MULTITYPE BRANCHING PROCESSES OBSERVING PARTICLES OF A GIVEN TYPE

CLAUDIA CECI,* *Università di Chieti*

ANNA GERARDI,** *Università dell'Aquila*

Abstract

A multitype branching process is presented in the framework of marked trees and its structure is studied by applying the strong branching property. In particular, the Markov property and the expression for the generator are derived for the process whose components are the numbers of particles of each type. The filtering of the whole population, observing the number of particles of a given type, is discussed. Weak uniqueness for the filtering equation and a recursive structure for the linearized filtering equation are proved under a suitable assumption on the reproduction law.

Keywords: Branching process; marked tree; filtering

2000 Mathematics Subject Classification: Primary 60J80
Secondary 60J75; 93E11

1. Introduction

Branching processes are usually used to model the behaviour of populations of 'living' particles [1] and have many applications in the physical and biological sciences. Multitype branching processes are a natural generalization of branching processes, allowing particles to be divided into a finite number of types and assuming that particles of different types have different probabilistic behaviours. These processes also arise in a variety of biological and physical applications: they could represent genetic types in an animal population; mutant types in a bacterial population; electrons, photons, and nucleons in a cosmic ray cascade; and so on.

To describe a multitype branching process, we assume that every particle lives an exponentially distributed time with the distribution's parameter depending on the particle's type. At its death, each particle splits into a random number of offspring according to a given reproduction law that also depends on particle type. Offspring particles have independent behaviours.

For the sake of simplicity, we consider two-type branching processes. To represent this model, in the first section of the paper, we choose a canonical space that is a modification of the space of trees marked by lifetimes [8], [14]. On this space we define two probability measures, labelled by the type of the ancestor, and we prove the branching property in a form that agrees with the characteristics of our model. This construction parallels that in [14], with some nontrivial modifications related to the dependence of the probabilistic behaviour of each particle on its type.

In the second section, by using the strong version of the branching property, we are able to describe the dynamics of the pair (X_t, Y_t), where, at time t, X_t is the number of type-1 particles

and Y_t is the number of type-2 particles. In particular, we prove that (X_t, Y_t) is a pure-jump Markov process and we find the structure of its generator.

When the behaviour of the whole population cannot be directly observed, and it is possible or convenient only to observe the behaviour of a subpopulation, the problem of obtaining information on the whole population, given the partial observation, has to be studied. In [5] a homogeneous population was considered, and the observed subpopulation consisted of the descendants of a given particle. In this paper, we deal with a heterogeneous population and we observe only the cardinality of the subpopulation consisting of type-2 particles. This problem reduces to the classical filtering problem of finding the conditional law of the cardinality of the whole population or, equivalently, the law of the process X_t given the past history of the process Y_t, when both the state process and the observation process are pure-jump processes with common jump times and unbounded jump intensities. Furthermore, the pair (X_t, Y_t) is a Markov process. The third section of this paper is devoted to the discussion of this filtering problem.

The required conditional law is a probability-measure-valued process, and it is a solution of the filtering equation known as the Kushner–Stratonovich equation [2]. Since Y_t is not a counting process, by following a method proposed in [3] and already used in [5], we introduce a multivariate point process whose past history coincides with the past history of the observation process. Thus, we are able to write down the filtering equation and then deal with the problem of establishing its uniqueness. To this end, we introduce the filtered martingale problem [13]. We first prove that any solution of the Kushner–Stratonovich equation provides a solution for the filtered martingale problem and then obtain uniqueness in law for the latter problem by adapting to our model a result given in [4].

Finally, let us observe that the filtering equation for the model studied in [5] was a linear one. This is not the case for a multitype model. This is the reason why, in order to obtain explicit representations for the filter, we use a linearization method that is a modification of the method proposed in [12] and generalized in [3]. By using the linearized equation, we exhibit two representations for the filter. The first one is obtained by the Feynman–Kac formula and the second one by a recursive algorithm, under a suitable assumption on the reproduction law.

2. The space of marked trees

We represent multitype branching processes on a space, introduced in [14] and [8], that is a modification of the space of marked trees. We refer the reader to those papers for any further details.

A *tree* is a subset $\tilde{\omega}$ of the denumerable set $U = \{\phi\} \cup \bigcup_{n=1}^{\infty} (\mathbb{N}^*)^n$, with $\mathbb{N}^* = \mathbb{N}/\{0\}$, and ϕ denotes the ancestor, such that

(i) $\phi \in \tilde{\omega}$,

(ii) $uv \in \tilde{\omega} \Rightarrow u \in \tilde{\omega}$ for all $u, v \in U$,

(iii) $uj \in \tilde{\omega} \Leftrightarrow 1 \leq j \leq \nu_u(\tilde{\omega})$ for all $u \in \tilde{\omega}$ and all $j \in \mathbb{N}^*$, with $\nu_u(\tilde{\omega}) := \nu_u^1(\tilde{\omega}) + \nu_u^2(\tilde{\omega})$, and $\nu_u^1(\tilde{\omega})$ and $\nu_u^2(\tilde{\omega})$ are $\mathbb{N}$-valued functions defined on U.

Part (iii) means that each particle u of $\tilde{\omega}$ has $\nu_u(\tilde{\omega})$ children – $\nu_u^1(\tilde{\omega})$ of type 1 and $\nu_u^2(\tilde{\omega})$ of type 2 – labelled by $u1, u2, \ldots, u\nu_u(\tilde{\omega})$. We let $\tilde{\Omega}$ be the space of trees and $\tilde{\Omega}_u = \{\tilde{\omega} \in \tilde{\Omega} : u \in \tilde{\omega}\}$ be the set of trees containing the particle u.

In order to consider a branching process in which each particle u has a lifetime σ_u, we need the notion of *marked tree*: a marked tree is a pair $\omega = (\tilde{\omega}, (\sigma_u)_{u \in \tilde{\omega}})$ with $\tilde{\omega} \in \tilde{\Omega}$ and

$\sigma_u \in (0, \infty)$, for all $u \in U$. We denote by Ω the set of marked trees, and we endow Ω with the σ-algebra

$$\mathcal{F} = \sigma\{\Omega_u; u \in U\}, \qquad \Omega_u := \{\omega \in \Omega : u \in \tilde{\omega}\}.$$

Hence, the functions

$$(v_u^1, v_u^2): \Omega_u \to \mathbb{N}^2 \quad \text{and} \quad \sigma_u: \Omega_u \to (0, \infty)$$

are random variables with respect to the σ-algebra $\mathcal{F} \cap \Omega_u$, and

$$\mathcal{F} = \sigma\{(v_u^1, v_u^2), \sigma_u; u \in U\}.$$

The birth time of each particle u is defined recursively as

$$\begin{aligned} S_\phi &= 0, \\ S_u &= S_v + \sigma_v, \end{aligned} \tag{2.1}$$

where v is the parent of u. Then, the line of the particles living at time t is given by

$$L_t = \{u : S_u \le t < S_u + \sigma_u\}$$

and the past of L_t is described by the σ-algebra

$$\mathcal{F}_t = \sigma\{L_s; s \le t\} = \sigma\{(v_u^1, v_u^2), \sigma_u; \ u \in U \text{ such that } S_u + \sigma_u \le t\}.$$

Finally, we define the type of each particle, denoting by γ_ϕ the type of the ancestor and setting

$$\gamma_{uj} = \begin{cases} 1, & j \le v_u^1, \\ 2, & v_u^1 < j \le v_u^1 + v_u^2 = v_u. \end{cases}$$

That is, we assume that the first v_u^1 children of the particle u (labelled $u1, u2, \ldots, uv_u^1$) are of type 1, while the remaining v_u^2 children are of type 2.

The *shifted tree* $T_{u,s}$ at u is the marked tree beginning at node u when u is s-old. More precisely, $T_{u,s}(\tilde{\omega}, (\sigma_v)_{v \in \tilde{\omega}})$, for $u \in \tilde{\omega}$ and $s < \sigma_u(\tilde{\omega})$, contains all the particles v such that $uv \in \tilde{\omega}$, and its marks are given by

$$\sigma_\phi \circ T_{u,s} = \sigma_u - s, \tag{2.2}$$

$$\sigma_v \circ T_{u,s} = \sigma_{uv}, \qquad v \ne \phi, \tag{2.3}$$

$$(v_v^1, v_v^2) \circ T_{u,s} = (v_{uv}^1, v_{uv}^2). \tag{2.4}$$

We now have to construct a probability measure on the space of marked trees in order to describe the dynamics of a two-type branching process. To this end, a reproduction law of a two-type branching process has to be given. This reproduction law is defined by the pair $(\{p^1(i, j)\}_{(i,j) \in \mathbb{N}^2}, \{p^2(i, j)\}_{(i,j) \in \mathbb{N}^2})$, where, for $\gamma = 1, 2$, $p^\gamma(i, j)$ is the probability that a particle of type γ has i offspring of type 1 and j offspring of type 2. We denote by $m_{h1} = \sum_{i,j \ge 0} i p^h(i, j)$ and $m_{h2} = \sum_{i,j \ge 0} j p^h(i, j)$, $h = 1, 2$, the expected number of type-1 particles borne by a particle of type h and the expected number of type-2 particles borne by a particle of type h, respectively. Throughout this paper, we assume $\{m_{hk}\}_{h,k=1,2}$ to be finite.

Theorem 2.1. *Given* $\alpha_1, \alpha_2 > 0$ *and two probability distributions*

$$\{p^1(i, j)\}_{(i,j) \in \mathbb{N}^2} \quad \text{and} \quad \{p^2(i, j)\}_{(i,j) \in \mathbb{N}^2}$$

on $\mathbb{N}^2$ *with finite first moments* $\{m_{hk}\}_{h,k=1,2}$, *for each* $\gamma \in \{1, 2\}$ *there exists a unique probability measure* P_γ *on* $(\Omega, \mathcal{F}_t, \mathcal{F})$ *such that the following statements hold.*

(i) *The pair (ν_ϕ^1, ν_ϕ^2) has law $\{p^\gamma(i, j)\}_{(i,j)\in\mathbb{N}^2}$; $\gamma_\phi = \gamma$; σ_ϕ is a random variable exponentially distributed with parameter α_γ; and the pair (ν_ϕ^1, ν_ϕ^2) and σ_ϕ are independent.*

(ii) *(Branching property.) The shifted trees $\{T_{u,t-S_u}\}_{u \in L_t}$ are conditionally independent with law P_{γ_u} given $\mathcal{F}_t$, i.e. for all bounded measurable functions f_u, $u \in U$, on Ω,*

$$E_\gamma\left(\prod_{u \in L_t} f_u \circ T_{u,t-S_u} \,\bigg|\, \mathcal{F}_t\right) = \prod_{u \in L_t} E_{\gamma_u}(f_u).$$

Proof. Consider the product space

$$\Omega^* = (\mathbb{N}^2 \times \mathbb{R}^+)^U$$

equipped with the σ-algebra $\mathcal{F}^*$ generated by the coordinates denoted by (ν_u^{1*}, ν_u^{2*}) and σ_u^*. Let S_u^* denote the random variables defined on Ω^* by the rule analogous to that proposed in (2.1), and let us set $\gamma_{uj}^* = \xi(j, \nu_u^{1*})$, where, for all $n > 0$,

$$\xi(j, n) = \begin{cases} 1 & \text{if } 0 < j \leq n, \\ 2 & \text{otherwise.} \end{cases}$$

We claim that, for each $\gamma \in \{1, 2\}$, a probability measure P_γ^* on $(\Omega^*, \mathcal{F}^*)$ can be defined as follows.

(i*) $P_\gamma^*((\nu_\phi^{1*}, \nu_\phi^{2*}) = (h, k), \sigma_\phi^* > t) = p^\gamma(h, k)e^{-\alpha_\gamma t}$ for $(h, k) \in \mathbb{N}^2$ and $t \geq 0$.

(ii*) $P_\gamma^*(\bigcap_{i=1}^n \{(\nu_{u_i}^{1*}, \nu_{u_i}^{2*}) = (h_i, k_i), \sigma_{u_i}^* > t_i\} \mid \mathcal{F}_{u_1,u_2,\ldots,u_n}^*) = \prod_{i=1}^n p^{\gamma_{u_i}^*}(h_i, k_i)e^{-\alpha_{\gamma_{u_i}^*} t_i}$ for all $(h_i, k_i) \in \mathbb{N}^2$ and $t_i \geq 0$, $i = 1, \ldots, n$. Here $u_1, u_2, \ldots, u_n$ are particles such that no one of them is an ancestor of any other, and

$$\mathcal{F}_{u_1,u_2,\ldots,u_n}^* = \sigma\{\nu_v^{1*}, \nu_v^{2*}, \sigma_v^* : v \text{ is an ancestor of some } u_i, i = 1, \ldots, n\}$$
$$\supseteq \sigma\{\gamma_{u_1}^*, \ldots, \gamma_{u_n}^*\}.$$

In particular, we can choose particles belonging to the same generation, i.e. particles such that $|u_i| = m$ for all $i = 1, \ldots, n$, where $|u|$ is the length of the sequence u. Writing

$$\mathcal{F}_m^* = \sigma\{(\nu_u^{1*}, \nu_u^{2*}), \sigma_u^* : |u| < m\},$$

we obtain

$$P_\gamma^*((\nu_{u_1}^{1*}, \nu_{u_2}^{2*}) = (h_1, k_1), \sigma_{u_1}^* > t_1, \ldots, (\nu_{u_n}^{1*}, \nu_{u_n}^{2*}) = (h_n, k_n), \sigma_{u_n}^* > t_n \mid \mathcal{F}_m^*)$$
$$= \prod_{i=1}^n p^{\gamma_{u_i}^*}(h_i, k_i)e^{-\alpha_{\gamma_{u_i}^*} t_i}$$

and, by recursivity on the generations, P_γ^* is completely defined on $\mathcal{F}^* = \bigvee_m \mathcal{F}_m^*$.

We now introduce the map $\psi : \Omega^* \to \Omega$ that associates to any $\omega^* \in \Omega^*$ the marked tree built with the coordinates (ν_u^{1*}, ν_u^{2*}) and σ_u^* (see [14]). The probability measure P_γ we are looking for in this theorem is then given by

$$P_\gamma(A) := P_\gamma^*(\psi^{-1}(A)) \quad \text{for all } A \in \mathcal{F}.$$

Condition (i) is clearly satisfied. To verify condition (ii), we define the shifted operators $T^*_{u,s} \colon \Omega^* \cap \{\sigma^*_u > s\} \to \Omega^*$, as in (2.2)–(2.4). By the definition of P^*_γ, we can prove that

$$\mathrm{E}^*_\gamma \left(\prod_{u \in L^*_t} f^*_u \circ T^*_{u,t-S^*_u} \ \Big| \ \mathcal{F}^*_t \right) = \prod_{u \in L^*_t} \mathrm{E}^*_{\gamma^*_u}(f^*_u), \tag{2.5}$$

where

$$L^*_t = \{u \colon S^*_u \leq t < S^*_u + \sigma^*_u\},$$
$$\mathcal{F}^*_t = \sigma\{L^*_s; s \leq t\} = \sigma\{(v^{1*}_u, v^{2*}_u), \sigma^*_u; u \in U \text{ such that } S^*_u + \sigma^*_u \leq t\}.$$

To obtain (2.5) we have to verify that, for any finite subset V of U,

$$\mathrm{E}^*_\gamma \left(\prod_{u \in L^*_t} \prod_{v \in V} \mathbf{1}_{\{(v^{1*}_{uv}, v^{2*}_{uv})=(h_{uv}, k_{uv}), \sigma^*_{uv} > t_{uv}\}} \ \Big| \ \mathcal{F}^*_t \right) = \prod_{u \in L^*_t} \mathrm{E}^*_{\gamma^*_u} \left(\prod_{v \in V} \mathbf{1}_{\{(v^{1*}_v, v^{2*}_v)=(h_{uv}, k_{uv}), \sigma^*_v > t_{uv}\}} \right)$$

but, for the sake of simplicity, we just consider the case in which $V = \{v\}$ for a fixed $v \in \mathbb{N}^*$. If we let

$$\mathcal{G}^{*v}_t = \mathcal{F}^*_t \vee \sigma\{\gamma^*_{uv}; u \in L^*_t\} \supseteq \sigma\{v^{1*}_u, v^{2*}_u, \sigma^*_u; u \in L^*_s, s \leq t\},$$

then

$$\mathrm{E}^*_\gamma \left(\prod_{u \in L^*_t} \mathbf{1}_{\{(v^{1*}_{uv}, v^{2*}_{uv})=(h_{uv}, k_{uv}), \sigma^*_{uv} > t_{uv}\}} \ \Big| \ \mathcal{F}^*_t \right)$$

$$= \mathrm{E}^*_\gamma \left(\mathrm{E}^*_\gamma \left(\prod_{u \in L^*_t} \mathbf{1}_{\{(v^{1*}_{uv}, v^{2*}_{uv})=(h_{uv}, k_{uv}), \sigma^*_{uv} > t_{uv}\}} \ \Big| \ \mathcal{G}^{*v}_t \right) \ \Big| \ \mathcal{F}^*_t \right)$$

$$= \mathrm{E}^*_\gamma \left(\prod_{u \in L^*_t} p^{\gamma^*_{uv}}(h_{uv}, k_{uv}) \mathrm{e}^{-\alpha_{\gamma^*_{uv}} t_{uv}} \ \Big| \ \mathcal{F}^*_t \right)$$

$$= \prod_{u \in L^*_t} \sum_{j \geq 0} p^{\xi(v,j)}(h_{uv}, k_{uv}) \mathrm{e}^{-\alpha_{\xi(v,j)} t_{uv}} \mathrm{P}^*_\gamma(v^{1*}_u = j \mid \mathcal{F}^*_t).$$

Since

$$\mathrm{P}^*_\gamma(v^{1*}_u = j \mid \mathcal{F}^*_t) = \mathrm{P}^*_\gamma(v^{1*}_u = j \mid \gamma^*_u) = \mathrm{P}^*_{\gamma^*_u}(v^{1*}_\phi = j),$$

we find that

$$\mathrm{E}^*_\gamma \left(\prod_{u \in L^*_t} \mathbf{1}_{\{(v^{1*}_{uv}, v^{2*}_{uv})=(h_{uv}, k_{uv}), \sigma^*_{uv} > t_{uv}\}} \ \Big| \ \mathcal{F}^*_t \right) = \prod_{u \in L^*_t} \mathrm{E}^*_{\gamma^*_u}(p^{\gamma^*_v}(h_{uv}, k_{uv}) \mathrm{e}^{-\alpha_{\gamma^*_v} t_{uv}})$$

$$= \prod_{u \in L^*_t} \mathrm{E}^*_{\gamma^*_u}(\mathbf{1}_{\{(v^{1*}_v, v^{2*}_v)=(h_{uv}, k_{uv}), \sigma^*_v > t_{uv}\}})$$

and we recover (2.5).

Furthermore,

$$E_\gamma\left(\prod_{u\in L_t} f_u \circ T_{u,t-S_u} \,\Big|\, \mathscr{F}_t\right)(\omega) = E_\gamma^*\left(\left(\prod_{u\in L_t} f_u \circ T_{u,t-S_u}\right)\circ \psi \,\Big|\, \psi^{-1}(\mathscr{F}_t)\right)(\psi^{-1}(\omega))$$

and, observing that

$$\psi \circ T_{u,t-S_u^*}^* = T_{u,t-S_u} \circ \psi, \qquad L_t^* = L_t \circ \psi, \qquad \psi^{-1}(\mathscr{F}_t) = \mathscr{F}_t^*,$$

we finally obtain

$$E_\gamma\left(\prod_{u\in L_t} f_u \circ T_{u,t-S_u} \,\Big|\, \mathscr{F}_t\right)(\omega) = E_\gamma^*\left(\prod_{u\in L_t^*} f_u \circ \psi \circ T_{u,t-S_u^*}^* \,\Big|\, \mathscr{F}_t^*\right)(\psi^{-1}(\omega))$$

$$= \prod_{u\in L_t^*(\psi^{-1}(\omega))} E_{\gamma_u^*(\psi^{-1}(\omega))}^*(f_u \circ \psi)$$

$$= \prod_{u\in L_t(\omega)} E_{\gamma_u(\omega)}(f_u).$$

As in [7] and [6], we can prove that this property (called the *strong branching property*) is also valid for the line of particles alive at an $\mathscr{F}_t$-stopping time.

Theorem 2.2. (Strong branching property.) *Let τ be an $\mathscr{F}_t$-stopping time, and let*

$$L_\tau = \{u\colon S_u \le \tau < S_u + \sigma_u\}$$

denote the line of particles living at time τ. Then, for all bounded measurable functions f_u, $u \in U$, on Ω,

$$E_\gamma\left(\prod_{u\in L_\tau} f_u \circ T_{u,\tau-S_u} \,\Big|\, \mathscr{F}_\tau\right) = \prod_{u\in L_\tau} E_{\gamma_u}(f_u).$$

3. Dynamics of the model

This section is devoted to proving that the process (X_t, Y_t), defined above, is a Markov process and to finding the structure of its generator; that is, to finding the operator L such that

$$f(X_t, Y_t) - f(X_0, Y_0) - \int_0^t Lf(X_s, Y_s)\, \mathrm{d}s$$

is a $(P_\gamma, \mathscr{F}_t)$-martingale ($\gamma = 1, 2$), for a suitable class of functions f.

To this end, we introduce the integer-valued random measure on $(0, \infty) \times \mathbb{N} \times \mathbb{N}$ associated with the jump process (X_t, Y_t), namely

$$m(\mathrm{d}t, \mathrm{d}x, \mathrm{d}y) = \sum_{n\ge 1} \delta_{\{\tau_n, X_{\tau_n}, Y_{\tau_n}\}}(\mathrm{d}t, \mathrm{d}x, \mathrm{d}y)\, \mathbf{1}_{\{\tau_n < \infty\}},$$

where $\delta_{\{t,x,y\}}$ is the Dirac measure concentrated at (t, x, y). We recall that the predictable projection of m is the random measure $p(\mathrm{d}t, \mathrm{d}x, \mathrm{d}y)$ such that [2], [11], for all nonnegative

$\mathcal{F}_t$-predictable $(\mathbb{N} \times \mathbb{N})$-indexed processes $H(s, x, y)$, the process

$$\int_0^t \int_{\mathbb{N} \times \mathbb{N}} H(s, x, y)(m(\mathrm{d}s, \mathrm{d}x, \mathrm{d}y) - p(\mathrm{d}s, \mathrm{d}x, \mathrm{d}y))$$

is a $(\mathrm{P}_\gamma, \mathcal{F}_t)$-martingale whenever

$$\mathrm{E}_\gamma \left(\int_0^t \int_{\mathbb{N} \times \mathbb{N}} H(s, x, y) p(\mathrm{d}s, \mathrm{d}x, \mathrm{d}y) \right) < \infty.$$

In particular, for any bounded measurable function f,

$$
\begin{aligned}
f(X_t, Y_t) \\
= f(X_0, Y_0) + \int_0^t \int_{\mathbb{N} \times \mathbb{N}} \big(f(x, y) - f(X_{s-}, Y_{s-}) \big) m(\mathrm{d}s, \mathrm{d}x, \mathrm{d}y) \qquad (3.1) \\
= f(X_0, Y_0) + \int_0^t \int_{\mathbb{N} \times \mathbb{N}} \big(f(x, y) - f(X_{s-}, Y_{s-}) \big) p(\mathrm{d}s, \mathrm{d}x, \mathrm{d}y) + (\mathrm{P}_\gamma, \mathcal{F}_t)\text{-martingale}.
\end{aligned}
$$

Thus, our problem reduces to computing the random measure $p(\mathrm{d}t, \mathrm{d}x, \mathrm{d}y)$, whose structure [11] is given by

$$p(\mathrm{d}t, \mathrm{d}x, \mathrm{d}y) = \sum_{n \geq 0} \frac{\mathrm{P}_\gamma(\tau_{n+1} \in \mathrm{d}t, X_{\tau_{n+1}} \in \mathrm{d}x, Y_{\tau_{n+1}} \in \mathrm{d}y \mid \mathcal{F}_{\tau_n})}{\mathrm{P}_\gamma(\tau_{n+1} - \tau_n \geq t - \tau_n \mid \mathcal{F}_{\tau_n})} \mathbf{1}_{\{\tau_n < t \leq \tau_{n+1}\}}. \qquad (3.2)$$

Here, $\{\tau_n\}_{n \geq 0}$ are the splitting times of the process $Z_t = X_t + Y_t$, representing the cardinality of the whole population at time t, and are defined as

$$\tau_0 = 0, \qquad \tau_{n+1} = \min\{S_u + \sigma_u : S_u + \sigma_u > \tau_n\}.$$

In order to perform this computation, we need two preliminary results, which are generalizations of Propositions 3.1 and 3.2 of [5].

Proposition 3.1. *The conditional law of the random variables $\tau_{n+1} - \tau_n, n \geq 0$, given $\mathcal{F}_{\tau_n}$, is an exponential law with parameter $\alpha_1 X_{\tau_n} + \alpha_2 Y_{\tau_n}$.*

  Proof. The decomposition

$$
\begin{aligned}
\{\tau_{n+1} \leq t\} &= \{\tau_n \leq t\} - \{\tau_n \leq t, \tau_{n+1} > t\} \\
&= \{\tau_n \leq t\} - \{\tau_n \leq t\} \bigcap_{v \in L_{\tau_n}} \{\sigma_\phi \circ T_{v, \tau_n - S_v} > t - \tau_n\}
\end{aligned}
$$

and the strong branching property allow us to compute

$$\mathrm{P}_\gamma(\tau_{n+1} \leq t \mid \mathcal{F}_{\tau_n}) = \begin{cases} 1 - \exp\{-(\alpha_1 X_{\tau_n} + \alpha_2 Y_{\tau_n})(t - \tau_n)\} & \text{if } \tau_n \leq t, \\ 0 & \text{if } \tau_n > t. \end{cases}$$

Proposition 3.2. *For any $n \geq 0$, the joint law of $(X_{\tau_{n+1}}, Y_{\tau_{n+1}}, \tau_{n+1})$, conditioned on $\mathscr{F}_{\tau_n}$, is given by*

$$
P_\gamma(X_{\tau_{n+1}} = x, Y_{\tau_{n+1}} = y, \tau_{n+1} \leq t \mid \mathscr{F}_{\tau_n})
$$

$$
= \frac{1 - \exp\{-(\alpha_1 X_{\tau_n} + \alpha_2 Y_{\tau_n})(t - \tau_n)\}}{\alpha_1 X_{\tau_n} + \alpha_2 Y_{\tau_n}} \mathbf{1}_{\{\tau_n \leq t\}}
$$

$$
\times \left\{ \alpha_1 X_{\tau_n} \sum_{h=1}^{x+1} \sum_{k=0}^{y} p^1(x - h + 1, y - k) \mathbf{1}_{\{X_{\tau_n} = h, Y_{\tau_n} = k\}} \right.
$$

$$
\left. + \alpha_2 Y_{\tau_n} \sum_{h=0}^{x} \sum_{k=1}^{y+1} p^2(x - h, y - k + 1) \mathbf{1}_{\{X_{\tau_n} = h, Y_{\tau_n} = k\}} \right\}. \tag{3.3}
$$

Proof. First observe that, in the event $\{Z_{\tau_n} = 0\}$, the claim is trivially true. Then, we perform the following computations in the event $\{Z_{\tau_n} > 0\}$.

$$
P_\gamma(X_{\tau_{n+1}} = x, Y_{\tau_{n+1}} = y, \tau_{n+1} \leq t \mid \mathscr{F}_{\tau_n})
$$

$$
= \sum_{u \in L_{\tau_n}} \sum_{h>0, k\geq 0} P_\gamma(S_u + \sigma_u = \tau_{n+1} \leq t, X_{\tau_{n+1}} - X_{\tau_n} = x - h,
$$

$$
Y_{\tau_{n+1}} - Y_{\tau_n} = y - k \mid \mathscr{F}_{\tau_n}) \mathbf{1}_{\{X_{\tau_n} = h, Y_{\tau_n} = k\}} \mathbf{1}_{\{\gamma_u = 1\}}
$$

$$
+ \sum_{u \in L_{\tau_n}} \sum_{h\geq 0, k>0} P_\gamma(S_u + \sigma_u = \tau_{n+1} \leq t, X_{\tau_{n+1}} - X_{\tau_n} = x - h,
$$

$$
Y_{\tau_{n+1}} - Y_{\tau_n} = y - k \mid \mathscr{F}_{\tau_n}) \mathbf{1}_{\{X_{\tau_n} = h, Y_{\tau_n} = k\}} \mathbf{1}_{\{\gamma_u = 2\}}.
$$

On the other hand,

$$
\gamma_u = 1 \Rightarrow X_{\tau_{n+1}} - X_{\tau_n} = v_u^1 - 1 \quad \text{and} \quad Y_{\tau_{n+1}} - Y_{\tau_n} = v_u^2,
$$

$$
\gamma_u = 2 \Rightarrow X_{\tau_{n+1}} - X_{\tau_n} = v_u^1 \quad \text{and} \quad Y_{\tau_{n+1}} - Y_{\tau_n} = v_u^2 - 1,
$$

meaning that

$$
P_\gamma(X_{\tau_{n+1}} = x, Y_{\tau_{n+1}} = y, \tau_{n+1} \leq t \mid \mathscr{F}_{\tau_n})
$$

$$
= \sum_{u \in L_{\tau_n}} \sum_{h=1}^{x+1} \sum_{k=0}^{y} P_\gamma(S_u + \sigma_u = \tau_{n+1} \leq t, v_u^1 - 1 = x - h, v_u^2 = y - k \mid \mathscr{F}_{\tau_n})
$$

$$
\times \mathbf{1}_{\{X_{\tau_n} = h, Y_{\tau_n} = k\}} \mathbf{1}_{\{\gamma_u = 1\}}
$$

$$
+ \sum_{u \in L_{\tau_n}} \sum_{h=0}^{x} \sum_{k=1}^{y+1} P_\gamma(S_u + \sigma_u = \tau_{n+1} \leq t, v_u^1 = x - h, v_u^2 - 1 = y - k \mid \mathscr{F}_{\tau_n})
$$

$$
\times \mathbf{1}_{\{X_{\tau_n} = h, Y_{\tau_n} = k\}} \mathbf{1}_{\{\gamma_u = 2\}},
$$

where, for $u \in L_{\tau_n}$, by introducing the shifted trees $\{T_{v, \tau_n - S_v}; v \in L_{\tau_n}\}$ and recalling the definitions given in (2.2)–(2.4), we have

$$
\{S_u + \sigma_u = \tau_{n+1} \leq t\} = \bigcap_{v \in L_{\tau_n}} \{S_u + \sigma_u \leq t, S_u + \sigma_u \leq S_v + \sigma_v\}
$$

$$
= \bigcap_{v \in L_{\tau_n}} \{\sigma_\phi \circ T_{u, \tau_n - S_u} \leq t - \tau_n, \sigma_\phi \circ T_{u, \tau_n - S_u} \leq \sigma_\phi \circ T_{v, \tau_n - S_v}\},
$$

$$
\{v_u^1 = h_1, v_u^2 = h_2\} = \{v_\phi^1 \circ T_{u, \tau_n - S_u} = h_1, v_\phi^2 \circ T_{u, \tau_n - S_u} = h_2\}.
$$

Thus, writing $\bar{P} = \bigotimes_{w \in L_{\tau_n}} P_{\gamma_w}$, we can use the strong branching property to obtain

$$P_\gamma(X_{\tau_{n+1}} = x, Y_{\tau_{n+1}} = y, \tau_{n+1} \le t \mid \mathcal{F}_{\tau_n})$$

$$= \sum_{u \in L_{\tau_n}} \sum_{h=1}^{x+1} \sum_{k=0}^{y} P_{\gamma_u}(v_\phi^1 - 1 = x - h, v_\phi^2 = y - k)$$

$$\times \bar{P}(\sigma(\omega_u) \le t - \tau_n, \sigma(\omega_u) \le \sigma(\omega_v) \text{ for all } v \in L_{\tau_n})$$

$$\times \mathbf{1}_{\{X_{\tau_n} = h, Y_{\tau_n} = k\}} \mathbf{1}_{\{\gamma_u = 1\}}$$

$$+ \sum_{u \in L_{\tau_n}} \sum_{h=0}^{x} \sum_{k=1}^{y+1} P_{\gamma_u}(v_\phi^1 = x - h, v_\phi^2 - 1 = y - k)$$

$$\times \bar{P}(\sigma(\omega_u) \le t - \tau_n, \sigma(\omega_u) \le \sigma(\omega_v) \text{ for all } v \in L_{\tau_n})$$

$$\times \mathbf{1}_{\{X_{\tau_n} = h, Y_{\tau_n} = k\}} \mathbf{1}_{\{\gamma_u = 2\}}.$$

The statement of the proposition follows, since

$$P_{\gamma_u}(v_\phi^1 - 1 = x - h, v_\phi^2 = y - k)\mathbf{1}_{\{\gamma_u = 1\}} = p^1(x - h + 1, y - k)\mathbf{1}_{\{\gamma_u = 1\}},$$

$$P_{\gamma_u}(v_\phi^1 = x - h, v_\phi^2 - 1 = y - k)\mathbf{1}_{\{\gamma_u = 2\}} = p^2(x - k, y - k + 1)\mathbf{1}_{\{\gamma_u = 2\}},$$

$$\bar{P}(\sigma(\omega_u) \le t - \tau_n, \sigma(\omega_u) \le \sigma(\omega_v) \text{ for all } v \in L_{\tau_n})$$

$$= \frac{\alpha_{\gamma_u}}{\alpha_1 X_{\tau_n} + \alpha_2 Y_{\tau_n}}(1 - \exp\{-(\alpha_1 X_{\tau_n} + \alpha_2 Y_{\tau_n})(t - \tau_n)\})\mathbf{1}_{\{\tau_n \le t\}}.$$

Furthermore, notice that, for any $n \ge 0$, the random variables $(X_{\tau_{n+1}}, Y_{\tau_{n+1}})$ and τ_{n+1} are conditionally independent, given $\mathcal{F}_{\tau_n}$, i.e.

$$P_\gamma(X_{\tau_{n+1}} = x, Y_{\tau_{n+1}} = y, \tau_{n+1} \le t \mid \mathcal{F}_{\tau_n})$$

$$= P_\gamma(X_{\tau_{n+1}} = x, Y_{\tau_{n+1}} = y \mid \mathcal{F}_{\tau_n}) P_\gamma(\tau_{n+1} \le t \mid \mathcal{F}_{\tau_n}).$$

Finally, we have all the tools to prove that the pair (X_t, Y_t) is a Markov process with generator given by

$$Lf(x, y) = \alpha_1 x \sum_{h,k \ge 0} p^1(h, k)\{f(x + h - 1, y + k) - f(x, y)\}$$

$$+ \alpha_2 y \sum_{h,k \ge 0} p^2(h, k)\{f(x + h, y + k - 1) - f(x, y)\}.$$

Theorem 3.1. *For any bounded and measurable function f on $\mathbb{N} \times \mathbb{N}$,*

$$f(X_t, Y_t) - f(X_0, Y_0) - \int_0^t Lf(X_s, Y_s)\,\mathrm{d}s$$

is a $(P_\gamma, \mathcal{F}_t)$-martingale $(\gamma = 1, 2)$.

Proof. Recalling the definition of $p(\mathrm{d}t, \mathrm{d}x, \mathrm{d}y)$ given by (3.2), by Propositions 3.1 and 3.2 we first obtain

$$\frac{\mathrm{P}_\gamma(\tau_{n+1} \in \mathrm{d}t, X_{\tau_{n+1}} \in \mathrm{d}x, Y_{\tau_{n+1}} \in \mathrm{d}y \mid \mathcal{F}_{\tau_n})}{\mathrm{P}_\gamma(\tau_{n+1} - \tau_n \geq t - \tau_n \mid \mathcal{F}_{\tau_n})} \mathbf{1}_{\{\tau_n < t \leq \tau_{n+1}\}}$$
$$= (\alpha_1 X_{\tau_n} + \alpha_2 Y_{\tau_n})$$
$$\times \sum_{(x', y') \in \mathbb{N} \times \mathbb{N}} \mathrm{P}_\gamma(X_{\tau_{n+1}} = x', Y_{\tau_{n+1}} = y' \mid \mathcal{F}_{\tau_n}) \delta_{\{x'\}}(\mathrm{d}x) \delta_{\{y'\}}(\mathrm{d}y) \, \mathrm{d}t \, \mathbf{1}_{\{\tau_n < t \leq \tau_{n+1}\}}.$$

Then,

$$p(\mathrm{d}t, \mathrm{d}x, \mathrm{d}y) = \sum_{n \geq 0} \mathbf{1}_{\{\tau_n < t \leq \tau_{n+1}\}} \left\{ \alpha_1 X_{\tau_n} \sum_{h=1}^{x'+1} \sum_{k=0}^{y'} p^1(x' - h + 1, y' - k) \, \mathbf{1}_{\{X_{\tau_n}=h, Y_{\tau_n}=k\}} \right.$$
$$\left. + \alpha_2 Y_{\tau_n} \sum_{h=0}^{x'} \sum_{k=1}^{y'+1} p^2(x' - h, y' - k + 1) \, \mathbf{1}_{\{X_{\tau_n}=h, Y_{\tau_n}=k\}} \right\}$$
$$\times \delta_{\{x'\}}(\mathrm{d}x) \delta_{\{y'\}}(\mathrm{d}y) \, \mathrm{d}t$$
$$= \alpha_1 X_{t-} \sum_{i,j \geq 0} p^1(i, j) \delta_{\{X_{t-}+i-1\}}(\mathrm{d}x) \delta_{\{Y_{t-}+j\}}(\mathrm{d}y) \, \mathrm{d}t$$
$$+ \alpha_2 Y_{t-} \sum_{i,j \geq 0} p^2(i, j) \delta_{\{X_{t-}+i\}}(\mathrm{d}x) \delta_{\{Y_{t-}+j-1\}}(\mathrm{d}y) \, \mathrm{d}t$$

and, recalling (3.1), the proof is complete.

Furthermore, the result given in Proposition 1.2 of [4] applies to our model. Thus, we can claim the following.

Proposition 3.3. *For any initial condition (X_0, Y_0) with finite first moments, the martingale problem for the operator L has a unique solution with sample paths in $D_{\mathbb{N} \times \mathbb{N}}[0, \infty)$, the space of càdlàg $(\mathbb{N} \times \mathbb{N})$-valued functions defined on $[0, \infty)$. The solution (X_t, Y_t) is a Markov process with finite first moments.*

Remark 3.1. As a consequence, if we denote by

$$L_X f(x, y) = \alpha_1 x \sum_{h,k \geq 0} p^1(k, h)\{f(x + k - 1) - f(x)\} + \alpha_2 y \sum_{h,k \geq 0} p^2(k, h)\{f(x + k) - f(x)\}$$

the operator L restricted to functions depending on x only, it is easy to prove that

$$m_t^f := f(X_t) - f(X_0) - \int_0^t L_X f(X_s, Y_s) \, \mathrm{d}s \tag{3.4}$$

is a $(\mathrm{P}_\gamma, \mathcal{F}_t)$-martingale, for any real-valued, bounded measurable function f.

4. The Kushner–Stratonovich equation

In this section, (X_t, Y_t) denotes the solution for the martingale problem for the operator L on the space $D_{\mathbb{N} \times \mathbb{N}}[0, \infty)$ endowed with its canonical filtration. The probability measure

P is the law of (X_t, Y_t) on $D_{\mathbb{N} \times \mathbb{N}}[0, \infty)$, for any initial condition (X_0, Y_0) with finite first moments.

The problem of finding the conditional law of the process X_t, given the history of Y_t, amounts to characterizing the filter of X_t with respect to the σ-algebra $\mathcal{F}_t^Y = \sigma\{Y_s; s \leq t\}$, i.e.

$$\pi_t(f) = \mathrm{E}(f(X_t) \mid \mathcal{F}_t^Y)$$

for any real-valued, bounded measurable function f. To this end, following a procedure suggested in [3] and already used in [5], we introduce the sequence of point processes $\{v_t^h\}_{h \geq 0}$ given by

$$v_t^h = \sum_{i \geq 0} \mathbf{1}_{\{S_i \leq t\}} \mathbf{1}_{\{Y_{S_i} = h\}},$$

where $\{S_n\}_{n \geq 0}$ denotes the sequence of jump times of the process Y_t. The following proposition can be proved as in [5], observing that

$$Y_t = Y_0 + \int_0^t \sum_{h \geq 0} (h - Y_{s-}) \, dv_s^h. \tag{4.1}$$

Proposition 4.1. *The σ-algebra $\mathcal{F}_t^Y$ coincides with the σ-algebra*

$$\bigvee_{h \geq 0} \mathcal{F}_t^{v^h} = \sigma\{v_s^h; s \leq t, h \geq 0\}.$$

Thus, $\pi_t(f) = \mathrm{E}(f(X_t) \mid \mathcal{F}_t^Y) = \mathrm{E}(f(X_t) \mid \bigvee_{h \geq 0} \mathcal{F}_t^{v^h})$. The process $v_t = \{v_t^h\}_{h \geq 0}$ is a multivariate point process. In order to obtain the filtering equation, we need the following lemma.

Lemma 4.1. *The $(\mathrm{P}, \mathcal{F}_t)$-predictable intensity of v_t^h is given, for all $h \geq 0$, by*

$$m_t^h = m^h(X_{t-}, Y_{t-}),$$

where

$$p_Y^i(k) = \sum_{h \in \mathbb{N}} p^i(h, k), \qquad i = 1, 2, \ k \in \mathbb{N},$$

$$m^h(x, y) = \alpha_1 x \, \mathbf{1}_{\{y < h\}} \, p_Y^1(h - y) + \alpha_2 y \, \mathbf{1}_{\{y \leq h+1\}} \, \mathbf{1}_{\{y \neq h\}} \, p_Y^2(h + 1 - y).$$

Proof. By the generalized Itô formula, we can compute the operator L^h describing the joint dynamics of (X_t, Y_t, v_t^h) in the sense that

$$F(X_t, Y_t, v_t^h) - F(X_0, Y_0, v_0^h) - \int_0^t L^h F(X_s, Y_s, v_s^h) \, ds$$

is a $(P, \mathcal{F}_t)$-martingale for any real-valued, bounded measurable function F. The operator L^h is defined by

$$
\begin{aligned}
L^h & F(x, y, v) \\
&= \alpha_1 x \sum_{i \geq 0} [F(x + i - 1, y, v) - F(x, y, v)] p^1(i, 0) \\
&\quad + \alpha_2 y \sum_{i \geq 0} [F(x + i, y, v) - F(x, y, v)] p^2(i, 1) \\
&\quad + \alpha_1 x \sum_{i \geq 0, \, j > 0} [F(x + i - 1, y + j, v) - F(x, y, v)] \mathbf{1}_{\{y+j \neq h\}} \, p^1(i, j) \\
&\quad + \alpha_2 y \sum_{i, \, j \geq 0, \, j \neq 1} [F(x + i, y + j - 1, v) - F(x, y, v)] \mathbf{1}_{\{y+j-1 \neq h\}} \, p^2(i, j) \\
&\quad + \alpha_1 x \sum_{i \geq 0, \, j > 0} [F(x + i - 1, y + j, v + 1) - F(x, y, v)] \mathbf{1}_{\{y+j = h\}} \, p^1(i, j) \\
&\quad + \alpha_2 y \sum_{i, \, j \geq 0, \, j \neq 1} [F(x + i, y + j - 1, v + 1) - F(x, y, v)] \mathbf{1}_{\{y+j-1 = h\}} \, p^2(i, j).
\end{aligned}
$$

By setting $F(x, y, v) = v$, we obtain

$$
m_t^{v^h} = v_t^h - v_0^h - \int_0^t m_s^h \, \mathrm{d}s,
$$

which is a $(P, \mathcal{F}_t)$-martingale.

Remark 4.1. The process

$$
N_t = \sum_{i \geq 0} \mathbf{1}_{\{S_i \leq t\}},
$$

which counts all the jumps of Y_t up to time t, has $(P, \mathcal{F}_t)$-predictable intensity

$$
\lambda_t \equiv \lambda(X_{t-}, Y_{t-}),
$$

where

$$
\lambda(x, y) = \sum_{h \geq 0} m^h(x, y) = \alpha_1 x(1 - p_Y^1(0)) + \alpha_2 y(1 - p_Y^2(1)).
$$

Theorem 4.1. *The filter $\pi_t(f)$ is a solution of the equation*

$$
\begin{aligned}
\pi_t(f) = \pi_0(f) &+ \int_0^t \pi_s(L_X f(\cdot, Y_s)) \, \mathrm{d}s \\
&+ \int_0^t \sum_{h \geq 0} \frac{1}{\pi_s(m^h(\cdot, Y_s))} \{\alpha_1 \, \mathbf{1}_{\{Y_s < h\}} \, p_Y^1(h - Y_s)[\pi_s(\varphi f) - \pi_s(\varphi)\pi_s(f)] \\
&\qquad\qquad + \pi_s(R_X^h f(\cdot, Y_s))\}|_{s=s-}(\mathrm{d}v_s^h - \pi_s(m^h(\cdot, Y_s)) \, \mathrm{d}s), \quad (4.2)
\end{aligned}
$$

where $\varphi(x) = x$ for all x, i.e. φ is the identity function, and R_X^h is defined in (4.3) below.

Proof. By the well-known innovation method [9], the filtering equation can be written as

$$\pi_t(f) = \pi_0(f) + \int_0^t \pi_s(L_X f(\cdot, Y_s)) \, ds$$

$$+ \int_0^t \sum_{h \geq 0} \pi_s^+(m^h(\cdot, Y_s))$$

$$\times [\pi_s(fm^h(\cdot, Y_s)) - \pi_s(f)\pi_s(m^h(\cdot, Y_s)) + \pi_s(R_X^h f(\cdot, Y_s))]|_{s=s-}$$

$$\times (dv_s^h - \pi_s(m^h(\cdot, Y_s)) \, ds),$$

where, as usual, $a^+ = (1/a) \, \mathbf{1}_{\{a>0\}}$, m_t^f is defined in (3.4), and $R_X^h(X_s, Y_s)$ is such that

$$\langle m^{v^h}, m^f \rangle_t = \int_0^t R_X^h f(X_s, Y_s) \, ds. \tag{4.3}$$

Taking into account the fact that

$$|\pi_s(fm^h(\cdot, Y_s)) - \pi_s(f)\pi_s(m^h(\cdot, Y_s)) + \pi_s(R_X^h f(\cdot, Y_s))| \leq \text{ const. } \|f\| \pi_s(m^h(\cdot, Y_s)),$$

where $\|f\|$ denotes the supremum norm of the function f, we recover (4.2).

Corollary 4.1. *The filtering equation can also be written as*

$$\pi_t(f) = \pi_0(f) + \int_0^t \{\pi_s(L_X^0 f(\cdot, Y_s)) \, ds - \alpha_1(1 - p_Y^1(0))[\pi_s(\varphi f) - \pi_s(\varphi)\pi_s(f)]\} \, ds$$

$$+ \int_0^t \sum_{h \geq 0} \frac{1}{\pi_s(m^h(\cdot, Y_s))}$$

$$\times \{\alpha_1 \, \mathbf{1}_{\{Y_s < h\}} \, p_Y^1(h - Y_s)[\pi_s(\varphi f) - \pi_s(\varphi)\pi_s(f)]$$

$$+ \pi_s(R_X^h f(\cdot, Y_s))\}|_{s=s-} \, dv_s^h,$$

where

$$L_X^0 f(x, y) = \alpha_1 x \sum_{i \geq 0} [f(x+i-1) - f(x)] p^1(i, 0) + \alpha_2 y \sum_{i \geq 0} [f(x+i) - f(x)] p^2(i, 1). \tag{4.4}$$

Proof. The joint dynamics of (X, Y, v^h) described by the operator L^h, and (4.3), give us

$$R_X^h f(x, y) = \alpha_1 x \sum_{i \geq 0, \, j > 0} [f(x+i-1) - f(x)] \mathbf{1}_{\{y+j=h\}} \, p^1(i, j)$$

$$+ \alpha_2 y \sum_{i, j \geq 0, \, j \neq 1} [f(x+i) - f(x)] \mathbf{1}_{\{y+j-1=h\}} \, p^2(i, j)$$

and

$$L_X^0 f(x, y) = L_X f(x, y) - \sum_{h \geq 0} R_X^h f(x, y).$$

In the remaining part of this section, we discuss uniqueness in law for the solutions of (4.2). We want to show that any two weak solutions of (4.2) have the same law, where a 'weak solution' is defined as follows.

Definition 4.1. The pair (π, Y), on the probability space $(\Omega, \mathcal{F}, P)$, is a weak solution of (4.2) if the following statements hold.

- Y is an $\mathbb{N}$-valued jump process with càdlàg trajectories.

- Process π is $\Pi(\mathbb{N})$-valued and $\mathcal{F}_t^Y$-adapted with càdlàg trajectories (where $\Pi(\mathbb{N})$ denotes the space of probability measures on $\mathbb{N}$). Furthermore, $\pi_t(|\varphi|) = \int_{\mathbb{N}} |x| \pi_t(dx) < \infty$ for all $t \geq 0$.

- If we set $v_t^h = \sum_{i \geq 0} \mathbf{1}_{\{S_i \leq t\}} \mathbf{1}_{\{Y_{S_i} = h\}}$ for any $h \geq 0$, where $\{S_n\}_{n \geq 0}$ denotes the sequence of jump times of Y, then the $(P, \mathcal{F}_t^Y)$-minimal intensity of v^h is given by $\pi_t(m^h(\cdot, Y_t))$.

- The triple (π, Y, v^h) satisfies (4.2), and all terms in this equation are well defined.

Weak uniqueness for the filtering equation can be obtained by the filtering martingale problem (FMP) approach, introduced in [13] and widely used in the filtering literature (see, for example, [4] and references therein). First we recall the well-known definition.

Definition 4.2. A process (π_t, Y_t) with sample paths in $D_{\Pi(\mathbb{N}) \times \mathbb{N}}[0, \infty)$ solves the FMP for L if π_t is $\mathcal{F}_t^Y$-adapted and

$$\pi_t(F(\cdot, Y_t)) - \int_0^t \pi_s(LF(\cdot, Y_s)) \, ds \tag{4.5}$$

is an $\mathcal{F}_t^Y$-local martingale for any real-valued, bounded measurable function F.

As in Proposition 2.3 of [4], by a direct computation we can see that any weak solution of (4.2) provides a solution of the FMP. Thus, uniqueness for the FMP with a given initial condition implies weak uniqueness of the filtering equation.

Proposition 4.2. *Let (π, Y) be a weak solution of the filtering equation. Then (π, Y) satisfies (4.5).*

Proof. Without loss of generality, we can choose $F(x, y) = f(x)g(y)$ in (4.5). Then $\pi_t(F(\cdot, Y_t)) = \pi_t(f)g(Y_t)$, where $\pi_t(f)$ is given by (4.2) and, as in (4.1),

$$g(Y_t) = g(Y_0) + \int_0^t \sum_{h \geq 0} (g(h) - g(Y_{s-})) \, dv_s^h$$

$$= g(Y_0) + \int_0^t \sum_{h \geq 0} (g(h) - g(Y_{s-})) \, ds$$

$$+ \int_0^t \sum_{h \geq 0} (g(h) - g(Y_{s-}))(dv_s^h - \pi_{s-}(m^h(\cdot, Y_{s-})) \, ds). \tag{4.6}$$

The statement of the proposition follows by applying the product formula to (4.2) and (4.6).

Finally, uniqueness for the FMP can be proved by observing that the process (X_t, Y_t) satisfies the assumptions (H1), (H3), and (H5) of Theorem 2.8 of [4].

5. Construction of the filter

In this section, we will try to find an explicit solution of (4.2). With this in mind, at a jump time S_i, we have

$$
\pi_{S_i}(f) = \begin{cases}
\pi_{S_i-}(f) & \text{if } \pi_{S_i-}(m^{Y_{S_i}}(\cdot, Y_{S_{i-1}})) = 0, \\[2mm]
\dfrac{\pi_{S_i-}(m^{Y_{S_i}}(\cdot, Y_{S_{i-1}})f) + \pi_{S_i-}(R_X^{Y_{S_i}} f(\cdot, Y_{S_{i-1}}))}{\pi_{S_i-}(m^{Y_{S_i}}(\cdot, Y_{S_{i-1}}))} & \text{otherwise;}
\end{cases}
$$

observe that π_{S_i} is completely determined by the knowledge of π_t, $t \in [S_{i-1}, S_i)$. The problem now arises of constructing π_t for $t \in [S_i, S_{i+1})$. By Corollary 4.1,

$$
\pi_t(f) = \pi_{S_i}(f) + \int_{S_i}^t \{\pi_s(L_X^0 f(\cdot, Y_s)) - \alpha_1(1 - p_Y^1(0))[\pi_s(\varphi f) - \pi_s(\varphi)\pi_s(f)]\}\, ds,
$$

where L_X^0 is as defined in (4.4).

The uniqueness in law for the filtering equation, proved in the last section, allows us to give a computable representation for the filter, following a method proposed in [12] and [3].

Lemma 5.1. *Let $\mathcal{M}(\mathbb{N})$ be the space of positive finite measures on $\mathbb{N}$ and let ρ_t^i be an $\mathcal{M}(\mathbb{N})$-valued process solving*

$$
\rho_t^i(f) = \pi_{S_i}(f) + \int_{S_i}^t \{\rho_s^i(L_X^0 f(\cdot, Y_s)) - \alpha_1(1 - p_Y^1(0))\rho_s^i(\varphi f)\}\, ds. \tag{5.1}
$$

Then, it is easy to verify that

$$
\pi_t(f) = \frac{\rho_t^i(f)}{\rho_t^i(1)}, \qquad t \in [S_i, S_{i+1}).
$$

We describe two methods of finding a solution of (5.1) belonging to $\mathcal{M}(\mathbb{N})$. The first one is based on the Feynman–Kac formula while the second one, under the assumption that $p^1(0, 0) = 0$, provides a recursive computation of the atoms of ρ_t^i.

Method 1. For any fixed $y \in \mathbb{N}$, let ξ_t be the (unique) solution to the martingale problem for the operator L_X^0, with initial condition (s, x), $s \in \mathbb{R}^+$, $x \in \mathbb{N}$. Let $P_{(s,x)}^y$ be its law on the space $D_{\mathbb{N}}[0, \infty)$ endowed with the canonical filtration $\mathcal{F}_t$. Setting

$$
\Phi_t(s, x, y)(f) = E_{(s,x)}^y \left(f(\xi_t) \exp\left\{ -\alpha_1(1 - p_Y^1(0)) \int_s^t \xi_u\, du \right\} \right)
$$

for any real-valued, bounded measurable function f, we can prove the following properties.

(i) For $f \geq 0$, we have $\Phi_t(s, x, y)(f) \geq 0$ and, for $f > 0$, we have $\Phi_t(s, x, y)(f) > 0$.

(ii) For fixed f, y, s, and x, $\Phi_t(s, x, y)(f)$ is càdlàg.

(iii) For any f, $\Phi_t(s, x, y)(f)$ is jointly measurable with respect to (s, x, y).

Properties (i) and (ii) are quite obvious, by the definition of Φ_t. To prove property (iii), we can apply Theorem 4.6, Chapter 4 of [10] to obtain measurability with respect to the initial

data (s, x). The claimed joint measurability then follows, recalling that y belongs to a discrete space. Finally, from properties (i), (ii), and (iii) we can show that

$$\rho_t^i(f) = \int_{\mathbb{N}} \Phi_t(S_i, x, Y_{S_i})(f)\pi_{S_i}(\mathrm{d}x)$$

is a solution of (5.1) belonging to $\mathcal{M}(\mathbb{N})$, as required in Lemma 5.1.

Method 2. When $p^1(0, 0) = 0$, we set $f_j(x) = \mathbf{1}_{\{x=j\}}$ and appeal to the following proposition.

Proposition 5.1. *When* $p^1(0, 0) = 0$, *the sequence* $\{\rho_t^i(f_j)\}_{i \geq 0, j \geq 0}$ *can be recursively computed.*

Proof. The sequence $\{\rho_t^i(f_j)\}_{j \geq 0}$, $t \in [S_i, S_{i+1})$, satisfies the system of equations

$$\rho_t^i(f_j) = \pi_{S_i}(j) + \int_{S_i}^t \left[\alpha_1 j(p^1(1, 0) - 1) - \alpha_2 Y_{S_i} \sum_{k \geq 1} p^2(k, 1)\right]\rho_s^i(f_j)\,\mathrm{d}s$$

$$+ \int_{S_i}^t \sum_{k=1}^j [\alpha_1(j - k)p^1(k + 1, 0) + \alpha_2 Y_{S_i} p^2(k, 1)]\rho_s^i(f_{j-k})\,\mathrm{d}s. \qquad (5.2)$$

Actually, for suitable functions $c(j, Y_{S_i})$ and $a(j, k, Y_{S_i})$, we can write the solution of (5.2) explicitly as

$$\rho_t^i(f_j) = \pi_{S_i}(f_j) \exp\{-c(j, Y_{S_i})(t - S_i)\}$$

$$+ \sum_{k=1}^j a(j, k, Y_{S_i}) \int_{S_i}^t \rho_s^i(f_{j-k}) \exp\{-c(j, Y_{S_i})(s - S_i)\}\,\mathrm{d}s.$$

Thus, $\rho_t^i(f_j) \geq \pi_{S_i}(f_j) \exp\{-c(j, Y_{S_i})(t - S_i)\}$, which, in turn, implies that $\rho_t^i(1) > 0$. Furthermore,

$$\rho_t^i(1) = 1 - \alpha_1(1 - p_Y^1(0)) \int_{S_i}^t \rho_s^i(\varphi)\,\mathrm{d}s \leq 1. \qquad (5.3)$$

If we set $\rho_t^i(f) = \sum_{j \in \mathbb{N}} f(j)\rho_t^i(f_j)$ then, by (5.2) and (5.3), we see that $\rho_t^i \in \mathcal{M}(\mathbb{N})$ and that Lemma 5.1 applies. Then $\rho_t^i(f)/\rho_t^i(1)$ solves (4.2) for $t \in [s_i, s_{i+1})$.

In Proposition 5.1, we claim that the filter can be computed recursively, at least from a theoretical point of view. Let us observe, however, that since the proposed computation requires increasing memory (storage capacity), it does not provide an efficient computational method from a practical, numerical point of view.

Acknowledgement

References

[1] ATHREYA, K. B. AND NEY, P. E. (1972). *Branching Processes*. Springer, New York.
[2] BRÉMAUD, P. (1981). *Point Processes and Queues*. Springer, New York.
[3] CALZOLARI, A. AND NAPPO, G. (2001). Filtering of a model with grouped data and counting observations. Preprint. Available at http://www.mat.uniroma1.it/people/nappo/nappo.html.

[4] CECI, C. AND GERARDI, A. (2001). Nonlinear filtering equation of a jump process with counting observations. *Acta Appl. Math.* **66,** 139–154.

[5] CECI, C. AND GERARDI, A. (2002). Conditional law of a branching process observing a subpopulation. *J. Appl. Prob.* **39,** 112–122.

[6] CECI, C. AND MAZLIAK, L. (1992). Une propriété forte de branchements. *C. R. Acad. Sci. Paris Sér. I Math.* **315,** 851–853.

[7] CECI, C., GERARDI, A. AND MAZLIAK, L. (1996). Some results about stopping times on the marked tree space. *Theory Prob. Appl.* **41,** 578–590.

[8] CHAUVIN, B. (1991). Product martingales and stopping lines for branching Brownian motion. *Ann. Prob.* **19,** 1195–1205.

[9] ELLIOTT, R. J. (1982). *Stochastic Calculus and Applications* (Appl. Math. **18**). Springer, New York.

[10] ETHIER, S. N. AND KURTZ, T. G. (1986). *Markov Processes. Characterization and Convergence*. John Wiley, New York.

[11] JACOD, J. (1975). Multivariate point processes: predictable projection, Radon–Nikodým derivatives, representation of martingales. *Z. Wahrscheinlichkeitsth.* **31,** 235–253.

[12] KLIEMANN, W., KOCH, G. AND MARCHETTI, F. (1990). On the unnormalized solution of the filtering problem with counting process observations. *IEEE Trans. Inf. Theory* **36,** 1415–1425.

[13] KURTZ, T. G. AND OCONE, D. (1988). Unique characterization of conditional distribution in nonlinear filtering. *Ann. Prob.* **16,** 80–107.

[14] NEVEU, J. (1986). Arbres et processus de Galton–Watson. *Ann. Inst. H. Poincaré Prob. Statist.* **22,** 199–207.