The Filtering Problem in a Model with Grouped Data and Counting Observation Times

A. Calzolari

Università di Roma "Tor Vergata"

G. Nappo

Università di Roma "La Sapienza"

1 Introduction

In this paper we consider a stochastic system (x_t, y_t, z_t). We assume that (x_t) is an unobservable process, while (z_t) is a counting process defined by $0 = S_0 < S_1 < \ldots < S_n < \ldots$, and it is completely observable. The process (y_t) can be observed only at the random times $\{S_i\}_{i \geq 0}$.

The information about (y_t) at time S_i can be either *precise*, or *grouped*, in the sense that either we can observe y_{S_i} exactly, or we can observe only whether y_{S_i} is in Γ_h, where $\{\Gamma_h\}$ is a partition of the state space of the $y-$component.

Such models can arise in different contexts.

A first example arises in stochastic finance modelling (see e.g. Frey and Runggaldier [9]). The process (x_t) represents the *historical volatility process*, while (y_t) represents the *shadow asset price process*, and finally S_i are the *transaction times*.

A second example arises in population modelling, when a branching process models a population growth and the size of a subpopulation is observed up to time t. Then the observation consists of the subpopulation size at its own branching times. The process (x_t) represents the population size, (y_t) represents the subpopulation size, while (z_t) counts the jump times of y_s for $s \in [0, t]$. In this sense it fits into the above situation. This example is studied by Ceci and Gerardi ([5]), when the information is *precise* in the above sense.

A third example can arise in a reliability context: a system is endowed of a pre-alarm mechanism (y_t) and of a control mechanism (z_t): both the control (z_t) and the pre-alarm (y_t) mechanisms are modelled as random clocks driven by the same random process (x_t). At the control times it is possible to know how many pre-alarms have been registered up to this time. For example, given a first critical value α_0 and a second critical value α_1 , with $\alpha_0 < \alpha_1$, the pre-alarm mechanism is "on" (or active) only when x_t takes values in (α_0, ∞), at a rate increasing with x_t, while the control mechanisms, beside being "on" at a constant rate, counts the pre-alarm times, when x_t is in (α_1, ∞). To fix ideas we write down a possible formal generator $\mathcal{L}$ for the system (x_t, y_t, z_t)

$$\mathcal{L}f(x,y,z) =$$

$$\begin{aligned}
&= b(x,y,z)\,\partial_x f(x,y,z) + \frac{1}{2}\,c(x,y,z)^2\,\partial_{xx} f(x,y,z) \\
&+ \lambda(x,y,z)\,\mathbb{I}_{\{x\in(\alpha_0,\alpha_1]\}}\,[f(x,y+1,z) - f(x,y,z)] \\
&+ \mu(x,y,z)\,\mathbb{I}_{\{x\in(\alpha_1,\infty)\}}\,[f(x,y+1,z+1) - f(x,y,z)] \\
&+ \gamma\,[f(x,y,z+1) - f(x,y,z)].
\end{aligned}$$

In [14] a similar model is considered and is given the filter of the random parameter θ of a Poisson process (y_t) observed at random times which generate a counting process (z_t). In this model the x-component is the random parameter θ, and in the above framework corresponds to the case when $b = 0$ and $c = 0$, $\lambda(x,y,z) = x$, $\alpha_0 = 0$, $\alpha_1 = \infty$ (and then $\mu(x,y,z)$ can be taken equal to zero), finally the intensity of (z_t) is not constant, but is linearly increasing with (y_t).

The grouped information case can arise when only rounded transaction prices are available, in the financial example. Considering the induced partition as a grid, the filter with grouped information could be interpreted as an approximation to the original filter, different from the one considered in [9].

In the remaining examples, the information is grouped when, for instance, the number of pre-alarms (or the size of the subpopulation) is given as the number of clusters of pre-alarms (or of individuals), each cluster being of a fixed size.

A final example comes from queueing theory. Let (Q_t) be an $M/M/1$ queue. Then it can be represented as

$$Q_t = Q_0 + A_t - \int_0^t \mathbb{I}_{\{Q_{s^-}\neq 0\}} dN_s \tag{1}$$

where (A_t), the arrival process, is a Poisson process with intensity λ and (N_t) is a Poisson process with intensity μ. The two Poisson processes are supposed independent. As usual, the departure process is denoted by (D_t), and defined as

$$D_t := \int_0^t \mathbb{I}_{\{Q_{s^-}\neq 0\}} dN_s. \tag{2}$$

Suppose that one cannot observe the queue completely, but that one can observe all the departure times, whether the queue is empty, and whether the queue is above or below a certain level h. A motivation for this model comes from the following situation. There is one server, and there is a waiting room of capacity h. However, when the room is full, the clients are not rejected, but they have to wait outside. The server can see only the client he is serving, and can get the information that the queue is continuing outside. However he cannot see how many people are waiting. Of course he can also see whether the queue is zero, that is when the room is empty. Moreover he can see when a client is leaving the queue, i.e. the departure times. This means that its observation times are a counting process (Z_t), namely

$$Z_t := \int_0^t \mathbb{I}_{\{Q_s \in \{1\} \cup \{h+1\}\}} dA_s \,+\, D_t, \tag{3}$$

or equivalently

$$Z_t := \int_0^t \mathbb{I}_{\{Q_{s^-} \in \{0\} \cup \{h\}\}} dA_s \,+\, D_t. \tag{4}$$

At any time we know whether the queue is 0, or in $\{1, \ldots, h\}$, or in the set $\{h+1, \ldots\}$. At the observation time S_i we also know whether S_i is an arrival time, or a departure time. This situation can also be modelled considering the bivariate point process

$$Z_t^A := \int_0^t \mathbb{I}_{\{Q_{s^-} \in \{0\} \cup \{h\}\}} dA_s, \quad Z_t^D := D_t, \tag{5}$$

with jump times $\{S_l^A\}$ and $\{S_l^D\}$. At the jump times $\{S_i\}$ we know in which set is Q_{S_i} and whether $S_i \in \{S_l^A\}$ or $S_i \in \{S_l^D\}$.

There is a peculiarity distinguishing the examples:

- in the first example there are no common jump times between the observation times of (z_t) and the jump times of the process (y_t),
- in the second example the observation times of (z_t) and the jump times of the process (y_t) coincide,
- in the third example we can assume that

$$y_t = N_t^{(1)} + N_t^{(3)}$$
$$z_t = N_t^{(2)} + N_t^{(3)}$$

  where $N_t^{(i)}, i \in \{1, 2, 3\}$, count the number of the "uncontrolled" pre-alarm times, of the "routine" control times, and of the pre-alarm that are also control times, respectively. It is important to note that we are assuming that we are not able to detect whether a control time is a "routine" control time or a pre-alarm time.
- In the last example there is no $x-$component, and the observation times are a subset of the jump times of the partially observed component. This makes no basic difference with respect to the third example. However, in the last example, we can detect which kind of jump has been observed, i.e. we can detect whether it is an arrival or a departure time. In the last Section we show how to handle with this difference.

In the first three situations the problem is to estimate the unobservable process (x_t) and more precisely to look for the conditional law of x_t, given the observation up to time t. As will be clear later, it is natural, and in a certain sense necessary, to derive instead the conditional law of (x_t, y_t). In the last situation we are directly interested in the law of the partially observed queue.

In this paper we assume a Markovian structure for the system, nevertheless this assumption is not necessary for the model. This problem can arise also for models with different assumptions. For instance, when the system is a (not necessarily Markovian) pure jump process with values in a countable space E, the system can be represented as a marked point process. Then the problem can be viewed as a filtering problem of a partially observed marked point process, and the techniques of Arjas, Haara and Norros ([1]) could be used.

We assume that (x_t) is a d-dimensional diffusion process with jumps, (y_t) is a k-dimensional diffusion process with jumps, and (z_t) is a counting process. The drift and diffusion coefficients and the jump intensities may depend on (x_t, y_t, z_t). It is important to note that common jump times are allowed for all the three component. The process (y_t) can also be a jump process with values in a discrete space $\mathcal{Y}$.

The kind of processes we have in mind are solutions of the martingale problem for the operator $\mathcal{L} \subset \mathbf{B}(\mathbb{R}^d \times \mathbb{R}^k \times \mathbb{N}) \times \mathbf{B}(\mathbb{R}^d \times \mathbb{R}^k \times \mathbb{N})$

$$\mathcal{L}f(x,y,z) = \tag{6}$$

$$\begin{aligned}
&= B(x,y,z)\nabla_{(x,y)}f(x,y,z) + \frac{1}{2}tr\{\nabla^2_{(x,y)}f(x,y,z)C(x,y,z)C^*(x,y,z)\} \\
&+ \int_{D_0(x,y,z)} [f(x+K_0(x,y,z;\zeta),y,z) - f(x,y,z)]\,\nu(d\zeta) \\
&+ \int_{D_1(x,y,z)} [f(x+K_0(x,y,z;\zeta),y+K_1(x,y,z;\zeta),z) - f(x,y,z)]\,\nu(d\zeta) \\
&+ \int_{D_2(x,y,z)} [f(x+K_0(x,y,z;\zeta),y,z+1) - f(x,y,z)]\,\nu(d\zeta) \\
&+ \int_{D_3(x,y,z)} [f(x+K_0(x,y,z;\zeta),y+K_3(x,y,z;\zeta),z+1) - f(x,y,z)]\,\nu(d\zeta)
\end{aligned}$$

where $(\Sigma, \mathbb{S})$ is a measurable space, $\nu(d\zeta)$ is a σ-finite measure on Σ, $D_l(x,y,z)$, for $l = 0,1,2,3$, are pairwise disjoint measurable subsets of Σ. We introduce the notation

$$K_2(x,y,z;\zeta) := \mathbb{I}_{D_2(x,y,z)}(\zeta),$$

so that K_2 is measurable. We assume that also the functions K_l are measurable, for the remaining values of l. Finally we assume that

$$D_l(x,y,z) = \{\zeta \in \Sigma : K_l(x,y,z;\zeta) \neq 0\}, \qquad l = 1,2,3,$$

and that the set where K_0 is different from zero can be represented as

$$D_0(x, y, z) \subset \{\zeta \in \Sigma : K_0(x, y, z; \zeta) \neq 0\} \subset \cup_{l=0,1,2,3} D_l(x, y, z).$$

The vector function B can be represented as $B = (b, \beta)$,with $b(\cdot)$ and $\beta(\cdot)$ taking values in $\mathbb{R}^d$ and $\mathbb{R}^k$, respectively. The matrix function C can be represented as $C = (c, \gamma)$, with $c(\cdot)$ and $\gamma(\cdot)$ taking values in the matrix spaces of dimension $d \times m$ and $k \times m$, respectively. This decompositions of B and C are unnecessary here, but will be used in the sequel.

Remark 1 *The case when (y_t) is a jump process can be recovered setting $\beta = 0$ and $\gamma = 0$. Moreover the case when (y_t) belongs to a discrete space can be recovered by using $\mathbb{Z}$ (or $\mathbb{Z}^k$) instead of $\mathbb{R}^k$. Also the case of a general discrete subset $\mathcal{Y}$ of $\mathbb{R}^k$ can be considered, when $y_0 \in \mathcal{Y}$, and the functions K_1 and K_3 are such that for all $x \in \mathbb{R}^d$, $z \in \mathbb{N}$ and $\zeta \in \Sigma$, the condition $y \in \mathcal{Y}$ implies $y + K_i(x, y, z; \zeta) \in \mathcal{Y}$, for $i = 1, 3$.*

Analogously, the case when (y_t) is a counting process can be recovered requiring that $\beta = 0$, $\gamma = 0$, $y_0 \in \mathbb{N}$, and that $K_i(x, y, z; \zeta) \in \{0, 1\}$, for $i = 1, 3$. In this case we can take $\mathcal{Y} = \mathbb{N}$ as the state space for the $y-$component.

The first aim of this paper is to explain how a general technique can be used to handle with this kind of problem when the system is Markovian. This technique is basically the generalization of the idea of Kliemann, Koch, and Marchetti [10], together with the Filtered Martingale Problem techniques of Kurtz and Ocone [12], and Kurtz [11]. The case considered in [10] is the case when only the random times $\{S_i\}$ are observed. This technique has been used also in Fan [8] when the state/observation system is a pure jump process with finite state space.

The second aim is to show how this technique can be applied to some particular systems.

The main idea of the paper consists of rewriting the original filtering problem as a filtering problem for a new Markov system (x_t, y_t, z_t, j_t). The new system is obtained by adding a process (j_t) to the original system.

In the case of precise information the process (j_t) represents the observation of (y_t) at the observation times, i.e.

$$j_s := y_{S_i}, \text{ when } s \in [S_i, S_{i+1}). \tag{7}$$

In the grouped information case the process (j_t) is defined as

$$j_s = h, \text{ when } s \in [S_i, S_{i+1}) \text{ and } y_{S_i} \in \Gamma_h. \tag{8}$$

Remark 2 *Observe that when (y_t) takes values in a denumerable (or finite) state space and the sets Γ_h are the singletons, then there is no difference between the* precise *and the* grouped *information.*

In the precise information case the enlarged system is solution of the martingale problem for the operator $\mathbb{L} \subset \mathbf{B}(\mathbb{R}^d \times \mathbb{R}^k \times \mathbb{N} \times \mathbb{R}^k) \times \mathbf{B}(\mathbb{R}^d \times \mathbb{R}^k \times \mathbb{N} \times \mathbb{R}^k)$ of the following form

$$\mathbb{L}F(x,y,z,j) = \tag{9}$$

$$\begin{aligned}
&= B(x,y,z)\nabla_{(x,y)}F(x,y,z,j) + \frac{1}{2}tr\{\nabla^2_{(x,y)}F(x,y,z,j)C(x,y,z)C^*(x,y,z)\} \\
&+ \int_{D_0(x,y,z)} [F(x+K_0(x,y,z;\zeta),y,z,j) - F(x,y,z,j)]\,\nu(d\zeta) \\
&+ \int_{D_1(x,y,z)} [F(x+K_0(x,y,z;\zeta),y+K_1(x,y,z;\zeta),z,j) - F(x,y,z,j)]\,\nu(d\zeta) \\
&+ \int_{D_2(x,y,z)} [F(x+K_0(x,y,z;\zeta),y,z+1,y) - F(x,y,z,j)]\,\nu(d\zeta) \\
&+ \int_{D_3(x,y,z)} [F(x+K_0(x,y,z;\zeta),y+K_3(x,y,z;\zeta),z+1,y+K_3(x,y,z;\zeta)) \\
&\qquad - F(x,y,z,j)]\,\nu(d\zeta)
\end{aligned}$$

In the grouped information case, when the partition is $\{\Gamma_h\}_{h\in\mathbb{Z}}$, the enlarged system is instead solution of the martingale problem for the operator $\mathbb{A} \subset \mathbf{B}(\mathbb{R}^d \times \mathbb{R}^k \times \mathbb{N} \times \mathbb{Z}) \times \mathbf{B}(\mathbb{R}^d \times \mathbb{R}^k \times \mathbb{N} \times \mathbb{Z})$ of the following form

$$\mathbb{A}F(x,y,z,j) = \tag{10}$$

$$\begin{aligned}
&= B(x,y,z)\nabla_{(x,y)}F(x,y,z,j) + \frac{1}{2}trC(x,y,z)C^*(x,y,z)\nabla^2_{(x,y)}F(x,y,z,j) \\
&+ \int_{D_0(x,y,z)} [F(x+K_0(x,y,z;\zeta),y,z,j) - F(x,y,z,j)]\,\nu(d\zeta) \\
&+ \int_{D_1(x,y,z)} [F(x+K_0(x,y,z;\zeta),y+K_1(x,y,z;\zeta),z,j) - F(x,y,z,j)]\,\nu(d\zeta) \\
&+ \sum_h \int_{D_2(x,y,z)} \mathbb{I}_{\{y\in\Gamma_h\}}\,[F(x+K_0(x,y,z;\zeta),y,z+1,h) - F(x,y,z,j)]\,\nu(d\zeta) \\
&+ \sum_h \int_{D_3^h(x,y,z)} [F(x+K_0(x,y,z;\zeta),y+K_3(x,y,z;\zeta),z+1,h) \\
&\qquad - F(x,y,z,j)]\,\nu(d\zeta)
\end{aligned}$$

where

$$D_3^h(x,y,z) := \{\zeta \in D_3(x,y,z) \mid y + K_3(x,y,z;\zeta) \in \Gamma_h\}. \tag{11}$$

Let $\xi_t := (x_t, y_t)$ and $\eta_t := (z_t, j_t)$, then the information up to time t is modelled by the σ -algebra

$$\mathcal{F}_t^\eta = \sigma\{\eta_s,\ s \leq t\} \vee \mathcal{N},$$

where $\mathcal{N}$ is the collection of $\mathbb{P}-$null events (*in the following, for any process* (U_t)*, we denote by* $\mathcal{F}_t^U$ *the* $\sigma-$*algebra generated by the process up to time* t*, completed with the null sets*).

Our aim is the computation of

$$\pi_t\phi := \mathbb{E}[\phi(\xi_t)|\mathcal{F}_t^\eta]. \tag{12}$$

With the above representation the filtering problem fits into the framework of the filtered martingale problem ([11], [12]).

Definition 1 *Let E_1 and E_2 be measurable spaces. Let $\mathcal{A}$ be a Markov operator on the product space $(E_1 \times E_2, \mathcal{E}_1 \times \mathcal{E}_2)$. Let (μ_t, U_t) be a process with trajectories in $D_{\Pi(E_1)\times E_2}[0,T]$, such that (μ_t) is $(\mathcal{F}_{t^+}^U)$-adapted, and*

$$\mu_t F(\cdot, U_t) - \int_0^t \mu_s \mathcal{A} F(\cdot, U_s) ds,$$

is a $(\mathcal{F}_{t^+}^U)$ - martingale, for all $F \in \mathcal{D}(\mathcal{A})$. Then (μ_t, U_t) is called a solution of the ***filtered martingale problem (FMP)*** *for $\mathcal{A}$.*

For the ease of the reader, we recall the link between the FMP and the usual filtering problem. Let (X_t, Y_t) be a Markov process on the space $E_1 \times E_2$, and let (X_t, Y_t) solve the martingale problem associated to the operator $\mathcal{A}$.

Let

$$\pi_t(\cdot) = \mathbb{P}(X_t \in \cdot|\mathcal{F}_t^Y) \tag{13}$$

be the filter of X_t given $\mathcal{F}_t^Y$. Clearly

$$\mathbb{E}[\pi_0 F(\cdot, Y_0)] = \mathbb{E}[F(X_0, Y_0)], \tag{14}$$

for all $F \in \mathbf{B}(E_1 \times E_2)$, and the pair process (π_t, Y_t) is a solution of the FMP associated to $\mathcal{A}$.

As a consequence, whenever the FMP has a unique (in law) solution in $D_{\Pi(E_1)\times E_2}[0,T]$ with initial condition μ_0 satisfying

$$\mathbb{E}[\mu_0 F(\cdot, U_0)] = \mathbb{E}[F(X_0, Y_0)], \tag{15}$$

then (π_t, Y_t) has the unique law, and this allows to identify the filter once a solution of the FMP is given, as it is explained in [12]. In [12], as well as in [11], sufficient conditions for uniqueness in law of the FMP can be found, when E_i are locally compact separable metric spaces or when they are complete separable metric spaces.

Back to our problems, we can easily see that they fit into the above framework. Indeed, in the precise information case $\mathcal{A} = \mathbb{L}$, $E_1 = \mathbb{R}^d \times \mathbb{R}^k$, $E_2 = \mathbb{N} \times \mathbb{R}^k$ and $(X_t, Y_t) = (\xi_t, \eta_t)$. While in the grouped data case $\mathcal{A} = \mathbb{A}$, $E_1 = \mathbb{R}^d \times \mathbb{R}^k$, $E_2 = \mathbb{N} \times \mathbb{Z}$, $(X_t, Y_t) = (\xi_t, \eta_t)$.

In Section 2 we discuss the properties of the original system (x_t, y_t, z_t) and of the enlarged system $(x_t, y_t, z_t, j_t) = (\xi_t, \eta_t)$. Section 3 is devoted to the filtering equation for the system (ξ_t, η_t). As in [10] and [8], the filtering equation is given also in a recursive form (see Theorem 2): between the observed jump times it is a deterministic non-linear integro-differential equation. The equation is parameterized by the number z_t of observations and the current value j_t of the observation. Its initial data is computed by means of an updating formula. In Section 4 we study the deterministic version of the filtering equation in the recursive form, and again following the same ideas of [10], we give a solution (see Theorem 3). In Section 5 we explain how a change of measure could help in proving that this solution is the filter (see Theorem 4 and Theorem 5); here we use the approach of the filtered martingale problem of [12]. Section 6 and Section 7 are devoted to general examples: the case when the partition is finite (see Theorem 6) and the case when (y_t) is a counting process (see Theorem 7), respectively. In both cases we assume that the jump intensities relative to (y_t, z_t) are uniformly bounded. The case when (y_t) has finite state space, can be considered also (see Remark 18). In these Sections we introduce two particular examples that are used as reference models. The first one concerns the explicit filter of a discrete state Markov system (see Lemma 3 and Remark 16) and the second one concerns the explicit filter of a Poisson process $(\widetilde{u}_t)$ observed at the jump times of another Poisson process $(\widetilde{z}_t)$, when

$$\widetilde{u}_t = \widetilde{\jmath}_0 + N_t^{(1)} + N_t^{(3)}$$

$$\widetilde{z}_t = N_t^{(2)} + N_t^{(3)}$$

where $((N_t^{(i)}), i \in \{1, 2, 3\}$ are independent Poisson processes (see Lemma 4 and Remark 19).

The last Section is devoted to the queue example, and an explicit version of the filter is given, for the case $h = 0$.

Notations *If E is a measurable space, then $\mathbf{B}(E)$ is the space of measurable bounded function on it. If E is a metric space, then $\mathcal{B}(E)$ are the Borel sets. We denote the expectation under any probability measure by the same symbol, $\mathbb{E}$. Given an operator $\mathcal{A}$ on $\mathcal{D}(\mathcal{A}) \subseteq \mathbf{B}(E_1 \times E_2)$, for all $e_2 \in E_2$ we define the operator $\mathcal{A}^{e_2}$ on $\{\phi \in \mathbf{B}(E_1)$ s.t. $F_\phi \in \mathcal{D}(\mathcal{A})$, where $F_\phi(e_1, e_2) := \phi(e_1)\ \forall e_1\, e_2\}$*

$$\mathcal{A}^{e_2}\phi(e_1) = \mathcal{A}F_\phi(e_1, e_2). \tag{16}$$

Usually (x_t) denotes a process or a deterministic trajectory of it, but in some cases in order to avoid misunderstanding we use $(x.)$.

For the ease of the reader we end the introduction by stating the sufficient conditions proved in [12]. First of all we need to recall some basic definitions.

Definition 2 *Let $\mathcal{A} \subset \boldsymbol{B}(E) \times \boldsymbol{B}(E)$. The* ***martingale problem*** *for $\mathcal{A}$ is* ***well posed*** *if and only if*

(i) *For any initial distribution ν* ***there exists a solution*** *for the martingale problem for $(\mathcal{A}, \nu)$, i.e. there exists a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ and there exists a process (Z_t), such that*

- *(Z_t) is measurable,*
- *Z_0 has distribution ν,*
- *for any $(f, g) \in \mathcal{A}$, the process $f(Z_t) - \int_0^t g(Z_s)ds$ is an $(\mathcal{F}_t)-$martingale*

(ii) *For any initial distribution ν there is* ***uniqueness*** *for the martingale problem for $(\mathcal{A}, \nu)$, that is if $(Z_t^{(1)})$ and $(Z_t^{(2)})$ are two solutions of the martingale problem for $(\mathcal{A}, \nu)$ or equivalently $(Z_t^{(k)})$ satisfy the conditions in* ***(i)****, then the finite dimensional distributions of $(Z_t^{(k)})$ coincide.*

Definition 3 *Let $\mathcal{A} \subset \boldsymbol{B}(E) \times \boldsymbol{B}(E)$. The* ***martingale problem*** *for $\mathcal{A}$ is* ***well posed in*** *$D_E([0, \infty))$ if and only if*

(i)' *For any initial distribution ν* ***there exists a solution*** *for the martingale problem for $(\mathcal{A}, \nu)$ in $D_E([0, \infty))$, i.e.*

- *Condition* ***(i)*** *of Definition 2 holds.*
- *For any initial distribution ν and for any solution (Z_t) of the martingale problem for $(\mathcal{A}, \nu)$, there is a modification with sample paths in $D_E([0, \infty))$.*

(ii)' *For any initial distribution ν there is* ***uniqueness*** *for the martingale problem for $(\mathcal{A}, \nu)$ in $D_E([0, \infty))$, i.e. if $(Z_t^{(1)})$ and $(Z_t^{(2)})$ are two solutions of the martingale problem for $(\mathcal{A}, \nu)$, with sample paths in $D_E([0, \infty))$, then the finite dimensional distributions of $(Z_t^{(k)})$ coincide (or equivalently the distributions of $(Z_t^{(k)})$ in $D_E([0, \infty))$ coincide).*

Remark 3 *There are examples of operators $\mathcal{A}$ such that uniqueness in $D_E([0, \infty))$ holds for martingale problem, and contemporary uniqueness (in the sense of Definition 2) fails (see Ethier and Kurtz [7] for examples of this kind, in particular Problems 21 and 22 in Chapter 4).*

Proposition 1 *Let E_1 and E_2 be locally compact, separable metric spaces, and let $E = E_1 \times E_2$. Let $\mathcal{A} \subset \widehat{C}(E) \times \widehat{C}(E)$, where $\widehat{C}(E)$ is the space of continuous function vanishing at infinity. Assume that*

$$\mathcal{D}(\mathcal{A}) = \{f \ : \ (f, g) \in \mathcal{A}\}$$

is an algebra and is dense in $\widehat{C}(E)$.

Assume that the martingale problem for $\mathcal{A}$ is well posed in $D_E([0, \infty))$.

Let (X_t, Y_t) be a solution of the martingale problem for $\mathcal{A}$, let (π_t) be the filter defined in (13). Let (μ_t, U_t) be a solution of the filtered martingale problem for $\mathcal{A}$ in a probability space $(\widehat{\Omega}, \widehat{\mathcal{F}}, \widehat{P})$.

Assume that the initial conditions (X_0, Y_0) and (μ_0, U_0) satisfy the analogous of (15), i.e.

$$\mathbb{E}^{\widehat{P}}[\mu_0 F(\cdot, U_0)] = \mathbb{E}[F(X_0, Y_0)]. \tag{17}$$

Then

- *For any $t > 0$ there exists a function H_t*

$$H_t : D_{E_2}([0, \infty)) \to \Pi(E_1)$$

satisfying

$$\mu_t = H_t(U(\cdot)) = H_t(U(\cdot \wedge s)) \qquad \widehat{P} - a.s. \tag{18}$$

for every $s > t$;

$$\pi_t = H_t(Y(\cdot)) \qquad \mathbb{P} - a.s. \tag{19}$$

for all t such that $\pi_{t^-} = \pi_t$, $\mathbb{P} - a.s.$, and $\mu_{t^-} = \mu_t$, $\widehat{P} - a.s.$.

- *The processes (μ_t, U_t) and (π_t, Y_t) have the same finite dimensional distributions, i.e. uniqueness in law for the FMP holds.*

Remark 4 *Note that uniqueness in law for the FMP and property (19) are linked by means of property (18), since (π_t) has cadlag sample paths, and therefore the set of times t such that $\mathbb{P}\{\pi_{t^-} \neq \pi_t\} > 0$ is at most denumerable. The same happens with (μ_t) and the set of times t such that $\widehat{P}\{\mu_{t^-} \neq \mu_t\} > 0$.*

Proposition 2 *Let E_1 and E_2 be complete separable metric spaces, and let $E = E_1 \times E_2$. Let $\mathcal{A} \subset \boldsymbol{B}(E) \times \boldsymbol{B}(E)$, where $\boldsymbol{B}(E)$ is the space of bounded measurable functions on E. Assume that the range of $\alpha - \mathcal{A}$*

$$\mathcal{R}(\alpha - \mathcal{A}) = \{\alpha f - g \ : \ (f, g) \in \mathcal{A}\}$$

is bp$-$dense in $\boldsymbol{B}(E)$, i.e. is dense with respect to the topology of bounded-pointwise convergence.

Assume furthermore that the martingale problem for $\mathcal{A}$ is well posed in $D_E([0, \infty))$.

Then the conclusions of Proposition 1 concerning H_t and the uniqueness in law of the FMP for $\mathcal{A}$ hold.

Remark 5 *If moreover $\mathcal{A}$ is linear and dissipative, then the condition on $\mathcal{R}(\alpha - \mathcal{A})$ implies uniqueness for the martingale problem for $\mathcal{A}$ in $D_E([0, \infty))$ (see [7] Corollary 4.4 Chapter 4).*

2 The partially observable system with grouped data

In this Section we discuss in more detail the original stochastic system $(x_t, y_t, z_t) \in \mathbb{R}^d \times \mathbb{R}^k \times \mathbb{N}$, and the enlarged one $(x_t, y_t, z_t, j_t) \in \mathbb{R}^d \times \mathbb{R}^k \times \mathbb{N} \times \mathbb{Z}$, in the grouped data case.

First of all we assume that the system (x_t, y_t, z_t) is realized in a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and that

$$
\begin{aligned}
x_t &= x_0 + \int_0^t b(\xi_s, z_s)\, d\, s + \int_0^t c(\xi_s, z_s) dW_s, \\
&\quad + \int_0^t \int_\Sigma K_0(\xi_{s^-}, z_{s^-}; \zeta) \mathcal{N}(ds, d\zeta) && (20)\\
y_t &= y_0 + \int_0^t \beta(\xi_s, z_s)\, d\, s + \int_0^t \gamma(\xi_s, z_s) dW_s, \\
&\quad + \int_0^t \int_{D_1(\xi_{s^-}, z_{s^-})} K_1(\xi_{s^-}, z_{s^-}; \zeta) \mathcal{N}(ds, d\zeta) \\
&\quad + \int_0^t \int_{D_3(\xi_{s^-}, z_{s^-})} K_3(\xi_{s^-}, z_{s^-}; \zeta) \mathcal{N}(ds, d\zeta), && (21)\\
z_t &= \int_0^t \int_{D_2(\xi_{s^-}, z_{s^-})} \mathcal{N}(ds, d\zeta) + \int_0^t \int_{D_3(\xi_{s^-}, z_{s^-})} \mathcal{N}(ds, d\zeta). && (22)
\end{aligned}
$$

Here

- (ξ_t) is a short notation for the couple (x_t, y_t),
- (W_t) is a standard $m-$dimensional Brownian motion,
- $\mathcal{N}(ds, d\zeta)$ is a Poisson measure on $\mathbb{R} \times \Sigma$ with mean measure $ds \times \nu(d\zeta)$,
- (x_0, y_0) is a random variable with values in $\mathbb{R}^d \times \mathbb{R}^k$ and probability distribution ν_0
- We assume that (x_0, y_0), (W_t) and $\mathcal{N}(ds, d\zeta)$ are independent of each other.

We recall the notation $K_2(x, y, z; \zeta) := \mathbb{I}_{D_2(x,y,z)}(\zeta)$, and the assumptions that $D_l(x, y, z) = \{\zeta \in \Sigma : K_l(x, y, z; \zeta) \neq 0\}$ for $l = 1, 2, 3$, that $D_0(x, y, z) \subset \{\zeta \in \Sigma : K_0(x, y, z; \zeta) \neq 0\} \subset \cup_{l=0,1,2,3} D_l(x, y, z)$, and that $D_l(x, y, z)$, for $l = 0, 1, 2, 3$, are pairwise disjoint measurable sets in Σ.

Under suitable hypotheses for the functions $B = (b, \beta)$, $C = (c, \gamma)$ and K_l (we discuss some of them at the end of this Section), the above system has a unique solution and (x_t, y_t, z_t) is a Markov process with formal generator $\mathcal{L}$.

As explained in the Introduction we add to the above system a new component (j_t) (see definition (8)). In this framework we can write

$$j_t = j_0 + \sum_h \int_0^t \int_{D_2(\xi_{s^-},z_{s^-}) \cup D_3(\xi_{s^-},z_{s^-})} \mathbb{I}_{\{y_s \in \Gamma_h\}} (h - j_{s^-}) \mathcal{N}(ds, d\zeta)$$

with

$$j_0 = \sum_h h \mathbb{I}_{\{y_0 \in \Gamma_h\}} \tag{23}$$

Taking into account the different behaviour of the $y-$component in D_2 and D_3, we can rewrite

$$\begin{aligned} j_t &= j_0 + \sum_h \int_0^t \int_{D_2(\xi_{s^-},z_{s^-})} \mathbb{I}_{\{y_{s^-} \in \Gamma_h\}} (h - j_{s^-}) \mathcal{N}(ds, d\zeta) \\ &+ \sum_h \int_0^t \int_{D_3(\xi_{s^-},z_{s^-})} \mathbb{I}_{\{y_{s^-} + K_3(\xi_{s^-},z_{s^-};\zeta) \in \Gamma_h\}} (h - j_{s^-}) \mathcal{N}(ds, d\zeta) \end{aligned}$$

or equivalently, using the definition (11) of $D_3^h(\xi, z)$,

$$\begin{aligned} j_t &= j_0 + \sum_h \int_0^t \int_{D_2(\xi_{s^-},z_{s^-})} \mathbb{I}_{\{y_{s^-} \in \Gamma_h\}} (h - j_{s^-}) \mathcal{N}(ds, d\zeta) \\ &\quad + \sum_h \int_0^t \int_{D_3^h(\xi_{s^-},z_{s^-})} (h - j_{s^-}) \mathcal{N}(ds, d\zeta) \end{aligned} \tag{24}$$

In order to derive the equation satisfied by the filter (12) we will use the marked point process p determined by the sequence

$$(S_i, j_{S_i})_{i \geq 1}. \tag{25}$$

Indeed the observation of the jump process (η_t) can be interpreted as the observation of the marked point process $p(dt \times dh)$, whose times are the jump times of (z_t), and whose marks are the values of the process (j_t) at these times. Formally this means that

$$\mathcal{F}_t^\eta = \mathcal{F}_t^p \vee \sigma\{j_0\}, \tag{26}$$

where $p(dt \times dh)$ is the counting random measure on $\mathbb{R}^+ \times \mathbb{Z}$, such that, for all $h \in \mathbb{Z}$,

$$p((0, t] \times \{h\}) = v_t^h, \tag{27}$$

with (v_t^h) the counting process defined by

$$v_t^h := \int_0^t \mathbb{I}_{\{j_s=h\}} dz_s = \int_0^t \mathbb{I}_{\{y_s \in \Gamma_h\}} dz_s. \tag{28}$$

The processes (v_t^h) count the jump times S_i of (z_t) when $y_{S_i} \in \Gamma_h$, that is when $j_{S_i} = h$, and

$$z_t = \sum_h v_t^h \tag{29}$$

$$j_t = j_0 + \sum_h \int_0^t (h - j_{s^-})\, dv_s^h. \tag{30}$$

The above observations show that $\mathcal{F}_t^\eta \subseteq \mathcal{F}_t^p \vee \sigma\{j_0\}$. The inverse inclusion is trivial and (26) follows. As a consequence, our filtering problem (12) can be viewed as the computation of

$$\pi_t \phi = \mathbb{E}\left[\, \phi(\xi_t) \right| \mathcal{F}_t^p \vee \sigma\{j_0\}] \tag{31}$$

for all $\phi \in \mathbf{B}(\mathbb{R}^d \times \mathbb{R}^k)$.

To this end we need the local characteristics of $p(dt, dh)$. They can be computed in terms of the intensities $(M_t^{(h)})$ of the processes (v_t^h). Since

$$v_t^h = \int_0^t \mathbb{I}_{\{y_s \in \Gamma_h\}} \mathbb{I}_{\{\Delta y_s = 0\}} dz_s + \int_0^t \mathbb{I}_{\{y_s \in \Gamma_h\}} \mathbb{I}_{\{\Delta y_s \neq 0\}} dz_s$$

we can also rewrite

$$\begin{aligned} v_t^h &= \int_0^t \mathbb{I}_{\{y_{s^-} \in \Gamma_h\}} \int_{D_2(\xi_{s^-}, z_{s^-})} \mathcal{N}(ds, d\zeta) \\ &+ \int_0^t \int_{D_3(\xi_{s^-}, z_{s^-})} \mathbb{I}_{\{y_{s^-} + K_3(\xi_{s^-}, z_{s^-}; \zeta) \in \Gamma_h\}} \mathcal{N}(ds, d\zeta) \\ &= \int_0^t \mathbb{I}_{\{y_{s^-} \in \Gamma_h\}} \int_{D_2(\xi_{s^-}, z_{s^-})} \mathcal{N}(ds, d\zeta) + \int_0^t \int_{D_3^h(\xi_{s^-}, z_{s^-})} \mathcal{N}(ds, d\zeta) \end{aligned}$$

where we have used the definition (11) of $D_3^h(\xi, z)$.

Using the last equality we can easily verify that the $(\mathcal{F}_t^{\xi,\eta})$ - intensity of (v_t^h) is

$$M_t^{(h)} = \mathbb{I}_{\{y_{t^-} \in \Gamma_h\}} \nu(D_2(\xi_{t^-}, z_{t^-})) + \nu(D_3^h(\xi_{t^-}, z_{t^-})). \tag{32}$$

It follows that the local characteristics $(\lambda_t, \Phi_t(\cdot))$ of the marked point process p

with respect to $(\mathcal{F}_t^{\xi,\eta})$ are

$$\lambda_t = \nu(D_2(\xi_{t^-}, z_{t^-})) + \nu(D_3(\xi_{t^-}, z_{t^-}))$$

$$\Phi_t(h) = \frac{M_t^{(h)}}{\nu(D_2(\xi_{t^-}, z_{t^-})) + \nu(D_3(\xi_{t^-}, z_{t^-})}.$$

Note that (λ_t) coincides with the $(\mathcal{F}_t^{x,y,z})$ - intensity of the process (z_t), as it was expected.

We end this Section discussing some sets of conditions. These conditions imply that the original system (20), (21), (22) has a unique solution and that the generator satisfies the hypotheses of Proposition 1 or Proposition 2. For the original system we consider the same kind of sufficient conditions used in [10]. The interesting fact is that these conditions are also sufficient conditions for the enlarged system (20), (21), (22), (24), as we are going to discuss. We will try to explain that they are sufficient in order to get uniqueness for the FMP for the enlarged system.

For the original system, as in [10], we use the results of Athreya, Kliemann and Koch [2] and consider the following sufficient conditions, which are divided in conditions on the drift and diffusion coefficients and on the jump coefficients.

FIRST SET OF CONDITIONS

Growth and Lipschitz conditions on the drift and diffusion coefficients

$$|B(\xi,z)|^2 + |C(\xi,z)C^*(\xi,z)|^2 \leq L\,(1 + |(\xi,z)|^2) \tag{33}$$

$$|B(\xi,z) - B(\xi',z')|^2 \leq L_B\,|(\xi,z) - (\xi',z')| \tag{34}$$

$$|C(\xi,z)C^*(\xi,z) - C(\xi',z')C^*(\xi',z')|^2 \leq L_C\,|(\xi,z) - (\xi',z')| \tag{35}$$

with L, L_B and L_C finite constant, for any $\xi = (x,y)$ and $\xi' = (x',y')$ in $\mathbb{R}^d \times \mathbb{R}^k$ and for any z and z' in $\mathbb{N}$.

Square integrable initial condition

$$\mathbb{E}[|(x_0,y_0)|^2] < \infty. \tag{36}$$

We are also assuming that $z(0) = 0$.

Regularity conditions on the diffusion part

For any $\xi = (x, y)$ in $\mathbb{R}^d \times \mathbb{R}^k$ and any z in $\mathbb{N}$, there exists a weak unique Feller solution $(\widetilde{\xi}_u, \widetilde{z}_u)$ for the diffusion part, that is for the equations

$$\widetilde{\xi}_t = \xi + \int_0^t B(\widetilde{\xi}_s, \widetilde{z}_s)ds + \int_0^t C(\widetilde{\xi}_s, \widetilde{z}_s)d\widetilde{W}_s \tag{37}$$

$$\widetilde{z}_t = z \tag{38}$$

satisfying

$$\sup_{\xi,z} \mathbb{E}[sup_{u\in[0,T]}|(\widetilde{\xi}_u, \widetilde{z}_u) - (\xi, z)|^4] < \infty. \tag{39}$$

Then the process $(\widetilde{\xi}_u, \widetilde{z}_u)$ has generator $\widetilde{\mathcal{L}}$, with domain $\mathcal{D}(\widetilde{\mathcal{L}})$ which is separating and $bp-$dense in $C_b(\mathbb{R}^d \times \mathbb{R}^k \times \mathbb{N})$.

Sublinear growth conditions on the intensities of jumps times

$$\lambda_l(x, y, z) := \nu(D_l(x, y, z)) \leq \overline{\lambda}\,(1 + |(x, y, z)|) \quad l = 0, 1, 2, 3. \tag{40}$$

Bounded jump size condition

$$\sup_{x,y,z,\zeta} |K_l(x, y, z; \zeta)| \leq \overline{K} < \infty \quad l = 0, 1, 2, 3, \tag{41}$$

where we recall that $K_2(x, y, z; \zeta) = \mathbb{I}_{D_2(x,y,z)}(\zeta)$

Remark 6 *Note that if the previous conditions (33), (34) and (35) are satisfied, then there is also strong existence and uniqueness for the system (37), (38). By an easy application of Ito's formula one can also show that $\widetilde{\xi}_u$ is square integrable and that for any function ϕ in $C^2_K(\mathbb{R}^d \times \mathbb{R}^k)$, the space of C^2-functions with compact support, and for any ψ in $\boldsymbol{B}(\mathbb{N}) = C_b(\mathbb{N})$, the function $f(\xi, z) = \phi(\xi)\psi(z)$ is in the domain of the bp−generator $\widetilde{\mathcal{L}}^{bp}$. The same holds for linear combinations of such functions. When the functions B and C are uniformly bounded, the function ϕ can be taken in the space $C^2_b(\mathbb{R}^d \times \mathbb{R}^k)$, moreover then (39) is satisfied.*

Remark 7 *Instead of considering system (37) and (38) one can consider the family of processes $(\widetilde{\xi}^{\xi,z}_t)$ such that*

$$\widetilde{\xi}^{\xi,z}_t = \xi + \int_0^t B(\widetilde{\xi}^{\xi,z}_s, z)ds + \int_0^t C(\widetilde{\xi}^{\xi,z}_s, z)d\widetilde{W}_s, \tag{42}$$

and (39) can be replaced by

$$\sup_{\xi,z} \mathbb{E}[sup_{u\in[0,T]}|\widetilde{\xi}^{\xi,z}_u - \xi|^4] < \infty. \tag{43}$$

When the jump intensities are bounded, then the conditions are simpler.

SECOND SET OF CONDITIONS

Growth and Lipschitz conditions on the drift and diffusion coefficients
(as above)

Square integrable initial condition
(as above)

Bounded intensities of jump times

$$sup_{x,y,z}\nu(D_l(x,y,z)) \leq \overline{\lambda}_l < \infty \quad l = 0,1,2,3 \tag{44}$$

Note that when the above condition (44) holds, then the regularity condition (39), or equivalently condition (43), is not necessary.

In [2] and [10] the above sets of conditions allow to construct a *cadlag* process $(\xi_t, z_t)_{t\in[0,T]}$ satisfying the original system. The trajectories of (ξ_t, z_t) have a finite number of discontinuity times in the interval $[0, T]$ and this process evolves like $(\widetilde{\xi}_u, \widetilde{z}_u)$ in between the discontinuity times.

As in [10], using the results of [2] and Çinlar, Jacod, Protter and Sharpe [6], one can show that the above conditions imply existence and uniqueness of the solution (x_t, y_t, z_t) of the original system. The process (x_t, y_t, z_t) is a time homogeneous Markov process with formal generator $\mathcal{L}$. Moreover $\mathcal{D}(\mathcal{L}) \supset \mathcal{D}(\widetilde{\mathcal{L}})$ and the domain of the resolvent $(\alpha I - \mathcal{L})^{-1}$ is separating and bounded pointwise dense in $\mathbf{B}(\mathbb{R}^d \times \mathbb{R}^k \times \mathbb{N})$ for each $\alpha > 0$.

In order to see that the generator of the enlarged system (20), (21), (22), (24) also has the same kind of properties one cannot use directly the fact that the enlarged system satisfies the same kind of conditions.

The conditions on the diffusive part of the system do not give any problem. Indeed the conditions (33), (34), (35), (39) on the system (37), (38) for the diffusive part of the original system and and the corresponding conditions for the diffusive part of the system enlarged with equation (24) are the same, since in either cases they reduce to the regularity condition (43) on the family of diffusions (42).

Also if we assume the bounded intensity condition (44) on the original system, then the same condition holds for the enlarged system.

On the contrary, if we assume the sublinear growth conditions on the intensities and the bounded jump size conditions for the original system, i.e. (40) and (41) respectively, then a problem arises. Although the corresponding sublinear growth conditions are clearly satisfied by the enlarged system, unfortunately the corresponding bounded jump size condition are not (in general) satisfied. Indeed the jump size of the $j-$component is not *a priori* bounded, since it depends on where the $y-$component is at the observed jump time.

Nevertheless one can easily handle with this problem when the partition $\{\Gamma_h\}$ is finite, since in this case we can use a finite set as state space for the $j-$component, instead of $\mathbb{Z}$, and the jump size of the $j-$component is necessarily bounded.

Also the general case can be recovered, but in this case one has to go through the details of the proof given in [2]. Indeed the same technique of [2] and [10] can be applied in order to construct the enlarged process (x_t, y_t, z_t, j_t), using the same construction and taking into account that the discontinuity times of (x_t, y_t, z_t, j_t) coincide with those of (x_t, y_t, z_t).

Then the other properties of the process (x_t, y_t, z_t, j_t) and of the formal generator $\mathbb{A}$ follow as in [10]. In particular for any function ϕ in $C_b^2(\mathbb{R}^d \times \mathbb{R}^k)$, the function $F(\xi, \eta) = \phi(\xi)$ is in the domain of the bp-generator $\mathbb{A}^{bp}$. Moreover the range of $(\alpha I - \mathbb{A}^{bp})$ is separating and bounded pointwise dense in $\mathbf{B}(\mathbb{R}^d \times \mathbb{R}^k \times \mathbb{N} \times \mathbb{Z})$ for each $\alpha > 0$.

3 The filter equation

In this section we derive (Theorem 2) the Kushner-Stratonovich equation for the filter (π_t) introduced in (12) or equivalently in (31). To this end we derive first the equation satisfied by

$$\pi_t F(\cdot, \eta_t) = \int F(\xi, \eta_t)\pi_t(d\xi) \tag{45}$$

for $F \in \mathbf{B}(\mathbb{R}^d \times \mathbb{R}^k \times \mathbb{N} \times \mathbb{Z})$.

In the rest of the paper we use the following notations

$$\lambda_l(\xi, z) := \nu(D_l(\xi, z)), \quad l = 0, \dots, 3 \qquad \lambda_3^h(\xi, z) := \nu(D_3^h(\xi, z)), \tag{46}$$

$$m^{(h)}(\xi, z) := \mathbb{I}_{\{y\in\Gamma_h\}}\lambda_2(\xi, z) + \lambda_3^h(\xi, z). \tag{47}$$

Then, recalling (32), the intensity of (v_t^h) can be expressed as

$$M_t^{(h)} = m^{(h)}(\xi_{t^-}, z_{t^-}).$$

Moreover we define for $F \in \mathbf{B}(\mathbb{R}^d \times \mathbb{R}^k \times \mathbb{N} \times \mathbb{Z})$

$$\mathcal{R}_2^{(h)} F(\xi, \eta) := \mathbb{I}_{\{y\in\Gamma_h\}} \int_{D_2(\xi,z)} [F(x+K_0(\xi, z; \zeta), y, z+1, h) - F(\xi, \eta)]\nu(d\zeta) \tag{48}$$

$$\mathcal{R}_3^{(h)} F(\xi, \eta) := \int_{D_3^{(h)}(\xi,z)} [F(x + K_0(\xi, z; \zeta), y, z + 1, h) - F(\xi, \eta)]\nu(d\zeta). \tag{49}$$

The interest of the following theorem relies in deriving the form of the filter equation. For the technique used here, one could even only guess its form, use this form in order to get a solution of the FMP and then use it in order to derive the filter, as in [12].

Theorem 1 *Assume that one of the sets of conditions discussed at the end of Section 2 are satisfied. Let $(\xi_t, \eta_t) = (x_t, y_t, z_t, j_t)$ be the solution of the enlarged system (20), (21), (22), (24), with initial conditions $\xi_0 = (x_0, y_0)$ and $\eta_0 = (0, j_0)$. Assume that (x_0, y_0) has distribution ν_0 and satisfies (36), and that $j_0 = \sum_h h\mathbb{I}_{\{y_0\in\Gamma_h\}}$.*

Then the filter (π_t) defined in (12) satisfies $\mathbb{P}$-a.s.

$$\begin{aligned}
\pi_t F(\cdot, \eta_t) \;&=\; \pi_0 F(\cdot, \eta_0) + \int_0^t \pi_s(\mathbb{A}F(\cdot, \eta_s))ds \\
&+ \sum_h \int_0^t [\pi_{s^-} m^{(h)}(\cdot, z_{s^-})]^+ \\
&\quad \Big[\pi_{s^-}(F m^{(h)})(\cdot, \eta_{s^-}) - \pi_{s^-} F(\cdot, \eta_{s^-})\pi_{s^-} m^{(h)}(\cdot, z_{s^-}) \\
&\qquad + \pi_{s^-}(\mathcal{R}_2^{(h)} + \mathcal{R}_3^{(h)})F(\cdot, \eta_{s^-})\Big] \left(dv_s^{(h)} - \pi_{s^-} m^{(h)}(\cdot, z_{s^-})ds\right),
\end{aligned} \tag{50}$$

for all $t \geq 0$, for all $F \in \mathcal{D}(\mathbb{A})$.

Here $m^{(h)}(\xi, z)$ are the function introduced in (47), and $[a]^+ = \frac{1}{a}$ when $a \neq 0$ while $[a]^+ = 0$ otherwise.

Moreover, $\mathbb{P} - a.s.$

$$\begin{aligned} \pi_0(dx, dy)(\omega) &= \sum_h{}' \mathbb{I}_{\{j_0(\omega)=h\}} \frac{\mathbb{I}_{\{\mathbb{R}^d \times \Gamma_h\}}(x,y)\nu_0(dx,dy)}{\nu_0(\mathbb{R}^d \times \Gamma_h)} \\ &+ \sum_h{}'' \mathbb{I}_{\{j_0(\omega)=h\}} \delta_{\{(0,0)\}}(dx, dy) \end{aligned} \tag{51}$$

where the first sum is restricted to the h such that $\nu_0(\mathbb{R}^d \times \Gamma_h) > 0$, and the second sum is restricted to the h such that $\nu_0(\mathbb{R}^d \times \Gamma_h) = 0$.

Proof The filter equation can be deduced using a version of the innovation method. In particular with standard method of stochastic calculus one can derive the innovation gain imposing that

$$\mathbb{E}\left[\{\pi_t(F(\cdot,\eta_t)) - \pi_0(F(\cdot,\eta_0))\}\, U_t\right] = \mathbb{E}\left[\{F(\xi_t,\eta_t) - F(\xi_0,\eta_0)\}\, U_t\right]$$

for all $F \in \mathcal{D}(\mathbb{A})$ and for all process (U_t) of the form

$$U_t = \int_0^t \mathcal{K}_s^{(h)} \left(dv_s^{(h)} - \pi_{s^-} m^{(h)}(\cdot, z_{s^-}) ds \right),$$

with $(\mathcal{K}_t^{(h)})$ an $\mathcal{F}_t^p \vee \sigma\{j_0\}$-predictable process.
Equality (51) is trivial, once we use (23) and observe that π_0 in (51) is a version of the conditional law of ξ_0 w.r.t. the σ-algebra generated by j_0. Clearly, instead of the measure concentrated in $\xi = (0,0)$, any other probability measure could be used.

Remark 8 *We note that $\mathcal{R}_2^{(h)} + \mathcal{R}_3^{(h)}$ represents the operator which drives the system (ξ_t, η_t) in the jump times of the process p with mark value h, i.e. in the jump times of $(v_t^{(h)})$.*

In the sequel we will need also the following observations.
Let $F \in \mathbf{B}(\mathbb{R}^d \times \mathbb{R}^k \times \mathbb{N} \times \mathbb{Z})$. Then for any $(\xi, \eta) \in \mathbb{R}^d \times \mathbb{R}^k \times \mathbb{N} \times \mathbb{Z}$, with $\eta = (z, j)$

$$\begin{aligned} \left| (\mathcal{R}_2^{(h)} + \mathcal{R}_3^{(h)}) F(\xi,\eta) \right| &\leq 2\, \|F\|_\infty \left(\mathbb{I}_{\{y \in \Gamma_h\}} \nu(D_2(\xi,z)) + \nu(D_3^{(h)}(\xi,z)) \right) \\ &= 2\, \|F\|_\infty\, m^{(h)}(\xi, z). \end{aligned}$$

Then for any $\eta \in \mathbb{N} \times \mathbb{Z}$ and for any probability measure π on $E = \mathbb{R}^d \times \mathbb{R}^k$

$$\left| \pi\, F(\cdot,\eta) m^{(h)}(\cdot,z) - \pi\, F(\cdot,\eta)\, \pi\, m^{(h)}(\cdot,z) + \pi\, \left(\mathcal{R}_2^{(h)} + \mathcal{R}_3^{(h)} \right) F(\cdot,\eta) \right| \leq$$

$$\leq 4 \|F\|_\infty \, \pi \, m^{(h)}(\cdot, z)$$

Therefore, when $\pi \, m^{(h)}(\cdot, z) = 0$*, the right hand side of the above inequality also is zero. As a consequence*

$$[\pi \, m^{(h)}(\cdot, z)]^+ \, \pi \, m^{(h)}(\cdot, z) \cdot$$
$$\cdot \left[\pi \, F(\cdot, \eta) m^{(h)}(\cdot, z) - \pi \, F(\cdot, \eta) \, \pi \, m^{(h)}(\cdot, z) + \pi \, \left(\mathcal{R}_2^{(h)} + \mathcal{R}_3^{(h)} \right) F(\cdot, \eta) \right]$$

is equal to

$$\pi \, F(\cdot, \eta) m^{(h)}(\cdot, z) - \pi \, F(\cdot, \eta) \, \pi \, m^{(h)}(\cdot, z) + \pi \, \left(\mathcal{R}_2^{(h)} + \mathcal{R}_3^{(h)} \right) F(\cdot, \eta),$$

and

$$[\pi \, m^{(h)}(\cdot, z)]^+ \, \left[\pi \, F(\cdot, \eta) m^{(h)}(\cdot, z) - \pi \, F(\cdot, \eta) \, \pi \, m^{(h)}(\cdot, z) \right.$$
$$\left. + \, \pi \, \left(\mathcal{R}_2^{(h)} + \mathcal{R}_3^{(h)} \right) F(\cdot, \eta) \right] \, \leq \, 4 \|F\|_\infty$$

An easy corollary of the previous theorem involves the filter equation for functions ϕ of the variable $\xi = (x, y)$.

Theorem 2 *Assume that one of the sets of conditions discussed at the end of Section 2 are satisfied. Let* i *be in* $\mathbb{N}$ *and let* G^i *be the operator*

$$\begin{aligned} G^i \phi(\xi) &= B(\xi, i) \nabla_\xi \phi(\xi) + \frac{1}{2} tr\{C(\xi, i) C^*(\xi, i) \nabla_x^2 \phi(\xi)\} \\ &+ \int_{D_0(\xi, i)} [\phi(x + K_0(\xi, i; \zeta), y) - \phi(\xi)] \, \nu(d\zeta) \\ &+ \int_{D_1(\xi, i)} [\phi(x + K_0(\xi, i; \zeta), y + K_1(\xi, i; \zeta)) - \phi(\xi)] \, \nu(d\zeta) \end{aligned} \tag{52}$$

on

$$\mathcal{D} = \{\phi \in \mathbf{B}(\mathbb{R}^d \times \mathbb{R}^k) \ s.t. \ F_\phi \in \mathcal{D}(\mathbb{A}), \ where \ F_\phi(\xi, \eta) := \phi(\xi) \, \forall \, \xi, \eta\}. \tag{53}$$

Let i *be in* $\mathbb{N}$*, let* h *be in* $\mathbb{Z}$ *and let* $Q^{i,h}$ *be the operator defined by*

$$\begin{aligned} Q^{i,h} \phi(\xi) &:= \mathbb{I}_{\{y \in \Gamma_h\}} \int_{D_2(\xi, i)} \phi(x + K_0(\xi, i; \zeta), y) \nu(d\zeta) \\ &\quad + \int_{D_3^{(h)}(\xi, i)} \phi(x + K_0(\xi, i; \zeta), y + K_3(\xi, i; \zeta)) \nu(d\zeta). \end{aligned} \tag{54}$$

Let $j_{S_i} = h_i$ *at the observation times* $t = S_i$*. Then the filter* (π_t) *satisfies the following recursive system* $\mathbb{P} - a.s.$*, for all* $\phi \in \mathcal{D}$*,*

$$\pi_0\phi = \pi_{S_0}\phi = \frac{\nu_0(\mathbb{I}_{\{\mathbb{R}^d\times\Gamma_{h_0}\}}\phi)}{\nu_0(\mathbb{R}^d\times\Gamma_{h_0})}, \tag{55}$$

i. e. π_0 *satisfies (51),*

$$\begin{aligned}\pi_t\phi &= \pi_{S_i}\phi + \int_{S_i}^t \pi_s G^i\phi(\cdot)ds \\ &\quad - \int_{S_i}^t [\pi_s(\phi(\lambda_2+\lambda_3))(\cdot,i) - \pi_s(\phi)\pi_s(\lambda_2+\lambda_3)(\cdot,i)]ds,\end{aligned} \tag{56}$$

for $S_i < t < S_{i+1}$ *and* $i \geq 0$.

At the observation times S_{i+1},

$$\pi_{S_{i+1}}\phi = \frac{\pi_{S_{i+1}^-}Q^{i,h_{i+1}}\phi}{\pi_{S_{i+1}^-}m^{(h_{i+1})}(\cdot,i)}, \tag{57}$$

for $i \geq 0$.

Remark 9 *The right-hand side of* (55) *is well-defined with probability 1, since*

$$\nu_0(\mathbb{R}^d\times\Gamma_{h_0}) > 0, \quad \mathbb{P}-a.s.$$

Indeed the probability of observing a value h_0 *of* j_0 *such that* $\nu_0(\mathbb{R}^d\times\Gamma_{h_0}) = \mathbb{P}(y_0\in\Gamma_{h_0}) = 0$ *is clearly zero.*

A similar property holds for the right-hand side of (57). *Indeed the predictable* $\mathcal{F}_t^p \vee \sigma\{j_0\}$*-intensity of* $(v_t^{(h)})$ *is* $\pi_{t^-}m^{(h)}(\cdot, z_{t^-})$*, and when it is evaluated at time* S_{i+1}*, its value is* $\pi_{S_{i+1}^-}m^{(h)}(\cdot,i)$*, since* $z_{S_{i+1}^-} = z_{S_i} = i$*. Moreover* S_{i+1} *is a jump time of the counting process* $(v_t^{(h_{i+1})})$*. Then the denominator of (57) is the intensity of* $(v_t^{(h_{i+1})})$ *evaluated at one of its jump times and cannot be zero.*

Remark 10 *Taking into account Remark 6 to the Regularity Conditions on the diffusion part (Section 2), under our assumptions*

- *when the jump intensities are uniformly bounded (condition (44)), then the domain* $\mathcal{D}$ *contains* $C_K^2(E)$*, where* $E = \mathbb{R}^d\times\mathbb{R}^k$*; when moreover* B *and* C *are uniformly bounded, then* $\mathcal{D}$ *contains* $C_b^2(E)$*,*
- *when the intensities grow sublinearly (condition (40)) and the jump sizes are bounded (condition (41)), then the domain* $\mathcal{D}$ *contains* $C_K^2(E)$*.*

Similarly, for the operators $Q^{i,h}$, we have that

- *under condition (44), then $Q^{i,h} \subset \boldsymbol{B}(E) \times \boldsymbol{B}(E)$,*
- *under conditions (40) and(41), then $Q^{i,h} \subset \boldsymbol{B}_K(E) \times \boldsymbol{B}(E)$, where $\boldsymbol{B}_K(E)$ is the subset of function in $\boldsymbol{B}(E)$ with compact support.*

It is also interesting to note that the operator G^i drives the system (x_t, y_t) between the observation times S_i and S_{i+1}, while the operator $Q^{i,h}$ takes into account what happens at the observation time S_{i+1}.

Proof of Theorem 2

Let ϕ be in $\mathcal{D}$, then using the notation (16), then , for $F(\xi,\eta) = F_\phi(\xi,\eta) = \phi(\xi)$, we can rewrite equation (50):

$$\begin{aligned}
\pi_t\phi &= \pi_0\phi + \int_0^t \pi_s \mathbb{A}^{z,j}\, \phi\big|_{z=z_{s^-},j=j_{s^-}}\, ds \\
&+ \sum_h \int_0^t \pi^+_{s^-} m^{(h)}(\cdot,z) \Big[\pi_{s^-}(\phi m^{(h)})(\cdot,z) - \pi_{s^-}\phi\pi_{s^-} m^{(h)}(\cdot,z) \\
&+ \pi_{s^-}(\mathcal{R}_2^{(h)} + \mathcal{R}_3^{(h)})^{z,j}\,\phi\Big] \Big(dv_s^{(h)} - \pi_{s^-} m^{(h)}(\cdot,z)ds\Big)\Big|_{z=z_{s^-},j=j_{s^-}}.
\end{aligned}$$

For $S_i < t < S_{i+1}$

$$\begin{aligned}
\pi_t\phi &= \pi_{S_i}\phi + \int_{S_i}^t \pi_s \mathbb{A}^{z,j}\,\phi\,\big|_{z=i,j=j_{S_i}}\, ds \qquad (58)\\
&- \sum_h \int_{S_i}^t \Big[\pi_s(\phi m^{(h)})(\cdot,z) - \pi_s\phi\pi_s m^{(h)}(\cdot,z) + \pi_{s^-}(\mathcal{R}_2^{(h)} + \mathcal{R}_3^{(h)})^{z,j}\,\phi\Big]\,\Big|_{z=i,j=j_{S_i}} ds.
\end{aligned}$$

Define

$$\mathcal{R}_2F(\xi,\eta) := \sum_h \mathcal{R}_2^{(h)}F(\xi,\eta)$$

$$\mathcal{R}_3F(\xi,\eta) := \sum_h \mathcal{R}_3^{(h)}F(\xi,\eta),$$

and observe that $\mathcal{R}_2F(\xi,\eta)$ and $\mathcal{R}_3F(\xi,\eta)$ coincide with the last two addends in the definition (10) of $\mathbb{A}$, so that

$$(\mathbb{A} - \mathcal{R}_2 - \mathcal{R}_3)F(\xi,\eta) = \qquad (59)$$

$$
\begin{aligned}
&= b(\xi,z)\nabla_x F(\xi,\eta)+\frac{1}{2}c(\xi,z)c^*(\xi,z)\nabla_\xi^2 F(\xi,\eta)\\
&+ \int_{D_0(\xi,z)}\left[F(x+K_0(\xi,z;\zeta),y,\eta)-F(\xi,\eta)\right]\nu(d\zeta)\\
&+ \int_{D_1(\xi,z)}\left[F(x+K_0(\xi,z;\zeta),y+K_1(\xi,z;\zeta),\eta)-F(\xi,\eta)\right]\nu(d\zeta).
\end{aligned}
$$

Finally observe that

$$\sum_h m^{(h)}(\xi,z)=\lambda_2(\xi,z)+\lambda_3(\xi,z). \tag{60}$$

With the previous observations in mind we can rewrite (58) as

$$
\begin{aligned}
\pi_t\phi &= \pi_{S_i}\phi+\int_{S_i}^t \pi_s(\mathbb{A}-\mathcal{R}_2-\mathcal{R}_3)^{z,j}\,\phi\,\big|_{z=i,j=j_{S_i}}\,ds \\
&- \int_{S_i}^t\left[\pi_s(\phi(\lambda_2+\lambda_3))(\cdot,i)-\pi_s(\phi)\pi_s(\lambda_2+\lambda_3)(\cdot,i)\right]\,ds.
\end{aligned} \tag{61}
$$

Note that $(\mathbb{A}-\mathcal{R}_2-\mathcal{R}_3)^{z,j}|_{z=i,j=j_{S_i}}$ represents the operator which drives the system between the observation times S_i and S_{i+1}, and note that

$$G^i\phi=(\mathbb{A}-\mathcal{R}_2-\mathcal{R}_3)^{z,j}\;\phi|_{z=i,j=j_{S_i}},$$

whichever is the value of $j=j_{S_i}$. Then (56) is proved between two jump times. At the observation times S_l, for $l\geq 1$, we easily verify that the filter satisfies

$$\pi_{S_l}\phi=\left.\frac{\pi_{S_l^-}\left[(\mathcal{R}_2^{(h)}+\mathcal{R}_3^{(h)})^{z,j}+\mathbb{I}_{\{y\in\Gamma_h\}}\lambda_2(\xi,z)+\lambda_3^{(h)}(\xi,z)\right]\phi}{\pi_{S_l^-}m^{(h)}(\cdot,z)}\right|_{z=z_{S_l^-},j=j_{S_{l-1}},h=j_{S_l}} \tag{62}$$

Taking into account (48) and (49) we immediately get that $(\mathcal{R}_2^{(h)}+\mathcal{R}_3^{(h)})^{z,j}$ does not depend on j and moreover

$$\left[(\mathcal{R}_2^{(h)}+\mathcal{R}_3^{(h)})^{z,j}+\mathbb{I}_{\{y\in\Gamma_h\}}\lambda_2(\xi,z)+\lambda_3^{(h)}(\xi,z)\right]\phi(\xi)=Q^{z,h}\phi(\xi)$$

Then the recursive equation (57) at jump time S_l, with observation $j_{S_l}=h_l$, i.e.

$$\pi_{S_l}\phi=\frac{\pi_{S_l^-}Q^{l-1,h_l}\phi}{\pi_{S_l^-}m^{(h_l)}(\cdot,l-1)}, \tag{63}$$

follows observing that $z_{S_l^-}=l-1$ and $j_{S_l}=h_l$.

4 The recursive algorithm

In this Section we define a deterministic algorithm that will be used in the next Section to identify the filter. We need to introduce the space $\mathcal{X}$ of history sets up to time T, whose elements are the history sets

$$\chi = \{(s_l, h_l) \;\; l = 0, 1, \dots, n \;\; \}$$

such that

$$n \geq 0, \quad 0 = s_0 < s_1 < s_2 < \dots < s_n < T, \;\; h_l \in \mathbb{Z}.$$

The space $\mathcal{X}$ is endowed with the topology induced by the map τ from $\mathcal{X}$ to the Skorohod space $D_{\mathbb{N}\times\mathbb{Z}}[0,T]$, obtained by mapping the history set $\chi = \{(s_l, h_l) \;\; l = 0, 1, \dots, n \;\; \}$ to the deterministic piece-wise constant functions $(\hat{z}_t, \hat{\jmath}_t) = \tau(\chi)$ defined as

$$\hat{z}_t = l \quad s_l \leq t < s_{l+1} \leq T,$$

$$\hat{\jmath}_t = h_l \quad s_l \leq t < s_{l+1} \leq T, \quad l = 0, \dots n,$$

with the convention $s_{n+1} = T$.

Let $\theta : D_{\mathbb{N}\times\mathbb{Z}}[0,T] \to \mathcal{X}$ be the map which associates to each deterministic trajectory $\hat{\eta}_\cdot$, i.e. $t \to \hat{\eta}_t = (\hat{z}_t, \hat{\jmath}_t)$, the set $\chi = \theta(\hat{\eta}_\cdot) = \{(s_l, h_l), l \geq 0\}$, where $s_0 = 0$, s_l are the ordered jump times of $(\hat{z}_t, \hat{\jmath}_t)$ and h_l are the values of $(\hat{\jmath}_t)$ at the jump times, that is $h_l = \hat{\jmath}_{s_l}$, for $l \geq 0$. Then θ is measurable.

Let $t \in [0,T]$ and $\chi = \{(s_l, h_l) \quad l = 0, 1, \dots, n\}$ be a history set, then we introduce the history set $[\chi]_t$ up to time t as in [1]:

$$[\chi]_t := \{(s_l, h_l) \quad l = 0, 1, \dots, i\},$$

when

$$s_i \leq t < s_{i+1},$$

and

$$[\chi]_T := \chi.$$

We now exploit the link between histories and our problem. Whenever the marked point process p defined in (25) (or in (27) and (28)) is non-explosive, and therefore the corresponding process $(\eta_t) = (z_t, j_t)$ takes values in $D_{\mathbb{N}\times\mathbb{Z}}[0,T]$, then, with the above notations, we can associate to it the history process $(\mathcal{H}_t)$ with values in $\mathcal{X}$

$$\mathcal{H}_t := [\theta(\eta_\cdot)]_t$$

i.e.

$$\mathcal{H}_t = \{(S_l, j_{S_l}), l = 0, 1, \dots, i\}$$

when the jump times of (η_t) satisfy $0 = S_0 < S_1 < \dots < S_i \leq t < S_{i+1}$.

Then the observation up to time t of the marked point process p coincides with the observation of the history process at time t, in the sense that

$$\mathcal{F}_t^p \vee \sigma\{j_0\} = \sigma\{\mathcal{H}_t\} = \sigma\{[\theta(\eta_\cdot)]_t\}.$$

We can read the system (56), (57) for the filter (π_t) as an equation in $D_{\Pi(\mathbb{R}^d\times\mathbb{R}^k)}[0,T]$ parameterized by random history sets χ.

The deterministic recursive nonlinear system

Let χ be a history set, and let μ be a probability measure in $\Pi(\mathbb{R}^d \times \mathbb{R}^k)$. Define $\mu_{s_0} = \mu_0 = \mu$, and

$$\begin{aligned} \mu_t\phi &= \mu_{s_i}\phi + \int_{s_i}^t \mu_s G^i \phi(\cdot) ds \\ &\quad - \int_{s_i}^t [\mu_s(\phi(\lambda_2+\lambda_3))(\cdot,i) - \mu_s(\phi)\mu_s(\lambda_2+\lambda_3)(\cdot,i)]\, ds, \end{aligned} \tag{64}$$

for $i \geq 0$ and $s_i < t < s_{i+1}$. Here G^i is the operator introduced in (102).

For $i \geq 1$, if $\mu_{s_i^-} m^{(h_i)}(\cdot, i-1) \neq 0$, define

$$\mu_{s_i}\phi = \frac{\mu_{s_i^-} Q^{i-1,h_i}\phi}{\mu_{s_i^-} m^{(h_i)}(\cdot, i-1)}, \tag{65}$$

where the operator $Q^{z,h}$ is introduced in (54).

On the other hand, if $\mu_{s_i^-} m^{(h_i)}(\cdot, i-1) = 0$, then

$$\mu_{s_i}\phi = \mu_{s_i^-}\phi. \tag{66}$$

Note that in (65) the denominator is the normalization factor since

$$m^{(h)}(\cdot, i-1) = Q^{i-1,h}\mathbb{I}_{\mathbb{R}^d\times\mathbb{Z}}.$$

Remark 11 *The above system does not use the value h_0. The value of h_0 plays a crucial role in connection with our filtering problem. Indeed, while the aim of this Section is to find* a *(not necessarily* the*) solution of the above deterministic system, the aim of the paper is to show that the filter $\pi_t(\omega)$ can be computed through this solution μ_t when $\chi = \theta(\eta_\cdot(\omega))$ and*

$$\mu = \pi_0(\omega) = \pi_{S_0}(\omega),$$

where $\pi_0(\omega)$ is the filter at time $t = 0$ and coincides $\mathbb{P}-a.s.$ with (51). We note that $\pi_0(\omega)$ can be rewritten in terms of the history sets and the map θ as $\pi_0(\omega) = \mu_0(\theta(\eta_\cdot(\omega)))$, $\mathbb{P}-a.s.$. Here, for $\chi = \{(s_l, h_l) \quad l = 0, 1, \dots, n\}$, we have used the following notation

$$\mu_0(\chi)\phi = \frac{\nu_0(\mathbb{I}_{\{\mathbb{R}^d \times \Gamma_{h_0}\}}\phi)}{\nu_0(\mathbb{R}^d \times \Gamma_{h_0})}, \tag{67}$$

when $\nu_0(\mathbb{R}^d \times \Gamma_{h_0}) > 0$, and $\mu_0(\chi)\phi = \delta_{\{(0,0)\}}\phi = \phi(0,0)$, otherwise. Instead of $\delta_{\{(0,0)\}}$ any other probability measure could be used (see the proof of Theorem 1 and Remark 9 to Theorem 2).

The deterministic recursive unnormalized system

In order to find a solution of the above recursive system for $t \in [0, T]$, we need an auxiliary system in the space of non-negative measures ρ on $\mathbb{R}^d \times \mathbb{R}^k$. This system is a modification of the deterministic recursive nonlinear system: the equation for $t \in (s_i, s_{i+1})$ is obtained from equation (64) by dropping the nonlinear term $\mu_s(\phi)\mu_s(\lambda_2 + \lambda_3)(\cdot, i)$ in the second integral

$$\rho_t\phi = \rho_{s_i}\phi + \int_{s_i}^t \rho_s G^i \phi(\cdot) ds - \int_{s_i}^t \rho_s(\phi(\lambda_2 + \lambda_3))(\cdot, i)(\cdot, i) ds; \tag{68}$$

the equations at times s_i are equal to (65) and (66), clearly using the non-negative measure $\rho_{s_i^-}$ instead of the probability measure $\mu_{s_i^-}$. A solution of the nonlinear system is then obtained by normalization (see Theorem 3, below).

Let μ be a probability measure on $\mathbb{R}^d \times \mathbb{R}^k$. Let $\chi = \{(s_l, h_l) \quad l = 0, 1, \dots, n\}$ be a history set.

Following the idea of [10], a solution $r(t, \chi) = r(t, \chi, \mu)$ (for the sake of simplicity, we drop the dependence from μ in the notation) of the unnormalized recursive system can be given via Feynman-Kac formula as follows:

For $t = s_0 = 0$

$$r(0, \chi)\phi = \mu\phi, \tag{69}$$

and for $t \in (0, s_1)$

$$r(t, \chi)\phi := \mathbb{E}[\phi(X_t^0(\chi), Y_t^0(\chi)) \exp\{-\int_0^t (\lambda_2 + \lambda_3)(X_u^0(\chi), Y_u^0(\chi), 0)\, du\}] \tag{70}$$

where $(X_u^0(\chi), Y_u^0(\chi))$ is a Markov process with values in $\mathbb{R}^d \times \mathbb{R}^k$, with generator G^0 and initial distribution

$$\widehat{\pi}_0(\chi) = r(0, \chi) = \mu. \tag{71}$$

When $t \in (s_i, s_{i+1})$

$$r(t,\chi)\phi := \mathbb{E}[\phi(X^i_{t-s_i}(\chi), Y^i_{t-s_i}(\chi)) \exp\{-\int_0^{t-s_i} (\lambda_2+\lambda_3)(X^i_u(\chi), Y^i_u(\chi), i)\, du\}] \tag{72}$$

where, for $i \geq 1$, $((X^i_u(\chi), Y^i_u(\chi))$ is a Markov process with values in $\mathbb{R}^d \times \mathbb{R}^k$, with generator G^i and initial distribution

$$\widehat{\pi}_i(\chi) = r(s_i, \chi). \tag{73}$$

The initial distributions are defined by induction as follows.

Let $r(s^-_{i+1}, \chi)$ be the limit (in the weak topology) of the measures $r(t,\chi)$, for t converging to s^-_{i+1}.

Then

- if $r(s^-_{i+1}, \chi) m^{(h_{i+1})}(\cdot, i) \neq 0$, then $\widehat{\pi}_{i+1}(\chi) = r(s_{i+1}, \chi)$, where

$$r(s_{i+1}, \chi)\phi = \frac{r(s^-_{i+1}, \chi) Q^{i,h_{i+1}}\phi}{r(s^-_{i+1}, \chi) m^{(h_{i+1})}(\cdot, i)}, \tag{74}$$

- if $r(s^-_{i+1}, \chi) m^{(h_{i+1})}(\cdot, i) = 0$, then $\widehat{\pi}_{i+1}(\chi) = r(s_{i+1}, \chi)$, where

$$r(s_{i+1}, \chi)\phi = \frac{r(s^-_{i+1}, \chi)\phi}{r(s^-_{i+1}, \chi)(\mathbb{R}^d \times \mathbb{R}^k)}. \tag{75}$$

Remark 12 *The measure $r(s^-_{i+1}, \chi)$ evaluated in any function $\psi \in \boldsymbol{B}(\mathbb{R}^d \times \mathbb{R}^k)$, is*

$$\mathbb{E}\left[\psi(X^i_{s_{i+1}-s_i}(\chi), Y^i_{s_{i+1}-s_i}(\chi)) \exp\{-\int_0^{s_{i+1}-s_i} (\lambda_2+\lambda_3)(X^i_u(\chi), Y^i_u(\chi), i)\, du\}\right].$$

Therefore the explicit expression of the numerator in (74) is

$$r(s^-_{i+1}, \chi) Q^{i,h_{i+1}}\phi :=$$

$$\mathbb{E}\left[\left\{\mathbb{I}_{\{Y^i_{s_{i+1}-s_i} \in \Gamma_h\}} \int_{D_2(\xi^i_{s_{i+1}-s_i}, i)} \phi\left(X^i_{s_{i+1}-s_i}(\chi) + K_0(\xi^i_{s_{i+1}-s_i}, i; \zeta), Y^i_{s_{i+1}-s_i}(\chi)\right) \nu(d\zeta)\right.\right.$$

$$\left. + \int_{D_3^{(h)}(\xi^i_{s_{i+1}-s_i}, i)} \phi\left(X^i_{s_{i+1}-s_i}(\chi) + K_0(\xi^i_{s_{i+1}-s_i}, i; \zeta), Y^i_{s_{i+1}-s_i}(\chi) + K_3(\xi^i_{s_{i+1}-s_i}, i; \zeta)\right) \nu(d\zeta)\right\}$$

$$\left. \exp\{-\int_0^{s_{i+1}-s_i} (\lambda_2+\lambda_3)(X^i_u(\chi), Y^i_u(\chi), i)\, du\}\right]$$

where $\xi^i_u = (X^i_u(\chi), Y^i_u(\chi))$

Remark 13 *For any* $\chi = \{(s_l, h_l) \quad l = 0, 1, \dots, n\}$ *and* $i \leq n$*, the process* $((X^i_u(\chi), Y^i_u(\chi))$ *does not depend on* (s_l, h_l) *for* $l > i$*. Then, for any* χ *and* t*, with* $s_i \leq t < s_{i+1}$*,* $r(t, \chi) = r(t, [\chi]_t)$*, and therefore* $r(t^-, \chi) = r(t^-, [\chi]_{t^-})$*.*

Moreover the evolution of the process $(X^i_u(\chi), Y^i_u(\chi))$ *is driven by* G^i *and therefore the evolution depends on* i*, but not on* h_i*, while the initial distribution* $\widehat{\pi}_i(\chi)$ *depends directly on* (s_i, h_i)*, and indirectly on* (s_l, h_l) *for* $l \leq i-1$*, since* $r(s_i^-, \chi) = r(s_i^-, [\chi]_{s_i-}) = r(s_i^-, [\chi]_{s_{i-1}})$*.*

Finally the solution $r(t, \chi)$ *depends also on the initial distribution* μ*. The same happens for the laws of the processes* $((X^i_u(\chi), Y^i_u(\chi))$*.*

Remark 14 *The definition (69) ... (75) of the unnormalized solution is well posed only if* $r(s_{i+1}^-, \chi)(\mathbb{R}^d \times \mathbb{R}^k) > 0$ *for any* $i \geq 0$*, since otherwise both (74) and (75) are meaningless, the denominators being both equal to zero.*

To this end we observe the following: let t *be a time in* (s_i, s_{i+1}) *and let* $r(t, \chi)(\mathbb{R}^d \times \mathbb{R}^k) = 0$*, then* $r(t', \chi)(\mathbb{R}^d \times \mathbb{R}^k) = 0$ *for all* $t \leq t' < s_{i+1}$*, and therefore* $r(s_{i+1}^-, \chi)(\mathbb{R}^d \times \mathbb{R}^k) = 0$*.*

Indeed, for a time , $r(t, \chi)(\mathbb{R}^d \times \mathbb{R}^k) = 0$ *if and only if*

$$\exp\{-\int_0^{t-s_i} (\lambda_2 + \lambda_3)(X^i_u(\chi), Y^i_u(\chi), i)\, du\} = 0 \qquad a.s.$$

This happens only when

$$\int_0^{t-s_i} (\lambda_2 + \lambda_3)(X^i_u(\chi), Y^i_u(\chi), i)\, du = \infty \qquad a.s.,$$

and in this case the same holds also for t'*.*

As a consequence, a necessary condition for the well-posedness of the solution is

$$\mathbb{P}\left(\int_0^{t-s_i} (\lambda_2 + \lambda_3)(X^i_u(\chi), Y^i_u(\chi), i)\, du < \infty\right) > 0, \qquad \text{for all } t \in (s_i, s_{i+1}),\ i \geq 0. \tag{76}$$

It turns out that the previous condition (76) for $t = s_{i+1}$ is the key for the proof that $r(t, \chi)$ is well-defined, as it is explained in the next two Lemmas.

Lemma 1 *Assume that the jump intensities* $\lambda_2(\xi, z)$ *and* $\lambda_3(\xi, z)$ *are bounded above, uniformly in* ξ*:*

$$\lambda_2(\xi, z) \leq \overline{\lambda}_2(z) \quad \lambda_3(\xi, z) \leq \overline{\lambda}_3(z), \tag{77}$$

for all $z \in \mathbb{N}$*.*

Then (69) ... (75) is well-defined and

$$r(t, \chi)(\mathbb{R}^d \times \mathbb{R}^k) > 0, \tag{78}$$

for every $t \in [0, T]$*.*

Proof Clearly condition (77) implies that

$$\int_0^v (\lambda_2 + \lambda_3)(\widehat{x}_u, \widehat{y}_u, i)\, du \le (\overline{\lambda}_2(i) + \overline{\lambda}_3(i))\, v$$

for any $v \ge 0$, and for any (measurable in time) function $(\widehat{x}_u, \widehat{y}_u)$. Then

$$\mathbb{P}\left(\int_0^v (\lambda_2 + \lambda_3)(X_u^i(\chi), Y_u^i(\chi), i)\, du < \infty \right) = 1 \tag{79}$$

for any $v \ge 0$, for any χ, for any i, and for any μ.

Then, under this assumption, the measures $r(t, \chi)$ could become zero, only at the times s_{i+1}, for $i \ge 0$. However this situation is ruled out, indeed:

a. the above relation (79) holds also with $v = s_{i+1} - s_i$,

b. $r(s_{i+1}^-, \chi)\phi = \lim_{t \to s_{i+1}^-} r(t, \chi)\phi$, for all $\phi \in C_b(\mathbb{R}^d \times \mathbb{R}^k)$

then, in view of Remark 12,

c. $r(s_{i+1}^-, \chi)(\mathbb{R}^d \times \mathbb{R}^k) = \mathbb{E}[\exp\{-\int_0^{s_{i+1}-s_i} (\lambda_2 + \lambda_3)(X_u^i(\chi), Y_u^i(\chi), i)\, du\}] \ne 0$

Finally, either we can apply (74) or we have to apply (75), $r(s_i, \chi)$ is a probability measure.

Lemma 2 *Assume that the jump intensities $\lambda_2(\xi, z)$ and $\lambda_3(\xi, z)$ are sub-linearly growing, uniformly in ξ:*

$$\lambda_2(\xi, z) \le \overline{\lambda}_2(z)(1 + |\xi|) \quad \lambda_3(\xi, z) \le \overline{\lambda}_3(z)(1 + |\xi|), \tag{80}$$

for all $z \in \mathbb{N}$.
Assume that the jump sizes K_0 and K_3 are bounded, uniformly in ξ:

$$K_0(\xi, z) \le \overline{K}_0(z) \quad K_3(\xi, z) \le \overline{K}_3(z), \tag{81}$$

for all $z \in \mathbb{N}$.

Let $(\mathcal{T}_u^i)$ be the semigroup generated by G^i, for any i.

Let $g_l(x, y) = |(x, y)|^l$ and assume that the semigroup $(\mathcal{T}_s^i)$ has the property:

$$\text{if } \nu g_l < \infty, \text{ then } \nu \mathcal{T}_s^i(g_l) \le C_i^l(\nu, T) < \infty, \text{ uniformly for } s \in [0, T].$$

Assume that the initial distribution μ has all moments finite.

Then (69) ... (75) is well-defined and condition (78) holds.

Proof We prove by induction that for all $i \ge 1$

$$\widehat{\pi}_i g_l < \infty \text{ for all } l,$$

and

$$r(s_i^-, \chi)(E) > 0,$$

where $E = \mathbb{R}^d \times \mathbb{R}^k$.

The properties of the semigroups $(\mathcal{T}_s^i)$ imply that $r(s_1^-, \chi)(E) > 0$. Indeed, taking into account that

$$r(s_{i+1}^-, \chi)(E) = \mathbb{E}\left[\exp\left\{-\int_0^{s_{i+1}-s_i} (\lambda_2 + \lambda_3)\,(X_u^i(\chi), Y_u^i(\chi), i)\;du\right\}\right],$$

and applying Jensen inequality to the function $\exp\{-r\}$ (so that for any r.v. R, $\mathbb{E}[\exp\{-R\}] \geq \exp\{-\mathbb{E}[R]\}$), we get

$$r(s_1^-, \chi)(E) \geq \exp\left\{-\mathbb{E}[\int_0^{s_1} (\lambda_2 + \lambda_3)\,(X_u^0(\chi), Y_u^0(\chi), 0)\;du]\right\}.$$

Then, by condition (80)

$$\mathbb{E}\left[\int_0^s (\lambda_2 + \lambda_3)\,(X_u^0(\chi), Y_u^0(\chi), 0)\;du\right] \leq \int_0^s \mathbb{E}\left[\left(\overline{\lambda}_2(0) + \overline{\lambda}_3(0)\right)(1 + |(X_u^0(\chi), Y_u^0(\chi))|)\right]\;du$$

$$= \int_0^s \mathbb{E}\left[\left(\overline{\lambda}_2(0) + \overline{\lambda}_3(0)\right)\left(1 + g_1(X_u^0(\chi), Y_u^0(\chi))\right)\right]\;du$$

The condition on the semigroups $(\mathcal{T}^i)$ implies then that the previous expression is bounded above uniformly for $s \in [0, T]$. Therefore

$$r(s_1^-, \chi)(E) > 0 \text{ for all } \chi.$$

Then, for the initial distribution $\widehat{\pi}_1(\chi)$ there is an alternative.

Either $r(s_1^-, \chi)m^{(h_1)}(\cdot, i) = 0$,

and then $\widehat{\pi}_1(\chi)g_l$ is finite if and only if the numerator in (75), for $i = 0$ and $\phi(\xi) = g_l(\xi)$, is finite, i.e.

$$r(s_1^-, \chi)g_l < \infty,$$

and the above inequality is assured by the property of the semigroups $(\mathcal{T}_s^i)$ since clearly

$$r(s_1^-, \chi)g_l \leq \mathbb{E}\left[g_l(X_{s_1}^0(\chi), Y_{s_1}^0(\chi))\right] = \mu\mathcal{T}_{s_1}^0(g_l) \leq C_0^l(\mu, T) < \infty,$$

or $r(s_1^-, \chi)m^{(h_1)}(\cdot, i) > 0$,

and then $\widehat{\pi}_1(\chi)g_l$ is finite if and only if the numerator in (74), for $i = 0$, is finite, i.e.

$$r(s_1^-, \chi)Q^{0,h_1}g_l < \infty,$$

and this follows by considering that, under our hypotheses, for any l, i, there exists a constant $\widetilde{C}_{l,i}$ such that

$$Q^{i,h} g_l \leq \widetilde{C}_{l,i} \left(g_{l+1} + 1 \right) \qquad \text{for all } h. \tag{82}$$

Before proving it , we remark that the above inequality (82), together with the property of the semigroups $(\mathcal{T}_s^i)$, implies that

$$\begin{aligned} r(s_1^-, \chi) Q^{0,h} g_l &\leq \widetilde{C}_{l,0} \left(\mathbb{E} \left[g_{l+1}(X_{s_1}^0(\chi), Y_{s_1}^0(\chi)) \right] + 1 \right) \\ &= \mu \mathcal{T}_{s_1}^0 \left(g_{l+1} \right) \leq C_0^{l+1}(\mu, T) < \infty. \end{aligned}$$

The proof of inequality (82) is easy, indeed

$$\begin{aligned} Q^{i,h} g_l(\xi) &:= \mathbb{I}_{\{y \in \Gamma_h\}} \int_{D_2(\xi,i)} \left| (x + K_0(\xi, i; \zeta), y) \right|^l \nu(d\zeta) \qquad (83) \\ &\quad + \int_{D_3^{(h)}(\xi,i)} \left| (x + K_0(\xi, i; \zeta), y + K_3(\xi, i; \zeta)) \right|^l \nu(d\zeta). \end{aligned}$$

$$Q^{i,h} g_l \leq \left(\overline{\lambda}_2(i) + \overline{\lambda}_3(i) \right) (1 + g_1) 2^l \left(g_l + \left(\overline{K}_0(i) + \overline{K}_3(i) \right)^l \right),$$

and therefore inequality (82) is proved.

Similarly, assuming that $r(s_i^-, \chi)(E) > 0$, and that

$$\widehat{\pi}_i g_l < \infty \text{ for all } l,$$

and using

$$r(s_{i+1}^-, \chi)(E) \geq \exp\{- \int_0^{s_{i+1} - s_i} \mathbb{E} \left[(\lambda_2 + \lambda_3) \left(X_u^i(\chi), Y_u^i(\chi), i \right) \right] \, du \}$$

we get that $r(s_{i+1}^-, \chi)(E) > 0$, and that

$$\widehat{\pi}_{i+1} g_l < \infty \text{ for all } l,$$

Indeed

$$\begin{aligned} r(s_{i+1}^-, \chi) g_l &\leq \mathbb{E} \left[g_l \left(X_{s_{i+1} - s_i}^i(\chi), Y_{s_{i+1} - s_i}^i(\chi) \right) \right] \\ &= \widehat{\pi}_i(\chi) \mathcal{T}_{s_{i+1} - s_i}^i (g_l) \leq C_i^l(\widehat{\pi}_i(\chi), T) < \infty, \end{aligned}$$

and

$$r(s_{i+1}^-, \chi) Q^{i, h_{i+1}} g_l < \infty \qquad \text{for all } l.$$

More in general, for any h,

$$r(s_{i+1}^-, \chi)Q^{i,h} g_l \leq \widetilde{C}_{l,i}\left(\mathbb{E}\left[g_l(X^i_{s_{i+1}-s_i}(\chi), Y^i_{s_{i+1}-s_i}(\chi))\right] + 1\right)$$
$$= \widehat{\pi}_i(\chi)\mathcal{T}^i_{s_{i+1}-s_i}(g_{l+1}) \leq C_i^{l+1}(\widehat{\pi}_i(\chi), T) < \infty,$$

and therefore, for all times $s \in [0, T]$

$$0 < r(s, \chi)(E) < \infty.$$

Observation 3 When one of the sets of conditions discussed at the end of Section 2 is satisfied, then, for any $i = 0, 1, \ldots$, the analogous conditions for the processes $(X^i_u(\chi), Y^i_u(\chi))_{u \geq 0}$ are satisfied, and then we can assume without loss of generality that their sample paths are in $D_{\mathbb{R}^d \times \mathbb{R}^k}[0, T]$.

Theorem 3 *Assume that one of the sets of conditions discussed at the end of Section 2 are satisfied. Let χ be a history set and let μ be a distribution on $\mathbb{R}^d \times \mathbb{R}^k$. Assume condition (78), i.e.*

$$r(t, \chi)(\mathbb{R}^d \times \mathbb{R}^k) > 0,$$

for every $t \in [0, T]$.

Define

$$H(t, \chi, \mu)\phi := \frac{r(t, \chi)\phi}{r(t, \chi)\mathbb{I}_{\mathbb{R}^d \times \mathbb{R}^k}}, \tag{84}$$

for every bounded measurable ϕ and $t \in [0, T]$. Then $H(t, \chi, \mu) \in D_{\Pi(\mathbb{R}^d \times \mathbb{R}^k)}[0, T]$ and moreover $(H(t, \chi, \mu))$ solves the system (64), (65), (66) with parameter χ and initial distribution μ.

Proof We note that $r(t, \chi)$ is a nonnegative non-trivial measure on $\mathbb{R}^d \times \mathbb{R}^k$ and therefore $H(t, \chi, \mu)$ is a probability measure. The continuity from the right with respect to t in the space $\Pi(\mathbb{R}^d \times \mathbb{R}^k)$, endowed with the weak topology, is assured by the continuity from the right of $r(t, \chi)\phi$, for any $\phi \in C_b(\mathbb{R}^d \times \mathbb{R}^k)$. Indeed for any such ϕ and for any i, the processes $(X^i_u(\chi), Y^i_u(\chi), i)_{u \geq 0}$ can be taken with paths in $D_{\mathbb{R}^d \times \mathbb{R}^k}[0, T]$ and

$$\phi(X^i_{t-s_i}(\chi), Y^i_{t-s_i}(\chi)) \exp\{-\int_0^{t-s_i} (\lambda_2 + \lambda_3)(X^i_u(\chi), Y^i_u(\chi), i)\, du\}$$

is uniformly bounded by $\|\phi\|_\infty$ and right continuous in time for $t \in [s_i, s_{i+1})$. Then right continuity for H follows. Analogously one can prove the existence of the limits from the left.

Let $\mu_t := H(t, \chi, \mu)$, then (64), (65), (66) follow by easy computation, using the linearized system.

The functional $H(t, \cdot, \cdot)$ defined in the previous Theorem, is the natural candidate of the functional H_t of Proposition 1. In the next Section we prove that one can recover the filter, when $H(t, \cdot, \cdot)$ is applied to the random history set generated by the observations and to the corresponding initial distribution.

5 The filter identification

In this section we follow the approach of [12] for the identification of the filter, and analogously to what was done in [10] in the case of counting observations, we prove that under suitable hypotheses the algorithm $H(t, \chi, \mu)$ of the previous section evaluated in $\chi = \theta(\eta.)$ and $\mu = \pi_0$ generates the filter.

Theorem 4 *Assume that one set of conditions of Section 2 is satisfied and that the jump intensities λ_2 and λ_3 are uniformly bounded above*

$$\lambda_l(\xi, \eta) \leq \overline{\lambda}_l, \quad \text{for } l = 2, 3.$$

Assume moreover that the range of $\alpha I - \mathbb{A}$ is separating and bounded pointwise dense in $\mathbf{B}(\mathbb{R}^d \times \mathbb{R}^k \times \mathbb{N} \times \mathbb{Z})$ for each $\alpha > 0$. Let $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{P})$ be a probability space and let $\widetilde{p}$ be a $\mathbb{Z}$-marked counting process on it. Let $(\widetilde{v}_t^{(h)})$ be the associated counting process relative to the mark h, i.e. $\widetilde{v}_t^{(h)} := \widetilde{p}((0, t], \{h\})$. Let $(\widetilde{z}_t)$ be a counting process and $(\widetilde{\jmath}_t)$ be a pure jump process defined on this space and assume that they verify the following relations:

$$\widetilde{z}_t = \sum_h \widetilde{v}_t^h \tag{85}$$

$$\widetilde{\jmath}_t = \widetilde{\jmath}_0 + \sum_h \int_0^t (h - \widetilde{\jmath}_{s^-}) \, d\widetilde{v}_s^h \tag{86}$$

where $\widetilde{\jmath}_0$ is a random variable with the same law of j_0, i.e.

$$\widetilde{P}(\widetilde{\jmath}_0 = h) = \nu_0(\mathbb{R}^d \times \Gamma_h).$$

Define $\widetilde{\eta}_t := (\widetilde{z}_t, \widetilde{\jmath}_t)$, and

$$\widetilde{\pi}_0(\widetilde{\omega})(dx, dy) := \mu_0\left(\theta(\widetilde{\eta}.(\widetilde{\omega}))\right)(dx, dy) = \frac{\mathbb{I}_{\{\mathbb{R}^d \times \Gamma_{\widetilde{\jmath}_0(\widetilde{\omega})}\}}(x, y)\nu_0(dx, dy)}{\nu_0(\mathbb{R}^d \times \Gamma_{\widetilde{\jmath}_0(\widetilde{\omega})})}, \tag{87}$$

where $\mu_0(\chi)$ is defined in (67).

Let $\widetilde{\mu}_t$ denote the probability measure

$$\widetilde{\mu}_t := H(t, \theta(\widetilde{\eta}.), \widetilde{\pi}_0), \tag{88}$$

where $H(t, \chi, \mu)$ defined in (84).

Finally assume that there exists a probability measure $\widetilde{P}_1 \ll \widetilde{P}|_{\mathcal{F}_T^{\widetilde{p}} \vee \sigma\{\widetilde{\jmath}_0\}}$ such that, for all $h \in \mathbb{Z}$, the predictable $(\widetilde{P}_1, \mathcal{F}_t^{\widetilde{p}} \vee \sigma\{\widetilde{\jmath}_0\})$-intensity of $(\widetilde{v}_t^{(h)})$ coincides with $H(t^-, \theta(\widetilde{\eta}.), \widetilde{\pi}_0) \, m^{(h)}(\cdot, \widetilde{z}_{t^-})$.

Then for any $\phi \in \mathbf{B}(\mathbb{R}^d \times \mathbb{Z})$ and for all $t \leq T$

$$\pi_t \phi = H(t, \theta(\eta.), \pi_0) \, \phi \quad \mathbb{P} - a.s. \tag{89}$$

Proof First of all note that the existence of the solution of the MP for $\mathbb{A}$ is guaranteed by the conditions of Section 2. The assumption on the range of $\alpha I - \mathbb{A}$ implies both uniqueness (in law) of the solution of the MP and of the FMP associated to the operator $\mathbb{A}$.

We consider the FMP with initial condition $\widetilde{\pi}_0$, and we note that

$$\mathbb{E}^{\widetilde{P}}[\widetilde{\pi}_0 F(\cdot, \widetilde{\eta}_0)] = \mathbb{E}[F(\xi_0, \eta_0)], \tag{90}$$

which is the analogous of (17) with $X_0 = \xi_0$ and $Y_0 = \eta_0$.

The process $(\widetilde{\mu}_t)$ is $\mathcal{F}_t^{\widetilde{p}} \vee \sigma\{\widetilde{\jmath}_0\}$-adapted with values in $\Pi(\mathbb{R}^d \times \mathbb{R}^k)$. By Theorem 3, for all $\widetilde{\omega}$, the trajectory $(\widetilde{\mu}_t(\widetilde{\omega}))$ satisfy the deterministic system (64), (65), (66), with initial condition $\mu = \widetilde{\pi}_0(\widetilde{\omega})$ and $\chi = \theta(\widetilde{\eta}_\cdot(\widetilde{\omega}))$. Thanks to (85) and (86), χ coincides with $\{(\widetilde{S}_l(\widetilde{\omega}), \widetilde{\jmath}_{S_l}(\widetilde{\omega})), l = 0, 1, \ldots, \widetilde{z}_T(\widetilde{\omega})\}$.

Under our notations, for all $\widetilde{\omega}$, $(\widetilde{\mu}_t)$ satisfies

$$\begin{aligned}
\widetilde{\mu}_t F(\cdot, \widetilde{\eta}_t) \;&=\; \widetilde{\mu}_0 F(\cdot, \widetilde{\eta}_0) + \int_0^t \widetilde{\mu}_s(\mathbb{A}F(\cdot, \widetilde{\eta}_s))ds \\
&+\; \sum_h \int_0^t [\widetilde{\mu}_{s^-} m^{(h)}(\cdot, \widetilde{z}_{s^-})]^+ \\
&\quad \Big[\widetilde{\mu}_{s^-}(Fm^{(h)})(\cdot, \widetilde{\eta}_{s^-}) - \widetilde{\mu}_{s^-} F(\cdot, \widetilde{\eta}_{s^-}) \widetilde{\mu}_{s^-} m^{(h)}(\cdot, \widetilde{z}_{s^-}) \\
&\qquad + \widetilde{\mu}_{s^-}(\mathcal{R}_2^{(h)} + \mathcal{R}_3^{(h)}) F(\cdot, \widetilde{\eta}_{s^-})\Big] \left(d\widetilde{v}_s^{(h)} - \widetilde{\mu}_{s^-} m^{(h)}(\cdot, \widetilde{z}_{s^-}) ds\right),
\end{aligned}$$

for any $F \in \mathcal{D}(\mathbb{A})$, with $F(\xi, \eta) = \phi(\xi)\psi(\eta)$ or a linear combination of such functions.

Indeed, the above identity is easy to prove in the first case, by using the product rule, and then it follows immediately for linear combinations.

Observe that by Remark 8 the integrands in $\left(d\widetilde{v}_s^{(h)} - \widetilde{\mu}_{s^-} m^{(h)}(\cdot, \widetilde{z}_{s^-}) ds\right)$ are bounded above by $4\|F\|_\infty$. Then the series is a martingale, since

$$\sum_h \widetilde{\mu}_{s^-} m^{(h)}(\cdot, \widetilde{z}_{s^-}) \leq \overline{\lambda}_2 + \overline{\lambda}_3$$

and therefore it is integrable. For a general $F \in \mathcal{D}(\mathbb{A})$ the martingale property follows by *bp*-convergence.

Moreover under $\widetilde{P}_1$ the intensity of $(\widetilde{v}_t^{(h)})$ is $H(t^-, \theta(\widetilde{\eta}_\cdot), \widetilde{\pi}_0)\, m^{(h)}(\cdot, \widetilde{z}_{t^-}) = \widetilde{\mu}_{t^-}\, m^{(h)}(\cdot, \widetilde{z}_{t^-})$. Then the second integral is a local martingale. Under our assumptions the intensity of $(\widetilde{z}_t)$ is bounded by a constant, therefore, taking into account (85), the second integral is a martingale. Then, under $\widetilde{P}_1$ the pair $(\widetilde{\mu}_t, \widetilde{\eta}_t)$ solves the FMP associated to our filtering problem with a *correct* initial condition (see (90) and compare with (17)).

By the assumption, the FMP is well-posed, and then, by Proposition 2, the processes (π_t, η_t) and $(\widetilde{\mu}_t, \widetilde{\eta}_t)$ are equal in law. Moreover (88) defines the functional H_t of (18), and therefore (19) implies (89).

Remark 15 *Formulas (85) and (86) are the analogous of formulas (29) and (30). They are used in the proof in order to get that* $\widetilde{\mu}_t$ *satisfies the above system. Moreover the hypothesis implies that* $\widetilde{\mu}_{t^-} m^{(h)}(\cdot, \widetilde{z}_{t^-})|_{t=\widetilde{S}_i, h=\widetilde{\jmath}_{\widetilde{S}_i}}$ *is positive* $\widetilde{P}_1$*-a.s., (see Remark 9 to Theorem 2).*

In the same framework of the previous theorem a simple sufficient condition implies the existence of the probability measure $\widetilde{P}_1$ absolutely continuous with respect to $\widetilde{P}$.

Theorem 5 *Assume the notations and the hypotheses of Theorem* 4 *up to condition (88). Let* $\widetilde{\mathcal{M}}_t^{(h)}$ *be the predictable* $(\widetilde{P}, \mathcal{F}_t^{\widetilde{p}} \vee \sigma\{\widetilde{\jmath}_0\})$*-intensity of* $(\widetilde{v}_t^{(h)})$.

Let us assume that for all $t \leq T$

$$\widetilde{\mathcal{M}}_t^{(h)}(\widetilde{\omega}) = 0 \quad \textit{implies} \quad \widetilde{\mu}_{t^-}(\widetilde{\omega})\, m^{(h)}(\cdot, \widetilde{z}_{t^-}(\widetilde{\omega})) = 0 \tag{91}$$

$$\int_0^t \sum_h \widetilde{\mathcal{M}}_s^{(h)} ds \leq A_t, \tag{92}$$

where A_t *is a deterministic increasing function, and*

$$\sup_{t \leq T} \sup_{h: \widetilde{\mathcal{M}}_t^{(h)} > 0} \frac{\widetilde{\mu}_{t^-} m^{(h)}(\cdot, \widetilde{z}_{t^-})}{\widetilde{\mathcal{M}}_t^{(h)}} \leq C \qquad \widetilde{P} - a.s. \tag{93}$$

for some positive constant C.

Then (89) holds.

Proof In order to use Theorem 4 we have to find a measure $\widetilde{P}_1 \ll \widetilde{P}$, such that the local characteristics $\left(\widetilde{\lambda}_t, \widetilde{\Phi}_t(\cdot)\right)$ of $\widetilde{p}$, under $\widetilde{P}$

$$\widetilde{\lambda}_t = \sum_h \widetilde{\mathcal{M}}_t^{(h)},$$

$$\widetilde{\Phi}_t(h) = \frac{\widetilde{\mathcal{M}}_t^{(h)}}{\sum_{h'} \widetilde{\mathcal{M}}_t^{(h')}},$$

are changed under $\widetilde{P}_1$ to $\left(\widetilde{\lambda}_{1,t}, \widetilde{\Phi}_{1,t}(\cdot)\right)$, where

$$\widetilde{\lambda}_{1,t} = \widetilde{\mu}_{t^-}(\lambda_2 + \lambda_3)(\cdot, \widetilde{z}_{t^-}),$$

$$\widetilde{\Phi}_{1,t}(h) = \frac{\widetilde{\mu}_{t^-} m^{(h)}(\cdot, \widetilde{z}_{t^-})}{\widetilde{\mu}_{t^-}(\lambda_2 + \lambda_3)(\cdot, \widetilde{z}_{t^-})},$$

(here we have used the convention that the normalization of a null measure is a fixed measure, e.g. a delta function in 0).

Note that condition (91) is necessary because the measure $\widetilde{\Phi}_{1,t}$ has to be absolutely continuous with respect to $\widetilde{\Phi}_t$.

The results by Boel, Varaiya and Wong [3] and by Jacod [13] (see also Theorems T10 and T11 of Chapter VIII in [4]), imply that $\widetilde{P}_1$ can be obtained by the change of measure

$$\left.\frac{d\widetilde{P}_1}{d\widetilde{P}}\right|_{\mathcal{F}_T^{\widetilde{p}}\vee\sigma\{\widetilde{\jmath}_0\}} = \widetilde{L}_T$$

where

$$\begin{aligned}\widetilde{L}_t \quad := \quad & \prod_{i\leq \widetilde{z}_t} \left.\frac{\widetilde{\mu}_{\widetilde{S}_i^-} m^{(h)}(\cdot, i-1)}{\widetilde{\mathcal{M}}^{(h)}_{\widetilde{S}_i}}\right|_{h=\widetilde{\jmath}_{\widetilde{S}_i}} \cdot \\ & \cdot \exp\left\{\int_0^t \sum_{h'} \left(1 - \frac{\widetilde{\mu}_{s^-} m^{(h')}(\cdot, \widetilde{z}_s^-)}{\widetilde{\mathcal{M}}^{(h')}_s} \mathbb{I}_{\{\widetilde{\mathcal{M}}^{(h')}_s>0\}}\right) \widetilde{\mathcal{M}}^{(h')}_s ds\right\}\end{aligned}$$

where $\widetilde{S}_i$ are the jump times of $(\widetilde{z}_t)$.

In fact first of all the boundedness from above of the functions λ_2 and λ_3 implies the following necessary condition:

$$\int_0^t \widetilde{\mu}_{s^-}(\lambda_2+\lambda_3)(\cdot, \widetilde{z}_s)ds < \infty, \qquad \text{for all } t\leq T, \qquad \widetilde{P}-a.s.$$

Moreover (93) implies the condition

$$\frac{1}{\sum_{h'} \widetilde{\mathcal{M}}^{(h')}_t} \sum_h \frac{\left(\widetilde{\mu}_{t^-} m^{(h)}(\cdot, \widetilde{z}_{t^-})\right)^\alpha}{\left(\widetilde{\mathcal{M}}^{(h)}_t\right)^{\alpha-1}} \mathbb{I}_{\{\widetilde{\mathcal{M}}^{(h)}_s>0\}} \leq C^\alpha \qquad \widetilde{P}-a.s.$$

for some $\alpha > 1$.
Then the right hand side of the previous inequality is less or equal to

$$C_1 + C_2\left(\widetilde{z}_t + \sum_h \widetilde{\mathcal{M}}^{(h)}_t t\right)$$

The above condition together with (92) implies that $(\widetilde{L}_t)$ is a $(\widetilde{P}, \mathcal{F}_t^{\widetilde{p}}\vee\sigma\{\widetilde{\jmath}_0\})$-martingale ([3], for a proof see, T11, chapter VIII, [4]).

6 The finite partition case

In this Section we consider the case when the partition $\{\Gamma_h\}$ is finite. We start the Section with a special model, that will be used as the reference model in Theorem 4 and Theorem 5. In the reference model we still allow a denumerable state space for the process $(\widetilde{\jmath}_t)$. Besides the processes $(\widetilde{z}_t)$ and $(\widetilde{\jmath}_t)$, in the next lemma another process $(\widetilde{u}_t)$ is considered. The process $(\widetilde{u}_t)$ is inspired by the marked process whose jump times are the jump times (T_i) of the process (y_t), and the mark takes the value h when y_{T_i} belongs to Γ_h.

Lemma 3 *Let $\widetilde{\lambda}^{(l)}$ be positive constants for $l = 1, 2, 3$. Let $(N^{(1)}_{h,h'}(t), t \le T)$ and $(N^{(3)}_{h,h'}(t), t \le T)$, h and $h' \in \mathbb{Z}$, be two families of independent Poisson processes defined on a probability space $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{P})$, with parameters $\widetilde{\lambda}^{(1)}_{h,h'} = \widetilde{\lambda}^{(1)}_{h'}$ and $\widetilde{\lambda}^{(3)}_{h,h'} = \widetilde{\lambda}^{(3)}_{h'}$, such that*

$$\widetilde{\lambda}^{(l)} := \sum_{h'} \widetilde{\lambda}^{(l)}_{h'} < \infty, \qquad \text{for } l = 1, 3. \tag{94}$$

On the same probability space let $N^{(2)}_t$ be an independent Poisson process with parameter $\widetilde{\lambda}^{(2)}$.

Let $(\widetilde{u}_t), (\widetilde{z}_t), (\widetilde{\jmath}_t)$ be defined as follows

$$\widetilde{u}_t = \widetilde{\jmath}_0 + \sum_{h,h'} \int_0^t \mathbb{I}_{\{\widetilde{u}_{s^-}=h\}} (h'-h) N^{(1)}_{h,h'}(ds) + \sum_{h,h'} \int_0^t \mathbb{I}_{\{\widetilde{u}_{s^-}=h\}} (h'-h) N^{(3)}_{h,h'}(ds)$$

$$\widetilde{z}_t = N^{(2)}_t + \sum_{h,h'} \int_0^t \mathbb{I}_{\{\widetilde{u}_{s^-}=h\}} N^{(3)}_{h,h'}(ds)$$

$$\widetilde{\jmath}_t = \widetilde{\jmath}_0 + \int_0^t (\widetilde{u}_{s^-} - \widetilde{\jmath}_{s^-}) N^{(2)}(ds) + \sum_{h,h'} \int_0^t \mathbb{I}_{\{\widetilde{u}_{s^-}=h\}} (h' - \widetilde{\jmath}_{s^-}) N^{(3)}_{h,h'}(ds),$$

where $\widetilde{\jmath}_0$ is a $\mathbb{Z}$-valued random variable, independent of the Poisson processes. Let $(\widetilde{v}^{(h)}_t)$, $h = 1, 2 \dots$, be the counting processes

$$\widetilde{v}^h_t := \int_0^t \mathbb{I}_{\{\widetilde{\jmath}_s = h\}} d\widetilde{z}_s \tag{95}$$

Let $\widetilde{\mathcal{M}}^{(h)}_t$ be the $(\mathcal{F}^{\widetilde{\eta}}_t)$-intensity of $(\widetilde{v}^{(h)}_t)$, where $\widetilde{\eta}_s = (\widetilde{z}_s, \widetilde{\jmath}_s)$, and let $\{\widetilde{S}_i\}$ be the sequence of the jump time of $(\widetilde{z}_t)$ and $\widetilde{h}_i := \widetilde{\jmath}_{\widetilde{S}_i}$. Then, for $\widetilde{S}_i < t \le \widetilde{S}_{i+1}$, $\widetilde{\mathcal{M}}^{(h)}_t$ satisfies

$$\widetilde{\mathcal{M}}^{(h)}_t = \widetilde{\lambda}^{(2)} \left[e^{-\widetilde{\lambda}^{(1)}(t-\widetilde{S}_i)} + \left(1 - e^{-\widetilde{\lambda}^{(1)}(t-\widetilde{S}_i)}\right) \frac{\widetilde{\lambda}^{(1)}_h}{\widetilde{\lambda}^{(1)}} \right] + \widetilde{\lambda}^{(3)}_h, \qquad \text{for } h = \widetilde{h}_i,$$

whereas

$$\widetilde{\mathcal{M}}_t^{(h)} = \widetilde{\lambda}^{(2)} \left(1 - e^{-\widetilde{\lambda}^{(1)}(t - \widetilde{S}_i)}\right) \frac{\widetilde{\lambda}_h^{(1)}}{\widetilde{\lambda}^{(1)}} + \widetilde{\lambda}_h^{(3)}, \qquad \textit{for } h \neq \widetilde{h}_i.$$

Proof Note that at the jump times $\widetilde{S}_i$ the value $\widetilde{h}_i$ of $\widetilde{\jmath}_{\widetilde{S}_i}$ coincides with the value of $\widetilde{u}_{\widetilde{S}_i}$. Then by definition

$$\widetilde{v}_t^{(h)} = \int_0^t \mathbb{I}_{\{\widetilde{u}_{s^-} = h\}} N^{(2)}(ds) + \sum_{h'} \int_0^t \mathbb{I}_{\{\widetilde{u}_{s^-} = h'\}} N^{(3)}_{\{h',h\}}(ds),$$

then, for all $h \in \mathbb{Z}$, under $\widetilde{P}$ the predictable $(\mathcal{F}_t^{\widetilde{u},\widetilde{z},\widetilde{\jmath}})$-intensity of $(\widetilde{v}_t^{(h)})$, is given by

$$\widetilde{\lambda}^{(2)} \mathbb{I}_{\{\widetilde{u}_{t^-} = h\}} + \sum_{h'} \widetilde{\lambda}_h^{(3)} \mathbb{I}_{\{\widetilde{u}_{t^-} = h'\}} = \widetilde{\lambda}^{(2)} \mathbb{I}_{\{\widetilde{u}_{t^-} = h\}} + \widetilde{\lambda}_h^{(3)}.$$

Then

$$\widetilde{M}_t^{(h)} = \widetilde{\lambda}^{(2)} \mathbb{I}_{\{\widetilde{u}_t = h\}} + \widetilde{\lambda}_h^{(3)}$$

is the optional version of the $(\mathcal{F}_t^{\widetilde{u},\widetilde{z},\widetilde{\jmath}})$-intensity.

Let us compute the predictable $(\mathcal{F}_t^{\widetilde{z},\widetilde{\jmath}})$-intensity of $(\widetilde{v}_t^{(h)})$. For $\widetilde{S}_i < t \leq \widetilde{S}_{i+1}$

$$\widetilde{\mathcal{M}}_t^{(h)} = \lim_{s \to t^-} \mathbb{E}^{\widetilde{P}} \left[\widetilde{M}_s^{(h)} \middle| \mathcal{F}_s^{\widetilde{z},\widetilde{\jmath}} \right] = \widetilde{\lambda}^{(2)} \lim_{s \to t^-} \widetilde{P}\left(\widetilde{u}_s = h \middle| \mathcal{F}_s^{\widetilde{z},\widetilde{\jmath}} \right) + \widetilde{\lambda}_h^{(3)}.$$

In order to compute the conditional expectation in the above formula, we observe that $\mathcal{F}_t^{\widetilde{z},\widetilde{\jmath}}$ is contained in $\mathcal{F}^N_{\widetilde{S}_i} \vee \sigma\{\widetilde{\jmath}_0\} \vee \mathcal{G}$, where $\mathcal{F}_t^N$ is the filtration generated by the Poisson processes, and

$$\mathcal{G} = \sigma\{N^{(2)}(\tau) - N^{(2)}(\widetilde{S}_i), N^{(3)}_{h,h'}(\tau) - N^{(3)}_{h,h'}(\widetilde{S}_i), \text{ for } \tau \geq \widetilde{S}_i\}.$$

Then, taking into account that the evolution of $(\widetilde{u}_s)$ between the observation times is due to the processes $(N^{(1)}_{h,h'}(\tau) - N^{(1)}_{h,h'}(\widetilde{S}_i))$,

$$\widetilde{P}\left(\widetilde{u}_s = h \middle| \mathcal{F}^N_{\widetilde{S}_i} \vee \sigma\{\widetilde{\jmath}_0\} \vee \mathcal{G} \right) = \mathbb{P}\left(U_\tau = h \middle| U_0 = h' \right)\Big|_{\tau = s - \widetilde{S}_i, h' = \widetilde{h}_i} \tag{96}$$

where (U_τ) is a process with with values in $\mathbb{Z}$ and generator

$$A_U \psi(h) = \sum_{h'} \widetilde{\lambda}_{h'}^{(1)} \left(\psi(h') - \psi(h) \right).$$

Then we get the thesis, since the r.h.s. of (96) is $\mathcal{F}_s^{\widetilde{z},\widetilde{\jmath}}$-measurable, and

$$\mathbb{P}\left(U_\tau = h | U_0 = h'\right) = e^{-\widetilde{\lambda}^{(1)}\tau} + \left(1 - e^{-\widetilde{\lambda}^{(1)}\tau}\right) \frac{\widetilde{\lambda}_h^{(1)}}{\widetilde{\lambda}^{(1)}}, \qquad \text{for } h = h',$$

whereas

$$\mathbb{P}\left(U_\tau = h | U_0 = h'\right) = \left(1 - e^{-\widetilde{\lambda}^{(1)}\tau}\right) \frac{\widetilde{\lambda}_h^{(1)}}{\widetilde{\lambda}^{(1)}} \qquad \text{for } h \neq h'.$$

Remark 16 *The system considered in Lemma 3 is an example of the original system considered in this paper, without the x-component, with the y-component corresponding to $(\widetilde{u}_t)$, and the z-component to $(\widetilde{z}_t)$. Moreover there is no diffusive part, the state space of the original system is $\mathbb{Z} \times \mathbb{N}$, then the grouped information and the precise information coincide. In the proof we have computed the corresponding filter of the process $(\widetilde{u}_t)$:*

$$\widetilde{P}\left(\widetilde{u}_t = h | \mathcal{F}_s^{\widetilde{z},\widetilde{\jmath}}\right) = e^{-\widetilde{\lambda}^{(1)}(t-\widetilde{S}_i)} \mathbb{I}_{\{h=\widetilde{h}_i\}} + \left(1 - e^{-\widetilde{\lambda}^{(1)}(t-\widetilde{S}_i)}\right) \frac{\widetilde{\lambda}_h^{(1)}}{\widetilde{\lambda}^{(1)}}$$

when $t \in (\widetilde{S}_i, \widetilde{S}_{i+1})$.

Moreover, since the intensity $\widetilde{\lambda}^{(1)}$ is constant, it is easy to see that the algorithm $H(t, \chi, \mu)$ applied to this case gives the same result: for any i the generator G^i coincides with A_U, and $\widehat{\pi}_i$ is the probability concentrated in $\widetilde{h}_i$.

It is interesting to note that for all h

$$\sum_h \widetilde{\mathcal{M}}_t^{(h)} \leq \widetilde{\lambda}^{(2)} + \widetilde{\lambda}^{(3)}, \quad \text{for all } t \leq T.$$

This observation is used in the proof of the next Theorem, which deals with finite partitions. When the partition is finite, in the previous Lemma, the indexes h and h' are elements of finite set, instead of $\mathbb{Z}$. Then assumption (94) is always verified, since it involves a finite sum. In this situation we can get the expression of the filter.

Theorem 6 *Let (ξ_t, η_t) be the solution, of system (20), (21), (22), (24). Let one of the set of hypotheses of Section 2 be satisfied.*

Assume that the partition $\{\Gamma_h\}$ is finite, and that

$$0 \leq \underline{\lambda}_l \leq \lambda_l(\xi, z) < \overline{\lambda}_l, \quad l = 1, 2, 3.$$

Then the filter (π_t) defined by (12) satisfies equality (89).

Proof We use the reference model introduced in the previous lemma with suitable parameters when the random variable $\widetilde{\jmath}_0$ has the distribution of j_0, as in Theorem 4, and then we apply Theorem 5. The value of the parameters $\widetilde{\lambda}_{h,h'}^{(1)}$ and $\widetilde{\lambda}^{(2)}$ is arbitrary, while we take

$$\widetilde{\lambda}_h^{(3)} = \overline{\lambda}_2 + \sup_{\xi,z} \lambda_3^h(\xi, z).$$

The hypotheses of Theorem 5 are verified:
Condition (91) is trivial, since with the above choice $\widetilde{\mathcal{M}}_t^{(h)}$ is always positive:

$$\widetilde{\mathcal{M}}_t^{(h)} \geq \widetilde{\lambda}_h^{(1)} C + \widetilde{\lambda}_h^{(3)} \geq \overline{\lambda}_2$$

where $C = (1 - e^{-\widetilde{\lambda}^{(1)}T})\widetilde{\lambda}^{(2)}/\widetilde{\lambda}^{(1)}$ is a constant depending on T.

Hypothesis (92) is verified with $A_t = \left(\widetilde{\lambda}_2 + \widetilde{\lambda}_3\right) t$.

Hypothesis (93) is verified since with the above choice, for all h and $t \leq T$

$$\frac{\widetilde{\mu}_{t^-} m^{(h)}(\cdot, \widetilde{z}_{t^-})}{\widetilde{\mathcal{M}}_t^{(h)}} \leq 1. \tag{97}$$

Indeed, in view of definition (88), for $\widetilde{S}_i < t \leq \widetilde{S}_{i+1}$, the numerator can be expressed in terms of the sequence of processes $\{(\widetilde{X}_u^i, \widetilde{Y}_u^i)_{u \geq 0}\}$, i.e. the sequence $\{(X_u^i(\chi), Y_u^i(\chi))_{u \geq 0}\}$ of Markov processes defined in Section 4, with generator G_i (102), and initial conditions $\widehat{\pi}_i(\chi)$ (defined in (71), (73)), for $\chi = \theta(\widetilde{\eta}_\cdot)$). Using the short notation $\widetilde{\Xi}_u^i = (\widetilde{X}_u^i, \widetilde{Y}_u^i)$ the numerator of (97) is

$$\mathbb{E}\left[\left(\mathbb{I}_{\{\widetilde{Y}_{t-s}^i \in \Gamma_h\}} \lambda_2(\widetilde{\Xi}_{t-s}^i, i) + \lambda_3^h(\widetilde{\Xi}_{t-s}^i, i)\right) \exp\left\{-\int_0^{t-s} (\lambda_2 + \lambda_3)(\widetilde{\Xi}_u^i, i)\, du\right\}\right]\Big|_{s=\widetilde{S}_i},$$

and is bounded above by

$$\left(\overline{\lambda}_2\, \mathbb{P}(\widetilde{Y}_{t-s}^i \in \Gamma_h)\Big|_{s=\widetilde{S}_i} + \sup_{\xi,z} \lambda_3^h(\xi, z)\right) \exp\left\{-(\underline{\lambda}_2 + \underline{\lambda}_3)(t - \widetilde{S}_i)\right\}.$$

As already observed, the denominator of (97) is bounded from below, and therefore

$$\begin{aligned}
\frac{\widetilde{\mu}_{t^-} m^{(h)}(\cdot, \widetilde{z}_{t^-})}{\widetilde{\mathcal{M}}_t^{(h)}} &\leq \frac{\overline{\lambda}_2\, \mathbb{P}(\widetilde{Y}_{t-s}^i \in \Gamma_h)\Big|_{s=\widetilde{S}_i} + \sup_{\xi,z} \lambda_3^h(\xi, z)}{\widetilde{\lambda}_h^{(1)} C + \widetilde{\lambda}_h^{(3)}} \\
&\leq \frac{\overline{\lambda}_2 + \sup_{\xi,z} \lambda_3^h(\xi, z)}{\widetilde{\lambda}_h^{(3)}} = 1.
\end{aligned}$$

Remark 17 *When the partition is not finite the previous proof does not work since, with the above choice, the condition $\sum_h \widetilde{\lambda}_h^{(3)} < \infty$ does not hold. The choice is due to the rough upper bound used in the last step of the proof. In order to consider the infinite case, one could try to estimate $\mathbb{P}(\widetilde{Y}_{t-s}^i \in \Gamma_h)\Big|_{s=\widetilde{S}_i}$ and $\mathbb{E}\left[\lambda_3^h(\widetilde{\Xi}_{t-s}^i, i)\right]\Big|_{s=\widetilde{S}_i}$ and then use the inequality*

$$\widetilde{\mu}_{t^-} m^{(h)}(\cdot, \widetilde{z}_{t^-}) \leq$$
$$\left(\overline{\lambda}_2 \mathbb{P}(\widetilde{Y}_{t-s}^i \in \Gamma_h) + \mathbb{E}\left[\lambda_3^h(\widetilde{\Xi}_{t-s}^i, i)\right]\right)\Big|_{s=\widetilde{S}_i} \exp\left\{-(\underline{\lambda}_2 + \underline{\lambda}_3)(t - \widetilde{S}_i)\right\}.$$

Maybe, with a sharp upper bound, there would be a chance to choose $\widetilde{\lambda}_h^{(3)}$ in such a way that the corresponding series is converging.

Remark 18 *In Theorem 6 we use the assumption that the infinitesimal parameters of the jumps of the* (y, z)*-component in the original system are uniformly bounded. Note that instead, for the* x*-component we allow sublinear growth conditions. As explained in Remark 1 we may assume that the* y*-component belongs to a discrete subset* $\mathcal{Y}$*. In the particular case when* $\mathcal{Y}$ *is finite and the partition is generated by the singletons of* $\mathcal{Y}$*, the above Theorem gives the representation of the filter when the observation is precise (see Remark 2).*

7 The case when (y_t) is counting process

As already explained in Remark 1, the case when (y_t) is counting process corresponds to the case when $\beta = 0$, $\gamma = 0$, $y_0 \in \mathbb{N}$, K_1 and K_3 take values in $\{0, 1\}$. Then the state space is $\mathbb{N}$. In this Section we assume moreover that the partition of the state space $\mathbb{N}$ is given by $\Gamma_h = \{h\}$, so that there is no difference between the *precise* and the *grouped* information (see Remark 2).

We start by introducing the *reference model* which we use in this case i.e. the probability space $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{P})$ and the marked point process $\widetilde{p}$, defined by means of a sequence of counting processes $(\widetilde{v}_t^{(h)})$, $h = 1, 2 \ldots$, as in the previous Section.

Lemma 4 *Let $\lambda^{(l)}$ be positive constants for $l = 1, 2, 3$. Let $(N_t^{(l)})$ be independent Poisson processes defined on a probability space $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{P})$, with parameters $\lambda^{(l)}$, for $l = 1, 2, 3$.*

Let $(\widetilde{u}_t), (\widetilde{z}_t), (\widetilde{\jmath}_t)$ be defined as follows

$$\widetilde{u}_t = \widetilde{\jmath}_0 + N_t^{(1)} + N_t^{(3)}$$

$$\widetilde{z}_t = N_t^{(2)} + N_t^{(3)}$$

$$\widetilde{\jmath}_t = \widetilde{\jmath}_0 + \int_0^t (\widetilde{u}_{s^-} - \widetilde{\jmath}_{s^-}) N^{(2)}(ds) + \int_0^t (\widetilde{u}_{s^-} + 1 - \widetilde{\jmath}_{s^-}) N^{(3)}(ds),$$

where $\widetilde{\jmath}_0$ is a $\mathbb{N}$-valued random variable independent of the Poisson processes. Let $(\widetilde{v}_t^{(h)})$, $h = 1, 2 \ldots$ be the counting processes defined as in (95). Let $\widetilde{\mathcal{M}}_t^{(h)}$ be the $(\mathcal{F}_t^{\widetilde{\eta}})$-intensity of $(\widetilde{v}_t^{(h)})$, where $\widetilde{\eta}_s = (\widetilde{z}_s, \widetilde{\jmath}_s)$, and let $\{\widetilde{S}_i\}$ be the sequence of the jump times of $(\widetilde{z}_t)$ and $\widetilde{h}_i := \widetilde{\jmath}_{\widetilde{S}_i}$. Then, for $\widetilde{S}_i < t \le \widetilde{S}_{i+1}$, $\widetilde{\mathcal{M}}_t^{(h)}$ satisfies

$$\widetilde{\mathcal{M}}_t^{(h)} = \lambda^{(2)} \mathbb{I}_{\{h - \widetilde{h}_i \ge 0\}} \frac{((t - \widetilde{S}_i)\lambda^{(1)})^{h - \widetilde{h}_i}}{(h - \widetilde{h}_i)!} e^{-\lambda^{(1)}(t - \widetilde{S}_i)} + \lambda^{(3)} \mathbb{I}_{\{h - \widetilde{h}_i - 1 \ge 0\}} \frac{((t - \widetilde{S}_i)\lambda^{(1)})^{h - \widetilde{h}_i - 1}}{(h - \widetilde{h}_i - 1)!} e^{-\lambda^{(1)}(t - \widetilde{S}_i)}. \tag{98}$$

Proof The proof is similar to that of Lemma 3, and could even be deduced by it (see Remark 19 below). Since by definition

$$\widetilde{v}_t^{(h)} = \int_0^t \mathbb{I}_{\{\widetilde{u}_{s^-} = h\}} N^{(2)}(ds) + \int_0^t \mathbb{I}_{\{\widetilde{u}_{s^-} + 1 = h\}} N^{(3)}(ds),$$

then for all $h \in \mathbb{Z}$ the optional $(\mathcal{F}_t^{\widetilde{u}, \widetilde{z}, \widetilde{\jmath}})$-intensity of $(\widetilde{v}_t^{(h)})$, is given by

$$\widetilde{M}_t^{(h)} = \lambda^{(2)} \mathbb{I}_{\{\widetilde{u}_t = h\}} + \lambda^{(3)} \mathbb{I}_{\{\widetilde{u}_t + 1 = h\}}.$$

The result then follows observing that

$$\widetilde{\mathcal{M}}_t^{(h)} = \lim_{s\to t^-} \mathbb{E}^{\widetilde{P}}\left[M_s^{(h)} \middle| \mathcal{F}_s^{\widetilde{z},\widetilde{\jmath}} \right]$$

$$= \lambda^{(2)} \lim_{s\to t^-} \widetilde{P}\left(\widetilde{u}_s = h | \mathcal{F}_s^{\widetilde{z},\widetilde{\jmath}} \right) + \lambda^{(3)} \lim_{s\to t^-} \widetilde{P}\left(\widetilde{u}_s = h-1 | \mathcal{F}_s^{\widetilde{z},\widetilde{\jmath}} \right)$$

and that for $\widetilde{S}_i < s < \widetilde{S}_{i+1}$

$$\begin{aligned}
\widetilde{P}\left(\widetilde{u}_s = h | \mathcal{F}_s^{\widetilde{z},\widetilde{\jmath}} \right) &= \mathbb{E}^{\widetilde{P}}\left[\widetilde{P}\left(\widetilde{u}_s = h | \mathcal{F}_s^{\widetilde{z},\widetilde{\jmath}} \vee \mathcal{F}_{\widetilde{S}_i}^{N^{(1)}} \right) \middle| \mathcal{F}_s^{\widetilde{z},\widetilde{\jmath}} \right] \\
&= \mathbb{E}^{\widetilde{P}}\left[\widetilde{P}\left(N_s^{(1)} - N_{\widetilde{S}_i}^{(1)} = h - \widetilde{h}_i \middle| \mathcal{F}_s^{\widetilde{z},\widetilde{\jmath}} \vee \mathcal{F}_{\widetilde{S}_i}^{N^{(1)}} \right) \middle| \mathcal{F}_s^{\widetilde{z},\widetilde{\jmath}} \right],
\end{aligned}$$

and finally that

$$\widetilde{P}(N_s^{(1)} - N_{\widetilde{S}_i}^{(1)} = h - \widetilde{h}_i | \mathcal{F}_s^{\widetilde{z},\widetilde{\jmath}} \vee \mathcal{F}_{\widetilde{S}_i}^{N^{(1)}}) = \frac{((s-\widetilde{S}_i)\lambda^{(1)})^{h-\widetilde{h}_i}}{(h-\widetilde{h}_i)!} e^{-\lambda^{(1)}(s-\widetilde{S}_i)}.$$

Remark 19 *From the proof of the Lemma 4 we get that the corresponding filter of the process* $(\widetilde{u}_t)$ *is*

$$\widetilde{P}\left(\widetilde{u}_t = h | \mathcal{F}_t^{\widetilde{z},\widetilde{\jmath}} \right) = \frac{((t-\widetilde{S}_i)\lambda^{(1)})^{h-\widetilde{h}_i}}{(h-\widetilde{h}_i)!} e^{-\lambda^{(1)}(t-\widetilde{S}_i)}, \qquad h \geq \widetilde{h}_i,$$

when $t \in (\widetilde{S}_i, \widetilde{S}_{i+1})$.

Theorem 7 *Let* (ξ_t, η_t) *be the solution, of system (20), (21), (22), (24). Let one of the set of hypotheses of Section 2 be satisfied. In the standing conditions of this Section, assume that*

$$0 \leq \underline{\lambda}_l \leq \lambda_l(\xi, z) < \overline{\lambda}_l, \quad l = 1, 2, 3.$$

Then the filter (π_t) *defined by (12) satisfies equality* (89).

Proof We show that the reference model introduced in the previous lemma with $\lambda^{(l)} = \overline{\lambda}_l, l = 1, 2, 3$, satisfies the hypotheses of Theorem 5. Hypothesis (92) is verified, since $\sum_h \widetilde{\mathcal{M}}_t^{(h)} = \lambda^{(2)} + \lambda^{(3)}$, for all $t \leq T$. Hypothesis (93) is verified since,

$$\frac{\widetilde{\mu}_{t^-} m^{(h)}(\cdot, \widetilde{z}_{t^-})}{\widetilde{\mathcal{M}}_t^{(h)}} \leq e^{(\overline{\lambda}_1 - \underline{\lambda}_1 - \underline{\lambda}_2 - \underline{\lambda}_3)T}, \tag{99}$$

for all $t \leq T$ and for all h such that $\widetilde{\mathcal{M}}_t^{(h)} > 0$.

First of all, in order to prove (99), we recall that, for $\widetilde{S}_i < t \le \widetilde{S}_{i+1}$, the intensity $\widetilde{\mathcal{M}}_t^{(h)}$ is different from zero if and only if $h \ge \widetilde{h}_i$. Let us observe that, when (y_t) is a counting process, definition (47) changes into

$$m^{(h)}(\xi, z) := \mathbb{I}_{\{y=h\}} \lambda_2(\xi, z) + \mathbb{I}_{\{y+1=h\}} \lambda_3(\xi, z).$$

Using the notation $(\widetilde{X}_u^i, \widetilde{Y}_u^i)$ as in the proof of Theorem 6, for $\widetilde{S}_i < t \le \widetilde{S}_{i+1}$, the numerator of the right hand side of inequality (99) is equal to

$$\mathbb{E}\left[m^{(h)}(\widetilde{X}_\tau^i, \widetilde{Y}_\tau^i, i) exp\{-\int_0^\tau (\lambda_2 + \lambda_3)(\widetilde{X}_u^i, \widetilde{Y}_u^i, i)\, du\} \right]\Bigg|_{\tau = t - \widetilde{S}_i}. \tag{100}$$

Since

$$m^{(h)}(\widetilde{X}_\tau^i, \widetilde{Y}_\tau^i, i) = \mathbb{I}_{\{\widetilde{Y}_\tau^i = h\}} \lambda_2(\widetilde{X}_\tau^i, \widetilde{Y}_\tau^i, i) + \mathbb{I}_{\{\widetilde{Y}_\tau^i + 1 = h\}} \lambda_3(\widetilde{X}_\tau^i, \widetilde{Y}_\tau^i, i),$$

and $\lambda_i(\xi, z) < \overline{\lambda}_i$, then expression (100) is bounded above by

$$\Big(\overline{\lambda}_2\, \mathbb{P}(\widetilde{Y}_\tau^i = h) + \overline{\lambda}_3\, \mathbb{P}(\widetilde{Y}_\tau^i = h-1)\Big)\Big|_{\tau = t - \widetilde{S}_i}\ e^{-(\underline{\lambda}_2 + \underline{\lambda}_3)(t - \widetilde{S}_i)}.$$

Being $\widetilde{Y}_0^i = \widetilde{h}_i$, then for any $k \in \mathbb{Z}$

$$\mathbb{P}\left(\widetilde{Y}_\tau^i = k\right) = \mathbb{P}\left(\widetilde{Y}_\tau^i - \widetilde{Y}_0^i = k - \widetilde{h}_i\right),$$

moreover

$$\mathbb{P}\left(\widetilde{Y}_\tau^i - \widetilde{Y}_0^i = k\right) \le \mathbb{I}_{\{k \ge 0\}} \frac{(\tau \overline{\lambda}_1)^k}{k!}\, e^{-\underline{\lambda}_1 \tau}. \tag{101}$$

Since $\underline{\lambda}_1 \le \lambda_1(\xi, z) \le \overline{\lambda}_1$, the above inequality can be easily proved by Girsanov Theorem (see Remark 20, below)

Finally we have that $\widetilde{\mu}_{t^-} m^{(h)}(\cdot, \widetilde{z}_{t^-})$ is bounded above by

$$\left(\overline{\lambda}_2\, \mathbb{I}_{\{h - \widetilde{h}_i \ge 0\}} \frac{((t - \widetilde{S}_i)\overline{\lambda}_1)^{h - \widetilde{h}_i}}{(h - \widetilde{h}_i)!} + \overline{\lambda}_3\, \mathbb{I}_{\{h - \widetilde{h}_i \ge 1\}} \frac{((t - \widetilde{S}_i)\overline{\lambda}_1)^{h - \widetilde{h}_i - 1}}{(h - \widetilde{h}_i - 1)!} \right) e^{-(\underline{\lambda}_1 + \underline{\lambda}_2 + \underline{\lambda}_3)(t - \widetilde{S}_i)}.$$

On the other hand, if $\lambda^{(l)} = \overline{\lambda}_l, l = 1, 2, 3$, then the $(\mathcal{F}_t^{\widetilde{z}, \widetilde{\jmath}})$-intensity $\widetilde{\mathcal{M}}_t^{(h)}$ (see (100)) if $\lambda^{(l)} = \overline{\lambda}_l, l = 1, 2, 3$, can be bounded below by the quantity in parentheses in the above bound multiplied by

$$e^{-\overline{\lambda}_1 (t - \widetilde{S}_i)}.$$

So the inequality (99) can be easily proved.

Remark 20 *Observe that the generator G^i can be rewritten as*

$$\begin{aligned} G^i \phi(\xi) \quad &= \quad B(\xi, i)\nabla_\xi \phi(\xi) + \frac{1}{2} tr\{C(\xi, i) C^*(\xi, i) \nabla_x^2 \phi(\xi)\} \qquad (102) \\ &+ \quad \lambda_0(x, y, i) \int_{\mathbb{R}^d} [\phi(x', y) - \phi(x, y)]\, \mu^{(0)}(x, y, i; dx') \\ &+ \quad \lambda_1(x, y, i) \int_{\mathbb{R}^d} [\phi(x', y+1) - \phi(x, y)]\, \mu^{(1)}(x, y, i; dx') \end{aligned}$$

where $\lambda_l(x,y,i)\mu^{(l)}(x,y,i;dx')$ *are the measures induced by* ν *restricted to* $D_l(x,y,i)$ *by means of the transformation* $\zeta \to x + K_l(x,y,i;\zeta)$*.*

Suppose that $(\Omega, \mathcal{F}, \mathbb{Q})$ *is a probability space, that* $(\widetilde{X}^i_u, \widetilde{Y}^i_u)$ *is a Markov process with initial distribution* $\widehat{\pi}_i$ *and generator obtained by* G^i *substituting* $\lambda_1(x,y,i)$ *in the last addend with* 1*.*

Then, under $\mathbb{Q}$*,* $\widetilde{Y}^i_\tau - \widetilde{Y}^i_0$ *is a standard Poisson process, with jump times* $\widetilde{T}_j$*. If we define the probability* $\mathbb{P}$ *as*

$$\left.\frac{d\mathbb{P}}{d\mathbb{Q}}\right|_{\mathcal{F}_\tau^{\widetilde{X}^i,\widetilde{Y}^i}} = \prod_{\widetilde{T}_j \le \tau} \lambda_1(\widetilde{X}^i_{\widetilde{T}_j}, \widetilde{Y}^i_{\widetilde{T}_j}, i) \exp\left\{ \int_0^\tau \left(1 - \lambda_1(\widetilde{X}^i_u, \widetilde{Y}^i_u, i)\right) du \right\},$$

then

$$\mathbb{P}\left(\widetilde{Y}^i_\tau - \widetilde{Y}^i_0 = k\right) = \int_{\{\widetilde{Y}^i_\tau - \widetilde{Y}^i_0 = k\}} \left.\frac{d\mathbb{P}}{d\mathbb{Q}}\right|_{\mathcal{F}_\tau^{\widetilde{X}^i,\widetilde{Y}^i}} d\mathbb{Q}$$

$$\le \left(\overline{\lambda}_1\right)^k \exp\left\{(1 - \underline{\lambda}_1)\ \tau\right\} \mathbb{Q}\left(\widetilde{Y}^i_\tau - \widetilde{Y}^i_0 = k\right),$$

and (101) follows.

8 An application to queues

In this section we discuss the queue example of the Introduction. We recall that we need to overcome the problem that we are able to detect which kind of jump we observe.

We could redo every step taking this fact into account, and similar techniques would apply. But we can rewrite the example in our framework, by adding an auxiliary component to the system. In a more precise sense we assume that at the jump times we can observe another component. Before we define the new component we recall that

$$\begin{aligned} Q_t &= Q_0 + A_t - \int_0^t \mathbb{I}_{\{Q_{s^-}\neq 0\}} dN_s \\ Z_t &= \int_0^t \mathbb{I}_{\{Q_{s^-}\in\{0\}\cup\{h\}\}} dA_s + \int_0^t \mathbb{I}_{\{Q_{s^-}\neq 0\}} dN_s. \end{aligned}$$

At time 0 we can observe whether $Q_0 = 0$, $1 \leq Q_0 \leq h$ or $Q_0 > h$. We can set $\Sigma := \{+1, -1\}$, the Poisson measure $\mathcal{N}(ds, +1) := A(ds)$, and $\mathcal{N}(ds, -1) := N(ds)$. Then the mean measure of the Poisson measure is $ds \times \nu(d\zeta)$, where $\nu\{1\} = \lambda$ and $\nu\{-1\} = \mu$. Using the above notation we can rewrite the queue and the counting process of the observation times as

$$\begin{aligned} Q_t &= Q_0 + \int_0^t \mathcal{N}(ds, +1) - \int_0^t \mathbb{I}_{\{Q_{s^-}\neq 0\}} \mathcal{N}(ds, -1) \\ &= Q_0 + \int_0^t \int_\Sigma \Big(\mathbb{I}_{\{\zeta=+1\}} - \mathbb{I}_{\{Q_{s^-}\neq 0\}}\mathbb{I}_{\{\zeta=-1\}}\Big) \mathcal{N}(ds, d\zeta) \\ Z_t &= \int_0^t \mathbb{I}_{\{Q_{s^-}\in\{0\}\cup\{h\}\}} \mathcal{N}(ds, +1) + \int_0^t \mathbb{I}_{\{Q_{s^-}\neq 0\}} \mathcal{N}(ds, -1) \\ &= \int_0^t \int_\Sigma \Big(\mathbb{I}_{\{Q_{s^-}\in\{0\}\cup\{h\}\}}\mathbb{I}_{\{\zeta=+1\}} + \mathbb{I}_{\{Q_{s^-}\neq 0\}}\mathbb{I}_{\{\zeta=-1\}}\Big) \mathcal{N}(ds, d\zeta) \end{aligned}$$

As already observed in the Introduction, we do not have the $x-$component. Then we can assume that $y_t^{(1)} = Q_t$. The observation times $\{S_i\}$ of (Z_t) form a subset of the jump times $\{T_j\}$ of (Q_t). At the observation time S_i the information can be expressed in terms of Q_{S_i} and of $Q_{S_i^-}$ (see below), and then it is natural to set

$$Y_t = Y_0 \quad \text{for } t \in [T_0, T_1), \qquad Y_t := Q_{T_{j-1}} \quad \text{for } t \in [T_j, T_{j+1}) \quad \text{for } j \geq 1.$$

as the new component y_t^0. Note that

$$\mathcal{Y} = \{(l', l), \text{ for } l' = 0, 1, \cdots \text{ and } l = 0, 1, \cdots, \text{ with } |l - l'| = 1\},$$

is an invariant set for $y_t = (y_t^0, y_t^1) = (Y_t, Q_t)$. Then a convenient choice for Y_0 is $Q_0 + 1$, so that $|Y_t - Q_t| = 1$ for any t, so that $\mathcal{Y}$ can be taken as its state space.

Then

$$Y_t = Y_0 + \int_0^t \int_\Sigma (Q_{s^-} - Y_{s^-})\, dQ_s$$

or equivalently

$$Y_t = Y_0 + \int_0^t \int_\Sigma (Q_{s^-} - Y_{s^-}) \left(\mathbb{I}_{\{\zeta=+1\}} + \mathbb{I}_{\{Q_{s^-} \neq 0\}} \mathbb{I}_{\{\zeta=-1\}} \right) \mathcal{N}(ds, d\zeta).$$

With the above notations we can rewrite

$$\begin{aligned}
Y_t &= Y_0 + \int_0^t \int_\Sigma (Q_{s^-} - Y_{s^-}) \left(\mathbb{I}_{\{\zeta=+1\}} + \mathbb{I}_{\{Q_{s^-} \neq 0\}} \mathbb{I}_{\{\zeta=-1\}} \right) \mathcal{N}(ds, d\zeta) \\
Q_t &= Q_0 + \int_0^t \int_\Sigma \left(\mathbb{I}_{\{\zeta=+1\}} - \mathbb{I}_{\{Q_{s^-} \neq 0\}} \mathbb{I}_{\{\zeta=-1\}} \right) \mathcal{N}(ds, d\zeta) \\
Z_t &= \int_0^t \int_\Sigma \left(\mathbb{I}_{\{Q_{s^-} \in \{0\} \cup \{h\}\}} \mathbb{I}_{\{\zeta=+1\}} + \mathbb{I}_{\{Q_{s^-} \neq 0\}} \mathbb{I}_{\{\zeta=-1\}} \right) \mathcal{N}(ds, d\zeta)
\end{aligned}$$

We need to define also the partition. In view of the possible kind of observations and taking into account that $y_{S_i} = (Y_{S_i}, Q_{S_i}) = (Q_{S_i^-}, Q_{S_i})$

0. $Q_{S_i} = 0$.

Then $Q_{S_i^-} = 1$ and the jump is due to a departure.

It is natural to define $\Gamma_0 = \{(1, 0)\}$

1. $Q_{S_i} \in \{1, \cdots, h\}$, and $Q_{S_i^-} = 0$.

Then $Q_{S_i} = 1$ and the jump is due to an arrival.

It is natural to define $\Gamma_1 = \{(0, 1)\}$.

2. $Q_{S_i} \in \{1, \cdots, h\}$, and $Q_{S_i^-} \in \{1, \cdots, h\}$.

Then the jump is due to a departure, and $Q_{S_i} = Q_{S_i^-} - 1$.

It is natural to define

$\Gamma_2 = \{(l', l),\ \text{for } l' = 2, \cdots, h \text{ and } l = 1, \cdots, h-1,\ \text{with } l = l' - 1\}$.

3. $Q_{S_i} \in \{1, \cdots, h\}$, and $Q_{S_i^-} \in \{h + 1, \ldots\}$.

Then the jump is due to a departure $Q_{S_i} = h$ and $Q_{S_i^-} = h + 1$.

It is natural to define $\Gamma_3 = \{(h + 1, h)\}$

4. $Q_{S_i} \in \{h + 1, \cdots\}$, and $Q_{S_i^-} \in \{1, \cdots, h\}$.

Then the jump is due to an arrival $Q_{S_i} = h + 1$ and $Q_{S_i^-} = h$.

It is natural to define $\Gamma_4 = \{(h, h + 1)\}$.

5. $Q_{S_i} \in \{h+1, \cdots\}$, and $Q_{S_i^-} \in \{h+1, \cdots\}$.

Then the jump is due to a departure.

It is natural to define

$\Gamma_5 = \{(l', l),\ \text{for}\ l' = h+2, \cdots\ \text{and}\ l = h+1, \cdots, \text{with}\ l = l' - 1\}$.

We can therefore consider the partition of $\mathcal{Y}$ generated by the sets $\{\Gamma_j\}$, for $j = 0, \cdots, 5$.

So far we have just written the above problem in our framework.

We now switch to a simpler case: the case when there is no waiting room, i.e. $h = 0$, and we will give the explicit expression of the filter. In this case the partition is simpler

0. $Q_{S_i} = 0$.

Then $Q_{S_i^-} = 1$, and the jump is due to a departure. It is natural to define $\Gamma_0 = \{(1, 0)\}$

1. $Q_{S_i} \in \{1, \cdots\}$, and $Q_{S_i^-} = 0$.

Then $Q_{S_i} = 1$ and the jump is due to an arrival.

It is natural to define $\Gamma_1 = \{(0, 1)\}$.

2. $Q_{S_i} \in \{1, \cdots\}$, and $Q_{S_i^-} \in \{1, \cdots\}$.

Then the jump is due to a departure, and $Q_{S_i} = Q_{S_i^-} - 1$.

It is natural to define

$\Gamma_2 = \{(l', l),\ \text{for}\ l' = 2, \cdots\ \text{and}\ l = 1, \cdots\ \text{with}\ l = l' - 1\}$.

Moreover,

$$\begin{aligned} Q_t &= Q_0 + A_t - \int_0^t \mathbb{I}_{\{Q_{s^-} \geq 1\}} dN_s \\ Z_t &= \int_0^t \mathbb{I}_{\{Q_{s^-} = 0\}} dA_s + \int_0^t \mathbb{I}_{\{Q_{s^-} \geq 1\}} dN_s. \end{aligned}$$

With the previous notations we can rewrite

$$\begin{aligned} Y_t &= Y_0 + \int_0^t \int_\Sigma (Q_{s^-} - Y_{s^-}) \left(\mathbb{I}_{\{Q_{s^-}=0\}} \mathbb{I}_{\{\zeta=+1\}} + \mathbb{I}_{\{Q_{s^-} \geq 1\}} \mathbb{I}_{\{\zeta=-1\}} \right) \mathcal{N}(ds, d\zeta) \\ Q_t &= Q_0 + \int_0^t \int_\Sigma \left(\mathbb{I}_{\{\zeta=+1\}} - \mathbb{I}_{\{Q_{s^-} \geq 1\}} \mathbb{I}_{\{\zeta=-1\}} \right) \mathcal{N}(ds, d\zeta) \\ Z_t &= \int_0^t \int_\Sigma \left(\mathbb{I}_{\{Q_{s^-}=0\}} \mathbb{I}_{\{\zeta=+1\}} + \mathbb{I}_{\{Q_{s^-} \geq 1\}} \mathbb{I}_{\{\zeta=-1\}} \right) \mathcal{N}(ds, d\zeta) \end{aligned}$$

and furthermore we can rewrite

$$
\begin{aligned}
Y_t &= Y_0 + \int_0^t \int_\Sigma (Q_{s^-} - Y_{s^-}) \Big(\mathbb{I}_{\{Q_{s^-}=0\}} \, \mathbb{I}_{\{\zeta=+1\}} \\
&\qquad + \mathbb{I}_{\{Q_{s^-}=1\}} \mathbb{I}_{\{\zeta=-1\}} + \mathbb{I}_{\{Q_{s^-}>1\}} \mathbb{I}_{\{\zeta=-1\}} \Big) \mathcal{N}(ds, d\zeta) \\
Q_t &= Q_0 + \int_0^t \int_\Sigma \Big(\mathbb{I}_{\{Q_{s^-}=0\}} \, \mathbb{I}_{\{\zeta=+1\}} + \mathbb{I}_{\{Q_{s^-}\geq 1\}} \, \mathbb{I}_{\{\zeta=+1\}} \\
&\qquad - \mathbb{I}_{\{Q_{s^-}=1\}} \mathbb{I}_{\{\zeta=-1\}} - \mathbb{I}_{\{Q_{s^-}>1\}} \mathbb{I}_{\{\zeta=-1\}} \Big) \mathcal{N}(ds, d\zeta) \\
Z_t &= \int_0^t \int_\Sigma \Big(\mathbb{I}_{\{Q_{s^-}=0\}} \mathbb{I}_{\{\zeta=+1\}} \\
&\qquad + \mathbb{I}_{\{Q_{s^-}=1\}} \mathbb{I}_{\{\zeta=-1\}} + \mathbb{I}_{\{Q_{s^-}>1\}} \mathbb{I}_{\{\zeta=-1\}} \Big) \mathcal{N}(ds, d\zeta)
\end{aligned}
$$

The process (J_t) can then be defined as follows:

$$
\begin{aligned}
J_t &= J_0 + \int_0^t \int_\Sigma \Big([1 - J_{s^-}] \, \mathbb{I}_{\{Q_{s^-}=0\}} \, \mathbb{I}_{\{\zeta=+1\}} \\
&\qquad + [0 - J_{s^-}] \, \mathbb{I}_{\{Q_{s^-}=1\}} \, \mathbb{I}_{\{\zeta=-1\}} \\
&\qquad + [2 - J_{s^-}] \, \mathbb{I}_{\{Q_{s^-}>1\}} \, \mathbb{I}_{\{\zeta=-1\}} \Big) \mathcal{N}(ds, d\zeta)
\end{aligned}
$$

where $J_0 = 0$ when $Q_0 = 0$, and $J_0 = 2$ when $Q_0 \geq 1$.

Then we can consider the process (Y_t, Q_t, Z_t, J_t) solution of the enlarged system as a process with values in $\mathcal{Y} \times \mathbb{N} \times \{0, 1, 2\}$.

The generator of the enlarged system is

$$
\begin{aligned}
\mathbb{A}F(y, q, z, j) &= \lambda \, \mathbb{I}_{\{q\geq 1\}} \left[F(q, q+1, z, j) - F(y, q, z, j) \right] \\
&+ \lambda \, \mathbb{I}_{\{q=0\}} \left[F(q, q+1, z+1, 1)) - F(y, q, z, j) \right] \\
&+ \mu \, \mathbb{I}_{\{q=1\}} \left[F(q, q-1, z+1, 0) - F(y, q, z, j) \right] \\
&+ \mu \, \mathbb{I}_{\{q>1\}} \left[F(q, q-1, z+1, 2) - F(y, q, z, j) \right],
\end{aligned}
$$

and initial distribution depending only on the random variable Q_0 as explained above $(Y_0 = Q_0 + 1, Z_0 = 0, J_0 = 2 \, \mathbb{I}_{\{Q_0 \geq 1\}})$.

Then, for all i, the generators $G^i = G$, where

$$
G\phi(y, q) = \lambda \, \mathbb{I}_{\{q \geq 1\}} \left[\phi(q, q+1) - \phi(y, q) \right].
$$

Then $(1, 0)$ is an absorbing state, while, when starting from another state, (Y^i_τ, Q^i_τ) are such that $Y^i_\tau = Q^i_\tau - 1$ and (Q^i_τ) grows like a Poisson process.

The updating operators can be easily computed from the above generator $\mathbb{A}$.

In this case

0. when $J_{S_i} = 0$, (so that $(Y_{S_i}, Q_{S_i}) = (1, 0)$, or equivalently $(Y_{S_i}, Q_{S_i}) \in \Gamma_0$), then clearly, for $S_i \le t < S_{i+1}$, we have that $(Y_t, Q_t) = (1, 0)$ and then, at time t, the conditional distribution π_t of the pair is concentrated in $(1, 0)$.

0a. then necessarily $(Y_{S_{i+1}}, Q_{S_{i+1}}) = (0, 1)$ and then, at time S_{i+1}, the conditional distribution $\pi_{S_{i+1}}$ of the pair is concentrated in $(0, 1)$.

1. when $J_{S_i} = 1$, (so that $(Y_{S_i}, Q_{S_i}) = (0, 1)$, or equivalently $(Y_{S_i}, Q_{S_i}) \in \Gamma_1$), then, at time S_i, the conditional distribution π_{S_i} of the pair is concentrated in $(0, 1)$, and more generally, for $S_i \le t < S_{i+1}$, the conditional distribution π_t of (Y_t, Q_t) is the distribution of $\left(A^i_\tau, 1 + A^i_\tau\right)\big|_{\tau=t-S_i}$, where (A^i_τ) is a Poisson process with parameter λ.

2. when $J_{S_i} = 2$, (or equivalently $(Y_{S_i}, Q_{S_i}) \in \Gamma_2$), then $Y_{S_i} = Q_{S_i} + 1$, the conditional distribution π^q_t of Q_t at time $t = S_i$ is concentrated in $\{1, \ldots\}$ and is proportional to the conditional distribution evaluated at time S_i^- and shifted by 1. More in general, for $S_i \le t < S_{i+1}$, is proportional to

$$\mathbb{P}(Q^i_0 + A^i_\tau - 1 = k) \quad \text{for } k \ge 1, \quad \text{evaluated in } \tau = t - S_i$$

where Q^i_0 is a random variable with distribution equal to $\pi^q_{S_i^-}$, the marginal of $\pi_{S_i^-}$, and where (A^i_τ) is a Poisson process with parameter λ.

Summarizing it is more convenient to consider the subsequence $\{S^A_i\}$ corresponding to the observation times when an arrival is observed, i.e. when a new busy period is starting; the subsequence $\{S^{D,0}_i\}$ corresponding to the observation times when a departure is observed and the queue takes the value 0, i.e. when a new idle period is starting, and finally the subsequence $\{S^{D,1}_i\}$, corresponding to the remaining departure times, observed during a busy period.

When $J_0 = 2$, i.e. when the queue at time 0 is strictly positive, we start by observing a departure time. We apply procedure **2.**, with $\pi_{S_0^-} = \pi_0$, until the first observation time. The same procedure is used until we observe times in $\{S^{D,1}_i\}$. Therefore, before observing a time in $\{S^{D,0}_i\}$, the conditional distribution π^q_t of Q_t is proportional to

$$\mathbb{P}(Q^0_0 + A^1_\tau - d = k) \quad \text{for } k \ge 1, \quad \text{evaluated in } d = D_t,$$

where Q^0_0 has the distribution of Q_0 conditioned to $\{Q_0 \ge 1\}$. Then, when a time $S^{D,0}_l$ is observed, we apply procedure **0.**. A time $S^{D,0}_l$ is always followed by the corresponding time S^A_l, which in its turn is followed by some observation times of $\{S^{D,1}_i\}$, and so on, so that during a busy period $[S^A_l, S^{D,0}_{l+1})$

$$\pi^q_t(\{k\}) \propto \mathbb{P}(1 + A^l_\tau - d = k)\big|_{d=D_t - D_{S^A_l},\, \tau=t-S^A_l} \qquad \text{for } k \ge 1.$$

When $J_0 = 0$, i.e. when the queue at time 0 is zero, then we start from procedure **0.**, until the first arrival time S^A_1 is observed. Then we apply procedure **1.**, etc.

In particular when at time $t = 0$ we start with an empty queue, with probability 1, then $\pi^q_t = \delta_{\{0\}}$ during the idle periods, $\pi^q_{S^A_i} = \delta_{\{1\}}$, while during a busy period $(S^A_i, S^{D,0}_i)$

$$\pi^q_t(\{k\}) \propto \mathbb{P}(1 + A^i_\tau - d = k)\big|_{d=D_t - D_{S^A_i},\, \tau = t - S^A_i} \qquad \text{for } k \geq 1,$$

and then

$$\pi^q_t(\{k\}) \propto \left.\frac{(\lambda\tau)^{k-1+d}}{(k-1+d)!}\right|_{d=D_t - D_{S^A_i},\, \tau = t - S^A_i} \qquad \text{for } k \geq 1.$$

Acknowledgements One of us, G. Nappo, wishes to thank Professor G. Koch for posing the question of grouped data when (y_t) is a counting process, some years ago. Both of us wish to thank Professor T. G. Kurtz for suggesting the queueing model.

References

[1] Arjas, E. Haara, P. Norros, I. *Filtering the histories of a partially observed marked point process*, Stochastic Processes and their Applications, 40, 225-250, 1992

[2] Athreya, K.B. Kliemann, W. Koch, G. *On sequential construction of solutions of stochastic differential equations with jump terms*, Systems and Control Letters, 10, 141-146, 1988

[3] Boel, R. Varaiya, P. Wong, E. *Martingales on jump processes. Part I: Representation results; Part II: Applications;* SIAM J. Control 13, 999-1061, 1975

[4] Bremaud, P. *Point Processes and Queues*, Springer Verlag, Berlin, 1980

[5] Ceci, C. Gerardi, A. *Conditional law of a branching process observing a subpopulation* J. Appl. Prob. (to appear)

[6] Çinlar, E. Jacod, J. Protter, P. Sharpe, M. J. *Semimartingales and Markov Processes*, Zeitschrft für Warsch., vol 54, pp. 161 - 219, 1980

[7] Ethier, S.N. Kurtz, T.G. *Markov Processes: Characterization and Convergence* Wiley, New York, 1986

[8] Fan, K. *On a new approach to the solution of the nonlinear filtering equation of jump processes*, Probab. in the Engineering and Informational Sciences, 10, 153-163, 1996

[9] Frey,R. Runggaldier,W. *A Nonlinear Filtering Approach to Volatility Estimation with a View Towards High Frequency Data*, Int. J. Theor. Appl. Finance, 4, 2, pp. 199-210, 2001

[10] Kliemann,W. Koch,G. Marchetti,F. *On the Unnormalized Solution of the Filtering Problem with Counting Observations*, IEEE Transactions on Information Theory, 36, 6, 1415-1425, 1990

[11] Kurtz,T.G. *Martingale problems for conditional distributions of Markov processes*, Electr. J. Probab. 3, 9, 1998

[12] Kurtz,T.G. Ocone,D. *Unique characterization of conditional distributions in nonlinear filtering*, Ann. of Prob., 16, 1, 80-107, 1988

[13] Jacod,J. *Multivariate point processes: predictable projection, Radon-Nikodym derivative, representation of martingales* Proc. Japanese-SSSR Symposium, Lecture Notes in Math. 330, Springer, 1975

[14] Nappo,G. *Filtro del parametro di un processo di Poisson con dati raggruppati,* unpublished manuscript