NONLINEAR FILTERING FOR
JUMP DIFFUSION OBSERVATIONS

CLAUDIA CECI * ** AND

KATIA COLANERI,* *** *University of Chieti-Pescara*

Abstract

We deal with the filtering problem of a general jump diffusion process, X, when the observation process, Y, is a correlated jump diffusion process having common jump times with X. In this setting, at any time t the σ-algebra $\mathcal{F}_t^Y$ provides all the available information about X_t, and the central goal is to characterize the filter, π_t, which is the conditional distribution of X_t given observations $\mathcal{F}_t^Y$. To this end, we prove that π_t solves the Kushner–Stratonovich equation and, by applying the filtered martingale problem approach (see Kurtz and Ocone (1988)), that it is the unique weak solution to this equation. Under an additional hypothesis, we also provide a pathwise uniqueness result.

Keywords: Filtering; jump diffusion process

2010 Mathematics Subject Classification: Primary 93E11; 60J75; 60J60

1. Introduction

In this paper we consider a partially observed system (or filtering model), (X, Y), defined on some filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,T]}, \mathrm{P})$, with $T \in (0, \infty)$, where the dynamics of the signal $X = (X_t)_{t\in[0,T]}$ form a general jump diffusion process and the observation $Y = (Y_t)_{t\in[0,T]}$ is a correlated jump diffusion process having common jump times with X. As usual in a filtering model, the signal X cannot be observed directly, but we can observe a stochastic process Y, related to X. At any time t, the σ-algebra $\mathcal{F}_t^Y := \sigma\{Y_s, \ s \le t\}$ generated by Y supplies all the available information about X_t. The central goal of solving a filtering problem is to characterize the conditional distribution of X_t given the observation $\mathcal{F}_t^Y$, which provides the most detailed description of our knowledge of X_t.

The two major approaches to nonlinear filtering are the innovation method and the reference probability method. The latter is usually employed when it is possible to find an equivalent probability measure that makes the state X and the observations Y independent and, in such a case, this approach is better suited for approximation results. In the partially observed model considered in this paper, X and Y are correlated processes that may also have common jump times, so here such a probability measure does not exist. Hence, we choose the innovation approach as it is more appropriate to take into account these two features. This method consists of characterizing the filter as the unique solution (in some sense) to a stochastic differential equation, the so-called Kushner–Stratonovich equation (KS equation). The innovation approach

has been developed by several authors in many works. For example, Lipster and Shiryaev [20, pp. 297–380] treated the diffusive case and Bremáud [2, pp. 83–107] derived the theory in the purely discontinuous observations setting. We also refer the reader to [8] for an extensive analysis of the topic.

Since its origins, filtering theory has been successfully used in a great variety of engineering problems, informational sciences, and, more recently, mathematical finance. Among all the possible applications, we mention the application that concerns the study of partially observable stochastic control problems, where, in particular, the KS equation plays a key role. Usually, in these situations it is convenient to introduce an equivalent control problem with complete observations, the so-called separated problem, in order to avoid a direct approach. In this new problem the state is represented by the pair filter observation, whose first component satisfies the controlled KS equation. Uniqueness for the solutions of this evolution equation is an essential tool to prove the equivalence between the original problem and the separated problem. For a more detailed survey on this subject, we refer the reader to [9] and [10] for the diffusive case and [5] and [21] for counting observations.

Filtering problems have been widely investigated in the literature, mainly in the cases when Y gives observations of X in additional Gaussian noise (see, for example, [16], [17], and [20, pp. 297–380]) and when Y is a counting process or a marked point process (see [2, pp. 83–107], [3], [6], [7], [11], [19], and the references therein). The case of mixed-type observations (marked point processes and diffusions) has been studied in [4], [12], and [13]. All the papers cited above analyze the situation in which the information flow has the structure $\mathcal{F}_t^m \vee \mathcal{F}_t^\eta$, where $m(\mathrm{d}t, \mathrm{d}x)$ is a marked point process whose dynamics are influenced by a stochastic factor X and η gives observations of X in additional Gaussian noise. This particular structure to the information flow has a financial motivation; nevertheless, in a general framework, it may be meaningful to consider the case where the observation flow is generated by a jump diffusion process as in the model developed in this paper that, to the authors' knowledge, has not been investigated yet in the existing filtering literature.

In [12], by assuming that the additional Gaussian noise is independent from X, the authors applied the reference probability approach even in the case of joint jumps between the state and observations. Moreover, the above assumption and the particular structure of the information flow allowed them to reduce the filtering problem with mixed observations to that with pure jump observations. This technique fails in the general model of correlated jump diffusion processes, as the one studied in this paper. Using the innovation method together with an $\mathcal{F}_t^Y$-martingale representation theorem, we characterize the filter in terms of the KS equation. Frey and Schmidt [13] studied the case where the state process X is modeled as a finite-state Markov chain without common jump times with $m(\mathrm{d}t, \mathrm{d}x)$ and Ceci [4] studied the case where the state process X is a jump diffusion process that allows common jump times with $m(\mathrm{d}t, \mathrm{d}x)$— as mentioned above, in both of these papers the structure of the information flow is different from that considered herein. However, the filtering equation we obtain is quite similar to that in [4], but we prove weak uniqueness for the solutions under weaker conditions and we also give a pathwise uniqueness result. This kind of uniqueness was not investigated in [4].

The paper is organized as follows. The filtering model is described in Section 2. The main result, which establishes a characterization of the filter as the unique solution to the KS equation, is given in Section 3. The proofs of weak and strong uniqueness for this equation are postponed to Appendix B, where we deduce these results from uniqueness for the filtered martingale problem (see [17] and [18]).

2. The filtering model

A partially observed system (X, Y), where X is the unobservable component (state process) and Y the observable component, on a stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0,T]}, P)$, is described by the following system of stochastic differential equations:

$$
\begin{aligned}
dX_t &= b_0(t, X_t)\, dt + \sigma_0(t, X_t)\, dW_t^0 + \int_Z K_0(t, X_{t^-}, \zeta) N(dt, d\zeta), \quad X_0 = x_0, \\
dY_t &= b_1(t, X_t, Y_t)\, dt + \sigma_1(t, Y_t)\, dW_t^1 + \int_Z K_1(t, X_{t^-}, Y_{t^-}, \zeta) N(dt, d\zeta), \quad Y_0 = y_0.
\end{aligned}
\tag{2.1}
$$

Here $x_0, y_0 \in \mathbb{R}$, $N(dt, d\zeta)$ is a Poisson random measure on $\mathbb{R}^+ \times Z$, and $\nu(d\zeta)\, dt$ represents its intensity. We denote by $\widetilde{N}(dt, d\zeta)$ the compound Poisson measure, i.e. $\widetilde{N}(dt, d\zeta) := N(dt, d\zeta) - \nu(d\zeta)\, dt$. Note that $\nu(d\zeta)$ is a σ-finite measure on a measurable space $(Z, \mathcal{Z})$. The processes W^0 and W^1 are correlated $(P, \mathcal{F}_t)$-standard Brownian motions with correlation coefficient $\rho \in [-1, 1]$. The $\mathbb{R}$-valued functions $b_0(t, x)$, $b_1(t, x, y)$, $\sigma_0(t, x) > 0$, $\sigma_1(t, y) > 0$, $K_0(t, x, \zeta)$, and $K_1(t, x, y, \zeta)$ are measurable functions of their arguments. Let us remark that in the dynamics of the observation process Y, the diffusive coefficient does not depend on the state process X, although the drift does.

From now on we will write $b_i(t), \sigma_i(t), K_i(t, \zeta), i = 0, 1$, for $b_0(t, X_t), b_1(t, X_t, Y_t),$ $\sigma_0(t, X_t), \sigma_1(t, Y_t), K_0(t, X_{t^-}, \zeta)$, and $K_1(t, X_{t^-}, Y_{t^-}, \zeta)$, respectively, unless it is necessary to underline the dependence on the processes involved.

We assume some requirements for (2.1) to be well defined: for $i = 0, 1$,

$$
E \int_0^T \int_Z |K_i(t, \zeta)| \nu(d\zeta)\, dt < \infty, \qquad E \int_0^T |b_i(t)|\, dt < \infty, \qquad E \int_0^T \sigma_i^2(t)\, dt < \infty. \tag{2.2}
$$

Under these constraints, both the processes X and Y have finite first moments.

We also assume strong existence and uniqueness for the system (2.1). Sufficient conditions are summarized in Appendix A. In particular, these assumptions imply that the pair (X, Y) is a $(P, \mathcal{F}_t)$-Markov process.

Denote by $(\mathcal{F}_t^Y)_{t \in [0,T]}$ the filtration generated by the observation process Y until time t. In the partially observed system considered in this paper, at any time t, the σ-algebra $\mathcal{F}_t^Y$ provides all the available information about the signal X_t.

By defining $\mathcal{P}(\mathbb{R})$ to be the space of the probability measures over $\mathbb{R}$, it is known that there exists a $\mathcal{P}(\mathbb{R})$-valued $\mathcal{F}_t^Y$-adapted process, π_t, such that

$$
\pi_t(f) = E[f(t, X_t) \mid \mathcal{F}_t^Y] \tag{2.3}
$$

for any measurable function $f(t, x)$ on $[0, T] \times \mathbb{R}$ such that $f(t, X)$ is integrable. Since X is a càdlàg process, there exists a version of π with càdlàg paths (see, for instance, [17]).

Throughout the rest of the paper, we will write $\widehat{R}$ for the $(P, \mathcal{F}_t^Y)$-optional projection of a progressively measurable process R, satisfying $E |R_t| < \infty$ for all $t \in [0, T]$, defined as the unique optional process such that, for any $\mathcal{F}_t^Y$-stopping time τ, $\widehat{R}_\tau = E[R_\tau \mid \mathcal{F}_\tau^Y]$ almost surely (a.s.) on $\{\tau < \infty\}$. With this notation we can write the $\mathcal{F}_t^Y$-optional projection of a process $f(t, X_t)$ as

$$
\widehat{f(t, X_t)} = \pi_t(f).
$$

In this case $\widehat{f(t, X_t)}$ has càdlàg trajectories (we will use both the notation $\widehat{f(t, X_t)}$ and $\pi_t(f)$ in the following).

Remark 2.1. In the sequel we will use two well-known facts: for every $(P, \mathcal{F}_t)$-martingale m, the projection $\widehat{m}$ is a $(P, \mathcal{F}_t^Y)$-martingale, and, for any progressively measurable process Ψ with $E \int_0^T |\Psi_t| \, dt < \infty$,

$$\widehat{\int_0^T \Psi_t \, dt} - \int_0^T \widehat{\Psi}_t \, dt$$

is a $(P, \mathcal{F}_t^Y)$-martingale. Note that this implies that $E \int_0^T \Psi_t \, dt = E \int_0^T \widehat{\Psi}_t \, dt$.

We will also need the following result.

Proposition 2.1. *Let $(m_t)_{t \geq 0}$ be a $(P, \mathcal{F}_t)$-local martingale. If there exists a localizing sequence $(\tau_n)_{n \in \mathbb{N}}$ of $\mathcal{F}_t^Y$-stopping times for m, then $\widehat{m}$ is a $(P, \mathcal{F}_t^Y)$-local martingale.*

Proof. By a standard calculation we obtain, for all $0 \leq s < t \leq T < \infty$,

$$
\begin{aligned}
E[\widehat{m}_{t \wedge \tau_n} \mid \mathcal{F}_s^Y] &= E[E[m_{t \wedge \tau_n} \mid \mathcal{F}_{t \wedge \tau_n}^Y] \mid \mathcal{F}_s^Y] \\
&= E[E[m_{t \wedge \tau_n} \mid \mathcal{F}_{t \wedge \tau_n}^Y] \mathbf{1}_{\{\tau_n > s\}} \mid \mathcal{F}_s^Y] + E[E[m_{t \wedge \tau_n} \mid \mathcal{F}_{t \wedge \tau_n}^Y] \mathbf{1}_{\{\tau_n \leq s\}} \mid \mathcal{F}_s^Y] \\
&= E[E[m_{t \wedge \tau_n} \mid \mathcal{F}_s] \mid \mathcal{F}_s^Y] \mathbf{1}_{\{\tau_n > s\}} + E[m_{\tau_n} \mid \mathcal{F}_{\tau_n}^Y] \mathbf{1}_{\{\tau_n \leq s\}} \\
&= \widehat{m}_{s \wedge \tau_n}.
\end{aligned}
$$

Hence, $\widehat{m}_{t \wedge \tau_n}$ is a $(P, \mathcal{F}_t^Y)$-martingale and this proves the statement. ∎

The integer-valued random measure associated to the jumps of the process Y is given by

$$m(dt, dx) = \sum_{\{s \,:\, \Delta Y_s \neq 0\}} \delta_{\{s, \Delta Y_s\}}(dt, dx),$$

where δ_a denotes the Dirac measure at point a. Note that the equality

$$\int_0^t \int_{\mathbb{R}} x \, m(ds, dx) = \int_0^t \int_Z K_1(s, \zeta) N(ds, d\zeta)$$

holds, and, in general, for any measurable function $g : \mathbb{R} \to \mathbb{R}$,

$$\int_0^t \int_{\mathbb{R}} g(x) m(ds, dx) = \int_0^t \int_Z \mathbf{1}_{\{K_1(s, \zeta) \neq 0\}} \, g(K_1(s, \zeta)) N(ds, d\zeta). \tag{2.4}$$

For all $t \in [0, T]$ and all $A \in \mathcal{B}(\mathbb{R})$, we define

$$
\begin{aligned}
d^0(t, x) &:= \{\zeta \in Z : K_0(t, x, \zeta) \neq 0\}, \\
d^1(t, x, y) &:= \{\zeta \in Z : K_1(t, x, y, \zeta) \neq 0\}, \\
d^A(t, x, y) &:= \{\zeta \in Z : K_1(t, x, y, \zeta) \in A \setminus \{0\}\} \subseteq d^1(t, x, y), \tag{2.5}
\end{aligned}
$$

and, finally,

$$D_t^A = d^A(t, X_{t-}, Y_{t-}) \subseteq D_t = d^1(t, X_{t-}, Y_{t-}), \qquad D_t^0 = d^0(t, X_{t-}).$$

Normally, $D_t^0 \cap D_t \neq \varnothing$, P-a.s. and this models the fact that the state process and the observation may have common jump times.

Under the assumption that

$$\mathrm{E} \int_0^T \nu(D_s)\, \mathrm{d}s < \infty, \tag{2.6}$$

in [3, Proposition 2.2] it was proved that the $(\mathrm{P}, \mathcal{F}_t)$-predictable projection (see [2] and [15]), $m^p(\mathrm{d}t, \mathrm{d}x)$, of the integer-valued measure $m(\mathrm{d}t, \mathrm{d}x)$ can be written as

$$m^p(\mathrm{d}t, \mathrm{d}x) = \lambda_t \phi_t(\mathrm{d}x)\, \mathrm{d}t, \tag{2.7}$$

where, for all $A \in \mathcal{B}(\mathbb{R})$,

$$m^p(\mathrm{d}t, A) = \lambda_t \phi_t(A)\, \mathrm{d}t = \nu(D_t^A)\, \mathrm{d}t.$$

This means that $\nu(D_t^A)$ provides the $(\mathrm{P}, \mathcal{F}_t)$-intensity of the point process $N_t(A) = m((0, t] \times A)$ that counts the jumps of the process Y until time t, whose widths belong to A. In particular, $\lambda_t = \nu(D_t)$ is the $(\mathrm{P}, \mathcal{F}_t)$-predictable intensity of the point process $N_t = m((0, t] \times \mathbb{R})$ which counts the jumps of the process Y until time t.

Remark 2.2. Equation (2.7) can also be written as

$$m^p(\mathrm{d}t, \mathrm{d}x) = \lambda_t \phi_t(\mathrm{d}x)\, \mathrm{d}t = \int_{D_t} \delta_{K_1(t,\zeta)}(\mathrm{d}x)\nu(\mathrm{d}\zeta)\, \mathrm{d}t. \tag{2.8}$$

We finally denote by $\nu^p(\mathrm{d}t, \mathrm{d}x)$ the $(\mathrm{P}, \mathcal{F}_t^Y)$-predictable projection of the integer-valued measure $m(\mathrm{d}t, \mathrm{d}x)$. The following proposition, proved in [3], gives a representation of $\nu^p(\mathrm{d}t, \mathrm{d}x)$ in terms of the filter.

Proposition 2.2. *The $(\mathrm{P}, \mathcal{F}_t^Y)$-predictable projection of the integer-valued measure $m(\mathrm{d}t, \mathrm{d}x)$ is given by*

$$\nu^p(\mathrm{d}t, \mathrm{d}x) = \widehat{\lambda_t \phi_t}(\mathrm{d}x)|_{t=t^-}\, \mathrm{d}t = \pi_{t^-}(\lambda_t \phi_t(\mathrm{d}x))\, \mathrm{d}t,$$

that is, for any $A \in \mathcal{B}(\mathbb{R})$,

$$\nu^p((0, t], A) = \int_0^t \pi_{s^-}(\lambda_s \phi_s(A))\, \mathrm{d}s = \int_0^t \pi_{s^-}(\nu(d^A(\cdot, Y_{s^-})))\, \mathrm{d}s,$$

where π_{t^-} denotes the left version of the process π_t.

The last part of this section focuses on finding a martingale representation theorem for $(\mathrm{P}, \mathcal{F}_t^Y)$-martingales which is an essential tool to derive the filtering equation. To this end, we introduce the $\mathcal{F}_t^Y$-compensated martingale random measure

$$m^\pi(\mathrm{d}t, \mathrm{d}x) = m(\mathrm{d}t, \mathrm{d}x) - \nu^p(\mathrm{d}t, \mathrm{d}x) = m(\mathrm{d}t, \mathrm{d}x) - \pi_{t^-}(\lambda_t \phi_t(\mathrm{d}x))\, \mathrm{d}t,$$

and, assuming that

$$\mathrm{E} \int_0^T \frac{b_1^2(t)}{\sigma_1^2(t)}\, \mathrm{d}t < \infty, \tag{2.9}$$

we define the innovation process

$$I_t = W_t^1 + \int_0^t \left[\frac{b_1(s)}{\sigma_1(s)} - \pi_s\left(\frac{b_1}{\sigma_1} \right) \right] \mathrm{d}s. \tag{2.10}$$

Note that, by Remark 2.1 and assumption (2.9),

$$\mathrm{E} \int_0^T \left| \pi_t\left(\frac{b_1}{\sigma_1}\right) \right| \mathrm{d}t \le \mathrm{E} \int_0^T \pi_t \left|\frac{b_1}{\sigma_1}\right| \mathrm{d}t = \mathrm{E} \int_0^T \frac{|b_1(t)|}{\sigma_1(t)} \, \mathrm{d}t < \infty.$$

By extending classical results from filtering theory (see [20]) to our setting we obtain the following result.

Proposition 2.3. *The random process $\{I_t\}_{t\in[0,T]}$ is a $(\mathrm{P}, \mathcal{F}_t^Y)$-Wiener process.*

We write $\mathcal{F}_t^m$ for the filtration generated by the random measure $m(\mathrm{d}t, \mathrm{d}x)$. Since the innovation process I and the random measure $m(\mathrm{d}t, \mathrm{d}x)$ are $\mathcal{F}_t^Y$-adapted, then $\mathcal{F}_t^m \vee \mathcal{F}_t^I \subseteq \mathcal{F}_t^Y$. In general, the inclusion is strict; however, we will prove a representation theorem for $(\mathrm{P}, \mathcal{F}_t^Y)$-martingales in terms of the $\mathcal{F}_t^Y$-compensated random martingale measure $m^\pi(\mathrm{d}t, \mathrm{d}x)$ and the innovation process I.

For this purpose, let us now consider the positive local martingale defined by

$$L_t = \mathcal{E}\left(-\int_0^t \frac{b_1(s)}{\sigma_1(s)} \, \mathrm{d}W_s^1\right) = \exp\left\{-\int_0^t \frac{b_1(s)}{\sigma_1(s)} \, \mathrm{d}W_s^1 - \frac{1}{2} \int_0^t \frac{b_1^2(s)}{\sigma_1^2(s)} \, \mathrm{d}s\right\},$$

where $\mathcal{E}$ denotes the Doléans-Dade exponential and we will make the following usual standing assumption.

Assumption A. *L is a $(\mathrm{P}, \mathcal{F}_t)$-martingale, that is, $\mathrm{E}[L_T] = 1$.*

Under Assumption A, we define a probability measure Q on $\mathcal{F}_T$ equivalent to P such that

$$\left.\frac{\mathrm{d}Q}{\mathrm{d}P}\right|_{\mathcal{F}_T} = L_T. \tag{2.11}$$

By Girsanov's theorem, the process

$$\widetilde{W}_t^1 = W_t^1 + \int_0^t \frac{b_1(s)}{\sigma_1(s)} \, \mathrm{d}s$$

is a $(\mathrm{Q}, \mathcal{F}_t)$-Wiener process and, by (2.10),

$$\widetilde{W}_t^1 = I_t + \int_0^t \pi_s\left(\frac{b_1}{\sigma_1}\right) \mathrm{d}s; \tag{2.12}$$

hence, $\widetilde{W}^1$ is a $(\mathrm{Q}, \mathcal{F}_t^Y)$-Wiener process, which in turn implies that

$$\widehat{L}_t = \mathrm{E}[L_t \mid \mathcal{F}_t^Y] = \left.\frac{\mathrm{d}Q}{\mathrm{d}P}\right|_{\mathcal{F}_t^Y} = \mathcal{E}\left(-\int_0^t \pi_s\left(\frac{b_1}{\sigma_1}\right) \mathrm{d}I_s\right). \tag{2.13}$$

Note that, by Jensen's inequality, $\pi_t^2(b_1/\sigma_1) \le \pi_t(b_1^2/\sigma_1^2)$ and, by Remark 2.1, the following integrability condition holds:

$$\mathrm{E} \int_0^T \pi_t\left(\frac{b_1^2}{\sigma_1^2}\right) \mathrm{d}t = \mathrm{E} \int_0^T \frac{b_1^2(t)}{\sigma_1^2(t)} \, \mathrm{d}t < \infty.$$

Before giving the claimed result, we want to underline the $(P, \mathcal{F}_t^Y)$-semimartingale representation of Y:

$$\mathrm{d}Y_t = \sigma_1(t)\,\mathrm{d}I_t + \left(\pi_t(b_1) + \int_{\mathbb{R}} x\widehat{\lambda_t\phi_t}(\mathrm{d}x)\right)\mathrm{d}t + \int_{\mathbb{R}} xm^\pi(\mathrm{d}t, \mathrm{d}x).$$

We can observe that, again by Remark 2.1 and the assumptions in (2.2),

$$\mathrm{E}\int_0^T |\pi_t(b_1)|\,\mathrm{d}t \leq \mathrm{E}\int_0^T \pi_t|b_1|\,\mathrm{d}t = \mathrm{E}\int_0^T |b_1(t)|\,\mathrm{d}t < \infty;$$

furthermore, taking into account (2.8),

$$\mathrm{E}\int_0^T \int_{\mathbb{R}} |x|\widehat{\lambda_t\phi_t}(\mathrm{d}x)\,\mathrm{d}t = \mathrm{E}\int_0^T \int_{\mathbb{R}} |x|\lambda_t\phi_t(\mathrm{d}x)\,\mathrm{d}t = \mathrm{E}\int_0^T \int_Z |K_1(t, \zeta)|\nu(\mathrm{d}\zeta)\,\mathrm{d}t < \infty.$$

Proposition 2.4. *Under (2.2), (2.6), (2.9), and Assumption A, every $(P, \mathcal{F}_t^Y)$-local martingale M admits the decomposition*

$$M = M_0 + \int_0^t \int_{\mathbb{R}} w(s, x)m^\pi(\mathrm{d}s, \mathrm{d}x) + \int_0^t h(s)\,\mathrm{d}I_s,$$

where $w(t, x)$ is an $\mathcal{F}_t^Y$-predictable process and $h(t)$ is an $\mathcal{F}_t^Y$-adapted process such that

$$\int_0^T \int_{\mathbb{R}} |w(t, x)|\pi_{t-}(\lambda_t\phi_t(\mathrm{d}x))\,\mathrm{d}t < \infty \quad \text{and} \quad \int_0^T h(t)^2\,\mathrm{d}t < \infty \quad P\text{-}a.s.$$

Proof. Let Q be the probability measure defined in (2.11). Then $\widetilde{W}^1$ is a $(Q, \mathcal{F}_t^Y)$-Brownian motion. Note that the following equality of σ-algebras holds:

$$\mathcal{F}_t^Y = \mathcal{F}_t^m \vee \mathcal{F}_t^{\widetilde{W}^1}. \tag{2.14}$$

As a matter of fact, recalling (2.12), we obtain the inclusion $\mathcal{F}_t^Y \supseteq \mathcal{F}_t^m \vee \mathcal{F}_t^{\widetilde{W}^1}$, while the other inclusion follows from the fact that Y solves the stochastic differential equation driven by $m(\mathrm{d}t, \mathrm{d}x)$ and W given by

$$\mathrm{d}Y_t = \int_{\mathbb{R}} xm(\mathrm{d}t, \mathrm{d}x) + \sigma_1(t, Y_t)\,\mathrm{d}\widetilde{W}_t^1.$$

Let us note that the Q-distribution of the pair $(m, \widetilde{W}_t^1)$ is uniquely determined by its $(Q, \mathcal{F}_t^m \vee \mathcal{F}_t^{\widetilde{W}^1})$-predictable characteristics (see Remark 3.2 of [1]), and, therefore, by the $(Q, \mathcal{F}_t^Y)$-predictable characteristics because of equality (2.14).

By applying Corollary III.4.31 of [15], every $(Q, \mathcal{F}_t^Y)$-local martingale, $\widetilde{M}$, has the representation property with respect to $(m, \widetilde{W}^1)$, which means that there exist two processes, $\widetilde{h}(t)$, $\mathcal{F}_t^Y$-adapted, and $\widetilde{w}(t, x)$, $\mathcal{F}_t^Y$-predictable, satisfying

$$\int_0^T \widetilde{h}^2(t)\,\mathrm{d}t < \infty \quad \text{and} \quad \int_0^T \int_{\mathbb{R}} |\widetilde{w}(t, x)|\pi_{t-}(\lambda_t\phi_t(\mathrm{d}x))\,\mathrm{d}t < \infty \quad Q\text{-}a.s.$$

such that

$$\widetilde{M}_t = \widetilde{M}_0 + \int_0^t \widetilde{h}(s)\,\mathrm{d}\widetilde{W}_s^1 + \int_0^t \int_{\mathbb{R}} \widetilde{w}(s, x)m^\pi(\mathrm{d}s, \mathrm{d}x). \tag{2.15}$$

Let M be a $(P, \mathcal{F}_t^Y)$-local martingale. By the Kallianpur–Striebel formula, $\widetilde{M}_t = M_t \widehat{L}_t^{-1}$ is a $(Q, \mathcal{F}_t^Y)$-local martingale, where $\widehat{L}$ is defined in (2.13). Thus, $M_t = \widetilde{M}_t \widehat{L}_t$ can be computed by the product formula:

$$\mathrm{d}M_t = \widetilde{M}_{t-}\,\mathrm{d}\widehat{L}_t + \widehat{L}_{t-}\,\mathrm{d}\widetilde{M}_t + \mathrm{d}\langle \widetilde{M}^c, \widehat{L}^c \rangle_t + \mathrm{d}\left(\sum_{s \leq t} \Delta \widetilde{M}_s \Delta \widehat{L}_s \right).$$

Note that $\mathrm{d}\widehat{L}_t = -\widehat{L}_t \pi_t(b_1/\sigma_1)\,\mathrm{d}I_t$. Then

$$\mathrm{d}M_t = -\widetilde{M}_t \widehat{L}_t \pi_t\left(\frac{b_1}{\sigma_1}\right)\mathrm{d}I_t + \widehat{L}_{t-}\left(\int_{\mathbb{R}} \widetilde{w}(t, x) m^\pi(\mathrm{d}t, \mathrm{d}x) + \widetilde{h}(t)\,\mathrm{d}\widetilde{W}_t^1 \right)$$

$$+ \mathrm{d}\left\langle \int_0^{\cdot} \widetilde{h}(s)\,\mathrm{d}\widetilde{W}_s^1, -\int_0^{\cdot} \widehat{L}_s \pi_s\left(\frac{b_1}{\sigma_1}\right)\mathrm{d}I_s \right\rangle_t$$

$$= -M_t \pi_t\left(\frac{b_1}{\sigma_1}\right)\mathrm{d}I_t + \widehat{L}_t \widetilde{h}(t)\left(\mathrm{d}I_t + \pi_t\left(\frac{b_1}{\sigma_1}\right)\mathrm{d}t \right) - \widetilde{h}(t)\widehat{L}_t \pi_t\left(\frac{b_1}{\sigma_1}\right)\mathrm{d}t$$

$$+ \int_{\mathbb{R}} \widehat{L}_{t-} \widetilde{w}(t, x) m^\pi(\mathrm{d}t, \mathrm{d}x)$$

$$= \left(-M_t \pi_t\left(\frac{b_1}{\sigma_1}\right) + \widehat{L}_t \widetilde{h}(t) \right)\mathrm{d}I_t + \int_{\mathbb{R}} \widehat{L}_{t-} \widetilde{w}(t, x) m^\pi(\mathrm{d}t, \mathrm{d}x).$$

Finally, we need to only define

$$w(t, x) = \widehat{L}_{t-} \widetilde{w}(t, x) \quad \text{and} \quad h(t) = -M_t \pi_t\left(\frac{b_1}{\sigma_1}\right) + \widehat{L}_t \widetilde{h}(t).$$

3. The filtering equation

Our purpose in this section is to characterize the filter that is the conditional distribution of the signal X given the observation $\mathcal{F}_t^Y$, which provides, as already stated, the most detailed description of our knowledge of X_t.

The literature concerning the filtering problem in the case of diffusion observations is quite rich: for textbook treatments, see, for instance, [16] and [20]. More recently, results for pure jump observations have been obtained (see [2], [3], [7], [11], and the references therein), while few results can be found for mixed-type information which involves pure jump processes and diffusions (see [4], [12], and [13]). To the authors' knowledge, this is the first time that the filtering problem is studied for the general jump diffusion system defined in (2.1).

First we want to recall a result proved in [7, Corollary 2.2].

Let us denote by $\mathcal{C}_b^{1,2}([0, T] \times \mathbb{R})$ and $\mathcal{C}_b^{1,2,2}([0, T] \times \mathbb{R} \times \mathbb{R})$ the sets of functions f respectively defined on $[0, T] \times \mathbb{R}$ and $[0, T] \times \mathbb{R} \times \mathbb{R}$ such that f, $\partial f/\partial t$, $\partial f/\partial x$, and $\partial^2 f/\partial x^2$ and f, $\partial f/\partial t$, and all the first and second derivatives with respect to (x, y), respectively, are bounded continuous functions.

Lemma 3.1. *Under the assumptions in (2.2) for $i = 0$ and*

$$\mathrm{E} \int_0^T \nu(D_t^0)\,\mathrm{d}t < \infty, \tag{3.1}$$

X is a $(P, \mathcal{F}_t)$-Markov process with generator

$$L^X f(t, x) = \frac{\partial f}{\partial t} + b_0(t, x)\frac{\partial f}{\partial x} + \frac{1}{2}\sigma_0^2(t, x)\frac{\partial^2 f}{\partial x^2} + \int_Z \{f(t, x + K_0(t, x, \zeta)) - f(t, x)\}\nu(\mathrm{d}\zeta).$$

More precisely, for any function $f(t, x) \in \mathcal{C}_b^{1,2}([0, T] \times \mathbb{R})$, the semimartingale decomposition

$$f(t, X_t) = f(0, x_0) + \int_0^t L^X f(s, X_s)\, \mathrm{d}s + m_t^f$$

holds, where m^f is the $(\mathrm{P}, \mathcal{F}_t)$-martingale given by

$$m_t^f = \int_0^t \frac{\partial f}{\partial x}(s, X_s)\sigma_0(s, X_s)\, \mathrm{d}W_s^0$$
$$+ \int_0^t \int_Z \{f(s, X_{s^-} + K_0(s, X_{s^-}, \zeta)) - f(s, X_{s^-})\}\widetilde{N}(\mathrm{d}s, \mathrm{d}\zeta). \tag{3.2}$$

The next theorem establishes the main result of this paper.

Theorem 3.1. *Under the assumptions of Proposition 2.4 and (3.1), the filter (2.3) is a solution of the KS equation, which is given for any function $f(t, x) \in \mathcal{C}_b^{1,2}([0, T] \times \mathbb{R})$ by*

$$\pi_t(f) = f(0, x_0) + \int_0^t \pi_s(L^X f)\, \mathrm{d}s + \int_0^t \int_{\mathbb{R}} w_s^{\pi}(f, x) m^{\pi}(\mathrm{d}s, \mathrm{d}x) + \int_0^t h_s^{\pi}(f)\, \mathrm{d}I_s, \tag{3.3}$$

where

$$w_t^{\pi}(f, x) = \frac{\mathrm{d}\pi_{t^-}(\lambda\phi f)}{\mathrm{d}\pi_{t^-}(\lambda\phi)}(x) - \pi_{t^-}(f) + \frac{\mathrm{d}\pi_{t^-}(\bar{L}f)}{\mathrm{d}\pi_{t^-}(\lambda\phi)}(x), \tag{3.4}$$

$$h_t^{\pi}(f) = \sigma_1^{-1}(t)[\pi_t(b_1 f) - \pi_t(b_1)\pi_t(f)] + \rho\pi_t\left(\sigma_0 \frac{\partial f}{\partial x}\right). \tag{3.5}$$

Here, by

$$\frac{\mathrm{d}\pi_{t^-}(\lambda\phi f)}{\mathrm{d}\pi_{t^-}(\lambda\phi)}(x) \quad and \quad \frac{\mathrm{d}\pi_{t^-}(\bar{L}f)}{\mathrm{d}\pi_{t^-}(\lambda\phi)}(x)$$

we mean the Radon–Nikodym derivatives of the measures $\pi_{t^-}(\lambda\phi f)$ and $\pi_{t^-}(\bar{L}f)$ with respect to $\pi_{t^-}(\lambda\phi)$, and the operator $\bar{L}f$ is defined as

$$\bar{L}f = \bar{L}f(\cdot, Y_{t^-}, \mathrm{d}z),$$

$$\bar{L}f(t, x, y, A) = \int_{d^A(t,x,y)} \{f(t, x + K_0(t, x, \zeta)) - f(t, x)\}\nu(\mathrm{d}\zeta) \quad for\ all\ A \in \mathcal{B}(\mathbb{R}).$$

Remark 3.1. Recall that $d^A(t, x, y)$ is defined in (2.5); hence, the operator $\bar{L}$ takes into account common jump times between the state X and the observations Y.

Remark 3.2. Let us observe that (3.3) is similar to the filtering equation derived in [4], even if a different partially observed system was considered there. More precisely, in [4], the information flow has the structure $\mathcal{F}_t^m \vee \mathcal{F}_t^{\eta}$, where $m(\mathrm{d}t, \mathrm{d}x)$ is a marked point process with dynamics affected by a stochastic factor X (whose dynamics are described by the first equation of (2.1)), and $\eta_t = \int_0^t \gamma(X_s)\, \mathrm{d}s + W_t^1$ for any bounded measurable function $\gamma(x)$. Nevertheless, let us point out that in [4] the derivation of the filtering equation required boundedness on $\lambda_t = \nu(D_t)$ and $\sigma_0(t, x)$.

Before giving the proof of Theorem 3.1, we ought to check that all the terms in (3.3) are well defined.

Remark 3.3. Since

$$\int_{\mathbb{R}} |w_s^\pi(f, x)| \pi_{s-}(\lambda_s \phi_s(\mathrm{d}x)) \le |\pi_{s-}(\lambda f)| + |\pi_{s-}(\lambda)\pi_{s-}(f)| + |\pi_{s-}(\bar{L}f)(\mathbb{R})|$$

$$\le 4\|f\|\pi_{s-}|\lambda|,$$

assumption (2.6) and Remark 2.1 imply that

$$\mathrm{E} \int_0^T \int_{\mathbb{R}} |w_s^\pi(f, x)| \pi_{s-}(\lambda_s \phi_s(\mathrm{d}x)) \, \mathrm{d}s \le 4\|f\| \, \mathrm{E} \int_0^T |\lambda_s| \, \mathrm{d}s < \infty. \tag{3.6}$$

Moreover, since, for any $f(t, x) \in \mathcal{C}_b^{1,2}([0, T] \times \mathbb{R})$,

$$(h_t^\pi(f))^2 \le B_f\{(\sigma_1^{-1}(t))^2 \pi_t^2(b_1) + \pi_t^2(\sigma_0)\}$$

with B_f a suitable positive constant, by Jensen's inequality and again by Remark 2.1, we obtain

$$\mathrm{E} \int_0^T (h_t^\pi(f))^2 \, \mathrm{d}t \le B_f \, \mathrm{E} \int_0^T \left(\frac{b_1^2(t)}{\sigma_1^2(t)} + \sigma_0^2(t) \right) \mathrm{d}t < \infty. \tag{3.7}$$

Thus, taking into account (3.6) and (3.7), the integrals in (3.3) with respect to the compensated martingale measure $m^\pi(\mathrm{d}t, \mathrm{d}x)$ and the innovation process I are $(\mathrm{P}, \mathcal{F}_t^Y)$-martingales. Finally, note that

$$|L^X f(t, X_t)| \le \widetilde{B}_f(1 + |b_0(t)| + |\sigma_0(t)|^2 + \nu(D_t^0))$$

for a suitable positive constant $\widetilde{B}_f$. Then

$$\mathrm{E} \int_0^T |\pi_t(L^X f)| \, \mathrm{d}t \le \mathrm{E} \int_0^T |L^X f(t, X_t)| \, \mathrm{d}t < \infty.$$

Proof of Theorem 3.1. We will consider the semimartingale

$$Z_t = f(t, X_t) = f(0, X_0) + \int_0^t L^X f(s, X_s) \, \mathrm{d}s + m_t^f,$$

where m_t^f is given in (3.2). To keep the formulae simple, we will leave out the dependence from the process X unless it is necessary, that is, $f_t = f(t, X_t)$, $L^X f_t = L^X f(t, X_t)$, and $\partial f(t)/\partial x = \partial f(t, X_t)/\partial x$.

Now we project the semimartingale Z onto $\mathcal{F}_t^Y$:

$$\widehat{Z}_t = \widehat{Z}_0 + \int_0^t \widehat{L^X f_s} \, \mathrm{d}s + \widehat{m}_t^f = \widehat{Z}_0 + \int_0^t \widehat{L^X f_s} \, \mathrm{d}s - \int_0^t \widehat{L^X f_s} \, \mathrm{d}s + \int_0^t \widehat{L^X f_s} \, \mathrm{d}s + \widehat{m}_t^f.$$

By Remark 2.1, $\widehat{Z}_t - \widehat{Z}_0 - \int_0^t \widehat{L^X f_s} \, \mathrm{d}s$ is an $\mathcal{F}_t^Y$-martingale. Proposition 2.4 ensures the existence of two processes h^π and w^π such that

$$\widehat{Z}_t - \widehat{Z}_0 - \int_0^t \widehat{L^X f}(s, X_s) \, \mathrm{d}s = \int_0^t \int_{\mathbb{R}} w_s^\pi(f, x) m^\pi(\mathrm{d}s, \mathrm{d}x) + \int_0^t h_s^\pi(f) \, \mathrm{d}I_s$$

with $\mathrm{E} \int_0^T \int_{\mathbb{R}} |w_s^\pi(f, x)| \pi_s(\lambda_s \phi_s(\mathrm{d}x)) \, \mathrm{d}s < \infty$ and $\mathrm{E} \int_0^T (h_s^\pi(f))^2 \, \mathrm{d}s < \infty$.

Now the strategy for proving the thesis consists of two steps. We will consider the $\mathcal{F}_t^Y$-adapted process $\widetilde{W}^1$, given by

$$\widetilde{W}_t^1 = W_t^1 + \int_0^t \frac{b_1(s)}{\sigma_1(s)}\,\mathrm{d}s = I_t + \int_0^t \pi_s\left(\frac{b_1}{\sigma_1}\right)\mathrm{d}s,$$

and a bounded process U of the form

$$U_t = \int_0^t \int_{\mathbb{R}} \Gamma(s,x) m(\mathrm{d}s,\mathrm{d}x),$$

where Γ is a bounded, $\mathcal{F}_t^Y$-predictable process.

Step 1. We will compute $\widehat{Z_t \widetilde{W}_t^1}$ and $\widehat{Z}_t \widetilde{W}_t^1$ separately; since $\widetilde{W}_t^1$ is $\mathcal{F}_t^Y$-adapted, the equality $\widehat{Z_t \widetilde{W}_t^1} = \widehat{Z}_t \widetilde{W}_t^1$ holds.

Step 2. We will compute $\widehat{Z_t U_t}$ and $\widehat{Z}_t U_t$, and, again, since U_t is $\mathcal{F}_t^Y$-adapted, the equality $\widehat{Z_t U_t} = \widehat{Z}_t U_t$ holds.

The two equalities in steps 1 and 2 will give us the structure of the processes h^π and w^π.

Step 1. By applying the product rule,

$$\begin{aligned}
\mathrm{d}(Z_t \widetilde{W}_t^1) &= Z_{t-}\,\mathrm{d}\widetilde{W}_t^1 + \widetilde{W}_{t-}^1\,\mathrm{d}Z_t + \mathrm{d}\langle Z^c, \widetilde{W}^1\rangle_t \\
&= Z_t\,\mathrm{d}W_t^1 + Z_{t-}\frac{b_1(t)}{\sigma_1(t)}\,\mathrm{d}t + \widetilde{W}_t^1 L^X f_t\,\mathrm{d}t + \frac{\partial f}{\partial x}(t)\sigma_0(t)\rho\,\mathrm{d}t + \mathrm{d}m_t^1,
\end{aligned}$$

where $m_t^1 = \int_0^t \widetilde{W}_s^1\,\mathrm{d}m_s^f$ is a $(P, \mathcal{F}_t)$-local martingale, and ρ is the correlation coefficient between the Brownian motions W^1 and W^0. Note that we can introduce an $\mathcal{F}_t^Y$-localizing sequence for m^1 as

$$\widetilde{\tau}_n = T \wedge \inf\{t : |\widetilde{W}_t^1| \geq n\}.$$

If we project $Z_t \widetilde{W}_t^1$ onto $\mathcal{F}_t^Y$, we will obtain, on $\{t \leq \widetilde{\tau}_n\}$,

$$\mathrm{d}(\widehat{Z_t \widetilde{W}_t^1}) = \left\{ Z_t \widehat{\frac{b_1(t)}{\sigma_1(t)}} + \widehat{\widetilde{W}_t^1 L^X f_t} + \widehat{\frac{\partial f(t)}{\partial x}}\sigma_0(t)\rho \right\}\mathrm{d}t + \widehat{Z_t\,\mathrm{d}W_t^1} + \mathrm{d}\widehat{m}_t^1 + \mathrm{d}\widetilde{m}_t^1,$$

where $\widetilde{m}^1$ is a $(P, \mathcal{F}_t^Y)$-martingale and, by Proposition 2.1, $\widehat{m_{t \wedge \widetilde{\tau}_n}^1}$ is a $(P, \mathcal{F}_t^Y)$-martingale. On the other hand,

$$\mathrm{d}(\widehat{Z}_t \widetilde{W}_t^1) = \left\{ \widehat{Z}_t \pi_t\left(\frac{b_1}{\sigma_1}\right) + \widehat{\widetilde{W}_t^1 L^X f_t} + h_t^\pi(f) \right\}\mathrm{d}t + \mathrm{d}m_t^2,$$

where $m_t^2 = \int_0^t \{\widetilde{W}_s^1 h_s^\pi(f) + \widehat{Z}_s\}\,\mathrm{d}I_s + \int_0^t \widetilde{W}_s^1 \int_{\mathbb{R}} w_s^\pi(f,x) m^\pi(\mathrm{d}s,\mathrm{d}x)$ is a $(P, \mathcal{F}_t^Y)$-local martingale.

Since $\widehat{Z_t \widetilde{W}_t^1} = \widehat{Z}_t \widetilde{W}_t^1$, they have the same limited variation parts, which means that

$$Z_t \widehat{\frac{b_1(t)}{\sigma_1(t)}} + \widehat{\widetilde{W}_t^1 L^X f_t} + \widehat{\frac{\partial f(t)}{\partial x}}\sigma_0(t)\rho = \widehat{Z}_t \pi_t\left(\frac{b_1}{\sigma_1}\right) + \widehat{\widetilde{W}_t^1 L^X f_t} + h_t^\pi(f) \quad \text{on } \{t \leq \widetilde{\tau}_n\}.$$

Equivalently,

$$h_t^\pi(f) = \pi_t\left(f\frac{b_1}{\sigma_1}\right) - \pi_t(f)\pi_t\left(\frac{b_1}{\sigma_1}\right) + \pi_t\left(\sigma_0\frac{\partial f}{\partial x}\right)\rho \quad \text{on } \{t \leq \widetilde{\tau}_n\}.$$

Now, when $n \to \infty$, $\widetilde{\tau}_n$ goes to T, P-a.s. and so the process $h^\pi(f)$ is completely defined.

Step 2. We now choose the bounded process

$$U_t := \int_0^t \int_{\mathbb{R}} \Gamma(s, x) m(\mathrm{d}s, \mathrm{d}x) = \int_0^t \int_Z \mathbf{1}_{D_s}(\zeta) \Gamma(s, K_1(s, \zeta)) N(\mathrm{d}s, \mathrm{d}\zeta)$$

(see (2.4) for the last equality). Then

$$
\begin{aligned}
\mathrm{d}(Z_t U_t) &= Z_{t-} \, \mathrm{d}U_t + U_{t-} \, \mathrm{d}Z_t + \mathrm{d}[Z, U]_t \\
&= \left\{ U_t L^X f_t + V_t + Z_t \int_{\mathbb{R}} \Gamma(t, x) \lambda_t \phi_t(\mathrm{d}x) \right\} \mathrm{d}t + \mathrm{d}m_t^3,
\end{aligned} \tag{3.8}
$$

where $V_t := \int_Z \mathbf{1}_{D_t}(\zeta) \{ f(t, X_{t-} + K_0(t, \zeta)) - f(t, X_{t-}) \} \Gamma(t, K_1(t, \zeta)) \nu(\mathrm{d}\zeta)$ and

$$
\begin{aligned}
m_t^3 &= \int_0^t \int_{\mathbb{R}} Z_{s-} \Gamma(s, x)(m(\mathrm{d}s, \mathrm{d}x) - \lambda_s \phi_s(\mathrm{d}x) \, \mathrm{d}s) + \int_0^t \frac{\partial f}{\partial x}(s) \sigma_0(s) U_s \, \mathrm{d}W_s^0 \\
&\quad + \int_0^t \int_Z \{ f(s, X_{s-} + K_0(s, \zeta)) - f(s, X_{s-}) \} \{ \mathbf{1}_{D_s}(\zeta) \Gamma(s, K_1(s, \zeta)) + U_{s-} \} \widetilde{N}(\mathrm{d}s, \mathrm{d}\zeta)
\end{aligned}
$$

is a $(\mathrm{P}, \mathcal{F}_t)$-martingale. By projecting onto $\mathcal{F}_t^Y$, (3.8) becomes

$$\mathrm{d}(\widehat{Z_t U_t}) = \left\{ \widehat{U_t L^X f_t} + \widehat{V_t} + \int_{\mathbb{R}} \Gamma(t, x) \widehat{Z_t \lambda_t \phi_t}(\mathrm{d}x) \right\} \mathrm{d}t + \mathrm{d}\widetilde{m}_t^3 \tag{3.9}$$

with $\widetilde{m}^3$ a $(\mathrm{P}, \mathcal{F}_t^Y)$-martingale.

On the other hand,

$$
\begin{aligned}
\mathrm{d}(\widehat{Z}_t U_t) &= \widehat{Z}_{t-} \, \mathrm{d}U_t + U_{t-} \, \mathrm{d}\widehat{Z}_t + \mathrm{d}[\widehat{Z}, U]_t \\
&= \left\{ \int_{\mathbb{R}} (\widehat{Z}_t + w_t^\pi(f, x)) \Gamma(t, x) \widehat{\lambda_t \phi_t}(\mathrm{d}x) + U_t \widehat{L^X f_t} \right\} \mathrm{d}t + \mathrm{d}m_t^4,
\end{aligned} \tag{3.10}
$$

where m^4 is the $(\mathrm{P}, \mathcal{F}_t^Y)$-martingale given by

$$m_t^4 = \int_0^t U_s h_s^\pi(f) \, \mathrm{d}I_s + \int_0^t \{ \widehat{Z}_{s-} \Gamma(s, x) + U_{s-} w_s^\pi(f, x) \} m^\pi(\mathrm{d}s, \mathrm{d}x).$$

As in step 1, the finite-variation parts in (3.9) and (3.10) must be equal, so

$$\int_{\mathbb{R}} w_t^\pi(f, x) \Gamma(t, x) \widehat{\lambda_t \phi_t}(\mathrm{d}x) = \int_{\mathbb{R}} \Gamma(t, x) \widehat{Z_t \lambda_t \phi_t}(\mathrm{d}x) + \widehat{V_t} - \int_{\mathbb{R}} \widehat{Z}_t \Gamma(t, x) \widehat{\lambda_t \phi_t}(\mathrm{d}x). \tag{3.11}$$

Now we look for $w^\pi(f, x)$ having the following structure:

$$w_t^\pi(f, x) = w_1(t, f, x) - w_2(t, f, x) + w_3(t, f, x).$$

We can always choose $w_2(t, f, x) = \widehat{Z}_{t-}$, and, by equality (3.11), w_1 and w_3 need to satisfy

$$\int_{\mathbb{R}} w_1(t, f, x) \Gamma(t, x) \widehat{\lambda_t \phi_t}(\mathrm{d}x) = \int_{\mathbb{R}} \Gamma(t, x) \widehat{Z_{t-} \lambda_t \phi_t}(\mathrm{d}x),$$

$$\int_{\mathbb{R}} w_3(t, f, x) \Gamma(t, x) \widehat{\lambda_t \phi_t}(\mathrm{d}x) = \widehat{V_t}.$$

Denoting by $\{T_n\}$ the sequence of jump times of Y (i.e. of $N_t = m([0, t) \times \mathbb{R})$), we select $\Gamma(t, x)$ of the form $\Gamma(t, x) = C_t \mathbf{1}_A(x) \mathbf{1}_{\{t \leq T_n \wedge T\}}$ with C any bounded, $\mathcal{F}_t^Y$-predictable, positive process and $A \in \mathcal{B}(\mathbb{R})$.

With this choice, the process $U_t := \int_0^t \int_{\mathbb{R}} \Gamma(s, x) m(ds, dx)$ is bounded since $|U_t| \leq \int_0^{T \wedge T_n} |C_s| \, dN_s \leq Dn$, where D is a suitable positive constant. Then, on $\{t \leq T_n \wedge T\}$,

$$V_t = \int_Z C_t \mathbf{1}_{D_t^A}(\zeta)\{f(t, X_{t^-} + K_0(t, X_{t^-}, \zeta)) - f(t, X_{t^-})\}\nu(d\zeta)$$

$$= C_t \int_{D_t^A} \{f(t, X_{t^-} + K_0(t, X_{t^-}, \zeta)) - f(t, X_{t^-})\}\nu(d\zeta).$$

If we let $\int_{d^A(t,x,y)}\{f(t, x + K_0(t, x, \zeta)) - f(t, x)\}\nu(d\zeta) =: \bar{L}f(t, x, y, A)$ then we obtain, for all $A \in \mathcal{B}(\mathbb{R})$ on $\{t \leq T_n \wedge T\}$,

$$\int_A w_3(t, f, x)\widehat{\lambda_t \phi_t}(dx) = \int_A \widehat{\bar{L}f}(X_{t^-}, Y_{t^-}, dx),$$

$$\int_A w_1(t, f, x)\widehat{\lambda_t \phi_t}(dx) = \int_A \widehat{Z_{t^-}\lambda_t \phi_t}(dx).$$

Thus, on $\{t \leq T_n \wedge T\}$,

$$w_1(t, f, x) - w_2(t, f, x) + w_3(t, f, x) = \frac{d\pi_{t^-}(\lambda\phi f)}{d\pi_{t^-}(\lambda\phi)}(x) - \pi_{t^-}(f) + \frac{d\pi_{t^-}(\bar{L}f)}{d\pi_{t^-}(\lambda\phi)}(x).$$

Now, since the counting process $N_t = m((0, t] \times \mathbb{R})$ is nonexplosive, T_n goes to ∞ with n and this concludes the proof.

It can be observed that the KS equation (3.3) can also be written as

$$\pi_t(f) = f(0, x_0) + \int_0^t \{\pi_s(L_0^X f) + \pi_s(f)\pi_s(\lambda_s) - \pi_s(f\lambda_s)\} \, ds$$

$$+ \int_0^t \int_{\mathbb{R}} w_s^\pi(f, x)m(ds, dx) + \int_0^t h_s^\pi(f) \, dI_s,$$

where

$$L_0^X f(t, x, y) = L^X f(t, x) - \bar{L}f(t, x, y, \mathbb{R})$$

$$= \frac{\partial f}{\partial t}(t, x) + b_0(t, x)\frac{\partial f}{\partial x} + \frac{1}{2}\sigma_0^2(t, x)\frac{\partial^2 f}{\partial x^2}$$

$$+ \int_{d^{1c}(t,x,y)} \{f(t, x + K_0(t, x, \zeta)) - f(t, x)\}\nu(d\zeta)$$

and $d^{1c}(t, x, y) = \{\zeta \in Z : K_1(t, x, y, \zeta) = 0\}$, and it has a natural recursive structure. It is easier to see it by writing the equation at jump times and between two consecutive jump times of Y. Indeed, if T_n is a jump time for the process Y that occurs before time T,

$$\pi_{T_n}(f) = \frac{d\pi_{T_n^-}(\lambda_{T_n}\phi_{T_n}f)}{d\pi_{T_n^-}(\lambda_{T_n}\phi_{T_n})}(Z_n) + \frac{d\pi_{T_n^-}(\bar{L}_{T_n}f)}{d\pi_{T_n^-}(\lambda_{T_n}\phi_{T_n})}(Z_n), \qquad Z_n = Y_{T_n} - Y_{T_{n-1}}.$$

Hence, $\pi_{T_n}(f)$ is completely determined by the observed data (T_n, Z_n) and by the knowledge of $\pi_t(f)$ for all $t \in [T_{n-1}, T_n)$, since $\pi_{T_n^-}(f) = \lim_{t \to T_n^-} \pi_t(f)$.

Then, for $t \in [T_n, T_{n+1} \wedge T)$,

$$\pi_t(f) = \pi_{T_n}(f) + \int_{T_n}^t \{\pi_s(L_0^X f) + \pi_s(f)\pi_s(\lambda_s) - \pi_s(f\lambda_s)\}\, \mathrm{d}s + \int_{T_n}^t h_s^\pi(f)\, \mathrm{d}I_s.$$

To show uniqueness for the solution to the KS equation, we want to proceed as in [17], but we need to know exactly the shape of the generator of the pair (X, Y). To this end, we give the following lemma.

Lemma 3.2. *Under (2.2), (2.6), and (3.1), (X, Y) is a $(\mathrm{P}, \mathcal{F}_t)$-Markov process with generator $L^{X,Y}$ defined by, for all $f \in \mathcal{C}_b^{1,2,2}([0, T] \times \mathbb{R} \times \mathbb{R})$,*

$$L^{X,Y} f(t, x, y) = \frac{\partial f}{\partial t} + b_0(t, x)\frac{\partial f}{\partial x} + b_1(t, x, y)\frac{\partial f}{\partial y} + \rho\sigma_0(t, x)\sigma_1(t, y)\frac{\partial^2 f}{\partial x \partial y}$$

$$+ \frac{1}{2}\sigma_0^2(t, x)\frac{\partial^2 f}{\partial x^2} + \frac{1}{2}\sigma_1^2(t, y)\frac{\partial^2 f}{\partial y^2}$$

$$+ \int_Z \{f(t, x + K_0(t, x, \zeta), y + K_1(t, x, y, \zeta)) - f(t, x, y)\}\nu(\mathrm{d}\zeta).$$

Proof. By the assumptions of existence and uniqueness for the solution of system (2.1), the martingale problem for the operator $L^{X,Y}$ is well posed and this implies that the pair (X, Y) is a $(\mathrm{P}, \mathcal{F}_t)$-Markov process. Then the proof consists of applying Itô's formula to a $\mathcal{C}_b^{1,2,2}([0, T] \times \mathbb{R} \times \mathbb{R})$ function $f(t, x, y)$:

$$\mathrm{d}f(t, X_t, Y_t) = L^{X,Y} f(t, X_t, Y_t)\, \mathrm{d}t + \sigma_0(t)\frac{\partial f(t, X_t, Y_t)}{\partial x}\, \mathrm{d}W_t^0 + \sigma_1(t)\frac{\partial f(t, X_t, Y_t)}{\partial y}\, \mathrm{d}W_t^1$$

$$+ \int_Z \{f(t, X_{t-} + K_0(t, \zeta), Y_{t-} + K_1(t, \zeta)) - f(t, X_{t-}, Y_{t-})\}\widetilde{N}(\mathrm{d}t, \mathrm{d}\zeta)$$

$$= L^{X,Y} f(t, X_t, Y_t)\, \mathrm{d}t + \mathrm{d}M_t^f. \tag{3.12}$$

Finally, by (2.2), (2.6), and (3.1), since

$$\mathrm{E}\int_0^T \int_Z |f(t, X_{t-} + K_0(t, \zeta), Y_{t-} + K_1(t, \zeta)) - f(t, X_{t-}, Y_{t-})|\nu(\mathrm{d}\zeta)\, \mathrm{d}t$$

$$\leq 2\|f\|\, \mathrm{E}\int_0^T \{\nu(D_t^0) + \nu(D_t)\}\, \mathrm{d}t$$

$$< \infty,$$

M^f is a $(\mathrm{P}, \mathcal{F}_t)$-martingale.

Remark 3.4. By projecting (3.12) onto $\mathcal{F}_t^Y$ we can state that

$$\pi_t(f(\cdot, Y_t)) - \int_0^t \pi_s(L^{X,Y} f(\cdot, Y_s))\, \mathrm{d}s$$

is a $(\mathrm{P}, \mathcal{F}_t^Y)$-martingale for each $f \in \mathcal{C}_b^{1,2,2}([0, T] \times \mathbb{R} \times \mathbb{R})$.

We want to use this martingale property to characterize the distribution of the pair (π, Y) by exploiting the idea given in [17]; therefore, we introduce the notion of the *filtered martingale problem*.

Definition 3.1. A process (μ, U) defined on a probability space $(\widetilde{\Omega}, \widetilde{\mathcal{F}}_t, \widetilde{P})$, with càdlàg trajectories and taking values in $\mathcal{P}(\mathbb{R}) \times \mathbb{R}$, is a solution of the filtered martingale problem FMP($L^{X,Y}$, x_0, y_0) if μ is $\mathcal{F}_t^U$-adapted and, for all $f \in \mathcal{C}_b^{1,2,2}([0, T] \times \mathbb{R} \times \mathbb{R})$,

$$\mu_t(f(\cdot, U_t)) - \int_0^t \mu_s(L^{X,Y} f(\cdot, U_s)) \, \mathrm{d}s$$

is a $(\widetilde{P}, \mathcal{F}_t^U)$-martingale and $\mathrm{E}^{\widetilde{P}}[\mu_0 f(\cdot, U_0)] = f(0, x_0, y_0)$.

Now we are ready to give the definition of a weak solution of the filtering equation.

Definition 3.2. A weak solution to the KS equation (3.3) is a process $(\mu, \widetilde{Y})$ defined on a probability space $(\widetilde{\Omega}, \widetilde{\mathcal{F}}_t, \widetilde{P})$ with càdlàg trajectories, taking values on $\mathcal{P}(\mathbb{R}) \times \mathbb{R}$, such that $\widetilde{Y}_0 = y_0$, $\widetilde{P}$-a.s., $\mathrm{E}^{\widetilde{P}}[\mu_0(f)] = f(0, x_0)$ for all $f \in \mathcal{C}_b^{1,2}([0, T] \times \mathbb{R})$, and the following statements hold.

(i) μ is $\mathcal{F}_t^{\widetilde{Y}}$-adapted.

(ii) The $(\widetilde{P}, \mathcal{F}_t^{\widetilde{Y}})$-predictable projection of the counting measure associated to the jumps of $\widetilde{Y}$, $\widetilde{m}(\mathrm{d}t, \mathrm{d}x)$, is given by $\mu_{t-}(\lambda_t \phi_t(\mathrm{d}x)) \, \mathrm{d}t$.

(iii) $\widetilde{I}_t = \int_0^t (1/\sigma_1(s))(\mathrm{d}\widetilde{Y}_s - \int_{\mathbb{R}} x \widetilde{m}(\mathrm{d}s, \mathrm{d}x))$ is a $(\widetilde{P}, \mathcal{F}_t^{\widetilde{Y}})$-Brownian motion.

(iv) The pair $(\mu, \widetilde{Y})$ solves the KS equation (3.3), with $m^\pi(\mathrm{d}t, \mathrm{d}x)$, I_t, $w_t^\pi(f, x)$, and $h_t^\pi(f)$ replaced by $m^\mu(\mathrm{d}t, \mathrm{d}x) = \widetilde{m}(\mathrm{d}t, \mathrm{d}x) - \mu_{t-}(\lambda_t \phi_t(\mathrm{d}x)) \, \mathrm{d}t$, $I_t^\mu = \widetilde{I}_t - \int_0^t \mu_s(b_1/\sigma_1) \, \mathrm{d}s$, $w_t^\mu(f, x)$, and $h_t^\mu(f)$, respectively.

(v) $\int_0^T \mu_t(b_2(\cdot, \widetilde{Y}_t)) \, \mathrm{d}t < \infty$, $\widetilde{P}$-a.s., with

$$b_2(t, x, y) = \lambda(t, x, y) + |b_0(t, x)| + \sigma_0^2(t, x) + \nu(d_0(t, x)) + \frac{b_1^2(t, x, y)}{\sigma_1^2(t, y)}.$$

Remark 3.5. In Definition 3.2(ii) we mean that, for all $A \in \mathcal{B}(\mathbb{R})$,

$$\mu_{t-}(\lambda_t \phi_t(A)) = \mu_{t-}(\nu(d^A(\cdot, \widetilde{Y}_{t-})))$$

is the $(\widetilde{P}, \mathcal{F}_t^{\widetilde{Y}})$-intensity of the counting process $\widetilde{m}((0, t] \times A)$. In particular,

$$\mu_{t-}(\lambda_t \phi_t(\mathbb{R})) = \mu_{t-}(\nu(d^1(\cdot, \widetilde{Y}_{t-}))) = \mu_{t-}(\lambda(\cdot, \widetilde{Y}_{t-})) = \mu_{t-}(\lambda)$$

is the $(\widetilde{P}, \mathcal{F}_t^{\widetilde{Y}})$-intensity of the point process $\widetilde{m}((0, t] \times \mathbb{R})$, where we have used the notation $\lambda(t, x, y) = \nu(d^1(t, x, y))$.

Remark 3.6. Taking into account Definition 3.2(v) we can prove that, for all $f \in \mathcal{C}_b^{1,2,2}([0, T] \times \mathbb{R} \times \mathbb{R})$,

$$\int_0^T \int_{\mathbb{R}} |w_t^\mu(f, x)| \mu_{t-}(\lambda_t \phi_t(\mathrm{d}x)) \, \mathrm{d}t \leq 4\|f\| \int_0^T \mu_t(\lambda) \, \mathrm{d}t$$

$$\leq 4\|f\| \int_0^T \mu_t(b_2) \, \mathrm{d}t$$

$$< \infty \quad \widetilde{P}\text{-a.s.,}$$

$$\int_0^T h_t^\mu(f)^2 \, dt \le B_f \int_0^T \left\{ \mu_t(\sigma_0^2) + \frac{\mu_t(b_1^2(\cdot, \widetilde{Y}_t))}{\sigma_1^2(t, \widetilde{Y}_t)} \right\} dt$$

$$\le B_f \int_0^T \mu_t(b_2) \, dt$$

$$< \infty \quad \widetilde{P}\text{-a.s.},$$

$$\int_0^T |\mu_t(L^X f)| \, dt \le \widetilde{B}_f \int_0^T \{ 1 + |\mu_t(b_0)| + \mu_t(\sigma_0)^2 + \mu_t(\nu(d_0)) \} \, dt$$

$$\le \widetilde{B}_f \int_0^T \mu_t(b_2) \, dt$$

$$< \infty \quad \widetilde{P}\text{-a.s.},$$

with B_f and $\widetilde{B}_f$ suitable positive constants. Thus, all the stochastic integrals in the KS equation considered in Definition 3.2(iv) are well defined, and those driven by I^μ and by $m^\mu(dt, dx) = \widetilde{m}(dt, dx) - \mu_{t-}(\lambda_t \phi_t(dx)) \, dt$ are $(\widetilde{P}, \mathcal{F}_t^{\widetilde{Y}})$-local martingales.

Remark 3.7. Note that the pair filter observation, (π, Y), is a weak solution to (3.3). As a matter of fact, Definition 3.2(i), (ii), (iv), and (v) are trivially verified. For Definition 3.2(iii), consider the probability measure Q defined in (2.11); then, by Girsanov's theorem, the process

$$\widetilde{W}_t^1 = W_t^1 + \int_0^t \frac{b_1(s)}{\sigma_1(s)} \, ds = I_t + \int_0^t \pi_s\left(\frac{b_1}{\sigma_1}\right) ds = \int_0^t \frac{1}{\sigma_1(s)} \left(dY_s - \int_{\mathbb{R}} x m(ds, dx) \right)$$

is a $(Q, \mathcal{F}_t^Y)$-Wiener process.

Now we can state a weak uniqueness result for the solution of the KS equation whose proof is postponed to Appendix B.

Theorem 3.2. *Under the hypotheses of Theorem 3.1, uniqueness for the solutions to* FMP($L^{X,Y}$, x_0, y_0) *implies that all weak solutions* $(\mu, \widetilde{Y})$ *of the KS equation have the same law. In particular, μ_t and π_t have the same law.*

In Appendix B we give a class of sufficient conditions that ensures uniqueness for the solution to the filtered martingale problem for $L^{X,Y}$ (see Proposition B.1).

In the remaining part of this section we discuss pathwise uniqueness for the solution of the KS equation. We start by giving the definition of the strong solution.

Definition 3.3. A strong solution for the KS equation is a $(P, \mathcal{F}_t^Y)$-adapted, càdlàg, $\mathcal{P}(\mathbb{R})$-valued process $\{\mu_t\}_{t \in [0,T]}$ such that $\int_0^T \mu_s(b_2) \, ds < \infty$, P-a.s. ($b_2$ is defined in Definition 3.2(v)) solves the KS equation, that is, for all $f \in \mathcal{C}_b^{1,2,2}([0, T] \times \mathbb{R} \times \mathbb{R})$ and all $t \le T$,

$$\mu_t(f) = \pi_0(f) + \int_0^t \mu_s(L^X f) \, ds + \int_0^t \int_{\mathbb{R}} w_s^\mu(f, x) m^\mu(ds, dx) + \int_0^t h_s^\mu(f) \, dI_s^\mu, \quad (3.13)$$

where

$$dI_t^\mu = dW_t^1 + \left\{ \frac{b_1(t)}{\sigma_1(t)} - \mu_t\left(\frac{b_1}{\sigma_1}\right) \right\} dt, \qquad m^\mu(dt, dx) = m(dt, dx) - \mu_{t-}(\lambda_t \phi_t(dx)) \, dt,$$

and $w^\mu(f, x)$ and $h^\mu(f)$ are defined respectively in (3.4) and (3.5), replacing π with μ.

Note that the condition $\int_0^T \mu_s(b_2) \, ds < \infty$, P-a.s. makes the integrals in (3.13) well defined.

Theorem 3.3. *Let (X, Y) be defined as in (2.1), and assume that uniqueness holds for the FMP($L^{X,Y}, x_0, y_0$). Let $\{\mu_t\}_{t \in [0,T]}$ be a strong solution of the KS equation such that $\mu_{t-}(\lambda_t \phi_t(dx)) \, dt$ and $\pi_{t-}(\lambda_t \phi_t(dx)) \, dt$ are equivalent measures over $[0, T] \times \mathbb{R}$. Then $\mu_t = \pi_t$, P-a.s. for all $t \leq T$.*

Proof. See Appendix B.

We conclude this section by considering a simplified model and giving a sufficient condition which implies that the additional hypothesis in Theorem 3.3 is satisfied.

Example 3.1. (*Observation dynamics driven by independent point processes.*) Suppose that there exists a finite set of measurable functions $K_1^i(t, y) \neq 0$ for all $(t, y) \in [0, T] \times \mathbb{R}$, $i = 1, \ldots, n$, such that

$$d^1(t, x, y) := \{\zeta \in Z : K_1(t, x, y, \zeta) \neq 0\} = \bigcup_{i=1}^n d_i^1(t, x, y)$$

and

$$d_i^1(t, x, y) \cap d_j^1(t, x, y) = \varnothing \quad \text{for all } i \neq j,$$

where $d_i^1(t, x, y) := \{\zeta \in Z : K_1(t, x, y, \zeta) = K_1^i(t, y)\}$. This implies that $K^1(t, X_{t-}, Y_{t-}, \zeta) = \sum_{i=1}^n K_i^1(t, Y_{t-}) \mathbf{1}_{D_t^i}(\zeta)$ with $D_t^i = d_i^1(t, X_{t-}, Y_{t-})$. It is not difficult to see that the observation process Y has the following dynamics:

$$dY_t = b_1(t, X_t, Y_t) \, dt + \sigma_1(t, Y_t) \, dW_t^1 + \sum_{i=1}^n K_1^i(t, Y_{t-}) \, dN_t^i. \qquad (3.14)$$

Here $N_t^i = N((0, t] \times D_t^i)$ for $i = 1, \ldots, n$ are independent counting processes with $(P, \mathcal{F}_t)$-intensities given by $\lambda_t^i = \nu(D_t^i)$. Let us point out that the signal X influences the drift and the intensities of the point process driving the observation dynamics but not the jump coefficients $K_t^i(t, Y_{t-})$ for $i = 1, \ldots, n$, which are observable. In such a model the counting measure $m(dt, dx)$ can be written as

$$m(dt, dx) = \sum_{\{s : \, \Delta Y_s \neq 0\}} \delta_{\{s, \Delta Y_s\}}(dt, dx) = \sum_{i=1}^n \delta_{K_1^i(t, Y_{t-})}(dx) \, dN_t^i,$$

and the $(P, \mathcal{F}_t)$-dual predictable projection of $m(dt, dx)$ becomes

$$
\begin{aligned}
\lambda_t \phi_t(dx) \, dt &= \int_{D_t} \delta_{K_1(t, \zeta)}(dx) \nu(d\zeta) \, dt \\
&= \sum_{i=1}^n \delta_{K_1^i(t, Y_{t-})}(dx) \int_{D_t^i} \nu(d\zeta) \, dt \\
&= \sum_{i=1}^n \delta_{K_1^i(t, Y_{t-})}(dx) \lambda_t^i \, dt.
\end{aligned}
$$

Of course, $\lambda_t = \nu(D_t) = \sum_{i=1}^n \lambda_t^i$ provides the $(P, \mathcal{F}_t)$-intensity of $N_t = m((0, t] \times \mathbb{R})$. We want to verify that, under the assumption that

$$\lambda^i(t, x, y) = \nu(d_i^1(t, x, y)) > 0 \quad \text{for all } (t, x, y) \in [0, T] \times \mathbb{R} \times \mathbb{R}, \, i = 1, \ldots, n, \quad (3.15)$$

for any $\mathcal{F}_t^Y$-adapted, càdlàg, $\mathcal{P}(\mathbb{R})$-valued process μ, the measures $\mu_{t-}(\lambda_t \phi_t(dx)) \, dt$ and

$\pi_{t^-}(\lambda_t\phi_t(dx))\,dt$ are equivalent. Note that we can write

$$\pi_{t^-}(\lambda_t\phi_t(dx))\,dt = \sum_{i=1}^{n} \delta_{K_1^i(t,Y_{t^-})}(dx)\pi_{t^-}(\lambda^i)\,dt,$$

$$\mu_{t^-}(\lambda_t\phi_t(dx))\,dt = \sum_{i=1}^{n} \delta_{K_1^i(t,Y_{t^-})}(dx)\mu_{t^-}(\lambda^i)\,dt,$$

because $\delta_{K_1^i(t,Y_{t^-})}(dx)$ for $i = 1,\ldots,n$ are $\mathcal{F}_t^Y$-measurable. Hypothesis (3.15) implies that $\pi_{t^-}(\lambda^i) > 0$ and $\mu_{t^-}(\lambda^i) > 0$, $i = 1,\ldots,n$, and the Radon–Nikodym derivative of $\mu_{t^-}(\lambda\phi_t(dx))\,dt$ with respect to $\pi_{t^-}(\lambda_t\phi_t(dx))\,dt$ becomes

$$\frac{d\mu_{t^-}(\lambda\phi)}{d\pi_{t^-}(\lambda\phi)}(x) = \frac{\sum_{i=1}^{n}\delta_{K_1^i(t,Y_{t^-})}(x)\mu_{t^-}(\lambda^i)}{\sum_{i=1}^{n}\delta_{K_1^i(t,Y_{t^-})}(x)\pi_{t^-}(\lambda^i)} = \sum_{i=1}^{n}\mathbf{1}_{\{K_1^i(t,Y_t)=x\}}\frac{\mu_{t^-}(\lambda^i)}{\pi_{t^-}(\lambda^i)}.$$

On the other hand, there also exists the Radon–Nikodym derivative of $\pi_{t^-}(\lambda_t\phi_t(dx))\,dt$ with respect to $\mu_{t^-}(\lambda_t\phi_t(dx))\,dt$ given by

$$\frac{d\pi_{t^-}(\lambda\phi)}{d\mu_{t^-}(\lambda\phi)}(x) = \sum_{i=1}^{n}\mathbf{1}_{\{K_1^i(t,Y_t)=x\}}\frac{\pi_{t^-}(\lambda^i)}{\mu_{t^-}(\lambda^i)},$$

and this means that these two measures are equivalent.

We now give the KS equation satisfied by the filter for this simplified model:

$$\pi_t(f) = f(0,x_0) + \int_0^t \pi_s(L^X f)\,ds$$

$$+ \int_0^t \left\{ \sigma_1(s)^{-1}[\pi_s(b_1 f) - \pi_s(b_1)\pi_s(f)] + \rho\pi_s\left(\sigma_0\frac{\partial f}{\partial x}\right)\right\} dI_s$$

$$+ \sum_{i=1}^{n}\int_0^t \pi_{s^-}(\lambda^i)^+[\pi_{s^-}(\lambda^i f) - \pi_{s^-}(f) + \pi_{s^-}(R^i f)](dN_s^i - \pi_{s^-}(\lambda^i)\,dt). \quad (3.16)$$

Here $a^+ := (1/a)\,\mathbf{1}_{\{a>0\}}$ and by R^i we mean the operator

$$R^i f(t,x,Y_{t^-}) = \int_{d_i^1(t,x,Y_{t^-})} \{f(t,x+K_0(t,x,\zeta)) - f(t,x)\}\nu(d\zeta).$$

Now we are in the position to state the following result

Proposition 3.1. *Let the state X be defined by the first equation of (2.1), and let the observations Y be defined by (3.14). Assume that (3.15) and uniqueness for $\mathrm{FMP}(L^{X,Y}, x_0, y_0)$ hold. Let $\{\mu_t\}_{t\in[0,T]}$ be a strong solution of the KS equation given by (3.16), replacing π by μ. Then $\mu_t = \pi_t$, P-a.s. for all $t \leq T$.*

Appendix A

In this appendix we give sufficient conditions (see, for instance, [7] and [14]) that ensure strong existence and strong uniqueness for solutions to system (2.1).

Assumption B. (i) *Let $b_0(t,x)$, $b_1(t,x,y)$, $\sigma_0(t,x)$, and $\sigma_1(t,y)$ be jointly continuous functions of their arguments, and let $K_0(t,x,\zeta)$ and $K_1(t,x,y,\zeta)$ be $\mathbb{R}$-valued, jointly continuous functions in (t,x,y).*

(ii) *Suppose that there exists a constant $C > 0$ such that, for all $t \in [0, T]$,*

$$|b_0(t, x)|^2 + |\sigma_0(t, x)|^2 + \int_Z |K_0(t, x, \zeta)|^2 \nu(\mathrm{d}\zeta) \leq C(1 + |x|^2),$$

$$|b_1(t, x, y)|^2 + |\sigma_1(t, y)|^2 + \int_Z |K_1(t, x, y, \zeta)|^2 \nu(\mathrm{d}\zeta) \leq C(1 + |x|^2 + |y|^2). \tag{A.1}$$

(iii) *For all $r > 0$, there exists a constant $L = L(r) > 0$ such that, for all $x, x', y, y' \in B_r(0) := \{z \in \mathbb{R} : |z| \leq r\}$,*

$$|b_0(t, x) - b_0(t, x')| + |\sigma_0(t, x) - \sigma_0(t, x')| \leq L|x - x'|, \tag{A.2a}$$

$$|b_1(t, x, y) - b_1(t, x', y)| + |\sigma_1(t, y) - \sigma_1(t, y')| \leq L(|x - x'| + |y - y'|), \tag{A.2b}$$

$$\int_Z |K_0(t, x, \zeta) - K_0(t, x', \zeta)|^2 \nu(\mathrm{d}\zeta) \leq L|x - x'|^2, \tag{A.2c}$$

$$\int_Z |K_1(t, x, y, \zeta) - K_1(t, x', y', \zeta)|^2 \nu(\mathrm{d}\zeta) \leq L(|x - x'|^2 + |y - y'|^2). \tag{A.2d}$$

We respectively refer to (A.1) and (A.2) as the *growth conditions* and *local Lipschitz conditions*.

Other classes of conditions which imply strong existence and weak uniqueness of solutions to system (2.1) without requiring continuity of K_i, $i = 0, 1$, can be found in [7, Appendix A].

Appendix B

Proof of Theorem 3.2. Let $(\mu, \widetilde{Y})$ be a weak solution to the KS equation. We will prove that $(\mu, \widetilde{Y})$ solves the stopped FMP($L^{X,Y}, x_0, y_0$). More precisely, we will show that there exists a sequence η_n of $\mathcal{F}_t^{\widetilde{Y}}$-stopping times, where η_n tends to ∞ with n, and probability measures $\widetilde{Q}_n$ equivalent to $\widetilde{P}$ such that

$$\mu_{t \wedge \eta_n}(F(\cdot, \widetilde{Y}_{t \wedge \eta_n})) - \int_0^{t \wedge \eta_n} \mu_s(L^{X,Y} F(\cdot, \widetilde{Y}_{s \wedge \eta_n})) \, \mathrm{d}s \tag{B.1}$$

is a $(\widetilde{Q}_n, \mathcal{F}_t^{\widetilde{Y}})$-martingale for each $F \in \mathcal{C}_b^{1,2,2}([0, T] \times \mathbb{R} \times \mathbb{R})$.

It is sufficient to prove (B.1) for functions of the type $F(t, x, y) = f(t, x)g(y)$. Recall that

$$\mathrm{d}\widetilde{Y}_t = \sigma_1(t) \, \mathrm{d}\widetilde{I}_t + \int_{\mathbb{R}} x\widetilde{m}(\mathrm{d}t, \mathrm{d}x).$$

By applying Itô's formula we obtain

$$\mathrm{d}g(\widetilde{Y}_t) = g'(\widetilde{Y}_{t-}) \, \mathrm{d}\widetilde{Y}_t + \frac{1}{2} g''(\widetilde{Y}_t)\sigma_1^2(t) \, \mathrm{d}t + \int_{\mathbb{R}} [g(\widetilde{Y}_{t-} + x) - g(\widetilde{Y}_{t-}) - g'(\widetilde{Y}_{t-})x]\widetilde{m}(\mathrm{d}t, \mathrm{d}x)$$

$$= g'(\widetilde{Y}_{t-})\sigma_1(t) \, \mathrm{d}\widetilde{I}_t + \frac{1}{2} g''(\widetilde{Y}_t)\sigma_1^2(t) \, \mathrm{d}t + \int_{\mathbb{R}} [g(\widetilde{Y}_{t-} + x) - g(\widetilde{Y}_{t-})]\widetilde{m}(\mathrm{d}t, \mathrm{d}x).$$

Since $(\mu, \widetilde{Y})$ is a weak solution to the KS equation, then

$$\mu_t(f) = f(0, x_0) + \int_0^t \mu_s(L^X f) \, \mathrm{d}s + \int_0^t \int_{\mathbb{R}} w_s^\mu(f, x) m^\mu(\mathrm{d}s, \mathrm{d}x) + \int_0^t h_s^\mu(f) \, \mathrm{d}I_s^\mu,$$

and, by the product rule,

$$
\begin{aligned}
\mathrm{d}(\mu_t(f)g(\widetilde{Y}_t)) \\
&= \mu_{t-}(f)\left(g'(\widetilde{Y}_t)\sigma_1(t)\,\mathrm{d}\widetilde{I}_t + \frac{1}{2}g''(\widetilde{Y}_t)\sigma_1^2(t)\,\mathrm{d}t + \int_{\mathbb{R}}[g(\widetilde{Y}_{t-}+x)-g(\widetilde{Y}_{t-})]\widetilde{m}(\mathrm{d}t,\mathrm{d}x)\right) \\
&\quad + g(\widetilde{Y}_{t-})\left(\mu_t(L^X f)\,\mathrm{d}t + \int_{\mathbb{R}}w_t^\mu(f,x)m^\mu(\mathrm{d}t,\mathrm{d}x) + h_t^\mu(f)\,\mathrm{d}I_t^\mu\right) \\
&\quad + \sigma_1(t)h_t^\mu(f)g'(\widetilde{Y}_t)\,\mathrm{d}t + \int_{\mathbb{R}}w_t^\mu(f,x)(g(\widetilde{Y}_{t-}+x)-g(\widetilde{Y}_{t-}))\widetilde{m}(\mathrm{d}t,\mathrm{d}x) \\
&= \left[g(\widetilde{Y}_t)\mu_t(L^X f) + g'(\widetilde{Y}_t)(\mu_t(f)\mu_t(b_1)+\sigma_1(t)h_t^\mu(f)) + \tfrac{1}{2}\mu_t(f)g''(\widetilde{Y}_t)\sigma_1^2(t)\right]\mathrm{d}t \\
&\quad + \int_{\mathbb{R}}(\mu_t(f)+w_t^\mu(f,x))[g(\widetilde{Y}_{t-}+x)-g(\widetilde{Y}_{t-})]\mu_t(\lambda_t\phi_t(\mathrm{d}x))\,\mathrm{d}t + \mathrm{d}M_t^{fg}, \qquad (\text{B.2})
\end{aligned}
$$

where we used the equalities

$$
\widetilde{m}(\mathrm{d}t,\mathrm{d}x) = m^\mu(\mathrm{d}t,\mathrm{d}x) + \mu_{t-}(\lambda_t\phi_t(\mathrm{d}x))\,\mathrm{d}t, \qquad \mathrm{d}\widetilde{I}_t = \mathrm{d}I_t^\mu + \mu_t\left(\frac{b_1}{\sigma_1}\right)\mathrm{d}t,
$$

and by $\mathrm{d}M_t^{fg}$ we mean

$$
\begin{aligned}
\mathrm{d}M_t^{fg} &= \int_{\mathbb{R}}\{(\mu_{t-}(f)+w_t^\mu(f,x))[g(\widetilde{Y}_{t-}+x)-g(\widetilde{Y}_{t-})]+g(\widetilde{Y}_{t-})w_t^\mu(f,x)\}m^\mu(\mathrm{d}t,\mathrm{d}x) \\
&\quad + \{\mu_t(f)\sigma_1(t)g'(\widetilde{Y}_t)+g(\widetilde{Y}_t)h_t^\mu(f)\}\,\mathrm{d}I_t^\mu.
\end{aligned}
$$

Now we want to introduce a probability measure equivalent to $\widetilde{\mathrm{P}}$ such that M^{fg} is a local martingale. To this end, define $\widetilde{L}_t = \mathcal{E}(\int_0^t \mu_s(b_1/\sigma_1)\,\mathrm{d}\widetilde{I}_s)$. Since $b_1(t)/\sigma_1(t)$ may be unbounded, $\widetilde{L}$ is only a $(\widetilde{\mathrm{P}},\mathcal{F}_t^Y)$-local martingale. Hence, we need to introduce the sequence of $\mathcal{F}_t^Y$-stopping times

$$
\eta_n = T \wedge \inf\left\{t: \int_0^t\left|\mu_s\left(\frac{b_1}{\sigma_1}\right)\right|\mathrm{d}s \geq n\right\} \wedge \inf\left\{t: \int_0^t\mu_s|b_2|\,\mathrm{d}s \geq n\right\},
$$

where $b_2(t,x,y)$ is given in Definition 3.2(v).

For any n, we build a new probability measure equivalent to $\widetilde{\mathrm{P}}$, $\widetilde{\mathrm{Q}}_n$, on $(\Omega,\mathcal{F}_T^Y)$ as

$$
\widetilde{L}_{\eta_n} = \frac{\mathrm{d}\widetilde{\mathrm{Q}}_n}{\mathrm{d}\widetilde{\mathrm{P}}} = \mathcal{E}\left(\int_0^{\eta_n}\mu_s\left(\frac{b_1}{\sigma_1}\right)\mathrm{d}\widetilde{I}_s\right) = \exp\left\{\int_0^{\eta_n}\mu_s\left(\frac{b_1}{\sigma_1}\right)\mathrm{d}\widetilde{I}_s - \frac{1}{2}\int_0^{\eta_n}\mu_s^2\left(\frac{b_1}{\sigma_1}\right)\mathrm{d}s\right\}.
$$

By Girsanov's theorem,

$$
I_t^\mu = \widetilde{I}_t - \int_0^{t\wedge\eta_n}\mu_s\left(\frac{b_1}{\sigma_1}\right)\mathrm{d}s
$$

is a $(\widetilde{\mathrm{Q}}_n,\mathcal{F}_t^Y)$-Brownian motion, and $M_{t\wedge\eta_n}^{fg}$ is a $(\widetilde{\mathrm{Q}}_n,\mathcal{F}_t^Y)$-martingale since the following estimations hold (see Remark 3.6):

$$
\mathrm{E}^{\widetilde{Q}_n}\int_0^{T\wedge\eta_n}\int_{\mathbb{R}}|w_t^\mu(f,x)|\mu_t(\lambda_t\phi_t(\mathrm{d}x))\,\mathrm{d}t \leq 4\|f\|\,\mathrm{E}^{\widetilde{Q}_n}\int_0^{T\wedge\eta_n}\mu_t(b_2)\,\mathrm{d}t \leq 4\|f\|n < \infty,
$$

$$
\mathrm{E}^{\widetilde{Q}_n}\int_0^{T\wedge\eta_n}h_t^\mu(f)^2\,\mathrm{d}t \leq B_f\,\mathrm{E}^{\widetilde{Q}_n}\int_0^{T\wedge\eta_n}\mu_t(b_2)\,\mathrm{d}t \leq B_f n < \infty.
$$

Finally, from the expressions of w^μ and h^μ, and the generator $L^{X,Y}$, (B.2) implies that

$$d\mu_{t\wedge\eta_n}(fg(\widetilde{Y}_{t\wedge\eta_n})) = \mu_{t\wedge\eta_n}(L^{X,Y}fg(\widetilde{Y}_{t\wedge\eta_n}))\,dt + dM^{fg}_{t\wedge\eta_n},$$

with $M^{fg}_{t\wedge\eta_n}$ a $(\widetilde{Q}_n, \mathcal{F}^Y_t)$-martingale, that is to say, the pair $(\mu, \widetilde{Y})$ solves the stopped FMP($L^{X,Y}$, x_0, y_0).

By Corollary 3.4 of [17], if uniqueness holds for the FMP($L^{X,Y}$, x_0, y_0) then there exists a measurable function $H_t \colon D_\mathbb{R}[0, T] \to \mathcal{P}(\mathbb{R})$ such that $\pi_t = H_t(Y)$, P-a.s. and $\mu_t\,\mathbf{1}_{\{t<\eta_n\}} = H_t(\widetilde{Y})\,\mathbf{1}_{\{t\leq\eta_n\}}$, $\widetilde{Q}_n$-a.s.

Since $\widetilde{Q}_n$ is equivalent to $\widetilde{P}$, the equality above becomes $\mu_t\,\mathbf{1}_{\{t<\eta_n\}} = H_t(\widetilde{Y})\,\mathbf{1}_{\{t<\eta_n\}}$, $\widetilde{P}$-a.s., and taking $n \to \infty$ we obtain $\mu_t = H_t(\widetilde{Y})$, $\widetilde{P}$-a.s.

Finally, since

$$dY_t = \sigma_1(t, Y_t)\,d\widetilde{W}^1_t + \int_\mathbb{R} xm(dt, dx)$$

and

$$d\widetilde{Y}_t = \sigma_1(t, \widetilde{Y}_t)\,d\widetilde{I}_t + \int_\mathbb{R} x\widetilde{m}(dt, dx)$$

under $\widetilde{P}$, the process $\widetilde{Y}$ has the same law as the process Y under P. Thus, $(\mu_t, \widetilde{Y}_t)$ and (π_t, Y_t) have the same law; in particular, μ_t and π_t have the same law.

Proof of Theorem 3.3. With the same arguments used to prove equality (B.2), it can be shown that

$$d(\mu_t(f)g(Y_t))$$
$$= [g(Y_t)\mu_t(L^X f) + (\sigma_1(t)h^\mu_t(f) + \mu_t(f)\mu_t(b_1))g'(Y_t) + \tfrac{1}{2}\mu_t(f)g''(Y_t)\sigma^2_1(t)]\,dt$$
$$+ \int_\mathbb{R} (\mu_t(f) + w^\mu_t(f, x))[g(Y_{t-} + x) - g(Y_{t-})]\mu_t(\lambda_t\phi_t(dx))\,dt + dm^{fg}_t,$$

where by m^{fg}_t we mean

$$dm^{fg}_t = \{\sigma_1(t)\mu_t(f)g'(Y_t) + g(Y_t)h^\mu_t(f)\}\,dI^\mu_t$$
$$+ \int_\mathbb{R}\{(\mu_{t-}(f) + w^\mu_t(f, x))[g(Y_{t-} + x) - g(Y_{t-})] + g(Y_{t-})w^\mu_t(f, x)\}m^\mu(dt, dx).$$

We need to define a new probability measure equivalent to P under which m^{fg} is a local martingale.

From the hypothesis of equivalence of the measures $\pi_{t-}(\lambda_t\phi_t(dx))\,dt$ and $\mu_{t-}(\lambda_t\phi_t(dx))\,dt$, there exists an $\mathcal{F}^Y_t$-predictable process $\Psi(t, x) > -1$, $\pi_t(\lambda_t\phi_t(dx))\,dt$-almost everywhere such that

$$(1 + \Psi(t, x))\pi_{t-}(\lambda_t\phi_t(dx))\,dt = \mu_{t-}(\lambda_t\phi_t(dx))\,dt.$$

Recalling that $I^\mu_t = I_t - \int_0^t\{\mu_s(b_1/\sigma_1) - \pi_s(b_1/\sigma_1)\}\,ds$, we define

$$\tau_n := \inf\left\{t \geq 0 \colon \int_0^t\left|\mu_s\left(\frac{b_1}{\sigma_1}\right) - \pi_s\left(\frac{b_1}{\sigma_1}\right)\right|^2\,ds \geq n\right\} \wedge \inf\left\{t \geq 0 \colon \int_0^t|\mu_s(b_2)|\,ds \geq n\right\}$$
$$\wedge \inf\left\{t \geq 0 \colon \int_0^t\int_\mathbb{R}|\Psi(s, x)|^2\pi_s(\lambda_s\phi_s(dx))\,ds \geq n\right\} \wedge T$$

and the change of measure

$$\frac{dQ_n}{dP}\bigg|_{\mathcal{F}_t^Y} = \Lambda_{t \wedge \tau_n},$$

where

$$\Lambda_{t \wedge \tau_n} = \mathcal{E}\left(\int_0^{t \wedge \tau_n} \left\{ \mu_s \left(\frac{b_1}{\sigma_1} \right) - \pi_s \left(\frac{b_1}{\sigma_1} \right) \right\} dI_s + \int_0^{t \wedge \tau_n} \int_{\mathbb{R}} \Psi(s, x) m^{\pi}(dt, dx) \right).$$

Girsanov's theorem implies that $I_t - \int_0^t \{ \mu_s(b_1/\sigma_1) - \pi_s(b_1/\sigma_1) \} \mathbf{1}_{\{s < \tau_n\}} ds$ is a $(Q_n, \mathcal{F}_t^Y)$-Brownian motion and that the $(Q_n, \mathcal{F}_t^Y)$-predictable projection of the measure $m(dx, dt)$ on $\{t < \tau_n\}$ is $\mu_{t^-}(\lambda_t \phi_t(dx)) dt$.

By performing similar computations as in the proof of Theorem 3.2 we find that $m_{t \wedge \tau_n}^{fg}$ is a $(Q_n, \mathcal{F}_t^Y)$-martingale and so the pair (μ, Y) solves the stopped FMP($L^{X,Y}, x_0, y_0$). Finally, by Corollary 3.4 of [17], there exists a functional H such that

$$\pi_t = H_t(Y) \quad \text{P-a.s.} \quad \text{and} \quad \mu_t \mathbf{1}_{\{t < \tau_n\}} = H_t(Y) \mathbf{1}_{\{t < \tau_n\}} \quad Q_n \text{-a.s.}$$

Nevertheless, Q_n and P are equivalent measures; therefore,

$$\mu_t \mathbf{1}_{\{t < \tau_n\}} = H_t(Y) \mathbf{1}_{\{t < \tau_n\}} \quad \text{P-a.s.}$$

with τ_n an increasing sequence, so there exists, P-a.s., $n(\omega)$ such that, for all $n > n(\omega)$, $\tau_n(\omega) = T$. Taking $n \to \infty$, we obtain $\mu_t = \pi_t$, P-a.s.

In the next proposition we provide sufficient conditions for the uniqueness of the solutions to FMP($L^{X,Y}, x_0, y_0$).

Proposition B.1. *Under Assumption B and either*

$$\sup_{t,x} \nu(d^0(t, x)) + \sup_{t,x,y} \nu(d^1(t, x, y)) < \infty \tag{B.3}$$

or

$$\sup_{t,x,y} \int_Z \{ |K_0(t, x, \zeta)| + |K_1(t, x, y, \zeta)| \} \nu(d\zeta) < \infty, \tag{B.4}$$

uniqueness holds for FMP($L^{X,Y} x_0, y_0$).

Proof. It is sufficient to apply Theorem 3.3 of [17] after having checked that the hypotheses are satisfied. By Assumption B, the martingale problem for $L^{X,Y}$ is well posed. Furthermore, we have to prove that we can choose as a domain for $L^{X,Y}$, a set of functions $\mathcal{D}_L \subset \mathcal{C}_b^{1,2,2}([0, T] \times \mathbb{R} \times \mathbb{R})$ such that, for $f \in \mathcal{D}_L$, $L^{X,Y} f \in \mathcal{C}_b([0, T] \times \mathbb{R} \times \mathbb{R})$. We choose as $\mathcal{D}_L$ the set of functions in $\mathcal{C}_b^{1,2,2}([0, T] \times \mathbb{R} \times \mathbb{R})$ having compact support with respect to (x, y) uniformly in t; then there exists $R_f > 0$ such that, for $|x| > R_f$ and $|y| > R_f$, $f(t, x, y) = 0$ for all $t \in [0, T]$.

Recalling the structure of the operator $L^{X,Y}$, i.e.

$$L^{X,Y} f(t, x, y) = \frac{\partial f}{\partial t} + b_0(t, x) \frac{\partial f}{\partial x} + b_1(t, x, y) \frac{\partial f}{\partial y}$$

$$+ \frac{1}{2} \left[\sigma_0^2(t, x) \frac{\partial^2 f}{\partial x^2} + 2\rho \sigma_0(t, x) \sigma_1(t, y) \frac{\partial^2 f}{\partial x \partial y} + \sigma_1^2(t, y) \frac{\partial^2 f}{\partial y^2} \right]$$

$$+ \int_Z (f(t, x + K_0(t, x, \zeta), y + K_1(t, x, y, \zeta)) - f(t, x, y)) \nu(d\zeta),$$

since

$$\left| \int_Z (f(t, x + K_0(t, x, \zeta), y + K_1(t, x, y, \zeta)) - f(t, x, y))\nu(\mathrm{d}\zeta) \right|$$
$$\leq 2\|f\|\nu(d^0(t, x, y) \cup d^1(t, x, y)),$$

under (B.3) and (A.1), we find that, for all $f \in \mathcal{D}_L$, there exists a constant $C_f > 0$ such that

$$\|L^{X,Y} f\| \leq \left\| \frac{\partial f}{\partial t} \right\| + C_f (1 + R_f^2) + 2\|f\| \sup_{t,x,y} \nu(d_0(t, x, y) \cup d_1(t, x, y)).$$

Hence, $L^{X,Y} f$ is bounded.

The same result can be obtained under (B.4). In fact,

$$\left| \int_Z (f(t, x + K_0(t, x, , \zeta), y + K_1(t, x, y, \zeta)) - f(t, x, y))\nu(\mathrm{d}\zeta) \right|$$
$$\leq \max\left\{ \left\| \frac{\partial f}{\partial x} \right\|, \left\| \frac{\partial f}{\partial y} \right\| \right\} \int_Z \{|K_0(t, x, \zeta)| + |K_1(t, x, y, \zeta)|\}\nu(\mathrm{d}\zeta).$$

Finally, in both cases, the continuity of $L^{X,Y} f(t, x, y)$ can be obtained by the dominated convergence theorem.

Acknowledgement

References

[1] BJÖRK, T., KABANOV, Y. AND RUNGGALDIER, W. (1997). Bond market structure in the presence of marked point processes. *Math. Finance* **7,** 211–239.

[2] BRÉMAUD, P. (1980). *Point Processes and Queues*. Springer, New York.

[3] CECI, C. (2006). Risk minimizing hedging for a partially observed high frequency data model. *Stochastics* **78,** 13–31.

[4] CECI, C. (2012). Utility maximization with intermediate consumption under restricted information for jump market models. To appear in *Internat. J. Theoret. Appl. Finance*.

[5] CECI, C. AND GERARDI, A. (1998). Partially observed control of a Markov jump process with counting observations: equivalence with the separated problem. *Stoch. Process. Appl.* **78,** 245–260.

[6] CECI, C. AND GERARDI, A. (2000). Filtering of a Markov jump process with counting observations. *Appl. Math. Optimization* **42,** 1–18.

[7] CECI, C. AND GERARDI, A. (2006). A mode for high frequency data under partial information: a filtering approach. *Internat. J. Theoret. Appl. Finance* **9,** 555–576.

[8] CRISAN, D. AND ROZOVSKIĬ, B. (eds) (2011). *The Oxford Handbook of Nonlinear Filtering*. Oxford University Press.

[9] EL KAROUI, N., HU NGUYEN, D. AND JEANBLANC-PIQUÉ, M. (1989). Existence of an optimal Markovian filter for the control under partial observations. *SIAM J. Control Optimization* **26,** 1025–1061.

[10] FLEMING, W. H. AND PARDOUX, É. (1982). Optimal control for partially observed diffusions. *SIAM J. Control Optimization* **20,** 261–285.

[11] FREY, R. AND RUNGGALDIER, W. J. (2001). A nonlinear filtering approach to volatility estimation with a view towards high frequency data. *Internat. J. Theoret. Appl. Finance* **4,** 199–210.

[12] FREY, R. AND RUNGGALDIER, W. (2010). Pricing credit derivatives under incomplete information: a nonlinear-filtering approach. *Finance Stoch.* **14,** 495–526.

[13] FREY, R. AND SCHMIDT, T. (2012). Pricing and hedging of credit derivatives via the innovation approach to nonlinear filtering. *Finance Stoch.* **16,** 105–133.

[14] GIHMAN, I. I. AND SKOROHOD, A. V. (1972). *Stochastic Differential Equations*. Springer, New York.

[15] JACOD, J. AND SHIRYAEV, A. N. (2003). *Limit Theorems for Stochastic Processes*, 2nd edn. Springer, Berlin.

[16] KALLIANPUR, G. (1980). *Stochastic Filtering Theory*. Springer, New York.

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

[18] KURTZ, T. G. AND NAPPO, G. (2010). The filtered martingale problem. In *The Oxford Handbook of Nonlinear Filtering*, eds D. Crisan and B. Rozovskiĭ, Oxford University Press, pp. 129–168.

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

[20] LIPSTER, R. S. AND SHIRYAEV, A. (1977). *Statistics of Random Processes I.* Springer, Berlin.

[21] MAZLIAK, L. (1993). Mixed control problem under partial observation. *Appl. Math. Optimization* **27,** 57–84.