SKELETAL STOCHASTIC DIFFERENTIAL EQUATIONS FOR CONTINUOUS-STATE BRANCHING PROCESSES

D. FEKETE,* ** *University of Exeter*

J. FONTBONA,*** *Universidad de Chile*

A. E. KYPRIANOU,**** *University of Bath*

Abstract

It is well understood that a supercritical continuous-state branching process (CSBP) is equal in law to a discrete continuous-time Galton–Watson process (the *skeleton* of *prolific individuals*) whose edges are dressed in a Poissonian way with immigration which initiates subcritical CSBPs (*non-prolific mass*). Equally well understood in the setting of CSBPs and superprocesses is the notion of a *spine* or *immortal particle* dressed in a Poissonian way with immigration which initiates copies of the original CSBP, which emerges when conditioning the process to survive eternally. In this article we revisit these notions for CSBPs and put them in a common framework using the well-established language of (coupled) stochastic differential equations (SDEs). In this way we are able to deal simultaneously with all types of CSBPs (supercritical, critical, and subcritical) as well as understanding how the skeletal representation becomes, in the sense of weak convergence, a spinal decomposition when conditioning on survival. We have two principal motivations. The first is to prepare the way to expand the SDE approach to the spatial setting of superprocesses, where recent results have increasingly sought the use of skeletal decompositions to transfer results from the branching particle setting to the setting of measure valued processes. The second is to provide a pathwise decomposition of CSBPs in the spirit of genealogical coding of CSBPs via Lévy excursions, albeit precisely where the aforesaid coding fails to work because the underlying CSBP is supercritical.

Keywords: Continuous-state branching process; stochastic differential equation; skeletal decomposition; spine decomposition

2010 Mathematics Subject Classification: Primary 60J80; 60H30; Secondary 60G99

1. Introduction

In this article we are interested in $X = (X_t, t \geq 0)$, a finite-mean, continuous-state branching process (CSBP). In particular, this means that X is a $[0, \infty)$-valued strong Markov process with absorbing state at zero and with law on $\mathbb{D}([0, \infty), \mathbb{R})$ (the space of càdlàg mappings from $[0, \infty)$ to $\mathbb{R}$) given by $\mathbb{P}_x$ for each initial state $x \geq 0$, such that $\mathbb{P}_{x+y} = \mathbb{P}_x * \mathbb{P}_y$. Here, $\mathbb{P}_{x+y} = \mathbb{P}_x * \mathbb{P}_y$ means that the sum of two independent processes, one issued from x and the other

issued from y, has the same law as the process issued from $x + y$. Its semigroup is characterised by the Laplace functional

$$\mathbb{E}_x(e^{-\theta X_t}) = e^{-xu_t(\theta)}, \qquad x, \theta, t \geq 0, \tag{1.1}$$

where $u_t(\theta)$ uniquely solves the evolution equation

$$u_t(\theta) + \int_0^t \psi(u_s(\theta)) \, ds = \theta, \qquad t \geq 0.$$

Here, we assume that the so-called branching mechanism ψ takes the form

$$\psi(\theta) = -\alpha\theta + \beta\theta^2 + \int_{(0,\infty)} (e^{-\theta x} - 1 + \theta x)\Pi(dx), \qquad \theta \geq 0, \tag{1.2}$$

where $\alpha \in \mathbb{R}$, $\beta \geq 0$, and Π is a measure concentrated on $(0, \infty)$ which satisfies $\int_{(0,\infty)} (x \wedge x^2)\Pi(dx) < \infty$. These restrictions on ψ are very mild and only exclude the possibility of having a non-conservative process or processes which have an infinite mean growth rate.

We also assume for convenience that $-\psi$ is not the Laplace exponent of a subordinator (i.e. a Bernstein function), thereby ruling out the case that X has monotone paths. It is easily checked that ψ is an infinitely smooth convex function on $(0, \infty)$ with at most two roots in $[0, \infty)$. More precisely, 0 is always a root; however, if $\psi'(0+) < 0$, then there is a second root in $(0, \infty)$.

The process X is henceforth referred to as a ψ-CSBP. It is easily verified that

$$\mathbb{E}_x[X_t] = xe^{-\psi'(0+)t}, \qquad t, x \geq 0. \tag{1.3}$$

The mean growth of the process is therefore characterised by $\psi'(0+)$, and accordingly we classify CSBPs by the value of this constant. We say that the ψ-CSBP is supercritical, critical, or subcritical accordingly as $-\psi'(0+) = \alpha$ is strictly positive, equal to zero, or strictly negative, respectively.

It is known that the process $(X, \mathbb{P}_x)$, $x > 0$, can also be represented as the unique strong solution to the stochastic differential equation (SDE)

$$X_t = x + \alpha \int_0^t X_{s-} \, ds + \sqrt{2\beta} \int_0^t \int_0^{X_{s-}} W(ds, du) + \int_0^t \int_0^\infty \int_0^{X_{s-}} r\tilde{N}(ds, dr, dv) \tag{1.4}$$

for $x > 0$ and $t \geq 0$, where $W(ds, du)$ is a white noise process on $(0, \infty)^2$ based on the Lebesgue measure $ds \otimes du$, and $N(ds, dr, dv)$ is a Poisson point process on $[0, \infty)^3$ with intensity $ds \otimes \Pi(dr) \otimes dv$. Moreover, we denote by $\tilde{N}(ds, dr, dv)$ the compensated measure of $N(ds, dr, dv)$. See [6, 7, 8] for this fact and further properties of the above SDEs.

Through the representation of a CSBP as either a strong Markov process whose semi-group is characterised by an integral equation, or as a solution to an SDE, there are three fundamental probabilistic decompositions that play a crucial role in motivating the main results in this paper. These concern CSBPs conditioned to die out, CSBPs conditioned to survive, and a path decomposition of the supercritical CSBPs.

1.1. CSBPs conditioned to die out

To understand what this means, let us momentarily recall that for all supercritical CSBPs (without immigration) the event $\{\lim_{t\to\infty} X_t = 0\}$ occurs with positive probability. Moreover, for all $x \geq 0$,

$$\mathbb{P}_x(\lim_{t\uparrow\infty} X_t = 0) = e^{-\lambda^* x},$$

where λ^* is the unique root on $(0, \infty)$ of the equation $\psi(\theta) = 0$. Note that ψ is strictly convex with the property that $\psi(0) = 0$ and $\psi(+\infty) = \infty$, thereby ensuring that the root $\lambda^* > 0$ exists; see Chapters 8 and 9 of [24] for further details. It is straightforward to show that the law of $(X, \mathbb{P}_x)$ conditional on the event $\{\lim_{t\uparrow\infty} X_t = 0\}$, say $\mathbb{P}_x^*$, agrees with the law of a ψ^*-CSBP, where

$$\psi^*(\theta) = \psi(\theta + \lambda^*).$$

See, for example, [45].

1.2. CSBPs conditioned to survive

The event $\{\lim_{t\to\infty} X_t = 0\}$ can be categorised further according to whether its intersection with $\{X_t > 0 \text{ for all } t \geq 0\}$ is empty or not. The classical work of Grey [21] distinguishes between these two cases according to an integral test. Indeed, the intersection is empty if and only if

$$\int^\infty \frac{1}{\psi(\theta)} \, d\theta < \infty. \tag{1.5}$$

If we additionally assume that $-\psi'(0+) = \alpha \leq 0$, that is to say, the process is critical or subcritical, then it is known that the notion of conditioning the process to stay positive can be made rigorous through a limiting procedure. More precisely, if we write

$$\zeta = \inf\{t > 0 : X_t = 0\},$$

then for all $A \in \mathcal{F}_t^X := \sigma(X_s : s \leq t)$ and $x > 0$,

$$\mathbb{P}_x^\uparrow(A) := \lim_{s\to\infty} \mathbb{P}_x(A \mid \zeta > t + s)$$

is well defined as a probability measure and satisfies the Doob h-transform

$$\left.\frac{d\mathbb{P}_x^\uparrow}{d\mathbb{P}_x}\right|_{\mathcal{F}_t^X} = e^{-\alpha t}\frac{X_t}{x}\mathbf{1}_{\{t<\zeta\}}.$$

In addition, $(X, \mathbb{P}_x^\uparrow)$, $x > 0$, has been shown to be equivalent in law to a process with a pathwise description which we give below. Before doing so, we need to introduce some more notation. To this end, define N^* to be a Poisson random measure on $[0, \infty)^2 \times \mathbb{D}([0, \infty), \mathbb{R})$ with intensity measure $ds \otimes r\Pi(dr) \otimes \mathbb{P}_r(d\omega)$. Moreover, when it exists, $\mathbb{Q}$ is the intensity, or 'excursion' measure, on the space $\mathbb{D}([0, \infty), \mathbb{R})$ which satisfies

$$\mathbb{Q}(1 - e^{-\theta\omega_t}) = -\frac{1}{x}\log\mathbb{E}_x(e^{-\theta X_t}) = u_t(\theta)$$

for $\theta, t \geq 0$. Here, the measure $\mathbb{Q}$ is the excursion measure on the space $\mathbb{D}([0, \infty), \mathbb{R})$ associated to $\mathbb{P}_x$, $x > 0$. See Theorems 3.10, 8.6, and 8.22 of [36] and [9, 12, 14, 30, 38] for further

details. We should note in particular that Theorem 3.10 in [36] gives a necessary and sufficient condition under which $\mathbb{Q}$ is well defined as an excursion entrance law, namely that $\lim_{\theta \to \infty} \psi'(\theta) = \infty$. This is automatically satisfied when $\beta > 0$. However, when $\beta = 0$, the reader will note that, in what we described below and elsewhere in the paper, the use of $\mathbb{Q}$ is not needed. We can accordingly build a Poisson point process N^c on $[0, \infty) \times \mathbb{D}([0, \infty), \mathbb{R})$ with intensity $2\beta \mathrm{d}s \otimes \mathbb{Q}(\mathrm{d}\omega)$. Then, for $x > 0$, $(X, \mathbb{P}_x^\uparrow)$ is equal in law to the stochastic process

$$\Lambda_t = X'_t + \int_0^t \int_{\mathbb{D}([0,\infty),\mathbb{R})} \omega_{t-s} N^c(\mathrm{d}s, \mathrm{d}\omega)$$
$$+ \int_0^t \int_0^\infty \int_{\mathbb{D}([0,\infty),\mathbb{R})} \omega_{t-s} N^*(\mathrm{d}s, \mathrm{d}r, \mathrm{d}\omega), \qquad t \geq 0, \qquad (1.6)$$

where X' has the law $\mathbb{P}_x$ and is independent of N^c and N^*, which are also independent of one another.

Intuitively, we can think of the process $(\Lambda_t, t \geq 0)$ as being the result of first running a subordinator

$$S_t = 2\beta t + \int_0^t \int_0^\infty r N^*(\mathrm{d}s, \mathrm{d}r), \qquad t \geq 0,$$

where we have slightly abused our notation and written $N^*(\mathrm{d}s, \mathrm{d}r)$, $s, r > 0$, in place of $\int_{\mathbb{D}([0,\infty),\mathbb{R})} N^*(\mathrm{d}s, \mathrm{d}r, \mathrm{d}\omega)$, $s, r > 0$. The subordinator $(S_t, t \geq 0)$ is usually referred to as the *spine*. To explain the formula (1.6), in a Poissonian way, we dress the spine with versions of X sampled under the excursion measure $\mathbb{Q}$. Moreover, at each jump of S we initiate an independent copy of X with initial mass equal to the size of the jump of S. See, for example (3.9) in [35] (4.3) in [34] (4.18) in [33], or the discussion in Section 12.3.2 of [24] or [36]. The reader is also referred, e.g., to [28, 29, 42] for further details of the notion of a spine.

It turns out that we may also identify the effect of the change of measure within the context of the SDE setting. In [19], it was shown that $(X, \mathbb{P}_x^\uparrow)$, $x > 0$, offers the unique strong solution to the SDE

$$X_t = x + \alpha \int_0^t X_{s-} \mathrm{d}s + \sqrt{2\beta} \int_0^t \int_0^{X_{s-}} W(\mathrm{d}s, \mathrm{d}u) + \int_0^t \int_0^\infty \int_0^{X_{s-}} r \tilde{N}(\mathrm{d}s, \mathrm{d}r, \mathrm{d}u)$$
$$+ \int_0^t \int_0^\infty r N^*(\mathrm{d}s, \mathrm{d}r) + 2\beta t, \qquad t \geq 0, \qquad (1.7)$$

where W, N, and $\tilde{N}$ are as in (1.4) and N^* is as above, and all noises are independent. See also [8] and [20].

1.3. Skeletal path decomposition of supercritical CSBPs

In [4, 5, 11] it was shown that the law of the process X, where X is defined by (1.4), can be recovered from a supercritical continuous-time Galton–Watson (GW) process, issued with a Poisson number of initial ancestors, and dressed in a Poissonian way using the law of the original process conditioned to become extinguished.

To be more precise, they showed that for each $x \geq 0$, $(X, \mathbb{P}_x)$ has the same law as the process $(\Lambda_t, t \geq 0)$ which has the following pathwise construction. First, sample from a continuous-time GW process with branching rate $q = \psi'(\lambda^*)$ and offspring distribution $\{p_k : k \geq 0\}$ such that its branching generator is given by

$$q\left(\sum_{k \geq 0} p_k r^k - r\right) = \frac{1}{\lambda^*} \psi(\lambda^*(1 - r)), \qquad r \in [0, 1]. \qquad (1.8)$$

This continuous-time GW process goes by the name of the *skeleton* and offers the genealogy of *prolific individuals*, that is, individuals who have infinite genealogical lines of descent (cf. [5]). With the particular branching generator given by (1.8), $p_0 = p_1 = 0$, and for $k \geq 2$, $p_k :=$ $p_k([0, \infty))$, where, for $r \geq 0$,

$$p_k(\mathrm{d}r) = \frac{1}{\lambda^* \psi'(\lambda^*)} \left\{ \beta(\lambda^*)^2 \delta_0(\mathrm{d}r) \mathbf{1}_{\{k=2\}} + (\lambda^*)^k \frac{r^k}{k!} \mathrm{e}^{-\lambda^* r} \Pi(\mathrm{d}r) \right\}.$$

If we denote the aforesaid GW process by $Z = (Z_t, t \geq 0)$ then we shall also insist that Z_0 has a Poisson distribution with parameter $\lambda^* x$. Next, thinking of the trajectory of Z as a graph, *dress* the life-lengths of Z in such a way that a ψ^*-CSBP is independently grafted on to each edge of Z at time t with rate

$$2\beta \mathrm{d}\mathbb{Q}^* + \int_0^\infty y \mathrm{e}^{-\lambda^* y} \Pi(\mathrm{d}y) \mathrm{d}\mathbb{P}_y^*.$$

Moreover, on the event that an individual dies and branches into $k \geq 2$ offspring, with probability $p_k(\mathrm{d}x)$, an additional independent ψ^*-CSBP is grafted on to the branching point with initial mass $x \geq 0$. The quantity Λ_t is now understood to be the total dressed mass present at time t together with the mass present at time t in an independent ψ^*-CSBP issued at time zero with initial mass x. Whilst it is clear that the pair (Z, Λ) is Markovian, it is less clear that Λ alone is Markovian. This must, however, be the case given the conclusion that Λ and X are equal in law. A key element in this respect is the non-trivial observation that, for each $t \geq 0$, the law of Z_t given Λ_t is that of a Poisson random variable with parameter $\lambda^* \Lambda_t$.

Such skeletal path decompositions for continuous-state branching processes, and spatial versions thereof, are by no means new. Examples include [4, 11, 15, 18, 22, 26, 27, 43, 44].

1.4. Overview

In this paper our objective is to understand the relationship between the skeletal decompositions of the type described above and the emergence of a spine on conditioning the process to survive. In particular, our tool of choice will be the use of SDE theory. The importance of this study is that it underlines a methodology that should carry over to the spatial setting of super-processes, where recent results have increasingly sought the use of skeletal decompositions to transfer results from the branching particle setting to the setting of measure-valued processes; cf. [13, 25, 39, 40]. In future work we hope to develop the SDE approach to skeletal decompositions in the spatial setting. We also expect this approach to be helpful in studying analogous decompositions in the setting of CSBPs with competition [3, 40]. Moreover, although our method takes inspiration from the genealogical coding of CSBPs by Lévy excursions, cf. [10], our approach appears to be applicable where the aforesaid method fails, namely supercritical processes.

2. Main results

In this section we summarise the main results of the paper. We have three main results. First, we provide a slightly more general family of skeletal decompositions in the spirit of [11], albeit with milder assumptions and that we use the language of SDEs. Second, taking lessons from this first result, we give a time-inhomogeneous skeletal decomposition, again using the language of SDEs, for both supercritical and (sub)critical CSBPs. Nonetheless, our proof will take inspiration from classical ideas on the genealogical coding of CSBPs through the exploration of associated excursions of reflected Lévy processes; see, for example, [10] and the references therein. Finally, our third main result shows that a straightforward limiting

procedure in the SDE skeletal decomposition for (sub)critical processes, which corresponds to conditioning on survival, reveals a weak solution to the SDE given in (1.7). It will transpire that conditioning the process to survive until later and later times is equivalent to 'thinning' the skeleton such that, in the limit, we get the spine decomposition. The limiting procedure also intuitively explains how the spine emerges in the conditioned process as a consequence of stretching out the skeleton in the SDE decomposition of the (sub)critical processes.

Before moving to the first main result, let us introduce some more notation. The reader will note that it is very similar to, but nonetheless subtly different from, previously introduced terms. Define the Esscher transformed branching mechanism $\psi_\lambda : \mathbb{R}_+ \to \mathbb{R}_+$ for $\theta \geq -\lambda$ and $\lambda \geq \lambda^*$ by

$$\psi_\lambda(\theta) = \psi(\theta + \lambda) - \psi(\lambda) = \psi'(\lambda)\theta + \beta\theta^2 + \int_{(0,\infty)} \left(e^{-\theta x} - 1 + \theta x\right) e^{-\lambda x} \Pi(\mathrm{d}x), \quad (2.1)$$

where

$$\psi'(\lambda) = -\alpha + 2\lambda\beta + \int_{(0,\infty)} \left(1 - e^{-\lambda x}\right) x \Pi(\mathrm{d}x) > 0.$$

This is the branching mechanism of a subcritical branching process on account of the fact that $-\psi'_\lambda(0+) = -\psi'(\lambda) < 0$. Heuristically speaking, given that $\lambda \mapsto \psi'(\lambda)$ is increasing, the ψ_λ-CSBP becomes more and more subcritical as λ increases.

Next, we need the continuous-time GW process parameterised by $\lambda \geq \lambda^*$, which has been seen before in, e.g., [11], and agrees with the process described by (1.8) when $\lambda = \lambda^*$. It branches at rate $\psi'(\lambda)$ and has a branching generator given by

$$F_\lambda(s) := \lambda^{-1} \psi((1 - s)\lambda), \qquad s \in [0, 1], \quad \lambda \geq \lambda^*.$$

That is to say, writing $F_\lambda(s)$ as in the left-hand side of (1.8), we now have $p_0 = \psi(\lambda)/\lambda\psi'(\lambda)$, $p_1 = 0$, and, for $k \geq 2$,

$$p_k = \frac{1}{\lambda\psi'(\lambda)} \left\{ \beta\lambda^2 \mathbf{1}_{\{k=2\}} + \int_{(0,\infty)} \frac{(\lambda r)^k}{k!} e^{-\lambda r} \Pi(\mathrm{d}r) \right\}.$$

We will also use the family $(\eta_k(\cdot))_{k\geq 0}$ of branch point immigration laws (conditional on the number of offspring at the branch point), where $\eta_1(\mathrm{d}r) = 0$, $x \geq 0$, and, otherwise,

$$\eta_k(\mathrm{d}r) = \frac{1}{p_k\lambda\psi'(\lambda)} \left\{ \psi(\lambda)\mathbf{1}_{\{k=0\}}\delta_0(\mathrm{d}r) + \beta\lambda^2\mathbf{1}_{\{k=2\}}\delta_0(\mathrm{d}r) + \mathbf{1}_{\{k\geq 2\}} \frac{(\lambda r)^k}{k!} e^{-\lambda r} \Pi(\mathrm{d}r) \right\} \quad (2.2)$$

for $r \geq 0$. Note, in particular, that when $\lambda > \lambda^*$ there is the possibility that no offspring are allowed. Since in this case some lines of descent are finite, the GW process no longer represents the prolific individuals.

Finally, we need to introduce a series of driving sources of randomness for the SDE which will appear in Theorem 2.1 below. Let $\mathbb{N}^0$ be a Poisson random measure on $[0, \infty)^3$ with intensity measure $\mathrm{d}s \otimes e^{-\lambda r}\Pi(\mathrm{d}r) \otimes \mathrm{d}v$, $\tilde{\mathbb{N}}^0$ be the associated compensated version of $\mathbb{N}^0$, $\mathbb{N}^1(\mathrm{d}s, \mathrm{d}r, \mathrm{d}j)$ be a Poisson point process on $[0, \infty)^2 \times \mathbb{N}$ with intensity $\mathrm{d}s \otimes re^{-\lambda r}\Pi(\mathrm{d}r) \otimes \sharp(\mathrm{d}j)$, and finally $\mathbb{N}^2(\mathrm{d}s, \mathrm{d}r, \mathrm{d}k, \mathrm{d}j)$ be a Poisson point process on $[0, \infty)^2 \times \mathbb{N}_0 \times \mathbb{N}$ with intensity $\psi'(\lambda)\mathrm{d}s \otimes \eta_k(\mathrm{d}r) \otimes p_k\sharp(\mathrm{d}k) \otimes \sharp(\mathrm{d}j)$, where $\mathbb{N}_0 = \{0\} \cup \mathbb{N}$ and $\sharp(\mathrm{d}\ell) = \sum_{i\in\mathbb{N}_0} \delta_i(\mathrm{d}\ell)$, $\ell \geq 0$,

denotes the counting measure on $\mathbb{N}_0$. As before, $W(\mathrm{d}s, \mathrm{d}u)$ will denote a white noise process on $(0, \infty)^2$ based on the Lebesgue measure $\mathrm{d}s \otimes \mathrm{d}u$.

Theorem 2.1. *Suppose that ψ corresponds to a supercritical branching mechanism (i.e. $\alpha > 0$) and $\lambda \geq \lambda^*$. Consider the coupled system of SDEs*

$$\begin{pmatrix} \Lambda_t \\ Z_t \end{pmatrix} = \begin{pmatrix} \Lambda_0 \\ Z_0 \end{pmatrix} - \psi'(\lambda) \int_0^t \begin{pmatrix} \Lambda_{s-} \\ 0 \end{pmatrix} \mathrm{d}s + \sqrt{2\beta} \int_0^t \int_0^{\Lambda_{s-}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} W(\mathrm{d}s, \mathrm{d}u)$$

$$+ \int_0^t \int_0^\infty \int_0^{\Lambda_{s-}} \begin{pmatrix} r \\ 0 \end{pmatrix} \tilde{\mathrm{N}}^0(\mathrm{d}s, \mathrm{d}r, \mathrm{d}\nu)$$

$$+ \int_0^t \int_0^\infty \int_1^{Z_{s-}} \begin{pmatrix} r \\ 0 \end{pmatrix} \mathrm{N}^1(\mathrm{d}s, \mathrm{d}r, \mathrm{d}j)$$

$$+ \int_0^t \int_0^\infty \int_0^\infty \int_1^{Z_{s-}} \begin{pmatrix} r \\ k-1 \end{pmatrix} \mathrm{N}^2(\mathrm{d}s, \mathrm{d}r, \mathrm{d}k, \mathrm{d}j)$$

$$+ 2\beta \int_0^t \begin{pmatrix} Z_{s-} \\ 0 \end{pmatrix} \mathrm{d}s, \qquad t \geq 0. \tag{2.3}$$

Equation (2.3) has a unique strong solution for arbitrary ($\mathcal{F}_0$-measurable) initial values $\Lambda_0 \geq 0$ and $Z_0 \in \mathbb{N}_0$ (where $\mathcal{F}_t := \sigma((\Lambda_s, Z_s) : s \leq t)$). Furthermore, under the assumption that Z_0 is an independent random variable which is Poisson distributed with intensity $\lambda \Lambda_0$, this unique solution satisfies the following:

(i) *For $t \geq 0$, conditional on $\mathcal{F}_t^\Lambda := \sigma(\Lambda_s : s \leq t)$, Z_t is Poisson distributed with intensity $\lambda \Lambda_t$.*

(ii) *The process $(\Lambda_t, t \geq 0)$ is Markovian and a weak solution to (1.4).*

(iii) *If $Z_0 = 0$, then $(\Lambda_t, t \geq 0)$ is a subcritical CSBP with branching mechanism ψ_λ.*

If we focus on the second element, Z, in the SDE (2.3), it can be seen that there is no dependency on the first element Λ. The converse is not true, however. Indeed, the stochastic evolution for Z is simply that of the continuous-time GW process with branching mechanism given by $F_\lambda(s)$, $s \in [0, 1]$. Given the evolution of Z, the process Λ here describes nothing more than the aggregation of a Poisson and branch-point dressing on Z together with an independent copy of a ψ_λ-CSBP. As is clear from (2.2), this results in the skeleton Z having the possibility of 'dead ends' (no offspring). Of course, if $\lambda = \lambda^*$ then this occurs with zero probability and the joint system of SDEs in (2.3) describes precisely the prolific skeleton decomposition. In the spirit of [11], albeit using different technology and in a continuum setting, Theorem 2.1 puts into a common framework a parametric family of skeletal decompositions for supercritical processes. Related work also appears in [1, 37].

Remark 2.1. Although we have assumed in the introduction that

$$\int_{(0,\infty)} (x \wedge x^2) \, \Pi(\mathrm{d}x) < \infty,$$

the reader can verify from the proof that this is in fact not needed. Indeed, suppose that we relax the assumption on Π to just $\int_{(0,\infty)} (1 \wedge x^2) \, \Pi(\mathrm{d}x) < \infty$, and we take the branching mechanism in the form

$$\psi(\theta) = -\alpha\theta + \beta\theta^2 + \int_{(0,\infty)} (\mathrm{e}^{-\theta x} - 1 + \theta x \mathbf{1}_{\{x<1\}}) \, \Pi(\mathrm{d}x), \qquad \theta \geq 0,$$

where $\psi'(0) < 0$ and

$$\int_{0+} \frac{1}{|\psi(\xi)|} \, d\xi = \infty$$

to ensure conservative supercriticality. Then the necessary adjustment we need to make occurs, for example, in (1.4), where jumps of size greater than or equal to 1 in the Poisson random measures N are separated out without compensation. However, the form of (2.3) remains the same, as all jumps of N^0 can be compensated.

Our objective, however, is to go further and demonstrate how the SDE approach can also apply in the finite-horizon setting. We do this below, but we should remark that the skeletal decomposition is heavily motivated by the description of the CSBP genealogy using the so-called height process in [10]. Indeed, for (sub)critical CSBPs we may consider the conclusion of Theorem 2.2 below as a rewording thereof. However, as the proof does not rely on the CSBP being (sub)critical, the same result holds in the supercritical case. Thus, Theorem 2.2 is also a time-inhomogeneous version of Theorem 2.1 for supercritical CSBPs, which setting was not discussed in [10].

Assume that ψ is a branching mechanism that satisfies Grey's condition (1.5). We fix a time marker $T > 0$ and we want to describe a coupled system of SDEs in the spirit of (2.3) in which the second component describes prolific genealogies to the time horizon T. In other words, our aim is to provide an SDE decomposition of the CSBP along those individuals in the population who have a descendant at time T.

To this end, recall that $(u_t(\theta), t \geq 0)$ is given by (1.1) and accordingly, for $t \geq 0$, $u_t(\infty) = -x^{-1} \log \mathbb{P}_x(X_t = 0)$ gives the rate at which extinction has occurred by time t. We need a Poisson random measure N_T^0 on $[0, T) \times [0, \infty)^2$ with intensity $ds \otimes e^{-u_{T-s}(\infty)r} \Pi(dr) \otimes dv$, a Poisson process N_T^1 on $[0, T) \times [0, \infty) \times \mathbb{N}_0$ with intensity $ds \otimes re^{-u_{T-s}(\infty)r} \Pi(dr) \otimes \sharp(dj)$, and a Poisson process $N_T^2(ds, dr, dk, dj)$ on $[0, T) \times [0, \infty) \times \mathbb{N}_0 \times \mathbb{N}$ with intensity

$$\left\{ \frac{u_{T-s}(\infty)\psi'(u_{T-s}(\infty)) - \psi(u_{T-s}(\infty))}{u_{T-s}(\infty)} \right\} ds \otimes \eta_k^{T-s}(dr) \otimes p_k^{T-s}\sharp(dk) \otimes \sharp(dj),$$

where, for $k \geq 2$,

$$\eta_k^{T-s}(dr) = \frac{\beta u_{T-s}^2(\infty)\mathbf{1}_{\{k=2\}}\delta_0(dr) + (u_{T-s}(\infty)r)^k \, e^{-u_{T-s}(\infty)r} \Pi(dr)/k!}{p_k^{T-s} \left(u_{T-s}(\infty)\psi'(u_{T-s}(\infty)) - \psi(u_{T-s}(\infty))\right)}, \qquad r \geq 0, \quad (2.4)$$

p_k^{T-s} is such that $p_0^{T-s} = p_1^{T-s} = 0$, and the remaining probabilities are computable by insisting that $\eta_k^{T-s}(\cdot)$ is itself a probability distribution for each $k \geq 2$.

Theorem 2.2. *Suppose that ψ corresponds to a branching mechanism which satisfies Grey's condition (1.5). Fix a time horizon $T > 0$ and consider the coupled system of SDEs*

$$\begin{pmatrix} \Lambda_t^T \\ Z_t^T \end{pmatrix} = \begin{pmatrix} \Lambda_0^T \\ Z_0^T \end{pmatrix} - \int_0^t \psi'(u_{T-s}(\infty)) \begin{pmatrix} \Lambda_{s-}^T \\ 0 \end{pmatrix} ds + \sqrt{2\beta} \int_0^t \int_0^{\Lambda_{s-}^T} \begin{pmatrix} 1 \\ 0 \end{pmatrix} W(ds, du)$$

$$+ \int_0^t \int_0^\infty \int_0^{\Lambda_{s-}^T} \begin{pmatrix} r \\ 0 \end{pmatrix} \tilde{N}_T^0(ds, dr, dv)$$

$$+ \int_0^t \int_0^\infty \int_1^{Z_{s-}^T} \begin{pmatrix} r \\ 0 \end{pmatrix} N_T^1(ds, dr, dj)$$

$$+ \int_0^t \int_0^\infty \int_0^\infty \int_1^{Z_{s-}^T} \binom{r}{k-1} \mathbb{N}_T^2(\mathrm{d}s, \mathrm{d}r, \mathrm{d}k, \mathrm{d}j)$$

$$+ 2\beta \int_0^t \binom{Z_{s-}^T}{0} \mathrm{d}s, \qquad 0 \le t < T. \tag{2.5}$$

Equation (2.5) *has a unique strong solution for arbitrary* ($\mathcal{F}_0^T$-*measurable*) *initial values* $\Lambda_0^T \ge 0$ *and* $Z_0^T \in \mathbb{N}_0$ (*where* $\mathcal{F}_t^T := \sigma((\Lambda_s^T, Z_s^T) : s \le t)$, $t < T$). *Furthermore, under the assumption that* Z_0^T *is an independent random variable which is Poisson distributed with intensity* $u_T(\infty)\Lambda_0^T$, *this unique solution satisfies the following:*

(i) *For* $T > t \ge 0$, *conditional on* $\mathcal{F}_t^{\Lambda^T} := \sigma(\Lambda_s^T : s \le t)$, Z_t^T *is Poisson distributed with intensity* $u_{T-t}(\infty)\Lambda_t^T$.

(ii) *The process* $(\Lambda_t^T, 0 \le t < T)$ *is Markovian and a weak solution to* (1.4).

(iii) *Conditional on* $\{Z_0^T = 0\}$, *the process* $(\Lambda_t^T, 0 \le t < T)$ *corresponds to a weak solution to* (1.4) *conditioned to become extinct by time T.*

The SDE evolution in Theorem 2.2 mimics the skeletal decomposition in (2.3), albeit that the different components in the decomposition are time dependent. Putting the SDE representation aside, such time-varying skeletons have been observed in, e.g., [10, 15]. We note that the underlying skeleton Z^T can be thought of as a time-inhomogeneous GW process (a T-prolific skeleton) such that, at time $s < T$, its branching rate is given by

$$q^{T-s} := \frac{u_{T-s}(\infty)\psi'(u_{T-s}(\infty)) - \psi(u_{T-s}(\infty))}{u_{T-s}(\infty)} \tag{2.6}$$

and its offspring distribution is given by $\{p_k^{T-s} : k \ge 0\}$. This has the feature that the branching rate explodes towards the time horizon T. To see why, we can appeal to (1.1), and note that

$$\mathbb{P}_x[X_t = 0] = \mathrm{e}^{-u_t(\infty)x}, \qquad x, t > 0,$$

and hence $\lim_{t \to 0} u_t(\infty) = \infty$. Moreover, we can easily verify from (1.2) that

$$\lim_{\lambda \to \infty} [\lambda \psi'(\lambda) - \psi(\lambda)]/\lambda = \infty.$$

Together, these facts imply the explosion of (2.6) as $s \to T$.

We also note from the integrals involving $\mathbb{N}_T^1$ and $\mathbb{N}_T^2$ that there is mass immigrating off the space-time trajectory of Z^T. Moreover, once mass has immigrated, the first four terms of (2.5) show that it evolves as a time-inhomogeneous CSBP.

Note that in the supercritical setting $u_{T-t}(\infty)$ converges to λ^* for all $t > 0$ as $T \to \infty$. This intuitively means that when T goes to ∞, one can recover the prolific skeleton decomposition of Theorem 2.1 from the time-inhomogeneous one of Theorem 2.2.

Finally, with the finite-horizon SDE skeletal decomposition in Theorem 2.2 we may now turn our attention to understanding what happens when we observe the solution to (2.5) in the (sub)critical case on a finite time horizon $[0, t_0]$, and we condition on there being at least one T-prolific genealogy while letting $T \to \infty$.

Theorem 2.3. *Suppose that* ψ *is a critical or subcritical branching mechanism such that Grey's condition* (1.5) *holds. Suppose, moreover, that* $((\Lambda_t^T, Z_t^T), 0 \le t < T)$ *is a weak solution to* (2.5) *and that* Z_0^T *is an independent random variable which is Poisson distributed with*

intensity $u_T(\infty)\Lambda_0^T$. Then, conditional on the event $Z_0^T > 0$, in the sense of weak convergence with respect to the Skorokhod topology on $\mathbb{D}([0, \infty), \mathbb{R}^2)$, for all $t_0 > 0$,

$$((\Lambda_t^T, Z_t^T), 0 \le t \le t_0) \to ((X_t^\uparrow, 1), 0 \le t \le t_0)$$

as $T \to \infty$, where $X^\uparrow$ is a weak solution to (1.7).

Theorem 2.3 puts the phenomena of spines and skeletons in the same framework. Roughly speaking, any subcritical branching population contains a naturally embedded skeleton which describes the 'fittest' genealogies. In our setting 'fittest' means surviving until time T, but other notions of fitness can be considered, especially when one introduces a spatial type to mass in the branching process. For example, in [22] a branching Brownian motion in a strip is considered, where 'fittest genealogies' pertains to those lines of descent which survive in the strip for all eternity. Having at least one line of descent in the skeleton corresponds to the event of survival. Thus, conditioning on survival as we make the survival event itself increasingly unlikely, e.g. by taking $T \to \infty$ in our model or taking the width of the strip down to a critical value in the branching Brownian motion model, the natural stochastic behaviour of the skeleton is to thin down to a single line of descent. This phenomenon was originally observed in [15], where the scaling limit of a GW process conditioned on survival is shown to converge to the immortal particle decomposition of the $(1 + \beta)$-superprocess conditioned on survival.

The remainder of the paper is structured as follows. In the next section we explain the heuristic behind how (1.4) can be decoupled into components that arise in (2.3). The heuristic is used in Section 4, where the proof of Theorem 2.1 is given. In this sense our proof of Theorem 2.1 has the feel of a 'guess-and-verify' approach. In Section 5, again in the spirit of a 'guess-and-verify' approach, we use ideas from the classical description of the exploration process of CSBPs in, e.g., [10] to provide the heuristic behind the mathematical structures that lie behind the proof of Theorem 2.2. Given the similarity of this proof to that of Theorem 2.1, it is sketched in Section 6. Finally, in Section 7 we provide the proof of Theorem 2.3.

3. Thinning of the CSBP SDE

In this section we will perform an initial manipulation of the SDE (1.4), which we will need in order to make comparative statements for Theorems 2.1 and 2.2. To this end, we will introduce some independent marks on the atoms of the Poisson process N driving (1.4) and use them to thin out various contributions to the SDE evolution.

Denote by $(t_i, r_i, v_i : i \in \mathbb{N})$ some enumeration of the atoms of N and recall that $\mathbb{N}_0 = \{0\} \cup \mathbb{N}$. By enlarging the probability space, we can introduce an additional mark to atoms of N, say $(k_i : i \in \mathbb{N})$, resulting in an 'extended' Poisson random measure,

$$\mathcal{N}(\mathrm{d}s, \mathrm{d}r, \mathrm{d}v, \mathrm{d}k) := \sum_{i\in\mathbb{N}} \delta_{(t_i, r_i, v_i, k_i)}(\mathrm{d}s, \mathrm{d}r, \mathrm{d}v, \mathrm{d}k)$$

on $[0, \infty)^3 \times \mathbb{N}_0$ with intensity

$$\mathrm{d}s \otimes \Pi(\mathrm{d}r) \otimes \mathrm{d}v \otimes \frac{(\lambda r)^k}{k!} \mathrm{e}^{-\lambda r} \sharp(\mathrm{d}k).$$

Now define three random measures by

$$N^0(\mathrm{d}s, \mathrm{d}r, \mathrm{d}\nu) = \mathcal{N}(\mathrm{d}s, \mathrm{d}r, \mathrm{d}\nu, \{0\}),$$

$$N^1(\mathrm{d}s, \mathrm{d}r, \mathrm{d}\nu) = \mathcal{N}(\mathrm{d}s, \mathrm{d}r, \mathrm{d}\nu, \{1\}),$$

$$N^2(\mathrm{d}s, \mathrm{d}r, \mathrm{d}\nu) = \mathcal{N}(\mathrm{d}s, \mathrm{d}r, \mathrm{d}\nu, \{k \geq 2\}).$$

Classical Poisson thinning now tells us that N^0, N^1, and N^2 are independent Poisson point processes on $[0, \infty)^3$ with respective intensities $\mathrm{d}s \otimes \mathrm{e}^{-\lambda r}\Pi(\mathrm{d}r) \otimes \mathrm{d}\nu$, $\mathrm{d}s \otimes (\lambda r)\mathrm{e}^{-\lambda r}\Pi(\mathrm{d}r) \otimes \mathrm{d}\nu$, and $\mathrm{d}s \otimes \sum_{k=2}^{\infty}(\lambda r)^k \mathrm{e}^{-\lambda r}\Pi(\mathrm{d}r)/k! \otimes \mathrm{d}\nu$.

With these thinned Poisson random measures in hand, we may start to separate out the different stochastic integrals in (1.4). We have that, for $t \geq 0$,

$$
\begin{aligned}
X_t = {} & x + \alpha \int_0^t X_{s-}\mathrm{d}s + \sqrt{2\beta} \int_0^t \int_0^{X_{s-}} W(\mathrm{d}s, \mathrm{d}u) + \int_0^t \int_0^{\infty} \int_0^{X_{s-}} r\tilde{N}^0(\mathrm{d}s, \mathrm{d}r, \mathrm{d}\nu) \\
& + \int_0^t \int_0^{\infty} \int_0^{X_{s-}} rN^1(\mathrm{d}s, \mathrm{d}r, \mathrm{d}\nu) + \int_0^t \int_0^{\infty} \int_0^{X_{s-}} rN^2(\mathrm{d}s, \mathrm{d}r, \mathrm{d}\nu) \\
& - \int_0^t \int_0^{\infty} X_{s-} \sum_{n=1}^{\infty} \frac{(\lambda r)^n}{n!}\mathrm{e}^{-\lambda r} r\,\Pi(\mathrm{d}r)\,\mathrm{d}s \\
= {} & x - \psi'(\lambda) \int_0^t X_s\mathrm{d}s + \sqrt{2\beta} \int_0^t \int_0^{X_{s-}} W(\mathrm{d}s, \mathrm{d}u) + \int_0^t \int_0^{\infty} \int_0^{X_{s-}} r\tilde{N}^0(\mathrm{d}s, \mathrm{d}r, \mathrm{d}\nu) \\
& + \int_0^t \int_0^{\infty} \int_0^{X_{s-}} rN^1(\mathrm{d}s, \mathrm{d}r, \mathrm{d}\nu) + 2\beta\lambda \int_0^t X_{s-}\,\mathrm{d}s \\
& + \int_0^t \int_0^{\infty} \int_0^{X_{s-}} rN^2(\mathrm{d}s, \mathrm{d}r, \mathrm{d}\nu),
\end{aligned}
\tag{3.1}
$$

where in the last equality we have used the easily derived fact that $-\int_{(0,\infty)}(1 - \mathrm{e}^{-\lambda r})r\,\Pi(\mathrm{d}r) = -\alpha + 2\beta\lambda - \psi'(\lambda)$. Recalling (2.1), the first line in the last equality of (3.1) corresponds to the dynamics of a subcritical CSBP with branching mechanism ψ_λ.

Inspecting the statement of Theorem 2.1, we see intuitively that in order to prove this result, our job is to show that the integrals on the right-hand side of (3.1) driven by N^1 and N^2 can be identified with the mass that immigrates off the skeleton.

4. λ-skeleton: proof of Theorem 2.1

We start by addressing the claim that (2.3) possesses a unique strong solution. Thereafter we prove claims (i) (ii), and (iii) of the theorem in order.

We can identify the existence of any weak solution to (2.3) with initial value $(\Lambda_0, Z_0) = (x, n)$, $x \geq 0$, $n \in \mathbb{N}_0$, by introducing additionally marked versions of the Poisson random measures $\mathbb{N}^1$ and $\mathbb{N}^2$, as well as an additional Poisson random measure $\mathbb{N}^*$. We will insist that $\mathbb{N}^1(\mathrm{d}s, \mathrm{d}r, \mathrm{d}j, \mathrm{d}\omega)$ has intensity $\mathrm{d}s \otimes r\mathrm{e}^{-\lambda r}\Pi(\mathrm{d}r) \otimes \sharp(\mathrm{d}j) \otimes \mathbb{P}_r^{(\lambda)}(\mathrm{d}\omega)$ on $[0, \infty)^2 \times \mathbb{N}_0 \times \mathbb{D}([0, \infty), \mathbb{R})$, $\mathbb{N}^2(\mathrm{d}s, \mathrm{d}r, \mathrm{d}k, \mathrm{d}j, \mathrm{d}\omega)$ has intensity $\psi'(\lambda)\mathrm{d}s \otimes \eta_k(\mathrm{d}r) \otimes p_k\sharp(\mathrm{d}k) \otimes \sharp(\mathrm{d}j) \otimes \mathbb{P}_r^{(\lambda)}(\mathrm{d}\omega)$ on $[0, \infty)^2 \times \mathbb{N}_0 \times \mathbb{N}_0 \times \mathbb{D}([0, \infty), \mathbb{R})$, and $\mathbb{N}^*(\mathrm{d}s, \mathrm{d}j, \mathrm{d}\omega)$ has intensity $2\beta\mathrm{d}s \otimes \sharp(\mathrm{d}j) \otimes \mathbb{Q}^{(\lambda)}(\mathrm{d}\omega)$ on $[0, \infty) \times \mathbb{N}_0 \times \mathbb{D}([0, \infty), \mathbb{R})$, where $\mathbb{P}_r^{(\lambda)}$ is the law of a ψ_λ-CSBP with initial value $r \geq 0$ (formally speaking, $\mathbb{P}_0^{(\lambda)}$ is the law of the null process) and $\mathbb{Q}^{(\lambda)}$ is the associated excursion measure.

Our proposed solution to to (2.3) will be to first define $(Z_t, t \geq 0)$ as the continuous-time GW process with branching rate $\psi'(\lambda)$ and offspring distribution given by $(p_k, k \geq 0)$. It is then easy to otherwise represent Z in the more complicated form

$$Z_t = \int_0^t \int_0^\infty \int_0^\infty \int_1^{Z_{s-}} \int_{\mathbb{D}([0,\infty),\mathbb{R})} (k-1) \, \mathrm{N}^2(\mathrm{d}s, \mathrm{d}r, \mathrm{d}k, \mathrm{d}j, \mathrm{d}\omega), \qquad t \geq 0.$$

Next we take

$$\Lambda_t = X_t^{(\lambda)} + D_t, \qquad t \geq 0, \tag{4.1}$$

where $X^{(\lambda)}$ is an autonomously independent copy of a ψ_λ-CSBP issued with initial mass x and, given N^1 and N^2, D_t, $t \geq 0$, is the uniquely identified (up to almost sure modification) 'dressed skeleton' described by

$$\begin{aligned}
D_t &= \int_0^t \int_0^\infty \int_1^{Z_{s-}} \int_{\mathbb{D}([0,\infty),\mathbb{R})} \omega_{t-s} \, \mathrm{N}^1(\mathrm{d}s, \mathrm{d}r, \mathrm{d}j, \mathrm{d}\omega) \\
&\quad + \int_0^t \int_0^\infty \int_0^\infty \int_1^{Z_{s-}} \int_{\mathbb{D}([0,\infty),\mathbb{R})} \omega_{t-s} \, \mathrm{N}^2(\mathrm{d}s, \mathrm{d}r, \mathrm{d}k, \mathrm{d}j, \mathrm{d}\omega) \\
&\quad + \int_0^t \int_1^{Z_{s-}} \int_{\mathbb{D}([0,\infty),\mathbb{R})} \omega_{t-s} \, \mathrm{N}^*(\mathrm{d}s, \mathrm{d}j, \mathrm{d}\omega).
\end{aligned}$$

To see why this provides a weak solution to (2.3), we may appeal to the martingale representation of the Markov pair (Λ, Z) described above. In particular, the generator of (Λ, Z) can be identified consistently with the generator of the process associated to (2.3); that is to say, their common generator is given by

$$\begin{aligned}
\mathbf{L}f(x, n) &= -\psi'(\lambda)x \frac{\partial f}{\partial x}(x, n) + \beta x \frac{\partial^2 f}{\partial x^2}(x, n) \\
&\quad + x \int_0^\infty \left[f(x+r, n) - f(x, n) - r \frac{\partial f}{\partial x}(x, n) \right] \mathrm{e}^{-r\lambda} \, \Pi(\mathrm{d}r) \\
&\quad + n \int_0^\infty \sum_{j \geq 1} [f(x+r, n+j-1) - f(x, n)] \frac{r^j \lambda^{j-1}}{j!} \mathrm{e}^{-r\lambda} \, \Pi(\mathrm{d}r) \\
&\quad + \beta\lambda n[f(x, n+1) - f(x, n)] + \frac{\psi(\lambda)}{\lambda} n[f(x, n-1) - f(x, n)] + 2\beta n \frac{\partial f}{\partial x}(x, n)
\end{aligned}$$

for $x \geq 0$, $n \in \mathbb{N}_0$, and for all non-negative, smooth, and compactly supported functions f. (Here, the penultimate term is understood to be zero when $n = 0$.) With this sense of commonality for their generators, it is then easy to verify the conditions of Theorem 2.3 of [23] and thus to conclude that (Λ, Z) provides a weak solution to (2.3).

Pathwise uniqueness is also relatively easy to establish. Indeed, suppose that Λ is the first component of any path solution to (2.3) with driving source of randomness N^0, N^1, N^2, and W, and suppose that we write it in the form

$$\Lambda_t =: \Lambda_0 + H(\Lambda_s, s<t) + I_t, \qquad t \geq 0,$$

where

$$H(\Lambda_s, s{<}t) = - \psi'(\lambda) \int_0^t \Lambda_{s-} \, \mathrm{d}s + \sqrt{2\beta} \int_0^t \int_0^{\Lambda_{s-}} W(\mathrm{d}s, \mathrm{d}u)$$

$$+ \int_0^t \int_0^\infty \int_0^{\Lambda_{s-}} r \tilde{\mathrm{N}}^0(\mathrm{d}s, \mathrm{d}r, \mathrm{d}v), \qquad t \geq 0,$$

and

$$I_t = \int_0^t \int_0^\infty \int_1^{Z_{s-}} r \mathrm{N}^1(\mathrm{d}s, \mathrm{d}r, \mathrm{d}j)$$

$$+ \int_0^t \int_0^\infty \int_0^\infty \int_1^{Z_{s-}} r \mathrm{N}^2(\mathrm{d}s, \mathrm{d}r, \mathrm{d}k, \mathrm{d}j) + 2\beta \int_0^t Z_{s-} \, \mathrm{d}s, \qquad t \geq 0.$$

Recalling that the almost sure path of Z is uniquely defined by N^2, it follows that, if $\Lambda^{(1)}$ and $\Lambda^{(2)}$ are two path solutions to (2.3) with the same initial value, then

$$\Lambda_t^{(1)} - \Lambda_t^{(2)} = H(\Lambda_s^{(1)}, s{<}t) - H(\Lambda_s^{(2)}, s{<}t), \qquad t \geq 0.$$

The reader will now note that the above equation is precisely the SDE obtained when looking at the path difference between two solutions of an SDE of the type given in (1.4). Since there is pathwise uniqueness for (1.4), we easily conclude that $\Lambda^{(1)} = \Lambda^{(2)}$ almost surely.

Finally, taking account of the existence of a weak solution and pathwise uniqueness, we may appeal to an appropriate version of the Yamada–Watanabe theorem (see, for example, Theorem 1.2 of [2]) to deduce that (2.3) possesses a unique strong solution. And since this holds for every fixed initial configuration x and n, it also holds when the initial values are independently randomised.

(i) This claim requires an analytical verification and, in some sense, is similar in spirit to the proof that, for $t \geq 0$, $Z_t \mid \Lambda_t$ is Poisson distributed with rate $\lambda^* \Lambda_t$ in the prolific skeletal decomposition found in [4]. A fundamental difference here is that we work with SDEs, and hence stochastic calculus, rather than integral equations for semigroups as in [4] and, moreover, the parameter λ need not be the minimal value, λ^*, in its range.

Standard arguments show that the solution to (2.3) is a strong Markov process, and accordingly we write $\mathbf{P}_{x,n}$, $x > 0$, $n \in \mathbb{N}_0$ for its probabilities. Moreover, with an abuse of notation we write, for $x > 0$,

$$\mathbf{P}_x(\cdot) = \sum_{n \geq 0} \frac{(\lambda x)^n}{n!} \mathrm{e}^{-\lambda x} \mathbf{P}_{x,n}(\cdot). \tag{4.2}$$

Define $f_t(\eta, \theta) := \mathbf{E}_x[\mathrm{e}^{-\eta \Lambda_t - \theta Z_t}]$, $x, \theta, \eta, t \geq 0$, and let $F_t := \mathrm{e}^{-\eta \Lambda_t - \theta Z_t}$, $t \geq 0$. Using Itô's formula for semi-martingales (see Theorem 32 of [40]), for $t \geq 0$,

$$\mathrm{d}F_t = - \eta F_{t-} \mathrm{d}\Lambda_t - \theta F_{t-} \mathrm{d}Z_t + \frac{1}{2} \eta^2 F_{t-} \mathrm{d}[\Lambda, \Lambda]_t^c + \frac{1}{2} \theta^2 F_{t-} \mathrm{d}[Z, Z]_t^c$$

$$+ \eta \theta F_{t-} \mathrm{d}[\Lambda, Z]_t^c + \Delta F_t + \eta F_{t-} \Delta \Lambda_t + \theta F_{t-} \Delta Z_t.$$

(Here and throughout the remainder of this paper, for any stochastic process Y we use the notation that $\Delta Y_t = Y_t - Y_{t-}$.) As Z is a pure jump process, we have that $[Z, Z]_t^c = [\Lambda, Z]_t^c = 0$. Taking advantage of the fact that

$$F_t = F_{t-}\mathrm{e}^{-\eta\Delta\Lambda_t-\theta\Delta Z_t},$$

we may thus write, in integral form,

$$\begin{aligned}
F_t = F_0 &- \eta \int_0^t F_{s-}\, \mathrm{d}\Lambda_s - \theta \int_0^t F_{s-}\, \mathrm{d}Z_s + \beta\eta^2 \int_0^t F_{s-}\Lambda_s\, \mathrm{d}s \\
&+ \sum_{s\le t} F_{s-} \left\{ \mathrm{e}^{-\eta\Delta\Lambda_s-\theta\Delta Z_s} - 1 + \eta\Delta\Lambda_s + \theta\Delta Z_s \right\},
\end{aligned}$$

where the sum is taken over the countable set of discontinuities of (Λ, Z). We can split up the sum of discontinuities according to the Poisson random measure in (2.3) that is responsible for the discontinuity. Hence, writing $\Delta^{(j)}$, $j = 0, 1, 2$, to mean an increment coming from each of the three Poisson random measures,

$$\begin{aligned}
F_t = F_0 &- \eta \int_0^t F_{s-}\, \mathrm{d}\Lambda_s - \theta \int_0^t F_{s-}\, \mathrm{d}Z_s + \beta\eta^2 \int_0^t F_{s-}\Lambda_{s-}\, \mathrm{d}s \\
&+ \sum_{s\le t} F_{s-} \left\{ \mathrm{e}^{-\eta\Delta^{(0)}\Lambda_s} - 1 + \eta\Delta^{(0)}\Lambda_s \right\} \\
&+ \sum_{s\le t} F_{s-} \left\{ \mathrm{e}^{-\eta\Delta^{(1)}\Lambda_s} - 1 + \eta\Delta^{(1)}\Lambda_s \right\} \\
&+ \sum_{s\le t} F_{s-} \left\{ \mathrm{e}^{-\eta\Delta^{(2)}\Lambda_s-\theta\Delta Z_s} - 1 + \eta\Delta^{(2)}\Lambda_s + \theta\Delta Z_s \right\}.
\end{aligned} \tag{4.3}$$

Now, note that we can rewrite the first element of the vectorial SDE (2.3) as

$$\begin{aligned}
\Lambda_t = \Lambda_0 &- \psi'(\lambda) \int_0^t \Lambda_{s-}\, \mathrm{d}s + \sqrt{2\beta} \int_0^t \int_0^{\Lambda_{s-}} W(\mathrm{d}s, \mathrm{d}u) + M_t \\
&+ \sum_{s\le t} \Delta^{(1)}\Lambda_s + \sum_{s\le t} \Delta^{(2)}\Lambda_s + 2\beta \int_0^t Z_{s-}\, \mathrm{d}s, \qquad t \ge 0,
\end{aligned}$$

where M_t is a zero-mean martingale corresponding to the integral in (2.3) with respect to $\tilde{\mathrm{N}}^0$. Therefore, performing the necessary calculus in (4.3) for the integral with respect to $\mathrm{d}\Lambda_t$, we get that

$$\begin{aligned}
F_t = &-F_0 - \eta(\psi'(\lambda) + \eta\beta) \int_0^t F_{s-}\Lambda_{s-}\, \mathrm{d}s - \sum_{s\le t} F_{s-}\left\{ \mathrm{e}^{-\eta\Delta^{(2)}\Lambda_s-\theta\Delta Z_s} - 1 \right\} \\
&+ 2\eta\beta \int_0^t F_{s-}Z_{s-}\, \mathrm{d}s - \sum_{s\le t} F_{s-}\left\{ \mathrm{e}^{-\eta\Delta^{(1)}\Lambda_s} - 1 \right\} \\
&- \sum_{s\le t} F_{s-}\left\{ \mathrm{e}^{-\eta\Delta^{(0)}\Lambda_s} - 1 + \eta\Delta^{(0)}\Lambda_s \right\}, \qquad t \ge 0,
\end{aligned}$$

is equal to a zero-mean martingale which is the sum of the previously mentioned M_t, $t \geq 0$, and the white noise integral. Taking expectations, we thus have

$$
\begin{aligned}
f_t(\eta, \theta) =& f_0(\eta, \theta) + \eta(\psi'(\lambda) + \eta\beta)\mathbf{E}_x \int_0^t [F_{s-}\Lambda_{s-}] \, \mathrm{d}s - 2\eta\beta\mathbf{E}_x \int_0^t [F_{s-}Z_{s-}] \, \mathrm{d}s \\
&+ \mathbf{E}_x \int_0^t [F_{s-}\Lambda_{s-}] \, \mathrm{d}s \left(\int_0^\infty (\mathrm{e}^{-\eta r} - 1 + \eta r)\mathrm{e}^{-\lambda r} \, \Pi(\mathrm{d}r) \right) \\
&+ \mathbf{E}_x \int_0^t [F_{s-}Z_{s-}] \, \mathrm{d}s \left(\int_0^\infty (\mathrm{e}^{-\eta r} - 1)r\mathrm{e}^{-\lambda r} \, \Pi(\mathrm{d}r) \right) \\
&+ \mathbf{E}_x \int_0^t \frac{1}{\lambda}[F_{s-}Z_{s-}] \, \mathrm{d}s \left(\sum_{k=0, k \neq 1}^\infty \int_0^\infty (\mathrm{e}^{-\eta r - \theta(k-1)} - 1) \left\{ \psi(\lambda)\mathbf{1}_{\{k=0\}}\delta_0(\mathrm{d}r) \right. \right. \\
&\qquad\qquad \left. \left. + \delta_0(\mathrm{d}r)\beta\lambda^2\mathbf{1}_{\{k=2\}} + \mathbf{1}_{\{k \geq 2\}}\frac{(\lambda r)^k}{k!}\mathrm{e}^{-\lambda r}\Pi(\mathrm{d}r) \right\} \right).
\end{aligned}
$$

Accumulating terms, we find that $f_t(\eta, \theta)$ satisfies the PDE

$$
\begin{aligned}
\frac{\partial}{\partial t}f_t(\eta, \theta) &= A_\lambda(\eta, \theta)\frac{\partial}{\partial \eta}f_t(\eta, \theta) + B_\lambda(\eta, \theta)\frac{\partial}{\partial \theta}f_t(\eta, \theta), \\
f_0(\eta, \theta) &= \mathrm{e}^{-(\eta + \lambda(1 - \mathrm{e}^{-\theta}))x},
\end{aligned}
\tag{4.4}
$$

where

$$
A_\lambda(\eta, \theta) = \eta(-\psi'(\lambda) - \eta\beta) - \int_0^\infty (\mathrm{e}^{-\eta r} - 1 + \eta r)\mathrm{e}^{-\lambda r} \, \Pi(\mathrm{d}r)
$$

$$
\begin{aligned}
B_\lambda(\eta, \theta) = 2\eta\beta - \sum_{k=0}^\infty \int_0^\infty (\mathrm{e}^{-\eta r - \theta(k-1)} - 1) &\left\{ \frac{\psi(\lambda)}{\lambda}\mathbf{1}_{\{k=0\}}\delta_0(\mathrm{d}r) + \delta_0(\mathrm{d}r)\beta\lambda\mathbf{1}_{\{k=2\}} \right. \\
&\left. + \mathbf{1}_{\{k \geq 1\}}\frac{\lambda^{k-1}r^k}{k!}\mathrm{e}^{-\lambda r}\Pi(\mathrm{d}r) \right\}.
\end{aligned}
$$

Standard theory for the linear partial differential Equation (4.4) (see, for example, Chapter 3 (Theorem 2, p. 107) of [17] and references therein) tells us that it has a unique local solution. Our aim now is to show that this solution is also represented by

$$
\mathbb{E}_x[\exp\{-(\eta + \lambda(1 - \mathrm{e}^{-\theta}))X_t\}] = \exp\{-u_t(\eta + \lambda(1 - \mathrm{e}^{-\theta}))x\}, \qquad x, t, \theta, \eta \geq 0, \tag{4.5}
$$

where we recall that X is the ψ-CSBP. To this end, let us define $\kappa = \eta + \lambda(1 - \mathrm{e}^{-\theta})$ and note that, for $x, t, \kappa \geq 0$, $g_t(\kappa) := \mathbb{E}_x[\exp\{-\kappa X_t\}]$ satisfies

$$
\begin{aligned}
\frac{\partial}{\partial t}g_t(\kappa) &= -\psi(\kappa)\frac{\partial}{\partial \kappa}g_t(\kappa), \\
g_0(\kappa) &= \mathrm{e}^{-\kappa x}
\end{aligned}
\tag{4.6}
$$

(see, for example, Exercise 12.2 in [24]). After a laborious amount of algebra we can verify that $-\psi(\kappa) = A_\lambda(\eta, \theta) + \lambda\mathrm{e}^{-\theta}B_\lambda(\eta, \theta)$, and hence we may develop the right-hand side of (4.6) and write, for $x, t, \eta, \theta \geq 0$,

$$
\begin{aligned}
\frac{\partial}{\partial t}g_t(\kappa) &= A_\lambda(\eta, \theta)\frac{\partial}{\partial \kappa}g_t(\kappa) + \lambda\mathrm{e}^{-\theta}B_\lambda(\eta, \theta)\frac{\partial}{\partial \kappa}g_t(\kappa) \\
&= A_\lambda(\eta, \theta)\frac{\partial}{\partial \eta}g_t(\kappa) + B_\lambda(\eta, \theta)\frac{\partial}{\partial \theta}g_t(\kappa).
\end{aligned}
$$

Now we choose $x = 1$. Then local uniqueness of the solution to (4.4), or equivalently the local uniqueness of (4.6), thus tells us that there exists $t_0 > 0$ such that $g_t(\eta + \lambda(1 - e^{-\theta})) = f_t(\eta, \theta)$ for all $\eta, \theta \geq 0$ and $t \in [0, t_0]$.

In conclusion, now that we have proved that for $t \in [0, t_0]$

$$\mathbb{E}_1[\exp\{-(\eta + \lambda(1 - e^{-\theta}))X_t\}] = \mathbf{E}_1[\exp\{-\eta\Lambda_t - \theta Z_t\}], \qquad \theta, \eta \geq 0, \qquad (4.7)$$

we can observe the following implications in turn. First, setting $\theta = 0$ and $\eta > 0$, we see that Λ_t under $\mathbf{P}_1$ has the same distribution as X_t under $\mathbb{P}_1$ for all $t \in [0, t_0]$. Next, setting both $\eta, \theta > 0$, we observe that (Λ_t, Z_t) under $\mathbf{P}_1$ has the same law as $(X_t, \mathrm{Po}(\lambda x)|_{x=X_t})$ under $\mathbb{P}_1$, where $\mathrm{Po}(\lambda x)$ is an autonomously independent Poisson random variable with rate λx. In particular, it follows that, for all $t \in [0, t_0]$, under $\mathbf{P}_1$, the law of Z_t given Λ_t is $\mathrm{Po}(\lambda \Lambda_t)$.

To get a global result we first show that the previous conclusions hold for any initial mass $x > 0$ on the time interval $[0, t_0]$, then, using the Markov property, we extend the results for any $t > 0$. First, from (4.5) we can observe that

$$\mathbb{E}_x\left[\exp\{-(\eta + \lambda(1 - e^{-\theta}))X_t\}\right] = \left(\mathbb{E}_1\left[\exp\{-(\eta + \lambda(1 - e^{-\theta}))X_t\}\right]\right)^x.$$

Thus, in order to extend the previous results to any $x > 0$ we only need to prove that

$$\mathbf{E}_x\left[e^{-\eta\Lambda_t - \theta Z_t}\right] = \left(\mathbf{E}_1\left[e^{-\eta\Lambda_t - \theta Z_t}\right]\right)^x.$$

Recalling the representation (4.1) and the notation (4.2), we can write, for $t \leq t_0$,

$$\begin{aligned}
\mathbf{E}_x\left[e^{-\eta\Lambda_t - \theta Z_t}\right] &= \sum_{n \geq 0} \frac{(\lambda x)^n}{n!} e^{-\lambda x} \mathbf{E}_{(x,n)}\left[e^{-\eta\Lambda_t - \theta Z_t}\right] \\
&= \sum_{n \geq 0} \frac{(\lambda x)^n}{n!} e^{-\lambda x} \mathbf{E}_x\left[e^{-\eta X_t^{(\lambda)}}\right] \mathbf{E}_{(0,n)}\left[e^{-\eta D_t - \theta Z_t}\right] \\
&= \left(\mathbf{E}_1\left[e^{-\eta X_t^{(\lambda)}}\right]\right)^x \sum_{n \geq 0} \frac{(\lambda x)^n}{n!} e^{-\lambda x} \left(\mathbf{E}_{(0,1)}\left[e^{-\eta D_t - \theta Z_t}\right]\right)^n \\
&= \left(\mathbf{E}_1\left[e^{-\eta X_t^{(\lambda)}}\right] \exp\left\{\lambda\left(\mathbf{E}_{(0,1)}\left[e^{-\eta D_t - \theta Z_t}\right] - 1\right)\right\}\right)^x \\
&= \left(\mathbf{E}_1\left[e^{-\eta\Lambda_t - \theta Z_t}\right]\right)^x,
\end{aligned}$$

as required.

Now take $t_0 < t \leq 2t_0$, and use the tower property to get

$$\begin{aligned}
\mathbf{E}_x\left[e^{-\eta\Lambda_t - \theta Z_t}\right] &= \mathbf{E}_x\left[\mathbf{E}_{\Lambda_{t_0}}\left[\exp\{-\eta\Lambda_{t-t_0} - \theta Z_{t-t_0}\}\right]\right] \\
&= \int_{\mathbb{R}_+} \mathbf{E}_y\left[\exp\{-\eta\Lambda_{t-t_0} - \theta Z_{t-t_0}\}\right] \mathbf{P}_x(\Lambda_{t_0} \in \mathrm{d}y),
\end{aligned}$$

and similarly

$$\mathbb{E}_x\left[\exp\{-(\eta + \lambda(1 - e^{-\theta}))X_t\}\right] = \int_{\mathbb{R}_+} \mathbb{E}_y\left[\exp\{-(\eta + \lambda(1 - e^{-\theta}))X_{t-t_0}\}\right] \mathbb{P}_x(X_{t_0} \in \mathrm{d}y).$$

Thus, using local uniqueness and the previously deduced implications on $[0, t_0]$, we see that

$$\mathbf{E}_x\left[e^{-\eta\Lambda_t - \theta Z_t}\right] = \mathbb{E}_x\left[\exp\{-(\eta + \lambda(1 - e^{-\theta}))X_t\}\right] \quad \text{on } t \in [0, 2t_0],$$

and by iterating the previous argument we get equality for any $t > 0$.

Finally, on account of the fact that (Λ_t, Z_t), $t \geq 0$, is a joint Markovian pair, this now global Poissonisation allows us to infer that Λ_t, $t \geq 0$, is itself Markovian. Indeed, for any bounded measurable and positive h and $s, t \geq 0$,

$$\mathbf{E}[h(\Lambda_{t+s})|\mathcal{F}_t^\Lambda] = \sum_{n \geq 0} \frac{(\lambda \Lambda_t)^n}{n!} e^{-\lambda \Lambda_t} \mathbf{E}_{x,n}[h(\Lambda_s)]_{x=\Lambda_t} = \mathbf{E}_x[h(\Lambda_s)]_{x=\Lambda_t}.$$

We may now conclude that for all $t \geq 0$ and $x > 0$, under $\mathbf{P}_x$, $Z_t \mid \mathcal{F}_t^\Lambda$ is $\mathrm{Po}(\lambda \Lambda_t)$ distributed, as required.

(ii) We have seen that the pair $((\Lambda_t, Z_t), t \geq 0)$ is a Markov process for any initial state (x, n), but, due to the dependence on Z, on its own $(\Lambda_t, t \geq 0)$ is not Markovian. However, considering (4.7) we see that after the Poissonisation of Z_0, $(\Lambda_t, t \geq 0)$ becomes a Markov process with a semi-group that agrees with that of $(X_t, t \geq 0)$. On account of the fact that X is the unique weak solution to (1.4), it automatically follows that Λ also represents the unique weak solution to (1.4).

(iii) Since the event $\{Z_0 = 0\}$ implies the event $\{Z_t = 0, t \geq 0\}$, the system (2.3) reduces to the SDE

$$\Lambda_t = x + \psi'(\lambda) \int_0^t \Lambda_{s-} \, \mathrm{d}s + \sqrt{2\beta} \int_0^t \int_0^{\Lambda_{s-}} W(\mathrm{d}s, \mathrm{d}u) + \int_0^t \int_0^\infty \int_0^{\Lambda_{s-}} r \tilde{N}^0(\mathrm{d}s, \mathrm{d}r, \mathrm{d}\nu),$$

which has the exact form of the SDE describing the evolution of a CSBP with branching mechanism ψ_λ. □

5. Exploration of subcritical CSBPs

The objective of this section is to give a heuristic description of how the notion of a prolific skeleton emerges in the subcritical case, and specifically why the structure of the SDE (2.5) is meaningful in this respect. We need to be careful about what one means by 'prolific', but nonetheless the inspiration for a decomposition can be gleaned by examining in more detail the description of subcritical CSBPs through the exploration process.

We assume throughout the conditions of Theorem 2.2. That is to say, X is a (sub)critical ψ-CSBP where ψ satisfies Grey's condition (1.5). Let $(\xi_t, t \geq 0)$ be a spectrally positive Lévy process with Laplace exponent ψ. Using the classical work of [31, 32] (see also [10, 30]) we can use generalised Ray—Knight-type theorems to construct X in terms of the so-called height process associated to ξ. For convenience, and to introduce more notation, we give a brief overview here.

Denote by $(\hat{\xi}_r^{(t)}, 0 \leq r \leq t)$ the time-reversed process at time t, that is, $\hat{\xi}_r^{(t)} := \xi_t - \xi_{(t-r)-}$, and let $\hat{S}_r^{(t)} := \sup_{s \leq r} \hat{\xi}_s^{(t)}$. We define H_t as the local time at level 0 at time t of the process $\hat{S}^{(t)} - \hat{\xi}^{(t)}$. Because the reversed process has a different point from which it is reversed at each time, the process H does not behave in a monotone way. The process $(H_t, t \geq 0)$ is called the ψ-height process, which, under assumption (1.5), is continuous. There exists a notion of local time up to time t of H at level $a \geq 0$, henceforth denoted by L_t^a. Specifically, the family $(L_t^a, a, t \geq 0)$ satisfies

$$\int_0^t g(H_s) \, \mathrm{d}s = \int_0^\infty g(a) L_t^a \, \mathrm{d}a \qquad t \geq 0,$$

where g is a non-negative measurable function.

For $x > 0$ let $T_x := \inf\{t \geq 0, \xi_t = -x\}$. Then the generalised Ray–Knight theorem for the ψ-CSBP process states that $(L^a_{T_x}, a \geq 0)$ has a càdlàg modification for which

$$(L^t_{T_x}, t \geq 0) \overset{\mathrm{d}}{=} (X, \mathbb{P}_x);$$

that is, the two processes are equal in law.

The height process also codes the genealogy of the ψ-CSBP. It can be shown that the excursions of H from 0 form a time-homogeneous Poisson point process of excursions with respect to local time at 0. We shall use n to denote its intensity measure.

If $X_0 = x$, then the total amount of local time of H accumulated at zero is x. Each excursion codes a real tree (see [10] for a precise meaning) such that the excursion that occurs after $u \leq x$ units of local time can be thought of as the descendants of the uth individual in the initial population. Here we are interested in the genealogy of the conditioned process and what we will call the embedded 'T-prolific' tree, that is, the tree of the individuals that survive up to time T. Conditioning the process on survival up to time T corresponds to conditioning the height process to have at least one excursion above level T. (We have the slightly confusing, but nonetheless standard, notational anomaly that a spatial height for an excursion corresponds to the spatial height in the tree that it codes, but that this may also be seen as a time into the forward evolution of the tree.) Let n_T denote the conditional probability $n(\cdot \mid \sup_{s \geq 0} \epsilon_s \geq T)$, where ϵ is a canonical excursion of H under n. Let $(Z^T_t, t \geq 0)$ be the process that counts the number of excursions above level t that hit level T within the excursion ϵ. Duquesne and Le Gall in [10] described the distribution of Z^T under n_T and proved the following.

Theorem 5.1. *Under* n_T *the process* $(Z^T_t, 0 \leq t < T)$ *is a time-inhomogeneous Markov process whose law is characterised by the following identities. For every* $\lambda > 0$,

$$n_T\left[\exp\{-\lambda Z^T_t\}\right] = 1 - \frac{u_t((1 - \mathrm{e}^{-\lambda})u_{T-t}(\infty))}{u_T(\infty)},$$

and if $0 \leq t < t' < T$,

$$n_T\left[\exp\{-\lambda Z^T_{t'}\} \mid Z^T_t\right] = \left(n_{T-t}\left[\exp\{-\lambda Z^{T-t}_{t'-t}\}\right]\right)^{Z^T_t}.$$

In essence, the second part of the above theorem shows that Z^T has the branching property. However, temporal inhomogeneity means that it is a time-dependent continuous-time GW process. In [10] it is moreover shown that, conditionally on L^t_σ under n_T, where σ is the length of the excursion ϵ, Z^T_t is Poisson distributed with intensity $u_{T-t}(\infty)L^t_\sigma$. Thinking of L^t_σ as the mass at time t in the tree of descendants of the prolific individual in the initial population that the excursion codes, we thus have a Poisson embedding of the number of prolific descendants of that one individual within the excursion.

The time-dependent continuous-time GW process in the theorem can also be characterised as follows. At time 0 we start with one individual. Then the law of the first branching time, γ_T, is given by

$$n_T(\gamma_T > t) = \frac{\psi(u_T(\infty))}{u_T(\infty)} \frac{u_{T-t}(\infty)}{\psi(u_{T-t}(\infty))}, \qquad t \in [0, T), \qquad (5.1)$$

and, conditionally on γ_T, the probability-generating function of the offspring distribution is

$$n_T\left[r^{Z^T_{\gamma_T}} \mid \gamma_T = t\right] = 1 + \frac{\psi((1 - r)u_{T-t}(\infty)) - (1 - r)u_{T-t}(\infty)\psi'(u_{T-t}(\infty))}{u_{T-t}(\infty)\psi'(u_{T-t}(\infty)) - \psi(u_{T-t}(\infty))}, \quad r \in [0, 1].$$

The offspring distribution when a split occurs at height t in the excursion (equivalently, time t in the underlying genealogical tree), say $(p_k^{T-t}, k \geq 0)$, is explicitly given by the following. We have $p_0^{T-t} = p_1^{T-t} = 0$, and for $k \geq 2$,

$$p_k^{T-t} = \frac{1}{u_{T-t}(\infty)\psi'(u_{T-t}(\infty)) - \psi(u_{T-t}(\infty))}$$
$$\times \left\{ \beta u_{T-t}^2(\infty)\mathbf{1}_{\{k=2\}} + \int_0^\infty \frac{(u_{T-t}(\infty)x)^k}{k!} e^{-u_{T-t}(\infty)x}\, \Pi(\mathrm{d}x) \right\}.$$

Using (5.1) we can compute the rate of branching at any height t in the excursion. First, it is not hard to see that

$$\mathrm{n}_T(\gamma_T > t + s \mid \gamma_T > t) = \mathrm{n}_{T-t}(\gamma_T > s), \qquad 0 \leq t + s < T.$$

Hence, the rate is

$$\frac{\mathrm{d}}{\mathrm{d}s}\mathrm{n}_{T-t}(\gamma_T > s)\Big|_{s=0} = \frac{u_{T-t}(\infty)\psi'(u_{T-t}(\infty)) - \psi(u_{T-t}(\infty))}{u_{T-t}(\infty)}, \qquad t \in [0, T).$$

Again thinking of L_σ^t, $t < T$, as the mass of the tree coded by the excursion, and noting that not all of this mass is prolific, we would like to characterise the non-T-prolific mass that has 'immigrated' along the path of the prolific tree. We expect this to be a CSBP conditioned to die before time T. Using (1.1), we know that the probability that the process dies up to time T is given by

$$\mathbb{P}[X_T = 0 \mid \mathcal{F}_t] = e^{-X_t u_{T-t}(\infty)},$$

where we assume Grey's condition (1.5) to ensure that the above conditioning makes sense. A simple application of the Markov property tells us that the law of X conditioned to die out by time T can be obtained by the following change of measure:

$$\frac{\mathrm{d}\mathbb{P}_x^T}{\mathrm{d}\mathbb{P}_x}\Big|_{\mathcal{F}_t} = \frac{e^{-X_t u_{T-t}(\infty)}}{e^{-x u_T(\infty)}}, \qquad t \geq 0, \quad x > 0. \tag{5.2}$$

Indeed, using the semigroup property of u_t, $t \geq 0$, it is not hard to verify that the right-hand side above is a martingale. We would like to understand how to characterise the evolution of the process $(X, \mathbb{P}_x^T)$, $x > 0$, a little better as the change of measure is time inhomogeneous.

To this end, we again appeal to Itô's formula. Denote $v(t) := u_{T-t}(\infty)$; then, for non-negative, twice-differentiable, and compactly supported functions f, after a routine, albeit lengthy, application of Itô's formula we get, for $t \geq 0$ and $x > 0$,

$$e^{-X_t v(t)}f(X_t) = e^{-xv(0)}f(x) + M_t$$
$$+ \int_0^t X_s e^{-X_s v(s)}\big(-f(X_s)v'(s) - \alpha v(s)f(X_s) + \alpha f'(X_s)$$
$$+ \beta v^2(s)f(X_s) - 2\beta v(s)f'(X_s) + \beta f''(X_s)\big)\mathrm{d}t$$
$$+ \sum_{s \leq t}\big(e^{-X_s v(s)}f(X_s) - e^{-X_{s-} v(s-)}f(X_{s-})$$
$$+ v(s-)e^{-X_{s-} v(s-)}f(X_{s-})\Delta X_s - e^{-X_{s-} v(s-)}f'(X_{s-})\Delta X_s\big),$$

where M_t, $t \geq 0$, represents the martingale terms. Taking expectations, we get

$$
\begin{aligned}
\mathbb{E}_x\big[e^{-X_t v(t)} f(X_t)\big] = {} & e^{-xv(0)} f(x) \\
& + \mathbb{E}_x\bigg[\int_0^t \big\{ X_s e^{-X_s v(s)} \big(-f(X_s) v'(s) - \alpha v(s) f(X_s) + \alpha f'(X_s) \\
& + \beta v^2(s) f(X_s) - 2\beta v(s) f'(X_s) + \beta f''(X_s) \big) \\
& + \int_0^\infty X_s \big(e^{-(X_s+y)v(s)} f(X_s + y) - e^{-X_s v(s)} f(X_s) \\
& + v(s) e^{-X_s v(s)} f(X_s) y - e^{-X_s v(s)} f'(X_s) y \big) \Pi(\mathrm{d}y) \big\} \mathrm{d}s \bigg].
\end{aligned}
$$

Gathering terms, making use of the expression for ψ in (1.2), and that

$$
-\frac{\partial}{\partial s} u_{T-s}(\infty) + \psi(u_{T-s}(\infty)) = 0,
$$

we have, for $t \geq 0$ and $x > 0$, the Dynkin formula

$$
\begin{aligned}
\mathbb{E}_x^T\big[f(X_t)\big] &= \mathbb{E}_x\bigg[\frac{e^{-u_{T-t}(\infty)X_t}}{e^{-u_T(\infty)x}} f(X_t) \bigg] \\
&= f(x) + \mathbb{E}_x\bigg[\int_0^t \frac{e^{-u_{T-s}(\infty)X_s}}{e^{-u_T(\infty)x}} \big(\hat{\mathcal{L}}_{T-s} f(X_s) \big) \mathrm{d}s \bigg] \\
&= f(x) + \mathbb{E}_x^T\bigg[\int_0^t \hat{\mathcal{L}}_{T-s} f(X_s) \, \mathrm{d}s \bigg],
\end{aligned}
$$

where the infinitesimal generator is given by

$$
\begin{aligned}
\hat{\mathcal{L}}_{T-t} f(x) = {} & \psi'(u_{T-t}(\infty)) x f'(x) + \beta x f''(x) \\
& + x \int_0^\infty (f(x+y) - f(x) - y f'(x)) e^{-u_{T-t}(\infty)y} \Pi(\mathrm{d}y).
\end{aligned}
$$

For comparison, consider the generator of a CSBP with Esscher-transformed branching mechanism ψ_λ, which is given by

$$
\mathcal{L}_\lambda f(x) = \psi'(\lambda) x f'(x) + \beta x f''(x) + x \int_0^\infty (f(x+y) - f(x) - y f'(x)) e^{-\lambda y} \Pi(\mathrm{d}y) \tag{5.3}
$$

for suitably smooth and integrable functions f. Recall that the CSBP with generator (5.3) is subcritical providing $\psi'(\lambda) > 0$ and, taking account of (1.3), the greater this value, the 'more subcritical' it becomes. It appears that $\hat{\mathcal{L}}_{T-t}$ has the form of an Esscher-transformed branching mechanism based on ψ where the parameter shift is controlled by $u_{T-t}(\infty)$, which explodes as $t \to T$. Said another way, if we define $V_t^T(\theta)$, $0 \leq t < T$, $x, \theta \geq 0$, as the exponent satisfying

$$
\mathbb{E}_x^T[e^{-\theta X_t}] = e^{-xV_t^T(\theta)},
$$

then

$$
V_t^T(\theta) = u_t(\theta + u_{T-t}(\infty)) - u_T(\infty). \tag{5.4}
$$

Recalling that we are assuming Grey's condition (1.5) for a (sub)critical process, we note from (1.1) that

$$\lim_{T \to \infty} u_T(\infty) = 0.$$

In that case, the density in (5.2) tends to unity as $T \to \infty$.

We conclude this section by returning to Theorem 2.2. The discussion in this section shows that in the (sub)critical case the components of the SDE (2.5) mimic precisely the description of the T-skeleton in the previous section. In particular, the first three integrals in (2.5) indicate that once mass is created in the SDE, it evolves as a time-dependent CSBP with generator $\hat{\mathcal{L}}_{T-t}$. Moreover, the evolution of the skeleton Z^T as described in (2.5), matches precisely the dynamics of the T-prolific skeleton described in the previous section (for which we preemptively used the same notation), which is a time-dependent continuous-time GW process. Indeed, the branching rate and the time-dependent offspring distribution of both match.

Remark 5.1. It is important to note that even though the time-dependent T-prolific skeleton is inspired by the height process, the description does not require ψ to be a (sub)critical branching mechanism. Indeed, only requiring Grey's condition to be satisfied ensures that the branching rate and offspring distribution of this section are well defined. Similarly, we can apply the change of measure in (5.2) for any CSBP satisfying (1.5), and get a time-dependent CSBP with generator $\hat{\mathcal{L}}_{T-t}$. Although the results of Theorem 2.2 were motivated by [10], we can see that the theorem can be stated in a more general setting, and thus extends the existing family of finite-horizon decompositions for CSBPs.

6. Finite-time horizon skeleton: proof of Theorem 2.2

Now that we understand that the mathematical structure of (2.5) is little more than a time-dependent version of (2.3), the reader will not be surprised by the claim that the proof of strong uniqueness to (2.5), as well as parts (i) and (ii) of Theorem 2.1, pass through almost verbatim, albeit needing some minor adjustments for additional time derivatives of $u_{T-t}(\infty)$ in, e.g. (4.3) and (4.6), which plays the role of λ. To avoid repetition we simply leave the proof of these two parts as an exercise for the reader.

On the event $\{Z_0 = 0\}$, which is concurrent with the event $\{Z_t = 0, \ 0 \le t < T\}$, close inspection of (2.5) allows us to note that Λ is generated by an SDE with time-varying coefficients. Indeed, standard arguments show that, conditional on $\{Z_0 = 0\}$, Λ is a time-inhomogeneous Markov process.

Suppose that we write $\mathbf{P}_{x,n}^T$, $x \ge 0$, $n \in \mathbb{N}_0$, for the law of the Markov probabilities corresponding to the solution of (2.5). Moreover, we will again abuse this notation in the spirit of (4.2) and write $\mathbf{P}_x^T$, $x \ge 0$, when Z_0^T is randomised to be an independent Poisson random variable with rate $u_T(\infty)x$. We can use parts (i) and (ii) of Theorem 2.2, together with (5.2), to deduce that

$$\begin{aligned}
\mathbf{E}_x^T\left[e^{-\eta \Lambda_t} \mid Z_0 = 0\right] &= \frac{\mathbf{E}_x^T\left[e^{-\eta \Lambda_t}, \ Z_0 = 0\right]}{\mathbf{P}_x^T(Z_0 = 0)} \\
&= \frac{\mathbf{E}_x^T\left[e^{-\eta \Lambda_t}, \ Z_t = 0\right]}{\mathbf{P}_x^T(Z_0 = 0)} \\
&= e^{u_T(\infty)x}\mathbf{E}_x^T\left[e^{-(\eta + u_{T-t}(\infty))\Lambda_t}\right] \\
&= \mathbb{E}_x^T\left[e^{-\eta X_t}\right].
\end{aligned}$$

This tells us that the semigroups of Λ conditional on $\{Z_0 = 0\}$ and X conditional to become extinct by time T agree. Part (iii) of Theorem 2.2 is thus proved. $\qquad\square$

7. Thinning the skeleton to a spine: proof of Theorem 2.3

The aim of this section is to recover the unique solution to (1.7) as a weak limit of (2.5) in the sense of Skorokhod convergence. To this end, we assume throughout the conditions of Theorem 2.3, in particular that ψ is a critical or subcritical branching mechanism and Grey's condition (1.5) holds.

There are three main reasons why we should expect this result, and these three reasons pertain to the three structural features of the skeleton decomposition: the feature of Poisson embedding, the Galton–Watson skeleton, and the branching immigration from the skeleton with an Esscher-transformed branching mechanism. Let us dwell briefly on these heuristics.

First, let us consider the behaviour of the skeleton $(Z_t^T, t{<}T)$ as $T \to \infty$. As we are assuming that ψ is a (sub)critical branching mechanism, it holds that $\lim_{T\to\infty} u_T(\infty) = 0$ as $T \to \infty$. Thus, recalling that $Z_0^T \sim \mathrm{Po}(u_T(\infty)x)$, i.e. independent and Poisson distributed with parameter $u_T(\infty)x$, and hence conditioning on survival to time T in the skeletal decomposition is tantamount to conditioning on the event $\{Z_0^T \geq 1\}$, we see that

$$\varpi_k^{x,T} := \mathbf{P}_x^T[Z_0 = k \mid Z_0 \geq 1] = \frac{(u_T(\infty)x)^k}{k!} \frac{\mathrm{e}^{-u_T(\infty)x}}{1 - \mathrm{e}^{-u_T(\infty)x}}, \qquad k \geq 1. \tag{7.1}$$

We thus see that the probabilities (7.1) all tend to zero unless $k = 1$, in which case the limit is unity. Moreover, Theorem 2.2(ii) and (iii) imply that the law of $(\Lambda_t^T, 0 \leq t{<}T)$ conditional on $(\mathcal{F}^{\Lambda_t^T} \cap \{Z_0^T \geq 1\}, 0 \leq t{<}T)$ corresponds to the law of the ψ-CSBP, X, conditioned to survive until time T. Intuitively, then, one is compelled to believe that, in law, there is asymptotically a single skeletal contribution to the law of X conditioned to survive.

Second, considering (5.1), it follows from l'Hospital's rule that the rate at which the aforementioned most common recent ancestor branches begins to slow down, since

$$\frac{\psi(u_T(\infty))}{u_T(\infty)} \frac{u_{T-t}(\infty)}{\psi(u_{T-t}(\infty))} \to 1 \quad \text{as } T \to \infty.$$

What we are thus observing is a thinning, in the weak sense, of the skeleton in terms of the number of branching events.

Third, we consider the mass that immigrates from the skeleton. For a fixed T, it evolves as a ψ-CSBP conditioned to die before time T. We recall that conditioning to die before time T is tantamount to the change of measure given in (5.2). It is easy to see that, as $T \to \infty$, the density in this change of measure converges to unity and hence immigrating mass, in the weak limit, should have the evolution of a ψ-CSBP.

With all this evidence in hand, Theorem 2.3 should now take on a natural meaning. We give its proof below.

Proof of Theorem 2.3. According to Theorem 2.5 on p. 167 of [16, Chapter 4], if E is a locally compact and separable metric space, $\mathcal{P}^T := (\mathcal{P}_t^T, t \geq 0)$, $T > 0$, is a sequence of Feller semigroups on $C_0(E)$ (the space of continuous functions on E vanishing at ∞, endowed with the supremum norm), $\mathcal{P} := (\mathcal{P}_t, t \geq 0)$ is a Feller semi-group on $C_0(E)$ such that, for $f \in C_0(E)$, with respect to the supremum norm on the space $C_0(E)$,

$$\lim_{T \to \infty} \mathcal{P}_t^T f = \mathcal{P}_t f, \qquad t \geq 0, \tag{7.2}$$

and moreover, $(\nu^T, T > 0)$ is a sequence of Borel probability measures on E such that $\lim_{T \to \infty} \nu^T = \nu$ weakly for some probability measure ν; then, with respect to the Skorokhod topology on $\mathbb{D}([0, \infty), E)$, Ξ^T converges weakly to Ξ, where $(\Xi^T, T > 0)$ are the strong Markov processes associated to $(\mathcal{P}^T, T > 0)$ with initial law $(\nu^T, T > 0)$ and Ξ is the strong Markov processes associated to $\mathcal{P}$ with initial law ν, respectively.

Note that such weak convergence results would normally require a tightness criterion; however, having the luxury of (7.2), where $\mathcal{P}$ is a Feller semi-group, removes this condition and this will be the setting in which we are able to apply the conclusion of the previous paragraph.

Fix $t_0 > 0$. We want to prove the weak convergence result in the finite time window $[0, t_0]$. In order to introduce the role of $\mathcal{P}^T$, $T > 0$, in our setting, we will abuse yet further the previous notation and define $\mathbf{P}^T_{x,n,s}, x \geq 0, n \geq 0, s \in [0, t_0]$ to be the Markov probabilities associated to the three-dimensional process $(\Lambda_t, Z_t, \tau_t), 0 \leq t \leq t_0$, whenever $t_0 < T$, where (Λ_t, Z_t), $0 \leq t \leq t_0$, is the weak solution to (2.5), and $\tau_t := t, 0 \leq t \leq t_0$. Consistently with the previous notation, we have $\mathbf{P}^T_{x,n,0} = \mathbf{P}^T_{x,n}, x \geq 0, n \geq 0$. Now define the associated time-dependent semigroup for the three-dimensional process, for $t \geq 0$ and $f \in C_0([0, \infty) \times \mathbb{N}_0 \times [0, \infty))$, such that $\mathbb{P}^T_t[f](x, n, s) = f(x, n, s)$ when $T \leq t_0$, and when $T > t_0$ we have

$$\mathbb{P}^T_t[f](x, n, s) := \mathbf{E}^T[f(\Lambda_{(t \vee s) \wedge t_0}, Z_{(t \vee s) \wedge t_0}, \tau_{(t \vee s) \wedge t_0}) \mid \Lambda_s = x, Z_s = n, \tau_s = s]$$

for $(x, n, s) \in [0, \infty) \times \mathbb{N}_0 \times [0, t_0]$, and $\mathbb{P}^T_t[f](x, n, s) := f(x, n, s)$ for $(x, n, s) \in [0, \infty) \times \mathbb{N}_0 \times (t_0, \infty)$. We take $\mathcal{P}^T = \mathbb{P}^T$. In order to verify the Feller property of $\mathbb{P}^T_t[f](x, n, s)$ we need to check two things (cf. Proposition 2.4 in Chapter III of [41]):

(i) For each $t \geq 0$, the function $(x, n, s) \mapsto \mathbb{P}^T_t[f](x, n, s)$ belongs to $C_0([0, \infty) \times \mathbb{N}_0 \times [0, \infty))$ for any f in that space.

(ii) For all $f \in C_0([0, \infty) \times \mathbb{N}_0 \times [0, \infty))$ and for each $(x, n, s) \in [0, \infty) \times \mathbb{N}_0 \times [0, \infty)$, we have $\lim_{t \downarrow s} \mathbb{P}^T_t[f](x, n, s) = f(x, n, s)$.

Note that when $T \leq t_0$, or $s \geq t_0$, or $t \leq s \leq t_0$, we have $\mathbb{P}^T_t[f](x, n, s) = f(x, n, s)$. Since $f \in C_0([0, \infty) \times \mathbb{N}_0 \times [0, \infty))$, both (i) and (ii) are trivially satisfied. We can also notice that the case when $T > t_0$, $s \leq t_0$ and $t \geq t_0$ reduces to the case of $t = t_0$, hence in order to show the Feller property of $\mathbb{P}^T_t[f](x, n, s)$ we can restrict ourselves to the case of $s \leq t \leq t_0 < T$.

By the denseness of the subalgebra generated by exponential functions (according to the uniform topology) in $C_0(E)$, it suffices to check, for (i), that

$$(x, n, s) \mapsto \mathbf{E}^T_{x,n,s} \left[e^{-\gamma \Lambda_t - \theta Z_t - \varphi \tau_t} \right], \qquad s \leq t \leq t_0 < T, \tag{7.3}$$

belongs to $C_0([0, \infty) \times \mathbb{N}_0 \times [0, \infty))$ and, for (ii), that

$$\lim_{t \downarrow s} \mathbf{E}^T_{x,n,s} \left[e^{-\gamma \Lambda_t - \theta Z_t - \varphi \tau_t} \right] = e^{-\gamma x - \theta n - \varphi s}, \qquad s \leq t \leq t_0 < T. \tag{7.4}$$

To this end, note that

$$\mathbf{E}^T_{x,n,s} \left[e^{-\gamma \Lambda_t - \theta Z_t - \varphi \tau_t} \right] = \mathbf{E}^{T-s}_{x,n} \left[e^{-\gamma \Lambda_{t-s} - \theta Z_{t-s}} \right] e^{-\varphi t}, \qquad s \leq t \leq t_0 < T. \tag{7.5}$$

In order to evaluate the expectation on the right-hand side above, we want to work with an appropriate representation of the unique weak solution to (2.5). We shall do so by following the example of how the weak solution to (2.3) was identified in the form (4.1).

As before, we need to introduce additionally marked versions of the Poisson random measures N^1_T and N^2_T, as well as an additional Poisson random measure N^*_T. We

will insist that Poisson random measure $\mathrm{N}_T^1(\mathrm{d}s, \mathrm{d}r, \mathrm{d}j, \mathrm{d}\omega)$ on $[0, T) \times [0, \infty) \times \mathbb{N}_0 \times \mathbb{D}([0, \infty), \mathbb{R})$ has intensity $\mathrm{d}s \otimes r\mathrm{e}^{-u_{T-s}(\infty)r}\Pi(\mathrm{d}r) \otimes \sharp(\mathrm{d}j) \otimes \mathbb{P}_r^{T-s}(\mathrm{d}\omega)$, Poisson random measure $\mathrm{N}_T^2(\mathrm{d}s, \mathrm{d}r, \mathrm{d}k, \mathrm{d}j, \mathrm{d}\omega)$ on $[0, T) \times [0, \infty) \times \mathbb{N}_0 \times \mathbb{N} \times \mathbb{D}([0, \infty), \mathbb{R})$ has intensity

$$q^{T-s}\mathrm{d}s \otimes \eta_k^{T-s}(\mathrm{d}r) \otimes p_k^{T-s}\sharp(\mathrm{d}k) \otimes \sharp(\mathrm{d}j) \otimes \mathbb{P}_r^{T-s}(\mathrm{d}\omega),$$

and Poisson random measure $\mathrm{N}_T^*(\mathrm{d}s, \mathrm{d}j, \mathrm{d}\omega)$ has intensity $2\beta\mathrm{d}s \otimes \sharp(\mathrm{d}j) \otimes \mathbb{Q}^{T-s}(\mathrm{d}\omega)$ on $[0, T) \times \mathbb{N}_0 \times \mathbb{D}([0, \infty), \mathbb{R})$, where $\mathbb{Q}^T$ is the excursion measure associated to $\mathbb{P}_r^T$, $r \geq 0$, satisfying

$$\mathbb{Q}^T(1 - \mathrm{e}^{-\gamma\omega_t}) = V_t^T(\gamma), \qquad \gamma > 0,$$

for $0 \leq t < T$, where V_t^T was defined in (5.4). To recall some of the notation used in these rates, see (2.4) and (2.6).

If the pair (Λ, Z) has law $\mathbf{P}_{x,n}^T$, then we can write

$$\Lambda_t = X_t + D_t, \qquad t < T, \tag{7.6}$$

where X is autonomously independent with law $\mathbb{P}_x^T$ and, given N^2, D is the uniquely identified (up to almost sure modification) 'dressed skeleton' described by

$$\begin{aligned}
D_t &= \int_0^t \int_0^\infty \int_1^{Z_{s-}} \int_{\mathbb{D}([0,\infty),\mathbb{R})} \omega_{t-s}\, \mathrm{N}_T^1(\mathrm{d}s, \mathrm{d}r, \mathrm{d}j, \mathrm{d}\omega) \\
&+ \int_0^t \int_0^\infty \int_0^\infty \int_1^{Z_{s-}} \int_{\mathbb{D}([0,\infty),\mathbb{R})} \omega_{t-s}\, \mathrm{N}_T^2(\mathrm{d}s, \mathrm{d}r, \mathrm{d}k, \mathrm{d}j, \mathrm{d}\omega) \\
&+ \int_0^t \int_1^{Z_{s-}} \int_{\mathbb{D}([0,\infty),\mathbb{R})} \omega_{t-s}\, \mathrm{N}_T^*(\mathrm{d}s, \mathrm{d}j, \mathrm{d}\omega),
\end{aligned}$$

where $Z_0 = n$. The verification of this claim follows almost verbatim that for (2.3), albeit with the obvious change to take account of the time-varying rates. We therefore omit the proof and leave it as an exercise for the reader.

With the representation (7.6), as Z is piecewise constant we can condition on the sigma-algebra generated by N_T^2 and show, using Campbell's formula in between the jumps of Z, that, for $0 \leq t < T$, $\gamma, \theta \geq 0$, $x \geq 0$, and $n \in \mathbb{N}_0$,

$$\begin{aligned}
\mathbf{E}_{x,n}^T \left[\mathrm{e}^{-\gamma\Lambda_t - \theta Z_t} \right] = \mathrm{e}^{-xV_t^T(\gamma)}\mathbf{E}_{0,n}^T &\left[\exp\left\{ -\theta Z_t - \int_0^t Z_v \phi_{u_{T-v}(\infty)}(V_{t-v}^{T-v}(\gamma))\mathrm{d}v \right\} \right. \\
&\left. \prod_{w \leq t} \left(\int_0^\infty \mathrm{e}^{-rV_{t-w}^{T-w}(\gamma)}\eta_{\Delta Z_w+1}^{T-w}(\mathrm{d}r) \right) \right], \tag{7.7}
\end{aligned}$$

where

$$V_t^T(\gamma) := u_t(\gamma + u_{T-t}(\infty)) - u_T(\infty), \qquad 0 \leq t < T,$$

and, for $\lambda, z \geq 0$,

$$\phi_\lambda(z) = 2\beta z + \int_0^\infty (1 - \mathrm{e}^{-zr})r\mathrm{e}^{-\lambda r}\,\Pi(\mathrm{d}r).$$

Given the identities (7.5) and (7.7), the two required verifications in (7.3) and (7.4) follow easily as a direct consequence of the continuity and bounded convergence in (7.7).

The target semigroup $\mathcal{P}$ on $f \in C_0([0, \infty) \times \mathbb{N}_0 \times [0, \infty))$ is defined as follows. For fixed $n \in \mathbb{N}_0$, $x \geq 0$, let $\mathbb{P}_x^{(n)}$ be the law of the homogeneous Markov process described by the weak solution to

$$X_t = x + \alpha \int_0^t X_{s-} \, ds + \sqrt{2\beta} \int_0^t \int_0^{X_{s-}} W(ds, du) + \int_0^t \int_0^\infty \int_0^{X_{s-}} r\tilde{N}(ds, dr, du)$$

$$+ \int_0^t \int_0^\infty rN^{(*,n)}(ds, dr) + 2n\beta t, \qquad t \geq 0,$$

with W, N, and $N^{(*,n)}$ Poisson random measures on $[0, \infty)^2 \times \mathbb{D}([0, \infty), \mathbb{R})$ with intensity measure $nds \otimes r\Pi(dr) \otimes \mathbb{P}_r(d\omega)$. Note, we have that, at no detriment to consistency, $\mathbb{P}_x^{(0)}$ can be replaced by $\mathbb{P}_x$. Then, we take the role of $\mathcal{P}_t$ played by the semi-group $\mathbb{P}_t^\uparrow$ given by

$$\mathbb{P}_t^\uparrow[f](x, n, s) := \mathbf{E}^{(n)}[f(X_{(t \vee s) \wedge t_0}, n, \tau_{(t \vee s) \wedge t_0}) \mid X_s = x, \tau_s = s]$$

for $(x, n, s) \in [0, \infty) \times \mathbb{N}_0 \times [0, t_0]$, and $\mathbb{P}_t^\uparrow[f](x, n, s) := f(x, n, s)$ otherwise. Here, $f \in C_0([0, \infty) \times \mathbb{N}_0 \times [0, \infty))$ and $\tau_t = t$, as above. Notice that $(X, \mathbb{P}_x^{(n)})$ is a branching process with immigration, whose Laplace transform is given by

$$\mathbf{E}_x^{(n)}(e^{-\gamma X_t}) = \exp\left\{ -xu_t(\gamma) - n \int_0^t \phi_0(u_{t-v}(\gamma)) \, dv \right\}, \qquad \gamma \geq 0.$$

From this, it is easily seen that $\mathbb{P}_t^\uparrow$ is Feller as well.

Lastly, for each $T \geq 0$ we take ν^T as the measure on $[0, \infty) \times \mathbb{N}_0 \times [0, \infty)$ given for each $x \geq 0$ by $\delta_x \otimes \pi^{T,x} \otimes \delta_0$, with $\pi^{T,x}(\cdot) = \sum_{n \geq 1} \varpi_n^{T,x} \delta_n(\cdot)$. Recall from (7.1) that $\pi^{T,x}$ converges weakly, as $T \to \infty$, to the measure $\delta_1(\cdot)$ on $\mathbb{N}_0$, hence ν^T converges weakly to $\nu := \delta_x \otimes \delta_1 \otimes \delta_0$. Thus, in order to invoke Theorem 2.5 of [16, Chapter 4], we just need to check the analogue of (7.2) in our setting.

To this end, notice first that we can restrict ourselves to $0 \leq s \leq t \leq t_0$, since when $s > t_0$, $\mathbb{P}_t^\uparrow[f](x, n, s) = \mathbb{P}_t^T[f](x, n, s)$ by definition, and the case when $0 \leq s \leq t_0 \leq t$ reduces to the case when $t = t_0$. Then, note from (2.6) that $q^T \to 0$ as $T \to \infty$, and this yields that, under $\mathbf{P}_{0,n}^T$, the process Z converges in probability uniformly on $[0, t]$ as $T \to \infty$ (cf. Theorem 6.1, Chapter 1, p. 28 of [16]) to the constant process $Z_s \equiv n$, $s \leq t$. Referring back to (7.7), the continuity in T of the deterministic quantities as they appear on the right-hand side and the previously mentioned uniform convergence of $(Z, \mathbf{P}_{0,n}^T)$ together imply that, for $x \geq 0$, $0 \leq s \leq t \leq t_0$, $n \in \mathbb{N}_0$,

$$\lim_{T \to \infty} \mathbb{P}_t^T[f_{\gamma,\theta,\varphi}](x, n, s) = e^{-\varphi t} \lim_{T \to \infty} \mathbf{E}_{x,n}^{T-s}\left[e^{-\gamma \Lambda_{t-s} - \theta Z_{t-s}} \right]$$

$$= \mathbb{P}_t^\uparrow[f_{\gamma,\theta,\phi}](x, n, s)$$

$$= \exp\left\{ -xu_{t-s}(\gamma) - \theta n - n \int_0^{t-s} \phi_0(u_{t-s-v}(\gamma)) \, dv - \varphi t \right\},$$

where

$$f_{\gamma,\theta,\varphi}(x, n, s) := e^{-\gamma x - \theta n - \varphi s}, \qquad \gamma, \theta, \varphi, x, s \geq 0, \quad n \in \mathbb{N}_0.$$

To conclude, it is thus enough to prove that this convergence holds uniformly in $x \geq 0$, $0 \leq s \leq t$, $n \in \mathbb{N}_0$, where $t \leq t_0$. Consider fixed $R > 0$ and $N \in \mathbb{N}$. Since $V_t^T(\gamma)$ defined above is

non-negative and, for each $n \in \mathbb{N}_0$, $Z_t \geq Z_0 = n$, $t \geq 0$, almost surely under $\mathbf{P}_{0,n}^T$, for all $T > 0$, using the triangle inequality, we have

$$\sup_{x \geq 0, s \leq t, n \in \mathbb{N}_0} |\mathbb{P}_t^T[f_{\gamma,\theta,\varphi}](x, n, s) - \mathbb{P}_t^{\uparrow}[f_{\gamma,\theta,\varphi}](x, n, s)| \leq A_R(T) + A^R(T) + B_N(T) + B^N,$$

where we have set:

$$A_R(T) := \sup_{x \leq R, s \leq t} \left| e^{-xV_{t-s}^{T-s}(\gamma)} - e^{-xu_{t-s}(\gamma)} \right|,$$

$$A^R(T) := \sup_{s \leq t} e^{-RV_{t-s}^{T-s}(\gamma)} + \sup_{s \leq t} e^{-Ru_{t-s}(\gamma)},$$

$$B_N(T) := \sup_{n \leq N, s \leq t} \left| \mathbf{E}_{0,n}^{T-s} \left[\exp \left\{ -\theta Z_{t-s} - \int_0^{t-s} Z_v \phi_{u_{T-s-v}(\infty)}(V_{t-s-v}^{T-s-v}(\gamma)) \, dv \right\} \right. \right.$$
$$\left. \left. \prod_{w \leq t-s} \left(\int_0^\infty e^{-rV_{t-s-w}^{T-s-w}(\gamma)} \eta_{\Delta Z_w + 1}^{T-s-w}(dr) \right) \right]$$
$$- \exp \left\{ -\theta n - n \int_0^{t-s} \phi_0(u_{t-s-v}(\gamma)) \, dv \right\} \right|,$$

and $B^N = 2e^{-\theta N}$. First, it is not hard to see that

$$A_R(T) \leq \sup_{s \leq t} R \left| V_{t-s}^{T-s}(\gamma) - u_{t-s}(\gamma) \right| = R \sup_{s \leq t} \left| u_{t-s}(\gamma + u_{T-t}(\infty)) - u_{T-s}(\infty) - u_{t-s}(\gamma) \right|.$$

The identity $\partial u_s(\theta)/\partial \theta = \exp \left\{ -\int_0^s \psi'(u_r(\theta)) \, dr \right\}$ (see (12.12) in [24, Chapter 12]) and the fact that $\psi'(\theta) \geq 0$ allows us to estimate $|u_{t-s}(\gamma + u_{T-t}(\infty)) - u_{t-s}(\gamma)|$ by $u_{T-t}(\infty)$. Recalling that $u_T(\infty) \to 0$ as $T \to \infty$, it follows that $A_R(T)$ tends to 0 as $T \to \infty$, for each $R > 0$. Next, since $(s, \gamma) \mapsto u_s(\gamma)$ is increasing in γ and decreasing in s, we have

$$V_{t-s}^{T-s}(\gamma) \geq u_{t-s}(\gamma + u_{T-t}(\infty)) - u_{T-t}(\infty) \geq \inf_{\lambda \leq u_{T-t}(\infty)} u_t(\gamma + \lambda) - \lambda,$$

which, for T sufficiently large, is bounded from below by $u_t(\gamma)/2 > 0$. Fix $\varepsilon > 0$. Choosing $R > 0$ such that $e^{-Ru_t(\gamma)/2} + e^{-Ru_t(\gamma)} \leq \varepsilon$, we thus get

$$\limsup_{T \to \infty} A^R(T) \leq \varepsilon.$$

With regard to the term $B_N(T)$, we have

$$B_N(T) \leq \max_{n \leq N} \sup_{s \leq t} \mathbf{E}_{0,n}^{T-s} \left[1 \wedge \left(\int_0^{t-s} \left| Z_v \phi_{u_{T-s-v}(\infty)}(V_{t-s-v}^{T-s-v}(\gamma)) \right. \right. \right.$$
$$\left. \left. \left. - n\phi_0(u_{t-s-v}(\gamma)) \right| \, dv \right) \right]$$
$$+ \max_{n \leq N} \sup_{s \leq t} \mathbf{E}_{0,n}^{T-s} \left[\left| e^{-\theta Z_{t-s}} - e^{-\theta n} \right| \right]$$
$$+ \max_{n \leq N} \sup_{s \leq t} \mathbf{E}_{0,n}^{T-s} \left[\left| 1 - \prod_{w \leq s} \left(\int_0^\infty e^{-rV_{t-s-w}^{T-s-w}(\gamma)} \eta_{\Delta Z_w + 1}^{T-s-w}(dr) \right) \right| \right]$$

$$\leq \max_{n \leq N} \sup_{s \leq t} \mathbf{E}_{0,n}^{T-s} \Bigg[\sup_{s' \leq t} 1 \wedge \Bigg(\int_0^{t-s'} \Big| Z_v \phi_{u_{T-s'-v}}(\infty)(V_{t-s-v}^{T-s'-v}(\gamma))$$

$$- n\phi_0(u_{t-s'-v}(\gamma)) \Big| \, \mathrm{d}v \Bigg) \Bigg]$$

$$+ \max_{n \leq N} \sup_{s \leq t} \mathbf{E}_{0,n}^{T-s} \Bigg[\sup_{s' \leq t} \Big| \mathrm{e}^{-\theta Z_{t-s'}} - \mathrm{e}^{-\theta n} \Big| \Bigg]$$

$$+ \max_{n \leq N} \sup_{s \leq t} \mathbf{E}_{0,n}^{T-s} \Bigg[\sup_{s' \leq t} \Bigg| 1 - \prod_{w \leq s'} \Bigg(\int_0^\infty \mathrm{e}^{-r V_{t-s'-w}^{T-s'-w}(\gamma)} \eta_{\Delta Z_w + 1}^{T-s'-w} (\mathrm{d}r) \Bigg) \Bigg| \Bigg]. \qquad (7.8)$$

The first term on the right-hand side above is bounded by

$$\max_{n \leq N} \sup_{s \leq t} \mathbf{P}_{0,n}^{T-s} \Bigg(\sup_{v \leq t} Z_v > n \Bigg) + 1 \wedge \Bigg(Nt \sup_{w \leq t} \Big| \phi_{u_{T-w}}(\infty)(V_{t-w}^{T-w}(\gamma)) - \phi_0(u_{t-w}(\gamma)) \Big| \Bigg),$$

and hence goes to 0 for each N as $T \to \infty$. On the other hand, as a function of $(Z_s, s \leq t)$, the expression inside the expectation in the second term of (7.8) is bounded and continuous with respect to the Skorokhod topology (recall that Skorokhod continuity is preserved for Z under the operation of supremum over finite time horizons). Moreover, it vanishes when $Z_s \equiv n$, $0 \leq s \leq t$. This implies that this term goes to 0 as well. Finally, the expression whose absolute value we take in the third term of (7.8) is bounded by 1, and vanishes unless Z jumps at least once on [0,s]. This shows that the last term is bounded by $\max_{n \leq N} \sup_{s \leq t} \mathbf{P}_{0,n}^{T-s}(\sup_{w \leq t} \Delta Z_w > 0)$, which goes to 0 when $T \to \infty$. Note that for all three terms in (7.8), we are using the fact that if $g(T) \geq 0$ is continuous in T and $\lim_{T \to \infty} g(T) = 0$ then, for each $\varepsilon > 0$ and for $0 < t \leq t_0$, by choosing T sufficiently large, we have $\sup_{s \leq t} g(T-s) < \varepsilon$. That is to say, $\lim_{T \to \infty} \sup_{s \leq t} g(T-s) = 0$.

Putting the pieces together and choosing $N \in \mathbb{N}_0$ large enough that $B^N \leq \varepsilon$, we thus get

$$\limsup_{T \to \infty} \sup_{x \geq 0, s \leq t, n \in \mathbb{N}_0} \Big| \mathbb{P}_t^T [f_{\gamma, \theta, \varphi}](x, n, s) - \mathbb{P}_t^\uparrow [f_{\gamma, \theta, \varphi}](x, n, s) \Big| \leq 2\varepsilon.$$

Since ε was arbitrary this shows the convergence of the semi-groups (7.2) in our setting which, together with the weak convergence of the initial configurations, gives the weak convergence of the associated processes on $[0, t_0]$. And since $t_0 > 0$ was chosen arbitrarily, this completes the proof of Theorem 2.3. $\qquad \square$

Acknowledgements

References

[1] ABRAHAM, R. AND DELMAS, J.-F. (2012). A continuum-tree-valued Markov process. *Ann. Prob.* **40**, 1167–1211.

[2] BARCZY, M., LI, Z. AND PAP, G. (2015). Yamada–Watanabe results for stochastic differential equations with jumps. *Internat. J. Stoch. Anal.* **2015**, 460472.

[3] BERESTYCKI, J., FITTIPALDI, M. C. AND FONTBONA, J. (2018). Ray–Knight representation of flows of branching processes with competition by pruning of Lévy trees. *Prob. Theory Relat. Fields* **172**, 725–788.

[4] BERESTYCKI, J., KYPRIANOU, A. E. AND MURILLO-SALAS, A. (2011). The prolific backbone for supercritical superprocesses. *Stoch. Process. Appl.* **121**, 1315–1331.

[5] BERTOIN, J., FONTBONA, J. AND MARTÍNEZ, S. (2008). On prolific individuals in a supercritical continuous-state branching process. *J. Appl. Prob.* **45**, 714–726.

[6] BERTOIN, J. AND LE GALL, J.-F. (2006). Stochastic flows associated to coalescent processes III: Limit theorems. *Illinois J. Math.* **50**, 147–181.

[7] DAWSON, D. A. AND LI, Z. (2006). Skew convolution semigroups and affine Markov processes. *Ann. Prob.* **34**, 1103–1142.

[8] DAWSON, D. A. AND LI, Z. (2012). Stochastic equations, flows and measure-valued processes. *Ann. Prob.* **40**, 813–857.

[9] DUQUESNE, T. AND LABBÉ, C. (2014). On the Eve property for CSBP. *Electron. J. Prob.* **19**, 31.

[10] DUQUESNE, T. AND LE GALL, J.-F. (2005). Random trees, Lévy processes and spatial branching processes. *Astérisque* **281**.

[11] DUQUESNE, T. AND WINKEL, M. (2007). Growth of Lévy trees. *Prob. Theory Relat. Fields* **139**, 313–371.

[12] DYNKIN, E. B. AND KUZNETSOV, S. E. (2004). ℕ-measures for branching exit Markov systems and their applications to differential equations. *Prob. Theory Relat. Fields* **130**, 135–150.

[13] ECKHOFF, M., KYPRIANOU, A. E. AND WINKEL, M. (2015). Spines, skeletons and the strong law of large numbers for superdiffusions. *Ann. Prob.* **43**, 2545–2610.

[14] EL KAROUI, N. AND ROELLY, S. (1991). Propriétés de martingales, explosion et représentation de Lévy–Khintchine d'une classe de processus de branchement à valeurs mesures. *Stoch. Process. Appl.* **38**, 239–266.

[15] ETHERIDGE, A. M. AND WILLIAMS, D. R. E. (2003). A decomposition of the $(1 + \beta)$-superprocess conditioned on survival. Proc. R. Soc. Edinburgh Sect. A 133, 829–847.

[16] ETHIER, S. N. AND KURTZ, T. G. (1986). Markov Processes. Wiley Series in Probability and Mathematical Statistics. John Wiley, New York.

[17] EVANS, L. C. (1998). *Partial Differential Equations*, Vol. **19**, Graduate Studies in Mathematics. American Mathematical Society, Providence, RI.

[18] EVANS, S. N. AND O'CONNELL, N. (1994). Weighted occupation time for branching particle systems and a representation for the supercritical superprocess. *Canad. Math. Bull.* **37**, 187–196.

[19] FITTIPALDI, M. C. AND FONTBONA, J. (2012). On SDE associated with continuous-state branching processes conditioned to never be extinct. *Electron. Commun. Probab.* **17**, 13.

[20] FU, Z. AND LI, Z. (2010). Stochastic equations of non-negative processes with jumps. *Stoch. Process. Appl.* **120**, 306–330.

[21] GREY, D. R. (1976). On possible rates of growth of age-dependent branching processes with immigration. *J. Appl. Prob.* **13**, 138–143.

[22] HARRIS, S. C., HESSE, M. AND KYPRIANOU, A. E. (2016). Branching Brownian motion in a strip: Survival near criticality. *Ann. Prob.* **44**, 235–275.

[23] KURTZ, T. (2011). Equivalence of stochastic equations and martingale problems. In Stochastic Analysis 2010, ed. D. CRISAN, Springer, New York.

[24] KYPRIANOU, A. E. (2014). *Fluctuations of Lévy Processes with Applications*, 2nd edn. Springer, Heidelberg.

[25] KYPRIANOU, A. E., MURILLO-SALAS, A. AND PÉREZ, J. L. (2012). An application of the backbone decomposition to supercritical super-Brownian motion with a barrier. *J. Appl. Prob.* **49**, 671–684.

[26] KYPRIANOU, A. E., PÉREZ, J.-L. AND REN, Y.-X. (2014). The backbone decomposition for spatially dependent supercritical superprocesses. In Séminaire de Probabilités XLVI, Vol. 2123, Lecture Notes in Mathematics, Springer, Berlin, pp. 33–59.

[27] KYPRIANOU, A. E. AND REN, Y.-X. (2012). Backbone decomposition for continuous-state branching processes with immigration. *Statist. Prob. Lett.* **82**, 139–144.

[28] LAMBERT, A. (2007). Quasi-stationary distributions and the continuous-state branching process conditioned to be never extinct. *Electron. J. Prob.* **12**, 420–446.

[29] LAMBERT, A. (2008). Population dynamics and random genealogies. *Stoch. Models* **24**, 45–163.

[30] LE GALL, J.-F. (1999). Spatial Branching Processes, Random Snakes and Partial Differential Equations. Lectures in Mathematics, ETH Zürich. Birkhäuser, Basel.

[31] LE GALL, J.-F. AND LE JAN, Y. (1998). Branching processes in Lévy processes: Laplace functionals of snakes and superprocesses. *Ann. Prob.* **26**, 1407–1432.

[32] LE GALL, J.-F. AND LE JAN, Y. (1998). Branching processes in Lévy processes: The exploration process. *Ann. Prob.* **26** 213–252.

[33] LI, CH.-KH. (2001). Skew convolution semigroups and related immigration processes. *Teor. Veroyat. Primen.* **46**, 247–274.

[34] LI, Z.-H. (1996). Immigration structures associated with Dawson–Watanabe superprocesses. *Stoch. Process. Appl.* **62**, 73–86.

[35] LI, Z.-H. (1998). Immigration processes associated with branching particle systems. *Adv. Appl. Prob.* **30**, 657–675.

[36] LI, Z. (2011). Measure-Valued Branching Markov Processes. Probability and its Applications. Springer, Heidelberg.

[37] LI, Z. (2014). Path-valued branching processes and nonlocal branching superprocesses. *Ann. Prob.* **42**, 41–79.

[38] LI, Z. (2019). Continuous-state branching processes with immigration. Preprint, arXiv:1901.03521.

[39] MILOS, P. (2012). Spatial CLT for the supercritical Ornstein–Uhlenbeck superprocess. Preprint, arXiv:1203.6661.

[40] PROTTER, P. (1990). *Stochastic Integration and Differential Equations*, Vol. **21** of Applications of Mathematics (New York). Springer, Berlin.

[41] REVUZ, D. AND YOR, M. (1999). *Continuous Martingales and Brownian Motion*, 3rd edn, Vol. **293** of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer, Berlin.

[42] ROELLY-COPPOLETTA, S. AND ROUAULT, A. (1989). Processus de Dawson–Watanabe conditionné par le futur lointain. *C. R. Acad. Sci. Paris Sér. I Math.* **309**, 867–872.

[43] SALISBURY, T. S. AND VERZANI, J. (1999). On the conditioned exit measures of super Brownian motion. *Prob. Theory Relat. Fields* **115**, 237–285.

[44] SALISBURY, T. S. AND VERZANI, J. (2000). Non-degenerate conditionings of the exit measures of super Brownian motion. *Stoch. Process. Appl.* **87**, 25–52.

[45] SHEU, Y.-C. (1997). Lifetime and compactness of range for super-Brownian motion with a general branching mechanism. *Stoch. Process. Appl.* **70**, 129–141.