# NONLINEAR FILTERING FOR JUMP DIFFUSION PROCESSES WITH A FINANCIAL APPLICATION

Claudia Ceci, Katia Colaneri

*Dipartimento di Scienze, Facoltá di Economia*
*Universitá di Chieti-Pescara, Viale Pindaro 42,*
*I-65127-Pescara, Italy*

**Abstract**

We deal with a filtering problem of a general jump-diffusion process, $X$, when the observations process, $Y$, is a correlated jump-diffusion having common jump times with $X$. In this frame, at any at time $t$ the $\sigma$-algebra $\mathcal{F}_t^Y$ provides all the available information about $X_t$ and the central goal is to characterize the filter, $\pi_t$, that is the conditional distribution of $X_t$ given observations $\mathcal{F}_t^Y$. To this aim, we prove that $\pi_t$ solves Kushner-Stratonovich equation and by applying the Filtered Martingale Problem approach ([18]), that it is the unique weak solution to this equation. Under an additional hypothesis we provide also a pathwise uniqueness result. As an application, we consider a financial market where $Y$ describes the logreturn process of a risky asset $S$ whose dynamics depends on an unobservable stochastic factor $X$. Investors acting on the market can access only to the information flow given by $\{\mathcal{F}_t^S\}_{t\in[0,T]} = \{\mathcal{F}_t^Y\}_{t\in[0,T]}$, generated by stock prices. Thus, we are in presence not only of an incomplete market situation but also of partial information. Assuming the price $S$ of the risky asset modeled directly under a martingale measure we study a risk-minimizing hedging problem, under restricted information, whose solution can be computed via the filter by using a projection result ([26]) .

*Keywords*: Filtering, Jump-Diffusion Processes, Hedging under restricted information, Risk minimization.
*2010 Mathematical Subject Classification*: 93E11; 60J75; 60J60; 91B28.

## 1. Introduction

In this paper we consider a partially observed system (or filtering model), $(X, Y)$, where the signal $X$ follows a general jump-diffusion process and the observation $Y$ is a correlated jump-diffusion having common jump times with $X$. As usual in a filtering model, the signal $X$ cannot be directly observed but we can only observe a stochastic process, $Y$, related to $X$. At any time $t$ the $\sigma$-algebra $\mathcal{F}_t^Y$ gives all the available information about $X_t$. The central goal of filtering theory 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 $X_t$. Filtering applications arise in a great variety of engineering problems, informational sciences and recently in mathematical finance. In particular in this note we deal with a financial application.

Filtering problems have been widely investigated in literature mainly in two cases: when $Y_t$ gives observations of $X_t$ in additional Gaussian noise (see for example [17, 21, 18]) and when $Y_t$ is a counting process or a marked point process (see [3, 20, 12, 6, 7, 4] and reference therein).

The case of mixed type observations (marked point processes and diffusions) has been studied in [13, 14, 5]. All these papers analyze the situation in which the information flow has the structure $\mathcal{F}_t^m \vee \mathcal{F}_t^\eta$, where $m(dt, dx)$ is a marked point process whose dynamics is influenced by a stochastic factor $X$ and $\eta$ gives observations of $X$ in additional Gaussian noise. Anyway, in this note it is considered the situation where the observation is a general jump-diffusion process, that to the authors' knowledge, has not been investigated into existing literature. In a credit derivatives framework, in [13] the marked point process component is given by the default indicator process and in [14] by the loss-state of the portfolio. Both of the models assume intensities of default times to be influenced by the stochastic factor $X$, and the additional Gaussian noise to be independent from $X$ and $m(dt, dx)$. The hypothesis of independence turns to be crucial in [13] for applying a reference probability approach in order to reduce the filtering problem to the case where the information flow consists only of the default history.

[0]Partially supported by PRIN 2008 "Probabilitá e Finanza"

As in [14, 5], we follow an alternative route based on the innovation method which allows us to take into account the case where the signal $X$ and the observation $Y$ are correlated processes. By the innovation method and an $\mathcal{F}_t^Y$-martingale representation theorem we write down the Kushner-Stratonovich equation (KS-equation) that the filter solves. The filtering equation is derived in [14] in the case where the state process $X$ is modeled as a finite-state Markov chain without common jump times with $m(dt, dx)$ and in [5] in the case where the state process $X$ is a jump-diffusion that allows common jump times with $m(dt, dx)$. Anyway, the filtering equation derived in [5] is quite similar to that obtained herein, even if the structure of the information flows are different. But, in this paper under weaker conditions we characterize the filter as the unique weak solution to KS-equation. Moreover, we give pathwise uniqueness under an additional constraint.

As an application we consider a financial market with a bond and a risky asset, whose price dynamics, $S$, follows a geometric jump-diffusion. We assume the logreturn process $Y$, to be affected by an unobservable stochastic factor, $X$, which may describe the activities of other markets, macroeconomics factors or microstructure rules that drive the market. The dynamics of $Y$ and $X$ may be strongly dependent, as provided in the filtering model $(X, Y)$ considered before, in particular these two processes are correlated and may have common jump times. This means that the model considered takes into account also the possibility of catastrophic events, which influence both the asset prices and the hidden state variable driving their dynamics. In this frame investors acting on the market get access only to the information flow, $\{\mathcal{F}_t^S\}_{t\in[0,T]} = \{\mathcal{F}_t^Y\}_{t\in[0,T]}$, generated by the stock price. We assume the risky asset $S$ to be modeled directly under a martingale measure and we deal with the hedging of a contingent claim by the risk minimization criterion which is well suited to deal with restricted information in such a setting (see [26]).

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 Kushner Stratonovich equation is given in Section 3. The proofs of weak and strong uniqueness for this equation are postponed in Appendix B. Here we deduce these results from uniqueness for the filtered martingale problem ([18]). The topic of Section 4 is a financial application of the filtering problem; we discuss the risk minimization hedging problem under restricted information.

## 2. The filtering model

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

$$\begin{cases} dX_t = b_0(t, X_t)dt + \sigma_0(t, X_t)dW_t^0 + \displaystyle\int_Z K_0(t, X_{t^-}, \zeta)N(dt, d\zeta); \qquad X_0 = x_0 \in \mathbb{R} \\ dY_t = b_1(t, X_t, Y_t)dt + \sigma_1(t, Y_t)dW_t^1 + \displaystyle\int_Z K_1(t, X_{t^-}, Y_{t^-}, \zeta)N(dt, d\zeta); \qquad Y_0 = y_0 \in \mathbb{R} \end{cases} \tag{2.1}$$

Here $N(dt, d\zeta)$ is a Poisson random measure on $\mathbb{R}^+ \times Z$, and $\nu(d\zeta)dt$ represents its intensity. Note that $\nu(d\zeta)$ is a $\sigma-$finite measure on a measurable space $(Z, \mathcal{Z})$. The processes $W_t^0$ and $W_t^1$ 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

$$\mathbb{E}\int_0^T \int_Z |K_i(t, \zeta)|\nu(d\zeta)dt < \infty, \quad \mathbb{E}\int_0^T |b_i(t)|dt < \infty, \quad \mathbb{E}\int_0^T \sigma_i^2(t)dt < \infty, \quad i = 0, 1, \tag{2.2}$$

under these constraints, both of the processes $X$ and $Y$ have finite first moment.

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.

We denote by $(\mathcal{F}_t^Y)_{t\geq 0}$ the filtration generated by the observation process $Y$ until time $t$. In the partially observed system considered in this note, at any time $t$, the $\sigma$-algebra $\mathcal{F}_t^Y$ provides all the available information about the signal $X_t$.

By defining $\mathcal{P}(\mathbb{R})$ 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, $\pi_t$, such that

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

for any bounded and measurable function $f(t,x)$ on $[0,T] \times \mathbb{R}$. Since $X_t$ is a càdlàg process, there exists a version of $\pi_t$ with càdlàg paths (see for instance [18]).

¿From now on we will write $\widehat{R}_t$ for the $(P, \mathcal{F}_t^Y)$-optional projection of a progressively measurable process $R_t$, satisfying $\mathbb{E}|R_t| < \infty$, defined as the unique optional process such that for any $\mathcal{F}_t^Y$-stopping time $\tau$, $\widehat{R}_\tau = \mathbb{E}[R_\tau | \mathcal{F}_\tau^Y]$ 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 (it can happen to use both of the notations, $\widehat{f(t, X_t)}$ and $\pi_t(f)$).

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

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

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

We will also need the following result.

**Proposition 2.2** *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 $m_t$, then $\widehat{m}_t$ is a $(P, \mathcal{F}_t^Y)$-local martingale.*

Proof.
By a standard calculation we get, $\forall\ 0 \leq s < t \leq T < \infty$

$$\begin{aligned}
\mathbb{E}\left[\widehat{m}_{t\wedge\tau_n} | \mathcal{F}_s^Y\right] &= \mathbb{E}\left[\mathbb{E}\left[m_{t\wedge\tau_n} | \mathcal{F}_{t\wedge\tau_n}^Y\right] | \mathcal{F}_s^Y\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[m_{t\wedge\tau_n} | \mathcal{F}_{t\wedge\tau_n}^Y\right] \mathbb{I}_{\tau_n > s} | \mathcal{F}_s^Y\right] + \mathbb{E}\left[\mathbb{E}\left[m_{t\wedge\tau_n} | \mathcal{F}_{t\wedge\tau_n}^Y\right] \mathbb{I}_{\tau_n \leq s} | \mathcal{F}_s^Y\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[m_{t\wedge\tau_n} | \mathcal{F}_s\right] | \mathcal{F}_s^Y\right] \mathbb{I}_{\tau_n > s} + \mathbb{E}\left[m_{\tau_n} | \mathcal{F}_{\tau_n}^Y\right] \mathbb{I}_{\tau_n \leq s} \\
&= \widehat{m}_{s\wedge\tau_n}
\end{aligned}$$

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

Let us introduce the integer-valued random measure associated to the jumps of the process $Y$

$$m(dt, dx) = \sum_{s:\Delta Y_s \neq 0} \delta_{\{s, \Delta Y_s\}}(dt, dx) \tag{2.4}$$

where $\delta_a$ denotes the Dirac measure at the point $a$. Note that the following equality holds

$$\int_0^t \int_{\mathbb{R}} x\ m(ds, dx) = \int_0^t \int_Z K_1(s, \zeta) N(ds, d\zeta) \tag{2.5}$$

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

$$\int_0^t \int_{\mathbb{R}} g(x)\, m(ds, dx) = \int_0^t \int_Z \mathbb{I}_{\{K_1(s,\zeta) \neq 0\}} g\left(K_1(s,\zeta)\right) N(ds, d\zeta). \tag{2.6}$$

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

$$d^0(t,x) := \{\zeta \in Z : K_0(t,x,\zeta) \neq 0\}, \qquad d^1(t,x,y) := \{\zeta \in Z : K_1(t,x,y,\zeta) \neq 0\},$$

$$d^A(t,x,y) := \{\zeta \in Z : K_1(t,x,y,\zeta) \in A \setminus \{0\}\} \subseteq d^1(t,x,y), \tag{2.7}$$

and finally,

$$D_t^A = d^A(t, X_{t^-}, Y_{t^-}) \subseteq D_t = d^1(t, X_{t^-}, Y_{t^-}), \quad D_t^0 = d^0(t, X_{t^-}). \tag{2.8}$$

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

Under the assumption

$$\mathbb{E} \int_0^T \nu(D_s)\, ds < \infty \tag{2.9}$$

in [4] (Proposition 2.2) it is proved that the $(P, \mathcal{F}_t)$-predictable projection ([16, 3]), $m^p(dt, dx)$, of the integer valued measure $m(dt, dx)$ can be written as

$$m^p(dt, dx) = \lambda_t \phi_t(dx) dt, \tag{2.10}$$

where $\forall A \in \mathcal{B}(\mathbb{R})$

$$m^p(dt, A) = \lambda_t \phi_t(A) dt = \nu(D_t^A) dt. \tag{2.11}$$

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

**Remark 2.3** *Equation (2.10) can be also written as*

$$m^p(dt, dx) = \lambda_t \phi_t(dx) dt = \int_{D_t} \delta_{K_1(t,\zeta)}(dx) \nu(d\zeta) dt. \tag{2.12}$$

We finally denote by $\nu^p(dt, dx)$ the $(P, \mathcal{F}_t^Y)$-predictable projection of the integer-valued measure $m(dt, dx)$. The following proposition, proved in [4], gives a representation of $\nu^p(dt, dx)$ in terms of the filter.

**Proposition 2.4** *The $(P, \mathcal{F}_t^Y)$-predictable projection of the integer-valued measure $m(dt, dx)$ is given by*

$$\nu^p(dt, dx) = \widehat{\lambda_t \phi_t}(dx)|_{t=t^-} dt = \pi_{t^-}(\lambda_t \phi_t(dx)) dt, \tag{2.13}$$

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

$$\nu^p((0,t], A) = \int_0^t \pi_{s^-}(\lambda_s \phi_s(A)) ds = \int_0^t \pi_{s^-}\big(\nu(d^A(., Y_{s^-}))\big) ds. \tag{2.14}$$

*where $\pi_{t^-}$ denotes the left version of the process $\pi_t$.*

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

$$m^\pi(dt, dx) = m(dt, dx) - \nu^p(dt, dx) = m(dt, dx) - \pi_{t^-}(\lambda_t \phi_t(dx)) dt, \tag{2.15}$$

and, assuming

$$\mathbb{E}\int_0^T \frac{b_1^2(t)}{\sigma_1^2(t)}dt < \infty, \tag{2.16}$$

we define the innovation process

$$I_t = W_t^1 + \int_0^t \left[\frac{b_1(s)}{\sigma_1(s)} - \pi_s\left(\frac{b_1}{\sigma_1}\right)\right] ds. \tag{2.17}$$

Let us notice that, by Remark 2.1 and assumptions (2.16),

$$\mathbb{E}\int_0^T \left|\pi_t\Big(\frac{b_1}{\sigma_1}\Big)\right| dt \le \mathbb{E}\int_0^T \pi_t\left|\frac{b_1}{\sigma_1}\right| dt = \mathbb{E}\int_0^T \frac{|b_1(t)|}{\sigma_1(t)}dt < \infty.$$

By extending classical results in filtering theory ([21]) to our frame we get:

**Proposition 2.5** *The random process $\{I_t\}_{t\in[0,T]}$ is a $(P, \mathcal{F}_t^Y)$-Wiener process.*

We write $\mathcal{F}_t^m$ for the filtration generated by the random measure $m(dt, dx)$. Since the innovation process $I_t$ and the random measure $m(dt, dx)$ 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 $(P, \mathcal{F}_t^Y)$-martingales in terms of the $\mathcal{F}_t^Y$-compensated random martingale measure $m^\pi(dt, dx)$ and the innovation process $I_t$.

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

$$L_t = \mathcal{E}\left(-\int_0^t \frac{b_1(s)}{\sigma_1(s)}dW_s^1\right) = exp\left\{-\int_0^t \frac{b_1(s)}{\sigma_1(s)}dW_s^1 - \frac{1}{2}\int_0^t \frac{b_1^2(s)}{\sigma_1^2(s)}ds\right\} \tag{2.18}$$

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

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

Under this last assumption we define a probability measure $Q$ on $\mathcal{F}_T$ equivalent to $P$ such that

$$\frac{dQ}{dP}\,|_{\mathcal{F}_T} = L_T. \tag{2.19}$$

By Girsanov Theorem the process

$$\widetilde{W}_t^1 = W_t^1 + \int_0^t \frac{b_1(s)}{\sigma_1(s)}ds \tag{2.20}$$

is a $(Q, \mathcal{F}_t)$-Wiener process and by (2.17),

$$\widetilde{W}_t^1 = I_t + \int_0^t \pi_s\left(\frac{b_1}{\sigma_1}\right) ds; \tag{2.21}$$

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

$$\widehat{L}_t = \mathbb{E}[L_t|\mathcal{F}_t^Y] = \frac{dQ}{dP}|_{\mathcal{F}_t^Y} = \mathcal{E}\Big(-\int_0^t \pi_s\Big(\frac{b_1}{\sigma_1}\Big)dI_s\Big). \tag{2.22}$$

Let us notice that, by Jensen's inequality, $\pi_t^2\left(\frac{b_1}{\sigma_1}\right) \le \pi_t\left(\frac{b_1^2}{\sigma_1^2}\right)$ and, by Remark 2.1, the following integrability condition holds $\mathbb{E}\displaystyle\int_0^T \pi_t\left(\frac{b_1^2}{\sigma_1^2}\right) dt = \mathbb{E}\int_0^T \frac{b_1^2(t)}{\sigma_1^2(t)}dt < \infty$.

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

$$dY_t = \sigma_1(t)dI_t + \left\{\pi_t(b_1) + \int_{\mathbb{R}} x\ \widehat{\lambda_t\phi_t}(dx)\right\}dt + \int_{\mathbb{R}} x\ m^\pi(dt, dx). \tag{2.23}$$

One can observe that again, by Remark 2.1 and assumptions (2.2),

$$\mathbb{E}\int_0^T |\pi_t(b_1)|dt \leq \mathbb{E}\int_0^T \pi_t|b_1|dt = \mathbb{E}\int_0^T |b_1(t)|dt < \infty;$$

furthermore taking into account (2.12),

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

**Proposition 2.6** *Under (2.2), (2.9), (2.16) and Assumption A, every $(P,\mathcal{F}_t^Y)$-local martingale $M_t$ admits the following decomposition*

$$M_t = M_0 + \int_0^t\int_{\mathbb{R}} w(s,x)m^\pi(ds,dx) + \int_0^t h(s)dI_s \tag{2.24}$$

*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*

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

Proof.
Let $Q$ be the probability measure defined in (2.19), then recalling the (2.21), $\widetilde{W}^1$ is a $(Q,\mathcal{F}_t^Y)$-Brownian motion. Note that the following equality of $\sigma$-algebras holds:

$$\mathcal{F}_t^Y = \mathcal{F}_t^m \vee \mathcal{F}_t^{\widetilde{W}^1}. \tag{2.25}$$

As a matter of fact, by (2.21) we get the inclusion $\mathcal{F}_t^Y \supseteq \mathcal{F}_t^m \vee \mathcal{F}_t^{\widetilde{W}^1}$, while the other one follows from the fact that $Y_t$ solves the stochastic differential equation driven by $m(dt,dx)$ and $\widetilde{W}_t^1$ given by

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

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 in [2]), and therefore by the $(Q,\mathcal{F}_t^Y)$-predictable characteristics because of the equality (2.25).

By applying Corollary III.4.31 of [16], every $(Q,\mathcal{F}_t^Y)$-local-martingale, $\widetilde{M}_t$, has the representation property with respect to $(m,\widetilde{W}^1)$, that 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

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

such that

$$\widetilde{M}_t = \widetilde{M}_0 + \int_0^t \widetilde{h}(s)\ d\widetilde{W}_s^1 + \int_0^t\int_{\mathbb{R}} \widetilde{w}(s,x)m^\pi(ds,dx). \tag{2.26}$$

Let $M_t$ be a $(P,\mathcal{F}_t^Y)$-local martingale. By Kallianpur-Striebel formula $\widetilde{M}_t = M_t\widehat{L}_t^{-1}$ is a $(Q,\mathcal{F}_t^Y)$-local martingale, where $\widehat{L}_t$ is defined in (2.22). Thus $M_t = \widetilde{M}_t\widehat{L}_t$ can be computed by the product formula:

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

Note that $d\widehat{L}_t = -\widehat{L}_t\pi_t\left(\frac{b_1}{\sigma_1}\right)dI_t$, and then

$$
\begin{aligned}
dM_t &= -\widetilde{M}_t \widehat{L}_t \pi_t \left(\frac{b_1}{\sigma_1}\right) dI_t + \widehat{L}_{t^-} \left\{ \int_{\mathbb{R}} \widetilde{w}(t,x) m^\pi(dt,dx) + \widetilde{h}(t) d\widetilde{W}^1_t \right\} + d\langle \int_0^\cdot \widetilde{h}(s) d\widetilde{W}^1_s, - \int_0^\cdot \widehat{L}_s \pi_s \left(\frac{b_1}{\sigma_1}\right) dI_s \rangle_t \\
&= -M_t \pi_t \left(\frac{b_1}{\sigma_1}\right) dI_t + \int_{\mathbb{R}} \widehat{L}_{t^-} \widetilde{w}(t,x) m^\pi(dt,dx) + \widehat{L}_t \widetilde{h}(t) \left( dI_t + \pi_t \left(\frac{b_1}{\sigma_1}\right) dt \right) - \widetilde{h}(t) \widehat{L}_t \pi_t \left(\frac{b_1}{\sigma_1}\right) dt \\
&= \left\{ -M_t \pi_t \left(\frac{b_1}{\sigma_1}\right) + \widehat{L}_t \widetilde{h}(t) \right\} dI_t + \int_{\mathbb{R}} \widehat{L}_{t^-} \widetilde{w}(t,x) m^\pi(dt,dx).
\end{aligned}
$$

Finally, we only need to define

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

## 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 $\mathcal{F}^Y_t$, which provides, as already said before, the most detailed description of our knowledge of $X_t$.

In the case of diffusion observations the filtering problem has been widely studied in literature: textbook treatments can be found for instance in Kallianpur [17] and Lipster Shiryaev [21]. More recently, results for pure-jump observations have been achieved (see [3, 7, 4, 12] and references therein), while few results can be found for mixed type information which involves pure-jump processes and diffusions ([13, 14, 5]) and, to the authors' knowledge, this is the first time that the filtering problem is studied for a general jump-diffusions system as the one defined in (2.1).

Several approaches have been considered in nonlinear filtering literature and among them we choose the innovation method which consists of deriving the dynamics of the filter, the so called Kushner-Stratonovich equation (KS-equation) and to characterize the filter itself as the unique solution to this equation. The KS-equation plays an essential role in the study of partially observable control problems by the Hamilton-Jacobi-Bellman approach (see for instance [8, 22, 1]).

First of all we want to recall a result proved in [7] (Corollary 2.2).
Let us denote by $\mathcal{C}^{1,2}_b([0,T]\times\mathbb{R})$ (resp. $\mathcal{C}^{1,2,2}_b([0,T]\times\mathbb{R}\times\mathbb{R})$) the set of the functions $f$ defined on $[0,T]\times\mathbb{R}$ (resp. $[0,T]\times\mathbb{R}\times\mathbb{R}$) such that $f$, $\frac{\partial f}{\partial t}$, $\frac{\partial f}{\partial x}$ and $\frac{\partial^2 f}{\partial x^2}$ (resp. $f$, $\frac{\partial f}{\partial t}$ and all the first and second derivatives with respect to $(x,y)$) are bounded continuous functions.

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

$$
\mathbb{E} \int_0^T \nu(D^0_t) dt < \infty \tag{3.1}
$$

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

$$
L^X f(t,x) = \frac{\partial f}{\partial t}(t,x) + b_0(t,x) \frac{\partial f}{\partial x} + \frac{1}{2}\sigma_0^2(t,x) \frac{\partial^2 f}{\partial x^2} + \int_Z \{ f(t, x + K_0(t,x,\zeta)) - f(t,x) \} \nu(d\zeta). \tag{3.2}
$$

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

$$
f(t, X_t) = f(0, x_0) + \int_0^t L^X f(s, X_s) ds + m^f_t \tag{3.3}
$$

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

$$
m^f_t = \int_0^t \frac{\partial f}{\partial x}(s, X_s) \sigma_0(s, X_s) dW^0_s + \int_0^t \int_Z \{ f(s, X_{s^-} + K_0(s, X_{s^-}, \zeta)) - f(s, X_{s^-}) \} \left( N(ds, d\zeta) - \nu(d\zeta) ds \right). \tag{3.4}
$$

Next theorem establishes the main result of this note.

**Theorem 3.2** *Under the same assumptions of Proposition 2.6, (3.1) and*

$$\mathbb{E}\int_0^T (\sigma_1^{-1}(t))^2 dt < \infty, \tag{3.5}$$

*the filter (2.3) is a solution of KS-equation, which is given, for any function $f(t,x) \in C_b^{1,2}([0,T]\times\mathbb{R})$ by*

$$\pi_t(f) = f(0,x_0) + \int_0^t \pi_s(L^X f)ds + \int_0^t \int_{\mathbb{R}} w_s^\pi(f,x) m^\pi(ds,dx) + \int_0^t h_s^\pi(f) dI_s \tag{3.6}$$

*where*

$$w_t^\pi(f,x) = \frac{d\pi_{t^-}(\lambda\phi f)}{d\pi_{t^-}(\lambda\phi)}(x) - \pi_{t^-}(f) + \frac{d\pi_{t^-}(\bar{L}f)}{d\pi_{t^-}(\lambda\phi)}(x) \tag{3.7}$$

$$h_t^\pi(f) = \sigma_1^{-1}(t)[\pi_t(f) - \pi_t(b_1)\pi_t(f)] + \rho\pi_t\left(\sigma_0\frac{\partial f}{\partial x}\right) \tag{3.8}$$

*Here by $\frac{d\pi_{t^-}(\lambda\phi f)}{d\pi_{t^-}(\lambda\phi)}(x)$ and $\frac{d\pi_{t^-}(\bar{L}f)}{d\pi_{t^-}(\lambda\phi)}(x)$ 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 follows:*

$$\bar{L}f = \bar{L}f(.,Y_{t^-},dz), \quad \forall A\in\mathcal{B}(\mathbb{R}) \quad \bar{L}f(t,x,y,A) = \int_{d^A(t,x,y)} [f(t,x+K_0(t,x,\zeta)) - f(t,x)]\nu(d\zeta).$$

*We recall that $d^A(t,x,y)$ is defined in (2.7) hence the operator $\overline{L}$ takes into account common jump times between the state $X$ and the observations $Y$.*

**Remark 3.3** *Let us observe that equation (3.6) is similar to the filtering equation derived in [5], even if a different partially observed system has been considered there. More precisely, in [5], the information flow has the structure $\mathcal{F}_t^m \vee \mathcal{F}_t^\eta$, where $m(dt,dx)$ is a marked point process with dynamics affected by a stochastic factor $X$ (whose dynamics is described by the first equation of (2.1)), and $\eta_t = \int_0^t \gamma(X_s)ds + W_t^1$, for any bounded measurable function with $\gamma(x)$. Nevertheless, let us point out that in [5] the filtering equation has been derived requiring boundness on $\lambda_t = \nu(D_t)$ and $\sigma_0(t,x)$.*

Before giving the proof of the above theorem, we ought to check that all the terms in (3.6) are well defined.

**Remark 3.4** *Since*

$$\int_{\mathbb{R}} |w_s^\pi(f,x)|\pi_{s^-}(\lambda_s\phi_s(dx)) \le |\pi_{s^-}(\lambda_s f)| + |\pi_{s^-}(\lambda_s)\pi_{s^-}(f)| + |\pi_{s^-}(\bar{L}f)(\mathbb{R})| \le 4\|f\|\pi_{s^-}|\lambda_s|$$

*assumption (2.9) and Remark 2.1 imply that*

$$\mathbb{E}\int_0^T \int_{\mathbb{R}} |w_s^\pi(f,x)|\pi_{s^-}(\lambda_s\phi_s(dx))ds \le 4\|f\|\mathbb{E}\int_0^T |\lambda_s|ds < \infty. \tag{3.9}$$

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

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

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

$$\mathbb{E}\int_0^T (h_t^\pi(f))^2 dt \le B_f\ \mathbb{E}\int_0^T \left\{(\sigma_1^{-1}(t))^2[1+b_1^2(t)] + \sigma_0^2(t)\right\}dt < \infty. \tag{3.10}$$

*Thus, taking into account (3.9) and (3.10), the integrals in (3.6) with respect to the compensated martingale measure $m^\pi(dt,dx)$ and to the innovation process $I_t$, are $(P,\mathcal{F}_t^Y)$-martingales. Finally note that*

$$\left|L^X f(t,X_t)\right| \le \widetilde{B}_f\left(1+|b_0(t)|+|\sigma_0(t)|^2+\nu(D_t^0)\right) \tag{3.11}$$

*for a suitable positive constant $\widetilde{B}_f$, and then $\mathbb{E}\int_0^T \left|\pi_t\left(L^X f\right)\right| dt \le \mathbb{E}\int_0^T \left|L^X f(t,X_t)\right| dt < \infty.$*

PROOF.
We shall consider the semimartingale

$$Z_t = f(t, X_t) = f(0, X_0) + \int_0^t L^X f(s, X_s)ds + m_t^f \tag{3.12}$$

where $m_t^f$ is given in (3.4). To keep formulas simpler to be read, we will leave out the dependence from the the process $X_t$ unless it is necessary, that is $f_t = f(t, X_t)$, $L^X f_t = L^X f(t, X_t)$ and $\frac{\partial f}{\partial x}(t) = \frac{\partial f}{\partial x}(t, X_t)$.
Now we project the semimartingale $Z_t$ on $\mathcal{F}_t^Y$

$$\widehat{Z}_t = \widehat{Z}_0 + \widehat{\int_0^t L^X f_s ds} + \widehat{m}_t^f = \widehat{Z}_0 + \int_0^t \widehat{L^X f_s} ds - \int_0^t \widehat{L^X f_s} ds + \widehat{\int_0^t L^X f_s ds} + \widehat{m}_t^f$$

and by Remark 2.1, $\widehat{Z}_t - \widehat{Z}_0 - \int_0^t \widehat{L^X f_s} ds$ is an $\mathcal{F}_t^Y$- martingale. Proposition 2.6 ensures us the existence of two processes $h^\pi, w^\pi$ such that

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

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

The strategy for proving the thesis consists on two steps. We will consider the $\mathcal{F}_t^Y$-adapted processes $\widetilde{W}_t^1$, whose expression is given by $\widetilde{W}_t^1 = W_t^1 + \int_0^t \frac{b_1(s)}{\sigma_1(s)} ds = I_t + \int_0^t \pi_s\left(\frac{b_1}{\sigma_1}\right) ds$, and a bounded process $U_t$ of the form $U_t = \int_0^t \int_{\mathbb{R}} \Gamma(s, x) m(ds, dx)$, where $\Gamma$ is a bounded $\mathcal{F}_t^Y$-predictable process and go on as follows:

**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 we have the equality $\widehat{Z_t U_t} = \widehat{Z}_t U_t$.

These two equalities will give us the shape of the processes $h^\pi, w^\pi$.

**Step 1.**
By applying the product rule

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

where $m_t^1 = \int_0^t \widetilde{W}_s^1 dm_s^f$ is a $(P, \mathcal{F}_t)$-local martingale and by $\rho$ we mean the correlation coefficient between the Brownian motions $W^1$ and $W^0$. Let us notice that one 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 $Z_t \widetilde{W}_t^1$ on $\mathcal{F}_t^Y$ we will get on $\{t \leq \widetilde{\tau}_n\}$

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

where $\widetilde{m}_t^1$ is a $(P, \mathcal{F}_t^Y)$-martingale (see Remark 2.1) and, by Proposition 2.2, $\widehat{m_{t \wedge \widetilde{\tau}_n}^1}$ is a $(P, \mathcal{F}_t^Y)$-martingale.

On the other side,

$$d(\widehat{Z}_t\widetilde{W}^1_t) = \left\{\widehat{Z}_t\pi_t\left(\frac{b_1}{\sigma_1}\right) + \widetilde{W}^1_t\widehat{L^X f_t} + h^\pi_t(f)\right\}dt + dm^2_t$$

where $m^2_t = \int_0^t \left\{\widetilde{W}^1_s h^\pi_s(f) + \widehat{Z}_s\right\} dI_s + \int_0^t \widetilde{W}^1_s \int_{\mathbb{R}} w^\pi_s(f,x)m^\pi(ds,dx)$ is a $(P,\mathcal{F}^Y_t)$-local martingale.
Since $\widehat{Z_t\widetilde{W}^1_t} = \widehat{Z}_t\widetilde{W}^1_t$, they have the same limited variation parts, that means

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

Equivalently,

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

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

**Step 2.**

We now choose the bounded process $U_t := \int_0^t \int_{\mathbb{R}} \Gamma(s,x)m(ds,dx) = \int_0^t \int_Z \mathbb{1}_{D_s}(\zeta)\Gamma(s,K_1(s,\zeta)))N(ds,d\zeta)$ (see (2.6) for last equality), then

$$d(Z_tU_t) = Z_{t^-}dU_t + U_{t^-}dZ_t + d[Z,U]_t = \left\{U_tL^X f_t + V_t + Z_t\int_{\mathbb{R}}\Gamma(t,x)\lambda_t\phi_t(dx)\right\}dt + dm^3_t \tag{3.13}$$

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

$$\begin{aligned} m^3_t \ =& \int_0^t\int_{\mathbb{R}} Z_{s^-}\Gamma(s,x)\big(m(ds,dx) - \lambda_s\phi_s(dx)ds\big) + \int_0^t \frac{\partial f}{\partial x}(s)\sigma_0(s)U_s dW^0_s + \\ &+ \int_0^t\int_Z \{f(s,X_{s^-}+K_0(s,\zeta)) - f(s,X_{s^-})\}\{\mathbb{1}_{D_s}(\zeta)\Gamma(s,K_1(s,\zeta)) + U_{s^-}\}\big(N(ds,d\zeta) - \nu(d\zeta)ds\big)\end{aligned}$$

is a $(P,\mathcal{F}_t)$-martingale. By projecting on $\mathcal{F}^Y_t$, the equation (3.13) becomes:

$$d(\widehat{Z_tU_t}) = \left\{\widehat{U_tL^X f_t} + \widehat{V}_t + \int_{\mathbb{R}}\Gamma(t,x)\widehat{Z_t\lambda_t\phi_t}(dx)\right\}dt + d\widetilde{m^3_t} \tag{3.14}$$

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

On the other hand

$$\begin{aligned} d(\widehat{Z}_tU_t) \quad &= \quad \widehat{Z}_{t^-}dU_t + U_{t^-}d\widehat{Z}_t + d[\widehat{Z},U]_t \\ &= \quad \left\{\int_{\mathbb{R}}\widehat{Z}_{t^-}\Gamma(t,x)\lambda_t\widehat{\phi_t(dx)} + U_t\widehat{L^X f_t} + \int_{\mathbb{R}}\Gamma(t,x)w^\pi_t(f,x)\lambda_t\widehat{\phi_t(dx)}\right\}dt + dm^4_t \end{aligned} \tag{3.15}$$

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

$$m^4_t = \int_0^t h^\pi_s(f)dI_s + \int_0^t \{\widehat{Z}_{s^-}(\Gamma(s,x) + w^\pi_s(f,x)) + U_{s^-}w^\pi_s(f,x)\}m^\pi(ds,dx).$$

As in **step 1.** the finite variation parts in (3.14) and (3.15) must be equal, so

$$\int_{\mathbb{R}} w^\pi_t(f,x)\Gamma(t,x)\lambda_t\widehat{\phi_t(dx)} = \int_{\mathbb{R}}\Gamma(t,x)\widehat{Z_t\lambda_t\phi_t}(dx) + \widehat{V}_t - \int_{\mathbb{R}}\widehat{Z}_t\Gamma(t,x)\lambda_t\widehat{\phi_t(dx)} \tag{3.16}$$

Now we are looking for $w^\pi_t(f,x)$ with the following shape:

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

We can always choose $w_2(t,f,x) = \widehat{Z}_{t^-}$ and by equality (3.16), $w_1, w_3$ need to satisfy:

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

$$\int_{\mathbb{R}} w_3(t,f,x)\Gamma(t,x)\widehat{\lambda_t\phi_t}(dx) = \widehat{V}_t.$$

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 $\Gamma(t,x) = C_t 1\!\!1_A(x) 1\!\!1_{\{t\le T_n\wedge T\}}$ with $C_t$ any bounded, $\mathcal{F}_t^Y$-predictable and 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 $|U_t| \le \int_0^{T\wedge T_n} |C_s|dN_s \le Dn$, with $D$ a suitable positive constant. Then on $\{t \le T_n \wedge T\}$

$$\begin{aligned} V_t &= \int_Z C_t 1\!\!1_{D_t^A}(\zeta)\{f(t,X_{t^-} + K_0(t,X_{t^-},\zeta)) - f(t,X_{t^-})\}\nu(d\zeta) \\ &= C_t \int_{D_t^A} \{f(t,X_{t^-} + K_0(t,X_{t^-},\zeta)) - f(t,X_{t^-})\}\nu(d\zeta). \end{aligned}$$

If we call $\int_{d^A(t,x,y)}\{f(t,x+K_0(t,x,\zeta)) - f(t,x)\}\nu(d\zeta) =: \overline{L}f(t,x,y,A)$, then we get $\forall A \in \mathcal{B}(\mathbb{R})$

$$\int_A w_3(t,f,x)\widehat{\lambda_t\phi_t}(dx) = \int_A \widehat{\overline{L}f}(X_{t^-},Y_{t^-},dx), \qquad \int_A w_1(t,f,x)\widehat{\lambda_t\phi_t}(dx) = \int_A \widehat{Z_{t^-}\lambda_t\phi_t}(dx) \quad \text{on} \quad \{t \le T_n\wedge T\}.$$

Thus

$$w_1(t,f,x) - w_2(t,f,x) + w_3(t,f,x) = \frac{d\pi_{t^-}(\lambda\phi f)}{d\pi_{t^-}(\lambda\phi)}(x) - \pi_{t^-}(f) + \frac{d\pi_{t^-}(\overline{L}f)}{d\pi_{t^-}(\lambda\phi)}(x) \quad \text{on} \quad \{t \le T_n \wedge T\}.$$

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

It can be observed that KS-equation (3.6) can be also written as

$$\pi_t(f) = f(0,x_0) + \int_0^t \{\pi_s(L_0^X f) + \pi_s(f)\pi_s(\lambda_s) - \pi_s(f\lambda_s)\}ds + \int_0^t \int_{\mathbb{R}} w_s^\pi(f,x)m(ds,dx) + \int_0^t h_s^\pi(f)dI_s \quad (3.17)$$

where

$$\begin{aligned} L_0^X f(t,x,y) &= L^X f(t,x) - \bar{L}f(t,x,y,\mathbb{R}) \\ &= \frac{\partial f}{\partial t}(t,x) + b_0(t,x)\frac{\partial f}{\partial x} + \frac{1}{2}\sigma_0^2(t,x)\frac{\partial^2 f}{\partial x^2} + \int_{d^{1^c}(t,x,y)} \{f(t,x+K_0(t,x,\zeta)) - f(t,x)\}\nu(d\zeta) \end{aligned}$$

($d^{1^c}(t,x,y) = \{\zeta \in Z : K_1(t,x,y,\zeta) = 0\}$) and it has a natural recursive structure. This can be seen if we write the equation at the jump times and between two consecutive jump times of $Y$. In fact, if $T_n$ is a jump time for the process $Y$ that occurs before time $T$,

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

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

$$\pi_t(f) = \pi_{T_n}(f) + \int_{T_n}^t \{\pi_s(L_0^X f) + \pi_s(f)\pi_s(\lambda_s) - \pi_s(f\lambda_s)\}ds + \int_{T_n}^t h_s^\pi(f)dI_s.$$

To show uniqueness for the solution to KS-equation we want to proceed as in [18], but we need to know exactly the shape of the generator of the pair $(X,Y)$. Then the following Lemma 3.5 helps us to reach the purpose.

**Lemma 3.5** *Under (2.2), (2.9) and (3.1), $(X_t, Y_t)$ is a $(P, \mathcal{F}_t)$-Markov process with generator $L^{X,Y}$ defined by, $\forall f \in C_b^{1,2,2}([0,T] \times \mathbb{R} \times \mathbb{R})$*

$$
\begin{aligned}
L^{X,Y} f(t,x,y) \;=\; & \frac{\partial f}{\partial t} + b_0(t,x)\frac{\partial f}{\partial x} + b_1(t,x,y)\frac{\partial f}{\partial y} + \frac{1}{2}\sigma_0^2(t,x)\frac{\partial^2 f}{\partial x^2} + \rho\sigma_0(t,x)\sigma_1(t,y)\frac{\partial^2 f}{\partial x \partial y} + \\
& + \frac{1}{2}\sigma_1^2(t,y)\frac{\partial^2 f}{\partial y^2} + \int_Z \left(f(t, x + K_0(t,x,\zeta), y + K_1(t,x,y,\zeta)) - f(t,x,y)\right)\nu(d\zeta)
\end{aligned}
\tag{3.18}
$$

Proof.
By the assumption of existence and uniqueness for the solution of the 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 $(P, \mathcal{F}_t)$-Markov process.

Then the proof consists of applying Itô's formula to a $C_b^{1,2,2}([0,T] \times \mathbb{R} \times \mathbb{R})$ function, $f(t,x,y)$,

$$
\begin{aligned}
df(t, X_t, Y_t) \;&=\; L^{X,Y} f(t, X_t, Y_t)dt + \sigma_0(t)\frac{\partial f}{\partial x}(t, X_t, Y_t)dW_t^0 + \sigma_1(t)\frac{\partial f}{\partial y}(t, X_t, Y_t)dW_t^1 + \\
&\quad + \int_Z \left(f(t, X_{t^-} + K_0(t,\zeta), Y_{t^-} + K_1(t,\zeta)) - f(t, X_{t^-}, Y_{t^-})\right)\left(N(dt, d\zeta) - \nu(d\zeta)dt\right) \\
&=\; L^{X,Y} f(t, X_t, Y_t)dt + dM_t^f.
\end{aligned}
\tag{3.19}
$$

Finally by (2.2), (2.9) and (3.1), since

$$
\mathbb{E}\int_0^T\int_Z |f(t, X_{t^-} + K_0(t,\zeta), Y_{t^-} + K_1(t,\zeta)) - f(t, X_{t^-}, Y_{t^-})|\nu(d\zeta)dt \le 2\|f\|\mathbb{E}\int_0^T \{\nu(D_t^0) + \nu(D_t)\}dt < \infty,
$$

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

**Remark 3.6** *By projecting equation* (3.19) *on $\mathcal{F}_t^Y$ we can state that $\pi_t\left(f(\cdot, Y_t)\right) - \int_0^t \pi_s\left(L^{X,Y} f(\cdot, Y_s)\right)ds$ is a $(P, \mathcal{F}_t^Y)$-martingale for each $f \in C_b^{1,2,2}([0,T] \times \mathbb{R} \times \mathbb{R})$.*

We want to use this martingale property to characterize the distribution of the pair $(\pi_t, Y_t)$ by exploiting the idea given in [18]; therefore we introduce the notion of *filtered martingale problem* (FMP).

**Definition 3.7** *We say that a process $(\mu_t, U_t)$ defined on a probability space $\left(\widetilde{\Omega}, \widetilde{\mathcal{F}}_t, \widetilde{P}\right)$, 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 $\mu$ is $\mathcal{F}_t^U$- adapted and*

$$
\mu_t\left(f(\cdot, U_t)\right) - \int_0^t \mu_s\left(L^{X,Y} f(\cdot, U_s)\right)ds
$$

*is a $(\widetilde{P}, \mathcal{F}_t^U)$-martingale for each $f \in C_b^{1,2,2}([0,T] \times \mathbb{R} \times \mathbb{R})$ and $\mathbb{E}^{\widetilde{P}}[\mu_0 f(\cdot, U_0)] = f(0, x_0, y_0)$.*

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

**Definition 3.8** *A weak solution to Kushner-Stratonovich equation* (3.6) *is a process $(\mu_t, \widetilde{Y}_t)$ 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., $\mathbb{E}^{\widetilde{P}}[\mu_0(f)] = f(0, x_0)$ $\forall f \in \mathcal{C}_b^{1,2}([0,T] \times \mathbb{R})$, and satisfying*

**(i)** *$\mu_t$ 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 $\widetilde{Y}$, $\widetilde{m}(dt, dx)$, is given by $\mu_{t^-}(\lambda_t \phi_t(dx))dt$*

**(iii)** $\widetilde{I}_t = \int_0^t \frac{1}{\sigma_1(s)} \left( d\widetilde{Y}_s - \int_{\mathbb{R}} x\, \widetilde{m}(ds, dx) \right)$ *is a* $(\widetilde{P}, \mathcal{F}_t^{\widetilde{Y}})$*-Brownian motion*

**(iv)** *the pair* $(\mu_t, \widetilde{Y}_t)$ *solves the Kushner-Stratonovich equation (3.6), with* $m^\pi(dt, dx)$, $I_t$, $w_t^\pi(f, x)$ *and* $h_t^\pi(f)$ *replaced by* $m^\mu(dt, dx) = \widetilde{m}(dt, dx) - \mu_{t^-}(\lambda_t \phi_t(dx))dt$, $I_t^\mu = \widetilde{I}_t - \int_0^t \mu_s(b_1/\sigma_1)ds$, $w_t^\mu(f, x)$ *and* $h_t^\mu(f)$, *respectively*

**(v)** $\int_0^T \mu_t(b_2(\cdot, \widetilde{Y}_t))dt < \infty$, $\widetilde{P}$*-a.s. with* $b_2(t, x, y) = \lambda(t, x, y) + |b_0(t, x)| + \sigma_0^2(t, x) + \nu(d_0(t, x)) + \frac{1 + b_1^2(t,x,y)}{\sigma_1^2(t,y)}$.

**Remark 3.9** *In* (**ii**) *we mean that* $\forall\, A \in \mathcal{B}(\mathbb{R})$, $\mu_{t^-}(\lambda_t \phi_t(A)) = \mu_{t^-}(\nu(d^A(\cdot, \widetilde{Y}_{t^-})))$ *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 used the notation* $\lambda(t, x, y) = \nu(d^1(t, x, y))$.

**Remark 3.10** *Taking into account* (**v**) *we can prove that for any* $f \in C_b^{1,2,2}([0, T] \times \mathbb{R} \times \mathbb{R})$

$$\int_0^T |w_t^\mu(f, x)| \mu_{t^-}(\lambda_t \phi_t(dx))dt \le 4\|f\| \int_0^T \mu_t(\lambda)dt \le 4\|f\| \int_0^T \mu_t(b_2)dt < \infty \quad \widetilde{P} - a.s. \tag{3.20}$$

$$\int_0^T h_t^\mu(f)^2 dt \le B_f \int_0^T \left\{ \mu_t(\sigma_0^2) + \frac{1 + \mu_t(b_1^2(\cdot, \widetilde{Y}_t)}{\sigma_1^2(t, \widetilde{Y}_t)} \right\} dt \le B_f \int_0^T \mu_t(b_2)dt < \infty \quad \widetilde{P} - a.s. \tag{3.21}$$

$$\int_0^T \left| \mu_t(L^X f) \right| dt \le \widetilde{B}_f \int_0^T \left\{ 1 + |\mu_t(b_0)| + \mu_t(\sigma_0)^2 + \mu_t(\nu(d_0)) \right\} dt \le \widetilde{B}_f \int_0^T \mu_t(b_2)dt < \infty \quad \widetilde{P} - a.s. \tag{3.22}$$

*with* $B_f$ *and* $\widetilde{B}_f$ *suitable positive constants. Thus all the stochastic integrals in KS-equation considered in* (**iv**) *are well defined and those driven by* $I_t^\mu$ *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 3.11** *Let us notice that the pair filter-observation,* $(\pi_t, Y_t)$, *is a weak solution to* (3.6). *As a matter of fact,* (**i**), (**ii**),(**iv**) *and* (**v**) *of Definition (3.8) are trivially verified; for* (**iii**) *consider the probability measure* $Q$ *defined in* (2.19), *then by Girsanov Theorem the process*

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

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

Now we can state a weak uniqueness result for the solution of Kushner-Stratonovich equation whose proof is postponed in Appendix B.

**Theorem 3.12** *Under the same hypotheses of Theorem 3.2, uniqueness for the solutions to FMP*$(L^{X,Y}, x_0, y_0)$ *implies that all weak solutions* $(\mu_t, \widetilde{Y}_t)$ *of Kushner-Stratonovich equation have the same law. In particular* $\mu_t$ *and* $\pi_t$ *have the same law.*

Again in Appendix B we will give a class of sufficient conditions that ensures uniqueness for the solution to the filtered martingale problem for $L^{X,Y}$ (see Proposition 6.1).

In the remaining part of the section we discuss pathwise uniqueness for the solution of Kushner-Stratonovich equation.

Firstly, we start by giving the definition of strong solution.

**Definition 3.13** *A strong solution for Kushner-Stratonovich equation is an $\mathcal{F}_t^Y-$adapted càdlàg $\mathcal{P}(\mathbb{R})-$valued process $\{\mu_t\}_{t\in[0,T]}$ such that $\int_0^T \mu_s(b_2)ds < \infty \quad P-a.s.$ ($b_2$ is defined in* (**v**) *of Definition 3.8), and solving Kushner-Stratonovich equation that is, $\forall f \in \mathcal{C}_b^{1,2,2}([0,T]\times\mathbb{R}\times\mathbb{R})$ and $\forall t \le T$*

$$\mu_t(f) = \pi_0(f) + \int_0^t \mu_s(L^X f)ds + \int_0^t \int_{\mathbb{R}} w_s^\mu(f,x) m^\mu(dt,dx) + \int_0^t h_s^\mu(f) dI_s^\mu \tag{3.23}$$

*where*

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

*and $w_t^\mu(f,x)$ and $h_t^\mu(f)$ defined respectively in (3.7),(3.8) replacing $\pi$ with $\mu$.*

Note that the condition $\int_0^T \mu_s(b_2)ds < \infty \quad P-a.s.$ makes the integrals in (3.23) well defined and those driven by $I^\mu$ and $m^\mu(dt,dx)$, $(P,\mathcal{F}_t^Y)$-local martingales (as already observed in Remark 3.10).

**Theorem 3.14** *Let $(X_t,Y_t)$ 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 for Kushner-Stratonovich 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 $t\le T$.*

The proof is postponed in Appendix B.

We conclude this section considering a simplified model and giving a sufficient condition which implies that the additional hypotesis in Theorem 3.14 is satisfied.

**Example 3.15 Observation dynamics driven by independent point processes.**

Suppose there exists a finite set of measurable functions $K_1^i(t,y) \neq 0$ for all $(t,y)\in[0,T]\times\mathbb{R}$, for $i=1,...,n$, such that

$$d^1(t,x,y) := \{\zeta \in Z : K_1(t,x,y,\zeta) \neq 0\} = \bigcup_{i=1}^n d_i^1(t,x,y) \quad \text{and} \quad d_i^1(t,x,y)\cap d_j^1(t,x,y) = \emptyset \quad \forall i \neq j$$

where $d_i^1(t,x,y) := \{\zeta \in Z : K_1(t,x,y,\zeta) = K_1^i(t,y)\}$. This implies that $K^1(t,X_{t^-},Y_{t^-},\zeta) = \sum_{i=1}^n K_i^1(t,Y_{t^-})\mathbb{1}_{D_t^i}(\zeta)$ 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:

$$dY_t = b_1(t,X_t,Y_t)dt + \sigma_1(t,Y_t)dW_t^1 + \sum_{i=1}^n K_1^i(t,Y_{t^-})dN_t^i \tag{3.24}$$

where $N_t^i = N((0,t]\times D_t^i)$, for $i=1,...,n$, turn to be 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 drift and the intensities of the point process driving the observation dynamics but not the jump coefficients $K_t^i(t,Y_{t^-})$ for $i=1,...,n$, 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

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

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

$$\lambda^i(t,x,y) = \nu\left(d_i^1(t,x,y)\right) > 0 \quad \forall (t,x,y) \in [0,T] \times \mathbb{R} \times \mathbb{R}, \quad i = 1, ..., n \tag{3.25}$$

for any $\mathcal{F}_t^Y$-adapted, càdlàg, $\mathcal{P}(\mathbb{R})$-valued process $\{\mu_t\}_{t\in[0,T]}$, 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

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

because $\delta_{K_1^i(t,Y_{t^-})}(dx)$ for $i = 1, ..., n$, are $\mathcal{F}_t^Y$-measurable. Note that (3.25) implies $\pi_{t^-}(\lambda^i) > 0$, $\mu_{t^-}(\lambda^i) > 0$, $i = 1, ..., n$, and the Radon-Nikodym derivative of $\mu_{t^-}(\lambda_t \phi_t(dx))dt$ with respect to $\pi_{t^-}(\lambda_t \phi_t(dx))dt$ becomes

$$\begin{aligned} \frac{d\mu_{t^-}(\lambda\phi)}{d\pi_{t^-}(\lambda\phi)}(x) &= \frac{\sum_{i=1}^{n} \delta_{K_1^i(t,Y_{t^-})}(x) \mu_{t^-}(\lambda^i)}{\sum_{i=1}^{n} \delta_{K_1^i(t,Y_{t^-})}(x) \pi_{t^-}(\lambda^i)} \\ &= \sum_{i=1}^{n} \mathbb{I}_{\left\{K_1^i(t,Y_t)=x\right\}} \frac{\mu_{t^-}(\lambda^i)}{\pi_{t^-}(\lambda^i)}. \end{aligned}$$

On the other side, there exists also 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

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

and this means that these two measures are equivalent. In this way we have just proved the following corollary for the simplified model of the example.

**Corollary 3.16** *Let $(X_t, Y_t)$ be the usual partially observed system of the* (2.1)*, where in particular the dynamics of $Y$ is given by* (3.24)*, and assume uniqueness for the $FMP(L^{X,Y}, x_0, y_0)$. Let $\{\mu_t\}_{t\in[0,T]}$ be a strong solution for KS-equation given by the* (3.6) $\Big($*with* $m^\pi(dt,dx) = \sum_{i=1}^{n} \delta_{K_1^i(t,Y_{t^-})}(dx)(dN_t^i - \pi_{t^-}(\lambda^i)dt)\Big)$ *replacing $\pi_t$ by $\mu_t$. Then $\mu_t = \pi_t$ $P - a.s.$ for all $t \leq T$.*

## 4. Application to finance: Risk minimizing hedging

On a probability space $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, P\right)$, we consider a financial market with a nonrisky asset, with price process normalized to unity, and one risky asset whose price process $S_t$ follows a geometric jump-diffusion process given by

$$S_t = S_0 e^{Y_t} \quad S_0 \in \mathbb{R}^+. \tag{4.1}$$

We assume that the dynamics of the logreturn process $Y_t$ depends on some unobservable stochastic factor $X$ and the pair $(X, Y)$ is the unique solution to the sistem (2.1).

Applying Itô's formula we get that $S_t$ solves the following differential equation

$$dS_t = S_{t^-} \left\{ b(t)dt + \sigma(t)dW_t^1 + \int_Z K(t,\zeta)N(dt,d\zeta) \right\} \tag{4.2}$$

where we wrote $b(t), \sigma(t)$ and $K(t,\zeta)$ for $b(t, X_t, S_t), \sigma(t, S_t), K(t, X_{t^-}, S_{t^-}, \zeta)$ respectively, and those functions are given by

$$
\begin{aligned}
b(t,x,s) &= b_1\left(t,x,\ln\frac{s}{S_0}\right)+\frac{1}{2}\sigma_1^2\left(t,\ln\frac{s}{S_0}\right) &(4.3)\\
\sigma(t,s) &= \sigma_1\left(t,\ln\frac{s}{S_0}\right) &(4.4)\\
K\left(t,x,s,\zeta\right) &= e^{K_1(t,x,\ln s/S_0,\zeta)}-1. &(4.5)
\end{aligned}
$$

In the following proposition we give the semimartingale structure of the process $S$ under general hypotheses.

**Proposition 4.1** *Assuming the following integrability conditions*

$$
\int_0^T |b_1(t)|dt<\infty;\quad \int_0^T \sigma^2(t)dt<\infty;\quad \int_0^T \nu(D_t)dt<\infty;\quad \int_0^T\int_Z |K(t,\zeta)|^2\nu(d\zeta)dt<\infty \quad P-a.s. \tag{4.6}
$$

*then $S_t$ is a $(P,\mathcal{F}_t)$-special semimartingale with unique decomposition*

$$
S_t = S_0 + M_t^S + A_t^S \tag{4.7}
$$

*where*

$$
A_t^S=\int_0^t S_r b(r)dr+\int_0^t\int_Z S_rK(r,\zeta)\nu(d\zeta)dr=\int_0^t S_rb(r)dr+\int_0^t\int_{\mathbb{R}} S_r(e^x-1)\lambda_r\phi_r(dx)dr \tag{4.8}
$$

*is a predictable process with bounded variation paths,*

$$
\begin{aligned}
M_t^S &= \int_0^t S_r\sigma(r)dW_r^1+\int_0^t\int_Z S_{r^-}K(r,\zeta)(N(dr,d\zeta)-\nu(d\zeta)dr)\\
&= \int_0^t S_r\sigma(r)dW_r^1+\int_0^t\int_{\mathbb{R}} S_{r^-}(e^x-1)(m(dr,dx)-\lambda_r\phi_r(dx)dr) &(4.9)
\end{aligned}
$$

*is a square-integrable local martingale whose angle process is given by*

$$
\langle M^S\rangle_t=\int_0^t S_r^2\sigma(r)^2dr+\int_0^t\int_Z S_r^2K(r,\zeta)^2\nu(d\zeta)dr=\int_0^t S_r^2\sigma(r)^2dr+\int_0^t\int_{\mathbb{R}} S_r^2(e^x-1)^2\lambda_r\phi_r(dx)dr. \tag{4.10}
$$

*If in addition*

$$
\forall t\in[0,T], x\in\mathbb{R}, s\in\mathbb{R}^+ \quad b(t,x,s)+\int_Z K(t,x,s,\zeta)\nu(d\zeta)=0 \tag{4.11}
$$

*$S$ is a square-integrable local martingale.*

Proof.
Recalling the definition in (4.1), we observe that the process $S_t$ is the exponential of the semimartingale $Y_t$, and therefore it is itself a semimartingale by Theorem 4.57 in [16]. Then if we integrate the equation (4.2), we get the explicit decomposition,

$$
\begin{aligned}
S_t &= S_0+\int_0^t S_{u^-}\left\{b(u)+\int_Z K(u,\zeta)\nu(d\zeta)\right\}du+\int_0^t S_u\sigma(u)\,dW_u^1+\int_0^t\int_Z S_{u^-}K(u,\zeta)\left(N(du,d\zeta)-\nu(d\zeta)du\right)\\
&= S_0+A_t^S+M_t^S
\end{aligned}
$$

where $A_t^S$ and $M_t^S$ are given by (4.8) and (4.9) respectively. This decomposition is unique since the process $A_t^S$ is predictable.

Besides $M_t^S$ is a square integrable local martingale (see [24], Theorem 1 page 102).
For the sharp brackets note that $M_t^S$ is the sum of a continuous local martingale and a purely discontinuous local martingale, therefore we get:

$$
d\langle M^S\rangle_t = S_t^2\sigma^2(t)dt+\int_Z S_t^2K^2(t,\zeta)\nu(d\zeta)dt
$$

which is equivalent to (4.10).

Finally, the last part of the proposition is a consequence of (4.8) and the decomposition of the semimartingale $S_t$. $\square$

The pair $(X, S)$ is a $(P, \mathcal{F}_t)$-Markov process whose generator is computed in the next lemma.

**Lemma 4.2** *Under the hypotheses of Proposition 4.1, and in addition*

$$\int_0^T \sigma_0^2(t, X_t)dt < \infty, \qquad \int_0^T \nu_0(D_t)dt < \infty \qquad P-a.s. \tag{4.12}$$

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

$$\begin{aligned} L^{X,S}f(t,x,s) =& \frac{\partial f}{\partial t} + b_0(t,x)\frac{\partial f}{\partial x} + b(t,x,s)\, s\frac{\partial f}{\partial s} + \frac{1}{2}\sigma_0^2(t,x)\frac{\partial^2 f}{\partial x^2} + \rho\, \sigma_0(t,x)\, \sigma(t,s)s\frac{\partial^2 f}{\partial x \partial s} + \\ &+ \frac{1}{2}\sigma^2(t,s)\, s^2\frac{\partial^2 f}{\partial s^2} + \int_Z \Big(f\big(t, x + K_0(t,x,\zeta), s(1 + K(t,x,s,\zeta))\big) - f(t,x,s)\Big)\nu(d\zeta). \end{aligned} \tag{4.13}$$

*More precisely, for any function $f(t,x,s) \in C_b^{1,2,2}([0,T]\times\mathbb{R}\times\mathbb{R}^+)$ the following semimartingale decomposition holds*

$$f(t, X_t, S_t) = f(t, x_0, S_0) + \int_0^t L^{X,S}f(r, X_r, S_r)dr + M_t^f \tag{4.14}$$

*where $M_t^f$ is a $(P, \mathcal{F}_t)$-local martingale.*

Proof.
By applying Itô's formula to the function $f(t, X_t, S_t)$ we get (4.14) with

$$dM_t^f = \sigma_0(t)\frac{\partial f}{\partial x}dW_t^0 + \sigma(t)S_t\frac{\partial f}{\partial s}dW_t^1 + \int_Z \left[f(t, X_{t^-} + K_0(t,\zeta), S_{t^-}(1 + K(t,\zeta)) - f(t, X_{t^-}, S_{t^-})\right](N(dt,d\zeta) - \nu(d\zeta)dt) \tag{4.15}$$

and by (4.12) and (4.6)

$$\int_0^T \sigma_0^2(t)\left(\frac{\partial f}{\partial x}\right)^2 dt < \infty, \qquad \int_0^T \sigma^2(t)S_t^2\left(\frac{\partial f}{\partial s}\right)^2 dt < \infty \quad P-a.s.$$

$$\int_0^T\int_Z |f(t, X_{t^-}+K_0(t,\zeta), S_{t^-}(1+K(t,\zeta)) - f(t, X_{t^-}, S_{t^-})|\nu(d\zeta)dt \leq 2\|f\| \int_0^T \{\nu(D_t^0)+\nu(D_t)\}dt < \infty \quad P-a.s.$$

that means that all the integrals in (4.14) are well defined and $M_t^f$ is a local martingale. $\square$

From now on we will assume that $P$ is a martingale measure for $S$ and the following stronger conditions

**Assumptions B**: (2.2), (2.9), (4.11), $\mathbb{E}\int_0^T \nu(D_t^0)dt < \infty$ and $\mathbb{E}\int_0^T\int_Z |K(t,\zeta)|^2\nu(d\zeta)dt < \infty$.

We consider a contingent claim whose final payoff is a function $H(S_T)$ such that $\mathbb{E}\left[H^2(S_T)\right] < \infty$.

In order to hedge against this claim, we want to use a portfolio strategy which involves the stock $S$ and the riskless bond (normalized to one), and which yield the random payoff $H(S_T)$ at the terminal time $T$.

More precisely, an $\mathcal{F}_t$-portfolio-strategy $\delta_t = (\delta_t^0, \delta_t^1)$ is a process representing the quantities invested in each title, such that $\delta_t^0$ is $\mathcal{F}_t$-adapted, and $\delta_t^1$ is $\mathcal{F}_t$-predictable.

The financial value of this portfolio $\delta$ is given by

$$V_t(\delta) = \delta_t^0 + \delta_t^1 S_t. \tag{4.16}$$

**Definition 4.3** *A strategy $\delta$ is H-admissible if $V_T(\delta) = H(S_T)$ $P-$a.s. and it verifies regularity conditions:*

$$\mathbb{E}\left(\int_0^T |\delta_t^1|^2 d\langle S\rangle_t\right) < \infty, \qquad \mathbb{E}\left(\sup_{t\le T} |V_t|^2\right) < \infty.$$

Since we are working in an incomplete market, the claim can not be replicated by a self-financing strategy (perfect duplication), it makes sense to define a cost process for the strategy $\delta$ as

$$C_t(\delta) = V_t(\delta) - \int_0^t \delta_r^1 dS_r. \tag{4.17}$$

The cost process is nothing but the difference between the portfolio value and the total gains from trade.

Let us make some comments about the structure of this process. First of all note that $\int_0^t \delta_r^1 dS_r$ is a square-integrable martingale even thought $S$ is just a local martingale, as it is shown in Lemma (2.1) of [26].

We are interested in the so called mean-self financing strategies, i.e. those strategies such that $C_t(\delta)$ is a martingale. Note that if $\delta$ is an H-admissible mean-self financing strategy, we get:

$$\mathbb{E}\left[C_T(\delta)|\mathcal{F}_t\right] := \mathbb{E}\left[V_T(\delta) - \int_0^T \delta_r^1 dS_r|\mathcal{F}_t\right] = V_t(\delta) - \int_0^t \delta_r^1 dS_r.$$

Since the part concerning the stochastic integral is a martingale, so is the portfolio value $V$. Now, the strategy is H-admissible, and this means that $V_T(\delta) = H(S_T)$. Putting together these two comments we get the following relation

$$\mathbb{E}\left[V_T(\delta)|\mathcal{F}_t\right] = \mathbb{E}\left[H(S_T)|\mathcal{F}_t\right] = V_t(\delta). \tag{4.18}$$

We say that an $H$-admissible strategy is risk minimizing if it minimizes the associated risk process

$$R_t(\delta) = \mathbb{E}\left[(C_t(\delta) - C_T(\delta))^2|\mathcal{F}_t\right]. \tag{4.19}$$

We now consider the situation where investors acting on the market can only keep past asset prices; the stochastic factor which affects the stock price dynamics is not directly observed. Thus the agent's information is described by the filtration $\{\mathcal{F}_t^S\}_{t\in[0,T]} = \{\mathcal{F}_t^Y\}_{t\in[0,T]}$.

In this partial observation background the hedger's decision $(\delta^0, \delta^1)$ must be adapted to the flow $\mathcal{F}_t^S$. Hence we define an $\mathcal{F}_t^S$-strategy $\delta$, as an $\mathcal{F}_t$-strategy such that the process $\delta_t^0$ is $\mathcal{F}_t^S$-adapted and $\delta_t^1$ is $\mathcal{F}_t^S$-predictable.

In this framework we give also the definition of the $\mathcal{F}_t^S$-risk-process as

$$R_t^S(\delta) = \mathbb{E}\left[(C_t(\delta) - C_T(\delta))^2|\mathcal{F}_t^S\right]. \tag{4.20}$$

The agent's aim is to find a strategy $\delta$ belonging to the set of $\mathcal{F}_t^S$ H-admissible strategies that minimizes the risk process $R_t^S$, and that we will call $\mathcal{F}_t^S$*-risk minimizing strategy.*

We want to proceed as in [26], so we first compute the risk minimizing strategy in complete information and then the $\mathcal{F}_t^S$-strategy can be found by projecting on $\mathcal{F}_t^S$.

**Proposition 4.4** *Under Assumptions B, let $g(t, X_t, S_t) := \mathbb{E}\left[H(S_T)|\mathcal{F}_t\right]$. If $g \in \mathcal{C}_b^{1,2,2}([0,T]\times\mathbb{R}\times\mathbb{R}^+)$ then $g$ solves*

$$\begin{cases} L^{X,S} g(t,x,s) = 0 \\ g(T,x,s) = H(s). \end{cases}$$

*Besides, the risk-minimizing strategy under complete information is*

$$\begin{cases} \delta_t^{0,*} = g(t, X_t, S_t) - \delta_t^{1,*} S_t \\ \delta_t^{1,*} = \dfrac{h(t, X_{t^-}, S_{t^-})}{S_{t^-}\Sigma(t, X_{t^-}, S_{t^-})} \end{cases} \tag{4.21}$$

*where*

$$h(t,x,s) = \rho\,\sigma_0(t,x)\sigma(t,s)\frac{\partial g}{\partial x} + s\,\sigma^2(t,s)\frac{\partial g}{\partial s} + \int_Z K(t,x,s,\zeta)\left[g\big(t, x+K_0(t,x,\zeta), s(1+K(t,x,s,\zeta))\big) - g(t,x,s)\right]\nu(d\zeta) \quad (4.22)$$

*and*

$$\Sigma(t,x,s) = \sigma(t,s)^2 + \int_Z K(t,x,s,\zeta)^2\nu(d\zeta). \quad (4.23)$$

Proof.
For the first part of the proposition, suppose that $g \in C_b^{1,2,2}([0,T]\times\mathbb{R}\times\mathbb{R}^+)$, by applying Itô's formula we get

$$g(t,X_t,S_t) = g(0,X_0,S_0) + \int_0^t L^{X,S}g(r,X_r,S_r)dr + M_t^g \quad (4.24)$$

where $M_t^g$ is a $(P,\mathcal{F})$-local martingale. Since, by definition, $g(t,X_t,S_t)$ is a $(P,\mathcal{F}_t)$-martingale, all finite variation terms in (4.24) have to vanish and this leads to equation (4.4).

Let us move on to the second part of the proposition. Since, by Proposition 4.1, $S$ is a square integrable local martingale, Kunita-Watababe decomposition allows us to write $H(S_T)$ as

$$H(S_T) = \mathbb{E}\left[H(S_T)\right] + \int_0^T \xi_r^H dS_r + L_T^H \quad (4.25)$$

where $\xi_t^H$ is an $\mathcal{F}_t$-predictable process with $\mathbb{E}\left[\int_0^T(\xi_r^H)^2 d\langle S\rangle_r\right] < \infty$, and $L^H$ is a square integrable martingale null at $t=0$, such that $\langle S, L^H\rangle_t = 0$ .
If $S$ is a local martingale, it has been proved in [27] that there exists a unique $H-$admissible mean-self financing risk minimizing strategy we denote by $(\delta^{0,*},\delta^{1,*})$, where $\delta^{1,*}$ is exactly given by the process $\xi^H$ in (4.25), then the value of this strategy $V(\delta^*)$ is a martingale according to the equation (4.18) and finally

$$V_t(\delta^*) = \mathbb{E}\left[V_T(\delta^*)|\mathcal{F}_t\right] = \mathbb{E}\left[H(S_T)|\mathcal{F}_t\right] = g(t,X_t,S_t). \quad (4.26)$$

By using the expression (4.25), the (4.26) becomes

$$\begin{aligned} V_t(\delta^*) &= \mathbb{E}\left[H(S_T)\right] + \mathbb{E}\left[\int_0^T \xi_r^H dS_r|\mathcal{F}_t\right] + \mathbb{E}\left[L_T^H|\mathcal{F}_t\right] \\ &= \mathbb{E}\left[H(S_T)\right] + \int_0^t \xi_r^H dS_r + L_t^H \end{aligned} \quad (4.27)$$

where the equality (4.27) follows from the fact that $L_t^H$ and $\int_0^t \xi_r^H dS_r$ are martingales.

Now, if we consider the sharp bracket between $V(\delta^*)$ and $S$, we have

$$\langle V(\delta^*), S\rangle_t = \langle\int_0^\cdot \xi_r^H\, dS_r, S\rangle_t + \langle L^H, S\rangle_t = \langle\int_0^\cdot \xi_r^H\, dS_r, \int_0^\cdot dS_r\rangle_t = \int_0^t \xi_r^H d\langle S\rangle_r$$

from which we obtain an expression for $\xi^H$ in terms of the Radon-Nikodym derivative

$$\xi_t^H = \frac{d\langle V(\delta^*), S\rangle_t}{d\langle S\rangle_t}.$$

We want to compute this derivative. Recall the shape of the martingale $M^g$, given in (4.15) replacing $f$ by $g$, and Proposition 4.1, we obtain

$$d\langle V(\delta^*), S\rangle_t = \int_0^t S_{r^-}h(r,X_{r^-},S_{r^-})dr \quad (4.28)$$

with $h(t, x, s)$ given in (4.22). On the other side, by (4.10)

$$\langle S\rangle_t = \int_0^t S_r^2 \left\{\sigma(r)^2 + \int_Z K(r,\zeta)^2\nu(d\zeta)\right\} dr = \int_0^t S_r^2\ \Sigma(r, X_{r^-}, S_{r^-})dr$$

hence the risk-minimizing strategy turns to be

$$\delta_t^{1,*} = \xi_t^H = \frac{h(t, X_{t^-}, S_{t^-})}{S_{t^-}\Sigma(t, X_{t-}, S_{t^-})}. \qquad \square$$

Note that Assumptions B imply integrability for the processes $h(t, X_{t^-}, S_{t^-})$ and $\Sigma(t, X_{t^-}, S_{t^-})$ and then allow us to apply Schweizer results (see [26]): this means that the risk minimizing strategy under partial information $\bar{\delta}^*$, for our model can be obtained by projecting the strategy $\delta^*$ over $\mathcal{F}_{t^-}^S = \mathcal{F}_{t^-}^Y$. More precisely

$$\begin{cases} \bar{\delta}_t^{0,*} = V_t(\bar{\delta}^*) - \bar{\delta}_t^{1,*} S_t \\ \bar{\delta}_t^{1,*} = \dfrac{\mathbb{E}\left[h(t, X_{t^-}, S_{t^-})|\mathcal{F}_{t^-}^Y\right]}{S_{t^-}\mathbb{E}\left[\Sigma(t, X_{t^-}, S_{t^-})|\mathcal{F}_{t^-}^Y\right]} \end{cases} \tag{4.29}$$

Recalling that the filter is defined as

$$\pi_t(f) = \mathbb{E}\left[f(t, X_t)|\mathcal{F}_t^Y\right]$$

and since it is a càdlàg process, the $(P, \mathcal{F}_t^S)$-predictable projection of $f(t, X_t)$ coincides with the left version of $\pi_t(f)$ (see [11]). Now we are in the position to state the announced result.

**Proposition 4.5** *Under Assumptions B, if $g \in \mathcal{C}_b^{1,2,2}([0,T] \times \mathbb{R} \times \mathbb{R})$, then the $\mathcal{F}_t^S$-risk-minimizing strategy is given by*

$$\begin{cases} \bar{\delta}_t^{0,*} = \pi_t\left[g(\cdot, S_t)\right] - \bar{\delta}_t^{1,*} S_t \\ \bar{\delta}_t^{1,*} = \dfrac{\pi_{t^-}\left[h(\cdot, S_{t^-})\right]}{S_{t^-}\pi_{t^-}\left[\Sigma(\cdot, S_{t^-})\right]}. \end{cases}$$

*where $h$ and $\Sigma$ are given in (4.22) and (4.23), respectively.*

## 5. Appendix A

We will give the following sufficient conditions (see for [7] and [15]) which ensure strong existence and strong uniqueness for solutions to system (2.1).

**Assumption C**

**(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 $K_0(t,x,\zeta), K_1(t,x,y,\zeta)$ $\mathbb{R}-$valued, jointly continuous functions in $(t,x,y)$.

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

$$\begin{aligned} |b_0(t,x)|^2 &\le C(1+|x|^2); & |\sigma_0(t,x)|^2 &\le C(1+|x|^2) \\ |b_1(t,x,y)|^2 &\le C(1+|x|^2+|y|^2); & |\sigma_1(t,y)|^2 &\le C(1+|y|^2) \\ \int_Z |K_0(t,x,\zeta)|^2\nu(d\zeta) &\le C(1+|x|^2); & \int_Z |K_1(t,x,y,\zeta)|^2\nu(d\zeta) &\le C(1+|x|^2+|y|^2) \end{aligned} \tag{5.1}$$

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

$$\begin{aligned} &|b_0(t,x) - b_0(t,x')| \le L|x-x'| \qquad |\sigma_0(t,x) - \sigma_0(t,x')| \le L|x-x'| \\ &|b_1(t,x,y) - b_1(t,x',y)| \le L(|x-x'| + |y-y'|) \qquad |\sigma_1(t,y) - \sigma_1(t,y')| \le L|y-y'| \\ &\int_Z |K_0(t,x,\zeta) - K_0(t,x',\zeta)|^2\nu(d\zeta) \le L|x-x'|^2 \\ &\int_Z |K_1(t,x,y,\zeta) - K_1(t,x',y',\zeta)|^2\nu(d\zeta) \le L(|x-x'|^2 + |y-y'|^2) \end{aligned} \tag{5.2}$$

We refer to (5.1) and (5.2) respectively as 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 deduced by those given in [7], Appendix A.

## 6. Appendix B

PROOF OF THEOREM 3.12

Let $(\mu, \widetilde{Y})$ be a weak solution to 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 $\eta_n$ of $\mathcal{F}_t^{\widetilde{Y}}$-stopping times, where $\eta_n$ tends to $\infty$ with $n$, and probability measures $\widetilde{Q}_n$ equivalent to $\widetilde{P}$ such that

$$\mu_{t\wedge\eta_n}\left(F(\cdot, \widetilde{Y}_{t\wedge\eta_n})\right) - \int_0^{t\wedge\eta_n} \mu_s\left(L^{X,Y}F(\cdot, \widetilde{Y}_{s\wedge\eta_n})\right) ds \tag{6.1}$$

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

It is sufficient to prove (6.1) for functions of the type $F(t, x, y) = f(t, x)g(y)$.

Remember that

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

by applying Itô's formula we get

$$\begin{aligned} dg(\widetilde{Y}_t) &= g'(\widetilde{Y}_{t-})d\widetilde{Y}_t + \frac{1}{2}g''(\widetilde{Y}_t)\sigma_1^2(t)dt + \int_{\mathbb{R}}\left[g(\widetilde{Y}_{t-} + x) - g(\widetilde{Y}_{t-})\right]\widetilde{m}(dt, dx) - g'(\widetilde{Y}_{t-})\int_{\mathbb{R}} x\,\widetilde{m}(dt, dx) \\ &= g'(\widetilde{Y}_{t-})\sigma_1(t)d\widetilde{I}_t + \frac{1}{2}g''(\widetilde{Y}_t)\sigma_1^2(t)dt + \int_{\mathbb{R}}\left[g(\widetilde{Y}_{t-} + x) - g(\widetilde{Y}_{t-})\right]\widetilde{m}(dt, dx). \end{aligned}$$

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

$$\mu_t(f) = f(0, x_0) + \int_0^t \mu_s(L^X f)ds + \int_0^t\int_{\mathbb{R}} w_s^\mu(f, x)m^\mu(ds, dx) + \int_0^t h_s^\mu(f)dI_s^\mu$$

and by the product rule

$$\begin{aligned} d(\mu_t(f)g(\widetilde{Y}_t)) &= \mu_{t^-}(f)\left(g'(\widetilde{Y}_t)\sigma_1(t)d\widetilde{I}_t + \frac{1}{2}g''(\widetilde{Y}_t)\sigma_1^2(t)dt + \int_{\mathbb{R}}\left[g(\widetilde{Y}_{t-} + x) - g(\widetilde{Y}_{t-})\right]\widetilde{m}(dt, dx)\right) + \\ &\quad + g(\widetilde{Y}_{t-})\left(\mu_t(L^X f)dt + \int_{\mathbb{R}} w_t^\mu(f, x)m^\mu(dt, dx) + h_t^\mu(f)dI_t^\mu\right) + \\ &\quad + \sigma_1(t)h_t^\mu(f)g'(\widetilde{Y}_t)dt + \int_{\mathbb{R}} w_t^\mu(f, x)\left(g(\widetilde{Y}_{t-} + x) - g(\widetilde{Y}_{t-})\right)\widetilde{m}(dt, dx) \\ &= \mu_{t^-}(f)g'(\widetilde{Y}_t)\mu_t(b_1)dt + \frac{1}{2}\mu_{t^-}(f)g''(\widetilde{Y}_t)\sigma_1^2(t)dt + g(\widetilde{Y}_{t-})\mu_t(L^X f)dt + +\sigma_1(t)h_t^\mu(f)g'(\widetilde{Y}_t)dt \\ &\quad + \int_{\mathbb{R}}(\mu_t(f) + w_t^\mu(f, x))\left[g(\widetilde{Y}_{t-} + x) - g(\widetilde{Y}_{t-})\right]\mu_t(\lambda_t\phi_t(dx))dt + dM_t^{fg} \end{aligned} \tag{6.2}$$

where in (6.2) we used the equality

$$\widetilde{m}(dt, dx) = m^\mu(dt, dx) + \mu_{t^-}(\lambda_t\phi_t(dx))dt, \quad d\widetilde{I}_t = dI_t^\mu + \mu_t\Big(\frac{b_1}{\sigma_1}\Big)dt$$

and by $dM_t^{fg}$ we mean

$$\begin{aligned} dM_t^{fg} &= \{\mu_t(f)\sigma_1(t)g'(\widetilde{Y}_t) + g(\widetilde{Y}_t)h_t^\mu(f)\}\, dI_t^\mu + \int_{\mathbb{R}}\{\mu_{t^-}(f) + w_t^\mu(f, x)\}\left[g(\widetilde{Y}_{t-} + x) - g(\widetilde{Y}_{t-})\right]m^\mu(dt, dx) + \\ &\quad + g(\widetilde{Y}_{t-})\int_{\mathbb{R}} w_t^\mu(f, x)\, m^\mu(dt, dx). \end{aligned} \tag{6.3}$$

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

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

where $b_2(t, x, y)$ is given in (**v**) of Definition 3.8.

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

$$\widetilde{L}_{\eta_n} = \frac{d\widetilde{Q}_n}{d\widetilde{P}} = \mathcal{E}\left(\int_0^{\eta_n} \mu_s\left(\frac{b_1}{\sigma_1}\right) d\widetilde{I}_s\right) = \exp\left(\int_0^{\eta_n} \mu_s\left(\frac{b_1}{\sigma_1}\right) d\widetilde{I}_s - \frac{1}{2}\int_0^{\eta_n} \mu_s^2\left(\frac{b_1}{\sigma_1}\right) ds\right). \tag{6.4}$$

By Girsanov's Theorem

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

is a $(\widetilde{Q}_n, \mathcal{F}_t^{\widetilde{Y}})$-Brownian motion and, $M_{t\wedge\eta_n}^{fg}$ is a $(\widetilde{Q}_n, \mathcal{F}_t^{\widetilde{Y}})$-martingale since the following estimations hold (see Remark 3.10):

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

$$\mathbb{E}^{\widetilde{Q}_n}\int_0^{T\wedge\eta_n} |h_t^\mu(f)|\mu_t(\lambda_t\phi_t(dx))dt \leq B_f\mathbb{E}^{\widetilde{Q}_n}\int_0^{T\wedge\eta_n} \mu_t\left(b_2\right)dt \leq B_f n < \infty$$

Finally, from the expressions of $w^\mu$, $h^\mu$ and the generator $L^{X,Y}$, equation (6.2) implies that

$$d\,\mu_{t\wedge\eta_n}\left(fg(\widetilde{Y}_{t\wedge\eta_n})\right) = \mu_{t\wedge\eta_n}\left(L^{X,Y}fg(\widetilde{Y}_{t\wedge\eta_n})\right)dt + dM_{t\wedge\eta_n}^{fg}, \tag{6.5}$$

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

By Corollary 3.4 in [18] if uniqueness holds for the FMP$(L^{X,Y}, x_0, y_0)$ then there exists a measurable function $H_t : D_{\mathbb{R}}[0,T] \to \mathcal{P}(\mathbb{R})$ such that $\pi_t = H_t(Y)$ $P$-a.s. and $\mu_t \mathbb{I}_{\{t<\eta_n\}} = H_t(\widetilde{Y})\mathbb{I}_{\{t<\eta_n\}}$ $\widetilde{Q}_n$-a.s.

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

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

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

Proof of Theorem 3.14

With the same passages we did for proving the equality (6.2) it can be shown that

$$\begin{aligned} d(\mu_t(f)g(Y)) &= \mu_t(f)g'(Y_t)\mu_t(b_1)dt + \frac{1}{2}\mu_t(f)g''(Y_t)\sigma_1^2(t)dt + g(Y_t)\mu_t(L^Xf)dt + \sigma_1(t)h_t^\mu(f)g'(Y_t)dt \\ &+ \int_{\mathbb{R}} \left(\mu_t(f) + w_t^\mu(f,x)\right)\left[g(Y_{t-}+x) - g(Y)_{t-})\right]\mu_t(\lambda_t\phi_t(dx))dt + dm_t^{fg} \end{aligned}$$

where by $m_t^{fg}$ we mean

$$\begin{aligned} dm_t^{fg} &= \{\sigma_1(t)\mu_t(f)g'(Y_t) + g(Y_t)h_t^\mu(f)\}\,dI_t^\mu + \int_{\mathbb{R}}\{\mu_{t-}(f) + w_t^\mu(f,x)\}\left[g(Y_{t-}+x) - g(Y_{t-})\right]m^\mu(dt,dx) + \\ &+ g(Y_{t-})\int_{\mathbb{R}} w_t^\mu(f,x)\,m^\mu(dt,dx). \end{aligned}$$

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

From the hypothesis of equivalence of the two measures $\pi_{t^-}(\lambda_t\phi_t(dx))dt$ ans $\mu_{t^-}(\lambda_t\phi_t(dx))dt$, there exists an $\mathcal{F}_t^Y$-predictable process $\Psi(t,x) > -1 \quad \pi_t(\lambda_t\phi_t(dx))dt - a.e.$ such that

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

Recalling that $I_t^\mu = I_t - \int_0^t \left\{\mu_s\left(\frac{b_1}{\sigma_1}\right) - \pi_s\left(\frac{b_1}{\sigma_1}\right)\right\} ds$, we define

$$\begin{aligned}\tau_n :=& T \wedge \inf\left\{t \geq 0 : \int_0^t \left|\mu_s\left(\frac{b_1}{\sigma_1}\right) - \pi_s\left(\frac{b_1}{\sigma_1}\right)\right|^2 ds \geq n\right\} \wedge \inf\left\{t\geq 0 : \int_0^t |\mu_s(b_2)|\, ds \geq n\right\} \\ &\wedge \inf\left\{t \geq 0 : \int_0^t \int_{\mathbb{R}} |\Psi(s,x)|^2\, \pi_s(\lambda_s\phi_s(dx))ds \geq n\right\}\end{aligned} \tag{6.6}$$

and the following change of measure

$$\frac{dQ_n}{dP}\Big|_{\mathcal{F}_t^Y} = \Lambda_{t\wedge\tau_n} \tag{6.7}$$

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

Girsanov theorem implies that $I_t - \displaystyle\int_0^t \left\{\mu_s\left(\frac{b_1}{\sigma_1}\right) - \pi_s\left(\frac{b_1}{\sigma_1}\right)\right\} \mathbb{I}_{s<\tau_n} ds$ 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 3.12 we get that $m^{fg}_{t\wedge\tau_n}$ is a $(Q_n, \mathcal{F}_t^Y)$-martingale and so the pair $(\mu_t, Y_t)$ solves the stopped $FMP(L^{X,Y}, x_0, y_0)$.

Finally, by Corollary 3.4 of [18], there exists a functional $H$ such that

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

Nevertheless $Q_n$ and $P$ are equivalent measures, therefore

$$\mu_t \mathbb{I}_{t<\tau_n} = H_t(Y)\mathbb{I}_{t<\tau_n} \qquad P - \text{a.s.}$$

$\tau_n$ is an increasing sequence, so there exists $P-$a.s. $n(\omega)$ such that $\forall n > n(\omega)$, $\tau_n(\omega) = T$.
Taking $n \to \infty$, we get $\mu_t = \pi_t$ $P-$a.s. □

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

**Proposition 6.1** *Under Assumptions C and one of the following conditions*

$$\sup_{t,x} \nu(d^0(t,x)) + \sup_{t,x,y} \nu(d^1(t,x,y)) < \infty \tag{6.8}$$

*or*

$$\sup_{t,x,y} \int_Z \{|K_0(t,x,\zeta)| + |K_1(t,x,y,\zeta)|\}\nu(d\zeta) < \infty \tag{6.9}$$

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

Proof. It is sufficient to apply Theorem 3.3 in [18] after having checked the hypotheses are satisfied. By Assumptions C 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 $\mathcal{D}_L \subset C_b^{1,2,2}([0,T]\times\mathbb{R}\times\mathbb{R})$, such that for $f \in \mathcal{D}_L$, $L^{X,Y} f \in C_b([0,T]\times\mathbb{R}\times\mathbb{R})$.
We choose as $\mathcal{D}_L$ the set of functions in $C_b^{1,2,2}([0,T]\times\mathbb{R}\times\mathbb{R})$ having compact support with respect to $(x,y)$ uniformly in $t$; then there exists $R_f > 0$ such that for $|x| > R_f$ and $|y| > R_f$, $f(t,x,y) = 0, \forall t \in [0,T]$.

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

$$\begin{aligned} L^{X,Y} f(t,x,y) \; &= \; \frac{\partial f}{\partial t} + b_0(t,x)\frac{\partial f}{\partial x} + b_1(t,x,y)\frac{\partial f}{\partial y} + \frac{1}{2}\sigma_0^2(t,x)\frac{\partial^2 f}{\partial x^2} + \rho\sigma_0(t,x)\sigma_1(t,y)\frac{\partial^2 f}{\partial x \partial y} + \\ &+ \frac{1}{2}\sigma_1^2(t,y)\frac{\partial^2 f}{\partial y^2} + \int_Z \left(f(t, x + K_0(t,x,\zeta), y + K_1(t,x,y,\zeta)) - f(t,x,y)\right)\nu(d\zeta) \end{aligned}$$

thus, under (6.8), since

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

we get, by (5.1) that $\forall f \in \mathcal{D}_L$ there exists a constant $C_f > 0$ such that

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

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

The same result can be obtained under (6.9). In fact

$$\left| \int_Z (f(t, x + K_0(t,x,,\zeta), y + K_1(t,x,y,\zeta)) - f(t,x,y)) \; \nu(d\zeta) \right| \leq$$

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

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

## References


[1] Bäuerle N. and Rieder U. (2007). Portfolio optimization with jumps and unobservable intensity process. *Mathematical Finance* **17** (2) 205-224.

[2] Bjork T., Kabanov Y. and Runggaldier W. (1997). Bond market structure in presence of marked point processes. *Mathematical Finance* **7** (2) 211-223.

[3] Brémaud P. (1980). Point Processes and Queues. *Springer-Verlag.*

[4] Ceci C. (2006). Risk minimizing hedging for a partially observed high frequency data model. *Stochastics: An International Journal of Probability and Stochastic Processes* **78** (1) 13-31.

[5] Ceci C. (2010). Utility maximization under restricted information for jump market models. Technical Report, http://www.sci.unich.it/tecrep.

[6] Ceci C. and Gerardi A. (2000). Filtering of a Markov Jump Process with Counting Observations. *Appl. Math. Optim.* **42** 1-18.

[7] Ceci C. and Gerardi A. (2006). A model for high frequency data under partial information: a filtering approach. *International Journal of Theoretical and Applied Finance* **9** (4) 1-22.

[8] Ceci C. and Gerardi A (1998). Partially observed control of a Markov jump process with counting observations: equivalence with the separated problem. *Stochastic Process. Appl.* **78** 245-260.

[9] Doléans-Dade C. (1970). Quelques applications de la formule de changement de variables pour le semi-martingales. *Z. fur W.* **16** 181-194.

[10] Ethier S.N. and Kurtz T.G. (1986). Markov Processes, Characterization, Convergence. *J. Wiley & sons.*

[11] Frey R. (2000). Risk minimization with incomplete information in a model for high-frequency data. *Mathematical Finance* **10** (2) 215-22.

[12] Frey R. and Runggaldier W. (2001). A nonlinear filtering approach to volatility estimation with a view towards high frequency data. *International Journal of Theoretical and Applied Finance* **4** (2) 199-210.

[13] Frey R. and Runggaldier W. (2010). Pricing Credit Derivatives under Incomplete Information: a Nonlinear-Filtering Approach. *Finance and Stochastics* **14** 495-526

[14] Frey R. and Schimdt T. (2011). Pricing and Hedging of Credit Derivatives via the Innovation Approach to Nonlinear Filtering. Preprint, dep. of math., Universität Leipzig, available from http://www.math.uni-leipzig.de/ frey/publications-frey.html to appear in *Finance and Stochastics.*

[15] Gihman I. J. and Skorohod A. V. (1972). Stochastic Differential Equations. *Springer-Verlag.*

[16] Jacod J. and Shiryaev A. (2003). Limit Theorems for Stochastic Processes (2nd ed.) *Springer.*

[17] Kallianpur G. (1980). Stochastic Filtering Theory. *Springer.*

[18] Kurtz T.G. and Ocone D. (1988). Unique characterization of condition distribution in nonlinear filtering. *Annals of Probability* **16** 80-107.

[19] Kurtz T.G. and Nappo G. (2010). The filtered martingale problem. In: The Oxford handbook of nonlinear filtering (Crisan D. and Rozovsky B. eds.). *Oxford University Press* 129-168.

[20] Kliemann W. H., Koch G. and Marchetti F. (1990). On the unnormalized Solution of the Filtering Problem with Counting Process Observations. *IEEE Transactions on Information Theory* **36** 1415-1425.

[21] Lipster R.S. and Shiryaev A. (1977). Statistics of Random Processes I. *Springer -Verlag.*

[22] Mazliak L. (1993). Mixed control problem under partial observations. *Applied Mathematics and Optimization* **27** 57-84.

[23] Pham H. and Quenez M.C. (2001). Optimal portfolio in partially observed stochastic volatility models. *The Annals of Applied Probability* **11** (1) 210-238

[24] Protter P. (1990). Stochastic Integration and Differential Equations: A New Approach (2nd ed.) *Springer-Verlag.*

[25] Sass J. (2007). Utility maximization with convex constraints and partial information. *Acta Appl. Math.* **97** 221-238.

[26] Schweizer M. (1994). Risk minimizing hedging strategies under partial information. *Mathematical Finance* **4** 327-342.

[27] Schweizer M. (2001). A guided tour through quadratic hedging approaches. In: Option Pricing, Interest Rate and Risk Management (Jouini E., Cvitanic J. and Musiela M. eds.). *Cambridge University Press* 538-574.