

HETEROGENEOUS POPULATION DYNAMICAL MODEL: A FILTERING PROBLEM

A. GERARDI* AND

P. TARDELLI,** *Università dell'Aquila*

Abstract

We consider a heterogeneous population of identical particles divided into a finite number of classes according to their level of health. The partition can change over time, and a suitable exchangeability assumption is made to allow for having identical items of different types. The partition is not observed; we only observe the cardinality of a particular class. We discuss the problem of finding the conditional distribution of particle lifetimes, given such observations, using stochastic filtering techniques. In particular, a discrete-time approximation is given.

Keywords: Heterogeneous population; exchangeability; stochastic filtering

2000 Mathematics Subject Classification: Primary 62N05; 93E11

Secondary 60H30; 60G35; 60J75

1. Introduction

Let us consider a population of identical particles $\{U_1, \dots, U_H\}$ divided, at any time t , into different classes $\{C_1(t), \dots, C_d(t)\}$ according to their 'health' level. The partition of the population can change, since the level of health of any single particle can increase, under some kind of maintenance, or decrease as time passes. In every class, each particle can die. The partition of the population cannot be observed; instead, the observation is just the number of dead particles up to time t . If we introduce the class $C_0(t)$ of dead particles up to time t , the observation coincides with the cardinality $Y(t)$ of this class. Our aim is to find the joint law of the lifetimes of any particle, given such observations.

In this paper, the model presented in [8] and [9] has been generalized. The main difference is that there, the unobservable partition of the population was supposed to be constant in time and the particles could only die. In [9], exchangeable random variables Z_1, \dots, Z_H were introduced in order to define the partition, setting $Z_i = k$ if and only if $U_i \in C_k$. The law of the lifetimes, given the variables Z_1, \dots, Z_H , was given as data and the lifetimes were assumed to be independent, given the partition of the population.

In place of this, in this paper, in order to define the class of each particle at every time t , we introduce a stochastic process $Z(t) = \{Z_i(t)\}_{1 \leq i \leq H}$ such that $Z_i(t) = k$ if and only if $U_i \in C_k(t)$, $i = 1, \dots, H$, $k = 0, 1, \dots, d$. Furthermore, an exchangeability assumption on the finite-dimensional distribution of $Z(t)$ will be made, in order to take into account the dynamics of the model and the fact that the particles are identical. The problem of constructing a process $Z(t)$ satisfying such an exchangeability assumption will be dealt with in Section 2,

Received 1 December 2003; revision received 10 November 2004.

* Postal address: Dipartimento di Ingegneria Elettrica, Facoltà di Ingegneria, Università dell'Aquila, Aquila, Italy.

** Email address: tardelli@ing.univaq.it

where we assume that $Z(t)$ is a Markov chain and where some general properties of the model are discussed, both in the discrete- and continuous-time cases.

In order to find the law of the lifetimes given $\{Y(s), s \leq t\}$, the number of particles belonging to the class C_0 up to time t , we first compute the law of the lifetimes given the variable $Z(t)$, and then find the filter of $Z(t)$ given $\{Y(s), s \leq t\}$. The continuous-time filtering problem is studied in Section 4. Strong uniqueness of the filtering equation is demonstrated, a linearized equation is introduced, and a useful representation of its solution is derived following a method which is a modification of that proposed in [10] and in [5].

In [9], an explicit solution for the filtering equation was given. In our model, an elementary recursive formula for the filter is set up, for the discrete-time model only, in Section 3. Then, in the final sections, we construct an approximate discrete-time model following some ideas given in [4]. Finally, we prove that the discrete-time model and its filter converge to the exact filter. The convergence of the filters is proven by using a linearization technique for the discrete-time filtering equation.

2. The model

Let us consider a finite or countable population $\mathcal{P} = \{U_j\}_{j \geq 1}$, where U_j are given particles, and, for H a positive integer, let $\mathcal{P}_H = \{U_j\}_{j=1, \dots, H}$ be a finite subpopulation of \mathcal{P} . Let \mathcal{P}_H be heterogeneous in that its elements can be of d different types, labelled by the natural numbers $1, \dots, d$. Given $t \in \mathbb{R}^+$, let $C_k(t)$, $k = 1, \dots, d$, be the subset of all individuals of type k at time t . Finally, let $C_0(t)$ denote the class of the particles dead by time t . Thus, $\mathcal{P}_H = \bigcup_{k=0,1, \dots, d} C_k(t)$ and we are interested in the case in which events of the form $\{U_i \in C_k(t)\}$, for $i = 1, \dots, H$ and for any $t > 0$, are measurable.

To this end, let us introduce the random variable $Z(t) = \{Z_i(t)_{i=1, \dots, H}\}$ defining the type of each particle:

$$\text{for all } t \in \mathbb{R}^+ \text{ and } k = 0, 1, \dots, d, \quad Z_i(t) = k \Leftrightarrow U_i \in C_k(t) \quad \text{for } i = 1, \dots, H.$$

By definition, $C_0(t)$ is an absorbing class in the sense that if a particle U_i enters $C_0(t)$, it can never leave. That is, for all $i = 1, \dots, H$, we assume that

$$Z_i(s) = 0 \Rightarrow Z_i(t) = 0 \quad \text{for all } t \geq s, \text{ almost surely.} \tag{2.1}$$

Then, for all $t \in \mathbb{R}^+$, $Z_i(t)$ is a random variable taking values in $\{0, 1, \dots, d\}$ and $Z(t) = (Z_1(t), \dots, Z_H(t))$ takes values in $\mathcal{H} = \{0, 1, \dots, d\}^H$.

In the construction of the model, we have to take into account the fact that different particles can have different labels even if they are considered indistinguishable. Thus, because we are dealing with a dynamical model, we introduce an exchangeability property involving the finite-dimensional distribution of the process $Z(t)$. That is, for all $n \geq 1$, all $t_1, \dots, t_n \in \mathbb{R}^+$ with $t_1 \leq \dots \leq t_n$, all permutations π of the set $\{1, \dots, H\}$, and all $k^{(1)}, \dots, k^{(n)} \in \mathcal{H}$,

$$P(Z(t_1) = k^{(1)}, \dots, Z(t_n) = k^{(n)}) = P(Z(t_1) = \pi k^{(1)}, \dots, Z(t_n) = \pi k^{(n)}), \tag{2.2}$$

where $\pi k^{(i)} = \{k_{\pi_1}^{(i)}, \dots, k_{\pi_H}^{(i)}\}$, $i = 1, \dots, n$, for $k^{(i)} = \{k_1^{(i)}, \dots, k_H^{(i)}\} \in \mathcal{H}$.

In particular, for any fixed t , $(Z_1(t), \dots, Z_H(t))$ is an exchangeable sequence; that is, for all $a_1, \dots, a_H \in \mathcal{H}$ and all finite permutations π of the indices,

$$P(Z_1(t) = a_1, \dots, Z_H(t) = a_H) = P(Z_1(t) = a_{\pi(1)}, \dots, Z_H(t) = a_{\pi(H)}), \tag{2.3}$$

which is a generalization of the exchangeability property given in [8] and [9].

2.1. The lifetimes

Later, we will construct a process verifying that conditions (2.1) and (2.2) are satisfied. Now, we are going to discuss some consequences of these conditions. The most important is related to lifetimes.

Let T_i be the lifetime of $U_i, i = 1, \dots, H$, defined as $T_i = \inf\{t \in \mathbb{R}^+ : Z_i(t) = 0\}$. Taking into account the structural properties of the model, as in [8] and [9], the lifetimes are a sequence of exchangeable random variables. The assumption (2.2), which is very natural in our context, implies such a result, while (2.3) does not.

Theorem 2.1. *Under the assumptions (2.1) and (2.2), $\{T_i\}_{i \geq 1}$ is a family of exchangeable random variables.*

Proof. Let us compute $P(T_1 \leq t_1, \dots, T_H \leq t_H)$ for all $t_1, \dots, t_H \in \mathbb{R}^+$. Without loss of generality, we assume that $t_1 \leq \dots \leq t_H$. Otherwise, we consider a permutation of indices such that $t_{1'} \leq \dots \leq t_{H'}$, and note that $\{T_1 \leq t_1, \dots, T_H \leq t_H\} = \{T_{1'} \leq t_{1'}, \dots, T_{H'} \leq t_{H'}\}$. Therefore, recalling (2.1), for all $n = 1, \dots, H$,

$$\begin{aligned}
 &P(T_1 \leq t_1, \dots, T_n \leq t_n) \\
 &= P(Z_1(t_1) = 0, Z_2(t_2) = 0, \dots, Z_n(t_n) = 0) \\
 &= P(Z_1(t_1) = 0, Z_1(t_2) = 0, \dots, Z_1(t_n) = 0, \\
 &\quad Z_2(t_2) = 0, \dots, Z_2(t_n) = 0, \\
 &\quad \quad \quad \vdots \\
 &\quad \quad \quad Z_n(t_n) = 0) \\
 &= \sum_{x_{i,j}; 2 \leq i \leq H; 1 \leq j \leq i-1} P(Z(t_1) = (0, x_{2,1}, \dots, x_{n,1}), \\
 &\quad \quad \quad Z(t_2) = (0, 0, x_{3,2}, \dots, x_{n,2}), \dots, Z(t_n) = (0, \dots, 0)) \\
 &= \sum_{x_{i,j}; 2 \leq i \leq H; 1 \leq j \leq i-1} P(Z(t_1) = \pi(0, x_{2,1}, \dots, x_{n,1}), \\
 &\quad \quad \quad Z(t_2) = \pi(0, 0, x_{3,2}, \dots, x_{n,2}), \dots, Z(t_n) = \pi(0, \dots, 0)) \\
 &= P(Z_{\pi^{-1}(1)}(t_1) = 0, Z_{\pi^{-1}(2)}(t_2) = 0, \dots, Z_{\pi^{-1}(n)}(t_n) = 0) \\
 &= P(T_{\pi^{-1}(1)} \leq t_1, \dots, T_{\pi^{-1}(n)} \leq t_n).
 \end{aligned}$$

In the sequel, we assume that $Z(t)$ has the Markov property. Therefore, we have to solve the problem of constructing the Markov process $Z(t)$ satisfying the conditions (2.1) and (2.2). As we pointed out in the introduction, we are going to study the properties of the Markov chain $Z(t)$, the partition of the subpopulation at time t , and find the law of the lifetimes, given $Z(t)$.

The constructions of the discrete-time model $Z(t)$ and of the continuous-time one $\zeta(t)$ are fairly standard. For the sake of self-consistency, and since we will use some definitions in the sequel, we are going to give a sketch of the construction.

2.2. The discrete-time model

In the discrete-time case ($t \in \mathbb{N}$), the existence of a process satisfying the conditions (2.1) and (2.2) is ensured by the Markov property. Let $\mu(a, b), a, b \in \mathcal{H}$, be a family of transition probabilities and let ν_0 be a probability measure on \mathcal{H} . Let $Z(t), t \in \mathbb{N}$, be the Markov chain with initial distribution ν_0 such that, for all $t > 0$ and all $a = (a_1, \dots, a_H), b = (b_1, \dots, b_H) \in \mathcal{H}$, we have $\mu(a, b) = P(Z(t) = b \mid Z(t - 1) = a)$.

We can easily verify that the following assumptions guarantee that the process $Z(t)$ satisfies (2.2) whenever $Z(0)$ satisfies (2.3):

- there exists an i such that $a_i = 0$ and $b_i \neq 0$ imply that $\mu(a, b) = 0$; (2.4)

- for all $a, b \in \mathcal{H}$ and all finite permutations π of the indices, $\mu(a, b) = \mu(\pi(a), \pi(b))$. (2.5)

2.3. The continuous-time model

We now introduce the continuous-time model by defining a continuous-time Markov process $\zeta(t)$ with generator L given, for real-valued, bounded measurable functions f on \mathcal{H} , by

$$Lf(z) = l(z) \sum_{z' \in \mathcal{H}} [f(z') - f(z)]p(z, z'), \tag{2.6}$$

where $\{p(z, z')\}$ is a family of transition probabilities satisfying (2.4) and (2.5), and $l(z) = l(\pi(z))$ is a positive function. Since \mathcal{H} is a finite set, there exist $\underline{l}, \bar{l} \in \mathbb{R}^+$ such that $0 < \underline{l} \leq l(z) \leq \bar{l}$ for all $z \in \mathcal{H}$.

By construction, L is clearly a bounded operator. Hence, for all $\nu_0 \in \Pi(\mathcal{H})$ (where $\Pi(\mathcal{H})$ is the space of probability measures on \mathcal{H}), there exists a unique Markov process $\zeta(t)$ with ν_0 as initial condition, L as generator, and sample paths in $D_{\mathcal{H}}[0, \infty)$, the space of \mathcal{H} -valued càdlàg functions defined on $[0, \infty)$.

Following a classical construction (see, for instance, [7]), we are going to provide a particular representation for $\zeta(t)$ that we will use to obtain the approximating model in Section 5. Given a probability space (Ω, \mathcal{F}, P) , let $\{Z(n)\}_{n \geq 0}$ be a Markov chain with initial distribution ν_0 and transition probabilities $p(z, z')$. Let $\{V_i\}_{i \geq 1}$ be a sequence of independent random variables exponentially distributed with parameter 1 and independent of $\{Z(n)\}_{n \geq 0}$. Finally, let

$$\tau_0 = 0, \quad \tau_n = \sum_{i=1}^n \frac{V_i}{l(Z(i-1))} \quad \text{for } n > 0, \quad \zeta(t) = \sum_{n \geq 0} Z(n) \mathbf{1}_{\{\tau_n \leq t < \tau_{n+1}\}}, \tag{2.7}$$

and $\mathcal{F}_t = \sigma\{\zeta(s), s \leq t\}$. Hence, on the space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, P)$, $\zeta(t)$ is a continuous-time pure-jump Markov process with generator L satisfying (2.1), and $\{\tau_i\}_{i \geq 1}$ is the sequence of its jump times. By using this particular construction, we can easily see that the process $\zeta(t)$ satisfies (2.2) whenever $\zeta(0)$ satisfies (2.3).

3. Conditional laws

In this section, we will obtain the conditional law of the lifetimes, given the history of the process $Y(t)$ (which is the cardinality of the class $C_0(t)$). To this end, as already observed, we first compute the law of the lifetimes, given $Z(t)$ (or $\zeta(t)$ in the continuous-time case), and then we set up a filtering problem to obtain the distribution of $Z(t)$ (or $\zeta(t)$), given the history of the process $Y(t)$. Setting $\Phi(z) = \sum_{i=1}^H \mathbf{1}_{\{z_i=0\}}$ for all $z \in \mathcal{H}$, we find that, for the discrete-time model, $Y(t) = \Phi(Z(t))$ and, for the continuous-time model, $Y(t) = \Phi(\zeta(t))$. The distribution of the lifetimes, given the partition of the population at time t , can be easily computed. If we again assume that $t_1 \leq \dots \leq t_H$ for the discrete-time model then, if t is such that $t_1 \leq t_2 \leq \dots \leq t_m \leq t \leq t_{m+1} \leq \dots \leq t_H$, for $P(Z(t) = k) \neq 0, k \in \mathcal{H}$, we obtain

$$\begin{aligned} &P(T_1 \leq t_1, T_2 \leq t_2, \dots, T_H \leq t_H \mid Z(t) = k) \\ &= \frac{P(Z_1(t_1) = 0, \dots, Z_m(t_m) = 0, Z(t) = k, Z_{m+1}(t_{m+1}) = 0, \dots, Z_H(t_H) = 0)}{P(Z(t) = k)}. \end{aligned}$$

A similar formula can be written down for the continuous-time model.

Here, the important point is that this distribution can be evaluated explicitly as a function of $k \in \mathcal{H}$. This is important because, in order to find the law of the lifetimes given the observed failures, only the law of $Z(t)$ (or $\zeta(t)$) given $\{Y(s), s \leq t\}$ is needed. In the next section, we will discuss the filtering problem in the continuous-time case. The last part of this section is devoted to the filtering problem for the discrete-time model.

We set $\pi_t(z) = P(Z(t) = z \mid \mathcal{F}_t^Y)$, where $\mathcal{F}_t^Y = \sigma\{Y(s), s \leq t\}$. Now, since $Z(t)$ is a Markov chain and $Y = \Phi(Z)$, the pair (Z, Y) is still Markov and $\pi_t(z)$, the filter, can be written using a Bayes formula as

$$\pi_t(z) = \frac{\sum_{z' \in \mathcal{H}} P(Z(t) = z, Y(t) \mid Z(t-1) = z', Y(t-1)) \pi_{t-1}(z')}{\sum_{z'', z''' \in \mathcal{H}} P(Z(t) = z'', Y(t) \mid Z(t-1) = z'', Y(t-1)) \pi_{t-1}(z''')},$$

where the law of $Z(0)$ is known. Therefore,

$$\begin{aligned} \pi_t(z) &\propto \sum_{z_{t-1}} \mathbf{1}_{\{Y(t-1) = \Phi(z_{t-1})\}} \mathbf{1}_{\{Y(t) = \Phi(z)\}} P(Z(t) = z \mid Z(t-1) = z_{t-1}) \pi_{t-1}(z_{t-1}) \\ &= \mathbf{1}_{\{Y(t) = \Phi(z)\}} \sum_{z_{t-1} \in \Phi^{-1}(Y(t-1))} \mu(z_{t-1}, z) \pi_{t-1}(z_{t-1}). \end{aligned} \tag{3.1}$$

By induction, we find that

$$\pi_t(z) \propto \mathbf{1}_{\{Y(t) = \Phi(z)\}} \sum_{z_{t-1} \in \Phi^{-1}(Y(t-1)), \dots, z_0 \in \Phi^{-1}(0)} \mu(z_0, z_1) \mu(z_1, z_2) \cdots \mu(z_{t-1}, z) P(Z(0) = z_0),$$

which could be also conjectured intuitively.

Remark 3.1. Summing up, we explicitly write

$$\begin{aligned} P(T_1 \leq t_1, T_2 \leq t_2, \dots, T_H \leq t_H \mid \mathcal{F}_t^Y) \\ = \sum_{z \in \mathcal{H}} P(T_1 \leq t_1, T_2 \leq t_2, \dots, T_H \leq t_H \mid Z(t) = z) \pi_t(z). \end{aligned}$$

An analogous result can be reached in the continuous-time model once $\pi_t(z)$ is computed, which will be done in the next section.

4. Filtering in the continuous-time case

Let us recall that the filter $\pi_t(f) = E[f(\zeta(t)) \mid \mathcal{F}_t^Y]$ satisfies a stochastic differential equation known as the Kushner–Stratonovich equation.

4.1. Filtering equation

In this subsection, we write down the Kushner–Stratonovich equation and, in the next one, we discuss uniqueness of its solutions.

Since, in general, $Y(t)$ is not a counting process, following an idea presented in [5] we introduce the multivariate point processes $U(t) = (U^1(t), \dots, U^H(t))$, defined as

$$U^j(t) := \sum_{i=1}^H \mathbf{1}_{\{T_{(i)} \leq t\}} \mathbf{1}_{\{Y(T_{(i)}) = j\}}, \quad j = 1, \dots, H,$$

where $T_{(1)}, \dots, T_{(H)}$, the order statistics of the lifetimes, give us the jump times of Y . The process $U^j(t)$ counts the number of jumps bringing $Y(t)$ to j , and

$$N_t = \sum_{j=1}^H U^j(t) = \sum_{i=1}^H \mathbf{1}_{\{T_{(i)} \leq t\}} \tag{4.1}$$

counts all the jumps of Y , up to time t . Since, by definition, $Y(t)$ is a nondecreasing process, $U^j(t)$ is $\{0, 1\}$ -valued for all j . Furthermore, the relation

$$Y(t) = Y(0) + \int_0^t \sum_{j=1}^H [j - Y(s-)] dU^j(s)$$

implies that $\mathcal{F}_t^Y = \mathcal{F}_t^U$, where, as usual,

$$\mathcal{F}_t^Y = \sigma\{Y(s), s \leq t\} \quad \text{and} \quad \mathcal{F}_t^U = \sigma\{U^1(s), \dots, U^H(s), s \leq t\}.$$

Therefore, our problem reduces to finding the filter that is the conditional law of $\zeta(t)$ given \mathcal{F}_t^U .

Proposition 4.1. *For the continuous-time model presented in this paper, the equation for the filter, for any real-valued function $f(z)$, $z \in \mathcal{H}$, can be written as*

$$\begin{aligned} \pi_t(f) = & v_0(f) + \int_0^t \{\pi_s(Gf) - \pi_s(mf) + \pi_s(m)\pi_s(f)\} ds \\ & + \sum_{j=1}^H \int_0^t (\pi_{s-}(m_j))^+ \{\pi_{s-}(m_j f) - \pi_{s-}(m_j)\pi_{s-}(f) + \pi_{s-}(R_j f)\} dU_s^j, \end{aligned} \tag{4.2}$$

where $a^+ = (1/a) \mathbf{1}_{\{a>0\}}$ as usual in filtering theory, and where

$$\begin{aligned} Gf(z) &= l(z) \sum_{z'} [f(z') - f(z)] \mathbf{1}_{\{\Phi(z')=\Phi(z)\}} p(z, z'), \\ R_j f(z) &= l(z) \sum_{z'} [f(z') - f(z)] \mathbf{1}_{\{\Phi(z') \neq \Phi(z)\}} \mathbf{1}_{\{\Phi(z')=j\}} p(z, z'), \\ m_j(z) &= l(z) \sum_{z'} \mathbf{1}_{\{\Phi(z') \neq \Phi(z)\}} \mathbf{1}_{\{\Phi(z')=j\}} p(z, z'), \\ m(z) &= \sum_{j=1}^H m_j(z) = l(z) \sum_{z'} \mathbf{1}_{\{\Phi(z') \neq \Phi(z)\}} p(z, z'). \end{aligned}$$

Proof. The classical innovation method allows us to write down an equation for the filter. In particular, our model is such that the results obtained in [2] apply and we refer to them for the consistency of our proof. Let us observe that the process (ζ, Y, U) is still Markov and that its joint generator is, for $z \in \mathcal{H}$, $u \in \{0, 1\}^H$, and $y \in \{0, 1, \dots, H\}$,

$$\mathcal{L}f(z, y, u) = \mathcal{L}_0 f(z, y, u) + \sum_{j=1}^H \mathcal{L}_1^j f(z, y, u),$$

where

$$\begin{aligned} \mathcal{L}_0 f(z, y, u) &= l(z) \sum_{z' \in \mathcal{H}} [f(z', y, u) - f(z, y, u)] \mathbf{1}_{\{\Phi(z') = \Phi(z)\}} p(z, z'), \\ \mathcal{L}_1^j f(z, y, u) &= l(z) \sum_{z' \in \mathcal{H}} [f(z', y + \Phi(z') - \Phi(z), u + e^j) - f(z, y, u)] \\ &\quad \times \mathbf{1}_{\{\Phi(z') \neq \Phi(z)\}} \mathbf{1}_{\{y + \Phi(z') - \Phi(z) = j\}} p(z, z'), \end{aligned}$$

and e^j is the vector such that $e_i^j = \delta_i^j, i = 1, \dots, H$. The generator \mathcal{L} , when restricted to the function depending only on the first variable z , coincides with the operator L given in (2.6). Moreover, the $(\mathbb{P}, \mathcal{F}_t)$ -intensity of U^j is given in terms of the process $\tilde{m}_j(\zeta(t), Y(t))$, where

$$\begin{aligned} \tilde{m}_j(z, y) &= l(z) \sum_{z' \in \mathcal{H}} [u_j + e_j^j - u_j] \mathbf{1}_{\{\Phi(z') \neq \Phi(z)\}} \mathbf{1}_{\{y + \Phi(z') - \Phi(z) = j\}} p(z, z') \\ &= l(z) \sum_{z' \in \mathcal{H}} \mathbf{1}_{\{\Phi(z') \neq \Phi(z)\}} \mathbf{1}_{\{y + \Phi(z') - \Phi(z) = j\}} p(z, z'). \end{aligned}$$

More precisely, the $(\mathbb{P}, \mathcal{F}_t^U)$ -intensity of U^j is given by $\pi_t(\tilde{m}_j) = E[\tilde{m}_j(Y(t), \zeta(t)) \mid \mathcal{F}_t^Y] = \pi_t(m_j)$. The latter equality holds since, by definition, $Y(t) = \Phi(\zeta(t))$ for all t and, so,

$$\begin{aligned} \tilde{m}_j(\zeta(t), Y(t)) &= l(\zeta(t)) \sum_{z' \in \mathcal{H}} \mathbf{1}_{\{\Phi(z') \neq \Phi(\zeta(t))\}} \mathbf{1}_{\{Y(t) + \Phi(z') - \Phi(\zeta(t)) = j\}} p(\zeta(t), z') \\ &= l(\zeta(t)) \sum_{z' \in \mathcal{H}} \mathbf{1}_{\{\Phi(z') \neq \Phi(\zeta(t))\}} \mathbf{1}_{\{\Phi(z') = j\}} p(\zeta(t), z') \\ &= m_j(\zeta(t)). \end{aligned}$$

Let us define

$$M(t) := f(\zeta(t)) - f(\zeta(0)) - \int_0^t Lf(\zeta(s)) ds,$$

which is a zero-mean $(\mathbb{P}, \mathcal{F}_t)$ -martingale. First, using [2, Theorem T1, Chapter IV, p. 87], we are able to write

$$\pi_t(f) = \nu_0(f) + \int_0^t \pi_s(Lf) ds + \hat{M}_t,$$

where \hat{M}_t is a zero-mean $(\mathbb{P}, \mathcal{F}_t^Y)$ -martingale. By the representation theorem [2, Theorem T17, Chapter III, p. 76], this can be written in turn as

$$\hat{M}_t = \sum_{j=1}^H \int_0^t K_s(j) (dU_s^j - \pi_{s-}(m_j) ds),$$

where $\{K_s(j)\}$ is a suitable sequence of $(\mathbb{P}, \mathcal{F}_t^Y)$ -predictable processes. Finally (see [2, Theorem T2, Chapter IV, p. 91]), the equation for the filter is then given by

$$\begin{aligned} \pi_t(f) &= \nu_0(f) + \int_0^t \pi_s(Lf) ds \\ &\quad + \sum_{j=1}^H \int_0^t \pi_{s-}(m_j)^+ \{ \pi_{s-}(m_j f) - \pi_{s-}(m_j) \pi_{s-}(f) + \pi_{s-}(R^j f) \} (dU_s^j - \pi_{s-}(m_j) ds) \end{aligned} \tag{4.3}$$

for any function $f(z), z \in \mathcal{H}$. In order to compute $R^j f$ and Gf , let us recall that

$$M(t) := f(\zeta(t)) - f(\zeta(0)) - \int_0^t Lf(\zeta(s)) \, ds \quad \text{and} \quad M^j(t) := U^j(t) - \int_0^t m_j(\zeta(s)) \, ds$$

are $(\mathbb{P}, \mathcal{F}_t)$ -martingales, and that $\langle M, M^j \rangle_t = \int_0^t R^j f(\zeta(s)) \, ds$. By standard stochastic calculus, we find that

$$R^j f(z) = l(z) \sum_{z'} [f(z') - f(z)] \mathbf{1}_{\{\Phi(z') \neq \Phi(z)\}} \mathbf{1}_{\{\Phi(z')=j\}} p(z, z').$$

Finally, we complete the proof by substituting

$$Gf(z) = Lf(z) - \sum_{j=1}^H R^j f(z) = l(z) \sum_{z'} [f(z') - f(z)] \mathbf{1}_{\{\Phi(z')=\Phi(z)\}} p(z, z')$$

into (4.3).

Remark 4.1. Let us remark that (4.2) depends only on the joint dynamics of the process (ζ, U) . This is in some sense intuitive, since $Y(t) = \Phi(\zeta(t))$ and $\Phi(\cdot)$ is a deterministic function.

4.2. Uniqueness

In general, (4.2) does not have a unique solution. Thus, in order to deduce the properties of the filter using (4.2), we need to find some kind of uniqueness property for it. Weak uniqueness could be obtained, as in [4], [6], [10], [11], and [12], by using the filtering martingale problem approach and, in particular, taking into account Proposition 2.4 of [6]. On the other hand, since this model has a finite state space, a stronger kind of uniqueness can be obtained. In fact, we are going to prove that there is pathwise uniqueness for the solutions of the filtering equation. The procedure used here has some similarities with that used in [1].

We note that in this procedure we do not require $\pi_t(m_j) > 0$.

Theorem 4.1. *Let π'_t be an \mathcal{F}_t^U -adapted, probability-measure-valued process with càdlàg trajectories, satisfying (4.2) driven by the process $U(t)$. Then π'_t coincides pathwise with the filter.*

Proof. The filter π_t satisfies (4.2). Thus, taking into account the fact that

$$\begin{aligned} & |\pi_{s-}(m_j)^+ - \pi'_{s-}(m_j)^+| \cdot |\pi'_{s-}(R^j f + m_j f) - \pi'_{s-}(m_j)\pi'_{s-}(f)| \\ &= |\pi_{s-}(m_j) - \pi'_{s-}(m_j)| \pi_{s-}(m_j)^+ \pi'_{s-}(m_j)^+ |\pi'_{s-}(R^j f + m_j f) - \pi'_{s-}(m_j)\pi'_{s-}(f)| \\ &\leq |\pi_{s-}(m_j) - \pi'_{s-}(m_j)| \pi_{s-}(m_j)^+ 2 \|f\| \bar{l}, \end{aligned}$$

for a suitable positive constant C we have

$$|\pi_t(f) - \pi'_t(f)| \leq C \|f\| \bar{l} \left\{ \int_0^t \|\pi_s - \pi'_s\| \, ds + \sum_{j=1}^H \int_0^t \|\pi_{s-} - \pi'_{s-}\| (\pi_{s-}(m_j))^+ \, dU_s^j \right\},$$

where $\|\pi - \pi'\| := \sup_{f \in B(\mathcal{H})} |\pi_t(f) - \pi'_t(f)| / \|f\|$, $B(\mathcal{H})$ denotes the space of real-valued, bounded measurable functions on \mathcal{H} , and $\|f\| := \sup_{z \in \mathcal{H}} |f(z)|$. Since $\pi_{s-}(m_j)$ is the predictable $(\mathbb{P}, \mathcal{F}_t^U)$ -intensity of U_s^j , we obtain

$$\begin{aligned} \mathbb{E} \left[\int_0^t \|\pi_{s-} - \pi'_{s-}\| (\pi_{s-}(m_j))^+ dU_s^j \right] &= \mathbb{E} \left[\int_0^t \|\pi_s - \pi'_s\| (\pi_s(m_j))^+ \pi_s(m_j) ds \right] \\ &\leq \mathbb{E} \left[\int_0^t \|\pi_s - \pi'_s\| ds \right]. \end{aligned}$$

Thus, for another suitable constant C' , we have

$$\mathbb{E} \|\pi_t - \pi'_t\| \leq C' \int_0^t \mathbb{E} \|\pi_s - \pi'_s\| ds$$

and, so, $\mathbb{E} \|\pi_t - \pi'_t\| = 0$. This, in turn, implies that $\mathbb{P}(\|\pi_t - \pi'_t\| > 0) = 0$ for all $t > 0$. Since $\|\pi_t - \pi'_t\|$ has càdlàg trajectories,

$$\mathbb{P} \left(\sup_{0 \leq t \leq T} \|\pi_t - \pi'_t\| > 0 \right) = \mathbb{P} \left(\sup_{\mathcal{T}} \|\pi_t - \pi'_t\| > 0 \right) \leq \sum_{t \in \mathcal{T}} \mathbb{P}(\|\pi_t - \pi'_t\| > 0) = 0$$

for any countable dense subset \mathcal{T} in $(0, T)$. Therefore, each solution to the filtering equation coincides pathwise with the filter.

Let us recall that the transition probabilities $p(z, z')$ have to satisfy conditions (2.4) and (2.5). From now on, we need the further assumption that,

$$\text{for all } z, z' \in \mathcal{H}, \text{ if both } z_i = 0 \text{ and } z'_i = 0 \text{ for some } i, \text{ then } p(z, z') > 0. \tag{4.4}$$

Roughly speaking, this means that any transition must be possible, provided that none of the dead particles have been brought back to life.

This assumption allows us to define the positive quantity

$$\underline{p} := \min\{p(z, z') : z, z' \in \mathcal{H} \text{ satisfy (4.4)}\}.$$

Remark 4.2. As a consequence of the last assumption, we find that, if $j > \Phi(z)$ (with $j = 1, \dots, H$) and $z \in \mathcal{H}$, then

$$m_j(z) \geq \underline{p} \sum_{z' \in \mathcal{H}} \mathbf{1}_{\{\Phi(z') > \Phi(z)\}} \mathbf{1}_{\{\Phi(z') = j\}} p(z, z') \geq \underline{p}.$$

4.3. A representation for the filter

Explicit expressions for the solutions to (4.2) are not available although, in the discrete-time model, a simple recursive formula is given by (3.1). For this reason, in the next section we will construct an approximate discrete-time model and, in the last section, prove that the filter of the approximate discrete-time model converges to the exact one.

To achieve this goal, it is useful to introduce a linearized equation, using a method that is a modification of that proposed in [10]. The linearized equation

$$\rho_t(f) = \nu_0(f) + \int_0^t \{\rho_s(Gf) - \rho_s(mf)\} ds + \sum_{j=0}^H \int_0^t \{\rho_{s-}(m_j f) - \rho_{s-}(f) + \rho_{s-}(R_j f)\} dU_s^j \tag{4.5}$$

is a modification of (4.2). With classical arguments, we can prove that (4.5) not only admits a unique solution in the weak sense but, moreover, by Lipschitz arguments, that it admits a unique strong solution that is necessarily \mathcal{F}_t^U -adapted. However, we restrict our attention to the results that we need in the sequel, which are given by the following proposition.

Proposition 4.2. *Equation (4.5) admits at least one \mathcal{F}_t^U -adapted solution. Such a solution $\rho_t(f)$, for any t , has the following properties:*

- (i) $\rho_t(f)$ is a finite positive measure;
- (ii) $e^{-t\bar{l}}(1 \wedge \underline{lp}) < \rho_t(1) \leq \bar{l} \vee 1$; and
- (iii) $\pi_t(f) = \rho_t(f)/\rho_t(1)$.

Proof. First we claim that, for any solution ρ_t to (4.5), $\rho_t(f)/\rho_t(1)$ provides a solution of (4.2) and coincides with the filter up to time $t_0 = \inf\{t \geq 0: \rho_t(1) = 0\}$. Then we construct a solution of (4.5) that has the required properties and such that $\rho_t(1) > 0$ for any t .

Let $X(t)$ be a process with initial condition (s, x) , $s \geq 0$, $x \in \mathcal{H}$, and generator G . Let $P_{s,x}$ be its law on $D_{\mathcal{H}}[s, T]$. Then, by the Feynman–Kac formula, for all $t \in [0, T_{(1)}]$,

$$\begin{aligned} \rho_t(f) &= \sum_{x \in \mathcal{H}} E^{P_{0,x}} \left[f(X_t) \exp \left\{ - \int_0^t m(X_u) du \right\} \right] \nu_0(\{x\}), \\ \rho_t(1) &= \sum_{x \in \mathcal{H}} E^{P_{0,x}} \left[\exp \left\{ - \int_0^t m(X_u) du \right\} \right] \nu_0(\{x\}) \geq e^{-t\bar{l}} > 0, \end{aligned}$$

and, for $t = T_{(1)}$,

$$\begin{aligned} \rho_{T_{(1)}}(f) &= \rho_{T_{(1)}-}(m_j f) + \rho_{T_{(1)}-}(R_j f)|_{j=Y_{T_{(1)}}}, \\ \rho_{T_{(1)}}(1) &= \rho_{T_{(1)}-}(m_j)|_{j=Y_{T_{(1)}}} \geq \underline{lp} \rho_{T_{(1)}-}(1), \end{aligned}$$

where the last inequality is a consequence of Remark 4.2. The statement of the proposition then follows by induction. In fact, for all $t \in [T_{(i)}, T_{(i+1)})$,

$$\begin{aligned} \rho_t(f) &= \sum_{x \in \mathcal{H}} E^{P_{s,x}} \left[f(X_t) \exp \left\{ - \int_s^t m(X_u) du \right\} \right] \Big|_{s=T_{(i)}} \frac{\rho_{T_{(i)}}(\{x\})}{\rho_{T_{(i)}}(1)}, \\ \rho_t(1) &= \sum_{x \in \mathcal{H}} E^{P_{s,x}} \left[\exp \left\{ - \int_s^t m(X_u) du \right\} \right] \Big|_{s=T_{(i)}} \frac{\rho_{T_{(i)}}(\{x\})}{\rho_{T_{(i)}}(1)} \geq e^{-(t-T_{(i)})\bar{l}} > 0, \end{aligned}$$

and, for $t = T_{(i+1)}$,

$$\begin{aligned} \rho_{T_{(i+1)}}(f) &= \rho_{T_{(i+1)}-}(m_j f) + \rho_{T_{(i+1)}-}(R_j f)|_{j=Y_{T_{(i+1)}}}, \\ \rho_{T_{(i+1)}}(1) &= \rho_{T_{(i+1)}-}(m_j)|_{j=Y_{T_{(i+1)}}} \geq \underline{lp} \rho_{T_{(i+1)}-}(1). \end{aligned}$$

Finally, observe that $\rho_t(1) \leq 1$ for $t \in [T_{(i)}, T_{(i+1)})$ and, for $t = T_{(i+1)}$, that

$$\rho_{T_{(i+1)}}(1) = \rho_{T_{(i+1)}-}(m_j)|_{j=Y_{T_{(i+1)}}} \leq \bar{l} \rho_{T_{(i+1)}-}(1) \leq \bar{l}.$$

This result will provide a central tool for the comparison between the approximate and the original filter, performed in the next sections.

5. The approximate model

The construction of the approximate process follows the same lines as in [4] and is related to the construction in Section 2. Let $h > 0$ be fixed and let

$$\begin{aligned} \theta_0^h &:= 0, & \theta_i^h &:= \left[\frac{V_i}{hl(Z(i-1))} \right] + 1 \quad \text{for } i > 0, \\ \tau_0^h &:= 0, & \tau_n^h &= h \sum_{i=1}^n \theta_i^h \quad \text{for } n > 0. \end{aligned}$$

Then, on a finite time horizon $[0, T]$ with $T > 0$, for $t = kh, k = 0, 1, \dots$, such that $kh \leq T$, the approximate process is defined as

$$\zeta^h(t) = \sum_{n \geq 0} Z(n) \mathbf{1}_{\{\tau_n^h \leq t < \tau_{n+1}^h\}},$$

where $Z(n)$ is the Markov chain defined in (2.7). Hence, on the space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P})$, $\zeta^h(t)$ is a discrete-time Markov chain and $\{\tau_i^h\}_{i \geq 1}$ is the sequence of its jump times. The following result is given in [4].

Proposition 5.1. *The process $\zeta^h(t)$ is a discrete-time Markov chain with transition probabilities given by*

$$\mu^h(z, z') = \mathbb{P}(\zeta^h((n+1)h) = z' \mid \zeta^h(nh) = z) = \delta_{\{z, z'\}} e^{-hl(z)} + p(z, z')(1 - e^{-hl(z)}). \tag{5.1}$$

We now introduce the discrete-time observations process, setting

$$Y^h(t) = \Phi(\zeta^h(t)), \quad U_t^{jh} = \sum_{k=1}^H \mathbf{1}_{\{T_k^h \leq t\}} \mathbf{1}_{\{Y^h(T_k^h) = j\}} \quad \text{and} \quad N_t^h = \sum_{j=0}^H U_t^{jh} = \sum_{k=1}^H \mathbf{1}_{\{T_k^h \leq t\}},$$

where $\{T_k^h\}$ is the sequence of jump times of $Y^h(t)$.

As in the continuous-time case, we can prove that $\mathcal{F}_t^{Y^h} = \mathcal{F}_t^{U^h}$, with $U^h = \{U_t^{jh}\}_{j=0,1,\dots,H}$. Hence, we have to solve a filtering problem in order to find the conditional law of $\zeta^h(t)$ given $\mathcal{F}_t^{U^h}$. We know that $\pi_t^h = \mathcal{L}(\zeta^h(t) \mid \mathcal{F}_t^{U^h})$ satisfies the equation

$$\begin{aligned} &\pi_{nh}^h(f) \\ &= v_0(f) + \sum_{k=1}^n (1 - \pi_{(k-1)h}^h(m^h))^+ \\ &\quad \times [\pi_{(k-1)h}^h(G^h f) + \pi_{(k-1)h}^h(m^h) \pi_{(k-1)h}^h(f) - \pi_{(k-1)h}^h(m^h f)] (1 - \Delta N_{kh}^h) \\ &+ \sum_{j=0}^H \sum_{k=1}^n \pi_{(k-1)h}^h(m_j^h)^+ \\ &\quad \times [\pi_{(k-1)h}^h(m_j^h f) - \pi_{(k-1)h}^h(m_j^h) \pi_{(k-1)h}^h(f) + \pi_{(k-1)h}^h(R_j^h f)] \Delta U_{kh}^{jh}, \tag{5.2} \end{aligned}$$

where

$$\begin{aligned} \Delta N_{kh}^h &= N_{kh}^h - N_{(k-1)h}^h, \\ \Delta U_{kh}^{jh} &= U_{kh}^{jh} - U_{(k-1)h}^{jh}, \\ G^h f(z) &= \sum_{z'} [f(z') - f(z)] \mathbf{1}_{\{\Phi(z')=\Phi(z)\}} \mu^h(z, z'), \\ R^h f(z) &= \sum_{j=1}^H R_j^h f(z) = \sum_{j=1}^H \sum_{z'} [f(z') - f(z)] \mathbf{1}_{\{\Phi(z')\neq\Phi(z)\}} \mathbf{1}_{\{\Phi(z')=j\}} \mu^h(z, z'), \\ m_j^h(z) &= \sum_{z'} \mathbf{1}_{\{\Phi(z')\neq\Phi(z)\}} \mathbf{1}_{\{\Phi(z')=j\}} \mu^h(z, z'), \\ m^h(z) &= \sum_{j=1}^H m_j^h(z) = \sum_{z'} \mathbf{1}_{\{\Phi(z')\neq\Phi(z)\}} \mu^h(z, z'). \end{aligned}$$

Equation (5.2) has a unique solution as a consequence of its recursive structure, taking into account (5.1) and the inequality

$$m_j^h(z) \leq m^h(z) \leq 1 - e^{-h\bar{l}} < 1, \quad j = 1, \dots, H. \tag{5.3}$$

Such a solution can be explicitly written down but, instead of doing so, we are going to provide a linearized version of (5.2) as a useful tool to prove the convergence of the discrete-time model to the continuous-time one.

Proposition 5.2. *The equation*

$$\begin{aligned} \rho_{nh}^h(f) &= v_0(f) + \sum_{k=1}^n \{ \rho_{(k-1)h}^h(G^h f) - \rho_{(k-1)h}^h(m^h f) \} (1 - \Delta N_{kh}^h) \\ &\quad + \sum_{j=1}^H \sum_{k=1}^n (1 - e^{-h})^+ \{ \rho_{(k-1)h}^h(m_j^h f) - (1 - e^{-h}) \rho_{(k-1)h}^h(f) + \rho_{(k-1)h}^h(R_j^h f) \} \Delta U_{kh}^{jh} \end{aligned} \tag{5.4}$$

admits a unique $\mathcal{F}_t^{U^h}$ -adapted solution. Such a solution $\rho_t^h(f)$, for any $t = nh$, has the following properties:

- (i) $\rho_t^h(f)$ is a finite positive measure;
- (ii) $0 < \rho_t^h(1) \leq (2\bar{l})^{N_t^h} \vee 1$ for all h such that $h(\bar{l} \vee 1) < \log 2$; and
- (iii) $\pi_t^h(f) = \rho_t^h(f) / \rho_t^h(1)$.

Proof. Here, we use the same argument as in Proposition 4.2. Then we only have to prove that $\rho_t^h(1)$ cannot vanish for $t \in [T_i^h, T_{i+1}^h)$. For $t = nh$, with $T_i^h/h \leq n < T_{i+1}^h/h$, by recalling (5.3) we obtain

$$\rho_{nh}^h(1) = \rho_{(n-1)h}^h(1) - \rho_{(n-1)h}^h(m^h) = \rho_{(n-1)h}^h(1 - m^h) \geq e^{-h\bar{l}} \rho_{(n-1)h}^h(1),$$

which, by induction, implies that $\rho^h(1) \geq e^{-T\bar{l}}$. At a jump time $T_i^h = kh$, we have $\Delta N_{kh}^h = 1$ and, since $m_j^h(z) = m_j(z)(1 - e^{-hl(z)})$, we have

$$\rho_{kh}^h(1) = (1 - e^{-h})^+ \rho_{(k-1)h}^h(m_j^h)|_{j=Y_{nh}^h} \geq \underline{l}p \frac{1 - e^{-hl}}{1 - e^{-h}} \rho_{(k-1)h}^h(1) \geq (\frac{1}{2}\underline{l}^2 p)^{N_{kh}^h}.$$

Moreover, for $\Delta N_{kh}^h = 0$, we have $\rho_{kh}^h(1) \leq \rho_{(k-1)h}^h(1)$ and, for $\Delta N_{kh}^h = 1$, we have

$$\rho_{kh}^h(1) \leq \frac{\rho_{(k-1)h}^h(m^h)}{1 - e^{-h}} \leq \frac{1 - e^{-\bar{l}h}}{1 - e^{-h}} \rho_{(k-1)h}^h(1).$$

6. Convergence

From now on, let $S^h := (\zeta^h, Y^h, U^h, N^h, \pi^h)$ denote the piecewise-constant, càdlàg, continuous-time interpolation of the processes introduced in the previous section. Moreover, we set $S := (\zeta, Y, U, N, \pi)$ and

$$\mathfrak{S} = \mathcal{H} \times \{0, 1, \dots, H\} \times \{0, 1\}^H \times \mathbb{N} \times \Pi(\mathcal{H}).$$

In this section, we prove the following theorem.

Theorem 6.1. *As $h \rightarrow 0$, the process S^h converges to the process S almost surely, with respect to the Skorokhod topology on the space $D_{\mathfrak{S}}[0, T]$.*

Throughout this section, we provide results that are true for almost all fixed ω . More precisely, let us observe that, in the model presented in this paper, τ_{N_T} is a continuous random variable and $P(\tau_{N_T} = T) = 0$ (recalling the definition of N_t given in (4.1)). We will prove that the convergence of S^h to S claimed in Theorem 6.1 holds for any $\omega \in \{\tau_{N_T} < T\}$. The proof of this theorem is a consequence of the next results. The first one follows from an argument similar to that in Section 4 of [3], but we give it for sake of completeness.

Proposition 6.1. *For $h < (T - \tau_{N_T})/N_T$, we have $\tau_{N_T}^h \leq T$, which in turn implies that $N_T = N_T^h$ almost surely.*

Proof. For any $\omega \in \{\tau_{N_T} < T\}$, we can choose h such that $0 < h \leq (T - \tau_{N_T})/N_T$. By definition, $\tau_k^h \leq \tau_k + kh$ for all k ; hence,

$$\tau_{N_T}^h \leq \tau_{N_T} + N_T h \leq T.$$

Moreover, since $\tau_k \leq \tau_k^h$ for all k , we have $N_T \geq N_T^h$, which means that

$$\tau_{N_T}^h \geq \tau_{N_T^h}^h.$$

However, observing that $\tau_{N_T^h}^h$ is the last jump time before T , we also have

$$\tau_{N_T^h}^h \leq T \leq \tau_{N_T}^h.$$

The combination of this last inequality with $T \geq \tau_{N_T}^h$ gives $\tau_{N_T^h}^h = \tau_{N_T}^h$, i.e. $N_T = N_T^h$ (recalling that $\{\tau_k^h\}_{k \geq 0}$ is a strictly monotone sequence).

In the event $\{N_T = N_T^h\}$, a function $\alpha_h(\cdot)$ of $[0, T]$ into itself can be defined such that

- (i) $\alpha_h(\cdot)$ is a piecewise-linear map and transforms the interval $[\tau_k^h, \tau_{k+1}^h)$ into $[\tau_k, \tau_{k+1})$ for all $k < N_T$, and transforms $[\tau_{N_T}^h, T)$ into $[\tau_{N_T}, T)$;

- (ii) $\sup_{t \in [0, T]} |\alpha_h(t) - t| = \max_{k \leq N_T} |\tau_k - \tau_k^h| \leq \max_{k \leq N_T} kh = N_T h$; and
- (iii) $|\zeta(\alpha_h(t)) - \zeta^h(t)| = |Y(\alpha_h(t)) - Y^h(t)| = |U(\alpha_h(t)) - U^h(t)| = |N_{\alpha_h(t)} - N_t^h| = 0$.

Thus, the convergence of the filters suffices to establish the result claimed in Theorem 6.1. Therefore, by definition of the Skorokhod topology, Theorem 6.1 is a consequence of Theorem 6.2.

Theorem 6.2. *Under the assumptions prevailing in this paper (in particular assumption (4.4)), for a suitable quantity C we have $\|\pi_{\alpha_h(t)} - \pi_t^h\| \leq Ch$.*

Recalling part (ii) of Proposition 4.2, the proof of Theorem 6.2 follows from Lemma 6.1 and Proposition 6.2.

Lemma 6.1. *The following inequality holds:*

$$\|\pi_{\alpha_h(t)} - \pi_t^h\| \leq \frac{2e^{T\bar{l}}}{1 \vee \underline{l}p} \|\rho_{\alpha_h(t)} - \rho_t^h\|.$$

Proof. The result follows by noting that

$$\begin{aligned} \|\pi_{\alpha_h(t)} - \pi_t^h\| &= \frac{1}{\rho_{\alpha_h(t)}(1)} \|\rho_{\alpha_h(t)} - \rho_t^h\| + \left\| \frac{\rho_t^h}{\rho_{\alpha_h(t)}(1)\rho_t^h(1)} (\rho_{\alpha_h(t)}(1) - \rho_t^h(1)) \right\| \\ &\leq \frac{2}{\rho_{\alpha_h(t)}(1)} \|\rho_{\alpha_h(t)} - \rho_t^h\| \end{aligned}$$

and recalling Proposition 4.2(ii).

Proposition 6.2. *For a suitable quantity $C \equiv C(T, N_T, \bar{l}) > 0$, we obtain*

$$\|\rho_{\alpha_h(t)} - \rho_t^h\| \leq (1 + 2\bar{l})^{N_t} e^{2\bar{l}} Ch.$$

Proof. By (4.5), since the first derivative $\alpha'_h(t)$ of $\alpha_h(t)$ exists for any $t \in [0, T]$ other than a finite number of points, and since $\alpha_h(T_{(i)}) = T_{(i)}$ for all i , we write

$$\begin{aligned} \rho_{\alpha_h(t)}(f) &= \nu_0(f) + \int_0^t \{\rho_{\alpha_h(s)}(Gf) - \rho_{\alpha_h(s)}(mf)\} \alpha'_h(s) ds \\ &\quad + \sum_{j=0}^H \int_0^t \{\rho_{\alpha_h(s)-}(m_j f) - \rho_{\alpha_h(s)-}(f) + \rho_{\alpha_h(s)-}(R_j f)\} dU_s^{jh}. \end{aligned}$$

Moreover, by (5.4),

$$\begin{aligned} \rho_t^h(f) &= \nu_0(f) + \int_0^t \frac{1}{h} \{\rho_s^h(G^h f) - \rho_s^h(m^h f)\} ds + \int_0^t \{\rho_{s-}^h(m^h f) - \rho_{s-}^h(G^h f)\} dN_s^h \\ &\quad + \sum_{j=0}^H \int_0^t (1 - e^{-h})^+ \{\rho_{s-}^h(m_j^h f) - (1 - e^{-h})\rho_{s-}^h(f) + \rho_{s-}^h(R_j^h f)\} dU_s^{jh} \end{aligned}$$

and, by standard computations, we successively obtain

$$\begin{aligned} & \left| \int_0^t (\alpha'_h(s) - 1) ds \right| \leq 2N_T h, \\ & |G^h f(x) - m^h(x) f(x)| \leq 2\|f\| \bar{l} h, \\ & \left| \int_0^t \{\rho_{s-}^h(m^h f) - \rho_{s-}^h(G^h f)\} dN_s^{Y^h} \right| \leq 2N_T \|f\| \bar{l} h, \\ & \left| \int_0^t \{\rho_{\alpha_h(s)}(Gf) - \rho_{\alpha_h(s)}(mf)\} \alpha'_h(s) ds - \int_0^t \frac{1}{h} \{\rho_s^h(G^h f) - \rho_s^h(m^h f)\} ds \right| \\ & \leq 4\|f\| \bar{l} N_T h \sup_{s \in [0,t]} \rho_s(1) + \|f\| \bar{l}^2 T h \sup_{s \in [0,t]} \rho_s(1) + 2\|f\| \bar{l} \int_0^t \|\rho_{\alpha_h(s)} - \rho_s^h\| ds. \end{aligned}$$

Moreover,

$$\begin{aligned} & \left| \sum_{j=0}^H \int_0^t \{\rho_{\alpha_h(s)-}(m_j f) - \rho_{\alpha_h(s)-}(f) + \rho_{\alpha_h(s)-}(R_j f)\} dU_s^{jh} \right. \\ & \quad \left. - \sum_{j=0}^H \int_0^t (1 - e^{-h})^+ \{\rho_{s-}^h(m_j^h f) - (1 - e^{-h}) \rho_{s-}^h(f) + \rho_{s-}^h(R_j^h f)\} dU_s^{jh} \right| \\ & \leq 2\bar{l} \|f\| \int_0^t \|\rho_{\alpha_h(s)-} - \rho_{s-}^h\| dN^h(s) + \bar{l} \|f\| N_T^h \bar{l} h \sup_{s \in [0,T]} \rho_s^h(1). \end{aligned}$$

Finally,

$$\begin{aligned} \|\rho_{\alpha_h(t)} - \rho_t^h\| & \leq 2h \bar{l} N_T + 4\bar{l} N_T h \sup_{s \in [0,t]} \rho_s(1) + \bar{l}^2 T h \sup_{s \in [0,t]} \rho_s(1) + 2\bar{l} \int_0^t \|\rho_{\alpha_h(s)} - \rho_s^h\| ds \\ & \quad + 2\bar{l} \int_0^t \|\rho_{\alpha_h(s)-} - \rho_{s-}^h\| dN_s^h + \bar{l}^2 N_T^h h \sup_{s \in [0,T]} \rho_s^h(1). \end{aligned}$$

Then, a suitable quantity $A \equiv A(T, N_T, \bar{l})$ can be found such that

$$\|\rho_{\alpha_h(t)} - \rho_t^h\| \leq Ah + 2\bar{l} \int_0^t \|\rho_{\alpha_h(s)} - \rho_s^h\| ds + 2\bar{l} \int_0^t \|\rho_{\alpha_h(s)-} - \rho_{s-}^h\| dN_s.$$

By using Gronwall’s lemma for $t \in [T_i, T_{(i+1)})$ for all i , and taking into account the update at the jump times, the proof is complete.

Acknowledgements

The authors are very grateful to Professor Fabio Spizzichino for suggesting the model and, along with Dr Barbara Torti, for many helpful discussions. We also thank the unknown referee, whose comments allowed us to improve the paper a great deal.

References

- [1] ARJAS, E., HAARA, P. AND NORROS, I. (1992). Filtering the histories of a partially observed marked point process. *Stoch. Process. Appl.* **40**, 225–250.
- [2] BRÉMAUD, P. (1980). *Point Processes and Queues*. Springer, New York.
- [3] CALZOLARI, A. AND NAPPO, G. (1996). A filtering problem with counting observations: approximations with error bounds. *Stoch. Stoch. Reports* **57**, 71–87.
- [4] CALZOLARI, A. AND NAPPO, G. (1997). A filtering problem with counting observations: error bounds due to the uncertainty on the infinitesimal parameters. *Stoch. Stoch. Reports* **61**, 1–19.
- [5] CALZOLARI, A. AND NAPPO, G. (2001). The filtering problem in a model with grouped data and counting observation times. Preprint. Available at <http://www.mat.uniroma1.it/people/nappo/nappo.html>.
- [6] CECI, C. AND GERARDI, A. (2001). Nonlinear filtering equation of a jump process with counting observation. *Acta Appl. Math.* **66**, 139–154.
- [7] ETHIER, S. N. AND KURTZ, T. G. (1986). *Markov Processes: Characterization and Convergence*. John Wiley, New York.
- [8] GERARDI, A., SPIZZICHINO, F. AND TORTI, B. (2000). Exchangeable mixture models for lifetimes: the role of occupation number. *Statist. Prob. Lett.* **49**, 365–375.
- [9] GERARDI, A., SPIZZICHINO, F. AND TORTI, B. (2000). Filtering equations for the conditional law of residual lifetimes from a heterogeneous population. *J. Appl. Prob.* **37**, 823–834.
- [10] 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.
- [11] KURTZ, T. G. (1998). Martingale problems for conditional distribution of Markov processes. *Electron. J. Probab.* **3**, 29pp.
- [12] KURTZ, T. G. AND OCONE, D. (1988). Unique characterization of conditional distributions in nonlinear filtering. *Ann. Probab.* **16**, 80–107.