Nonlinear filtering for Markov systems with delayed observations *

Antonella Calzolari † Patrick Florchinger ‡ Giovanna Nappo §

Abstract

This paper deals with nonlinear filtering problems with delay, i.e. we consider a system $(\boldsymbol{X}, \boldsymbol{Y})$, which can be represented by means of a system $(\boldsymbol{X}, \hat{\boldsymbol{Y}})$, in the sense that $Y_t = \hat{Y}_{a(t)}$, where $a(t)$ is a delayed time transformation. We start with $\boldsymbol{X}$ being a Markov process, and then study Markovian systems, not necessarily diffusive, with correlated noises. The interest is focused on existence of explicit representations of the corresponding filters as functionals depending on the observed trajectory. Different assumptions on the function $a(t)$ are considered.

Key Words: Nonlinear Filtering, Jump Processes, Diffusion processes, Markov Processes, Delay Stochastic Differential Equation

AMS Classification: 93E11, 60G35, 60J75, 60H20, 62M20

1 Introduction

Let $(\boldsymbol{X}, \boldsymbol{Y}) = (X_t, Y_t)_{t \geq 0}$ be a partially observed stochastic system. That is, assume that the *state process* $\boldsymbol{X} = (X_t)_{t \geq 0}$ of the system cannot be directly observed, while the other component $\boldsymbol{Y} = (Y_t)_{t \geq 0}$ is completely observable, and therefore is referred to as the *observation process*. The aim of stochastic nonlinear filtering is to compute the conditional law π_t of the state process

*Partially supported by MURST project *"Processi Stocastici, Modelli Stocastici ed Applicazioni"*

†Dipartimento di Matematica - Università di Roma "Tor Vergata", via della Ricerca Scientifica 1, I 00133 Roma, Italy

‡Département de Mathématiques, Université de Metz, 23 Allée des Oeillets F 57160 Moulins les Metz, France

§Dipartimento di Matematica - Università di Roma "La Sapienza", piazzale A. Moro 2, I 00185 Roma, Italy

at time t, given the observation process up to time t, i.e. the computation of

$$\pi_t(\varphi) = E\big[\varphi(X_t)/\mathcal{F}_t^Y\big], \tag{1}$$

for all functions φ belonging to a determining class, i.e. the best estimate of $\varphi(X_t)$ given the $\sigma-$algebra of the observations up to time t, $\mathcal{F}_t^Y = \sigma\{Y_s,\, s \leq t\}$.

A classical model of partially observed system arises when the system is a $k \times d-$dimensional Markov diffusion process, with state $\boldsymbol{\xi} = (\xi_t)_{t\geq 0}$

$$\xi_t = \xi_0 + \int_0^t b(\xi_s, \eta_s)ds + \int_0^t \sigma(\xi_s, \eta_s)d\beta_s + \int_0^t \tilde{\sigma}(\xi_s, \eta_s)d\omega_s, \qquad t \geq 0, \tag{2}$$

and observation $\boldsymbol{\eta} = (\eta_t)_{t\geq 0}$

$$\eta_t = \int_0^t h(\xi_s)ds + \omega_t, \qquad t \geq 0, \tag{3}$$

where $\boldsymbol{\beta} = (\beta_t)_{t\geq 0}$ and $\boldsymbol{\omega} = (\omega_t)_{t\geq 0}$ are independent Wiener processes and ξ_0 is a random variable independent of $\boldsymbol{\beta}$ and $\boldsymbol{\omega}$.

Under suitable hypotheses on the coefficients, one can prove that the filter $\pi_t^\xi(\varphi) = E\big[\varphi(\xi_t)/\mathcal{F}_t^\eta\big]$ solves a stochastic partial differential equation known as the Kushner-Stratonovich equation and that the unnormalized filter solves a linear stochastic partial differential equation, the Zakai equation (see e.g. Pardoux [17] and the references therein).

We stress that in this model the state process is not necessarily Markovian, while the overall system is Markovian. The same holds for the model studied by Kliemann, Koch and Marchetti [14], where the state is a jump-diffusion process and the observation is a counting process. Recently non-linear filtering has been applied in financial problems in the framework of Bayesian analysis. In particular we quote the papers by Zeng [21], and by Cvitanić, Liptser and Rozovskii [8], in which the observation is a marked point process.

In this paper we consider the filtering problem for a partially observable system $\big(\boldsymbol{X}, \boldsymbol{Y}\big)$, with delayed observations, i.e. such that there exists a process $\hat{\boldsymbol{Y}}$ such that the observation process $\boldsymbol{Y}$ satisfies

$$Y_t = \hat{Y}_{a(t)}, \qquad t \geq 0, \tag{4}$$

where the function $a(\cdot) : [0, \infty) \to [0, \infty)$ is a ***delayed time transformation***, i.e. is non decreasing, with $0 \leq a(t) \leq t$ for all $t \geq 0$. In the following we will use the short notation $\boldsymbol{Y} = \hat{\boldsymbol{Y}} \circ \boldsymbol{\mathcal{A}}$ for (4).

The inspiring example is the simple delayed diffusion model considered in [4], where the state is a Markov diffusion and the observation is available with a fixed delay τ; this corresponds to the choice of $a(t) = (t - \tau)^+$.

The main results of this paper (Theorem 2.3 and Theorem 3.1) are given under the condition that the system $\big(\boldsymbol{X}, \hat{\boldsymbol{Y}}\big)$ is a Markov process for which there exists a feasible filter, i.e. an explicit representation of the filter as a functional depending on the observed trajectory up to time t (see (9)). We stress that we are not necessarily assuming that the signal $\boldsymbol{X}$ itself is Markovian, and that we distinguish between continuous and piecewise constant time transformations. Since for delayed time transformations

$$\mathcal{F}_t^Y \subseteq \mathcal{F}_{a(t)}^{\hat{Y}} \subseteq \mathcal{F}_t^{\hat{Y}},$$

the filtering problem with delayed observations we are dealing with in this paper is connected with the extrapolation (or prediction) problem for the system $\big(\boldsymbol{X}, \hat{\boldsymbol{Y}}\big)$. This problem has been largely studied, see e.g. Liptser and Shiryayev [16] and Pardoux [17], in the case when the observation process $\hat{\boldsymbol{Y}}$ is a diffusion and the signal is a semimartingale. Though our hypotheses imply that the signal $\boldsymbol{X}$ itself is a semimartingale, and in this respect our assumptions are more restrictive, however we are not assuming that the observation process $\hat{\boldsymbol{Y}}$ is a diffusion, and in this respect our assumptions are less restrictive than the usual ones. Moreover, the main concern of extrapolation results is the characterization of the optimal nonlinear extrapolation by means of Kushner-Stratonovitch and/or Zakai type equations. On the contrary we focus on the explicit expression of the filter for the system $\big(\boldsymbol{X}, \boldsymbol{Y}\big)$ with delayed observation, in terms of the feasible version of the filter of the partially observed Markov system $\big(\boldsymbol{X}, \hat{\boldsymbol{Y}}\big)$ and of its associated semigroup. To obtain an explicit representation of the filter is interesting in its own, and moreover it plays a key role in the connected filtering approximation problem (see [5]).

The results concerning continuous time transformations are given in Section 2. The continuity assumption on the function $a(\cdot)$ is crucial in the proofs of Theorem 2.2 and Theorem 2.3 since it implies that

$$\mathcal{F}_t^Y = \mathcal{F}_{a(t)}^{\hat{Y}}, \tag{5}$$

whenever (4) holds. These results allow us to manage different situations illustrated by examples considering both diffusive and jump systems. These examples highlight the differences between the two results; furthermore as an example of a system with correlated noises we study the cubic sensor model (see (15) and (16)), for which we give explicitly the robust Zakai equation for the unnormalized filter (see (17) and (18)).

We conclude Section 2 with a brief discussion of a case which is intermediate between continuous and piecewise constant time transformations, i.e. when the information "arrives by packets", in the sense that the information up to time t is

$$\mathcal{G}_t = \mathcal{F}^Y_{t_i}, \qquad \text{for } t \in [t_i, t_{i+1}),$$

with $\{t_i;\ i \geq 0\}$ a fixed increasing sequence of times (see Remark 2.4). This situation arises when we can observe the trajectory of $\boldsymbol{Y}|_{s\leq t}$ only at the times $t = t_i$; the delayed time transformation being continuous, this corresponds to observing the trajectory of $\hat{\boldsymbol{Y}}|_{s\leq r}$ only at the times $r = a(t_i)$, i.e. $\mathcal{G}_t = \mathcal{F}^Y_{t_i} = \mathcal{F}^{\hat{Y}}_{a(t_i)}$, for $t \in [t_i, t_{i+1})$. This kind of filtration is considered by Schweizer in [18] as an example of delayed information for a financial model.
Section 3 treats the filtering problem with delayed observations when the time transformation $a(\cdot)$ is a step function, i.e. $a(t) = a(t_i)$ for $t \in [t_i, t_{i+1})$, for an increasing sequence of times $t_i < t_{i+1}$. In this case the situation is completely different: whenever (4) holds, the observation process is a (random) step function, the information available during the interval of time $[t_i, t_{i+1})$ is $\mathcal{F}^Y_{[t_i,t_{i+1})} = \sigma(Y_{t_i})$, and therefore

$$\mathcal{F}^Y_t = \sigma(Y_{t_i} = \hat{Y}_{a(t_i)}, i:\ t_i \leq t),$$

which is clearly strictly contained in $\mathcal{F}^{\hat{Y}}_{a(t)}$. Under suitable regularity assumptions on the semigroup associated to $\big(\boldsymbol{X}, \hat{\boldsymbol{Y}}\big)$, the problem can be reduced to a combination of a discrete time filter with the evolution of the associated semigroup (Theorem 3.1).

In the Appendix we first recall the method initiated by Clark [7] and Davis [9] to obtain the robust Zakai equation for partially observed diffusion systems with uncorrelated noises. Then we derive the robust Zakai equation for the cubic sensor problem with correlated noises by applying the results established in [11]. To our knowledge [11] is the only paper in the literature dealing with the robust Zakai equation for partially observed diffusion systems with correlated noises. Note that, in the latter case, robust filters (i.e. feasible filters continuous with respect to the trajectory of the observation process) have also been studied by Elliott and Kohlmann [10].

2 Continuous delayed time transformation

In this section we consider continuous time transformations. The first result of this section (Lemma 2.1) plays a key role in our analysis since it implies the equality (5). After giving the definition of the feasible filter we state our main results, Theorem 2.2 and Theorem 2.3.

Lemma 2.1. *Assume that the function $a(\cdot)$ is a continuous delayed time transformation, and that $Y_t = \hat{Y}_{a(t)}$, for all $t \geq 0$. Then*

$$Y_{a^{-1}(s)} = \hat{Y}_s, \qquad s \geq 0, \tag{6}$$

where

$$a^{-1}(s) = \inf\{u : a(u) \geq s\} \tag{7}$$

is the generalized inverse of $a(\cdot)$.

In the following (6) will be written in the shorter way as $(Y \circ \mathcal{A}^{-1})_s := Y_{a^{-1}(s)}$, $s \geq 0$, or

$$\boldsymbol{\hat{Y}} = \boldsymbol{Y} \circ \boldsymbol{\mathcal{A}^{-1}}. \tag{8}$$

Proof. The proof of (6) is immediate by observing that $Y_{a^{-1}(s)} = \hat{Y}_{a(a^{-1}(s))} = \hat{Y}_s$, since $a(a^{-1}(s)) = s$, $a(\cdot)$ being a nondecreasing continuous function.

The continuity property is crucial, since, together with the fact that a is nondecreasing, with $0 \leq a(t) \leq t$, it implies that $a(0) = 0$ and $Im(a|_{[0,T]}) = [a(0), a(T)] = [0, a(T)]$. Moreover, by the definition (7) of $a^{-1}(s)$, there exists a sequence u_n such that $a(u_n) \geq s$ and $u_n \searrow a^{-1}(s)$. By right continuity $a(u_n) \searrow a(a^{-1}(s))$ (moreover a is nondecreasing), and therefore $a(a^{-1}(s)) \geq s$. Seeking a contradiction suppose that $a(a^{-1}(s)) > s$, then for every $s_0 \in (s, a(a^{-1}(s)))$ it cannot exists a t_0 such that $a(t_0) = s_0 > s$, since otherwise, for n sufficiently large $u_n \leq t_0$ and therefore $a(u_n) \leq a(t_0) = s_0$. Then $Im(a|_{[0,T]})$ does not contain $(s, a(a^{-1}(s)))$, which contradicts the continuity condition on the function $a(\cdot)$.

□

An important feature in nonlinear filtering is to obtain a feasible filter: for the system $(\boldsymbol{X}, \boldsymbol{\hat{Y}})$ we mean that there exists a functional $\hat{U}_s$ for which $\hat{U}_s(\psi|\boldsymbol{y}) = \hat{U}_s(\psi|y(\cdot \wedge s))$ a.s. with respect to the law of $\boldsymbol{\hat{Y}}$, and such that the conditional law $\hat{\pi}_s$ of X_s given $\mathcal{F}_s^{\hat{Y}}$ may be expressed as

$$\hat{\pi}_s(\psi) = E\big[\psi(X_s)/\mathcal{F}_s^{\hat{Y}}\big] = \hat{U}_s(\psi|\hat{Y}_{\cdot\wedge s}). \tag{9}$$

In the following we refer to the above situation by saying that the system $(\boldsymbol{X}, \boldsymbol{\hat{Y}})$ admits a ***feasible filter***; furthermore we identify the functional $\hat{U}_s$ with its underlying measure. For the diffusion case this problem, initiated by Clark [7] and Davis [9] when considering feasible filters continuous with respect to the trajectory of the observation process (i.e. robust filters), has been studied by many authors in different frameworks. When dealing with counting observation this problem has been studied by Brémaud (see [3]) for the doubly stochastic case, and by Kliemann et al. in [14] for more general systems.

Theorem 2.2. *Suppose that the state process* $\boldsymbol{X}$ *is a Markov process with generator* $\boldsymbol{A}$*, and that the observation process* $\boldsymbol{Y}$ *satisfies* $Y_t = \hat{Y}_{a(t)}$*, where* $\hat{\boldsymbol{Y}}$ *is adapted to the filtration* $\mathcal{F}_t^X \vee \mathcal{H}$*, with* $\mathcal{H}$ *a* σ*-algebra independent of* $\mathcal{F}_\infty^X$*, and where the function* $a(\cdot)$ *is a continuous delayed time transformation. Then*

$$\pi_t(\varphi) = E\big[\exp\{\boldsymbol{A}(t - a(t))\}\varphi(X_{a(t)})/\mathcal{F}_{a(t)}^{\hat{Y}}\big]. \tag{10}$$

Furthermore, if the system $(\boldsymbol{X}, \hat{\boldsymbol{Y}})$ *admits a feasible filter, then*

$$\pi_t(\varphi) = \hat{U}_{a(t)}\big(\exp\{\boldsymbol{A}(t - a(t))\}\varphi \,|\, (Y \circ \mathcal{A}^{-1})_{\cdot\wedge a(t)}\big). \tag{11}$$

Proof. The continuity of the function $a(\cdot)$ implies (5) and therefore

$$\pi_t(\varphi) = E\big[\varphi(X_t)/\mathcal{F}_{a(t)}^{\hat{Y}}\big] = E\big[E\big[\varphi(X_t)/\mathcal{F}_{a(t)}^X \vee \mathcal{H}\big]/\mathcal{F}_{a(t)}^{\hat{Y}}\big],$$

which coincides with $E\big[E\big[\varphi(X_t)/\mathcal{F}_{a(t)}^X\big]/\mathcal{F}_{a(t)}^{\hat{Y}}\big]$ by the independence property of $\mathcal{H}$, and then assertion (10) follows. Since the filter is feasible (see (9)), assertion (11) follows immediately by Lemma 2.1.

□

Note that in the previous Theorem equality (11) is more interesting than (10) since it expresses the filter in term of the observed trajectory $\boldsymbol{Y}$, instead of the underlying process $\hat{\boldsymbol{Y}}$.

Before giving some examples of application of the previous result we consider the case when $(\boldsymbol{X}, \hat{\boldsymbol{Y}})$ is a Markov system.

Theorem 2.3. *Assume that* $(\boldsymbol{X}, \hat{\boldsymbol{Y}})$ *is a Markov process with generator* $\boldsymbol{L}$*, and that the observation process* $\boldsymbol{Y}$ *satisfies* $Y_t = \hat{Y}_{a(t)}$*, where the function* $a(\cdot)$ *is a continuous delayed time transformation. Then*

$$\pi_t(\varphi) = E\big[\big(\exp\{\boldsymbol{L}(t - a(t))\}\phi\big)(X_{a(t)}, \hat{Y}_{a(t)})/\mathcal{F}_{a(t)}^{\hat{Y}}\big],$$

where $\phi(x, y) = \varphi(x)$*. Moreover, if* $(\boldsymbol{X}, \hat{\boldsymbol{Y}})$ *admits a feasible filter, then*

$$\pi_t(\varphi) = \hat{U}_{a(t)}\big((\exp\{\boldsymbol{L}(t - a(t))\}\phi)(\cdot, Y_t) \,|\, (Y \circ \mathcal{A}^{-1})_{\cdot\wedge a(t)}\big). \tag{12}$$

Proof. The proof is similar to that of Theorem 2.2: indeed, since $a(\cdot)$ is continuous,

$$\pi_t(\varphi) = E\big[\varphi(X_t)/\mathcal{F}^{\hat{Y}}_{a(t)}\big]$$

and furthermore, for any $r \leq t$

$$\begin{aligned} E\big[\varphi(X_t)/\mathcal{F}^{\hat{Y}}_r\big] &= E\big[E\big[\varphi(X_t)/\mathcal{F}^{X,\hat{Y}}_r\big]/\mathcal{F}^{\hat{Y}}_r\big] \\ &= E\big[\big(\exp\{\boldsymbol{L}(t-r)\}\phi\big)(X_r, \hat{Y}_r)/\mathcal{F}^{\hat{Y}}_r\big]. \end{aligned} \tag{13}$$

□

As a first example consider $\boldsymbol{X}$ being a Markov diffusion, i.e. $\boldsymbol{X} = \boldsymbol{\xi}$, where $\boldsymbol{\xi}$ is given by (2) with the coefficients depending only on the first variable, $\tilde{\sigma} = 0$, and $\boldsymbol{Y} = \hat{\boldsymbol{Y}} \circ \boldsymbol{\mathcal{A}}$, with

$$\hat{Y}_t = \int_0^t h(X_s)ds + \hat{W}_t, \tag{14}$$

where $\hat{\boldsymbol{W}} = (\hat{W}_t)_{t\geq 0}$ is a Wiener process, independent of $\boldsymbol{X}$. This example (already considered in [4]) satisfies the first conditions of Theorem 2.2 with $\mathcal{H} = \mathcal{F}^{\hat{W}}_\infty$. By using the techniques initiated by Clark and Davis one can prove easily that the filter is robust.

If instead of (14) one considers

$$\hat{Y}_t = \int_0^t H\big(X_s, \textstyle\int_0^s \alpha(X_u)du + V_s\big)\, ds + \hat{W}_t,$$

where $\boldsymbol{V}$ and $\hat{\boldsymbol{W}}$ are two independent Wiener processes, both independent of $\boldsymbol{X}$, then the first conditions of Theorem 2.2 are satisfied with $\mathcal{H} = \mathcal{F}^{V,\hat{W}}_\infty$.

It is interesting to note that Theorem 2.3 cannot be applied directly in this framework since $(\boldsymbol{X}, \hat{\boldsymbol{Y}})$ is not a Markov process. Nevertheless it can be applied if we introduce the auxiliary process

$$Z_t = \int_0^t \alpha(X_u)du + V_t$$

and consider the filter of the Markov diffusion process $(\boldsymbol{X}, \boldsymbol{Z})$ given $\hat{\boldsymbol{Y}}$. Using the techniques initiated by Clark and Davis, this filter can be characterized by a functional, from which the functional $\hat{U}_s$ in (9) can be easily obtained by projection, and therefore all the results of Theorem 2.2 hold.

The next example concerns the cubic sensor with correlated noises and delayed observation, i.e. the case when $(\boldsymbol{X}, \hat{\boldsymbol{Y}}) = (\boldsymbol{\xi}, \boldsymbol{\eta})$, where

$$\xi_t = \xi_0 + \sigma\,\beta_t + \tilde{\sigma}\,\omega_t, \tag{15}$$

$$\eta_t = \int_0^t \xi_s^3\,ds + \omega_t, \tag{16}$$

with $\sigma > 0$ and $\tilde{\sigma} \geq 0$. When $\tilde{\sigma} = 0$ the above filtering problem has been studied by Sussmann in [19]. When $\tilde{\sigma} \neq 0$ this system does not satisfies the hypotheses of Theorem 2.2, since, though the state process is Markovian, the noises are correlated. Nevertheless all the hypotheses of Theorem 2.3 are satisfied, since the system is Markovian and admits a robust filter: Let p_0^ξ be the density of ξ_0, then the functional $\hat{U}_t$ is given by

$$\hat{U}_t(dx|\boldsymbol{y}) \propto e^{H(y_t, x - \tilde{\sigma} y_t)}\,\hat{q}_t(x - \tilde{\sigma} y_t|\boldsymbol{y}, p_0^\xi)\,dx, \tag{17}$$

where $H(t,x) = \dfrac{1}{4\,\tilde{\sigma}}\big((x + \tilde{\sigma}\,t)^4 - x^4\big)$, and $\hat{q}_t(x|\boldsymbol{y}, p_0^\xi)$ solves the following robust Zakai equation established in the Appendix

$$\begin{aligned} q_t(x) = p_0^\xi(x) + \int_0^t \tfrac{e^{-H(y_s,x)}}{2} \Big[\sigma^2\,\tfrac{d^2}{dx^2} + 2\,\tilde{\sigma}(x + \tilde{\sigma} y_s)^3\,\tfrac{d}{dx} \\ + \big(3\,\tilde{\sigma}(x + \tilde{\sigma} y_s)^2 - (x + \tilde{\sigma} y_s)^6\big)\Big]\,\big(e^{H(y_s,x)} q_s(x)\big)\,ds. \end{aligned} \tag{18}$$

Finally we point out that Theorem 2.3 can also be applied to the jump-diffusion model with counting observations considered in [14]. In the latter paper the authors have shown that, under suitable conditions, these systems admit a feasible filter which can be represented by means of a recursive algorithm. In general the feasible filter cannot be computed explicitly, and an approximation may be necessary. This approximation problem has been studied in [5] for the jump case, i.e. when $(\boldsymbol{X}, \hat{\boldsymbol{Y}})$ is a Markov process with generator $\boldsymbol{L}$ of the form

$$\begin{aligned} \boldsymbol{L}\phi(x,y) = \lambda_0(x,y) \int \big(\phi(x', y) - \phi(x,y)\big)\mu_0(x,y;dx') \\ + \lambda_1(x,y) \int \big(\phi(x', y+1) - \phi(x,y)\big)\mu_1(x,y;dx'), \end{aligned} \tag{19}$$

where λ_i are measurable functions and μ_i are probability kernels, for $i = 0, 1$.

Remark 2.4. *When the information "arrives by packets", in the sense explained in the Introduction, that is when the information up to time t is $\mathcal{G}_t = \mathcal{F}^{\hat{Y}}_{a(t_i)}$ for $t \in [t_i, t_{i+1})$, assuming we are in the setting of Theorem 2.3 we obtain that the filter is given by*

$$E[\varphi(X_t)/\mathcal{G}_t] = \hat{U}_{a(t_i)}\big(\exp\{\boldsymbol{L}(t - a(t_i))\}\phi|(Y \circ \mathcal{A}^{-1})_{\cdot \wedge a(t_i)}\big), \tag{20}$$

for $t \in [t_i, t_{i+1})$, with $\phi(x,y) = \varphi(x)$. Note that in formula (20), for $t \in [t_i, t_{i+1})$, one uses the trajectory of $\boldsymbol{Y}$ up to time t_i.

As recalled in the Introduction, Schweizer [18] has considered an example of delayed information for a financial model by taking a similar filtration. More precisely in [18] the state $\boldsymbol{X}$ is a Markov diffusion with generator $\boldsymbol{A}$, the information available at time t is $\mathcal{G}_t = \mathcal{F}^{\hat{Y}}_{\tilde{a}(t)}$ where $\tilde{a}(\cdot)$ is a càdlàg delayed time transformation. In this case it corresponds to take $\tilde{a}(t) = a(t_i)$ for $t \in [t_i, t_{i+1})$.

3 Piecewise constant delayed time transformations

As explained in Section 2 (see Lemma 2.1), the continuity assumption on the function $a(\cdot)$ is crucial, since $\mathcal{F}^Y_t = \mathcal{F}^{\hat{Y}}_{a(t)}$. The situation is completely different when the time transformation $a(\cdot)$ is a step function, i.e. $a(t) = a(t_i)$ for $t \in [t_i, t_{i+1})$, for a strictly increasing sequence of times t_i, with $t_0 = 0$. When dealing with this problem in the setting of Theorem 2.3, except for the continuity assumption on $a(\cdot)$, which is substituted by a step-wise assumption, we get for any measurable bounded function φ, when $t_k \le t < t_{k+1}$

$$\pi_t(\varphi) = E\big[\exp\{\boldsymbol{L}(t - a(t_k))\}\phi(X_{a(t_k)}, \hat{Y}_{a(t_k)})/\sigma(\hat{Y}_{a(t_i)}, i \le k)\big],$$

where $\phi(x,y) = \varphi(x)$. Indeed, since $\mathcal{F}^Y_t \subset \mathcal{F}^{\hat{Y}}_{a(t)}$, by (13) and the chain rule for conditional expectations, we have

$$\pi_t(\varphi) = E\big[\exp\{\boldsymbol{L}(t - a(t))\}\phi(X_{a(t)}, \hat{Y}_{a(t)})/\sigma(\hat{Y}_{a(t_i)}, i:\ t_i \le t)\big].$$

As a consequence, when $t_k \le t < t_{k+1}$, we can rewrite the filter $\pi_t(\varphi)$ as

$$\breve{\pi}_{a(t_k)}(\exp\{\boldsymbol{L}(t - a(t_k))\}\phi(\cdot, \hat{Y}_{a(t_k)})) = \breve{\pi}_{s_k}\big(\exp\{\boldsymbol{L}(t - s_k)\}\phi(\cdot, \hat{Y}_{s_k})\big), \quad (21)$$

where $s_k = a(t_k)$ and $\breve{\pi}_{s_k}$ denotes the discrete time filter for the system $\{(X_{s_k}, \hat{Y}_{s_k});\ k \ge 0\}$.

To compute the above quantities one could use the results established by Joannides and Le Gland in [13], with a slight modification. However, our case is much simpler than the one considered in [13], and a representation of the filter can be obtained directly.

Theorem 3.1. *Assume that $(\boldsymbol{X}, \hat{\boldsymbol{Y}})$ is a Markov process with generator $\boldsymbol{L}$ and that the observation process $\boldsymbol{Y}$ satisfies*

$$Y_t = \hat{Y}_{a(t)},$$

where the delayed time transformation $a(t)$ is a step function.
Assume further that the semigroup $\exp\{\boldsymbol{L}t\}$ *of the Markov process* $(\boldsymbol{X}, \hat{\boldsymbol{Y}})$ *has the property that whenever the initial distribution of* $(X_0, \hat{Y}_0)$ *is*

$$\mu(dx, dy) = p(x)\, dx\, \delta_{\hat{y}}(dy),$$

then the distribution of $(X_u, \hat{Y}_u)$ *at time* u *has a joint density* $\hat{p}_u$ *given by*

$$\hat{p}_u(x, y|p, \hat{y})\, dx\, dy = \big(\exp\{\boldsymbol{L}^* u\}\mu\big)(dx, dy),$$

where $\boldsymbol{L}^*$ *is the adjoint of* $\boldsymbol{L}$.
Assume finally that the distribution of X_0 *is* $p_0^X(x)dx$, $\hat{Y}_0 = y_0$, *and denote*

$$p_0(x) = p_0^X(x), \quad p_{k+1}(x) = \frac{\hat{p}_{a(t_{k+1})-a(t_k)}(x, Y_{t_{k+1}}|p_k, Y_{t_k})}{\int \hat{p}_{a(t_{k+1})-a(t_k)}(\xi, Y_{t_{k+1}}|p_k, Y_{t_k}) d\xi}, \qquad k \geq 0.$$

Then, for any t, *the filter* π_t *is given by* $\pi_0(dx) = p_0^X(x)dx$, *and*

$$\pi_t(dx) = \hat{p}^X_{t-a(t_k)}(x|p_k, Y_{t_k})dx, \qquad \textit{for } t_k \leq t < t_{k+1}, \qquad k \geq 0,$$

where

$$\hat{p}_u^X(x|p, \hat{y}) := \int \hat{p}_u(x, y|p, \hat{y})\, dy.$$

Proof. Taking (21) into account, we get

$$\pi_t(\cdot) = \int \big(\exp\{\boldsymbol{L}^*(t - s_k)\}\mu_k\big)(\cdot, dy), \tag{22}$$

with

$$\mu_k(dx, dy) = \check{\pi}_{s_k}(dx)\delta_{\hat{Y}_{s_k}}(dy),$$

and, as a consequence, we only need to compute the discrete time filter

$$\check{\pi}_{s_k}(dx) = P\big[X_{s_k} \in dx/\sigma(\hat{Y}_{s_i}, i \leq k)\big].$$

To this end, we evaluate the quantities

$$P\big[(X_u, \hat{Y}_u) \in (dx, dy)/\sigma(\hat{Y}_{s_i}, i:\ s_i \leq u)\big],$$

by the following procedure:

For $0 < u < s_1$, since X_0 has a density p_0,

$$\begin{aligned} P\big[(X_u, \hat{Y}_u) \in (dx, dy)/\sigma(\hat{Y}_{s_i}, i:\ s_i \leq u)\big] &= P\big[(X_u, \hat{Y}_u) \in (dx, dy)\big] \\ &= \hat{p}_u(x, y|p_0, y_0)dx\, dy, \end{aligned}$$

and for $u = s_1$

$$P\big[X_{s_1} \in dx/\sigma(\hat{Y}_{s_1})\big] = P\big[X_{s_1} \in dx/\hat{Y}_{s_1}\big]$$
$$= \frac{\hat{p}_{s_1}(x, \hat{Y}_{s_1}|p_0, y_0)}{\int \hat{p}_{s_1}(\xi, \hat{Y}_{s_1}|p_0, y_0)d\xi}\ dx =: p_1(x)dx.$$

Then, for $s_1 < u < s_2$

$$P\big[(X_u, \hat{Y}_u) \in (dx, dy)/\sigma(\hat{Y}_{s_i}, i:\ s_i \leq u)\big] = P\big[(X_u, \hat{Y}_u) \in (dx, dy)/\sigma(\hat{Y}_{s_1})\big]$$
$$= \hat{p}_{u-s_1}(x, y|p_1, \hat{Y}_{s_1})dx\, dy,$$

and for $u = s_2$

$$P\big[X_{s_2} \in dx/\sigma(\hat{Y}_{s_i},\ i \leq 2)\big] = P\big[X_{s_2} \in dx/\hat{Y}_{s_2}, \hat{Y}_{s_1}\big]$$
$$= \frac{\hat{p}_{s_2-s_1}(x, \hat{Y}_{s_2}|p_1, \hat{Y}_{s_1})}{\int \hat{p}_{s_2-s_1}(\xi, \hat{Y}_{s_2}|p_1, \hat{Y}_{s_1})d\xi}\ dx =: p_2(x)dx.$$

Therefore, all the quantities we need can be easily computed by iterating these steps.
Recalling (22) and that

$$\hat{p}_u(x, y|p, \hat{y})dx\ dy = \big(\exp\{\boldsymbol{L}^* u\}\mu\big)(dx, dy)$$

for $\mu(dx, dy) = p(x)dx\, \delta_{\hat{y}}(dy)$, we get the thesis.

$\square$

Note that, for $t \in [t_k, t_{k+1})$, the filter π_t, as given in the theorem, depends explicitly on Y_{t_k}, but also indirectly on $Y_{t_1}, \cdots, Y_{t_k}$, through the density p_k.

It is also interesting to note that if $\boldsymbol{X}$ is a Markov process with generator $\boldsymbol{A}$, with the property that whenever the initial distribution of X_0 has a density, then the distribution of X_u at time u has a density, we have

$$\hat{p}^X_u(x|p, \hat{y}) = \int \hat{p}_u(x, y|p, \hat{y})dy = \big(\exp\{\boldsymbol{A}^* u\}\mu^X\big)(dx),$$

with $\mu^X(dx) = p(x)dx$, and therefore the computation of the filter becomes much easier, and furthermore, for $t \in [t_k, t_{k+1})$, the explicit dependence on Y_{t_k} of the filter π_t disappears.

4 Appendix

The purpose of this section is to compute the robust filter for the cubic sensor model with correlated noises. With this aim we first recall how to compute the robust filter when dealing with the classical model of a partially observed diffusive system $(\boldsymbol{\xi}, \boldsymbol{\eta})$ given by (2) and (3). In this case the generator $\boldsymbol{L}$ is

$$\boldsymbol{L}f(x,y) = (b(x,y), h(x)) \cdot \nabla f(x,y) + \frac{1}{2} tr\{\nabla^2 f(x,y)\Sigma(x,y)\Sigma^*(x,y)\},$$

where $\Sigma(x,y) = \begin{pmatrix} \sigma(x,y) & \tilde{\sigma}(x,y) \\ 0 & Id \end{pmatrix}$.

Assuming that all the coefficients are bounded one can prove (see for example Pardoux [17]) that the filter $\pi_t^\xi(\varphi) = E\big[\varphi(\xi_t)/\mathcal{F}_t^\eta\big]$ can be obtained via the Kallianpur-Striebel formula

$$\pi_t^\xi(\varphi) = \frac{\rho_t^\xi(\varphi)}{\rho_t^\xi(\mathbf{1})},$$

where $\mathbf{1}(x) = 1$, and ρ_t^ξ is the so-called unnormalized filter. The latter solves the linear stochastic partial differential equation known as the Zakai equation (see [20])

$$\rho_t^\xi(\varphi) = \mu_0^\xi(\varphi) + \int_0^t \rho_s^\xi(\boldsymbol{A}_{\eta_s}\varphi)ds + \int_0^t \rho_s^\xi(\boldsymbol{\Lambda}_{\eta_s}\varphi)d\eta_s$$

where μ_0^ξ is the distribution of ξ_0, $\boldsymbol{A}_y\varphi(x) = \boldsymbol{L}\phi(x,y)$, where $\phi(x,y) = \varphi(x)$, i.e. $\boldsymbol{A}_y$ is the second order differential operator defined by

$$\begin{aligned}\boldsymbol{A}_y\varphi(x) = b(x,y) \cdot \nabla\varphi(x) &+ \frac{1}{2} tr\{\nabla^2\varphi(x)\sigma(x,y)\sigma^*(x,y)\} \\ &+ \frac{1}{2} tr\{\nabla^2\varphi(x)\tilde{\sigma}(x,y)\tilde{\sigma}^*(x,y)\} \end{aligned} \tag{23}$$

and $\boldsymbol{\Lambda}_y$ is the first order differential operator defined by

$$\boldsymbol{\Lambda}_y\varphi(x) = h(x)\varphi(x) + \tilde{\sigma}(x,y)\nabla\varphi(x). \tag{24}$$

Remark 4.1. *When $\boldsymbol{\xi}$ is a Markov diffusion the above Zakai equation can also be obtained, under some additional hypotheses, when h is unbounded by means of the same arguments when $\tilde{\sigma} = 0$ (see Hopkins [12] or Baras, Blankenship and Hopkins [1]) and by different techniques when $\tilde{\sigma}$ does not depend on y and $\boldsymbol{\eta}$ is a one-dimensional process (see Florchinger [11]).*

Furthermore note that, under suitable hypotheses, a Zakai equation can be obtained when the state process $\boldsymbol{\xi}$ is a general Markov process (not necessarily given by (2)), and the observation process $\boldsymbol{\eta}$ is a diffusion process given by (3), with $\boldsymbol{\omega}$ independent of $\boldsymbol{\xi}$ (see for example Bhatt, Kallianpur and Karandikar [2]).

If the density p^ξ_t of the unnormalized filter ρ^ξ_t exists and is regular enough, one can easily deduce from the above Zakai equation that it solves the following linear stochastic partial differential equation

$$p^\xi_t = p^\xi_0 + \int_0^t \boldsymbol{A}^*_{\eta_s} p^\xi_s ds + \int_0^t \boldsymbol{\Lambda}^*_{\eta_s} p^\xi_s d\eta_s,$$

where p^ξ_0 is the density of ξ_0.

Starting from the above equation one can get the robust Zakai equation. First, assume that $\tilde{\sigma} = 0$ and set

$$q^\xi_t(x) = p^\xi_t(x) e^{-h(x)\eta_t}.$$

Then q^ξ_t solves the robust Zakai equation (see Clark [7] and Davis [9]), i.e. the deterministic equation with random coefficients

$$q^\xi_t(x) = p^\xi_0(x) + \int_0^t \left[e^{-h(x)\eta_s}\, \boldsymbol{A}^*_{\eta_s} \left(q^\xi_s(\cdot)\, e^{h(\cdot)\eta_s} \right)(x) - \tfrac{1}{2}\, h^2(x) q^\xi_s(x) \right] ds. \tag{25}$$

Now, assume that all the coefficients do not depend on y and that $\boldsymbol{\eta}$ is a one-dimensional process. In this correlated case, the robust Zakai equation has been obtained in [11] as follows.

Let Φ_t be the flow associated with the function $\tilde{\sigma}$, i.e. the unique solution of $\Phi_t(x) = x + \int_0^t \tilde{\sigma}(\Phi_s(x))\, ds$, and H be the function defined on $\mathbb{R} \times \mathbb{R}^k$ by

$$H(t,x) = \int_0^t h(\Phi_s(x))\, ds.$$

Then by setting

$$q^\xi_t(x) = p^\xi_t(\Phi_{\eta_t}(x))\ |J\Phi_{\eta_t}(x)|\ e^{-H(\eta_t,x)},$$

where $J\psi$ denotes the Jacobian of a regular function ψ, one gets, by applying the generalization of Itô formula proved by Kunita (see Theorem 8.1 in [15]), the following robust Zakai equation

$$q^\xi_t(x) = p^\xi_0(x) + \int_0^t e^{-H(\eta_s,x)}\ |J\Phi_{\eta_s}(x)|\ \boldsymbol{C^h} \left(|J\Phi_{\eta_s}(\cdot)|^{-1}\ e^{H(\eta_s,\cdot)}\, q^\xi_s(\cdot) \right)(x)\, ds,$$

where $\boldsymbol{C^h}$ is a second order differential operator, which, when also the signal process is one-dimensional, is given by

$$\begin{aligned} \boldsymbol{C^h}\psi(x) &= \boldsymbol{A}^*\psi(x) + \frac{1}{2}\left[\tilde{h}'(x)\,\tilde{\sigma}(x) - \tilde{h}^2(x) \right] \psi(x) \\ &\quad + \left[\tilde{h}(x)\tilde{\sigma}(x) - \frac{1}{2}\,\tilde{\sigma}'(x)\tilde{\sigma}(x) \right]\ \psi'(x) - \frac{1}{2}\,\tilde{\sigma}^2(x)\,\psi''(x), \end{aligned}$$

with $\tilde{h}(x) = h(x) - \tilde{\sigma}'(x)$.

We now explain how to get the functional $\hat{U}_t$ for the model considered. For any continuous (deterministic) function $\boldsymbol{y}$, and for any probability density $\hat{p}_0$ denote by $\hat{q}_t(x|\boldsymbol{y};\hat{p}_0)$ the solution of

$$\begin{aligned} q_t(x) &= \hat{p}_0(x) \\ &+ \int_0^t e^{-H(y_s,x)} \, |J\Phi_{y_s}(x)| \, \boldsymbol{C}^{\boldsymbol{h}} \Big(|J\Phi_{y_s}(\cdot)|^{-1} \, e^{H(y_s,\cdot)} q_s(\cdot) \Big)(x) \, ds, \end{aligned} \tag{26}$$

$$\hat{\rho}_t(dx|\boldsymbol{y};\hat{p}_0) := \hat{q}_t(\Phi_{y_t}^{-1}(x)|\boldsymbol{y};\hat{p}_0) \, |J\Phi_{y_t}(\cdot)|^{-1} \, e^{H(y_t,\cdot)} \, dx,$$

and by

$$\hat{U}_t(\varphi|\boldsymbol{y}) := \frac{\hat{\rho}_t(\varphi|\boldsymbol{y};\hat{p}_0)}{\hat{\rho}_t(\mathbf{1}|\boldsymbol{y};\hat{p}_0)}. \tag{27}$$

Note that $\hat{q}_t$, $\hat{\rho}_t$, and $\hat{U}_t$ depend on the trajectory $\boldsymbol{y}$ restricted to the interval $[0,t]$. Then, with the above notations

$$\rho_t^{\xi}(dx) = p_t^{\xi}(x)dx = \hat{\rho}_t(dx|\boldsymbol{\eta};p_0^{\xi}) = \hat{q}_t(\Phi_{\eta_t}^{-1}(x)|\boldsymbol{\eta};p_0^{\xi}) \, |J\Phi_{\eta_t}(\cdot)|^{-1} \, e^{H(\eta_t,\cdot)} \, dx,$$

and consequently

$$\pi_t^{\xi}(\varphi) = \hat{U}_t(\varphi|\boldsymbol{\eta}) = \frac{\hat{\rho}_t(\varphi|\boldsymbol{\eta};p_0^{\xi})}{\hat{\rho}_t(\mathbf{1}|\boldsymbol{\eta};p_0^{\xi})}.$$

We end by observing that when $\tilde{\sigma} = 0$, then $\Phi_t(x) = x$ so that $\hat{\rho}_t(dx|\boldsymbol{y};\hat{p}_0) = \hat{q}_t(x|\boldsymbol{y};\hat{p}_0)\, e^{h(x)y_t} dx$, and the equation for $\hat{q}_t(x|\boldsymbol{y};\hat{p}_0)$ simplifies to the Zakai equation (25) in this setting.

Remark 4.2. *When the observation coefficient h is unbounded, and the noises are correlated, the filter can be characterized as the solution of the above robust Zakai equation by using the results by Florchinger [11] and by Cannarsa and Vespri [6].*

The cubic sensor model with correlated noises considered in (15) and (16) falls in the models considered in the above Remarks (the growth restriction on h stated in [11] is satisfied in the polynomial case). For this system one gets $\Phi_t(x) = x + \tilde{\sigma}\, t$ and $H(t,x) = \dfrac{1}{4\,\tilde{\sigma}} \big((x+\tilde{\sigma}\, t)^4 - x^4\big)$, and therefore

$$q_t^{\xi}(x) = p_t^{\xi}(x + \tilde{\sigma}\, \eta_t)\, e^{-H(\eta_t,x)}$$

satisfy the following robust Zakai equation

$$\begin{aligned} q_t^{\xi}(x) = p_0^{\xi}(x) + \int_0^t \tfrac{e^{-H(\eta_s,x)}}{2} \Big[\sigma^2 \tfrac{d^2}{dx^2} + 2\,\tilde{\sigma}\Phi_{\eta_s}^3(x) \tfrac{d}{dx} \\ + \big(3\,\tilde{\sigma}\Phi_{\eta_s}^2(x) - \Phi_{\eta_s}^6(x)\big)\Big] \big(e^{H(\eta_s,x)} q_s^{\xi}(x)\big)\, ds. \end{aligned}$$

In this example (26) reduces to (18), and then by (27) one gets the functional $\hat{U}_t$ in (17).

References

[1] Baras, J. S., Blankenship, G. L., and Hopkins, Jr., W. E. Existence, uniqueness, and asymptotic behavior of solutions to a class of Zakai equations with unbounded coefficients. *IEEE Trans. Automat. Control 28*, 2 (1983), 203–214.

[2] Bhatt, A. G., Kallianpur, G., and Karandikar, R. L. Uniqueness and robustness of solution of measure-valued equations of nonlinear filtering. *Ann. Probab. 23*, 4 (1995), 1895–1938.

[3] Brémaud, P. *Point processes and queues. Martingale dynamics.* Springer-Verlag, New York, 1981. Springer Series in Statistics.

[4] Calzolari, A., Florchinger, P., and Nappo, G. Nonlinear filtering for Markov diffusion systems with delayed observations. In *Proceedings of the 42nd Conference on Decision and Control (Maui, Hawaii, December 2003)* (2003), Inst. Electr. Electron. Engrs., New York, pp. 1404 – 1405.

[5] Calzolari, A., Florchinger, P., and Nappo, G. Approximation of nonlinear filters for Markov systems with delayed observations. *SIAM J. Control Optim. 45*, 2 (2006), 599–633 (electronic).

[6] Cannarsa, P., and Vespri, V. Existence and uniqueness of solutions to a class of stochastic partial differential equations. *Stochastic Anal. Appl. 3*, 3 (1985), 315–339.

[7] Clark, J. M. C. The design of robust approximations to the stochastic differential equations of nonlinear filtering. In *Communication systems and random process theory (Proc. 2nd NATO Advanced Study Inst., Darlington, 1977).* Sijthoff & Noordhoff, Alphen aan den Rijn, 1978, pp. 721–734. NATO Advanced Study Inst. Ser., Ser. E: Appl. Sci., No. 25.

[8] Cvitanić, J., Liptser, R., and Rozovskii, B. A filtering approach to tracking volatility from prices observed at random times. *Ann. Appl. Probab. 16*, 3 (2006), 1633–1652.

[9] Davis, M. H. A. A pathwise solution of the equations of nonlinear filtering. *Teor. Veroyatnost. i Primenen. 27*, 1 (1982), 160–167.

[10] Elliott, R. J., and Kohlmann, M. Robust filtering for correlated multidimensional observations. *Math. Z. 178*, 4 (1981), 559–578.

[11] Florchinger, P. Zakai equation of nonlinear filtering with unbounded coefficients. The case of dependent noises. *Systems Control Lett. 21*, 5 (1993), 413–422.

[12] Hopkins, Jr., W. E. *Nonlinear filtering of nondegenerate diffusions with unbounded coefficients.* PhD thesis, University of Maryland at College Park, 1982.

[13] Joannides, M., and Le Gland, F. Nonlinear filtering with perfect discrete time observations. In *Proceedings of the 34th Conference on Decision and Control (New Orleans, LA, December 1995)* (1995), Inst. Electr. Electron. Engrs., New York, pp. 4012–4017.

[14] Kliemann, W., Koch, G., and Marchetti, F. On the unnormalized solution of the filtering problem with counting observations. *IEEE Trans. Inform. Theory 316*, 6 (1990), 1415–1425.

[15] Kunita, H. Stochastic differential equations and stochastic flows of diffeomorphisms. In *École d'été de probabilités de Saint-Flour, XII—1982*, vol. 1097 of *Lecture Notes in Math.* Springer, Berlin, 1984, pp. 143–303.

[16] Liptser, R. S., and Shiryaev, A. N. *Statistics of random processes. I*, expanded ed., vol. 5 of *Applications of Mathematics.* Springer-Verlag, Berlin, 1977. General theory, Translated from the 1974 Russian original by A. B. Aries, Stochastic Modelling and Applied Probability.

[17] Pardoux, E. Filtrage non linéaire et équations aux dérivées partielles stochastiques associées. In *Ecole d'Eté de Probabilités de Saint-Flour XIX—1989*, vol. 1464 of *Lecture Notes in Math.* Springer, Berlin, 1991, pp. 67–163.

[18] Schweizer, M. Risk-minimizing hedging strategies under restricted information. *Math. Finance 4*, 4 (1994), 327–342.

[19] Sussmann, H. J. Rigorous results on the cubic sensor problem. In *Stochastic systems: the mathematics of filtering and identification and applications (Les Arcs, 1980)*, vol. 78 of *NATO Adv. Study Inst. Ser. C: Math. Phys. Sci.* Reidel, Dordrecht, 1981, pp. 637–648.

[20] Zakai, M. On the optimal filtering of diffusion processes. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 11* (1969), 230–243.

[21] Zeng, Y. A partially observed model for micromovement of asset prices with Bayes estimation via filtering. *Math. Finance 13*, 3 (2003), 411–444.