Scientiae Mathematicae Japonicae Online, **e-2006**, 915–932

CONDITIONAL LAW OF LIFETIMES FOR A HETEROGENEOUS POPULATION OF LIVING PARTICLES VIA FILTERING TECHNIQUES

Paola Tardelli

Received January 25, 2006

Abstract. A model is given for the evolution of a heterogeneous population of identical living particles, divided into different classes. The model is dynamic, since the partition of the population changes as the time goes on. The aim of this paper is to determine the law of the lifetime of each particle given the number of dead particles up to time t. This program is achieved by introducing the occupancy numbers, which are defined as the cardinality of each class. Assuming that the partition is non-observable, a filtering problem arises and the observation coincides with the cardinality of the class of dead particles. A discussion is performed about discrete time approximations of the filter.

1. Introduction.

For H positive integer, let H identical living particles be simultaneously in the same environment and subjected to the same source of stress. At every time $t \in \mathbb{R}^+$, the type $Z_i(t)$ of any given particle, for $1 \leq i \leq H$, is identified by an integer number between 0 and d, $d \in \mathbb{N}$. This means that the population is divided in $d+1$ classes depending on the different types.

The partition of the population is supposed to be dynamic, since the type of each particle, which can be seen as the level of "health" of any single particle, can change, while the time goes on, depending on some kind of treatment used. Let the partition of the population be non-observable. The observations is just the number of particles dead up to time t, namely the number of particles having type equal to 0. The type 0 is absorbing, in the sense that if a particle has level 0, it stays that way indefinitely.

This model is a generalization of that presented in [7] and [8]. The main difference relies on the fact that there, the partition of the population does not change during the time and the particles could only die. In [8], a family of exchangeable random variables $Z_1, \dots, Z_H$ were introduced in order to define the partition, setting $Z_i = k$ if and only if the particle labeled by i belongs to the class characterized by the level k. The law of the lifetimes given the variables $Z_1, \dots, Z_H$ was given as a data and the lifetimes were assumed to be independent, given the partition of the population. Then, an exchangeability property for the lifetimes was proven, although they were supposed to be conditionally independent but not conditionally identically distributed.

In [9], in order to preserve the property of the particles to be identical and to let the partition to be dynamic, a suitable exchangeable assumption is made. This is a generalization of the exchangeability property, for fixed t, assumed in [8]. Therefore, in [9], the family of the lifetimes was proven to be an exchangeable sequence of random variables, when the dynamics depend on the types of the particles belonging to the whole population.

2000 *Mathematics Subject Classification.* 60G09, 62N05, 60H30, 60G35, 60J75.

Key words and phrases. Heterogeneous populations, exchangeability, stochastic filtering, jump processes.

In this paper, the occupancy numbers, defined as the cardinalities of each class and already introduced in [7] and [8], are considered. Assuming that the dynamics depend only on the crowding of the classes, namely depends on the cardinality of the classes, a one-to-one relation is stated between the law of the process Z and the law of the occupancy numbers process. Another important consequence of the previous assumption, takes us to prove that the occupancy numbers are a Markov process if and only if Z is a Markov process. The aim is to provide the conditional law of the lifetimes given the observations. To this end, a first step consists in finding the conditional law of the lifetimes given the occupancy numbers process, then, a second step bring us to deal with a filtering problem to get the conditional law of the occupancy numbers given the observations.

In order to solve a filtering problem, we recall that the filter satisfies a stochastic differential equation known as the Kushner-Stratonovich equation. Therefore, to deduce the properties of the filter from this equation, we need to find some kind of uniqueness. Weak uniqueness could be obtained, as in [12] and [13], by using the filtering martingale problem approach. But, taking into account the peculiarity of this model, a stronger kind of uniqueness can be reached, namely, path-wise uniqueness, [1].

Moreover, a hand-able representation for the filter is given by a linearized equation and a Feynmann Kac's formula. Then, a discrete time approximation is presented and the convergence in the Skorohod topology is proven, following a method which is a modification of that suggested in [11] and [4].

As a final remark, let us observe that this model applies to a population of living particles as well as to study the failure times of a set of items under maintenance.

2. General setting.

For H positive integer, let $\mathcal{P}_H$ be a finite sub-population of $\mathcal{P}$, where $\mathcal{P} = \{U_j\}_{j\geq 1}$ is a finite or countable population and U_j are given identical particles. Each elements U_j can be of d different types, labeled by the natural numbers $1, \ldots, d$, consequently, $\mathcal{P}_H$ is an heterogeneous population. The subset of all particles of type k, $k = 1, \ldots, d$, at time $t \in \mathbb{R}^+$, is denoted by $C_k(t)$. Since particles of any type can die, let $C_0(t)$ be the class of the particles dead up to time t, thus, $\mathcal{P}_H = \bigcup_{k=0,1,\ldots,d} C_k(t)$.

Let $Z(t) = \{Z_i(t)\}_{1\leq i\leq H}$, for $t \in \mathbb{R}^+$, be defined assuming that $Z_i(t) = k$ if and only if $U_i \in C_k(t)$, for $k = 0, 1, ..., d$ and for $i = 1, \ldots, H$. Let Z be a stochastic process, then, for $t \in \mathbb{R}^+$, $Z_i(t)$ is a random variable taking value in $\{0, 1, \ldots, d\}$ and $Z(t)$ takes value in $\mathcal{H} = \{0, 1, \ldots, d\}^H$. In [7] and [8], the partition of the population does not change with the time. Instead of this, in order to have a dynamic model, as in [9], two conditions appear very natural in this context.

Assumption 2.1 *For $i = 1, \ldots H$ and for all time $t \geq s$,* $\quad Z_i(s) = 0 \implies Z_i(t) = 0 \quad$ *a.s..*

Assumption 2.2 *Choosing $t_i \in \mathbb{R}^+$, $i = 1, \ldots, n$, with $t_1 \leq \ldots \leq t_n$, for all β permutation on $\{1, \ldots, H\}$ and for all $k^{(1)}, \ldots, k^{(n)} \in \mathcal{H}$,*

$$P\left(Z(t_1) = k^{(1)}, \ldots, Z(t_n) = k^{(n)}\right) = P\left(Z(t_1) = \beta k^{(1)}, \ldots, Z(t_n) = \beta k^{(n)}\right)$$

where $k^{(i)} = \{k_1^{(i)}, \ldots, k_H^{(i)}\}$, $\beta k^{(i)} = \{k_{\beta_1}^{(i)}, \ldots, k_{\beta_H}^{(i)}\}$.

Assumption 2.1 describes the particular property of the class $C_0(t)$, particles can enter in $C_0(t)$ but cannot go out of it. Assumption 2.2 is a generalization of the exchangeability property given in [7] and [8] and it is, in some sense, an exchangeability property of the trajectories. In particular, it implies that $Z(t) = (Z_1(t), \ldots Z_H(t))$ is an exchangeable sequence, for any fixed $t \geq 0$.

Definition 2.3 *Let* $T_i := \inf\{t \in I\!R^+ : Z_i(t) = 0\}$ *be the lifetime of* U_i*, for* $i = 1, \dots H$.

We are going to prove, as in [7] and [8], that the lifetimes are a sequence of exchangeable random variables. In our dynamic model the assumption of exchangeability of $Z(t)$, for a fixed t, is not enough. But a stronger kind of property is necessary and Assumption 2.2 is the required tool together with Assumption 2.1. The proof of Theorem 2.4 below can be found in [9], we just recall a sketch of it, for sake of completeness.

Theorem 2.4 *Under Assumption 2.1 and Assumption 2.2, the sequence* $\{T_i\}_{i\geq 1}$ *is a family of exchangeable random variables.*

Proof. In order to compute $P(T_1 \leq t_1, \dots, T_H \leq t_H)$, $\forall t_1, \dots, t_H \in I\!R^+$, without loss of generality, let $t_1 \leq \dots \leq t_H$. If this is not the case, we can consider a permutation of index such that: $t_{1'} \leq \dots \leq t_{H'}$ and let us 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 Assumption 2.1, $\forall n = 1, \dots, H$,

$$
\begin{aligned}
P(T_1 \leq t_1, \dots, T_n \leq t_n) &= P\big(Z_1(t_1) = 0, Z_2(t_2) = 0, \dots, Z_n(t_n) = 0\big) = \\
&= P\big(Z_1(t_1) = 0, Z_1(t_2) = 0, \dots, Z_1(t_n) = 0 \\
&\qquad\qquad Z_2(t_2) = 0, \dots, Z_2(t_n) = 0 \\
&\qquad\qquad \dots \\
&\qquad\qquad \dots, Z_n(t_n) = 0\big) =
\end{aligned}
$$

$$
\begin{aligned}
&= \sum_{x_{i,j}, 2\leq i\leq H, 1\leq j\leq i-1} P\big(Z(t_1) = (0, x_{2,1}, \dots, x_{n,1}), Z(t_2) = (0, 0, x_{3,2}, \dots, x_{n,2}), \dots\big) = \\
&= \sum_{x_{i,j}, 2\leq i\leq H, 1\leq j\leq i-1} P\big(Z(t_1) = \pi(0, x_{2,1}, \dots, x_{n,1}), Z(t_2) = \pi(0, 0, x_{3,2}, \dots, x_{n,2}), \dots\big) = \\
&= P\big(Z_{\pi^{-1}(1)}(t_1) = 0, Z_{\pi^{-1}(2)}(t_2) = 0, \dots, Z_{\pi^{-1}(n)}(t_n) = 0\big) = \\
&= P(T_{\pi^{-1}(1)} \leq t_1, \dots, T_{\pi^{-1}(n)} \leq t_n) \qquad \square
\end{aligned}
$$

The existence of a process Z, verifying Assumption 2.1 and Assumption 2.2, is obtained in discrete time, $t \in I\!N$, by the Markov property. In fact, let $\mu(a, b)$, for $a, b \in \mathcal{H}$, be a transition probability family and let ν_0 be a probability measure on $\mathcal{H}$.

Assumption 2.5 *Let us assume that both the following conditions hold*

i) *If there exists an index* i *such that* $a_i = 0$ *and* $b_i \neq 0$ *then* $\mu(a, b) = 0$.

ii) *For* $a, b \in \mathcal{H}$ *and for each* β*, permutation of index, then* $\mu(a, b) = \mu(\beta a, \beta b)$.

Let $Z(t)$, $t \in I\!N$, be the Markov chain with initial law ν_0 and with transition probabilities defined as $P(Z(t) = b|Z(t-1) = a) := \mu(a, b)$, for $t > 0$ where $a = (a_1, \dots a_H)$ and $b = (b_1, \dots b_H) \in \mathcal{H}$. Then, the existence of a process Z, in discrete time, follows from Proposition 2.6 below, which can be proven by direct computations.

Proposition 2.6 *If Assumption 2.5 holds and if* $Z(0) = Z_1(0), \dots, Z_H(0)$ *is an exchangeable sequence of random variables, then the process* Z *satisfies Assumption 2.1 and Assumption 2.2.*

Proof. The thesis is achieved recalling that any subsequence of an exchangeable sequence of random variables still is an exchangeable sequence of random variables and noting that, for all β permutation on $\{1,\dots,H\}$ and $\forall n \geq 1$,

$$\begin{aligned}
&P(Z(0)=a_0, Z(1)=a_1, Z(2)=a_2,\dots,Z(n)=a_n) = \\
&\quad = \mu(a_0,a_1)\cdot\ldots\cdot\mu(a_{n-1},a_n)\cdot P(Z(0)=a_0) = \\
&\quad = \mu(\beta(a_0),\beta(a_1))\cdot\ldots\cdot\mu(\beta(a_{n-1}),\beta(a_n))\cdot P(Z(0)=\beta(a_0)) = \\
&\quad = P(Z(0)=\beta(a_0), Z(1)=\beta(a_1), Z(2)=\beta(a_2),\dots,Z(n)=\beta(a_n)). \qquad \square
\end{aligned}$$

Remark 2.7 *Let* $\mu^{(m)}(a,b) = P(Z(t+m) = b|Z(t) = a)$, $\forall t, m \in I\!N$. *If the second condition of Assumption 2.5 holds, then* $\mu^{(m)}(a,b) = \mu^{(m)}(\beta(a),\beta(b))$.

In continuous time, $t \in I\!R^+$, Z can be constructed as a continuous time Markov process with generator given by

$$L^z f(z) = l(z) \sum_{z'\in\mathcal{H}} [f(z') - f(z)]\, p(z,z') \tag{2.1}$$

where $l(z)$ is a positive function and $\{p(z,z')\}$ is a family of transition probabilities.

Assumption 2.8 *For any* β*, permutation of index on* $\{1,\dots,H\}$*, let* $l(z) = l(\beta z)$ *and let* $\{p(z,z')\}$ *be a family of transition probabilities verifying Assumption 2.5.*

Since $\mathcal{H}$ is finite, there exist $\underline{l}, \bar{l} \in I\!R^+$ such that $0 < \underline{l} \leq l(z) \leq \bar{l}$, $\forall z \in \mathcal{H}$. By construction, the generator L^z, (2.1), is bounded. Then, [6], for all ν_0, probability measure on $\mathcal{H}$, there exists a unique Markov process Z, with sample paths in $D_{\mathcal{H}}\ [0,+\infty)$ (the space of right continuous $\mathcal{H}$ -valued functions on $[0,\infty)$ having left limits), initial condition ν_0 and generator L^z. Furthermore, such a process verifies Assumption 2.2, whenever $Z(0)$ is an exchangeable sequence of random variables.

In order to clarify the previous claim, let us recall a particular realization of the process Z, suggested in [6]. On a probability space $(\Omega,\mathcal{F},P)$, let $\{Z_n\}_{n\geq 0}$ be a Markov chain defined by ν_0 and $p(z,z')$. Let $\{V_i\}_{i\geq 1}$ be a sequence of random variables independent and exponentially distributed with parameter 1 and independent of $\{Z_n\}_{n\geq 0}$ and let

$$\tau_0 = 0, \qquad \tau_n = \sum_{i=1}^{n} \frac{V_i}{l(Z_{i-1})}, \quad \textit{for} \quad n > 0, \qquad Z(t) = \sum_{n\geq 0} Z_n \mathbf{1}_{\{\tau_n \leq t < \tau_{n+1}\}} \tag{2.2}$$

and $\mathcal{F}_t = \sigma\{Z(s), s \leq t\}$. Hence, on $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\geq 0},P)$, Z is a continuous time pure jump Markov process with generator L^z verifying Assumption 2.1 and $\{\tau_i\}_{i\geq 1}$ is the sequence of its jump times. Noting that Assumption 2.2 just involves the law of the process Z, this particular realization allows us to prove Theorem 2.9 below.

Theorem 2.9 *Under Assumption 2.8 and if* $Z(0) = Z_1(0),\dots,Z_H(0)$ *is an exchangeable sequence of random variables*

i) $\forall t$*, fixed,* $Z(t) = \{Z_i(t)\}_{1\leq i\leq H}$ *is an exchangeable sequence of random variables,*

ii) for $s \leq t$, $P\big(Z(t) = k|Z(s) = h\big) = P(Z(t) = \beta(k)|Z(s) = \beta(h)\big)$ *and*

iii) for $t_0 \leq t_1 \leq \ldots \leq t_n$*, and* $a_0, a_1 \ldots a_n \in \mathcal{H}$,

$$P\big(Z(t_0) = a_0,\dots,Z(t_n) = a_n\big) = P(Z(t_0) = \beta(a_0),\dots,Z(t_n) = \beta(a_n)),$$

that is Assumption 2.2 holds.

Proof. For the first step, setting $k = (k^{(1)}, \ldots, k^{(H)})$, recalling (2.2) and recalling that the random variables $\{V_i\}_{i\geq 1}$ and the Markov chain $\{Z_n\}_{i\geq 0}$ are independent, we get that

$$P(Z(t) = k) = \sum_{n\geq 0} P\big(Z(t) = k, \tau_n \leq t < \tau_{n+1}\big) =$$
$$= P\Big(Z_0 = k, 0 \leq t < \tau_1\Big) + \sum_{n>0} P\Big(Z_n = k,\ \tau_n \leq t < \tau_n + \frac{V_{n+1}}{l(Z_n)}\Big) =$$
$$= P(Z_0 = k) \cdot e^{-l(k)t} + \sum_{n>0} \sum_{h_0,\ldots,h_{n-1}\in\mathcal{H}} P\big(Z_0 = h_0, \ldots, Z_{n-1} = h_{n-1}, Z_n = k\big) \cdot$$
$$\cdot P\Big(\sum_{i=1}^{n} \frac{V_i}{l(h_{i-1})} \leq t < \sum_{i=1}^{n} \frac{V_i}{l(h_{i-1})} + \frac{V_{n+1}}{l(k)},\Big)$$

and the thesis is achieved by Assumption 2.8 and by the exchangeability of the sequence $Z(0)$. For the second step, setting

$$P\big(Z(t) = k, Z(s) = h\big) = \sum_{n\geq 0} P\big(Z(t) = k, Z(s) = h, \tau_n \leq s \leq t < \tau_{n+1}\big) +$$
$$+ \sum_{n\geq 0} P\big(Z(t) = k, Z(s) = h, \tau_n \leq s \leq \tau_{n+1} \leq t < \tau_{n+2}\big) + \tag{2.3}$$
$$+ \sum_{n\geq 0, p>1} P\big(Z(t) = k, Z(s) = h, \tau_n \leq s \leq \tau_{n+1} \leq \tau_{n+p} \leq t < \tau_{n+p+1}\big)$$

and recalling, again, Proposition 2.6, we have that the first term in the right hand side of (2.3) is equal to

$$\sum_{n\geq 0} P(Z(t) = k, Z(s) = h, \tau_n \leq s \leq t < \tau_{n+1}) =$$
$$= \sum_{n\geq 0} P\big(Z_n = k, Z_n = h, \tau_n \leq s \leq t < \tau_{n+1}\big) =$$
$$= \delta_{h,k}\Big\{P\big(Z_0 = k, 0 \leq s \leq t < \frac{V_1}{l(Z_0)}\big) +$$
$$+ \sum_{\substack{n\geq 1,\\ h_0,\ldots,h_{n-1}\in\mathcal{H}}} P\Big(Z_0 = h_0, \ldots, Z_n = k, \sum_{i=1}^{n} \frac{V_i}{l(Z_{i-1})} \leq s \leq t < \sum_{i=1}^{n+1} \frac{V_i}{l(Z_{i-1})}\Big)\Big\} =$$
$$= \delta_{h,k}\Big\{P(Z_0 = k)\ P\left(0 \leq s \leq t < \frac{V_1}{l(k)}\right) + \sum_{\substack{n\geq 1,\\ h_0,\ldots,h_{n-1}\in\mathcal{H}}} P\big(Z_0 = h_0, \ldots, Z_n = k\big) \cdot$$
$$\cdot P\Big(\sum_{i=1}^{n} \frac{V_i}{l(h_{i-1})} \leq s \leq t < \sum_{i=1}^{n} \frac{V_i}{l(h_{i-1})} + \frac{V_{n+1}}{l(k)}\Big)\Big\} =$$
$$= \sum_{n\geq 0} P\big(Z(t) = \beta(k), Z(s) = \beta(h), \tau_n \leq s \leq t < \tau_{n+1}\big).$$

In an analogous way, the second and third term in (2.3) can be computed and the thesis is achieved by the first step. The third step immediately follows by the Markov property and the previous results. □

3. The occupancy numbers

One area where the occupancy number process can be useful is not only in modeling the behavior of biological populations but also in queueing models.

Definition 3.1 *Let* $\Phi = (\Phi_1, \ldots, \Phi_d, \Phi_0)$, $\Phi_i(z) = \sum_{j=1,\ldots,H} \mathbf{1}_{z_j=i}$, *for each* $i = 1, \ldots, d$ *be a deterministic function.*

The pair (X, Y) *is the occupancy number process, where* $X := (X^1, \ldots, X^d)$ *and, for* $t \in I\!R^+$, $X^i(t) = \#C_i(t) = \Phi_i(Z(t))$ *and where* $Y(t) = \#C_0(t) = \Phi_0(Z(t))$.

As a consequence of Definition 3.1 above, the process (X, Y) takes values in

$$\mathcal{K} := \{(x_1, \ldots, x_d, y) : x_i, y \in I\!N \cup \{0\}, \forall i;\ x_1 + \cdots + x_d + y = H\}$$

while the components X and Y are such that X takes values in

$$\mathcal{X} := \{(x_1, \ldots, x_d) :\ x_i \in I\!N \cup \{0\},\ \forall i,\ x_1 + \ldots + x_d \leq H\}$$

and Y is, almost surely, non-decreasing with respect to t, Assumption 2.1.

The aim of this paper is to find the conditional law of the lifetimes given the past history of the process Y. To this end, in [9], the authors follow a two step procedure. First, they evaluated the law of the lifetimes given $Z(t)$, and, then, they studied a filtering problem to obtain the distribution of $Z(t)$, given the observations. In our model, again a two step procedure will be followed, but, previously, some kind of relation between the law of Z, $\mathcal{L}(Z)$, and the law of the occupancy numbers X and Y, $\mathcal{L}(X, Y)$ have to be established in order to deduce the joint law of (X, Y) and its Markov property.

For any fixed time t, as already observed in [7], since $(X(t), Y(t)) = \Phi(Z(t))$ and Φ is a deterministic function not necessarily one-to-one, there exists a one-to-one correspondence between the law of $(X(t), Y(t))$, $\mathcal{L}(X(t), Y(t))$, and the law of $Z(t)$, $\mathcal{L}(Z(t))$.

Proposition 3.2 *For* $k \in \{0, 1, \ldots, d\}^H$ *and* $\forall t \in I\!R^+$, *if* $Z(t)$, *for fixed* t, *is an exchangeable sequence of random variables, then*

$$P(Z(t) = k) = \frac{\Phi_0(k)! \cdot \Phi_1(k)! \cdot \ldots \cdot \Phi_d(k)!}{H!} \cdot P((X(t), Y(t)) = \Phi(k)) \tag{3.1}$$

and for each $s \in \mathcal{K}$

$$P\Big((X(t), Y(t)) = s\Big) = P\left(Z(t) \in \Phi^{-1}(s)\right). \tag{3.2}$$

As a first remark, observe that since $\Phi(k) = \Phi(\beta k)$, for all β permutation on $\{1, \ldots, H\}$ and for all $k \in \{0, 1, \ldots, d\}^H$, this last result agrees with Assumption 2.2. Furthermore, immediately, the result given in (3.2) can be easily generalized and we get that for each $n > 0$ and for each $s^{(1)}, \ldots, s^{(n)} \in \mathcal{K}$, then

$$P\left((X(t_1), Y(t_1)) = s^{(1)}, \ldots, (X(t_n), Y(t_n)) = s^{(n)}\right) = \\ = P\left(Z(t_1) \in \Phi^{-1}(s^{(1)}), \ldots, Z(t_n) \in \Phi^{-1}(s^{(n)})\right).$$

On the other hand, the result given in (3.1) cannot be generalized in a dynamic context, recalling that the function Φ is not necessarily one-to-one. To this end, a further assumption is needed and, in particular, we choose to assume that the dynamics of the process $Z(t)$ just depend on the number of particles belonging to each class.

Assumption 3.3 *Let* $\gamma(h,\widetilde{h}) := \prod_{i=1}^{H}(\mathbf{1}_{h_i\neq 0} + \mathbf{1}_{h_i=0,\widetilde{h}_i=0})$, *for* $h,\widetilde{h} \in \{0,1,\ldots,d\}^H$. *Recalling that* $\Phi = (\Phi_1,\ldots,\Phi_d,\Phi_0)$ *and that* $\Phi_i(z) = \sum_{j=1,\ldots,H} \mathbf{1}_{z_j=i}$, $i = 0,1,\ldots,d$, *for each* $n > 0$, *for all* $t_1 \leq \ldots \leq t_n$, $\forall(h^{(1)},\ldots,h^{(n)})$, $\forall(h'^{(1)},\ldots,h'^{(n)})$, $h^{(i)},h'^{(i)} \in \{0,1,\ldots,d\}^H$ *such that*

$$\begin{array}{lll} i. & \Phi(h^{(i)}) = \Phi(h'^{(i)}) & i = 1,\ldots,n \\ ii. & \gamma(h^{(i)},h^{(i+1)}) = \gamma(h'^{(i)},h'^{(i+1)}) & i = 1,\ldots,n-1 \end{array}$$

we assume that

$$P(Z(t_1) = h^{(1)},\ldots,Z(t_n) = h^{(n)}) = P(Z(t_1) = h'^{(1)},\ldots,Z(t_n) = h'^{(n)}). \tag{3.3}$$

This technical assumption becomes more "friendly" just observing that if the vectors $h^{(i)}$ and $h'^{(i)}$, for $i = 1,\ldots,n$, produce the same vector of occupancy numbers, then the finite dimensional distributions of Z coincide. Moreover, let us observe that Assumption 3.3 implies both Assumption 2.1 and Assumption 2.2 while vice-versa is not true.

Proposition 3.4 *Under Assumption 3.3, we get that, for* $n \geq 1$, *for* $t_1 \leq \ldots \leq t_n$ *and for* $h^{(1)},\ldots,h^{(n)} \in \{0,1,\ldots,d\}^H$,

$$\begin{aligned} &P(Z(t_1) = h^{(1)},\ldots,Z(t_n) = h^{(n)}) = \\ &\quad = A(h^{(1)},\ldots,h^{(n)}) \cdot P\Big((X(t_1),Y(t_1)) = \Phi(h^{(1)}),\ldots,(X(t_n),Y(t_n)) = \Phi(h^{(n)})\Big) \end{aligned}$$

where $A(h^{(1)},\ldots,h^{(n)})$ *are deterministic quantities given by*

$$\begin{aligned} A(h^{(1)},\ldots,h^{(n)}) &= \frac{\Phi_0(h^{(1)})!\Phi_1(h^{(1)})!\ldots\Phi_d(h^{(1)})!}{H!}\cdot \\ &\cdot\prod_{j=1}^{n-1}\frac{(\Phi_0(h^{(j+1)})-\Phi_0(h^{(j)}))!\Phi_1(h^{(j+1)})!\ldots\Phi_d(h^{(j+1)})!}{(H-\Phi_0(h^{(j)}))!}\cdot\gamma(h^{(j)},h^{(j+1)}). \end{aligned}$$

Proposition 3.4 above, readily obtained by combinatorial techniques, shows us that, under Assumption 3.3, the dynamics of the process Z just depend on the number of particles belonging to each class. But this assumption has another important consequence about the Markov property. In general, if Z is a Markov process this does not implies that such is the occupancy numbers process (X,Y), but, in our case, Proposition 3.5 below, let us to deduce a necessary and sufficient condition.

Proposition 3.5 *Under Assumption 3.3,* (X,Y) *is a Markov process if and only if* Z *is a Markov process.*

Proof. If Z is a Markov process, setting $\mathcal{F}_t^Z = \sigma\{Z(s),\ s \leq t\}$, then, for $h^{(1)},\ldots,h^{(n)},h \in \{0,1,\ldots,d\}^N$, for $t_1 \leq \ldots \leq t_n \leq s \leq t$,

$$\begin{aligned} &P\Big((X(t),Y(t)) = m\big|Z(t_1) = h^{(1)},\ldots,Z(t_n) = h^{(n)},Z(s) = h\Big) = \\ &\quad = \sum_{k:\gamma(h,k)=1,\Phi(k)=m} P\Big(Z(t) = k\big|Z(s) = h\Big) = \\ &\quad = P\Big((X(t),Y(t)) = m\big|(X(s),Y(s)) = \Phi(h)\Big). \end{aligned}$$

Vice-versa, since $\mathcal{F}_s^{(X,Y)} \subset \mathcal{F}_s^Z$, by (3.3) and Proposition 3.4, $\forall h^{(1)}, \dots, h^{(n)} \in \{0, 1, \dots, d\}^N$,

$$P\Big(Z(t) = k | Z(t_1) = h^{(1)}, \dots, Z(t_n) = h^{(n)}, Z(s) = h\Big) =$$
$$= A(h^{(1)}, \dots, h^{(n)}) \cdot P\Big((X(t), Y(t)) = \Phi(k) | (X(s), Y(s)) = \Phi(h)\Big) =$$
$$= P(Z(t) = k | Z(s) = h). \qquad \square$$

The Markov property is the key for the existence of a process Z satisfying Assumption 3.3, since, in such a case, (X, Y) can be constructed as the solution of a suitable Martingale problem associated with the generator

$$L^{x,y} f(x, y) = \Lambda(x, y) \sum_{(x',y') \in \mathcal{K}} [f(x', y') - f(x, y)] M((x, y); (x', y')), \tag{3.4}$$

where Λ and M are function such that

$$\Lambda(\Phi(z)) = l(z) \qquad \text{and} \qquad M(\Phi(z), \Phi(z')) = p(z, z') \tag{3.5}$$

and their existence is assured by Assumption 3.3. The Martingale problem associated with the generator $L^{x,y}$ is well posed since by construction, the generator given in (3.4) is a bounded operator and its solution (X, Y) is a Markov process. Therefore, (3.5) provides the law of the process Z which, by construction, verifies Assumption 3.3.

4. The filtering problem

Recalling that the partition of the population is assumed to be non-observable, the observation is the cardinality $Y(t)$ of the class $C_0(t)$. Then, setting $\mathcal{F}_t^Y = \sigma\{Y(s), s \le t\}$, our aim is to find $\mathcal{L}(T_1, \dots, T_H | \mathcal{F}_t^Y)$, the conditional law of the lifetimes given the history of the process $Y(t)$.

To achieved this program, the first step consists in computing the conditional law of the lifetimes given the process $(X(t), Y(t))$, $\mathcal{L}(T_1, \dots, T_H | X(t), Y(t))$. This law can be deduced by the conditional law of the lifetimes given the process $Z(t)$, $\mathcal{L}(T_1, \dots, T_H | Z(t))$, and the relation between the law of Z, $\mathcal{L}(Z)$, and the law of $(X(t), Y(t))$, $\mathcal{L}(X, Y)$, is given previously. Then, we need to find $\mathcal{L}(T_1, \dots, T_H | Z(t))$ and, to this end, we recall the results reached in [9]. Assuming that $t_1 \le \dots \le t_H$, if t is a time such that $t_1 \le t_2 \le \dots \le t_m \le t \le t_{m+1} \le \dots \le t_H$, for $P(Z(t) = k) \neq 0$, $k \in \mathcal{H}$,

$$P(T_1 \le t_1, T_2 \le t_2, \dots, T_H \le t_H | Z(t) = k) = \frac{1}{P(Z(t) = k)} \cdot \mathbf{1}_{k_1 = \dots = k_m = 0} \cdot \tag{4.1}$$
$$\cdot P(Z_1(t_1) = 0, \dots, Z_m(t_m) = 0, Z(t) = k, Z_{m+1}(t_{m+1}) = 0, \dots, Z_H(t_H) = 0).$$

Remark 4.1 *Summing up, once $\widetilde{\pi}_t(z) = P(Z(t) = z | \mathcal{F}_t^Y)$ is computed,*

$$P(T_1 \le t_1, T_2 \le t_2, \dots, T_H \le t_H | \mathcal{F}_t^Y) =$$
$$= \sum_{z \in \mathcal{H}} P(T_1 \le t_1, T_2 \le t_2, \dots, T_H \le t_H | Z(t) = z) \cdot \widetilde{\pi}_t(z)$$

and the problem of determining a filter $\widetilde{\pi}_t(z)$ is discussed, exhaustively, in [9].

4.1. Conditional law of lifetimes

Setting, $|x| = x_1 + \dots + x_d$, for each $x \in \mathcal{X}$, note that, since $Y = H - |X|$, $\mathcal{L}(T_1, \dots, T_H | X(t))$ coincides with $\mathcal{L}(T_1, \dots, T_H | X(t), Y(t))$.

Proposition 4.2 *Under Assumption 3.3, assuming that* $t_1 \leq \ldots \leq t_H$*, if* t *is a time such that* $t_1 \leq t_2 \leq \ldots \leq t_m \leq t \leq t_{m+1} \leq \ldots \leq t_H$*, for* $P(X(t) = x) \neq 0$, $x \in \mathcal{X}$,

$$P(T_1 \leq t_1, T_2 \leq t_2, \ldots, T_H \leq t_H | X(t) = x) = \\ = \frac{\sum_{k \in \Phi^{-1}(x, H-|x|)} P(T_1 \leq t_1, T_2 \leq t_2, \ldots, T_H \leq t_H | Z(t) = k)}{\sum_{h \in \Phi^{-1}(x, H-|x|)} P(Z(t) = h)} \tag{4.2}$$

and then $\mathcal{L}(T_1, \ldots, T_H | X(t))$ *can be computed by (4.1).*

Proof. Since $Y(t) = H - |X(t)|$, then the thesis is achieved noting that

$$P\Big(T_1 \leq t_1, \ldots, T_H \leq t_H | X(t) = x\Big) = P\left(T_1 \leq t_1, \ldots, T_H \leq t_H \Big| Z(t) \in \Phi^{-1}(x, H - |x|)\right). \square$$

4.2. Filtering equation

The process Y is not, in general, a counting process, since the component of the process Z are not independent. To overcome this difficulty, as in [5], let us introduce the multivariate point process $U = (U^1, \ldots, U^H)$ be defined as

$$U^j(t) := \sum_{i \geq 1} \mathbf{1}_{\{\tau_i \leq t\}} \, \mathbf{1}_{\{Y(\tau_i) = j\}} \qquad j = 1, \ldots, H, \tag{4.3}$$

where $\{\tau_i\}_{i \geq 1}$ is the sequence of the jump times of Y and it is a subset of $T_{(1)}, \ldots, T_{(H)}$, an order statistic of the lifetimes. Since

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

then $\mathcal{F}_t^Y = \mathcal{F}_t^U$, where $\mathcal{F}_t^U = \sigma\{U^1(s), \ldots, U^H(s), s \leq t\}$ and our problem reduces to find the conditional law of $X(t)$ given $\mathcal{F}_t^U$, $\pi_t(f) = \mathbb{E}[f(X(t)) | \mathcal{F}_t^U]$, $\forall f$ real valued function. In our frame, X is a $\mathcal{X}-$valued Markov process and its generator is given by

$$Lf(x) = L_0 f(x) + \sum_{j=1,\ldots,H} L_1^j f(x), \tag{4.4}$$

$$\begin{aligned} L_0 f(x) &= \lambda(x) \sum_{x' \in \mathcal{X}} [f(x') - f(x)] \mu_0(x, x'), \\ L_1^j f(x) &= \lambda(x) \sum_{x' \in \mathcal{X}} [f(x') - f(x)] \, \mathbf{1}_{|x'| = H - j} \cdot \mu_1(x, x'), \end{aligned}$$

where $\lambda(x) = \Lambda(x, H - |x|)$ and

$$\begin{aligned} \mu_0(x, x') &= M((x, H - |x|); (x', H - |x'|)) \mathbf{1}_{|x'| = |x| > 0}, \\ \mu_1(x, x') &= M((x, H - |x|); (x', H - |x'|)) \mathbf{1}_{0 \leq |x'| < |x|}. \end{aligned}$$

This implies that there are only two kind of jumps, for $X(t) = x$, $X(t)$ jumps following a transition function $\mu_0(x, x')$ and $Y(t)$ does not jump. Otherwise, $X(t)$ jumps following a transition function $\mu_1(x, x')$ and in this case, $Y(t) := H - |X(t)|$ increases. But, only the second kind of jumps are registered as observations.

With this setting, the filtering problem reduces to find the conditional law of $X(t)$ given the observations $Y(t)$, counting all the jumps of $|X(t)|$. Recalling that the filter is one of the solutions of the Kushner-Stratonovich equation, we are going to write it down in Theorem 4.3 below.

Theorem 4.3 *For any real function $f(x)$, $x \in \mathcal{X}$, the Kushner-Stratonovich equation is*

$$\begin{aligned}(4.5)\qquad \pi_t(f) = \nu_0(f) + \int_0^t [\pi_s(L_0 f) + \pi_{s-}(f)\pi_{s-}(\lambda_1) - \pi_{s-}(\lambda_1 f)] \cdot ds + \\ + \sum_{j=1,\dots,H} \int_0^t (\pi_{s-}(\lambda^j))^+ \cdot [\pi_{s-}(L_1^j f) - \pi_{s-}(f)\pi_{s-}(\lambda^j) + \pi_{s-}(\lambda^j f)] \cdot dU^j(s),\end{aligned}$$

where the $(P, \mathcal{F}_t)$-intensity function of the process $U^j(t)$, is given by

$$(4.6)\qquad \lambda^j(x) := \lambda(x) \sum_{x' \in \mathcal{X}} \mathbf{1}_{|x'|=H-j} \cdot \mu_1(x, x'), \qquad \forall j = 1, \dots, H,$$

and $\lambda_1(x) = \sum_{j=1,\dots,H} \lambda^j(x) = \lambda(x) \sum_{x' \in \mathcal{X}} \mu_1(x, x')$.

Proof. Equation (4.5) is obtained by applying the classical innovation method, for instance looking at [2], and a sketch of the proof is given for sake of self-consistency.

To achieve this program, noting that the process (X, Y, U) is still Markov, for $x \in \mathcal{X}$, $u \in \{0, 1\}^H$, $y \in \{0, 1, \dots, H\}$, the joint generator of (X, Y, U) is given by

$$L^{x,y,u} f(x, y, u) = L_0^{x,y,u} f(x, y, u) + \sum_{j=1}^H L_{j,1}^{x,y,u} f(x, y, u),$$

where

$$\begin{aligned} L_0^{x,y,u} f(x, y, u) &= \lambda(x) \sum_{x' \in \mathcal{X}} [f(x', y, u) - f(x, y, u)] \cdot \mu_0(x, x'), \\ L_{j,1}^{x,y,u} f(x, y, u) &= \lambda(x) \sum_{x' \in \mathcal{X}} [f(x', y - |x'| + |x|, u + e^j) - f(x, y, u)] \cdot \mathbf{1}_{|x'|=H-j} \cdot \mu_1(x, x') \end{aligned}$$

and e^j is the vector such that $e_i^j = \delta_i^j$, $i = 1, \dots, H$. The generator $L^{x,y,u}$, restricted to a function just depending on the first variable x, coincides with the operator L given in equation (4.4). Furthermore, for $j = 1, \dots, H$, the $(P, \mathcal{F}_t)$-intensity of U^j is given by the process $\lambda^j(x)$ where

$$\lambda^j(x) = \lambda(x) \sum_{x' \in \mathcal{X}} [u_j + e_j^j - u_j] \cdot \mathbf{1}_{|x'|=H-j} \cdot \mu_1(x, x') = \lambda(x) \sum_{x' \in \mathcal{X}} \mathbf{1}_{|x'|=H-j} \cdot \mu_1(x, x')$$

and the $(P, \mathcal{F}_t^U)$-intensity of U^j is given by $\pi_t(\lambda^j) = \mathbb{E}[\lambda^j(x) | \mathcal{F}_t^Y]$. Let M_t be the 0-mean $(P, \mathcal{F}_t)$-martingale defined as

$$(4.7)\qquad M(t) := f(X(t)) - f(X(0)) - \int_0^t Lf(X(s)) ds.$$

First of all, Theorem IV, T1 of [2] applies and, setting $\widehat{M}_t$ a 0-mean $(P, \mathcal{F}_t^Y)$-martingale, then

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

By the representation theorem (III, T17) of [2], $\widehat{M}_t$ can be written as

$$\widehat{M}_t = \sum_{j=1}^H \int_0^t K_{s-}(j) \cdot (dU_s^j - \pi_{s-}(\lambda^j)\ ds).$$

Finally, (IV, T2 of [2]) the equation for the filter, for any function $f(x)$, $x \in \mathcal{X}$, is

$$(4.8)\quad \pi_t(f) = \nu_0(f) + \int_0^t \pi_s(Lf)\, ds + \\ + \sum_{j=1}^{H} \int_0^t \pi_{s-}(\lambda^j)^+ \{\pi_{s-}(L_1^j f) - \pi_{s-}(\lambda^j)\pi_{s-}(f) + \pi_{s-}(\lambda^j f)\}(dU_s^j - \pi_{s-}(\lambda^j)\, ds).$$

In order to compute $L_1^j f$, let $M^j(t) := U^j(t) - \int_0^t \lambda^j(X(s))ds$, recall that $M(t)$ and $M^j(t)$ are $(P, \mathcal{F}_t)$-martingales and that $< M, M^j >_t = \int_0^t L_1^j f(X(s))\, ds$. By a standard stochastic calculus, we get that

$$L_1^j f(x) = \lambda(x) \sum_{x' \in \mathcal{X}} [f(x') - f(x)]\, \mathbf{1}_{|x'|=H-j} \cdot \mu_1(x, x'). \qquad \square$$

As far as uniqueness for the solution to (4.5) is concerned, we give the following result.

Theorem 4.4 *Equation (4.5) has a unique path-wise solution, necessarily, $\mathcal{F}_t^Y$-adapted.*

Proof. At any jump time τ_n, the filter is uniquely determined by the knowledge of $\pi_{\tau_n -}$. In fact, there exists a unique value j, $j = 1, \dots, H$, such that $\pi_t(\lambda^j) > 0$ a.s., i.e.

$$(4.9)\qquad \pi_{\tau_n}(f) = \frac{\pi_{\tau_n -}(L_1^j f) + \pi_{\tau_n -}(\lambda^j f)}{\pi_{\tau_n -}(\lambda^j)}\Big|_{j=Y_{\tau_n}}.$$

For $t \in [\tau_n, \tau_{n+1})$,

$$(4.10)\qquad \pi_t(f) = \pi_{\tau_n}(f) + \int_{\tau_n}^t [\pi_s(L_0 f) + \pi_s(f)\pi_s(\lambda_1) - \pi_s(\lambda_1 f)]ds.$$

Moreover, any two solutions π and π' of (4.10), with $\pi_{\tau_n}(f) = \pi'_{\tau_n}(f)$, are such that

$$\|\pi_t - \pi'_t\| \le C \int_{\tau_n}^t \|\pi_s - \pi'_s\| ds,$$

for a constant C explicitly computable. Then, Equation (4.10) is Lipschitz with respect to the bounded variation norm, which means that the solution is unique and necessarily $\mathcal{F}_t^Y$-adapted. $\square$

The Kushner-Stratonovich equation has a natural recursive structure, following the jump times $\{\tau_n\}_{n \ge 0}$ and this can be easily seen if we write it in the form (4.5). The structure of this equation shows that $\pi_t(f)$ has a deterministic behavior between two consecutive jumps times of Y. At any jump time of Y, say τ_n, $\pi_t(f)$ jumps and its jump-size is given by (4.9). We want to stress that Equation (4.9) shows that at a jump time τ_n, $\pi_{\tau_n}(f)$ is completely determined by the knowledge of $\pi_t(f)$ in the interval $[\tau_{n-1}, \tau_n)$. In fact, for any function f,

$$\pi_{\tau_n^-}(f) = \lim_{t \to \tau_n^-} \pi_t(f).$$

To obtain an useful representation for the filter, for $t \in [\tau_n, \tau_{n+1})$, in the next section, we are going to introduce a linearization procedure and to this end, recalling that, by construction, $0 < \underline{l} \le \lambda(x) \le \overline{l}$, $\forall x \in \mathcal{X}$, from now on, we need the further condition.

Assumption 4.5

$$\forall x, x' \in \mathcal{X}: \quad |x| > |x'| \geq 0 \quad \textit{then} \quad \mu_1(x, x') > 0. \tag{4.11}$$

Roughly speaking, if the number of dead particles increases, every transition must be possible, i.e. any of such transition probability is strictly positive. Then, this assumptions allows us to define the positive quantity $\underline{\mu} := \min_{x,x' \in \mathcal{X}} \{\mu_1(x, x') : |x| > |x'| \geq 0\}$.

Remark 4.6 *As a consequence of the last assumption, recalling (4.6), then*

$$\lambda^j(x) \geq \underline{l} \sum_{x' \in \mathcal{X}} \mathbf{1}_{|x'|=H-j} \cdot \mu_1(x, x') \geq \underline{l} \cdot \underline{\mu}.$$

4.3. A representation for the filter.

The behavior of the filter π_t for $t \in [\tau_n, \tau_{n+1})$, where the Kushner-Stratonovich equation (4.5) reduces to Equation (4.10), is a nonlinear one. Thus, an explicit expression for the solutions of the equation (4.10) is, in general, not available, but it is possible to provide a hand-able representation for it. This representation, described with a method which is a modification of that proposed in [11], is an essential tool to achieve the approximating discrete time model which we are going to give later on. Let us introduce the linearized equation

$$\begin{aligned} \rho_t(f) &= \nu_0(f) + \int_0^t \{\rho_s(L_0 f) - \rho_s(\lambda_1 f)\} ds + \\ &\quad + \sum_{j=1}^H \int_0^t \{\rho_{s-}(\lambda^j f) - \rho_{s-}(f) + \rho_{s-}(L_1^j f)\} dU_s^j, \end{aligned} \tag{4.12}$$

which is obtained by (4.5) dropping out the nonlinear terms. Equation (4.12) not only admits a unique solution in the weak sense, but, by Lipschitz arguments, it admits a unique path-wise solution which is necessarily $\mathcal{F}_t^U$-adapted. This last claim can be proven with a procedure similar to that used in Theorem 4.4.

Proposition 4.7 *Equation (4.12) admits at least one solution $\mathcal{F}_t^U$-adapted. In addition such solution $\rho_t(f)$, $\forall t$, is a finite positive measure, $e^{-t \cdot \underline{l}}(1 \wedge \underline{l} \cdot \underline{\mu}) < \rho_t(1) \leq \bar{l} \vee 1$, and $\pi_t(f) = \dfrac{\rho_t(f)}{\rho_t(1)}$.*

Proof. First we claim that, for ρ_t any solution of (4.12), $\dfrac{\rho_t(f)}{\rho_t(1)}$ provides a solution of (4.5) and then coincides with the filter up to time $t_0 = \inf\{t \geq 0 : \rho_t(1) = 0\}$. In fact, assuming that $\rho_t(1) > 0$, it is easy to verify that $\pi_t(f) = \frac{\rho_t(f)}{\rho_t(1)}$ solves equation (4.5). In particular, for $t \in [\tau_n, \tau_{n+1})$, the thesis is achieved by computing $\frac{d}{dt} \frac{\rho_t(f)}{\rho_t(1)}$, taking into account Equation (4.12) and observing that such derivative coincides with $\frac{d}{dt} \pi_t(f)$.

Then, let us to construct a solution of (4.12) which has the required properties and such that $\rho_t(1) > 0$ for any t. Let $X_{s,x}(t)$ be a process, with initial condition (s, x), $s \geq 0$, $x \in \mathcal{X}$, and generator L, given by (4.4). Let $P_{s,x}$ its law on $D_{\mathcal{X}}[s, T]$. Then, by the Feynman Kac's formula,

$$\forall t \in \big[\tau_i, \tau_{i+1}\big) \qquad \begin{cases} \rho_t(f) = \sum_{x\in\mathcal{X}} E^{P_{s,x}} \Big[f(X_{s,x}(t))\ exp \Big\{ -\int_s^t \lambda_1(X_{s,x}(u))\ du \Big\} \Big] \Big|_{s=\tau_i} \dfrac{\rho_{\tau_i}(\{x\})}{\rho_{\tau_i}(1)} \\ \\ \rho_t(1) = \sum_{x\in\mathcal{X}} E^{P_{s,x}} \Big[exp \Big\{ -\int_s^t \lambda_1(X_{s,x}(u))\ du \Big\} \Big] \Big|_{s=\tau_i} \dfrac{\rho_{\tau_i}(\{x\})}{\rho_{\tau_i}(1)} \geq e^{-(t-\tau_i)\cdot \bar{l}} > 0 \end{cases}$$

$$t = \tau_{i+1} \qquad \begin{cases} \rho_{\tau_{i+1}}(f) = \rho_{\tau_{i+1}-}(\lambda^j f) + \rho_{\tau_{i+1}-}(L_1^j f)\Big|_{j=Y_{\tau_{i+1}}} \\ \\ \rho_{\tau_{i+1}}(1) = \rho_{\tau_{i+1}-}(\lambda^j)\Big|_{j=Y_{\tau_{i+1}}} \geq \underline{l} \cdot \underline{\mu} \cdot \rho_{\tau_{i+1}-}(1), \end{cases}$$

where the last inequality is a consequence of Remark 4.6 and by induction the thesis is reached. Finally observe that $\rho_t(1) \leq 1$ for $t \in \big[\tau_i, \tau_{i+1}\big)$ and for $t = \tau_{i+1}$

$$\rho_{\tau_{i+1}}(1) = \rho_{\tau_{i+1}-}(\lambda^j)\Big|_{j=Y_{\tau_{i+1}}} \leq \bar{l} \cdot \rho_{\tau_{i+1}-}(1) \leq \bar{l}. \qquad \square$$

5. An approximating discrete time model.

This construction provides a discrete time approximating model strongly convergent to the original one, [4].

On a probability space $(\Omega, \mathcal{F}, P)$, let $\{X_n\}_{n\geq 0}$ be the Markov chain defined by initial law ν_0 and transition function $\mu(x,x') = \mu_0(x,x') + \mu_1(x,x')$. Let $\{V_i\}_{i\geq 1}$ be a sequence of random variables independent and exponentially distributed with parameter 1 and independent on $\{X_n\}_{n\geq 0}$. Let $h > 0$ be fixed and set, successively,

$$\theta_0^h := 0 \qquad \text{and} \qquad \theta_n^h = h \sum_{i=1}^{n} \left[\frac{V_i}{h\ \lambda(X_{i-1})} \right] + nh \quad \text{for} \quad n > 0,$$

where $[a]$ denotes the integer part of a. Then, on a finite time horizon $[0,T]$, with $T > 0$, for $t = kh$, with $k = 0, 1, \ldots$, such that $kh \leq T$, the approximating process is defined as

$$X^h(t) = \sum_{n\geq 0} X_n\ \mathbf{1}_{\{\theta_n^h \leq t < \theta_{n+1}^h\}}.$$

Hence, on the space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, P)$, X^h is a discrete time Markov chain, $\{\theta_n^h\}_{n\geq 1}$ is the sequence of its jump times. The family of the transition probabilities of the process X^h can be obtained by Proposition 5.1 below, just proved in [4].

Proposition 5.1 *The process X^h is a discrete time Markov chain with transition probabilities given by*

$$\begin{aligned} \mu^h(x,x') &= P\big(X^h((n+1)h) = x' \big| X^h(nh) = x\big) = \\ &= \delta_{\{x,x'\}} e^{-h\lambda(x)} + (\mu_0(x,x') + \mu_1(x,x'))(1 - e^{-h\lambda(x)}). \end{aligned} \tag{5.1}$$

We introduce now the discrete time observations process setting

$$Y^h(t) = H - |X^h(t)|, \quad U_t^{jh} = \sum_{k\geq 1} \mathbf{1}_{\{\tau_k^h \leq t\}}\ \mathbf{1}_{Y^h(\tau_k^h)=j} \quad \text{and} \quad N_t^h = \sum_{j=0}^{H} U_t^{jh} = \sum_{k\geq 1} \mathbf{1}_{\{\tau_k^h \leq t\}},$$

where $\{\tau_k^h\}_{k\geq 0}$ is the sequence of the jump times of the process Y^h and it is a subsequence of $\{\theta_k^h\}_{k\geq 0}$, the sequence of the jump times of the process X^h.

Again, as in the continuous time case, $\mathcal{F}_t^{Y^h} = \mathcal{F}_t^{U^h}$, with $U^h = \{U^{jh}\}_{j=0,1,\ldots,H}$. Then, in order to find the conditional law of $X^h(t)$ given $\mathcal{F}_t^{U^h}$, a discrete time filtering problem has to be considered and $\pi_t^h = \mathcal{L}(X^h(t)|\mathcal{F}_t^{U^h})$ satisfies the equation

$$
\begin{aligned}
(5.2)\quad \pi_{nh}^h(f) = \nu_0(f) + \sum_{k=1}^{n} \Big[\pi_{(k-1)h}^h(L_0^h f) + \pi_{(k-1)h}^h(\lambda^h)\pi_{(k-1)h}^h(f) - \pi_{(k-1)h}^h(\lambda^h f)\Big] \cdot \\
\cdot(1-\pi_{(k-1)h}^h(\lambda^h))^+ \cdot (1-\Delta N_{kh}^h) + \\
+\sum_{j=0}^{H}\sum_{k=1}^{n}\Big[\pi_{(k-1)h}^h(\lambda_j^h f) - \pi_{(k-1)h}^h(\lambda_j^h)\ \pi_{(k-1)h}^h(f) + \pi_{(k-1)h}^h(L_1^{jh} f)\Big] \cdot \\
\cdot\pi_{(k-1)h}^h(\lambda_j^h)^+ \cdot \Delta U_{kh}^{jh},
\end{aligned}
$$

where

$$
\begin{aligned}
L_0^h f(x) &= \sum_{x'\in\mathcal{X}} [f(x') - f(x)]\ \mathbf{1}_{|x|=|x'|>0}\ \mu^h(x,x'), \\
L_1^h f(x) &= \sum_{j=1}^{H} L_1^{jh} f(x) = \sum_{j=1}^{H}\sum_{x'\in\mathcal{X}} [f(x')-f(x)]\cdot \mathbf{1}_{|x'|\neq|x|}\cdot \mathbf{1}_{|x'|=H-j}\cdot \mu^h(x,x'), \\
\lambda_j^h(x) &= \sum_{x'\in\mathcal{X}} \mathbf{1}_{|x'|\neq|x|}\cdot \mathbf{1}_{|x'|=H-j}\cdot \mu^h(x,x') \qquad \text{and} \\
\lambda^h(x) &= \sum_{j=1}^{H} \lambda_j^h(x) = \sum_{x'\in\mathcal{X}} \mathbf{1}_{|x'|\neq|x|}\cdot \mu^h(x,x').
\end{aligned}
$$

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

$$
(5.3)\qquad \lambda_j^h(x) \leq \lambda^h(x) \leq 1 - e^{-h\cdot\bar{l}} < 1, \qquad \forall j = 1,\ldots,H.
$$

Such solution can be explicitly written down. In place of this, we are going to provide a linearized version of (5.2) as a useful tool to prove the convergence of the approximating discrete time model to the continuous time one.

Proposition 5.2 *The equation*

$$
\begin{aligned}
(5.4)\quad \rho_{nh}^h(f) = \nu_0(f) + \sum_{k=1}^{n}\{\rho_{(k-1)h}^h(L_0^h f) - \rho_{(k-1)h}^h(\lambda^h f)\}(1-\Delta N_{kh}^h) + \\
+\sum_{j=1}^{H}\sum_{k=1}^{n}(1-e^{-h})^+\Big\{\rho_{(k-1)h}^h(\lambda_j^h f) - (1-e^{-h})\rho_{(k-1)h}^h(f) + \rho_{(k-1)h}^h(L_1^{jh} f)\Big\}\Delta U_{kh}^{jh}
\end{aligned}
$$

admits a unique solution $\mathcal{F}_t^U$-adapted. Such solution $\rho_t^h(f)$, for any $t = nh$, is a finite positive measure, $0 < \rho_t^h(1) \leq (2\cdot\bar{l})^{N_t^h} \vee 1$, $\forall\ h\cdot(\underline{l}\vee 1) < \log 2$ and $\pi_t^h(f) = \dfrac{\rho_t^h(f)}{\rho_t^h(1)}$.

Proof. Let us use, here, the same argument as in Proposition 4.7. Then we have just to prove that $\rho^h_t(1)$ cannot vanish for $t \in [\tau^h_i, \tau^h_{i+1})$. For $t = nh$, with $\tau^h_i \leq nh < \tau^h_{i+1}$, recalling the inequality (5.3),

$$\rho^h_{nh}(1) = \rho^h_{(n-1)h}(1) - \rho^h_{(n-1)h}(\lambda^h) = \rho^h_{(n-1)h}(1-\lambda^h) \geq e^{-h\cdot\bar{l}}\ \rho^h_{(n-1)h}(1),$$

which by induction implies

$$\rho^h_t(1) \geq e^{-T\cdot\bar{l}}. \tag{5.5}$$

At a jump time $\tau^h_i = kh$, $\Delta N^h_{kh} = 1$ and, since $\lambda^h_j(x) = \lambda_j(x)\cdot(1-e^{-h\lambda(x)})$,

$$\rho^h_{kh}(1) = (1-e^{-h})^{+}\rho^h_{(k-1)h}(\lambda^h_j)\big|_{j=Y^h_{nh}} \geq \underline{l}\cdot\underline{\mu}\cdot\frac{1-e^{-h\underline{l}}}{1-e^{-h}}\cdot\rho^h_{(k-1)h}(1) \geq \left(\frac{1}{2}\,\underline{l}^2\cdot\underline{\mu}\right)^{N^h_{kh}}.$$

Moreover, for $\Delta N^h_{kh} = 0$, $\rho^h_{kh}(1) \leq \rho^h_{(k-1)h}(1)$ and for $\Delta N^h_{kh} = 1$,

$$\rho^h_{kh}(1) \leq \frac{\rho^h_{(k-1)h}(\lambda^h)}{1-e^{-h}} \leq \frac{1-e^{-\bar{l}\cdot h}}{1-e^{-h}}\ \rho^h_{(k-1)h}(1). \qquad \square$$

Let $S^h := (X^h, Y^h, U^h, N^h, \pi^h)$ denote the piecewise constant cadlag continuous time interpolation of the processes above introduced. Moreover, let $S := (X, Y, U, N, \pi)$, and $\mathcal{S} = \mathcal{X} \times \{0, 1, \ldots, H\} \times \{0,1\}^H \times I\!N \times \Pi(\mathcal{X})$, where $\Pi(\mathcal{X})$ is the space of probability measure on $\mathcal{X}$.

Theorem 5.3 *The process S^h converges to the process S a.s., as $h \to 0$, with respect to the Skorohod topology on the space $D_{\mathcal{S}}[0,T]$.*

The proof of Theorem 5.3 follows, in some sense, the same line of Theorem 6.1 in [9] and we recall it for sake of self-completeness.

Let us focus that overall this Section the results are true for almost all fixed ω. More precisely, let $N(t) := \sum_{i\geq 1} \mathbf{1}_{\{\tau_i \leq t\}}$ be the process counting the jump of Y, then τ_{N_T} is a continuous random variable and $P(\tau_{N_T} = T) = 0$. Note that the convergence of S^h to S claimed in Theorem 5.3 holds for any $\omega \in \{\tau_{N_T} < T\}$. The proof of this theorem is a consequence of the next results and the proof follows the same lines as in [3] Section 4.

Proposition 5.4 *For $h < \dfrac{T-\tau_{N_T}}{N_T}$, then $\tau^h_{N_T} \leq T$, which in turn implies $N_T = N^h_T$ a.s..*

Proof. For any $\omega \in \{\tau_{N_T} < T\}$, we can choose h such that $0 < h \leq \dfrac{T-\tau_{N_T}}{N_T}$. By definition $\tau^h_k \leq \tau_k + kh, \forall k$, then $\tau^h_{N_T} \leq \tau_{N_T} + N_T h \leq T$. Moreover, since $\tau_k \leq \tau^h_k, \forall k$, then $N_T \geq N^h_T$ which means that $\tau^h_{N^h_T} \leq \tau^h_{N_T}$. But observing that $\tau^h_{N^h_T}$ is the last jump time before T, $\tau^h_{N^h_T} \leq T \leq \tau^h_{N_T}$. This last inequality joint with $T \geq \tau^h_{N_T}$, give us that $\tau^h_{N^h_T} = \tau^h_{N_T}$, i.e. $N_T = N^h_T$, (recalling that $\{\tau^h_k\}_{k\geq 0}$ is a strictly monotone sequence). $\square$

On the event $\left(N_T = N^h_T\right)$, 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 intervals $[\tau^h_k, \tau^h_{k+1})$ into $[\tau_k, \tau_{k+1})$, $\forall k < N_T$, and $[\tau^h_{N_T}, T)$ into $[\tau_{N_T}, T)$.

(ii) $\sup_{t\in[0,T]} |\alpha_h(t) - t| = \max_{k\leq N_T} |\tau_k - \tau^h_k| \leq \max_{k\leq N_T} kh = N_T h.$

(iii) $|X(\alpha_h(t)) - X^h(t)| = |Y(\alpha_h(t)) - Y^h(t)| = |\, U(\alpha_h(t)) - U^{jh}(t)| = |N_{\alpha_h(t)} - N^h_t| = 0.$
Then, in order to reach the result claimed in Theorem 5.3, the convergence of the filters is just enough. Therefore, by definition of the Skorohod topology, such Theorem is a consequence of Theorem 5.5 below.

Theorem 5.5 *Under the assumptions prevailing in this paper, (in particular (4.11)),*

i) $\|\pi_{\alpha_h(t)} - \pi^h_t\| \le \dfrac{2 \cdot e^{T\cdot \bar{l}}}{1 \vee \underline{l} \cdot \underline{\mu}} \cdot \left\|\rho_{\alpha_h(t)} - \rho^h_t\right\|$ *and*

ii) $\left\|\rho_{\alpha_h(t)} - \rho^h_t\right\| \le (1{+}2\,\bar{l})^{N_t}\, e^{2t\cdot\bar{l}}\, Ch$, *for a suitable quantity* $C = C(T, N_T, \bar{l}) > 0,$

namely $\|\pi_{\alpha_h(t)} - \pi^h_t\| \le C \cdot h.$

Proof. First of all, noting that $\pi_t(f) = \dfrac{\rho_t(f)}{\rho_t(1)}$ and $\pi^h_t(f) = \dfrac{\rho^h_t(f)}{\rho^h_t(1)}$, by Proposition 4.7 and by Proposition 5.2, respectively, then

$$\begin{aligned}\|\pi_{\alpha_h(t)} - \pi^h_t\| &= \frac{1}{\rho_{\alpha_h(t)}(1)}\left\|\rho_{\alpha_h(t)} - \rho^h_t\right\| + \left\|\frac{\rho^h_t}{\rho_{\alpha_h(t)}(1)\rho^h_t(1)}(\rho_{\alpha_h(t)}(1) - \rho^h_t(1))\right\| \le \\ &\le \frac{2}{\rho_{\alpha_h(t)}(1)}\left\|\rho_{\alpha_h(t)} - \rho^h_t\right\|\end{aligned}$$

and since, by Proposition 4.7, $e^{-t\cdot\underline{l}}(1 \wedge \underline{l}\cdot\underline{\mu}) < \rho_t(1) \le \bar{l} \vee 1$, the first result is achieved. For the second claim, since the first derivative of $\alpha_h(t)$, say $\alpha'_h(t)$, exists for any $t \in [0,T]$ but a finite number of points and $\alpha_h(\tau_i) = \tau_i, \forall i$, (4.12) implies that

$$\begin{aligned}\rho_{\alpha_h(t)}(f) &= \nu_0(f) + \int_0^t \{\rho_{\alpha_h(s)}(L_0 f) - \rho_{\alpha_h(s)}(\lambda_1 f)\}\alpha'_h(s)ds \\ &\quad + \sum_{j=1,\dots,H}\int_0^t \{\rho_{\alpha_h(s)-}(\lambda^j f) - \rho_{\alpha_h(s)-}(f) + \rho_{\alpha_h(s)-}(L^j_1 f)\}dU^{jh}_s\end{aligned}$$

and, (5.4) implies that

$$\begin{aligned}\rho^h_t(f) &= \nu_0(f) + \int_0^t \frac{1}{h}\{\rho^h_s(L^h_0 f) - \rho^h_s(\lambda^h f)\}ds + \int_0^t \{\rho^h_{s-}(\lambda^h f) - \rho^h_{s-}(L^h_0 f)\}dN^h_s \\ &\quad + \sum_{j=1,\dots,H}\int_0^t (1-e^{-h})^+\{\rho^h_{s-}(\lambda^h_j f) - (1-e^{-h})\rho^h_{s-}(f) + \rho^h_{s-}(L^{jh}_1 f)\}dU^{jh}_s.\end{aligned}$$

Thus, since by tedious but straightforward computations, successively we get

$$|L_0 f(x) - \lambda_1(x) f(x)| \le 3\,\bar{l}\,\|f\|,$$

$$\left|\int_0^t (\alpha'_h(s) - 1)\,ds\right| \le 3N_T h \qquad \text{and} \qquad \left|\frac{L^h_0 f(x) - \lambda^h(x) f(x)}{h}\right| \le 2\,\bar{l}\,\|f\|,$$

then

$$\int_0^t \left|\{\rho_{\alpha_h(s)}(L_0 f) - \rho_{\alpha_h(s)}(\lambda_1 f)\}\alpha'_h(s) - \frac{1}{h}\{\rho^h_s(L^h_0 f) - \rho^h_s(\lambda^h f)\}\right| ds \le$$

$$\le 6\, N_T\, \bar{l}\,\|f\| \cdot h \sup_{s\in[0,T]} \rho_s(1) + 3\,\bar{l}^2\,\|f\|\,T\ \cdot h \sup_{s\in[0,T]} \rho_s(1) + 2\,\bar{l}\,\|f\| \int_0^t \left\|\rho_{\alpha_h(s)} - \rho^h_s\right\| ds.$$

Moreover

$$\sum_{j=1,\dots,H}\int_0^t \Big|\rho_{\alpha_h(s)-}(\lambda^j f) - \rho_{\alpha_h(s)-}(f) + \rho_{\alpha_h(s)-}(L_1^j f) - \rho_{s-}^h(\lambda^h f) + \rho_{s-}^h(L_0^h f)$$
$$-(1-e^{-h})^+\{\rho_{s-}^h(\lambda_j^h f) - (1-e^{-h})\rho_{s-}^h(f) + \rho_{s-}^h(L_1^{jh} f)\}\Big|\, dU_s^{jh} \le$$
$$\le \|f\|\;(2+\bar l)\;\bar l\cdot h\cdot \sup_{s\in[0,T]}\rho_s^h(1)\cdot N_T^h + (1+\bar l)\;\|f\|\;\int_0^t \big\|\rho_{\alpha_h(s)} - \rho_s^h\big\|\, dN_s^h.$$

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

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

By using Gronwall Lemma for $t \in [\tau_i, \tau_{i+1}), \forall i$, and taking into account the update at the jump times, finally, the thesis is reached. □

Acknowledgment. The author thanks the chairman and all the scientific and local committee for organizing the Biocomp2005, Diffusion Processes in Neurobiology and Subcellular Biology, Vietri sul Mare, Italy, 12-16 December 2005.

References

[1] E. Arjas, P. Haara and I. Norros, Filtering the histories of a partially observed marked point process, Stochastic Process. Appl., **40**, (1992), p. 225 – 250.

[2] P. Brémaud. **Point processes and queues** Springer-Verlag, (1980).

[3] A. Calzolari and G. Nappo, A filtering problem with counting observations: approximations with error bounds. Stochastics Stochastics Rep., **57**, (1996) 71 – 87.

[4] A. Calzolari and G. Nappo, A filtering problem with counting observations: error bounds due to the uncertainty on the infinitesimal parameters, Stochastics Stochastics Rep., **61**, (1997), p. 1 – 19.

[5] A. Calzolari and G. Nappo, The filtering problem in a model with grouped data and counting observation times, preprint Dipartimento di Matematica G. Castelnuovo, Universitá di Roma La Sapienza, (2001).

[6] S. N. Ethier and T. G. Kurtz, **Markov processes: characterization and convergence** J. Wiley (1986).

[7] A. Gerardi, F. Spizzichino and B. Torti. Exchangeable mixture models for lifetimes: the role of "occupation number", Statist. Probab. Lett. **49**, (2000), p. 365 – 375.

[8] A. Gerardi, F. Spizzichino and B. Torti. Filtering equations for the conditional law of residual lifetimes from a heterogeneous population, J. Appl. Probab. **37**, (2000), p. 823 – 834.

[9] A. Gerardi and P. Tardelli. Heterogeneous population dynamical model: a filtering problem. Journal of Applied Probability, **42**, (2005), p. 346 – 361.

[10] F. P. Kelly. **Reversibility and stochastic networks** Wiley, (1994).

[11] W. Kliemann, G. Koch, F. Marchetti. On the unnormalized solution of the filtering problem with counting process observations, IEEE Trans. Inf. Theory, **36**, 6, (1990), p. 1415 – 1425.

[12] T.G. Kurtz. Martingale problems for conditional distribution of Markov processes, Electronic J. Probab. **3**, 9, (1998), p. 1 – 29.

[13] T.G. Kurtz and D. Ocone. Unique characterization of conditional distributions in nonlinear filtering, Ann. Probab. **16**, (1988), p. 80 – 107.

Paola Tardelli
Dipartimento di Ingegneria Elettrica, Università dell'Aquila
I-67040, Poggio di Roio, L'Aquila, Italy
E-mail: tardelli@ing.univaq.it