SDEs for embedded successful genealogies

submitted by

Dorottya Fekete

for the degree of Doctor of Philosophy

of the

University of Bath

Department of Mathematical Sciences

July 2019

Signature of Author ....

Dorottya Fekete

Summary

This thesis concerns different skeletal decompositions of various branching processes. First we develop a stochastic differential equation (SDE) approach to describe the fitness of certain sub-populations in asexual high-density stochastic population models. Initially we only consider continuous-state branching processes, then we extend the SDE approach to the spatial setting of superprocesses. In both cases the SDE can be used to simultaneously describe the total mass, and those embedded genealogies that propagate prolific traits in surviving populations, where 'survival' can be interpreted in different ways. For example, it can mean survival beyond a certain time-horizon, but it can also mean survival according to some spatial criteria.

In the second part of the thesis we construct the prolific backbone decomposition of multitype superprocesses by extending the semigroup approach already available for one-type branching processes.

Acknowledgements

I'm grateful to so many people who one way or another helped me over the course of my years in Bath.

First of all I would like to thank my supervisor, Andreas Kyprianou for his friendship, continuous guidance and support; your enthusiasm towards research, and the friendly atmosphere of our meetings always got me excited about academia, no matter how I'd felt hours before. I'm especially grateful for taking me to all those conferences around the world, where apart from learning a lot I also got to know a lot of inspiring probabilists, many of whom I'm lucky to call my friend now. These trips played a massive role in the completion of this thesis, in fact, the most productive work during my PhD was probably done on those 10 hour long flights to Mexico. A massive thanks goes to my second supervisor, Juan Carlos Pardo, whose support, just like Andreas's, went way beyond the boundaries of academia. Thank you for all the fruitful academic discussions, and for finding the time to go for a coffee, whenever I needed it the most. I feel incredibly lucky to have worked with people who supported not only my academic development, but also cared for me on a personal level. I would also like to thank Peter Mörters, who while being my supervisor, always made time for me despite his extremely busy schedule.

A big thank you to my other co-authors, Joaquín Fontbona, José Luis Pérez and Sandra Palau from whom I have learnt a lot. I would also like to thank Sandra for being such a good friend; I enjoyed every minute we have spent together. I would also like to acknowledge all the members of Prob-L@b for providing an inspiring and friendly environment. Many thanks goes to my examiners Matt Roberts and Jean-François Delmas for taking the time to read the thesis and making very valuable suggestions for the final version.

I would like to give a huge thank you to Paul, Anna, Jess and especially Susie for having your door open all time, putting up with my personal drama, and truly going out of your way to make my SAMBa years more enjoyable. I will never forget the birthday cake you gave me in my first year being abroad, it meant a lot. I am also grateful to EPSRC for the four years of financial support.

A very special thank you to Emma, Tsogii and Alice for their friendship, encouragement and making my time in Bath unforgettable, and to Ramses for always being there for me during my time in Mexico, muchas gracias! I would also like to thank The Climbing Academy and all my friends there for providing a pleasant escape from the PhD struggles.

My warmest thanks I owe to my family, especially my parents Katalin and Sándor, for their unconditional love, enduring effort and endless support. Words cannot express how grateful I am for believing in me and working so hard to make this journey possible. This thesis is dedicated to you.

Contents

Chapter 1

Introduction

Historically branching processes have been used as stochastic models for the evolution of a population. They were originally introduced by Watson and Galton [32] in the 19th century to investigate the extinction of British aristocratic family names. Since then the original model has led to many generalisations, and branching processes have been a subject of extensive research. Their popularity is due to the fact that they can be widely applied in the physical and biological sciences; in modelling nuclear chain reactions, cosmic radiation, or bacterial reproduction, to mention just a few. Even though many aspects of branching processes are now well understood, survival remains a central topic of interest.

As we will see "survival" can be interpreted in many different ways; it might simply mean the existence of individuals that have an infinite genealogical line of descent, but in a different setup "survival" can also mean the existence of genealogies that stay in a strip, or visit a compact domain infinitely often. A common thread is that on the event of "survival", certain genealogies do something infinitely often while others do not, and thus we can split the individuals of the population into two categories. The so called prolific individuals are part of the surviving genealogies, and as such they form the backbone of the process. The rest of the population can be thought of as a "dressing" of this backbone, since they can be seen as genealogies grafted onto the trajectories of the prolific skeleton. Individuals of the dressing have no long term contribution to "survival", and therefore many results are true for the original process if and only if they are true for the backbone. But since the latter is usually a stochastically "thinner" object with almost sure survival, results are generally easier to prove for the backbone.

Decomposing different branching processes along their infinite genealogies has been an intensely researched topic in the recent years, which unavoidably led to not entirely consistent terminology. The expressions "backbone" and "skeleton" both appeared in research papers, signifying the same object. As in this thesis we study several different branching processes, while drawing inspiration from the vast literature already available on the topic, we will use the terms "skeleton" and "backbone" interchangeably.

To motivate our results, we start by giving some examples for stochastic processes and their backbones that already exist in the literature. We try to cover a wide range of processes by discussing branching processes in both discrete and continuous setup, as well as the spatial extension of these. But as our main aim here is to give an intuitive picture, we won't be presenting the results in full rigour, for more detail the reader is referred to the referenced literature.

1.1 Examples of backbones in literature

1.1.1 Galton-Watson process

Let us start with considering a Galton-Watson (GW) process $Z = (Z_t, t \geq 0)$ with offspring distribution A, and initial population $Z_0 = 1$. Here by survival we mean the event $\{\forall t \geq 0 : Z_t > 0\}$, that is when we have at least one individual in the population at any time t. It is well known that the probability of this event is positive if and only if the mean offspring number $m = \mathbb{E}[A]$ is greater than one, in which case we call the process supercritical. When $m \leq 1$, then the process goes extinct almost surely, and Z is said to be (sub)critical (critical when $m = 1$, and subcritical when $m < 1$).

Since GW processes are the most elementary examples of branching processes, it is not surprising that a prolific backbone was first described in terms of a supercritical GW-process. Even though the term "backbone" was not used at the time, it was shown that on the event of survival there exist a pathwise decomposition of the process into a GW process which has almost sure survival, and which is dressed with copies of the original GW process conditioned on extinction. To motivate the subsequent results we give a brief summary of this decomposition here; for a detailed description see for example Part D of Athreya and Ney [1].

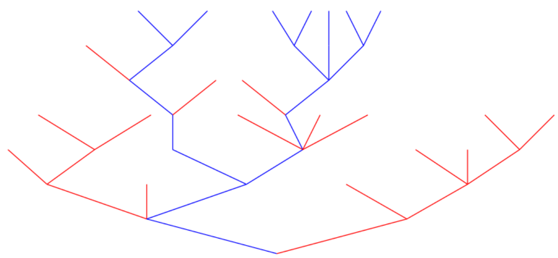

Figure 1-1: Backbone decomposition of a Galton-Watson process on the event of survival.

Consider a supercritical Galton-Watson process which might go extinct with positive probability. To get a visual picture of the backbone decomposition imagine that we colour blue every individual with an infinite line of descent, and colour red all the remaining individuals as shown in Figure 1-1. On the event of survival we see an infinite blue tree which is dressed with finite red trees, while on the event of extinction we only have a finite red tree. One can show that the two components, namely the red and blue trees are both GW processes, and in each case the generating function of the offspring distribution can be obtained as a transformation of the offspring generating function of the original process. To see what these transformations are let us denote the offspring generating function of the original GW process by F. Under our conditions F is strictly convex, and the extinction probability is given by the unique solution of $F(s) = s$ on $[0, 1)$, which we denote by q. Then the generating function of the blue tree is given by

$$F_b(s) := \frac{F((1-q)s + q) - q}{1 - q}, \quad s \in [0, 1],$$

while the generating function of the red tree is

$$F_r(s) := q^{-1}F(qs), \quad s \in [0,1].$$

Notice that $F_b'(1) = F'(1)$ and $F_b(0) = 0$, that is the blue process has the same mean as the original GW process, but at every branching event it produces at least one offspring. Without the possibility of having zero children all the genealogies of this process are infinite. On the other hand it is not hard to see that the GW process associated to F_r is subcritical, and as such it dies out almost surely.

We can also identify F_b and F_r from the graph of F. In particular, if Figure 1-2 is the graph of F, then we can get F_b by stretching out the square with opposing corners (q,q) and $(1,1)$ into the unit square. Similarly, if we take the graph within the square with opposing corners $(0,0)$ and (q,q), multiply it by q^{-1}, and yet again stretch it out into the unit square, we get the generating function of the red tree.

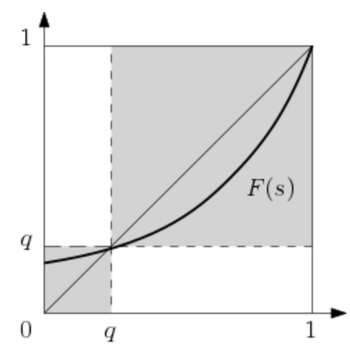

Figure 1-2: Graph of the generating function F split at q; rescaling the lower-left quadrant and upper-right quadrant gives the generating functions F_r and F_b respectively.

Without knowing the history of the process, we can find the prolific individuals at time t by making a Bernoulli selection. Each individual is prolific, independently of each other, with probability $1-q$. Non-prolific individuals can only give birth to other non-prolific individuals, where the number of children is determined by the subcritical offspring distribution F_r. Prolific individuals on the other hand can have both prolific and non-prolific children, however they always have at least one prolific child. More precisely, a blue particle gives birth to B blue children and R red ones with generating function

$$\mathbb{E}\left[\ell^B s^R\right] = \frac{F((1-q)\ell + qs) - F(qs)}{1-q}, \quad \ell, s \in [0,1].$$

When $B = 1$ we call the event an immigration, instead of branching, since while new red particles branch off the blue tree, each of them initiating a red tree, the number of prolific individuals remains the same.

It is also well known that the law of a tree initiated from a red particle is that of the original process conditioned to go extinct, while a blue individual initiates a process which has the same law as that of the original process conditioned to survive.

Next we discuss a spatial extension of the GW process, before moving onto processes with continuous state space. We will see that the main observations we made in this section will remain true in all of these cases. Namely that the characteristics of the prolific individuals and the immigration can be written as a transformation of the characteristics of the original process; that the law of the immigration is that of the original process conditioned on extinction, while the law of the dressed prolific tree is that of the process conditioned to survive; and that at any given time we can find the prolific individuals by making an appropriate random selection.

1.1.2 Branching Brownian motion in a strip

Spatial branching processes combine the branching phenomenon with spatial motion, where we monitor not only the size of the population, but also the location of the individuals. Here, we consider a branching particle system in $\mathbb{R}^d$, where the reproduction law follows a continuous time Galton-Watson process. More precisely, each particle or individual of the population moves according to some Markov process in $\mathbb{R}^d$, and after some exponential time with parameter $\beta > 0$ it dies and is replaced by a random number of children according to a distribution $(p_k)_{k\in\mathbb{N}_0}$. The offspring particles have starting position given by their birth location, they behave independently of each other, and they copy the behaviour of their parent.

If we think about the particles as atoms having mass one, then at a given time, the process can be described as the sum of Dirac measures at the locations of the particles that are alive at that time. Thus, we can say that the branching particle system takes its values in the space of finite atomic measures on $\mathbb{R}^d$, and as such it is also an example of a measure valued process.

A branching particle system can be characterised by two parameters, its motion and its branching generator. When particles move according to a Brownian motion (possibly with drift), the process is called a branching Brownian motion. Harris, Hesse and Kyprianou [18] studied a one-dimensional version of this process, where particles are killed on exiting the interval $(0, K)$. They were interested in the evolution of the process on the event of survival. As a result of the killing, being prolific now comes with an extra spatial constraint, since in order to have a surviving, infinite genealogy all the particles of that line of descent have to stay in the strip $(0, K)$.

The visual picture associated to the backbone decomposition is similar to what we saw in the previous section, see Figure 1-3. That is, on the event of survival, after colouring blue all the prolific individuals, and red all the remaining particles, the resulting image is an infinite blue tree dressed with finite red trees. The main difference here is that the colouring induces a spatially dependent bias not only on the offspring distribution, but also on the motion of the blue and red particles. This is intuitively not surprising, since surviving genealogies can never reach the killing boundary. Survival is thus guaranteed by eliminating the possibility of zero offspring, and having diffusion with space dependent drift term as the single-particle motion, where the drift pushes the particles away from the boundary.

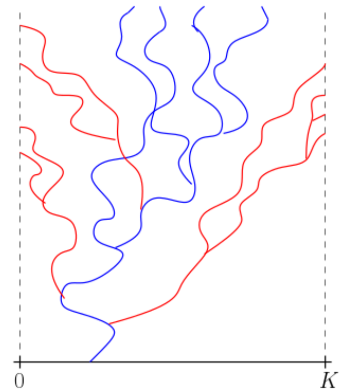

Figure 1-3: Skeletal decomposition of a BBM in the strip $(0,K)$ *on the event of survival.*

Just as in the Galton-Watson case we have that the red tree has the law of the process conditioned on extinction, while the law of the dressed blue tree is that of the process conditioned on survival. To make this statement slightly more precise we need to introduce some notation. Let $Z = (Z_t, t \geq 0)$ be a branching Brownian motion with branching generator $G(s) = \beta(\sum_{k\in\mathbb{N}_0} p_k s^k - s)$, $s \in [0,1]$. Each particle moves like a Brownian motion with drift $-\mu$, for $\mu > 0$, and is killed when it hits 0 or K. The infinitesimal generator of such motion is given by

$$\mathcal{L} = \frac{1}{2}\frac{\mathrm{d}^2}{\mathrm{d}x^2} - \mu\frac{\mathrm{d}}{\mathrm{d}x}, \quad x \in (0,K),$$

which is defined for all $u \in C^2(0,K)$ (the space of twice continuously differentiable functions on $(0,K)$), with $u(0+) = u(K-) = 0$. Note that the boundary conditions on u correspond to the killing at 0 and K respectively. We denote the law of the process Z by $\mathbb{P}^K$. Furthermore for each $x \in (0,K)$ let $q_K(x)$ be defined as the probability of survival when the process is started from a single particle with initial location x.

Recall that the blue tree consists of all the prolific individuals, that is individuals with an infinite line of descent, while the red trees consist of all the remaining particles. Clearly a process started from a single red particle will become extinct almost surely, since by definition, this initial particle has to be part of a finite line of descent. Thus, describing the behaviour of a red tree also gives us the evolution of the original process Z conditioned on extinction. To understand the evolution of the process conditioned on survival, we first need to describe the behaviour of its two building blocks, that is the purely red and purely blue trees. Proposition 11 and Theorem 12 in [18] identify the branching generators and the single particle motions of the red and blue branching diffusions. In particular, the branching generator of the red tree is given by the following, space-dependent transformation of the branching generator G,

$$G^R(s,x) = \frac{1}{1-q_K(x)}\left(G(s(1-q_K(x))) - sG(1-q_K(x))\right), \quad s \in [0,1],\ x \in (0,K),$$

and the red particles move according to the infinitesimal generator

$$\mathcal{L}^R = \frac{1}{2}\frac{\mathrm{d}^2}{\mathrm{d}x^2} - \left(\mu + \frac{q_K'(x)}{1-q_K(x)}\right)\frac{\mathrm{d}}{\mathrm{d}x}, \quad x \in (0,K),$$

for $u \in C^2(0,K)$ with $u(0+) = u(K-) = 0$. On the other hand, the branching activity of the blue tree is governed by the space-dependent branching generator

$$G^B(s,x) = \frac{1}{q_K(x)}(G(sq_K(x) + (1-q_K(x))) - (1-s)G(1-q_K(x))), \qquad s \in [0,1],\ x \in (0,K),$$

and the single particle motion is the one associated to the infinitesimal generator

$$\mathcal{L}^B = \frac{1}{2}\frac{\mathrm{d}^2}{\mathrm{d}x^2} - \left(\mu - \frac{q_K'(x)}{q_K(x)}\right)\frac{\mathrm{d}}{\mathrm{d}x}, \quad x \in (0,K),$$

for $u \in C^2(0,K)$ with $u(0+) = u(K-) = 0$.

Just as in the case of GW-processes, the branching generators of the blue and red trees can be recovered from the original branching generators G. For each $x \in (0,K)$, we can split the graph of G in two at $1 - q_K(x)$, which is the probability of extinction of the process started from a single individual at position x. After some scaling the left hand side gives the branching generator of the red tree, while the right hand side corresponds to the branching generator of the blue tree (see Figure 1-4). It is not hard to see that G^R is indeed a subcritical branching generator, while G^B is supercritical, where the probability of having no offspring is zero.

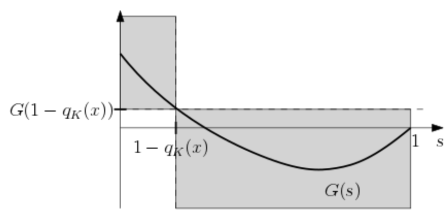

Figure 1-4: Graph of the branching generator G split at $1 - q_K(x)$; rescaling the upper-left quadrant and lower-right quadrant gives the branching generators G^R and G^B respectively.

To give the intuition behind how we get the space-dependent drift terms for the red and blue particles let us denote by $\xi = (\xi_t, t \geq 0)$ a Brownian motion with drift $-\mu$ started from $x \in (0,K)$, and killed upon exiting $(0,K)$. We denote by Π_x^K the law of ξ started from $x \in (0,K)$, and set $\mathcal{F}_t := \sigma(\xi_s, s \leq t)$. In Section 3 of [18] the authors show that conditioning the process on extinction induces a martingale change of measure on the level of single particle motion, having density

$$\frac{1-q_K(\xi_t)}{1-q_K(x)}\mathrm{e}^{\int_0^t \frac{G(1-q_K(\xi_s))}{1-q_K(\xi_s)}\mathrm{d}s}, \quad t \geq 0, \tag{1.1}$$

with respect to Π^K. Similarly conditioning on survival induces a change of measure where the density with respect to Π^K is given by

$$\frac{q_K(\xi_t)}{q_K(x)}\mathrm{e}^{-\int_0^t \frac{G(1-q_K(\xi_s))}{q_K(\xi_s)}\mathrm{d}s}\mathbf{1}_{\{T_{(0,K)}>t\}}, \tag{1.2}$$

where $T_{(0,K)}$ denotes the first time the initial particle exits $(0, K)$. To see how this affects the infinitesimal generator of the motion term we recall the following version of Revuz and Yor [28], VIII Proposition 3.4, which was adapted by Hesse et al. [18] for the case of branching Brownian motion in a strip.

Lemma 1.1.1. *Let $x \in (0, K)$. Let $h \in C^2(0, K)$ and suppose that*

$$\frac{h(\xi_t)}{h(x)} \exp\left\{-\int_0^t \frac{\mathcal{L}h(\xi_s)}{h(\xi_s)} \mathrm{d}s\right\}, \quad t \geq 0, \tag{1.3}$$

is a Π_x^K-martingale. Then under $\widehat{\Pi}_x^K$, which is the probability measure having Radon-Nikodym derivative (1.3) with respect to Π_x^K on $\mathcal{F}_t$, ξ has infinitesimal generator

$$\mathcal{L} + \frac{h'(y)}{h(y)} \mathrm{d}y,$$

for functions $u \in C^2(0, K)$ with $u(0+) = u(K-) = 0$.

Furthermore Remark 10 in [18] implies that $\mathcal{L}(1 - q_K) + G(1 - q_K) = 0$, and also that $\mathcal{L}q_K - G(1 - q_K) = 0$. This means that the martingale change of measure characterised by the density (1.1) is equivalent to an h-transform of the infinitesimal generator $\mathcal{L}$ with $h = 1 - q_K$. And similarly using (1.2) as a density is once again an h-transform of $\mathcal{L}$, this time with $h = q_K$.

We have seen that the law of the process Z conditioned on extinction is that of the red branching diffusion. Conditioning the initial particle to have colour blue is equivalent to conditioning the process on survival. In this case what we see is an infinite blue tree dressed with independent copies of the red branching diffusion. Conditional on the blue branching diffusion we have that red particles, initiating the finite red trees, immigrate continuously along the trajectories of the blue particles, where an immigration with $n \geq 1$ immigrants occurs at some space-dependent rate $\beta_n^1(x)$, $x \in (0, K)$. In addition to the continuous immigration we also have that when the blue tree branches at position $x \in (0, K)$, and produces $k \geq 2$ offspring, an additional n red particles immigrate at the same position with some probability $p_{n,k}(x)$. (For the exact expressions for $\beta_n^1(x)$ and $p_{n,k}(x)$ see Theorem 12 in [18].)

Finally, at any given time t, given the positions of the particles, the prolific ones can be found by making a Bernoulli selection where, for a particle position x, the Bernoulli trial has a probability of success $q_K(x)$.

1.1.3 Continuous-state branching processes

Continuous-state branching processes (CSBP) are the continuous time and space analogues of Galton-Watson processes. They were introduced by Jirina [19] and Lamperti [23], although continuous population models had already been used previously by Feller [17] to study large populations. He argued that continuous approximations give more accurate results than using large discrete systems with inevitable simplifications, and thus constructed a continuous process (now known as Feller's branching diffusion) as a high density limit of GW processes. For a general background on CSBPs see Chapter 12 of [20] or Chapter 3 of [26].

Since we will only consider finite mean branching processes, we introduce CSBPs in this setting. A (finite-mean) CSBP is a $[0,\infty)$-valued strong Markov process $X = (X_t, t \geq 0)$ with probabilities $(\mathbb{P}_x, x \geq 0)$, that has càdlàg paths, and satisfies the branching property, that is for any $x, y \geq 0$ the law of the process started from $x+y$ is that of the sum of two independent CSBPs, one issued from x, and the other issued from y.

Its semigroup is characterised by the Laplace functional

$$\mathbb{E}_x\left[\mathrm{e}^{-\theta X_t}\right] = \mathrm{e}^{-x u_t(\theta)}, \quad x, \theta, t \geq 0,$$

where $u_t(\theta)$ is the unique solution to the evolution equation

$$u_t(\theta) + \int_0^t \psi\left(u_s(\theta)\right) \mathrm{d}s = \theta, \quad t \geq 0.$$

We call ψ the branching mechanism of X, and it satisfies the Lévy-Khintchine formula

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

where $\alpha \in \mathbb{R}$, $\beta \geq 0$, and Π is a measure concentrated on $(0,\infty)$ such that $\int_{(0,\infty)} \left(x \wedge x^2\right) \Pi(\mathrm{d}x)$ is finite.

For convenience we assume that $-\psi$ is not the Laplace exponent of a subordinator, which rules out the case that X has monotone paths.

The branching mechanism plays the role of the branching generator of a discrete branching process, in a sense that it completely characterises the dynamics of the process. For instance, it is not too hard to see that

$$\mathbb{E}_x\left[X_t\right] = \mathrm{e}^{-\psi'(0+)t}, \quad x, t \geq 0,$$

that is the mean growth is fully determined by $\psi'(0+)$. Similarly to the discrete setting we call the process supercritical, critical or subcritical accordingly as $-\psi'(0+)$ is strictly positive, equal to zero or strictly negative. Here we assume that $\psi(\infty) = \infty$, which ensures that the event of extinguishing $\{\lim_{t\to\infty} X_t = 0\}$ has positive probability. When the process is critical or subcritical the probability of this event is one. Note, that the reason for the slightly different terminology (we use *extinguishing* and not *extinction*) is that, since a CSBP is $[0,\infty)$-valued, we can have that even though the process does go to zero as $t \to \infty$, X_t remains strictly positive for any finite t. This clearly cannot happen when the branching process is integer valued.

A supercritical CSBP has a positive probability of survival, which allows for the existence of infinite genealogies. Recall, that we called the surviving genealogies prolific. Bertoin et al. [3] were interested in the process formed by the prolific individuals, and showed that its characteristics can be recovered from the branching mechanism of the

CSBP, similarly to the Galton-Watson case. A decomposition of the CSBP along its embedded prolific skeleton was first given by Duquesne and Winkel [7], who used the framework of Lévy trees, and later by Berestycki et al. [2]. They studied the prolific backbone decomposition of superprocesses (which we will cover in the next section), and got the backbone decomposition of CSBPs as a special case of their results.

Just as in the case of branching Brownian motion we first give the characterisation of the prolific individuals and the immigration process, then we explain how to recover the law of the supercritical CSBP from these two processes. Note that ψ is strictly convex and under our conditions the equation $\psi(\lambda) = 0$ has exactly one root on $(0, \infty)$, which we denote by λ^*. We can think of λ^* as the rate of survival since for all $x \geq 0$,

$$\mathbb{P}_x \left(\lim_{t \to \infty} X_t = 0 \right) = \mathrm{e}^{-\lambda^* x}.$$

To get the equivalent of the red tree we saw in the previous sections, we need to characterise the law of the process conditioned on extinguishing, which we denote by $\mathbb{P}^*_x$. One can show (see e.g. [30]) that X under $\mathbb{P}^*_x$ is a subcritical CSBP with branching mechanism ψ^*, where

$$\psi^*(\theta) = \psi(\theta + \lambda^*), \quad \theta \geq 0.$$

The copies of this subcritical process is what we immigrate on the event of survival.

Next we describe the analogue of the previously seen blue tree. Even though we are in the continuous world, under our conditions, the number of those individuals who contribute on a long term is only discrete. More precisely, the prolific individuals form a supercritical continuous-time Galton-Watson process Z with branching generator

$$F(s) = \frac{1}{\lambda^*} \psi(\lambda^*(1 - s)), \quad s \in [0, 1]. \tag{1.4}$$

Using (1.4) it is easy to verify that the probability of having no offspring is zero, thus all the genealogies of Z are indeed infinite.

As both ψ^* and F are expressed in terms of the branching mechanism of X, it is not too surprising that we can recover these characteristics from ψ. Splitting the graph of ψ at λ^*, as in Figure 1-5, we can see that the right-hand side gives the branching mechanism of X conditioned to become eventually extinguished, while the left-hand side corresponds to the branching generator of the embedded prolific tree.

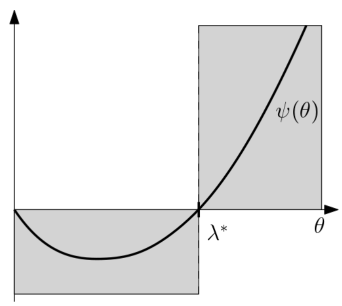

Figure 1-5: Graph of the branching mechanism ψ split at λ^*; rescaling the lower-left quadrant and upper-right quadrant gives the branching generator F and branching mechanism ψ^* respectively.

Recall, that with probability $e^{-\lambda^* x}$ the process becomes extinguished, in which case we see the dynamics of a CSBP with branching mechanism ψ^*. On the event of survival the picture is slightly more complicated. What we see in distribution is the sum of two independent processes. One is an independent copy of X under $\mathbb{P}_x^*$, while the other is the accumulated mass we get after running a supercritical Galton-Watson process, with branching generator F, and dressing it in a Poissonian way with independent copies of X conditioned to become extinguished. In particular we have three types of immigration which give the dressing of the tree. Along the trajectories of Z independent copies of X under $\mathbb{P}_x^*$ immigrate continuously and discontinuously. Furthermore every time Z branches, an additional independent copy of X conditioned to become extinguished is grafted onto the branchpoint of Z, where the initial mass of the immigrant depends on the number of offspring at the branching event. Figure 1-6 gives an intuitive picture of this decomposition, where the light grey mass represents the immigration along the trajectories (the accumulation of the mass coming from the continuous and discontinuous immigration), while the dark grey mass corresponds to the branch point immigration. Finally on the left had side we have the initial burst of subcritical mass. For the precise construction of the dressed tree, with the exact rates for each type of immigration, we refer the reader to Chapter 2, where we also give an alternative proof of the backbone decomposition of CSBPs.

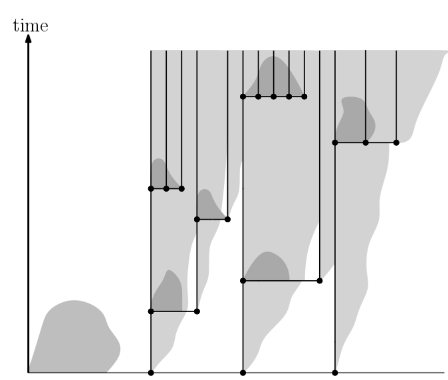

Figure 1-6: Backbone decomposition of a supercritical CSBP on the event of survival.

Even though the skeleton does not contribute towards the size of the population, at a given time $t \geq 0$, we can find the prolific individuals by running a Poisson point process with rate λ^* along the half line $[0, \infty)$ and noting the marks that fall into $[0, X_t)$. Thus given X_t, the number of prolific individuals at time t is Poisson with parameter $\lambda^* X_t$. This is intuitively not surprising considering a CSBP can always be constructed as a weak limit of rescaled GW processes (see [24] or [26]), thus what we see here is the analogue of the Bernoulli selection we had in the discrete setting.

1.1.4 Superprocesses

Just as branching Brownian motion was a natural spatial extension of a Galton-Watson process, superprocesses are spatial versions of continuous-state branching processes. Similarly to CSBPs, superprocesses first appeared as high density limits of discrete branching particle systems (see [31]), and as such were used as continuous approximations of large spatial populations. Recall that a branching Browian motion is a measure valued process, where each particle has unit mass. In the continuum world these atoms are replaced by a cloud of infinitesimal particles, the evolution of which results in a finite measure valued process (which is no longer atomic). To characterise the evolution of these infinitesimal particles we need two parameters, namely the motion and the branching mechanism. As in Chapter 3 we will consider a superprocess with spatially dependent branching mechanism, in this section we introduce superprocesses in this setting. For a general overview, and for futher details of what we present below, we refer the reader to the books of Dynkin [8, 9], Etheridge [12], Le Gall [25] and Li [26].

Let E be a domain in $\mathbb{R}^d$, and let us denote by $\mathcal{M}(E)$ the space of finite Borel measures on E. In this section we are interested in a strong Markov process X on E that takes values in $\mathcal{M}(E)$. The evolution of X can be characterised by two quantities $\mathcal{P}$ and ψ. Here, $(\mathcal{P})_{t\geq 0}$ is a conservative diffusion semigroup on E, and the so-called branching mechanism ψ takes the form

$$\psi(x, z) = -\alpha(x)z + \beta(x)z^2 + \int_{(0,\infty)} \left(\mathrm{e}^{-zu} - 1 + zu\right) m(x, \mathrm{d}u), \quad x \in E, z \geq 0, \quad (1.5)$$

where α and $\beta \geq 0$ are bounded measurable mappings from E to $\mathbb{R}$ and $[0, \infty)$ respectively, and for each $x \in E$, $m(x, \mathrm{d}u)$ is a measure concentrated on $(0, \infty)$, such that $x \mapsto (u \wedge u^2)m(x, \mathrm{d}u)$ is bounded and measurable. We denote by $\mathbb{P}_\mu$ the law of X issued from $\mu \in \mathcal{M}(E)$. Then, under suitable regularity conditions, the semigroup of X can be characterised as follows. For each $\mu \in \mathcal{M}(E)$ and non-negative, bounded measurable test function f we have

$$\mathbb{E}_\mu \left[\mathrm{e}^{-\langle f, X_t\rangle}\right] = \mathrm{e}^{-\langle u_f(\cdot,t),\mu\rangle}, \quad t \geq 0,$$

where $u_f(x, t)$ is the unique non-negative solution to the integral equation

$$u_f(x, t) = \mathcal{P}_t[f](x) - \int_0^t \mathrm{d}s \cdot \mathcal{P}_s[\psi(\cdot, u_f(\cdot, t - s))](x), \quad x \in E, t \geq 0,$$

and $\langle f, \mu \rangle$ denotes the integral of f with respect to the measure μ. That is, intuitively, the infinitesimal particles of the superprocess X are moving independently in E according to the diffusion semigroup $\mathcal{P}$ and they continuously reproduce, where the branching phenomenon is governed by ψ.

Just as in the previous examples, we can decompose a superprocess along its prolific individuals. Such a decomposition for superprocesses first appeared in Evans and O'Connell [13], who considered superprocesses with quadratic branching mechanism. They showed that the distribution of this process can be written as the sum of two independent processes. The first one is a copy of the original process conditioned on extinction. The second is produced by running a dyadic branching particle system, which has initial configuration given by a Poisson random measure, and along its trajectories immigrating continuously independent copies of the original process conditioned on extinction. Later this decomposition was extended to the spatially dependent case by Engländer and Pinsky [11]. Berestycki et al. [2] gave a pathwise prolific backbone decomposition of a superprocess with general, but spatially independent branching mechanism, and later Kyprianou et al. [22] and Eckhoff et al. [10] extended this pathwise decomposition to the spatially dependent case that we consider here.

In this section we use the results of [22] to describe the backbone decomposition of the previously introduced superprocess X. Let $\mathcal{E} = \{\langle 1, X_t \rangle = 0 \text{ for some } t > 0\}$ the event of extinction, and for each $x \in E$ define

$$w(x) = -\log \mathbb{P}_{\delta_x}(\mathcal{E}).$$

To get a non-trivial backbone decomposition we need the probability of extinction to be strictly positive, but less than one. Note that when the branching mechanism is spatially independent, the total mass process $\langle 1, X_t \rangle$ is a CSBP with branching mechanism ψ, and thus criticality can be expressed in terms of the largest root of ψ. In the spatially dependent case however w cannot be identified as the root of the branching mechanism. Thus, in order to proceed, we make the assumption that w is locally bounded away from 0 and ∞. Being locally bounded away from 0 ensures that the superprocess is supercritical in the sense that the extinction probability is not unity, while the second condition forces this probability to be greater than zero. Implicitly this also means that the extinguishing event, which we discussed in the previous section, coincides with the event of extinction.

Kyprianou et al. [22] showed that under these assumptions the population does have infinite genealogies, the evolution of which can be described by a branching particle system, which we denote by Z. Recall that to characterise such a process we need to specify its branching generator and its motion, and just as in the previous examples, we can construct these parameters from the components on the superprocess X.

Due to the spatial dependence of the branching mechanism, distinguishing the infinite genealogies alters the spatial motion of the corresponding particles. This bias is similar to what we saw in the branching Brownian motion case, where the constraint of staying in a strip was responsible for the spatial dependence. Hence, what we have here is again an h-transform of the motion generator, this time with $h = w$. We denote the associated semigroup by $\mathcal{P}^w$. In the branching Brownian motion case, the h-transform introduced a spatially dependent drift that kept the particles away from the boundary, here the dynamics of $\mathcal{P}^w$ encourages the motion to visit domains where the global survival rate is high.

The branching generator of Z is given by

$$G(x,s) = q(x) \sum_{k \geq 0} p_k(x) \left(s^k - s \right), \quad s \in [0,1],$$

where q and p_k, $k \geq 0$ can be built from the branching mechanism of X. In particular,

$$q(x) = \psi'(x, w(x)) - \frac{\psi(x, w(x))}{w(x)}, \quad x \in E,$$

and $p_0(x) = p_1(x) = 0$, and for $k \geq 2$,

$$p_k(x) = \frac{1}{w(x)q(x)} \left\{ \beta(x) w^2(x) \mathbf{1}_{\{k=2\}} + w^k(x) \int_{(0,\infty)} \frac{u^k}{k!} \mathrm{e}^{-w(x)u} m(x, \mathrm{d}u) \right\}, \quad x \in E,$$

thus, once again, the possibility of zero offspring is eliminated. Hence the particles of Z are encouraged to visit domains with high global survival rate, and when they branch, at least two offspring particles are produced who then follow the behaviour of their parent. This is the process that we dress up with immigration in order to recover the law of the original superprocess. What we immigrate is a copy of X conditioned on extinction. In [22] the authors show that this process is again a superprocess with motion semigroup $\mathcal{P}$, and branching mechanism ψ^*, where

$$\psi^*(x,z) = \psi(x, z + w(x)) - \psi(x, w(x)), \quad x \in E, z \geq 0.$$

It is not hard to see that the branching generator can be rewritten as

$$G(x,s) = \frac{1}{w(x)} \left(\psi(x, w(x)(1-s)) - (1-s)\psi(x, w(x)) \right), \quad x \in E, s \in [0,1],$$

after which, for a fixed $x \in E$, we can easily identify the two components from the graph of $\psi(x,z)$. In particular, splitting Figure 1-7 at $w(x)$, the lower part corresponds to the branching generator of the prolific branching particle system, while the upper part translates to the branching mechanism of the immigration.

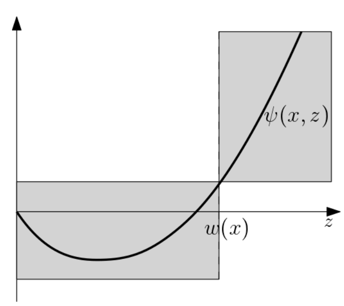

Figure 1-7: Graph of the branching mechanism $z \mapsto \psi(x,z)$ split at $w(x)$; rescaling the lower-left quadrant and upper-right quadrant gives the branching generator $z \mapsto G(x,z)$ and branching mechanism $z \mapsto \psi^*(x,z)$ respectively.

Immigration occurs similarly to the CSBP case, that is backbone particles throw mass off continuously and discontinuously along their spatial trajectory, and every time they branch, additional mass is immigrated at the branch point, where the initial mass depends on the number of offspring at the branching event.

Then X under $\mathbb{P}_\mu$ has the same law as the sum of two independent processes. One is an independent copy of X conditioned on extinction and having initial configuration μ. The other one is the accumulation of mass produced by the dressed branching particle system, which has initial configuration consisting of a Poisson random field of particles in E with intensity $w(x)\mu(\mathrm{d}x)$. This Poissonian structure of the backbone is preserved in time, in the sense that at any time t, the backbone particles are given by a Poisson random measure having intensity $w(x)X_t(\mathrm{d}x)$.

Above, we focused on those backbone decompositions that are relevant to the results presented in this thesis, but by no means was this an exhaustive overview of the existing literature. For example, to mention just a few, backbones have been identified for CSBPs with immigration [21], superprocesses with non-local branching mechanism [27], and for super Brownian motion whose exit measure is conditioned to hit a number of specified point on the boundary of the domain [29].

The aim of this thesis is to deepen the understanding of skeletal decompositions even further, by developing a new approach for their study, as well as widening the class of processes with available backbone decomposition. In the next section we give a brief overview of our results.

1.2 Summary of chapters

Each chapter in this thesis contains a research article. The first two articles ([14] and [15]) resulted from collaboration with Andreas Kyprianou and Joaquín Fontbona, while the article in Chapter 4 [16] is a result of collaboration with Sandra Palau, José-Luis Pérez and Juan-Carlos Pardo.

1.2.1 Skeletal stochastic differential equations for continuous-state branching processes (Chapter 2).

It is well known that a finite-mean CSBP can be represented as the unique strong solution to

$$X_t = x + \alpha \int_0^t X_{s-}\mathrm{d}s + \sqrt{2\beta}\int_0^t \int_0^{X_{s-}} W(\mathrm{d}s,\mathrm{d}u) + \int_0^t \int_0^\infty \int_0^{X_{s-}} r\tilde{N}(\mathrm{d}s,\mathrm{d}r,\mathrm{d}\nu), \tag{1.6}$$

for $x > 0, t \geq 0$, where $W(\mathrm{d}s,\mathrm{d}u)$ is a white noise process on $(0,\infty)^2$ based on the Lebesgue measure $\mathrm{d}s \otimes \mathrm{d}u$ and $N(\mathrm{d}s,\mathrm{d}r,\mathrm{d}\nu)$ is a Poisson point process on $(0,\infty)^3$ with intensity $\mathrm{d}s \otimes \Pi(\mathrm{d}r) \otimes \mathrm{d}\nu$. Moreover, $\tilde{N}(\mathrm{d}s,\mathrm{d}r,\mathrm{d}\nu)$ is the compensated measure of $N(\mathrm{d}s,\mathrm{d}r,\mathrm{d}\nu)$. We refer the reader to [4] for the properties of this SDE.

In this article we use the language of stochastic differential equations (SDEs) to study different backbone decompositions of a CSBP. In particular, we write down a coupled SDE which simultaneously describes the evolution of the backbone and the total accumulated mass, and show that the total mass process is a weak solution to (1.6).

Our first result concerns the decomposition of supercritical CSBPs. The coupled SDE in this case provides a common framework for the parametric family of decompositions constructed by Duquesne and Winkel [7]. They showed that an embedded skeleton can be given for each parameter $\lambda \geq \lambda^*$, where λ^* is the largest root of the branching mechanism. The choice of λ^* corresponds to the prolific backbone, while choosing a parameter greater than λ^* selects additional genealogical lines to the prolific ones when constructing the skeleton. Since a prolific backbone contains all infinite genealogical lines, the additional ones are necessarily only finite. In other words, the skeleton has the possibility of 'dead ends' (no offspring).

Next, we concentrate on another family of decompositions. We have seen that (sub)critical CSBPs go extinct almost surely, thus the population has no prolific individuals in the classic sense. However if we fix some finite time $T > 0$, the population has a positive probability of survival until that fixed time. Exploiting this fact we can define a time-inhomogeneous skeleton consisting of all those individuals who have a descendant at time T. We call these individuals T-prolific. Thus, up to time T we can describe a decomposition of the CSBP, where the T-prolific skeleton is given by a GW process with time dependent branching rate and offspring distribution, and the immigration process is a CSBP conditioned to go extinct by time T. Note, that as a result of this conditioning, the branching mechanism of the immigrants is also time-dependent. This finite time horizon decomposition was heavily motivated by the work of Duquesne and Le Gall [6], who gave a description of the (sub)critical CSBP genealogy using the so-called height process. The SDE approach does not rely on the existence of the height-process, and as such it extends the finite horizon setting for supercritical CSBPs.

It is clear that conditioning a (sub)critical CSBP to survive until time T is equivalent to conditioning on having at least one T-prolific individual in the initial population. And due to the Poissonian nature of the backbone, the latter is the same as conditioning a Poisson random variable to be at least one. Our last objective is to study what happens to the coupled SDE when we make this conditioning and take the time horizon to infinity. What we can see is that as $T \to \infty$, the branching rate of the backbone slows down, less and less offspring are produced, and eventually we end up with a single line of descent. In the literature this one infinite genealogy is called the spine, and it emerges when a (sub)cirtical CSBP is conditioned to survive forever. Using our SDE representation, we show how the T-prolific skeleton thins down to become the spine, when we condition the CSBP to survive until larger and larger times.

To appear in Journal of Applied Probability 56.4 (December 2019).

1.2.2 Skeletal stochastic differential equations for superprocesses (Chapter 3).

In this article we extend the previously seen coupled SDE approach to superprocesses, demonstrating the robustness of the method. We study supercritical superprocesses with space-dependent branching mechanism, as seen in [22] and [10]. More precisely, let E be a domain in $\mathbb{R}^d$, and consider a superprocess $(X, \mathbb{P}_\mu)$ with branching mechanism given by (1.5), and motion given by the Feller diffusion semigroup $(\mathcal{P}_t)_{t\geq 0}$. The infinitesimal generator of $\mathcal{P}$ can be written in the form

$$\mathcal{L} = \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{d} b_i(x)\frac{\partial}{\partial x_i},$$

for some suitable functions a and b. Let us denote by $\mathcal{M}(E)$ the space of finite Borel measures on E, and let $\mathcal{M}(E)^\circ := \mathcal{M}(E)\setminus\{0\}$, where 0 is the null measure. Then, if f is in the domain of $\mathcal{L}$, the superprocess X has the representation

$$\begin{aligned}\langle f, X_t\rangle = \langle f, \mu\rangle + \int_0^t \langle \alpha f, X_s\rangle \mathrm{d}s + M_t^c(f) + \int_0^t \int_{\mathcal{M}(E)^\circ} \langle f, \nu\rangle \tilde{N}(\mathrm{d}s, \mathrm{d}\nu) \\ + \int_0^t \langle \mathcal{L} f, X_s\rangle \mathrm{d}s, \quad t \geq 0.\end{aligned} \tag{1.7}$$

Here $M_t^c(f)$ is a continuous local martingale with quadratic variation $2X_t(bf^2)\mathrm{d}t$, and $\tilde{N}$ is an optional random measure on $[0,\infty)\times\mathcal{M}(E)^\circ$ with predictable compensator $\hat{N}(\mathrm{d}s,\mathrm{d}\nu) = \mathrm{d}sK(X_{s-},\mathrm{d}\nu)$, where

$$\int_{\mathcal{M}(E)^\circ} \langle f, \nu\rangle K(\mu, \mathrm{d}\nu) = \int_E \mu(\mathrm{d}x)\int_{(0,\infty)} \langle f, u\delta_x\rangle m(x, \mathrm{d}u).$$

For more details see Chapter 7 in [26].

This representation is the basis of our coupled SDE, which simultaneously describes the branching particle system formed by the prolific particles, and the evolution of the total mass that is created by the immigration processes and the initial burst of mass. We once again show that this total mass has the same law as that of the superprocess. Even though we had to develop a different method for the proof, our results show that the coupled SDE approach is not unique to the non-spatial case. The article only deals with the supercritical superprocesses, leaving the subcritical case open for the time being.

Preprint: arXiv:1904.05966

1.2.3 Backbone decomposition of multitype superprocesses (Chapter 4).

In the final chapter we consider the backbone decomposition of multitype superprocesses with a finite number of types. Technically a multitype superprocess can be defined as a one-type superprocess on a state space that is the mixture of a continuous and a discrete space, and whose branching has both local and non-local elements. Informally each individual has an assigned type which specifies its branching mechanism and motion semigroup. The individual keeps its type throughout its life, but when it branches it has the possibility to give birth to children of different types (hence the non-local element of the branching mechanism). While backbone decompositions are fairly well-known for one-type CSBPs and superprocesses that are supported on a domain of $\mathbb{R}^d$, before this article no such decompositions or even description of prolific genealogies had been given for the previously described multitype case. We should mention however that other decompositions of multitype superprocesses have been constructed, see e.g. [5], where the authors give the Williams' decomposition of a multitype superprocess, where the genealogy is decomposed with respect to the last individual alive.

We adapt the methods developed in [2] to construct the prolific backbone decomposition of a supercritical multitype superprocess. The prolific skeleton is a multitype branching diffusion process, with type dependent branching rate, offspring generator and motion semigroup. Comparing the form of the offspring distribution between the one-type case and the multitype case, the main difference is that now we are allowed to have one offspring at a branching event. However in this case, that offspring has to have a different type from its parent. Just as in the one-type case this process is dressed up with immigration, where the immigration process is the original process conditioned on extinction. Once again we have continuous and discontinuous immigration along the trajectories of the skeleton particles, and branch point based immigration, all of which depend on the type of the backbone particle.

By showing that the total mass vector of a multitype superprocess is a multitype CSBP, it is clear that after turning the movement off, we get the prolific backbone decomposition for multitype CSBPs as a special case of our results.

Preprint: arXiv:1803.09620

This thesis is presented in an alternative thesis format which includes publications. These were developed independently of the introduction and are supposed to be self-contained. Hence, it is inevitable to have some inconsistency in notations and redundant contents to the introduction chapter.

Bibliography

[1] K. B. Athreya and P. E. Ney. *Branching Processes.* Springer-Verlag Berlin Heidelberg, 1972.

[2] J. Berestycki, A.E. Kyprianou, and A. Murillo-Salas. The prolific backbone for supercritical superprocesses. *Stochastic Processes and their Applications*, 121(6):1315–1331, jun 2011.

[3] Jean Bertoin, Joaquin Fontbona, and Servet Martnez. On prolific individuals in a supercritical continuous-state branching process. *Journal of Applied Probability*, 45(3):714–726, 2008.

[4] Donald A. Dawson and Zenghu Li. Stochastic equations, flows and measure-valued processes. *Ann. Probab.*, 40(2):813–857, 03 2012.

[5] Jean-Franois Delmas and Olivier Hnard. A williams decomposition for spatially dependent superprocesses. *Electron. J. Probab.*, 18:43 pp., 2013.

[6] T. Duquesne and J.-F. Le Gall. Random Trees, Lévy Processes and Spatial Branching Processes. *ArXiv Mathematics e-prints*, September 2005.

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

[8] E. B. Dynkin. *An introduction to branching measure-valued processes*, volume 6 of *CRM Monograph Series.* American Mathematical Society, Providence, RI, 1994.

[9] E. B. Dynkin. *Diffusions, superprocesses and partial differential equations*, volume 50. American Mathematical Society, Colloquium Publications, Providence, RI, 2002.

[10] Maren Eckhoff, Andreas E. Kyprianou, and Matthias Winkel. Spines, skeletons and the strong law of large numbers for superdiffusions. *Annals of Probability*, 43(5):2545 – 2610, 9 2015.

[11] János Engländer and Ross G. Pinsky. On the construction and support properties of measure-valued diffusions on $d \subseteq \mathbb{R}^d$ with spatially dependent branching. *Ann. Probab.*, 27(2):684–730, 04 1999.

[12] Alison M. Etheridge. *An introduction to superprocesses*, volume 20 of *University Lecture Series.* American Mathematical Society, Providence, RI, 2000.

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

[14] D. Fekete, J. Fontbona, and A. E. Kyprianou. Skeletal stochastic differential

equations for continuous-state branching process. *To appear in Journal of Applied Probability*, 56(4), December 2019.

[15] Dorottya Fekete, Joaquin Fontbona, and Andreas E. Kyprianou. Skeletal stochastic differential equations for superprocesses. *arXiv e-prints*, page arXiv:1904.05966, Apr 2019.

[16] Dorottya Fekete, Sandra Palau, Juan Carlos Pardo, and José Luis Pérez. Backbone decomposition of multitype superprocesses. *arXiv e-prints*, page arXiv:1803.09620, Mar 2018.

[17] William Feller. Diffusion processes in genetics. In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, pages 227–246, Berkeley, Calif., 1951. University of California Press.

[18] Simon Harris, Marion Hesse, and Andreas Kyprianou. Branching brownian motion in a strip: survival near criticality. *Annals of Probability*, 44(1):235–275, 1 2016.

[19] Miloslav Ji?ina. Stochastic branching processes with continuous state space. *Czechoslovak Mathematical Journal*, 08(2):292–313, 1958.

[20] A. E. Kyprianou. *Fluctuations of Lévy processes with applications*. Universitext. Springer, Heidelberg, second edition, 2014. Introductory lectures.

[21] A.E. Kyprianou and Y.-X. Ren. Backbone decomposition for continuous-state branching processes with immigration. *Statistics & Probability Letters*, 82(1):139–144, 2012.

[22] Andreas Kyprianou, J.-L. Pérez, and Y X Ren. *The backbone decomposition for spatially dependent supercritical superprocesses*, pages 33 – 59. Lecture Notes in Mathematics. Springer, 10 2014.

[23] John Lamperti. Continuous state branching processes. *Bull. Amer. Math. Soc.*, 73(3):382–386, 05 1967.

[24] John Lamperti. The limit of a sequence of branching processes. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 7(4):271–288, Aug 1967.

[25] Jean-Francois Le Gall. *Spatial Branching Processes, Random Snakes and Partial Differential Equations*. Lectures in Mathematics. ETH Zrich. Birkhäuser Basel, 1999.

[26] Zenghu Li. *Measure-valued branching Markov processes*. Probability and its Applications (New York). Springer, Heidelberg, 2011.

[27] Antonio Murillo-Salas and José Luis Pérez. The backbone decomposition for superprocesses with non-local branching. In *XI Symposium on Probability and Stochastic Processes*, pages 199–216, Cham, 2015. Springer International Publishing.

[28] D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion*. Grundlehren der mathematischen Wissenschaften. Springer Berlin Heidelberg, 2004.

[29] Thomas S. Salisbury and John Verzani. On the conditioned exit measures of super brownian motion. *Probability Theory and Related Fields*, 115(2):237–285, Sep 1999.

[30] Y.-C. Sheu. Lifetime and compactness of range for super-Brownian motion with a general branching mechanism. *Stochastic Process. Appl.*, 70(1):129–141, 1997.

[31] Shinzo Watanabe. A limit theorem of branching processes and continuous state branching processes. *J. Math. Kyoto Univ.*, 8(1):141–167, 1968.

[32] H. W. Watson and F. Galton. On the probability of the extinction of families. *The Journal of the Anthropological Institute of Great Britain and Ireland*, 4:138–144, 1875.

Appendix 6B: Statement of Authorship

This declaration concerns the article entitled:	
Skeletal stochastic differential equations for continuous-state branching process	
Publication status (tick one)	
Draft manuscript [ ] **Submitted** [ ] **In review** [ ] **Accepted** [X] **Published** [ ]	
Publication details (reference)	To appear in *Journal of Applied Probability* 56.4 (December 2019).
Copyright status (tick the appropriate statement)	
I hold the copyright for this material [ ]	Copyright is retained by the publisher, but I have been given permission to replicate the material here [X]
Candidate's contribution to the paper (provide details, and also indicate as a percentage)	*All research has been conducted in equal partnership with collaborators and supervisors. It is unwise to try and measure percentages of intellectual contribution as research scholarship lies as much in the escalation of ideas through mathematical discourse as it does with the seed of ideas themselves. That said, the mathematical content of this thesis is, as an entire piece of work, inextricably associated to its author through intellectual ownership. Formulation of ideas: The candidate played an integral and fully collaborative role in the formulation of ideas. Design of methodology: The candidate played an integral and fully collaborative role in the design of methodology. Experimental work: N/A Presentation of data in journal format: N/A *The wording in this box follows the advice and approval of my supervisor, Professor Kyprianou.
Statement from Candidate	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature.
Signed	
Date	16/07/2019

Last update: Feb 2019

Chapter 2

Skeletal stochastic differential equations for continuous-state branching process

Dorottya Fekete [1] , Joaquin Fontbona [2], Andreas E. Kyprianou [1]

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) SDEs (cf. [7, 8, 6]). 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; cf. [26, 14, 40]. The second is to provide a pathwise decomposition of CSBPs in the spirit of genealogical coding of CSBPs via Lévy excursions in Duquesne and LeGall [10] albeit precisely where the aforesaid coding fails to work because the underlying CSBP is supercritical.

[1]Department of Mathematical Sciences, University of Bath, Claverton Down, Bath, BA2 7AY, United Kingdom. Email: d.fekete@bath.ac.uk, a.kyprianou@bath.ac.uk

[2]Centre for Mathematical Modelling, DIM CMM, UMI 2807 UChile-CNRS, Universidad de Chile, Santiago, Chile.. Email: fontbona@dim.uchile.cl

2.1 Introduction

In this article we are interested in $X = (X_t, t \geq 0)$ a continuous-state, finite-mean 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

$$\mathbb{E}_x(\mathrm{e}^{-\theta X_t}) = \mathrm{e}^{-x u_t(\theta)}, \qquad x, \theta, t \geq 0, \tag{2.1}$$

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

$$u_t(\theta) + \int_0^t \psi(u_s(\theta))\mathrm{d}s = \theta, \qquad t \geq 0. \tag{2.2}$$

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

$$\psi(\theta) = -\alpha\theta + \beta\theta^2 + \int_{(0,\infty)} (\mathrm{e}^{-\theta x} - 1 + \theta x)\Pi(\mathrm{d}x),\ \theta \geq 0, \tag{2.3}$$

where $\alpha \in \mathbb{R}$, $\beta \geq 0$ and Π is a measure concentrated on $(0,\infty)$ which satisfies $\int_{(0,\infty)}(x \wedge x^2)\Pi(\mathrm{d}x) < \infty$. 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

$$\mathbb{E}_x[X_t] = x\mathrm{e}^{-\psi'(0+)t}, \qquad t, x \geq 0. \tag{2.4}$$

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)

$$X_t = x + \alpha\int_0^t X_{s-}\mathrm{d}s + \sqrt{2\beta}\int_0^t\int_0^{X_{s-}} W(\mathrm{d}s,\mathrm{d}u) + \int_0^t\int_0^\infty\int_0^{X_{s-}} r\tilde{N}(\mathrm{d}s,\mathrm{d}r,\mathrm{d}\nu), \tag{2.5}$$

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

Through the representation of a CSBP as either a strong Markov process whose semigroup 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.

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

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

where λ^* is the unique root on $(0,\infty)$ of the equation $\psi(\theta) = 0$. Note that under our conditions ψ is strictly convex with the property that $\psi(0) = 0$ and $\psi(+\infty) = \infty$, thereby ensuring that the root $\lambda^* > 0$ exists; see Chapter 8 and 9 of [25] 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

$$\psi^*(\theta) = \psi(\theta + \lambda^*). \tag{2.6}$$

See for example [46].

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 [22] distinguishes between these two cases according to an integral test. Indeed, the intersection is empty if and only if

$$\lim_{t\to\infty} \int_t^\infty \frac{1}{\psi(\theta)} \mathrm{d}\theta < \infty. \tag{2.7}$$

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 $A \in \mathcal{F}^X_t := \sigma(X_s : s \leq t)$ and $x > 0$,

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

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

$$\left.\frac{\mathrm{d}\mathbb{P}^\uparrow_x}{\mathrm{d}\mathbb{P}_x}\right|_{\mathcal{F}^X_t} = \mathrm{e}^{-\alpha t} \frac{X_t}{x} \mathbf{1}_{\{t<\zeta\}}. \tag{2.8}$$

In addition, $(X, \mathbb{P}^{\uparrow}_x)$, $x > 0$, has been shown to be equivalent in law to a process which has 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 $\mathrm{d}s \otimes r\Pi(\mathrm{d}r) \otimes \mathbb{P}_r(\mathrm{d}\omega)$. Moreover, $\mathbb{Q}$ is the intensity, or 'excursion' measure on the space $\mathbb{D}([0,\infty),\mathbb{R})$ which satisfies

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

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 [37] and [15, 31, 13, 9, 39] for further details. We should note in particular that Theorem 3.10 in [37] gives the necessary and sufficient conditions 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 $x > 0$, $(X, \mathbb{P}^{\uparrow}_x)$ is equal in law to the stochastic process

$$\begin{aligned}\Lambda_t = X'_t + \int_0^t \int_{\mathbb{D}([0,\infty),\mathbb{R})} \omega_{t-s} N^{\mathrm{c}}(\mathrm{d}s, \mathrm{d}\omega) \\ + \int_0^t \int_0^\infty \int_{\mathbb{D}([0,\infty),\mathbb{R})} \omega_{t-s} N^*(\mathrm{d}s, \mathrm{d}r, \mathrm{d}\omega), \qquad t \geq 0,\end{aligned} \tag{2.9}$$

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

Intuitively, one 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 (2.9), 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 [36], (4.3) in [35], (4.18) in [34] or the discussion in Section 12.3.2 of [25] or [37]. The reader is also referred to e.g. [43] or [29, 30] for further details of the notion of a spine.

It turns out that one may also identify the effect of the change of measure within the context of the SDE setting. In [20], it was shown that $(X, \mathbb{P}^{\uparrow}_x)$, $x > 0$, offers the unique strong solution to the SDE

$$\begin{aligned}X_t = x + \alpha\int_0^t X_{s-}\mathrm{d}s + \sqrt{2\beta}\int_0^t \int_0^{X_{s-}} W(\mathrm{d}s, \mathrm{d}u) + \int_0^t \int_0^\infty \int_0^{X_{s-}} r\tilde{N}(\mathrm{d}s, \mathrm{d}r, \mathrm{d}u) \\ + \int_0^t \int_0^\infty r N^*(\mathrm{d}s, \mathrm{d}r) + 2\beta t, \qquad t \geq 0,\end{aligned} \tag{2.10}$$

where W, N and $\tilde{N}$ are as in (2.5) and N^* is as above, and all noises are independent. See also [8] and [21].

Skeletal path decomposition of supercritical CSBPs. In [12, 5] and [4] it was shown that the law of the process X, where X is defined by (2.5), can be recovered from a supercritical continuous-time Galton–Watson process (GW), 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 Galton–Watson process with branching generator

$$q\left(\sum_{k\geq 0} p_k r^k - r\right) = \frac{1}{\lambda^*}\psi(\lambda^*(1-r)), \qquad r \in [0,1]. \tag{2.11}$$

Note that in the above generator, we have that $q = \psi'(\lambda^*)$ is the rate at which individuals reproduce and $\{p_k : k \geq 0\}$ is the offspring distribution. This continuous-time Galton–Watson 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 (2.11), $p_0 = p_1 = 0$, 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

$$2\beta\mathrm{d}\mathbb{Q}^* + \int_0^\infty y\mathrm{e}^{-\lambda^* y}\Pi(\mathrm{d}y)\mathrm{d}\mathbb{P}^*_y. \tag{2.12}$$

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 [19, 45, 44, 16, 12, 4, 23, 28, 27].

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 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; cf. [26, 14, 40, 41]. 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 continuous state branching processes with competition [3, 41]. Moreover, although our method takes inspiration from the genealogical coding of CSBPs by Lévy excursions, cf. Duquesne and LeGall [10], our approach appears to be applicable where the aforesaid method fails, namely supercritical processes.

2.2 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 [12], 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, both for 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 (2.10). 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 but, nonetheless, subtly different to 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

$$\psi_\lambda(\theta) = \psi(\theta+\lambda) - \psi(\lambda) = \psi'(\lambda)\theta + \beta\theta^2 + \int_{(0,\infty)} \left(\mathrm{e}^{-\theta x} - 1 + \theta x\right) \mathrm{e}^{-\lambda x}\Pi(\mathrm{d}x), \quad (2.13)$$

where

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

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 Galton Watson process parameterised by $\lambda \geq \lambda^*$, which has been seen before in e.g. [12] and agrees with the process described by (2.11) when $\lambda = \lambda^*$. It branches at rate $\psi'(\lambda)$ and has branching generator given by

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

That is to say, writing $F_\lambda(s)$ as in the left-hand side of (2.11), we now have $p_0 = \psi(\lambda)/\lambda\psi'(\lambda)$, $p_1 = 0$ and for $k \geq 2$,

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

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 $\eta_1(\mathrm{d}r) = 0$, $r \geq 0$, and, otherwise,

$$\eta_k(\mathrm{d}r) = \frac{1}{p_k\lambda\psi'(\lambda)}\left\{\psi(\lambda)\mathbf{1}_{\{k=0\}}\delta_0(\mathrm{d}r) + \beta\lambda^2\mathbf{1}_{\{k=2\}}\delta_0(\mathrm{d}r) + \mathbf{1}_{\{k\geq 2\}}\frac{(\lambda r)^k}{k!}\mathrm{e}^{-\lambda r}\Pi(\mathrm{d}r)\right\}, \tag{2.14}$$

for $r \geq 0$. Note in particular that, when $\lambda > \lambda^*$, there is the possibility that no offspring at a branching event. Since in this case some lines of descent are finite, the Galton-Watson 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 2.2.1 below. Let $\mathtt{N}^0$ be a Poisson random measure on $[0,\infty)^3$ with intensity measure $\mathrm{d}s \otimes \mathrm{e}^{-\lambda r}\Pi(\mathrm{d}r) \otimes \mathrm{d}\nu$, $\tilde{\mathtt{N}}^0$ be the associated compensated version of $\mathtt{N}^0$, $\mathtt{N}^1(\mathrm{d}s, \mathrm{d}r, \mathrm{d}j)$ be a Poisson point process on $[0,\infty)^2 \times \mathbb{N}$ with intensity $\mathrm{d}s \otimes r\mathrm{e}^{-\lambda r}\Pi(\mathrm{d}r) \otimes \sharp(\mathrm{d}j)$, and finally let $\mathtt{N}^2(\mathrm{d}s, \mathrm{d}r, \mathrm{d}k, \mathrm{d}j)$ 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 2.2.1. *Suppose that ψ corresponds to a supercritical branching mechanism (i.e. $\alpha > 0$) and $\lambda \geq \lambda^*$. Consider the coupled system of SDEs*

$$\begin{aligned}
\begin{pmatrix} \Lambda_t \\ Z_t \end{pmatrix} &= \begin{pmatrix} \Lambda_0 \\ Z_0 \end{pmatrix} - \psi'(\lambda)\int_0^t \begin{pmatrix} \Lambda_{s-} \\ 0 \end{pmatrix} \mathrm{d}s + \sqrt{2\beta}\int_0^t\int_0^{\Lambda_{s-}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} W(\mathrm{d}s, \mathrm{d}u) \\
&\quad + \int_0^t\int_0^\infty\int_0^{\Lambda_{s-}} \begin{pmatrix} r \\ 0 \end{pmatrix} \tilde{\mathtt{N}}^0(\mathrm{d}s, \mathrm{d}r, \mathrm{d}\nu) \\
&\quad + \int_0^t\int_0^\infty\int_1^{Z_{s-}} \begin{pmatrix} r \\ 0 \end{pmatrix} \mathtt{N}^1(\mathrm{d}s, \mathrm{d}r, \mathrm{d}j) \\
&\quad + \int_0^t\int_0^\infty\int_0^\infty\int_1^{Z_{s-}} \begin{pmatrix} r \\ k-1 \end{pmatrix} \mathtt{N}^2(\mathrm{d}s, \mathrm{d}r, \mathrm{d}k, \mathrm{d}j) \\
&\quad + 2\beta\int_0^t \begin{pmatrix} Z_{s-} \\ 0 \end{pmatrix} \mathrm{d}s, \qquad t \geq 0.
\end{aligned} \tag{2.15}$$

The equation (2.15) 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 (2.5);*

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

If one focuses on the second element, Z, in the SDE (2.15), 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 (2.14) 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 (2.15) describes precisely the prolific skeleton decomposition. In the spirit of [12], albeit using different technology and in a continuum setting, Theorem 2.2.1 puts into a common framework a parametric family of skeletal decompositions for supercritical processes. Related work also appears in [1, 38].

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

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

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

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

to ensure conservative supercriticality. Then the necessary adjustment one needs to make occurs, for example, in (2.5), where jumps of size greater than equal to 1 in the Poisson random measures N is separated out without compensation. However, the form of (2.15) 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 Duquesne and Le Gall [11]. Indeed, for (sub)critical CSBPs one may consider the conclusion of Theorem 2.2.2, 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 2.2.2 is also a time-inhomogeneous version of Theorem 2.2.1 for supercritical CSBPs, which setting was not discussed in [11].

Assume that ψ is a branching mechanism that satisfies Grey's condition (2.7). We fix a time marker $T > 0$ and we want to describe a coupled system of SDEs in the spirit of (2.15) 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 descendent at time T. Once again we need to introduce some more notation. We need a Poisson random measure $\mathtt{N}^0_T$ on $[0,T) \times [0,\infty)^2$ with intensity $\mathrm{d}s \otimes \mathrm{e}^{-u_{T-s}(\infty)r}\Pi(\mathrm{d}r) \otimes \mathrm{d}\nu$, a Poisson process $\mathtt{N}^1_T$ on $[0,T) \times [0,\infty) \times \mathbb{N}_0$ with intensity $\mathrm{d}s \otimes r\mathrm{e}^{-u_{T-s}(\infty)r}\Pi(\mathrm{d}r) \otimes \sharp(\mathrm{d}j)$, and a Poisson process $\mathtt{N}^2_T(\mathrm{d}s, \mathrm{d}r, \mathrm{d}k, \mathrm{d}j)$ on $[0,T) \times [0,\infty) \times \mathbb{N}_0 \times \mathbb{N}$ with intensity

$$\left\{ \frac{u_{T-s}(\infty)\psi'(u_{T-s}(\infty)) - \psi(u_{T-s}(\infty))}{u_{T-s}(\infty)} \right\} \mathrm{d}s \otimes \eta_k^{T-s}(\mathrm{d}r) \otimes p_k^{T-s}\sharp(\mathrm{d}k) \otimes \sharp(\mathrm{d}j),$$

where, for $k \geq 2$,

$$\eta_k^{T-s}(\mathrm{d}r) = \frac{\beta u^2_{T-s}(\infty)\mathbf{1}_{\{k=2\}}\delta_0(\mathrm{d}r) + (u_{T-s}(\infty)r)^k\, \mathrm{e}^{-u_{T-s}(\infty)r}\Pi(\mathrm{d}r)/k!}{p_k^{T-s}\left(u_{T-s}(\infty)\psi'(u_{T-s}(\infty)) - \psi(u_{T-s}(\infty))\right)}, \qquad r \geq 0, \tag{2.16}$$

and 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 2.2.2. *Suppose that ψ corresponds to a branching mechanism which satisfies Grey's condition (2.7). Fix a time horizon $T > 0$ and consider the coupled system of SDEs*

$$\begin{aligned}
\begin{pmatrix} \Lambda_t^T \\ Z_t^T \end{pmatrix} &= \begin{pmatrix} \Lambda_0^T \\ Z_0^T \end{pmatrix} - \int_0^t \psi'(u_{T-s}(\infty)) \begin{pmatrix} \Lambda_{s-}^T \\ 0 \end{pmatrix} \mathrm{d}s + \sqrt{2\beta} \int_0^t \int_0^{\Lambda_{s-}^T} \begin{pmatrix} 1 \\ 0 \end{pmatrix} W(\mathrm{d}s, \mathrm{d}u) \\
&\quad + \int_0^t \int_0^\infty \int_0^{\Lambda_{s-}^T} \begin{pmatrix} r \\ 0 \end{pmatrix} \tilde{\mathtt{N}}^0_T(\mathrm{d}s, \mathrm{d}r, \mathrm{d}\nu) \\
&\quad + \int_0^t \int_0^\infty \int_1^{Z_{s-}^T} \begin{pmatrix} r \\ 0 \end{pmatrix} \mathtt{N}^1_T(\mathrm{d}s, \mathrm{d}r, \mathrm{d}j) \\
&\quad + \int_0^t \int_0^\infty \int_0^\infty \int_1^{Z_{s-}^T} \begin{pmatrix} r \\ k-1 \end{pmatrix} \mathtt{N}^2_T(\mathrm{d}s, \mathrm{d}r, \mathrm{d}k, \mathrm{d}j) \\
&\quad + 2\beta \int_0^t \begin{pmatrix} Z_{s-}^T \\ 0 \end{pmatrix} \mathrm{d}s, \qquad 0 \leq t < T.
\end{aligned} \tag{2.17}$$

The equation (2.17) has a unique strong solution for arbitrary ($\mathcal{F}_0^T$-measurable) initial values $\Lambda_0^T \geq 0$ and $Z_0^T \in \mathbb{N}_0$ (where $\mathcal{F}_t^T := \sigma((\Lambda_s^T, Z_s^T) : s \leq t), t < T$). 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 $T > t \geq 0$, conditional on $\mathcal{F}_t^{\Lambda^T} := \sigma(\Lambda_s^T, s \leq t)$, 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 (2.5);*

(iii) *Conditional on* $\{Z_0^T = 0\}$*, the process* $(\Lambda_t^T, 0 \le t < T)$ *corresponds to a weak solution to (2.5) conditioned to become extinct by time* T*.*

The SDE evolution in Theorem 2.2.2 mimics the skeletal decomposition in (2.15), albeit that the different components in the decomposition are time-dependent. In particular, we note that the underlying skeleton Z^T can be thought of as a time-inhomogenous Galton–Watson process (a T-prolific skeleton) such that, at time $s < T$, its branching rate is given by

$$q^{T-s} := \frac{u_{T-s}(\infty)\psi'(u_{T-s}(\infty)) - \psi(u_{T-s}(\infty))}{u_{T-s}(\infty)} \tag{2.18}$$

and 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 (2.1), 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, one easily verifies from (2.3) that $\lim_{\lambda\to\infty}[\lambda\psi'(\lambda) - \psi(\lambda)]/\lambda = \infty$. Together, these facts imply the explosion of (2.18) as $s \to T$.

We also note from the integrals involving $\mathtt{N}_T^1$ and $\mathtt{N}_T^2$ that there is mass immigrating off the space-time trajectory of Z^T. Moreover, once mass has immigrated, the first four terms of (2.17) show that it evolves as a time-inhomogenous CSBP.

Note, that in the supercritical setting $u_{T-t}(\infty)$ converges to λ^* for all fixed $t > 0$ as $T \to \infty$. This intuitively means that when T goes to ∞, one can recover the prolific skeleton decomposition of Theorem 2.2.1 from the time-inhomogeneous one of Theorem 2.2.2.

Finally with the finite-horizon SDE skeletal decomposition in Theorem 2.2.2, we may now turn our attention to understanding what happens when we observe the solution to (2.17) 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 2.2.3. *Suppose that* ψ *is a critical or subcritical branching mechanism such that Grey's condition (2.7) holds. Suppose, moreover, that* $((\Lambda_t^T, Z_t^T), 0 \le t < T)$ *is a weak solution to (2.17) 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 (2.10).*

Theorem 2.2.3 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 [23] 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 decent. This phenomenon was originally observed in [16], where the scaling limit of a Galton–Watson processes 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 followed. In the next section we explain the heuristic behind how (2.5) can be decoupled into components that arise in (2.15). The heuristic is used in Section 2.4 where the proof of Theorem 2.2.1 is given. In this sense our proof of Theorem 2.2.1 has the feel of a 'guess-and-verify' approach. In Section 2.5, 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. [11] to provide the heuristic behind the mathematical structures that lie behind the proof of Theorem 2.2.2. Given the similarity of this proof to that of Theorem 2.2.1, it is sketched in Section 2.6. Finally in Section 2.7 we provide the proof of Theorem 2.2.3.

2.3 Thinning of the CSBP SDE

In this section, we will perform an initial manipulation of the SDE (2.5), which we will need in order to make comparative statements for Theorems 2.2.1 and 2.2.2. To this end, we will introduce some independent marks on the atoms of the Poisson process N driving (2.5) and use them to thin out various contributions to the SDE evolution.

Denote by $(t_i, r_i, \nu_i : i \in \mathbb{N})$ 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,

$$\mathcal{N}(\mathrm{d}s, \mathrm{d}r, \mathrm{d}\nu, \mathrm{d}k) := \sum_{i \in \mathbb{N}} \delta_{(t_i, r_i, \nu_i, k_i)}(\mathrm{d}s, \mathrm{d}r, \mathrm{d}\nu, \mathrm{d}k) \tag{2.19}$$

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

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

Now define three random measures by

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

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

and

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

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 (2.5). We have that, for $t\geq 0$,

$$\begin{aligned}
X_t &= x+\alpha\int_0^t X_{s-}\mathrm{d}s+\sqrt{2\beta}\int_0^t\int_0^{X_{s-}}W(\mathrm{d}s,\mathrm{d}u)+\int_0^t\int_0^\infty\int_0^{X_{s-}} r\tilde{N}^0(\mathrm{d}s,\mathrm{d}r,\mathrm{d}\nu)\\
&\quad+\int_0^t\int_0^\infty\int_0^{X_{s-}} rN^1(\mathrm{d}s,\mathrm{d}r,\mathrm{d}\nu)+\int_0^t\int_0^\infty\int_0^{X_{s-}} rN^2(\mathrm{d}s,\mathrm{d}r,\mathrm{d}\nu)\\
&\quad-\int_0^t\int_0^\infty X_{s-}\sum_{n=1}^{\infty}\frac{(\lambda r)^n}{n!}\mathrm{e}^{-\lambda r}r\Pi(\mathrm{d}r)\mathrm{d}s\\
&= x-\psi'(\lambda)\int_0^t X_{s-}\mathrm{d}s+\sqrt{2\beta}\int_0^t\int_0^{X_{s-}}W(\mathrm{d}s,\mathrm{d}u)+\int_0^t\int_0^\infty\int_0^{X_{s-}} r\tilde{N}^0(\mathrm{d}s,\mathrm{d}r,\mathrm{d}\nu)\\
&\quad+\int_0^t\int_0^\infty\int_0^{X_{s-}} rN^1(\mathrm{d}s,\mathrm{d}r,\mathrm{d}\nu)+2\beta\lambda\int_0^t X_{s-}\mathrm{d}s\\
&\quad+\int_0^t\int_0^\infty\int_0^{X_{s-}} rN^2(\mathrm{d}s,\mathrm{d}r,\mathrm{d}\nu),
\end{aligned}\tag{2.20}$$

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 (2.13), the first line in the last equality of (2.20) corresponds to the dynamics of a subcritical CSBP with branching mechanism ψ_λ.

Inspecting the statement of Theorem 2.2.1, we see intuitively that in order to prove this result, our job is to show that the integrals on the right-hand side of (2.20) driven by N^1 and N^2, and the integral $2\beta\lambda\int_0^t X_{s-}\mathrm{d}s$ can be identified with the mass that immigrates off the skeleton.

2.4 λ-Skeleton: Proof of Theorem 2.2.1

We start by addressing the claim that (2.15) 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 (2.15) 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 $\mathtt{N}^1$ and $\mathtt{N}^2$, as well as an additional Poisson random measure $\mathtt{N}^*$. We will insist that $\mathtt{N}^1(\mathrm{d}s,\mathrm{d}r,\mathrm{d}j,\mathrm{d}\omega)$ has intensity $\mathrm{d}s\otimes r\mathrm{e}^{-\lambda r}\Pi(\mathrm{d}r)\otimes\sharp(\mathrm{d}j)\otimes\mathbb{P}^{(\lambda)}_r(\mathrm{d}\omega)$ on $[0,\infty)^2\times\mathbb{N}_0\times\mathbb{D}([0,\infty),\mathbb{R})$, $\mathtt{N}^2(\mathrm{d}s,\mathrm{d}r,\mathrm{d}k,\mathrm{d}j,\mathrm{d}\omega)$ 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 $\mathtt{N}^*(\mathrm{d}s, \mathrm{d}j, \mathrm{d}\omega)$ 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 (2.15) will be to first define $(Z_t, t \geq 0)$ as the continuous-time Galton–Watson process with branching rate $\psi'(\lambda)$ and offspring distribution given by $(p_k, k \geq 0)$. It is then easy to otherwise Z in the more complicated form

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

Next we take

$$\Lambda_t = X_t^{(\lambda)} + D_t, \qquad t \geq 0, \tag{2.21}$$

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

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

To see why this provides a weak solution to (2.15), we may appeal to the Martingale representation of weak solutions (see Theorem 2.3 of [24]) and note that the pair (Λ, Z) described above are Markovian and that its generator can be identified consistently with the generator of the process associated to (2.15); that it to say, their common generator is given by

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

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 $n = 0$.)

Pathwise uniqueness is also relatively easy to establish. Indeed, suppose that Λ is the first component of any path solution to (2.15) with driving source of randomness $\mathtt{N}^0$,

$\mathtt{N}^1$, $\mathtt{N}^2$ 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

$$H(\Lambda_s, s < t) = -\psi'(\lambda)\int_0^t \Lambda_{s-}\mathrm{d}s + \sqrt{2\beta}\int_0^t\int_0^{\Lambda_{s-}} W(\mathrm{d}s, \mathrm{d}u) \\ + \int_0^t\int_0^\infty\int_0^{\Lambda_{s-}} r\tilde{\mathtt{N}}^0(\mathrm{d}s, \mathrm{d}r, \mathrm{d}\nu), \qquad t \geq 0,$$

and

$$I_t = \int_0^t\int_0^\infty\int_1^{Z_{s-}} r\mathtt{N}^1(\mathrm{d}s, \mathrm{d}r, \mathrm{d}j) \\ + \int_0^t\int_0^\infty\int_0^\infty\int_1^{Z_{s-}} r\mathtt{N}^2(\mathrm{d}s, \mathrm{d}r, \mathrm{d}k, \mathrm{d}j) + 2\beta\int_0^t Z_{s-}\mathrm{d}s, \qquad t \geq 0.$$

Recalling that the almost sure path of Z is uniquely defined by $\mathtt{N}^2$, it follows that, if $\Lambda^{(1)}$ and $\Lambda^{(2)}$ are two path solutions to (2.15) 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 one obtains when looking at the path difference between two solutions of an SDE of the type given in (2.5). Since there is pathwise uniqueness for (2.5), 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 (2.15) 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 $t \geq 0$, $Z_t|\Lambda_t$ 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 (2.15) 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 $x > 0$,

$$\mathbf{P}_x(\cdot) = \sum_{n\geq 0}\frac{(\lambda x)^n}{n!}\mathrm{e}^{-\lambda x}\mathbf{P}_{x,n}(\cdot). \tag{2.22}$$

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, cf. Theorem 32 of [41], for $t\geq 0$,

$$
\begin{aligned}
\mathrm{d}F_t =& -\eta F_{t-}\mathrm{d}\Lambda_t - \theta F_{t-}\mathrm{d}Z_t + \frac{1}{2}\eta^2 F_{t-}\mathrm{d}[\Lambda,\Lambda]^c_t + \frac{1}{2}\theta^2 F_{t-}\mathrm{d}[Z,Z]^c_t \\
&+ \eta\theta F_{t-}\mathrm{d}[\Lambda,Z]^c_t + \Delta F_t + \eta F_{t-}\Delta\Lambda_t + \theta F_{t-}\Delta Z_t.
\end{aligned}
$$

(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 $[Z,Z]^c_t = [\Lambda,Z]^c_t = 0$. 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

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

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 (2.15) 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,

$$
\begin{aligned}
F_t =& F_0 - \eta\int_0^t F_{s-}\mathrm{d}\Lambda_s - \theta\int_0^t F_{s-}\mathrm{d}Z_s + \beta\eta^2\int_0^t F_{s-}\Lambda_{s-}\mathrm{d}s \\
&+ \sum_{s\leq t} F_{s-}\left\{\mathrm{e}^{-\eta\Delta^{(0)}\Lambda_s} - 1 + \eta\Delta^{(0)}\Lambda_s\right\} \\
&+ \sum_{s\leq t} F_{s-}\left\{\mathrm{e}^{-\eta\Delta^{(1)}\Lambda_s} - 1 + \eta\Delta^{(1)}\Lambda_s\right\} \\
&+ \sum_{s\leq t} F_{s-}\left\{\mathrm{e}^{-\eta\Delta^{(2)}\Lambda_s-\theta\Delta Z_s} - 1 + \eta\Delta^{(2)}\Lambda_s + \theta\Delta Z_s\right\}. \qquad (2.23)
\end{aligned}
$$

Now, note that we can re-write the first element of the vectorial SDE (2.15) as

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

where M_t is a zero-mean martingale corresponding to the integral in (2.15) with respect to $\tilde{\mathtt{N}}^0$. Therefore performing the necessary calculus in (2.23) for the integral with respect to $\mathrm{d}\Lambda_t$, we get that

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

for $t \geq 0$, 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

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

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

$$
\begin{aligned}
\frac{\partial}{\partial t}f_t(\eta,\theta) &= A_\lambda(\eta,\theta)\frac{\partial}{\partial\eta}f_t(\eta,\theta) + B_\lambda(\eta,\theta)\frac{\partial}{\partial\theta}f_t(\eta,\theta), \\
f_0(\eta,\theta) &= \mathrm{e}^{-(\eta+\lambda(1-\mathrm{e}^{-\theta}))x},
\end{aligned}
\tag{2.24}
$$

where

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

Standard theory for linear partial differential equation (2.24), see for example Chapter 3 (Theorem 2, p107) of [18] 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

$$
\mathbb{E}_x[\mathrm{e}^{-(\eta+\lambda(1-\mathrm{e}^{-\theta}))X_t}] = \mathrm{e}^{-u_t(\eta+\lambda(1-\mathrm{e}^{-\theta}))x}, \qquad x,t,\theta,\eta \geq 0, \tag{2.25}
$$

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

$$
\begin{aligned}
\frac{\partial}{\partial t}g_t(\kappa) &= -\psi(\kappa)\frac{\partial}{\partial\kappa}g_t(\kappa), \\
g_0(\kappa) &= \mathrm{e}^{-\kappa x}.
\end{aligned}
\tag{2.26}
$$

See for example Exercise 12.2 in [25]. After a laborious amount of algebra one 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 (2.26) and write, for $x, t, \eta, \theta \geq 0$,

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

Now we fix the initial value of (2.24) and (2.26) by choosing $x = 1$. Then local uniqueness of the solution to (2.24) (or equivalently the local uniqueness of (2.26)) thus tells us that there exists $t_0 > 0$ such that $g_t(\eta + \lambda(1 - \mathrm{e}^{-\theta})) = f_t(\eta, \theta)$ for all $\eta, \theta \geq 0$ and $t \in [0, t_0]$.

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

$$\mathbb{E}_1[\mathrm{e}^{-(\eta+\lambda(1-\mathrm{e}^{-\theta}))X_t}] = \mathbf{E}_1[\mathrm{e}^{-\eta\Lambda_t - \theta Z_t}], \qquad \theta, \eta \geq 0, \tag{2.27}$$

we can sequentially observe the following implications. Firstly, 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 $(X_t, \texttt{Po}(\lambda x)|_{x=X_t})$ under $\mathbb{P}_1$, where $\texttt{Po}(\lambda x)$ 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 $\texttt{Po}(\lambda \Lambda_t)$.

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 Markov property, we extend the results for any $t > 0$. First, from (2.25) we can observe that

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

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

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

Recalling the representation (2.21) and the notation (2.22) we can write, for $t \leq t_0$,

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

as required.

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

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

and similarly

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

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

$$\mathbf{E}_x\left[\mathrm{e}^{-\eta\Lambda_t-\theta Z_t}\right] = \mathbb{E}_x\left[\mathrm{e}^{-(\eta+\lambda(1-\mathrm{e}^{-\theta}))X_t}\right]$$

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), $t\ge 0$, is a joint Markovian pair, this now global Poissonisation allows us to infer that Λ_t, $t\ge 0$, is itself Markovian. Indeed, for any bounded measurable and positive h and $s,t\ge 0$,

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

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

(ii) We have seen that the pair $((\Lambda_t, Z_t), t\ge 0)$ is a Markov process for any initial state (x,n) but, due to the dependence on Z, on its own $(\Lambda_t, t\ge 0)$ is not Markovian. However considering (2.27) we see that after the Poissonisation of Z_0, $(\Lambda_t, t\ge 0)$ becomes a Markov process with semi-group that agrees with that of $(X_t, t\ge 0)$. On account of the fact that X is the unique weak solution to (2.5), it automatically follows that Λ also represents the unique weak solution to (2.5).

(iii) Since the event $\{Z_0 = 0\}$ implies the event $\{Z_t = 0, t\ge 0\}$, the system (2.15) 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 as the SDE describing the evolution of a CSBP with branching mechanism ψ_λ. □

2.5 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 (2.17) 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 2.2.2. That is to say, X a (sub)critical ψ-CSBP where ψ satisfies Grey's condition (2.7). Let $(\xi_t, t \geq 0)$ be a spectrally positive Lévy process with Laplace exponent ψ. Using the classical work of [32, 33] (see also [11, 31]) 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 $(\hat{\xi}^{(t)}_r, 0 \leq r \leq t)$ the time reversed process at time t, that is $\hat{\xi}^{(t)}_r := \xi_t - \xi_{(t-r)-}$, and let $\hat{S}^{(t)}_r := \sup_{s \leq r} \hat{\xi}^{(t)}_s$. 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 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 (2.7), is continuous. There exists a notion of local time up to time t of H at level $a \geq 0$, henceforth denoted by L^a_t. Specifically, the family $(L^a_t, a, t \geq 0)$ satisfies

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

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

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

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 $\mathtt{n}$ 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 [11] 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 'u-th' 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 $\mathtt{n}_T$ denote the conditional probability $\mathtt{n}(\cdot | \sup_{s\geq 0} \epsilon_s \geq T)$ where ϵ is a canonical excursion of H under $\mathtt{n}$. 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 [11] describe the distribution of Z^T under $\mathtt{n}_T$ and prove the following.

Theorem 2.5.1. *Under* $\mathtt{n}_T$ *the process* $(Z_t^T, 0 \le t < T)$ *is a time-inhomogeneous Markov process whose law is characterised by the following identities. For every* $\lambda > 0$

$$\mathtt{n}_T\left[\exp\{-\lambda Z_t^T\}\right] = 1 - \frac{u_t((1-\mathrm{e}^{-\lambda})u_{T-t}(\infty))}{u_T(\infty)},$$

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

$$\mathtt{n}_T\left[\exp\{-\lambda Z_{t'}^T\}|Z_t^T\right] = \left(\mathtt{n}_{T-t}\left[\exp\{-\lambda Z_{t'-t}^{T-t}\}\right]\right)^{Z_t^T}.$$

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 Galton–Watson process. In [11] it is moreover shown that, conditionally on L_σ^t under $\mathtt{n}_T$, where σ is the length of the excursion ϵ, Z_t^T is Poisson distributed with intensity $u_{T-t}(\infty)L_\sigma^t$. Thinking of L_σ^t 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 Galton–Watson 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

$$\mathtt{n}_T(\gamma_T > t) = \frac{\psi(u_T(\infty))}{u_T(\infty)}\frac{u_{T-t}(\infty)}{\psi(u_{T-t}(\infty))}, \qquad t \in [0, T), \tag{2.28}$$

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

$$\mathtt{n}_T\left[r^{Z_{\gamma_T}^T}|\gamma_T = t\right] = 1 + \frac{\psi((1-r)u_{T-t}(\infty)) - (1-r)u_{T-t}(\infty)\psi'(u_{T-t}(\infty))}{u_{T-t}(\infty)\psi'(u_{T-t}(\infty)) - \psi(u_{T-t}(\infty))}, \qquad r \in [0,1].$$

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

$$\begin{aligned} p_k^{T-t} = {} & \frac{1}{u_{T-t}(\infty)\psi'(u_{T-t}(\infty)) - \psi(u_{T-t}(\infty))} \\ & \times \left\{\beta u_{T-t}^2(\infty)\mathbf{1}_{\{k=2\}} + \int_0^\infty \frac{(u_{T-t}(\infty)x)^k}{k!}\mathrm{e}^{-u_{T-t}(\infty)x}\Pi(\mathrm{d}x)\right\}. \end{aligned} \tag{2.29}$$

Using (2.28) we can compute the rate of branching at any height t in the excursion. First it is not hard to see that

$$\mathtt{n}_T(\gamma_T > t + s|\gamma_T > t) = \mathtt{n}_{T-t}(\gamma_T > s), \qquad 0 \le t + s < T.$$

Hence, the rate is

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

Again thinking of L^t_σ, $t<T$ 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 (2.1), we know that the probability that the process dies up to time T is given by:

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

where we assume Grey's condition (2.7) 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

$$\left.\frac{\mathrm{d}\mathbb{P}^T_x}{\mathrm{d}\mathbb{P}_x}\right|_{\mathcal{F}_t} = \frac{\mathrm{e}^{-X_t u_{T-t}(\infty)}}{\mathrm{e}^{-x u_T(\infty)}}, \qquad t \geq 0, x > 0. \tag{2.30}$$

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 $(X, \mathbb{P}^T_x)$, $x > 0$, 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 $x > 0$,

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

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

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

Gathering terms, making use of the expression for ψ in (2.3) 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

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

where the infinitesimal generator is given by

$$
\begin{aligned}
\hat{\mathcal{L}}_{T-t} f(x) = \psi'(u_{T-t}(\infty)) x f'(x) + \beta x f''(x) + \\
x\int_0^\infty (f(x+y) - f(x) - y f'(x))\mathrm{e}^{-u_{T-t}(\infty)y}\Pi(\mathrm{d}y). \quad (2.31)
\end{aligned}
$$

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

$$
\mathcal{L}_\lambda f(x) = \psi'(\lambda) x f'(x) + \beta x f''(x) + x\int_0^\infty (f(x+y) - f(x) - y f'(x))\mathrm{e}^{-\lambda y}\Pi(\mathrm{d}y) \quad (2.32)
$$

for suitably smooth and integrable functions f. Recall that the CSBP with generator (2.32) is subcritical providing $\psi'(\lambda) > 0$ and, taking account of (2.4), 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

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

then

$$V_t^T(\theta) = u_t(\theta + u_{T-t}(\infty)) - u_T(\infty). \tag{2.33}$$

Recalling that we are assuming Grey's condition (2.7) for a (sub)critical process, we note from (2.1) that

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

In that case, the density in (2.30) tends to unity as $T \to \infty$.

We conclude this section by returning to Theorem 2.2.2. The discussion in this section shows that in the (sub)critical case, the components of the SDE (2.17) mimic precisely the description of the T-skeleton in the previous section. In particular, the first three integrals in (2.17) 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 (2.17), matches precisely the dynamics of the T-prolific skeleton described in the previous section (for which we pre-emptively used the same notation), which is a time-dependent continuous-time Galton–Watson process. Indeed, the branching rate and the time-dependent offspring distribution of both match.

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 (2.30) for any CSBP satisfying (2.7), and get a time-dependent CSBP with generator $\hat{\mathcal{L}}_{T-t}$. Although the results of Theorem 2.2.2 were motived by Duquesne and Le Gall [11], 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.

2.6 Finite-time horizon Skeleton: Proof of Theorem 2.2.2

Now that we understand that the mathematical structure of (2.17) is little more than a time-dependent version of (2.15), the reader will not be surprised by the claim that the proof of strong uniqueness to (2.17) as well as parts (i) and (ii) of Theorem 2.2.1 pass through almost verbatim, albeit needing some minor adjustments for additional time derivatives of $u_{T-t}(\infty)$, 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 (2.17) 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 processes.

Suppose that we write $\mathbf{P}^T_{x,n}$, $x \ge 0, n \in \mathbb{N}_0$ for the law of the Markov probabilities corresponding to the solution of (2.17). Moreover, we will again abuse this notation in the spirit of (2.22) and write $\mathbf{P}^T_x$, $x \ge 0$, when Z^T_0 is randomised to be an independent

Poisson random variable with rate $u_T(\infty)x$. We can use part (i) and (ii) of Theorem 2.2.2, together with (2.30) to deduce that

$$\begin{aligned}
\mathbf{E}_x^T\left[\mathrm{e}^{-\eta\Lambda_t}|Z_0=0\right] &= \frac{\mathbf{E}_x^T\left[\mathrm{e}^{-\eta\Lambda_t},\ Z_0=0\right]}{\mathbf{P}_x^T(Z_0=0)}\\
&= \frac{\mathbf{E}_x^T\left[\mathrm{e}^{-\eta\Lambda_t},\ Z_t=0\right]}{\mathbf{P}_x^T(Z_0=0)}\\
&= \mathrm{e}^{u_T(\infty)x}\mathbf{E}_x^T\left[\mathrm{e}^{-(\eta+u_{T-t}(\infty))\Lambda_t}\right]\\
&= \mathbb{E}_x^T\left[\mathrm{e}^{-\eta X_t}\right]. \qquad (2.34)
\end{aligned}$$

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 2.2.2 is thus proved. □

2.7 Thinning the skeleton to a spine: Proof of Theorem 2.2.3

The aim of this section is to recover the unique solution to (2.10) as a weak limit of (2.17) in the sense of Skorokhod convergence. To this end, we assume throughout the conditions of Theorem 2.2.3, in particular that ψ is a critical or subcritical branching mechanism and Grey's condition (2.7) 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 $Z_0^T\sim \mathtt{Po}(u_T(\infty)x)$, 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

$$\varpi_k^{x,T} := \mathbf{P}_x^T[Z_0=k|Z_0\geq 1] = \frac{(u_T(\infty)x)^k}{k!}\frac{\mathrm{e}^{-u_T(\infty)x}}{1-\mathrm{e}^{-u_T(\infty)x}}, \qquad k\geq 1. \qquad (2.35)$$

We thus see that the probabilities (2.35) all tend to zero unless $k=1$ in which case the limit is unity. Moreover, Theorem 2.2.2 (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 (2.28), it follows from l'Hospital's rule that the rate at which aforementioned most common recent ancestor branches begins to slow down since

$$\frac{\psi(u_T(\infty))}{u_T(\infty)}\frac{u_{T-t}(\infty)}{\psi(u_{T-t}(\infty))}\to 1,$$

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.

Thirdly, 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 (2.30). 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 2.2.3 should now take on a natural meaning. We give its proof below.

Proof of Theorem 2.2.3. According to Theorem 2.5 on p.167 of [17, 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 semigroup on $C_0(E)$ such that, for $f\in C_0(E)$, with respect to the supremum norm on the space $C_0(E)$,

$$\lim_{T\to\infty}\mathcal{P}^T_t f = \mathcal{P}_t f, \quad t\geq 0, \tag{2.36}$$

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 (2.36), where $\mathcal{P}$ is a Feller semigroup, 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 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 (Λ_t, Z_t, τ_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 (2.17), and $\tau_t := t$, $0\leq t\leq t_0$. Consistently with previous related 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 $\mathtt{P}^T_t[f](x,n,s)=f(x,n,s)$ when $T\leq t_0$, and when $T>t_0$ we have

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

for $(x,n,s)\in[0,\infty)\times\mathbb{N}_0\times[0,t_0]$, and $\mathtt{P}_t^T[f](x,n,s):=f(x,n,s)$ for $(x,n,s)\in[0,\infty)\times\mathbb{N}_0\times(t_0,\infty)$. We take $\mathcal{P}^T=\mathtt{P}^T$. In order to verify the Feller property of $\mathtt{P}_t^T[f](x,n,s)$ we need to check two things (cf. Proposition 2.4 in Chapter III of [42]):

(i) For each $t\geq 0$, the function $(x,n,s)\mapsto \mathtt{P}_t^T[f](x,n,s)$ 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 $\lim_{t\downarrow s}\mathtt{P}_t^T[f](x,n,s)=f(x,n,s)$.

Note that when $T\leq t_0$, or $s\geq t_0$, or $t\leq s\leq t_0$, we have $\mathtt{P}_t^T[f](x,n,s)=f(x,n,s)$. 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 $\mathtt{P}_t^T[f](x,n,s)$ we can restrict ourselves to the case of $s\leq t\leq t_0<T$.

By denseness of the sub-algebra generated by exponential functions (according to the uniform topology) in $C_0([0,\infty)\times\mathbb{N}_0\times[0,\infty))$, it suffices to check, for (i), that

$$(x,n,s)\mapsto \mathbf{E}_{x,n,s}^T\left[\mathrm{e}^{-\gamma\Lambda_t-\theta Z_t-\varphi\tau_t}\right],\quad s\leq t\leq t_0<T, \tag{2.37}$$

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

$$\lim_{t\downarrow s}\mathbf{E}_{x,n,s}^T\left[\mathrm{e}^{-\gamma\Lambda_t-\theta Z_t-\varphi\tau_t}\right]=\mathrm{e}^{-\gamma x-\theta n-\varphi s},\quad s\leq t\leq t_0<T. \tag{2.38}$$

To this end, note that

$$\mathbf{E}_{x,n,s}^T\left[\mathrm{e}^{-\gamma\Lambda_t-\theta Z_t-\varphi\tau_t}\right]=\mathbf{E}_{x,n}^{T-s}\left[\mathrm{e}^{-\gamma\Lambda_{t-s}-\theta Z_{t-s}}\right]\mathrm{e}^{-\varphi t},\qquad s\leq t\leq t_0<T. \tag{2.39}$$

In order to evaluate expectation on the right-hand side above, we want to work with an appropriate representation of the unique weak solution to (2.17). We shall do so by following the example of how the weak solution to (2.15) was identified in the form (2.21). As before, we need to introduce additionally marked versions of the Poisson random measures $\mathtt{N}_T^1$ and $\mathtt{N}_T^2$, as well as an additional Poisson random measure $\mathtt{N}_T^*$. We will insist that Poisson random measure $\mathtt{N}_T^1(\mathrm{d}s,\mathrm{d}r,\mathrm{d}j,\mathrm{d}\omega)$ 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 $\mathtt{N}_T^2(\mathrm{d}s,\mathrm{d}r,\mathrm{d}k,\mathrm{d}j,\mathrm{d}\omega)$ 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 $\mathtt{N}_T^*(\mathrm{d}s,\mathrm{d}j,\mathrm{d}\omega)$ 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

$$\mathbb{Q}^T(1-\mathrm{e}^{-\gamma\omega_t})=V_t^T(\gamma),\qquad\gamma>0, \tag{2.40}$$

for $0\leq t<T$, where V_t^T was defined in (2.33). To recall some of the notation used in these rates, see (2.16) and (2.18).

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

$$\Lambda_t = X_t + D_t, \qquad t < T, \tag{2.41}$$

where X is autonomously independent with law $\mathbb{P}^T_x$ and, given $\mathtt{N}^2$, D is the uniquely identified (up to almost sure modification) 'dressed skeleton' described by

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

where $Z_0 = n$. The verification of this claim follows almost verbatim the same as for (2.15) albeit with 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 (2.41), as Z is piecewise constant, we can condition on the sigma-algebra generated by $\mathtt{N}^2_T$ and show, using Campbell's formula in between the jumps of Z, that, for $0 \le t < T$, $\gamma, \theta \ge 0$, $x \ge 0$ and $n \in \mathbb{N}_0$,

$$\begin{aligned} &\mathbf{E}^T_{x,n}\left[\mathrm{e}^{-\gamma\Lambda_t - \theta Z_t}\right] \\ &= \mathrm{e}^{-xV^T_t(\gamma)} \mathbf{E}^T_{0,n}\left[\mathrm{e}^{-\theta Z_t - \int_0^t Z_v \phi_{u_{T-v}(\infty)}(V^{T-v}_{t-v}(\gamma))\mathrm{d}v} \prod_{w \le t}\left(\int_0^\infty \mathrm{e}^{-rV^{T-w}_{t-w}(\gamma)} \eta^{T-w}_{\Delta Z_w + 1}(\mathrm{d}r)\right)\right] \end{aligned} \tag{2.42}$$

where

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

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

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

Given the identities (2.39) and (2.42), the two required verifications in (2.37) and (2.38) follow easily as direct consequence of continuity and bounded convergence in (2.42).

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 $n \in \mathbb{N}_0$, $x \ge 0$, let $\mathbb{P}^{(n)}_x$ be the law of the homogeneous Markov process described by the weak solution to

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

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

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

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

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

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

Lastly, for each $T \geq 0$ we take ν^T 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 (2.35) 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 in [17, Chapter 4], we just need to check the analogue of (2.36) in our setting.

To this end, notice first that we can restrict ourselves to $0 \leq s \leq t \leq t_0$, since when $s > t_0$, $\mathtt{P}_t^{\uparrow}[f](x,n,s) = \mathtt{P}_t^T[f](x,n,s)$ 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 (2.18) that $q^T \to 0$ as $T \to \infty$, and this yields that, under $\mathbf{P}_{0,n}^T$, process Z converge in probability uniformly on $[0,t]$, as $T \to \infty$ (cf. Theorem 6.1, Chapter 1, p28 of [17]) to the constant process $Z_s \equiv n$, $s \leq t$. Referring back to (2.42), 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 $x \geq 0, 0 \leq s \leq t \leq t_0$, $n \in \mathbb{N}_0$,

$$\begin{aligned}
\lim_{T\to\infty} \mathtt{P}_t^T[f_{\gamma,\theta,\varphi}](x,n,s) &= \mathrm{e}^{-\varphi t} \lim_{T\to\infty} \mathbf{E}_{x,n}^{T-s}\left[\mathrm{e}^{-\gamma\Lambda_{t-s} - \theta Z_{t-s}}\right] \\
&= \mathtt{P}_t^{\uparrow}[f_{\gamma,\theta,\varphi}](x,n,s) \\
&= \mathrm{e}^{-xu_{t-s}(\gamma) - \theta n - n\int_0^{t-s}\phi_0(u_{t-s-v}(\gamma))\mathrm{d}v - \varphi t},
\end{aligned}$$

where

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

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 nonnegative and, for each $n \in \mathbb{N}_0$, $Z_t \geq Z_0 = n$, $t \geq 0$, a.s. under $\mathbf{P}_{0,n}^T$, for all $T > 0$, using the triangle inequality, we have

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

where we have set:

$$A_R(T) := \sup_{x\leq R, s\leq t} |\mathrm{e}^{-xV_{t-s}^{T-s}(\gamma)} - \mathrm{e}^{-xu_{t-s}(\gamma)}|, \quad A^R(T) := \sup_{s\leq t} \mathrm{e}^{-RV_{t-s}^{T-s}(\gamma)} + \sup_{s\leq t} \mathrm{e}^{-Ru_{t-s}(\gamma)},$$

$$B_N(T) := \sup_{n\le N, s\le t}\left|\mathbf{E}^{T-s}_{0,n}\left[\mathrm{e}^{-\theta Z_{t-s}-\int_0^{t-s} Z_v\phi_{u_{T-s-v}(\infty)}(V^{T-s-v}_{t-s-v}(\gamma))\mathrm{d}v}\right.\right.$$
$$\left.\left.\prod_{w\le t-s}\left(\int_0^\infty \mathrm{e}^{-rV^{T-s-w}_{t-s-w}(\gamma)}\eta^{T-s-w}_{\Delta Z_w+1}(\mathrm{d}r)\right)\right] - \mathrm{e}^{-\theta n - n\int_0^{t-s}\phi_0(u_{t-s-v}(\gamma))\mathrm{d}v}\right|$$

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

$$A_R(T) \le \sup_{s\le t} R|V^{T-s}_{t-s}(\gamma) - u_{t-s}(\gamma)| = R\sup_{s\le t}|u_{t-s}(\gamma + u_{T-t}(\infty)) - u_{T-s}(\infty) - u_{t-s}(\gamma)|.$$

The identity $\partial u_s(\theta)/\partial\theta = \mathrm{e}^{-\int_0^s \psi'(u_r(\theta))\mathrm{d}r}$ (see (12.12) in [25, Chapter 12]) and the fact that $\psi'(\theta)\ge 0$ 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

$$V^{T-s}_{t-s}(\gamma) \ge u_{t-s}(\gamma + u_{T-t}(\infty)) - u_{T-t}(\infty) \ge \inf_{\lambda\le u_{T-t}(\infty)} u_t(\gamma+\lambda) - \lambda,$$

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

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

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

$$\begin{aligned}
B_N(T) \le{}& \max_{n\le N}\sup_{s\le t}\mathbf{E}^{T-s}_{0,n}\Bigg[1\wedge \\
&\qquad\left(\int_0^{t-s}|Z_v\phi_{u_{T-s-v}(\infty)}(V^{T-s-v}_{t-s-v}(\gamma)) - n\phi_0(u_{t-s-v}(\gamma))|\mathrm{d}v\right)\Bigg] \\
&+ \max_{n\le N}\sup_{s\le t}\mathbf{E}^{T-s}_{0,n}\left[\left|\mathrm{e}^{-\theta Z_{t-s}} - \mathrm{e}^{-\theta n}\right|\right] \\
&+ \max_{n\le N}\sup_{s\le t}\mathbf{E}^{T-s}_{0,n}\left[\left|1 - \prod_{w\le s}\left(\int_0^\infty \mathrm{e}^{-rV^{T-s-w}_{t-s-w}(\gamma)}\eta^{T-s-w}_{\Delta Z_w+1}(\mathrm{d}r)\right)\right|\right] \\
\le{}& \max_{n\le N}\sup_{s\le t}\mathbf{E}^{T-s}_{0,n}\Bigg[\sup_{s'\le t}1\wedge \\
&\qquad\left(\int_0^{t-s'}|Z_v\phi_{u_{T-s'-v}(\infty)}(V^{T-s'-v}_{t-s-v}(\gamma)) - n\phi_0(u_{t-s'-v}(\gamma))|\mathrm{d}v\right)\Bigg] \\
&+ \max_{n\le N}\sup_{s\le t}\mathbf{E}^{T-s}_{0,n}\left[\sup_{s'\le t}\left|\mathrm{e}^{-\theta Z_{t-s'}} - \mathrm{e}^{-\theta n}\right|\right] \\
&+ \max_{n\le N}\sup_{s\le t}\mathbf{E}^{T-s}_{0,n}\left[\sup_{s'\le t}\left|1 - \prod_{w\le s'}\left(\int_0^\infty \mathrm{e}^{-rV^{T-s'-w}_{t-s'-w}(\gamma)}\eta^{T-s'-w}_{\Delta Z_w+1}(\mathrm{d}r)\right)\right|\right].
\end{aligned} \tag{2.43}$$

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

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

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 (2.43) 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 the third term of (2.43) is bounded by 1, and vanishes unless Z jumps at least once on $[0,s]$. This shows that the last term is bounded by $\max_{n\leq N}\sup_{s\leq t}\mathbf{P}^{T-s}_{0,n}(\sup_{w\leq t}\Delta Z_w>0)$, which goes to 0 when $T\to\infty$. Note that for all three terms in (2.43), 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 $0<t\leq t_0$, 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 such that $B^N\leq\varepsilon$, we thus get

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

Since ε was arbitrary this shows the convergence of the semigroups (2.36) 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 also completes the proof of Theorem 2.2.3. □

Bibliography

[1] Romain Abraham and Jean-François Delmas. A continuum-tree-valued Markov process. *Ann. Probab.*, 40(3):1167–1211, 2012.

[2] M. Barczy, Z. Li, and G. Pap. Yamada-Watanabe results for stochastic differential equations with jumps. *Int. J. Stoch. Anal.*, pages Art. ID 460472, 23, 2015.

[3] J. Berestycki, M. C. Fittipaldi, and J. Fontbona. Ray-Knight representation of flows of branching processes with competition by pruning of Lévy trees. *Probab. Theory Relat. Fields*, 172(3):725–788, 2018.

[4] J. Berestycki, A. E. Kyprianou, and A. Murillo-Salas. The prolific backbone for supercritical superprocesses. *Stochastic Process. Appl.*, 121(6):1315–1331, 2011.

[5] J. Bertoin, J. Fontbona, and S. Martínez. On prolific individuals in a supercritical continuous-state branching process. *J. Appl. Probab.*, 45(3):714–726, 2008.

[6] J. Bertoin and J-F. Le Gall. Stochastic flows associated to coalescent processes iii: limit theorems. *Illinois J. Math.*, 50(1):147–181, 2006.

[7] D. A. Dawson and Z. Li. Skew convolution semigroups and affine Markov processes. *Annals of Probability*, 34:1103?1142, 2006.

[8] D. A. Dawson and Z. Li. Stochastic equations, flows and measure-valued processes. *Annals of Probability*, 40(2):813–857, 2012.

[9] T. Duquesne and C. Labbé. On the eve property for csbp. *Electron. J. Probab.*, 19:31 pp., 2014.

[10] T. Duquesne and J.-F. Le Gall. Random trees, Lévy processes and spatial branching processes. *Astérisque*, (281):vi+147, 2002.

[11] T. Duquesne and J.-F. Le Gall. Random Trees, Lévy Processes and Spatial Branching Processes. *ArXiv Mathematics e-prints*, September 2005.

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

[13] E. B. Dynkin and S. E. Kuznetsov. $\mathbb{N}$-measures for branching exit Markov systems and their applications to differential equations. *Probab. Theory Related Fields*, 130(1):135–150, 2004.

[14] M. Eckhoff, A. E. Kyprianou, and M. Winkel. Spines, skeletons and the strong law of large numbers for superdiffusions. *Ann. Probab.*, 43(5):2545–2610, 2015.

[15] N. El Karoui and S. Roelly. Propriétés de martingales, explosion et représentation de Lévy-Khintchine d'une classe de processus de branchement à valeurs mesures. *Stochastic Process. Appl.*, 38(2):239–266, 1991.

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

[17] S. N. Ethier and T. G. Kurtz. *Markov processes*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, 1986. Characterization and convergence.

[18] L. C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 1998.

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

[20] M. C. Fittipaldi and J. Fontbona. On SDE associated with continuous-state branching processes conditioned to never be extinct. *Electron. Commun. Probab.*, 17:no. 49, 13, 2012.

[21] Zongfei Fu and Zenghu Li. Stochastic equations of non-negative processes with jumps. *Stochastic Process. Appl.*, 120(3):306–330, 2010.

[22] D. R. Grey. On possible rates of growth of age-dependent branching processes with immigration. *J. Appl. Probability*, 13(1):138–143, 1976.

[23] S. C. Harris, M. Hesse, and A. E. Kyprianou. Branching Brownian motion in a strip: survival near criticality. *Ann. Probab.*, 44(1):235–275, 2016.

[24] T. Kurtz. *Stochastic Analysis 2010*. Springer, 2011. Equivalence of stochastic equations and martingale problems.

[25] A. E. Kyprianou. *Fluctuations of Lévy processes with applications*. Universitext. Springer, Heidelberg, second edition, 2014. Introductory lectures.

[26] A. E. Kyprianou, A. Murillo-Salas, and J. L. Pérez. An application of the backbone decomposition to supercritical super-Brownian motion with a barrier. *J. Appl. Probab.*, 49(3):671–684, 2012.

[27] A. E. Kyprianou, J.-L. Pérez, and Y.-X. Ren. The backbone decomposition for spatially dependent supercritical superprocesses. In *Séminaire de Probabilités XLVI*, volume 2123 of *Lecture Notes in Math.*, pages 33–59. Springer, Cham, 2014.

[28] A. E. Kyprianou and Y.-X. Ren. Backbone decomposition for continuous-state branching processes with immigration. *Statist. Probab. Lett.*, 82(1):139–144, 2012.

[29] A. Lambert. Quasi-stationary distributions and the continuous-state branching process conditioned to be never extinct. *Electron. J. Probab.*, 12:no. 14, 420–446, 2007.

[30] A. Lambert. Population dynamics and random genealogies. *Stoch. Models*, 24(suppl. 1):45–163, 2008.

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

[32] J.-F. Le Gall and Y. Le Jan. Branching processes in lévy processes: Laplace functionals of snakes and superprocesses. *Ann. Probab.*, 26(4):1407–1432, 1998.

[33] J.-F. Le Gall and Y. Le Jan. Branching processes in lévy processes: The exploration process. *Ann. Probab.*, 26(1):213–252, 1998.

[34] Ch.-Kh. Li. Skew convolution semigroups and related immigration processes. *Teor. Veroyatnost. i Primenen.*, 46(2):247–274, 2001.

[35] Zeng-Hu Li. Immigration structures associated with Dawson-Watanabe superprocesses. *Stochastic Process. Appl.*, 62(1):73–86, 1996.

[36] Zeng-Hu Li. Immigration processes associated with branching particle systems. *Adv. in Appl. Probab.*, 30(3):657–675, 1998.

[37] Zenghu Li. *Measure-valued branching Markov processes*. Probability and its Applications (New York). Springer, Heidelberg, 2011.

[38] Zenghu Li. Path-valued branching processes and nonlocal branching superprocesses. *Ann. Probab.*, 42(1):41–79, 2014.

[39] Zenghu Li. Continuous-state branching processes with immigration. *arXiv e-prints*, page arXiv:1901.03521, Jan 2019.

[40] P. Milos. Spatial central limit theorem for the supercritical ornstein-uhlenbeck superprocess. *Preprint*, 2012.

[41] P. Protter. *Stochastic integration and differential equations*, volume 21 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, 1990. A new approach.

[42] D. Revuz and M. Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, third edition, 1999.

[43] S. Roelly-Coppoletta and A. Rouault. Processus de Dawson-Watanabe conditionné par le futur lointain. *C. R. Acad. Sci. Paris Sér. I Math.*, 309(14):867–872, 1989.

[44] T. S. Salisbury and J. Verzani. On the conditioned exit measures of super Brownian motion. *Probab. Theory Related Fields*, 115(2):237–285, 1999.

[45] T. S. Salisbury and J. Verzani. Non-degenerate conditionings of the exit measures of super Brownian motion. *Stochastic Process. Appl.*, 87(1):25–52, 2000.

[46] Y.-C. Sheu. Lifetime and compactness of range for super-Brownian motion with a general branching mechanism. *Stochastic Process. Appl.*, 70(1):129–141, 1997.

Concluding remarks

In this chapter we have considered various skeletal decompositions of CSBPs, including a parametric family of decompositions for supercritical CSBPs, and finite-time horizon decompositions for CSBPs that satisfy Grey's condition. We have developed a (coupled) SDE approach that allowed us to put these decompositions into a common framework. The second co-ordinate of the coupled SDE describes the skeleton process, and the first co-ordinate gives the total mass in the system, while the structure of the SDE also reveals how this mass is immigrated off the skeleton. We have shown that upon projecting onto the appropriate filtration, the total mass is equal in law to the original CSBP.

In the finite time-horizon case we have used the SDE representation to observe what happens with the skeletal decomposition of a (sub)critical CSBP as we condition on survival up to larger and larger times. By doing so we have linked two well knows decompositions of (sub)critical CSBP, namely the skeletal decomposition and the so-called spine decomposition, which emerges when the CSBP is conditioned to survive eternally.

In the next chapter, in order the show the robustness of this method, we extend the SDE approach to the spatial setting of superprocesses.

Appendix 6B: Statement of Authorship

This declaration concerns the article entitled:			
Skeletal stochastic differential equations for superprocesses			
Publication status (tick one) **Draft manuscript** [ ] **Submitted** [X] **In review** [ ] **Accepted** [ ] **Published** [ ]			
Publication details (reference)	arXiv:1904.05966		
Copyright status (tick the appropriate statement) I hold the copyright for this material [ ] Copyright is retained by the publisher, but I have been given permission to replicate the material here [ ]			
Candidate's contribution to the paper (provide details, and also indicate as a percentage)	*All research has been conducted in equal partnership with collaborators and supervisors. It is unwise to try and measure percentages of intellectual contribution as research scholarship lies as much in the escalation of ideas through mathematical discourse as it does with the seed of ideas themselves. That said, the mathematical content of this thesis is, as an entire piece of work, inextricably associated to its author through intellectual ownership. Formulation of ideas: The candidate played an integral and fully collaborative role in the formulation of ideas. Design of methodology: The candidate played an integral and fully collaborative role in the design of methodology. Experimental work: N/A Presentation of data in journal format: N/A *The wording in this box follows the advice and approval of my supervisor, Professor Kyprianou.		
Statement from Candidate	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature.		
Signed		**Date**	16/07/2019

Last update: Feb 2019

Chapter 3

Skeletal stochastic differential equations for superprocesses

Dorottya Fekete [1] , Joaquin Fontbona [2], Andreas E. Kyprianou [1]

Abstract

It is well understood that a supercritical superprocess is equal in law to a discrete Markov branching process whose genealogy is dressed in a Poissonian way with immigration which initiates subcritial superprocesses. The Markov branching process corresponds to the genealogical description of *prolific individuals*, that is individuals who produce eternal genealogical lines of decent, and is often referred to as the *skeleton* or *backbone* of the original superprocess. The Poissonian dressing along the skeleton may be considered to be the remaining non-prolific genealogical mass in the superprocess. Such skeletal decompositions are equally well understood for continuous-state branching processes (CSBP).

In a previous article we developed an SDE approach to study the skeletal representation of CSBPs, which provided a common framework for the skeletal decompositions of supercritical and (sub)critical CSBPs. It also helped us to understand how the skeleton thins down onto one infinite line of descent when conditioning on survival until larger and larger times, and eventually forever.

Here our main motivation is to show the robustness of the SDE approach by expanding it to the spatial setting of superprocesses. The current article only considers supercritical superprocesses, leaving the subcritical case open.

3.1 Introduction

In this paper we revisit the notion of the so-called skeletal decomposition of superprocesses. It is well-known that when the survival probability is not 0 or 1, then non-trivial infinite genealogical lines of descent, which we call *prolific*, can be identified on the event

[1]Department of Mathematical Sciences, University of Bath, Claverton Down, Bath, BA2 7AY, United Kingdom. Email: d.fekete@bath.ac.uk, a.kyprianou@bath.ac.uk

[2]Centre for Mathematical Modelling, DIM CMM, UMI 2807 UChile-CNRS, Universidad de Chile, Santiago, Chile.. Email: fontbona@dim.uchile.cl

of survival. By now it is also well understood that the process itself can be decomposed along its prolific genealogies, where non-prolific mass is immigrated in a Poissonian way along the stochastically 'thinner' prolific skeleton. This fundamental phenomenon was first studied by Evans and O'Connell [14] for superprocesses with quadratic branching mechanism. They showed that the distribution of the superprocess at time $t \geq 0$ can be written as the sum of two independent processes. The first is a copy of the original process conditioned on extinction, while the second process is understood as the superposition of mass that has immigrated continuously along the trajectories of a dyadic branching particle diffusion, which is initiated from a Poisson number of particles. This distributional decomposition was later extended to the spatially dependent case by Engländer and Pinsky [10].

A pathwise decomposition for superprocesses with general branching mechanism was provided by Berestycki et al. [2]. Here the role of the skeleton is played by a branching particle diffusion that has the same motion generator as the superprocess, and the immigration is governed by three independent Poisson point processes. The first one results in what we call continuous immigration along the skeleton, where the so-called excursion measure plays the central role, and it assigns zero initial mass to the immigration process. The second point process discontinuously grafts independent copies of the original process conditioned on extinction on to the path of the skeleton. Finally, additional copies of the original process conditioned on extinction are immigrated off the skeleton at its branch points, where the initial mass of the immigrant depends on the number of offspring at the branch point. The spatially dependent version of this decomposition was considered in [22] and [9].

Other examples of skeletal decompositions for superprocesses include [33, 11, 23, 28, 16].

In a previous article [15] we developed a stochastic differential equation (SDE) approach to study the skeletal decomposition of continuous state branching processes (CSBPs). These decompositions were by no means new; prolific genealogies for both supercritical and subcritical CSBPs had been described, albeit in the latter case we have to be careful what we mean by 'prolific'. In particular, in [3], [5] and [23] specifically CSBPs were considered, but since the total mass process of a superprocess with spatially independent branching mechanism is a CSBP, skeletal decompositions for CSBPs also appear as a special case of some of the previously mentioned results.

The results in [15] were motivated by the work of Duquesne and Winkel [5], and Duquesne and Le Gall [4]. Duquesne and Winkel, in the context of Lévy trees, provided a parametric family of decompositions for finite-mean supercritical CSBPs that satisfy Grey's condition. They showed that one can find a decomposition of the CSBP for a whole family of embedded skeletons, where the 'thinnest' one is the prolific skeleton with all the infinite genealogical lines of descent, while the other embedded skeletons not only contain the infinite genealogies, but also some finite ones grafted on to the prolific tree. On the other hand, Duquesne and Le Gall studied subcritical CSBPs, and using the height process gave a description of those genealogies who survive until some fixed time $T > 0$. It is well known that a subcritical CSBP goes extinct almost surely, thus prolific individuals, in the classic sense, do not exist in the population. But since it is possible that the process survives until some fixed time T, individuals who have at least one descendent at time T can be found with positive probability. We call these individuals T-prolific.

The SDE approach provides a common framework for the parametric family of decompositions of Duquesne and Winkel, as well as for the time-inhomogeneous decompositions we get, when we decompose the process along its T-prolific genealogies. We note that these finite-horizon decompositions exist for both supercritical and subcritical process. In the subcritical case the SDE representation can be used to observe the behaviour of the system when we condition on survival up to time T, then take T to infinity. Conditioning a subcritical CSBP to survive eternally results in what is known as a spine decomposition, where independent copies of the original process are grafted on to one infinite line of descent, that we call the spine (for more details, we refer the reader to [32, 24, 25, 17, 1]). And indeed, in [15] we see how the skeletal representation becomes, in the sense of weak convergence, a spinal decomposition when conditioning on survival, and in particular how the skeleton thins down to become the spine as $T \to \infty$.

In this paper our objective is to demonstrate the robustness of this aforementioned method by expanding the SDE approach to the spatial setting of superprocesses. We consider supercritical superprocesses with space dependent branching mechanism, but in future work we hope to extend results to the time-inhomogeneous case of subcritical processes.

The rest of this paper is organised as follows. In the remainder of this section we introduce our model and fix some notation. Then in Section 3.2 we remind the reader of some key existing results relevant to the subsequent exposition, in particular we recall the details of the skeletal decomposition of superprocesses with spatially dependent branching mechanism, as appeared in [22] and [9]. The main result of the paper is stated in Section 3.3, where we reformulate the result of Section 3.2 by writing down a coupled SDE, whose second coordinate corresponds to the skeletal process, while the first coordinate describes the evolution of the total mass in system. Finally, in Section 3.4 we give the proof of our results.

Superprocess. Let E be a domain of $\mathbb{R}^d$, and denote by $\mathcal{M}(E)$ the space of finite Borel measures on E. Furthermore let $\mathcal{M}(E)^\circ := \mathcal{M}(E) \setminus \{0\}$, where 0 is the null measure. We are interested in a strong Markov process X on E taking values in $\mathcal{M}(E)$. The process is characterised by two quantities $\mathcal{P}$ and ψ. Here $(\mathcal{P}_t)_{t\geq 0}$ is the semigroup of an $\mathbb{R}^d$-valued diffusion killed on exiting E, and ψ is the so-called branching mechanism. The latter takes the form

$$\psi(x,z) = -\alpha(x)z + \beta(x)z^2 + \int_{(0,\infty)} \left(\mathrm{e}^{-zu} - 1 + zu\right) m(x, \mathrm{d}u), \quad x \in E, \ z \geq 0, \quad (3.1)$$

where α and $\beta \geq 0$ are bounded measurable mappings from E to $\mathbb{R}$ and $[0,\infty)$ respectively, and $(u \wedge u^2)m(x, \mathrm{d}u)$ is a bounded kernel from E to $(0,\infty)$.

For technical reasons we assume that $\mathcal{P}$ is a Feller semigroup whose generator takes the form

$$\mathcal{L} = \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{d} b_i(x)\frac{\partial}{\partial x_i}, \tag{3.2}$$

where $a : E \to \mathbb{R}^{d\times d}$ is the diffusion matrix that takes values in the set of symmetric, positive definite matrices, and $b : E \to \mathbb{R}^d$ is the drift term.

Then the one-dimensional distributions of X can be characterised as follows. For all $\mu \in \mathcal{M}(E)$ and $f \in B^+(E)$, where $B^+(E)$ denotes the non-negative measurable functions on E, we have

$$\mathbb{E}_\mu\left[\mathrm{e}^{-\langle f, X_t\rangle}\right] = \exp\left\{-\int_E u_f(x,t)\mu(\mathrm{d}x)\right\}, \quad t \geq 0,$$

where $u_f(x,t)$ is the unique non-negative, locally bounded solution to the integral equation

$$u_f(x,t) = \mathcal{P}_t[f](x) - \int_0^t \mathrm{d}s \cdot \mathcal{P}_s[\psi(\cdot, u_f(\cdot, t-s))](x), \quad x \in E,\ t \geq 0. \tag{3.3}$$

Here we use the notation

$$\langle f, \mu\rangle = \int_E f(x)\mu(\mathrm{d}x), \quad \mu \in \mathcal{M}(E),\ f \in B^+(E).$$

For each $\mu \in \mathcal{M}(E)$ we denote by $\mathbb{P}_\mu$ the law of the process X issued from $X_0 = \mu$. The process $(X, \mathbb{P}_\mu)$ is called a $(\mathcal{P}, \psi)$-superprocess.

For more details on the above see Fitzsimmons [18]; for a general overview on superprocesses we refer the reader to the books of Dynkin [6, 7], Etheridge [12], Le Gall [26] and Li [27].

Next, we recall the SDE representation of $(X, \mathbb{P}_\mu)$ (for more details see Chapter 7 of [27]). To this end let us first define $H(x, \mathrm{d}\nu)$ as the natural extension of m from $(0,\infty)$ to $\mathcal{M}(E)^\circ$. More precisely H is concentrated on measures of the form $\{u\delta_x\}$ and it satisfies the integrability condition

$$\sup_{x\in E}\int_{\mathcal{M}(E)^\circ} (\nu(1) \wedge \nu(1)^2) H(x, \mathrm{d}\nu) < \infty.$$

Let $C_0(E)^+$ denote the space of non-negative continuous functions on E vanishing at infinity. We assume that α and β are continuous, furthermore $x \mapsto (\nu(1)\wedge\nu(1)^2)H(x, \mathrm{d}\nu)$ is continuous in the sense of weak convergence on $\mathcal{M}(E)^\circ$, and

$$f \mapsto \int_{\mathcal{M}(E)^\circ} (\nu(f) \wedge \nu(f)^2) H(x, \mathrm{d}\nu).$$

maps $C_0(E)^+$ into itself.

Next let $N(\mathrm{d}s,\mathrm{d}\nu)$ be the optional random measure on $[0,\infty)\times\mathcal{M}(E)$ defined by

$$N(\mathrm{d}s,\mathrm{d}\nu)=\sum_{s>0}\mathbf{1}_{\{\Delta X_s\neq 0\}}\delta_{(s,\Delta X_s)}(\mathrm{d}s,\mathrm{d}\nu),$$

where $\Delta X_s = X_s - X_{s-}$, and let $\hat{N}(\mathrm{d}s,\mathrm{d}\nu)$ denote the predictable compensator of $N(\mathrm{d}s,\mathrm{d}\nu)$. It can be shown that $\hat{N}(\mathrm{d}s,\mathrm{d}\nu)=\mathrm{d}sK(X_{s-},\mathrm{d}\nu)$ with

$$K(\mu,\mathrm{d}\nu)=\int_E \mu(\mathrm{d}x)H(x,\mathrm{d}\nu).$$

If we denote the compensated measure by $\tilde{N}(\mathrm{d}s,\mathrm{d}\nu)$, then for any $f\in D_0(\mathcal{L})$ (the set of functions in $C_0(E)$ that are also in the domain of $\mathcal{L}$) we have

$$X_t(f)=X_0(f)+M_t^c(f)+M_t^d(f)+\int_0^t X_s(\mathcal{L}f+\alpha f)\mathrm{d}s,\quad t\geq 0, \tag{3.4}$$

where $t\mapsto M_t^c(f)$ is a continuous local martingale with quadratic variation $2X_{t-}(\beta f^2)\mathrm{d}t$ and

$$t\mapsto M_t^d(f)=\int_0^t\int_{\mathcal{M}(E)^\circ}\nu(f)\tilde{N}(\mathrm{d}s,\mathrm{d}\nu),\quad t\geq 0,$$

is a purely discontinuous local martingale.

The representation (3.4) is what we will use in Section 3.3 when developing the SDE approach to the skeletal decomposition of $(X,\mathbb{P}_\mu)$. However before we could proceed with this line of analysis, we first need to recall the details of this skeletal decomposition, as it not only motivates our results, but also proves to be helpful in understanding the structure of our SDE.

3.2 Skeletal decomposition

Recall, that the main idea behind the skeletal decomposition is that under certain conditions we can identify prolific genealogies in the population, and by immigrating non-prolific mass along the trajectories of these prolific genealogies we can recover the law of the original superprocess. The infinite genealogies are described by a Markov branching process whose initial state is given by a Poisson random measure, while traditionally the immigrants are independent copies of the original process conditioned to become extinct.

In this section we first characterise the two components, then explain how to construct the skeletal decomposition from these building blocks. The results of this section are lifted from [22] and [9].

As we have mentioned the skeleton is often constructed using the event of extinction, that is the event $\mathcal{E}_{\text{fin}}=\{\langle 1,X_t\rangle=0 \text{ for some } t>0\}$. This guides the skeleton particles into regions where the survival probability is high. If we write $w(x) = -\log \mathbb{P}_{\delta_x}(\mathcal{E}_{\text{fin}})$, and assume that $\mu \in \mathcal{M}(E)$ is such that $\langle w, \mu\rangle < \infty$, then it is not hard to see that

$$\mathbb{P}_\mu(\mathcal{E}_{\text{fin}}) = \exp\left\{-\int_E w(x)\mu(\mathrm{d}x)\right\}.$$

Furthermore, by conditioning $\mathcal{E}_{\text{fin}}$ on $\mathcal{F}_t := \sigma(X_s, s \le t)$ we get that

$$\mathbb{E}_\mu\left(\mathrm{e}^{-\langle w, X_t\rangle}\right) = \mathrm{e}^{-\langle w,\mu\rangle}.$$

In [9] the authors point out that in order to construct a skeletal decomposition along those genealogies that avoid the behaviour specified by w (in this case 'extinction'), all we need is that the function w gives rise to a multiplicative martingale $\left(\left(\mathrm{e}^{-\langle w, X_t\rangle}\right)_{t\ge 0}, \mathbb{P}_\mu\right)$. In particular, a skeletal decomposition is given for any choice of a martingale function w which satisfies the following conditions.

- For all $x \in E$ we have $w(x) > 0$ and $\sup_{x\in E} w(x) < \infty$, and
- $\mathbb{E}_\mu\left(\mathrm{e}^{-\langle w, X_t\rangle}\right) = \mathrm{e}^{-\langle w,\mu\rangle}$ for all $\mu \in \mathcal{M}_c(E),\ t \ge 0$. (Here $\mathcal{M}_c(E)$ denotes the set of finite, compactly supported measures on E).

The condition $w(x) > 0$ implicitly hides the notion of supercriticality, as it ensures that survival happens with positive probability. Note however that 'survival' can be interpreted in many different ways. For example, the choice of $\mathcal{E}_{\text{fin}}$ results in skeleton particles that are simply part of some infinite genealogical line of descent, but we could also define surviving genealogies as those who visit a compact domain in E infinitely often.

Remark 3.1. *Note, that the authors in [22] and [9] show the existence of the skeletal decomposition under a slightly more general setup, where w is only locally bounded from above. Our SDE approach however forces us to be a bit more restrictive, and assume global boundedness.*

For reasons that will become apparent in the next section we make the additional assumption that w is in the domain of the generator $\mathcal{L}$.

Skeleton. First we identify the branching particle system that takes the role of the skeleton in the decomposition of the superprocess. In general, a Markov branching process $Z = (Z_t, t \ge 0)$ takes values in $\mathcal{M}_a(E)$ (the set of finite, atomic measures in E), and it can be characterised by the pair $(\mathcal{P}, F)$, where $\mathcal{P}$ is the semigroup of a diffusion and F is the branching generator which takes the form

$$F(x, s) = q(x)\sum_{n\ge 0} p_n(x)(s^n - s), \quad x \in E,\ s \in [0, 1].$$

Here q is a bounded, measurable mapping from E to $[0, \infty)$, and $\{p_n(x), n \ge 0\}$, $x \in E$ are measurable sequences of probability distributions. For $\nu \in \mathcal{M}_a(E)$ we denote the law of the process Z issued from ν by $\mathbf{P}_\nu$. Then we can describe $(Z, \mathbf{P}_\nu)$ as follows. We start with initial state $Z_0 = \nu$. Particles move according to $\mathcal{P}$, and at a spatially dependent rate $q(x)\mathrm{d}t$ a particle is killed and is replaced by n offspring with probability $p_n(x)$. The offspring particles then behave independently and according to the same law as their parent.

In order to specify the parameters of Z we first need to introduce some notation. Let $\xi = (\xi_t, t \geq 0)$ be the diffusion process on $E \cup \{\dagger\}$ (the one-point compactification of E with a cemetery state) corresponding to $\mathcal{P}$, and let us denote its probabilities by $\{\Pi_x, x \in E\}$. (Note that the previously defined martingale function w can be extended to $E \cup \{\dagger\}$ by defining $w(\dagger) = 0$). Then for all $x \in E$

$$\frac{w(\xi_t)}{w(x)} \exp\left\{-\int_0^t \frac{\psi(\xi_s, w(\xi_s))}{w(\xi_s)} \mathrm{d}s\right\}, \quad t \geq 0,$$

is a positive local martingale, and hence a supermartingale. (To see why this is true we refer the reader to the discussion in Section 2.1.1. of [9]). Now let $\tau_E = \inf\{t > 0 : \xi_t \in \{\dagger\}\}$, and consider the following change of measure

$$\left.\frac{\mathrm{d}\Pi_x^w}{\mathrm{d}\Pi_x}\right|_{\sigma(\xi_s, s\in[0,t])} = \frac{w(\xi_t)}{w(x)} \exp\left\{-\int_0^t \frac{\psi(\xi_s, w(\xi_s))}{w(\xi_s)} \mathrm{d}s\right\}, \quad \text{on } \{t < \tau_E\},\ x \in E, \quad (3.5)$$

which uniquely determines a family of (sub)probability measures $\{\Pi_x^w, x \in E\}$ (see for example [13]).

If we denote by $\mathcal{P}^w$ the semigroup of the $E \cup \{\dagger\}$ valued process whose probabilities are $\{\Pi^w, x \in E\}$, then it can be shown that the generator corresponding to $\mathcal{P}^w$ is given by

$$\mathcal{L}^w := \mathcal{L}_0^w - w^{-1}\mathcal{L}w = \mathcal{L}_0^w - w^{-1}\psi(\cdot, w),$$

where $\mathcal{L}_0^w u = w^{-1}\mathcal{L}(wu)$ whenever u is in the domain of $\mathcal{L}$. Note that $\mathcal{L}^w$ is also called an h-transform of the generator $\mathcal{L}$ with $h = w$. The theory of h-transforms for measure-valued diffusions was developed in [10].

Intuitively if

$$w(x) = -\log \mathbb{P}_{\delta_x}(\mathcal{E}) \quad (3.6)$$

defines a martingale function with the previously introduced conditions for some tail event $\mathcal{E}$, then the motion associated to $\mathcal{L}^w$ forces the particles to avoid the behaviour specified by $\mathcal{E}$. In particular when $\mathcal{E} = \mathcal{E}_{\text{fin}}$ then $\mathcal{P}^w$ encourages ξ to visit domains where the global survival rate is high.

Now we can characterise the skeleton process of $(X, \mathbb{P}_\mu)$ associated to w. In particular, $Z = (Z_t, t \geq 0)$ is a Markov branching process with diffusion semigroup $\mathcal{P}^w$ and branching generator

$$F(x, s) = q(x) \sum_{n\geq 0} p_n(x)(s^n - s), \quad x \in E,\ s \in [0, 1],$$

where

$$q(x) = \psi'(x, w(x)) - \frac{\psi(x, w(x))}{w(x)}, \quad (3.7)$$

and $p_0(x) = p_1(x) = 0$, and for $n \geq 2$

$$p_n(x) = \frac{1}{w(x)q(x)} \left\{ \beta(x) w^2(x) \mathbf{1}_{\{n=2\}} + w^n(x) \int_{(0,\infty)} \frac{y^n}{n!} \mathrm{e}^{-w(x)y} m(x, \mathrm{d}y) \right\}. \tag{3.8}$$

Here we used the notation

$$\psi'(x, w(x)) := \left.\frac{\partial}{\partial z} \psi(x, z)\right|_{z=w(x)}, \quad x \in E.$$

We refer to the process Z as the $(\mathcal{P}^w, F)$ skeleton.

Immigration. Next we characterise the process that we immigrate along the previously introduced branching particle system. To this end let us define the following function

$$\psi^*(x, z) = \psi(x, z + w(x)) - \psi(x, w(x)), \quad x \in E,$$

which can be written as

$$\psi^*(x, z) = -\alpha^*(x) z + \beta(x) z^2 + \int_{(0,\infty)} (\mathrm{e}^{-zu} - 1 + zu) m^*(x, \mathrm{d}u), \quad x \in E, \tag{3.9}$$

where

$$\alpha^*(x) = \alpha(x) - 2\beta(x) w(x) - \int_{(0,\infty)} (1 - \mathrm{e}^{-w(x)u}) u \, m(x, \mathrm{d}u) = -\psi'(x, w(x)),$$

and

$$m^*(x, \mathrm{d}u) = \mathrm{e}^{-w(x)u} m(x, \mathrm{d}u).$$

Note that under our assumptions ψ^* is a branching mechanism of the form (3.1). We denote the probabilities of the $(\mathcal{P}, \psi^*)$-superprocess by $(\mathbb{P}^*_\mu)_{\mu \in \mathcal{M}(E)}$.

If $\mathcal{E}$ is the event associated with w (see (3.6)), and $\langle w, \mu \rangle < \infty$, then we have

$$\mathbb{P}^*_\mu(\cdot) = \mathbb{P}_\mu(\cdot | \mathcal{E}).$$

In particular, when $\mathcal{E} = \mathcal{E}_{\text{fin}}$, then $\mathbb{P}^*_\mu$ is the law of the superprocess conditioned to become extinct.

Skeletal path decomposition. Here we give the precise construction of the skeletal decomposition that we introduced in a heuristic way at the beginning of this section. Let $\mathbb{D}([0, \infty) \times \mathcal{M}(E))$ denote the space of measure valued càdlàg function. Suppose that $\mu \in \mathcal{M}(E)$, and let Z be a $(\mathcal{P}^w, F)$-Markov branching process with initial configuration consisting of a Poisson random field of particles in E with intensity $w(x)\mu(\mathrm{d}x)$. Next, dress the branches of the spatial tree that describes the trajectory of Z in such a way that a particle at the space-time position $(x, t) \in E \times [0, \infty)$ has an independent $\mathbb{D}([0, \infty) \times \mathcal{M}(E))$-valued trajectory grafted on to it with rate

$$2\beta(x) \mathrm{d}\mathbb{Q}^*_x + \int_{(0,\infty)} y \mathrm{e}^{-w(x)y} m(x, \mathrm{d}y) \times \mathrm{d}\mathbb{P}^*_{y\delta_x}. \tag{3.10}$$

Here $\mathbb{Q}_x^*$ is the excursion measure on the space $\mathbb{D}([0,\infty)\times\mathcal{M}(E))$ which satisfies

$$\mathbb{Q}_x^*\left(1-\mathrm{e}^{-\langle f,X_t\rangle}\right)=u_f^*(x,t)$$

for $x\in E,\ t\geq 0$ and $f\in B_b^+(E)$ (the space of non-negative, bounded measurable functions on E), where $u_f^*(x,t)$ is the unique solution to (3.3) with the branching mechanism ψ replaced by ψ^*. (For more details on excursion measures see [8]). Moreover, when a particle in Z dies and gives birth to $n\geq 2$ offspring at spatial position $x\in E$, with probability $\eta_n(x,\mathrm{d}y)\mathbb{P}^*_{y\delta_x}$ an additional independent $\mathbb{D}([0,\infty)\times\mathcal{M}(E))$-valued trajectory is grafted on to the space-time branching point, where

$$\eta_n(x,\mathrm{d}y)=\frac{1}{w(x)q(x)p_n(x)}\left\{\beta(x)w^2(x)\delta_0(\mathrm{d}y)\mathbf{1}_{\{n=2\}}+w^n(x)\frac{y^n}{n!}\mathrm{e}^{-w(x)y}m(x,\mathrm{d}y)\right\}. \tag{3.11}$$

Note, that overall we have three different types of immigration processes that contribute to the dressing of the skeleton. In particular, the first term of (3.10) is what we call 'continuous immigration' along the skeleton, while the second term is referred to as the 'discontinuous immigration', and finally (3.11) corresponds to the so-called 'branch-point immigration'.

Now we define Λ_t as the total mass from the dressing present at time t together with the mass present at time t of an independent copy of $(X,\mathbb{P}^*_\mu)$ issued at time 0. We denote the law of (Λ,Z) by $\mathbf{P}_\mu$. Then in [22] the authors showed that $(\Lambda,\mathbf{P}_\mu)$ is Markovian and has the same law to $(X,\mathbb{P}_\mu)$. Furthermore, under $\mathbf{P}_\mu$, conditionally on Λ_t, the measure Z_t is a Poisson random measure with intensity $w(x)\Lambda_t(\mathrm{d}x)$.

3.3 SDE representation of the dressed tree

Recall that our main motivation is to reformulate the skeletal decomposition of superprocesses using the language of SDEs. Thus in this section, after giving an SDE representation of the skeletal process, we derive the coupled SDE for the dressed skeleton, which simultaneously describes the evolution of the skeleton and the total mass in the system.

SDE of the skeleton. We use the arguments on page 3 of [34] to derive the SDE for the branching particle diffusion, that will act as the skeleton. Let $(\xi_t,t\geq 0)$ be the diffusion process corresponding to $\mathcal{P}$. Since the generator of the motion is given by (3.2), the process ξ satisfies

$$\mathrm{d}\xi_t=b(\xi_t)\mathrm{d}t+\sigma(\xi_t)\mathrm{d}B_t,$$

where $\sigma:\mathbb{R}^d\to\mathbb{R}^d$ is such that $\sigma(x)\sigma^T(x)=a(x)$, and $(B_t,t\geq 0)$ is a d-dimensional Brownian motion (see for example Chapter 1 of [30]).

It is easy to verify that if $(\tilde{\xi}_t,t\geq 0)$ is the diffusion process under $\mathcal{P}^w$, then it satisfies

$$\mathrm{d}\tilde{\xi}_t=\left(b(\tilde{\xi}_t)+\frac{\nabla w(\tilde{\xi}_t)}{w(\tilde{\xi}_t)}a(\tilde{\xi}_t)\right)\mathrm{d}t+\sigma(\tilde{\xi}_t)\mathrm{d}B_t,$$

where ∇w is the gradient of w. To simplify computations, define the function $\tilde{b}$ on E given by

$$\tilde{b}(x) := b(x) + \frac{\nabla w(x)}{w(x)} a(x).$$

For $h \in C_b^2(E)$ (the space of bounded, twice differentiable continuous functions on E), using Itô's formula (see e.g. section 8.3 of [29]) we get

$$\mathrm{d}h(\tilde{\xi}_t) = (\nabla h(\tilde{\xi}_t))^{\mathtt{t}} \tilde{b}(\tilde{\xi}_t)\mathrm{d}t + \frac{1}{2}\mathrm{Tr}\left[\sigma^{\mathtt{t}}(\tilde{\xi}_t) H_h(\tilde{\xi}_t)\sigma(\tilde{\xi}_t)\right]\mathrm{d}t + (\nabla h(\tilde{\xi}_t))^{\mathtt{t}}\sigma(\tilde{\xi}_t)\mathrm{d}B_t,$$

where $x^{\mathtt{t}}$ denotes the transpose of x, Tr is the trace operator, and H_h is the Hessian of h with respect to x, that is $H_h(x)_{i,j} = \frac{\partial^2}{\partial x_i \partial x_j} h(x)$.

Next, summing over all the particles alive at time t, the collection of which we denote by $\mathcal{I}_t$, gives

$$\mathrm{d}\langle h, Z_t\rangle = \left\langle \nabla h(\cdot)\cdot\tilde{b}(\cdot), Z_t \right\rangle \mathrm{d}t + \left\langle \frac{1}{2}\mathrm{Tr}\left[\sigma^{\mathtt{t}}(\cdot)H_h(\cdot)\sigma(\cdot)\right], Z_t \right\rangle \mathrm{d}t + \sum_{\alpha\in\mathcal{I}_t} (\nabla h(\xi_t^\alpha))^{\mathtt{t}}\sigma(\xi_t^\alpha)\mathrm{d}B_t^\alpha, \tag{3.12}$$

where for each α, B^α is an independent copy of B.

If an individual branches at time t then we have

$$\langle h, Z_t - Z_{t-}\rangle = \sum_{\alpha:\text{death time of }\alpha = t} (k_\alpha - 1)h(\xi_t^\alpha). \tag{3.13}$$

Here k_α is the number of children of individual α, which has distribution $\{p_k, k = 0, 1, \dots\}$.

Simple algebra shows that

$$\mathrm{Tr}\left[\sigma^{\mathtt{t}}(x)H_h(x)\sigma(x)\right] = \sum_{ij} a_{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j} h(x).$$

Thus if we denote by $N^\dagger(\mathrm{d}s, \mathrm{d}\rho)$ the optional random measure

$$N^\dagger(\mathrm{d}s, \mathrm{d}\rho) = \sum_{s>0} \mathbf{1}_{\{\Delta Z_s \neq 0\}}\delta_{(s,\Delta Z_s)}(\mathrm{d}s, \mathrm{d}\rho),$$

on $[0,\infty)\times\mathcal{M}_a(E)$, where $\Delta Z_s = Z_s - Z_{s-}$, then by combining (3.12) and (3.13) we get

$$\langle h, Z_t\rangle = \langle h, Z_0\rangle + \int_0^t \langle \mathcal{L}^w h, Z_s\rangle \mathrm{d}s + V_t^c + \int_0^t \int_{\mathcal{M}_a(E)} \langle h, \rho\rangle N^\dagger(\mathrm{d}s, \mathrm{d}\rho), \tag{3.14}$$

where V_t^c is a continuous local martingale given by

$$V_t^c = \int_0^t \sum_{\alpha\in\mathcal{I}_s} (\nabla h(\xi_s^\alpha))^{\mathtt{t}}\sigma(\xi_s^\alpha)\mathrm{d}B_s^\alpha, \tag{3.15}$$

and the predictable compensator of the optional random measure $N^\dagger$ is given by

$$\hat{N}^\dagger(\mathrm{d}s, \mathrm{d}\rho) = q(x)\mathrm{d}s \int_E Z_{s-}(\mathrm{d}x) p_k(x)\pi(x, \mathrm{d}k),$$

where q, $p_k(\mathrm{d}x)$ are given by (3.7), (3.8). And $\pi(x, \mathrm{d}k)$ takes the form $\#(\mathrm{d}(k-1))\delta_x$, by which we mean that $\langle h, \pi(x, \mathrm{d}k)\rangle = (k-1)h(x)$.

Note that from (3.15) it is easy to see that the quadratic variation of V_t^c is

$$\langle V^c\rangle_t = \int_0^t \sum_{\alpha\in\mathcal{I}_s} (\nabla h(\xi_s^\alpha))^{\mathsf{t}} \sigma(\xi_s^\alpha)\sigma(\xi_s^\alpha)^{\mathsf{t}} \nabla h(\xi_s^\alpha)\mathrm{d}s = \int_0^t \langle (\nabla h)^{\mathsf{t}} a \nabla h, Z_s\rangle \mathrm{d}s.$$

Thinning of the SDE. Now we will see how to modify the SDE given by (3.4) in order to separate out the different types of immigration processes. We use ideas developed in [15].

Recall that the SDE describing the superprocess $(X, \mathbb{P}_\mu)$ takes the following form

$$\begin{aligned}\langle f, X_t\rangle = \langle f, \mu\rangle + \int_0^t \langle \alpha f, X_s\rangle \mathrm{d}s + M_t^c(f) + \int_0^t \int_{\mathcal{M}(E)^\circ} \langle f, \nu\rangle \tilde{N}(\mathrm{d}s, \mathrm{d}\nu)\\ + \int_0^t \langle \mathcal{L}f, X_s\rangle \mathrm{d}s, \quad t \geq 0.\end{aligned} \tag{3.16}$$

Here $M_t^c(f)$ is as in (3.4), and $\tilde{N}$ is an optional random measure on $[0,\infty)\times\mathcal{M}(E)^\circ$ with predictable compensator $\hat{N}(\mathrm{d}s, \mathrm{d}\nu) = \mathrm{d}s K(X_{s-}, \mathrm{d}\nu)$, where

$$\int_{\mathcal{M}(E)^\circ} \langle f, \nu\rangle K(\mu, \mathrm{d}\nu) = \int_E \mu(\mathrm{d}x) \int_{(0,\infty)} \langle f, u\delta_x\rangle m(x, \mathrm{d}u).$$

Denote by $(s_i, \nu_i : i \in \mathbb{N})$ some enumeration of the atoms of N. Next we introduce independent marks to the atoms of N, that is we define the random measure

$$\mathcal{N}(\mathrm{d}s, \mathrm{d}\nu, \mathrm{d}k) = \sum_{i\in\mathbb{N}} \delta_{(s_i,\nu_i,k_i)}(\mathrm{d}s, \mathrm{d}\nu, \mathrm{d}k),$$

whose compensator $\mathrm{d}s\mathcal{K}(X_{s-}, \mathrm{d}\nu, \mathrm{d}k)$ is given by

$$\int_{\mathcal{M}(E)^\circ} \langle f, \nu\rangle \mathcal{K}(\mu, \mathrm{d}\nu, \mathrm{d}k) = \int_E \mu(\mathrm{d}x) \int_{(0,\infty)} \langle f, u\delta_x\rangle \frac{(w(x)u)^k}{k!} \mathrm{e}^{-w(x)u} \#(\mathrm{d}k) m(x, \mathrm{d}u).$$

Now we can define three random measures by

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

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

and

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

Using Proposition (10.47) of [20] we see that N^0, N^1 and N^2 are also optional random measures and their compensators $\mathrm{d}sK^0(X_{s-},\mathrm{d}\nu)$, $\mathrm{d}sK^1(X_{s-},\mathrm{d}\nu)$ and $\mathrm{d}sK^2(X_{s-},\mathrm{d}\nu)$ satisfy

$$\int_{\mathcal{M}(E)^\circ}\langle f,\nu\rangle K^0(\mu,\mathrm{d}\nu) = \int_E \mu(\mathrm{d}x)\int_{(0,\infty)}\langle f,u\delta_x\rangle \mathrm{e}^{-w(x)u}m(x,\mathrm{d}u),$$

$$\int_{\mathcal{M}(E)^\circ}\langle f,\nu\rangle K^1(\mu,\mathrm{d}\nu) = \int_E \mu(\mathrm{d}x)\int_{(0,\infty)}\langle f,u\delta_x\rangle w(x)u\mathrm{e}^{-w(x)u}m(x,\mathrm{d}u)$$

and

$$\int_{\mathcal{M}(E)^\circ}\langle f,\nu\rangle K^2(\mu,\mathrm{d}\nu) = \int_E \mu(\mathrm{d}x)\int_{(0,\infty)}\langle f,u\delta_x\rangle \sum_{k=2}^{\infty}\frac{(w(x)u)^k}{k!}\mathrm{e}^{-w(x)u}m(x,\mathrm{d}u).$$

Using these processes we can rewrite (3.16), so we get

$$\begin{aligned}
\langle f,X_t\rangle &= \langle f,\mu\rangle + \int_0^t\langle \alpha f,X_s\rangle \mathrm{d}s + M_t^c(f) + \int_0^t\int_{\mathcal{M}(E)^\circ}\langle f,\nu\rangle \tilde{N}^0(\mathrm{d}s,\mathrm{d}\nu) + \int_0^t\langle \mathcal{L}f,X_s\rangle \mathrm{d}s\\
&\quad + \int_0^t\int_{\mathcal{M}(E)^\circ}\langle f,\nu\rangle N^1(\mathrm{d}s,\mathrm{d}\nu) + \int_0^t\int_{\mathcal{M}(E)^\circ}\langle f,\nu\rangle N^2(\mathrm{d}s,\mathrm{d}\nu)\\
&\quad - \int_0^t\left\langle \int_{(0,\infty)} uf(\cdot)\left(1-\mathrm{e}^{-uw(\cdot)}\right)m(\cdot,\mathrm{d}u), X_{s-}\right\rangle \mathrm{d}s\\
&= \langle f,\mu\rangle - \int_0^t\langle \psi'(\cdot,w(\cdot,s))f(\cdot),X_s\rangle \mathrm{d}s + M_t^c(f) + \int_0^t\int_{\mathcal{M}(E)^\circ}\langle f,\nu\rangle \tilde{N}^0(\mathrm{d}s,\mathrm{d}\nu)\\
&\quad + \int_0^t\langle \mathcal{L}f,X_s\rangle \mathrm{d}s + \int_0^t\int_{\mathcal{M}(E)^\circ}\langle f,\nu\rangle N^1(\mathrm{d}s,\mathrm{d}\nu)\\
&\quad + \int_0^t\int_{\mathcal{M}(E)^\circ}\langle f,\nu\rangle N^2(\mathrm{d}s,\mathrm{d}\nu) + \int_0^t\langle 2\beta(\cdot)w(\cdot)f(\cdot),X_{s-}\rangle \mathrm{d}s,
\end{aligned} \tag{3.17}$$

where we have used the fact that $\alpha(x)-\int_{(0,\infty)}(1-\mathrm{e}^{-w(x)u})um(x,\mathrm{d}u) = -\psi'(x,w(x)) + 2\beta(x)w(x)$. Recalling (3.9) we see that the first line of the right-hand side of (3.17) corresponds to the dynamics of a $(\mathcal{P},\psi^*)$-superprocess. Our aim now is to link the remaining three terms to the three types of immigration along the skeleton, and write down a system of SDEs that describe the skeleton and total mass simultaneously. Heuristically this system of SDEs will combine the appropriate terms of equations (3.14) and (3.17).

Coupled SDE. Following the ideas of the previous sections we introduce the following independent driving sources of randomness that we will use in the construction of our coupled SDE.

- Let $\mathtt{N}^0_\mu(\mathrm{d}s,\mathrm{d}\nu)$ be an optional random measure on $[0,\infty)\times\mathcal{M}(E)^\circ$ for some $\mu\in\mathcal{M}(E)$, whose predictable compensator can be written as $\hat{\mathtt{N}}^0_\mu(\mathrm{d}s,\mathrm{d}\nu) =$

$\mathrm{d}s\mathtt{K}^0(\mathrm{d}\mu, \mathrm{d}\nu)$, where

$$\mathtt{K}^0(\mathrm{d}\mu, \mathrm{d}\nu) = \int_E \mu(\mathrm{d}x)\mathrm{e}^{-w(x)u} m(x, \mathrm{d}u). \tag{3.18}$$

Let $\tilde{\mathtt{N}}^0_\mu(\mathrm{d}s, \mathrm{d}\nu)$ be its compensated version,

- $\mathtt{N}^1_\tau(\mathrm{d}s, \mathrm{d}\nu)$ be an optional random measure on $[0, \infty) \times \mathcal{M}(E)^\circ$ for some $\tau \in \mathcal{M}_a(E)$, whose predictable compensator can be written as $\hat{\mathtt{N}}^1_\tau(\mathrm{d}s, \mathrm{d}\nu) = \mathrm{d}s\mathtt{K}^1(\mathrm{d}\tau, \mathrm{d}\nu)$, where

$$\mathtt{K}^1(\mathrm{d}\nu, \mathrm{d}\tau) = \int_E \tau(\mathrm{d}x)u\mathrm{e}^{-w(x)u} m(x, \mathrm{d}u), \tag{3.19}$$

- and $\mathtt{N}^2_\tau(\mathrm{d}s, \mathrm{d}\rho, \mathrm{d}\nu)$ an optional random measure on $[0, \infty) \times \mathcal{M}_a(E) \times \mathcal{M}(E)^\circ$ for some $\tau \in \mathcal{M}_a(E)$, whose predictable compensator can be written in the form $\hat{\mathtt{N}}^2_\tau(\mathrm{d}s, \mathrm{d}\rho, \mathrm{d}\nu) = \mathrm{d}s\mathtt{K}^2(\mathrm{d}\tau, \mathrm{d}\rho, \mathrm{d}\nu)$, where

$$\mathtt{K}^2(\mathrm{d}\tau, \mathrm{d}\rho, \mathrm{d}\nu) = q(x) \int_E \tau(\mathrm{d}x)p_k(x)\pi(x, \mathrm{d}k)\eta_k(x, \mathrm{d}u), \tag{3.20}$$

where q, $p_k(\mathrm{d}x)$ and $\eta_k(x, \mathrm{d}u)$ are given by (3.7), (3.8) and (3.11). And $\pi(x, \mathrm{d}k)$ takes the form $\#(\mathrm{d}(k-1))\delta_x$ by which, just as before, we mean that for a suitable test function h we have $\langle h, \pi(x, \mathrm{d}k)\rangle = (k-1)h(x)$.

Now we can state our main result.

Theorem 3.3.1. *Consider the following system of SDEs for $f, h \in D_0(\mathcal{L})$,*

$$\begin{aligned}
\begin{pmatrix} \langle f, \Lambda_t \rangle \\ \langle h, Z_t \rangle \end{pmatrix} &= \begin{pmatrix} \langle f, \Lambda_0 \rangle \\ \langle h, Z_0 \rangle \end{pmatrix} - \int_0^t \begin{pmatrix} \langle \psi'(\cdot, w(\cdot))f(\cdot), \Lambda_{s-} \rangle \\ 0 \end{pmatrix} \mathrm{d}s + \begin{pmatrix} U^c_t(f) \\ V^c_t(h) \end{pmatrix} \\
&\quad + \int_0^t \int_{\mathcal{M}(E)^\circ} \begin{pmatrix} \langle f, \nu \rangle \\ 0 \end{pmatrix} \tilde{\mathtt{N}}^0_{\Lambda_{s-}}(\mathrm{d}s, \mathrm{d}\nu) + \int_0^t \begin{pmatrix} \langle \mathcal{L}f, \Lambda_{s-} \rangle \\ \langle \mathcal{L}^w h, Z_{s-} \rangle \end{pmatrix} \mathrm{d}s \\
&\quad + \int_0^t \int_{\mathcal{M}(E)^\circ} \begin{pmatrix} \langle f, \nu \rangle \\ 0 \end{pmatrix} \mathtt{N}^1_{Z_{s-}}(\mathrm{d}s, \mathrm{d}\nu) \\
&\quad + \int_0^t \int_{\mathcal{M}_a(E)} \int_{\mathcal{M}(E)^\circ} \begin{pmatrix} \langle f, \nu \rangle \\ \langle h, \rho \rangle \end{pmatrix} \mathtt{N}^2_{Z_{s-}}(\mathrm{d}s, \mathrm{d}\rho, \mathrm{d}\nu) \\
&\quad + \int_0^t \begin{pmatrix} \langle 2\beta(\cdot)f(\cdot), Z_{s-} \rangle \\ 0 \end{pmatrix} \mathrm{d}s, \qquad t \geq 0,
\end{aligned} \tag{3.21}$$

where $\Lambda_0 \in \mathcal{M}(E)$ is given and fixed. In the above equation $(U^c_t(f), t \geq 0)$ is a continuous local martingale with quadratic variation $2\langle \beta f^2, \Lambda_{t-} \rangle \mathrm{d}t$, and $(V^c_t(h), t \geq 0)$ is a continuous local martingale with quadratic variation $\langle (\nabla h)^{\mathtt{t}} a \nabla h, Z_{t-} \rangle \mathrm{d}t$. Note, that in the optional random measure terms the randomness in the compensators are given by $\mu = \Lambda_{s-}$ for the term $\mathtt{N}^0$ (using the notation of (3.18)), and by $\tau = Z_{s-}$ for the terms $\mathtt{N}^1$ and $\mathtt{N}^2$ (using the notation of (3.19) and (3.20)).

Then equation (3.21) has a unique weak solution, and under the assumption that Z_0 is a Poisson random measure with intensity $w(x)\Lambda_0(\mathrm{d}x)$ we have the following:

(i) $Z_t|\mathcal{F}_t^\Lambda$ *(where* $\mathcal{F}_t^\Lambda = \sigma(\Lambda_s : s \leq t)$*) is a Poisson random measure with intensity* $w(x)\Lambda_t(\mathrm{d}x)$*;*

(ii) Conditional on $(\mathcal{F}_t^\Lambda, t \geq 0)$*, the process* $(\Lambda_t, t \geq 0)$ *is Markovian and a weak solution to (3.16).*

3.4 Sketch proof

In the proof we are working towards showing that the unique weak solution exists, and that

$$\mathbf{E}_\mu\left[\mathrm{e}^{-\langle f,\Lambda_t\rangle-\langle h,Z_t\rangle}\right] = \mathbb{E}_\mu\left[\mathrm{e}^{-\langle f+w(1-\mathrm{e}^{-h}),X_t\rangle}\right], \tag{3.22}$$

where X satisfies (3.16). We will see that from (3.22) claims (i) and (ii) follow easily.

The idea of this latter part is to fix $T > 0$ and $f, h \in D_0(\mathcal{L})$, and choose time-dependent test functions f^T and h^T in a way that the processes

$$F_t^T = \mathrm{e}^{-\langle f^T(\cdot,T-t),\Lambda_t\rangle-\langle h^T(\cdot,T-t),Z_t\rangle},$$

and

$$G_t^T = \mathrm{e}^{-\langle f^T(\cdot,T-t)+w(1-\mathrm{e}^{-h^T(\cdot,T-t)}),X_t\rangle},$$

have constant expectations on $[0, T]$. The test functions are defined as solutions to some partial differential equations with final value conditions $f^T(x, T) = f(x)$ and $h^T(x, T) = h(x)$. This, together with the fact that $\Lambda_0 = X_0 = \mu$, and that Z_0 is a Poisson random measure with intensity $w(x)\Lambda_0(\mathrm{d}x)$, then will give us (3.22).

Thus to prove the main result we need the existence of solutions of two differential equations. Recall from Section 3.2 that in the skeletal decomposition of superprocesses the total mass present at time t has two main components. The first one corresponds to an initial burst of subcritical mass, which is independent copy of $(X, \mathbb{P}_\mu^*)$, and the second one is the accumulated mass from the dressing of the skeleton. As we will see in the main part of the proof, one can associate the first differential equation, that is the equation defining f^T, to $(X, \mathbb{P}_\mu^*)$, while the equation defining h^T has an intimate relation to the dressed tree. To help with the presentation of the main ideas we first derive these equations, then give the proof of Theorem 3.3.1.

Preliminary results. First we consider the initial burst of subcritical mass. Recall that $(X, \mathbb{P}_\mu^*)$ is a $(\psi^*, \mathcal{P})$-superprocess, and as such its law can be characterised through an integral equation. More precisely, for all $\mu \in \mathcal{M}(E)$ and $f \in B^+(E)$, we have

$$\mathbb{E}_\mu\left[\mathrm{e}^{-\langle f,X_t^*\rangle}\right] = \exp\left\{-\int_E u_f^*(x,t)\mu(\mathrm{d}x)\right\}, \quad t \geq 0,$$

where $u_f^*(x, t)$ is the unique non-negative solution to the integral equation

$$u_f^*(x,t) = \mathcal{P}_t[f](x) - \int_0^t \mathrm{d}s \cdot \mathcal{P}_s[\psi^*(\cdot, u_f^*(\cdot, t-s))](x), \quad x \in E,\ t \geq 0. \tag{3.23}$$

Li (Theorem 7.11 of [27]) showed that this integral equation is equivalent to the following differential equation

$$\begin{aligned}\frac{\partial}{\partial t} u^*_f(x,t) &= \mathcal{L} u^*_f(x,t) - \psi^*(x, u^*_f(x,t)), \\ u^*_f(x,0) &= f(x).\end{aligned} \tag{3.24}$$

Thus (3.24) also has a unique non-negative solution.

As we will see later, what we need in the main proof is a time-reversed version of (3.24). In particular, if for each fixed $T > 0$ we define $f^T(x,t) = u^*_f(x, T-t)$, then it is not hard to see that the following lemma holds.

Lemma 3.4.1. *Fix $T > 0$, and let $f \in D_0(\mathcal{L})$. Then the following differential equation has a unique non-negative solution*

$$\begin{aligned}\frac{\partial}{\partial t} f^T(x,t) &= -\mathcal{L} f^T(x,t) + \psi^*(x, f^T(x,t)), \quad 0 \le t \le T, \\ f^T(x,T) &= f(x),\end{aligned} \tag{3.25}$$

where ψ^* is given by (3.9).

In order to get the differential equation that defines h^T we need to derive similar results for the dressed tree. To this end consider the pair (Λ, Z), where Z is a $(\mathcal{P}^w, F)$ branching Markov process with $Z_0 = \nu$ for some $\nu \in \mathcal{M}_a(E)$, and whose jumps are coded by the coordinates of the random measure $\mathtt{N}^2$. Furthermore we define $\Lambda_t = X^*_t + D_t$, where X^* is an independent copy of the $(\mathcal{P}, \psi^*)$-superprocess with initial value $X^*_0 = \mu$, $\mu \in \mathcal{M}(E)$, and the process $(D_t, t \ge 0)$ is described by

$$\begin{aligned}\langle f, D_t\rangle = &\int_0^t \int_{\mathcal{M}(E)^\circ} \int_{\mathbb{D}([0,\infty)\times\mathcal{M}(E))} \langle f, \omega_{t-s}\rangle \mathtt{N}^1(\mathrm{d}s, \mathrm{d}\nu, \mathrm{d}\omega) \\ &+ \int_0^t \int_{\mathcal{M}_a(E)} \int_{\mathcal{M}(E)^\circ} \int_{\mathbb{D}([0,\infty)\times\mathcal{M}(E))} \langle f, \omega_{t-s}\rangle \mathtt{N}^2(\mathrm{d}s, \mathrm{d}\rho, \mathrm{d}\nu, \mathrm{d}\omega) \\ &+ \int_0^t \int_{\mathcal{M}_a(E)} \int_{\mathbb{D}([0,\infty)\times\mathcal{M}(E))} \langle f, \omega_{t-s}\rangle \mathtt{N}^*(\mathrm{d}s, \mathrm{d}\rho, \mathrm{d}\omega),\end{aligned} \tag{3.26}$$

where $f \in D_0(\mathcal{L})$ and with a slight abuse of notation

- $\mathtt{N}^1$ is an optional random measure on $[0,\infty) \times \mathcal{M}(E)^\circ \times \mathbb{D}([0,\infty) \times \mathcal{M}(E))$ whose predictable compensator can be written as $\hat{\mathtt{N}}^1(\mathrm{d}s, \mathrm{d}\nu, \mathrm{d}\omega) = \mathrm{d}s \mathtt{K}^1(\mathrm{d}Z_{s-}, \mathrm{d}\nu, \mathrm{d}\omega)$ where
$$\mathtt{K}^1(\mathrm{d}\tau, \mathrm{d}\nu, \mathrm{d}\omega) = \int_E \tau(\mathrm{d}x) u \mathrm{e}^{-w(x)u} m(x, \mathrm{d}u) \mathbb{P}^*_{u\delta_x}(\mathrm{d}\omega),$$

- $\mathtt{N}^2$ is an optional random measure on $[0,\infty) \times \mathcal{M}_a(E) \times \mathcal{M}(E)^\circ \times \mathbb{D}([0,\infty) \times \mathcal{M}(E))$ whose predictable compensator can be written as $\hat{\mathtt{N}}^2(\mathrm{d}s, \mathrm{d}\rho, \mathrm{d}\nu, \mathrm{d}\omega) = \mathrm{d}s \mathtt{K}^2(\mathrm{d}Z_{s-}, \mathrm{d}\rho, \mathrm{d}\nu, \mathrm{d}\omega)$ where
$$\mathtt{K}^2(\mathrm{d}\tau, \mathrm{d}\rho, \mathrm{d}\nu, \mathrm{d}\omega) = \int_E q(x) \tau(\mathrm{d}x) p_k(x) \pi(x, \mathrm{d}k) \eta_k(x, \mathrm{d}u) \mathbb{P}^*_{u\delta_x}(\mathrm{d}\omega),$$

- $\mathtt{N}^*$ is an optional random measure on $[0,\infty)\times\mathcal{M}_a(E)\times\mathbb{D}([0,\infty)\times\mathcal{M}(E))$ whose predictable compensator can be written as $\hat{\mathtt{N}}^*(\mathrm{d}s,\mathrm{d}\rho,\mathrm{d}\omega)=\mathrm{d}s\mathtt{K}^*(\mathrm{d}Z_{s-},\mathrm{d}\rho,\mathrm{d}\omega)$

$$\mathtt{K}^*(\mathrm{d}\tau,\mathrm{d}\rho,\mathrm{d}\omega)=\int_E 2\beta(x)\mathrm{d}s\tau(\mathrm{d}x)\mathbb{Q}_x(\mathrm{d}\omega),$$

where

$$\mathbb{Q}(1-\mathrm{e}^{-\langle f,\omega_t\rangle})=-\log\mathbb{E}^*_{\delta_x}(\mathrm{e}^{-\langle f,X_t\rangle})=u^*_f(x,t).$$

Denote the probabilities of (Λ,Z) by $\mathbf{P}_{(\mu,\nu)}$. Then we have the following result, an equivalent of which also appeared in [2] and [22].

Proposition 3.4.1. *For every $\mu\in\mathcal{M}(E)$, $\nu\in\mathcal{M}_a(E)$ and $f,h\in B^+_b(E)$ we have*

$$\mathbf{E}_{(\mu,\nu)}\left[\mathrm{e}^{-\langle f,D_t\rangle-\langle h,Z_t\rangle}\right]=\mathrm{e}^{-\langle v_{f,h}(\cdot,t),\nu\rangle},$$

where $\exp\{-v_{f,h}(x,t)\}$ *is the unique $[0,1]$-valued solution to the following integral equation*

$$\begin{aligned}w(x)\mathrm{e}^{-v_{f,h}(x,t)}=&\,\mathcal{P}_t\left[w(\cdot)\mathrm{e}^{-h(\cdot)}\right](x)\\&+\int_0^t\mathrm{d}s\cdot\mathcal{P}_s\left[\psi^*\left(\cdot,-w(\cdot)\mathrm{e}^{-v_{f,h}(\cdot,t-s)}+u^*_f(\cdot,t-s)\right)-\psi^*(\cdot,u^*_f(\cdot,t-s))\right](x),\end{aligned}\tag{3.27}$$

and u^*_f is the unique non-negative solution to (3.23).

Since the proof of Proposition 3.4.1 is a straightforward adaptation of the proof of Theorem 2 in [22], we leave this to the reader.

Next we show that the integral equation (3.27) is equivalent to the following differential equation

$$\begin{aligned}\mathrm{e}^{-v_{f,h}(x,t)}w(x)\frac{\partial}{\partial t}v_{f,h}(x,t)=&-\mathcal{L}\left[w(\cdot)\mathrm{e}^{-v_{f,h}(\cdot,t)}\right](x)\\&-\left(\psi^*\left(x,-w(x)\mathrm{e}^{-v_{f,h}(x,t)}+u^*_f(x,t)\right)-\psi^*(x,u^*_f(x,t))\right),\end{aligned}\tag{3.28}$$

$$v_{f,h}(x,0)=h(x).$$

Lemma 3.4.2. *Let $f,h\in D_0(\mathcal{L})$. If $v_{f,h}$ is a solution to (3.27), then it also solves (3.28). Conversely, if $v_{f,h}$ solves (3.28), then it also satisfies (3.27).*

Proof. We first prove the claim that the integral equation implies the differential equation. To this end consider (3.27). Note that since $\mathcal{P}$ is a Feller semigroup the right hand side is differentiable in t, and thus $v_{f,h}(x,t)$ is also differentiable in t. To find the differential version of the equation, we can use the standard technique of propagating the derivative at zero using the semigroup property of $v_{f,h}$ and u^*_f. Indeed, on one hand the semigroup property can easily be verified using

$$\begin{aligned}
\mathbf{E}_{(\mu,\nu)}\left[\mathrm{e}^{-\langle f,\Lambda_{t+s}\rangle-\langle h,Z_{t+s}\rangle}\right] &= \mathbf{E}_{(\mu,\nu)}\left[\mathbf{E}\left[\mathrm{e}^{-\langle f,\Lambda_{t+s}\rangle-\langle h,Z_{t+s}\rangle}\middle|\mathcal{F}_s\right]\right]\\
&= \mathbf{E}_{(\mu,\nu)}\left[\mathbf{E}_{(\Lambda_s,Z_s)}\left[\mathrm{e}^{-\langle f,\Lambda_t\rangle-\langle h,Z_t\rangle}\right]\right]\\
&= \mathbf{E}_{(\mu,\nu)}\left[\mathrm{e}^{-\langle u^*_f(\cdot,t),\Lambda_s\rangle-\langle v_{f,h}(\cdot,t),Z_s\rangle}\right]\\
&= \mathrm{e}^{-\left\langle u^*_{u^*_f(\cdot,t)}(\cdot,s),\mu\right\rangle-\left\langle v_{u^*_f(\cdot,t),v_{f,h}(\cdot,t)}(\cdot,s),\nu\right\rangle},
\end{aligned}$$

that is we have $v_{u^*_f(\cdot,t),v_{f,h}(\cdot,t)}(\cdot,s) = v_{f,h}(\cdot,t+s)$, and $u^*_{u^*_f(\cdot,s)}(\cdot,t) = u^*_f(\cdot,t+s)$. This implies

$$\frac{\partial}{\partial t}u^*_f(x,t) = \frac{\partial}{\partial s}u_{u^*_f(\cdot,t)}(x,s)\bigg|_{s\downarrow 0} = \frac{\partial}{\partial s}u_{u^*_f(\cdot,t)}(x,0+),$$

and

$$\frac{\partial}{\partial t}v_{f,h}(\cdot,t) = \frac{\partial}{\partial s}v_{u^*_f(\cdot,t),v_{f,h}(\cdot,t)}(x,s)\bigg|_{s\downarrow 0}.$$

On the other hand differentiating in t and taking $t\downarrow 0$ gives

$$\begin{aligned}
-w(x)\mathrm{e}^{-v_{f,h}(x,0+)}\,\frac{\partial}{\partial t}v_{f,h}(x,t)\bigg|_{t=0+} =&\mathcal{L}\left[w(\cdot)\mathrm{e}^{-h(\cdot)}\right](x)\\
&+\psi^*\left(x,-w(x)\mathrm{e}^{-v_{f,h}(x,0+)}+u^*_f(x,0+)\right)\\
&-\psi^*(x,u^*_f(x,0+)).
\end{aligned}\tag{3.29}$$

which, recalling $v_{f,h}(x,0)=h(x)$ and $u^*_f(x,0)=f(x)$, can be rewritten as

$$\begin{aligned}
\frac{\partial}{\partial t}v_{f,h}(x,0+) =& -\frac{1}{w(x)}\mathrm{e}^{h(x)}\mathcal{L}\left[w(\cdot)\mathrm{e}^{-h(\cdot)}\right](x)\\
&-\frac{1}{w(x)}\mathrm{e}^{h(x)}\psi^*\left(x,-w(x)\mathrm{e}^{-h(x)}+f(x)\right)+\frac{1}{w(x)}\mathrm{e}^{h(x)}\psi^*(x,f(x)).
\end{aligned}$$

Hence combining the previous observations we get

$$\begin{aligned}
\frac{\partial}{\partial t}v_{f,h}(x,t) =& -\frac{1}{w(x)}\mathrm{e}^{v_{f,h}(x,t)}\mathcal{L}\left[w(\cdot)\mathrm{e}^{-v_{f,h}(x,t)}\right](x)\\
&-\frac{1}{w(x)}\mathrm{e}^{v_{f,h}(x,t)}\psi^*\left(x,-w(x)\mathrm{e}^{-v_{f,h}(x,t)}+u^*_f(x,t)\right)\\
&+\frac{1}{w(x)}\mathrm{e}^{v_{f,h}(x,t)}\psi^*(x,u^*_f(x,t)).
\end{aligned}$$

To see why the differential equation implies the integral equation define

$$g(x,s) = \mathcal{P}_{t-s}\left(w(x)\mathrm{e}^{-v_{f,h}(x,s)}\right),\quad 0\le s\le t.$$

Then differentiating with respect to the time parameter gives

$$\begin{aligned}\frac{\partial}{\partial s}g(x,s) &= -\mathcal{P}_{t-s}w(x)\mathrm{e}^{v_{f,h}(x,s)}\frac{\partial}{\partial s}v_{f,h}(x,s) - \mathcal{P}_{t-s}\mathcal{L}\left(w(x)\mathrm{e}^{-v_{f,h}(x,s)}\right)\\ &= \mathcal{P}_{t-s}\left[\psi^*\left(x,-w(x)\mathrm{e}^{-v_{f,h}(x,s)}+u_f^*(x,s)\right)-\psi^*(x,u_f^*(x,s))\right],\end{aligned}$$

which then we can integrate over $[0,t]$ to get (3.27). □

Finally, just as in the previous section, we fix $T>0$, and define $h^T(x,t) := v_{f,h}(x,T-t)$ to get the following result.

Lemma 3.4.3. *Fix $T>0$, and take $f,h\in D_0(\mathcal{L})$. If f^T if the unique solution to (3.25), then the following differential equation has a unique non-negative solution*

$$\begin{aligned}\mathrm{e}^{-h^T(x,t)}w(x)\frac{\partial}{\partial t}h^T(x,t) =&\mathcal{L}\left(w(x)\mathrm{e}^{-h^T(x,t)}\right)\\ &+\left(\psi^*\left(x,-w(x)\mathrm{e}^{-h^T(x,t)}+f^T(x,t)\right)-\psi^*(x,f^T(x,t))\right),\\ h^T(x,T)=&h(x),\end{aligned}\tag{3.30}$$

where ψ^* is given by (3.9), and w is a martingale function that satisfies the conditions in Section 3.2.

Proof of main result. Finally we give the proof of Theorem 3.3.1. The techniques we use here to prove parts (i) and (ii) of the theorem are similar in spirit to those in the proof of Theorem 2.1 in [15], in a sense that we use stochastic calculus to show the equality (3.22); however instead of arriving to a partial differential equation, we use stochastic analysis to prove the equality of expectations.

Proof. We start by showing that there exists a (weak) solution to the coupled SDE (3.21). Uniqueness will follow from later steps of the proof.

Take $f,h\in\mathcal{D}_0(\mathcal{L})$. To show the existence of solution we once again consider the process we constructed in the previous section. That is, we take Z to be a $(\mathcal{P}^w,F)$ branching Markov process with $Z_0=\nu$ for some $\nu\in\mathcal{M}_a(E)$, and whose jumps are coded by the coordinates of the random measure $\mathtt{N}^2$. Furthermore we define $\Lambda_t = X_t^*+D_t$, where X^* is an independent copy of the $(\mathcal{P},\psi^*)$-superprocess with initial value $X_0^*=\mu$, $\mu\in\mathcal{M}(E)$, and the process $(D_t, t\geq 0)$ is described as in equation (3.26). To see why the pair $(\langle f,\Lambda_t\rangle,\langle h,Z_t\rangle)$ satisfies (3.26), we may appeal to the Martingale representation of weak solutions (in the spirit of e.g Theorem 2.3 of [21], Theorem 7.13 of [27], or the arguments in the proof of Theorem 2.1 in [15]). The above pair is Markovian, and its generator can be identified as

$$\begin{aligned}
\mathcal{A}G(\langle f,\mu\rangle,\langle h,\nu\rangle) &= \langle -\psi'(\cdot,w(\cdot))f(\cdot)+\mathcal{L}f,\mu\rangle \,\frac{\partial}{\partial x_1}G(x_1,x_2)\bigg|_{x_1=\langle f,\mu\rangle,x_2=\langle h,\nu\rangle}\\
&+ \langle \beta f^2,\mu\rangle \,\frac{\partial^2}{\partial x_1^2}G(x_1,x_2)\bigg|_{x_1=\langle f,\mu\rangle,x_2=\langle h,\nu\rangle} + \langle 2\beta f,\nu\rangle \,\frac{\partial}{\partial x_1}G(x_1,x_2)\bigg|_{x_1=\langle f,\mu\rangle,x_2=\langle h,\nu\rangle}\\
&+ \langle \mathcal{L}^w h,\nu\rangle \,\frac{\partial}{\partial x_2}G(x_1,x_2)\bigg|_{x_1=\langle f,\mu\rangle,x_2=\langle h,\nu\rangle}\\
&+ \frac{1}{2}\left\langle (\nabla h)^{\mathrm{t}} a\nabla h,\nu\right\rangle \,\frac{\partial^2}{\partial x_2^2}G(x_1,x_2)\bigg|_{x_1=\langle f,\mu\rangle,x_2=\langle h,\nu\rangle}\\
&+ \int_E\left[\int_{(0,\infty)}\Bigg(G(\langle f,\mu+u\delta_x\rangle,\langle h,\nu\rangle)-G(\langle f,\mu\rangle,\langle h,\nu\rangle)\right.\\
&\qquad\left. -\,\langle f,u\delta_x\rangle\,\frac{\partial}{\partial x_1}G(x_1,x_2)\bigg|_{x_1=\langle f,\mu\rangle,x_2=\langle h,\nu\rangle}\Bigg)\mathrm{e}^{-w(x)u}m(x,\mathrm{d}u)\right]\mu(\mathrm{d}x)\\
&+ \int_E\left[\int_{(0,\infty)}\left(G(\langle f,\mu+u\delta_x\rangle,\langle h,\nu\rangle)-G(\langle f,\mu\rangle,\langle h,\nu\rangle)\right)u\mathrm{e}^{-w(x)u}m(x,\mathrm{d}u)\right]\nu(\mathrm{d}x)\\
&+ \int_E\left[\int_{(0,\infty)}\sum_{k=2}^{\infty}\left(G(\langle f,\mu+u\delta_x\rangle,\langle h,\nu+(k-1)\delta_x\rangle)\right)-G(\langle f,\mu\rangle,\langle h,\nu\rangle)\right.\\
&\qquad\left. w^{k-1}(x)\frac{u^k}{k!}\mathrm{e}^{-w(x)u}m(x,\mathrm{d}u)\right]\nu(\mathrm{d}x)\\
&+ \int_E \beta(x)w(x)\left(G(\langle f,\mu\rangle,\langle h,\nu+\delta_x\rangle)\right)\nu(\mathrm{d}x).
\end{aligned}$$

for $\mu\in\mathcal{M}(E)$, $\nu\in\mathcal{M}_a(E)$, and for all non-negative, smooth, compactly supported G. It is easy to see that this coincides with the generator of the SDE (3.21), hence we have identified a weak solution to (3.21).

Next we move onto the proof of parts (i) and (ii) of the theorem. Uniqueness of the solution will follow from these steps.

Fix $T>0$, and let f^T be the unique non-negative solution to (3.25), and h^T be the unique non-negative solution to (3.30). Define $F_t^T := \mathrm{e}^{-\langle f^T(\cdot,t),\Lambda_t\rangle-\langle h^T(\cdot,t),Z_t\rangle}$, $t\le T$. Using stochastic calculus, we first verify that our choice of f^T and h^T results in the process F_t^T, $t\le T$, having constant expectation on $[0,T]$. In the definition of F^T both $\langle f^T(\cdot,t),\Lambda_t\rangle$ and $\langle h^T(\cdot,t),Z_t\rangle$ are semi-martingales, thus we can use Itô's formula (see e.g. Theorem 32 in [31]) to get

$$\begin{aligned}
\mathrm{d}F_t^T =& -F_{t-}^T\mathrm{d}\Lambda_t^{f^T}-F_{t-}^T\mathrm{d}Z_t^{h^T}+\frac{1}{2}F_{t-}^T\mathrm{d}\left[\Lambda^{f^T},\Lambda^{f^T}\right]_t^c+\frac{1}{2}F_{t-}^T\mathrm{d}\left[Z^{h^T},Z^{h^T}\right]_t^c\\
&+F_{t-}^T\mathrm{d}\left[\Lambda^{f^T},Z^{h^T}\right]_t^c+\Delta F_t^T+F_{t-}^T\Delta\Lambda_t^{f^T}+F_{t-}^T\Delta Z_t^{h^T},\quad 0\le t\le T,
\end{aligned}$$

where $\Delta\Lambda_t^{f^T}=\langle f^T(\cdot,t),\Lambda_t-\Lambda_{t-}\rangle$, and to avoid heavy notation we have written $\Lambda_t^{f^T}$

instead of $\langle f^T(\cdot,t),\Lambda_t\rangle$, and $Z_t^{h^T}$ instead of $\langle h^T(\cdot,t),Z_t\rangle$. Note that without the movement Z is a pure jump process, and since the interaction between Λ and Z is limited to the time of the immigration events, we have that $\left[\Lambda^{f^T},Z^{h^T}\right]_t^c=0$. Taking advantage of

$$F_t^T=F_{t-}^T\mathrm{e}^{-\Delta\Lambda_t^{f^T}-\Delta Z_t^{h^T}},$$

we may thus write in integral form

$$\begin{aligned}F_t^T&=F_0^T-\int_0^t F_{s-}^T\mathrm{d}\Lambda_s^{f^T}-\int_0^t F_{s-}^T\mathrm{d}Z_s^{h^T}+\int_0^t F_{s-}^T\langle\beta(\cdot)(f^T(\cdot,s))^2,\Lambda_{s-}\rangle\mathrm{d}s\\&\quad+\frac{1}{2}\int_0^t F_{s-}^T\langle(\nabla h^T(\cdot,s))^{\mathsf{t}}a\nabla h^T(\cdot,s),Z_{s-}\rangle\mathrm{d}s+\sum_{s\le t}\left\{\Delta F_s^T+F_{s-}^T\Delta\Lambda_s^{f^T}+F_{s-}^T\Delta Z_s^{h^T}\right\}.\end{aligned}$$

To simplify the notation we used that both $f^T(x,t)$ and $h^T(x,t)$ are continuous in t, thus $f^T(x,t)=f^T(x,t-)$ and $h^T(x,t)=h^T(x,t-)$.

We can split up the last term, that is the sum of discontinuities according to the optional random measure in (3.21) responsible for this discontinuity. Thus, writing $\Delta^{(i)},\ i=0,1,2$, to mean an increment coming from each of the three random measures,

$$\begin{aligned}F_t^T=&F_0^T-\int_0^t F_{s-}^T\mathrm{d}\Lambda_s^{f^T}-\int_0^t F_{s-}^T\mathrm{d}Z_s^{h^T}+\int_0^t F_{s-}^T\langle\beta(\cdot)(f^T(\cdot,s))^2,\Lambda_{s-}\rangle\mathrm{d}s\\&+\frac{1}{2}\int_0^t F_{s-}^T\langle(\nabla h^T(\cdot,s))^{\mathsf{t}}a\nabla h^T(\cdot,s),Z_{s-}\rangle\mathrm{d}s+\sum_{s\le t}F_{s-}^T\left\{\mathrm{e}^{-\Delta^{(0)}\Lambda_s^{f^T}}-1+\Delta^{(0)}\Lambda_s^{f^T}\right\}\\&+\sum_{s\le t}F_{s-}^T\left\{\mathrm{e}^{-\Delta^{(1)}\Lambda_s^{f^T}}-1+\Delta^{(1)}\Lambda_s^{f^T}\right\}\\&+\sum_{s\le t}F_{s-}^T\left\{\mathrm{e}^{-\Delta^{(2)}\Lambda_s^{f^T}-\Delta Z_s^{h^T}}-1+\Delta^{(2)}\Lambda_s^{f^T}+\Delta Z_s^{h^T}\right\}.\end{aligned}$$

Next, plugging in $\mathrm{d}\Lambda_s^{f^T}$ and $\mathrm{d}Z_s^{h^T}$ gives

$$\begin{aligned}F_t^T=&F_0^T+\int_0^t F_{s-}^T\langle\psi'(\cdot,w(\cdot))f^T(\cdot,s),\Lambda_{s-}\rangle\mathrm{d}s+\int_0^t F_{s-}^T\langle\beta(\cdot)(f^T(\cdot,s))^2,\Lambda_{s-}\rangle\mathrm{d}s\\&-\int_0^t F_{s-}^T\langle\mathcal{L}f^T(\cdot,s),\Lambda_{s-}\rangle\mathrm{d}s-\int_0^t F_{s-}^T\left\langle\frac{\partial}{\partial s}f^T(\cdot,s),\Lambda_{s-}\right\rangle\mathrm{d}s\\&-\int_0^t F_{s-}^T\left\langle\frac{\partial}{\partial s}h^T(\cdot,s),Z_{s-}\right\rangle\mathrm{d}s-\int_0^t F_{s-}^T\langle\mathcal{L}^w h^T(\cdot,s),Z_{s-}\rangle\mathrm{d}s\\&-\int_0^t F_{s-}^T\langle 2\beta(\cdot)f^T(\cdot,s),Z_{s-}\rangle\mathrm{d}s+\sum_{s\le t}F_{s-}^T\left\{\mathrm{e}^{-\Delta^{(0)}\Lambda_s^{f}}-1+\Delta^{(0)}\Lambda_s^{f^T}\right\}\\&+\sum_{s\le t}F_{s-}^T\left\{\mathrm{e}^{-\Delta^{(1)}\Lambda_s^{f^T}}-1\right\}+\sum_{s\le t}F_{s-}^T\left\{\mathrm{e}^{-\Delta^{(2)}\Lambda_s^{f^T}-\Delta Z_s^{h^T}}-1\right\}\\&+\frac{1}{2}\int_0^t F_{s-}^T\langle(\nabla h^T(\cdot,s))^{\mathsf{t}}a\nabla h^T(\cdot,s),Z_{s-}\rangle\mathrm{d}s+M_t^{loc},\end{aligned}\tag{3.31}$$

where M_t^{loc} is a local martingale corresponding to the terms $U_t^c(f^T)$, $V_t^c(h^T)$ and the integral with respect to the random measure $\tilde{\mathtt{N}}^0$ in (3.21). Note that the two terms with the time-derivative are due to the extra time dependence of the test-functions in the integrals $\langle f^T(\cdot,s),\Lambda_s\rangle$ and $\langle h^T(\cdot,s),Z_s\rangle$. In particular a change in $\langle f^T(s,\cdot),\Lambda_s\rangle$ corresponds to either a change in Λ_s or a change in $f^T(\cdot,s)$.

Next we show that the local martingale term is in fact a real martingale, which will then disappear when we take expectations. First note that due to the boundedness of the drift and diffusion coefficients of the branching mechanism, and the conditions we had on its Lévy measure, the branching of the superprocess can be stochastically dominated by a finite mean CSBP. This means that the CSBP associated to the Esscher-transformed branching mechanism ψ^* is almost surely finite on any finite time interval $[0,T]$, and thus the function f^T is bounded on $[0,T]$. Using the boundedness of f^T and the drift coefficient β, the quadratic variation of the integral

$$\int_0^t F_{s-}^T \mathrm{d}U_s^c(f^T) \tag{3.32}$$

can be bounded from above as follows

$$\begin{aligned}\int_0^t 2F_{s-}^T\langle\beta(\cdot)(f^T(\cdot,s))^2,\Lambda_{s-}\rangle\mathrm{d}s &\le \int_0^t \mathrm{e}^{-\langle f^T(\cdot,s),\Lambda_{s-}\rangle}\langle C(f^T(\cdot,s))^2,\Lambda_{s-}\rangle\mathrm{d}s\\ &\le \int_0^t \mathrm{e}^{-\widetilde{C}||\Lambda_{s-}||}\widehat{C}||\Lambda_{s-}||\mathrm{d}s,\end{aligned}$$

where $C,\widehat{C}$ and $\widetilde{C}$ are finite constants. Since the function $x\mapsto \mathrm{e}^{-\widetilde{C}x}x$ is bounded on $[0,\infty)$, the previous quadratic variation is finite, and so the process (3.32) is a martingale on $[0,T]$.

To show the martingale nature of the stochastic integral

$$\int_0^t F_{s-}\mathrm{d}V_s^c(h^T) \tag{3.33}$$

we note that due to construction, $h^T\in\mathcal{D}_0(\mathcal{L})$, and is bounded on $[0,T]$. Thus, $V_t^c(h^T)$ is in fact a martingale on $[0,T]$, and since $F_{s-}\le 1$, $s\in[0,T]$, the quadratic variation of (3.33) is also finite, which gives the martingale nature of (3.33) on $[0,T]$.

Finally, we consider the integral

$$\int_0^t\int_{\mathcal{M}(E)^\circ} F_{s-}^T\langle f^T(\cdot,s),\nu\rangle\tilde{N}^0(\mathrm{d}s,\mathrm{d}\nu). \tag{3.34}$$

Note that

$$\begin{aligned}
Q_t := \int_0^t \int_{\mathcal{M}(E)^\circ} \left(F_{s-}^T \langle f^T(\cdot,s),\nu\rangle\right)^2 \hat{N}^0(\mathrm{d}s,\mathrm{d}\nu) \\
&= \int_0^t \int_E \int_{(0,\infty)} \left(F_{s-}^T u f^T(x,s)\right)^2 \mathrm{e}^{-w(x)u} m(x,\mathrm{d}u)\Lambda_{s-}(\mathrm{d}x)\mathrm{d}s \\
&\leq \int_0^t \mathrm{e}^{-2C||\Lambda_{s-}||} C \left\langle \int_{(0,\infty)} u^2 \mathrm{e}^{-w(x)u} m(x,\mathrm{d}u), \Lambda_{s-} \right\rangle \mathrm{d}s \\
&\leq \int_0^t \mathrm{e}^{-\widetilde{C}||\Lambda_{s-}||} \widehat{C} ||\Lambda_{s-}|| \mathrm{d}s,
\end{aligned}$$

where $C, \widetilde{C}$ and $\widehat{C}$ are finite constants. Thus $\mathbb{E}[Q_t] < \infty$ on $[0,T]$, and we can refer to page 63 of [19] to conclude that the process (3.34) is indeed a martingale on $[0,T]$.

Thus, after taking expectations and gathering terms, we can write the previous equation in the following form

$$\begin{aligned}
\mathbf{E}_\mu\left[F_t^T\right] = \mathbf{E}_\mu\left[F_0^T\right] &+ \int_0^t \mathbf{E}_\mu\left[F_{s-}^T \langle A(\cdot, f^T(\cdot,s)), \Lambda_{s-}\rangle\right] \mathrm{d}s \\
&- \int_0^t \mathbf{E}_\mu\left[F_{s-}^T \left\langle \frac{\partial}{\partial s} f^T(\cdot,s), \Lambda_{s-}\right\rangle\right] \mathrm{d}s \\
&+ \int_0^t \mathbf{E}_\mu[F_{s-}^T \langle B(\cdot, h^T(\cdot,s), f^T(\cdot,s)), Z_{s-}\rangle] \mathrm{d}s \\
&- \int_0^t \mathbf{E}_\mu\left[F_{s-}^t \left\langle \frac{\partial}{\partial s} h^T(\cdot,s), Z_{s-}\right\rangle\right] \mathrm{d}s, \quad 0 \leq t \leq T,
\end{aligned} \tag{3.35}$$

where

$$\begin{aligned}
A(x,f) &= \psi'(x,w(x))f + \beta(x)f^2 - \mathcal{L}f + \int_{(0,\infty)} \left(\mathrm{e}^{-uf} - 1 + uf\right) \mathrm{e}^{-w(x)u} m(x,\mathrm{d}u) \\
&= -\mathcal{L}f + \psi^*(x,f),
\end{aligned} \tag{3.36}$$

and

$$\begin{aligned}
B(x,h,f) = \frac{1}{2}(\nabla h)^{\mathtt{t}} a \nabla h - \mathcal{L}^w h - 2\beta(x) f + \int_{(0,\infty)} (\mathrm{e}^{-uf} - 1) u \mathrm{e}^{-w(x)u} m(x,\mathrm{d}u) \\
+ \sum_{k=2}^\infty \int_{(0,\infty)} \left(\mathrm{e}^{-uf-(k-1)h} - 1\right) \frac{1}{w(x)} \Big\{\beta(x) w^2(x) \delta_0(\mathrm{d}u) \mathbf{1}_{\{k=2\}} \\
+ w^k(x) \frac{u^k}{k!} \mathrm{e}^{-w(x)u} m(x,\mathrm{d}u)\Big\}.
\end{aligned} \tag{3.37}$$

We can see immediately that $A(x, f^T(x,t))$ is exactly what we have on the right-hand side of (3.25). Furthermore, using that

$$\frac{1}{2}(\nabla h)^{\mathtt{t}} a \nabla h - \mathcal{L}^w h = \mathrm{e}^h \frac{1}{w} \mathcal{L}\left(w \mathrm{e}^{-h}\right) - \frac{1}{w}\psi(\cdot, w), \tag{3.38}$$

we can also verify that

$$B(x,h,f) = \mathrm{e}^h \frac{1}{w}\mathcal{L}\left(w\mathrm{e}^{-h}\right) + \mathrm{e}^h \frac{1}{w}\left(\psi^*\left(x, -w(x)\mathrm{e}^{-h} + f\right) - \psi^*(x,f)\right),$$

that is, $B(x, h^T(x,t), f^T(x,t))$ equals to the right-hand side of (3.30). Hence, recalling the defining equations of f^T (3.25) and h^T (3.30), we get that the last four terms of (3.35) cancel, and thus $\mathbf{E}_\mu[F_t^T] = \mathbf{E}_\mu[F_0^T]$ for $t \in [0,T]$, as required. In particular, using the boundary conditions for f^T and h^T, we get that

$$\mathbf{E}_\mu\left[F_T^T\right] = \mathbf{E}_\mu\left[\mathrm{e}^{-\langle f(\cdot),\Lambda_T\rangle - \langle h(\cdot), Z_T\rangle}\right] = \mathbf{E}_\mu\left[\mathrm{e}^{-\langle f^T(\cdot,0),\Lambda_0\rangle - \langle h^T(\cdot,0), Z_0\rangle}\right] = \mathbf{E}_\mu\left[F_0^T\right]. \tag{3.39}$$

Before moving on, we note that equation (3.39) relates the joint Laplace functional of the solution to two objects both of which are characterised as a unique solution to a partial differential equation. Since this uniquely identifies the Laplace function, we get that the weak solution to the SDE (3.21) is indeed unique.

Next, we can notice that by construction we can relate the right-hand side of the expression (3.39) to the superprocess. In particular, using the Poissonian nature of Z_0, and that $X_0 = \Lambda_0 = \mu$ is deterministic we have

$$\mathbf{E}_\mu\left[\mathrm{e}^{-\langle f^T(\cdot,0),\Lambda_0\rangle - \langle h^T(\cdot,0), Z_0\rangle}\right] = \mathbb{E}_\mu\left[\mathrm{e}^{-\left\langle f^T(\cdot,0) + w(\cdot)\left(1-\mathrm{e}^{-h^T(\cdot,0)}\right), X_0\right\rangle}\right], \tag{3.40}$$

where X_t is a solution to (3.17). Thus, by choosing the right test-functions, we could equate the value of F_t^T at T to its initial value, which in turn gave a connection with the superprocess. The next step is to show that the process

$$\mathrm{e}^{-\left\langle f^T(\cdot,t) + w(\cdot)\left(1-\mathrm{e}^{-h^T(\cdot,t)}\right), X_t\right\rangle}, \quad t\in[0,T],$$

has constant expectation on $[0,T]$, which would then allow us to deduce

$$\mathbf{E}_\mu\left[\mathrm{e}^{-\langle f(\cdot),\Lambda_T\rangle - \langle h(\cdot), Z_T\rangle}\right] = \mathbb{E}_\mu\left[\mathrm{e}^{-\left\langle f(\cdot) + w(\cdot)\left(1-\mathrm{e}^{-h(\cdot)}\right), X_T\right\rangle}\right].$$

To simplify the notation let $\kappa^T(x,t) := f^T(x,t) + w(x)\left(1-\mathrm{e}^{-h^T(x,t)}\right)$, and define $G_t^T := \mathrm{e}^{-\langle \kappa^T(\cdot,t), X_t\rangle}$. As the argument here is the exact copy of the previous analysis, we only give the main steps of the calculus, and leave it to the reader to fill in the gaps.

Since $\langle \kappa^T(\cdot,t), X_t\rangle$, $t \le T$, is a semi-martingale, we can use Itô's formula to get

$$\begin{aligned}
G_t^T = G_0^T &+ \int_0^t G_{s-}^T \left\langle \psi'(\cdot, w(\cdot))\kappa^T(s,\cdot), X_{s-}\right\rangle \mathrm{d}s + \int_0^t G_{s-}^T \langle \beta(\cdot)(\kappa^T(\cdot,s))^2, X_{s-}\rangle \mathrm{d}s \\
&- \int_0^t G_{s-}^T \langle 2\beta(\cdot)w(\cdot)\kappa^T(\cdot,s), X_{s-}\rangle \mathrm{d}s - \int_0^t G_{s-}^T \langle \mathcal{L}\kappa^T(\cdot,s), X_{s-}\rangle \mathrm{d}s \\
&+ \int_0^t G_{s-}^T \left\langle \int_0^\infty \left(\mathrm{e}^{-u\kappa^T(\cdot,s)} - 1 + u\kappa^T(\cdot,s)\right)\mathrm{e}^{-w(\cdot)u} m(\cdot,\mathrm{d}u), X_{s-}\right\rangle \mathrm{d}s \\
&+ \int_0^t G_{s-}^t \left\langle \int_0^\infty \left(\mathrm{e}^{-u\kappa^T(\cdot,s)} - 1\right) w(\cdot) u \mathrm{e}^{-w(\cdot)u} m(\cdot,\mathrm{d}u), X_{s-}\right\rangle \mathrm{d}s \\
&+ \int_0^t G_{s-}^T \left\langle \int_0^\infty \left(\mathrm{e}^{-u\kappa^T(\cdot,s)} - 1\right) \sum_{k=2}^\infty \frac{(w(\cdot)u)^k}{k!} \mathrm{e}^{-w(\cdot)u} m(\cdot,\mathrm{d}u), X_{s-}\right\rangle \mathrm{d}s \\
&- \int_0^t G_{s-}^T \left\langle \frac{\partial}{\partial s}\kappa^T(\cdot,s), X_{s-}\right\rangle \mathrm{d}s + M_t^{loc}.
\end{aligned} \tag{3.41}$$

where M_t^{loc} is a local martingale corresponding to the term $M_t^c(f)$, and the integral with respect to the random measure $\tilde{N}^0$ in (3.17). Note that the reasoning that lead to the martingale nature of the local martingale term of (3.31) can also be applied here, which gives that M_t^{loc} in (3.41) is in fact a true martingale on $[0,T]$, which we denote by M_t.

Next we plug in κ^T, and after some laborious amount of algebra get

$$\begin{aligned}
G_t^T = &G_0^T + \int_0^t G_{s-}^T \langle \psi'(\cdot, w(\cdot)) f^T(\cdot,s), X_{s-}\rangle \mathrm{d}s + \int_0^t G_{s-}^T \langle \beta(\cdot)(f^T(\cdot,s))^2, X_{s-}\rangle \mathrm{d}s \\
&- \int_0^t G_{s-}^T \langle \mathcal{L} f^T(\cdot,s), X_{s-}\rangle \mathrm{d}s \\
&+ \int_0^t G_{s-}^T \left\langle \int_{(0,\infty)} (\mathrm{e}^{-uf^T(\cdot,s)} - 1 + uf^T(\cdot,s))\mathrm{e}^{-w(\cdot)u} m(\cdot,\mathrm{d}u), X_{s-}\right\rangle \mathrm{d}s \\
&- \int_0^t G_{s-}^T \langle 2\beta(\cdot) f^T(\cdot,s)\mathrm{e}^{-h^T(\cdot,s)} w(\cdot), X_{s-}\rangle \mathrm{d}s \\
&+ \int_0^t G_{s-}^T \left\langle \int_{(0,\infty)} (\mathrm{e}^{-uf^T(\cdot,s)} - 1) u \mathrm{e}^{-w(\cdot)u} m(\cdot,\mathrm{d}u)\mathrm{e}^{-h^T(\cdot,s)} w(\cdot), X_{s-}\right\rangle \mathrm{d}s \\
&+ \int_0^t G_{s-}^T \Bigg\langle \sum_{k=2}^\infty \int_{(0,\infty)} (\mathrm{e}^{-uf^T(\cdot,s)-(k-1)h^T(\cdot,s)} - 1)\frac{1}{w(\cdot)} \\
&\qquad \left\{\beta(\cdot) w^2(\cdot)\delta_0(\mathrm{d}u)\mathbf{1}_{\{k=2\}} + w^k(\cdot)\frac{u^k}{k!}\mathrm{e}^{-w(\cdot)u} m(\cdot,\mathrm{d}u)\right\} \mathrm{e}^{-h^T(\cdot)} w(\cdot), X_{s-}\Bigg\rangle \mathrm{d}s \\
&+ \int_0^t G_{s-}^T \left\langle (1 - \mathrm{e}^{-h^T(\cdot,s)})\psi(\cdot, w(\cdot)) - \mathcal{L}w(\cdot)(1 - \mathrm{e}^{-h^T(\cdot,s)}), X_{s-}\right\rangle \mathrm{d}s \\
&- \int_0^t G_{s-}^T \left\langle \frac{\partial}{\partial s}\kappa^T(\cdot,s), X_{s-}\right\rangle \mathrm{d}s + M_t.
\end{aligned}$$

Using once again the identity (3.38), and taking expectations give

$$\begin{aligned}\mathbb{E}_\mu[G_t^T] = \mathbb{E}_\mu[G_0^T] &+ \int_0^t \mathbb{E}_\mu[G_{s-}^T\langle A(\cdot, f^T(\cdot,s)), X_{s-}\rangle]\mathrm{d}s \\ &+ \int_0^t \mathbb{E}_\mu[G_{s-}^T\langle \mathrm{e}^{-h^T(\cdot,s)}w(\cdot)B(\cdot, h^T(\cdot,s), f^T(\cdot,s), X_{s-}\rangle]\mathrm{d}s \\ &- \int_0^t \mathbb{E}_\mu\left[G_{s-}^T\left\langle \frac{\partial}{\partial s}\kappa^t(s,\cdot), X_{s-}\right\rangle\right]\mathrm{d}s,\end{aligned} \tag{3.42}$$

where A and B are given by (3.36) and (3.37). Finally, noting

$$\frac{\partial}{\partial s}\kappa^T(x,s) = \frac{\partial}{\partial s}f^T(x,s) + w(x)\mathrm{e}^{-h^T(x,s)}\frac{\partial}{\partial s}h^T(x,s),$$

gives

$$\frac{\partial}{\partial s}\kappa^T(s,x) = -A(x, f^T(x,s)) - w(x)\mathrm{e}^{-h^T(x,s)}B(x, h^T(x,s), f^T(x,s)),$$

which results in the cancellation of the last three terms in (3.42), and hence verifies the constant expectation of G_t^T on $[0,T]$. In particular, we have proved that

$$\begin{aligned}\mathbb{E}_\mu[G_T^T] &= \mathbb{E}_\mu\left[\mathrm{e}^{-\left\langle f(\cdot)+w(\cdot)\left(1-\mathrm{e}^{-h(\cdot)}\right), X_T\right\rangle}\right] \\ &= \mathbb{E}_\mu\left[\mathrm{e}^{-\left\langle f^T(\cdot,0)+w(\cdot)\left(1-\mathrm{e}^{-h^T(\cdot,0)}\right), X_0\right\rangle}\right] = \mathbb{E}_\mu[G_0^T].\end{aligned} \tag{3.43}$$

In conclusion, combining the previous observations (3.39) and (3.40) with (3.43) gives

$$\mathbf{E}_\mu\left[\mathrm{e}^{-\langle f(\cdot),\Lambda_T\rangle - \langle h(\cdot), Z_T\rangle}\right] = \mathbb{E}_\mu\left[\mathrm{e}^{-\left\langle f(\cdot)+w(\cdot)\left(1-\mathrm{e}^{-h(\cdot)}\right), X_T\right\rangle}\right].$$

Since $T > 0$ was arbitrary this equality holds for any time $T > 0$.

Then we have the following implications. First, choosing h and f not identical to zero, we get that the pair (Λ_t, Z_t) under $\mathbf{P}_\mu$ has the same law as $(X_t, \mathrm{Po}(w(x)X_t(\mathrm{d}x)))$ under $\mathbb{P}_\mu$, where $\mathrm{Po}(w(x)X_t(\mathrm{d}x))$ is an autonomously independent Poisson random measure with intensity $w(x)X_t(\mathrm{d}x)$, thus Z_t given Λ_t is indeed a Poisson random measure with intensity $w(x)\Lambda_t(\mathrm{d}x)$. Since the solution to an SDE is a Markov process this Poissonisation allows us to infer that $\Lambda_T, T \geq 0$ is itself Markovian.

Next, by setting $h = 0$ we find that

$$\mathbf{E}_\mu\left[\mathrm{e}^{-\langle f(\cdot),\Lambda_T\rangle}\right] = \mathbb{E}_\mu\left[\mathrm{e}^{-\langle f(\cdot), X_T\rangle}\right],$$

which shows that under $\mathbf{P}_\mu$, Λ_T has the same distribution as X_T under $\mathbb{P}_\mu$, and hence proves that $(\Lambda_T, T \geq 0)$ is indeed a weak solution to (3.16). □

Bibliography

[1] Romain Abraham and Jean-Franois Delmas. A continuum-tree-valued markov process. *Ann. Probab.*, 40(3):1167–1211, 05 2012.

[2] J. Berestycki, A.E. Kyprianou, and A. Murillo-Salas. The prolific backbone for supercritical superprocesses. *Stochastic Processes and their Applications*, 121(6):1315–1331, jun 2011.

[3] Jean Bertoin, Joaquin Fontbona, and Servet Martínez. On prolific individuals in a supercritical continuous-state branching process. *Journal of Applied Probability*, 45(3):714–726, 2008.

[4] T. Duquesne and J.-F. Le Gall. Random Trees, Lévy Processes and Spatial Branching Processes. *ArXiv Mathematics e-prints*, September 2005.

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

[6] E. B. Dynkin. *An introduction to branching measure-valued processes*, volume 6 of *CRM Monograph Series*. American Mathematical Society, Providence, RI, 1994.

[7] E. B. Dynkin. *Diffusions, superprocesses and partial differential equations*, volume 50. American Mathematical Society, Colloquium Publications, Providence, RI, 2002.

[8] E. B. Dynkin and S. E. Kuznetsov. $\mathbb{N}$-measures for branching exit Markov systems and their applications to differential equations. *Probab. Theory Related Fields*, 130(1):135–150, 2004.

[9] M. Eckhoff, A. E. Kyprianou, and M. Winkel. Spines, skeletons and the strong law of large numbers for superdiffusions. *Ann. Probab.*, 43(5):2545–2610, 09 2015.

[10] János Engländer and Ross G. Pinsky. On the construction and support properties of measure-valued diffusions on $d \subseteq \mathbb{R}^d$ with spatially dependent branching. *Ann. Probab.*, 27(2):684–730, 04 1999.

[11] A. M. Etheridge and D. R. E. Williams. A decomposition of the $(1+\beta)$-superprocess conditioned on survival. *Proceedings of the Royal Society of Edinburgh: Section A Mathematics*, 133(04):829, aug 2003.

[12] Alison M. Etheridge. *An introduction to superprocesses*, volume 20 of *University Lecture Series*. American Mathematical Society, Providence, RI, 2000.

[13] S. N. Ethier and T. G. Kurtz. *Markov processes*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, 1986. Characterization and convergence.

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

[15] D. Fekete, J. Fontbona, and A. E. Kyprianou. Skeletal stochastic differential equations for continuous-state branching process. *ArXiv e-prints*, February 2017.

[16] Dorottya Fekete, Sandra Palau, Juan Carlos Pardo, and José Luis Pérez. Backbone decomposition of multitype superprocesses. *arXiv e-prints*, page arXiv:1803.09620, March 2018.

[17] M. Fittipaldi and J. Fontbona. On sde associated with continuous-state branching processes conditioned to never be extinct. *Electron. Commun. Probab.*, 17:no. 49, 1–13, 2012.

[18] P. J. Fitzsimmons. Construction and regularity of measure-valued markov branching processes. *Israel Journal of Mathematics*, 64(3):337–361, Dec 1988.

[19] N. Ikeda and S. Watanabe. *Stochastic differential equations and diffusion processes.* North-Holland mathematical library. North-Holland Pub. Co., 1989.

[20] J. Jacod. *Calcul Stochastique et Problèmes de Martingales.* Number no. 714 in Lecture Notes in Mathematics. Springer, 1979.

[21] T. Kurtz. *Stochastic Analysis 2010.* Springer, 2011. Equivalence of stochastic equations and martingale problems.

[22] A. Kyprianou, J.-L. Pérez, and Y.-X. Ren. *The backbone decomposition for spatially dependent supercritical superprocesses*, volume 2123 of *Lecture Notes in Mathematics*, pages 33 – 59. Springer International Publishing, Switzerland, 2014.

[23] A.E. Kyprianou and Y.-X. Ren. Backbone decomposition for continuous-state branching processes with immigration. *Statistics & Probability Letters*, 82(1):139–144, 2012.

[24] A. Lambert. Quasi-stationary distributions and the continuous-state branching process conditioned to be never extinct. *Electron. J. Probab.*, 12:no. 14, 420–446, 2007.

[25] A. Lambert. Population dynamics and random genealogies. *Stoch. Models*, 24(suppl. 1):45–163, 2008.

[26] Jean-Francois Le Gall. *Spatial Branching Processes, Random Snakes and Partial Differential Equations.* Lectures in Mathematics. ETH Zürich. Birkhäuser Basel, 1999.

[27] Z. Li. *Measure Valued Branching Markov Processes.* Springer-Verlag Berlin Heidelberg, 2011.

[28] A. Murillo-Salas and J.-L. Pérez. *The Backbone Decomposition for Superprocesses with Non-local Branching*, pages 199–216. Springer International Publishing, Cham, 2015.

[29] B. Øksendal. *Stochastic Differential Equations: An Introduction with Applications.* Hochschultext / Universitext. Springer, 2003.

[30] R.G. Pinsky. *Positive Harmonic Functions and Diffusion.* Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2008.

[31] P. Protter. *Stochastic integration and differential equations*, volume 21 of *Applications of Mathematics (New York).* Springer-Verlag, Berlin, 1990. A new approach.

[32] S. Roelly-Coppoletta and A. Rouault. Processus de Dawson-Watanabe conditionné par le futur lointain. *C. R. Acad. Sci. Paris Sér. I Math.*, 309(14):867–872, 1989.

[33] Thomas S. Salisbury and John Verzani. On the conditioned exit measures of super brownian motion. *Probability Theory and Related Fields*, 115(2):237–285, Sep 1999.

[34] J. Xiong. *Three Classes of Nonlinear Stochastic Partial Differential Equations.* World Scientific, 2013.

Concluding remarks

In this chapter we have extended the coupled SDE approach developed in Chapter 2 to the spatial setting of superprocesses. By considering a superprocess with spatially dependent branching mechanism we are allowed to have various interpretations of what successful genealogies mean. In particular, the martingale function used in the definition of the skeleton process can be associated with the event of finite time extinction, but we can also have some additional spatial constraints. In both cases the skeleton particles avoid the behaviour specified by this martingale function.

Just as in Chapter 2, the coupled SDE simultaneously describes the embedded skeleton and the total mass in the system. The second co-ordinate of the SDE gives the Markov branching process which acts as the skeleton, while the first co-ordinate reveals how this skeleton is dressed with immigration. We have showed that the process defined by the first co-ordinate, upon projected to the appropriate filtration, has the same law as the original superprocess.

Even though we have only considered supercritical superprocesses, we believe that by using a time-dependent h-transform to define the skeleton process, our method can be extended to the time-inhomogeneous case of subcritical superprocesses. Nevertheless, the paper leaves this case open.

In the next chapter we will take a slightly different approach to skeletal decompositions, and construct the prolific backbone decomposition of a multitype superprocess using a semigroup approach, which has been widely used in the one-type case. Our results in Chapter 4 expand the class of branching processes for which prolific genealogies have been described.

Appendix 6B: Statement of Authorship

This declaration concerns the article entitled:			
Backbone decomposition of multitype superprocesses			
Publication status (tick one) **Draft manuscript** [ ] **Submitted** [X] **In review** [ ] **Accepted** [ ] **Published** [ ]			
Publication details (reference)	arXiv:1803.09620		
Copyright status (tick the appropriate statement) I hold the copyright for this material [ ] Copyright is retained by the publisher, but I have been given permission to replicate the material here [ ]			
Candidate's contribution to the paper (provide details, and also indicate as a percentage)	*All research has been conducted in equal partnership with collaborators and supervisors. It is unwise to try and measure percentages of intellectual contribution as research scholarship lies as much in the escalation of ideas through mathematical discourse as it does with the seed of ideas themselves. That said, the mathematical content of this thesis is, as an entire piece of work, inextricably associated to its author through intellectual ownership. Formulation of ideas: The candidate played an integral and fully collaborative role in the formulation of ideas. Design of methodology: The candidate played an integral and fully collaborative role in the design of methodology. Experimental work: N/A Presentation of data in journal format: N/A *The wording in this box follows the advice and approval of my supervisor, Professor Kyprianou.		
Statement from Candidate	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature.		
Signed		**Date**	16/07/2019

Chapter 4

Backbone decomposition of multi-type superprocesses

Dorottya Fekete [1] , Sandra Palau [2], Juan Carlos Pardo [3] , Jose-Luis Pérez [3]

Abstract

In this paper, we provide a construction of the so-called backbone decomposition for multitype supercritical superprocesses. While backbone decompositions are fairly well-known for both continuous-state branching processes and superprocesses in the one-type case, so far no such decompositions or even description of prolific genealogies have been given for the multitype cases.

Here we focus on superprocesses, but by turning the movement off, we get the prolific backbone decomposition for multitype continuous-state branching processes as an easy consequence of our results.

4.1 Introduction and main results.

Motivated by the distributional decomposition of supercritical superprocesses with quadratic branching mechanism presented in Evans and O'Connell, [11] and the path-wise decomposition of Duquesne and Winkel [6] of continuous-state branching processes (CB-processes), Berestycki et al. [3] provided a pathwise construction of the so-called backbone decomposition for supercritical superprocesses. The authors in [3] showed that the superprocess can be written as the sum of two independent processes. The first one is an initial burst of subcritical mass, while the second one is subcritical mass immigrating continuously and discontinuously along the path of a branching particle system called the *backbone* that we explain briefly below.

[1]Department of Mathematical Sciences, University of Bath, Claverton Down, Bath, BA2 7AY, United Kingdom. Email: d.fekete@bath.ac.uk

[2]Instituto de Investigaciones en Matemáticas Aplicadas y Sistemas, Circuto Escolar 3000, C.U., 04510 Ciudad de Mx́ico, CDMX, Mexico.. Email: sandra@sigma.iimas.unam.mx

[3]Centro de Investigación en Matemáticas A.C. Calle Jalisco s/n. 36240 Guanajuato, México. E-mail: jcpardo@cimat.mx, jl.garmendia@cimat.mx.

In Evans and O'Connell [11] a distributional decomposition of supercritical superprocesses with quadratic spatially independent branching mechanism, as sum of two independent processes, was given. Later Engländer and Pinsky [9] provided a similar decomposition for the spatially dependent case. In both constructions, the first process is a copy of the original process conditioned on extinction. The second process is understood as the aggregate accumulation of mass that has immigrated *continuously* along the path of an auxiliary dyadic branching particle diffusion which starts with a Poisson number of particles. Such an embedded branching particle system was introduced as the *backbone*.

A pathwise backbone decomposition appears in Salisbury and Verzani [22], who consider the case of conditioning a super-Brownian motion as it exits a given domain such that the exit measure contains at least n pre-specified points in its support. There it was found that the conditioned process has the same law as the superposition of mass that immigrates in a Poissonian way along the spatial path of a branching particle motion which exits the domain with precisely n particles at the pre-specified points. Another pathwise backbone decomposition for branching particle systems is given in Etheridge and Williams [10], which is used in combination with a limiting procedure to prove another version of Evans' immortal particle picture.

Duquesne and Winkel [6], in the context of Lévy trees and with no spatial motion, considered a similar decomposition for CB-processes whose branching mechanism ψ satisfies that $0 \leq -\psi'(0+) < \infty$ and the so-called Grey's condition

$$\int^{\infty} \frac{\mathrm{d}u}{\psi(u)} < \infty.$$

In this case the *backbone* corresponds to a continuous-time Galton-Watson process, and the general nature of the branching mechanism induces three different sorts of immigration. The *continuous immigration* is described by a Poisson point process of independent processes along the backbone, and the immigration mechanism is given by the so-called excursion measure which assigns zero initial mass and finite length to the immigration processes. The *discontinuous immigration* is provided by two sources of immigration. The first one is described again by a Poisson point process of independent processes along the backbone where the immigration mechanism is given by the law of the original process conditioned on extinction, and with initial mass randomised by an infinite measure. The second source of discontinuous immigration is given by independent copies of the original process conditioned on extinction, which are added to the backbone at its branching times, with randomly distributed initial mass that depends on the number of offspring at the branch point. Other decomposition of Lévy trees include Duquesne and Wang [5], who provide a composition of Lévy trees according to their diameter.

In Berestycki et al. [3], a similar decomposition is provided for a class of superprocesses whose branching mechanisms satisfy the same conditions as those considered by Duquesne and Winkel. It is important to note that the authors in [3] also considered supercritical CB-processes that, with positive probability, may die out without this ever happening in a finite time. This also allows the inclusion of branching mechanisms which are associated to CB-processes with paths of bounded variation which were excluded in [6]. Kyprianou and Ren [17] look at the case of a CB-process with immigration for which a similar backbone decomposition to [3] can be given. Finally, backbone decompositions have also been considered for superprocesses with spatially dependent branching mechanisms which are local, see Kyprianou et al. [16], Eckhoff et al. [8], and Chen et al. [4], and non-local, see Murillo-Salas and Pérez [19]. Note that in the aforementioned articles the superprocess is supported on a domain of $\mathbb{R}^d$, thus the state space is always continuous.

In this paper, we offer a similar construction for multitype superprocesses whose branching mechanisms are general, but with the restriction of being spatially independent and having a finite number of types. Technically such a process can be defined as a one-type superprocess on a state space that is the mixture of a continuous and a discrete space, and whose branching has both local and non-local elements. While backbone decompositions in the one-type case are fairly well-known for both CB-processes and superprocesses whose state-space is continuous, so far no such decompositions or even description of prolific genealogies (i.e. those individuals with infinite line of descent) have been given for multitype processes. Here we focus on superprocesses, but by turning the movement off, we get the prolific backbone decomposition for multitype continuous-state branching processes (MCB-processes) as an easy consequence of our results.

Multitype superprocesses were first studied by Gorostiza and Lopez-Mimbela [12] for the particular case of quadratic branching. Later Li [18] extended the notion of multitype superprocesses to more general branching mechanisms (see also Section 6.2 in the monograph of Li [20]). Roughly speaking, the dynamics of the superprocesses introduced by Li are as follows. The movement of mass of a given type is a Borel right process, the death and birth of mass of each type are associated with a spectrally positive Lévy process. From a given type, the creation of mass of other types is given by the law of a subordinator, and is distributed according to a discrete distribution that depends on the type. We are interested in a slightly more general superprocess where the discrete distributions are randomly chosen by a probability kernel that depends on the type. Thus the locations of non-locally displaced offspring involve two sources of randomness. One of the advantages of taking this general branching mechanism is that if there is no spatial motion, we recover the MCB-process studied by Kyprianou et al. [15], which was properly defined by Li in Example 2.2 in [20].

Kyprianou et al. [15] studied the almost sure growth of supercritical MCB-processes and implicitly described a spine decomposition. In [15], the authors show that a MCB-process conditioned to never become extinct is equal in law to the sum of an independent copy of the original process and three different sources of immigration along a spine (continuous, discontinuous and in the times when the spine jumps). More precisely, the spine is given by a Markov chain, the continuous and discontinuous immigrations are described by a Poisson point process along the spine, where MCB-processes with the original branching mechanism are immigrating with zero initial mass and with randomised initial mass, respectively. Due to the non-local nature of the branching mechanism, an additional phenomenon occurs; a positive random amount of mass immigrates off the spine each time it jumps from one state to another. Moreover, the distribution of the immigrating mass depends on where the spine jumped from and where it jumped to.

The backbone and spine decompositions are quite different. In the backbone decomposition, the object that we dress is a multitype branching diffusion while in the spine decomposition, this object is a Markov chain which does not branch. Another difference is related to the immigration processes. In the spine decomposition, these are independent copies of the original process while in the backbone decomposition they are independent copies of the process conditioned to become extinct. In other words, we can think of the backbone as all the particles that have an infinite genealogical line of descent, and of the spine as just one infinite line of descent.

4.1.1 Multitype superprocesses.

Before we introduce multitype superprocesses and some of their properties, we first recall some basic notation. Let $\ell \in \mathbb{N}$ be a natural number, and set $S = \{1, 2, \cdots, \ell\}$. We denote by $\mathcal{M}(\mathbb{R}^d)$, $\mathcal{B}(\mathbb{R}^d)$ and $\mathcal{B}^+(\mathbb{R}^d)$ the respective spaces of finite Borel measures, bounded Borel functions and positive bounded Borel functions on $\mathbb{R}^d$. The space $\mathcal{M}(\mathbb{R}^d)$ is endowed with the topology of weak convergence.

For $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^\ell$, we introduce $[\boldsymbol{u}, \boldsymbol{v}] = \sum_{j=1}^{\ell} u_j v_j$, and $\boldsymbol{u} \cdot \boldsymbol{v}$ as the vector with entries $(\boldsymbol{u} \cdot \boldsymbol{v})_j = u_j v_j$. For a matrix A, we denote by $A^{\mathtt{t}}$ its transpose. For any $\boldsymbol{f} = (f_1, \ldots, f_\ell)^{\mathtt{t}} \in \mathcal{B}(\mathbb{R}^d)^\ell$ and $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_\ell)^{\mathtt{t}} \in \mathcal{M}(\mathbb{R}^d)^\ell$, we define

$$\langle \boldsymbol{f}, \boldsymbol{\mu} \rangle := \sum_{i=1}^{\ell} \int_{\mathbb{R}^d} f_i(x) \mu_i(\mathrm{d}x).$$

Furthermore, we also use $|\boldsymbol{u}| := [\boldsymbol{u}, \boldsymbol{u}]^{1/2}$ for the Euclidian norm of any $\boldsymbol{u} \in \mathbb{R}^\ell$, and $\|\boldsymbol{\mu}\| := \langle \mathbf{1}, \boldsymbol{\mu} \rangle$ for the total mass of the measure $\boldsymbol{\mu}$.

Suppose that for any $i \in S$, the process $\xi^{(i)} = (\xi^{(i)}_t, t \geq 0)$ is a diffusion with conservative transition semigroup $(\mathtt{P}^{(i)}_t, t \geq 0)$ on $\mathbb{R}^d$. We also introduce a vectorial function $\boldsymbol{\psi} : S \times \mathbb{R}^\ell_+ \to \mathbb{R}^\ell$ such that

$$\psi(i, \boldsymbol{\theta}) := -[\boldsymbol{\theta}, \boldsymbol{B}\boldsymbol{e}_i] + \beta_i \theta_i^2 + \int_{\mathbb{R}^\ell_+} \left(\mathrm{e}^{-[\boldsymbol{\theta}, \boldsymbol{y}]} - 1 + \theta_i y_i \right) \Pi(i, \mathrm{d}\boldsymbol{y}), \qquad \boldsymbol{\theta} \in \mathbb{R}^\ell_+, i \in S, \tag{4.1}$$

where $\boldsymbol{B}$ is an $\ell \times \ell$ real valued matrix such that $B_{ij} \mathbf{1}_{\{i \neq j\}} \in \mathbb{R}_+$, $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_\ell\}$ is the natural basis in $\mathbb{R}^\ell$, $\beta_i \in \mathbb{R}_+$, and Π is a measure satisfying the following integrability condition

$$\int_{\mathbb{R}^\ell_+ \setminus \{\mathbf{0}\}} \left((|\boldsymbol{y}| \wedge |\boldsymbol{y}|^2) + \sum_{j \in S} \mathbf{1}_{\{j \neq i\}} y_j \right) \Pi(i, \mathrm{d}\boldsymbol{y}) < \infty, \qquad \text{for} \quad i \in S.$$

We call the function $\boldsymbol{\psi}$ the branching mechanism and we also refer to Π as its associated Lévy measure. Intuitively, the conditions on the entries of $\boldsymbol{B}$ and the above integrability condition hides the fact that creation of mass into own type is governed by a spectrally positive Lévy process, while creation of mass into different types can be associated to a subordinator.

The first result that we present here says that multitype superprocesses associated to the branching mechanism $\boldsymbol{\psi}$ and the diffusions $\{\xi^{(i)}, i \in S\}$ are well-defined. Its proof is based on similar arguments as those used to prove Theorem 6.4 in Li [20], but for completeness we present its proof in Section 4.2.

Proposition 4.1.1. *There is a strong Markov process* $\boldsymbol{X} = (\boldsymbol{X}_t, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P}_{\boldsymbol{\mu}})$ *with state space* $\mathcal{M}(\mathbb{R}^d)^\ell$ *and transition probabilities defined by*

$$\mathbb{E}_{\boldsymbol{\mu}}\left[\mathrm{e}^{-\langle \boldsymbol{f}, \boldsymbol{X}_t\rangle}\right] = \exp\Big\{ - \langle \boldsymbol{V}_t \boldsymbol{f}, \boldsymbol{\mu}\rangle\Big\}, \qquad \boldsymbol{\mu} \in \mathcal{M}(\mathbb{R}^d)^\ell, \tag{4.2}$$

where $\boldsymbol{f} \in \mathcal{B}^+(\mathbb{R}^d)^\ell$ *and* $\boldsymbol{V}_t\boldsymbol{f}(x) = (V_t^{(1)}\boldsymbol{f}(x), \cdots, V_t^{(\ell)}\boldsymbol{f}(x))^{\mathtt{t}} : \mathbb{R}^d \to \mathbb{R}_+^\ell$ *is the unique locally bounded solution to the integral equation*

$$V_t^{(i)}\boldsymbol{f}(x) = \mathtt{P}_t^{(i)} f_i(x) - \int_0^t \mathrm{d}s \int_{\mathbb{R}^d} \psi(i, \boldsymbol{V}_{t-s}\boldsymbol{f}(y))\mathtt{P}_s^{(i)}(x, \mathrm{d}y), \qquad i \in S. \tag{4.3}$$

Definition 1. *The process* $\boldsymbol{X}$ *is called a* $(\mathtt{P}, \boldsymbol{\psi})$*-mutitype superprocess with* ℓ *types and with law given by* $\mathbb{P}_{\boldsymbol{\mu}}$ *for each initial configuration* $\boldsymbol{\mu} \in \mathcal{M}(\mathbb{R}^d)^\ell$.

Our definition is consistent with the multitype superprocesses that appear in the literature. Indeed, we observe that the multitype superprocesses considered by Gorostiza and Lopez-Mimbela [12] are associated with the branching mechanism

$$\psi(i, \boldsymbol{\theta}) = -d_i[\boldsymbol{\theta}, \boldsymbol{\pi}^{(i)}] + \beta_i\theta_i^2,$$

where $d_i, \beta_i \in \mathbb{R}_+$, $\boldsymbol{\pi}^{(i)} = \{\pi_j^{(i)}, j \in S\}$ is a probability distribution on S, and the spatial movement is driven by the family $\{\xi^{(i)}, i \in S\}$ of symmetric stable processes. Li [18] (see also Section 6.2 in [20]) introduced multitype superprocesses with spatial movement driven by Borel right processes and whose branching mechanism is of the form

$$\psi(i, \boldsymbol{\theta}) = b_i\theta_i + \beta_i\theta_i^2 - d_i[\boldsymbol{\theta}, \boldsymbol{\pi}^{(i)}] + \int_{\mathbb{R}_+}\Big(\mathrm{e}^{-u\theta_i} - 1 + \theta_i u\Big)\, l(i, \mathrm{d}u) + \int_{\mathbb{R}_+}\Big(\mathrm{e}^{-u[\boldsymbol{\theta}, \boldsymbol{\pi}^{(i)}]} - 1\Big)\, n(i, \mathrm{d}u),$$

where $b_i, d_i, \beta_i \in \mathbb{R}_+$, $\boldsymbol{\pi}^{(i)} = \{\pi_j^{(i)}, j \in S\}$ is a probability distribution on S, and $l(i, \mathrm{d}u)$, $n(i, \mathrm{d}u)$ are measures on $\mathbb{R}_+$ satisfying

$$\int_{\mathbb{R}_+} (u \wedge u^2) l(i, \mathrm{d}u) < \infty \qquad \text{and} \qquad \int_{\mathbb{R}_+} u\, n(i, \mathrm{d}u) < \infty,$$

that represent the local and non-local kernels, respectively. The latter branching mechanism can be rewritten in the form of (4.1) by taking $B_{ji} := -b_i \mathbf{1}_{\{i=j\}} + d_i \pi_j^{(i)}$, and

$$\Pi(i, \mathrm{d}\boldsymbol{y}) = \mathbf{1}_{\{\boldsymbol{y}=u\boldsymbol{e}_i\}} l(i, \mathrm{d}u) + \mathbf{1}_{\{\boldsymbol{y}=u\boldsymbol{\pi}^{(i)}\}} n(i, \mathrm{d}u).$$

It is important to note that if the branching mechanism is given as in (4.1) and there is no spatial movement, then the associated total mass of a superprocess is a MCB-process, see for instance Example 2.2 in [20]. Indeed, it is not difficult to see that the total mass vector of a spatially homogeneous multitype superprocess, whose underlying spatial motion is conservative, is a MCB-process. Recall that an ℓ-type MCB-process $\boldsymbol{Y} = (\boldsymbol{Y}_t, t \geq 0)$ with branching mechanism $\boldsymbol{\psi}$ can be characterised through its Laplace transform. If we denote by $\mathbf{P}_{\boldsymbol{y}}$ the law of such a process with initial state $\boldsymbol{y} \in \mathbb{R}_+^\ell$, then

$$\mathbf{E}_{\boldsymbol{y}} \left[\mathrm{e}^{-[\boldsymbol{\theta}, \boldsymbol{Y}_t]} \right] = \exp\big\{ -[\boldsymbol{y}, \boldsymbol{v}_t(\boldsymbol{\theta})] \big\}, \qquad \text{for} \quad \boldsymbol{\theta} \in \mathbb{R}_+^\ell, t \geq 0, \tag{4.4}$$

where

$$t \mapsto \boldsymbol{v}_t(\boldsymbol{\theta}) = (\boldsymbol{v}_t(1, \boldsymbol{\theta}) \ldots, \boldsymbol{v}_t(\ell, \boldsymbol{\theta}))^{\mathtt{t}}$$

is the unique locally bounded solution, with non-negative entries, to the system of integral equations

$$\boldsymbol{v}_t(i, \boldsymbol{\theta}) = \theta_i - \int_0^t \psi(i, \boldsymbol{v}_{t-s}(\boldsymbol{\theta})) \mathrm{d}s, \quad i \in S. \tag{4.5}$$

Suppose that $(\boldsymbol{X}_t, \mathbb{P}_{\boldsymbol{\mu}})_{t \geq 0}$ is a $(\mathtt{P}, \boldsymbol{\psi})$-multitype superprocess and define the total mass vector as $\boldsymbol{Y} = (\boldsymbol{Y}_t, t \geq 0)$ with entries

$$Y_t(i) = X_t(i, \mathbb{R}^d) = \int_{\mathbb{R}^d} X_t(i, \mathrm{d}x), \qquad t \geq 0,$$

and initial vector $\boldsymbol{\mu} = (\mu_1(\mathbb{R}^d), \cdots, \mu_\ell(\mathbb{R}^d))^{\mathtt{t}}$. Let $\boldsymbol{\theta} \in \mathbb{R}_+^\ell$, and take $f_i(x) = \theta_i$ for each $i \in S, x \in \mathbb{R}^d$. Since the branching mechanism and the vector $\boldsymbol{\theta}$ are spatially independent, the system of functions $\boldsymbol{V}_t\boldsymbol{\theta}$ that satisfies (4.3) does not depend on $x \in \mathbb{R}^d$. In other words

$$\begin{aligned} V_t^{(i)}\boldsymbol{\theta} &= \mathtt{P}_t^{(i)}\theta_i - \int_0^t \mathrm{d}s \int_{\mathbb{R}^d} \psi(i, \boldsymbol{V}_{t-s}\boldsymbol{\theta}) \mathtt{P}_s^{(i)}(x, \mathrm{d}y) \\ &= \theta_i - \int_0^t \psi(i, \boldsymbol{V}_{t-s}\boldsymbol{\theta}) \mathrm{d}s, \quad i \in S. \end{aligned}$$

Recall that the previous system of equations has a unique solution, therefore $\boldsymbol{V}_t\boldsymbol{\theta} = \boldsymbol{v}_t(\boldsymbol{\theta})$ for any $x \in \mathbb{R}^d$. By (4.2) and the relationship between $\boldsymbol{X}$ and $\boldsymbol{Y}$, the total mass vector is indeed a MCB-process.

Since the total mass vector of a multitype superprocess is a MCB-process, we can determine its asymptotic behaviour through its first moment, similarly to the one-type case. More precisely, denote by $\boldsymbol{M}(t)$ the $\ell \times \ell$ matrix with elements

$$M(t)_{ij} = \mathbb{E}_{\boldsymbol{e}_i\delta_x}\Big[\langle \boldsymbol{e}_j, \boldsymbol{X}_t \rangle\Big], \quad i, j \in S,$$

where $\boldsymbol{e}_i\delta_x$ denotes a measure valued vector that has unit mass at position $x \in \mathbb{R}^d$, in the i-th coordinate, and zero mass everywhere else.

Barczy et al. [1] (see Lemma 3.4) proved that the mean matrix $\boldsymbol{M}(t)$ can be written in terms of the branching mechanism $\boldsymbol{\psi}$. In other words, for all $t > 0$

$$\boldsymbol{M}(t) = \mathrm{e}^{t\widetilde{\boldsymbol{B}}^{\mathtt{t}}},$$

where the matrix $\widetilde{\boldsymbol{B}}$ is given by

$$\widetilde{B}_{i,j} = B_{i,j} + \int_{\mathbb{R}_+^\ell} (y_i - \delta_{i,j} y_j)\Pi(j, \mathrm{d}\boldsymbol{y}).$$

Moreover, after straightforward computations (see for instance the computations after identity (2.15) in [1]) we observe that the branching mechanism $\boldsymbol{\psi}$ can be rewritten as follows

$$\psi(i,\boldsymbol{\theta}) := -[\boldsymbol{\theta}, \widetilde{\boldsymbol{B}}\boldsymbol{e}_i] + \beta_i\theta_i^2 + \int_{\mathbb{R}_+^\ell} \left(\mathrm{e}^{-[\boldsymbol{\theta},\boldsymbol{y}]} - 1 + [\boldsymbol{\theta}, \boldsymbol{y}]\right)\Pi(i, \mathrm{d}\boldsymbol{y}), \qquad \boldsymbol{\theta} \in \mathbb{R}_+^\ell, i \in S. \tag{4.6}$$

In the sequel, we assume that the matrix $\widetilde{\boldsymbol{B}}^{\mathtt{t}}$ is irreducible, and therefore the mean matrix $\boldsymbol{M}$ is irreducible as well. Then a Perron-Frobenius-type result (see Appendix A of [2]) guarantees that there exists a unique leading eigenvalue Γ, and right and left eigenvectors $\mathtt{u}, \mathtt{v} \in \mathbb{R}_+^\ell$, whose coordinates are strictly positive such that, for $t \geq 0$,

$$\boldsymbol{M}(t)\mathtt{u} = \mathrm{e}^{\Gamma t}\mathtt{u}, \qquad \widetilde{\boldsymbol{B}}^{\mathtt{t}}\mathtt{u} = \Gamma\mathtt{u}, \qquad \mathtt{v}^{\mathtt{t}}\boldsymbol{M}(t) = \mathrm{e}^{\Gamma t}\mathtt{v}^{\mathtt{t}}, \qquad \text{and} \qquad \mathtt{v}^{\mathtt{t}}\widetilde{\boldsymbol{B}}^{\mathtt{t}} = \Gamma\mathtt{v}^{\mathtt{t}}.$$

It is important to note, that since the branching mechanism is spatially independent, the value of Γ does not depend on the spatial variable.

Moreover, Γ determines the long term behaviour of $\boldsymbol{X}$. Indeed, employing the same terminology as in the one-type case, we call the process supercritical, critical or subcritical accordingly as Γ is strictly positive, equal to zero or strictly negative. In Kyprianou and Palau [14], the authors show that when $\Gamma \leq 0$ the total mass goes to zero almost surely. Barczy and Pap [2] show that if $\Gamma > 0$, then the total mass process satisfies

$$\lim_{t\to\infty} \mathrm{e}^{-\Gamma t}\, \mathbf{E}_{\boldsymbol{e}_i}[\boldsymbol{Y}_t] = \boldsymbol{e}_i^{\mathtt{t}}\mathtt{u}\mathtt{v}^{\mathtt{t}}, \qquad \text{for } i \in S,$$

which is a non-zero vector. Furthermore, Kyprianou and Palau in [14] also prove that, when $\Gamma > 0$ and the following "$x \log x$" condition holds

$$\sum_{i=1}^{\ell} \int_{1 \leq \langle \mathbf{1}, \boldsymbol{y}\rangle < \infty} \langle \mathbf{1}, \boldsymbol{y}\rangle \log(\langle \mathbf{1}, \boldsymbol{y}\rangle)\Pi(i, \mathrm{d}\boldsymbol{y}) < \infty, \tag{4.7}$$

then

$$\mathbb{P}_{\boldsymbol{e}_i\delta_x}\Big(\lim_{t\to\infty} \|\boldsymbol{X}_t\| = 0\Big) < 1, \qquad \text{for } i \in S, \quad x \in \mathbb{R}^d. \tag{4.8}$$

See also Theorem 1.4 in [15].

Here, we are also interested in the case that extinction occurs in finite time. More precisely, let us define $\mathcal{E} := \{\|\boldsymbol{X}_t\| = 0 \text{ for some } t > 0\}$, the event of *extinction* and take $w_i : \mathbb{R}^d \mapsto \mathbb{R}_+$ to be the function such that

$$w_i(x) := -\log \mathbb{P}_{\boldsymbol{e}_i\delta_x}(\mathcal{E}), \quad i \in S. \tag{4.9}$$

Since the branching mechanism is spatially independent, and the total mass vector of $\boldsymbol{X}$ is a MCB-process, we get that $w_i(x) = w_i$, for all $x \in \mathbb{R}^d$, for some constant w_i.

In what follows, we assume

$$0 < w_i < \infty, \qquad \text{for all } i \in S. \tag{4.10}$$

Assumption (4.10) or similar assumptions have been used in most of the cases where backbones have been constructed. For instance in [3] and [6], the authors assume Grey's condition which is equivalent to w_i being finite. In [4, 8, 15, 19] a very similar condition appears for the spatially dependent case. Assumption (4.10) is not only necessary for the construction of the multitype superprocess conditioned on extinction but also for the construction of the so-called Dynkin-Kuznetsov measure, as we will see below.

On the other hand, assumption (4.10) is not very restrictive. For instance, it is satisfied if $\Gamma > 0$, condition (4.7) holds and $\beta := \inf_{i\in S} \beta_i > 0$. Indeed from (4.8), we see that

$$\mathbb{P}_{\boldsymbol{e}_i\delta_x}(\mathcal{E}) \le \mathbb{P}_{\boldsymbol{e}_i\delta_x}\Big(\lim_{t\to\infty} \|\boldsymbol{X}_t\| = 0\Big) < 1.$$

From (4.4) and the fact that the total mass is a MCB-process, it is clear that

$$\mathbb{P}_{\boldsymbol{e}_i\delta_x}\Big(\|\boldsymbol{X}_t\| = 0\Big) = \exp\left\{-\lim_{\boldsymbol{\theta}\hookrightarrow\infty} v_t(i,\boldsymbol{\theta})\right\},$$

where $\boldsymbol{v}_t(i,\boldsymbol{\theta})$ is given by (4.5) and $\boldsymbol{\theta} \hookrightarrow \infty$ means that each coordinate of $\boldsymbol{\theta}$ goes to ∞. In other words, if we show that

$$\lim_{t\to\infty}\lim_{\boldsymbol{\theta}\hookrightarrow\infty} v_t(i,\boldsymbol{\theta}) < \infty \qquad \text{for all } i \in S,$$

then we have that (4.10) holds. In order to prove that the above limit is finite, we introduce

$$A_t(\boldsymbol{\theta}) := \sup_{i\in S} \frac{v_t(i,\boldsymbol{\theta})}{\mathtt{u}_i},$$

where $\mathtt{u}_i$ denotes the i-th coordinate of the right eigenvector associated to Γ. Since the supremum of finitely many continuously differentiable functions is differentiable except at most countably many isolated points, we may fix $t \ge 0$ such that $A_t(\boldsymbol{\theta})$ is differentiable at t and select i in such a way that $A_t(\boldsymbol{\theta})\mathtt{u}_i = v_t(i,\boldsymbol{\theta})$. Then by using (4.5) and (4.6) we can deduce that

$$\begin{aligned}\mathtt{u}_i\frac{\mathrm{d}}{\mathrm{d}t}A_t(\boldsymbol{\theta}) = \frac{\mathrm{d}}{\mathrm{d}t}v_t(i,\boldsymbol{\theta}) =& \sum_{j\in S}\widetilde{B}_{ji}\mathtt{u}_j\frac{\boldsymbol{v}_t(j,\boldsymbol{\theta})}{\mathtt{u}_j} - \beta_i(\boldsymbol{v}_t(i,\boldsymbol{\theta}))^2\\ &- \int_{\mathbb{R}_+^\ell}\Big(\mathrm{e}^{-[\boldsymbol{v}_t(\boldsymbol{\theta}),\boldsymbol{y}]} - 1 + [\boldsymbol{v}_t(\boldsymbol{\theta}),\boldsymbol{y}]\Big)\Pi(i,\mathrm{d}\boldsymbol{y}).\end{aligned}$$

Since $1-x-\mathrm{e}^{-x}\le 0$, for all $x>0$, $\widetilde{B}_{i,j}\mathbf{1}_{\{i\neq j\}}>0$ and $A_t(\boldsymbol{\theta})\mathtt{u}_i=v_t(i,\boldsymbol{\theta})$, we have

$$\mathtt{u}_i\frac{\mathrm{d}}{\mathrm{d}t}A_t(\boldsymbol{\theta})\le A_t(\boldsymbol{\theta})\sum_{j\in S}\widetilde{B}_{ji}\mathtt{u}_j-\beta_i(A_t(\boldsymbol{\theta})\mathtt{u}_i)^2=A_t(\boldsymbol{\theta})(\widetilde{\boldsymbol{B}}^{\mathtt{t}}\mathtt{u})_i-\beta_i(A_t(\boldsymbol{\theta})\mathtt{u}_i)^2.$$

Next, we use that $\mathtt{u}$ is an eigenvector of $\widetilde{\boldsymbol{B}}^{\mathtt{t}}$ to get

$$\mathtt{u}_i\frac{\mathrm{d}}{\mathrm{d}t}A_t(\boldsymbol{\theta})\le A_t(\boldsymbol{\theta})\Gamma\mathtt{u}_i-\beta_i(A_t(\boldsymbol{\theta})\mathtt{u}_i)^2.$$

By defining $\underline{\mathtt{u}}:=\inf_{i\in S}\mathtt{u}_i$ and recalling the definition of β, the previous identity implies

$$\frac{\mathrm{d}}{\mathrm{d}t}A_t(\boldsymbol{\theta})\le A_t(\boldsymbol{\theta})\Gamma-\beta\underline{\mathtt{u}}(A_t(\boldsymbol{\theta}))^2.$$

Since, Γ,β and $\underline{\mathtt{u}}$ are strictly positive, an integration by parts allow us to deduce that

$$A_t(\boldsymbol{\theta})\le\frac{\Gamma\mathrm{e}^{\Gamma t}}{\frac{\Gamma}{A_0(\boldsymbol{\theta})}+\beta\underline{\mathtt{u}}(\mathrm{e}^{\Gamma t}-1)}.$$

Finally, if we define $\bar{\mathtt{u}}:=\sup_{i\in S}\mathtt{u}_i$, the previous computations lead to

$$w_i=\lim_{t\to\infty}\lim_{\boldsymbol{\theta}\hookrightarrow\infty}v_t(i,\boldsymbol{\theta})\le\bar{\mathtt{u}}\lim_{t\to\infty}\lim_{\boldsymbol{\theta}\hookrightarrow\infty}A_t(\boldsymbol{\theta})\le\frac{\bar{\mathtt{u}}\Gamma}{\beta\underline{\mathtt{u}}}<\infty.$$

The following result is also needed for constructing the associated Dynkin-Kuznetsov measures which provide a way to dress the backbone.

Proposition 4.1.2. *Suppose that condition (4.10) holds. Then $\boldsymbol{\psi}(\boldsymbol{w})=\mathbf{0}$. Moreover, for $x\in\mathbb{R}^d$, $i\in S$ and $t>0$, we have that*

$$\mathbb{P}_{\boldsymbol{e}_i\delta_x}(\|\boldsymbol{X}_t\|=0)>0.$$

For simplicity of exposition, the proof of this result is presented in Section 4.2.

As we said before, our aim is to describe the backbone decomposition of $\boldsymbol{X}$. According to Berestycki et al. [3] a one-type supercritical superprocess can be decomposed into an initial burst of subcritical mass and three types of immigration processes along the backbone, which are two types of Poissonian immigrations and branch point immigrations. In order to use the same idea in the multitype case, we need to determine the components of this decomposition. These are the multitype branching diffusion process, that gives the prolific genealogies, and copies of the original multitype superprocess conditioned on extinction.

4.1.2 The multitype supercritical superdiffusion conditioned on extinction.

It is well known that under some conditions a supercritical CB-process can be conditioned to become extinct by conditioning the associated spectrally positive Lévy process to drift to $-\infty$. Such a conditioning appears as an Esscher transform on the underlying Lévy process in the Lamperti transform, where the shift parameter is given by the largest root of the branching mechanism. Here we show that a similar result still holds in the multitype case. In particular we have the following result.

Proposition 4.1.3. *For each $\boldsymbol{\mu} \in \mathcal{M}(\mathbb{R}^d)^\ell$, define the law of $\boldsymbol{X}$ with initial configuration $\boldsymbol{\mu}$ conditioned on becoming extinct by $\mathbb{P}^\dagger_{\boldsymbol{\mu}}$, and let $\mathcal{F}_t := \sigma(X_s, s \le t)$. Specifically, for all events A, measurable with respect to $\mathcal{F}$,*

$$\mathbb{P}^\dagger_{\boldsymbol{\mu}}(A) = \mathbb{P}_{\boldsymbol{\mu}}\left(A \,|\, \mathcal{E}\right).$$

Then, for all $\boldsymbol{f} \in \mathcal{B}^+(\mathbb{R}^d)^\ell$

$$\mathbb{E}^\dagger_{\boldsymbol{\mu}}\left[\mathrm{e}^{-\langle \boldsymbol{f}, \boldsymbol{X}_t\rangle}\right] = \exp\Big\{-\langle \boldsymbol{V}_t^\dagger \boldsymbol{f}, \boldsymbol{\mu}\rangle\Big\},$$

where

$$V_t^{\dagger,(i)}\boldsymbol{f}(x) := V_t^{(i)}(\boldsymbol{f} + \boldsymbol{w})(x) - w_i, \qquad i \in S,$$

is the unique locally bounded solution of

$$V_t^{\dagger,(i)}\boldsymbol{f}(x) = \mathtt{P}_t^{(i)} f_i(x) - \int_0^t \mathrm{d}s \int_{\mathbb{R}^d} \psi^\dagger(i, \boldsymbol{V}^\dagger_{t-s}\boldsymbol{f}(y))\mathtt{P}_s^{(i)}(x, \mathrm{d}y), \qquad i \in S, \tag{4.11}$$

where $\boldsymbol{\psi}^\dagger(\boldsymbol{\lambda}) := \boldsymbol{\psi}(\boldsymbol{\lambda} + \boldsymbol{w})$ and $\boldsymbol{w}$ is given by (4.9). In other words, $(\boldsymbol{X}, \mathbb{P}^\dagger_{\boldsymbol{\mu}})$ is a $(\mathtt{P}, \boldsymbol{\psi}^\dagger)$-multitype superprocess.

For simplicity of exposition, the proof of this result is presented in Section 4.2.

4.1.3 Dynkin-Kuznetsov measure.

As we mentioned before, a key ingredient in the construction of the backbone, or even the spine decomposition for superprocesses, is the so-called Dynkin-Kuznetsov measure. It is important to note that the existence of such measures was taken for granted in most of the references that appear in the literature, in particular in [3, 8, 15, 19]. Fortunately, from the assumptions and the way the dressing processes are constructed this omission does not play an important role on the validity of their results. Here, we provide a rigorous argument for their existence. See also Ren et al. [21] for the study of Dynkin-Kutznetsov measures for one-type superprocesses with non-local branching mechanism.

Let us denote by $\mathcal{X}$ the space of càdlàg paths from $[0, \infty)$ to $\mathcal{M}(\mathbb{R}^d)^\ell$.

Proposition 4.1.4. *Let $\boldsymbol{X}$ be a $(\mathtt{P}, \boldsymbol{\psi})$-multitype superprocess satisfying (4.10). For $x \in \mathbb{R}^d$, there exists a measure $\mathbb{N}_{x\boldsymbol{e}_i}$ on the space $\mathcal{X}$ satisfying*

$$\mathbb{N}_{x\boldsymbol{e}_i}\left(1 - \mathrm{e}^{-\langle \boldsymbol{f}, \boldsymbol{X}_t\rangle}\right) = -\log \mathbb{E}_{\delta_x \boldsymbol{e}_i}\left[\mathrm{e}^{-\langle \boldsymbol{f}, \boldsymbol{X}_t\rangle}\right], \tag{4.12}$$

for all $\boldsymbol{f} \in \mathcal{B}^+(\mathbb{R}^d)^\ell$ and $t \ge 0$.

Again, for simplicity of exposition we provide the proof of this Proposition in Section 4.2.

Following the same terminology as in the literature, we call $\{(\mathbb{N}_{x\boldsymbol{e}_i}, x \in \mathbb{R}^d), i \in S\}$ the Dynkin-Kuznetsov measures. We denote by $\mathbb{N}^\dagger$ the Dynkin-Kuznetsov measures associated to the multitype superprocess conditioned on extinction, which are also well defined (see the discussion after the proof of Proposition 4.1.4).

4.1.4 Prolific individuals.

Here, we consider those individuals of the superprocess who are responsible for the infinite growth of the process. In our case, we have that the so-called prolific individuals, i.e. those with an infinite genealogical line of descent, form a branching particle diffusion where the particles move according to the same motion semigroup as the superprocess itself, and their branching generator can be expressed in terms of the branching mechanism of the superprocess. Let $\boldsymbol{Z} = (\boldsymbol{Z}_t, t \geq 0)$ be a multitype branching diffusion process (MBDP) with ℓ types, where the movement of each particle of type $i \in S$ is given by the semigroup $\mathtt{P}^{(i)}$. The branching rate $\boldsymbol{q} \in \mathbb{R}_+^\ell$ takes the form

$$q_i = \frac{\partial}{\partial x_i}\psi(i, \boldsymbol{x})\bigg|_{\boldsymbol{x}=\boldsymbol{w}}, \qquad i \in S, \tag{4.13}$$

where $\boldsymbol{w}$ was defined in (4.9).

The offspring distribution $(p^{(i)}_{j_1,\dots,j_\ell})_{(j_1,\dots,j_\ell)\in\mathbb{N}^\ell}$ satisfies

$$p^{(i)}_{j_1,\dots,j_\ell} = \frac{1}{w_i q_i}\left(\beta_i w_i^2 \mathbf{1}_{\{\boldsymbol{j}=2\boldsymbol{e}_i\}} + \left(B_{ki}w_k + \int_{\mathbb{R}_+^\ell} w_k y_k \mathrm{e}^{-[\boldsymbol{w},\boldsymbol{y}]}\Pi(i, \mathrm{d}\boldsymbol{y})\right)\mathbf{1}_{\{\boldsymbol{j}=\boldsymbol{e}_k\}}\mathbf{1}_{\{i\neq k\}} \right.$$
$$\left. + \int_{\mathbb{R}_+^\ell} \frac{(w_1y_1)^{j_1}\dots(w_\ell y_\ell)^{j_\ell}}{j_1!\dots j_\ell!}\mathrm{e}^{-[\boldsymbol{w},\boldsymbol{y}]}\Pi(i,\mathrm{d}\boldsymbol{y})\mathbf{1}_{\{j_1+\dots+j_\ell\geq 2\}}\right), \tag{4.14}$$

where $\boldsymbol{j} = (j_1, \cdots, j_\ell)$. Note that $p^{(i)}_{j_1,\dots,j_\ell}$ is a probability distribution. Indeed, since

$\boldsymbol{\psi}(\boldsymbol{w}) = \boldsymbol{0}$, for each $i \in S$ we get that

$$\begin{aligned}
w_i q_i &= w_i q_i - \psi(i, \boldsymbol{w}) \\
&= w_i \left(-B_{ii} + 2\beta_i w_i + \int_{\mathbb{R}_+^\ell} \left(1 - \mathrm{e}^{-[\boldsymbol{w},\boldsymbol{y}]}\right) y_i \Pi(i, \mathrm{d}\boldsymbol{y}) \right) \\
&\qquad\qquad\qquad + [\boldsymbol{w}, \boldsymbol{B}\boldsymbol{e}_i] - \beta_i w_i^2 - \int_{\mathbb{R}_+^\ell} \left(\mathrm{e}^{-[\boldsymbol{w},\boldsymbol{y}]} - 1 + w_i y_i\right) \Pi(i, \mathrm{d}\boldsymbol{y}) \\
&= \sum_{j \neq i} B_{ji} w_j + \beta_i w_i^2 + \int_{\mathbb{R}_+^\ell} \mathrm{e}^{-[\boldsymbol{w},\boldsymbol{y}]} \left(\mathrm{e}^{[\boldsymbol{w},\boldsymbol{y}]} - 1 - w_i y_i\right) \Pi(i, \mathrm{d}\boldsymbol{y}) \\
&= \sum_{j \neq i} \left(B_{ji} w_j + \int_{\mathbb{R}_+^\ell} w_j y_j \mathrm{e}^{-[\boldsymbol{w},\boldsymbol{y}]} \Pi(i, \mathrm{d}\boldsymbol{y}) \right) \\
&\qquad\qquad\qquad + \beta_i w_i^2 + \int_{\mathbb{R}_+^\ell} \mathrm{e}^{-[\boldsymbol{w},\boldsymbol{y}]} \left(\mathrm{e}^{[\boldsymbol{w},\boldsymbol{y}]} - 1 - [\boldsymbol{w}, \boldsymbol{y}]\right) \Pi(i, \mathrm{d}\boldsymbol{y}) \\
&= \sum_{j \neq i} \left(B_{ji} w_j + \int_{\mathbb{R}_+^\ell} w_j y_j \mathrm{e}^{-[\boldsymbol{w},\boldsymbol{y}]} \Pi(i, \mathrm{d}\boldsymbol{y}) \right) + \beta_i w_i^2 \\
&\qquad\qquad\qquad + \int_{\mathbb{R}_+^\ell} \sum_{j_1 + \cdots + j_\ell \geq 2} \frac{(w_1 y_1)^{j_1} \ldots (w_\ell y_\ell)^{j_\ell}}{j_1! \ldots j_\ell!} \mathrm{e}^{-[\boldsymbol{w},\boldsymbol{y}]} \Pi(i, \mathrm{d}\boldsymbol{y}),
\end{aligned}$$

where in the last row we have used the multinomial theorem, i.e.

$$\sum_{n=2}^{\infty} \frac{[\boldsymbol{x}, \boldsymbol{y}]^n}{n!} = \sum_{n=2}^{\infty} \frac{1}{n!} \sum_{j_1 + \cdots + j_\ell = n} \binom{n}{j_1, \ldots, j_\ell} \prod_{k=1}^{\ell} (x_k y_k)^{j_k} = \sum_{j_1 + \cdots + j_\ell \geq 2} \frac{(x_1 y_1)^{j_1} \ldots (x_\ell y_\ell)^{j_\ell}}{j_1! \ldots j_\ell!}. \tag{4.15}$$

Let $\boldsymbol{F}(\boldsymbol{s}) = (F_1(\boldsymbol{s}), \ldots, F_\ell(\boldsymbol{s}))^{\mathtt{t}}$, $\boldsymbol{s} \in [0, 1]^\ell$, be the branching mechanism of $\boldsymbol{Z}$, which is given by

$$F_i(\boldsymbol{s}) = q_i \sum_{\boldsymbol{j} \in \mathbb{N}^\ell} (s_1^{j_1} \ldots s_\ell^{j_\ell} - s_i) p^{(i)}_{j_1, \ldots, j_\ell} = \frac{1}{w_i} \psi(i, \boldsymbol{w} \cdot (\boldsymbol{1} - \boldsymbol{s})), \qquad i \in S, \tag{4.16}$$

where we recall that $\boldsymbol{1}$ denotes the vector with value 1 in each coordinate and $\boldsymbol{u} \cdot \boldsymbol{v}$ is the element-wise multiplication of the vectors $\boldsymbol{u}$ and $\boldsymbol{v}$. The intuition behind the process $\boldsymbol{Z}$ is as follows. A particle of type i from its birth executes a $\mathsf{P}^{(i)}$ motion, and after an independent and exponentially distributed random time with parameter q_i dies and gives birth at its death position to an independent number of offspring with distribution $\{p^{(i)}_{\boldsymbol{j}}, \boldsymbol{j} \in \mathbb{N}^\ell\}$. We call $\boldsymbol{Z}$ the *backbone* of the multitype superprocess $\boldsymbol{X}$, and denote its initial distribution by $\boldsymbol{\nu} \in \mathcal{M}_a(\mathbb{R}^d)^\ell$, where $\mathcal{M}_a(\mathbb{R}^d)$ denotes the space of atomic measures on $\mathbb{R}^d$. Comparing the form of the offspring distribution between the one-type case and the multitype case, the main difference is that now we are allowed to have one offspring at a branching event. However in this case, that offspring has to have a different type from its parent.

4.1.5 The backbone decomposition.

Our primary aim is to give a decomposition of the $(\mathtt{P},\boldsymbol{\psi})$-multitype superprocess along its embedded backbone $\boldsymbol{Z}$. The main idea is to dress the process $\boldsymbol{Z}$ with immigration, where the processes we immigrate are copies of the $(\mathtt{P},\boldsymbol{\psi}^\dagger)$-multitype superprocess. The dressing relies on three different types of immigration mechanisms. These are two types of Poissonian immigrations along the life span of each prolific individual, and an additional creation of mass at the branch points of the embedded particle system. In the first case, we immigrate independent copies of the $(\mathtt{P},\boldsymbol{\psi}^\dagger)$-multitype superprocess, where the immigration rate along a particle of type $i\in S$ is related to a subordinator in $\mathbb{R}_+^\ell$, whose Laplace exponent is given by

$$\phi(i,\boldsymbol{\lambda}) = \frac{\partial}{\partial x_i}\psi^\dagger(i,\boldsymbol{x})\bigg|_{\boldsymbol{x}=\boldsymbol{\lambda}} - \frac{\partial}{\partial x_i}\psi^\dagger(i,\boldsymbol{x})\bigg|_{\boldsymbol{x}=\boldsymbol{0}} = \frac{\partial}{\partial x_i}\psi(i,x)\bigg|_{\boldsymbol{x}=\boldsymbol{\lambda}+\boldsymbol{w}} - \frac{\partial}{\partial x_i}\psi(i,\boldsymbol{x})\bigg|_{\boldsymbol{x}=\boldsymbol{w}},$$

which can be rewritten as

$$\phi(i,\boldsymbol{\lambda}) = 2\beta_i\lambda_i + \int_{\mathbb{R}_+^\ell}\left(1-\mathrm{e}^{-[\boldsymbol{\lambda},\boldsymbol{y}]}\right)y_i\mathrm{e}^{-[\boldsymbol{w},\boldsymbol{y}]}\Pi(i,\mathrm{d}\boldsymbol{y}). \tag{4.17}$$

When an individual of type $i\in S$ has branched and its offspring is given by $\boldsymbol{j}=(j_1,\dots,j_\ell)\in\mathbb{N}^\ell$, we immigrate an independent copy of the $(\mathtt{P},\boldsymbol{\psi}^\dagger)$-multitype superprocess where the initial mass has distribution

$$\begin{aligned}\eta_{\boldsymbol{j}}^{(i)}(\mathrm{d}\boldsymbol{y}) = \frac{1}{w_iq_ip_{\boldsymbol{j}}^{(i)}}\Bigg(&\beta_iw_i^2\mathbf{1}_{\{\boldsymbol{j}=2\boldsymbol{e}_i\}}\delta_{\boldsymbol{0}}(\mathrm{d}\boldsymbol{y})+\\ &\left(B_{ki}w_k\delta_{\boldsymbol{0}}(\mathrm{d}\boldsymbol{y})+w_ky_k\mathrm{e}^{-[\boldsymbol{w},\boldsymbol{y}]}\Pi(i,\mathrm{d}\boldsymbol{y})\right)\mathbf{1}_{\{\boldsymbol{j}=\boldsymbol{e}_k\}}\mathbf{1}_{\{i\neq k\}}\\ &+\frac{(w_1y_1)^{j_1}\dots(w_\ell y_\ell)^{j_\ell}}{j_1!\dots j_\ell!}\mathrm{e}^{-[\boldsymbol{w},\boldsymbol{y}]}\Pi(i,\mathrm{d}\boldsymbol{y})\mathbf{1}_{\{j_1+\dots+j_\ell\geq 2\}}\Bigg).\end{aligned} \tag{4.18}$$

Before we state our main results, we recall and introduce some notation. Recall that $\mathcal{X}$ denotes the space of càdlàg paths. Similarly to the one-type case, we use an Ulam-Harris labelling to reference the particles, and we denote the obtained tree by $\mathcal{T}$. For a particle $u\in\mathcal{T}$ let γ_u denote the type of the particle, τ_u its birth time, σ_u its death time, and $z_u(t)$ its spatial position at time t (whenever $\tau_u\leq t<\sigma_u$).

Definition 2. *For $\boldsymbol{\nu}\in\mathcal{M}_a(\mathbb{R}^d)^\ell$, let $\boldsymbol{Z}$ be a MBDP with initial configuration $\boldsymbol{\nu}$, and let $\widetilde{\boldsymbol{X}}$ be an independent copy of $\boldsymbol{X}$ under $\mathbb{P}_{\boldsymbol{\mu}}^\dagger$. We define the stochastic process $\boldsymbol{\Lambda}=(\boldsymbol{\Lambda}_t,t\geq 0)$ on $\mathcal{M}(\mathbb{R}^d)^\ell$ by*

$$\boldsymbol{\Lambda} = \widetilde{\boldsymbol{X}} + \boldsymbol{I}^{\mathbb{N}^\dagger} + \boldsymbol{I}^{\mathbb{P}^\dagger} + \boldsymbol{I}^{\eta},$$

where the processes $\boldsymbol{I}^{\mathbb{N}^\dagger}=(\boldsymbol{I}_t^{\mathbb{N}^\dagger},t\geq 0)$, $\boldsymbol{I}^{\mathbb{P}^\dagger}=(\boldsymbol{I}_t^{\mathbb{P}^\dagger},t\geq 0)$, and $\boldsymbol{I}^{\eta}=(\boldsymbol{I}_t^{\eta},t\geq 0)$ are independent of $\widetilde{\boldsymbol{X}}$ and, conditionally on $\boldsymbol{Z}$, are mutually independent. Moreover, these three processes are described pathwise as follows.

i) ***Continuous immigration.*** *The process $\boldsymbol{I}^{\mathbb{N}^\dagger}$ is $\mathcal{M}(\mathbb{R}^d)^\ell$-valued such that*

$$\boldsymbol{I}_t^{\mathbb{N}^\dagger} = \sum_{u\in\mathcal{T}} \sum_{t\wedge\tau_u \le r < t\wedge\sigma_u} \boldsymbol{X}_{t-r}^{(1,u,r)},$$

where, given $\boldsymbol{Z}$, independently for each $u \in \mathcal{T}$ such that $\tau_u < t$, the processes $\boldsymbol{X}^{(1,u,r)}$ are countable in number and correspond to $\mathcal{X}$-valued Poissonian immigration along the space-time trajectory $\{(z_u(r), r), r \in [\tau_u, t\wedge\sigma_u)\}$ with rate $2\beta_{\gamma_u}\mathrm{d}r \times \mathrm{d}\mathbb{N}^\dagger_{z_u(r)\boldsymbol{e}_{\gamma_u}}$.

ii) ***Discontinuous immigration.*** *The process $\boldsymbol{I}^{\mathbb{P}^\dagger}$ is $\mathcal{M}(\mathbb{R}^d)^\ell$-valued such that*

$$\boldsymbol{I}_t^{\mathbb{P}^\dagger} = \sum_{u\in\mathcal{T}} \sum_{t\wedge\tau_u \le r < t\wedge\sigma_u} \boldsymbol{X}_{t-r}^{(2,u,r)}$$

where, given $\boldsymbol{Z}$, independently for each $u \in \mathcal{T}$ such that $\tau_u \le t$, the processes $\boldsymbol{X}_\cdot^{(2,u,r)}$ are countable in number and correspond to $\mathcal{X}$-valued, Poissonian immigration along the space-time trajectory $\{(z_u(r), r), r \in [\tau_u, t\wedge\sigma_u)\}$ with rate

$$\mathrm{d}r \times \int_{\boldsymbol{y}\in\mathbb{R}_+^\ell} y_{\gamma_u} \mathrm{e}^{-[\boldsymbol{w},\boldsymbol{y}]}\Pi(\gamma_u, \mathrm{d}\boldsymbol{y}) \times \mathrm{d}\mathbb{P}^\dagger_{\boldsymbol{y}\delta_{z_u(r)}}.$$

iii) ***Branch point based immigration.*** *The process $\boldsymbol{I}^\eta$ is $\mathcal{M}(\mathbb{R}^d)^\ell$-valued such that*

$$\boldsymbol{I}_t^\eta = \sum_{u\in\mathcal{T}} \mathbf{1}_{\{\sigma_u\le t\}} \boldsymbol{X}_{t-\sigma_u}^{(3,u)}$$

where, given $\boldsymbol{Z}$, independently for each $u \in \mathcal{T}$ such that $\sigma_u \le t$, the process $\boldsymbol{X}_\cdot^{(3,u)}$ is an independent copy of $\boldsymbol{X}$ issued at time σ_u with law $\mathbb{P}_{\boldsymbol{Y}_u\delta_{z_u(\sigma_u)}}$ where $\boldsymbol{Y}_u$ is an independent random variable with distribution $\eta^{(\gamma_u)}_{\mathcal{N}_1^u,\dots,\mathcal{N}_\ell^u}(\mathrm{d}\boldsymbol{y})$. Here $(\mathcal{N}_1^u, \dots, \mathcal{N}_\ell^u)$ is the offspring of u, i.e. $\mathcal{N}_i^u$ is the number of offspring of type i.

Moreover, we denote the law of the pair $(\boldsymbol{\Lambda}, \boldsymbol{Z})$ by $\widehat{\mathbb{P}}_{(\boldsymbol{\mu},\boldsymbol{\nu})}$.

Since $\boldsymbol{Z}$ is a MBDP and, given $\boldsymbol{Z}$, immigrating mass occurs independently according to a Poisson point process or at the splitting times of $\boldsymbol{Z}$, we can deduce that the process $((\boldsymbol{\Lambda}, \boldsymbol{Z}), \widehat{\mathbb{P}}_{(\boldsymbol{\mu},\boldsymbol{\nu})})$ is Markovian. It is important to note that the mass which has immigrated up to a fixed time evolves in a Markovian way thanks to the branching property.

Now we are ready to state the main results of the paper. Our first result determines the law of the couple $(\boldsymbol{\Lambda}, \boldsymbol{Z})$, and in particular shows that $\boldsymbol{\Lambda}$ is conservative.

Theorem 4.1.5. *For $\boldsymbol{\mu} \in \mathcal{M}(\mathbb{R}^d)^\ell$, $\boldsymbol{\nu} \in \mathcal{M}_a(\mathbb{R}^d)^\ell$, $\boldsymbol{f}, \boldsymbol{h} \in \mathcal{B}^+(\mathbb{R}^d)^\ell$, and $t \ge 0$ we have*

$$\widehat{\mathbb{E}}_{(\boldsymbol{\mu},\boldsymbol{\nu})}\left[\mathrm{e}^{-\langle \boldsymbol{f},\boldsymbol{\Lambda}_t\rangle - \langle \boldsymbol{h},\boldsymbol{Z}_t\rangle}\right] = \exp\left\{-\langle \boldsymbol{V}_t^\dagger \boldsymbol{f}, \boldsymbol{\mu}\rangle - \langle \boldsymbol{U}_t^{(\boldsymbol{f})}\boldsymbol{h}, \boldsymbol{\nu}\rangle\right\}, \tag{4.19}$$

where $\boldsymbol{V}^{\dagger}$ *is defined in (4.11), and* $\exp\{-\boldsymbol{U}_t^{(\boldsymbol{f})}\boldsymbol{h}(x)\} = (\exp\{-U_t^{(\boldsymbol{f},1)}\boldsymbol{h}(x)\},\cdots,\exp\{-U_t^{(\boldsymbol{f},\ell)}\boldsymbol{h}(x)\})^{\mathtt{t}} : \mathbb{R}^d \to \mathbb{R}_+^{\ell}$ *is the unique* $[0,1]^{\ell}$*-valued solution to the system of integral equations*

$$\mathrm{e}^{-U_t^{(\boldsymbol{f},i)}\boldsymbol{h}(x)} = \mathtt{P}_t^{(i)}\mathrm{e}^{-h_i(x)} + \frac{1}{w_i}\int_0^t \mathrm{d}s \int_{\mathbb{R}^d} \Big[\psi^{\dagger}\left(i, -\boldsymbol{w}\cdot\mathrm{e}^{-\boldsymbol{U}_{t-s}^{(\boldsymbol{f})}\boldsymbol{h}(y)} + \boldsymbol{V}_{t-s}^{\dagger}\boldsymbol{f}(y)\right) - \psi^{\dagger}(i, \boldsymbol{V}_{t-s}^{\dagger}\boldsymbol{f}(y))\Big]\,\mathtt{P}_s^{(i)}(x,\mathrm{d}y) \tag{4.20}$$

for $x \in \mathbb{R}^d$*, and* $t \geq 0$*. In particular, for each* $t \geq 0$*,* $\boldsymbol{\Lambda}_t$ *has almost surely finite mass.*

Finally, we state the main result of this paper which, actually, is a consequence of Theorem 4.1.5. To be more precise, we consider a randomised version of the law $\mathbb{P}_{(\boldsymbol{\nu},\boldsymbol{\mu})}$ by replacing the deterministic choice of $\boldsymbol{\nu}$ in such a way that for each $i \in S$, ν_i is a Poisson random measure in $\mathbb{R}^d$ having intensity $w_i\mu_i$. The resulting law is denoted by $\widehat{\mathbb{P}}_{\boldsymbol{\mu}}$.

Theorem 4.1.6. *For any* $\boldsymbol{\mu} \in \mathcal{M}(\mathbb{R}^d)^{\ell}$ *the process* $(\boldsymbol{\Lambda}, \widehat{\mathbb{P}}_{\boldsymbol{\mu}})$ *is Markovian and has the same law as* $(\boldsymbol{X}, \mathbb{P}_{\boldsymbol{\mu}})$*.*

The remainder of this paper is devoted to the proofs of all the results presented in the Introduction.

4.2 Proofs

We first present the proofs of Propositions 4.1.1,4.1.2 and 4.1.4 which are devoted to the construction of the multitype superprocess $\boldsymbol{X}$ and its associated Dynkin-Kuznetsov measures.

Proof of Proposition 4.1.1. Recall that $(\mathtt{P}_t^{(i)}, t \geq 0)$ denotes the semigroup of the diffusion $(\xi_t^{(i)}, t \geq 0)$. We introduce $\Xi = (\Xi_t, t \geq 0)$ a Markov process in the product space $\mathbb{R}^d \times S$ whose transition semigroup $(\mathtt{T}_t, t \geq 0)$ is given by

$$\mathtt{T}_t f(x,i) = \int_{\mathbb{R}^d} f(y,i)\mathtt{P}_t^{(i)}(x,\mathrm{d}y) \qquad \text{for} \quad x \in \mathbb{R}^d, \tag{4.21}$$

where f is a bounded Borel function on $\mathbb{R}^d \times S$. We denote the aforementioned set of functions by $\mathcal{B}(\mathbb{R}^d \times S)$ and we use $\mathcal{M}(\mathbb{R}^d \times S)$ for the space of finite Borel measures on $\mathbb{R}^d \times S$, endowed with the topology of weak convergence.

For each $f \in \mathcal{B}(\mathbb{R}^d \times S)$, we introduce the operator

$$\Psi(x,i,f) = \psi(i,(f(x,1),\cdots,f(x,\ell))).$$

Recall that for a measure $\mu \in \mathcal{M}(\mathbb{R}^d \times S)$, we use the notation

$$\langle f, \mu\rangle = \int_{\mathbb{R}^d\times S} f(x,i)\mu(\mathrm{d}(x,i)).$$

Following the theory developed in the monograph of Li [20], we observe that the operator Ψ satisfies equation (2.26) in [20], and that the assumptions of Theorems 2.21 and 5.6, in the same monograph, are fulfilled. Therefore there exits a strong Markov superprocess $\mathcal{Z} = (\mathcal{Z}_t, \mathcal{G}_t, \mathbb{Q}_\mu)$ with state space $\mathcal{M}(\mathbb{R}^d \times S)$, and transition probabilities determined by

$$\mathbb{Q}_\mu\left[\mathrm{e}^{-\langle f, \mathcal{Z}_t\rangle}\right] = \exp\Big\{ - \langle \mathtt{V}_t f, \mu\rangle\Big\}, \qquad t \geq 0,$$

where $f \in \mathcal{B}(\mathbb{R}^d \times S)$ and $t \mapsto \mathtt{V}_t f$ is the unique locally bounded positive solution to

$$\mathtt{V}_t f(x,i) = \mathtt{T}_t f(x,i) - \int_0^t \mathrm{d}s \int_{\mathbb{R}^d \times S} \Psi(y, j, \mathtt{V}_{t-s} f)\mathtt{T}_s(x, i, \mathrm{d}(y,j)).$$

For $i \in S$ and $\mu \in \mathcal{M}(\mathbb{R}^d \times S)$, we define $\mathtt{U}_i\mu \in \mathcal{M}(\mathbb{R}^d)$ by $\mathtt{U}_i\mu(B) = \mu(B \times \{i\})$ for $B \in \mathcal{B}(\mathbb{R}^d)$, the Borel sets in $\mathbb{R}^d$. Observe that $\mu \mapsto (\mathtt{U}_i\mu)_{i\in S}$ is a homeomorphism between $\mathcal{M}(\mathbb{R}^d \times S)$ and $\mathcal{M}(\mathbb{R}^d)^\ell$. In other words, we can define a strong Markov process $\boldsymbol{X} \in \mathcal{M}(\mathbb{R}^d)^\ell$ associated with $\mathcal{Z}$ and $(U_i)_{i\in S}$ as follows. For each $i \in S$, we define $X_t(i, \mathrm{d}x) := \mathtt{U}_i\mathcal{Z}_t(\mathrm{d}x) = \mathcal{Z}_t(\mathrm{d}x \times \{i\})$ with probabilities $\mathbb{P}_{\boldsymbol{\mu}} := \mathbb{Q}_\mu$, where $\boldsymbol{\mu} = (\mu_1, \cdots, \mu_\ell) \in \mathcal{M}(\mathbb{R}^d)^\ell$, and each $\mu_i = \mathtt{U}_i\mu$. In a similar way, there is a homeomorphism between $\mathcal{B}(\mathbb{R}^d)^\ell$ and $\mathcal{B}(\mathbb{R}^d \times S)$; that is to say for $\boldsymbol{f} \in \mathcal{B}(\mathbb{R}^d)^\ell$ we define $f(x,i) = f_i(x)$. By applying the aforementioned homeomorphisms, we deduce that $(\boldsymbol{X}_t, \mathbb{P}_{\boldsymbol{\mu}})$ satisfies (4.2), and (4.3) has a unique locally bounded solution. $\square$

We now prove Proposition 4.1.2, which will be very useful for the existence of Dynkin-Kutznetsov measures.

Proof of Proposition 4.1.2. By (4.9) and the branching property of $\boldsymbol{X}$ we have

$$\mathbb{P}_{\boldsymbol{\mu}}(\mathcal{E}) = \mathrm{e}^{-\langle \boldsymbol{w}, \boldsymbol{\mu}\rangle}. \tag{4.22}$$

Furthermore by conditioning the event $\mathcal{E}$ on $\mathcal{F}_t$ and using the Markov property, we obtain that

$$\mathrm{e}^{-\langle \boldsymbol{w}, \boldsymbol{\mu}\rangle} = \mathbb{E}_{\boldsymbol{\mu}}\Big[\mathbb{E}[\mathbf{1}_{\mathcal{E}}|\mathcal{F}_t]\Big] = \mathbb{E}_{\boldsymbol{\mu}}\Big[\mathbb{E}_{\boldsymbol{X}_t}[\mathbf{1}_{\mathcal{E}}]\Big] = \mathbb{E}_{\boldsymbol{\mu}}\left[\mathrm{e}^{-\langle \boldsymbol{w}, \boldsymbol{X}_t\rangle}\right].$$

Thus from (4.3) and the assumption (4.10) we also get that $\boldsymbol{\psi}(\boldsymbol{w}) = \mathbf{0}$.

For the second part of the statement, we recall the definition of the total mass vector $\boldsymbol{Y} = (\boldsymbol{Y}_t, t \geq 0)$ whose entries satisfy $Y_t(i) = X_t(i, \mathbb{R}^d)$. From identity (4.4) and assumption (4.10), we know that for each $i \in S$, there exists a positive deterministic time T_i such that

$$\mathbf{P}_{\boldsymbol{e}_i}(\|\boldsymbol{Y}_t\| = 0) = \mathrm{e}^{-\lim_{\boldsymbol{\theta} \hookrightarrow \infty} v_t(i, \boldsymbol{\theta})} \begin{cases} = 0 & \text{for } t < T_i, \\ > 0 & \text{for } t > T_i, \end{cases}$$

where $\boldsymbol{v}_t(i, \boldsymbol{\theta})$ is given by (4.5), and we recall that $\boldsymbol{\theta} \hookrightarrow \infty$ means that each coordinate of $\boldsymbol{\theta}$ goes to ∞.

Next, we define the sets $S_1 := \{i \in S : T_i = 0\}$ and $S_2 := \{i \in S : T_i > 0\}$. For a vector $\boldsymbol{y} = (y_1, \cdots, y_\ell)$, we denote its support by $\mathrm{supp}(\boldsymbol{y}) := \{i \in S : y_i \neq 0\}$. Thus, the proof will be completed if we show that $S_2 = \emptyset$. We proceed by contradiction.

Let us assume that $S_2 \neq \emptyset$ and define $T := \inf\{T_i : i \in S_2\}$ which is strictly positive by definition. Take $i \in S_2$ and observe from the Markov property that

$$\begin{aligned} 0 &= \mathbf{P}_{\boldsymbol{e}_i}\Big(\|\boldsymbol{Y}_{3T/4}\| = 0\Big) \\ &\geq \mathbf{P}_{\boldsymbol{e}_i}\Big(\|\boldsymbol{Y}_{3T/4}\| = 0, \mathrm{supp}(\boldsymbol{Y}_{T/2}) \subset S_1\Big) \\ &= \mathbf{E}_{\boldsymbol{e}_i}\Big[\mathbf{P}_{\boldsymbol{Y}_{T/2}}\Big(\|\boldsymbol{Y}_{T/4}\| = 0\Big), \mathrm{supp}(\boldsymbol{Y}_{T/2}) \subset S_1\Big]. \end{aligned}$$

By the branching property, if $\boldsymbol{y}$ is a vector such that $\mathrm{supp}(\boldsymbol{y}) \subset S_1$ then $\mathbf{P}_{\boldsymbol{y}}(\|\boldsymbol{Y}_t\| = 0) > 0$, for all $t > 0$. Therefore, we necessarily have

$$0 = \mathbf{P}_{\boldsymbol{e}_i}\Big(\mathrm{supp}(\boldsymbol{Y}_{T/2}) \subset S_1\Big),$$

and implicitly

$$1 = \mathbf{P}_{\boldsymbol{e}_i}\Big(\mathrm{supp}(\boldsymbol{Y}_{T/2}) \cap S_2 \neq \emptyset\Big) = \mathbf{P}_{\boldsymbol{e}_i}\Big(\|\boldsymbol{Y}_{T/2}\| > 0\Big), \qquad \text{for all } i \in S_2.$$

Hence, using the branching property again, if $\boldsymbol{y}$ is a vector such that $\mathrm{supp}(\boldsymbol{y}) \cap S_2 \neq \emptyset$, we have

$$1 = \mathbf{P}_{\boldsymbol{y}}\Big(\|\boldsymbol{Y}_{T/2}\| > 0\Big) = \mathbf{P}_{\boldsymbol{y}}\Big(\mathrm{supp}(\boldsymbol{Y}_{T/2}) \cap S_2 \neq \emptyset\Big).$$

Finally, we use the Markov property recursively and the previous equality, to deduce that for all $k \geq 1$,

$$\mathbf{P}_{\boldsymbol{y}}\Big(\|\boldsymbol{Y}_{kT/2}\| > 0\Big) = 1 \qquad \text{for all } i \in S_2,$$

which is inconsistent with the definitions of T and T_i. In other words, $S_2 = \emptyset$. This completes the proof. $\square$

We now prove the existence of the Dynkin-Kuznetsov measures.

Proof of Proposition 4.1.4. Let us denote by $\mathcal{M}^0(\mathbb{R}^d \times S) := \mathcal{M}(\mathbb{R}^d \times S) \setminus \{0\}$, where 0 is the null measure. Consider the Markov superprocess $\mathcal{Z}$ introduced in the previous proof. Let $(\mathtt{Q}_t, t \geq 0)$ and $(\mathtt{V}_t, t \geq 0)$ be the transition and cumulant semigroups associated with $\mathcal{Z}$. By Theorem 1.36 in [20], $\mathtt{V}_t$ has the following representation

$$\mathtt{V}_t f(x,i) = \int_{\mathbb{R}^d \times S} f(y,j)\Lambda_t(x,i,\mathrm{d}(y,j)) + \int_{\mathcal{M}^0(\mathbb{R}^d\times S)} \Big(1 - \mathrm{e}^{-\langle f,\nu\rangle}\Big) L_t(x,i,\mathrm{d}\nu), \quad t \geq 0,$$

where f is a positive Borel function on $\mathbb{R}^d \times S$, $\Lambda_t(x,i,\mathrm{d}(y,j))$ is a bounded kernel on $\mathbb{R}^d \times S$, and $(1 \wedge \langle 1, \nu\rangle)L_t(x,i,\mathrm{d}\nu)$ is a bounded kernel from $\mathbb{R}^d \times S$ to $\mathcal{M}^0(\mathbb{R}^d \times S)$.

Let $\widetilde{\mathcal{X}}^+$ be the space of càdlàg paths $t \to \widetilde{w}_t$ from $[0,\infty)$ to $\mathcal{M}(\mathbb{R}^d \times S)$ having the null measure as a trap. Let $(\mathtt{Q}^0_t, t \geq 0)$ be the restriction of $(\mathtt{Q}_t, t \geq 0)$ to $\mathcal{M}^0(\mathbb{R}^d \times S)$ and

$$E_0 := \Big\{(x,i) \in \mathbb{R}^d \times S : \Lambda_t(x,i,\mathbb{R}^d \times S) = 0, \ \text{ for all } \ t > 0\Big\}.$$

By Proposition 2.8 in [20], for all $(x,i)\in E_0$ the family of measures $(L_t(x,i,\cdot), t\geq 0)$ on $\mathcal{M}^0(\mathbb{R}^d\times S)$ constitutes an entrance law for $(\mathtt{Q}^0_t, t\geq 0)$. Therefore, by Theorem A.40 of [20] for all $(x,i)\in E_0$ there exists a unique σ-finite measure $\widetilde{\mathbb{N}}_{(x,i)}$ on $\widetilde{\mathcal{X}}^+$ such that $\widetilde{\mathbb{N}}_{(x,i)}(\{0\})=0$, and for any $0<t_1<\cdots<t_n<\infty$

$$\widetilde{\mathbb{N}}_{(x,i)}(\mathcal{Z}_{t_1}\in \mathrm{d}\nu_1, \mathcal{Z}_{t_2}\in \mathrm{d}\nu_2,\dots,\mathcal{Z}_{t_n}\in \mathrm{d}\nu_n) \\ = L_{t_1}(x,i,\mathrm{d}\nu_1)\mathtt{Q}^0_{t_2-t_1}(\nu_1,\mathrm{d}\nu_2)\dots \mathtt{Q}^0_{t_n-t_{n-1}}(\nu_{n-1},\mathrm{d}\nu_n).$$

It follows that for all $t>0$, $(x,i)\in E_0$, and $f\in\mathcal{B}(\mathbb{R}^d\times S)$ positive, we have

$$\widetilde{\mathbb{N}}_{(x,i)}\left(1-\mathrm{e}^{-\langle f,\mathcal{Z}_t\rangle}\right)=\int_{\mathcal{M}^0(\mathbb{R}^d\times S)}\left(1-\mathrm{e}^{-\langle f,\nu\rangle}\right)L_t(x,i,\mathrm{d}\nu)=\mathtt{V}_t f(x,i).$$

Recall the homeomorphism $\mu\mapsto (U_i\mu)_{i\in S}$ and the definition of the superprocess $\boldsymbol{X}$ from the proof of Proposition 4.1.1. By taking the constant function $f(x,i)=\lambda\in\mathbb{R}$, and using the definitions of $\mathtt{V}_t,\mathtt{Q}_t$, we deduce that

$$-\log \mathbb{E}_{\boldsymbol{e}_i\delta_x}\left[\mathrm{e}^{-\lambda\langle \mathbf{1},\boldsymbol{X}_t\rangle}\right]=\lambda\Lambda_t(x,i,\mathbb{R}^d\times S)+\int_{\mathcal{M}^0(\mathbb{R}^d\times S)}\left(1-\mathrm{e}^{-\lambda\langle 1,\nu\rangle}\right)L_t(x,i,\mathrm{d}\nu).$$

If we take λ goes to infinity, the left hand side of the above identity converges to $-\log\mathbb{P}_{\boldsymbol{e}_i\delta_x}(\|\boldsymbol{X}_t\|=0)$ which is finite by Proposition (4.1.2). Henceforth, $\Lambda_t(x,i,\mathbb{R}^d\times S)=0$ and $(x,i)\in E_0$.

Next, recall that $\mathcal{X}$ denotes the space of càdlàg paths from $[0,\infty)$ to $\mathcal{M}(\mathbb{R}^d)^\ell$. Then $(\mathtt{U}_i)_{i\in S}$ induces a homeomorphism between $\widetilde{\mathcal{X}}$ and $\mathcal{X}$. More precisely, the homeomorphism $\mathcal{U}:\widetilde{\mathcal{X}}\to\mathcal{X}$ is given by $\widetilde{w}_t\to\boldsymbol{w}_t=(w_t(1),\cdots,w_t(\ell))$ where for all $i\in S$ the measure in the ith coordinate is given by $w_t(i,B)=\widetilde{w}_t(B\times\{i\})$. This implies that for all $(x,i)\in\mathbb{R}^d\times S$ we can define the measures $\mathbb{N}_{x\boldsymbol{e}_i}$ on $\mathcal{X}$ given by $\mathbb{N}_{x\boldsymbol{e}_i}(B):=\widetilde{\mathbb{N}}_{(x,i)}(\mathcal{U}^{-1}(B))$. In other words, we obtain

$$\mathbb{N}_{x\boldsymbol{e}_i}\left(1-\mathrm{e}^{-\langle \boldsymbol{f},\boldsymbol{X}_t\rangle}\right)=-\log\mathbb{E}_{\boldsymbol{e}_i\delta_x}\left[\mathrm{e}^{-\langle \boldsymbol{f},\boldsymbol{X}_t\rangle}\right],$$

for all $\boldsymbol{f}\in\mathcal{B}(\mathbb{R}^d)^\ell$ and $t\geq 0$.

□

It is important to note that the Dynkin-Kuznetsov measures $\mathbb{N}^\dagger$ associated to the multi-type superprocess conditioned on extinction are also well defined since $|\log\mathbb{P}^\dagger_{\delta_x\boldsymbol{e}_i}(\mathcal{E})|<\infty$.

We now prove Proposition 4.1.3.

Proof of Proposition 4.1.3. Using (4.22), (4.10) and the Markov property, we have for

$\boldsymbol{f} \in \mathcal{B}^+(\mathbb{R}^d)^\ell$

$$\begin{aligned}
\mathbb{E}^{\dagger}_{\boldsymbol{\mu}}\left[\mathrm{e}^{-\langle \boldsymbol{f}, \boldsymbol{X}_t\rangle}\right] &= \mathrm{e}^{\langle \boldsymbol{w}, \boldsymbol{\mu}\rangle}\mathbb{E}_{\boldsymbol{\mu}}\left[\mathrm{e}^{-\langle \boldsymbol{f}, \boldsymbol{X}_t\rangle}\mathbf{1}_{\mathcal{E}}\right] \\
&= \mathrm{e}^{\langle \boldsymbol{w}, \boldsymbol{\mu}\rangle}\mathbb{E}_{\boldsymbol{\mu}}\left[\mathrm{e}^{-\langle \boldsymbol{f}, \boldsymbol{X}_t\rangle}\mathbb{P}_{\boldsymbol{X}_t}(\mathcal{E})\right] \\
&= \mathrm{e}^{\langle \boldsymbol{w}, \boldsymbol{\mu}\rangle}\mathbb{E}_{\boldsymbol{\mu}}\left[\mathrm{e}^{-\langle \boldsymbol{f}, \boldsymbol{X}_t\rangle}\mathrm{e}^{-\langle \boldsymbol{w}, \boldsymbol{X}_t\rangle}\right] \\
&= \mathrm{e}^{-\langle \boldsymbol{V}_t(\boldsymbol{f}+\boldsymbol{w})-\boldsymbol{w}, \boldsymbol{\mu}\rangle}.
\end{aligned}$$

Since $\boldsymbol{V}_t(\boldsymbol{f}+\boldsymbol{w})$ satisfies (4.3), using the definitions of $\boldsymbol{V}^{\dagger}_t\boldsymbol{f}$ and $\boldsymbol{\psi}^{\dagger}$ we obtain that $\boldsymbol{V}^{\dagger}_t\boldsymbol{f}$ satisfies (4.11). Recalling that $\boldsymbol{\psi}(\boldsymbol{w}) = \mathbf{0}$ and computing $\boldsymbol{\psi}(\boldsymbol{\theta}+\boldsymbol{w}) - \boldsymbol{\psi}(\boldsymbol{w})$, we deduce that

$$\psi^{\dagger}(i, \boldsymbol{\theta}) = -[\boldsymbol{\theta}, \boldsymbol{B}^{\dagger}\boldsymbol{e}_i] + \beta_i\theta_i^2 + \int_{\mathbb{R}^\ell_+}\left(\mathrm{e}^{-[\boldsymbol{\theta},\boldsymbol{y}]} - 1 + \theta_i y_i\right)\mathrm{e}^{-[\boldsymbol{w},\boldsymbol{y}]}\Pi(i, \mathrm{d}\boldsymbol{y}), \tag{4.23}$$

where

$$B^{\dagger}_{ij} = B_{ij} - \left(2\beta_i w_i + \int_{\mathbb{R}^\ell_+}\left(1 - \mathrm{e}^{-[\boldsymbol{y},\boldsymbol{w}]}\right) y_i\Pi(i, \mathrm{d}\boldsymbol{y})\right)\mathbf{1}_{\{j=i\}}. \tag{4.24}$$

This implies that $\boldsymbol{\psi}^{\dagger}$ is a branching mechanism and therefore the solution of (4.11) is unique. In other words, $\boldsymbol{X}$ under $\mathbb{P}^{\dagger}_{\boldsymbol{\mu}}$ is a multitype superprocess with branching mechanism given by $\boldsymbol{\psi}^{\dagger}(\boldsymbol{\theta})$. □

In order to proceed with the proof of Theorem 4.1.5, the following two lemmas are necessary.

Lemma 4.2.1. *For each $\boldsymbol{f} \in \mathcal{B}(\mathbb{R}^d)^\ell$, $\boldsymbol{\nu} \in \mathcal{M}_a(\mathbb{R}^d)^\ell$, $\boldsymbol{\mu} \in \mathcal{M}(\mathbb{R}^d)^\ell$, and $t \geq 0$ we have*

$$\widehat{\mathbb{E}}_{(\boldsymbol{\mu},\boldsymbol{\nu})}\left[\mathrm{e}^{-\langle \boldsymbol{f}, \boldsymbol{I}^{\mathbb{N}^{\dagger}}_t + \boldsymbol{I}^{\mathbb{P}^{\dagger}}_t\rangle}\middle| (\boldsymbol{Z}_s, s \leq t)\right] = \exp\left\{-\int_0^t \langle \boldsymbol{\phi}(\boldsymbol{V}^{\dagger}_{t-r}\boldsymbol{f}), \boldsymbol{Z}_r\rangle \mathrm{d}r\right\}, \tag{4.25}$$

where $\boldsymbol{\phi}$ is given by (4.17) and $\boldsymbol{V}^{\dagger}_t\boldsymbol{f}$ satisfies (4.11).

Proof. As the different immigration mechanisms are independent given the backbone, we may look at the Laplace functional of the continuous and discontinuous immigrations separately. For the continuous immigration, we can condition on $\boldsymbol{Z}$, use Campbell's formula, then equation (4.12) for $\mathbb{N}^{\dagger}$, and finally the definition of $\boldsymbol{V}^{\dagger}_t\boldsymbol{f}(x) = (V^{\dagger,(1)}_t\boldsymbol{f}(x), \cdots, V^{\dagger,(\ell)}_t\boldsymbol{f}(x))^{\mathbf{t}}$ to obtain

$$\begin{aligned}
&\widehat{\mathbb{E}}_{(\boldsymbol{\mu},\boldsymbol{\nu})}\left[\exp\{-\langle \boldsymbol{f}, \boldsymbol{I}^{\mathbb{N}^{\dagger}}_t\rangle\}\middle| (\boldsymbol{Z}_s, s \leq t)\right] \\
&\qquad = \exp\left\{-\sum_{u\in\mathcal{T}} 2\beta_{\gamma_u}\int_{t\wedge\tau_u}^{t\wedge\sigma_u}\mathrm{d}r\mathbb{N}^{\dagger}_{z_u(r)\boldsymbol{e}_{\gamma_u}}\left(1 - \mathrm{e}^{-\langle \boldsymbol{f}, \boldsymbol{X}_{t-r}\rangle}\right)\right\} \\
&\qquad = \exp\left\{-\sum_{u\in\mathcal{T}} 2\beta_{\gamma_u}\int_{t\wedge\tau_u}^{t\wedge\sigma_u}\mathrm{d}r V^{\dagger,(\gamma_u)}_{t-r}\boldsymbol{f}(z_u(r))\right\}.
\end{aligned}$$

In a similar way, for the discontinuous immigration, by conditioning on $\boldsymbol{Z}$, using Campbell's formula and the definition of $\boldsymbol{V}_t^{\dagger}\boldsymbol{f}$ we get

$$\begin{aligned}
&\widehat{\mathbb{E}}_{(\boldsymbol{\mu},\boldsymbol{\nu})}\left[\exp\{-\langle \boldsymbol{f}, \boldsymbol{I}_t^{\mathbb{P}^{\dagger}}\rangle\}\Big| (\boldsymbol{Z}_s, s\le t)\right] \\
&\qquad = \exp\left\{-\sum_{u\in\mathcal{T}}\int_{t\wedge\tau_u}^{t\wedge\sigma_u} \mathrm{d}r \int_{\mathbb{R}_+^{\ell}} y_{\gamma_u}\mathrm{e}^{-[\boldsymbol{w},\boldsymbol{y}]}\mathbb{E}^{\dagger}_{\boldsymbol{y}\delta_{z_u(r)}}\left[1-\mathrm{e}^{-\langle \boldsymbol{f},\boldsymbol{X}_{t-r}\rangle}\right]\Pi(\gamma_u,\mathrm{d}\boldsymbol{y})\right\} \\
&\qquad = \exp\left\{-\sum_{u\in\mathcal{T}}\int_{t\wedge\tau_u}^{t\wedge\sigma_u} \mathrm{d}r \int_{\mathbb{R}_+^{\ell}} y_{\gamma_u}\mathrm{e}^{-[\boldsymbol{w},\boldsymbol{y}]}\left(1-\mathrm{e}^{-[\boldsymbol{V}^{\dagger}_{t-r}\boldsymbol{f}(z_u(r)),\boldsymbol{y}]}\right)\Pi(\gamma_u,\mathrm{d}\boldsymbol{y})\right\}.
\end{aligned}$$

Therefore, by putting the pieces together we obtain the following

$$\begin{aligned}
&\widehat{\mathbb{E}}_{(\boldsymbol{\mu},\boldsymbol{\nu})}\left[\exp\left\{-\langle \boldsymbol{f}, \boldsymbol{I}_t^{\mathbb{N}^{\dagger}}+\boldsymbol{I}_t^{\mathbb{P}^{\dagger}}\rangle\right\}\Big| (\boldsymbol{Z}_s, s\le t)\right] \\
&\qquad\qquad = \exp\left\{-\sum_{u\in\mathcal{T}}\int_{t\wedge\tau_u}^{t\wedge\sigma_u}\phi(\gamma_u,\boldsymbol{V}^{\dagger}_{t-r}\boldsymbol{f}(z_u(r)))\mathrm{d}r\right\},
\end{aligned}\tag{4.26}$$

where $\phi(i,\boldsymbol{\lambda})$ is given by formula (4.17). The previous equation is in terms of the tree $\mathcal{T}$. We want to rewrite it in terms of the multitype branching diffusion, thus

$$\begin{aligned}
\sum_{u\in\mathcal{T}}\int_{t\wedge\tau_u}^{t\wedge\sigma_u}\phi(\gamma_u,\boldsymbol{V}^{\dagger}_{t-r}\boldsymbol{f}(z_u(r)))\mathrm{d}r &= \sum_{i\in S}\sum_{u\in\mathcal{T},\gamma_u=i}\int_{t\wedge\tau_u}^{t\wedge\sigma_u}\phi(i,\boldsymbol{V}^{\dagger}_{t-r}\boldsymbol{f}(z_u(r)))\mathrm{d}r \\
&= \int_0^t\sum_{i\in S}\sum_{u\in\mathcal{T},z_u=i}\phi(i,\boldsymbol{V}^{\dagger}_{t-r}\boldsymbol{f}(z_u(r)))\mathbf{1}_{\{r\in[t\wedge\tau_u,t\wedge\sigma_u)\}}\mathrm{d}r \\
&= \int_0^t\left\langle\boldsymbol{\phi}(\boldsymbol{V}^{\dagger}_{t-r}),\boldsymbol{Z}_r\right\rangle\mathrm{d}r.
\end{aligned}$$

□

Observe that the processes $\boldsymbol{I}^{\mathbb{N}^{\dagger}}=(\boldsymbol{I}_t^{\mathbb{N}^{\dagger}},t\ge 0)$, $\boldsymbol{I}^{\mathbb{P}^{\dagger}}=(\boldsymbol{I}_t^{\mathbb{P}^{\dagger}},t\ge 0)$ and $\boldsymbol{I}^{\eta}=(\boldsymbol{I}_t^{\eta},t\ge 0)$ are initially zero-valued $\widehat{\mathbb{P}}_{(\boldsymbol{\mu},\boldsymbol{\nu})}$-a.s. In order to study the rest of the immigration along the backbone we have the following result.

Lemma 4.2.2. *Suppose that* $\boldsymbol{f},\boldsymbol{h}\in\mathcal{B}^+(\mathbb{R}^d)^{\ell}$ *and* $\boldsymbol{g}_s(x)\in\mathcal{B}^+(\mathbb{R}\times\mathbb{R}^d)^{\ell}$. *Define the vectorial function* $\mathrm{e}^{-\boldsymbol{W}_t(x)}=(\mathrm{e}^{-W_t^{(1)}(x)},\cdots,\mathrm{e}^{-W_t^{(\ell)}(x)})$ *as follows*

$$\mathrm{e}^{-W_t^{(i)}(x)}:=\widehat{\mathbb{E}}_{(\boldsymbol{\mu},\boldsymbol{e}_i\delta_x)}\left[\exp\left\{-\langle\boldsymbol{f},\boldsymbol{I}_t^{\eta}\rangle-\langle\boldsymbol{h},\boldsymbol{Z}_t\rangle-\int_0^t\langle\boldsymbol{g}_{t-s},\boldsymbol{Z}_s\rangle\mathrm{d}s\right\}\right].$$

Then, $\mathrm{e}^{-\boldsymbol{W}_t(x)}$ *is a locally bounded solution to the integral system*

$$\begin{aligned}
\mathrm{e}^{-W_t^{(i)}(x)} = \mathtt{P}_t^{(i)}\mathrm{e}^{-h_i(x)}+\frac{1}{w_i}\int_0^t\mathrm{d}s\int_{\mathbb{R}^d}&\left[H_{t-s}^{(i)}\left(y,\boldsymbol{w}\cdot\mathrm{e}^{-\boldsymbol{W}_{t-s}(y)}\right)\right. \\
&\left.-w_i g_{t-s}^i(y)\mathrm{e}^{-W_{t-s}^{(i)}(y)}\right]\mathtt{P}_s^{(i)}(x,\mathrm{d}y),
\end{aligned}\tag{4.27}$$

where

$$H_s^{(i)}(x,\boldsymbol{\theta}) = [\boldsymbol{\theta}, \boldsymbol{B}^\dagger \boldsymbol{e}_i] + \beta_i\theta_i^2 + \int_{\mathbb{R}_+^\ell} \left(\mathrm{e}^{[\boldsymbol{\theta},\boldsymbol{y}]} - 1 - \theta_i y_i\right) \mathrm{e}^{-[\boldsymbol{w}+\boldsymbol{V}_s^\dagger \boldsymbol{f}(x),\boldsymbol{y}]}\Pi(i,\mathrm{d}\boldsymbol{y}). \quad (4.28)$$

In the latter formula $\boldsymbol{B}^\dagger$ *is given by (4.24) and* $\boldsymbol{V}_t^\dagger \boldsymbol{f}$ *is the unique solution to (4.11).*

It is important to note that $\boldsymbol{W}$ depends on the functions $\boldsymbol{f}, \boldsymbol{h}$ and $\boldsymbol{g}$ but for simplicity on exposition we suppress this dependency.

Proof. Recall that $\boldsymbol{Z}$ is a multitype branching diffusion, where the motion of each particle with type $i \in S$ is given by the semigroup $\mathtt{P}^{(i)}$ and its branching generator is given by (4.13). For simplicity, we denote by $\mathtt{P}_x^{(i)}$ the law of the diffusion $\xi^{(i)}$ starting at x. By conditioning on the time of the first branching event of $\boldsymbol{Z}$ we get

$$\begin{aligned}\mathrm{e}^{-W_t^{(i)}(x)} =& \mathtt{E}_x^{(i)}\left[\mathrm{e}^{-q_i t}\mathrm{e}^{-\int_0^t g_{t-r}^i(\xi_r^{(i)})\mathrm{d}r}\mathrm{e}^{-h_i(\xi_t^{(i)})}\right] \\ &+ \mathtt{E}_x^{(i)}\left[\int_0^t q_i \mathrm{e}^{-q_i s}\mathrm{e}^{-\int_0^s g_{s-r}^i(\xi_r^{(i)})\mathrm{d}r}\right. \\ &\left.\sum_{\boldsymbol{j}\in\mathbb{N}^\ell} p_{\boldsymbol{j}}^{(i)} \mathrm{e}^{-\sum_{k\in S} j_k W_{t-s}^{(k)}(\xi_s^{(i)})} \int_{\mathbb{R}_+^\ell} \eta_{\boldsymbol{j}}^{(i)}(\mathrm{d}\boldsymbol{y})\mathrm{e}^{-[\boldsymbol{V}_{t-s}^\dagger \boldsymbol{f}(\xi_t^{(i)}),\boldsymbol{y}]}\mathrm{d}s\right],\end{aligned}$$

where $\boldsymbol{j} = (j_1, \ldots, j_\ell)$. On the other hand, by Proposition 2.9 in [20], we see that $\mathrm{e}^{-W_t^{(i)}(x)}$ also satisfies

$$\begin{aligned}\mathrm{e}^{-W_t^{(i)}(x)} =& \mathtt{E}_x^{(i)}\left[\mathrm{e}^{-h_i(\xi_t^{(i)})}\right] - \mathtt{E}_x^{(i)}\left[\int_0^t q_i \mathrm{e}^{-W_{t-s}^{(i)}(x)}\mathrm{d}s\right] - \mathtt{E}_x^{(i)}\left[\int_0^t g_{t-s}^i(\xi_s^{(i)})\mathrm{e}^{-W_{t-s}^{(i)}(x)}\mathrm{d}s\right] \\ &+ \mathtt{E}_x^{(i)}\left[\int_0^t q_i \sum_{\boldsymbol{j}\in\mathbb{N}^\ell} p_{\boldsymbol{j}}^{(i)} \mathrm{e}^{-\sum_{k\in S} j_k W_{t-s}^{(k)}(\xi_s^{(i)})} \int_{\mathbb{R}_+^\ell} \eta_{\boldsymbol{j}}^{(i)}(\mathrm{d}\boldsymbol{y})\mathrm{e}^{-[\boldsymbol{V}_{t-s}^\dagger \boldsymbol{f}(\xi_t^{(i)}),\boldsymbol{y}]}\mathrm{d}s\right].\end{aligned}$$

By substituting the definitions of $p_{\boldsymbol{j}}^{(i)}$ and $\eta_{\boldsymbol{j}}^{(i)}$ (see (4.14) and (4.18)), we get that for all $x \in \mathbb{R}^d$

$$
\begin{aligned}
R(x) :&= \sum_{\boldsymbol{j}\in\mathbb{N}^\ell} p_{\boldsymbol{j}}^{(i)} \mathrm{e}^{-[\boldsymbol{j},\boldsymbol{W}_{t-s}(x)]} \int_{\mathbb{R}_+^\ell} \eta_{\boldsymbol{j}}^i(\mathrm{d}\boldsymbol{y}) \mathrm{e}^{-[\boldsymbol{V}_{t-s}^\dagger \boldsymbol{f}(x),\boldsymbol{y}]} \\
&= \frac{1}{w_i q_i} \sum_{\boldsymbol{j}\in\mathbb{N}^\ell} \Bigg[\beta_i w_i^2 \mathrm{e}^{-[\boldsymbol{j},\boldsymbol{W}_{t-s}(x)]} \mathbf{1}_{\{\boldsymbol{j}=2\boldsymbol{e}_i\}} + \Big(B_{ki} w_k \mathrm{e}^{-[\boldsymbol{j},\boldsymbol{W}_{t-s}(x)]} \\
&+ \int_{\mathbb{R}_+^\ell} w_k y_k \mathrm{e}^{-[\boldsymbol{w},\boldsymbol{y}]} \mathrm{e}^{-[\boldsymbol{j},\boldsymbol{W}_{t-s}(x)]} \mathrm{e}^{-[\boldsymbol{V}_{t-s}^\dagger \boldsymbol{f}(x),\boldsymbol{y}]} \Pi(i,\mathrm{d}\boldsymbol{y}) \Big) \mathbf{1}_{\{\boldsymbol{j}=\boldsymbol{e}_k\}} \mathbf{1}_{\{k\neq i\}} \\
&+ \int_{\mathbb{R}_+^\ell} \frac{(w_1y_1)^{j_1}\dots(w_\ell y_\ell)^{j_\ell}}{j_1!\dots j_\ell!} \mathrm{e}^{-[\boldsymbol{w},\boldsymbol{y}]} \mathrm{e}^{-[\boldsymbol{j},\boldsymbol{W}_{t-s}(x)]} \mathrm{e}^{-[\boldsymbol{V}_{t-s}^\dagger \boldsymbol{f}(x),\boldsymbol{y}]} \Pi(i,\mathrm{d}\boldsymbol{y}) \mathbf{1}_{\{j_1+\dots+j_\ell\geq 2\}} \Bigg] \\
&= \frac{1}{w_i q_i} \Bigg[\beta_i \left(w_i \mathrm{e}^{-W_{t-s}^{(i)}(x)} \right)^2 \\
&+ \sum_{k\in S, k\neq i} \mathrm{e}^{-W_{t-s}^{(k)}(x)} \left(B_{ki} w_k + \int_{\mathbb{R}_+^\ell} w_k y_k \mathrm{e}^{-[\boldsymbol{w},\boldsymbol{y}]} \mathrm{e}^{-[\boldsymbol{V}_{t-s}^\dagger \boldsymbol{f}(x),\boldsymbol{y}]} \Pi(i,\mathrm{d}\boldsymbol{y}) \right) \\
&+ \int_{\mathbb{R}_+^\ell} \sum_{n\geq 2} \frac{[\boldsymbol{w}\cdot \mathrm{e}^{-\boldsymbol{W}_{t-s}(x)},\boldsymbol{y}]^n}{n!} \mathrm{e}^{-[\boldsymbol{w},\boldsymbol{y}]} \mathrm{e}^{-[\boldsymbol{V}_{t-s}^\dagger \boldsymbol{f}(x),\boldsymbol{y}]} \Pi(i,\mathrm{d}\boldsymbol{y}) \Bigg],
\end{aligned}
$$

where in the last row we have used (4.15). By merging the two integrals, we get

$$
\begin{aligned}
R(x) =& \frac{1}{w_i q_i} \Bigg[\beta_i \left(w_i \mathrm{e}^{-W_{t-s}^{(i)}(x)} \right)^2 + \sum_{k\in S, k\neq i} B_{ki} w_k \mathrm{e}^{-W_{t-s}^{(k)}(x)} \\
&+ \int_{\mathbb{R}_+^\ell} \left(\mathrm{e}^{[\boldsymbol{w}\cdot \mathrm{e}^{-\boldsymbol{W}_{t-s}(x)},\boldsymbol{y}]} - 1 - w_i \mathrm{e}^{-W_{t-s}^{(i)}(x)} y_i \right) \mathrm{e}^{-[\boldsymbol{w}+\boldsymbol{V}_{t-s}^\dagger \boldsymbol{f}(x),\boldsymbol{y}]} \Pi(i,\mathrm{d}\boldsymbol{y}) \Bigg].
\end{aligned}
$$

So, putting the pieces together and using the definitions of q_i, $\boldsymbol{B}^\dagger$ and $H^{(i)}$, (see identities (4.13),(4.24) and (4.28)) we deduce that

$$
\begin{aligned}
\mathrm{e}^{-W_t^{(i)}(x)} = \mathtt{E}_x^{(i)} \Bigg[& \mathrm{e}^{-h_i(\xi_t^{(i)})} - \int_0^t g_{t-s}^i(\xi_s^{(i)}) \mathrm{e}^{-W_{t-s}^{(i)}(\xi_s^{(i)})} \mathrm{d}s \\
& + \frac{1}{w_i} \int_0^t H_{t-s}^{(i)}(\xi_s^{(i)}, \boldsymbol{w}\cdot \mathrm{e}^{-\boldsymbol{W}_{t-s}(\xi_s^{(i)})}) \mathrm{d}s \Bigg],
\end{aligned}
$$

as stated. Therefore, $\mathrm{e}^{-\boldsymbol{W}_t(x)}$ satisfies (4.27). □

Proof of Theorem 4.1.5. Since $\widetilde{\boldsymbol{X}}$ is an independent copy of $\boldsymbol{X}$ under $\mathbb{P}_{\boldsymbol{\mu}}^\dagger$, it is enough to show that for $\boldsymbol{\mu} \in \mathcal{M}(\mathbb{R}^d)^\ell$, $\boldsymbol{\nu} \in \mathcal{M}_a(\mathbb{R}^d)^\ell$, $\boldsymbol{f}, \boldsymbol{h} \in \mathcal{B}^+(\mathbb{R}^d)^\ell$, the vectorial function $\mathrm{e}^{-\boldsymbol{U}_t^{(\boldsymbol{f})}\boldsymbol{h}(x)}$ defined by

$$
\mathrm{e}^{-U_t^{(\boldsymbol{f},i)}\boldsymbol{h}(x)} := \widehat{\mathbb{E}}_{\boldsymbol{\mu},\boldsymbol{e}_i\delta_x} \left[\mathrm{e}^{-\langle \boldsymbol{f}, \boldsymbol{I}_t^{\mathbb{N}^\dagger} + \boldsymbol{I}_t^{\mathbb{P}^\dagger} + \boldsymbol{I}_t^{\eta} \rangle - \langle \boldsymbol{h}, \boldsymbol{Z}_t \rangle} \right],
$$

is a solution to (4.20) and that this solution is unique. By its definition, it is clear that $\mathrm{e}^{-\boldsymbol{U}_t^{(\boldsymbol{f})}\boldsymbol{h}(x)} \in [0,1]^\ell$ for all $x \in \mathbb{R}^d$ and $t \geq 0$. On the other hand from Lemma 4.2.1, we observe that

$$\mathrm{e}^{-U_t^{(\boldsymbol{f},i)}\boldsymbol{h}(x)} = \widehat{\mathbb{E}}_{\boldsymbol{\mu},\boldsymbol{e}_i\delta_x}\left[\exp\left\{-\langle \boldsymbol{f}, \boldsymbol{I}_t^\eta\rangle - \langle \boldsymbol{h}, \boldsymbol{Z}_t\rangle - \int_0^t \langle \boldsymbol{\phi}(\boldsymbol{V}_{t-r}^\dagger \boldsymbol{f}), \boldsymbol{Z}_r\rangle \mathrm{d}r\right\}\right].$$

Therefore Lemma 4.2.2 implies that the vectorial function $\mathrm{e}^{-\boldsymbol{U}_t^{(\boldsymbol{f})}\boldsymbol{h}(x)}$ satisfies

$$\mathrm{e}^{-U_t^{(\boldsymbol{f},i)}\boldsymbol{h}(x)} = \mathtt{E}_x^{(i)}\Bigg[\mathrm{e}^{-h_i(\xi_t^{(i)})} + \frac{1}{w_i}\int_0^t \Bigg(H_{t-s}^{(i)}(\xi_s^{(i)}, \boldsymbol{w}\cdot \mathrm{e}^{-\boldsymbol{U}_{t-s}^{(\boldsymbol{f},i)}\boldsymbol{h}(\xi_s^{(i)})}) \\ - \phi(i, \boldsymbol{V}_{t-r}^\dagger \boldsymbol{f}(\xi_s^{(i)}))w_i \mathrm{e}^{-U_{t-s}^{(\boldsymbol{f},i)}\boldsymbol{h}(\xi_s^{(i)})}\Bigg)\mathrm{d}s\Bigg],$$

where $H^{(i)}$ is given as in (4.28). Using the definitions of $\boldsymbol{\psi}^\dagger$, $\boldsymbol{\phi}$ and H (see identities (4.17) (4.23) and (4.28)), we observe for all $i \in S$, $x \in \mathbb{R}^d$ and $\boldsymbol{\theta} \in \mathbb{R}_+^l$, that

$$H_t^{(i)}(x, \boldsymbol{\theta}) - \phi(i, \boldsymbol{V}_t^\dagger \boldsymbol{f}(x))\theta_i = \psi^\dagger\left(i, -\boldsymbol{\theta} + \boldsymbol{V}_t^\dagger \boldsymbol{f}(x)\right) - \psi^\dagger(i, \boldsymbol{V}_t^\dagger \boldsymbol{f}(x)).$$

Therefore, $\mathrm{e}^{-\boldsymbol{U}_t^{(\boldsymbol{f})}\boldsymbol{h}(x)}$ is a solution to (4.20).

In order to finish the proof, we show that the solution to (4.20) is unique. Our arguments use Gronwall's lemma and similar ideas to those used in the monograph of Li [20] and in Proposition 4.1.1. With this purpose in mind, we first deduce some additional inequalities. Recall that the function $\psi^\dagger(i, \boldsymbol{\theta})$ defined in (4.23) is a branching mechanism. Using similar notation as in Proposition 4.1.1, we introduce the operator

$$\Psi^\dagger(x, i, f) = \psi^\dagger(i, (f(x,1), \cdots, f(x,\ell))),$$

for $f \in \mathcal{B}(\mathbb{R}^d \times S)$, and observe that it satisfies identity (2.26) in [20]. Therefore, following line by line the arguments in the proof of Proposition 2.20 in [20], we may deduce that $\Psi^\dagger$ satisfies Condition 2.11 in [20]. In other words, for all $a \geq 0$, there exists $L_a > 0$ such that

$$\sup_{(x,i)\in\mathbb{R}^d\times S} |\Psi^\dagger(x,i,f) - \Psi^\dagger(x,i,g)| \leq L_a\|f-g\|, \qquad \text{for} \quad f, g \in \mathcal{B}_a(\mathbb{R}^d \times S), \quad (4.29)$$

where $\|f\| := \sup_{(x,i)\in\mathbb{R}^d\times S}|f(x,i)|$ and $\mathcal{B}_a(\mathbb{R}^d \times S) := \{f \in \mathcal{B}(\mathbb{R}^d \times S) : \|f\| \leq a\}$.

On the other hand by Proposition 2.21 in [20], for all $f \in \mathcal{B}(\mathbb{R}^d \times S)$, there exists $t \mapsto \mathtt{V}_t^\dagger f$ a unique locally bounded positive solution to

$$\mathtt{V}_t^\dagger f(x,i) = \mathtt{T}_t f(x,i) - \int_0^t \mathrm{d}s \int_{\mathbb{R}^d\times S} \Psi^\dagger(y, j, \mathtt{V}_{t-s}^\dagger f)\mathtt{T}_s(x, i, \mathrm{d}(y,j)),$$

where the semigroup $\mathtt{T}_t$ is given as in (4.21). Moreover, by Proposition 2.14 in [20], for all $T > 0$ there exists $C(T)$ such that

$$\sup_{0\leq s\leq T}\ \sup_{(x,i)\in\mathbb{R}^d\times S} |\mathtt{V}_s^\dagger f(x,i)| \leq C(T)\|f\|.$$

Hence using the homeomorphism between $\mathcal{B}(\mathbb{R}^d)^\ell$ and $\mathcal{B}(\mathbb{R}^d \times S)$ which was defined in the proof of Proposition 4.1.1 (i.e. for $\boldsymbol{f} \in \mathcal{B}(\mathbb{R}^d)^\ell$, we define $f(x,i) = f_i(x)$) and the previous inequality, we deduce that

$$\sup_{0\leq s\leq T} \sup_{x\in\mathbb{R}^d} \sup_{i\in S} \left| V_s^{\dagger,(i)} \boldsymbol{f}(x) \right| \leq C(T)\|\boldsymbol{f}\| \qquad \text{for} \quad \boldsymbol{f} \in \mathcal{B}^+(\mathbb{R}^d)^\ell, \tag{4.30}$$

where $\|\boldsymbol{f}\| = \sup_{x\in\mathbb{R}^d} \sup_{i\in S} |f_i(x)|$ and $\boldsymbol{V}^\dagger \boldsymbol{f}$ is given by (4.11).

Next, we take $\mathrm{e}^{-\boldsymbol{W}_t(x)}$ and $\mathrm{e}^{-\widetilde{\boldsymbol{W}}_t(x)}$, two solutions of (4.20), and observe that for all $i \in S$

$$w_i \mathrm{e}^{-W_t^{(i)}(x)} - w_i \mathrm{e}^{-\widetilde{W}_t^{(i)}(x)} = \int_0^t \mathrm{d}s \int_{\mathbb{R}^d} \Big[\psi^\dagger\left(i, -\boldsymbol{w}\cdot \mathrm{e}^{-\boldsymbol{W}_{t-s}(y)} + \boldsymbol{V}_{t-s}^\dagger \boldsymbol{f}(y)\right) \\ -\psi^\dagger\left(i, -\boldsymbol{w}\cdot \mathrm{e}^{-\widetilde{\boldsymbol{W}}_{t-s}(y)} + \boldsymbol{V}_{t-s}^\dagger \boldsymbol{f}(y)\right)\Big]\, \mathtt{P}_s^{(i)}(x, \mathrm{d}y).$$

Since $\mathrm{e}^{-\boldsymbol{W}_t(x)} \in [0,1]^\ell$ and $\boldsymbol{V}^\dagger f$ satisfies (4.30), we have, for all $s \leq T$, that

$$\left\| -\boldsymbol{w}\cdot \mathrm{e}^{-\boldsymbol{W}_s(x)} + \boldsymbol{V}_s^\dagger f(x) \right\| \leq \|\boldsymbol{w}\| + C(T)\|f\| := a(T),$$

and the same inequality holds for $\mathrm{e}^{-\widetilde{\boldsymbol{W}}_t(x)}$. Therefore, by the definition of $\Psi^\dagger$ and (4.29), there exists $L_T > 0$ such that we obtain, for all $t \leq T$, the following inequality

$$\left| w_i \mathrm{e}^{-W_t^{(i)}(x)} - w_i \mathrm{e}^{-\widetilde{W}_t^{(i)}(x)} \right| \leq \int_0^t \mathrm{d}s \int_{\mathbb{R}^d} L_T \left\| \boldsymbol{w}\cdot \mathrm{e}^{-\boldsymbol{W}_{t-s}(x)} - \boldsymbol{w}\cdot \mathrm{e}^{-\widetilde{\boldsymbol{W}}_{t-s}(x)} \right\| \mathtt{P}_s^{(i)}(x, \mathrm{d}y).$$

The latter implies the following inequality

$$\left\| \boldsymbol{w}\cdot \mathrm{e}^{-\boldsymbol{W}_t(x)} - \boldsymbol{w}\cdot \mathrm{e}^{-\widetilde{\boldsymbol{W}}_t(x)} \right\| \leq L_T \int_0^t \left\| \boldsymbol{w}\cdot \mathrm{e}^{-\boldsymbol{W}_s(x)} - \boldsymbol{w}\cdot \mathrm{e}^{-\widetilde{\boldsymbol{W}}_s(x)} \right\| \mathrm{d}s, \qquad \text{for all } t \leq T.$$

Thus by Gronwall's inequality, we deduce that

$$\boldsymbol{w}\cdot \mathrm{e}^{-\boldsymbol{W}_s(x)} = \boldsymbol{w}\cdot \mathrm{e}^{-\widetilde{\boldsymbol{W}}_s(x)} \qquad \text{for all} \quad s \leq T.$$

Finally, because $T > 0$ was arbitrary, we get the uniqueness of the solution to (4.20). □

Proof of Theorem 4.1.6. Recall that $((\boldsymbol{\Lambda}, \boldsymbol{Z}), \widehat{\mathbb{P}}_{(\boldsymbol{\mu},\boldsymbol{\nu})})$ is a Markov process and that $\widehat{\mathbb{P}}_{\boldsymbol{\mu}}$ is defined as $\widehat{\mathbb{P}}_{(\boldsymbol{\mu},\widetilde{\boldsymbol{\nu}})}$, where $\widetilde{\boldsymbol{\nu}}$ is such that $\widetilde{\nu}_i$ is a Poisson random measure with intensity $w_i\mu_i$, for all $i \in S$. Therefore, for $s, t \geq 0$, we see that

$$\widehat{\mathbb{E}}_{\boldsymbol{\mu}}\Big[f(\boldsymbol{\Lambda}_{t+s})\Big|(\boldsymbol{\Lambda}_u, u \leq s)\Big] = \widehat{\mathbb{E}}_{\boldsymbol{\mu},\widetilde{\boldsymbol{\nu}}}\Big[f(\boldsymbol{\Lambda}_{t+s})\Big|(\boldsymbol{\Lambda}_u, u \leq s)\Big] = \widehat{\mathbb{E}}_{(\boldsymbol{\Lambda}_s, \boldsymbol{Z}_s)}\Big[f(\boldsymbol{\Lambda}_t)\Big].$$

Then, in order to deduce that $(\boldsymbol{\Lambda}, \widehat{\mathbb{P}}_{\boldsymbol{\mu}})$ is Markovian, we need to show that each coordinate of $\boldsymbol{Z}_t = (Z_t^1, \ldots, Z_t^\ell)$ given $\boldsymbol{\Lambda}_t = (\Lambda_t^1, \ldots, \Lambda_t^\ell)$ is a Poisson random measure with intensity $w_i\Lambda_t^i$. From Campbell's formula for Poisson random measures (see for instance Section 3.2 of [13]), the latter is equivalent to showing that for all $\boldsymbol{h} \in \mathcal{B}^+(\mathbb{R}^d)^\ell$

$$\widehat{\mathbb{E}}_{\boldsymbol{\mu}}\left[\mathrm{e}^{-\langle \boldsymbol{h}, \boldsymbol{Z}_t\rangle}\middle| \boldsymbol{\Lambda}_t\right] = \exp\left\{-\langle \boldsymbol{w}\cdot(\boldsymbol{1}-\mathrm{e}^{\boldsymbol{h}}), \boldsymbol{\Lambda}_t\rangle\right\},$$

or equivalently, that for all $\boldsymbol{f}, \boldsymbol{h} \in \mathcal{B}^+(\mathbb{R}^d)^\ell$

$$\widehat{\mathbb{E}}_{\boldsymbol{\mu}}\left[\mathrm{e}^{-\langle \boldsymbol{f}, \boldsymbol{\Lambda}_t\rangle - \langle \boldsymbol{h}, \boldsymbol{Z}_t\rangle}\right] = \widehat{\mathbb{E}}_{\boldsymbol{\mu}}\left[\mathrm{e}^{-\langle \boldsymbol{w}\cdot(\boldsymbol{1}-\mathrm{e}^{-\boldsymbol{h}})+\boldsymbol{f}, \boldsymbol{\Lambda}_t\rangle}\right]. \tag{4.31}$$

Using (4.19) with $\widetilde{\boldsymbol{\nu}}$, we find

$$\widehat{\mathbb{E}}_{\boldsymbol{\mu}}\left[\mathrm{e}^{-\langle \boldsymbol{f}, \boldsymbol{\Lambda}_t\rangle - \langle \boldsymbol{h}, \boldsymbol{Z}_t\rangle}\right] = \exp\left\{-\langle \boldsymbol{V}_t^\dagger \boldsymbol{f} + \boldsymbol{w}\cdot(\boldsymbol{1}-\mathrm{e}^{-\boldsymbol{U}_t^{(\boldsymbol{f})}\boldsymbol{h}}), \boldsymbol{\mu}\rangle\right\}.$$

Similarly, considering (4.19) again with $\widetilde{\boldsymbol{\nu}}$, $\widetilde{\boldsymbol{f}} = \boldsymbol{w}\cdot(\boldsymbol{1}-\mathrm{e}^{-\boldsymbol{h}}) + \boldsymbol{f}$ and $\widetilde{\boldsymbol{h}} = \boldsymbol{0}$, we get that

$$\begin{aligned}&\widehat{\mathbb{E}}_{\boldsymbol{\mu}}\left[\mathrm{e}^{-\langle \boldsymbol{w}\cdot(\boldsymbol{1}-\mathrm{e}^{-\boldsymbol{h}})+\boldsymbol{f}, \boldsymbol{\Lambda}_t\rangle}\right] \\ &\qquad = \exp\left\{-\left\langle \boldsymbol{V}_t^\dagger(\boldsymbol{w}\cdot(\boldsymbol{1}-\mathrm{e}^{-\boldsymbol{h}})+\boldsymbol{f}) + \boldsymbol{w}\cdot(\boldsymbol{1}-\mathrm{e}^{-\boldsymbol{U}_t^{(\boldsymbol{w}\cdot(\boldsymbol{1}-\mathrm{e}^{-\boldsymbol{h}})+\boldsymbol{f})}\boldsymbol{0}}), \boldsymbol{\mu}\right\rangle\right\}.\end{aligned}$$

Hence, if we prove that for any $\boldsymbol{f}, \boldsymbol{h} \in \mathcal{B}^+(\mathbb{R}^d)^\ell$, $x \in \mathbb{R}^d$, and $i \in S$, the following identity holds

$$\begin{aligned}V_t^{\dagger(i)}\boldsymbol{f}(x) + w_i(1-\mathrm{e}^{-U_t^{(\boldsymbol{f},i)}\boldsymbol{h}(x)}) &= V_t^{\dagger(i)}(\boldsymbol{w}\cdot(\boldsymbol{1}-\mathrm{e}^{-\boldsymbol{h}})+\boldsymbol{f})(x) \\ &\quad + w_i\left(1-\mathrm{e}^{-U_t^{((\boldsymbol{w}\cdot(\boldsymbol{1}-\mathrm{e}^{-\boldsymbol{h}})+\boldsymbol{f}),i)}\boldsymbol{0}(x)}\right),\end{aligned} \tag{4.32}$$

we can deduce (4.31).

In order to obtain (4.32), we first observe that identities (4.11) and (4.20) together with the definition of $\boldsymbol{\psi}^\dagger$ allow us to see that both left and right hand sides of (4.32) solve (4.3) with initial condition $\boldsymbol{f} + \boldsymbol{w}\cdot(\boldsymbol{1}-\mathrm{e}^{-\boldsymbol{h}})$. Since (4.3) has a unique solution, namely $\boldsymbol{V}_t(\boldsymbol{f} + \boldsymbol{w}\cdot(\boldsymbol{1}-\mathrm{e}^{-\boldsymbol{h}}))$, we conclude that (4.32) holds and it is equal to $V_t^{(i)}(\boldsymbol{f} + \boldsymbol{w}\cdot(\boldsymbol{1}-\mathrm{e}^{-\boldsymbol{h}}))(x)$. Hence, we can finally deduce that $(\boldsymbol{\Lambda}, \widehat{\mathbb{P}}_{\boldsymbol{\mu}})$ is a Markov process. Moreover, we have

$$\widehat{\mathbb{E}}_{\boldsymbol{\mu}}\left[\mathrm{e}^{-\langle \boldsymbol{f}, \boldsymbol{\Lambda}_t\rangle - \langle \boldsymbol{h}, \boldsymbol{Z}_t\rangle}\right] = \mathrm{e}^{-\left\langle \boldsymbol{V}_t(\boldsymbol{f}+\boldsymbol{w}\cdot(\boldsymbol{1}-\mathrm{e}^{-\boldsymbol{h}})), \boldsymbol{\mu}\right\rangle} = \mathbb{E}_{\boldsymbol{\mu}}\left[\mathrm{e}^{-\langle \boldsymbol{f}+\boldsymbol{w}\cdot(\boldsymbol{1}-\mathrm{e}^{-\boldsymbol{h}}), \boldsymbol{X}_t\rangle}\right],$$

and if, in particular, we take $\boldsymbol{h} = \boldsymbol{0}$ the above identity is reduced to

$$\widehat{\mathbb{E}}_{\boldsymbol{\mu}}\left[\mathrm{e}^{-\langle \boldsymbol{f}, \boldsymbol{\Lambda}_t\rangle}\right] = \mathbb{E}_{\boldsymbol{\mu}}\left[\mathrm{e}^{-\langle \boldsymbol{f}, \boldsymbol{X}_t\rangle}\right].$$

This completes the proof. □

Bibliography

[1] Barczy, M., Li, Z., and Pap, G. (2015). Stochastic differential equation with jumps for multi-type continuous state and continuous time branching processes with immigration. *ALEA Lat. Am. J. Probab. Math. Stat.* **12**(1), 129 – 169.

[2] Barczy, M. and Pap, G. (2016). Asymptotic behavior of critical, irreducible multi-type continuous state and continuous time branching processes with immigration. *Stochastics and Dynamics.* **16**(4).

[3] Berestycki, J., Kyprianou, A.E., Murillo-Salas, A. (2011). The prolific backbone for supercritical superdiffusions. *Stoch. Proc. Appl.* **121**, 1315 – 1331.

[4] Chen, Z-Q, Ren, Y-X, . and Yang T. Skeleton decomposition and law of large numbers for supercritical superprocesses. Preprint 2017.

[5] Duquesne, T., Wang, M. Decomposition of Lévy trees along their diameter. *Ann. Inst. H. Poincar Probab. Statist.* **53** (2017), no. 2, 539–593.

[6] Duquesne, T., Winkel, M. (2007).Growth of Lévy trees. *Probab. Theory Related Fields.* **139**, 313 – 371.

[7] Dynkin, E.B. (2002). Diffusions, Superdiffusion and Partial Differential Equations. *American Mathematical Society, Colloquium Publications.* **50**.

[8] Eckhoff, M., Kyprianou, A.E., Winkel, M. (2015). Spines, skeletons and the strong law of large numbers for superdiffusions. *Ann. Probab.* **43**, no. 5, 2545–2610.

[9] Engländer, J., Pinsky, R.G. (1999). On the construction and support properties of measure-valued diffusions on $D \subset \mathbb{R}^d$ with spatially dependent branching. *Ann. Probab.*, **27**, 684-730.

[10] Etheridge, A., Williams, D.R.E. (2003). A decomposition of the $(1+\beta)$-superprocess conditioned on survival. *Proc. Royal. Soc. Edin.* **133A**, 829–847.

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

[12] Gorostiza, L., López-Mimbela, J.A. (1990) The multitype measure branching process. *Adv. in Appl. Probab.* **22**(1), 49 – 67.

[13] Kingman, J. (1993). *Poisson Processes.* Oxford University Press.

[14] Kyprianou, A.E., Palau, S. (2017). Extinction properties of multi-type continuous-state branching processes. To appear in *Stochastic Process. Appl.*

[15] Kyprianou, A.E., Palau, S., Ren, Y. (2018).Almost sure growth of supercritical multi-type continuous-state branching process. To appear in *ALEA Lat. Am. J. Probab. Math. Stat.*

[16] Kyprianou, A.E., Pérez, J.L., S., Ren, Y. (2015). The backbone decomposition for spatially dependent supercritical superprocesses. *Séminaire de Probabilités XLVI,* 33–60.

[17] Kyprianou, A.E., Ren, Y-X. (2012) Backbone decomposition for continuous-state branching processes with immigration. *Statist. Probab. Lett.* **82**, 139–144.

[18] Li, Z. (1992). A note on the multitype measure branching process. *Adv. in Appl. Probab.* **24**(2), 496 – 498

[19] Murillo-Salas, A., Pérez, J.L. (2015). The backbone decomposition for superprocesses with non-local branching. *XI Symposium on Probability and Stochastic Processes, Progr. Probab.*, **69**, 199–216, Birkhuser/Springer, Cham.

[20] Li, Z. (2011). *Measure-Valued Branching Markov Processes*, Springer-Verlag Berlin Heidelberg.

[21] Ren, Y-X, Song R. and Yang T. Spine decomposition and L log L criterion for superprocesses with non-local branching mechanisms. Preprint 2016.

[22] Salisbury, T., Verzani, J. (1999). On the conditioned exit measures of super Brownian motion. *Probab. Theory Relat. Fields* **115**, 237–285.

Concluding remarks

The main aim of this chapter was to extend the prolific backbone decomposition of high-density populations to the multitype case. Unlike in Chapter 2 and 3, where the main tool to study skeletal decompositions was stochastic analysis, here we have used the more classic semigroup approach to prove our results.

We have considered multitype superprocesses, but since the branching was governed by a spatially independent branching mechanism, we also get the prolific backbone decomposition of multitype CSBPs by turning the movement off.