

# Decompositions of CSBPs via Poissonian Intertwining

Clément Foucart\* and Olivier Hénard†

March 30, 2025

## Abstract

We revisit certain decompositions of continuous-state branching processes (CSBPs), commonly referred to as skeletal decompositions, through the lens of intertwining of semigroups. Precisely, we associate to a CSBP  $X$  with branching mechanism  $\psi$  a family of  $\mathbb{R}_+ \times \mathbb{Z}_+$ -valued branching processes  $(X^\lambda, L^\lambda)$ , indexed by a parameter  $\lambda \in (0, \infty)$ , that satisfies an intertwining relationship with  $X$  through the Poisson kernel with parameter  $\lambda x$ . The continuous component  $X^\lambda$  has the same law as  $X$ , while the discrete component  $L^\lambda$ , conditionally on  $X_t^\lambda$ , has a Poisson distribution with parameter  $\lambda X_t^\lambda$ . The law of  $(X^\lambda, L^\lambda)$  depends on the position of  $\lambda$  within  $[0, \infty) = [0, \rho) \cup [\rho, \infty)$ , where  $\rho$  is the largest positive root of  $\psi$ . When  $\lambda \geq \rho$ , various well-known results concerning skeleton decompositions are recovered. In the supercritical case ( $\rho > 0$ ), when  $\lambda < \rho$ , a novel phenomenon arises: a birth term appears in the skeleton, corresponding to a one-unit proportional immigration from the continuous to the discrete component. Along the way, the class of continuous-time branching processes taking values in  $\mathbb{R}_+ \times \mathbb{Z}_+$  is constructed.

**Keywords.** Continuous-state branching process, Intertwining, Esscher transform, Two-type branching process, Explosion

## 1 Introduction

This paper investigates how Continuous-State Branching Processes (CSBPs) can be decomposed into  $\mathbb{R}_+ \times \mathbb{Z}_+$ -valued branching processes using intertwining relationships between Markovian semigroups. An intertwining relationship is a property of commutation of semigroups, or generators, through some kernel. Formally, two semigroups  $P_t$  and  $Q_t$  are said to be intertwined with respect to a kernel  $\Lambda$  if they satisfy the following relationship

$$\Lambda P_t = Q_t \Lambda, \quad \forall t \geq 0. \quad (1)$$

Intertwining is intimately connected to the problem of determining whether a function of a Markov process remains Markovian, as explored by Pitman and Rogers [PR81]. It also arises in Markovian filtering, see e.g. Kurtz and Ocone [KO88], Kurtz [Kur98], and Kurtz and Nappo [KN11], where Markov mapping theorems are established in the setting of martingale problems. Numerous examples of intertwined Markov semigroups have been discovered, including those presented by Carmona et al. [CPY98] and Pal and Shkolnikov [PS13].

The intertwining theory has been further developed in various directions, becoming a fundamental tool in stochastic processes. They play for instance an important role in the study of strong stationary times, see for instance Diaconis and Fill [DF90], Miclo [Mic20] and Arnaudon et al. [ACPM24].

---

\*CMAP, Ecole Polytechnique. Email: clement.foucart@polytechnique.edu

†LMO, Université Paris-Saclay. Email: olivier.henard@universite-paris-saclay.fr

Intertwining likewise appears in analyzing certain interacting particle systems, see e.g. Floreani et al. [FJRW24]. It also plays a key role in the so-called *lookdown* construction of the genealogy of branching populations, see for instance Kurtz [Kur18] and Etheridge and Kurtz [EK19] for recent works in this direction.

The concept of decomposition of branching processes into “skeletons” dates back to Harris [Har48], who showed that any supercritical Bienaymé-Galton-Watson process with positive probability of extinction can be embedded in a two-type branching process, see also Athreya-Ney’s book [AN04, Chapter I-Section 12]. In this framework, one type represents individuals with infinite line of descent (called prolific or immortal individuals), while the other corresponds to mortal individuals. This decomposition has become a key tool in branching process theory. In the context of Continuous State Branching Processes (CSBPs), the prolific individuals form a discrete branching process, while the non-prolific ones form a continuum, as explored by Bertoin et al. [BFM08], see also Lambert and Uribe Bravo, [LUB24], for a recent work in this direction in the context of splitting trees. The framework with spatial motion has also garnered considerable attention, we refer for instance to Eckhoff et al. [EKW15], Fekete et al. [FFK20] and the references therein.

In a seminal paper, Duquesne and Winkel [DW07] introduced a nested family of discrete trees with edge lengths that is consistent under Poissonian sampling of the leaves; they proved that any such family embeds a Lévy real tree encoding the genealogy of a CSBP. In a similar spirit, Abraham and Delmas [AD12] have constructed a tree-valued Markov process, where evolving skeletons serve as the primary objects. See also Abraham and Delmas [AD09] and Abraham et al. [ADV10] for closely related studies. More recently, these decompositions have been revisited using Poissonization techniques applied to the stochastic differential equation with jumps solved by a CSBP, as discussed by Fekete et al. [FFK19]. In all these works, Esscher transforms of the branching mechanism play a central role. To summarize, given an eventually positive branching mechanism  $\psi$ , and denoting by  $\rho$  its largest positive root, a skeleton decomposition consists in “grafting” subcritical continuous masses — evolving with a branching mechanism given by an Esscher transform of  $\psi$  at the right of  $\rho$  — onto a discrete branching process, so that the total continuous mass is a CSBP governed by  $\psi$ .

Although not always highlighted in the works previously cited – see however [BFM08, Page 722] and [LUB24, Page 1283] – a common feature in such skeletal decompositions lies on the fact that the joint process encoding both the discrete component and the continuous mass satisfies the branching property. Two-type branching processes taking values in  $\mathbb{R}_+ \times \mathbb{Z}_+$  come therefore naturally into play. To our knowledge, no general treatment of this class of processes has been presented before. We summarize their fundamental properties in Section 2, see Theorem 2.1 and Proposition 2.3, and recall well-known facts about one-dimensional branching processes. The  $\mathbb{R}_+ \times \mathbb{Z}_+$ -valued branching processes are constructed in Section 4 by using classical results from the theory of Markov processes.

Our main results are presented in Section 3. We start by establishing an intertwining relationship, through the Poisson kernel with parameter  $\lambda x > 0$ , between the generator of a CSBP (which may be immortal, namely its branching mechanism  $\psi$  can be negative) and that of a specific  $\mathbb{R}_+ \times \mathbb{Z}_+$ -valued branching process  $(X^\lambda, L^\lambda)$ , see Theorem 3.1. We show then that a relationship of the form (1) holds between  $P_t$  the semigroup of  $(X^\lambda, L^\lambda)$  and  $Q_t$  the semigroup of the CSBP( $\psi$ ), see Theorem 3.4. The process  $(L_t^\lambda, t \geq 0)$  is the so-called *skeleton*. The decomposition of the CSBP follows from an application of Pitman-Rogers theorem, see Theorem A, and states the following: if  $X^\lambda$  starts from  $x > 0$  and the initial distribution of the skeleton  $L^\lambda$  has a Poisson law with parameter  $\lambda x$ , the first coordinate projection  $X^\lambda$  is a CSBP( $\psi$ ) started from  $x$  and for any  $t \geq 0$ , the law of  $L_t^\lambda$ , conditionally on  $X_t^\lambda$ , is Poisson with parameter  $\lambda X_t^\lambda$ .

The dynamics of  $(X^\lambda, L^\lambda)$  depends on whether  $\lambda > \rho$ ,  $\lambda = \rho$  or  $\lambda < \rho$ . When  $\lambda = \rho$ , we recover the two-type process studied in [BFM08] for which the skeleton is the prolific discrete tree. When  $\lambda > \rho$ , a death term arises along the skeleton and various results from [DW07] are reobtained. The intertwining approach also enables us to investigate the setting  $\lambda < \rho$ , the decomposition here involves the Esscher transform *at the left* of  $\rho$ , a scenario that, to our knowledge, has not been previously studied. As we shall see, there is a significant change in the dynamics: the skeleton is no longer autonomous, and the continuous mass now generates discrete-type individuals through a proportional birth term.

Next, we observe that if the CSBP explodes without being killed, it does so simultaneously with any of its skeletons, see Proposition 3.7. Last, we establish that, for any branching mechanism (including the explosive and immortal ones), the skeletons, once rescaled by  $\lambda$ , converge weakly in the Skorokhod topology toward the CSBP, as  $\lambda$  goes to  $\infty$ , see Theorem 3.12.

**Notation.** We set  $\mathbb{R}_+ := [0, \infty)$  and let  $\mathbb{N}$  and  $\mathbb{Z}_+$  be respectively the sets of positive integers and non-negative integers. Denote by  $C_0(\mathbb{R}_+)$  and  $C_0(\mathbb{R}_+ \times \mathbb{Z}_+)$  the spaces of continuous functions on  $\mathbb{R}_+$  and  $\mathbb{R}_+ \times \mathbb{Z}_+$ , respectively, that are vanishing at infinity. Denote by  $C_0^2(\mathbb{R}_+)$ , the functions that are twice-differentiable with first two derivatives in  $C_0(\mathbb{R}_+)$ . For any function  $f$  on  $\mathbb{R}_+ \times \mathbb{Z}_+$ , such that  $x \mapsto f(x, \ell) \in C_0^2(\mathbb{R}_+)$ , the first two derivatives with respect to  $x$  are denoted by  $f'(x, \ell)$  and  $f''(x, \ell)$ .

## 2 Preliminaries

### 2.1 $\mathbb{R}_+ \times \mathbb{Z}_+$ -valued continuous-time branching processes

CSBPs have been introduced by Jirina [Jir58], Lamperti [Lam67a, Lam67b] and Silverstein [Sil68]. They are the scaling limits of Bienaymé-Galton-Watson processes and represent the random evolution of a *continuous* population. Two-dimensional branching processes with continuous-state space  $\mathbb{R}_+ \times \mathbb{R}_+$  have been defined by Watanabe [Wat69].

They are specific affine processes, see Duffie et al. [DFS03] and Caballero et al. [CGB17] and are also known to be strong solutions to certain stochastic differential equations with jumps, see Barczy et al. [BLP15].

Two-type branching processes taking values in  $\mathbb{R}_+ \times \mathbb{Z}_+$  do not form a subclass of the branching processes with values in  $\mathbb{R}_+^2$  studied in [Wat69], for the same reason that branching processes taking values in  $\mathbb{Z}_+$  do not form a subclass of the branching processes taking values in  $\mathbb{R}_+$  (the CSBPs): the former may have jumps of size  $-1$ , and not the latter.

The case where one component evolves in  $\mathbb{Z}_+$  is thus structurally distinct and, to our knowledge, not reducible to any case studied in the literature.

While their general form will certainly not come as a surprise to the reader, they are central to this article and might prove useful in other contexts. We therefore begin by presenting them in detail.

Let  $\gamma \in \mathbb{R}$  and  $b, \sigma, d, \kappa, k \in \mathbb{R}_+$ . Let  $\pi(dy, dk)$  and  $\rho(dy, dk)$  be two measures on  $\mathbb{R}_+ \times \mathbb{Z}_+$  such that

$$\int_0^\infty (1 \wedge y^2) \pi(dy, \{0\}) + \sum_{k \geq 1} \int_0^\infty \pi(dy, dk) < \infty, \quad \int_0^\infty (1 \wedge y) \rho(dy, \{0\}) + \sum_{k \geq 1} \int_0^\infty \rho(dy, dk) < \infty. \quad (2)$$

We call *admissible* such parameters. Define the operator

$$\begin{aligned} \mathcal{L}f(x, \ell) &:= \gamma x f'(x, \ell) + b f'(x, \ell) + \frac{\sigma^2}{2} x f''(x, \ell) - \kappa x f(x, \ell) - k f(x, \ell) \\ &+ x \sum_{k \geq 0} \int_{\mathbb{R}_+} (f(x + y, \ell + k) - f(x, \ell) - y \mathbf{1}_{(0,1)}(y) f'(x, \ell)) \pi(dy, dk) \\ &+ \ell \sum_{k \geq 0} \int_{\mathbb{R}_+} (f(x + y, \ell + k) - f(x, \ell)) \rho(dy, dk) + d \ell (f(x, \ell - 1) - f(x, \ell)), \end{aligned} \quad (3)$$

with  $(x, \ell) \in \mathbb{R}_+ \times \mathbb{Z}_+$  and  $f : (x, \ell) \mapsto f(x, \ell)$  a function in the space  $\mathcal{D}$ :

$$\begin{aligned} \mathcal{D} := \left\{ f : (x, \ell) \in \mathbb{R}_+ \times \mathbb{Z}_+ \mapsto f(x, \ell) \in \mathbb{R}, \text{ such that:} \right. \\ \left. \forall \ell \in \mathbb{Z}_+, x \mapsto f(x, \ell) \in C_0^2(\mathbb{R}_+), \quad \lim_{x \rightarrow \infty} x (|f(x, \ell)| + |f'(x, \ell)| + |f''(x, \ell)|) = 0, \right. \\ \left. \text{and } \forall x \in \mathbb{R}_+, \lim_{\ell \rightarrow \infty} \ell (|f(x, \ell)| + |f'(x, \ell)| + |f''(x, \ell)|) = 0 \right\}. \quad (4) \end{aligned}$$

Consider the one-point compactification of  $E := \mathbb{R}_+ \times \mathbb{Z}_+$ ,

$$E^\Delta := (\mathbb{R}_+ \times \mathbb{Z}_+) \cup \{\Delta\}, \text{ with } \Delta := \{(x, \ell) : x + \ell = \infty\} \text{ and set } \mathcal{L}f(\Delta) := 0.$$

**Theorem 2.1.** *For any admissible parameters  $(\gamma, b, \sigma, \kappa, k, \pi, \rho)$ , there exists a unique  $E^\Delta$ -valued càdlàg strong Markov process  $\mathbf{X} = (X, L)$ , with cemetery state  $\Delta$ , solution to the martingale problem*

$$\text{MP}_{\mathbf{X}}(\mathcal{L}, \mathcal{D}) : \quad \forall f \in \mathcal{D}, \left( f(\mathbf{X}_t) - \int_0^t \mathcal{L}f(\mathbf{X}_s) ds, t \geq 0 \right) \text{ is a martingale.} \quad (5)$$

Moreover, the semigroup of  $\mathbf{X}$  satisfies for any  $(x, n)$ ,  $t \geq 0$  and  $(q, r) \in (0, \infty) \times (0, 1)$

$$\mathbb{E}_{(x, n)}[e^{-qX_t} r^{L_t}] = e^{-xu_t(q, r)} f_t(q, r)^n, \quad (6)$$

with  $t \mapsto F_t(q, r) := (u_t(q, r), f_t(q, r))$ , the unique solution to the two-dimensional o.d.e

$$\frac{d}{dt} F_t(q, r) = -\Psi(F_t(q, r)), \quad (7)$$

$$F_0(q, r) = (u_0(q, r), f_0(q, r)) = (q, r), \quad (8)$$

where  $\Psi = (\Psi_c, -\Psi_d)$ , with for any  $q \in (0, \infty)$  and  $r \in [0, 1)$ ,

$$\Psi_c(q, r) := \sum_{k \geq 0} \int_{\mathbb{R}_+} (e^{-qy} r^k - 1 + qy \mathbf{1}_{(0,1)}(y)) \pi(dy, dk) - \gamma q + \frac{\sigma^2}{2} q^2 - \kappa, \quad (9)$$

$$\Psi_d(q, r) := \sum_{k \geq 0} \int_{\mathbb{R}_+} (e^{-qy} r^{k+1} - r) \rho(dy, dk) - bqr + d(1 - r) - kr. \quad (10)$$

*Remark 2.2.* The killing terms in (3) with parameters  $\kappa$  and  $k$  can be interpreted as single jumps to the boundary  $\infty$ . Indeed for any  $f \in \mathcal{D}$ ,

$$-\kappa x f(x, \ell) = \kappa x (f(\infty, \ell) - f(x, \ell)) \text{ and } -k\ell f(x, \ell) = k\ell (f(x, \infty) - f(x, \ell)).$$

**Proposition 2.3.** *The semigroup  $(P_t)$  of the process  $\mathbf{X}$  satisfying (6) is Feller with absorbing state  $\Delta$ . For all  $(q, r) \in (0, \infty) \times (0, 1)$ , set  $f_{q,r} : (x, \ell) \mapsto e^{-qx} r^\ell$ . The space*

$$D := \text{Vect} \{f_{q,r}, q \in (0, \infty), r \in (0, 1)\}, \quad (11)$$

satisfies  $P_t D \subset D$  and is a core for the generator  $\mathcal{L}$  of  $\mathbf{X}$ . Moreover,  $\mathbf{X}$  possesses the branching property. Specifically, its transition kernel  $P_t((x, \ell), \cdot)$  satisfies:

$$\forall x, y \in [0, \infty), \forall n, m \in \mathbb{Z}_+, \quad P_t((x + y, n + m), \cdot) = P_t((x, n), \cdot) \star P_t((y, m), \cdot), \quad (12)$$

where  $\star$  stands for the convolution of measures.

Theorem 2.1 and Proposition 2.3 are established in Section 4. We call  $\Psi$  the *joint branching mechanism* of  $\mathbf{X}$ .

*Remark 2.4.* Any  $\mathbb{R}_+ \times \mathbb{Z}_+$ -valued Markov process satisfying the branching property (12) and the following condition on its semigroup

$$t \mapsto \mathbb{E}_{(x, n)}[e^{-qX_t} r^{L_t}] \text{ is differentiable at } 0,$$

will fall into the class studied here. We do not enter into this study here and refer for instance to Gihman and Skorokhod's book [GS83, Chapter V].

Theorem 2.1 covers the two classical settings, namely the discrete-state branching processes and those in continuous-state space, as we now show. We recall some well-known facts about them. More background can be found in the books by Harris [Har02], Kyprianou [Kyp14] and Li [Li22].

### 2.1.1 One-dimensional discrete-state branching processes.

Let  $\Psi_d$  as in (10), with  $b \equiv 0$  and a measure  $\rho$  shrinking to  $\rho(dy, dk) = \delta_0(dy)\rho_d(dk)$ , for some finite measure  $\rho_d(dk)$  on  $\mathbb{N}$ . In this case there is no dependence on the variable  $q$  in  $\Psi_d$  and the function

$$\varphi : [0, 1) \ni r \rightarrow \Psi_d(0, r) = \Psi_d(q, r) = \sum_{k \geq 1} (r^{k+1} - r)\mu(k) + d(1 - r) - kr,$$

is the mechanism of a continuous-time Markov branching process with reproduction measure  $\mu := d\delta_{-1} + \rho_d + k\delta_\infty$ . Furthermore in the case  $\Psi_c \equiv 0$ , the component  $X$  in the process  $(X, L)$  is degenerated to the constant process and  $L$  is a classical discrete branching process with reproduction measure  $\mu$ . Notice that when  $k = 0$ ,

$$\varphi'(1-) = \sum_{k \in \mathbb{Z}_+ \cup \{-1\}} k\mu(k) = \sum_{k \in \mathbb{N}} k\mu(k) - d \in (-\infty, \infty],$$

and  $L$  is supercritical if and only if  $\varphi'(1-) > 0$ . We say that  $L$  is immortal if there is no death in its dynamics, i.e.  $d \equiv 0$ . A necessary and sufficient condition for the process  $L$  to be non-explosive is  $\int_1^\infty \frac{dx}{|\varphi(x)|} = \infty$ , see [Har48].

### 2.1.2 One-dimensional continuous-state branching processes

Let  $\Psi_c$  as in (9) with  $\pi(dy, dk) = \delta_0(dk)\nu(dy)$ , for a certain Lévy measure  $\nu$  on  $(0, \infty)$ , such that  $\int_0^\infty (1 \wedge y^2)\nu(dy) < \infty$ . In this case there is no dependence on the variable  $r$  in  $\Psi_c$ , and the function

$$\psi : (0, \infty) \ni q \mapsto \Psi_c(q, 1) = \Psi_c(q, r) = \frac{\sigma^2}{2}q^2 - \gamma q + \int_0^\infty (e^{-qy} - 1 + qy\mathbf{1}_{(0,1)}(y))\nu(dy) - \kappa, \quad (13)$$

is the branching mechanism of a CSBP with parameters  $(\frac{\sigma^2}{2}, \gamma, \nu, \kappa)$ . In this case the process  $L$  is *autonomous* as its dynamics does not depend on the component  $X$ .

Furthermore in the case  $\Psi_d \equiv 0$ , the component  $L$  degenerates to a constant process and  $X$  is a classical CSBP( $\psi$ ). When  $\kappa = 0$ , we recall that  $X$  is said to be supercritical if  $\psi'(0+) \in [-\infty, 0)$ , critical if  $\psi'(0+) = 0$  and subcritical when  $\psi'(0+) > 0$ . The CSBP  $X$  is immortal, i.e.  $X_t \xrightarrow[t \rightarrow \infty]{} \infty$  a.s. if and only if  $\psi \leq 0$ , see e.g. [Kyp14, Chapter 12].

Theorem 2.1 and Proposition 2.3 applied in this special setting ensure the following facts: the CSBP( $\psi$ ),  $X$ , is a  $[0, \infty]$ -valued càdlàg Feller process, with  $\infty$  as absorbing state and its generator is

$$\mathcal{G}f(x) := \frac{\sigma^2}{2}xf''(x) + \gamma xf'(x) + x \int_0^\infty (f(x+y) - f(x) - yf'(x)\mathbf{1}_{(0,1)}(y))\nu(dy) - \kappa xf(x), \quad (14)$$

acting on

$$\mathcal{D}_c := \left\{ f \in C_0^2 : \lim_{x \rightarrow \infty} x(|f(x)| + |f'(x)| + |f''(x)|) = 0 \right\}. \quad (15)$$

Denote by  $Q_t$  the semigroup associated with  $X$ . Setting  $e_q(x) := e^{-qx}$  for all  $q > 0$  and  $x \in [0, \infty)$ , one has  $\mathcal{G}e_q(x) = x\psi(q)e_q(x)$  and

$$Q_t e_q(x) = e^{-xu_t(q)} \text{ with } \frac{d}{dt}u_t(q) = -\psi(u_t(q)), u_0(q) = q.$$

Moreover, the space

$$D_c := \text{Vect}\{e_q(\cdot), q \in (0, \infty)\} \quad (16)$$

is a core and the semigroup  $(Q_t)$  uniquely satisfies the backward Kolmogorov equation

$$\forall f \in D_c, \quad \frac{d}{dt}Q_t f = \mathcal{G}Q_t f.$$

When  $\kappa = 0$ , we recall that  $X$  is said to be supercritical if  $\psi'(0+) < 0$ , critical if  $\psi'(0+) = 0$  and subcritical when  $\psi'(0+) > 0$ . The CSBP  $X$  is immortal, i.e.  $X_t \xrightarrow[t \rightarrow \infty]{} \infty$  a.s. if and only if  $\psi \leq 0$ , see e.g. [Kyp14, Chapter 12]. A necessary and sufficient condition for the continuous-state process  $X$  to be non-explosive is  $\int_0^{\infty} \frac{dx}{|\psi(x)|} = \infty$ , see Grey [Gre74].

We now give natural conditions under which the first and second coordinates are autonomous (discrete or continuous) branching processes.

**Proposition 2.5** (Explosion and autonomous coordinates).

1. In case  $\pi(dy, dk) = \nu(dy)\delta_0(dk)$  for a measure  $\nu(dy)$  on  $(0, \infty)$  such that  $\int_0^{\infty} (1 \wedge y^2)\nu(dy) < \infty$ , and under the condition  $\int_0^1 \frac{dq}{|\Psi_c(q, 1)|} = \infty$ , the coordinate  $(L_t)$  is a discrete branching process with branching mechanism

$$\Psi_d(0, r) = \sum_{k \geq 1} (r^{k+1} - r)\rho(\mathbb{R}_+, \{k\}) + d(1 - r) - kr.$$

2. In case  $\rho(dy, dk) = \delta_0(dy)\rho_d(dk)$  for a finite measure  $\rho_d(dk)$  on  $\mathbb{Z}_+$ ,  $b = 0$ , and under the condition  $\int_0^1 \frac{dr}{|\Psi_d(0, r)|} = \infty$ , the coordinate  $(X_t)$  is a CBSP with branching mechanism

$$\Psi_c(q, 1) = \frac{\sigma^2}{2}q^2 - \gamma q + \int_0^{\infty} (e^{-qy} - 1 + qy\mathbf{1}_{(0,1)}(y))\pi(dy, \mathbb{Z}_+) - \kappa.$$

The integral conditions in Proposition 2.5 ensure that the autonomous coordinate does not explode, hence preventing the situation where the other coordinate stops evolving by being in the cemetery point. The proposition is proved at the end of Section 4.

### 3 Main results

Let  $\psi$  be a branching mechanism with quadruplet by  $(\sigma, \gamma, \nu, \kappa)$ , see (13). The largest root of  $\psi$  is denoted by

$$\rho = \psi^{-1}(0) := \sup\{x > 0 : \psi(x) < 0\} \in [0, \infty].$$

As explained in the Introduction, we shall see hereafter that the CSBP( $\psi$ ) hides a family of  $\mathbb{R}_+ \times \mathbb{Z}_+$ -valued branching processes.

For any  $\lambda \in (0, \infty)$ , the Esscher transform of  $\psi$  at  $\lambda$ , is given by  $\psi_\lambda(\cdot) := \psi(\lambda + \cdot) - \psi(\lambda)$ , see Figure 1. This defines a new branching mechanism with no killing term, i.e.  $\psi_\lambda(0+) = 0$ . Specifically  $\psi_\lambda$  takes the explicit form

$$\psi_\lambda(q) = \frac{\sigma^2}{2}q^2 + \psi'(\lambda)q + \int_0^{\infty} ye^{-\lambda y}\nu(dy) (e^{-qy} - 1 + qy). \quad (17)$$

We denote by  $\mathcal{G}^{\psi_\lambda}$  the infinitesimal generator of the CSBP( $\psi_\lambda$ ).

Define the Poisson kernel  $K$ :

$$K(x, \ell) := e^{-\lambda x} \frac{(\lambda x)^\ell}{\ell!}, \forall x \in [0, \infty), \forall \ell \in \mathbb{Z}_+. \quad (18)$$

For any function  $f : (x, \ell) \mapsto f(x, \ell)$ , belonging to  $\mathcal{D}$ , we set

$$Kf(x) := \sum_{\ell=0}^{\infty} e^{-\lambda x} \frac{(\lambda x)^\ell}{\ell!} f(x, \ell), \text{ and } \mathcal{G}^{\psi_\lambda} f(x, l) := \mathcal{G}^{\psi_\lambda} f_\ell(x) \text{ with } f_\ell : x \mapsto f(x, \ell).$$

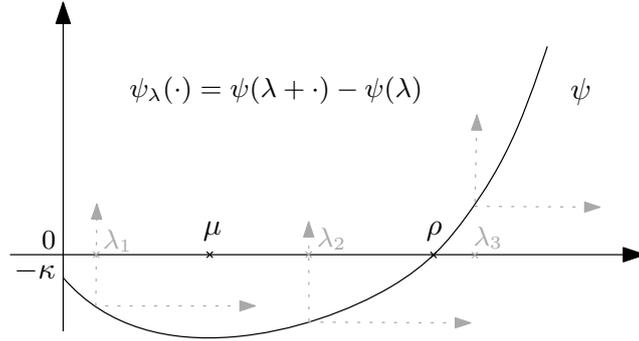


Figure 1: A supercritical mechanism and its Esscher transforms

We will introduce an operator  $\mathcal{H}$ , acting on  $\mathcal{D}$ , that will satisfy an *algebraic* intertwining relationship with the one-dimensional generator  $\mathcal{G}$  via the kernel  $K$ .

To begin, we define an operator  $\mathcal{J}$  that plays a central role. For any  $f \in \mathcal{D}$  and  $(x, \ell) \in \mathbb{R}_+ \times \mathbb{Z}_+$ , set

$$\begin{aligned} \mathcal{J}f(x, \ell) := & \ell \sum_{k \geq 0} \int_{(0, \infty)} (f(x+y, \ell+k) - f(x, \ell)) s(dy, dk) \\ & + \frac{\sigma^2}{2} \lambda \ell [f(x, \ell+1) - f(x, \ell)] + \sigma^2 \ell f'(x, \ell), \end{aligned} \quad (19)$$

with

$$s(dy, dk) := ye^{-\lambda y} \frac{(\lambda y)^k}{(k+1)!} \nu(dy) \mathbf{1}_{\mathbb{Z}_+}(k). \quad (20)$$

**Theorem 3.1** (Algebraic intertwining). *Let  $\mathcal{H}$  be the operator defined on  $\mathcal{D}$  as follows:*

i) *When  $\lambda \geq \rho$ , set*

$$\mathcal{H}f(x, \ell) := -\kappa x f(x, \ell) + \mathcal{G}^{\psi_\lambda} f(x, \ell) + \mathcal{J}f(x, \ell) + \ell \frac{\psi(\lambda)}{\lambda} [f(x, \ell-1) - f(x, \ell)]. \quad (21)$$

ii) *When  $\lambda \leq \rho$ , set*

$$\mathcal{H}f(x, \ell) := -\kappa x f(x, \ell) + \mathcal{G}^{\psi_\lambda} f(x, \ell) + \mathcal{J}f(x, \ell) - x \psi(\lambda) [f(x, \ell+1) - f(x, \ell)]. \quad (22)$$

*In both cases, one has*

$$K\mathcal{H}f(x) = \mathcal{G}Kf(x), \quad \forall x \geq 0. \quad (23)$$

The proof of Theorem 3.1 is provided in Section 5. Notice that in case i),  $\psi(\lambda) \geq 0$  while in case ii)  $\psi(\lambda) \leq 0$ . This ensures that the multiplicative factor preceding the last term in (21) and (22) is non-negative.

The operator  $\mathcal{H}$  given by (21)-(22) is the generator of a  $\mathbb{R}_+ \times \mathbb{Z}_+$ -valued branching process  $\mathbf{X} := (X^\lambda, L^\lambda)$ . We represent and explain further its dynamics in Figure 2. The jump measures  $\rho$  and  $\pi$ , as they appear in the general form of the generator (3), are given as follows

$$\rho(dy, dk) := s(dy, dk) + \frac{\sigma^2}{2} \lambda \delta_0(dy) \delta_1(dk) \quad (24)$$

and

$$\pi(dy, dk) := \begin{cases} ye^{-\lambda y} \nu(dy) \delta_0(dk) & \text{if } \lambda \geq \rho, \\ ye^{-\lambda y} \nu(dy) \delta_0(dk) - \psi(\lambda) \delta_0(dy) \delta_1(dk) & \text{if } \lambda < \rho. \end{cases} \quad (25)$$

The following corollary gives an analytic expression for the joint branching mechanism  $\Psi = (\Psi_c, -\Psi_d)$  of  $(X^\lambda, L^\lambda)$ , thereby elucidating its connection to  $\psi$ .

**Corollary 3.2** (Joint branching mechanism). *For any  $\lambda > 0$ ,*

$$\psi(q + \lambda(1 - r)) = \Psi_c(q, r) + \lambda\Psi_d(q, r), \quad \forall (q, r) \in (0, \infty) \times (0, 1). \quad (26)$$

Moreover,

$$i) \text{ When } \lambda \geq \rho: \quad \Psi_c(q, r) = \psi_\lambda(q) - \kappa, \quad (27)$$

$$\Psi_d(q, r) = \sum_{k \in \mathbb{Z}_+} \int_{(0, \infty)} (e^{-qy} r^{k+1} - r) \rho(dy, dk) + \frac{\psi(\lambda)}{\lambda} (1 - r) - \sigma^2 qr \quad (28)$$

$$= \frac{\psi(q + \lambda(1 - r)) - \psi(q + \lambda) + \psi(\lambda) + \kappa}{\lambda}. \quad (29)$$

$$ii) \text{ When } \lambda \leq \rho: \quad \Psi_c(q, r) = \psi_\lambda(q) - \kappa - \psi(\lambda)(r - 1), \quad (30)$$

$$\Psi_d(q, r) = \sum_{k \in \mathbb{Z}_+} \int_{(0, \infty)} (e^{-qy} r^{k+1} - r) \rho(dy, dk) - \sigma^2 qr \quad (31)$$

$$= \frac{\psi(q + \lambda(1 - r)) - \psi(q + \lambda) + r\psi(\lambda) + \kappa}{\lambda}. \quad (32)$$

The proof of Corollary 3.2 is in Section 6.1.

*Remark 3.3.* Notice that when  $\rho \in (0, \infty)$ , the map  $\lambda \mapsto (\Psi_c, \Psi_d)$  is continuous (for the uniform norm) at  $\lambda = \rho$ , since  $\psi$  is continuous and  $\psi(\rho) = 0$ .

For any  $\lambda > 0$  and  $x \in [0, \infty)$ , denote by  $\text{Poi}(\lambda x)$  the Poisson law with parameter  $\lambda x$  and notice that  $K(x, \ell) = \mathbb{P}(\text{Poi}(\lambda x) = \ell)$  for all  $\ell \in \mathbb{Z}_+$ . Let  $(P_t)$  be the semigroup of the branching process  $(X^\lambda, L^\lambda)$  with mechanism  $(\Psi_c, -\Psi_d)$ , and  $(Q_t)$  be the semigroup of a CSBP( $\psi$ ). Call  $(\mathcal{F}_t^{X^\lambda})$  the usual augmentation of the natural filtration of  $X^\lambda$ , see e.g. Revuz-Yor's book [RY05, pages 45 and 93].

**Theorem 3.4** (Intertwined semigroups and skeleton decomposition). *For any function  $f \in C_0(\mathbb{R}_+ \times \mathbb{Z}_+)$ , one has*

$$\Lambda P_t f(x) = Q_t \Lambda f(x), \quad \forall t, x \geq 0, \quad (33)$$

where for all  $x \in \bar{\mathbb{R}}_+$  and  $\ell \in \bar{\mathbb{Z}}_+$ ,

$$\Lambda(x, (x, \ell)) := K(x, \ell) \mathbf{1}_{\{x < \infty\}} + \delta_\Delta(x, \ell) \mathbf{1}_{\{x = \infty\}}. \quad (34)$$

Moreover, if  $(X^\lambda, L^\lambda)$  has for initial law  $X_0^\lambda = x$  and  $L_0^\lambda \stackrel{\text{law}}{=} \text{Poi}(\lambda x)$ , then the following holds:

1. The process  $X^\lambda$  is a CSBP( $\psi$ ) started from  $x$ .
2. For all  $\ell \in \mathbb{Z}_+$  and any  $(\mathcal{F}_t^{X^\lambda})$ -stopping time  $\tau$ ,

$$\mathbb{P}(L_\tau^\lambda = \ell | \mathcal{F}_\tau^{X^\lambda}) = K(X_\tau^\lambda, \ell) \text{ a.s. on the set } \{X_\tau < \infty\}.$$

Theorem 3.4 is established in Section 6.2 and relies on Pitman-Rogers theorem, see Theorem A.

Let us provide a more detailed explanation of the statement of Theorem 3.4. Consider first the case of a supercritical branching mechanism  $\psi$ , i.e.  $\psi'(0+) < 0$ , with no killing ( $\kappa = 0$ ) and with  $\rho \in (0, \infty)$ . Regarding the event of asymptotic extension, one has:

$$\mathbb{P}_x(X_t \xrightarrow[t \rightarrow \infty]{} 0) = e^{-x\rho} = \mathbb{P}(L_0^\rho = 0) \in (0, 1),$$

see e.g. [Kyp14, Theorem 12.7] for the first equality. In the case  $\lambda = \rho$ , the process  $(X^\lambda, L^\lambda)$  coincides with the process studied in [BFM08, pages 722-723]. It is a branching process with two types: the *prolific* individuals, which are represented along the discrete component  $(L_t^\rho, t \geq 0)$  and the non-prolific ones, evolving in  $\mathbb{R}_+$ . Since  $\psi(\rho) = 0$ , the discrete branching process  $(L_t^\rho, t \geq 0)$  has no death term,

see (28), and is thus immortal. The event of having no prolific individuals in the CSBP,  $\{L_0^\rho = 0\}$ , also matches with the event of asymptotic extinction. Furthermore, the continuous dynamics is governed by  $\psi_\rho$ , which is the mechanism of a CSBP( $\psi$ ) conditioned on its extinction, see [Kyp14, Exercice 12.4]. Finally, Theorem 3.4 asserts that the projection onto the continuous component results in a CSBP( $\psi$ ).

More generally, for any branching mechanism  $\psi$  such that  $\rho \in [0, \infty)$ , when  $\lambda \geq \rho$ , case i), the branching mechanism  $\psi_\lambda$  is *subcritical*, see Figure 1. The CSBPs( $\psi_\lambda$ ), with generator  $\mathcal{G}^{\psi_\lambda}$ , are thus converging towards 0 almost surely. Notice that  $\Psi_c$ , see (27) does not depend on the variable  $r$ . Thus, the discrete component  $L^\lambda$  is autonomous. In particular, conditionally on  $\{L_0^\lambda = 0\}$ ,  $L_t^\lambda = 0$  for all  $t \geq 0$  a.s. and  $X^\lambda$  is a CSBP( $\psi_\lambda$ ) started from  $x$ .

We now explain the dynamics of the process  $\mathbf{X} = (X^\lambda, L^\lambda)$  for both scenarios:  $\lambda > \rho$  and  $\lambda < \rho$ , see Figure 2. We call *skeleton*, the discrete component  $L^\lambda$ , and refer to the process  $\mathbf{X}$  as a skeleton decomposition.

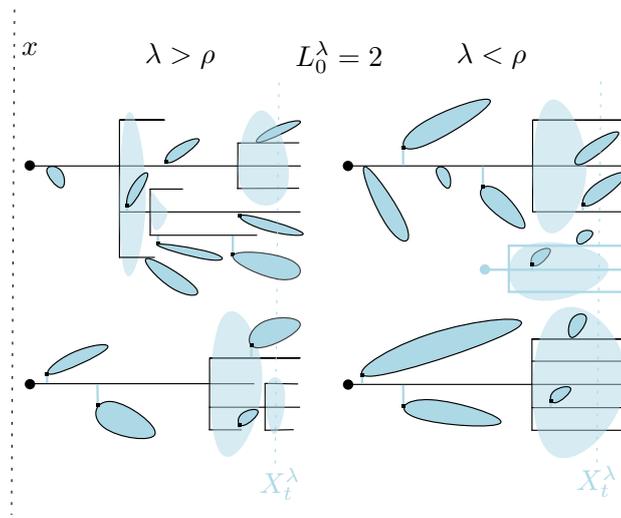


Figure 2: Schematic representation of the intertwined two-type branching process

The skeleton is depicted by lines, while the blue parts represent the continuous mass. The bubbles encircled in black represent CSBPs( $\psi_\lambda$ ) that are grafted along each line. The CSBP starting from a black square starts from a macroscopic mass governed by the Lévy measure<sup>1</sup>  $s(dy, \{0\}) = ye^{-\lambda y}\nu(dy)$ . The bubbles that stick to the skeleton only exist if  $\sigma > 0$ . They arrive on each line at rate  $\sigma^2$  and represent CSBPs started from an infinitesimal mass<sup>2</sup>.

At any splitting time of the skeleton, say into  $k \geq 1$  new lines, a continuous mass, governed by  $s(dy, \{k\})$ , is spread and starts evolving as a CSBP( $\psi_\lambda$ ). This is depicted by the blue shadow behind the lines. In the case  $\psi(0+) = -\kappa < 0$ , the continuous mass is furthermore killed at rate  $\kappa x$ . In other words, we graft CSBPs with mechanism  $\psi_\lambda - \kappa$ . In any case, the skeleton and the grafted CSBPs accommodate in a way so that the first coordinate process  $X^\lambda$  keeps the law of a CSBP( $\psi$ ).

- i) When  $\lambda > \rho$ , a death term, with rate  $\frac{\psi(\lambda)}{\lambda}$ , in the skeleton  $L^\lambda$  emerges, see (28). The latter has therefore leaves as represented in Figure 2. Heuristically, since  $\psi'_\lambda(0+) = \psi'(\lambda)$  and  $\psi(\lambda)/\lambda$  increase on  $[\rho, \infty)$ , the greater  $\lambda$  is, the more subcritical are the CSBP( $\psi_\lambda$ ) and the more leaves has the skeleton.
- ii) When  $\lambda < \rho$ , case ii), the dynamics changes as the continuous mass  $X^\lambda$  also begets discrete individuals. It is acting on the skeleton through an additional linear birth term with positive

<sup>1</sup>One has  $\int_0^1 y s(dy, \{0\}) < \infty$

<sup>2</sup>A rigorous formalisation requires to work with the canonical measure, see e.g. [AD09]. We shall not need it in our construction here

rate  $-\psi(\lambda)$ . In this setting, the skeleton is immortal (it has no leaf) and is not autonomous, as  $\Psi_c$  depends on the variable  $r$ , see (30).

Let  $\mu := \operatorname{argmin} \psi \in [0, \infty]$  be the location of the minimum of  $\psi$ , see Figure 1. The branching mechanism  $\psi_\lambda$  is subcritical, critical or supercritical when respectively  $\lambda > \mu$ ,  $\lambda = \mu$  and  $\lambda < \mu$ . When  $\lambda < \mu$ , not only does the continuous mass generate new lines, but the grafted CSBPs must also be supercritical to ensure that the total continuous mass process  $X^\lambda$  retains the law of the CSBP( $\psi$ ).

Notice that when the CSBP( $\psi$ ) is *immortal*, i.e. it tends to  $\infty$  a.s., one has  $\rho = \mu = \infty$  and only the decomposition (ii), with supercritical grafted CSBPs, makes sense. A classical example in this framework is the stable mechanism  $\psi(\lambda) = -c\lambda^\alpha$  with  $\alpha \in (0, 1)$ ,  $c > 0$ .

*Remark 3.5* (Conditional law of  $X^\lambda$  given  $L^\lambda$ ). Let  $n \geq 0$  and let  $A$  be a Borel set. Since conditionally on  $X_t^\lambda$ ,  $L_t^\lambda \stackrel{\text{law}}{=} \operatorname{Poi}(\lambda X_t^\lambda)$ , we see that:

$$\mathbb{P}(X_t^\lambda \in A | L_t^\lambda = n) = \frac{\mathbb{P}(X_t^\lambda \in A, L_t^\lambda = n)}{\mathbb{P}(L_t^\lambda = n)} = \frac{\mathbb{E}[e^{-\lambda X_t^\lambda} (X_t^\lambda)^n \mathbf{1}_A(X_t^\lambda)]}{\mathbb{E}[e^{-\lambda X_t^\lambda} (X_t^\lambda)^n]}.$$

*Remark 3.6* (A one-dimensional intertwining). The identities (23) and (33) encapsulate an intertwining relationship between the CSBP and its skeleton alone. Let  $G^\lambda$  and  $Q_t^\lambda$  denote the infinitesimal generator and the semigroup of the process  $L^\lambda$ . For any bounded function  $f : \mathbb{Z}_+ \mapsto \mathbb{R}_+$ , set  $Kf(x) = \sum_{\ell \geq 0} f(\ell) \frac{(\lambda x)^\ell}{\ell!} e^{-\lambda x}$ , one has for all  $x \geq 0$  and  $t \geq 0$ ,

$$KG^\lambda f(x) = \mathcal{G}Kf(x) \text{ and } KQ_t^\lambda f(x) = Q_t Kf(x).$$

We emphasize that these identities alone do not establish any direct coupling between the processes governed by the semigroups  $Q_t^\lambda$  and  $Q_t$ . In general, finding such a coupling is challenging - see the pioneering article [DF90] and for instance [Mic20]. Within the framework of Theorem 3.4, the intertwining relationship is directly related to a function (the projection on the first coordinate) and Pitman-Rogers theorem, see Theorem A, applies.

As an application of Theorem 3.1, we now study the phenomenon of explosion of a CSBP.

Notice that for all  $\lambda \in (0, \infty)$ ,  $\psi'_\lambda(0+) = \psi'(\lambda) \in (-\infty, \infty)$ , the grafted CSBPs have thus finite mean. In particular, none is exploding. In the case  $\kappa = 0$ , if the CSBP( $\psi$ ) can explode, i.e.  $\int_0 \frac{dx}{|\psi(x)|} < \infty$ , see e.g. [Kyp14, Theorem 12.3], its explosion occurs simultaneously with any of its skeleton  $L^\lambda$ . Heuristically, the only way to reach infinity by accumulation of large jumps is to graft CSBPs on an infinite number of lines. This is stated in the next proposition whose proof will appear Theorem 3.4.

**Proposition 3.7** (Simultaneous explosion). *If  $(X^\lambda, L^\lambda)$  has for initial law  $X_0^\lambda = x$  and  $L_0^\lambda \stackrel{\text{law}}{=} \operatorname{Poi}(\lambda x)$ , then the following identity of events holds a.s.*

$$\{L_t^\lambda = \infty\} = \{X_t^\lambda = \infty\}, \text{ for all } t \geq 0.$$

*In particular, when  $\int_0 \frac{dx}{|\psi(x)|} < \infty$ , explosion has positive probability and occurs simultaneously in the CSBP( $\psi$ ),  $X^\lambda$ , and the skeleton  $L^\lambda$ .*

*Remark 3.8.* Notice that the event of simultaneous extinction of both types is not almost sure, indeed for any  $\lambda > 0$ , when  $L_0^\lambda \stackrel{\text{law}}{=} \operatorname{Poi}(\lambda x)$  and  $X_0^\lambda = x$ , we have by Theorem 3.4(2):

$$\mathbb{P}(L_t^\lambda = 0) = \mathbb{E}[e^{-\lambda X_t^\lambda}] > \mathbb{P}(X_t^\lambda = 0), \forall t \geq 0.$$

The proof of Proposition 3.7 is in Section 6.3.

We study now the family of skeletons, with  $\lambda \geq \rho$ , and will establish that suitably renormalized, they converge in the sense of Skorokhod, towards a CSBP( $\psi$ ), see Theorem 3.12 below.

We start by providing explicitly the reproduction law of the autonomous skeletons. The following formula can be found in [DW07, FFK19, LUB24] in the case  $\kappa = 0$ . We will shed some light on the calculations behind.

**Proposition 3.9** (Offspring distribution of skeletons). *Assume  $\lambda \geq \rho$ . In this case, the skeleton  $(L_t^\lambda, t \geq 0)$  is autonomous and its branching mechanism is given by*

$$\varphi_\lambda(r) := \Psi_d(0, r) = \frac{1}{\lambda} \left( \psi(\lambda(1-r)) + \kappa \right) = \psi'(\lambda) \sum_{k=-1}^{\infty} (r^{k+1} - r) p_k^\lambda, \quad (35)$$

with  $(p_k^\lambda, k \geq -1)$  the probability measure given by

$$p_{-1}^\lambda = \frac{1}{\lambda \psi'(\lambda)} \psi(\lambda), \quad p_0^\lambda = 0, \quad p_k^\lambda = \frac{1}{\lambda \psi'(\lambda)} \left( \frac{\sigma^2 \lambda^2}{2} \mathbf{1}_{\{k=1\}} + \int_0^\infty e^{-\lambda y} \frac{(\lambda y)^{k+1}}{(k+1)!} \nu(dy) \right), \quad \forall k \geq 1. \quad (36)$$

The proof of Proposition 3.9 is in Section 7.1.

*Remark 3.10.* Notice that  $\varphi'_\lambda(1-) = -\psi'(0+)$ . In particular,  $L^\lambda$  is (sub)critical if and only if the CSBP( $\psi$ ) is (sub)critical. Observe also that when  $\lambda < \rho$ , the formula (36) cannot define an offspring distribution as at least one of the following holds  $\psi'(\lambda) < 0$  or  $\psi(\lambda) < 0$ , which violates the condition  $p_k \geq 0$  for all  $k$ . The formula (36) is reminiscent to the coalescence rates of ancestral lineages in CSBPs backwards in time, see Foucart et al. [FMM19, Theorem 5.10] and Johnston and Lambert [JL23].

*Remark 3.11* (Binomial intertwining). The family of continuous-time Galton-Watson processes  $(L^\lambda, \lambda \geq \rho)$  satisfies an intertwining relationship with one another through a binomial kernel. This can be seen as a counterpart of the Bernoulli leaf coloring investigated in [DW07, Section 4.1]. Recall  $G^\lambda$  and  $Q_t^\lambda$ , the generator and semigroup of  $L^\lambda$ . For all  $\mu > \lambda > \rho$  and any bounded function  $f : \mathbb{Z}_+ \rightarrow \mathbb{R}$ , we have for all  $n \geq 0$  and  $t \geq 0$ ,

$$G^\mu B_{\lambda/\mu} f(n) = B_{\lambda/\mu} G^\lambda f(n) \text{ and } Q_t^\mu B_{\lambda/\mu} f(n) = B_{\lambda/\mu} Q_t^\lambda f(n)$$

with  $B_{\lambda/\mu} f(n) := \sum_{k=0}^n \binom{n}{k} (\lambda/\mu)^k (1 - \lambda/\mu)^{n-k} f(k)$ .

The next theorem shows that for any branching mechanism  $\psi$ , the skeletons converge towards the CSBP( $\psi$ ) in the Skorokhod sense.

**Theorem 3.12.** *Let  $x \in [0, \infty)$ . Denote by  $(X_t(x), t \geq 0)$  a CSBP( $\psi$ ) starting from  $x$ , and let  $(L_t^\lambda(x), t \geq 0)$  be a discrete branching Markov process with mechanism  $\varphi_\lambda$ , see (35), starting from an independent Poisson number of individuals with parameter  $\lambda x$ . Then, as  $\lambda$  go to  $\infty$ , we have,*

$$\left( \frac{1}{\lambda} L_t^\lambda(x), t \geq 0 \right) \xrightarrow{\mathbb{D}} (X_t(x), t \geq 0). \quad (37)$$

Theorem 3.12 has been established in [DW19] in the setting of non-immortal and non-explosive CSBPs, i.e.  $\rho < \infty$  and  $\int_0^\infty \frac{dx}{|\psi(x)|} = \infty$  respectively. The main argument in [DW19] relies on a work of Helland [Hel78] about random time-change transformations. We will follow here a different approach allowing us to treat all CSBPs.

The proofs of Proposition 3.9 and Theorem 3.12 are given in Section 7.

*Remark 3.13.* The fact that we considered the continuous-time Galton-Watson processes  $L^\lambda$  with a Poisson number of initial individuals simplifies many calculations. Note that by Theorem 3.4,  $Z_t^\lambda$  has a Poisson law with parameter  $\lambda X_t$  where  $X$  is a CSBP( $\psi$ ). The convergence of the one-dimensional law can therefore be seen as an application of the weak law of large numbers. The convergence in the finite-dimensional sense stays true if one considers instead the processes  $(L_t^\lambda/\lambda, t \geq 0)$  with  $L_0^\lambda = [\lambda x]$ . We refer the reader to [LG99, pages 20-21] for calculations in this direction.

## 4 Proofs of Theorem 2.1 and Proposition 2.3

In this section, we investigate the class of branching processes, potentially explosive, comprising one continuous and one discrete component. The construction is made in three steps. First, we establish

the existence of a càdlàg solution to the martingale problem  $\text{MP}_{\mathbf{X}}(\mathcal{L}, \mathcal{D})$  (at this stage the process is not yet known to be markovian). Second, we observe a duality relationship between the operator  $\mathcal{L}$  and a system of o.d.es through the functions indexed by  $(q, r) \in (0, \infty) \times (0, 1)$  and defined by

$$f_{q,r}(x, \ell) := e^{-qx} r^\ell, \text{ for all } (x, \ell) \in E = \mathbb{R}_+ \times \mathbb{Z}. \quad (38)$$

We show existence and uniqueness of a global solution to this system. Last, standard results, see for instance Ethier and Kurtz's book [EK09, Chapter 4], will apply and ensure that the duality holds true at the level of the one-dimensional laws, entailing uniqueness of the solution to the martingale problem and consequently the strong Markov property. Recall  $\Delta = \{(x, \ell) : x + \ell = \infty\}$  the cemetery point.

**Lemma 4.1.** *There exists a càdlàg solution  $(\mathbf{X}_t)_{t \geq 0} = (X_t, L_t)_{t \geq 0}$  to the martingale problem associated to  $(\mathcal{L}, \mathcal{D})$ , taking values in  $E \cup \{\Delta\}$  and killed at the first explosion time,  $\zeta := \inf\{t > 0 : (X_{t-}, L_{t-}) = \Delta \text{ or } (X_t, L_t) = \Delta\}$ .*

*Proof.* We apply here [EK09, Theorem 5.4 page 199]. Recall the space of functions  $\mathcal{D}$  defined in (4). Notice that by definition  $C_0(\mathbb{R}_+ \times \mathbb{Z}) = \{f \text{ continuous s.t. } f(x, \ell) \rightarrow 0, \text{ as } (x, \ell) \rightarrow \Delta\}$ . We must first check that  $\mathcal{L}$  is a linear operator on  $C_0(\mathbb{R}_+ \times \mathbb{Z})$ , see [EK09, Page 8]. Plainly  $\mathcal{D} \subset C_0(\mathbb{R}_+ \times \mathbb{Z})$  and we only have to verify that the range of  $\mathcal{L}$ , i.e.  $\{\mathcal{L}f : f \in \mathcal{D}\}$ , is a subset of  $C_0(\mathbb{R}_+ \times \mathbb{Z})$ . Recall  $\mathcal{L}$  in (3). We focus on the term of small jumps, the others are treated along similar arguments:

$$\begin{aligned} & \left| x \sum_{k \geq 0} \int_0^1 (f(x+y, \ell+k) - f(x, \ell) - yf'(x, \ell)) \pi(dy, dk) \right| \\ & \leq \int_0^1 x |f(x+y, \ell) - f(x, \ell) - yf'(x, \ell)| \pi(dy, \{0\}) \\ & \quad + \sum_{k \geq 1} \int_0^1 x |f(x+y, \ell+k) - f(x, \ell) - yf'(x, \ell)| \pi(dy, dk). \end{aligned}$$

Since  $f \in \mathcal{D}$ , the integrand in the first integral above is continuous in  $x$  and tends to zero as  $x + \ell$  goes to  $\infty$ . It is moreover dominated by  $Cy^2/2$  with  $C := \sup_{(x, \ell) \in E} |f''(x, \ell)|$ , which is integrable on  $(0, 1)$  with respect to the measure  $\pi(dy, \{0\})$ , see (2). Hence, by Lebesgue theorem, the first integral term is continuous in  $x$  and vanishes. This is also true for the second integral term, since the measure  $\pi(dy, dk)$ , restricted to  $\mathbb{R}_+ \times \mathbb{N}$ , is finite, see (2), and the integrand is continuous vanishing as  $(x, \ell) \rightarrow \Delta$ . We conclude that  $\mathcal{L}f \in C_0(\mathbb{R}_+ \times \mathbb{Z}_+)$ .

We argue now that  $\mathcal{D}$  is dense in  $C_0(\mathbb{R}_+ \times \mathbb{Z}_+)$  for the uniform norm. The linear span of the functions  $f_{q,r}$ ,  $D = \text{Vect}\{f_{q,r}, q > 0, r < 1\}$  is a subset of  $\mathcal{D}$  and forms a subalgebra of  $C_0(\mathbb{R}_+ \times \mathbb{Z}_+)$ , that is separating  $\mathbb{R}_+ \times \mathbb{Z}_+$ . Therefore by Stone-Weierstrass theorem, see e.g. [Rud76, Theorem 7.32 page 162], the latter is dense in  $C_0(\mathbb{R}_+ \times \mathbb{Z}_+)$ . Next, we show the positive maximum principle, see [EK09, page 165], namely we check that for any  $f \in \mathcal{D}$ , if  $(x_0, \ell_0) \in E = \mathbb{R}_+ \times \mathbb{Z}_+$  is such that  $\sup f(x, \ell) = f(x_0, \ell_0) \geq 0$ , then

$$\mathcal{L}f(x_0, \ell_0) \leq 0.$$

First, plainly since  $\kappa \geq 0$  and  $k \geq 0$ , one has  $-\kappa x_0 f(x_0, \ell_0) \leq 0$  and  $-k \ell_0 f(x_0, \ell_0) \leq 0$ . Next, if  $\ell_0 \in \mathbb{Z}_+$  and  $x_0 \in (0, \infty)$  then  $f(\ell_0 + k, x_0) \leq f(\ell_0, x_0)$ , similarly  $f(\ell_0 + k, x_0 + y) \leq f(\ell_0, x_0)$  and necessarily  $f'(x_0, \ell_0) = 0$ , so that the compensation term and the drift term vanish. Last, if  $x_0 = 0$ , then necessarily  $f'(0, \ell_0) \leq 0$  and since  $b \geq 0$ , the drift term is also non-positive. In both cases  $x_0 > 0$  and  $x_0 = 0$ , the second derivative term in  $x_0$ , coming from the diffusive part, is nonpositive since  $x_0$  is a maximum and  $\sigma \geq 0$ .  $\square$

The next lemma states an *algebraic* duality relationship for the operator  $\mathcal{L}$ .

**Lemma 4.2.** *For all  $(x, \ell) \in \mathbb{R}_+ \times \mathbb{Z}_+$ , define on  $(0, \infty) \times (0, 1)$ , the map*

$$g_{x, \ell} : (q, r) \mapsto e^{-qx} r^\ell.$$

One has

$$\mathcal{L}f_{q,r}(x, \ell) = xe^{-qx}r^\ell\Psi_c(q, r) + e^{-qx}\ell r^{\ell-1}\Psi_d(q, r) \quad (39)$$

$$= -\Psi_c(q, r)\frac{\partial}{\partial q}g_{x,\ell}(q, r) + \Psi_d(q, r)\frac{\partial}{\partial r}g_{x,\ell}(q, r) \quad (40)$$

$$= -\Psi \cdot \nabla g_{x,\ell}(q, r) \quad (41)$$

where  $\Psi = (\Psi_c, -\Psi_d)$ .

*Proof.* Recall  $\mathcal{L}$  in (3), one has

$$\begin{aligned} \mathcal{L}f_{q,r}(x, \ell) &= -\gamma x q e^{-qx} r^\ell - b \ell q e^{-qx} r^\ell + \frac{\sigma^2}{2} x q^2 e^{-qx} r^\ell - \kappa x e^{-qx} r^\ell - \kappa \ell e^{-qx} r^\ell \\ &\quad + x \sum_{k \geq 0} \int_{\mathbb{R}_+} \left( e^{-q(x+y)} r^{\ell+k} - e^{-qx} r^\ell + y \mathbf{1}_{(0,1)}(y) q e^{-qx} r^\ell \right) \pi(dy, dk) \\ &\quad + \ell \sum_{k \geq 0} \int_{\mathbb{R}_+} \left( e^{-q(x+y)} r^{\ell+k} - e^{-qx} r^\ell \right) \rho(dy, dk) + \underbrace{d\ell(e^{-qx} r^{\ell-1} - e^{-qx} r^\ell)}_{=d\ell r^{\ell-1} e^{-qx} (1-r)}. \end{aligned}$$

By rearranging everything in order to make appear  $\Psi_c$  and  $\Psi_d$ , as defined in (9) and (10), we get

$$\mathcal{L}f_{q,r}(x, \ell) = xe^{-qx}r^\ell\Psi_c(q, r) + e^{-qx}\ell r^{\ell-1}\Psi_d(q, r).$$

Using the facts that

$$xe^{-qx}r^\ell = -\frac{\partial}{\partial q}g_{x,\ell}(q, r) \text{ and } \ell e^{-qx}r^{\ell-1} = \frac{\partial}{\partial r}g_{x,\ell}(q, r),$$

we have finally

$$\mathcal{L}f_{q,r}(x, \ell) = -\Psi_c(q, r)\frac{\partial}{\partial q}g_{x,\ell}(q, r) + \Psi_d(q, r)\frac{\partial}{\partial r}g_{x,\ell}(q, r).$$

□

**Lemma 4.3.** *There exists a unique solution  $t \mapsto F_t(q, r) := (u_t(q, r), f_t(q, r))$  to*

$$\frac{d}{dt}u_t(q, r) = -\Psi_c(u_t(q, r), f_t(q, r)), \quad \frac{d}{dt}f_t(q, r) = \Psi_d(u_t(q, r), f_t(q, r)) \quad (42)$$

$$u_0(q, r) = q, \quad f_0(q, r) = r. \quad (43)$$

Furthermore, for all  $t \in [0, \infty)$ ,  $r \in (0, 1)$  and  $q \in (0, \infty)$ ,

$$0 < u_t(q, r) < \infty \text{ and } 0 < f_t(q, r) < 1.$$

*Proof.* All first partial derivatives of the function  $\Psi(q, r) = (\Psi_c(q, r), -\Psi_d(q, r))$  are continuous and bounded on domains of the form  $(a, b) \times (l, r) \subset (0, \infty) \times (0, 1)$ . This entails that  $\Psi$  is locally Lipschitz on  $(0, \infty) \times (0, 1)$ , see e.g. [BR78, Chapter 6, page 174]. Cauchy-Lipschitz theorem ensures the existence of a local solution. We now show that it is bounded below and above by positive and finite quantities.

We first find an upper bound for  $\Psi_c$ . Rewrite the integrand of (9) with the help of the identity

$$e^{-qy}r^k - 1 + qy\mathbf{1}_{(0,1)}(y) = e^{-qy}r^k - e^{-qy} + e^{-qy} - 1 + qy\mathbf{1}_{(0,1)}(y), \quad \forall y \in (0, \infty), k \geq 0. \quad (44)$$

Since  $e^{-qy}r^k - e^{-qy} \leq 0$ ,  $e^{-qy} - 1 \leq 0$  and there is  $C > 0$  such that for all  $y \in (0, 1)$ ,  $0 \leq e^{-qy} - 1 + qy \leq C\frac{(qy)^2}{2}$ , we get

$$\begin{aligned} \Psi_c(q, r) &\leq \sum_{k \in \mathbb{N}} \int_{(0, \infty)} (e^{-qy} - 1 + qy\mathbf{1}_{(0,1)}(y)) \pi(dy, dk) - \gamma q + \frac{\sigma^2}{2} q^2 - \kappa \\ &\leq \sum_{k \in \mathbb{N}} \int_{(0, 1)} (e^{-qy} - 1 + qy) \pi(dy, dk) - \gamma q + \frac{\sigma^2}{2} q^2 \\ &\leq c_{\sigma, \pi} q^2 - \gamma q \end{aligned} \quad (45)$$

with  $c_{\sigma,\pi} := \frac{\sigma^2}{2} + C \int_0^1 y^2 \pi(dy, \mathbb{Z}_+)$ . By comparison, we see from the o.d.e

$$\frac{d}{dt} u_t(q, r) = -\Psi_c(u_t(q, r), f_t(q, r))$$

that

$$u_t(q, r) \geq \frac{q\gamma e^{-\gamma t}}{\gamma + \frac{qc_{\sigma,\pi}}{2}(e^{-\gamma t} - 1)} > 0.$$

We now show that  $f_t(q, r) < 1$ . Recall (10). Plainly  $e^{-qy}r^{k+1} - r = r(e^{-qy}r^k - 1) \leq 0$  and  $b, d \geq 0$ , thus  $\Psi_d(q, r) \leq d(1 - r)$ . Hence,

$$\frac{d}{dt} f_t(q, r) = \Psi_d(u_t(q, r), f_t(q, r)) \leq d(1 - f_t(q, r)).$$

By setting  $k_t(q, r) = 1 - f_t(q, r)$ , we get  $\frac{d}{dt} k_t(q, r) \geq -dk_t(q, r)$ , so  $k_t(q, r) \geq re^{-dt} > 0$  and  $f_t(q, r) < 1$  for all  $t \geq 0$ .

It remains to establish that for all  $(q, r) \in (0, \infty) \times (0, 1)$ ,

$$u_t(q, r) < \infty \text{ and } f_t(q, r) > 0, \forall t \geq 0.$$

Recall

$$\frac{d}{dt} u_t(q, r) = -\Psi_c(u_t(q, r), f_t(q, r))$$

with

$$\begin{aligned} -\Psi_c(q, r) &= -\sum_{k \geq 0} \int_{\mathbb{R}_+} (e^{-qy}r^k - 1 + qy\mathbf{1}_{(0,1)}(y)) \pi(dy, dk) + \gamma q - \frac{\sigma^2}{2} q^2 + \kappa \\ &\leq \int_{\mathbb{R}_+} (1 - e^{-qy} - qy\mathbf{1}_{(0,1)}(y)) \pi(dy, \{0\}) + \sum_{k \geq 1} \int_{\mathbb{R}_+} \pi(dy, dk) + \gamma q + \kappa \\ &\leq \int_1^\infty \pi(dy, \{0\}) + \sum_{k \geq 1} \int_{\mathbb{R}_+} \pi(dy, dk) + \gamma q + \kappa =: c_{\pi,\kappa} + \gamma q \end{aligned}$$

hence  $\frac{d}{dt} u_t(q, r) \leq c_{\pi,\kappa} + \gamma u_t(q, r)$ , which in turn implies, recalling  $u_0(q, r) = q$ ,

$$u_t(q, r) \leq \frac{(c_{\pi,\kappa} + \gamma q)e^{\gamma t} - c_{\pi,\kappa}}{\gamma}. \quad (46)$$

For any  $\eta \in (0, \infty)$ , define  $u_t^d(q, \eta)$  such that

$$e^{-u_t^d(q, \eta)} = f_t(q, e^{-\eta}). \quad (47)$$

Set

$$\bar{\Psi}_d(q, \eta) := e^\eta \Psi_d(q, e^{-\eta}).$$

We see from the o.d.e (42) solved by  $f_t(q, r)$  that

$$\frac{d}{dt} u_t^d(q, \eta) = -\bar{\Psi}_d(u_t(q, e^{-\eta}), u_t^d(q, \eta)).$$

By plugging  $r = e^{-\eta}$  in the expression (10) of  $\Psi_d(q, r)$ , we get

$$\begin{aligned}
-\bar{\Psi}_d(q, \eta) &= \sum_{k=0}^{\infty} \int_{\mathbb{R}_+} (1 - e^{-qy} e^{-\eta k}) \rho(dy, dk) + bq - d(e^\eta - 1) - k \\
&\leq \int_{\mathbb{R}_+} (1 - e^{-qy}) \rho(dy, \{0\}) + \sum_{k=1}^{\infty} \int_{\mathbb{R}_+} \rho(dy, \{k\}) + bq \\
&\leq \left( \int_0^1 y \rho(dy, \{0\}) + b \right) q + \int_1^\infty \rho(dy, \{0\}) + \sum_{k=1}^{\infty} \int_{\mathbb{R}_+} \rho(dy, \{k\}) =: c_{\rho,1} q + c_{\rho,2},
\end{aligned} \tag{48}$$

hence  $\frac{d}{dt} u_t^d(q, \eta) = c_{\rho,1} u_t(q, e^{-\eta}) + c_{\rho,2}$ , and given the bound (46) above on  $u_t(q, r)$ , we get

$$u_t^d(q, \eta) \leq \eta + c_{\rho,1} \frac{(c_{\pi,\kappa} + \gamma q) e^{\gamma t} - \gamma c_{\pi,\kappa} t}{\gamma^2} + c_{\rho,2} t,$$

which ensures  $u_t^d(q, \eta) < \infty$ , and in turn, by (47),  $f_t(q, r) > 0$ .

Let  $(q, r) \in (0, \infty) \times (0, 1)$ . We have just established that up to any time  $t \geq 0$ , any solution  $[0, t] \ni s \mapsto (u_s(q, r), f_s(q, r))$  to (42) stays in a domain  $D \subset (0, \infty) \times (0, 1)$ . The function  $\Psi$  being Lipschitz on such domain  $D$ , there is a unique solution to the equation up to time  $t$ . The latter being arbitrary, the solution is global, namely it is defined on the whole half-line.  $\square$

We now characterize the one-dimensional law of the solution  $(X, L)$  with the help of  $t \mapsto F_t(q, r)$ .

**Lemma 4.4.** *For all  $(q, r) \in (0, \infty) \times [0, 1)$ ,*

$$\mathbb{E}_{(x,n)}[e^{-qX_t r^{L_t}}] = e^{-xu_t(q,r)} f_t(q, r)^n \quad \forall (x, n) \in \mathbb{R}_+ \times \mathbb{Z}_+, \forall t \geq 0. \tag{49}$$

*Proof.* The task here is to establish the duality relationship between the solution to  $\text{MP}(\mathcal{L}, \mathcal{D})$ ,  $(X, L)$  provided by Lemma 4.1, and the deterministic process  $t \mapsto F_t(q, r) = (u_t(q, r), f_t(q, r))$ . Having noticed the algebraic relationship (41), it remains to apply [EK09, Theorem 4.11 page 192], and thus to verify its condition (4.50). The latter shrinks here to the following. Let

$$g((x, \ell), (u, f)) := x \Psi_c(u, f) e^{-ux} f^\ell + e^{-ux} \ell f^{\ell-1} \Psi_d(u, f),$$

one has to check that for any  $T > 0$ ,

$$S_T := \sup_{s, t \leq T} |g((X_s, L_s), (u_t, f_t))|,$$

is integrable, with  $(u_t, f_t) = F_t(q, r)$ . Using the inequalities

$$x e^{-ux} \leq 1/u \text{ and } \ell f^\ell = \ell e^{-\ell \ln(1/f)} \leq 1/\ln(1/f),$$

we get for all  $(x, \ell) \in \mathbb{R}_+ \times \mathbb{Z}_+$ , all  $q \in (0, \infty)$ ,  $r \in (0, 1)$  and all  $t \geq 0$

$$|g((x, \ell), (u_t, f_t))| \leq \left| \frac{\Psi_c(u_t, f_t)}{u_t} \right| + \left| \frac{\Psi_d(u_t, f_t)}{f_t \ln(1/f_t)} \right|.$$

According to Lemma 4.3,  $u_t \in (0, \infty)$  and  $f_t \in (0, 1)$  for all  $t$ . By continuity, their extrema on the compact time interval  $[0, T]$  belongs to  $(0, \infty)$  and  $(0, 1)$ , therefore  $S_T$  is bounded.  $\square$

*Proof of Theorem 2.1.* It follows by applying [EK09, Theorem 4.2 page 184].  $\square$

We establish now Proposition 2.3 and verify among other things the branching property of  $\mathbf{X}$ .

*Proof of Proposition 2.3.* We start by verifying the Feller property and exhibiting a core. Recall  $D$  the space generated by the linear combinations of the functions  $f_{q,r}$ . Since as  $t$  goes to 0,  $u_t(q,r) \rightarrow q$  and  $f_t(q,r) \rightarrow r$ , one has for any  $(x,\ell) \in \mathbb{R}_+ \times \mathbb{Z}_+$ ,  $P_t f_{q,r}(x,\ell) \xrightarrow{t \rightarrow 0} f_{q,r}(x,\ell)$ . Furthermore,  $D \subset \mathcal{D}$  and  $P_t D \subset C_0(\mathbb{R}_+ \times \mathbb{Z}_+)$ . Therefore, by density of  $D$  in  $C_0(\mathbb{R}_+ \times \mathbb{Z}_+)$ , we see that for all  $f \in C_0(\mathbb{R}_+ \times \mathbb{Z}_+)$ ,

$$P_t f \in C_0(\mathbb{R}_+ \times \mathbb{Z}_+) \text{ and } P_t f(x,\ell) \xrightarrow{t \rightarrow 0} f(x,\ell).$$

Last, since  $D$  is a subset of the domain of  $\mathcal{L}$ , dense in  $C_0(\mathbb{R}_+ \times \mathbb{Z}_+)$  and  $P_t D \subset D$ , this is a core, see e.g. [Kal02, Proposition 19.9]. The branching property (12) readily follows from (6), indeed

$$\begin{aligned} \mathbb{E}_{(x+y, n+m)}[e^{-qX_t, rL_t}] &= e^{-xu_t(q,r)} f_t(q,r)^n e^{-yu_t(q,r)} f_t(q,r)^m \\ &= \mathbb{E}_{(x,n)}[e^{-qX_t, rL_t}] \mathbb{E}_{(y,m)}[e^{-qX_t, rL_t}]. \end{aligned}$$

□

We now establish Proposition 2.5 where conditions are given for the coordinates of the bi-type branching process  $(X, L)$  to be autonomous.

*Proof of Proposition 2.5.* We only give the proof in case 1, since the two proofs are the same. We thus assume that  $\pi(dy, dk) = \nu(dy)\delta_0(dk)$  for a measure  $\nu(dy)$  on  $(0, \infty)$  such that  $\int_0^\infty (1 \wedge y^2)\nu(dx) < \infty$ . In this case, the map  $(q, r) \mapsto \Psi_c(q, r)$  does not depend on the variable  $r$  and can therefore be simply denoted by  $q \mapsto \Psi_c(q, 1)$ . By (7), the map  $t \mapsto u_t(0+, r)$  is solution of the o.d.e

$$\frac{d}{dt} u_t = -\Psi_c(u_t, 1), \quad u_0 = 0.$$

The condition  $\int_0 \frac{dx}{|\Psi_c(x, 1)|} = \infty$  is necessary and sufficient for this o.d.e. to have no solution other than the null function. Hence under this condition,  $u_t(0+, r) = 0$  for all  $t \geq 0$  and  $(L_t)$  then satisfies for any  $n \in \mathbb{Z}^+$ ,

$$\mathbb{E}_n[r^{L_t}] = f_t(0+, r)^n,$$

where  $f_t(0+, r)$  is the unique solution to the o.d.e:

$$\frac{d}{dt} f_t(0+, r) = -\Psi_d(0+, f_t(0+, r)), \quad f_0(0+, r) = r,$$

meaning that  $(L_t)$  is a discrete branching process with branching mechanism

$$\varphi : [0, 1) \ni r \rightarrow \Psi_d(0+, r) = \sum_{k \geq 1} (r^{k+1} - r)\rho(\mathbb{R}_+, \{k\}) + d(1-r) - kr.$$

□

## 5 Proof of Theorem 3.1

Let  $\psi$  be a branching mechanism, see (13). Recall the infinitesimal generator  $\mathcal{G}$  of the CSBP( $\psi$ ), see (14). We will establish in this section the algebraic intertwining relationship (23) between  $\mathcal{G}$  and  $\mathcal{H}$ . We start by identifying the triplet of the Esscher transform of  $\psi$ ,  $\psi_\lambda(\cdot) = \psi(\lambda + \cdot) - \psi(\lambda)$ , for any  $\lambda \geq 0$ , see (17).

**Lemma 5.1.** *For any  $\lambda \in [0, \infty)$ ,*

$$\psi_\lambda(q) = \frac{\sigma^2}{2} q^2 + \int_{(0, \infty)} \nu(dy) e^{-\lambda y} [e^{-qy} - 1 + qy] + \psi'(\lambda)q,$$

with

$$\psi'(\lambda) = \sigma^2 \lambda - \gamma + \int_{(0, 1)} (1 - e^{-\lambda y}) y \nu(dy) - \int_1^\infty \nu(dy) y e^{-\lambda y}.$$

*Proof.*

$$\psi_\lambda(q) := \psi(\lambda + q) - \psi(\lambda) \quad (50)$$

$$\begin{aligned} &= \sigma^2(\lambda + q)^2/2 - \gamma(\lambda + q) - (\sigma^2\lambda^2/2 - \gamma\lambda) \\ &\quad + \int_{(0,\infty)} \nu(dy) [e^{-(q+\lambda)y} - 1 + (q + \lambda)y \mathbf{1}_{(0,1)}(y)] \end{aligned} \quad (51)$$

$$= \frac{\sigma^2}{2}q^2 + \int_{(0,\infty)} \nu(dy) e^{-\lambda y} [e^{-qy} - 1 + qy] + (\lambda\sigma^2 - \gamma)q \quad (52)$$

$$+ \left( \int_{(0,\infty)} \nu(dy) y \mathbf{1}_{(0,1)}(y) (1 - e^{-\lambda y}) - \int_1^\infty ye^{-\lambda y} \nu(dy) \right) q, \quad (53)$$

$$= \frac{\sigma^2}{2}q^2 + \int_{(0,\infty)} \nu(dy) e^{-\lambda y} [e^{-qy} - 1 + qy] + \psi'(\lambda)q. \quad (54)$$

□

Recall the Poisson kernel  $K$  with parameter  $\lambda$  in (18) and the space  $\mathcal{D}$ , see (4). One can<sup>3</sup> easily check that for any  $f \in \mathcal{D}$ ,  $Kf \in D_c$ , see (15).

Recall the operator  $\mathcal{H}$  in (21)-(22) and define

$$\mathcal{R}f(x, \ell) := \begin{cases} \frac{\psi(\lambda)}{\lambda} \ell [f(x, \ell - 1) - f(x, \ell)] & , \text{ if } \lambda \geq \rho, \\ -\psi(\lambda)x [f(x, \ell + 1) - f(x, \ell)] & , \text{ if } \lambda \leq \rho. \end{cases} \quad (55)$$

Observe that the factor term in  $\mathcal{R}$  lying in front of the incremental term,  $f(x, \ell - 1) - f(x, \ell)$  or  $f(x, \ell + 1) - f(x, \ell)$ , is always non-negative. This will be important when interpreting those terms as jump rates. At this stage of the study, at which an algebraic relation is targeted, this plays however essentially no role. Indeed the following lemma shows that after Poissonization with parameter  $\lambda x$ , the form of the operator  $\mathcal{R}$  does not depend on whether  $\lambda \leq \rho$  or  $\lambda > \rho$ . This will allow us to study both cases simultaneously.

**Lemma 5.2.** *For any  $\lambda \in (0, \infty)$ ,*

$$\begin{aligned} K\mathcal{R}f(x) &= \sum_{\ell \geq 0} e^{-\lambda x} \frac{(\lambda x)^\ell}{\ell!} (-\psi(\lambda)x) [f(x, \ell + 1) - f(x, \ell)] \\ &= \sum_{\ell \geq 1} e^{-\lambda x} \frac{(\lambda x)^\ell}{\ell!} \frac{\psi(\lambda)}{\lambda} \ell [f(x, \ell - 1) - f(x, \ell)]. \end{aligned}$$

*Proof of Lemma 5.2.*

$$\begin{aligned} &\sum_{\ell \geq 0} e^{-\lambda x} \frac{(\lambda x)^\ell}{\ell!} (-x\psi(\lambda)) [f(x, \ell + 1) - f(x, \ell)] \\ &= x\psi(\lambda) \sum_{\ell \geq 0} e^{-\lambda x} \frac{(\lambda x)^\ell}{\ell!} [f(x, \ell) - f(x, \ell + 1)] \\ &= \psi(\lambda) \sum_{\ell \geq 0} \frac{1}{\lambda} e^{-\lambda x} \frac{(\lambda x)^{\ell+1}}{(\ell+1)!} (\ell+1) [f(x, \ell) - f(x, \ell + 1)] \\ &= \frac{\psi(\lambda)}{\lambda} \sum_{\ell \geq 1} e^{-\lambda x} \frac{(\lambda x)^\ell}{\ell!} \ell [f(x, \ell - 1) - f(x, \ell)]. \end{aligned}$$

□

<sup>3</sup>see the forthcoming Equations (60)-(61) for the calculation of the two first derivatives.

By linearity of the operator  $\mathcal{G}$ , we can separate the study into its local and non-local parts. We start therefore with the generator of the Feller diffusion.

### 5.1 The case of the Feller diffusion ( $\Psi(q) = \frac{\sigma^2}{2}q^2 - \gamma q$ )

We focus here on the setting of a pure diffusive CSBP process. The integral term in  $\mathcal{G}$  vanishes and the branching mechanism shrinks to  $\psi(q) = \frac{\sigma^2}{2}q^2 - \gamma q$  with  $\gamma \in \mathbb{R}$ . Notice that  $\frac{\psi(\lambda)}{\lambda} = \frac{\sigma^2}{2}\lambda - \gamma$ ,  $\psi'(\lambda) = \sigma^2\lambda - \gamma$  and  $\rho = \frac{2\gamma}{\sigma^2} \vee 0$ . One has

$$\mathcal{G}^{\psi_\lambda} f(x, \ell) = \mathcal{G}^{\psi_\lambda} f_\ell(x) = x \left[ \frac{\sigma^2}{2} f''(x, \ell) - (\sigma^2\lambda - \gamma) f'(x, \ell) \right].$$

Introduce the operator  $\mathcal{J}^L$  (the superscript  $L$  is for *local*):

$$\mathcal{J}^L f(x, \ell) := \sigma^2 \ell f'(x, \ell) + \ell \frac{\sigma^2}{2} \lambda [f(x, \ell + 1) - f(x, \ell)]. \quad (56)$$

One has

$$\mathcal{H}f(x, \ell) = \mathcal{G}^{\psi_\lambda} f(x, \ell) + \mathcal{J}^L f(x, \ell) + \mathcal{R}f(x, \ell).$$

We now establish that  $\mathcal{H}$  intertwines  $\mathcal{G}$ , namely we show that the following relationship holds:

$$\mathcal{G}Kf(x) = K\mathcal{H}f(x).$$

To prove this, we shall expand the following expression

$$\mathcal{G}Kf(x) = x \left( \frac{\sigma^2}{2} (Kf)''(x) + \gamma (Kf)'(x) \right)$$

and compare it with

$$K\mathcal{H}f(x) = \sum_{\ell \geq 0} e^{-\lambda x} \frac{(\lambda x)^\ell}{\ell!} \left[ \underbrace{x \left( \frac{\sigma^2}{2} f''(x, \ell) - (\sigma^2\lambda - \gamma) f'(x, \ell) \right)}_{\mathcal{G}^{\psi_\lambda} f(x, \ell)} \right] \quad (57)$$

$$+ \sum_{\ell \geq 0} e^{-\lambda x} \frac{(\lambda x)^\ell}{\ell!} \left( \underbrace{\ell \sigma^2 f'(x, \ell) + \ell \frac{\sigma^2}{2} \lambda [f(x, \ell + 1) - f(x, \ell)]}_{\mathcal{J}^L f(x, \ell)} \right) \quad (58)$$

$$+ \underbrace{\sum_{\ell \geq 1} e^{-\lambda x} \frac{(\lambda x)^\ell}{\ell!} \left[ \ell \left( \frac{\sigma^2}{2} \lambda - \gamma \right) \right]}_{K\mathcal{R}f(x)} (f(x, \ell - 1) - f(x, \ell)). \quad (59)$$

Expanding  $\mathcal{G}Kf(x)$  requires to compute the first two derivatives of the expression  $(Kf)(x)$ :

$$(Kf)'(x) = \lambda \sum_{\ell \geq 0} e^{-\lambda x} \left[ \frac{(\lambda x)^{\ell-1}}{(\ell-1)!} - \frac{(\lambda x)^\ell}{\ell!} \right] f(x, \ell) + \sum_{\ell \geq 0} e^{-\lambda x} \frac{(\lambda x)^\ell}{\ell!} f'(x, \ell). \quad (60)$$

$$(Kf)''(x) = \lambda^2 \sum_{\ell \geq 0} e^{-\lambda x} \left[ \frac{(\lambda x)^{\ell-2}}{(\ell-2)!} - 2 \frac{(\lambda x)^{\ell-1}}{(\ell-1)!} + \frac{(\lambda x)^\ell}{\ell!} \right] f(x, \ell) \\ + 2\lambda \sum_{\ell \geq 0} e^{-\lambda x} \left[ \frac{(\lambda x)^{\ell-1}}{(\ell-1)!} - \frac{(\lambda x)^\ell}{\ell!} \right] f'(x, \ell) + \sum_{\ell \geq 0} e^{-\lambda x} \frac{(\lambda x)^\ell}{\ell!} f''(x, \ell). \quad (61)$$

Next, the expression for  $\mathcal{G}Kf(x)$  can be splitted into three terms, that we distinguish according to the partial derivative of  $f(x, \ell)$  (zeroth, first or second) that is used:

$\mathcal{G}Kf(x)$

$$= x \frac{\sigma^2}{2} \lambda^2 \sum_{\ell \geq 0} e^{-\lambda x} \left[ \left( \frac{(\lambda x)^{\ell-2}}{(\ell-2)!} - 2 \frac{(\lambda x)^{\ell-1}}{(\ell-1)!} + \frac{(\lambda x)^\ell}{\ell!} \right) + \gamma \lambda \left( \frac{(\lambda x)^{\ell-1}}{(\ell-1)!} - \frac{(\lambda x)^\ell}{\ell!} \right) \right] f(x, \ell) \quad (62)$$

$$+ x \sum_{\ell \geq 0} e^{-\lambda x} \left[ \sigma^2 \lambda \left( \frac{(\lambda x)^{\ell-1}}{(\ell-1)!} - \frac{(\lambda x)^\ell}{\ell!} \right) + \gamma \frac{(\lambda x)^\ell}{\ell!} \right] f'(x, \ell) \quad (63)$$

$$+ x \left[ \frac{\sigma^2}{2} \sum_{\ell \geq 0} e^{-\lambda x} \frac{(\lambda x)^\ell}{\ell!} \right] f''(x, \ell). \quad (64)$$

In (64), the term in factor of  $f''(x, \ell)$  identifies with the similar term in  $K\mathcal{H}f(x)$  in (57) so there is nothing to do here.

To transform (62), the idea is to separate it into three sums and reindex them in a convenient manner. We rewrite (62) as

$$x \sum_{\ell \geq 0} e^{-\lambda x} \left[ \frac{\sigma^2}{2} \lambda^2 \left( \frac{(\lambda x)^{\ell-2}}{(\ell-2)!} - \frac{(\lambda x)^{\ell-1}}{(\ell-1)!} \right) + \frac{\sigma^2}{2} \lambda^2 \left( \frac{(\lambda x)^\ell}{\ell!} - \frac{(\lambda x)^{\ell-1}}{(\ell-1)!} \right) + \gamma \lambda \left( \frac{(\lambda x)^{\ell-1}}{(\ell-1)!} - \frac{(\lambda x)^\ell}{\ell!} \right) \right] f(x, \ell).$$

Therefore

$$\begin{aligned} (62) &= x \sum_{\ell \geq 0} e^{-\lambda x} \left[ -\frac{\sigma^2}{2} \lambda^2 \frac{(\lambda x)^\ell}{\ell!} [f(x, \ell+1) - f(x, \ell)] + \frac{\sigma^2}{2} \lambda^2 \frac{(\lambda x)^{\ell-1}}{(\ell-1)!} [f(x, \ell+1) - f(x, \ell)] \right. \\ &\quad \left. + \gamma \lambda \frac{(\lambda x)^\ell}{\ell!} [f(x, \ell+1) - f(x, \ell)] \right] \\ &= x \sum_{\ell \geq 0} e^{-\lambda x} \frac{(\lambda x)^\ell}{\ell!} \left( \gamma \lambda - \frac{\sigma^2}{2} \lambda^2 \right) [f(x, \ell+1) - f(x, \ell)] + x \frac{\sigma^2}{2} \lambda^2 \sum_{\ell \geq 0} e^{-\lambda x} \frac{(\lambda x)^{\ell-1}}{(\ell-1)!} [f(x, \ell+1) - f(x, \ell)]. \\ &= x \underbrace{\sum_{\ell \geq 0} e^{-\lambda x} \frac{(\lambda x)^\ell}{\ell!} \left( \gamma \lambda - \frac{\sigma^2}{2} \lambda^2 \right) [f(x, \ell+1) - f(x, \ell)]}_{\mathcal{KR}f(x)} + \sum_{\ell \geq 0} e^{-\lambda x} \frac{(\lambda x)^\ell}{\ell!} \frac{\sigma^2}{2} \lambda \ell [f(x, \ell+1) - f(x, \ell)]. \end{aligned}$$

In the same way:

$$\begin{aligned} (63) &= \sum_{\ell \geq 0} e^{-\lambda x} \frac{(\lambda x)^\ell}{\ell!} \left[ \sigma^2 \lambda x \left( \frac{\ell}{\lambda x} - \ell \right) + \gamma x \right] f'(x, \ell) \\ &= \sum_{\ell \geq 0} e^{-\lambda x} \frac{(\lambda x)^\ell}{\ell!} \sigma^2 \ell f'(x, \ell) + \sum_{\ell \geq 0} e^{-\lambda x} \frac{(\lambda x)^\ell}{\ell!} (\gamma - \sigma^2 \lambda) x f'(x, \ell). \end{aligned}$$

Recall  $\mathcal{J}^L$  and  $\mathcal{G}^{\psi_\lambda}$ . We get

$$\begin{aligned}
\mathcal{G}Kf(x) &= (62) + (63) + (64) \\
&= \mathcal{K}\mathcal{R}f(x) + \overbrace{\sum_{\ell \geq 0} e^{-\lambda x} \frac{(\lambda x)^\ell}{\ell!} \frac{\sigma^2}{2} \lambda \ell [f(x, \ell + 1) - f(x, \ell)] + \sum_{\ell \geq 0} e^{-\lambda x} \frac{(\lambda x)^\ell}{\ell!} \sigma^2 \ell f'(x, \ell)}^{\mathcal{K}\mathcal{J}^L f(x)} \\
&\quad + \underbrace{\sum_{\ell \geq 0} e^{-\lambda x} \frac{(\lambda x)^\ell}{\ell!} (\gamma - \sigma^2 \lambda) x f'(x, \ell) + \sum_{\ell \geq 0} e^{-\lambda x} \frac{(\lambda x)^\ell}{\ell!} \frac{\sigma^2}{2} x f''(x, \ell)}_{\mathcal{K}\mathcal{G}^{\psi_\lambda} f(x)} \\
&= K(\mathcal{R} + \mathcal{J}^L + \mathcal{G}^{\psi_\lambda})f(x) \\
&= K\mathcal{H}f(x).
\end{aligned}$$

## 5.2 The case of pure jump CSBPs

We now deal with CSBP processes with no local part and assume  $\sigma = \gamma = 0$ . Define

$$\mathcal{J}^{NL}f(x, \ell) := \ell \sum_{k \geq 0} \int_{(0, \infty)} \nu(dy) y e^{-\lambda y} \frac{(\lambda y)^k}{(k+1)!} [f(x+y, \ell+k) - f(x, \ell)]. \quad (65)$$

Notice that  $\mathcal{H} = c + \mathcal{G}^{\psi_\lambda} + \mathcal{J}^{NL} + \mathcal{R}$ , with  $cf(x, \ell) := -\kappa x f(x, \ell)$ . We are going to compute

$$\mathcal{G}Kf(x) = x \int_{(0, \infty)} \nu(dy) [Kf(x+y) - Kf(x) - y \mathbf{1}_{(0,1)}(y) (Kf)'(x)] - \kappa x Kf(x). \quad (66)$$

Plainly,  $-\kappa x Kf(x) = -\sum_{\ell=0}^{\infty} \frac{(\lambda x)^\ell}{\ell!} \kappa x f(x, \ell) = K(cf)(x)$ , so that the killing term in  $\mathcal{G}$  matches with a killing term along the continuous component of  $\mathcal{H}$ . We now study the integrand in (66). Write the Poisson kernel evaluated at  $x+y$ :

$$Kf(x+y) = \sum_{\ell \geq 0} e^{-\lambda(x+y)} \frac{(\lambda(x+y))^\ell}{\ell!} f(x+y, \ell)$$

and its derivative evaluated at  $x$ :

$$(Kf)'(x) = -\sum_{\ell \geq 0} \lambda e^{-\lambda x} \frac{(\lambda x)^\ell}{\ell!} f(x, \ell) + \sum_{\ell \geq 1} \lambda e^{-\lambda x} \frac{(\lambda x)^{\ell-1}}{(\ell-1)!} f(x, \ell) + \sum_{\ell \geq 0} e^{-\lambda x} \frac{(\lambda x)^\ell}{\ell!} f'(x, \ell).$$

The expression for  $x[Kf(x+y) - Kf(x) - y \mathbf{1}_{(0,1)}(y) (Kf)'(x)]$  then splits in three basic blocks that are:

$$x[Kf(x+y) - Kf(x) - y \mathbf{1}_{(0,1)}(y) (Kf)'(x)] \quad (67)$$

$$= \sum_{\ell \geq 0} e^{-\lambda x} x \left[ e^{-\lambda y} \frac{(\lambda(x+y))^\ell}{\ell!} f(x+y, \ell) - \frac{(\lambda x)^\ell}{\ell!} f(x, \ell) \right] \quad (68)$$

$$- xy \mathbf{1}_{(0,1)}(y) \left[ -\sum_{\ell \geq 0} \lambda e^{-\lambda x} \frac{(\lambda x)^\ell}{\ell!} f(x, \ell) + \sum_{\ell \geq 1} \lambda e^{-\lambda x} \frac{(\lambda x)^{\ell-1}}{(\ell-1)!} f(x, \ell) \right] \quad (69)$$

$$- \sum_{\ell \geq 0} e^{-\lambda x} \frac{(\lambda x)^\ell}{\ell!} xy \mathbf{1}_{(0,1)}(y) f'(x, \ell). \quad (70)$$

We expand (68) using Newton binomial formula :

$$\begin{aligned} & \sum_{\ell \geq 0} e^{-\lambda x} x \left[ e^{-\lambda y} \frac{\lambda^\ell \sum_{0 \leq k \leq \ell} \binom{\ell}{k} x^{\ell-k} y^k}{\ell!} f(x+y, \ell) - (e^{-\lambda y} + 1 - e^{-\lambda y}) \frac{(\lambda x)^\ell}{\ell!} f(x, \ell) \right] \\ &= \sum_{\ell \geq 0} e^{-\lambda x} \frac{(\lambda x)^\ell}{\ell!} x e^{-\lambda y} [f(x+y, \ell) - f(x, \ell)] \end{aligned} \quad (71)$$

$$+ \sum_{\ell \geq 1} \sum_{1 \leq k \leq \ell} e^{-\lambda x} \frac{(\lambda x)^{\ell+1-k}}{(\ell+1-k)!} (\ell+1-k)y e^{-\lambda y} \frac{(\lambda y)^{k-1}}{k!} f(x+y, \ell) \quad (72)$$

$$- \sum_{\ell \geq 0} e^{-\lambda x} \frac{(\lambda x)^\ell}{\ell!} x (1 - e^{-\lambda y}) f(x, \ell). \quad (73)$$

The second term (72) can be rewritten by a reindexation of the double sum as follows:

$$\begin{aligned} & \sum_{\ell \geq 1} \sum_{1 \leq k \leq \ell} e^{-\lambda x} \frac{(\lambda x)^{\ell+1-k}}{(\ell+1-k)!} (\ell+1-k)y e^{-\lambda y} \frac{(\lambda y)^{k-1}}{k!} f(x+y, \ell) \\ &= \sum_{\ell \geq 0, k \geq 0} e^{-\lambda x} \frac{(\lambda x)^\ell}{\ell!} \ell y e^{-\lambda y} \frac{(\lambda y)^k}{(k+1)!} f(x+y, \ell+k). \end{aligned} \quad (74)$$

Recollecting our findings so far, one has

$$x[Kf(x+y) - Kf(x) - y\mathbf{1}_{(0,1)}(y)(Kf)'(x)]$$

$$= \sum_{\ell \geq 0} e^{-\lambda x} \frac{(\lambda x)^\ell}{\ell!} x e^{-\lambda y} [f(x+y, \ell) - f(x, \ell)] \quad (75)$$

$$+ \sum_{\ell \geq 0} e^{-\lambda x} \frac{(\lambda x)^\ell}{\ell!} \ell \sum_{k \geq 0} y e^{-\lambda y} \frac{(\lambda y)^k}{(k+1)!} f(x+y, \ell+k) \quad (76)$$

$$- \sum_{\ell \geq 0} e^{-\lambda x} \frac{(\lambda x)^\ell}{\ell!} x (1 - e^{-\lambda y}) f(x, \ell) \quad (77)$$

$$- xy\mathbf{1}_{[0,1]}(y) \left( - \sum_{\ell \geq 0} \lambda e^{-\lambda x} \frac{(\lambda x)^\ell}{\ell!} f(x, \ell) + \sum_{\ell \geq 1} \lambda e^{-\lambda x} \frac{(\lambda x)^{\ell-1}}{(\ell-1)!} f(x, \ell) \right) \quad (78)$$

$$- \sum_{\ell \geq 0} e^{-\lambda x} \frac{(\lambda x)^\ell}{\ell!} xy\mathbf{1}_{[0,1]}(y) f'(x, \ell). \quad (79)$$

Reordering the terms by gathering (75) and (79), (76) with the second part of (78) and (77) with the first part, we get:

$$x[Kf(x+y) - Kf(x) - y\mathbf{1}_{(0,1)}(y)(Kf)'(x)]$$

$$= \sum_{\ell \geq 0} e^{-\lambda x} \frac{(\lambda x)^\ell}{\ell!} x e^{-\lambda y} [f(x+y, \ell) - f(x, \ell)] \quad (80)$$

$$- \sum_{\ell \geq 0} e^{-\lambda x} \frac{(\lambda x)^\ell}{\ell!} xy \mathbf{1}_{(0,1]}(y) f'(x, \ell) \quad (81)$$

$$+ \sum_{\ell \geq 0} e^{-\lambda x} \frac{(\lambda x)^\ell}{\ell!} \ell \sum_{k \geq 0} y e^{-\lambda y} \frac{(\lambda y)^k}{(k+1)!} f(x+y, \ell+k) \quad (82)$$

$$- \sum_{\ell \geq 1} \lambda e^{-\lambda x} xy \mathbf{1}_{[0,1)}(y) \frac{(\lambda x)^{\ell-1}}{(\ell-1)!} f(x, \ell) \quad (83)$$

$$- \sum_{\ell \geq 0} e^{-\lambda x} \frac{(\lambda x)^\ell}{\ell!} x (1 - e^{-\lambda y} - \lambda y \mathbf{1}_{[0,1)}(y)) f(x, \ell). \quad (84)$$

Equation (81) can be rewritten as

$$\begin{aligned} (81) &= -x \sum_{\ell \geq 0} e^{-\lambda x} \frac{(\lambda x)^\ell}{\ell!} (e^{-\lambda y} + 1 - e^{-\lambda y}) y \mathbf{1}_{(0,1]}(y) f'(x, \ell) \\ &= -x \sum_{\ell \geq 0} e^{-\lambda x} \frac{(\lambda x)^\ell}{\ell!} (e^{-\lambda y} y \mathbf{1}_{(0,1]}(y) f'(x, \ell) + (1 - e^{-\lambda y}) y \mathbf{1}_{(0,1]}(y) f'(x, \ell)). \end{aligned}$$

Integrating (80)+(81) with respect to  $\nu$  yields

$$x \sum_{\ell \geq 0} e^{-\lambda x} \frac{(\lambda x)^\ell}{\ell!} \int_{(0, \infty)} e^{-\lambda y} \nu(dy) (f(x+y, \ell) - f(x, \ell) - y \mathbf{1}_{(0,1]}(y) f'(x, \ell)) + \psi'(\lambda) f'(x, \ell) = KG^{\psi_\lambda} f(x).$$

Notice that

$$(83) = - \sum_{\ell \geq 0} e^{-\lambda x} y \mathbf{1}_{(0,1)}(y) \ell \frac{(\lambda x)^\ell}{\ell!} f(x, \ell).$$

Using the following identity

$$y e^{-\lambda y} \sum_{k \geq 0} \frac{(\lambda y)^k}{(k+1)!} = e^{-\lambda y} \frac{1}{\lambda} \sum_{k \geq 1} \frac{(\lambda y)^k}{k!} = \frac{1}{\lambda} (1 - e^{-\lambda y}).$$

one can rewrite (82)+(83) as follows:

$$(82) + (83) = \sum_{\ell \geq 0} e^{-\lambda x} \frac{(\lambda x)^\ell}{\ell!} \ell \left[ \sum_{k \geq 0} y e^{-\lambda y} \frac{(\lambda y)^k}{(k+1)!} f(x+y, \ell+k) - y \mathbf{1}_{(0,1]}(y) f(x, \ell) \right] \quad (85)$$

$$= \sum_{\ell \geq 0} e^{-\lambda x} \frac{(\lambda x)^\ell}{\ell!} \ell \sum_{k \geq 0} y e^{-\lambda y} \frac{(\lambda y)^k}{(k+1)!} (f(x+y, \ell+k) - f(x, \ell)) \quad (86)$$

$$+ \sum_{\ell \geq 0} e^{-\lambda x} \frac{(\lambda x)^\ell}{\ell!} \ell \left( \frac{1 - e^{-\lambda y}}{\lambda} - y \mathbf{1}_{(0,1]}(y) \right) f(x, \ell) \quad (87)$$

$$= (86) + x \sum_{\ell' \geq 0} e^{-\lambda x} \frac{(\lambda x)^{\ell'}}{\ell'!} [1 - e^{-\lambda y} - \lambda y \mathbf{1}_{(0,1]}(y)] f(x, \ell' + 1). \quad (88)$$

We sum (82)+(83)+(84) and obtain:

$$(82) + (83) + (84) = (86) - x \left( e^{-\lambda y} - 1 + \lambda y 1_{(0,1]}(y) \right) \sum_{\ell \geq 0} e^{-\lambda x} \frac{(\lambda x)^\ell}{\ell!} [f(x, \ell + 1) - f(x, \ell)]. \quad (89)$$

By integrating with respect to  $\nu$ , the term (86), one has

$$\sum_{\ell \geq 0} e^{-\lambda x} \frac{(\lambda x)^\ell}{\ell!} \ell \sum_{k \geq 0} \int_{(0, \infty)} \nu(dy) y e^{-\lambda y} \frac{(\lambda y)^k}{(k+1)!} [f(x+y, \ell+k) - f(x, \ell)],$$

and we recognize the cross jump term  $\mathcal{J}^{NL}f(x, \ell)$ , see (65).

By integrating with respect to  $\nu$ , the second part of the right-hand side in (89), we get

$$-x\psi(\lambda) \sum_{\ell \geq 0} e^{-\lambda x} \frac{(\lambda x)^\ell}{\ell!} [f(x, \ell + 1) - f(x, \ell)] = K\mathcal{R}f(x). \quad (90)$$

Finally, we have established

$$\mathcal{G}Kf(x) = K(c + \mathcal{G}^{\psi_\lambda} + \mathcal{J}^{NL} + \mathcal{R})f(x) = K\mathcal{H}f(x).$$

*Proof of Theorem 3.1.* Only remains to gather the diffusive part and the jump part, this is a direct consequence of linearity and the fact that  $\mathcal{J}^L + \mathcal{J}^{NL} = \mathcal{J}$ . One has indeed

$$\mathcal{H} = c + \mathcal{G}^{\psi_\lambda} + \mathcal{J}^L + \mathcal{J}^{NL} + \mathcal{R},$$

thus

$$K\mathcal{H}f = K(c + \mathcal{G}^{\psi_\lambda} + \mathcal{J}^L + \mathcal{J}^{NL} + \mathcal{R})f = \mathcal{G}Kf.$$

□

## 6 Proofs of Corollary 3.2, Theorem 3.4 and Proposition 3.7

Let  $\psi$  be a branching mechanism. Let  $\lambda > 0$  and  $(X^\lambda, L^\lambda)$  be the two-type branching process with generator  $\mathcal{H}$ , see (21) and (22). We start by studying the joint branching mechanism  $\Psi = (\Psi_c, \Psi_d)$  and then establish the skeleton decomposition, that is to say Theorem 3.4. Proposition 3.7 about the explosion will be a consequence and is proved at the end of the section. We stress to the reader that the proof of Theorem 3.4 does not appeal to Corollary 3.2 but only to Theorem 3.1 and Theorem 2.1.

### 6.1 Proof of Corollary 3.2: study of the joint branching mechanism

Notice that  $\Psi$  depends on  $\lambda$ . For any  $q \in (0, \infty)$ , define on  $[0, \infty)$ , the map  $e_q(x) := e^{-qx}$ . The identities (27) and (30):

$$\Psi_c(q, r) = \psi_\lambda(q) - \kappa \text{ if } \lambda \geq \rho \text{ and } \Psi_c(q, r) = \psi_\lambda(q) - \psi(\lambda)(r-1) - \kappa \text{ if } \lambda < \rho,$$

follows readily from the definition of the generator  $\mathcal{H}$ , see (21)-(22) and the fact that  $\mathcal{G}^{\psi_\lambda}e_q(x) = x\psi_\lambda(q)e_q(x) = -\psi_\lambda(q)\frac{d}{dq}e_q(x)$ . The extra birth term when  $\lambda < \rho$  comes from the last term in (22). The expression for  $\Psi_d$  given by (28)-(31) according whether  $\lambda \geq \rho$  or  $\lambda < \rho$  follows by definition of the measure  $\rho(dy, dk)$ . The extra death term when  $\lambda \geq \rho$  comes from the last term in (21). The identity

$$\psi(q + \lambda(1-r)) = \Psi_c(q, r) + \lambda\Psi_d(q, r). \quad (91)$$

is a consequence of the algebraic intertwining relationship (23) applied to the function  $f_{q,r}(x, \ell) = e^{-qx}r^\ell$ . Indeed, one has  $Kf_{q,r}(x) = e^{-(q+\lambda(1-r))x} = e_{q+\lambda(1-r)}(x)$ , hence on the one hand

$$\mathcal{G}Kf_{q,r}(x) = \mathcal{G}e_{q+\lambda(1-r)}(x) = x\psi(q + \lambda(1-r))e_{q+\lambda(1-r)}(x).$$

On the other hand, recall

$$\mathcal{H}f_{q,r}(x, \ell) = x\Psi_c(q, r)e^{-qx}r^\ell + e^{-qx}\ell r^{\ell-1}\Psi_d(q, r),$$

see (39), and thus

$$\begin{aligned} K\mathcal{H}f_{q,r}(x) &= \sum_{\ell \geq 0} e^{-\lambda x} \frac{(\lambda x)^\ell}{\ell!} \mathcal{H}f_{q,r}(x, \ell) \\ &= \Psi_c(q, r)x e^{-(\lambda(1-r)+q)x} + e^{-qx}\Psi_d(q, r)e^{-\lambda x} \frac{d}{dr} e^{\lambda x r} \\ &= x e_{q+\lambda(1-r)}(x) (\Psi_c(q, r) + \lambda \Psi_d(q, r)). \end{aligned}$$

Now the equality  $\mathcal{G}Kf_{q,r} = K\mathcal{H}f_{q,r}$  entails (91). The identities (29) and (32) follow readily by this relationship and (27) and (30).  $\square$

## 6.2 Proof of Theorem 3.4: intertwining of semigroups

Denote by  $(Q_t)$  the semigroup of a CSBP( $\psi$ ). Call  $(P_t)$  the semigroup of the process  $(X^\lambda, L^\lambda)$ .

**Lemma 6.1** (Intertwining of semigroups). *For all  $f \in C_0(\mathbb{R}_+ \times \mathbb{Z}_+)$ ,*

$$KP_t f(x) = Q_t Kf(x), \quad \forall t, x \geq 0. \quad (92)$$

*Proof.* Recall Proposition (2.3) and the definition of the cores  $D$  and  $D_c$  in (11) and (16) respectively. Let  $f \in D$ . Note that  $Kf \in D_c$  and the map  $(t, x) \mapsto Q_t(Kf)(x)$  is the unique solution to the *backward* Kolmogorov equation:

$$\frac{d}{dt} Q_t Kf = \mathcal{G}Q_t Kf, \quad Q_0 Kf = Kf,$$

see Subsection 2.1.2. Similarly, the semigroup  $(P_t)$  of  $\mathbf{X} = (X^\lambda, L^\lambda)$  satisfies the *backward* Kolmogorov equation:

$$\frac{d}{dt} P_t f = \mathcal{L}P_t f, \quad P_0 f = f.$$

Proposition 2.3 ensures that  $P_t f \in D \subset \mathcal{D}$ . Then the algebraic relationship given by Theorem 3.1 ensures that for any  $t \geq 0$ ,  $K\mathcal{L}P_t f = \mathcal{G}KP_t f$ . The derivative being in a uniform sense, one can interchange it with the Poisson kernel  $K$  and we see that

$$\frac{d}{dt} KP_t f = K \frac{d}{dt} P_t f = K\mathcal{L}P_t f = \mathcal{G}KP_t f.$$

Last, since  $KP_0 f = Kf$ , by uniqueness of the solution of the backward equation solved by  $Q_t$ , we have for all  $t \geq 0$ ,

$$KP_t f = Q_t Kf.$$

The fact that this holds for any  $f \in C_0(\mathbb{R}_+ \times \mathbb{Z}_+)$  follows by density.  $\square$

The intertwining of semigroups immediately extends to the kernel  $\Lambda$  defined in (34).

**Lemma 6.2** (Intertwining of semigroups). *For all  $f \in C_0(\mathbb{R}_+ \times \mathbb{Z}_+)$ ,*

$$\Lambda P_t f(x) = Q_t \Lambda f(x), \quad \forall t \in \mathbb{R}_+, x \in \mathbb{R}_+ \cup \{\infty\}. \quad (93)$$

*Proof.* From the definition of  $\Lambda$  in term of  $K$  in (34), we get:

$$\Lambda P_t f(x) = \sum_{\ell \in \mathbb{Z}_+} \Lambda(x, (x, \ell)) P_t f(x, \ell) + \Lambda(x, \Delta) P_t f(\Delta) = \mathbf{1}_{[0, \infty)}(x) KP_t f(x) + \mathbf{1}_{\{\infty\}}(x) f(\Delta)$$

while

$$Q_t \Lambda f(x) = Q_t \left( \sum_{\ell \in \mathbb{Z}_+} \Lambda(x, (x, \ell)) f(x, \ell) + \Lambda(x, \Delta) f(\Delta) \right) = \mathbf{1}_{[0, \infty)}(x) Q_t K f(x) + \mathbf{1}_{\{\infty\}}(x) f(\Delta)$$

and (93) now follows from (92).  $\square$

*Proof of Theorem 3.4.* The first statement is given by Lemma 6.1. The two other assertions follow by applying Pitman-Rogers theorem, see Theorem A. Precisely, we set  $S = (\mathbb{R}_+ \times \mathbb{Z}_+) \cup \{\Delta\}$  and  $S' = \mathbb{R}_+ \cup \{\infty\} = \bar{\mathbb{R}}_+$ , then the map  $\phi : S \rightarrow S'$  and the kernel  $\Lambda$  from  $S'$  to  $S$  are defined by:

$$\phi(x, \ell) := x, \quad \Lambda(x, (x, \ell)) := K(x, \ell) \mathbf{1}_{[0, \infty)}(x) + \delta_\Delta(x, \ell) \mathbf{1}_{\{\infty\}}(x).$$

Last  $P_t$  and  $Q_t$  are the Markov semigroups defined on  $S$  and  $S'$  before Lemma 6.1, and  $\Phi$  be defined as the operator that acts on bounded measurable functions on  $S'$  by right composition of  $\phi$ , namely  $\Phi(f) = f \circ \phi$ . With these notations,  $\Lambda \Phi$  is indeed the identity kernel on  $S'$ , since  $\Lambda(x, \phi^{-1}(x)) = 1$ . Second, Lemma 6.1 together with Remark 7.3 ensure that  $\Lambda P_t = Q_t \Lambda$  is satisfied for each  $t \geq 0$ , which in turn, composing on the right by  $\Phi$  ensures that  $Q_t$  is indeed defined from  $P_t$  by  $Q_t = K P_t \Phi$  for each  $t \geq 0$ . The assumptions of Pitman-Rogers criterion, Theorem A, are thus met. The first statement of Theorem 3.4 follows from Theorem A-(1). Its second statement follows by Theorem A-(2), see (100), with  $A = \{\ell\} \times \bar{\mathbb{R}}_+$ .  $\square$

We now study the phenomenon of explosion.

### 6.3 Proof of Proposition 3.7: explosion along skeletons

Recall that our aim is to establish that when the CSBP  $X^\lambda$  explodes continuously (i.e. not by a single jump to  $\infty$ ), it does simultaneously as any of its skeleton. We work under the assumption  $\psi(0) = -\kappa = 0$ , so that no killing is allowed. Recall that the cemetery point of the process  $(X^\lambda, L^\lambda)$  is  $\Delta = \{(x, \ell) : x + \ell = \infty\}$ . Since  $f_{q,r}(x, \ell) = e^{-qx} r^\ell \xrightarrow{(q,r) \rightarrow (0,1)} \mathbf{1}_{\mathbb{R}_+ \times \mathbb{Z}_+}(x, \ell)$  and  $K \mathbf{1}_{\mathbb{R}_+ \times \mathbb{Z}_+} = \mathbf{1}_{\mathbb{R}_+}$ , we see from (92) that

$$K(P_t \mathbf{1}_{\mathbb{R}_+ \times \mathbb{Z}_+})(x) = Q_t(\mathbf{1}_{\mathbb{R}_+})(x). \quad (94)$$

At a probabilistic level, set  $\mathbb{P} := \mathbb{P}_{x, \text{Poi}(\lambda x)}$ , define  $\zeta := \inf\{t > 0 : (X_{t-}^\lambda, L_{t-}^\lambda) = \Delta\}$  and  $\zeta_c := \inf\{t > 0 : X_{t-}^\lambda = \infty\}$ . By definition, one has  $\zeta \leq \zeta_c$  a.s. and (94) entails the equality

$$\mathbb{P}(\zeta > t) = \mathbb{P}(\zeta_c > t), \quad \forall t > 0,$$

which in turn ensures that  $\mathbb{P}(\zeta = \zeta_c) = 1$ . We now show that  $\zeta = \zeta_d := \inf\{t > 0 : L_{t-}^\lambda = \infty\}$ . Denote by  $\zeta_n^+ := \inf\{t > 0 : X_t^\lambda > n\}$ , the first passage time above  $n$ . Notice that for all  $n \geq 1$ ,  $\zeta_n^+ < \zeta_c = \lim_{n \rightarrow \infty} \uparrow \zeta_n^+$ . By Theorem 3.4, conditionally on  $\zeta_n^+ < \infty$  and  $X_{\zeta_n^+}^\lambda, L_{\zeta_n^+}^\lambda$  has law  $\text{Poi}(\lambda X_{\zeta_n^+}^\lambda)$ . Thus, for any  $\ell \geq 0$ ,

$$\mathbb{P}(L_{\zeta_n^+}^\lambda \leq \ell | X_{\zeta_n^+}^\lambda) = e^{-\lambda X_{\zeta_n^+}^\lambda} \sum_{k=0}^{\ell} \frac{(\lambda X_{\zeta_n^+}^\lambda)^k}{k!} \leq 2^\ell e^{-\frac{\lambda}{2} X_{\zeta_n^+}^\lambda} \leq 2^\ell e^{-n\lambda/2} \text{ a.s. on } \{\zeta_c < \infty\}.$$

By taking expectation and then letting  $n$  go to  $\infty$ , we see that for all  $\ell \in \mathbb{Z}_+$

$$\mathbb{P}(L_{\zeta_c-}^\lambda \leq \ell, \zeta_c < \infty) = 0.$$

We conclude that  $\mathbb{P}$ -almost surely  $\zeta_d \leq \zeta_c$ . Therefore,  $\mathbb{P}(\zeta = \zeta_d) = 1$  and the proof is achieved.  $\square$

## 7 Proofs of Proposition 3.9 and Theorem 3.12

We study in this section the skeletons when  $\lambda \geq \rho$ . Recall that in this case, they are autonomous. We start by identifying their branching mechanisms. We establish next that once rescaled by  $\lambda$  they converge in the Skorohod sense towards the CSBP( $\psi$ ).

## 7.1 Proof of Proposition 3.9: skeletons offspring distributions

First, by Corollary 3.2, we see, by letting  $q$  go to 0, in (29), that when  $\lambda \geq \rho$ ,  $\Psi_d(0, r) = \frac{\psi(\lambda(1-r)) + \kappa}{\lambda}$ . We now look for the offspring distribution. The form of  $\Psi_d$  in (28) entails

$$\Psi_d(0, r) = \sum_{k \geq 1} (r^{k+1} - r) \int_0^\infty y e^{-\lambda y} \frac{(\lambda y)^k}{(k+1)!} \nu(dy) + \frac{\sigma^2}{2} \lambda (r^2 - r) + \frac{\psi(\lambda)}{\lambda} (1 - r).$$

It remains to compute the total rate of branching. Namely

$$B := \sum_{k \geq 1} \int_0^\infty y e^{-\lambda y} \frac{(\lambda y)^k}{(k+1)!} \nu(dy) + \frac{\sigma^2}{2} \lambda + \frac{\psi(\lambda)}{\lambda} =: B_1 + \frac{\sigma^2}{2} \lambda + \frac{\psi(\lambda)}{\lambda}.$$

Plainly

$$\begin{aligned} B_1 &= \frac{1}{\lambda} \int_0^\infty e^{-\lambda y} \sum_{k=1}^\infty \frac{(\lambda y)^{k+1}}{(k+1)!} \nu(dy) = \frac{1}{\lambda} \int_0^\infty (1 - e^{-\lambda y} - \lambda y e^{-\lambda y}) \nu(dy) \\ &= \frac{1}{\lambda} \left( \int_0^1 (1 - e^{-\lambda y} - \lambda y) \nu(dy) + \int_0^1 \lambda y (1 - e^{-\lambda y}) \nu(dy) \right. \\ &\quad \left. + \int_1^\infty (1 - e^{-\lambda y}) \nu(dy) - \int_1^\infty \lambda y e^{-\lambda y} \nu(dy) \right). \end{aligned} \quad (95)$$

One has

$$\frac{\psi(\lambda)}{\lambda} = \frac{\sigma^2}{2} \lambda - \gamma + \frac{1}{\lambda} \left( \int_0^1 (e^{-\lambda y} - 1 + \lambda y) \nu(dy) + \int_1^\infty (e^{-\lambda y} - 1) \nu(dy) \right). \quad (96)$$

Thus, adding (95) and (96) and canceling the common terms, we obtain

$$B = B_1 + \frac{\sigma^2}{2} \lambda + \frac{\psi(\lambda)}{\lambda} = \sigma^2 \lambda - \gamma + \int_0^1 y (1 - e^{-\lambda y}) \nu(dy) - \int_1^\infty y e^{-\lambda y} \nu(dy) = \psi'(\lambda).$$

□

We now establish Theorem 3.12.

## 7.2 Proof of Theorem 3.12: scaling limits of skeletons

Denote by  $[0, \infty]$ , the extended half-line endowed with the metric  $d(x, y) := |e^{-x} - e^{-y}|$ , with convention  $e^{-\infty} = 0$ . Call  $\mathbb{D}$  the Skorohod space of  $[0, \infty]$ -valued càdlàg paths. Denote by  $\mathbf{P}_x$  the probability law on  $\mathbb{D}$  associated to the skeleton process  $(L_t^\lambda, t \geq 0)$  issued from a Poisson number of individuals:  $L_0^\lambda \stackrel{\text{law}}{=} \text{Poi}(\lambda x)$ .

**Lemma 7.1.** *For all  $q \in (0, \infty)$ ,*

$$\mathbf{E}_x \left[ e^{-q L_t^\lambda / \lambda} \right] \xrightarrow{\lambda \rightarrow \infty} e^{-x u_t(q)}, \text{ uniformly in } x \in [0, \infty).$$

*Proof.* Recall  $\varphi_\lambda$  in (35). Define  $f_t^\lambda(q) := \mathbb{E} \left( e^{-q L_t^\lambda / \lambda} | L_0^\lambda = 1 \right)$ . One has

$$\begin{aligned} \mathbf{E}_x \left[ e^{-q L_t^\lambda / \lambda} \right] &= \sum_{k=0}^\infty \frac{(\lambda x)^k}{k!} e^{-\lambda x} \mathbb{E} \left[ e^{-q L_t^\lambda / \lambda} | L_0^\lambda = k \right] \\ &= \sum_{k=0}^\infty \frac{(\lambda x f_t^\lambda(e^{-q/\lambda}))^k}{k!} e^{-\lambda x} \\ &= e^{-\lambda x} \left( 1 - f_t^\lambda(e^{-q/\lambda}) \right). \end{aligned} \quad (97)$$

The map  $t \mapsto v_t^\lambda(q) := \lambda(1 - f_t^\lambda(e^{-q/\lambda}))$  satisfies the o.d.e

$$\begin{aligned}\frac{d}{dt}v_t^\lambda(q) &= -\lambda\varphi_\lambda(f_t^\lambda(e^{-q/\lambda})) = -\psi(v_t^\lambda(q)) \\ v_0^\lambda(q) &= \lambda(1 - e^{-q/\lambda}).\end{aligned}$$

Therefore, by uniqueness,

$$v_t^\lambda(q) = u_t(\lambda(1 - e^{-q/\lambda})), \text{ for all } t \geq 0.$$

By continuity of  $\eta \mapsto u_t(\eta)$ , we see that  $v_t^\lambda(q) \xrightarrow{\lambda \rightarrow \infty} u_t(q)$ . Thus, for any  $x \geq 0$ ,

$$\mathbf{E}_x \left[ e^{-qL_t^\lambda/\lambda} \right] \xrightarrow{\lambda \rightarrow \infty} e^{-xu_t(q)}.$$

It remains to establish the uniform convergence, and, as we shall see, its order can be determined as well:

$$\begin{aligned}\sup_{x \geq 0} \left| \mathbf{E}_x \left[ e^{-qL_t^\lambda/\lambda} \right] - e^{-xu_t(q)} \right| &= \sup_{x \geq 0} \left| e^{-xu_t(\lambda(1 - e^{-q/\lambda}))} - e^{-xu_t(q)} \right| \\ &= \sup_{x \geq 0} e^{-xu_t(q)} \left| e^{-x(u_t(\lambda(1 - e^{-q/\lambda})) - u_t(q))} - 1 \right| \\ &\leq \sup_{x \geq 0} x e^{-xu_t(q)} \left| u_t(\lambda(1 - e^{-q/\lambda})) - u_t(q) \right| \\ &= \frac{1}{u_t(q)} \left| u_t(\lambda(1 - e^{-q/\lambda})) - u_t(q) \right| \\ &\leq \frac{1}{u_t(q)} \frac{\partial}{\partial q} u_t(q) \left( q - \lambda(1 - e^{-q/\lambda}) \right) = o(1/\lambda),\end{aligned}$$

where  $o(1/\lambda)$  is a positive function  $g$  such that  $\lim_{\lambda \rightarrow \infty} \lambda g(\lambda) = 0$ . □

Denote by  $\mathbf{Q}_t^{1/\lambda}$  the semigroup of  $(L_t^\lambda/\lambda, t \geq 0)$  under the probability laws  $(\mathbf{P}_x, x \in [0, \infty))$ . Recall  $Q_t$  the semigroup of the CSBP( $\psi$ ). We denote below the supremum norm by  $\|\cdot\|_\infty$ .

**Lemma 7.2.** *For any  $f \in C_0(\bar{\mathbb{R}}_+)$ ,*

$$\|\mathbf{Q}_t^{1/\lambda} f - Q_t f\|_\infty \xrightarrow{\lambda \rightarrow \infty} 0.$$

*Proof.* Let  $f \in C_0(\bar{\mathbb{R}}_+)$  and  $\epsilon > 0$ . By density, there exists  $g$  in  $D_c$ , the linear span of  $\{e_q(\cdot), q > 0\}$ , such that  $\|f - g\|_\infty \leq \epsilon/4$ . Recall that  $Q_t e_q(x) = e^{-xu_t(q)}$ . By Lemma 7.1, one can choose  $\lambda$  large enough so that  $\|\mathbf{Q}_t^{1/\lambda} g - Q_t g\|_\infty \leq \epsilon/2$  and one has

$$\begin{aligned}\|\mathbf{Q}_t^{1/\lambda} f - Q_t f\|_\infty &= \|\mathbf{Q}_t^{1/\lambda} f - \mathbf{Q}_t^{1/\lambda} g + \mathbf{Q}_t^{1/\lambda} g + Q_t g - Q_t f - Q_t g\|_\infty \\ &\leq 2\|f - g\|_\infty + \|\mathbf{Q}_t^{1/\lambda} g - Q_t g\|_\infty \leq \epsilon.\end{aligned}$$

□

*Proof of Theorem 3.12.* This is obtained by combining Lemma 7.2 and [EK09, Theorem 2.5 p.167], which states that uniform convergence of a family of Feller semigroups, together with the convergence of the initial laws, entails the convergence in Skorohod's sense of the associated processes. □

## Appendix

Let  $(S, \mathcal{S})$ ,  $(S', \mathcal{S}')$  be two measurable spaces, with their respective set of bounded and measurable functions denoted by  $b\mathcal{S}$  and  $b\mathcal{S}'$ . Let  $(P_t)_{t \geq 0}$  be a Markov semigroup on  $(S, \mathcal{S})$ , and  $\phi : S \rightarrow S'$  be measurable. The action of  $\phi$  by right-composition defines an operator  $\Phi : b\mathcal{S}' \rightarrow b\mathcal{S}$ ,  $f \mapsto \Phi(f) := f \circ \phi$ . Denote by  $\mathbf{X}$  the Markov process with semigroup  $(P_t)$  and denote its image by the mapping  $\phi$ ,  $\phi(\mathbf{X}) := (\phi(\mathbf{X}_t), t \geq 0)$ . Call  $\mathcal{F}_t^{\phi(\mathbf{X}), 0} = \sigma(\phi(\mathbf{X}_s), 0 \leq s \leq t)$  the natural filtration generated by  $\phi(\mathbf{X})$ . Let  $\mathcal{F}_t^{\phi(\mathbf{X})}$  be its usual augmentation.

**Theorem A** (Pitman-Rogers criterion, Theorem 2 in [PR81]). *Suppose there is a Markov kernel  $\Lambda$  from  $(S', \mathcal{S}')$  to  $(S, \mathcal{S})$  such that*

- $\Lambda\Phi = I$ , the identity kernel on  $S'$ .
- for each  $t \geq 0$ , the Markov kernel  $Q_t := \Lambda P_t \Phi$  from  $S'$  to  $S'$  satisfies the identity

$$\Lambda P_t = Q_t \Lambda. \quad (98)$$

1. If  $\mathbf{X}$  has initial distribution  $\Lambda(y, \cdot)$ , with  $y \in S'$ , then for each  $t \geq 0$ ,  $A \in \mathcal{S}$ ,

$$\mathbb{P}(\mathbf{X}_t \in A | \mathcal{F}_t^{\phi(\mathbf{X}), 0}) = \Lambda(\phi(\mathbf{X}_t), A) \quad a.s., \quad (99)$$

holds, and  $\phi(\mathbf{X})$  is Markov with starting state  $y$  and transition semigroup  $(Q_t)$ . Additionally, (99) again holds with  $\mathcal{F}_t^{\phi(\mathbf{X}), 0}$  replaced by  $\mathcal{F}_t^{\phi(\mathbf{X})}$ .

2. If  $\mathbf{X}$  has a.s. càdlàg sample paths,  $\phi(\mathbf{X})$  is Feller and  $x \mapsto \Lambda(x, \cdot)$  is weakly continuous, then for each  $\mathcal{F}_T^{\phi(\mathbf{X})}$  stopping time  $T$ , it holds:

$$\mathbb{P}(\mathbf{X}_T \in A | \mathcal{F}_T^{\phi(\mathbf{X})}) = \Lambda(\phi(\mathbf{X}_T), A) \quad a.s. \text{ on the event } \{T < \infty\}. \quad (100)$$

Only the second statement in point (1) and point (2) are formally new. We give a complete proof of the theorem for the benefit of the reader. Regarding the second assumption, we stress that *any* Markov kernel  $(Q_t)$  satisfying the intertwining relationship (98) necessarily satisfies, under the assumption  $\Lambda\Phi = I$ , that  $Q_t = \Lambda P_t \Phi$ .

*Proof.* We first establish (1). Let  $y \in S'$ . Given  $f \in b\mathcal{S}'$  and  $g \in b\mathcal{S}$ , and using that  $f(\phi(x)) = f(y)$ ,  $\Lambda(y, \cdot)$ -a.s., we get (6) as follows

$$\Lambda(\Phi f)g(y) = \int_S \Lambda(y, dx) f(\phi(x))g(x) = \int_S \Lambda(y, dx) f(y)g(x) = (f\Lambda g)(y).$$

Now, for  $y \in S'$ , the quantity  $\Lambda P_t(\Phi f)g$  evaluated at  $y$  has to be interpreted as  $\Lambda P_t(\Phi f)g(y) = \int_S \Lambda(y, dx) \int_S P_t(x, dx')(f(\phi(x'))g(x'))$  and, using that  $\Lambda P_t = Q_t \Lambda$ , we arrive at the following equality, which corresponds to Eq (7) p. 574 in [PR81]:

$$\begin{aligned} (\Lambda P_t(\Phi f)g)(y) &= (\Lambda P_t)(f(\phi)g)(y) \\ &= (Q_t \Lambda)(f(\phi)g)(y) \\ &= Q_t(\Lambda(f(\phi)g))(y) \\ &= Q_t(\Lambda((\Phi f)g))(y) \\ &= Q_t(f\Lambda g)(y). \end{aligned}$$

As for the induction step, Eq (8), setting  $h(x) = P_{t_n}(f_n \circ \phi \cdot g)(x)$ , and assuming the property holds at step  $n - 1$ , we get:

$$\begin{aligned} \Lambda P_{t_1}(\Phi f_1)P_{t_2}(\Phi f_2) \dots P_{t_n}(\Phi f_n)g &= \Lambda P_{t_1}(\Phi f_1) \dots P_{t_{n-1}}(\Phi f_{n-1})h \\ &= Q_{t_1} f_1 \dots Q_{t_{n-1}} f_{n-1} \Lambda h \\ &= Q_{t_1} f_1 \dots Q_{t_{n-1}} f_{n-1} \Lambda P_{t_n}(f_n \circ \phi \cdot g) \\ &= Q_{t_1} f_1 \dots Q_{t_{n-1}} f_{n-1} \Lambda P_{t_n}(\Phi(f_n)g) \\ &= Q_{t_1} f_1 \dots Q_{t_{n-1}} f_{n-1} Q_{t_n}(f_n) \Lambda g \end{aligned}$$

Now, writing  $\mathbb{E}_{\Lambda(y, \cdot)}$  for  $\int_x \Lambda(y, dx) \mathbb{E}_x$ , we get:

$$\begin{aligned} \mathbb{E}_{\Lambda(y, \cdot)}[f_1(\phi(\mathbf{X}_{t_1}))f_2(\phi(\mathbf{X}_{t_1+t_2}))f_{n+1}(\phi(\mathbf{X}_{t_1+\dots+t_n}))g(\mathbf{X}_{t_1+\dots+t_n})] &= (\Lambda P_{t_1}(\Phi f_1)P_{t_2}(\Phi f_2)P_{t_n}(\Phi f_n)g)(y) \\ &= (Q_{t_1}f_1 \dots Q_{t_{n-1}}f_{n-1}Q_{t_n}f_n \Lambda g)(y) \end{aligned}$$

implies, with  $g = 1$ , that  $\phi \circ X$  starting from initial measure  $\Lambda(y, \cdot)$  is Markov with semi-group  $Q_t$ , while the identity:

$$\begin{aligned} &\mathbb{E}_{\Lambda(y, \cdot)}\left[f_1(\phi(\mathbf{X}_{t_1}))f_2(\phi(\mathbf{X}_{t_1+t_2})) \dots f_n(\phi(\mathbf{X}_{t_1+\dots+t_n}))g(\mathbf{X}_{t_1+\dots+t_n})\right] \\ &= \mathbb{E}_{\Lambda(y, \cdot)}\left[f_1(\phi(\mathbf{X}_{t_1}))f_2(\phi(\mathbf{X}_{t_1+t_2})) \dots f_n(\phi(\mathbf{X}_{t_1+\dots+t_n})) \int \Lambda(\mathbf{X}_{t_1+\dots+t_n}, dx)g(x)\right] \end{aligned}$$

and this in turn gives by a monotone class theorem that for each bounded or positive random variable  $Z$ , measurable with respect to  $\mathcal{F}_t^{\phi(\mathbf{X}), 0}$ ,

$$\mathbb{E}_{\Lambda(y, \cdot)}\left[Zg(\mathbf{X}_t)\right] = \mathbb{E}_{\Lambda(y, \cdot)}\left[Z \int \Lambda(\mathbf{X}_t, dx)g(x)\right].$$

This entails that the conditional distribution of  $\mathbf{X}_t$  given  $\mathcal{F}_t^{\phi(\mathbf{X}), 0}$  is  $\Lambda(\mathbf{X}_t, dx)$ . Let now  $Z$  be a positive  $\mathcal{F}_t^{\phi(\mathbf{X})}$ -measurable random variable. By definition of the augmented filtration, there are two  $\mathcal{F}_t^{\phi(\mathbf{X}), 0}$ -measurable random variables  $Z', Z''$  such that  $Z' \leq Z \leq Z''$  and  $\mathbb{P}_{\Lambda(y, \cdot)}(Z'' - Z' > 0) = 0$ . Hence,  $\mathbb{E}_{\Lambda(y, \cdot)}[\mathbf{1}_{\{Z'' - Z' > 0\}}g(\mathbf{X}_t)] = 0$ . With no loss of generality, assuming  $g$  non-negative, one has

$$\mathbb{E}_{\Lambda(y, \cdot)}[Z'g(\mathbf{X}_t)] \leq \mathbb{E}_{\Lambda(y, \cdot)}[Zg(\mathbf{X}_t)] \leq \mathbb{E}_{\Lambda(y, \cdot)}[Z''g(\mathbf{X}_t)].$$

The lower and upper bounds are  $\mathbb{P}_{\Lambda(y, \cdot)}$ -a.s. equal to  $\int \Lambda(X_t, dx)g(x)$ , which ends the proof of (1).

For statement (2). We start by considering a stopping time  $T$  taking values in a countable set  $D$ . One has plainly,

$$\begin{aligned} \mathbf{1}_{\{T < \infty\}} \mathbb{E}_{\Lambda(y, \cdot)}[g(\mathbf{X}_T) | \mathcal{F}_T^{\Phi(\mathbf{X})}] &= \mathbf{1}_{\{T < \infty\}} \sum_{d \in D} \mathbf{1}_{\{T=d\}} \mathbb{E}_{\Lambda(y, \cdot)}[g(\mathbf{X}_d) | \mathcal{F}_d^{\Phi(\mathbf{X})}] \\ &= \mathbf{1}_{\{T < \infty\}} \sum_{d \in D} \mathbf{1}_{\{T=d\}} \int \Lambda(\mathbf{X}_d, dx)g(x) \\ &= \mathbf{1}_{\{T < \infty\}} \int \Lambda(\mathbf{X}_T, dx)g(x). \end{aligned}$$

Setting  $T_n := ((2^n T) + 1)/2^n$  defines a sequence of stopping times taking their values in a countable set and decreasing to  $T$ . By assumption,  $\mathbf{X}$  has càdlàg sample paths, hence  $\mathbf{X}_{T_n} \xrightarrow[n \rightarrow \infty]{} \mathbf{X}_T$  a.s. on  $\{T < \infty\}$ . Assume  $g$  continuous bounded. Then by combining [RY05, Corollary (2.4), Chap. II] and the weak continuity of the kernel  $y \mapsto \Lambda(y, \cdot)$ , one gets:

$$\mathbf{1}_{\{T < \infty\}} \mathbb{E}_{\Lambda(y, \cdot)}[g(\mathbf{X}_T) | \bigcap_{n \geq 1} \mathcal{F}_{T_n}^{\Phi(\mathbf{X})}] = \lim_{n \rightarrow \infty} \mathbf{1}_{\{T_n < \infty\}} \int \Lambda(\mathbf{X}_{T_n}, dx)g(x) = \mathbf{1}_{\{T < \infty\}} \int \Lambda(\mathbf{X}_T, dx)g(x).$$

Since by assumption  $\phi(\mathbf{X})$  is Feller, the usual augmented filtration is right-continuous, see [RY05, Proposition 2.10, p. 93], and

$$\bigcap_{n \geq 1} \mathcal{F}_{T_n}^{\Phi(\mathbf{X})} = \mathcal{F}_{T^+}^{\Phi(\mathbf{X})} = \mathcal{F}_T^{\Phi(\mathbf{X})},$$

see e.g. [RY05, Exercice 4.17, Chap.I] and finally (100) holds by a standard argument, see e.g. Billingsley's book [Bil99, Theorem 1.2]. □

*Remark 7.3* (Restriction to compactly supported continuous functions). The identity (98) is an identity between Markov kernels on  $S'$ , or equivalently between operators on the class  $bS'$  of bounded measurable functions on  $S'$ . Under the assumption that  $S'$  is locally compact and separable, it is enough to check that the two operators coincide on the class of compactly supported continuous functions by the uniqueness part of Riesz theorem, see Rudin [Rud87, Theorem 2.14 p.40].

*Remark 7.4* (Submarkovian intertwining). The setting where  $(P_t)$  is a sub-Markovian semigroup only can be recasted in the Markovian setting by the adjunction of cemetery points as follows. Precisely, assume that all the assumptions of Theorem A hold except for the fact that  $(P_t)$  and consequently also  $(Q_t)$  are sub-Markovian. We extend  $\phi$  by requiring  $\bar{\phi} : \bar{S} := S \cup \{\delta\} \rightarrow \bar{S}' := S' \cup \{\delta\}$  to satisfy  $\bar{\phi}(\delta) = \delta'$ , and extend  $\Lambda$  to  $\bar{\Lambda} : \bar{S}' \rightarrow \bar{S}$  by requiring  $\bar{\Lambda}(\delta', \{\delta\}) = 1$ . The definitions of  $P_t$  and  $Q_t$  are accordingly modified to accommodate functions defined on  $\delta$  and  $\delta'$  by setting  $\bar{P}_t f(\delta) = f(\delta)$  for  $f \in \bar{S}$  and  $\bar{Q}_t f(\delta') = f(\delta')$  for  $f \in \bar{S}'$ . Then one easily checks that the three identities

$$\bar{\Lambda}\bar{\Phi} = I, \quad \bar{Q}_t = \bar{\Lambda}\bar{P}_t\bar{\Phi}, \quad \bar{\Lambda}\bar{P}_t = \bar{Q}_t\bar{\Lambda}$$

again hold (for the obvious definition of  $\bar{\Phi}$  from  $\bar{\phi}$ , and as consequence, the identity (99) again holds for  $A \in \bar{S}$  with the extended kernel  $\bar{\Lambda}$ ). Last,  $\bar{\phi} \circ \bar{X}$  is Markov on  $\bar{S}'$  with semi-group  $\bar{Q}_t$ .

**Acknowledgements.** Clément Foucart is supported by the European Union (ERC, SINGER, 101054787). Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them.

## References

- [ACPM24] Marc Arnaudon, Koléhè Coulibaly-Pasquier, and Laurent Miclo. Couplings of Brownian motions with set-valued dual processes on Riemannian manifolds. *J. Éc. Polytech., Math.*, 11:473–522, 2024.
- [AD09] Romain Abraham and Jean-François Delmas. Changing the branching mechanism of a continuous state branching process using immigration. *Ann. Inst. Henri Poincaré, Probab. Stat.*, 45(1):226–238, 2009.
- [AD12] Romain Abraham and Jean-François Delmas. A continuum-tree-valued Markov process. *Ann. Probab.*, 40(3):1167–1211, 2012.
- [ADV10] Romain Abraham, Jean-François Delmas, and Guillaume Voisin. Pruning a Lévy continuum random tree. *Electron. J. Probab.*, 15:1429–1473, 2010. Id/No 46.
- [AN04] Krishna B. Athreya and Peter E. Ney. *Branching processes*. Mineola, NY: Dover Publications, reprint of the 1972 original edition, 2004.
- [BFM08] Jean Bertoin, Joaquin Fontbona, and Servet Martínez. On prolific individuals in a supercritical continuous-state branching process. *Journal of applied probability*, 45(3):714–726, 2008.
- [Bil99] Patrick Billingsley. *Convergence of probability measures*. Wiley Ser. Probab. Stat. Chichester: Wiley, 2nd ed. edition, 1999.
- [BLP15] Máttyás Barczy, Zenghu Li, and Gyula Pap. Stochastic differential equation with jumps for multi-type continuous state and continuous time branching processes with immigration. *ALEA, Lat. Am. J. Probab. Math. Stat.*, 12(1):129–169, 2015.
- [BR78] Garrett Birkhoff and Gian-Carlo Rota. *Ordinary differential equations*. 3rd ed. New York etc.: John Wiley & Sons. XI., 1978.
- [CGB17] Maria Emilia Caballero, José Luis Pérez Garmendia, and Gerónimo Uribe Bravo. Affine processes on  $\mathbb{R}_+^m \times \mathbb{R}^n$  and multiparameter time changes. *Ann. Inst. Henri Poincaré, Probab. Stat.*, 53(3):1280–1304, 2017.

- [CPY98] Philippe Carmona, Frédérique Petit, and Marc Yor. Beta-gamma random variables and intertwining relations between certain markov processes. *Revista Matemática Iberoamericana*, 14(2):311–367, 1998.
- [DF90] Persi Diaconis and James Allen Fill. Strong stationary times via a new form of duality. *Ann. Probab.*, 18(4):1483–1522, 1990.
- [DFS03] Darrell Duffie, Damir Filipović, and Walter Schachermayer. Affine processes and applications in finance. *The Annals of Applied Probability*, 13(3):984–1053, 2003.
- [DW07] Thomas Duquesne and Matthias Winkel. Growth of lévy trees. *Probability Theory and Related Fields*, 139(3):313–371, 2007.
- [DW19] Thomas Duquesne and Matthias Winkel. Hereditary tree growth and Lévy forests. *Stochastic Processes Appl.*, 129(10):3690–3747, 2019.
- [EK09] Stewart N Ethier and Thomas G Kurtz. *Markov processes: characterization and convergence*. John Wiley & Sons, 2009.
- [EK19] Alison M. Etheridge and Thomas G. Kurtz. Genealogical constructions of population models. *Ann. Probab.*, 47(4):1827–1910, 2019.
- [EKW15] Maren Eckhoff, Andreas E. Kyprianou, and Matthias Winkel. Spines, skeletons and the strong law of large numbers for superdiffusions. *Ann. Probab.*, 43(5):2545–2610, 2015.
- [FFK19] Dorottya Fekete, Joaquin Fontbona, and Andreas E Kyprianou. Skeletal stochastic differential equations for continuous-state branching processes. *Journal of Applied Probability*, 56(4):1122–1150, 2019.
- [FFK20] Dorottya Fekete, Joaquin Fontbona, and Andreas E. Kyprianou. Skeletal stochastic differential equations for superprocesses. *J. Appl. Probab.*, 57(4):1111–1134, 2020.
- [FJRW24] Simone Floreani, Sabine Jansen, Frank Redig, and Stefan Wagner. Intertwining and duality for consistent Markov processes. *Electron. J. Probab.*, 29:34, 2024. Id/No 67.
- [FMM19] Clément Foucart, Chunhua Ma, and Bastien Mallein. Coalescences in continuous-state branching processes. *Electron. J. Probab.*, 24:52, 2019. Id/No 103.
- [Gre74] D. R. Grey. Asymptotic behaviour of continuous time, continuous state-space branching processes. *J. Appl. Probab.*, 11:669–677, 1974.
- [GS83] Iosif I. Gihman and Anatoli. V. Skorohod. *The theory of stochastic processes II. Transl. from the Russian by S. Kotz. Repr. of the 1st ed. 1975*, volume 218 of *Grundlehren Math. Wiss.* Springer, Cham, 1983.
- [Har48] Theodore E. Harris. Branching processes. *Ann. Math. Stat.*, 19:474–494, 1948.
- [Har02] Theodore E. Harris. *The theory of branching processes*. Mineola, NY: Dover Publications, corrected reprint of the 1963 original edition, 2002.
- [Hel78] Inge S. Helland. Continuity of a class of random time transformations. *Stochastic Processes and their Applications*, 7(1):79–99, 1978.
- [Jir58] Miloslav Jirina. Stochastic branching processes with continuous state space. *Czech. Math. J.*, 8:292–312, 1958.
- [JL23] Samuel G. G. Johnston and Amaury Lambert. The coalescent structure of uniform and Poisson samples from multitype branching processes. *Ann. Appl. Probab.*, 33(6A):4820–4857, 2023.
- [Kal02] Olav Kallenberg. *Foundations of modern probability*. Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2002.

- [KN11] Thomas G. Kurtz and Giovanna Nappo. The filtered Martingale problem. In *The Oxford handbook of nonlinear filtering*, pages 129–165. Oxford: Oxford University Press, 2011.
- [KO88] Thomas G. Kurtz and Daniel L. Ocone. Unique characterization of conditional distributions in nonlinear filtering. *Ann. Probab.*, 16(1):80–107, 1988.
- [Kur98] Thomas G. Kurtz. Martingale problems for conditional distributions of Markov processes. *Electron. J. Probab.*, 3:1–29, 1998. Id/No 9.
- [Kur18] Thomas G. Kurtz. Genealogical constructions and asymptotics for continuous-time Markov and continuous-state branching processes. *Adv. Appl. Probab.*, 50(A):197–209, 2018.
- [Kyp14] Andreas E. Kyprianou. *Fluctuations of Lévy processes with applications*. Universitext. Springer, Heidelberg, second edition, 2014. Introductory lectures.
- [Lam67a] John Lamperti. Continuous state branching processes. *Bull. Amer. Math. Soc.*, 73:382–386, 1967.
- [Lam67b] John Lamperti. The limit of a sequence of branching processes. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 7:271–288, 1967.
- [LG99] Jean-François Le Gall. *Spatial branching processes, random snakes and partial differential equations*. Springer Science & Business Media, 1999.
- [Li22] Zenghu Li. *Measure-valued branching Markov processes*, volume 103 of *Probab. Theory Stoch. Model.* Berlin: Springer, 2nd edition edition, 2022.
- [LUB24] Amaury Lambert and Gerónimo Uribe Bravo. Totally ordered measured trees and splitting trees with infinite variation. II: Prolific skeleton decomposition. *ALEA, Lat. Am. J. Probab. Math. Stat.*, 21(2):1275–1307, 2024.
- [Mic20] Laurent Miclo. On the construction of measure-valued dual processes. *Electron. J. Probab.*, 25:64, 2020. Id/No 6.
- [PR81] Jim Pitman and Chris Rogers. Markov functions. *The Annals of Probability*, pages 573–582, 1981.
- [PS13] Soumik Pal and Mykhaylo Shkolnikov. Intertwining diffusions and wave equations. Preprint, arXiv:1306.0857 [math.PR], 2013.
- [Rud76] Walter Rudin. Principles of mathematical analysis. 3rd ed. International Series in Pure and Applied Mathematics. Düsseldorf etc.: McGraw-Hill Book Company. X, 342 p. DM 47.80 (1976)., 1976.
- [Rud87] Walter Rudin. *Real and complex analysis*. New York, NY: McGraw-Hill, 3rd ed. edition, 1987.
- [RY05] Daniel Revuz and Marc Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren Math. Wiss.* Berlin: Springer, 3rd ed., 3rd. corrected printing edition, 2005.
- [Sil68] Martin L. Silverstein. A new approach to local times. *J. Math. Mech.*, 17:1023–1054, 1968.
- [Wat69] Shinzo Watanabe. On two dimensional Markov processes with branching property. *Transactions of the American Mathematical Society*, 136:447–466, 1969.