)$ for some *offspring distribution* q on $\mathbb{N}_0 = \{0, 1, 2, \ldots\}$; recursively, each individual in the population produces children independently and according to the same distribution q. We represent this model by a *graph-theoretic tree* rooted at the progenitor, where each individual is a vertex and the *parent-child relation* specifies edges $v \to w$ between parent v and child w. Vertices related to just their parent vertex and to no child vertices are called *leaves*. More precisely, we will follow Neveu [36] to distinguish individuals (see Section 2.1). We will consider in this paper the following well-known and/or natural variants of Galton-Watson trees (see e.g. Jagers [24]):

- GW(q)-trees as the most basic model just described;
- GW(q,κ)-trees as GW(q)-trees, where each individual is marked by an independent *lifetime* with distribution κ on $(0,\infty)$; this includes the case $\kappa = \mathrm{Exp}(c)$ of the exponential lifetime distribution with rate parameter $c \in (0,\infty)$;
- GW(q,κ,β)-bushes as *bushes* (finite sequences) of a random number N of GW(q,κ)-trees, where N is Poisson distributed with parameter $\beta \in [0,\infty)$, in shorthand: $N \sim \mathrm{Poi}(\beta)$;
- GWI(q,κ,η,χ)-forests as *forests* (point processes on the *forest floor* $[0,\infty)$) of independent bushes of N_i GW(q,κ)-trees at the locations S_i of a renewal process with inter-renewal distribution χ on $(0,\infty)$, where each N_i, $i \geq 1$, has distribution η on $\mathbb{N} = \{1, 2, \ldots\}$.

*University of Oxford; email: cao@stats.ox.ac.uk

†University of Oxford; email: winkel@stats.ox.ac.uk

With the common interpretation that individuals give birth only at the end of their life and that renewal locations are *immigration times*, all but the first model give rise to continuous-time processes counting the number Y_t of individuals in the population at time $t \geq 0$. In these continuous-time models it is natural to take $q(1) = 0$, since an individual producing a single child at its death time can be viewed as continuing to live instead of being replaced by its child.

Reduction by Bernoulli leaf colouring was studied in [11] and reads as follows in our setting:

- independently mark each leaf of a tree T (with lifetimes), or of a bush $B = (T_{(1)}, \ldots, T_{(N)})$ or of a forest $F = (B(t), t \geq 0)$ by a *black* colour mark with probability $1 - p \in (0, 1)$, by a *red* colour mark otherwise; for this to be non-trivial, let $q(0) > 0$, as T then has leaves;
- if there are any black leaves, consider, as illustrated in Figure 1,
  - the *p-reduced subtree* $T_{\text{sub}}^{p-\text{rdc}}$ of T as the subtree of T spanned by the root and the black leaves (with lifetime marks inherited);
  - and the *p-reduced tree* $T^{p-\text{rdc}}$ derived from $T_{\text{sub}}^{p-\text{rdc}}$ by identifying vertices via the equivalence relation generated by $v \equiv w$ for vertices in $T_{\text{sub}}^{p-\text{rdc}}$ if $v \to w$ and w is the only child of v in $T_{\text{sub}}^{p-\text{rdc}}$ (marked by the sum of lifetimes in each equivalence class);
  - or the *p-reduced bush* $B^{p-\text{rdc}} = (T_{(I_1)}^{p-\text{rdc}}, \ldots, T_{(I_{N^{p-\text{rdc}}})}^{p-\text{rdc}})$ as the p-reduced trees associated with the subsequence $(I_1, \ldots, I_{N^{p-\text{rdc}}})$ of trees in B that have black leaves;
  - or the *p-reduced forest* $F^{p-\text{rdc}} = (B^{p-\text{rdc}}(t), t \geq 0)$ of p-reduced bushes.

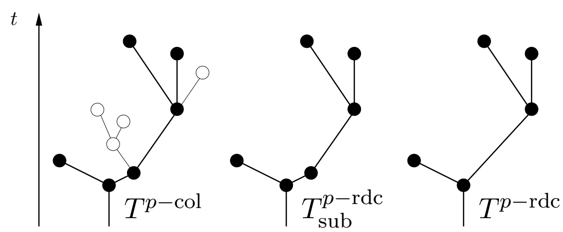

Figure 1: Black vertices are represented by solid circles and red ones by circle lines.

It is easily seen that if T is a Galton-Watson tree, then given that there are any black leaves, the p-reduced subtree $T_{\text{sub}}^{p-\text{rdc}}$ and the p-reduced tree $T^{p-\text{rdc}}$ are also Galton-Watson trees [11]. By standard thinning properties of Poisson point processes, the p-reduced bush $B^{p-\text{rdc}}$ associated with a GW(q, κ, β)-bush B is also a Galton-Watson bush. We refer to the offspring distribution $q^{p-\text{rdc}}$, the lifetime distribution $\kappa^{p-\text{rdc}}$ and the Poisson parameter $\beta^{p-\text{rdc}}$ of $B^{p-\text{rdc}}$ as the *p-reduced triplet* $(q^{p-\text{rdc}}, \kappa^{p-\text{rdc}}, \beta^{p-\text{rdc}})$ *associated with* (q, κ, β), similarly for forests etc. It is not hard to find offspring distributions $\widehat{q}$ that do *not* arise as p-reduced offspring distributions for any q (e.g. $\widehat{q}(0) = \widehat{q}(3) = 1/2$). More precisely, [11] obtained the following characterisation.

Theorem 1 (Theorem 4.2 of [11]). *For an offspring distribution q, the following are equivalent:*

(i) *There is a family $(q_\lambda)_{\lambda \geq 0}$ of offspring distributions with $q_1 = q$ such that q_μ is the $(1-\mu/\lambda)$-reduced offspring distribution associated with q_λ, for all $0 \leq \mu < \lambda < \infty$.*

(ii) *The generating function φ_q of q satisfies*

$$\varphi_q(s) = \sum_{i=0}^{\infty} s^i q(i) = s + \widetilde{\psi}(1 - s), \qquad 0 \leq s \leq 1, \tag{1}$$

where for some $\widetilde{b} \in \mathbb{R}$, $\widetilde{a} \geq 0$ and a measure $\widetilde{\Pi}$ on $(0, \infty)$ with $\int_{(0,\infty)} (1 \wedge x^2) \widetilde{\Pi}(dx) < \infty$,

$$\widetilde{\psi}(r) = \widetilde{b}r + \widetilde{a}r^2 + \int_{(0,\infty)} (e^{-rx} - 1 + rx\mathbf{1}_{\{x<1\}}) \widetilde{\Pi}(dx), \qquad r \geq 0. \tag{2}$$

In the setting of (i) *and* (ii), *a consistent family* $(B_\lambda)_{\lambda\ge 0}$ *of* $\mathrm{GW}(q_\lambda, \mathrm{Exp}(c_\lambda), \beta_\lambda)$*-bushes can be constructed such that* $(B_\mu, B_\lambda) \overset{(\mathrm{d})}{=} (B_\lambda^{(1-\mu/\lambda)-\mathrm{rdc}}, B_\lambda)$ *for all* $0 \le \mu < \lambda < \infty$. *For each* $c = c_1 \in (0,\infty)$ *and* $\beta = \beta_1 \in (0,\infty)$, *the family* $(q_\lambda, c_\lambda, \beta_\lambda)_{\lambda\ge 0}$ *is unique,* $(q_\lambda)_{\lambda\ge 0}$ *does not depend on* (c,β), *while* $(c_\lambda)_{\lambda\ge 0}$ *depends on* q *but not on* β *and* $(\beta_\lambda)_{\lambda\ge 0}$ *on* q *but not on* c.

In [11], this result is a key step in the construction of Lévy trees as genealogies of Markovian continuous-state branching processes with branching mechanism ψ, where ψ is a linear transformation of $\widetilde{\psi}$ that we recall in Section 2.3.1. In the present paper we establish characterisations analogous to Theorem 1 for the other variants of Galton-Watson trees, bushes and forests.

Theorem 2. *For a pair* (q,κ) *of offspring and lifetime distributions, the following are equivalent:*

(i) *There are families* $(q_\lambda, \kappa_\lambda)_{\lambda\ge 0}$ *with* $q_1 = q$ *and* $\kappa_1 = \kappa$ *such that* (q_μ, κ_μ) *is the* $(1-\mu/\lambda)$*-reduced pair associated with* $(q_\lambda, \kappa_\lambda)$, *for all* $0 \le \mu < \lambda < \infty$.

(ii) *The generating function* φ_q *of* q *satisfies* $\varphi_q(s) = s + \widetilde{\psi}(1-s)$, *where* $\widetilde{\psi}$ *is of the form* (2). *Moreover,* κ *is geometrically divisible in that there is a family* $(X_\alpha^{(j)}, j \ge 1)$ *of independent identically distributed random variables and* $G^{(\alpha)} \sim \mathrm{Geo}(\alpha)$ *independent geometric with parameter* α, *i.e.* $\mathbb{P}(G^{(\alpha)} = k) = \alpha(1-\alpha)^{k-1}$, $k \in \mathbb{N}$, *such that* $X_\alpha^{(1)} + \cdots + X_\alpha^{(G^{(\alpha)})} \sim \kappa$

- *for all* $\alpha > 1/\widetilde{\psi}'(\infty)$ *if* $\widetilde{\psi}'(\infty) < \infty$, *where* $\widetilde{\psi}'(\infty)$ *means* $\lim_{r\to\infty} \widetilde{\psi}'(r)$;
- *for all* $\alpha > 0$ *if* $\widetilde{\psi}'(\infty) = \infty$.

In the setting of (i) *and* (ii), *a consistent family* $(B_\lambda)_{\lambda\ge 0}$ *of* $\mathrm{GW}(q_\lambda, \kappa_\lambda, \beta_\lambda)$*-bushes can be constructed such that* $(B_\mu, B_\lambda) \overset{(\mathrm{d})}{=} (B_\lambda^{(1-\mu/\lambda)-\mathrm{rdc}}, B_\lambda)$ *for all* $0 \le \mu < \lambda < \infty$. *For each* $\beta = \beta_1 \in (0,\infty)$, *the family* $(q_\lambda, \kappa_\lambda, \beta_\lambda)_{\lambda\ge 0}$ *is unique,* $(q_\lambda)_{\lambda\ge 0}$ *does not depend on* (κ,β) *while* $(\kappa_\lambda)_{\lambda\ge 0}$ *depends on* q *but not on* β *and* $(\beta_\lambda)_{\lambda\ge 0}$ *on* q *but not on* κ.

The requirement on κ set in the second bullet point is referred to as geometric infinite divisibility in the literature, see [30], also Section 2.2 here. Since the distribution $\kappa = \mathrm{Exp}(c)$ is geometrically infinitely divisible, Theorem 2 is an extension of Theorem 1.

Theorem 3. *For a pair* (q,η) *of offspring and immigration distributions, the following are equivalent:*

(i) *There are families* $(q_\lambda, \eta_\lambda)_{\lambda\ge 0}$ *with* $q_1 = q$ *and* $\eta_1 = \eta$ *such that* (q_μ, η_μ) *is the* $(1-\mu/\lambda)$*-reduced pair associated with* $(q_\lambda, \eta_\lambda)$, *for all* $0 \le \mu < \lambda < \infty$.

(ii) *The generating function* φ_q *of* q *satisfies* $\varphi_q(s) = s + \widetilde{\psi}(1-s)$, *where* $\widetilde{\psi}$ *is of the form* (2). *Moreover, the generating function* φ_η *of* η *satisfies*

$$\varphi_\eta(s) = \sum_{i=1}^{\infty} s^i \eta(i) = 1 - \widetilde{\phi}(1-s), \qquad 0 \le s \le 1,$$

where for some $\widetilde{d} \in \mathbb{R}$, *and a measure* $\widetilde{\Lambda}$ *on* $(0,\infty)$ *with* $\int_{(0,\infty)} (1 \wedge x)\widetilde{\Lambda}(dx) < \infty$,

$$\widetilde{\phi}(r) = \widetilde{d}r + \int_{(0,\infty)} (1 - e^{-rx})\widetilde{\Lambda}(dx), \qquad r \ge 0. \tag{3}$$

In the setting of (i) *and* (ii), *a consistent family* $(F_\lambda)_{\lambda\ge 0}$ *of* $\mathrm{GWI}(q_\lambda, \mathrm{Exp}(c_\lambda), \eta_\lambda, \mathrm{Exp}(h_\lambda))$*-forests can be constructed such that* $(F_\mu, F_\lambda) \overset{(\mathrm{d})}{=} (F_\lambda^{(1-\mu/\lambda)-\mathrm{rdc}}, F_\lambda)$ *for all* $0 \le \mu < \lambda < \infty$. *For each* $c = c_1 \in (0,\infty)$ *and* $h = h_1 \in (0,\infty)$, *the family* $(q_\lambda, c_\lambda, \eta_\lambda, h_\lambda)_{\lambda\ge 0}$ *is unique,* $(q_\lambda)_{\lambda\ge 0}$ *does not depend on* (c,η,h), *while* $(c_\lambda)_{\lambda\ge 0}$ *depends on* q *but not on* (η,h), $(\eta_\lambda)_{\lambda\ge 0}$ *depends on* q *but not on* (c,h) *and* $(h_\lambda)_{\lambda\ge 0}$ *depends on* (q,η) *but not on* c.

The binary special case with single immigrants, where for some $\theta \geq 0$ and all $\lambda \geq 0$

$$q_\lambda(0) = \frac{1}{2} + \frac{1}{2\sqrt{\theta^2 + 2\lambda}}, \quad q_\lambda(2) = \frac{1}{2} - \frac{1}{2\sqrt{\theta^2 + 2\lambda}}, \qquad c_\lambda = \sqrt{\theta^2 + 2\lambda},$$

$$\eta_\lambda(1) = 1, \qquad h_\lambda = \sqrt{\theta^2 + 2\lambda} - \theta,$$

leads to the setting of [38], where $(F_\lambda)_{\lambda \geq 0}$ was shown to have independent "increments" expressed by a composition rule, and to converge to the forest in Brownian motion with drift $-\theta$.

Theorems 2 and 3 describe in the same way the genealogy of associated continuous-state branching processes (CSBP) as Theorem 1. Specifically, for Theorem 2 the continuous-state processes are Sagitov's age-dependent $\mathrm{CSBP}(K, \psi)$ based on a branching mechanism ψ and the distribution of a local time process K, i.e. either an inverse subordinator or an inverse increasing random walk, see [27, 40] and Section 3.3.2 here; for Theorem 3, they are CSBP with immigration, $\mathrm{CBI}(\psi, \phi)$, where ϕ is an immigration mechanism, see [29, 31] and Section 2.2 here.

Proposition 4. *Let $(Z_t^\lambda, t \geq 0)$ be the population size process in the setting of Theorem 2. If $\widetilde{\psi}'(0) > -\infty$, then*

$$\frac{Z_t^\lambda}{\psi^{-1}(\lambda)} \to Z_t \qquad \text{almost surely as } \lambda \to \infty, \text{ for all } t \geq 0,$$

where $(Z_t, t \geq 0)$ is a $\mathrm{CSBP}(K, \psi)$ with $Z_0 = \beta$, for some ψ linear transformation of $\widetilde{\psi}$ and $K = (K_s, s \geq 0)$ such that $\inf\{s \geq 0 \colon K_s > V_\lambda\} \sim \kappa_\lambda$ for $V_\lambda \sim \mathrm{Exp}(c_\lambda)$ with c_λ as in Theorem 1.

Proposition 5. *Let $(Y_t^\lambda, t \geq 0)$ be the population size process in the setting of Theorem 3. Then*

$$\frac{Y_t^\lambda}{\psi^{-1}(\lambda)} \to Y_t \qquad \text{in distribution as } \lambda \to \infty, \text{ for all } t \geq 0,$$

where $(Y_t, t \geq 0)$ is a $\mathrm{CBI}(\psi, \phi)$ with $Y_0 = 0$, for ψ and ϕ linear transformations of $\widetilde{\psi}$ and $\widetilde{\phi}$. If furthermore $\psi'(0) > -\infty$ and $\phi'(0) < \infty$, then the convergence holds in the almost sure sense.

These convergence results should be seen in the context of the large literature on space-time scaling limits of branching processes in discrete or continuous time, see [10, 12, 27, 29, 34, 37, 40]. Convergence in distribution holds under much weaker assumptions on the families $(q_\lambda, \kappa_\lambda, \beta_\lambda)_{\lambda \geq 0}$ or $(q_\lambda, c_\lambda, \eta_\lambda, h_\lambda)_{\lambda \geq 0}$ and invariance principles in varying degrees of generality have been obtained. It is also well-known that the convergence in distribution at a fixed time for Markovian branching processes implies the convergence in distribution of the whole process in the Skorohod sense of convergence of right-continuous functions with left limits. In [12], joint convergence of processes and their genealogical trees is shown, also for a wider class of families $(B_\lambda)_{\lambda \geq 0}$ suitably converging to bushes of Lévy trees. The main contribution of the present work is to provide almost sure approximations of more general classes of continuous-state processes and consistent families of trees that contain full information about the genealogy of the population of the limiting continuous-state process, which is not contained in the limiting process itself nor in the approximating discrete-state branching processes.

The structure of this paper is as follows. In Section 2 we formally set up the framework in which we represent trees, we recall preliminaries from Duquesne and Winkel [11] and develop a bit further some aspects that readily transfer and serve in the more general context here. We also provide some background about continuous-state branching processes with immigration, and about geometric infinite divisibility. Section 3 presents the theory around Theorem 2 and Proposition 4, while Section 4 deals with Theorem 3 and Proposition 5. In each setting we provide explicit formulas for offspring distributions, lifetime distributions and immigration distributions as appropriate; we also provide explicit reconstruction procedures that reverse the reduction for the consistent families of bushes and forests and establish connections with backbone decompositions (Theorem 22) and Lévy trees. We finally deduce generalisations combining lifetimes and immigration.

2 Preliminaries

2.1 Discrete trees with edge lengths and colour marks

2.1.1 Discrete trees and the Galton-Watson branching property

Following Neveu [36], Chauvin [8] and others, we let

$$\mathbb{U} = \bigcup_{n\geq 0} \mathbb{N}^n = \{\emptyset, 1, 2, \ldots, 11, 12, \ldots, 21, 22, \ldots, 111, 112, \ldots\}$$

be the set of *integer words*, where $\mathbb{N} = \{1, 2, 3, \ldots\}$ and where $\mathbb{N}^0 = \{\emptyset\}$ has the *empty word* $\emptyset$ as its only element. For $u = u_1u_2\cdots u_n \in \mathbb{U}$, and $v = v_1v_2\cdots v_m \in \mathbb{U}$, we denote by $|u|$ the *length* of u, e.g. $|u_1u_2\cdots u_n| = n$, and by $uv = u_1u_2\cdots u_nv_1v_2\cdots v_m$ the *concatenation* of words in $\mathbb{U}$.

Definition 6. A subset $\mathbf{t} \subset \mathbb{U}$ is called a *tree* if

- $\emptyset \in \mathbf{t}$; we refer to $\emptyset$ as the *progenitor* of $\mathbf{t}$;
- for all $u \in \mathbb{U}$ and $j \in \mathbb{N}$ with $uj \in \mathbf{t}$, we have $u \in \mathbf{t}$; we refer to u as the *parent* of uj;
- for all $u \in \mathbf{t}$, there exists $\nu_u(\mathbf{t}) \in \mathbb{N}_0 = \{0, 1, 2, \ldots\}$ such that $uj \in \mathbf{t} \Leftrightarrow 1 \leq j \leq \nu_u(\mathbf{t})$; we refer to $\nu_u(\mathbf{t})$ as the *number of children* of u.

We refer to the length $|u|$ of a word $u \in \mathbf{t}$ as the *generation* of the *individual* u in the *genealogical tree* $\mathbf{t}$. An element $u \in \mathbf{t}$ is called a *leaf* of $\mathbf{t}$ if and only if $\nu_u(\mathbf{t}) = 0$. We consider the lexicographical order $\leq$ on $\mathbb{U}$ and its restriction to $\mathbf{t}$ as the canonical *total order*. For $u, v \in \mathbb{U}$, we write $u \preceq uv$, defining a *partial order* on $\mathbb{U}$, whose restriction to $\mathbf{t}$ is the *genealogical order* on $\mathbf{t}$. The partial order $\preceq$ is compatible with the total order $\leq$ in that $u \preceq v \Rightarrow u \leq q$. A tree $\mathbf{t}$ in the sense of Definition 6 can be represented graphically as in Figure 2.

Let $\mathbb{T}$ be the space of all such trees, and let $\mathbb{T}_u = \{\mathbf{t} \in \mathbb{T} \colon u \in \mathbf{t}\}$. Then ν_u is a map defined on $\mathbb{T}_u$ taking values in $\mathbb{N}_0$. Note that $\mathbb{T}$ is uncountable. A sigma-algebra on $\mathbb{T}$ is defined as

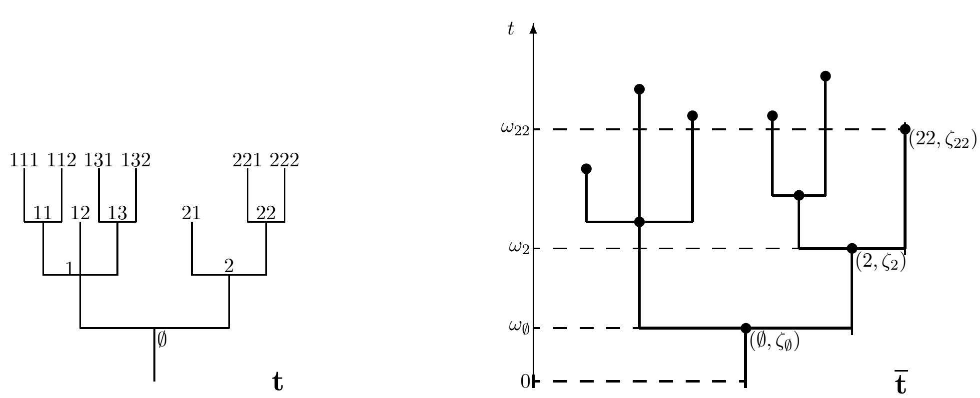

Figure 2: On the left, $\mathbf{t} = \{\emptyset, 1, 2, 11, 12, 13, 21, 22, 111, 112, 131, 132, 221, 222\}$, and on the right $\overline{\mathbf{t}} = \{(\emptyset, \zeta_\emptyset), (1, \zeta_1), (11, \zeta_{11}), (12, \zeta_{12}), (13, \zeta_{13}), (2, \zeta_2), (21, \zeta_{21}), (211, \zeta_{211}), (212, \zeta_{212}), (22, \zeta_{22})\}$.

$\mathcal{F} = \sigma\{\mathbb{T}_u, u \in \mathbb{U}\}$. We also specify the nth generation $\pi_n(\mathbf{t}) = \{u \in \mathbf{t} \colon |u| = n\} = \mathbf{t} \cap \mathbb{N}^n$ and set $\mathcal{F}_n = \sigma\{\mathbb{T}_u, |u| \leq n\} = \sigma\{\pi_m, m \leq n\}$.

We define the shift map/operator θ_u that assigns to a tree $\mathbf{t}$ its *subtree* $\mathbf{t}_u = \theta_u \mathbf{t}$ *above* $u \in \mathbf{t}$:

$$\theta_u \colon \mathbb{T}_u \to \mathbb{T}, \qquad \mathbf{t} \mapsto \theta_u \mathbf{t} = \mathbf{t}_u = \{v \in \mathbb{U} \colon uv \in \mathbf{t}\}.$$

Clearly $\mathbb{T}_\emptyset = \mathbb{T}$ and $\{\nu_u \geq j\} \cap \mathbb{T}_u = \mathbb{T}_{uj}$, also $\theta_u^{-1}(\mathbb{T}_v) = \mathbb{T}_{uv}$ and

$$\mathbb{T}_v = \{\mathbf{t} \in \mathbb{T} \colon \nu_{v_1 v_2 \cdots v_k}(\mathbf{t}) \geq v_{k+1} \text{ for all } 0 \leq k < m\} \qquad \text{for } v = v_1 v_2 \cdots v_m \in \mathbb{U}.$$

These relations allow us to consider a *random* tree τ whose distribution is a probability measure $\mathbb{P}_q$ on $\mathbb{T}$, under which the numbers of children ν_u of the individuals u in the random tree are independent random variables with distribution q. More formally, $\mathbb{P}_q$ is characterised as follows:

GW(q)-trees and their branching property (see e.g. Neveu [36])

(a) For any probability measure q on $\mathbb{N}_0$, there exists a unique probability measure $\mathbb{P}_q$ on $(\mathbb{T}, \mathcal{F})$ such that $\mathbb{P}_q(\nu_\emptyset = j) = q(j)$, and conditionally on $\{\nu_\emptyset = j\}$ for any $j \geq 1$ with $q(j) > 0$, the subtrees θ_i, $1 \leq i \leq j$, above the first generation are independent with distribution $\mathbb{P}_q$. A random tree τ with distribution $\mathbb{P}_q$ is called a GW(q)-tree.

(b) Under $\mathbb{P}_q(\,\cdot\,|\mathcal{F}_n, \pi_n = A)$, the subtrees θ_u, $u \in A$, above the nth generation are independent and with distribution $\mathbb{P}_q$, for all finite $A \subset \mathbb{N}^n$ and $n \geq 1$ with $\mathbb{P}(\pi_n = A) > 0$.

For finite trees, in particular in the *(sub)critical* case $\mathbb{E}_q(\nu_\emptyset) = \sum_{j \in \mathbb{N}_0} j q(j) \leq 1$, the measure $\mathbb{P}_q$ can be expressed as

$$\mathbb{P}_q(\{\mathbf{t}\}) = \prod_{v \in \mathbf{t}} q(\nu_v(\mathbf{t})), \qquad \text{for all } \mathbf{t} \in \mathbb{T},$$

but this does not specify the measure $\mathbb{P}_q$ in the *supercritical* case $\mathbb{E}_q(\nu_\emptyset) > 1$, where $\mathbb{P}_q$ assigns positive measure to infinite trees. Here, $\mathbb{E}_q$ is the expectation operator associated with $\mathbb{P}_q$. For a GW(q)-tree τ, the process $G_n = \#\pi_n(\tau)$, $n \geq 0$, is known as a GW(q)-branching process.

2.1.2 Marked trees and discrete branching processes in continuous time

Let $(\mathbb{H}, \mathcal{H})$ be a measurable space of marks. We can attach a mark $\xi_u \in \mathbb{H}$ to each vertex u of a given tree $\mathbf{t}$. Formally, a marked tree is a subset

$$\overline{\mathbf{t}} \subset \mathbb{U} \times \mathbb{H} \quad \text{such that} \quad \mathbf{t} = \{u \in \mathbb{U} \colon (u, \xi_u) \in \overline{\mathbf{t}} \text{ for some } \xi_u \in \mathbb{H}\} \in \mathbb{T} \tag{4}$$

and where $\overline{\mathbf{t}} \cap \{u\} \times \mathbb{H} = \{(u, \xi_u)\}$, i.e. the map $\xi \colon \mathbf{t} \to \mathbb{H}$ is unique. So a marked tree has the form $\overline{\mathbf{t}} = \{(u, \xi_u) \in \mathbb{U} \times \mathbb{H} \colon u \in \mathbf{t}\}$. We denote by $\mathbb{T}^{\mathbb{H}}$ the set of marked trees. We consider the set of trees $\mathbb{T}_u^{\mathbb{H}} = \{\overline{\mathbf{t}} \in \mathbb{T}^{\mathbb{H}} \colon u \in \mathbf{t}\}$ containing individual u and note that $\xi_u \colon \mathbb{T}_u^{\mathbb{H}} \to \mathbb{H}$ is a map.

For marked trees, we set $\nu_u(\overline{\mathbf{t}}) = \nu_u(\mathbf{t})$ and $\overline{\mathbf{t}}_u = \overline{\theta}_u \overline{\mathbf{t}} = \{(v, \xi_v) \colon v \in \mathbf{t}_u\} = \{(v, \xi_v) \colon uv \in \mathbf{t}\}$. As sigma-algebra on $\mathbb{T}^{\mathbb{H}}$ we take one that makes $\overline{\mathbf{t}} \mapsto \mathbf{t}$ in (4) and $\overline{\mathbf{t}} \mapsto \xi_u(\overline{\mathbf{t}})$ measurable:

$$\mathcal{F}^{\mathbb{H}} = \sigma\{\mathbb{T}_{u,H}^{\mathbb{H}}, u \in \mathbb{U}, H \in \mathcal{H}\}, \text{ where } \mathbb{T}_{u,H}^{\mathbb{H}} = \{\overline{\mathbf{t}} \in \mathbb{T}_u^{\mathbb{H}} \colon \xi_u \in H\}.$$

For example, for $\mathbb{H} = (0, \infty)$, the marks can represent the *lifetimes* of individuals. We will later use $\mathbb{H} = (0, \infty) \times \{0, 1\}$ so that $\xi_u = (\zeta_u, \gamma_u)$ consists of a lifetime mark $\zeta_u \in (0, \infty)$ and a colour mark $\gamma_u \in \{0, 1\}$. We consider a model when $u \in \mathbf{t}$ will produce children at the moment of its death. The birth and death times α_u and ω_u of each individual $u \in \mathbf{t}$, are then defined recursively by

$$\begin{cases} \alpha_\emptyset = 0, \ \omega_\emptyset = \zeta_\emptyset, \\ \alpha_{uj} = \omega_u, \ \omega_{uj} = \alpha_{uj} + \zeta_{uj}, \quad 1 \leq j \leq \nu_u, u \in \mathbf{t}. \end{cases}$$

We denote by $\overline{\pi}_t(\overline{\mathbf{t}}) = \{u \in \mathbf{t}\colon \alpha_u < t \leq \omega_u\}$ the set of individuals alive at time $t \geq 0$ and define $\mathcal{F}_t^{\mathbb{H}} = \sigma\{\overline{\pi}_s, s \leq t\}$. For $u \in \overline{\pi}_t(\overline{\mathbf{t}})$, we denote by

$$\overline{\theta}_{u,t}(\overline{\mathbf{t}}) = \{(\emptyset, \omega_u - t)\} \cup \{(v, \zeta_{uv})\colon uv \in \mathbf{t}\}$$

the subtree of individual u above t. Figure 2 gives a graphical representation of a tree $\overline{\mathbf{t}}$ where lifetimes are shown as edge lengths.

In this formalism, we can define and study GW(q,κ)-trees as GW(q) trees with independent lifetimes distributed according to a measure κ on $(0,\infty)$:

GW(q,κ)-trees and their branching property (see Neveu [36], Chauvin [8])

(a) For any probability measure $\mathbb{Q}$ on $\mathbb{N}_0 \times \mathbb{H}$, there exists a unique probability measure $\mathbb{P}_{\mathbb{Q}}$ on $(\mathbb{T}^{\mathbb{H}}, \mathcal{F}^{\mathbb{H}})$, such that $(\nu_\emptyset, \xi_\emptyset) \sim \mathbb{Q}$ and conditionally on $\{\nu_\emptyset = j, \xi_\emptyset \in H\}$ for any $j \geq 1$, $H \in \mathcal{H}$, with $\mathbb{Q}(\{j\} \times H) > 0$, the subtrees $\overline{\theta}_i$, $1 \leq i \leq j$, are independent with distribution $\mathbb{P}_{\mathbb{Q}}$. For $\mathbb{H} = (0,\infty)$ and $\mathbb{Q} = q \otimes \kappa$ a random tree T with distribution $\mathbb{P}_{\mathbb{Q}}$ is called a GW(q,κ)-tree.

(b) Under $\mathbb{P}_{q\otimes\kappa}(\,\cdot\,|\mathcal{F}_t^{\mathbb{H}}, \overline{\pi}_t = A)$, the subtrees $\overline{\theta}_{u,t}$, $u \in A$, above time t are independent and distributed like $\overline{\theta}_{\emptyset,s}$ under $\mathbb{P}_{q\otimes\kappa}(\,\cdot\,|\zeta_\emptyset > s)$, where $s = t - \alpha_u$ is the ($\mathcal{F}_t^{\mathbb{H}}$-measurable) age of u at time t, for all finite $A \subset \mathbb{U}$ with $\mathbb{P}(\overline{\pi}_t = A) > 0$.

For further details including strong branching properties, we refer to [6, 7]. For a GW(q,κ)-tree T, the process $Z_t = \#\overline{\pi}_t(T)$, $t \geq 0$, is known as a Bellman-Harris branching process [2, 3]. The Markovian special case for $\kappa = \mathrm{Exp}(c)$ is also called a continuous-time Galton-Watson process.

2.1.3 Coloured leaves, coloured trees and a two-colours branching property

On a suitable probability space $(\Omega, \mathcal{A}, \mathbb{P})$, let $T\colon (\Omega, \mathcal{A}, \mathbb{P}) \to (\mathbb{T}^{(0,\infty)}, \mathcal{F}^{(0,\infty)}, \mathbb{P}_{q\otimes\kappa})$ be a GW(q,κ)-tree. We assume $q(0) > 0$, i.e. T has leaves, and $q(1) = 0$, as an individual producing a single child can be viewed as continuing to live instead of being replaced by its child.

Following [11], we independently mark the *leaves* u of T with one of two colours, say *red*, $\mathbb{P}(\gamma_u(T) = 1|\nu_u(T) = 0) = p$, or *black*, $\mathbb{P}(\gamma_u(T) = 0|\nu_u(T) = 0) = 1-p$, for some given $p \in (0,1)$. It will be convenient to also mark each non-leaf individual in black if the subtree above it has at least one black leaf, red otherwise. Such a marked tree $T^{p-\mathrm{col}}$ is a random element of $\mathbb{T}^{\mathbb{H}}$ for $\mathbb{H} = (0,\infty) \times \{0,1\}$. We denote its distribution by $\mathbb{P}_{q\otimes\kappa}^{p-\mathrm{col}}$. Note that it is *not* of the form $\mathbb{P}_{\mathbb{Q}}$ introduced in the previous section, because marks for non-leaf individuals will not be independent. We set

$$g(p) = \mathbb{P}_{q\otimes\kappa}^{p-\mathrm{col}}(\gamma_\emptyset = 1) = \mathbb{P}(T^{p-\mathrm{col}} \text{ has only red colour marks}) = \mathbb{E}[p^{\#\{u\in T\colon \nu_u(T)=0\}}]. \quad (5)$$

Branching properties of coloured GW(q,κ)-trees (cf. Duquesne and Winkel [11])

(a) For all Borel-measurable $k\colon (0,\infty) \to [0,\infty)$, $j \geq 2$, $\varepsilon_i \in \{0,1\}$ and $\mathcal{F}^{\mathbb{H}}$-measurable $f_i\colon \mathbb{T}^{\mathbb{H}} \to [0,\infty)$, $i = 1,\ldots,j$, we have

$$\mathbb{E}_{q\otimes\kappa}^{p-\mathrm{col}}\left[k(\zeta_\emptyset)\prod_{i=1}^{j} f_i(\overline{\theta}_i); \nu_\emptyset = j; (\gamma_1,\ldots,\gamma_j) = (\varepsilon_1,\ldots,\varepsilon_j)\right]$$
$$= \int_{(0,\infty)} k(z)\kappa(dz)\; q(j)\; g(p)^{j_r}(1-g(p))^{j_b}\; \prod_{i=1}^{j} \mathbb{E}_{q\otimes\kappa}^{p-\mathrm{col}}[f_i|\gamma_\emptyset = \varepsilon_i]$$

where $j_r = \varepsilon_1 + \cdots + \varepsilon_j$ and $j_b = j - j_r$ are the numbers of red and black colour marks.

(b) For all $t \geq 0$ and $\mathcal{F}^{\mathbb{H}}$-measurable $f_u \colon \mathbb{T}^{\mathbb{H}} \to [0, \infty)$

$$\mathbb{E}^{p-\mathrm{col}}_{q\otimes\kappa}\left[\prod_{u\in\overline{\pi}_t} f_u(\overline{\theta}_{u,t})\middle|\mathcal{F}^{\mathbb{H}}_t\right] = \prod_{u\in\overline{\pi}_t} \mathbb{E}^{p-\mathrm{col}}_{q\otimes\kappa}[f_u(\overline{\theta}_{\emptyset,s})|\zeta_\emptyset > s]\Big|_{s=t-\alpha_u}$$

In the exponential case $\kappa = \mathrm{Exp}(c)$, this simplifies to

$$\mathbb{E}^{p-\mathrm{col}}_{q\otimes\mathrm{Exp}(c)}\left[\prod_{u\in\overline{\pi}_t} f_u(\overline{\theta}_{u,t})\middle|\mathcal{F}^{\mathbb{H}}_t\right] = \prod_{u\in\overline{\pi}_t} \mathbb{E}^{p-\mathrm{col}}_{q\otimes\mathrm{Exp}(c)}[f_u] \tag{6}$$

Reduction procedure to identify the "black tree" in a two-colours tree

- We can extract $\widetilde{T}^{p-\mathrm{rdc}}_{\mathrm{sub}} = \{(u, \zeta_u) \in \mathbb{U} \times (0,\infty) \colon (u, \zeta_u, 1) \in T^{p-\mathrm{col}}\}$, the individuals of $T^{p-\mathrm{col}}$ with black colour marks. If $\widetilde{T}^{p-\mathrm{rdc}}_{\mathrm{sub}} \neq \emptyset$, we rename the individuals of $\widetilde{T}^{p-\mathrm{rdc}}_{\mathrm{sub}}$ by the unique injection

$$\iota \colon \widetilde{\tau}^{p-\mathrm{rdc}}_{\mathrm{sub}} = \{u \in \mathbb{U} \colon (u, \zeta_u) \in \widetilde{T}^{p-\mathrm{rdc}}_u\} \to \mathbb{U},$$

  that is increasing for the lexicographical total order on $\mathbb{U}$, maps onto an element $\tau^{p-\mathrm{rdc}}_{\mathrm{sub}}$ of $\mathbb{T}$ and is compatible with the genealogical partial orders. We refer to the image tree $T^{p-\mathrm{rdc}}_{\mathrm{sub}} = \{(\iota(u), \zeta_u) \colon u \in \widetilde{\tau}^{p-\mathrm{rdc}}_{\mathrm{sub}}\}$ as the *p-reduced subtree of T*.

- As a further reduction, we remove single-child individuals and add their lifetimes to the child's lifetime. Formally, we define $\widetilde{\tau}^{p-\mathrm{rdc}} = \{v \in \tau^{p-\mathrm{rdc}}_{\mathrm{sub}} \colon \nu_v(\tau^{p-\mathrm{rdc}}_{\mathrm{sub}}) \neq 1\}$, and

$$\widetilde{\zeta}_u = \sum_{i=J_u}^{n} \zeta_{u_1\cdots u_i}(\tau^{p-\mathrm{rdc}}_{\mathrm{sub}}), \text{ where } J_u = \sup\{j \colon \nu_{u_1\cdots u_i}(\tau^{p-\mathrm{rdc}}_{\mathrm{sub}}) = 1 \text{ for all } i \in \{j, \ldots, n-1\}\},$$

  for all $u = u_1 \cdots u_n \in \widetilde{\tau}^{p-\mathrm{rdc}}$. Again, there is a unique injection $\iota' \colon \widetilde{\tau}^{p-\mathrm{rdc}} \to \mathbb{U}$ that is increasing for the lexicographical total order on $\mathbb{U}$, maps onto an element $\tau^{p-\mathrm{rdc}}$ of $\mathbb{T}$ and is compatible with the genealogical partial orders. We refer to the image tree $T^{p-\mathrm{rdc}} = \{(\iota'(u), \widetilde{\zeta}_u) \colon u \in \widetilde{\tau}^{p-\mathrm{rdc}}\}$ as the *p-reduced tree* (or as the *black tree*).

Figure 1 in the Introduction illustrates the reduction procedure.

Remark 7. (a) The reduction procedure is transitive in that for independent colouring, we have $(T^{(1-\overline{p}_1)-\mathrm{rdc}})^{(1-\overline{p}_2)-\mathrm{rdc}} \overset{(\mathrm{d})}{=} T^{(1-\overline{p}_1\overline{p}_2)-\mathrm{rdc}}$. In particular, colouring for $T^{(1-\overline{p}_1)-\mathrm{rdc}}$ and $T^{(1-\overline{p}_3)-\mathrm{rdc}}$ for $\overline{p}_3 < \overline{p}_1$ can be coupled such that $T^{(1-\overline{p}_3)-\mathrm{rdc}} = (T^{(1-\overline{p}_1)-\mathrm{rdc}})^{(1-\overline{p}_3/\overline{p}_1)-\mathrm{rdc}}$.

(b) Although we have used notation for a random $\mathrm{GW}(q, \kappa)$-tree $T^{p-\mathrm{col}}$ with leaves coloured independently, note that the reduction of $T^{p-\mathrm{col}}$ to a black tree is a purely deterministic procedure. Our focus here has been on the technical framework and how it is used to formulate relevant examples. We postpone to Section 2.3 the review of further developments, notably of the reconstruction/growth procedures that reverse the reduction.

2.1.4 Bushes and forests – models with several progenitors and immigration

Branching processes with immigration have been studied widely (see e.g. Athreya and Ney [1] and Jagers [24]). We consider the model, where immigrants arrive at the times S_i, $i \geq 1$, of a renewal process $J_t = \#\{i \geq 1 \colon S_i \leq t\}$, i.e. where $S_0 = 0$ and the interarrival times $S_i - S_{i-1}$, $i \geq 1$, are independent and identically distributed random variables with a common distribution on $(0, \infty)$ that we denote by χ. At each immigration time S_i the number N_i of immigrants is independent and has a common distribution η on $\mathbb{N}$. Each immigrant produces offspring independently according to the rules of $\mathrm{GW}(q,\kappa)$-trees.

Denoting by $Z^{(i)}_{t-S_i}$ the size at time t of the population of immigrants arriving at time S_i,

$$Y_t = \sum_{i=1}^{J_t} Z^{(i)}_{t-S_i} \tag{7}$$

is the total population size at time $t \geq 0$. Here, $Z^{(i)}$ are independent sums of N_i independent Bellman-Harris processes with offspring and lifetime distributions q and κ as in Section 2.1.2.

To capture the genealogical trees of the population, we use the notion of a *bush* as a random sequence $B = (T_{(1)}, \ldots, T_{(N)})$ of independent trees, and the notion of a *forest* as a point process $F = (B(t), t \geq 0)$ of independent bushes

$$B(S_i) = B^{(i)}, \quad i \geq 0, \qquad B(t) = \partial, \quad t \notin \{S_i, i \geq 1\} \quad \text{for a cemetery point } \partial.$$

$\mathrm{GW}(q,\kappa,\beta)$**-bushes and** $\mathrm{GWI}(q,\kappa,\eta,\chi)$**-forests**

- A $\mathrm{GW}(q,\kappa,\beta)$-bush is a bush $B = (T_{(1)}, \ldots, T_{(N)})$ of independent $\mathrm{GW}(q,\kappa)$-trees $T_{(j)}$, where $N \sim \mathrm{Poi}(\beta)$.
- A $\mathrm{GWI}(q,\kappa,\eta,\chi)$-forest is a forest $F = (B(t), t \geq 0)$ where each bush $B(S_i) = B^{(i)}$ is associated with immigration at the times S_i of a renewal process with inter-arrival distribution χ and consists of an independent η-distributed number N_i of trees $T^{(i)}_{(j)}$.

It is straightforward to transfer the notions of colouring and reduction to the setting of bushes and forests, since they apply tree by tree. We will slightly abuse notation and write $u \in B$ to refer to individuals in a bush, instead of writing formally $u = (i, u'')$ with $u'' \in T_{(i)}$. Similarly, $u \in F$ means $u = (t, u')$ with $u' \in B(t)$ in the sense just defined. We will also abuse notation ν_u, ζ_u and γ_u accordingly for $u \in B$ and $u \in F$.

2.2 Continuous-state branching processes and immigration

We have looked at branching processes with immigration in the discrete state space $\mathbb{N}_0 = \{0, 1, 2, \ldots\}$ in continuous time. In this section we will recall (Markovian) continuous-state branching processes with immigration, where the state space (population size) will no longer be $\mathbb{N}_0$ but $[0, \infty)$, and time is also continuous.

2.2.1 Subordinators and geometric infinite divisibility

Definition 8. An increasing right-continuous stochastic process $\sigma = (\sigma(t), t \geq 0)$ in $[0, \infty)$ is called a *subordinator* if it has stationary independent increments, i.e. if for every $u, t \geq 0$, the *increment* $\sigma(t+u) - \sigma(t)$ is independent of $(\sigma(s), s \leq t)$ and $\sigma(t+u) - \sigma(t) \overset{(\mathrm{d})}{=} \sigma(u)$.

The distribution of a subordinator σ on the space $\mathbb{D}([0,\infty),[0,\infty))$ of functions $f\colon [0,\infty) \to [0,\infty)$ that are right-continuous and have left limits equipped with the Borel sigma-algebra generated by Skorohod's topology, see e.g. [23], is specified by the Laplace transforms of its one-dimensional distributions. For every $t \geq 0$ and $r \geq 0$,

$$\mathbb{E}(\exp\{-r\sigma(t)\}) = \exp\{-t\phi(r)\} \tag{8}$$

where the function $\phi\colon [0,\infty) \to [0,\infty)$ is called the *Laplace exponent* of σ. There exist a unique real number $d \geq 0$ and a unique measure Λ on $(0,\infty)$ with $\int(1 \wedge x)\Lambda(dx) < \infty$, such that for every $r \geq 0$

$$\phi(r) = dr + \int_{(0,\infty)} (1 - e^{-rx})\Lambda(dx). \tag{9}$$

Conversely, any function ϕ of the form (9) is the Laplace exponent of a subordinator, which can be constructed as $(dt + \sum_{s\leq t} \Delta\sigma_s, t \geq 0)$ for a Poisson point process $(\Delta\sigma_s, s \geq 0)$ in $(0,\infty)$ with intensity measure Λ. Equation (9) is referred to as the Lévy-Khintchine representation of ϕ. We refer to Bertoin [5] for an introduction to subordinators and their applications.

Definition 9. A random variable X is *geometrically infinitely divisible* (g.i.d.) if for all $\alpha \in (0,1)$

$$X \overset{(\mathrm{d})}{=} \sum_{j=1}^{G^{(\alpha)}} X_\alpha^{(j)}, \qquad j \geq 1, \tag{10}$$

for a sequence $X_\alpha^{(j)}$, $j \geq 1$, of independent identically distributed (i.i.d.) random variables and an independent $G^{(\alpha)} \sim \mathrm{Geo}(\alpha)$:

$$\mathbb{P}(G^{(\alpha)} = k) = \alpha(1-\alpha)^{k-1}, \qquad k \geq 1.$$

For example, an exponential random variable $X \sim \mathrm{Exp}(c)$ is g.i.d. since (10) holds for $X_\alpha^{(j)} \sim \mathrm{Exp}(c/\alpha)$. The class of g.i.d. distributions can be characterised as follows.

Lemma 10 ([30])**.** *A random variable X is g.i.d. if and only if it can be expressed as $X \overset{(\mathrm{d})}{=} \sigma(V)$, where $\sigma = (\sigma(t), t \geq 0)$ is a subordinator and $V \sim \mathrm{Exp}(c)$ independent of σ for one equivalently all $c \in (0,\infty)$.*

Indeed we can then express (V and hence) X as

$$V = \sum_{j=1}^{G^{(\alpha)}} V_\alpha^{(j)} \qquad \text{and} \qquad X = \sigma_V = \sum_{j=1}^{G^{(\alpha)}} \left(\sigma\left(V_\alpha^{(1)} + \cdots + V_\alpha^{(j)}\right) - \sigma\left(V_\alpha^{(1)} + \cdots + V_\alpha^{(j-1)}\right)\right).$$

Lemma 11. *If X is such that (10) holds for some $\alpha \in (0,1)$, the distribution of $X_\alpha^{(j)}$ is unique.*

Proof. $\mathbb{E}(e^{-rX}) = \dfrac{\mathbb{E}(e^{-rX_\alpha^{(1)}})\alpha}{1-(1-\alpha)\mathbb{E}(e^{-rX_\alpha^{(1)}})} \quad \Rightarrow \quad \mathbb{E}(e^{-rX_\alpha^{(1)}}) = \dfrac{\mathbb{E}(e^{-rX})}{\alpha + (1-\alpha)\mathbb{E}(e^{-rX})}.$ □

2.2.2 Continuous-state branching processes

Continuous-state branching processes were introduced by Jirina [26] and Lamperti [33]. They are the limiting processes of sequences of rescaled Galton-Watson processes as the initial population size tends to infinity and the mean lifetime tends to zero. In this section we follow Le Gall [35].

Definition 12. A continuous-state branching process (CSBP) is a right-continuous Markov process $(Z_t, t \geq 0)$ in $[0,\infty)$, whose transition kernels $P_t(x, dz)$ are such that for every $t \geq 0$, $x \geq 0$ and $x' \geq 0$ we have

$$P_t(x + x', \cdot) = P_t(x, \cdot) * P_t(x', \cdot),$$

where $*$ denotes convolution. In other words, if for a given transition kernel we denote, for each $x \geq 0$, by Z^x a CSBP starting from $Z_0^x = x$, then for $\widetilde{Z}^{x'} \overset{(\mathrm{d})}{=} Z^{x'}$ independent, we require $Z_t^x + \widetilde{Z}_t^{x'} \overset{(\mathrm{d})}{=} Z_t^{x+x'}$.

The transition kernel is specified by the Laplace transforms

$$\mathbb{E}(\exp\{-rZ_t\}|Z_0 = x) = \exp\{-xu_t(r)\} \tag{11}$$

where $u_t\colon [0,\infty) \to [0,\infty)$. In fact, $(t,r) \mapsto u_t(r)$ is necessarily the unique non-negative solution of

$$u_t(r) + \int_0^t \psi(u_s(r))ds = r \qquad \text{or} \qquad \frac{\partial u_t(r)}{\partial t} = -\psi(u_t(r)) \tag{12}$$

with $u_0(r) = r$, for some $\psi\colon [0,\infty) \to \mathbb{R}$ of the form

$$\psi(r) = br + ar^2 + \int_{(0,\infty)} (e^{-rx} - 1 + rx\mathbf{1}_{\{x<1\}})\Pi(dx), \tag{13}$$

where $b \in \mathbb{R}$, $a \geq 0$ and Π is a measure on $(0,\infty)$ with $\int(1 \wedge x^2)\Pi(dx) < \infty$ and where ψ satisfies the non-explosion condition $\int_{0+} |\psi(r)|^{-1}dr = \infty$, see [20]. Equation (13) is referred to as the Lévy-Khintchine representation of ψ. The process $(Z_t, t \geq 0)$ is then called a CSBP *with branching mechanism* ψ, or a CSBP(ψ). We denote by $\mathbb{P}_\psi^x$ the distribution of $(Z_t^x, t \geq 0)$ on $\mathbb{D}([0,\infty),[0,\infty))$. There also exists a sigma-finite measure Θ_ψ on $\mathbb{D}([0,\infty),[0,\infty))$, such that

$$(Z_t^x, t \geq 0) \stackrel{(d)}{=} \left(\sum_{0 \leq y \leq x} E_t(y), t \geq 0 \right), \qquad x \geq 0,$$

where $(E(y), y \geq 0)$ is a Poisson point process in $\mathbb{D}([0,\infty),[0,\infty))$ with intensity measure Θ_ψ.

2.2.3 Continuous-state branching processes with immigration

Similarly, a discrete-state branching process with immigration has a continuous analogue, the continuous-state branching process with immigration, CBI for short. Following Kawazu and Watanabe [29], see also [10, 31], besides the branching mechanism ψ for a CSBP, we also have an immigration mechanism ϕ of the form (9) for the CBI, which we then refer to as CBI(ψ,ϕ).

A CBI(ψ,ϕ) is a Markov process $(Y_t, t \geq 0)$ on $[0,\infty)$ whose transition kernels are characterized by their Laplace transform, which in terms of ϕ and $u_t(r)$ as in (12) satisfy

$$\mathbb{E}(\exp\{-rY_t\}|Y_0 = x) = \exp\left\{ -xu_t(r) - \int_0^t \phi(u_s(r))ds \right\}.$$

In fact, a subordinator $\sigma = (\sigma(t), t \geq 0)$ with Laplace exponent ϕ can be viewed as a pure immigration process CBI(ϕ, 0). Indeed, a general CBI(ψ,ϕ) is such that by time $t \geq 0$ a population of total size $\sigma(t)$ has immigrated and evolved like a CSBP(ψ); specifically, consider a Poisson point process $(E^s, s \geq 0)$ in $\mathbb{D}([0,\infty),[0,\infty))$ with intensity measure $d\Theta_\psi + \int_{(0,\infty)} \mathbb{P}_\psi^x \Lambda(dx)$, then in analogy with (7)

$$Y_t = \sum_{s \leq t} E_{t-s}^s, \qquad t \geq 0,$$

is a CBI(ψ,ϕ). This follows by the exponential formula and properties of Poisson point processes.

Examples of continuous-state branching processes with immigration include (sub)critical CSBP conditioned on survival, see [32] and literature therein.

2.3 Growth of Galton-Watson bushes with exponential edge lengths

In Theorem 1, we consider families of $\mathrm{GW}(q_\lambda, \mathrm{Exp}(c_\lambda), \beta_\lambda)$-bushes B_λ, $\lambda \geq 0$, with the consistency property that any two bushes, for parameters $\mu < \lambda$ say, are related by p-reduction as formally defined in Section 2.1.3, for $p = 1 - \mu/\lambda$. The choice of p is dictated (up to a positive power for the ratio) by the consistency requirement that the relation holds for *all* μ and *all* λ (cf. Remark 7(a)).

The equivalence of (i) and (ii) in Theorem 1 is a statement purely about offspring distributions $(q_\lambda, \lambda \geq 0)$. The reason for including the other two parameter sequences $(c_\lambda, \lambda \geq 0)$ and $(\beta_\lambda, \lambda \geq 0)$ in the remainder of the statement is simplicity. Specifically, if we consider trees without lifetimes and hence without embedding in time, the removal of single-child individuals will be more artificial as it reduces the height of the trees; if we look at trees instead of bushes, the reduced tree will only be GW if we condition on the existence of at least one black leaf, and this does not lead to consistent families of random trees on the same probability space.

In [11], the bushes of Theorem 1 are used to construct Lévy trees as a strong representation of the genealogy of the limiting $\mathrm{CSBP}(\psi)$ under some extra conditions on ψ. In a weaker sense, the family $(B_\lambda, \lambda \geq 0)$ itself is already a representation of the genealogy of $\mathrm{CSBP}(\psi)$, under no conditions on ψ other than $\psi(\infty) = \infty$ to exclude the case of increasing CSBPs corresponding to "no death" i.e. "no leaves". Roughly, B_λ is the genealogy of a Poisson sample chosen among all individuals with intensity proportional to λ; as $\lambda \to \infty$, the set of individuals included becomes dense.

2.3.1 Consistent families of $\mathrm{GW}(q_\lambda, \mathrm{Exp}(c_\lambda), \beta_\lambda)$-bushes

In [11], the parameters in a consistent family of Galton-Watson bushes are represented in terms of a branching mechanism ψ that is just required to satisfy $\psi(\infty) = \infty$, so that ψ is eventually increasing and has a right inverse $\psi^{-1}\colon [0, \infty) \to [\psi^{-1}(0), \infty)$:

$$\varphi_{q_\lambda}(s) = s + \frac{\psi(\psi^{-1}(\lambda)(1-s))}{\psi^{-1}(\lambda)\psi'(\psi^{-1}(\lambda))}, \qquad c_\lambda = \psi'(\psi^{-1}(\lambda)), \qquad \beta_\lambda = \beta\psi^{-1}(\lambda). \tag{14}$$

It follows from the derivation there that this ψ and $(\widetilde{\psi}, c)$ in the statement of Theorem 1 are related in a linear way as

$$\psi(r) = k_1 \widetilde{\psi}(k_2 r), \tag{15}$$

where $k_1 = 1/\widetilde{\psi}(1) = \psi^{-1}(1)\psi'(\psi^{-1}(1))$ and $k_2 = c\widetilde{\psi}(1) = 1/\psi^{-1}(1)$. For the underlying CSBPs, the relationship (15) just means $Z_t = k_2 \widetilde{Z}_{k_1 k_2 t}$, so ψ and $\widetilde{\psi}$ essentially refer to the same CSBP. With this parameterisation, (5) can be expressed more explicitly for $q = q_\lambda$ and $p = 1 - \mu/\lambda$ as

$$g_\lambda(1 - \mu/\lambda) = \mathbb{P}^{(1-\mu/\lambda)-\mathrm{col}}_{q_\lambda \otimes \mathrm{Exp}(c_\lambda)}(\gamma_\emptyset = 1) = 1 - \frac{\widetilde{\psi}^{-1}(\mu)}{\widetilde{\psi}^{-1}(\lambda)} = 1 - \frac{\psi^{-1}(\mu)}{\psi^{-1}(\lambda)}. \tag{16}$$

2.3.2 Analysis of the reduction procedure for $\mathrm{GW}(q_\lambda, \mathrm{Exp}(c_\lambda), \beta_\lambda)$-bushes

In this section we will study some key points in the reduction procedure in the setting of Theorem 1. These will be important for the proofs of Theorems 2 and 3. There are three steps in the reduction procedure from λ to $\mu < \lambda$: colouring with $p = 1 - \mu/\lambda$, passage to the p-reduced sub-bush and passage to the p-reduced bush.

In the last step, the lifetime $\zeta_u^{(1-\mu/\lambda)-\mathrm{rdc}}$ of individual $u \in T_\lambda^{(1-\mu/\lambda)-\mathrm{rdc}}$ is obtained combining a number $G_u^{(\alpha)}$ of lifetimes of T_λ when removing the single-child individuals. By the two-colours branching property in Section 2.1.3, the random numbers $G_u^{(\alpha)}$ are $\mathrm{Geo}(\alpha)$, where α is the probability that an individual in $T_{\lambda,\mathrm{sub}}^{(1-\mu/\lambda)-\mathrm{rdc}}$ produce zero or more than two children. We can express α in terms of ψ:

Lemma 13. *Let $0 \le \mu < \lambda < \infty$. Given that $u \in B_\lambda^{(1-\mu/\lambda)-\mathrm{rdc}}$, we have $G_u^{(\alpha)} \sim \mathrm{Geo}(\alpha)$ with*

$$\alpha = \mathbb{P}\left(\nu_\emptyset\left(T_{\lambda,\mathrm{sub}}^{(1-\mu/\lambda)-\mathrm{rdc}}\right) \neq 1 \,\middle|\, \gamma_\emptyset\left(T_\lambda^{(1-\mu/\lambda)-\mathrm{col}}\right) = 0\right) = \frac{\psi'(\psi^{-1}(\mu))}{\psi'(\psi^{-1}(\lambda))}. \tag{17}$$

Proof. According to the definition of α and the two-colours branching property of Section 2.1.3,

$$\begin{aligned}
1-\alpha &= \mathbb{P}\left(\nu_\emptyset\left(T_{\lambda,\mathrm{sub}}^{(1-\mu/\lambda)-\mathrm{rdc}}\right) = 1 \,\middle|\, \gamma_\emptyset\left(T_\lambda^{(1-\mu/\lambda)-\mathrm{col}}\right) = 0\right) \\
&= \frac{1}{1-g_\lambda(1-\mu/\lambda)} \sum_{j=2}^{\infty} \binom{j}{1} q_\lambda(j) g_\lambda(1-\mu/\lambda)^{j-1}(1-g_\lambda(1-\mu/\lambda)) \\
&= \varphi'_{q_\lambda}(g_\lambda(1-\mu/\lambda))
\end{aligned}$$

and by (14) and (16) we obtain $1-\alpha = 1-\psi'(\psi^{-1}(\mu))/\psi'(\psi^{-1}(\lambda))$. □

In the reconstruction procedure reversing the reduction procedure, we will therefore subdivide each lifetime in the $\mathrm{GW}(q_\mu, \mathrm{Exp}(c_\mu), \beta_\mu)$-bush into a geometric random number of $G_u^{(\alpha)}$ parts.

Remark 14. By the branching property of coloured $\mathrm{GW}(q,\kappa)$-trees in Section 2.1.3, lifetime marks are independent of colour marks. Therefore, Lemma 13 also holds for general $B_\lambda \sim \mathrm{GW}(q_\lambda, \kappa_\lambda, \beta_\lambda)$-bushes, not just for $\mathrm{GW}(q_\lambda, \mathrm{Exp}(c_\lambda), \beta_\lambda)$-bushes.

In the case of $\mathrm{Exp}(c_\mu)$ lifetimes, given the lifetime $\zeta_u(T_\mu)$ the conditional distribution of the random variable $N_u = G_u^{(\alpha)} - 1$ follows a Poisson distribution:

Proposition 15. *Let ζ be a random variable having distribution $\zeta \sim \mathrm{Exp}(\alpha c)$. Suppose that ζ is subdivided into $G^{(\alpha)}$ independent parts, $\zeta = \zeta_1 + \cdots + \zeta_{G^{(\alpha)}}$, where $\zeta_i \sim \mathrm{Exp}(c)$ and $G^{(\alpha)} \sim \mathrm{Geo}(\alpha)$ are independent. Then given $\zeta = z$ for $z \ge 0$, we have*

$$\mathbb{P}(G^{(\alpha)} = k \mid \zeta = z) = \frac{((1-\alpha)cz)^{k-1} e^{-(1-\alpha)cz}}{(k-1)!}. \tag{18}$$

Proof. This is, of course, well-known in the context of Poisson processes, but let us give a direct argument and write the left hand side as a conditional expectation $\mathbb{P}(G^{(\alpha)} = k \mid \zeta = z) = \mathbb{E}(\mathbf{1}_{\{G^{(\alpha)}=k\}} | \zeta = z)$. We also set $g_k(z) = ((1-\alpha)cz)^{k-1} e^{-(1-\alpha)cz}/(k-1)!$.

Claim: $\mathbb{E}(f(\zeta)g_k(\zeta)) = \mathbb{E}(f(\zeta)\mathbf{1}_{\{G^{(\alpha)}=k\}})$ for all measurable $f \ge 0$.

As $\zeta \sim \mathrm{Exp}(\alpha c)$, $\mathbb{E}(f(\zeta)g_k(\zeta)) = \int_0^\infty f(z) g_k(z) \alpha c e^{-\alpha c z} dz = \int_0^\infty f(z) \frac{\alpha c((1-\alpha)cz)^{k-1} e^{-cz}}{(k-1)!} dz$. On the other hand, since $\zeta = \zeta_1 + \cdots + \zeta_{G^{(\alpha)}}$, the conditional distribution of ζ given $G^{(\alpha)} = k$ is $\mathrm{Gamma}(c,k)$. Therefore,

$$\mathbb{E}(f(\zeta)\mathbf{1}_{\{G^{(\alpha)}=k\}}) = \mathbb{P}(G^{(\alpha)} = k) \int_0^\infty f(z) \frac{z^{k-1} c^k e^{-cz}}{(k-1)!} dz = \int_0^\infty f(z) \frac{\alpha c((1-\alpha)cz)^{k-1} e^{-cz}}{(k-1)!} dz,$$

and so $g_k(z)$ is a version of the conditional probability $\mathbb{P}(G^{(\alpha)} = k \mid \zeta = z)$, as claimed. □

Moreover, we can also find the conditional joint distribution of the lifetimes $(\zeta_1, \ldots, \zeta_{G^{(\alpha)}-1})$ given $\zeta = z$ and $G^{(\alpha)} = k$ as follows:

$$f_{\zeta_1,\zeta_2,\ldots,\zeta_{k-1} | G^{(\alpha)}=k, \zeta=z}(y_1, y_2, \ldots, y_{k-1}) = \frac{(k-1)!}{z^{k-1}} \tag{19}$$

for all $y_1>0,\ldots,y_{k-1}>0$ such that $y_1+\cdots+y_{k-1}<z$.

In the middle step of the reduction procedure, all red individuals are removed. [11] noted that it is a consequence of the two-colours branching property, see Section 2.1.3 here, that they form independent $\mathrm{GW}(q_{\lambda,\mathrm{red}}^{(1-\mu/\lambda)-\mathrm{col}},\mathrm{Exp}(c_\lambda))$-trees, "red trees", where

$$\varphi_{q_{\lambda,\mathrm{red}}^{(1-\mu/\lambda)-\mathrm{col}}}(s)=s+\frac{\psi_\mu(\psi_\mu^{-1}(\lambda-\mu)(1-s))}{\psi_\mu(\lambda-\mu)\psi_\mu'(\psi_\mu^{-1}(\lambda-\mu))},\qquad\text{with } \psi_\mu(r)=\psi(\psi^{-1}(\mu)+r)-\mu. \tag{20}$$

Furthermore, the numbers of red trees removed at the branchpoints are conditionally independent, and given that a branchpoint has $m\ge 1$ subtrees containing black leaves, the generating function of the number of red trees can be expressed in terms of the mth derivative $\psi_\mu^{(m)}$ of ψ_μ:

$$\frac{\psi_\mu^{(m)}(\psi_\mu^{-1}(\lambda-\mu)(1-s))}{\psi_\mu^{(m)}(0)}\text{ for } m\ge 2 \text{ or } \frac{\psi_\mu'(\psi_\mu^{-1}(\lambda-\mu)(1-s))-\psi_\mu'(\psi_\mu^{-1}(\lambda-\mu))}{\psi_\mu'(\psi_\mu^{-1}(\lambda-\mu)-\psi_\mu'(0)}\text{ for } m=1. \tag{21}$$

2.3.3 Reconstruction procedure for $\mathrm{GW}(q_\lambda,\mathrm{Exp}(c_\lambda),\beta_\lambda)$-bushes

Let $B_\mu=(T_\mu^{(1)},T_\mu^{(2)},\ldots,T_\mu^{(N_\mu)})$ be a $\mathrm{GW}(q_\mu,\mathrm{Exp}(c_\mu),\beta_\mu)$-bush. We will construct B_λ.

1. For each individual u in $T_\mu^{(i)}$, given the lifetime $\zeta_u^{(i)}=z$, subdivide into a random number $G_u^{(\alpha,i)}$ of parts with distribution (18), where α is as in (17) and the parts $(\zeta_{u,1}^{(i)},\ldots,\zeta_{u,G_u^{(\alpha,i)}}^{(i)})$ have joint distribution (19). Now for each $i\in\{1,\ldots,N_\mu\}$, there is a unique injection
$$(\iota_i')^{-1}\colon\tau_\mu^{(i)}=\{u\in\mathbb{U}\colon(u,\zeta_u^{(i)})\in T_\mu^{(i)}\}\to\mathbb{U}$$
such that $(\iota_i')^{-1}(\emptyset)=1^k$ with $k=G_\emptyset^{(\alpha,i)}-1$ and $(\iota_i')^{-1}(uj)=(\iota_i')^{-1}(u)j1^k$ with $k=G_{uj}^{(\alpha,i)}-1$, for all $uj\in\tau_\mu^{(i)}$, where 1^k is a string of k letters 1. We define
$$\widehat{T}_\mu^{(i)}=\{(1^{n-1},\zeta_{\emptyset,n}^{(i)})\colon 1\le n\le G_\emptyset^{(\alpha,i)}\}\cup\{((\iota_i')^{-1}(u)j1^{n-1},\zeta_{u,n}^{(i)})\colon uj\in\tau_\mu^{(i)},1\le n\le G_u^{(\alpha,i)}\}.$$

2. For each individual u in $\widehat{T}_\mu^{(i)}$, given $\widehat{\nu}_u^{(i)}=m\ge 1$ children, consider a random number $C_u^{(i)}$ of further children with distribution (21) and a uniform random permutation $\varrho_u^{(i)}$ among the $(m+C_u^{(i)})!/m!$ permutations with $\varrho_u^{(i)}(1)<\cdots<\varrho_u^{(i)}(m)$. Let $T^{(u,1,i)},\ldots,T^{(u,k,i)}$ with $k=C_u^{(i)}$ be independent $\mathrm{GW}(q_{\lambda,\mathrm{red}}^{(1-\mu/\lambda)-\mathrm{rdc}},\mathrm{Exp}(c_\lambda))$-trees with offspring distribution (20). Then for each $i\in\{1,\ldots,N_\mu\}$, there is a unique injection
$$(\iota_i)^{-1}\colon\widehat{\tau}_\mu^{(i)}=\{u\in\mathbb{U}\colon(u,\widehat{\zeta}_u^{(i)})\in\widehat{T}_\mu^{(i)}\}\to\mathbb{U}$$
such that $(\iota_i)^{-1}(\emptyset)=\emptyset$ and $(\iota_i)^{-1}(uj)=(\iota_i)^{-1}(u)\varrho_u^{(i)}(j)$ for all $uj\in\widehat{\tau}_\mu^{(i)}$. We define
$$\begin{aligned}\widehat{T}_\lambda^{(i)}\;=\;&\{((\iota_i)^{-1}(u),\widehat{\zeta}_u^{(i)})\colon u\in\widehat{\tau}_\mu^{(i)}\}\\ &\cup\{(\iota_i)^{-1}(u)\varrho_u(\widehat{\nu}_u^{(i)}+j)v,\zeta_v^{(u,j,i)})\colon u\in\widehat{\tau}_\mu^{(i)},1\le j\le C_u^{(i)},v\in\tau^{(u,j,i)}\}.\end{aligned}$$

3. Given $N_\mu=n$, consider a random number $N_\lambda^{\mathrm{red}}\sim\mathrm{Poi}(\beta_\lambda-\beta_\mu)$ of further progenitors and a uniform random permutation ϱ among the $(n+N_\lambda^{\mathrm{red}})!/n!$ permutations with $\varrho(1)<\cdots<\varrho(n)$. Let $\widehat{T}_\lambda^{(n+1)},\ldots,\widehat{T}^{(n+k)}$ with $k=N_\lambda^{\mathrm{red}}$ be independent $\mathrm{GW}(q_{\lambda,\mathrm{red}}^{(1-\mu/\lambda)-\mathrm{rdc}},\mathrm{Exp}(c_\lambda))$-trees with offspring distribution (20). Then we finally define
$$B_\lambda=(\widehat{T}_\lambda^{\varrho^{-1}(1)},\ldots,\widehat{T}_\lambda^{\varrho^{-1}(n+k)})\qquad\text{with } n=N_\mu \text{ and } k=N_\lambda^{\mathrm{red}}.$$

Remark 16. (a) The constructed bush is indeed a $\mathrm{GW}(q_\lambda, \mathrm{Exp}(c_\lambda), \beta_\lambda)$-bush and the pair obtained (B_μ, B_λ) has the same distribution as $(B_\lambda^{(1-\mu/\lambda)-\mathrm{rdc}}, B_\lambda)$. The intermediate trees $\widehat{T}_\mu^{(i)}$ have the same distribution as the $(1-\mu/\lambda)$-reduced subtrees, also jointly with the pair. If the split into two parts in the definition of $\widehat{T}_\lambda^{(i)}$ is used to assign black colour marks to the first part and red colour marks to the second part, then the resulting trees have the same distribution as the $(1-\mu/\lambda)$-coloured trees $B_\lambda^{(1-\mu/\lambda)-\mathrm{col}}$. We refer to Figure 1 as a graphical illustration of the reduction and hence the reconstruction procedure.

(b) In [11], this reconstruction procedure is also formulated for representations of the trees in a space of $\mathbb{R}$-trees, tree-like metric spaces that we briefly address in Section 3.3.

(c) It is a simple consequence of the reduction procedure and/or the reconstruction procedure that $(B_\lambda, \lambda \geq 0)$ is an inhomogeneous *Markov* process in a suitable space of finite sequences of $(0,\infty)$-marked trees. Similarly, (6) can be strengthened in the present context to

$$\mathbb{E}\left(\prod_{u\in\overline{\pi}_t(B_\lambda)} f_u\left(\overline{\theta}_{u,t}\left(B_\lambda^{(1-\mu/\lambda)-\mathrm{col}}\right)\right)\middle|\, \mathcal{G}_{\lambda,t}\right) = \prod_{u\in\overline{\pi}_t(B_\lambda)} \mathbb{E}\left(f_u\left(B_\lambda^{(1-\mu/\lambda)-\mathrm{col}}\right)\right),$$

where $\mathcal{G}_{\lambda,t} = \sigma\left\{\overline{\pi}_s(B_{\lambda'}), \lambda' \geq \lambda, s \leq t\right\}$, $\lambda \geq 0$, $t \geq 0$, and $B_\lambda^{(1-\mu/\lambda)-\mathrm{col}}$ is as in (a).

2.3.4 Limiting behaviour as $\lambda \to \infty$

In the context of limiting results as $\lambda \to \infty$, we record the following corollary of Lemma 13.

Corollary 17. *If we fix $\mu > 0$ in the setting of Lemma 13, we have*

$$\alpha = \alpha(\mu,\lambda) = \frac{\psi'(\psi^{-1}(\mu))}{\psi'(\psi^{-1}(\lambda))} \to \frac{\psi'(\psi^{-1}(\mu))}{\psi'(\infty)} \qquad \text{as } \lambda \to \infty,$$

where $\psi'(\infty)$ means $\lim_{\lambda\to\infty}\psi'(\lambda)$, in the following sense:

- *If $\psi'(\infty) < \infty$, then $\alpha(\mu,\lambda) \to \alpha_0(\mu) = \psi'(\psi^{-1}(\mu))/\psi'(\infty)$.*
- *If $\psi'(\infty) = \infty$, then $\alpha(\mu,\lambda) \to \alpha_0(\mu) = 0$.*

Note that this means that as $\lambda \to \infty$, lifetimes are cut into finite $\mathrm{Geo}(\alpha_0(\mu))$-distributed numbers of pieces in the first case, but into infinitely many pieces in the second case.

Lemma 18. *Let $(Z_t^\lambda, t \geq 0)$, $\lambda \geq 0$, be continuous-time Galton-Watson processes associated with a consistent family $(B_\lambda)_{\lambda\geq 0}$, of $\mathrm{GW}(q_\lambda, \mathrm{Exp}(c_\lambda), \beta_\lambda)$-bushes as in Theorem 1. Then*

$$\frac{1}{\psi^{-1}(\lambda)} Z_t^\lambda \to Z_t \qquad \text{in distribution, as } \lambda \to \infty, \text{ for all } t \geq 0,$$

where $(Z_t, t \geq 0)$ is a $\mathrm{CSBP}(\psi)$ starting from $Z_0 = \beta$, with ψ as in (14). If furthermore $\psi'(0) > -\infty$, then the convergence holds in the almost sure sense.

Proof. First consider $t = 0$. The initial population sizes are Poisson distributed with

$$\mathbb{E}(\exp\{-rZ_0^\lambda/\psi^{-1}(\lambda)\}) = \exp\{\beta\psi^{-1}(\lambda)(e^{-r/\psi^{-1}(\lambda)} - 1)\} \to e^{-\beta r},$$

as $\lambda \to \infty$, since $\psi^{-1}(\lambda) \to \infty$.

For $t>0$, the desired limiting distribution is characterised by (11) in terms of $u_t(r)$ and for $x=\beta$. If we integrate (12), we can identify $u_t(r)$ as the unique solution of

$$\int_{u_t(r)}^{r}\frac{dv}{\psi(v)}=t. \tag{22}$$

Consider $Z_t^\lambda/\psi^{-1}(\lambda)$ as a sum of a Poisson number of independent $\mathrm{GW}(q_\lambda,\mathrm{Exp}(c_\lambda))$-processes. Set $s=e^{-r/\psi^{-1}(\lambda)}$ and apply the branching property at the first branching time of a single $\mathrm{GW}(q_\lambda,\mathrm{Exp}(c_\lambda))$-tree to obtain for its population size X_t^λ at time t with $w_t^\lambda(s)=\mathbb{E}(s^{X_t^\lambda})$

$$w_t^\lambda(s)=se^{-c_\lambda t}+\int_0^t\sum_{k=0}^{\infty}q_\lambda(k)(\mathbb{E}(s^{X_{t-y}^\lambda}))^kc_\lambda e^{-c_\lambda y}dy=se^{-c_\lambda t}+\int_0^t\varphi_{q_\lambda}(w_z^\lambda(s))c_\lambda e^{-c_\lambda(t-z)}dz.$$

Now apply (14), multiply by $e^{c_\lambda t}$, differentiate with respect to t and rearrange to get

$$1=\frac{\psi^{-1}(\lambda)\frac{\partial}{\partial t}w_t^\lambda(s)}{\psi(\psi^{-1}(\lambda)(1-w_t^\lambda(s)))}\quad\Rightarrow\quad t=\int_{\psi^{-1}(\lambda)(1-w_t^\lambda(s))}^{\psi^{-1}(\lambda)(1-s)}\frac{dv}{\psi(v)}. \tag{23}$$

For $s=e^{-r/\psi^{-1}(\lambda)}$, we have $\psi^{-1}(\lambda)(1-s)\to r$ and by (22) also $\psi^{-1}(\lambda)(1-w_t^\lambda(s))\to u_t(r)$ and then

$$\mathbb{E}(\exp\{-rZ_t^\lambda/\psi^{-1}(\lambda)\})=\exp\{-\beta\psi^{-1}(\lambda)(1-w_t^\lambda(s))\}\to e^{-\beta u_t(r)}\qquad\text{as }\lambda\to\infty,$$

as required.

For the proof of almost sure convergence, recall our notation $\overline{\pi}_t(T)\subset\mathbb{U}$ for the population alive at time t of a tree T, which we will slightly abuse and also apply to bushes. Let $0\le\mu<\lambda$, $p=1-\mu/\lambda$ and $\mathcal{G}_{\lambda,t}=\sigma\left\{\overline{\pi}_s(B_{\lambda'}),\lambda'\ge\lambda,s\le t\right\}$. Now note that $Z_t^\lambda/\psi^{-1}(\lambda)$ is $\mathcal{G}_{\lambda,t}$-measurable, and that $\psi'(0)>-\infty$ ensures that $\mathbb{E}(Z_t^\mu)<\infty$. Then, a.s.,

$$\mathbb{E}\left(Z_t^\mu\middle|\,\mathcal{G}_{\lambda,t}\right)=\mathbb{E}\left(\sum_{u\in\overline{\pi}_t(B_\lambda)}\mathbf{1}_{\{\gamma_u(B_\lambda^{p-\mathrm{col}})=0\}}\middle|\,\mathcal{G}_{\lambda,t}\right)=\sum_{u\in\overline{\pi}_t(B_\lambda)}\mathbb{E}\left(\mathbf{1}_{\{\gamma_u(B_\lambda^{p-\mathrm{col}})=0\}}\middle|\,\mathcal{G}_{\lambda,t}\right),$$

since $Z_t^\mu=\#\overline{\pi}_t(B_\mu)$. By the branching property in Remark 16(c), applied to functions $f_u(\cdot)=\mathbf{1}_{\{\gamma_\emptyset(\cdot)=0\}}$, and $f_v(\cdot)\equiv1$ for $v\ne u$, we obtain, a.s.,

$$\sum_{u\in\overline{\pi}_t(B_\lambda)}\mathbb{E}\left(\mathbf{1}_{\{\gamma_u(B_\lambda^{p-\mathrm{col}})=0\}}\middle|\,\mathcal{G}_{\lambda,t}\right)=Z_t^{\lambda,\beta}\,\mathbb{P}\left(\gamma_\emptyset(B_\lambda^{p-\mathrm{col}})=0\right)=(1-g_\lambda(p))Z_t^{\lambda,\beta}.$$

According to (16), this shows that

$$\mathbb{E}\left(\frac{Z_t^\mu}{\psi^{-1}(\mu)}\middle|\,\mathcal{G}_{\lambda,t}\right)=\frac{Z_t^\lambda}{\psi^{-1}(\lambda)}\qquad\text{a.s.,}$$

which is the martingale property in the (decreasing) filtration $(\mathcal{G}_{\lambda,t},\lambda\ge0)$ that implies that $Z_t^\lambda/\psi^{-1}(\lambda)\to Z_t$ almost surely as $\lambda\to\infty$, see e.g. [42]. □

3 Growth of $\mathrm{GW}(q_\lambda,\kappa_\lambda,\beta_\lambda)$-bushes: lifetimes

Theorem 2 considers Galton-Watson bushes B_λ with general independent κ_λ-distributed lifetimes. The main statement beyond the exponential case of Theorem 1 is that consistency of a full family $(B_\lambda,\lambda\ge0)$ under Bernoulli leaf colouring requires κ_λ to be geometrically divisible.

3.1 Proof of Theorem 2

(i) $\Rightarrow$ (ii): Suppose, (i) holds. In particular, q_μ is then the $(1-\mu/\lambda)$-reduced offspring distribution associated with q_λ, for all $\mu<\lambda$. By Theorem 1, $\varphi_q(s)=s+\widetilde{\psi}(1-s)$, where $\widetilde{\psi}$ has the form (2). To be specific, let $c=1$, $\beta\in(0,\infty)$ and parametrise $(q_\lambda,\beta_\lambda)$ using ψ as in Section 2.3.1.

Consider a $\mathrm{GW}(q_\lambda,\kappa_\lambda)$-tree T_λ. By Remark 14, Lemma 13 applies. Using notation similar to the reduction procedure of Section 2.1.3 for $T=T_\lambda$, $p=1-1/\lambda$, $\iota=\iota_\lambda$, $\lambda\geq 1$, we can write $\zeta_\emptyset\sim\kappa_1$ as

$$\zeta_\emptyset=\sum_{v\in\iota_\lambda^{-1}(\emptyset)}\zeta_v^\lambda\overset{(\mathrm{d})}{=}\sum_{j=1}^{G^{(\alpha)}}X_\alpha^{(j)},\tag{24}$$

for $X_\alpha^{(j)}\sim\kappa_\lambda$, $j\geq 1$, and $G^{(\alpha)}\sim\mathrm{Geo}(\alpha)$ with $\alpha=\psi'(\psi^{-1}(1))/\psi'(\psi^{-1}(\lambda))$ independent. Specifically, consider $\Gamma_n=\{G_\emptyset^{(\alpha)}=n\}$, where $G_\emptyset^{(\alpha)}=\iota_\lambda^{-1}(\emptyset)$ in the notation of Section 2.3.2. On Γ_n, write $v_0=\emptyset$ and for $1\leq j\leq n-1$, denote by v_j the unique black child of v_{j-1}, then an n-fold inductive application of the two-colours branching property of Section 2.1.3, also summing over all offspring numbers and colour combinations as in Lemma 13, yields for $p=1-1/\lambda$

$$\begin{aligned}\mathbb{E}^{p-\mathrm{col}}_{q_\lambda\otimes\kappa_\lambda}\left[\prod_{j=0}^{n-1}k_j(\zeta_{v_j});\Gamma_n\,\middle|\,\gamma_\emptyset=0\right]&=\int_{(0,\infty)}k_0(z)\kappa_\lambda(dz)\varphi'_{q_\lambda}(g_\lambda(p))\mathbb{E}^{p-\mathrm{col}}_{q_\lambda\otimes\kappa_\lambda}\left[\prod_{j=0}^{n-2}k_{j+1}(\zeta_{v_j});\Gamma_{n-1}\,\middle|\,\gamma_\emptyset=0\right]\\&=\left(\prod_{j=0}^{n-1}\int_{(0,\infty)}k_j(z)\kappa_\lambda(dz)\right)(1-\alpha)^{n-1}\alpha.\end{aligned}$$

By Corollary 17, $\alpha\downarrow\psi'(\psi^{-1}(1))/\psi'(\infty)=1/\widetilde{\psi}'(\infty)$ as $\lambda\to\infty$. Therefore, we can write $\zeta_\emptyset\sim\kappa_1$ as in (24) for all $\alpha>1/\widetilde{\psi}'(\infty)$, with the convention $1/\widetilde{\psi}'(0)=0$ if $\widetilde{\psi}'(0)=\infty$. This yields (ii).

(ii) $\Rightarrow$ (i): Suppose, (ii) holds. By Theorem 1, the family $(q_\lambda,\lambda\geq 0)$ exists as required. By Theorem 1 and Remark 14, we can express φ_{q_λ} and α in terms of ψ as in Section 2.3.1 and Lemma 13, choosing $c=1$.

For $\lambda>1$, geometric divisibility of κ permits us to define κ_λ as the distribution of $X_\alpha^{(j)}$ for $\alpha=\psi'(\psi^{-1}(1))/\psi'(\psi^{-1}(\lambda))=1/\widetilde{\psi}'(\widetilde{\psi}^{-1}(\lambda\widetilde{\psi}(1)))>1/\widetilde{\psi}'(\infty)$. Now consider a $\mathrm{GW}(q_\lambda,\kappa_\lambda)$-tree and $p=1-1/\lambda$. Use notation from (i) $\Rightarrow$ (ii) and also set $v^*=v_{G_\emptyset^{(\alpha)}-1}$. On $\Upsilon_j=\{\nu_{v^*}-\gamma_{v^*1}-\cdots-\gamma_{v^*\nu_{v^*}}=j\}$ for $j\geq 2$, denote by $w_1,\ldots,w_j$ the black children of v^*. Then, by Remark 14 and repeated application of the two-colours branching property, we obtain

$$\begin{aligned}&\mathbb{E}^{p-\mathrm{col}}_{q_\lambda\otimes\kappa_\lambda}\left[\exp\left\{-r\sum_{m=0}^{G_\emptyset^{(\alpha)}-1}\zeta_{v_m}\right\}\prod_{i=1}^{j}f_i(\overline{\theta}_{w_i});\Upsilon_j\,\middle|\,\gamma_\emptyset=0\right]\\&=\sum_{n=1}^{\infty}\left(\int_{(0,\infty)}e^{-rz}\kappa_\lambda(dz)\right)^n(1-\alpha)^{n-1}\ \varphi^{(j)}_{q_\lambda}(g_\lambda(p))\frac{(1-g_\lambda(p))^{j-1}}{j!}\ \prod_{i=1}^{j}\mathbb{E}^{p-\mathrm{col}}_{q_\lambda\otimes\kappa_\lambda}\left[f_i\middle|\,\gamma_\emptyset=0\right].\end{aligned}$$

This is the branching property characterizing $\mathrm{GW}(q,\kappa)$, because the first term is the Laplace transform of a geometric sum with distribution κ, up to a factor of α, and for the middle term we identify the offspring distribution q using all the cancellations due to (14), (16) and (17)

$$\begin{aligned}&\sum_{j=2}^{\infty}s^j\frac{1}{\alpha}\frac{\varphi^{(j)}_{q_\lambda}(g_\lambda(p))}{j!}(1-g_\lambda(p))^{j-1}+\left(1-\sum_{j=2}^{\infty}\frac{1}{\alpha}\frac{\varphi^{(j)}_{q_\lambda}(g_\lambda(p))}{j!}(1-g_\lambda(p))^{j-1}\right)\\&=1+\frac{\varphi_{q_\lambda}(g_\lambda(p)+s(1-g_\lambda(p)))-1-(1-s)\varphi'_{q_\lambda}(g_\lambda(p))(1-g_\lambda(p))}{\alpha(1-g_\lambda(p))}=s+\frac{\psi(\psi^{-1}(1)(1-s))}{\psi^{-1}(1)\psi'(\psi^{-1}(1))}\end{aligned}$$

confirming that (q,κ) is the $(1-1/\lambda)$-reduced pair associated with $(q_\lambda,\kappa_\lambda)$. For $\mu<1$, set $\alpha=\psi'(\psi^{-1}(\mu))/\psi'(\psi^{-1}(1))$ and define κ_μ to be the distribution of

$$\sum_{j=1}^{G^{(\alpha)}} X^{(j)}, \qquad \text{for independent } X^{(j)}\sim\kappa,\ j\geq 1, \text{ independent of } G^{(\alpha)}\sim\mathrm{Geo}(\alpha).$$

As above, (q_μ,κ_μ) is the $(1-\mu)$-reduced pair associated with (q,κ). The reduction relation for $0\leq\mu<\lambda<\infty$ follows using transitivity of colouring reduction (see Remark 7(a)) for $0\leq\nu<\mu<\lambda<\infty$. Specifically, for $\mu=1$, this yields that $(B_\lambda^{(1-\nu/\lambda)-\mathrm{rdc}},B_\lambda)\overset{(\mathrm{d})}{=}(B_\nu,B_\lambda)$. For $\nu=1<\mu<\lambda$ and $\nu<\mu<\lambda=1$, this argument can be combined with the uniqueness of the divisor distribution (Lemma 11). This completes the proof of (ii) $\Rightarrow$ (i).

We identified $(\beta_\lambda,\lambda\geq 0)$ in (i) $\Rightarrow$ (ii). The same reasoning as in Remark 14 allows us to combine (i) here and Theorem 1 to see that (q,κ,β) is the $(1-1/\lambda)$-reduced triplet associated with $(q_\lambda,\kappa_\lambda,\beta_\lambda)$. The existence of $(B_\lambda,\lambda\geq 0)$ now follows from Kolmogorov's consistency theorem. Uniqueness of the families $(q_\lambda,\lambda\geq 0)$, $(\kappa_\lambda,\lambda\geq 0)$ and $(\beta_\lambda,\lambda\geq 0)$ for each $\beta=\beta_1\in(0,\infty)$ follows from the uniqueness results in Theorem 1 and as shown in (ii) $\Rightarrow$ (i). □

3.2 Reconstruction procedures and backbone decomposition

If $\kappa_\lambda(dz)=f_\lambda(z)dz$ is absolutely continuous for all $\lambda\geq 0$ and $\zeta=\zeta_1+\cdots+\zeta_{G^{(\alpha)}}\sim\kappa_\mu$ for independent $\zeta_j\sim\kappa_\lambda$ and $G^{(\alpha)}\sim\mathrm{Geo}(\alpha)$ for α as in Section 2.3.2, we find conditional joint distributions as in (19)

$$f_{\zeta_1,\ldots,\zeta_{n-1}|G^{(\alpha)}=n,\zeta=z}(y_1,\ldots,y_{n-1})=\frac{f_\lambda(z-\sum_{j=1}^{n-1}y_j)\prod_{j=1}^{n-1}f_\lambda(y_j)}{f_\lambda^{*(n)}(z)} \tag{25}$$

for $y_j>0$ with $y_1+\cdots+y_{n-1}<z$, $n\geq 1$, where $f_\lambda^{*(n)}$ is the nth convolution power of f_λ.

3.2.1 Reconstruction procedure for $\mathrm{GW}(q_\lambda,\kappa_\lambda,\beta_\lambda)$-bushes

For $B_\mu\sim\mathrm{GW}(q_\mu,\kappa_\mu,\beta_\mu)$ the procedure in Section 2.3.3 with $\mathrm{Exp}(c_\lambda)$ replaced by κ_λ and (19) replaced by (25) constructs $B_\lambda\sim\mathrm{GW}(q_\lambda,\kappa_\lambda,\beta_\lambda)$.

In the general case, one could use regular conditional distributions. Alternatively, we can adapt the procedure in Section 2.3.3 using the subordinator (random walk) representation of geometrically infinitely (finitely) divisible distributions as explained below.

3.2.2 Reconstruction procedure with subordinators in the case where κ is g.i.d.

In the g.i.d. case, let $\mathbb{H}=\bigcup_{\zeta\in(0,\infty)}\{\zeta\}\times\mathbb{D}([0,\zeta],[0,\infty))$, where $\mathbb{D}([0,\zeta],[0,\infty))$ is the set of functions $f\colon[0,\zeta]\to[0,\infty)$ that are right-continuous with left limits, equipped with the Borel sigma-algebra generated by the metric topology induced by

$$d((\zeta_1,f_1),(\zeta_2,f_2))=|\zeta_1-\zeta_2|+d_{\mathrm{Sk}}(f_1(\cdot\wedge\zeta_1),f_2(\cdot\wedge\zeta_2)),$$

where d_{Sk} is a metric that generates Skorohod's topology on $\mathbb{D}([0,\infty),[0,\infty))$, see e.g. [35, Section IV.1]. Consider a subordinator $(\sigma(t),t\geq 0)$, under $\mathbb{P}$, such that $\sigma(V)\sim\kappa_1$ for an independent $V\sim\mathrm{Exp}(1)$. Define the following measure on $\mathbb{N}_0\times\mathbb{H}$

$$\mathbb{Q}_\mu(\{j\}\times H)=q_\mu(j)\int_{(0,\infty)}\mathbb{P}((z,(\sigma(t),0\leq t\leq z))\in H)c_\mu e^{-c_\mu z}dz. \tag{26}$$

Now consider a bush $\overline{B}_\mu$ of $N_\mu \sim \mathrm{Poi}(\beta_\mu)$ random trees with distribution $\mathbb{P}_{\mathbb{Q}_\mu}$ as defined in the branching property of Section 2.1.2 for the measure $\mathbb{Q}_\mu$ just defined. Then the reconstruction procedure of Section 2.3.3 can be applied subdividing subordinator lifetimes $\zeta_u^{(i)} \sim \mathrm{Exp}(c_\mu)$ rather than directly the population lifetimes $\sigma_u^{(i)}(\zeta_u^{(i)}) \sim \kappa_\mu$. Also define

$$\sigma_{u,m}^{(i)}(t) = \sigma_u^{(i)}(\zeta_{u,1}^{(i)} + \cdots + \zeta_{u,m-1}^{(i)} + t) - \sigma_u^{(i)}(\zeta_{u,1}^{(i)} + \cdots + \zeta_{u,m-1}^{(i)}), \quad 0 \le t \le \zeta_{u,m}^{(i)}, 1 \le m \le G_u^{(\alpha,i)}.$$

The remainder of the procedure is easily adapted. The resulting $\overline{B}_\lambda$ is a bush of $N_\lambda \sim \mathrm{Poi}(\beta_\lambda)$ random trees with distribution $\mathbb{P}_{\mathbb{Q}_\lambda}$.

3.2.3 Reconstruction procedure with random walks in the case where $\psi'(\infty) < \infty$.

In the case where κ is geometrically divisible up to $\alpha_0(1) = \psi'(\psi^{-1}(1))/\psi'(\infty) > 0$, let $\mathbb{H} = \bigcup_{n\in\mathbb{N}}\{n\} \times [0,\infty)^{n+1}$ be the space of random walk paths. On $\mathbb{N}_0 \times \mathbb{H}$, consider

$$\mathbb{Q}_\mu^{\mathrm{RW}}(\{j\} \times H) = q_\mu(j) \sum_{n=1}^{\infty} \mathbb{P}((n, (\sigma(k), 0 \le k \le n)) \in H)(1 - \alpha_0(\mu))^{n-1}\alpha_0(\mu),$$

where, under $\mathbb{P}$, $(\sigma(k), k \ge 0)$ is a random walk with $\sigma(k+1) - \sigma(k) \sim \kappa_\infty$ i.i.d. In fact, in this case, the distribution κ_∞ exists since the Laplace transforms of κ_λ converge as $\lambda \to \infty$ to a completely monotone function continuous at zero. Then the reconstruction procedure of Section 2.3.3 can be applied subdividing geometric "random walk lifetimes" $G_u^{(i)} \sim \mathrm{Geo}(\alpha_0(\mu))$ into $G_u^{(i)} = G_{u,1}^{(i)} + \cdots + G_{u,G_u^{(\alpha,i)}}$, for independent $G_{u,m}^{(i)} \sim \mathrm{Geo}(\alpha_0(\lambda))$, $m \ge 1$, and $G_u^{(\alpha,i)} \sim \mathrm{Geo}(\alpha)$. Also define

$$\sigma_{u,m}^{(i)}(k) = \sigma_u^{(i)}(G_{u,1}^{(i)} + \cdots + G_{u,m-1}^{(i)} + k) - \sigma_u^{(i)}(G_{u,1}^{(i)} + \cdots + G_{u,m-1}^{(i)}), \qquad 0 \le k \le G_{u,m}^{(i)}.$$

The remainder of the procedure is easily adapted. The resulting $\overline{B}_\lambda$ is a bush of $N_\lambda \sim \mathrm{Poi}(\beta_\lambda)$ random trees with distribution $\mathbb{P}_{\mathbb{Q}_\lambda^{\mathrm{RW}}}$.

3.2.4 Backbone decomposition of supercritical Bellman-Harris processes

The reconstruction procedures that build a $\mathrm{GW}(q_\lambda, \kappa_\lambda, \beta_\lambda)$-bush B_λ from a $\mathrm{GW}(q_\mu, \kappa_\mu, \beta_\mu)$-bush give rise to decompositions of the associated Bellman-Harris process $Z_t^\lambda = \#\overline{\pi}_t(B_\lambda)$ along the $\mathrm{GW}(q_\mu, \kappa_\mu, \beta_\mu)$-bush B_μ. In the sequel, we will write $\mathrm{BH}(q_\lambda, \kappa_\lambda)$ for such a Bellman-Harris process and when we specify its initial distribution, all these individuals are taken with zero age.

In the supercritical case $\psi'(0) < 0$, note that the $\mathrm{GW}(q_{\lambda,\mathrm{red}}^{(1-\mu/\lambda)-\mathrm{rdc}}, \kappa_\lambda)$-trees with offspring distribution as in (20) that are grafted onto B_μ are subcritical for all $0 \le \mu < \lambda$. The case $\mu = 0$ in is at the heart of many decompositions in various settings, mainly continuous analogues with and without spatial motion, see [4, 11, 15, 16]. As an immediate consequence of our reconstruction procedures, we obtain a version of the backbone decomposition for Bellman-Harris processes.

Corollary 19. *Let ψ be a supercritical branching mechanism, B_0 a bush of $N_0 \sim \mathrm{Poi}(\beta\psi^{-1}(0))$ random trees with distribution $\mathbb{P}_{\mathbb{Q}_0}$ as in* (26)*. Subdivide each subordinator lifetime as in Section 3.2.2 to get a bush $\widehat{B}_0$. For each $u \in \widehat{B}_0$ independently, given $\widehat{\nu}_u = m$ children, consider a $\mathrm{BH}(q_{\lambda,\mathrm{red}}^{1-\mathrm{rdc}}, \kappa_\lambda)$-process $Z^{(u)}$ with $Z_0^{(u)}$ of distribution* (21)*. Also consider a $\mathrm{BH}(q_{\lambda,\mathrm{red}}^{1-\mathrm{rdc}}, \kappa_\lambda)$-process Z^{root} with $Z_0^{\mathrm{root}} \sim \mathrm{Poi}(\beta(\psi^{-1}(\lambda) - \psi^{-1}(0)))$. Then the process*

$$Z_t = \#\overline{\pi}_t(B_0) + Z_t^{\mathrm{root}} + \sum_{u\in\widehat{B}_0 :\, \omega_u \le t} Z_{t-\omega_u}^{(u)}$$

is a $\mathrm{BH}(q_\lambda, \kappa_\lambda)$-process with $Z_0 \sim \mathrm{Poi}(\beta\psi^{-1}(\lambda))$.

3.3 Limiting trees and branching processes as $\lambda \to \infty$

3.3.1 Convergence of trees: $\mathbb{R}$-tree representations, Lévy trees and snakes

A random marked tree $\overline{T}_\lambda$ with distribution $\mathbb{P}_{\mathbb{Q}_\lambda}$ as in Section 3.2.2 specifies marks $\zeta_u \sim \mathrm{Exp}(c_\lambda)$ as well as $\sigma_u(\zeta_u) \sim \kappa_\lambda$. Therefore, we can associate coupled trees

$$\begin{aligned} & T_\lambda^\circ = \{(u,\zeta_u)\colon (u,\zeta_u,\sigma_u) \in \overline{T}_\lambda\} \sim \mathrm{GW}(q_\lambda, \mathrm{Exp}(c_\lambda), \beta_\lambda) \\ \text{and} \quad & T_\lambda^\bullet = \{(u,\sigma_u(\zeta_u))\colon (u,\zeta_u,\sigma_u) \in \overline{T}_\lambda\} \sim \mathrm{GW}(q_\lambda, \kappa_\lambda, \beta_\lambda), \end{aligned}$$

that only differ in their lifetimes. In the same way, we obtain coupled bushes and consistent families $(B_\lambda^\circ, \lambda \geq 0)$ and $(B_\lambda^\bullet, \lambda \geq 0)$. Several other representations of $(T_\lambda^\circ, T_\lambda^\bullet)$ are natural. For an $\mathbb{R}$-tree representation

$$\mathcal{T}_\lambda^\circ = \{\rho\} \cup \bigcup_{u \in \tau_\lambda} \{u\} \times (\alpha_u^\circ, \omega_u^\circ]$$

of T_λ°, with root $\rho = (\emptyset, 0)$ and metric d given by $\mathrm{d}((v,s),(w,t)) = |t-s|$ for $v \preceq w$ or $w \preceq v$,

$$\mathrm{d}((uiv,s),(ujw,t)) = s + t - 2\omega_u^\circ \qquad \text{for } u,v,w \in \mathbb{U} \text{ and } i,j \in \mathbb{N},\ i \neq j,$$

see [11, Sect. 3.3], we can consider the measure

$$\mathcal{W}_\lambda(\{u\} \times (a,b]) = \sigma_u(b - \alpha_u^\circ) - \sigma_u(a - \alpha_u^\circ), \qquad \alpha_u^\circ \leq a < b \leq \omega_u^\circ, u \in \tau_\lambda,$$

which for $\sigma_u \overset{(\mathrm{d})}{=} \sigma$ as in (8)-(9), $u \in \tau_\lambda$, is of the form $\mathcal{W}_\lambda = d\mathrm{Leb} + \mathcal{R}_\lambda$, where $\mathcal{R}_\lambda$ is an infinitely divisible independently scattered random measure on $\mathcal{T}_\lambda^\circ$, in the sense of [28, 39, 41], with F-measure $\mathrm{Leb}\otimes\Lambda$, where Leb denotes Lebesgue measure on $\mathcal{T}_\lambda^\circ$. For technical convenience, we can embed $(\mathcal{T}_\lambda^\circ, \mathrm{d}, \rho)$ as a metric subspace of $\ell_1(\mathbb{N}) = \{x = (x_n)_{n \geq 1}\colon x_n \in [0,\infty), n \geq 1; ||x||_1 < \infty\}$, where $||x||_1 = \sum |x_n|$ is the ℓ_1-norm, with root $0 \in \ell_1(\mathbb{N})$. We can embed bushes $(\mathcal{B}_\lambda^\circ, \lambda \geq 0)$ consistently following [11, Remarks 4.9-4.10]: in the (sub)critical case, this family starts from $\varnothing$, is piecewise constant and evolves by adding single branches, so we can represent the jth branch, of length L_j° say, as $[[x(j), x(j) + L_j^\circ e_j]]$ for some $x(j) = (x_1(j), \ldots, x_{j-1}(j), 0, \ldots) \in \mathcal{B}_{\lambda_{j-1}}^\circ$ and $e_j = (0, \ldots, 0, 1, 0, \ldots)$ the jth coordinate vector in $\ell_1(\mathbb{N})$. Similarly, we can embed $(\mathcal{B}_\lambda^\bullet, \lambda \geq 0)$ for certain $L_j^\bullet = \sigma_j(L_j^\circ)$. The supercritical case can be handled using [11, Proposition 3.7].

For the $\ell_1(\mathbb{N})$-embedding, we can consider $\mathcal{W}_\lambda$ as a random measure on $\ell_1(\mathbb{N})$ (with support included in the embedded $\mathcal{B}_\lambda^\circ$). Then $\mathcal{B}_\lambda^{\circ,\mathrm{wt}} = (\mathcal{B}_\lambda^\circ, \mathcal{W}_\lambda)$ is a weighted $\mathbb{R}$-tree in the sense of [17, 19], in the (sub)critical case, but weak or vague convergence as $\lambda \to \infty$ are not appropriate since in the limit the measures become infinite on any ball around a point in some $\mathcal{B}_\lambda^\circ$. Nevertheless, consistency implies that in the case where convergence to a locally compact and separable Lévy bush $\mathcal{B}^\circ$ occurs in the Gromov-Hausdorff sense [11, Theorem 5.1], the random measures $\mathcal{W}_\lambda$ consistently build an infinitely divisible independently scattered random measure $\mathcal{W}$ on $\mathcal{B}^\circ \subset \ell_1(\mathbb{N})$ whose Poissonian component still has F-measure $\mathrm{Leb}|_{\mathcal{B}^\circ} \otimes \Lambda$, while the continuous component is still $d\mathrm{Leb}|_{\mathcal{B}^\circ}$ and where Leb is one-dimensional Lebesgue measure on $\ell_1(\mathbb{N})$.

For $x \in \mathcal{B}^\circ$ consider the path $[[0,x]] = \{y \in \mathcal{B}^\circ\colon y_n \leq x_n, n \geq 1\} = \{f_x(t)\colon 0 \leq t \leq ||x||_1\}$, where $||f_x(t)||_1 = t$, $0 \leq t \leq ||x||_1$. Then $\mathcal{S}\colon \mathcal{B}^\circ \to \mathbb{H}$ with $\mathbb{H}$ as in Section 3.2.2 given by $\mathcal{S}(x) = (||x||_1, (\mathcal{W}([[0, f_x(t)]]), 0 \leq t \leq ||x||_1))$ can be seen as a snake in the sense of [35, Section I.3.2], where the spatial motion here is a subordinator with characteristic pair (d, Λ) as in (9). See [25] for another setting where discontinuous snakes appear naturally.

Proposition 20. *In the setting of Theorem 2, if $\mathcal{B}^\circ$ is separable and locally compact, then $\mathcal{B}^\bullet$, the closure of $\bigcup \mathcal{B}_\lambda^\bullet$ in $\ell_1(\mathbb{N})$, is separable, but locally compact only if $\Lambda = 0$. Also, if $\mathcal{B}^\circ$ is furthermore bounded, then $\mathcal{B}^\bullet$ is bounded if and only if $\Lambda = 0$.*

Proof. Separability of $\mathcal{B}^\bullet$ is trivial as $\mathcal{B}^\bullet_\lambda$ is separable. Now argue conditionally given $(\mathcal{B}^\circ_\lambda)_{\lambda\ge 0}$. To show that boundedness and local compactness fail unless $\Lambda = 0$, first consider $\Lambda = \delta_h$, the Dirac measure in $h > 0$. It is not hard to show that the subtree of $\mathcal{B}^\circ$ above every vertex, except leaves, has infinite total length. As a consequence, there will be $\lambda_1 > 0$ for which $\mathcal{W}_{\lambda_1}$ has an atom x_1, and the subtree $\{y \in \mathcal{B}^\circ : x_1 \in [[0, y[[\}$ above x_1 has infinite length. Inductively, we find an increasing sequence $x_1 \prec x_2 \prec \cdots$ of atoms of $\mathcal{W}$ showing that $\mathcal{B}^\bullet$ is not bounded.

Now assume that $\mathcal{B}^\bullet$ is locally compact. By the Hopf-Rinow theorem [21], closed balls are then compact. Since $\mathcal{B}^\circ$ is locally compact, [11, Remark 5.1] implies that [11, Theorem 5.1] applies. In particular, $\psi'(\infty) = \infty$, so that $\mathcal{B}^\circ \setminus [[0, L^\circ_1 e_1]]$ has infinitely many connected components, each of infinite total length, so that they all contain atoms of $\mathcal{W}$. However, $L^\bullet_1 = \sigma_1(L^\circ_1)$ is bounded and for a large enough ball with $B(r) = \{x \in \mathcal{B}^\bullet : ||x||_1 \le r\} \supset [[0, L^\bullet_1 e_1]]$, the set $B(r+h)\setminus[[0, L^\bullet_1 e_1]]$ also has infinitely many connected components each exceeding diameter h. Considering a cover of $B(r+h)$ by open balls of radii less than $h/2$, there is no finite subcover, as each connected component needs at least one ball that does not intersect any other connected component, which contradicts the compactness of $B(r + h)$. So $\mathcal{B}^\bullet$ is not locally compact.

For any $\Lambda \ne 0$ and $h > 0$ with $\Lambda((h, \infty)) > 0$, let $\widetilde{\Lambda} = \Lambda((h, \infty))\delta_h$, couple $\mathcal{B}^\bullet$ and $\widetilde{\mathcal{B}}^\bullet$ and argue as above that $\widetilde{\mathcal{B}}^\bullet$ is not bounded nor locally compact, then deduce the same for $\mathcal{B}^\bullet$. □

Since $\mathcal{B}^\bullet$ is separable for a large class of branching mechanisms and lifetime subordinators, the framework of [19] can be used to further study these trees.

3.3.2 Superprocesses, backbones and convergence of Bellman-Harris processes

Sagitov [40] studied convergence of Bellman-Harris processes to certain non-Markovian CSBP whose distribution is best described via Markovian superprocesses that record residual lifetimes, see also [27]. Specifically, a (ξ, K, ψ)-superprocess [13] is a Markov process $M = (M_t, t \ge 0)$ on the space of finite Borel measures $\mathbb{M}([0, \infty))$ with transition semigroup characterised by

$$\mathbb{E}\left(\exp\left\{-\int_{[0,\infty)} f(z)M_t(dz)\right\}\middle|\, M_0 = m\right) = \exp\left\{-\int_{[0,t]} u_{t-z}(f)m(dz) - \int_{(t,\infty)} f(z-t)m(dz)\right\}, \quad (27)$$

for all $f : [0, \infty) \to [0, \infty)$ bounded continuous, where $u_t(f)$ is the unique (at least if $\psi'(0) > -\infty$) nonnegative solution of

$$u_t(f) + \int_{[0,t]} \psi(u_{t-s}(f))dH_s = \mathbb{E}(f(\xi_t)) \quad (28)$$

and $H_s = \mathbb{E}(K_s)$ is the renewal function of a strictly increasing subordinator or random walk σ, here in terms of the local time process $K_s = \inf\{t \ge 0 : \sigma(t) > s\}$, which is an additive functional of our particular choice $\xi_t = \sigma(K_t) - t$ of Markovian spatial motion. Since branching only occurs at the origin, M is a catalytic superprocess [9, 14]. We call the associated population size process $Z_t = M_t([0, \infty))$ a $\mathrm{CSBP}(K, \psi)$. Such processes appear as limits of Bellman-Harris processes, also in our setting. Unless stated otherwise, we understand $Z_0 = \beta$ as $M_0 = \beta\delta_0$.

Proposition 21. *Let $Z^\lambda_t = \#\overline{\pi}_t(B_\lambda)$ for a consistent family $(B_\lambda)_{\lambda\ge 0}$ of $\mathrm{GW}(q_\lambda, \kappa_\lambda, \beta_\lambda)$-bushes with branching mechanism ψ as in Theorem 2. Suppose that ψ is subcritical or critical, i.e. $\psi'(0) \ge 0$, and that $\psi'(\infty) = \infty$. Let σ be a subordinator as in Section 3.2.2. Then, for all $t \ge 0$*

$$\frac{1}{\psi^{-1}(\lambda)} Z^\lambda_t \longrightarrow Z_t \qquad \textit{almost surely, as } \lambda \to \infty,$$

where $(Z_t, t \ge 0)$ is a $\mathrm{CSBP}(K, \psi)$ starting from $Z_0 = \beta$.

Proof. The almost sure convergence follows from the same martingale argument as in Lemma 18. The identification of the limiting process follows from [40, Theorem 1], see [27] for a proof. Specifically, we consider the limit along the subsequence $(\lambda_n, n \geq 1)$ for which $\psi^{-1}(\lambda_n) = n$. $\square$

Proposition 4 states that the conclusion of Proposition 21 holds in the supercritical as well as $\psi'(\infty) < \infty$ cases. Sagitov [40] announces his convergence result to include the (finite-mean) supercritical case, but the proof in Kaj and Sagitov [27] only treats the subcritical and critical cases. They say that the supercritical case would require some further assumptions and additional work. In our less general setting, this is not difficult – our proof does not rely on [27].

Proof of Proposition 4. In the g.i.d. case $\psi'(\infty) = \infty$, we use Lemma 10 to represent κ_λ in terms of a strictly increasing subordinator σ as $\kappa_\lambda = \mathbb{P}(\sigma(V_\lambda) \in \cdot)$ for $V_\lambda \sim \mathrm{Exp}(\psi'(\psi^{-1}(\lambda)))$. We introduce Markovian measure-valued branching processes M^λ that do not record residual lifetimes of Z^λ, but what we may think of as limiting residual lifetimes (in the limit $\lambda \to \infty$),

$$M_t^\lambda = \sum_{u \in \overline{\pi}_t(B_\lambda^\bullet)} \delta_{\sigma_u(K^u_{t-\alpha_u^\bullet})-(t-\alpha_u^\bullet)} = \sum_{x \in \mathcal{B}_\lambda^\circ \colon \sigma(x-) \leq t \leq \sigma(x)} \delta_{\sigma(x)-t}, \qquad \text{where } \sigma(x-) = \mathcal{W}_\lambda([[0,x[[),$$

using notation as in Section 3.3.1 and also $K_s^u = \inf\{t \geq 0 \colon \sigma_u(t) > s\}$. By [27, Lemma 3] or [13, Formula (1.5)], for the Markov process $\xi_t = \sigma(K_t) - t$, the semigroup of M^λ is such that

$$\mathbb{E}\left(\exp\left\{-\int_{[0,\infty)} f(z) M_t^\lambda(dz)\right\}\middle| M_0^\lambda = m\right) = \exp\left\{-\int_{[0,t]} v_{t-z}^\lambda(f) m(dz) + \int_{(t,\infty)} f(z-t) m(dz)\right\},$$

for all $f \colon [0,\infty) \to [0,\infty)$ bounded continuous, where $v_t^\lambda(f)$ satisfies

$$1 - e^{-v_t^\lambda(f)} = \mathbb{E}\left(1 - e^{-f(\sigma(K_t)-t)}\right) - \int_{[0,t]} \left(\varphi_\lambda(e^{-v_{t-s}^\lambda(f)}) - e^{-v_{t-s}^\lambda(f)}\right) d\mathbb{E}(N_s^\lambda),$$

and where $N_s^\lambda = \#\{k \geq 1 \colon R_k^\lambda \leq s\}$ is the renewal process associated with a random walk with $R_k^\lambda - R_{k-1}^\lambda \sim \kappa_\lambda$. It is easily checked that $\mathbb{E}(N_s^\lambda) = H_s$. The remainder is straightforward (cf. [13, Section 1.2]). We apply (15) to see that $u_t^\lambda(f) = \psi^{-1}(\lambda)(1 - \exp\{-v_t^\lambda(f/\psi^{-1}(\lambda))\})$ satisfies

$$u_t^\lambda(f) + \int_{[0,t]} \psi(u_{t-s}^\lambda(f)) dH_s = \mathbb{E}\left(\psi^{-1}(\lambda)\left(1 - \exp\left\{-\frac{f(\sigma(K_t)-t)}{\psi^{-1}(\lambda)}\right\}\right)\right).$$

Uniqueness in (28) means that $u_t^\lambda(f) = u_t(f_\lambda)$, where $f_\lambda = \psi^{-1}(\lambda)(1 - e^{-f/\psi^{-1}(\lambda)}) \uparrow f$. By the Monotone Convergence Theorem, this implies for $N_\lambda \sim \mathrm{Poi}(\beta\psi^{-1}(\lambda))$ and $M_0^\lambda = N_\lambda \delta_0$ that

$$\begin{aligned}
\mathbb{E}\left(\exp\left\{-\int_{[0,\infty)} f(z) \frac{M_t^\lambda(dz)}{\psi^{-1}(\lambda)}\right\}\right) &= \exp\left\{-\beta\psi^{-1}(\lambda)(1 - v_t^\lambda(f/\psi^{-1}(\lambda)))\right\} = \exp\{-\beta u_t(f_\lambda)\} \\
&= \mathbb{E}\left(\exp\left\{-\int_{[0,\infty)} f_\lambda(z) M_t(dz)\right\}\middle| M_0 = \beta\delta_0\right) \\
&\to \mathbb{E}\left(\exp\left\{-\int_{[0,\infty)} f(z) M_t(dz)\right\}\middle| M_0 = \beta\delta_0\right) = \exp\{-\beta u_t(f)\}.
\end{aligned}$$

In particular, for $Z_t^\lambda = M_t^\lambda([0,\infty)$, we obtain $Z_t^\lambda/\psi^{-1}(\lambda) \to Z_t$, where Z is a $\mathrm{CSBP}(K,\psi)$. The martingale argument of Lemma 18 establishes almost sure convergence in the case $\psi'(0) > -\infty$.

In the finitely geometrically divisible case, we use bushes $\overline{B}_\lambda$ based on measures $\mathbb{Q}_\lambda^{\mathrm{RW}}$ and

$$M_t^\lambda = \sum_{u \in \overline{\pi}(B_\lambda^\bullet)} \delta_{\sigma_u(K^u_{t-\alpha_u^\bullet})-(t-\alpha_u^\bullet)}, \qquad \text{where } K_s^u = \inf\{k \geq 1 \colon \sigma_u(k) > s\}.$$

Then the argument above is easily adapted. $\square$

As an application, let us derive a backbone decomposition. This should be useful to deduce more general supercritical Bellman-Harris convergence results from subcritical results. Our present paper is not about superprocesses nor convergence of general triangular arrays, so we do not push for highest generality nor assumptions as in [27], but we would like to mention a now natural approach – in a sense to be made precise, convergence of supercritical processes is equivalent to convergence of backbones and convergence of associated subcritical processes.

We write $\mathbb{P}^r_{K,\psi}$ for the distribution of a CSBP(K,ψ) starting from $r \geq 0$. Just as for CSBP(ψ) in Section 2.2.2, we consider the sigma-finite measure $\Theta_{K,\psi}$ such that $\mathbb{P}^r_{K,\psi}$ is the distribution of a sum over a Poisson point process with intensity measure $r\Theta_{K,\psi}$.

Theorem 22 (Backbone decomposition for CSBP(K,ψ))**.** *Let ψ be a (non-explosive) supercritical, $\psi_0(r) = \psi(r + \psi^{-1}(0))$ the associated subcritical branching mechanism. Let $\overline{B}_0$ be a bush of $N_0 \sim \mathrm{Poi}(\beta\psi^{-1}(0))$ trees with distribution $\mathbb{P}_{\mathbb{Q}_0}$ as in (26), and, as in Section 3.3.1, $(\mathcal{B}_0^\circ, \mathcal{W}_0)$ a representation as a weighted $\mathbb{R}$-tree, $\sigma(x) = \mathcal{W}_0([[0,x]])$. Given $(\mathcal{B}_0^\circ, \mathcal{W}_0)$, consider*

- *points $(Z^x, x \in \mathcal{P})$ of a Poisson point process in $\mathbb{D}([0,\infty),[0,\infty))$ with intensity measure*

$$Q(df,dx) = \left(2a\Theta_{K,\psi_0}(df) + \int_{(0,\infty)} \mathbb{P}^r_{K,\psi_0}(df) r e^{-r\psi^{-1}(0)}\Pi(dr)\right) \mathrm{Leb}|_{\mathcal{B}_0^\circ}(dx),$$

- *extra points $(Z^x, x \in \mathrm{Br}(\mathcal{B}_0^\circ))$ independent of $(Z^x, x \in \mathcal{P})$ with distribution*

$$Q^{(l(x))}(df) = \frac{2a1_{\{l(x)=2\}}}{|\psi_0^{(l(x))}(0)|}\delta_0(df) + \int_{(0,\infty)} \mathbb{P}^r_{K,\psi_0}(df)\frac{r^{l(x)}e^{-r\psi^{-1}(0)}}{|\psi_0^{(l(x))}(0)|}\Pi(dr),$$

  where $l(x)+1$ is the number of connected components of $\mathcal{B}_0^\circ \setminus \{x\}$ and $\mathrm{Br}(\mathcal{B}_0^\circ)$ is the set of branchpoints $\{x \in \mathcal{B}_0^\circ \colon l(x) \geq 2\} \setminus \{0\}$,

- *and an extra point Z^0 independent of $(Z^x, x \in \mathcal{P} \cup \mathrm{Br}(\mathcal{B}_0^\circ))$ with distribution $\mathbb{P}^\beta_{K,\psi_0}$.*

Then the process $Z_t = \displaystyle\sum_{x \in \mathcal{P} \cup \mathrm{Br}(\mathcal{B}_0^\circ) \cup \{0\}} Z^x_{t-\sigma(x)}$ is a CSBP(K,ψ) starting from $Z_0 = \beta$.

Proof. Since Z is not a Markov process, we will deduce the theorem from the richer structure of a Markovian (ξ, K, ψ)-superprocesses M starting from $M_0 = m$. Slightly abusing notation, we consider M also under $\Theta_{K,\psi}$ and $\mathbb{P}^r_{K,\psi}$ and Q. Then the intensity measures and distributions in the bullet points specify a point process $(M^x, x \in \mathcal{P} \cup \mathrm{Br}(\mathcal{B}_0^\circ) \cup \{0\})$. From the exponential formula for Poisson point processes and from (27) for the subcritical branching mechanism ψ_0 with $u_t^{\mathrm{sub}}(f)$ associated via the analogue of (28), it is not hard to calculate

$$\mathbb{E}\left(\exp\left\{-\sum_{x\in\mathcal{P}}\int_{[0,\infty)} f(z)M^x_{t-\sigma(x)}(dz)\right\}\middle|\, \mathcal{B}_0^\circ, \mathcal{W}_0\right) = \exp\left\{-\int_{\mathcal{B}_0^\circ}(\psi_0'(u^{\mathrm{sub}}_{t-\sigma(x)}(f)) - \psi_0'(0))\mathrm{Leb}(dx)\right\},$$

$$\mathbb{E}\left(\exp\left\{-\sum_{x\in\mathrm{Br}(\mathcal{B}_0^\circ)}\int_{[0,\infty)} f(z)M^x_{t-\sigma(x)}(dz)\right\}\middle|\, \mathcal{B}_0^\circ, \mathcal{W}_0\right) = \prod_{x\in\mathrm{Br}(\mathcal{B}_0^\circ)} \frac{\psi_0^{(l)}(u^{\mathrm{sub}}_{t-\sigma(x)}(f))}{\psi_0^{(l)}(0)},$$

$$\mathbb{E}\left(\exp\left\{-\int_{[0,\infty)} f(z)M^0_t(dz)\right\}\middle|\, \mathcal{B}_0^\circ, \mathcal{W}_0\right) = \exp\left\{-\int_{[0,t]} u^{\mathrm{sub}}_{t-z}(f)m(dz) - \int_{(t,\infty)} f(z-t)m(dz)\right\},$$

using the convention $u_s^{\mathrm{sub}}(f) = 0$ for $s < 0$. It now suffices to show that the backbone decomposition of approximations M^λ of M, cf. Corollary 19, appropriately converges to these quantities.

Let us formulate that backbone decomposition in the current setting. Let $\mathbb{P}^{r,\lambda}_{\kappa,\psi_0}$ be the distribution of M^λ given $M_0^\lambda = N_\lambda^r\delta_0$ with $N_\lambda^r \sim \mathrm{Poi}(\psi^{-1}(\lambda)r)$. Then, given $(\mathcal{B}_0^\circ, \mathcal{W}_0)$, consider points $(M^{x,\lambda}, x \in \mathcal{P}^\lambda)$ of a Poisson point process with intensity measure

$$2a(\psi^{-1}(\lambda)-\psi^{-1}(0))\mathbb{P}(M^\lambda \in \cdot | M_0^\lambda = \delta_0) + \int_{(0,\infty)} \mathbb{P}^{r,\lambda}_{\kappa,\psi_0} r e^{-r\psi^{-1}(0)}\Pi(dr),$$

and

$$M^{x,\lambda} \sim \frac{2a1_{\{l(x)=2\}}}{|\psi_0^{(l(x))}(0)|}\delta_0 + \int_{(0,\infty)} \mathbb{P}^{r,\lambda}_{K,\psi_0} \frac{r^{l(x)}e^{-r\psi^{-1}(0)}}{|\psi_0^{(l(x))}(0)|}\Pi(dr) \quad \text{for } x \in \mathrm{Br}(\mathcal{B}_0^\circ) \text{ and } M^{0,\lambda} \sim \mathbb{P}^{\beta,\lambda}_{K,\psi_0}.$$

Then the analogous calculations yield for $u_t^{\lambda,\mathrm{sub}}(f) = \psi_0^{-1}(\lambda)(1-e^{-v_t^{\lambda,\mathrm{sub}}(f/\psi^{-1}(\lambda))})$

$$\mathbb{E}\left(\exp\left\{-\sum_{x\in\mathcal{P}^\lambda}\int_{[0,\infty)} f(z)\frac{M^{x,\lambda}_{t-\sigma(x)}(dz)}{\psi^{-1}(\lambda)}\right\}\middle| \mathcal{B}_0^\circ, \mathcal{W}_0\right) = \exp\left\{-\int_{\mathcal{B}_0^\circ}(\psi_0'(u^{\lambda,\mathrm{sub}}_{t-\sigma(x)}(f)) - \psi_0'(0))\mathrm{Leb}(dx)\right\},$$

$$\mathbb{E}\left(\exp\left\{-\sum_{x\in\mathrm{Br}(\mathcal{B}_0^\circ)}\int_{[0,\infty)} f(z)\frac{M^{x}_{t-\sigma(x)}(dz)}{\psi^{-1}(\lambda)}\right\}\middle| \mathcal{B}_0^\circ, \mathcal{W}_0\right) = \prod_{x\in\mathrm{Br}(\mathcal{B}_0^\circ)} \frac{\psi_0^{(l)}(u^{\lambda,\mathrm{sub}}_{t-\sigma(x)}(f))}{\psi_0^{(l)}(0)},$$

$$\mathbb{E}\left(\exp\left\{-\int_{[0,\infty)} f(z)\frac{M^0_t(dz)}{\psi^{-1}(\lambda)}\right\}\middle| \mathcal{B}_0^\circ, \mathcal{W}_0\right) = \exp\left\{-\int_{[0,t]} u^{\lambda,\mathrm{sub}}_{t-z}(f)m(dz) - \int_{(t,\infty)} f(z-t)m(dz)\right\},$$

But from the proof of Proposition 4, we know that $u_t^{\lambda,\mathrm{sub}}(f) \to u_t^{\mathrm{sub}}(f)$ as $\lambda \to \infty$, and this completes the proof. $\square$

We can specialise this backbone decomposition to the case $K_t = t$, when a $\mathrm{CSBP}(K,\psi)$ is simply a Markovian $\mathrm{CSBP}(\psi)$ and $\sigma(x) = d(0,x)$ is just the height of $x \in \mathcal{B}_0^\circ$. In this framework, and even with a spatial motion added, this decomposition was obtained recently by Berestycki et al. [4], generalising an analogous result of [11, Theorem 5.6] formulated in a context of Lévy trees.

4 Growth of $\mathrm{GWI}(q_\lambda, \kappa_\lambda, \eta_\lambda, \chi_\lambda)$-forests: immigration

Theorem 3 is about forests F_λ of $\mathrm{GW}(q_\lambda, \mathrm{Exp}(c_\lambda))$-trees arising from immigration of independent η_λ-distributed numbers of immigrants at $\mathrm{Exp}(h_\lambda)$-spaced times. The main statement beyond the no-immigration case of Theorem 1 is that consistency of a family $(F_\lambda, \lambda \geq 0)$ under Bernoulli leaf colouring relates η_λ to a continuous-state immigration mechanism ϕ. After some more general remarks, we focus on the Markovian case of exponential lifetimes and inter-immigration times.

4.1 A two-colours regenerative property and associated forest reduction

Let $F = (B(t), t \geq 0)$ be a $\mathrm{GWI}(q,\kappa,\eta,\chi)$-forest as defined in Section 2.1.4, specifically denote by $S_1 = \inf\{t \geq 0\colon B(t) \neq \partial\} \sim \chi$ the first immigration time, by $B(S_1) = (T^{(1)}_{(1)}, \ldots, T^{(1)}_{(N_1)})$ the bush of independent genealogical trees $T^{(1)}_{(j)} \sim \mathrm{GW}(q,\kappa)$, $j \geq 1$, of the $N_1 \sim \eta$ time-S_1 immigrants and by $F_{\mathrm{post}} = (B^{\mathrm{post}}(t), t \geq 0)$ the post-S_1 forest given by $B^{\mathrm{post}}(0) = \partial$ and $B^{\mathrm{post}}(t) = B(S_1 + t)$ for $t > 0$. It is immediate from the definition that F satisfies a *regenerative property* at S_1 in that $(S_1, B(S_1))$ is independent of F_{post} and $F_{\mathrm{post}} \overset{(\mathrm{d})}{=} F$, and that the distribution of $(S_1, B(S_1))$ as above together with this regenerative property characterises the distribution of F. Since colouring and reduction apply tree by tree, we obtain for the associated forest $F^{p-\mathrm{col}}$ of coloured trees $T^{(i),p-\mathrm{col}}_{(j)} \sim \mathbb{P}^{p-\mathrm{col}}_{q\otimes\kappa}$, $1 \leq j \leq N_i$, $i \geq 1$, a

Regenerative property of coloured $\mathrm{GWI}(q,\kappa,\eta,\chi)$**-forests**

(a) For all $n\geq 1$, $\varepsilon_j\in\{0,1\}$, and measurable functions k, f_j and G, $1\leq j\leq n$, we have

$$\mathbb{E}\left(k(S_1)G(F^{p-\mathrm{col}}_{\mathrm{post}})\prod_{j=1}^n f_j(T^{p-\mathrm{col}}_{(j)});N_1=n;\Big(\gamma_\emptyset(T^{p-\mathrm{col}}_{(1)}),\ldots,\gamma_\emptyset(T^{p-\mathrm{col}}_{(n)})\Big)=(\varepsilon_1,\ldots,\varepsilon_n)\right)$$

$$=\int_{(0,\infty)}k(z)\chi(dz)\eta(n)g(p)^{n_r}(1-g(p))^{n_b}\mathbb{E}(G(F^{p-\mathrm{col}}))\prod_{j=1}^n\mathbb{E}^{p-\mathrm{col}}_{q\otimes\kappa}[f_j|\gamma_\emptyset=\varepsilon_j],$$

where $n_r=\varepsilon_1+\cdots+\varepsilon_n$ and $n_b=n-n_r$ are the numbers of red and black colours.

(b) For $t\geq 0$, consider the post-t forest $F^{p-\mathrm{col}}_{\mathrm{post}-t}=(B^{p-\mathrm{col}}(t+s),s\geq 0)$ and the pre-t sigma-algebra $\mathcal{F}_t=\sigma\{B^{p-\mathrm{col}}(r),r\leq t\}$. Then for all measurable functions f_u, $u\in\mathbb{U}$, and G,

$$\mathbb{E}\left(G(F^{p-\mathrm{col}}_{\mathrm{post}-t})\prod_{u\in\overline{\pi}_t(F)}f_u(\overline{\theta}_{u,t}(F^{p-\mathrm{col}}))\middle|\mathcal{F}_t\right)=\mathbb{E}(G(F^{p-\mathrm{col}}))\prod_{u\in\overline{\pi}_t(F)}\mathbb{E}^{p-\mathrm{col}}_{q\otimes\kappa}[f_u(\overline{\theta}_{\emptyset,s})|\zeta_\emptyset>s]\Big|_{s=t-\alpha_u}.$$

As a trivial application of (a), we can calculate the probability that all immigrants are red

$$\mathbb{P}\left(\Big(\gamma_\emptyset(T^{p-\mathrm{col}}_{(1)}),\ldots,\gamma_\emptyset(T^{p-\mathrm{col}}_{(N_1)})\Big)=(1,\ldots,1)\right)=\sum_{n\geq 1}\eta(n)g(p)^n=\varphi_\eta(g(p)).$$

We deduce the distribution of the number of red immigrants given that all immigrants are red

$$\eta^{p-\mathrm{col}}_{\mathrm{red}}(m)=\begin{cases}\frac{\eta(m)g(p)^m}{\varphi_\eta(g(p))}, & \text{if } m\geq 1,\\ 0, & \text{if } m=0,\end{cases}\qquad\text{with generating function } \varphi_{\eta^{p-\mathrm{col}}_{\mathrm{red}}}(s)=\frac{\varphi_\eta(sg(p))}{\varphi_\eta(g(p))}.\quad(29)$$

Also by the regenerative property (a), the number $\widetilde{G}$ of immigrations until we see the first black immigrant is geometrically distributed with parameter $1-\varphi_\eta(g(p))$, i.e.

$$\mathbb{P}(\widetilde{G}=j)=\varphi_\eta(g(p))^{j-1}(1-\varphi_\eta(g(p))),\qquad j\geq 1.\tag{30}$$

Conditioning on having at least one black immigrant (probability $1-\varphi_\eta(g(p))$), we get for $\ell\geq 1$

$$\eta^{p-\mathrm{rdc}}(\ell)=\frac{1}{1-\varphi_\eta}(g(p))\sum_{m\geq 0}\binom{m+\ell}{m}\eta(m+\ell)(g(p))^m(1-g(p))^\ell=\frac{(1-g(p))^\ell\varphi_\eta^{(\ell)}(g(p))}{\ell!(1-\varphi_\eta(g(p))}.$$

Similarly, conditioning on having ℓ black immigrants, $\ell\geq 1$, we get for the number of red ones

$$\widetilde{\eta}_\ell(m)=\eta(m+\ell)\frac{(m+\ell)!}{m!}(g(p))^m\frac{1}{\varphi_\eta^{(\ell)}(g(p))},\qquad\text{for } m\geq 0.\tag{31}$$

These distributions have generating functions that we can express in terms of φ_η, for $s\in[0,1]$

$$\varphi_{\eta^{p-\mathrm{rdc}}}(s)=\frac{\varphi_\eta(g(p)+s(1-g(p)))-\varphi_\eta(g(p))}{1-\varphi_\eta(g(p))}\qquad\text{and}\qquad\varphi_{\widetilde{\eta}_\ell}(s)=\frac{\varphi_\eta^{(\ell)}(sg(p))}{\varphi_\eta^{(\ell)}(g(p))},\quad\ell\geq 1,$$

and as $\varphi_{\eta^{p-\mathrm{rdc}}}$ and φ_η are analytic, we can extend φ_η analytically to $[-g(p)/(1-g(p)),1]$. Evaluating at $s=v_p=-g(p)/(1-g(p))$, we get $\varphi_\eta(g(p))=-\varphi_{\eta^{p-\mathrm{rdc}}}(v_p)/(1-\varphi_{\eta^{p-\mathrm{rdc}}}(v_p))$, so

$$\varphi_\eta(r)=\frac{\varphi_{\eta^{p-\mathrm{rdc}}}(v_p+r(1-v_p))-\varphi_{\eta^{p-\mathrm{rdc}}}(v_p)}{1-\varphi_{\eta^{p-\mathrm{rdc}}}(v_p)},\qquad r\in[0,1].\tag{32}$$

As reduction preserves the Galton-Watson property for trees [11], we now see that a p-reduced $\mathrm{GWI}(q,\kappa,\eta,\chi)$-forest is a $\mathrm{GWI}(q^{p-\mathrm{rdc}},\kappa^{p-\mathrm{rdc}},\eta^{p-\mathrm{rdc}},\chi^{p-\mathrm{rdc}})$-forest. Specifically, $\eta^{p-\mathrm{rdc}}$ is as above, $\chi^{p-\mathrm{rdc}}$ the distribution of a geom$(1-\varphi_\eta(g(p)))$ sum of independent χ-distributed variables.

4.2 Growth of $\mathrm{GWI}(q_\lambda, \mathrm{Exp}(c_\lambda), \eta_\lambda, \mathrm{Exp}(h_\lambda))$-forests

4.2.1 Proof of Theorem 3

(i) $\Rightarrow$ (ii): Suppose, (i) holds. In particular, q_μ is then the $(1-\mu/\lambda)$-reduced offspring distribution associated with q_λ, for all $0 \le \mu < \lambda < \infty$. By Theorem 1, $\varphi_q(s) = s + \widetilde{\psi}(1-s)$, where $\widetilde{\psi}$ has the form (2). To be specific, let $c = 1$ and parametrise (q_λ, c_λ) using ψ as in Section 2.3.1.

Now consider the relationship between $\eta = \eta_\mu$ and η_λ for $\lambda > \mu = 1$. By the discussion above (32), we can extend φ_η analytically to $[-g_\lambda(1-1/\lambda)/(1-g_\lambda(1-1/\lambda)), 1]$, where, expressing as in (16), we have $g_\lambda(1-1/\lambda) = 1 - \psi^{-1}(1)/\psi^{-1}(\lambda) \to 1$ as $\lambda \to \infty$. Differentiating (32), we see that φ'_η has positive derivatives on $(-\infty, 1)$. Setting $\widetilde{\phi}(r) = 1 - \varphi_\eta(1-r)$, $r \ge 0$, the derivative $\widetilde{\phi}'$ is completely monotone on $(0, \infty)$ and, by Bernstein's theorem (see e.g. [18]), there exists a Radon measure $\widetilde{\Lambda}^*$ on $[0, \infty)$ such that

$$\widetilde{\phi}'(r) = \int_{[0,\infty)} e^{-rx} \widetilde{\Lambda}^*(dx) < \infty, \qquad r > 0.$$

From $\widetilde{\phi}(0) = 0$, we get integrability $\int_{(1,\infty)} x^{-1} \widetilde{\Lambda}^*(dx) < \infty$ and

$$\widetilde{\phi}(u) = \widetilde{\phi}(0) + \int_0^u \widetilde{\phi}'(r)dr = \widetilde{\Lambda}^*(\{0\})u + \int_{(0,\infty)} \frac{1-e^{-ux}}{x} \widetilde{\Lambda}^*(dx),$$

and, in particular, setting $\widetilde{d} = \widetilde{\Lambda}^*(\{0\})$ and $\widetilde{\Lambda}(dx) = x^{-1}\widetilde{\Lambda}^*|_{(0,\infty)}(dx)$ yields (ii).

(ii) $\Rightarrow$ (i): Now suppose that (ii) holds. According to Theorem 1, the family of offspring distribution $(q_\lambda, \lambda \ge 0)$ exists as required; furthermore, we can express φ_{q_λ} in terms of ψ as in Section 2.3.1, choosing $c = 1$.

By (16), we have $g_\lambda(1-\mu/\lambda) = 1 - \psi^{-1}(\mu)/\psi^{-1}(\lambda)$, so $v_{\mu,\lambda} = -g_\lambda(1-\mu/\lambda)/(1-g_\lambda(1-\mu/\lambda)) = 1 - \psi^{-1}(\lambda)/\psi^{-1}(\mu)$ for all $0 \le \mu < \lambda < \infty$. By (32) and the discussion above (32), the required immigration distributions must be of the following form, respectively for $\mu < 1$ and $\lambda > 1$

$$\begin{aligned} \varphi_{\eta_\mu}(s) = \frac{\varphi_\eta(g_1(1-\mu) + s(1-g_1(1-\mu))) - \varphi_\eta(g_1(1-\mu))}{1-\varphi_\eta(g_1(1-\mu))} &= 1 - \frac{\widetilde{\phi}((1-s)\psi^{-1}(\mu)/\psi^{-1}(1))}{\widetilde{\phi}(\psi^{-1}(\mu)/\psi^{-1}(1))}, \\ \varphi_{\eta_\lambda}(r) = \frac{\varphi_\eta(v_{1,\lambda} + r(1-v_{1,\lambda})) - \varphi_\eta(v_{1,\lambda})}{1-\varphi_\eta(v_{1,\lambda})} &= 1 - \frac{\widetilde{\phi}((1-r)\psi^{-1}(\lambda)/\psi^{-1}(1)}{\widetilde{\phi}(\psi^{-1}(\lambda)/\psi^{-1}(1))}. \end{aligned}$$

Since $\widetilde{\phi}'$ is completely monotone, simple differentiation yields that these functions are indeed generating functions of immigration distributions. Furthermore, η_μ is the $(1-\mu/\lambda)$-reduced immigration distribution of η_λ for all $0 \le \mu < \lambda < \infty$, by the transitivity of colouring reduction noted in Remark 7(a), which also applies to forests, since colouring and reduction are defined tree by tree. The full statement of (i) can now be obtained formally as in the proof of Theorem 2, with the simpler regenerative property here taking the role of the branching property there.

In the setting of (i) and (ii) for $c = c_1 \in (0, \infty)$, $h = h_1 \in (0, \infty)$, Kolmogorov's consistency theorem allows us to set up a consistent family $(F_\lambda, \lambda \ge 0)$ of $\mathrm{GWI}(q_\lambda, \mathrm{Exp}(c_\lambda), \eta_\lambda, \mathrm{Exp}(h_\lambda))$-forests. Uniqueness of (q_λ, c_λ), $\lambda \ge 0$, follows from Theorem 1. Uniqueness of $(\eta_\lambda, \lambda \ge 0)$ was noted in (ii) $\Rightarrow$ (i). Uniqueness of $(h_\lambda, \lambda \ge 0)$ follows from the relationship between inter-immigration times as geometric sums, where we calculate h_λ from (30) for $\lambda > 1 > \mu$ as

$$h_\lambda = (1 - \varphi_{\eta_\lambda}(g_\lambda(1-1/\lambda)))h = \frac{h}{\widetilde{\phi}(\psi^{-1}(\lambda)/\psi^{-1}(1))} \quad \text{and} \quad h_\mu = \frac{h}{\widetilde{\phi}(\psi^{-1}(\mu)/\psi^{-1}(1))}.$$

□

4.2.2 Freedom in parameterisation and standard choice

In analogy to Section 2.3.1, we can use a single function ϕ to replace $(\widetilde{\phi}, h)$ of Theorem 3 and parametrise $(\eta_\lambda, \mathrm{Exp}(h_\lambda))$, $\lambda \geq 0$, such that

$$\varphi_{\eta_\lambda}(v) = 1 - \frac{\phi(\psi^{-1}(\lambda)(1-v))}{\phi(\psi^{-1}(\lambda))} \qquad \text{and} \qquad h_\lambda = \phi(\psi^{-1}(\lambda)), \tag{33}$$

where ϕ is a linear transformation $\phi(s) = k_3\widetilde{\phi}(k_4 s)$ of $\widetilde{\phi}$. Specifically, we choose $k_3 = h$ and $k_4 = 1/\psi^{-1}(1)$. It is easy to check that this works, using $\widetilde{\phi}(1) = 1 - \eta(0) = 1$.

In this parameterisation, we can also express in terms of ψ and ϕ the remaining quantities studied in Section 4.1. E.g. (33) and (29) now yield the generating function of the pure-red immigration distribution

$$\varphi_{\eta^{(1-\mu/\lambda)-\mathrm{col}}_{\lambda,\mathrm{red}}}(s) = \frac{\phi(\psi^{-1}(\lambda)) - \phi(\psi^{-1}(\lambda) - s(\psi^{-1}(\lambda) - \psi^{-1}(\mu)))}{\psi^{-1}(\lambda) - \psi^{-1}(\mu)}. \tag{34}$$

The parameters of the geometric distributions (30) take a simple form that leads to a distinction of finite/infinite immigration rate $\alpha^{\mathrm{imm}}_{\mu,\lambda} = 1 - \varphi_{\eta_\lambda}(g_\lambda(1-\mu/\lambda)) = \phi(\psi^{-1}(\mu))/\phi(\psi^{-1}(\lambda))$, and

- $\alpha^{\mathrm{imm}}_{\mu,\lambda} \to \frac{\phi(\psi^{-1}(\mu))}{\phi(\infty)} > 0$ as $\lambda \to \infty$, if $\phi(\infty) < \infty$;
- $\alpha^{\mathrm{imm}}_{\mu,\lambda} \to 0$ as $\lambda \to \infty$, if $\phi(\infty) = \infty$.

With the formulas above we can formulate explicitly a reconstruction procedure.

4.2.3 Reconstruction procedure for $\mathrm{GWI}(q_\lambda, \mathrm{Exp}(c_\lambda), \eta_\lambda, \mathrm{Exp}(h_\lambda))$-forests

For $F_\mu \sim \mathrm{GWI}(q_\mu, \mathrm{Exp}(c_\mu), \eta_\mu, \mathrm{Exp}(h_\mu))$, we modify the steps of Section 2.3.3 to construct F_λ.

1. In every tree of F_μ, subdivide lifetimes as in Section 2.3.3 and hence construct a forest $\widehat{F}_\mu$.
2. In every tree of $\widehat{F}_\mu$, add further children and independent red trees as in Section 2.3.3 and hence construct a forest $\widehat{F}_\lambda$.
3. At every immigration time, given that there are $N^{(i)}_\mu = \ell$ immigrants in $\widehat{F}_\mu$, consider a random number $N^{(i),\mathrm{red}}_\lambda \sim \widetilde{\eta}_\ell$ of further immigrants as in (31), proceed as in Section 2.3.3 and superpose a further independent $\mathrm{GWI}(q^{(1-\mu/\lambda)-\mathrm{col}}_{\lambda,\mathrm{red}}, \mathrm{Exp}(c_\lambda), \eta^{(1-\mu/\lambda)-\mathrm{col}}_{\lambda,\mathrm{red}}, \mathrm{Exp}(h_\lambda - h_\mu))$-forest with distributions as in (20) and (34) to finally obtain F_λ.

4.2.4 Convergence of the population sizes: proof of Proposition 5

For convergence in distribution, we calculate the Laplace transform of $Y^\lambda_t/\psi^{-1}(\lambda)$, where

$$Y^\lambda_t = \overline{\pi}_t(F_\lambda) = \sum_{i=1}^{J^\lambda_t} \overline{\pi}_{t-S^\lambda_i}\left(B_\lambda(S^\lambda_i)\right) = \sum_{i=1}^{J^\lambda_t}\sum_{j=1}^{N^\lambda_i} \overline{\pi}_{t-S^\lambda_i}\left(T^{(i),\lambda}_{(j)}\right)$$

with notation as in and around (7), but with all quantities λ-dependent. We exploit that

$$E^\lambda(S^\lambda_i) = Z^{(i),\lambda} = \left(\overline{\pi}_t(B_\lambda(S^\lambda_i)), t \geq 0\right), \qquad E^\lambda(s) = 0, \quad s \notin \left\{S^\lambda_i, i \geq 1\right\},$$

is a Poisson point process with intensity measure h_λ times the distribution of a $\mathrm{GW}(q_\lambda, \mathrm{Exp}(c_\lambda))$-process $Z^{(1),\lambda}$ starting from $Z_0^{(1),\lambda} \sim \eta_\lambda$. By the exponential formula for Poisson point processes,

$$\mathbb{E}\left(\exp\left\{-rY_t^\lambda/\psi^{-1}(\lambda)\right\}\right) = \exp\left\{-h_\lambda \int_0^t \sum_{m=1}^\infty \eta_\lambda(m) \left(1 - \left(\mathbb{E}\left(s^{Z_{t-v}^\lambda}\right)\right)^m\right) dv\right\},$$

where $s = e^{-r/\psi^{-1}(\lambda)}$ and Z^λ is the population size of a single $\mathrm{GW}(q_\lambda, \mathrm{Exp}(c_\lambda))$-tree as in the proof of Lemma 18. Using notation and asymptotics from there, as well as (33), this equals

$$\exp\left\{-h_\lambda \int_0^t \left(1 - \varphi_{\eta_\lambda}\left(w_{t-v}^\lambda(s)\right)\right) dv\right\} = \exp\left\{-\int_0^t \phi\left(\psi^{-1}(\lambda)\left(1 - w_{t-v}^\lambda(s)\right)\right) dv\right\}.$$

Since $\psi^{-1}(\lambda)\left(1 - w_{t-v}^\lambda(s)\right) \to u_{t-v}(r)$, and, by (22) and (23), all these quantities are bounded by $\max\{r, \psi^{-1}(0)\}$, dominated convergence completes the proof of convergence in distribution.

Almost sure convergence follows by martingale arguments as in Lemma 18, using a version of the regenerative property (ii) of Section 4.1 rather than the version of the branching property (6) that we presented in Remark 16(c). $\square$

From Proposition 5 and Lemma 18 we deduce the analogous convergence result for GWI-processes starting from initial population sizes $Y_0^\lambda \sim \mathrm{Poi}(\beta_\lambda)$. The limiting CBI then has $Y_0 = \beta$.

4.3 Analogous results for $\mathrm{GWI}(q_\lambda, \kappa_\lambda, \eta_\lambda, \chi_\lambda)$-forests

Finally, let us combine Theorems 2 and 3 into a single statement and also deduce the analogous pattern for general inter-immigration distributions that now emerges naturally.

Corollary 23. *For a tuple (q, κ, η, χ) of offspring, lifetime, immigration and inter-immigration distributions, the following are equivalent:*

(i) *There are $(q_\lambda, \kappa_\lambda, \eta_\lambda, \chi_\lambda)_{\lambda\geq 0}$ with $(q_1, \kappa_1, \eta_1, \chi_1) = (q, \kappa, \eta, \chi)$ such that $(q_\mu, \kappa_\mu, \eta_\mu, \chi_\mu)$ is the $(1 - \mu/\lambda)$-reduced tuple associated with $(q_\lambda, \kappa_\lambda, \eta_\lambda, \chi_\lambda)$, for all $0 \leq \mu < \lambda < \infty$.*

(ii) *The generating functions φ_q of q and φ_η of η satisfy $\varphi_q(s) = s + \widetilde{\psi}(1-s)$ for some $\widetilde{\psi}$ of the form* (2) *and $\varphi_\eta(s) = 1 - \widetilde{\phi}(1 - s)$ for some $\widetilde{\phi}$ of the form* (3)*; κ is geometrically divisible for all $\alpha > 1/\widetilde{\psi}'(\infty)$ if $\widetilde{\psi}'(\infty) < \infty$, or for all $\alpha > 0$ if $\widetilde{\psi}'(\infty) = \infty$. Moreover, χ is also geometrically divisible*

- *for all $\alpha > 1/\widetilde{\phi}(\infty)$ if $\widetilde{\phi}(\infty) < \infty$;*
- *for all $\alpha > 0$ if $\widetilde{\phi}(\infty) = \infty$.*

In the setting of (i) *and* (ii)*, a consistent family $(F_\lambda)_{\lambda\geq 0}$ of $\mathrm{GWI}(q_\lambda, \kappa_\lambda, \eta_\lambda, \chi_\lambda)$-forests can be constructed such that $(F_\mu, F_\lambda) \stackrel{(\mathrm{d})}{=} (F_\lambda^{(1-\mu/\lambda)-\mathrm{rdc}}, F_\lambda)$ for all $0 \leq \mu < \lambda < \infty$.*

We omit the proof which is a straightforward combination of the proofs of Theorems 2 and 3.

Similarly, the convergence results of Proposition 4 and 5 find their analogue in this setting:

Corollary 24. *Let $(Y_t^\lambda, t \geq 0)$ be the population size process in the setting of Corollary 23. Then*

$$\frac{Y_t^\lambda}{\psi^{-1}(\lambda)} \to Y_t \qquad \text{in distribution as } \lambda \to \infty, \text{ for all } t \geq 0,$$

where $(Y_t, t \geq 0)$ is a $\mathrm{CBI}(K, \psi, \widehat{K}, \phi)$ with $Y_0 = 0$, for

- *branching mechanism ψ a linear transformations of $\widetilde{\psi}$ as in (15)*

- $K = (K_s, s \geq 0)$ *with* $\inf\{s \geq 0\colon K_s > V_\lambda\} \sim \kappa_\lambda$ *for* $V_\lambda \sim \mathrm{Exp}(c_\lambda)$ *with* c_λ *of Theorem 3,*
- *immigration mechanism* ϕ *a linear transformation of* $\widetilde{\phi}$ *as in (33)*
- $\widehat{K} = (\widehat{K}_s, s \geq 0)$ *with* $\inf\{s \geq 0\colon \widehat{K}_s > \widehat{V}_\lambda\} \sim \chi_\lambda$ *for* $\widehat{V}_\lambda \sim \mathrm{Exp}(h_\lambda)$ *with* h_λ *of Theorem 3.*

If furthermore $\psi'(0) > -\infty$ *and* $\phi'(0) < \infty$*, then the convergence holds in the almost sure sense.*

Like a $\mathrm{CBI}(\psi, \phi)$, a $\mathrm{CBI}(K, \psi, \widehat{K}, \phi)$ can be constructed from a Poisson point process $(E^s, s \geq 0)$ in $\mathbb{D}([0,\infty),[0,\infty))$ with intensity measure $d\Theta_{K,\psi} + \int_{(0,\infty)} \mathbb{P}^x_{K,\psi}\Lambda(dx)$, where (d, Λ) are the characteristics of ϕ in (3). Also consider an independent subordinator $\widehat{\sigma}$ or increasing random walk, in the $\phi(\infty) < \infty$ case, in fact $\widehat{K}_s = \inf\{t \geq 0\colon \widehat{\sigma}(t) > s\}$ clarifies the notation and the meaning of $\widehat{K}$. Then

$$Y_t = \sum_{s \leq \widehat{K}_t} E^s_{t-\widehat{\sigma}(s)}$$

is a $\mathrm{CBI}(K, \psi, \hat{K}, \phi)$ with $Y_0 = 0$. Like a $\mathrm{CSBP}(K, \psi)$, a $\mathrm{CBI}(K, \psi, \widehat{K}, \phi)$ is non-Markovian, but admits a Markovian representation (M, ϑ) that records residual lifetimes as well as residual times to the next immigration, with values in $\mathbb{M}([0,\infty)) \times [0,\infty)$, where $\vartheta_t = \widehat{\sigma}(\widehat{K}_t) - t$, so that

$$\begin{aligned}
&\mathbb{E}\left(\exp\left\{-\int_{[0,\infty)} f(z)M_t(dz) - r\vartheta_t\right\}\middle| M_0 = m, \vartheta_0 = s\right) \\
&\quad = \mathbb{E}\Bigg(\exp\Bigg\{-\int_{[0,t]} u_{t-z}(f)m(dz) - \int_{(t,\infty)} f(z-t)m(dz) \\
&\qquad\qquad - \int_{[0,t-s]} \phi(u_{t-s-z}(f))d\widehat{K}_z - r\left(\widehat{\sigma}(\widehat{K}_{(t-s)^+}) - (t-s)\right)\Bigg\}\Bigg),
\end{aligned}$$

where $u_t(f)$ is the unique nonnegative solution of (28). Then the process $Y_t = M_t([0,\infty))$ is a $\mathrm{CBI}(K, \psi, \hat{K}, \phi)$. We leave any further details including the proof of Corollary 24 to the reader.

References

[1] K. B. Athreya and P. E. Ney. *Branching processes*. Springer-Verlag, New York, 1972. Die Grundlehren der mathematischen Wissenschaften, Band 196.

[2] R. Bellman and T. Harris. On age-dependent binary branching processes. *Ann. of Math. (2)*, 55:280–295, 1952.

[3] R. Bellman and T. E. Harris. On the theory of age-dependent stochastic branching processes. *Proc. Nat. Acad. Sci. U. S. A.*, 34:601–604, 1948.

[4] J. Berestycki, A. E. Kyprianou, and A. Murillo. The prolific backbone for supercritical superdiffusions, 2009. Preprint, arXiv:0912.4736v1 [Math.PR].

[5] J. Bertoin. Subordinators: examples and applications. In *Lectures on probability theory and statistics (Saint-Flour, 1997)*, volume 1717 of *Lecture Notes in Math.*, pages 1–91. Springer, Berlin, 1999.

[6] C. Ceci and A. Gerardi. Conditional law of a branching process observing a subpopulation. *J. Appl. Probab.*, 39(1):112–122, 2002.

[7] C. Ceci, A. Gerardi, and L. Mazliak. Some results about stopping times of the marked tree space. *Teor. Veroyatnost. i Primenen.*, 41(3):578–590, 1996.

[8] B. Chauvin. Arbres et processus de Bellman-Harris. *Ann. Inst. H. Poincaré Probab. Statist.*, 22(2):209–232, 1986.

[9] D. Dawson and Fleischmann. A super-Brownian motion with a single point catalyst. *Stoch. Proc. Appl.*, 49:3–40, 1994.

[10] T. Duquesne. Continuum random trees and branching processes with immigration. *Stochastic Process. Appl.*, 119(1):99–129, 2009.

[11] T. Duquesne and M. Winkel. Growth of Lévy trees. *Probab. Theory Related Fields*, 139(3-4):313–371, 2007.

[12] T. Duquesne and M. Winkel. General growth by hereditary properties and an invariance principle for Galton-Watson real trees, 2010. Work in progress.

[13] E. Dynkin. Branching particle systems and superprocesses. *Ann. Probab.*, 19:1157–1194, 1991.

[14] E. Dynkin. Branching with a single point catalyst. *Proceedings of Symposia in Pure Mathematics*, 57:423–426, 1995.

[15] A. M. Etheridge and D. R. E. Williams. A decomposition of the $(1+\beta)$-superprocess conditioned on survival. *Proc. Roy. Soc. Edinburgh Sect. A*, 133(4):829–847, 2003.

[16] S. N. Evans and N. O'Connell. Weighted occupation time for branching particle systems and a representation for the supercritical superprocess. *Canad. Math. Bull.*, 37(2):187–196, 1994.

[17] S. N. Evans and A. Winter. Subtree prune and regraft: a reversible real tree-valued Markov process. *Ann. Probab.*, 34(3):918–961, 2006.

[18] W. Feller. *An introduction to probability theory and its applications. Vol. II.* Second edition. John Wiley & Sons Inc., New York, 1971.

[19] A. Greven, P. Pfaffelhuber, and A. Winter. Convergence in distribution of random metric measure spaces (Λ-coalescent measure trees). *Probab. Theory Related Fields*, 145(1-2):285–322, 2009.

[20] D. R. Grey. Asymptotic behaviour of continuous time, continuous state-space branching processes. *J. Appl. Probability*, 11:669–677, 1974.

[21] M. Gromov. *Metric structures for Riemannian and non-Riemannian spaces.* Modern Birkhäuser Classics. Birkhäuser Boston Inc., Boston, MA, english edition, 2007. Based on the 1981 French original, With appendices by M. Katz, P. Pansu and S. Semmes, Translated from the French by Sean Michael Bates.

[22] T. E. Harris. *The theory of branching processes.* Die Grundlehren der Mathematischen Wissenschaften, Bd. 119. Springer-Verlag, Berlin, 1963.

[23] J. Jacod and A. N. Shiryaev. *Limit theorems for stochastic processes*, volume 288 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences].* Springer-Verlag, Berlin, second edition, 2003.

[24] P. Jagers. Age-dependent branching processes allowing immigration. *Theor. Probability Appl.*, 13:225–236, 1968.

[25] S. Janson and J.-F. Marckert. Convergence of discrete snakes. *J. Theor. Probab.*, 18:615–645, 2005.

[26] M. Jiřina. Stochastic branching processes with continuous state space. *Czechoslovak Math. J.*, 8 (83):292–313, 1958.

[27] I. Kaj and S. Sagitov. Limit processes for age-dependent branching particle systems. *J. Theoret. Probab.*, 11(1):225–257, 1998.

[28] O. Kallenberg. *Foundations of Modern Probability.* Springer-Verlag, Berlin, second edition, 2001.

[29] K. Kawazu and S. Watanabe. Branching processes with immigration and related limit theorems. *Teor. Verojatnost. i Primenen.*, 16:34–51, 1971.

[30] L. B. Klebanov, G. M. Maniya, and I. A. Melamed. A problem of V. M. Zolotarev and analogues of infinitely divisible and stable distributions in a scheme for summation of a random number of random variables. *Teor. Veroyatnost. i Primenen.*, 29(4):757–760, 1984.

[31] A. Lambert. The genealogy of continuous-state branching processes with immigration. *Probab. Theory Related Fields*, 122(1):42–70, 2002.

[32] A. Lambert. Quasi-stationary distributions and the continuous-state branching process conditioned to be never extinct. *Elec. J. Probab.*, 12:420–446, 2007.

[33] J. Lamperti. Continuous state branching processes. *Bull. Amer. Math. Soc.*, 73:382–386, 1967.

[34] J. Lamperti. The limit of a sequence of branching processes. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 7:271–288, 1967.

[35] J.-F. Le Gall. *Spatial branching processes, random snakes and partial differential equations.* Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 1999.

[36] J. Neveu. Arbres et processus de Galton-Watson. *Ann. Inst. H. Poincaré Probab. Statist.*, 22(2):199–207, 1986.

[37] M. Pinski. Limit theorems for continuous state branching processes with immigration. *Bull. Amer. Math. Soc.*, 78:242–244, 1972.

[38] J. Pitman and M. Winkel. Growth of the Brownian forest. *Ann. Probab.*, 33(6):2188–2211, 2005.

[39] B. S. Rajput and J. Rosiński. Spectral representations of infinitely divisible processes. *Probab. Theory Related Fields*, 82(3):451–487, 1989.

[40] S. M. Sagitov. A Bellman-Harris branching process that starts with a large number of particles. *Dokl. Akad. Nauk SSSR*, 314(1):128–131, 1990.

[41] K. Urbanik and W. A. Woyczyński. A random integral and Orlicz spaces. *Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys.*, 15:161–169, 1967.

[42] D. Williams. *Probability with martingales.* Cambridge Mathematical Textbooks. Cambridge University Press, Cambridge, 1991.