Weierstraß-Institut
für Angewandte Analysis und Stochastik

im Forschungsverbund Berlin e.V.

Preprint ISSN 0946 – 8633

Trimmed trees and embedded particle systems

Klaus Fleischmann[1], Jan Swart[2]

submitted: 05 December 2002

1 Weierstrass-Institute
for Applied Analysis and Stochastics
Mohrenstr. 39
D-10117 Berlin,
Germany
E-Mail: fleischmann@wias-berlin.de

2 University Erlangen-Nuremberg
Bismarckstr. $1\frac{1}{2}$
D-91054 Erlangen
Germany
E-Mail: swart@mi.uni-erlangen.de

No. 793
Berlin 2002

1991 *Mathematics Subject Classification.* 60J80 60G57 60J60 60K35.

Key words and phrases. (Historical) superprocess, binary branching, Poissonization, embedded particle system, trimmed tree, compensated h-transform, finite ancestry property.

Work sponsored by the German Science Foundation; in part by its graduate program "Probabilistic Analysis and Stochastic Processes".

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Mohrenstraße 39
D — 10117 Berlin
Germany

Fax: + 49 30 2044975
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

Abstract

In a supercritical branching particle system, the trimmed tree consists of those particles which have descendants at all times. We develop this concept in the superprocess setting. For a class of continuous superprocesses, we identify the trimmed tree, which turns out to be a binary splitting particle system with a new underlying motion that is a compensated h-transform of the old one. We show how trimmed trees may be estimated from above by embedded binary branching particle systems.

MSC 2000. 60J80, 60G57, 60J60, 60K35.
Keywords. (Historical) superprocess, binary branching, Poissonization, embedded particle system, trimmed tree, compensated h-transform, finite ancestry property.
Acknowledgments. Work sponsored by the German Science Foundation; in part by its graduate program "Probabilistic Analysis and Stochastic Processes".

Contents

1 Introduction and main results

1.1 Introduction

It frequently happens that a superprocess $\mathcal{X} = (\mathcal{X}_t)_{t\geq 0}$, taking values in the space $\mathcal{M}(E)$ of finite measures on some space E, and a branching particle process $X = (X_t)_{t\geq 0}$ are related by the formula

$$P^{\mathrm{Pois}(\mu)}[X_t \in \cdot\,] = P^{\mu}[\mathrm{Pois}(\mathcal{X}_t) \in \cdot\,] \qquad (t \geq 0,\ \mu \in \mathcal{M}(E)). \tag{1.1}$$

Here $\mathrm{Pois}(\mathcal{X}_t)$ denotes a Poisson point measure with random intensity $\mathcal{X}_t$ and $P^{\mathrm{Pois}(\mu)}$ denotes the law of the process X, started with initial law $\mathcal{L}(X_0) = \mathcal{L}(\mathrm{Pois}(\mu))$. For example, (1.1) holds when $\mathcal{X}$ is the standard, critical super-Brownian motion in $\mathbb{R}^d$, corresponding to the evolution equation $\frac{\partial}{\partial t}u_t = \frac{1}{2}\Delta u_t - u_t^2$, and X is a system of binary branching Brownian motions with branching rate one and death rate one. Loosely speaking, X can be obtained from $\mathcal{X}$ by *Poissonization*. Poissonization relations of the form (1.1) have been exploited by various authors, for example [GRW90, Formula (8)], [Kle98, Formula (4.19)], and [Win02, Formula (1.23)].

In the present paper, we investigate Poissonization relations for a class of continuous superprocesses with Feller underlying motion. We give conditions implying that a superprocess $\mathcal{X}$ and a branching particle system X may be coupled as processes, such that

$$P[X_t \in \cdot\,|(\mathcal{X}_s)_{0\leq s\leq t}] = P[\mathrm{Pois}(h\mathcal{X}_t) \in \cdot\,|\mathcal{X}_t] \quad \text{a.s.} \quad \forall t \geq 0, \tag{1.2}$$

where h is a sufficiently smooth density (Theorem 4). Formula (1.2) says that the conditional law of X_t, given $(\mathcal{X}_s)_{0\leq s\leq t}$, is the law of a Poisson point measure with intensity $h\mathcal{X}_t$.

The *weighted superprocess* $(h\mathcal{X}_t)_{t\geq 0}$ occuring in (1.2) is a superprocess itself, which compared to $\mathcal{X}$ has a new branching mechanism and a new underlying motion, the latter being a 'compensated' h-transform of the old one. For the case that $\mathcal{X}$ is a superdiffusion, this fact was proved and exploited by Engländer and Pinsky in [EP99].

Let $\mathcal{X}$ and X be related by (1.2), let $\mathcal{A} := \{\exists\tau < \infty \text{ such that } \mathcal{X}_t = 0\ \forall t \geq \tau\}$ denote the event that $\mathcal{X}$ becomes extinct after some random time τ, and let $A := \{\exists\tau < \infty \text{ such that } X_t = 0\ \forall t \geq \tau\}$. Since $P[X_t = 0\,|\,\mathcal{X}_t = 0] = 1$ $(t \geq 0)$, one clearly has $\mathcal{A} \subset A$ a.s. We investigate when the converse inclusion holds, i.e., when the extinction of X implies the extinction of $\mathcal{X}$. In particular, for a supercritical superprocess $\mathcal{X}$, we construct a binary splitting particle system X that, heuristically, corresponds to those infinitesimal bits of mass of $\mathcal{X}$ which have descendants at all times. More precisely, when $\hat{\mathcal{X}}$ and $\hat{X}$ are the historical processes associated with $\mathcal{X}$ and X, respectively, we give conditions on $\mathcal{X}$ and X such that

$$\forall t \geq 0\ \exists\tau < \infty \ \text{ s.t. } \ \forall r \geq \tau \quad \mathrm{supp}(\hat{X}_t) = \mathrm{supp}(\hat{\mathcal{X}}_r \circ \pi_{[0,t]}^{-1}) \qquad \text{a.s.} \tag{1.3}$$

(Theorem 7). Here $\pi_{[0,t]}$ denotes projection on the space $\mathcal{D}_E[0,t]$ of cadlag paths from $[0,t]$ into E. Informally, $\hat{\mathcal{X}}_t$ is a random measure on paths of length t, measuring how much each *line of descent* contributes to the population at time t; likewise, $\hat{X}_t$ counts how often each line of descent contributes to X_t. Thus, (1.3) says that eventually, all mass of the superprocess $\mathcal{X}$ descends from finitely many lines of descent, which are given by $\mathrm{supp}(\hat{X}_t)$.

We call X the *trimmed tree* of $\mathcal{X}$. The *reduced tree* of a branching process describes the family relations between all particles alive at a fixed time, and of their ancestors (neglecting those lines of descent that died earlier). Thus, our trimmed tree can be viewed as the limit of reduced trees as time tends to infinity. Reduced trees have been studied intensively in the branching literature. For a historical background, see, e.g., the last paragraph in Section 12.1 from [Daw93, p. 201].

When X is the trimmed tree of $\mathcal{X}$, then $\mathcal{X}$ and X are related by a Poissonization formula of the form (1.2) where $h = p$, the *infinitesimal survival probability* of $\mathcal{X}$, given by

$$p(x) = \frac{\partial}{\partial \varepsilon} P^{\varepsilon\delta_x}[\mathcal{X}_t > 0 \quad \forall t \geq 0]\Big|_{\varepsilon=0} \qquad (x \in E). \tag{1.4}$$

It is worth to mention that the weighted superprocess $(p\mathcal{X}_t)_{t\geq 0}$ with p as in (1.4) plays an important role in the work of Engländer and Pinsky [EP99], who investigate support properties (such as recurrence) of superdiffusions by analytic tools. Weighted superprocesses and embedded particle systems also played an important role in [FS02], which motivated our present article.

The paper is organised as follows. In Sections 1.2–1.4, we introduce our objects of interest together with some of their elementary properties in more detail. Sections 1.5 and 1.6 contain our main results, while Section 1.7 is devoted to discussion. Proofs are deferred to Section 2.

1.2 Poissonization of superprocesses

Let E be a compact metrizable space and let $B(E)$, $\mathcal{C}(E)$ denote the spaces of bounded measurable real functions and continuous real functions on E, respectively. We set $B_+(E) := \{f \in B(E) : f \geq 0\}$, $B_{[0,1]}(E) := \{f \in B(E) : 0 \leq f \leq 1\}$, and define $\mathcal{C}_+(E), \mathcal{C}_{[0,1]}(E)$ similarly. $\mathcal{M}(E)$ denotes the space of finite measures on E, equipped with the topology of weak convergence. If $\mu \in \mathcal{M}(E)$ and $f \in B(E)$ then $\langle \mu, f\rangle := \int_E f \, \mathrm{d}\mu$ denotes the integral of f with respect to μ. By $\mathcal{N}(E) \subset \mathcal{M}(E)$ we denote the space of finite point measures, i.e., measures ν of the form $\sum_{i=1}^n \delta_{x_i}$ with $x_i \in E$. We interpret such a point measure as a collection of n particles, situated at positions $x_1, \ldots, x_n$. For $f \in B_{[0,1]}(E)$ and $\nu = \sum_{i=1}^n \delta_{x_i} \in \mathcal{N}(E)$ we use the notation $f^\nu := \prod_{i=1}^n f(x_i)$ (where $f^0 := 1$). If μ is a random variable taking values in $\mathcal{M}(E)$, then $\mathrm{Pois}(\mu)$ denotes an $\mathcal{N}(E)$-valued random variable such that conditioned on μ, $\mathrm{Pois}(\mu)$ is a Poisson point measure with intensity μ. If ν is a random variable taking values in $\mathcal{N}(E)$ and $f \in B_{[0,1]}(E)$, then $\mathrm{Thin}_f(\nu)$ denotes a point measure obtained from ν by independently thinning the particles in ν, where a particle at x is kept with probability $f(x)$. Note that

$$\begin{array}{lll} \text{(i)} & P[\mathrm{Pois}(f\mu) = 0 \,|\, \mu] = e^{-\langle \mu, f\rangle} & (f \in B_+(E)), \\ \text{(ii)} & P[\mathrm{Thin}_f(\nu) = 0 \,|\, \nu] = (1-f)^\nu & (f \in B_{[0,1]}(E)). \end{array} \tag{1.5}$$

It is well-known that

$$\mathrm{Thin}_f(\mathrm{Thin}_g(\nu)) = \mathrm{Thin}_{fg}(\nu) \quad \text{and} \quad \mathrm{Thin}_f(\mathrm{Pois}(\mu)) = \mathrm{Pois}(f\mu). \tag{1.6}$$

Let G be the generator of a Feller process $\xi = (\xi_t)_{t\geq 0}$ on E and let $\alpha \in \mathcal{C}_+(E)$, $\beta \in \mathcal{C}(E)$. Then, for each $f \in B_+(E)$, an appropriate integrated version (see formula (2.8) below) of the semilinear Cauchy problem

$$\begin{cases} \frac{\partial}{\partial t} u_t = G u_t + \beta u_t - \alpha u_t^2 & (t \geq 0), \\ u_0 = f, \end{cases} \tag{1.7}$$

has a unique solution $u_t =: \mathcal{U}_t f$ $(t \geq 0)$ in $B_+(E)$. Moreover, there exists a unique (in law) Markov process $\mathcal{X}$ with continuous sample paths in $\mathcal{M}(E)$, defined by its Laplace functionals

$$E^\mu\big[e^{-\langle \mathcal{X}_t, f\rangle}\big] = e^{-\langle \mu, \mathcal{U}_t f\rangle} \qquad (t \geq 0, \ \mu \in \mathcal{M}(E), \ f \in B_+(E)). \tag{1.8}$$

$\mathcal{X}$ is called the *superprocess* in E with *underlying motion generator* G, *activity* α and *growth parameter* β (the last two terms are our terminology), or shortly the (G,α,β)*-superprocess.* $(\mathcal{U}_t)_{t\geq 0} = \mathcal{U} = \mathcal{U}(G,\alpha,\beta)$ is called the *log-Laplace semigroup* of $\mathcal{X}$. In fact, $\mathcal{U}_t f$ can be defined unambiguously for any measurable $f : E \to [0,\infty]$ such that (1.8) holds (where $e^{-\infty} := 0$). $\mathcal{X}$ can be constructed in several ways and is nowadays standard; see, e.g., [Fit88, Fit91, Fit92]. We can think of $\mathcal{X}$ as describing a population where mass flows with generator G, and during a time interval $\mathrm{d}t$ a bit of mass $\mathrm{d}m$ at position x produces offspring with mean $(1+\beta(x)\mathrm{d}t)\mathrm{d}m$ and finite variance $2\alpha(x)\mathrm{d}t\,\mathrm{d}m$. For basic facts on superprocesses we refer to [Daw93, Eth00].

Similarly, when G is (again) the generator of a Feller process in a compact metrizable space E and $b, d \in \mathcal{C}_+(E)$, then, for any $f \in B_{[0,1]}(E)$, an integrated version of the semilinear Cauchy problem

$$\begin{cases} \frac{\partial}{\partial t}u_t = Gu_t + bu_t(1-u_t) - du_t & (t \geq 0), \\ u_0 = f, \end{cases} \tag{1.9}$$

has a unique solution $u_t =: U_t f$ $(t \geq 0)$ in $B_{[0,1]}(E)$. Moreover, there exists a unique Markov process X with cadlag sample paths in $\mathcal{N}(E)$, defined by its generating functionals

$$E^{\nu}\big[(1-f)^{X_t}\big] = (1-U_t f)^{\nu} \qquad (t \geq 0,\ \nu \in \mathcal{N}(E),\ f \in B_{[0,1]}(E)). \tag{1.10}$$

We call X the *binary branching particle system* in E with underlying motion generator G, branching rate b and death rate d, or shortly the (G,b,d)*-bin-bra-process.* $(U_t)_{t\geq 0} = U = U(G,b,d)$ is called the *generating semigroup* of X. The particles in X perform independent motions with generator G and additionally, a particle branches with local rate b into two new particles, created at the position of the old one, and particles die with local rate d. If the death rate is zero, we also speak about *binary splitting* instead of binary branching.

Because of (1.5), formulas (1.8) and (1.10) can be rewritten as

$$\begin{array}{lll} \text{(i)} & P^{\mu}[\mathrm{Pois}(f\mathcal{X}_t) = 0] = P[\mathrm{Pois}((\mathcal{U}_t f)\mu) = 0] & (t \geq 0,\ \mu \in \mathcal{M}(E),\ f \in B_+(E)), \\ \text{(ii)} & P^{\nu}[\mathrm{Thin}_f(X_t) = 0] = P[\mathrm{Thin}_{\mathcal{U}_t f}(\nu) = 0] & (t \geq 0,\ \nu \in \mathcal{N}(E),\ f \in B_{[0,1]}(E)). \end{array} \tag{1.11}$$

The following lemma is now an easy observation.

Lemma 1 (Poissonization of superprocesses) *Let $\mathcal{X}$ be the (G,α,β)-superprocess, assume that $\alpha \geq \beta$ and let X be the $(G,\alpha,\alpha-\beta)$-bin-bra-process. Then*

$$P^{\mathrm{Pois}(\mu)}[X_t \in \cdot\,] = P^{\mu}[\mathrm{Pois}(\mathcal{X}_t) \in \cdot\,] \qquad (t \geq 0,\ \mu \in \mathcal{M}(E)). \tag{1.12}$$

Proof Let $\mathcal{U} = \mathcal{U}(G,\alpha,\beta)$ and $U = U(G,\alpha,\alpha-\beta)$ denote the log-Laplace semigroup of $\mathcal{X}$ and the moment generating semigroup of X, respectively. Comparing the Cauchy problems (1.7) and (1.9) we see that $\mathcal{U}_t f = U_t f$ for all $f \in B_{[0,1]}(E)$ and $t \geq 0$. It follows that for any $f \in B_{[0,1]}(E)$, $\mu \in \mathcal{M}(E)$, and $t \geq 0$,

$$\begin{aligned} &P^{\mathrm{Pois}(\mu)}[\mathrm{Thin}_f(X_t) = 0] = P[\mathrm{Thin}_{U_t f}(\mathrm{Pois}(\mu)) = 0] = P[\mathrm{Pois}((U_t f)\mu) = 0] \\ &\quad = P^{\mu}[\mathrm{Pois}(f\mathcal{X}_t) = 0] = P^{\mu}[\mathrm{Thin}_f(\mathrm{Pois}(\mathcal{X}_t)) = 0]. \end{aligned} \tag{1.13}$$

Since this holds for arbitrary $f \in B_{[0,1]}(E)$, the law of X_t, when X is started with initial law $\mathcal{L}(X_0) = \mathcal{L}(\mathrm{Pois}(\mu))$, coincides with the law of $\mathrm{Pois}(\mathcal{X}_t)$, when $\mathcal{X}$ is started in $\mathcal{X}_0 = \mu$. ∎

Remark (Locally compact spaces) Let E be a locally compact but not compact, separable, metrizable space, G the generator of a Feller process $\xi = (\xi_t)_{t\geq 0}$ on E, whose semigroup maps the space $\mathcal{C}_0(E)$ of continuous real functions vanishing at infinity into itself, and let α, β be bounded continuous functions on E, $\alpha \geq 0$. Then the (G, α, β)-superprocess may be defined as follows. First, E may be embedded in a compact metrizable space $\overline{E}$ such that E is an open dense subset of $\overline{E}$ and such that the functions α, β can be extended to continuous functions $\overline{\alpha}, \overline{\beta}$ on $\overline{E}$. Second, ξ may be extended to a Feller process in $\overline{E}$ (with generator denoted by $\overline{G}$) by putting $P^x[\xi_t = x \ \forall t \geq 0] := 1$ for $x \in \overline{E} \backslash E$. Identifying $\mathcal{M}(E)$ with the space $\{\mu \in \mathcal{M}(\overline{E}) : \mu(\overline{E} \backslash E) = 0\}$, the $(\overline{G}, \overline{\alpha}, \overline{\beta})$-superprocess $\overline{\mathcal{X}}$ satisfies $P^\mu[\overline{\mathcal{X}}_t \in \mathcal{M}(E) \ \forall t \geq 0] = 1$ for all $\mu \in \mathcal{M}(E)$. The (G, α, β)-superprocess may then be defined as the restriction of $\overline{\mathcal{X}}$ to $\mathcal{M}(E)$. In this way, the results in this paper can be applied, for example, to the usual super-Brownian motion (with finite initial mass). To keep notation simple, we formulate our results for superprocesses in a compact space E in the rest of this paper.

1.3 Historical superprocesses and branching particle systems

Let E be a compact metrizable space as before and let $\mathcal{D}_E[0, \infty), \mathcal{D}_E[0, t]$ denote the spaces of cadlag paths $w : [0, \infty) \to E$ and $w : [0, t] \to E$, respectively, equipped with the Skorohod topology. Let ξ be a Feller process in E. Then the *path process* $\hat{\xi}$ associated with ξ is a time-inhomogeneous Markov process with time-dependent state space $\mathcal{D}_E[0, t]$, defined as follows. If ξ^x is the process ξ started in $\xi_0^x = x \in E$, then $\hat{\xi}_t^{s,w}$, the path process $\hat{\xi}$ started at time $s \geq 0$ in $w \in \mathcal{D}_E[0, s]$ and evaluated at time $t \geq s$, is defined as

$$\hat{\xi}_t^{s,w}(r) := \begin{cases} w(r) & \text{if} \quad 0 \leq r \leq s \\ \xi_{r-s}^{w(s)} & \text{if} \quad s \leq r \leq t. \end{cases} \tag{1.14}$$

For $t \geq 0$, we identify the space $\mathcal{D}_E[0, t]$ with the space $\{w \in \mathcal{D}_E[0, \infty) : w(u) = w(t) \ \forall u \geq t\}$ of paths stopped at time t. With this identification, $\hat{\xi}^{s,w} : [s, \infty) \to \mathcal{D}_E[0, \infty)$ has cadlag sample paths. Note that $\hat{\xi}_t^{0,x}$, the path process started at time zero in $x \in \mathcal{D}_E\{0\} \cong E$ and evaluated at time $t \geq 0$, records the path followed by ξ^x up to time t.

If $\mathcal{X}$ is a (G, α, β)-superprocess in E as defined in the last section, then by definition the *historical superprocess* $\hat{\mathcal{X}}$ associated with $\mathcal{X}$ is the time-inhomogeneous superprocess with time-dependent state space $\mathcal{M}(\mathcal{D}_E[0, t])$, with underlying motion $\hat{\xi}$, time-dependent activity $\hat{\alpha}_t(w) := \alpha(w(t))$, and time-dependent growth parameter $\hat{\beta}_t(w) := \beta(w(t))$. We call $\hat{\mathcal{X}}$ the *historical* (G, α, β)*-superprocess*. We identify as before $\mathcal{D}_E[0, t]$ with the subspace of $\mathcal{D}_E[0, \infty)$ consisting of paths stopped at time t, and in this identification $\mathcal{X} : [0, \infty) \to \mathcal{M}(\mathcal{D}_E[0, \infty))$ has continuous sample paths. For the technical complications arising from the fact that the underlying motion is time-inhomogeneous and the space $\mathcal{D}_E[0, \infty)$ is not locally compact, we refer to to [DP91]; see also Section 2.2 below for more details. If $\hat{\mathcal{X}}$ is started at time zero in $\hat{\mathcal{X}}_0 = \mu \in \mathcal{M}(\mathcal{D}_E\{0\}) \cong \mathcal{M}(E)$ and $\pi_t(w) := w(t)$ denotes projection on the endpoint of a path $w \in \mathcal{D}_E[0, t]$, then the projection

$$\mathcal{X}_t := \hat{\mathcal{X}}_t \circ \pi_t^{-1} \qquad (t \geq 0) \tag{1.15}$$

gives back the original (G, α, β)-superprocess $\mathcal{X}$ started in $\mathcal{X}_0 = \mu$.

Likewise, if X is a (G, b, d)-bin-bra-process in E as defined in the last section, then the *historical binary branching particle system* $\hat{X}$ associated with X is defined as the time-inhomogeneous binary branching particle system with time-dependent state space $\mathcal{N}(\mathcal{D}_E[0, t])$, with underlying motion $\hat{\xi}$, time-dependent branching rate $\hat{b}(t, w) := b(w(t))$, and time-dependent death rate $\hat{d}(t, w) := d(w(t))$. We call $\hat{X}$ the *historical* (G, b, d)*-bin-bra-process*. Viewed as a process in $\mathcal{N}(\mathcal{D}_E[0,\infty))$, $\hat{X}$ has cadlag sample paths. If $\hat{X}$ is started at time zero in $\hat{X}_0 = \nu \in \mathcal{N}(\mathcal{D}_E\{0\}) \cong \mathcal{N}(E)$ then the analogue of (1.15) gives back the (non-historical) (G, b, d)-bin-bra-process X started in $X_0 = \nu$.

1.4 Weighted superprocesses and compensated h-transforms

Let ξ be a Feller process in E. Let G be the generator of ξ, i.e., $Gf := \lim_{t\to 0} t^{-1}(P_t f - f)$ where $P_t f(x) := E^x[f(\xi_t)]$ is the semigroup associated with ξ and the domain $\mathcal{D}(G)$ of G consists of all functions $f \in \mathcal{C}(E)$ for which the limit exists in the supremum norm. The following lemma, the proof of which can be found in Section 2.3.3 below, introduces *compensated h-transforms* of Feller processes.

Lemma 2 (Compensated h-transform of a Feller process) *Let G be the generator of a Feller process ξ in a compact metrizable space E and assume that $h \in \mathcal{D}(G)$ satisfies $h > 0$. Then the operator*

$$G^h f := \frac{1}{h}\Big(G(hf) - (Gh)f\Big), \tag{1.16}$$

with domain $\mathcal{D}(G^h) := \{f \in \mathcal{C}(E) : hf \in \mathcal{D}(G)\}$ is the generator of a Feller process ξ^h on E. The laws of ξ^h and ξ are related by

$$P^x\big[(\xi^h_s)_{s\in[0,t]} \in \mathrm{d}w\big] = \tfrac{h(w_t)}{h(x)} e^{-\int_0^t \frac{Gh}{h}(w_s)\mathrm{d}s}\, P^x\big[(\xi_s)_{s\in[0,t]} \in \mathrm{d}w\big] \qquad (t > 0,\ x \in E). \tag{1.17}$$

Remark (h-transforms) Doob's h-transform of a Feller process is the process with generator $\tilde{G}^h f := \frac{1}{h} G(hf)$. (See, for example, [Doo84, Section 2.VI.13] or [Sha88, Formula (62.23)].) Here h is superharmonic, i.e., $Gh \le 0$, and the h-transformed process has an additional local killing rate Gh/h. In our set-up, it is natural to compensate this killing by adding the term $-Gh/h$ in the definition of G^h. In this case we can allow h to be any positive function in the domain of G.

Let E, G, α and β be as in Section 1.2 and let $\mathcal{X}$ be the (G, α, β)-superprocess. The following lemma, which was proved in a non-historical setting for superdiffusions in [EP99], describes the relation between weighted superprocesses and compensated h-transforms.

Lemma 3 (Weighted superprocess) *Let $\hat{\mathcal{X}}$ be the historical (G, α, β)-superprocess and assume that $h \in \mathcal{D}(G)$, $h > 0$. Then the weighted process $\hat{\mathcal{X}}^h$ defined by*

$$\hat{\mathcal{X}}^h_t(\mathrm{d}w) := h(w_t)\hat{\mathcal{X}}_t(\mathrm{d}w) \qquad (t \ge 0) \tag{1.18}$$

is the historical $(G^h, h\alpha, \beta + \frac{Gh}{h})$-superprocess.

In particular, by formula (1.15), if $\mathcal{X}$ is the (G, α, β)-superprocess, then $\mathcal{X}^h_t(\mathrm{d}x) := h(x)\mathcal{X}_t(\mathrm{d}x)$ $(t \ge 0)$ is the $(G^h, h\alpha, \beta + \frac{Gh}{h})$-superprocess. The proof of Lemma 3 can be found in Section 2.3.4 below. The proof of Lemma 3 is deferred to Section 2.3.4.

1.5 Main results

We are ready to state our first result.

Theorem 4 (Embedded particle system) *Let E be a compact metrizable space, G the generator of a Feller process in E and $\alpha \in \mathcal{C}_+(E)$, $\beta \in \mathcal{C}(E)$. Assume that $h \in \mathcal{D}(G)$ satisfies $h > 0$ and, for some $\gamma \in \mathcal{C}_+(E)$,*

$$Gh + \beta h - \alpha h^2 = -\gamma h. \tag{1.19}$$

Then the historical (G, α, β)-superprocess $\hat{\mathcal{X}}$ started in $\hat{\mathcal{X}}_0 = \mu \in \mathcal{M}(E)$ and the historical $(G^h, h\alpha, \gamma)$-bin-bra-process $\hat{X}$ started in $\hat{X}_0 = \mathrm{Pois}(h\mu)$ can be coupled as processes such that

$$P\big[\hat{X}_t \in \cdot \,\big|(\hat{\mathcal{X}}_s)_{0\leq s\leq t}\big] = P\big[\mathrm{Pois}((h\circ\pi_t)\hat{\mathcal{X}}_t) \in \cdot \,\big|\hat{\mathcal{X}}_t\big] \quad \text{a.s.} \quad \forall t \geq 0. \tag{1.20}$$

It follows from (1.15) that the associated non-historical processes $\mathcal{X}$ and X are related by (1.2). The phrase 'coupled as processes' means that $(\hat{\mathcal{X}}_t)_{t\geq 0}$ and $(\hat{X}_t)_{t\geq 0}$ can be defined on the same probability space in such a way that (1.20) holds.

If $\hat{\mathcal{X}}$ and $\hat{X}$ are related by (1.20), then clearly the extinction of $\hat{\mathcal{X}}$ implies the extinction of $\hat{X}$ a.s. We now investigate when the converse conclusion can be drawn, i.e., when eventually all mass of the superprocess $\mathcal{X}$ descends from particles in X. Set

$$p(x) := -\log P^{\delta_x}\big[\mathcal{X}_t = 0 \quad t\text{-eventually}\big] \qquad (x \in E). \tag{1.21}$$

Here, $-\log 0 := \infty$ and we write '*t-eventually*' behind an event, depending on t, to denote the existence of a (random) time $\tau < \infty$ such that the event holds for all $t \geq \tau$. (Thus, $\{\mathcal{X}_t = 0 \ t\text{-eventually}\} := \bigcup_{\tau<\infty}\bigcap_{t\geq\tau}\{\mathcal{X}_t = 0\}$.) It is not hard to check that p, defined by (1.21), satisfies (1.4). Therefore, we call p the *infinitesimal survival probability* of $\mathcal{X}$. Note that

$$P^{\delta_x}[\mathcal{X}_t = 0] = E^{\delta_x}\big[e^{-\langle\mathcal{X}_t, \infty\rangle}\big] = e^{-\mathcal{U}_t\infty(x)} \qquad (t \geq 0,\ x \in E). \tag{1.22}$$

The following proposition is proved in Section 2.4.1.

Proposition 5 (Properties of the infinitesimal survival probability) *Assume that $\sup_{x\in E}\mathcal{U}_t\infty(x) < \infty$ for some $t > 0$. Then one has the following:*

(a) *Pointwise $\mathcal{U}_t\infty \downarrow p$ as $t \uparrow \infty$ and $\lim_{t\to\infty}\mathcal{U}_t f = p$ for all $f \in \mathcal{C}_+(E)$ with $f > 0$.*

(b) *$\mathcal{U}_t p = p$ for all $t \geq 0$.*

(c) *A function $f \in \mathcal{C}_+(E)$ satisfies $\mathcal{U}_t f = f$ for all $t \geq 0$ if and only if $f \in \mathcal{D}(G)$ and f solves*

$$Gf + \beta f - \alpha f^2 = 0. \tag{1.23}$$

(d) *If $\inf_{x\in E} p(x) > 0$, then p is continuous and p is the unique positive solution to (1.23).*

We now formulate our main theorem, which gives sufficient conditions for all mass of the superprocess $\mathcal{X}$ to descend eventually from particles in an embedded particle system X. We write $\pi_{[0,s]}$ to denote projection on $\mathcal{D}_E[0, s]$. By definition, the support $\mathrm{supp}(\mu)$ of a measure μ is the smallest closed set such that $\mu(\mathrm{supp}(\mu)^c) = 0$.

Theorem 6 (Eventual descent from an embedded particle system) *Let $\hat{\mathcal{X}}$, $\hat{X}$, and h be as in Theorem 4, and assume that $\mathcal{U} = \mathcal{U}(G, \alpha, \beta)$ satisfies $\sup_{x\in E}\mathcal{U}_t\infty(x) < \infty$ for some $t > 0$. Then $p \leq h$, and $\hat{\mathcal{X}}$ and $\hat{X}$ may be coupled as processes such that (1.20) holds and additionally*

$$\mathrm{supp}(\hat{X}_t) \supset \mathrm{supp}(\hat{\mathcal{X}}_r \circ \pi^{-1}_{[0,t]}) \quad r\text{-eventually} \quad \forall t \geq 0 \quad \text{a.s.} \tag{1.24}$$

If moreover $\inf_{x\in E} p(x) > 0$, then by Proposition 5 we may take $h = p$ in Theorem 4. In this case one has the following:

Theorem 7 (Trimmed tree of a superprocess) *Let E be a compact metrizable space, G the generator of a Feller process in E and $\alpha \in \mathcal{C}_+(E)$, $\beta \in \mathcal{C}(E)$. Assume that $\mathcal{U} = \mathcal{U}(G,\alpha,\beta)$ satisfies $\sup_{x\in E} \mathcal{U}_t\infty(x) < \infty$ for some $t > 0$ and $\inf_{x\in E} p(x) > 0$. Then the historical (G,α,β)-superprocess $\hat{\mathcal{X}}$ started in $\hat{\mathcal{X}}_0 = \mu \in \mathcal{M}(E)$ and the historical $(G^p, p\alpha, 0)$-bin-bra-process $\hat{X}$ started in $\hat{X}_0 = \mathrm{Pois}(p\mu)$ can be coupled as processes such that*

$$P\big[\hat{X}_t \in \cdot \,\big|(\hat{\mathcal{X}}_s)_{0\leq s\leq t}\big] = P\big[\mathrm{Pois}((p\circ\pi_t)\hat{\mathcal{X}}_t) \in \cdot \,\big|\hat{\mathcal{X}}_t\big] \quad \text{a.s.} \quad \forall t \geq 0, \tag{1.25}$$

and

$$\mathrm{supp}(\hat{X}_t) = \mathrm{supp}(\hat{\mathcal{X}}_r \circ \pi^{-1}_{[0,t]}) \quad r\textit{-eventually} \quad \forall t \geq 0 \quad \text{a.s.} \tag{1.26}$$

If $\hat{\mathcal{X}}$ and $\hat{X}$ are coupled as in Theorem 7, then we say that $\hat{X}$ is the *trimmed tree* of $\hat{\mathcal{X}}$. If $\mathcal{X}_t = \hat{\mathcal{X}}_t \circ \pi_t^{-1}$ and $X_t = \hat{X}_t \circ \pi_t^{-1}$ are the associated non-historical processes, then we also call X the trimmed tree of $\mathcal{X}$. Note that the death rate of X is zero, i.e., X is a binary splitting process.

Remark (Checking the assumptions on $\mathcal{U}_t\infty$ and p) Upper bounds on $\mathcal{U}_t\infty$ and lower bounds on p can be found, in practical situations, by finding solutions to an appropriate differential inequality, see Lemmas 9 and 23 below.

1.6 Finite ancestry

In this section, we investigate the assumption, in Theorems 6 and 7, that $\sup_{x\in E} \mathcal{U}_t\infty(x) < \infty$ for some $t > 0$. In particular, we show that this assumption is equivalent, in some sense, to the statement that all mass of the superprocess $\mathcal{X}$ descends eventually from finitely many ancestors.

In order to do this, we need to equip the historical (G,α,β)-superprocess $\hat{\mathcal{X}}$ with some additional structure, that makes it possible to distinguish different ancestors. To this aim, set $E' := E \times [0,1]$. Define a Feller process $\xi' = (\xi, \eta)$ on E', where for given initial conditions $(x,y) \in E\times[0,1]$, ξ is the Feller process with generator G started in x, and $\eta_t := y$ $(t \geq 0)$. Put $\alpha'(x,y) := \alpha(x)$ and $\beta'(x,y) := \beta(x)$. Let $\hat{\mathcal{X}}'$ denote the historical (G', α', β')-superprocess. Then the formula

$$\hat{\mathcal{X}}_t := \hat{\mathcal{X}}'_t \circ \psi_t^{-1} \qquad (t \geq 0) \tag{1.27}$$

gives back the original historical (G,α,β)-superprocess $\hat{\mathcal{X}}$, where ψ_t denotes the projection from $\mathcal{D}_{E\times[0,1]}[0,t]$ to $\mathcal{D}_E[0,t]$. (For the proof of (1.27), see Lemma 15 below.) The following lemma is proved in Section 2.4.2.

Lemma 8 (Finite ancestry) *Let $\hat{\mathcal{X}}$ be the historical (G,α,β)-superprocess, let $\hat{\mathcal{X}}'$ be the extended historical superprocess just defined, and let $\mathcal{U} = \mathcal{U}(G,\alpha,\beta)$. Let ℓ denote Lebesgue measure on $[0,1]$. Then one has the relations* (i) $\Leftrightarrow$ (ii) $\Rightarrow$ (iii), *where*

$$\begin{aligned}
&\text{(i)} \quad \sup_{x\in E} \mathcal{U}_t\infty(x) < \infty \quad \textit{for some } t > 0,\\
&\text{(ii)} \quad P^{0,\mu\otimes\ell}[\mathrm{supp}(\hat{\mathcal{X}}'_t \circ \pi_0^{-1}) \textit{ is finite } t\textit{-eventually}] = 1 \quad \forall \mu \in \mathcal{M}(E),\\
&\text{(iii)} \quad P^{0,\mu}[\mathrm{supp}(\hat{\mathcal{X}}_t \circ \pi_0^{-1}) \textit{ is finite } t\textit{-eventually}] = 1 \quad \forall \mu \in \mathcal{M}(E).
\end{aligned} \tag{1.28}$$

We interpret $\mathrm{supp}(\hat{\mathcal{X}}'_t \circ \pi_0^{-1})$ as the ancestors at time 0 of the population of $\mathcal{X}$ at time t. We have extended the underlying space E to make sure that different ancestors live a.s. on different positions. Note that if E is finite, then (iii) is always trivially fulfilled even when (i) fails.

For many superprocesses, it is actually the case that

$$\sup_{x\in E} \mathcal{U}_t\infty(x) < \infty \qquad \forall t > 0. \tag{1.29}$$

If a superprocess $\mathcal{X}$ satisfies (1.29) then in view of Lemma 8 and formula (1.31.i) below we say that $\mathcal{X}$ has the *finite ancestry property.* A sufficient, but not necessary condition for (1.29) is that $\alpha > 0$. One has the following bound (see e.g. [FS02, Lemma 11]).

Lemma 9 (Extinction estimate) *Set* $\underline{\alpha} := \inf_{x\in E} \alpha(x)$ *and* $\overline{\beta} := \sup_{x\in E} \beta(x)$. *If* $\underline{\alpha} > 0$, *then*

$$\overline{\mathcal{U}}_t\infty \leq \frac{\overline{\beta}}{\underline{\alpha}\,(1 - e^{-\overline{\beta}t})} \quad (\overline{\beta} \neq 0) \quad \textit{and} \quad \overline{\mathcal{U}}_t\infty \leq \frac{1}{\underline{\alpha}\,t} \quad (\overline{\beta} = 0). \tag{1.30}$$

On the other hand, it is possible for a (G, α, β)-superprocess to satisfy (1.29) while $\underline{\alpha} = 0$, see [FS02, Lemmas 5 and 6].

The following consequence of the finite ancestry property is proved in Section 2.4.2 below.

Lemma 10 (Finite ancestry and preserved past property) *If* $\mathcal{X}$ *has the finite ancestry property, then*

$$\begin{array}{lll} \text{(i)} & \mathrm{supp}(\hat{\mathcal{X}}_r \circ \pi_{[0,t]}^{-1}) \textit{ is finite} & \forall\, 0 \leq t < r \quad \text{a.s.,} \\ \text{(ii)} & \mathrm{supp}(\hat{\mathcal{X}}_r \circ \pi_{[0,t]}^{-1}) \supset \mathrm{supp}(\hat{\mathcal{X}}_{r'} \circ \pi_{[0,t]}^{-1}) & \forall\, 0 \leq t < r \leq r' \quad \text{a.s.} \end{array} \tag{1.31}$$

Note that (1.31.ii) says that lines of descent (up to a given time s) can get extinct, but no new ones are created. Here, as elsewhere in this paper, the order of the statements $\forall 0 \leq t < r \leq r'$ a.s. means that the same zero set works for all times t, r, r' such that $0 \leq t < r \leq r'$. In particular, if the superprocess $\mathcal{X}$ in Theorem 7 has the finite ancestry property, then a.s. the sets $\mathrm{supp}(\hat{\mathcal{X}}_r \circ \pi_{[0,t]}^{-1})$ are finite for all $r > t$, and decrease to $\mathrm{supp}(\hat{X}_t)$ as $r \uparrow \infty$.

1.7 Methods, discussion, and outline of the proofs

The main ideas behind our proofs of Theorems 4, 6 and 7 are the simple observations about Poissonization and weighting of superprocesses in Lemmas 1 and 3. Our strategy is to construct a version of the superprocess with so much additional structure that one can distinguish all ancestors of the population alive at a given time. For such a sufficiently enriched process, we then explicitly identify the trimmed tree, and check that it is a binary splitting particle system. This is done in Proposition 35 and Lemma 36 from Section 2.4.3 below. The essential step, where a coupling of $\mathcal{X}_t$ and X_t for fixed t is improved to a coupling of $\mathcal{X}$ and X as processes, occurs in the proof of Lemma 36. Forgetting step by step some of the added structure, we then arrive at Theorems 4, 6 and 7.

Interesting side results of this approach are a number of lemmas about the 'lines of descent' of a superprocess, in particular Lemma 10, which plays an essential role in the proofs. On the other hand, our approach does not make any statements about the transition probabilities of the joint process $(\mathcal{X}_t, X_t)_{t\geq 0}$, when $\mathcal{X}$ and X (and their historical counterparts) are coupled as in Theorem 4. Another possible approach to our Theorem 4 (not followed in this paper) would be to specify a joint Markov evolution for $(\mathcal{X}, X)$ and then show that if the process is started in a state such that $X_0 = \mathrm{Pois}(h\mathcal{X}_0)$, then $X_t = \mathrm{Pois}(h\mathcal{X}_t)$ for all $t \geq 0$. Here, X

would be an autonomous binary branching particle system, while $\mathcal{X}$ would be a superprocess with an additional mass creation on the positions of the particles in X.

Our results can be generalized in several directions. If the space E is not compact but locally compact, then generalizations of our results can be derived using the compactification technique sketched at the end of Section 1.2. This requires, however, that the functions h and p can be extended to positive continuous functions on some compactification of E, i.e., that h and p are uniformly bounded away from zero. Truly local versions of our results, where h and p are only required to be locally bounded away from zero, are somewhat more subtle.

A lot of our proofs work for superprocesses whose underlying motion is a general Hunt process on a Polish space and whose activity and growth parameter are bounded and measurable, but we don't know how to treat compensated h-transforms and weighted superprocesses (Lemmas 2 and 3) in this context.

The proofs are organized as follows. After settling some notational and topological issues in Section 2.1, we formally introduce historical processes in Section 2.2 and collect some of their elementary properties. Section 2.3 treats compensated h-transforms and weighted superprocesses. In Section 2.4, finally, we prove our main results.

2 Proofs

2.1 Topological preliminaries

Let E be a Polish space (i.e., E is a separable topological space and there exists a complete metric generating the topology). We always equip E with the Borel-σ-field $\mathcal{B}(E)$. We let $B(E)$, $B_+(E)$, and $B_{[0,1]}(E)$ denote the spaces of bounded, bounded nonnegative, and $[0,1]$-valued, real measurable functions on E, respectively. Recall that if $\{f_i : i \in \mathbb{N}\} \subset B(E)$ separates points, then $\mathcal{B}(E) = \sigma(f_i : i \in \mathbb{N})$ (this result is known as Fernique's theorem). We also remind the reader of the fact that a subspace F of a Polish space E is itself Polish in the induced topology, if and only if F is a G_δ-subset of E, i.e., a countable intersection of open sets [Bou58, §6 No. 1, Theorem. 1].

Let $\mathcal{C}_b(E)$ denote the space of bounded continuous real functions on E. We write $\mathcal{M}(E)$ for the space of finite measures on E, equipped with the topology of weak convergence (with weak convergence denoted by $\Rightarrow$), under which $\mathcal{M}(E)$ is a Polish space [EK86, Theorem 3.1.7]. Recall that $\mu_n \Rightarrow \mu$ iff $\langle \mu_n, f\rangle \to \langle \mu, f\rangle$ for all $f \in \mathcal{C}_b(E)$ (see e.g. [EK86, Theorem 3.3.1]) and note that the topology on $\mathcal{M}(E)$ does not depend on the choice of the metric on E. The Borel-σ-field on $\mathcal{M}(E)$ is generated by the mappings $\mu \mapsto \mu(A)$, $A \in \mathcal{B}(E)$. If $F \subset E$ is measurable, we identify $\mathcal{M}(F)$ with the space $\{\mu \in \mathcal{M}(E) : \mu(E\backslash F) = 0\}$. In particular, when F is a G_δ-subset of E (and therefore Polish in the induced topology), then the topology of weak convergence on $\mathcal{M}(F)$ coincides with the induced topology from its embedding in $\mathcal{M}(E)$. By $\mathcal{M}_1(E) \subset \mathcal{M}(E)$ we denote the space of probability measures and $\mathcal{N}(E) \subset \mathcal{M}(E)$ denotes the space of finite point measures on E.

We denote by $\mathcal{D}_E[0,\infty)$ the space of cadlag (i.e., right-continuous with existing left limits) functions $w : [0,\infty) \to E$, equipped with the Skorohod topology. This is the J_1 topology defined in [Sko56]. The space $\mathcal{D}_E[0,\infty)$ is Polish [EK86, Theorem 3.5.6]. One has $w_n \to w$ in $\mathcal{D}_E[0,\infty)$ if and only if for each $T > 0$ there exists a sequence of strictly increasing, continuous $\lambda_n : [0,T] \to [0,\infty)$ with $\lambda_n(0) = 0$, such that

$$\lim_{n\to\infty} \sup_{t\in[0,T]} |\lambda_n(t) - t| = 0, \tag{2.1}$$

and such that (compare [EK86, Proposition 3.5.3])

$$\left.\begin{array}{lll} w_n(\lambda_n(t_n)) \to w(t) & \text{whenever} & t_n \downarrow t \\ w_n(\lambda_n(t_n)) \to w(t-) & \text{whenever} & t_n \uparrow t \end{array}\right\} \quad (t_n, t \in [0,T]). \tag{2.2}$$

Note that the topology on $\mathcal{D}_E[0,\infty)$ does not depend on the choice of the metric on E.

2.2 Historical processes

2.2.1 Hunt processes

Let E be a Polish space and let $(P_t)_{t\geq 0}$ be a measurable transition probability on E, i.e., $(t,x) \mapsto P_t(x,\cdot)$ is a (Borel) measurable map from $[0,\infty) \times E$ into $\mathcal{M}_1(E)$, $P_0(x,\cdot) = \delta_0$ for all $x \in E$ and $P_t P_s f = P_{t+s} f$ for all $s,t \geq 0$, $f \in B(E)$, where we adopt the notation

$$P_t f(x) := \int_E P_t(x, \mathrm{d}y) f(y) \qquad (x \in E,\ f \in B(E)). \tag{2.3}$$

Assume that $(P_t)_{t\geq 0}$ is the transition probability of a Markov process with cadlag sample paths in E, i.e., for every $\mu \in \mathcal{M}_1(E)$ there exists a $\mathcal{D}_E[0,\infty)$-valued random variable ξ, unique in distribution, such that $\mathcal{L}(\xi_0) = \mu$ and

$$E[f(\xi_t)|\mathcal{F}_s] = (P_{t-s}f)(\xi_s) \quad \text{a.s.} \qquad (0 \leq s \leq t,\ f \in B(E)), \tag{2.4}$$

where $(\mathcal{F}_t)_{t\geq 0}$ denotes the filtration generated by ξ. By definition, the Markov process with transition probability $(P_t)_{t\geq 0}$ is a *Hunt process* if, for every $\mathcal{D}_E[0,\infty)$-valued random variable ξ satisfying (2.4), the following statements hold (see [Sha88, Theorem (I.7.4) and Definition (V.47.3)]):

(i) **(Right property)** For every $t > 0$ and $f \in B(E)$, the map $[0,t) \ni s \mapsto P_{t-s} f(\xi_s)$ is a.s. right-continuous.

(ii) **(Quasi left-continuity)** For every increasing sequence of $\mathcal{F}_{\cdot+}$-stopping times $\tau_n \uparrow \tau$, one has $\xi_{\tau_n} \to \xi_\tau$ a.s. on $\{\tau < \infty\}$.

(2.5)

Here $\mathcal{F}_{\cdot+} = (\mathcal{F}_{t+})_{t\geq 0}$ denotes the right-continuous modification of $(\mathcal{F}_t)_{t\geq 0}$. The right property implies the strong Markov property [Sha88, Theorem (I.7.4)]. Conditions (2.5.i) and (2.5.ii) are properties of the law of ξ only and therefore being a Hunt process is a property of the transition probability. It suffices to check (2.5.i) for all $f \in \mathcal{C}_{\mathrm{b}}(E)$ [Sha88, Theorem (I.7.4)]. A Feller process on a compact metrizable space is a Hunt process (see [Sha88, Theorem (I.9.26) and Exercise (I.9.27)] or [Get75, (9.11)]). We identify a Hunt process with the collection of random variables $(\xi^x)^{x\in E}$, where ξ^x denotes the process started in $x \in E$.

We will also need time-inhomogeneous Hunt processes with a time-dependent state space E_t. We assume that the E_t are (or can be identified with) subsets of some Polish space E and that the set $\dot{E} := \{(t,x) \in [0,\infty) \times E : x \in E_t\}$ is a G_δ-subset of $[0,\infty) \times E$ (and therefore Polish in the induced topology). Let $W_{[s,\infty)} := \{w \in \mathcal{D}_E[s,\infty) : w_t \in E_t\ \forall t \geq s\}$ denote the space of all paths the process can follow after time s. Generalizing our previous definition, we say that a collection of random variables $(\xi^{s,x})^{(s,x)\in\dot{E}}$, where $\xi^{s,x}$ takes values in $W_{[s,\infty)}$, is a *time-inhomogeneous Hunt process*, if the collection of random variables $(\dot{\xi}^{(s,x)})^{(s,x)\in\dot{E}}$ defined by

$$\dot{\xi}^{(s,x)}_t := (s+t, \xi^{s,x}_{s+t}) \qquad ((s,x) \in \dot{E},\ t \geq 0) \tag{2.6}$$

is a (time-homogeneous) Hunt process in $\dot{E}$. If $(\xi^{s,x})^{(s,x)\in\dot{E}}$ is a time-inhomogeneous Hunt process then we write $P_{s,t}(x,\cdot) := P[\xi^{s,x}_t \in \cdot]$ and we let $P_{s,t} : B(E_t) \to B(E_s)$ denote the operator

$$P_{s,t}f(x) := \int_{E_t} P_{s,t}(x, \mathrm{d}y) f(y) \qquad (x \in E_s,\ f \in B(E_t)). \tag{2.7}$$

We call $(P_{s,t})_{t\geq s\geq 0}$ the semigroup associated with ξ.

2.2.2 Superprocesses with Hunt underlying motion

Let ξ be a (time-homogeneous) Hunt process in a Polish space E with semigroup $(P_t)_{t\geq 0}$ and assume that $\alpha \in B_+(E)$, $\beta \in B(E)$. Then, for every $f \in B_+(E)$, there exists a unique $\mathcal{B}([0,\infty)\times E)$-measurable nonnegative function u which is bounded on $[0,T]\times E$ for all $T>0$, solving the Cauchy integral equation

$$u_t = P_t f + \int_0^t P_{t-s}\big(\beta u_s - \alpha u_s^2\big)\,\mathrm{d}s \qquad (t \geq 0). \tag{2.8}$$

Moreover, it is shown in [Fit88, Fit91, Fit92] that there exists a unique (in law) Hunt process $(\mathcal{X}^\mu)^{\mu\in\mathcal{M}(E)}$, with continuous sample paths, such that

$$E^\mu[e^{-\langle \mathcal{X}_t, f\rangle}] = e^{-\langle \mu, \mathcal{U}_t f\rangle} \qquad (t \geq 0,\ \mu \in \mathcal{M}(E),\ f \in B_+(E)). \tag{2.9}$$

We call $\mathcal{X}$ the superprocess with underlying motion ξ, activity α and growth paramater β, or shortly the (ξ,α,β)-superprocess and we call $\mathcal{U} = \mathcal{U}(\xi,\alpha,\beta)$ its log-Laplace semigroup. In fact, $\mathcal{U}_t f$ can be defined unambiguously such that (2.9) holds for any measurable $f : E \to [0,\infty]$ [FS02, Lemma 9].

We list some elementary properties of (ξ,α,β)-superprocesses that we will need later. The following lemma is an easy consequence of (2.9).

Lemma 11 (Branching property) *Let $\mu_1,\mu_2 \in \mathcal{M}(E)$ and let $\mathcal{X}^{\mu_1}, \mathcal{X}^{\mu_2}$ be independent copiess of the (ξ,α,β)-superprocess started in μ_1,μ_2, respectively. Then*

$$\mathcal{X}^{\mu_1+\mu_2}_t := \mathcal{X}^{\mu_1}_t + \mathcal{X}^{\mu_2}_t \qquad (t \geq 0) \tag{2.10}$$

is the (ξ,α,β)-superprocess started in $\mu_1+\mu_2$.

The following lemma is proved in [Fit88].

Lemma 12 (Moment formulas) *For every $f \in B(E)$, there exists a unique $\mathcal{B}([0,\infty)\times E)$-measurable function v which is bounded on $[0,T]\times E$ for all $T>0$, such that*

$$v_t = P_t f + \int_0^t P_{t-s}(\beta v_s)\mathrm{d}s \qquad (t \geq 0) \tag{2.11}$$

and the formula $\mathcal{V}_t f := v_t$ defines a (linear) semigroup $(\mathcal{V}_t)_{t\geq 0}$ on $B(E)$. One has

$$\mathcal{V}_t f(x) = E^x\big[f(\xi_t)e^{\int_0^t \beta(\xi_s)\mathrm{d}s}\big] \qquad (t \geq 0,\ x \in E,\ f \in B(E)). \tag{2.12}$$

Moreover, for all $t \geq 0,\ f,g \in B(E)$,

$$\begin{array}{ll} \text{(i)} & E^\mu[\langle \mathcal{X}_t, f\rangle] = \langle \mu, \mathcal{V}_t f\rangle \\ \text{(ii)} & \mathrm{Cov}^\mu(\langle \mathcal{X}_t, f\rangle, \langle \mathcal{X}_t, g\rangle) = 2\displaystyle\int_0^t \mathrm{d}s\,\big\langle \mu, \mathcal{V}_s\big(\alpha\,(\mathcal{V}_{t-s}f)(\mathcal{V}_{t-s}g)\big)\big\rangle. \end{array} \tag{2.13}$$

The following lemma is an easy consequence of Lemma 12 and the fact that $0 \leq \mathcal{V}_t f \leq e^{\|\beta\|t}\|P_t f\|$ for all $f \in B_+(E)$ (where $\|\cdot\|$ denotes the supremum norm).

Lemma 13 (Absolute continuity of moment measures) *Let μ be a probability measure on E and $m \geq 0$. Then*

$$\left.\begin{array}{ll} \text{(i)} & E^{m\mu}[\mathcal{X}_t] \ll P^\mu[\xi_t \in \cdot\,] \\ \text{(ii)} & E^{m\mu}[\mathcal{X}_t \otimes \mathcal{X}_t] \ll P^\mu[\xi_t \in \cdot\,] \otimes P^\mu[\xi_t \in \cdot\,] + Q^\mu_t, \end{array}\right\} \quad (t \geq 0) \tag{2.14}$$

where Q^μ_t is the measure on $E \times E$ defined as

$$Q^\mu_t := \int_0^t \mathrm{d}s \int_E P^\mu[\xi_s \in \mathrm{d}x]\Big(P^x[\xi_{t-s} \in \cdot\,] \otimes P^x[\xi_{t-s} \in \cdot\,]\Big) \qquad (t \geq 0). \tag{2.15}$$

A measure $\mu \in \mathcal{M}(E)$ is atomless (i.e., $\mu(\{x\}) = 0$ for all $x \in E$) if and only if[1]

$$\mu \otimes \mu(\{(x_1, x_2) \in E \times E : x_1 = x_2\}) = 0. \tag{2.16}$$

The following lemma follows from formulas (2.14.ii) and (2.16).

Lemma 14 (Atomless superprocess) *Assume that $P^x[\xi_t \in \cdot\,]$ is atomless for every $t > 0$ and $x \in E$. Then $\mathcal{X}_t$ is atomless a.s. for every $t > 0$ and initial state $\mu \in \mathcal{M}(E)$.*

Our next lemma is the following:

Lemma 15 (Image property) *Let E, F be Polish spaces, let $\psi : E \to F$ be continuous and let $\xi = (\xi^x)^{x \in E}$ and $\eta = (\eta^y)^{y \in F}$ be Hunt processes in E, F, respectively, satisfying*

$$\psi(\xi^x_t) = \eta^{\psi(x)}_t \qquad (x \in E,\ t \geq 0). \tag{2.17}$$

Assume that $\alpha_F \in B_+(F)$, $\beta_F \in B(F)$ and let $\alpha_E \in B_+(E)$, $\beta_E \in B(E)$ be given by

$$\alpha_E := \alpha_F \circ \psi \quad \textit{and} \quad \beta_E := \beta_F \circ \psi. \tag{2.18}$$

Let $\mathcal{X}$ be the (ξ, α_E, β_E)-superprocess with initial state $\mu \in \mathcal{M}(E)$. Then

$$\mathcal{Y}_t := \mathcal{X}_t \circ \psi^{-1} \qquad (t \geq 0) \tag{2.19}$$

is the $(\eta, \alpha_F, \beta_F)$-superprocess with initial state $\mu \circ \psi^{-1}$.

Proof Let P^E and P^F denote the semigroups associated with the processes ξ and η, respectively. Formula (2.17) implies that $P^E_t(f \circ \psi) = (P^F_t f) \circ \psi$ for all $f \in B(F)$. Using this fact and (2.18) it is not hard to show that also $\mathcal{U}^E_t(f \circ \psi) = (\mathcal{U}^F_t f) \circ \psi$ for all $f \in B_+(F)$, where $\mathcal{U}^E = \mathcal{U}(\xi, \alpha_E, \beta_E)$ and $\mathcal{U}^F = \mathcal{U}(\eta, \alpha_F, \beta_F)$ are the log-Laplace semigroups of $\mathcal{X}$ and $\mathcal{Y}$, respectively. Let $(\mathcal{F}_t)_{t \geq 0}$ be the filtration generated by $\mathcal{X}$. Then, for all $0 \leq s \leq t$

$$\begin{aligned} E[e^{-\langle \mathcal{X}_t \circ \psi^{-1}, f\rangle} \,|\, \mathcal{F}_s] &= E[e^{-\langle \mathcal{X}_t, f \circ \psi\rangle} \,|\, \mathcal{F}_s] = e^{-\langle \mathcal{X}_s, U^E_{t-s}(f \circ \psi)\rangle} \\ &= e^{-\langle \mathcal{X}_s, (U^F_{t-s} f) \circ \psi\rangle} = e^{-\langle \mathcal{X}_s \circ \psi^{-1}, U^F_{t-s} f\rangle} \qquad (f \in B_+(F)). \end{aligned} \tag{2.20}$$

This shows that $(\mathcal{X}_t \circ \psi^{-1})_{t \geq 0}$ is a Markov process and that its transition probabilities coincide with those of the $(\eta, \alpha_F, \beta_F)$-superprocess. Since ψ is continuous, $\mathcal{X}_t \circ \psi^{-1}$ has continuous sample paths. ∎

The following simple observation will be useful later.

[1] If μ has an atom then (2.16) is obviously violated. Conversely, if μ has no atoms then, using tightness, it is not hard to show that for every $\varepsilon \geq 0$ there exists a Borel measurable partition $\{A_1, \ldots, A_n\}$ of E such that $\mu(A_i) \leq \varepsilon$ $(i = 1, \ldots, n)$, which in turn implies (2.16).

Lemma 16 (Preserved sets)
(a) *If* $F \subset E$ *is measurable and* $P^x[\xi_t \in F] = 1\ \forall t \geq 0$ $(x \in F)$*, then*

$$P^\mu[\mathcal{X}_t \in \mathcal{M}(F)] = 1 \quad \forall t \geq 0 \qquad (\mu \in \mathcal{M}(F)). \tag{2.21}$$

(b) *If* $F \subset E$ *is a* G_δ*-set and* $P^x[\xi_t \in F,\ \forall t \geq 0,\ \xi_{t-} \in F\ \forall t > 0] = 1$ $(x \in F)$*, then*

$$P^\mu[\mathcal{X}_t \in \mathcal{M}(F)\ \forall t \geq 0] = 1 \qquad (\mu \in \mathcal{M}(F)). \tag{2.22}$$

Proof Statement (a) follows from (2.14.i) while (b) follows by applying Lemma 15 to the inclusion map $F \subset E$, where we use that the restriction of ξ to F is again a Hunt process. The assumption that F is a G_δ-set guarantees that F is a Polish space and that the event $\{\mathcal{X}_t \in \mathcal{M}(F)\ \forall t \geq 0\}$ is Borel measurable. ∎

We conclude this section by constructing superprocesses with time-inhomogeneous underlying motion. Let $\xi = (\xi^{s,x})^{(s,x)\in\dot{E}}$ be a time-inhomogeneous Hunt process as defined at the end of the last section, and assume that $\dot{\alpha} \in B_+(\dot{E})$, $\dot{\beta} \in B(\dot{E})$. Let $\dot{\xi}$ be the time-homogeneous Hunt process in (2.6) and let $\dot{\mathcal{X}}$ denote the $(\dot{\xi}, \dot{\alpha}, \dot{\beta})$-superprocess. Using Lemma 13 we see that $\dot{\mathcal{X}}_t^{\delta_s\otimes\mu}$ is concentrated on $\{s+t\} \times E_{s+t}$ a.s. $\forall t \geq 0$. Since $\dot{\mathcal{X}}^{\delta_s\otimes\mu}$ has continuous sample paths and since $\{\delta_t \otimes \mu : t \geq 0,\ \mu \in \mathcal{M}(E_t)\} \subset \mathcal{M}(\dot{E})$ is closed, there exists a process $\mathcal{X}^{s,\mu}$ with continuous sample paths in $\mathcal{M}(E)$ such that $\mathcal{X}_{s+t}^{s,\mu} \in \mathcal{M}(E_{s+t})$ for all $t \geq 0$ and

$$\dot{\mathcal{X}}_t^{\delta_s\otimes\mu} = \delta_{s+t} \otimes \mathcal{X}_{s+t}^{s,\mu}. \tag{2.23}$$

Set $\dot{\mathcal{M}} := \{(t,\mu) \in [0,\infty) \times \mathcal{M}(E) : \mu \in \mathcal{M}(E_t)\}$. It is not hard to check that $\mathcal{X} = (\mathcal{X}^{s,\mu})^{(s,\mu)\in\dot{\mathcal{M}}}$ is a time-inhomogeneous Hunt process with continuous sample paths, and

$$E^{s,\mu}[e^{-\langle \mathcal{X}_t, f\rangle}] = e^{-\langle \mu, \mathcal{U}_{s,t} f\rangle} \qquad (t \geq s \geq 0,\ \mu \in \mathcal{M}(E_s),\ f \in B_+(E_t)), \tag{2.24}$$

where $(\mathcal{U}_{s,t}f)_{s\in[0,t]} =: u \in B_+(\{(s,x) \in [0,t] \times E : x \in E_s\})$ solves the equation

$$u_s = P_{s,t}f + \int_s^t P_{s,r}\big(\beta_r u_r - \alpha_r u_r^2\big)\,\mathrm{d}r \qquad (s \in [0,t]). \tag{2.25}$$

Here $\alpha_t(x) := \dot{\alpha}(t,x)$, $\beta_t(x) := \dot{\beta}(t,x)$ $((t,x) \in \dot{E})$, and $(P_{s,t})_{t\geq s\geq 0}$ is the (time-inhomogeneous) semigroup associated with ξ. We call $\mathcal{X}$ the (time-inhomogeneous) (ξ, α_t, β_t)-superprocess and we call $(\mathcal{U}_{s,t})_{t\geq s\geq 0}$ the (time-inhomogeneous) log-Laplace semigroup associated with $\mathcal{X}$.

2.2.3 Historical superprocesses

Let $\xi = (\xi^x)^{x\in E}$ be a Hunt process in a Polish space E and let $\hat{\xi} = (\hat{\xi}^{s,w})^{s\geq 0,\ w\in\mathcal{D}_E[0,s]}$ be the associated path process, defined as in (1.14). Identify, as usual, $\mathcal{D}_E[0,s]$ with the subspace of $\mathcal{D}_E[0,\infty)$ consisting of paths stopped at time s, and define $\tilde{E} \subset [0,\infty) \times \mathcal{D}_E[0,\infty)$ by

$$\tilde{E} := \{(s,w) : s \geq 0,\ w \in \mathcal{D}_E[0,s]\}. \tag{2.26}$$

Then $(\hat{\xi}^{s,w})^{(s,w)\in\tilde{E}}$ is a time-inhomogeneous Hunt process (see [DP91, Proposition 2.1.2]). If $\mathcal{X}$ is a (ξ,α,β)-superprocess, then by definition the historical (ξ,α,β)-superprocess $\hat{\mathcal{X}}$ is the (time-inhomogeneous) $(\hat{\xi}, \hat{\alpha}_t, \hat{\beta}_t)$-superprocess, where $\hat{\alpha}_t(w) := \alpha(w(t))$ and $\hat{\beta}_t(w) := \beta(w(t))$

($(t,w)\in\tilde E$). We are now in a situation where we can prove some of the elementary properties of historical superprocesses mentioned in the Section 1.

Proof of formula (1.15) If $\hat\xi$ is the path process associated with a Hunt process ξ, started at time $s\geq 0$ in $w\in\mathcal{D}_E[0,s]$, then $\xi_t:=\pi_{s+t}(\hat\xi_{s+t})$ $(t\geq 0)$ gives back the original Hunt process ξ started in $\pi_s(\hat\xi_s)$. Moreover, the map $(t,w)\mapsto w(t)$ from $\tilde E$ into E is continuous. (Note that this is true even though the map $w\mapsto w(t)$ from $\mathcal{D}_E[0,\infty)$ into E is in general discontinuous.) Therefore, Lemma 15 (the image property of superprocesses) shows that if $(\hat{\mathcal{X}}_t)_{t\geq s}$ is the historical (ξ,α,β)-superprocess started at time $s\geq 0$ in $\hat\mu\in\mathcal{D}_E[0,s]$, then

$$\mathcal{X}_t:=\hat{\mathcal{X}}_{s+t}\circ\pi_{s+t}^{-1}\qquad(t\geq 0)\tag{2.27}$$

is the (nonhistorical) (ξ,α,β)-superprocess started in $\hat\mu\circ\pi_s^{-1}$. ∎

One of the driving ideas behind the development of historical superprocesses has been the desire to have a means for distinguishing those parts of the population that descend from different ancestors. However, all that a path in $\mathcal{D}_E[0,t]$ tells you is where in space these ancestors have lived in the past. Let us say that the underlying motion ξ has the *distinct path property* if the law of $(\xi_s)_{s\in[0,t]}$ (considered as a $\mathcal{D}_E[0,t]$-valued random variable) is atomless for every $t>0$ and for every initial state $\xi_0=x\in E$. This is called 'Property S' in [Daw93, Definition (12.2.2.6)], and occurs as formula (3.18) in [DP91]. In this case, one imagines that different ancestors follow a.s. different paths, and therefore it should be possible to recognize an ancestor from its path. As an immediate consequence of Lemma 14, one has the following: (An analogue of this result in a spatially homogeneous setting, but for more general branching mechanisms, can be found in [DP91, Proposition 4.1.8 (b)].)

Lemma 17 (Atomless historical superprocesses) *If ξ has the distinct path property, then $\hat{\mathcal{X}}_t$ is atomless a.s. $\forall t>0$.*

The following lemma, which is an immediate consequence of Lemma 15 (the image property), shows that one can always extend the space of a superprocess with Feller underlying motion, such that the new underlying motion has the distinct path property, and the historical superprocess is atomless. (Compare [DP91, Remark below Proposition 3.5].)

Lemma 18 (Extended historical superprocess) *Let G be the generator of a Feller process ξ in a compact metrizable space E, and let $\alpha\in\mathcal{C}_+(E)$, $\beta\in\mathcal{C}(E)$. Let η be a Feller process in a compact metrizable space F such that η has the distinct path property. Let G' denote the generator of the Feller process (ξ,η) in $E\times F$, where for given initial conditions, ξ and η evolve independently. Put $\alpha'(x,y):=\alpha(x)$ and $\beta'(x,y):=\beta(x)$. Let ψ_t denote the projection from $\mathcal{D}_{E\times F}[0,t]$ to $\mathcal{D}_E[0,t]$. Let $\hat\mu,\hat\rho$ be finite measures on $\mathcal{D}_E[0,s],\mathcal{D}_F[0,s]$, respectively, and assume that $\hat\rho$ is atomless. If $\hat{\mathcal{X}}'$ is the historical (G',α',β')-superprocess started at time s in $\hat\mu\otimes\hat\rho$, then*

$$\hat{\mathcal{X}}_t:=\hat{\mathcal{X}}'_t\circ\psi_t^{-1}\qquad(t\geq s)\tag{2.28}$$

is the historical (G,α,β)-superprocess started at time s in $\hat\mu$. Moreover, $\hat{\mathcal{X}}'_t$ is atomless a.s. $\forall t\geq s$ and its underlying motion has the distinct path property.

(For example, one may take for η Brownian motion on the unit circle.)

We return to the more general case of Hunt underlying motion.

Lemma 19 (Finite dimensional projections) *Let $\mathcal{X}$ be a (ξ,α,β)-superprocess with log-Laplace semigroup $\mathcal{U}=\mathcal{U}(\xi,\alpha,\beta)$ and let $\hat{\mathcal{X}}$ be the associated historical (ξ,α,β)-superprocess. Then, for all $n\geq 0$, $0=t_0<t_1<\cdots<t_{n+1}$, and $f\in B_+(E^{n+2})$,*

$$\begin{aligned} &E^{t_n,\hat{\mu}}\Big[e^{-\int_{\mathcal{D}_E[0,t_{n+1}]}\hat{\mathcal{X}}_{t_{n+1}}(\mathrm{d}w)f(w_{t_0},\ldots,w_{t_{n+1}})}\Big] \\ &\qquad = e^{-\int_{\mathcal{D}_E[0,t_n]}\hat{\mu}(\mathrm{d}w)\mathcal{U}_{t_{n+1}-t_n}f(w_{t_0},\ldots,w_{t_n},\,\cdot\,)(w_{t_n})}. \end{aligned} \tag{2.29}$$

Conversely, any Markov process $\hat{\mathcal{X}}$ with time-dependent state space $\mathcal{D}_E[0,t]$ and continuous sample paths, satisfying (2.29), is the historical (ξ,α,β)-superprocess.

Proof The fact that $\hat{\mathcal{X}}$ satisfies (2.29) can be found in [DP91, Theorem 2.2.5 (b)] or [Daw93, Theorem 12.3.4]. Conversely, if a Markov process $\hat{\mathcal{X}}$ satisfies (2.29), then, for all $0\leq k\leq n$,

$$\begin{aligned} &E^{t_k,\hat{\mu}}\Big[e^{-\int_{\mathcal{D}_E[0,t_{n+1}]}\hat{\mathcal{X}}_{t_{n+1}}(\mathrm{d}w)f(w_{t_0},\ldots,w_{t_{n+1}})}\Big] \\ &\qquad = e^{-\int_{\mathcal{D}_E[0,t_k]}\hat{\mu}(\mathrm{d}w)f_k(w_{t_0},\ldots,w_{t_k})}, \end{aligned} \tag{2.30}$$

where we have inductively defined functions $f_l\in B_+(E^{l+1})$ by

$$\begin{aligned} f_{n+1}(x_0,\ldots,x_{n+1}) &:= f(x_0,\ldots,x_{n+1}), \\ f_l(x_0,\ldots,x_l) &:= \mathcal{U}_{t_{l+1}-t_l}f_{l+1}(x_0,\ldots,x_l,\,\cdot\,)(x_l) \qquad (k\leq l\leq n). \end{aligned} \tag{2.31}$$

The expectations in (2.30) clearly determine the transition probabilities of $\hat{\mathcal{X}}$ uniquely. ∎

Note that formula (2.30) says that, if $F(w):=f(w_{t_0},\ldots,w_{t_{n+1}})$ and $\hat{\mathcal{U}}$ is the (time-inhomogeneous) log-Laplace semigroup of $\hat{\mathcal{X}}$, then

$$\hat{\mathcal{U}}_{t_k,t_{n+1}}F(w)=f_k(w_{t_0},\ldots,w_{t_k}). \tag{2.32}$$

Lemma 20 (Mean of historical superprocess) *Let $\hat{\mathcal{X}}$ be the historical (ξ,α,β)-superprocess. Then, for any $\mu\in\mathcal{M}_1(E)$ and $m\geq 0$,*

$$E^{m\mu}[\hat{\mathcal{X}}_t](\mathrm{d}w)=m\,e^{\int_0^t\beta(w_s)\mathrm{d}s}\,P^\mu[(\xi_s)_{s\in[0,t]}\in\mathrm{d}w] \qquad (t\geq 0). \tag{2.33}$$

In particular, if $\alpha=0$ then $\hat{\mathcal{X}}_t$ is deterministic and given by the right-hand side of (2.33).

Proof By Lemma 12, the mean of a superprocess does not depend on the activity. Therefore, it suffices to prove that the historical $(\xi,0,\beta)$-superprocess is deterministic and given by the right-hand side of (2.33). Define $\hat{\mathcal{X}}_t(\mathrm{d}w)$ $(t\geq 0)$ by the right-hand side of (2.33). Let $\mathcal{U}=\mathcal{U}(\xi,0,\beta)$ denote the log-Laplace semigroup of the (non-historical) $(\xi,0,\beta)$-superprocess. Since $\alpha=0$, $\mathcal{U}$ coincides with the linear semigroup $\mathcal{V}$ in formula (2.12). It follows that, for $n\geq 0$, $0=t_0<t_1<\cdots<t_{n+1}$, and $f\in B_+(E^{n+2})$,

$$\begin{aligned} &\int_{\mathcal{D}_E[0,t_{n+1}]}\hat{\mathcal{X}}_{t_{n+1}}(\mathrm{d}w)f(w_{t_0},\ldots,w_{t_{n+1}}) \\ &\quad=\int_{\mathcal{D}_E[0,t_{n+1}]}m\,e^{\int_0^{t_{n+1}}\beta(w_s)\mathrm{d}s}\,f(w_{t_0},\ldots,w_{t_{n+1}})P^\mu[(\xi_s)_{s\in[0,t]}\in\mathrm{d}w] \\ &\quad=mE^\mu\Big[e^{\int_0^{t_{n+1}}\beta(\xi_s)\mathrm{d}s}f(\xi_{t_0},\ldots,\xi_{t_{n+1}})\Big] \\ &\quad=mE^\mu\Big[e^{\int_0^{t_n}\beta(\xi_s)\mathrm{d}s}E\big[e^{\int_{t_n}^{t_{n+1}}\beta(\xi_s)\mathrm{d}s}f(\xi_{t_0},\ldots,\xi_{t_{n+1}})\big|(\xi_s)_{s\in[0,t_n]}\big]\Big] \\ &\quad=mE^\mu\Big[e^{\int_0^{t_n}\beta(\xi_s)\mathrm{d}s}\tilde{f}(\xi_{t_0},\ldots,\xi_{t_n})\Big]=\int_{\mathcal{D}_E[0,t_n]}\hat{\mathcal{X}}_{t_n}(\mathrm{d}w)\tilde{f}(w_{t_0},\ldots,w_{t_n}), \end{aligned} \tag{2.34}$$

where

$$\tilde{f}(x_0,\ldots,x_n) := \mathcal{U}_{t_{n+1}-t_n} f(x_0,\ldots,x_n,\,\cdot\,)(x_n). \tag{2.35}$$

Thus, $\hat{\mathcal{X}}$ satisfies (2.29). Since $\hat{\mathcal{X}}$ is a Markov process with continuous sample paths it follows from Lemma 19 that $(\hat{\mathcal{X}}_t)_{t\geq 0}$ is the historical $(\xi,0,\beta)$-superprocess started at time 0 in $m\mu$. ∎

The following lemma will be important in the proof of Lemma 10.

Lemma 21 (Preserved past property) *Let $\hat{\mathcal{X}}$ be the historical (ξ,α,β)-superprocess started at time $s\geq 0$ in $\hat{\mu}\in\mathcal{D}_E[0,s]$.*
(a) *If $F\subset\mathcal{D}_E[0,s]$ is measurable, then*

$$P^{s,\hat{\mu}}[\hat{\mathcal{X}}_t\circ\pi^{-1}_{[0,s]}\in\mathcal{M}(F)]=1 \quad \forall t\geq s \qquad (\hat{\mu}\in\mathcal{M}(F)). \tag{2.36}$$

(b) *If $F\subset\mathcal{D}_E[0,s]$ is a G_δ-set, then*

$$P^{s,\hat{\mu}}[\hat{\mathcal{X}}_t\circ\pi^{-1}_{[0,s]}\in\mathcal{M}(F)\ \forall t\geq s]=1 \qquad (\hat{\mu}\in\mathcal{M}(F)). \tag{2.37}$$

(c) *If $F,F^{\rm c}\subset\mathcal{D}_E[0,s]$ are G_δ-sets, then*

$$P^{s,\hat{\mu}}\big[1_{\{\hat{\mathcal{X}}_{t'}\circ\pi^{-1}_{[0,s]}(F)>0\}}\leq 1_{\{\hat{\mathcal{X}}_t\circ\pi^{-1}_{[0,s]}(F)>0\}}\ \forall t'\geq t\geq s\big]=1. \tag{2.38}$$

Proof Recall the definition of $\tilde{E}$ in (2.26) and set $\tilde{F}:=\{(t,w)\in\tilde{E}: t\geq s,\ \pi_{[0,s]}(w)\in F\}$. If F is measurable then $\tilde{F}$ is measurable. Moreover, since $\pi_{[0,s]}$ is the pointwise limit of a sequence of continuous functions (compare [EK86, Proposition 3.7.1]), $\tilde{F}$ is a G_δ-set when F is a G_δ-set. The path process $\hat{\xi}$ satisfies

$$P^{s',w}[(t,\hat{\xi}_t)\in\tilde{F}\ \forall t\geq s',\ (t,\hat{\xi}_{t-})\in\tilde{F}\ \forall t>s']=1 \quad ((s',w)\in\tilde{F}). \tag{2.39}$$

Therefore (a) follows from Lemma 16 (a) and (b) follows from Lemma 16 (b). To prove (c), use the branching property (Lemma 11) to write

$$\hat{\mathcal{X}}^{s,\hat{\mu}}_t=\hat{\mathcal{X}}^{s,1_F\hat{\mu}}_t+\hat{\mathcal{X}}^{s,1_{F^{\rm c}}\hat{\mu}}_t \quad \forall t\geq s \quad \text{a.s.} \tag{2.40}$$

Then, applying (b) to F and $F^{\rm c}$,

$$\hat{\mathcal{X}}^{s,\hat{\mu}}_t\circ\pi^{-1}_{[0,s]}(F)=\hat{\mathcal{X}}^{s,1_F\hat{\mu}}_t\circ\pi^{-1}_{[0,s]}(F)+\hat{\mathcal{X}}^{s,1_{F^{\rm c}}\hat{\mu}}_t\circ\pi^{-1}_{[0,s]}(F)=\langle\hat{\mathcal{X}}^{s,1_F\hat{\mu}}_t\circ\pi^{-1}_{[0,s]},1\rangle+0 \quad \forall t\geq s \quad \text{a.s.} \tag{2.41}$$

Applying the strong Markov property to the stopping time $\inf\{t\geq s:\hat{\mathcal{X}}^{s,1_F\hat{\mu}}_t=0\}$ it is not hard to see that

$$1_{\{\hat{\mathcal{X}}^{s,1_F\hat{\mu}}_{t'}\circ\pi^{-1}_{[0,s]}>0\}}\leq 1_{\{\hat{\mathcal{X}}^{s,1_F\hat{\mu}}_t\circ\pi^{-1}_{[0,s]}>0\}} \quad \forall t'\geq t\geq s \quad \text{a.s.}, \tag{2.42}$$

which proves (c). ∎

2.2.4 Historical binary branching particle systems

Historical binary branching particle systems can be introduced in much the same way as historical superprocesses. First, binary branching particle systems whose underlying motion is a Hunt process ξ with cadlag sample paths in a Polish space E, are defined through their generating semigroup, which in turn is defined via the unique solution to a Cauchy integral equation of the form (2.8). If ξ is such a Hunt process and $b, d \in B_+(E)$, then the historical (ξ, b, d)-bin-bra-process $\hat{X}$ is the (time-inhomogeneous) $(\hat{\xi}, \hat{b}, \hat{d})$-bin-bra-process, where $\hat{\xi}$ is the path process associated with ξ and $\hat{b}(t, w) := b(w(t))$, $\hat{d}(t, w) := d(w(t))$. Because this is very similar to what we have already seen, we skip the details.

Many of the elemantary properties of historical superprocesses have analogues for historical binary branching particle systems. For example, if the underlying motion has the distinct path property, then the historical binary branching particle system at time $t > 0$ is a.s. a simple point measure. (One way to prove this is to use Poissonization and Lemma 17.) Also the formula for the finite dimensional projections of a historical superprocess (Lemma 19) has a straightforward analogue for particle systems.

2.3 Compensated h-transforms and weighted superprocesses

2.3.1 Preliminaries from semigroup theory

Let E be a compact metrizable space and let $\mathcal{C}(E)$ be the Banach space of continuous real functions on E, equipped with the supremum norm, denoted by $\|\cdot\|$. Let $S = (S_t)_{t\geq 0}$ be a semigroup of bounded linear operators on $\mathcal{C}(E)$. By definition, S is strongly continuous if $\lim_{t\to 0} \|S_t f - f\| = 0$ for all $f \in \mathcal{C}(E)$. S is positive if $f \geq 0$ implies $S_t f \geq 0$ $(t \geq 0)$. For $\lambda \in \mathbb{R}$, let us say that S is *λ-contractive* if $\|S_t f\| \leq e^{\lambda t}\|f\|$ $(t \geq 0)$. The following version of the Hille-Yosida theorem can easily be derived from [EK86, Theorem 4.2.2 and Proposition 1.1.5.(b)].

Lemma 22 (Hille-Yosida theorem) *A linear operator G on $\mathcal{C}(E)$ with domain $\mathcal{D}(G)$ is the generator of a strongly continuous, positive, λ-contractive semigroup S on $\mathcal{C}(E)$, with $\lambda \in \mathbb{R}$, if and only if*

(i) *G is closed,*

(ii) *$\mathcal{D}(G)$ is dense in $\mathcal{C}(E)$,*

(iii) *$Gf(x) \leq \lambda f(x)$ whenever $f \in \mathcal{D}(G)$ assumes its maximum over E in a point $x \in E$ with $f(x) \geq 0$,*

(iv) *For all $f \in \mathcal{D}(G)$ there exists a continuously differentiable $u : [0, \infty) \to \mathcal{C}(E)$ such that $u_0 = f$, $u_t \in \mathcal{D}(G)$, and $\frac{\partial}{\partial t} u_t = G u_t \quad (t \geq 0)$.*
$$(2.43)$$

The function u in (iv) is unique and given by $S_t f = u_t$ $(t \geq 0,\ f \in \mathcal{D}(G))$.

If S is a strongly continuous, positive semigroup, and instead of the λ-contractivity, S satisfies the stronger requirement that $S_t 1 = 1$ $(t \geq 0)$, then S is called a Feller semigroup.

Let G be the generator of a strongly continuous, positive, λ-contractive semigroup on $\mathcal{C}(E)$ and let $\alpha \in \mathcal{C}_+(E)$, $\beta \in \mathcal{C}(E)$. By definition, a *mild* solution to the Cauchy problem (1.7) is a continuous function $u : [0, \infty) \to \mathcal{C}(E)$ satisfying

$$u_t = S_t f + \int_0^t S_{t-s}\big(\beta u_s - \alpha u_s^2\big)\, \mathrm{d}s \qquad (t \geq 0). \tag{2.44}$$

By definition, u is a *classical* solution to (1.7) if $t \mapsto u_t$ is continuously differentiable in $\mathcal{C}(E)$, $u_t \in \mathcal{D}(G) \cap \mathcal{C}(E)$ for all $t \geq 0$, and (1.7) holds. Every classical solution is a mild solution. For classical solutions one has the following comparison result.

Lemma 23 (Sub- and supersolutions) *Fix $T > 0$ and assume that u is a classical solution to (1.7) on $[0,T]$. Assume that $\tilde{u} : [0,T] \to \mathcal{C}(E)$ is continuously differentiable, $\tilde{u}_t \in \mathcal{D}(G)$ for all $t \in [0,T]$ and*

$$\begin{cases} \frac{\partial}{\partial t}\tilde{u}_t \leq G\tilde{u}_t + \beta\tilde{u}_t - \alpha\tilde{u}_t^2 & (t \in [0,T]), \\ \tilde{u}_0 \leq f, \end{cases} \tag{2.45}$$

where $f \in \mathcal{C}(E)$. Then $\tilde{u}_T \leq u_T$. The same holds with all inequality signs reversed.

Proof This is a standard application of the maximum principle. Set $g_t := G\tilde{u}_t + \beta\tilde{u}_t - \alpha\tilde{u}_t^2 - \frac{\partial}{\partial t}\tilde{u}_t$ and $\Delta_t := u_t - \tilde{u}_t$ $(t \in [0,T])$. Then Δ solves

$$\begin{cases} \frac{\partial}{\partial t}\Delta_t = G\Delta_t + \beta\Delta_t - \alpha\,(u_t + \tilde{u}_t)\Delta_t + g_t & (t \in [0,T]), \\ \Delta_0 = f - \tilde{u}_0. \end{cases} \tag{2.46}$$

We will show that $\Delta_T \geq 0$. Let γ be a constant such that $\lambda + \beta(x) - \alpha(x)\,(u_t(x) + \tilde{u}_t(x)) + \gamma < 0$ for all $(t,x) \in [0,T] \times E$, where λ is the constant in property (2.43.iii). Then $\tilde{\Delta}_t := e^{\gamma t}\Delta_t$ solves

$$\begin{cases} \frac{\partial}{\partial t}\tilde{\Delta}_t = (G - \lambda)\tilde{\Delta}_t + \{\lambda + \beta - \alpha\,(u_t + \tilde{u}_t) + \gamma\}\tilde{\Delta}_t + g_t e^{\gamma t} & (t \in [0,T]), \\ \tilde{\Delta}_0 = f - \tilde{u}_0. \end{cases} \tag{2.47}$$

Imagine that $\Delta_t(x) < 0$ for some $x \in E$. Then $\tilde{\Delta}$ must assume a (strictly) negative minimum in some point $(t,x) \in (0,T] \times E$. But in such a point one would have $\frac{\partial}{\partial t}\tilde{\Delta}_t(x) \leq 0$ while $(G - \lambda)\tilde{\Delta}_t(x) \geq 0$ by property (2.43.iii) and $\{\lambda + \beta(x) - \alpha(x)\,(u_t(x) + \tilde{u}_t(x)) + \gamma(x)\}\tilde{\Delta}_t(x) > 0$, which contradicts (2.47). The same proof works with all inequality signs reversed. ∎

Existence of solutions to (1.7) is guaranteed by the following lemma.

Lemma 24 (Classical and mild solutions to a semilinear Cauchy problem) *For each $f \in \mathcal{C}(E)$ there exists a unique mild solution u of (1.7) up to an 'explosion time' T, with $\lim_{t\uparrow T}\|u_t\| = \infty$ if T is finite. $\mathcal{U}_t f := u_t$ $(t < T)$ defines a continuous map from $\mathcal{C}(E)$ into itself. If $f \in \mathcal{D}(G)$ then the mild solution to (1.7) is a classical solution. The time T is infinite if $f \geq 0$, in which case also $u \geq 0$, or if $\alpha = 0$.*

Proof The statements about mild solutions follow from [Paz83, Theorems 6.1.2 and 6.1.4] and the statement about classical solutions from [Paz83, Theorems 6.1.5]. If $f \in \mathcal{D}(G) \cap \mathcal{C}_+(E)$, then using Lemma 23 it is easy to prove that the classical solution to (1.7) satisfies $0 \leq u \leq e^{(\lambda + \|\beta\|)t}\|f\|$. Since $\mathcal{D}(G)$ is dense, $\mathcal{C}_+(E)$ is the closure of its interior, and $\mathcal{U}_t$ is continuous, the same bounds hold for mild solutions. The fact that solutions do not explode in the linear case $\alpha = 0$ follows from [Paz83, Theorem 6.1.2]. ∎

2.3.2 Superprocesses with Feller underlying motion

Let E be compact and metrizable, G the generator of a Feller semigroup $(P_t)_{t\geq 0}$ on $\mathcal{C}(E)$, $\alpha \in \mathcal{C}_+(E)$ and $\beta \in \mathcal{C}(E)$. Then one has the following:

Lemma 25 (Feller property of superprocess) *Let $\mathcal{X}$ be the (G,α,β)-superprocess with log-Laplace semigroup $\mathcal{U} = \mathcal{U}(G,\alpha,\beta)$. Then $\mathcal{X}$ is a Feller process. For each $f \in \mathcal{C}_+(E)$, the map $(t,x) \mapsto \mathcal{U}_t f(x)$ from $[0,\infty) \times E$ into $[0,\infty)$ is continuous.*

Proof Since E is compact, the space $\mathcal{M}(E)$ is locally compact. By [Paz83, Theorem 6.1.4], $(t,x) \mapsto \mathcal{U}_t f(x)$ is jointly continuous in t and x whenever $f \in \mathcal{C}_+(E)$. Therefore, and by (1.8)

$$E^{\mu_n}\big[e^{-\langle \mathcal{X}_{t_n}, f\rangle}\big] \to E^{\mu}\big[e^{-\langle \mathcal{X}_t, f\rangle}\big] \quad \text{as} \quad \mu_n \Rightarrow \mu,\ t_n \to t \qquad (f \in \mathcal{C}_+(E)). \tag{2.48}$$

By the Stone-Weierstrass Theorem, the linear span of all functions of the form $\mu \mapsto e^{-\langle \mu, f\rangle}$ with $f \in \mathcal{C}(E)$ and $f > 0$ is dense in the space $\mathcal{C}_0(\mathcal{M}(E))$ of continuous functions on $\mathcal{M}(E)$ vanishing at infinity. Thus, (2.48) implies that $\mathcal{L}^{\mu_n}(\mathcal{X}_{t_n}) \Rightarrow \mathcal{L}^{\mu}(\mathcal{X}_t)$ whenever $\mu_n \Rightarrow \mu$, $t_n \to t$. It is not hard to see that the semigroup of $\mathcal{X}$ maps functions vanishing at infinity into functions vanishing at infinity, and therefore $\mathcal{X}$ is a Feller process. ∎

2.3.3 Compensated h-transforms of Feller processes

Proof of Lemma 2 (first part) Let $(P_t)_{t\geq 0}$ be the Feller semigroup with generator G. Define a linear semigroup $S = (S_t)_{t\geq 0}$ on $\mathcal{C}(E)$ by

$$S_t f := \frac{1}{h} P_t(hf) \qquad (t \geq 0,\ f \in \mathcal{C}(E)). \tag{2.49}$$

Since h is bounded away from zero and $(P_t)_{t\geq 0}$ is a Feller semigroup, $\lim_{t\to 0} \|S_t f - f\| = 0$ for all $f \in \mathcal{C}(E)$, i.e., S is strongly continuous. Set $\lambda := \|\frac{Gh}{h}\|$. An elementary comparison argument based on Lemma 23 shows that $P_t h \leq e^{\lambda t} h$ $(t \geq 0)$. It follows that $\|S_t f\| = \|\frac{1}{h} P_t(hf)\| \leq \|f\| \|\frac{1}{h} P_t(h)\| \leq e^{\lambda t}\|f\|$, so S is λ-contractive. Obviously, S is positive. It is easy to see that the generator $\tilde{G}$ of S is given by

$$\tilde{G} f := \frac{1}{h} G(hf) \quad \text{with} \quad \mathcal{D}(\tilde{G}) := \{f \in \mathcal{C}(E) : hf \in \mathcal{D}(G)\}. \tag{2.50}$$

Since S is a strongly continuous, positive, λ-contractive semigroup, its generator $\tilde{G}$ satisfies the properties (2.43.i–iv). Let G^h be the operator defined in (1.16), i.e., $G^h = \tilde{G} - \frac{Gh}{h}$ with $\mathcal{D}(G^h) = \mathcal{D}(\tilde{G})$. Using Lemma 22 we see that G^h is the generator of a 2λ-contractive semigroup. Indeed, if $f \in \mathcal{D}(G^h)$ assumes a nonnegative maximum in x, then $G^h f(x) = (\tilde{G} - \frac{Gh}{h}) f(x) \leq 2\lambda f(x)$, and therefore G^h satisfies property (2.43.iii). Moreover, G^h satisfies property (2.43.iv) by Lemma 24. It is easy to check that $u_t := 1$ $(t \geq 0)$ solves $\frac{\partial}{\partial t} u_t = G^h u_t$ and therefore G^h generates a Feller semigroup.

The fact that the laws of the Feller processes ξ, ξ^h with generators G, G^h are related by (1.17) will be proved in the next section.

2.3.4 Weighted superprocesses

Proof of Lemma 3 Denote the log-Laplace semigroups of $\mathcal{X}$ and $\mathcal{X}^h$ by $\mathcal{U} = \mathcal{U}(G,\alpha,\beta)$ and $\mathcal{U}^h = \mathcal{U}(G^h, h\alpha, \beta + \frac{Gh}{h})$, respectively. By Lemma 24, for every $f \in \mathcal{D}(G^h) \cap \mathcal{C}_+(E)$, the function $u_t = \mathcal{U}_t(hf)$ is a classical solution to the Cauchy problem

$$\begin{cases} \frac{\partial}{\partial t} u_t = G u_t + \beta u_t - \alpha u_t^2 & (t \geq 0), \\ u_0 = hf. \end{cases} \tag{2.51}$$

A little calculation shows that $u^h_t := \frac{1}{h}u_t$ is a classical solution to the Cauchy problem

$$\left\{\begin{array}{l} \frac{\partial}{\partial t}u^h_t = G^h u^h_t + (\beta + \frac{Gh}{h})u^h_t - h\alpha(u^h_t)^2 \qquad (t \geq 0), \\ u^h_0 = f, \end{array}\right. \tag{2.52}$$

and therefore $\mathcal{U}^h_t f = \frac{1}{h}\mathcal{U}_t(hf)$ for all $f \in \mathcal{D}(G^h) \cap \mathcal{C}_+(E)$. Since $\mathcal{D}(G^h)$ is dense in $\mathcal{C}(E)$ and $\mathcal{C}_+(E)$ is the closure of its interior, it follows that

$$\mathcal{U}^h_t(f) = \frac{1}{h}\mathcal{U}_t(hf) \qquad (t \geq 0,\ f \in \mathcal{C}_+(E)). \tag{2.53}$$

It is clear that the process $\hat{\mathcal{X}}^h$ defined in (1.18) is a Markov process with continuous sample paths. To see that $\hat{\mathcal{X}}^h$ is the historical $(G^h, h\alpha, \beta + \frac{Gh}{h})$-superprocess, by Lemma 19, it suffices to check that $\hat{\mathcal{X}}^h$ satisfies (2.29) for the log-Laplace semigroup $\mathcal{U}^h$. This is easily done, since we have

$$\begin{aligned} & E\Big[e^{-\int_{\mathcal{D}_E[0,t_{n+1}]}\hat{\mathcal{X}}^h_{t_{n+1}}(\mathrm{d}w)f(w_{t_0},\ldots,w_{t_{n+1}})}\,\Big|\,\hat{\mathcal{X}}^h_{t_n} = \mu\Big] \\ &= E\Big[e^{-\int_{\mathcal{D}_E[0,t_{n+1}]}h(w_{t_{n+1}})\hat{\mathcal{X}}_{t_{n+1}}(\mathrm{d}w)f(w_{t_0},\ldots,w_{t_{n+1}})}\,\Big|\,(h\circ\pi_{t_n})\hat{\mathcal{X}}_{t_n} = \mu\Big] \\ &= E\Big[e^{-\int_{\mathcal{D}_E[0,t_{n+1}]}\hat{\mathcal{X}}_{t_{n+1}}(\mathrm{d}w)h(w_{t_{n+1}})f(w_{t_0},\ldots,w_{t_{n+1}})}\,\Big|\,\hat{\mathcal{X}}_{t_n} = (h\circ\pi_{t_n})^{-1}\mu\Big] \\ &= e^{-\int_{\mathcal{D}_E[0,t_n]}h(w_{t_n})^{-1}\mu(\mathrm{d}w)\mathcal{U}_{t_{n+1}-t_n}\{h(\cdot)f(w_{t_0},\ldots,w_{t_n},\cdot)\}(w_{t_n})} \\ &= e^{-\int_{\mathcal{D}_E[0,t_n]}\mu(\mathrm{d}w)\mathcal{U}^h_{t_{n+1}-t_n}f(w_{t_0},\ldots,w_{t_n},\cdot)(w_{t_n})}. \end{aligned} \tag{2.54}$$

∎

Proof of Lemma 2 (continued) We need to prove formula (1.17). Let $\hat{\mathcal{X}}$ be the (deterministic) historical $(G, 0, 0)$-superprocess started in $\hat{\mathcal{X}}_0 = \delta_x$ and set

$$\hat{\mathcal{X}}^h_t(\mathrm{d}w) := h(w_t)\hat{\mathcal{X}}_t(\mathrm{d}w) \qquad (t \geq 0). \tag{2.55}$$

By Lemma 3, $\hat{\mathcal{X}}^h$ is the historical $(G^h, 0, \frac{Gh}{h})$-superprocess started in $\hat{\mathcal{X}}_0 = h(x)\delta_x$ and therefore, by Lemma 20,

$$\begin{array}{ll} \text{(i)} & \hat{\mathcal{X}}_t(\mathrm{d}w) = P^x[(\xi_s)_{s\in[0,t]} \in \mathrm{d}w], \\ \text{(ii)} & \hat{\mathcal{X}}^h_t(\mathrm{d}w) = h(x)\, e^{\int_0^t \frac{Gh}{h}(w_s)\mathrm{d}s}\ P^x[(\xi^h_s)_{s\in[0,t]} \in \mathrm{d}w]. \end{array} \tag{2.56}$$

Combining (2.55) and (2.56) we arrive at (1.17). ∎

2.4 Proofs of the main results

2.4.1 The infinitesimal survival probability

In this section we prove Proposition 5.

Lemma 26 (Eventual extinction) *One has $\mathcal{U}_t\infty \downarrow p$ as $t \uparrow \infty$. If $\sup_{x\in E}\mathcal{U}_t\infty(x) < \infty$ for some $t > 0$, then*

$$P^\mu\big[\mathcal{X}_t = 0 \quad t\text{-eventually}\big] = e^{-\langle \mu, p\rangle} \qquad (\mu \in \mathcal{M}(E)) \tag{2.57}$$

and $\mathcal{U}_t p = p$ for all $t \geq 0$.

Proof Since the zero measure is an absorbing state, $\{\mathcal{X}_t = 0\} = \{\mathcal{X}_r = 0 \ \forall r \geq t\}$ a.s. and therefore $\{\mathcal{X}_{t_n} = 0\} \uparrow \{\mathcal{X}_t = 0 \ t\text{-eventually}\}$ a.s. as $t_n \uparrow \infty$. Thus, taking the limit in (1.22), we see that $\mathcal{U}_t\infty \downarrow p$. If $\sup_{x\in E}\mathcal{U}_t\infty(x) < \infty$ for some $t > 0$, then $\langle\mu, \mathcal{U}_t\infty\rangle \downarrow \langle\mu, p\rangle$ for all $\mu \in \mathcal{M}(E)$ and therefore, taking the limit in $P^\mu[\mathcal{X}_t = 0] = e^{-\langle\mu,\mathcal{U}_t\infty\rangle}$, we arrive at (2.57). Formula (1.8) shows that $\mathcal{U}_t$ is continuous with respect to bounded decreasing sequences and therefore $\mathcal{U}_t p = \mathcal{U}_t(\lim_{s\uparrow\infty}\mathcal{U}_s\infty) = \lim_{s\uparrow\infty}\mathcal{U}_{t+s}\infty = p$ for all $t \geq 0$. ∎

Lemma 27 (Extinction versus explosion) *If* $\sup_{x\in E}\mathcal{U}_t\infty(x) < \infty$ *for some* $t > 0$, *then*

$$P^\mu\big[\lim_{t\to\infty}\langle\mathcal{X}_t, 1\rangle = \infty \ \text{or} \ \mathcal{X}_t = 0 \ t\text{-eventually}\big] = 1 \qquad (\mu \in \mathcal{M}(E)), \tag{2.58}$$

and

$$\lim_{t\to\infty}\mathcal{U}_t f(x) = p(x) \qquad \forall x \in E, \ f \in \mathcal{C}(E), \ f > 0. \tag{2.59}$$

Proof The proof of [FS02, Lemma 12] carries over to our situation and shows that (2.58) holds. This implies that for any $x \in E$ and $f \in \mathcal{C}(E)$ with $f > 0$,

$$\lim_{t\to\infty}\mathcal{U}_t f(x) = \lim_{t\to\infty} -\log E^{\delta_x}[e^{-\langle\mathcal{X}_t,f\rangle}] = -\log P^{\delta_x}[\mathcal{X}_t = 0 \ t\text{-eventually}] = p(x). \tag{2.60}$$

∎

Even though the underlying motion has the Feller property and α, β are continuous functions, p need not be continuous in general, as is illustrated by the following examples, which we give without proof.

Example 28 (Discontinuous infinitesimal survival probability) *Let* ξ *be the deterministic Feller process in* $[-1,1]$ *given by the differential equation*

$$\tfrac{\partial}{\partial t}\xi_t = 1 - (\xi_t)^2 \qquad (t \geq 0). \tag{2.61}$$

Let $\mathcal{X}$ *be the superprocess in* $[-1,1]$ *with underlying motion* ξ, *activity* $\alpha(x) := 1$, *and growth parameter* $\beta(x) := -x$. *Then*

$$-\log P^{\delta_x}\big[\mathcal{X}_t = 0 \quad t\text{-eventually}\big] = \begin{cases} 1 & \text{if } x = -1 \\ 0 & \text{if } x \in (-1,1]. \end{cases} \tag{2.62}$$

Let $\mathcal{Y}$ *be the superprocess in* $[-1,1]$ *with underlying motion* ξ, *activity* $\alpha(x) := x \vee 0$, *and growth parameter* $\beta(x) := x \vee 0$. *Then*

$$-\log P^{\delta_x}\big[\mathcal{Y}_t = 0 \quad t\text{-eventually}\big] = \begin{cases} \infty & \text{if } x = -1 \\ 1 & \text{if } x \in (-1,1]. \end{cases} \tag{2.63}$$

Nevertheless, one has the following:

Lemma 29 (Continuity of the infinitesimal survival probability) *If* $\sup_{x\in E}\mathcal{U}_t\infty(x) < \infty$ *for some* $t > 0$ *and* $\inf_{x\in E} p(x) > 0$, *then* p *is continuous.*

Proof Our strategy will be to prove that the event that $\mathcal{X}$ gets extinct depends in a continuous way on the path of $\mathcal{X}$, and therefore, by the Feller property, on the initial condition. To do this, we show that by observing $\mathcal{X}$ for a finite time, one can be almost certain whether $\mathcal{X}$ will get extinct or not.

Set

$$\underline{p} := \inf_{x\in E} p(x) \quad \text{and} \quad \overline{p} := \sup_{x\in E} p(x). \tag{2.64}$$

Note that by (2.57),

$$e^{-\langle \mu, 1\rangle \overline{p}} \leq P^{\mu}\big[\mathcal{X}_t = 0 \quad t\text{-eventually}\big] \leq e^{-\langle \mu, 1\rangle \underline{p}} \qquad (\mu \in \mathcal{M}(E)). \tag{2.65}$$

Fix $x_0 \in E$. We will show that p is continuous at x_0. Let $0 < c < C < \infty$ and $\varepsilon', \varepsilon'' > 0$ be arbitrary. Choose continuous functions f_0, f_1, f_∞ from $[0,\infty)$ into $[0,1]$, summing up to one, such that $1_{[0,c/2]} \leq f_0 \leq 1_{[0,c]}$, $1_{[c,C]} \leq f_1 \leq 1_{[c/2,2C]}$, and $1_{[2C,\infty)} \leq f_\infty \leq 1_{[C,\infty)}$. By Lemma 27, there exists a $T > 0$ such that

$$E^{\delta_{x_0}}[f_1(\langle \mathcal{X}_T, 1\rangle)] \leq \varepsilon'. \tag{2.66}$$

Let d be a metric generating the topology on E. By Lemma 25, we can choose $\delta > 0$ such that for all $x \in E$ with $d(x, x_0) \leq \delta$:

$$\Big|E^{\delta_{x_0}}[f_r(\langle \mathcal{X}_T, 1\rangle)] - E^{\delta_x}[f_r(\langle \mathcal{X}_T, 1\rangle)]\Big| \leq \varepsilon'' \qquad (d(x,x_0) \leq \delta,\ r = 0,1). \tag{2.67}$$

Write

$$\begin{aligned} P^{\delta_x}\big[\mathcal{X}_t = 0 \quad t\text{-eventually}\big] &= E^{\delta_x}\Big[\sum_{r=0,1,\infty} f_r(\langle \mathcal{X}_T, 1\rangle) 1_{\{\mathcal{X}_t = 0 \quad t\text{-eventually}\}}\Big] \\ = \sum_{r=0,1,\infty} & E^{\delta_x}\Big[f_r(\langle \mathcal{X}_T, 1\rangle) P^{\mathcal{X}_T}\big[\mathcal{X}_t = 0 \quad t\text{-eventually}\big]\Big]. \end{aligned} \tag{2.68}$$

Then it follows from (2.65) that

$$\begin{aligned} & E^{\delta_x}\big[f_0(\langle \mathcal{X}_T, 1\rangle)\big] - (1 - e^{-c\overline{p}}) \leq E^{\delta_x}\big[f_0(\langle \mathcal{X}_T, 1\rangle)\big] e^{-c\overline{p}} \\ & \quad \leq P^{\delta_x}\big[\mathcal{X}_t = 0 \quad t\text{-eventually}\big] \\ & \quad \leq E^{\delta_x}\big[f_0(\langle \mathcal{X}_T, 1\rangle)\big] + E^{\delta_x}\big[f_1(\langle \mathcal{X}_T, 1\rangle)\big] + E^{\delta_x}\big[f_\infty(\langle \mathcal{X}_T, 1\rangle)\big] e^{-C\underline{p}} \\ & \quad \leq E^{\delta_x}\big[f_0(\langle \mathcal{X}_T, 1\rangle)\big] + (\varepsilon' + \varepsilon'') + e^{-C\underline{p}} \qquad (d(x,x_0) \leq \delta). \end{aligned} \tag{2.69}$$

Therefore, for all $x \in E$ with $d(x, x_0) \leq \delta$:

$$\begin{aligned} & \Big|P^{\delta_{x_0}}\big[\mathcal{X}_t = 0 \quad t\text{-eventually}\big] - P^{\delta_x}\big[\mathcal{X}_t = 0 \quad t\text{-eventually}\big]\Big| \\ & \quad \leq \Big|P^{\delta_{x_0}}\big[\mathcal{X}_t = 0 \quad t\text{-eventually}\big] - E^{\delta_{x_0}}\big[f_0(\langle \mathcal{X}_T, 1\rangle)\big]\Big| \\ & \qquad + \Big|E^{\delta_{x_0}}\big[f_0(\langle \mathcal{X}_T, 1\rangle)\big] - E^{\delta_x}\big[f_0(\langle \mathcal{X}_T, 1\rangle)\big]\Big| \\ & \qquad + \Big|E^{\delta_x}\big[f_0(\langle \mathcal{X}_T, 1\rangle)\big] - P^{\delta_x}\big[\mathcal{X}_t = 0 \quad t\text{-eventually}\big]\Big| \\ & \quad \leq \big((1 - e^{-c\overline{p}}) + (\varepsilon' + \varepsilon'' + e^{-C\underline{p}})\big) + \varepsilon'' + \big((1 - e^{-c\overline{p}}) + (\varepsilon' + \varepsilon'' + e^{-C\underline{p}})\big). \end{aligned} \tag{2.70}$$

Since $0 < c < C < \infty$ and $\varepsilon', \varepsilon'' > 0$ are arbitrary, the right-hand side of (2.70) can be made arbitrarily small. Thus, we have shown that for each $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$\big|e^{-p(x_0)} - e^{-p(x)}\big| \leq \varepsilon \qquad \forall x \in E \text{ with } d(x, x_0) \leq \delta. \tag{2.71}$$

This shows that p is continuous at x_0. ∎

Proof of Proposition 5 Parts (a) and (b) follow from Lemmas 26 and 27. To prove part (c), note that if $f \in \mathcal{C}_+(E)$ satisfies $\mathcal{U}_t f = f$ for all $t \geq 0$, then $u_t := f$ $(t \geq 0)$ is a mild solution to (1.7), i.e.,

$$f = P_t f + \int_0^t P_s\big(\beta f - \alpha f^2\big)\,\mathrm{d}s \qquad (t \geq 0). \tag{2.72}$$

Thus,

$$\lim_{t\to 0} t^{-1}(P_t f - f) = -\lim_{t\to 0} t^{-1}\int_0^t P_s\big(\beta f - \alpha f^2\big)\,\mathrm{d}s = -\beta f + \alpha f^2, \tag{2.73}$$

which proves that $f \in \mathcal{D}(G)$ and that (1.23) holds. Conversely, if $f \in \mathcal{D}(G) \cap \mathcal{C}_+(E)$ solves (1.23), then $u_t := f$ is a classical solution to (1.7) and therefore $\mathcal{U}_t f = f$ for all $t \geq 0$.

If $\inf_{x\in E} p(x) > 0$, finally, then p is continuous by Lemma 29 and therefore p solves (1.23) by parts (b) and (c). Moreover, part (a) shows that in this case there exists only one positive fixed point of $\mathcal{U}$. ∎

Lemma 8 shows that the assumption that $\sup_{x\in E} \mathcal{U}_t\infty(x) < \infty$ for some $t > 0$ cannot be dropped from Theorems 6 and 7. However, the reader may wonder if this condition is not implied by the simpler-looking condition $\sup_{x\in E} p(x) < \infty$. To show that this is not the case, we include the following example.

Example 30 (Nonuniform convergence of $\mathcal{U}_t\infty$) *There exists a generator G of a Feller process in a compact metrizable space E, and $\alpha \in \mathcal{C}_+(E)$, such that $\mathcal{U} = \mathcal{U}(G, \alpha, 0)$ satisfies*

$$\begin{array}{rll} \text{(i)} & \mathcal{U}_t\infty(x) < \infty & \forall x \in E,\ t > 0, \\ \text{(ii)} & \mathcal{U}_t\infty \downarrow 0 & \textit{as } t \uparrow \infty, \\ \text{(iii)} & \sup_{x\in E} \mathcal{U}_t\infty(x) = \infty & \forall t \geq 0. \end{array} \tag{2.74}$$

Proof Take $E := [0,1]^2$. Define a Feller process $\xi = (\xi^x)^{x\in E}$ in E by

$$\begin{array}{lll} \xi^{(x,y)}_t := (x, ye^{-t}) & & ((x,y) \in [0,1]\times[0,1)), \\ \xi^{(x,1)}_t := \begin{cases} (x,1) & (t \leq \tau_x), \\ (x, e^{-(t-\tau_x)}) & (t > \tau_x) \end{cases} & & (x \in [0,1]), \end{array} \tag{2.75}$$

where τ_x $(x \in (0,1])$ is an exponentially distributed random variable with mean x, and $\tau_0 := 0$. It is not hard to see that ξ is a Feller process. Let G denote its generator. Choose $\alpha \in \mathcal{C}_+(E)$ such that $\alpha(0,1) = 0$ and $\alpha > 0$ elsewhere. Set

$$\underline{\alpha(x,\cdot)} := \inf\{\alpha(x,y) : y \in [0,1]\} \qquad (x \in [0,1]). \tag{2.76}$$

For fixed $x \in [0,1]$, the process ξ restricted to $\{x\}\times[0,1]$ is an autonomous Feller process, and $\underline{\alpha(x,\cdot)} > 0$ for $x > 0$. Therefore, using (1.30), one has

$$\mathcal{U}_t\infty(x,y) \leq \frac{1}{\underline{\alpha(x,\cdot)}t} \qquad (t > 0,\ (x,y) \in (0,1]\times[0,1]). \tag{2.77}$$

The superprocess $\mathcal{X}$ started in $\delta_{(0,y)}$ $(y \in [0,1])$ is concentrated on $(0, ye^{-t})$ at time t, if it survives. Therefore, applying (1.30) to the process $(\mathcal{X}_t)_{t\geq\varepsilon}$, we have for each $\varepsilon > 0$ that

$$\mathcal{U}_t\infty(0,y) \leq \frac{1}{\delta(t-\varepsilon)} \quad (t > \varepsilon), \quad \text{where} \quad \delta := \inf\{\alpha(0, e^{-t}) : t \in [\varepsilon, \infty]\}. \tag{2.78}$$

This proves (2.74.i) and (2.74.ii). Now consider the process $(1_{(0,1]\times\{1\}}\mathcal{X}_t)_{t\geq 0}$. It is not too hard to see that this is an autonomous superprocess without (i.e., with constant) underlying motion, activity $\alpha(\cdot,1)$ and growth parameter $\beta(x) := -\frac{1}{x}$. Therefore, (see (1.30)),

$$\mathcal{U}_t(\infty 1_{(0,1]\times\{1\}})(x,1) = \frac{\beta(x)}{\alpha(x,1)(1-e^{-\beta(x)t})} = \frac{x^{-1}}{\alpha(x,1)(e^{t/x}-1)} \qquad (t>0,\ x\in(0,1]). \tag{2.79}$$

We can additionally choose $\alpha(x,1) := e^{-1/x^2}$ $(x\in(0,1])$. Then

$$\lim_{x\to 0}\mathcal{U}_t(\infty 1_{(0,1]\times\{1\}})(x,1) = \infty \qquad (t>0). \tag{2.80}$$

It follows that $\sup_{x\in E}\mathcal{U}_t\infty(x) \geq \sup_{x\in E}\mathcal{U}_t(\infty 1_{(0,1]\times\{1\}})(x) = \infty$, which proves (2.74.iii). ∎

2.4.2 Surviving lines of descent

In this section we prove Lemma 8. To prepare for this, we need some facts about Poisson point measures. Let E be a Polish space. Then, for every $\mu\in\mathcal{M}(E)$, there exists a unique (in distribution) random variable $\mathrm{Pois}(\mu)$ with values in $\mathcal{N}(E)$, such that

$$E\big[(1-f)^{\mathrm{Pois}(\mu)}\big] = e^{-\langle\mu,f\rangle} \qquad (f\in B_+(E)). \tag{2.81}$$

If μ is atomless, then $\mathrm{Pois}(\mu)$ a.s. takes values in the space $\mathcal{N}^*(E) := \{\nu\in\mathcal{N}(E) : \nu(\{x\})\leq 1\ \forall x\in E\}$ of simple point measures on E. Note that $\mathcal{N}^*(E)$ is an open subset of $\mathcal{N}(E)$, and therefore a Polish space in the induced topology. We identify $\mathcal{N}^*(E)$ with the space of finite subsets of E. If $\mu\in\mathcal{M}(E)$ is atomless, then a $\mathcal{N}^*(E)$-valued random variable ν is a Poisson point measure with intensity μ if and only if (see [MKM78, Proposition 1.4.7])

$$P[\nu(A)=0] = e^{-\mu(A)} \qquad (A\in\mathcal{B}(E)). \tag{2.82}$$

It is not hard to see that the event $\{\mathrm{supp}(\mu) \text{ is finite}\}\subset\mathcal{M}(E)$ is measurable and that $\mu\mapsto\mathrm{supp}(\mu)$ is a measurable map from $\{\mathrm{supp}(\mu) \text{ is finite}\}$ into $\mathcal{N}^*(E)$. We start with a technical lemma.

Lemma 31 (Finitely supported measures) *Let E be a Polish space, let μ be an atomless measure on E, and let $\mathcal{Z}$ be an $\mathcal{M}(E)$-valued random variable such that*

$$P[\mathcal{Z}(A)=0] = e^{-\mu(A)} \qquad (A\in\mathcal{B}(E)). \tag{2.83}$$

Then

$$P[\mathrm{supp}(\mathcal{Z}) \textit{ is finite}] = \begin{cases} 1 & \textit{if} \quad \mu(E)<\infty \\ 0 & \textit{if} \quad \mu(E)=\infty. \end{cases} \tag{2.84}$$

Moreover, if $\mu(E)<\infty$, then $\mathrm{supp}(\mathcal{Z})$ is a Poisson point measure with intensity μ.

Proof Assume that $\mu(E)<\infty$. Choose a metric, compatible with the topology on E, such that the completion of E in this metric is compact. (This is possible on any separable metrizable space.) Then E is totally bounded in this metric and therefore, for each n, we can choose a finite covering of E with open balls of radius $\frac{1}{n}$. By taking differences we can find, for each

$n \geq 1$, a finite measurable partition $A^{(n)} = \{A_1^{(n)}, \ldots, A_{k_n}^{(n)}\}$ of E such that each set $A_i^{(n)}$ has radius $\leq \frac{1}{n}$. It follows that

$$E\Big[\sum_{i=1}^{k_n} 1_{\{\mathcal{Z}(A_i^{(n)})>0\}}\Big] = \sum_{i=1}^{k_n} \big(1 - e^{-\mu(A_i^{(n)})}\big) \leq \mu(E). \tag{2.85}$$

We can choose our partitions such that $A^{(n+1)}$ is a refinement of $A^{(n)}$. In this case

$$\sum_{i=1}^{k_n} 1_{\{\mathcal{Z}(A_i^{(n)})>0\}} \uparrow |\text{supp}(\mathcal{Z})| \quad \text{as } n \uparrow \infty \quad \text{a.s.} \tag{2.86}$$

Combining this with (2.85) we see that $\text{supp}(\mathcal{Z})$ is finite a.s. Moreover, formula (2.82) shows that $\text{supp}(\mathcal{Z})$ is a Poisson point measure with intensity μ.

Assume, on the other hand, that $\mu(E) = \infty$. Since μ is atomless, there exist[2] measurable disjoint sets $(B_i)_{i\geq 0}$ such that $\mu(B_i) \geq 1$. Formula (2.83) shows that the events $\{\mathcal{Z}(B_i) > 0\}$ are independent and that

$$\sum_{i=1}^{\infty} P[\mathcal{Z}(B_i) > 0] = \sum_{i=1}^{\infty} \big(1 - e^{-\mu(B_i)}\big) = \infty. \tag{2.87}$$

Therefore, by Borel-Cantelli,

$$P[\mathcal{Z}(B_i) > 0 \text{ for infinitely many } i] = 1, \tag{2.88}$$

which proves that $P[\text{supp}(\mathcal{Z}) \text{ is finite}] = 0$. ∎

The following lemma gives a historical version of formula (1.11.i). Moreover, it shows that the particles in $\text{Pois}((\mathcal{U}_t f)\mu)$ are, in a sense, the ancestors of the particles in $\text{Pois}(f\mathcal{X}_t)$.

Lemma 32 (Poissonization of historical superprocess) *Let $\hat{\mathcal{X}}$ be the historical (G, α, β)-superprocess started at time $s \geq 0$ in $\hat{\mu} \in \mathcal{M}(\mathcal{D}_E[0, s])$. Assume that $\hat{\mu}$ is atomless. If $\hat{\nu}$ is a $\mathcal{N}(\mathcal{D}_E[0, s+t])$-valued random variable such that, for a given $f \in B_+(E)$ and $t \geq 0$,*

$$P[\hat{\nu} \in \cdot\, |(\hat{\mathcal{X}}_r)_{s\leq r\leq s+t}] = P[\text{Pois}((f \circ \pi_{s+t})\hat{\mathcal{X}}_{s+t}) \in \cdot\, |\hat{\mathcal{X}}_{s+t}] \quad \text{a.s.}, \tag{2.89}$$

then $\hat{\nu} \circ \pi_{[0,s]}^{-1}$ is a Poisson point measure with intensity $(\mathcal{U}_t f \circ \pi_s)\hat{\mu}$.

Proof Since $\hat{\mu}$ is atomless, by (2.82), it suffices to show that for all $A \in \mathcal{B}(\mathcal{D}_E[0, s])$

$$P[\hat{\nu} \circ \pi_{[0,s]}^{-1}(A) = 0] = e^{-(\mathcal{U}_t f \circ \pi_s)\hat{\mu}\,(A)}. \tag{2.90}$$

By (2.89),

$$P[\hat{\nu} \circ \pi_{[0,s]}^{-1}(A) = 0] = E^{s,\hat{\mu}}\big[e^{-(f \circ \pi_{s+t})\hat{\mathcal{X}}_{s+t} \circ \pi_{[0,s]}^{-1}(A)}\big]. \tag{2.91}$$

[2]To see this, choose partitions $A^{(n)}$ as above. For each n, there must exists an i_n such that $\mu(A_{i_n}^{(n)}) = \infty$. Since $A^{(n+1)}$ is a refinement of $A^{(n)}$, we can organize it so that $A_{i_n}^{(n)} \supset A_{i_{n+1}}^{(n+1)}$. Set $A^\infty := \bigcap_{n=1}^\infty A_{i_n}^{(n)}$, which may be the empty set or a set consisting of one point. Then $\mu(E\backslash A_{i_n}^{(n)}) \uparrow \mu(E\backslash A^\infty) = \infty$ since μ is atomless. Thus we can find an n such that $\mu(E\backslash A_{i_n}^{(n)}) \geq 1$ and $\mu(A_{i_n}^{(n)}) = \infty$. Repeating this procedure, we see that the set $A_{i_n}^{(n)}$ may be further split into a piece with mass ≥ 1 and a piece with mass infinity, and the statement follows by induction.

By the branching property (Lemma 11) and by Lemma 21 (a), we can rewrite the right-hand side of this equation as

$$\begin{aligned}&E^{s,1_A\hat{\mu}}\big[e^{-(f\circ\pi_{s+t})\hat{\mathcal{X}}_{s+t}\circ\pi^{-1}_{[0,s]}(A)}\big]E^{s,1_{A^c}\hat{\mu}}\big[e^{-(f\circ\pi_{s+t})\hat{\mathcal{X}}_{s+t}\circ\pi^{-1}_{[0,s]}(A)}\big]\\&\quad=E^{s,1_A\hat{\mu}}\big[e^{-\langle(f\circ\pi_{s+t})\hat{\mathcal{X}}_{s+t},1\rangle}\big]\cdot 1.\end{aligned}\tag{2.92}$$

From the relation between a historical superprocess and its associated superprocess (2.27) it is obvious that

$$E^{s,1_A\hat{\mu}}\big[e^{-\langle\hat{\mathcal{X}}_{s+t}\circ\pi^{-1}_{s+t},f\rangle}\big]=e^{-\langle(1_A\hat{\mu})\circ\pi_s^{-1},\mathcal{U}_tf\rangle}.\tag{2.93}$$

It follows that

$$E^{s,\hat{\mu}}\big[e^{-(f\circ\pi_{s+t})\hat{\mathcal{X}}_{s+t}\circ\pi^{-1}_{[0,s]}(A)}\big]=e^{-(\mathcal{U}_tf\circ\pi_s)\hat{\mu}(A)}.\tag{2.94}$$

Combining this with (2.91) we see that (2.90) holds. ∎

The proof of Lemma 32 has the following corollary.

Corollary 33 (Surviving lines of descent) *Let $\hat{\mathcal{X}}$ be the historical (G,α,β)-superprocess started at time $s\geq 0$ in $\hat{\mu}\in\mathcal{M}(\mathcal{D}_E[0,s])$. Assume that $\hat{\mu}$ is atomless. Then, for any $t>0$,*

$$P[\mathrm{supp}(\hat{\mathcal{X}}_{s+t}\circ\pi^{-1}_{[0,s]})\text{ is finite}]=1\ \Leftrightarrow\ \langle\hat{\mu}\circ\pi_s^{-1},\mathcal{U}_t\infty\rangle<\infty.\tag{2.95}$$

Moreover, if $\langle\hat{\mu}\circ\pi_s^{-1},\mathcal{U}_t\infty\rangle<\infty$ then $\mathrm{supp}(\hat{\mathcal{X}}_{s+t}\circ\pi^{-1}_{[0,s]})$ is a Poisson point measure with intensity $(\mathcal{U}_t\infty\circ\pi_s)\hat{\mu}$.

Proof Letting $f\uparrow\infty$ in (2.94) we see that

$$P^{s,\hat{\mu}}[\hat{\mathcal{X}}_{s+t}\circ\pi^{-1}_{[0,s]}(A)=0]=e^{-(\mathcal{U}_t\infty\circ\pi_s)\hat{\mu}(A)}\qquad(A\in\mathcal{B}(\mathcal{D}_E[0,s])).\tag{2.96}$$

Now the statements follow from Lemma 31. ∎

Proof of Lemma 8 If $\sup_{x\in E}\mathcal{U}_t\infty(x)<\infty$ for some $t>0$, then $\langle\mu,\mathcal{U}_t\infty\rangle<\infty$ for all $\mu\in\mathcal{M}(E)$. On the other hand, if $\sup_{x\in E}\mathcal{U}_t\infty(x)=\infty$ for all $t\geq 0$, then we can find $\mu\in\mathcal{M}(E)$ such that $\langle\mu,\mathcal{U}_t\infty\rangle=\infty$ for all $t\geq 0$. To see this, choose strictly positive $(\varepsilon_n)_{n\geq 0}$ such that $\sum_{n\geq 0}\varepsilon_n=1$. Choose $t_n\uparrow\infty$ and $x_n\in E$ such that $\mathcal{U}_{t_n}\infty(x_n)\geq\varepsilon_n^{-1}$ and choose $\mu:=\sum_{n\geq 0}\varepsilon_n\delta_{x_n}$. Then $\langle\mu,\mathcal{U}_{t_n}\infty\rangle\geq\sum_{m\geq n}\varepsilon_m\mathcal{U}_{t_n}(x_m)\geq\sum_{m\geq n}\varepsilon_m\mathcal{U}_{t_m}(x_m)=\infty$.

The log-Laplace semigroup $\mathcal{U}'=\mathcal{U}(G',\alpha',\beta')$ satisfies $\mathcal{U}'_t(f\circ\psi)=(\mathcal{U}_tf)\circ\psi$ where ψ denotes the projection from E' to E (see Lemma 15). Therefore (i) implies that $\langle\mu\otimes\ell,\mathcal{U}'_t\infty\rangle<\infty$ for some $t>0$, which by Corollary 33 implies (ii). On the other hand, if (i) does not hold, then there exists a $\mu\in\mathcal{M}(E)$ such that $\langle\mu\otimes\ell,\mathcal{U}'_t\infty\rangle=\infty$ for all $t\geq 0$, and in this case Corollary 33 shows that (ii) does not hold. Finally, since $\hat{\mathcal{X}}_t=\hat{\mathcal{X}}'_t\circ\psi_t^{-1}$, (ii) implies (iii). ∎

Proof of Lemma 10 We prove the following, slightly more general result.

Lemma 34 (Immortal lines of descent) *Let $\hat{\mathcal{X}}$ be the historical (G,α,β)-superprocess started at time 0 in $\mu\in\mathcal{M}(E)$. Assume that $q\geq 0$ and that $\sup_{x\in E}\mathcal{U}_t\infty(x)<\infty$ for all $t>q$, where $\mathcal{U}=\mathcal{U}(G,\alpha,\beta)$. Then*

$$\begin{array}{lll}\text{(i)} & \mathrm{supp}(\hat{\mathcal{X}}_r\circ\pi^{-1}_{[0,t]})\textit{ is finite} & \forall t,r\geq 0\textit{ such that }t+q<r\quad\text{a.s.}\\ \text{(ii)} & \mathrm{supp}(\hat{\mathcal{X}}_r\circ\pi^{-1}_{[0,t]})\supset\mathrm{supp}(\hat{\mathcal{X}}_{r'}\circ\pi^{-1}_{[0,t]}) & \forall t,r,r'\geq 0\textit{ such that }t+q<r\leq r'\quad\text{a.s.}\end{array}\tag{2.97}$$

Proof Let us introduce the shorthand

$$\hat{\mathcal{X}}_{t,r} := \hat{\mathcal{X}}_r \circ \pi_{[0,t]}^{-1} \qquad (0 \leq t \leq r). \tag{2.98}$$

Let $D \subset [0,\infty)$ be countable and dense. The implication $\Leftarrow$ in (2.95) also holds if $\hat{\mu}$ is not atomless; this can be proved by extending the space E as in Lemma 8. Therefore,

$$\mathrm{supp}(\hat{\mathcal{X}}_{t,r}) \text{ is finite} \quad \forall t,r \in D,\ t+q<r \quad \text{a.s.} \tag{2.99}$$

Let $\mathcal{O}$ be a countable basis for the topology on $\mathcal{D}_E[0,t]$. Conditioning on $\hat{\mathcal{X}}_t$ and applying Lemma 21 (c), we see that

$$1_{\{\hat{\mathcal{X}}_{t,r'}(O)>0\}} \leq 1_{\{\hat{\mathcal{X}}_{t,r}(O)>0\}} \quad \forall r,r' \geq 0,\ t \in D,\ O \in \mathcal{O},\ t \leq r \leq r' \quad \text{a.s.} \tag{2.100}$$

It follows that

$$\mathrm{supp}(\hat{\mathcal{X}}_{t,r'}) \subset \mathrm{supp}(\hat{\mathcal{X}}_{t,r}) \quad \forall r,r' \geq 0,\ t \in D,\ t \leq r \leq r' \quad \text{a.s.} \tag{2.101}$$

This implies that

$$\mathrm{supp}(\hat{\mathcal{X}}_{t,r'}) \subset \mathrm{supp}(\hat{\mathcal{X}}_{t,r}) \quad \forall r' \geq 0,\ t,r \in D,\ t+q<r \leq r' \quad \text{a.s.}, \tag{2.102}$$

where the right-hand side is finite by (2.99). Therefore (2.99) can be sharpened to

$$\mathrm{supp}(\hat{\mathcal{X}}_{t,r'}) \text{ is finite} \quad \forall r' \geq 0,\ t \in D,\ \ t+q<r' \quad \text{a.s.} \tag{2.103}$$

Since

$$\mathrm{supp}(\hat{\mathcal{X}}_{t',r'}) = \pi_{[0,t']}(\mathrm{supp}(\hat{\mathcal{X}}_{t,r'})) \quad \forall t',r' \geq 0,\ t \in D,\ \ t'+q \leq t+q<r' \quad \text{a.s.}, \tag{2.104}$$

formula (2.103) can be further sharpened to

$$\mathrm{supp}(\hat{\mathcal{X}}_{t',r'}) \text{ is finite} \quad \forall t',r' \geq 0,\ t'+q<r' \quad \text{a.s.} \tag{2.105}$$

This proves (2.97.i). Moreover, by (2.101) and (2.105)

$$\begin{aligned} \mathrm{supp}(\hat{\mathcal{X}}_{t',r'}) = \pi_{[0,t']}(\mathrm{supp}(\hat{\mathcal{X}}_{t,r'})) \subset \pi_{[0,t']}(\mathrm{supp}(\hat{\mathcal{X}}_{t,r})) = \mathrm{supp}(\hat{\mathcal{X}}_{t',r}) \\ \forall t',r,r' \geq 0,\ t \in D,\ t'+q \leq t+q<r \leq r' \quad \text{a.s.} \end{aligned} \tag{2.106}$$

Therefore (2.102) can be sharpened to

$$\mathrm{supp}(\hat{\mathcal{X}}_{t',r'}) \subset \mathrm{supp}(\hat{\mathcal{X}}_{t',r}) \quad \forall t',r,r' \geq 0,\ t'+q<r \leq r' \quad \text{a.s.}, \tag{2.107}$$

which proves (2.97.ii). ∎

2.4.3 Proof of the main theorems

Our first and crucial proposition in this section shows that it is possible to embed a collection I of 'immortal' lines of descent in certain historical superprocesses. We then identify these immortal lines of descent as a historical binary branching particle system. Finally, we generalize our results in a number of steps, until we arrive at the statements in Section 1.5.

Recall the definition of the distinct path property before Lemma 17.

Proposition 35 (Embedded tree) *Let $\hat{\mathcal{X}}$ be the historical (G,α,α)-superprocess started at time 0 in $\mu \in \mathcal{M}(E)$. Assume that μ is atomless and that the Feller process with generator G has the distinct path property. Then $\hat{\mathcal{X}}$ may be coupled to a random set $I \subset \mathcal{D}_E[0,\infty)$, such that the random sets $I_t := \{\pi_{[0,t]}(w) : w \in I\}$ are finite for all $t \geq 0$ and satisfy*

$$P[I_t \in \cdot\,|(\hat{\mathcal{X}}_s)_{0\leq s\leq t}] = P[\mathrm{Pois}(\hat{\mathcal{X}}_t) \in \cdot\,|\hat{\mathcal{X}}_t] \quad \text{a.s.} \quad \forall t \geq 0. \tag{2.108}$$

If in addition, $\mathcal{U} = \mathcal{U}(G,\alpha,\alpha)$ satisfies $\sup_{x\in E} \mathcal{U}_t\infty(x) < \infty$ for some $t > 0$, then $p := \lim_{t\uparrow\infty} \mathcal{U}_t\infty = 1$ and I may be chosen such that moreover

$$I_t = \mathrm{supp}(\hat{\mathcal{X}}_r \circ \pi^{-1}_{[0,t]}) \quad r\text{-eventually} \quad \forall t \geq 0 \quad \text{a.s.} \tag{2.109}$$

Proof Identify, as usual, finite subsets and simple point measures. For each $T \geq 0$, let $I^{(T)}$ be a random finite subset of $\mathcal{D}_E[0,T]$ such that

$$P[I^{(T)} \in \cdot\,|(\hat{\mathcal{X}}_t)_{0\leq t\leq T}] = P[\mathrm{Pois}(\hat{\mathcal{X}}_T) \in \cdot\,|\hat{\mathcal{X}}_T]. \tag{2.110}$$

Put

$$I^{(T)}_t := \{\pi_{[0,t]}(w) : w \in I^{(T)}\} = \mathrm{supp}(I^{(T)} \circ \pi^{-1}_{[0,t]}) \qquad (0 \leq t \leq T). \tag{2.111}$$

Using the fact that $\hat{\mathcal{X}}_t$ is a.s. atomless by Lemma 17, conditioning on $(\hat{\mathcal{X}}_s)_{0\leq s\leq t}$, applying Lemma 32 and the fact that the function 1 is a fixed point of $\mathcal{U}(G,\alpha,\alpha)$, we find that

$$P[I^{(T)}_t \in \cdot\,|(\hat{\mathcal{X}}_s)_{0\leq s\leq t}] = P[\mathrm{Pois}(\hat{\mathcal{X}}_t) \in \cdot\,|\hat{\mathcal{X}}_t] \quad \text{a.s.} \quad \forall 0 \leq t \leq T. \tag{2.112}$$

Thus, we can satisfy (2.108) up to a finite time horizon T. To let $T \uparrow \infty$, we need to take a projective limit. For $0 \leq S \leq T$, define a map $\psi_{S,T} : \mathcal{N}^*(\mathcal{D}_E[0,T]) \to \mathcal{N}^*(\mathcal{D}_E[0,S])$ by

$$\psi_{S,T}(J) := \{\pi_{[0,S]}(w) : w \in J\} \qquad (J \in \mathcal{N}^*(\mathcal{D}_E[0,T])). \tag{2.113}$$

Then (2.112) shows that the random variables $((\hat{\mathcal{X}}_t)_{0\leq t\leq T}, I^{(T)})_{T\geq 0}$ satisfy the consistency relation $\mathcal{L}((\hat{\mathcal{X}}_t)_{0\leq t\leq S}, \psi_{S,T}(I^{(T)})) = \mathcal{L}((\hat{\mathcal{X}}_t)_{0\leq t\leq S}, I^{(S)})$ $(0 \leq S \leq T)$. Note that $((\hat{\mathcal{X}}_t)_{0\leq t\leq T}, I^{(T)})$ takes values in the Polish space $\mathcal{C}_{\mathcal{D}_E[0,\infty)}[0,T] \times \mathcal{N}^*(\mathcal{D}_E[0,T])$. Let $\mathcal{N}^{(\infty)}$ be the space of all countable subsets $I \subset \mathcal{D}_E[0,\infty)$ such that $\psi_{T,\infty}(I) := \{\pi_{[0,T]}(w) : w \in I\}$ is finite for all $T \geq 0$. Equip $\mathcal{N}^{(\infty)}$ with the σ-field generated by the mappings $\psi_{T,\infty} : \mathcal{N}^{(\infty)} \to \mathcal{N}^*(\mathcal{D}_E[0,T])$ $(T \geq 0)$. Taking the projective limit of the variables $((\hat{\mathcal{X}}_t)_{0\leq t\leq T}, I^{(T)})_{T\geq 0}$, we can construct [3] a random variable $(\tilde{\mathcal{X}}, I)$ with values in $\mathcal{C}_{\mathcal{D}_E[0,\infty)}[0,\infty) \times \mathcal{N}^{(\infty)}$ such that $((\tilde{\mathcal{X}}_t)_{0\leq t\leq T}, \psi_{T,\infty}(I))$

[3] Let $(E_n)_{n\geq 1}$ be Polish spaces and let $\psi_{n,m} : E_n \to E_m$ be measurable maps satisfying $\psi_{m,k} \circ \psi_{n,m} = \psi_{n,k}$ $(k \leq m \leq n)$. Let E_∞ be an arbitrary set and let $\psi_{\infty,n} : E_\infty \to E_n$ $(n \geq 1)$ be maps satisfying $\psi_{\infty,n} \circ \psi_{n,m} = \psi_{\infty,m}$ $(1 \leq m \leq n)$. Assume that for each sequence $(x_n)_{n\geq 1}$ of points in $(E_n)_{n\geq 1}$ satisfying $\psi_{n,m}(x_n) = x_m$ $(1 \leq m \leq n)$, there exists a unique point $x \in E_\infty$ such that $x_n = \psi_{\infty,n}(x)$ $(n \geq 1)$. Equip E_∞ with the σ-field generated by the maps $\{\psi_{\infty,n} : n \geq 1\}$. Let $(\mu_n)_{n\geq 1}$ be probability measures on the spaces $(E_n)_{n\geq 1}$

is equal in distribution to $((\hat{\mathcal{X}}_t)_{0\le t\le T}, I^{(T)})$ for all $T \ge 0$. It follows that $\tilde{\mathcal{X}}$ is the historical (G,α,α)-superprocess started at time 0 in $\mu \in \mathcal{M}(E)$ and I is a random set satisfying (2.108).

Assume that $\sup_{x\in E}\mathcal{U}_t\infty(x) < \infty$ for some $t > 0$. We must show that we can choose I such that moreover (2.109) holds. First note that the function 1 is a positive solution to (1.23) and therefore, by Proposition 5 (a), $p = 1$. Choose $q \ge 0$ such that $\sup_{x\in E}\mathcal{U}_t\infty(x) < \infty$ for all $t > q$. Then, by Lemma 34, the random sets $\mathrm{supp}(\hat{\mathcal{X}}_r \circ \pi^{-1}_{[0,t]})$ are finite and nonincreasing in $r > t + q$ for all $t \ge 0$ a.s. Define random finite subsets $I_t \subset \mathcal{D}_E[0, t]$ by

$$I_t := \bigcap_{r>t+q} \mathrm{supp}(\hat{\mathcal{X}}_r \circ \pi^{-1}_{[0,t]}) \quad \forall t \ge 0 \quad \text{a.s.} \tag{2.114}$$

Then (2.109) is fulfilled. Define $I \subset \mathcal{D}_E[0, \infty)$ by

$$I := \{w \in \mathcal{D}_E[0,\infty) : \pi_{[0,t]}(w) \in I_t \ \forall t \ge 0\}. \tag{2.115}$$

Then

$$I_t = \{\pi_{[0,t]}(w) : w \in I\} \quad \forall t \ge 0 \quad \text{a.s.} \tag{2.116}$$

By Corollary 33,

$$P\big[\mathrm{supp}(\hat{\mathcal{X}}_r\circ\pi^{-1}_{[0,t]}) \in \cdot \,\big|(\hat{\mathcal{X}}_s)_{0\le s\le t}\big] = P\big[\mathrm{Pois}\big((\mathcal{U}_{r-t}\infty\circ\pi_t)\hat{\mathcal{X}}_t\big) \in \cdot\,\big|\hat{\mathcal{X}}_t\big] \ \text{ a.s.} \quad \forall t, r \ge 0,\ t+q < r. \tag{2.117}$$

Taking the limit $r \uparrow \infty$ we see that also (2.108) holds. ∎

Our next step is to identify the embedded tree I in Proposition 35 as a binary splitting particle system. For $t \ge 0$, define equivalence relations $\overset{t-}{\sim}$ and $\overset{t+}{\sim}$ on I by

$$\begin{aligned} w \overset{t-}{\sim} v \quad &\text{iff} \quad \pi_{[0,t)}(w) = \pi_{[0,t)}(v), \\ w \overset{t+}{\sim} v \quad &\text{iff} \quad \pi_{[0,t+\varepsilon]}(w) = \pi_{[0,t+\varepsilon]}(v) \quad \text{for some } \varepsilon > 0, \end{aligned} \tag{2.118}$$

and let I_{t-}, I_{t+} denote the collections of $\overset{t-}{\sim}$, $\overset{t+}{\sim}$ equivalence classes in I, respectively. Define counting measures $\hat{X}_{t-}, \hat{X}_{t+}$ on $\mathcal{D}_E[0, t]$ by

$$\begin{aligned} \hat{X}_{t-} &:= \sum_{w\in I_{t-}} \delta_{\pi_{[0,t]}(w)} \qquad (t > 0), \\ \hat{X}_t &:= \sum_{w\in I_{t+}} \delta_{\pi_{[0,t]}(w)} \qquad (t \ge 0). \end{aligned} \tag{2.119}$$

It is not hard to see that $\hat{X} = (\hat{X}_t)_{t\ge0}$ has right-continuous sample paths with left limits given by $\hat{X}_{t-}$, and that

$$I_t = \hat{X}_t \quad \text{a.s.} \quad \forall t \ge 0. \tag{2.120}$$

Note that the 'a.s.' and '$\forall t \ge 0$' cannot be interchanged here, since $\hat{X}_t$ is not a simple point measure at those (random) times when $|I_t| < |I_{t+}|$, i.e., when splitting occurs.

---

such that $\mu_m = \mu_n \circ \psi^{-1}_{n,m}$ $(1 \le m \le n)$. Then there exists a unique probability measure μ_∞ on E_∞ such that $\mu_\infty \circ \psi^{-1}_{\infty,n} = \mu_n$ $(n \ge 1)$. To see this, construct the product space $\prod_{n=1}^\infty E_n$ and let π_k $(k \ge 1)$ denote the projection from $\prod_{n=1}^\infty E_n$ to $\prod_{n=1}^k E_n$. Define $\phi_k : E_k \to \prod_{n=1}^k E_n$ by $\phi_k := (\psi_{k,1}, \ldots, \psi_{k,k})$. Define measures $\tilde{\mu}_k$ on $\prod_{n=1}^k E_n$ by $\tilde{\mu}_k := \mu_k \circ \phi_k^{-1}$. The measures $(\tilde{\mu}_k)_{k\ge1}$ are consistent, and therefore, by Kolmogorov's extension theorem, there exists a unique probability measure $\tilde{\mu}_\infty$ on $\prod_{n=1}^\infty E_n$ such that $\tilde{\mu}_\infty \circ \pi_k = \tilde{\mu}_k$ $(k \ge 1)$. Set $\tilde{E}_N := \{(x_n)_{n\ge1} \in \prod_{n=1}^\infty E_n : \psi_{l,k}(x_k) = x_l \ \forall 1 \le l \le k \le N\}$ and $\tilde{E}_\infty : \bigcap_{N\ge1} \tilde{E}$. It is easy to see that $\tilde{\mu}_\infty$ is concentrated on $\tilde{E}_N$ for each $N \ge 1$ and therefore $\tilde{\mu}_\infty$ is a measure on $\tilde{E}_\infty$. There exists a natural measurable bijection from $\tilde{E}_\infty$ to E_∞ and $\tilde{\mu}_\infty$ induces a measure μ_∞ on E_∞ with the desired properties.

Lemma 36 (Identification of the embedded tree) *$\hat{X}$ is the $(G,\alpha,0)$-bin-bra-process started at time 0 in $\mathrm{Pois}(\mu)$.*

Proof By (2.120) and (2.108),

$$P[\hat{X}_t \in \cdot\,|(\hat{\mathcal{X}}_s)_{0\leq s\leq t}] = P[\mathrm{Pois}(\hat{\mathcal{X}}_t) \in \cdot\,|\hat{\mathcal{X}}_t] \quad \text{a.s.} \quad \forall t \geq 0. \tag{2.121}$$

Let $\hat{X}'$ denote the $(G,\alpha,0)$-bin-bra-process started at time 0 in $\mathrm{Pois}(\mu)$. The log-Laplace semigroup $(\hat{\mathcal{U}}_{s,t})_{0\leq s\leq t}$ of the historical (G,α,α)-superprocess $\hat{\mathcal{X}}$ and the generating semigroup $(\hat{U}_{s,t})_{0\leq s\leq t}$ of the historical $(G,\alpha,0)$-bin-bra-process $\hat{X}'$ are defined by the same Cauchy integral equation, hence

$$\hat{\mathcal{U}}_{s,t}f = \hat{U}_{s,t}f \qquad (0 \leq s \leq t,\ f \in B_{[0,1]}(\mathcal{D}_E[0,t])). \tag{2.122}$$

Therefore, we may reason exactly as in the proof of Lemma 1 to see that

$$P^{0,\mathrm{Pois}(\mu)}[\hat{X}'_t \in \cdot\,] = P^{0,\mu}[\mathrm{Pois}(\hat{\mathcal{X}}_t) \in \cdot\,] \qquad (t \geq 0,\ \mu \in \mathcal{M}(E)). \tag{2.123}$$

Combining (2.121) and (2.123) we see that

$$P[\hat{X}_t \in \cdot\,] = P[\hat{X}'_t \in \cdot\,] \qquad (t \geq 0). \tag{2.124}$$

It follows from our definition of $\hat{X}$ that

$$\hat{X}_s = \mathrm{supp}(\hat{X}_t \circ \pi^{-1}_{[0,s]}) \quad \text{a.s.} \quad \forall\, 0 \leq s \leq t. \tag{2.125}$$

By a straightforward analogue of Lemma 21 (a) for historical particle systems, $\mathrm{supp}(\hat{X}'_t \circ \pi^{-1}_{[0,s]}) \subset \mathrm{supp}(\hat{X}'_s)$ a.s. $\forall 0 \leq s \leq t$. Since the death rate of $\hat{X}'$ is zero, particles cannot get extinct, and therefore in fact $\mathrm{supp}(\hat{X}'_t \circ \pi^{-1}_{[0,s]}) = \mathrm{supp}(\hat{X}'_s)$ a.s. $\forall 0 \leq s \leq t$. Since $\hat{X}'_s$ is a.s. a simple point measure (which follows from (2.124) and the fact that $\hat{X}_s$ is a.s. a simple point measure), X' satisfies, in analogy with (2.125),

$$\hat{X}'_s = \mathrm{supp}(\hat{X}'_t \circ \pi^{-1}_{[0,s]}) \quad \text{a.s.} \quad \forall 0 \leq s \leq t. \tag{2.126}$$

It follows from (2.124), (2.125) and (2.126) that

$$P[(\hat{X}_{t_1},\ldots,\hat{X}_{t_n}) \in \cdot\,] = P[(\hat{X}'_{t_1},\ldots,\hat{X}'_{t_n}) \in \cdot\,] \qquad (0 \leq t_1 < t_2 < \cdots < t_n). \tag{2.127}$$

Since $\hat{X}$ and $\hat{X}'$ have right-continuous sample paths, $\hat{X}$ and $\hat{X}'$ are equal in distribution. ∎

Proposition 37 (Generalization to $\alpha \geq \beta$) *Let $\hat{\mathcal{X}}$ be the historical (G,α,β)-superprocess started at time 0 in $\mu \in \mathcal{M}(E)$. Assume that μ is atomless and that the Feller process with generator G has the distinct path property. Assume that $\gamma := \alpha - \beta \geq 0$. Then $\hat{\mathcal{X}}$ can be coupled to the historical (G,α,γ)-bin-bra-process $\hat{X}$ started in $\hat{X}_0 = \mathrm{Pois}(\mu)$ such that*

$$P\big[\hat{X}_t \in \cdot\,\big|(\hat{\mathcal{X}}_s)_{0\leq s\leq t}\big] = P\big[\mathrm{Pois}(\hat{\mathcal{X}}_t) \in \cdot\,\big|\hat{\mathcal{X}}_t\big] \quad \text{a.s.} \quad \forall t \geq 0. \tag{2.128}$$

If in addition $\mathcal{U} = \mathcal{U}(G,\alpha,\beta)$ satisfies $\sup_{x\in E}\mathcal{U}_t\infty(x) < \infty$ for some $t > 0$, then $p := \lim_{t\uparrow\infty}\mathcal{U}_t\infty \leq 1$ and the coupling may be chosen such that moreover

$$\mathrm{supp}(\hat{X}_t) \supset \mathrm{supp}(\hat{\mathcal{X}}_r \circ \pi^{-1}_{[0,t]}) \quad r\text{-eventually} \quad \forall t \geq 0 \quad \text{a.s.} \tag{2.129}$$

If in addition $\gamma = 0$, then $p = 1$ and the coupling may be chosen such that equality holds r-eventually in (2.129).

Proof For $\gamma = 0$ the statements follow from Proposition 35 and Lemma 36. To treat the case $\gamma \geq 0$, set $E^\dagger := E \cup \{\dagger\}$ where $\dagger$ is an isolated cemetary point not belonging to E. Define a linear operator $G^\dagger$ on $\mathcal{C}(E^\dagger)$ by

$$\begin{aligned} G^\dagger f(x) &:= Gf(x) + \gamma(x)(f(\dagger) - f(x)) \qquad (x \in E), \\ G^\dagger f(\dagger) &:= 0, \end{aligned} \tag{2.130}$$

where $\mathcal{D}(G^\dagger)$ consists of those $f \in \mathcal{C}(E^\dagger)$ such that the restriction of f to E is in $\mathcal{D}(G)$. Set, moreover,

$$\begin{aligned} \alpha^\dagger(x) &:= \alpha(x) \qquad (x \in E), \\ \alpha^\dagger(\dagger) &:= 1. \end{aligned} \tag{2.131}$$

Let $\hat{\mathcal{X}}^\dagger$ denote the historical $(G^\dagger, \alpha^\dagger, \alpha^\dagger)$-superprocess started at time 0 in $\mu \in \mathcal{M}(E)$ and let $\hat{X}^\dagger$ denote the historical $(G^\dagger, \alpha^\dagger, 0)$-bin-bra-process started at time 0 in Pois(μ). For $t \geq 0$, let $\hat{\mathcal{X}}_t$ and $\hat{X}_t$ denote the restrictions of $\hat{\mathcal{X}}^\dagger_t$ and $\hat{X}^\dagger_t$ to $\mathcal{D}_E[0,t]$, respectively. We will show that $(\hat{\mathcal{X}}_t)_{t\geq 0}$, so defined, is the historical (G, α, β)-superprocess, and that $(\hat{X}_t)_{t\geq 0}$ is the historical (G, α, γ)-bin-bra-process.

Note, first of all, that E is a closed subset of $E^\dagger$ and $\mathcal{D}_E[0,\infty)$ is a closed subset of $\mathcal{D}_{E^\dagger}[0,\infty)$, and therefore $\hat{\mathcal{X}}_t$ and $\hat{X}_t$ have continuous and cadlag sample paths, respectively.

Let $(\mathcal{F}_t)_{t\geq 0}$ denote the filtration generated by $\hat{\mathcal{X}}^\dagger$. Fix $t \geq 0$ and $F \in B(\mathcal{D}_E[0,t])$ and define $F^\dagger \in B(\mathcal{D}_{E^\dagger}[0,t])$ by $F^\dagger(w) := F(w)$ if $w \in \mathcal{D}_E[0,t]$ and $F^\dagger(w) := 0$ otherwise. Let $\hat{\mathcal{U}}^\dagger$ and $\hat{\mathcal{U}}$ denote the log-Laplace semigroups of the historical $(G^\dagger, \alpha^\dagger, \alpha^\dagger)$-superprocess and the historical (G, α, β)-superprocess, respectively, defined as in (2.24). We need to show that, for all $0 \leq s \leq t$,

$$E\big[e^{-\langle \hat{\mathcal{X}}_t, F\rangle}\big|\mathcal{F}_s\big] = E\big[e^{-\langle \hat{\mathcal{X}}^\dagger_t, F^\dagger\rangle}\big|\mathcal{F}_s\big] = e^{-\langle \hat{\mathcal{X}}^\dagger_s, \hat{\mathcal{U}}^\dagger_{s,t}F^\dagger\rangle} = e^{-\langle \hat{\mathcal{X}}_s, \hat{\mathcal{U}}_{s,t}F\rangle} \quad \text{a.s.}, \tag{2.132}$$

which shows that $\hat{\mathcal{X}}$ is a Markov process with the same transition probabilities as the historical (G, α, β)-superprocess. Thus, we need to show that

$$\hat{\mathcal{U}}^\dagger_{s,t}F^\dagger(w) = \hat{\mathcal{U}}_{s,t}F(w) \qquad (0 \leq s \leq t,\ w \in \mathcal{D}_E[0,s],\ F \in B_+(\mathcal{D}_E[0,t])). \tag{2.133}$$

It suffices to check (2.133) for functions F that depend only on the value of the path at finitely many times and therefore, by (2.32), it suffices to show that

$$\mathcal{U}^\dagger_t f^\dagger(x) = \mathcal{U}_t f(x) \qquad (t \geq 0,\ f \in B_+(E)), \tag{2.134}$$

where $f^\dagger(x) := f(x)$ for $x \in E$ and $f^\dagger(\dagger) := 0$. If additionally $f \in \mathcal{D}(G)$ then $f^\dagger \in \mathcal{D}(G^\dagger)$ and the function $u^\dagger$ defined as $u^\dagger_t := \mathcal{U}_t f^\dagger$ is a classical solution to the Cauchy problem

$$\begin{cases} \frac{\partial}{\partial t}u^\dagger_t = G^\dagger u^\dagger_t + \beta^\dagger u^\dagger_t - \alpha^\dagger (u^\dagger_t)^2 \qquad (t \geq 0), \\ u^\dagger_0 = f^\dagger. \end{cases} \tag{2.135}$$

The Feller process $\xi^\dagger$ with generator $G^\dagger$ satisfies $P^\dagger[\xi^\dagger_t = \dagger] = 1$ for all $t \geq 0$ and therefore, by Lemma 16 (a), for $s \geq 0$ fixed $\mathcal{X}^\dagger_s$ is concentrated on $\dagger$ when $\mathcal{X}^\dagger_0 = \delta_\dagger$. It follows that $u^\dagger_t(\dagger) = 0$ for all $t \geq 0$. Using this fact and (2.135), it is easy to see that the restriction u_t of $u^\dagger_t$ to E solves the Cauchy problem

$$\begin{cases} \frac{\partial}{\partial t}u_t = Gu_t + \beta u_t - \alpha u_t^2 \qquad (t \geq 0), \\ u_0 = f, \end{cases} \tag{2.136}$$

which shows that (2.134) holds for all $f \in \mathcal{D}(G) \cap \mathcal{C}_+(E)$. Using the continuity of log-Laplace semigroups with respect to bounded pointwise limits, this generalizes to $f \in B_+(E)$.

The fact that $(\hat{X}_t)_{t\geq 0}$ is the historical (G, α, γ)-bin-bra-process is proved in the same way, using the generating semigroups $U(G^\dagger, \alpha^\dagger, 0)$ and $U(G, \alpha, \gamma)$. Note that $\mathcal{U}(G^\dagger, \alpha^\dagger, \alpha^\dagger)$ and $U(G^\dagger, \alpha^\dagger, 0)$ coincide on $B_{[0,1]}(E)$ and likewise $\mathcal{U}(G, \alpha, \beta)$ and $U(G, \alpha, \gamma)$ coincide on $B_{[0,1]}(E)$, so that this comes down to the same fact about Cauchy problems that we have already checked.

Applying Proposition 35 and Lemma 36 we see that $\hat{\mathcal{X}}^\dagger$ and $\hat{X}^\dagger$ may be coupled such that

$$P[\hat{X}^\dagger_t \in \cdot\, | (\hat{\mathcal{X}}^\dagger_s)_{0\leq s\leq t}] = P[\mathrm{Pois}(\hat{\mathcal{X}}^\dagger_t) \in \cdot\, | \hat{\mathcal{X}}^\dagger_t] \quad \text{a.s.} \quad \forall t \geq 0, \tag{2.137}$$

which implies (2.128). If in addition $\sup_{x\in E} \mathcal{U}^\dagger_t \infty(x) < \infty$ for some $t > 0$, then $p^\dagger := \lim_{t\uparrow\infty} \mathcal{U}^\dagger_t \infty = 1$ and the coupling may be chosen such that moreover

$$\mathrm{supp}(\hat{X}^\dagger_t) = \mathrm{supp}(\hat{\mathcal{X}}^\dagger_r \circ \pi^{-1}_{[0,t]}) \quad r\text{-eventually} \quad \forall t \geq 0 \quad \text{a.s.} \tag{2.138}$$

Let $W := \{w \in \mathcal{D}_{E^\dagger}[0, \infty) : 1_{\{w(r)=\dagger\}} \leq 1_{\{w(r')=\dagger\}}\ \forall 0 \leq r \leq r'\}$ denote the space of paths that are trapped in $\dagger$, once they reach $\dagger$. By Lemma 16 (b) and the fact that $\dagger$ is a trap for the underlying motion, $\hat{\mathcal{X}}^\dagger$ is concentrated on paths that are trapped in $\dagger$, once they reach $\dagger$, and therefore

$$\mathrm{supp}(\hat{X}_t) = \mathrm{supp}(\hat{X}^\dagger_t) \cap \mathcal{D}_E[0, t] = \mathrm{supp}(\hat{\mathcal{X}}^\dagger_r \circ \pi^{-1}_{[0,t]}) \cap \mathcal{D}_E[0, t] \supset \mathrm{supp}(\hat{\mathcal{X}}_r \circ \pi^{-1}_{[0,t]}) \tag{2.139}$$

$\forall 0 \leq t \leq r$ a.s. Formulas (2.138) and (2.139) imply (2.129). Finally, for all $x \in E$,

$$p(x) = -\log P^{\delta_x}\big[\mathcal{X}_t = 0 \quad t\text{-eventually}\big] \leq -\log P^{\delta_x}\big[\mathcal{X}^\dagger_t = 0 \quad t\text{-eventually}\big] = p^\dagger(x) = 1. \tag{2.140}$$

■

In order to prove Theorems 4, 6, and 7, we only need to show the following:

Proposition 38 (Generalization to $h \neq 1$ and measures with atoms) *Let $\hat{\mathcal{X}}$ be the historical (G, α, β)-superprocess started at time 0 in $\mu \in \mathcal{M}(E)$. Assume that $h \in \mathcal{D}(G)$ satisfies $h > 0$ and, for some $\gamma \in \mathcal{C}_+(E)$,*

$$Gh + \beta h - \alpha h^2 = -\gamma h. \tag{2.141}$$

Then $\hat{\mathcal{X}}$ can be coupled to the historical $(G^h, h\alpha, \gamma)$-bin-bra-process $\hat{X}$ started in $\hat{X}_0 = \mathrm{Pois}(h\mu)$ such that

$$P\big[\hat{X}_t \in \cdot\, \big| (\hat{\mathcal{X}}_s)_{0\leq s\leq t}\big] = P\big[\mathrm{Pois}((h \circ \pi_t)\hat{\mathcal{X}}_t) \in \cdot\, \big| \hat{\mathcal{X}}_t\big] \quad \text{a.s.} \quad \forall t \geq 0. \tag{2.142}$$

If in addition $\mathcal{U} = \mathcal{U}(G, \alpha, \beta)$ satisfies $\sup_{x\in E} \mathcal{U}_t\infty(x) < \infty$ for some $t > 0$, then $p := \lim_{t\uparrow\infty} \mathcal{U}_t\infty \leq h$ and the coupling may be chosen such that moreover

$$\mathrm{supp}(\hat{X}_t) \supset \mathrm{supp}(\hat{\mathcal{X}}_r \circ \pi^{-1}_{[0,t]}) \quad r\text{-eventually} \quad \forall t \geq 0 \quad \text{a.s.} \tag{2.143}$$

If in addition $\gamma = 0$ then $p = h$ and and the coupling may be chosen such that equality holds r-eventually in (2.143).

Proof Assume that μ is atomless and that the Feller process with generator G has the distinct path property. Set $\hat{\mathcal{X}}^h_t(\mathrm{d}w) := h(w_t)\hat{\mathcal{X}}_t(\mathrm{d}w)$ $(t \geq 0)$. By Lemma 3, $\hat{\mathcal{X}}^h$ is the historical (G^h, α^h, β^h)-superprocess, where G^h is defined in (1.16) and $\alpha^h := h\alpha$, $\beta^h := \beta + \frac{Gh}{h}$. Formula (2.141) implies that

$$-\gamma = \beta^h - \alpha^h \leq 0. \tag{2.144}$$

Therefore the statements follow from Proposition 37. In order to drop the assumption that μ is atomless and that the Feller process with generator G has the distinct path property, extend $\hat{\mathcal{X}}$ as in Lemma 18, and note that all functions $(h, p, \ldots)$ do not depend on the extra coordinate. The statements then follow by projection. ∎

References

[Bou58] N. Bourbaki. *Éléments de Mathématique. VIII. Part. 1: Les Structures Fondamentales de l'Analyse. Livre III: Topologie Générale. Chap. 9: Utilisation des Nombres Réels en Topologie Générale. 2 iéme éd.* Actualités Scientifiques et Industrielles 1045. Hermann & Cie, Paris, 1958.

[Daw93] D.A. Dawson. Measure-valued Markov processes. In P.L. Hennequin, editor, *École d'été de probabilités de Saint Flour XXI–1991*, volume 1541 of *Lecture Notes in Mathematics*, pages 1–260. Springer, Berlin, 1993.

[DP91] D.A. Dawson and E.A. Perkins. *Historical Processes.* Mem. Amer. Math. Soc., Volume 93 No. 454, 1991.

[Doo84] J.L. Doob. *Classical Potential Theory and Its Probabilistic Counterpart.* Springer, New York, 1984.

[Eth00] A.M. Etheridge. *An Introduction to Superprocesses.* Volume 20 of *University Lecture Series*, American Mathematical Society (AMS), Providence, 2000.

[EK86] S.N. Ethier and T.G. Kurtz. *Markov Processes; Characterization and Convergence.* John Wiley & Sons, New York, 1986.

[EP99] J. Engländer and R.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(1): 684–730, 1999.

[Fit88] P.J. Fitzsimmons. Construction and regularity of measure-valued branching processes. *Isr. J. Math.*, 64(3): 337–361, 1988.

[Fit91] P.J. Fitzsimmons. Correction to "Construction and regularity of measure-valued branching processes". *Isr. J. Math.*, 73(1): 127, 1991.

[Fit92] P.J. Fitzsimmons. On the martingale problem for measure-valued Markov branching processes. pp 39–51 in: *Seminar on Stochastic Processes, 1991.* Progress in Probability 29, Birkhäuser, Boston, 1992.

[FS02] K. Fleischmann and J. Swart. Extinction versus explosion in a supercritical super-Wright-Fisher diffusion. WIAS Berlin, Preprint No. 752: 1–22, 2002.

[Get75] R.K. Getoor. *Markov processes: Ray Processes and Right Processes.* Lecture Notes in Mathematics 440. Springer, Berlin, 1975.

[GRW90] L.G. Gorostiza, S. Roelly-Coppoletta, and A. Wakolbinger. Sur la persistence du process de Dawson Watanabe stable (l'inversion de la limite en temps et de la renormalisation). *Seminaire de Probabiliés XXIV. Lecture Notes in Math.*, 1426: 275–281, Springer, Berlin, 1990.

[Kle98] A. Klenke. Clustering and invariant measures for spatial branching models with infinite variance. *Ann. Probab.* 26(3): 1057–1087, 1998.

[MKM78] K. Matthes, J. Kerstan, and J. Mecke. *Infinitely Divisible Point Processes.* Wiley, Chichester, 1978.

[Paz83] A. Pazy. *Semigroups of Linear Operators and Applications to Partial Differential Equations.* Springer, New York, 1983.

[Sha88] M. Sharpe. *General Theory of Markov Processes.* Academic Press, Boston, 1988.

[Sko56] A.V. Skorohod. Limit theorems for stochastic processes. *Theory Probab. Appl.*, 1: 261–290, 1956.

[Win02] A. Winter. Multiple scale analysis of spatial branching processes under the Palm distribution. Electr. J. Prob. 7: Nr. 12, 1–70, 2002.