International Journal of Pure and Applied Mathematics

Volume 27 No. 3 2006, 305-318

NONLINEAR FILTERING WITH CORRELATED NOISES IN THE CASE OF HIGH SIGNAL-TO-NOISE RATIO

Jang Schiltz

Applied Mathematics Unit
University of Luxembourg
162A, Avenue de la Faïencerie, Luxembourg, L-1511, LUXEMBOURG
e-mail: jang.schiltz@uni.lu

Abstract: This paper deals with the problem of estimating a state process, the measurements of which are corrupted by an independent but correlated white noise of order ε. We derive a finite dimensional filter, obtained as the solution of a stochastic differential equation which is asymptotically optimal as ε tends to 0.

AMS Subject Classification: 60G35, 60H07, 93E11, 60G12
Key Words: nonlinear filtering, suboptimal filters, stochastic differential equations, Malliavin calculus

1. Introduction

The necessity to estimate a state process which is only partially observed occures in many applied problems and has been widely studied for many years. Adequate mathematical structures have been established and we dispose today of plenty of theoretical results about this subject.

In this article we assume that the process we want to estimate (signal) is a Markovian diffusion process and that it is observed through a function of this process perturbated by a white noise. In fact, we try to find the best possible estimation of the signal at a time t, knowing the observation up to time t. It is however by now well known that, except in some particular cases, the differential equations which allow to solve this problem, for instance the Zakai equation (cf. [14]), are infinite dimensional. Hence, every procedure that brings out suboptimal filters of finite dimension becomes interesting.

Received: February 10, 2006

We assume here that the perturbating white noise is of order ε and that for $\varepsilon = 0$, the signal is exactly observed. This problem has been treated in detail in the case of noncorrelated noises. A.Z. Jazwinski [3] gave the semimartingale decomposition of the filter in a special case in order to construct some second-order filters, B.Z. Bobrovsky and M. Zakai [2] have found some estimates of the conditioned moments of the signal, whereas R. Katzur, B.Z. Bobrovsky and Z. Schuss [6] have used singular perturbation techniques on Zakai's equation to derive formally an asymptotic expansion of the conditional density and they deduced approximate finite-dimensional filters. J. Picard [10] has obtained some suboptimal filters by other methods. He used two basic concepts of the nonlinear filtering theory, namely the Kallianpur-Striebel [4] formula and the decomposition of the optimal filter in semimartingales, as well as some results on the time reversal of diffusion processes. A. Bensoussan [1] has proved a part of this results by more elementary methods and J. Picard [11] has then generalized some of this results to multidimensional signals. Finally, J. Picard [12] has treated the multidimensional case by working under slightly weaker assumptions and by studying also the smoothing problem. He did this without involving Zakai's equation, by simple probabilistic consideration and using the stochastic calculus of variations.

On the other hand, J. Picard [13] studied a filtering system in the case of a high signal-to-noise ratio with correlated noises. He proved that this problem can be viewed as a linear filtering problem with randomly time-varying parameters, and that the filter is auto-adaptive with respect to changes of parameters. The study is based on time discretization; the main tools are an averaging principle and an application of the ordinary differential equation method for the study of stochastic algoritms.

The aim of this paper too is to study this problem for a nonlinear filtering problem with correlated noises, but from a completely different point of view. It consists of three sections organized as follows.

In the second section we introduce the filtering system. We assume that all the processes are $\mathbb{R}^m$-valued. We define a class of suboptimal filters and we show that they do not differ from the optimal filter by more than order $\sqrt{\varepsilon}$. In Section 3, we precise the estimate that we will use and we show that this estimate approches the optimal filter by order ε. The essential tools of the proof are an adequate change of probability and some results of stochastic calculus of variations (Malliavin calculus). The reader will find details about this subject in the books of P. Malliavin [7] and D. Nualart [8] and the references therein.

2. Setting of the Problem and a Preliminary Result

Let $(\Omega, \mathcal{F}, P)$ be a complete probability space and $(\mathcal{F}_t)_{t\in[0,T]}$ a right-continuous increasing family of sub-σ-algebras of $\mathcal{F}$. Let w and v be two independent $\mathcal{F}_t$- Brownian motions with values in $\mathbb{R}^d$ and $\mathbb{R}^m$. If x_t is a semimartingale on $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$, $\circ dx_t$ (respectively dx_t) denotes its Stratonovitch (respectively Itô) differential.

Let us consider the nonlinear filtering problem associated with the system signal/observation pair $(x_t, y_t) \in (\mathbb{R}^m)^2$ solution of the following stochastic differential system:

$$\begin{cases} x_t = x_0 + \int_0^t b(x_s)\, ds + \int_0^t \sigma(x_s)\, dw_s + \int_0^t g(x_s)\, dv_s\,, \\ \qquad y_t = \int_0^t h(x_s)\, ds + \varepsilon\, v_t, \end{cases} \tag{1}$$

verifying the following hypotheses:

(H_1) x_0 is an $\mathbb{R}^m$-valued, $\mathcal{F}_0$-measurable random variable independent of w and v with finite moments of all orders.

(H_2) b and h are $\mathcal{C}^2(\mathbb{R}^m, \mathbb{R}^m)$ functions with bounded first and second derivatives.

(H_3) σ and g are bounded $\mathcal{C}^2(\mathbb{R}^m, \mathbb{R}^m \otimes \mathbb{R}^d)$ respectively $\mathcal{C}^2(\mathbb{R}^m, \mathbb{R}^m \otimes \mathbb{R}^m)$ functions with bounded first derivatives.

(H_4) The function $a = \sigma\sigma^\tau + gg^\tau$ is uniformly elliptic.

(H_5) The functions $a^{-\frac{1}{2}}b$ and $h'b$ are Lipschitzian.

Remark. Condition (H_5) implies that the function σ is inversable.

As usually in nonlinear filtering problems we are then able to define the filter associated with the system (1).

Definition 2.1. For all t in $[0,T]$ denote by π_t the filter associated with the system (1), defined for all functions ψ in $\mathcal{C}_b(\mathbb{R}^m, \mathbb{R})$ by

$$\pi_t\psi = E\big[\psi(x_t)/\mathcal{Y}_t\big], \tag{2}$$

where $\mathcal{Y}_t = \sigma(y_s\,/\,0 \le s \le t)$.

Moreover we consider the following class of suboptimal filters:

$$m_t = m_0 + \int_0^t b(m_s)\, ds + \frac{1}{\varepsilon}\int_0^t h'^{-1}(m_s)\, K_s\, \big(dy_s - h(m_s)\, ds\big), \tag{3}$$

where $m_0 \in \mathbb{R}^m$ is arbitrary and $\{K_t,\, t \leq 0\}$ is a $\mathcal{Y}_t$-progressively measurable bounded process such that for all (t, w) in $[0,T] \times \Omega$, $K_t(w)$ is a uniformly elliptic bounded function.

Remarks. (i) The definition of the sub-optimal filters implies that the signal and the observation are of the same dimension.

(ii) In the following, if f is a vectorial function of the variable $x \in \mathbb{R}^m$, f' will denote the Jacobian matrix of f; if f is scalar, f' will be a line vector; this notation is extended to functions which may also depend upon other parameters.

For this filters, we then have a first estimation of the commited error.

Proposition 2.2. *For each $t_0 > 0$ and $p \geq 1$, we have:*

$$\sup_{t \geq t_0} \|x_t - m_t\|_p = O(\sqrt{\varepsilon}) \tag{4}$$

Proof. Itô's formula implies that

$$h(x_t) = h(x_0) + \int_0^t Lh(x_s)ds + \int_0^t (h'\sigma)(x_s)dw_s + \int_0^t (h'g)(x_s)dv_s,$$

where L is the second order differential operator defined for any function f in $\mathcal{C}^2(\mathbb{R}^m, \mathbb{R}^m)$ by

$$\begin{aligned}(Lf)^l(x) = b^i(x)\frac{\partial f^l}{\partial x_i} \quad &+ \frac{1}{2}\sum_{k=1}^{d} \big(\sigma(x)\big)_k^i \big(\sigma(x)\big)_k^j \frac{\partial^2 f^l}{\partial x_i \partial x_j}(x) \\ &+ \frac{1}{2}\sum_{k=1}^{m} \big(g(x)\big)_k^i \big(g(x)\big)_k^j \frac{\partial^2 f^l}{\partial x_i \partial x_j}(x).\end{aligned}$$

Likewise,

$$h(m_t) = h(m_0) + \int_0^t \tilde{L}_s h(m_s)ds + \frac{1}{\varepsilon}\int_0^t K_s\big(h(x_s) - h(m_s)\big)ds + \int_0^t K_s dv_s,$$

where $\tilde{L}_t$ is the second order differential operator defined for any function f in $\mathcal{C}^2(\mathbb{R}^m, \mathbb{R}^m)$ by

$$(\tilde{L}_t f)^l(x) = b^i(x)\frac{\partial f^l}{\partial x_i} + \frac{1}{2}\sum_{k=1}^{m} \big(h'^{-1}(x)K_t\big)_k^i \big(h'^{-1}(x)K_t\big)_k^j \frac{\partial^2 f^l}{\partial x_i \partial x_j}(x).$$

Hence,

$$\begin{aligned} h(x_t) - h(m_t) \\ &= h(x_0) - h(m_0) + \int_0^t \big(Lh(x_s) - \tilde{L}_s h(m_s)\big)\, ds + \int_0^t (h'\sigma)(x_s)\, dw_s \\ &\quad + \int_0^t (h'g)(x_s)\, dv_s - \frac{1}{\varepsilon}\int_0^t K_s dv_s - \frac{1}{\varepsilon}\int_0^t K_s\big(h(x_s) - h(m_s)\big) ds. \end{aligned}$$

Since K_t is uniformly elliptic, we can prove (4) as in [11].

By applying Itô's formula, we compute $|x_t - m_t|^k$ for even integers k, then, having taken the expectation of the two members of the obtained relation, we use some estimates on ordinary inequations. □

Corollary 2.3. *For any $t_0 > 0$ and $p \geq 1$, we have:*

$$\sup_{t \geq t_0} \|m_t - \pi_t I\|_p = O(\sqrt{\varepsilon}) \tag{5}$$

and

$$\sup_{t \geq t_0} \|x_t - \pi_t I\|_p = O(\sqrt{\varepsilon}), \tag{6}$$

where I denotes the identity matrix on $\mathbb{R}^m$.

3. The Main Result

We now choose an estimate from the previously defined class of filters for which the error will be of order ε.

Let γ be the function defined by $\gamma = \big(h'ah'^{\tau}\big)^{\frac{1}{2}}$. For any t in $[0, T]$, we set

$$K_t = \gamma(m_t). \tag{7}$$

We can then announce the main result of this paper.

Theorem 3.1. *For any $t_0 > 0$ and $p \geq 1$, we have:*

$$\sup_{t \geq t_0} \|m_t - \pi_t I\|_p = O(\varepsilon). \tag{8}$$

Proof. At first we define some stochastic processes which will be useful later on.

Consider the process

$$\overline{w}_t = \int_0^t (\gamma^{-1}h'\sigma)(x_s)\, dw_s + \int_0^t (\gamma^{-1}h'g)(x_s)\, dv_s. \tag{9}$$

Levy's Theorem (e.g. [5]) then implies that $\overline{w}_t$ is an $\mathbb{R}^m$-valued $(\mathcal{F}_t, P)$- Brownian motion and

$$dx_t = b(x_t)dt + h'^{-1}\gamma(x_t)\, d\overline{w}_t. \tag{10}$$

For any t in $[0,T]$, set

$$Z_t = \exp\Big(\frac{1}{\varepsilon}\int_0^t h^\tau(x_s)\, dy_s - \frac{1}{2\varepsilon^2}\int_0^t |h(x_s)|^2\, ds\Big) \tag{11}$$

and

$$\Lambda_t = \exp\Big(-\frac{1}{\varepsilon}\int_0^t \big(h(x_s) - h(m_s)\big)^\tau\, d\overline{w}_s + \frac{1}{2\varepsilon^2}\int_0^t |h(x_s) - h(m_s)|^2\, ds\Big). \tag{12}$$

Novikov's criterion then implies that the processes Z_t^{-1} and Λ_t^{-1} are exponential martingales. So we can apply Girsanov's Theorem and define some reference probability measures which allow us to show the desirated estimates via Kallianpur-Striebel's formula.

So, let us define the probability measures $\mathring{P}$ and $\tilde{P}$ by the Radon-Nicodym derivatives

$$\frac{d\mathring{P}}{dP}\bigg|_{\mathcal{F}_t} = Z_t^{-1}, \tag{13}$$

and

$$\frac{d\tilde{P}}{d\mathring{P}}\bigg|_{\mathcal{F}_t} = \Lambda_t^{-1}. \tag{14}$$

Hence, by Girsanov's Theorem, under $\tilde{P}$, $\tilde{w}_t = \overline{w}_t - \frac{1}{\varepsilon}\int\limits_0^t \big(h(x_s) - h(m_s)\big)\, ds$ and $y_{\frac{t}{\varepsilon}}$ are two independent standard Brownian motions.

Moreover, the stochastic process x_t can be written as the solution of the stochastic differential equation

$$dx_t = \frac{1}{\varepsilon}\,(h'^{-1}\gamma)(x_t)\big(h(x_t) - h(m_t)\big)\, dt + b(x_t)\, dt + (h'^{-1}\gamma)(x_t)\, d\tilde{w}_t. \tag{15}$$

We have on the other hand

$$Z_t\Lambda_t = \exp\Big(-\frac{1}{\varepsilon}\int_0^t \big(h(x_s) - h(m_s)\big)^\tau d\overline{w}_s + \frac{1}{\varepsilon^2}\int_0^t h^\tau(x_s)\big(dy_s - h(m_s)\, ds\big) + \frac{1}{2\varepsilon^2}\int_0^t |h(m_s)|^2 ds\Big).$$

Let F be the function defined for each x, m in $\mathbb{R}^m$, by

$$F(x,m) = \big(h(x) - h(m)\big)^\tau \gamma^{-1}(m)\big(h(x) - h(m)\big). \tag{16}$$

Then, Itô's formula implies that

$$\begin{aligned} F(x_t, m_t) &= F(x_0, m_0) + 2\int_0^t \big(h(x_s) - h(m_s)\big)^\tau \gamma^{-1}(m_s) h'(x_s) dx_s \\ &+ \int_0^t \big(h(x_s) - h(m_s)\big)^\tau \frac{\partial \gamma^{-1}}{\partial m_i}(m_s)\big(h(x_s) - h(m_s)\big)\, dm_s^i \\ &- 2\int_0^t \big(h(x_s) - h(m_s)\big)^\tau (\gamma^{-1} h')(m_s)\, dm_s \\ &+ \int_0^t (A_x F + A_{x,m} F + A_m F)(x_s, m_s)\, ds, \end{aligned}$$

where $A_x, A_{x,m}$ and A_m are the second order differential operators defined for any function f in $\mathcal{C}^2\big((\mathbb{R}^m)^2\big)$ by

$$\begin{aligned} A_x f(x,m) = & \frac{1}{2}\sum_{k=1}^{d} \big(\sigma(x)\big)_k^i \big(\sigma(x)\big)_k^j \frac{\partial^2 f}{\partial x_i \partial x_j}(x,m) \\ & + \frac{1}{2}\sum_{k=1}^{m} \big(g(x)\big)_k^i \big(g(x)\big)_k^j \frac{\partial^2 f}{\partial x_i \partial x_j}(x,m), \end{aligned}$$

$$A_{x,m} f(x,m) = \frac{1}{2}\sum_{k=1}^{d} \big(\gamma(m)\big)_k^i \big(\sigma(x)\big)_k^j \frac{\partial^2 f}{\partial m_i \partial x_j}(x,m)$$

and

$$A_m f(x,m) = \frac{1}{2}\sum_{k=1}^{d} \big(\gamma(x)\big)_k^i \big(\gamma(x)\big)_k^j \frac{\partial^2 f}{\partial m_i \partial m_j}(x,m).$$

Consequently,

$$\begin{aligned} F(x_t, m_t) &= F(x_0, m_0) + 2\int_0^t \big(h(x_s) - h(m_s)\big)^\tau \gamma^{-1}(m_s) h'(x_s) dx_s \\ & -\frac{2}{\varepsilon}\int_0^t \big(h(x_s) - h(m_s)\big)^\tau [dy_s - h(m_s)\, ds] \\ & -2\int_0^t \big(h(x_s) - h(m_s)\big)^\tau (\gamma^{-1} h' b)(m_s)\, ds \\ & +\int_0^t \big(h(x_s) - h(m_s)\big)^\tau \frac{\partial \gamma^{-1}}{\partial m_i}(m_s)\big(h(x_s) - h(m_s)\big)\, dm_s^i \end{aligned}$$

$$+\int_0^t (A_xF + A_{x,m}F + A_mF)(x_s, m_s)\, ds.$$

Hence,

$$\begin{aligned}
&F(x_t, m_t) \\
&\quad = F(x_0, m_0) - 2\int_0^t \big(h(x_s) - h(m_s)\big)^\tau \big(\gamma^{-1}(x_s) - \gamma^{-1}(m_s)\big) h'(x_s) dx_s \\
&\quad + 2\int_0^t \big(h(x_s) - h(m_s)\big)^\tau\, d\overline{w}_s - \frac{2}{\varepsilon}\int_0^t \big(h(x_s) - h(m_s)\big)^\tau \big[dy_s - h(m_s) ds\big] \\
&\qquad + 2\int_0^t \big(h(x_s) - h(m_s)\big)^\tau \big((\gamma^{-1}h'b)(x_s) - (\gamma^{-1}h'b)(m_s)\big)\, ds \\
&\qquad\quad + \int_0^t (A_xF + A_{x,m}F + A_mF)(x_s, m_s)\, ds \\
&\qquad\quad + \int_0^t \big(h(x_s) - h(m_s)\big)^\tau \frac{\partial \gamma^{-1}}{\partial m_i}(m_s)\big(h(x_s) - h(m_s)\big)\, dm_s^i.
\end{aligned}$$

So, we finally have,

$$\begin{aligned}
Z_t\Lambda_t = \exp\Big(&-\frac{1}{2\varepsilon}\big(F(x_t, m_t) - F(x_0, m_0)\big) + \frac{1}{\varepsilon}\int_0^t \chi_1(x_s, m_s)\, ds \\
&+ \frac{1}{\varepsilon}\int_0^t \chi_2^\tau(x_s, m_s)\, dm_s + \frac{1}{\varepsilon}\int_0^t \chi_3^\tau(x_s, m_s)\, dx_s \\
&+ \frac{1}{\varepsilon^2}\int_0^t h^\tau(m_s)\, dy_s - \frac{1}{2\varepsilon^2}\int_0^t |h(m_s)|^2\, ds\Big),
\end{aligned} \tag{17}$$

with

$$\begin{aligned}
\chi_1(x, m) &= \big(h(x) - h(m)\big)^\tau \big((\gamma^{-1}h'b)(x) - (\gamma^{-1}h'b)(m)\big) \\
&\qquad + \frac{1}{2}(A_xF + A_{x,m}F + A_mF)(x, m), \\
\chi_2^i(x, m) &= \frac{1}{2}\big(h(x) - h(m)\big)^\tau \frac{\partial\gamma^{-1}}{\partial m_i}(m)\big(h(x) - h(m)\big), \quad i = 1, ..., m, \\
\chi_3(x, m) &= -(h')^\tau(x)\big(\gamma^{-1}(x) - \gamma^{-1}(m)\big)^\tau \big(h(x) - h(m)\big).
\end{aligned}$$

In the following, if f denotes a function of (x, m), we denote $\frac{\partial f}{\partial x}$ by f'.

By the rules of the Malliavin Calculus (cf. [8] or [7]), if $\overline{D}$ (respectively $\tilde{D}$) denotes the derivation operator in the direction of $\overline{w}$ (respectively $\tilde{w}$), (15) implies that for all $0 \le s \le t$,

$$\tilde{D}_s x_t = \zeta_{st}\, ({h'}^{-1}\gamma)(x_s), \tag{18}$$

where $\{\zeta_{st},\, t \geq s\}$ is the solution of the stochastic differential equation

$$\begin{aligned}
&\zeta_{st} \\
&= 1 + \frac{1}{\varepsilon}\int_s^t \zeta_{sr}\,(h'^{-1}\gamma)'(x_r)\big(h(x_r) - h(m_r)\big)\,dr + \frac{1}{\varepsilon}\int_s^t \zeta_{sr}(h'^{-1}\gamma h')(x_r)\,dr \\
&\qquad + \int_s^t \zeta_{sr}\,b'(x_r)\,dr + \int_s^t \zeta_{sr}\,(h'^{-1}\gamma)'(x_r)\,d\tilde{w}_r + \int_s^t \zeta_{sr}\,g'(x_r)\,dv_r.
\end{aligned}$$

Since

$$\tilde{D}_s y_t = \tilde{D}_s m_t = 0 \tag{19}$$

it follows from (17) and the chain rule that

$$\begin{aligned}
\tilde{D}_s\,\log(Z_t\Lambda_t) \;=\; & -\frac{1}{2\varepsilon}F'(x_t, m_t)\tilde{D}_s x_t + \frac{1}{\varepsilon}\int_s^t \chi_1'(x_r, m_r)\tilde{D}_s x_r\,dr \\
& + \frac{1}{\varepsilon}\int_s^t dm_r^\tau \chi_2'(x_r, m_r)\tilde{D}_s x_r + \frac{1}{\varepsilon}\int_s^t dx_r^\tau \chi_3'(x_r, m_r)\tilde{D}_s x_r \\
& + \frac{1}{\varepsilon}\int_s^t \chi_3(x_r, m_r)\tilde{D}_s x_r\,dr.
\end{aligned}$$

Hence

$$\begin{aligned}
&\frac{1}{2}F'\,(x_t, m_t) = -\varepsilon\tilde{D}_s\,\log(Z_t\Lambda_t)(\tilde{D}_s x_t)^{-1} + \int_s^t \chi_1'(x_r, m_r)\tilde{D}_s x_r(\tilde{D}_s x_t)^{-1}\,dr \\
&\quad + \int_s^t dm_r^\tau \chi_2'(x_r, m_r)\tilde{D}_s x_r(\tilde{D}_s x_t)^{-1} + \int_s^t dx_r^\tau \chi_3'(x_r, m_r)\tilde{D}_s x_r(\tilde{D}_s x_t)^{-1} \\
&\quad + \int_s^t \chi_3(x_r, m_r)\tilde{D}_s x_r(\tilde{D}_s x_t)^{-1}\,dr.
\end{aligned}$$

By integrating that last equality between 0 and t, we get

$$\begin{aligned}
\frac{1}{2}F'(x_t, m_t) \;=\; & -\frac{\varepsilon}{t}\int_0^t \tilde{D}_s\,\log(Z_t\Lambda_t)(\tilde{D}_s x_t)^{-1}\,ds \\
& + \frac{1}{t}\int_0^t\int_s^t \chi_1'(x_r, m_r)\tilde{D}_s x_r(\tilde{D}_s x_t)^{-1}\,dr\,ds \\
& + \frac{1}{t}\int_0^t\int_s^t dm_r^\tau\,\chi_2'(x_r, m_r)\tilde{D}_s x_r(\tilde{D}_s x_t)^{-1}\,ds \\
& + \frac{1}{t}\int_0^t\int_s^t dx_r^\tau\,\chi_3'(x_r, m_r)\tilde{D}_s x_r(\tilde{D}_s x_t)^{-1}\,ds
\end{aligned} \tag{20}$$

$$+\frac{1}{t}\int_0^t\int_s^t \chi_3(x_r,m_r)\tilde{D}_s x_r(\tilde{D}_s x_t)^{-1}\,dr\,ds\,.$$

On the other hand, by the definition of the function F,

$$\frac{1}{2}F'(x_t,m_t)=\big(h(x_t)-h(m_t)\big)^\tau\,\gamma^{-1}(m_t)h'(x_t), \tag{21}$$

so

$$E\Big[\frac{1}{2}F'(x_t,m_t)/\mathcal{Y}_t\Big]=E[\big(h(x_t)-h(m_t)\big)^\tau/\mathcal{Y}_t]\,(\gamma^{-1}h')(m_t)$$
$$+\,E[\big(h(x_t)-h(m_t)\big)^\tau\gamma^{-1}(m_t)\big(h'(x_t)-h'(m_t)\big)/\mathcal{Y}_t].$$

Consequently, Proposition 1.2 implies that

$$E\Big[\frac{1}{2}F'(x_t,m_t)/\mathcal{Y}_t\Big]=E[\big(h(x_t)-h(m_t)\big)^\tau/\mathcal{Y}_t]\,(\gamma^{-1}h')(m_t)+O(\varepsilon).$$

Hence, it follows from the assumptions (H_3) and (H_4) that the theorem will be proved if we show that for any $t_0>0$, $p\ge 1$,

$$\sup_{t\ge t_0}\big\|E\big[\frac{1}{2}F'(x_t,m_t)/\mathcal{Y}_t\big]\big\|_p=O(\varepsilon). \tag{22}$$

So, by equality (24) it suffices to show:

(i) $\sup_{t\ge t_0}\frac{1}{t}\Big\|E\Big[\int_0^t\tilde{D}_s\,\log(Z_t\Lambda_t)(\tilde{D}_s x_t)^{-1}\,ds/\mathcal{Y}_t\Big]\Big\|_p\le c_p$,

(ii) $\sup_{t\ge t_0}\frac{1}{t}\Big\|E\Big[\int_0^t\int_s^t\chi_1'(x_r,m_r)\tilde{D}_s x_r(\tilde{D}_s x_t)^{-1}\,dr\,ds/\mathcal{Y}_t\Big]\Big\|_p=O(\varepsilon)$,

(iii) $\sup_{t\ge t_0}\frac{1}{t}\Big\|E\Big[\int_0^t\int_s^t dm_r^\tau\,\chi_2'(x_r,m_r)\tilde{D}_s x_r(\tilde{D}_s x_t)^{-1}\,ds/\mathcal{Y}_t\Big]\Big\|_p=O(\varepsilon)$,

(iv) $\sup_{t\ge t_0}\frac{1}{t}\Big\|E\Big[\int_0^t\int_s^t dx_r^\tau\,\chi_3'(x_r,m_r)\tilde{D}_s x_r(\tilde{D}_s x_t)^{-1}\,ds/\mathcal{Y}_t\Big]\Big\|_p=O(\varepsilon)$,

(v) $\sup_{t\ge t_0}\frac{1}{t}\Big\|E\Big[\int_0^t\int_s^t\chi_3(x_r,m_r)\tilde{D}_s x_r(\tilde{D}_s x_t)^{-1}\,dr\,ds/\mathcal{Y}_t\Big]\Big\|_p=O(\varepsilon)$.

To prove (i) to (ii) it is necessary to prove the following two lemmas.

Lemma 3.2. *For each $\varepsilon>0$, $p\ge 1$, there exist strictly positive constants $a(p)$ and $\tilde{a}(p)$ such that*

$$\|\zeta_{ts}\|_p\le a(p)\,\exp\Big[-\frac{\tilde{a}(p)}{\varepsilon}(t-s)\Big],\quad 0\le s\le t. \tag{23}$$

Proof. By (18) and the definition of $\tilde{w}_t$, $\{\zeta_{st},\,t\ge s\}$ is the unique solution of the stochastic differential equation

$$\zeta_{st}=1\quad+\quad\frac{1}{\varepsilon}\int_s^t\zeta_{sr}(h'^{-1}\gamma h')(x_r)\,dr+\int_s^t\zeta_{sr}\,b'(x_r)\,dr$$

$$+ \sum_{i=1}^{m} \int_s^t \zeta_{sr} \frac{\partial}{\partial x_i}(h'^{-1}\gamma)(x_r)\, d\overline{w}_r^i.$$

Hence,

$$\begin{aligned} E\left[\sup_{t\in[0,T]} |\zeta_{st}|^p\right] &\le E\Big\{1 + \sup_{t\in[0,T]}\Big(\big|\frac{1}{\varepsilon}\int_s^t \zeta_{sr}(h'^{-1}\gamma h')(x_r)\, dr\big|^p + \big|\int_s^t \zeta_{sr}\, b'(x_r)\, dr\big|^p \\ &\qquad + \big|\sum_{i=1}^{m}\int_s^t \zeta_{sr}\frac{\partial}{\partial x_i}(h'^{-1}\gamma)(x_r)\, d\overline{w}_r^i\big|^p \\ &\le c\Big\{E\Big[\frac{1}{\varepsilon}\int_s^T \big|\zeta_{sr}(h'^{-1}\gamma h')(x_r)\big|^p dr\Big] + E\Big[\int_s^T \big|\zeta_{sr}\, b'(x_r)\big|^p dr\Big] \\ &\qquad + E\Big(\sum_{i=1}^{m}\int_s^T \big|\zeta_{sr}\frac{\partial}{\partial x_i}(h'^{-1}\gamma)(x_r)\big|^2 dr\Big)^{\frac{p}{2}}\Big\}, \end{aligned}$$

by Burkhölder's inequatlity. Since γ is hypoelliptic, the assumptions $(H_2)-(H_4)$ imply that there exist some strictly positive constants c et c' independant of ε such that

$$\|\zeta_{st}\|_p \le \frac{c}{\varepsilon}\int_s^T \|\zeta_{sr}\|_p\, dr + c'\int_s^T \|\zeta_{sr}\|_p\, dr. \tag{24}$$

The lemma then follows from Gronwall's Theorem and the fact that $\zeta_{ts} = \zeta_{st}^{-1}$. □

Lemma 3.3. *For any $p \ge 1$, there exists a constant $c(p)$ such that*

$$\|\overline{D}_s\zeta_{t0}\|_p \le c(p), \tag{25}$$

and

$$\|\tilde{D}_s\zeta_{t0}\|_p \le c(p). \tag{26}$$

Proof. For any t in $[0,T]$,

$$\begin{aligned} \zeta_{0t} = 1 \quad &+ \quad \frac{1}{\varepsilon}\int_0^t \zeta_{0r}(h'^{-1}\gamma h')(x_r)\, dr + \int_0^t \zeta_{0r}\, b'(x_r)\, dr \\ &+ \quad \sum_{i=1}^{m}\int_0^t \zeta_{0r}\frac{\partial}{\partial x_i}(h'^{-1}\gamma)(x_r)\, d\overline{w}_r^i. \end{aligned}$$

Itô's formula then implies

$$\begin{aligned}\zeta_{t0} = 1 \quad &- \quad \frac{1}{\varepsilon}\int_0^t (h'^{-1}\gamma h')(x_r)\,\zeta_{r0} dr - \int_0^t b'(x_r)\,\zeta_{r0}\,dr \\ &- \quad \sum_{i=1}^m \int_0^t \frac{\partial}{\partial x_i}(h'^{-1}\gamma)(x_r)\,\zeta_{r0}\,d\overline{w}_r^i \\ &+ \quad \sum_{i=1}^m \int_0^t \Big(\frac{\partial}{\partial x_i}(h'^{-1}\gamma h')\big(\frac{\partial}{\partial x_i}(h'^{-1}\gamma h')\big)^\tau(x_r)\Big)\zeta_{r0}\,dr.\end{aligned}$$

Consequently,

$$\begin{aligned}\overline{D}_s\zeta_{ts} \quad = \quad & 1 - \frac{1}{\varepsilon}\int_0^t (h'^{-1}\gamma h')'(x_r)\,\zeta_{r0}\,dr - \frac{1}{\varepsilon}\int_0^t (h'^{-1}\gamma h')(x_r)\,\overline{D}_s\zeta_{r0}\,dr \\ & - \int_0^t b''(x_r)\,\zeta_{r0}\,dr - \int_0^t b'(x_r)\,\overline{D}_s\zeta_{r0}\,dr \\ & + \sum_{i=1}^m \int_0^t \Big(\frac{\partial}{\partial x_i}(h'^{-1}\gamma h')\big(\frac{\partial}{\partial x_i}(h'^{-1}\gamma h')\big)^\tau(x_r)\Big)'\zeta_{r0}\,dr \\ & + \sum_{i=1}^m \int_0^t \Big(\frac{\partial}{\partial x_i}(h'^{-1}\gamma h')\big(\frac{\partial}{\partial x_i}(h'^{-1}\gamma h')\big)^\tau(x_r)\Big)\overline{D}_s\zeta_{r0}\,dr.\end{aligned}$$

By using Lemma 1.5, the assumptions $(H_2)-(H_4)$ and Burkhölder's inequality, one then can show like in the proof of the previous lemma that there exist some strictly positive constants c and c' such that

$$\|\overline{D}_s\zeta_{t0}\|_p \leq c' + c\int_0^t \|\overline{D}_s\zeta_{r0}\|_p\,dr \tag{27}$$

Hence relation (25). Relation (26) can be obtained by similar computations. □

Let us now pass to the proof of the expressions (i) to (vii).

Similar computations as the ones on p. 153 from [9] imply

$$\begin{aligned}&\frac{1}{t}E\Big[\int_0^t \tilde{D}_s\,\log(Z_t\Lambda_t)\,(\tilde{D}_s x_t)^{-1}\,ds/\mathcal{Y}_t\Big] \\ &\qquad = \quad \frac{1}{t}E\Big[\int_0^t (\gamma^{-1}h')(x_s)\zeta_{0s}\big(\overline{D}_s\zeta_{t0} - \tilde{D}_s\zeta_{t0}\big)\,ds/\mathcal{Y}_t\Big] \\ &\qquad\quad -\frac{1}{\varepsilon t}E\Big[\int_0^t \big(h(x_s) - h(m_s)\big)\zeta_{ts}(\gamma^{-1}h')\,ds/\mathcal{Y}_t\Big],\end{aligned}$$

and (i) then follows from the Lemma 3.2 and Lemma 3.3.

Moreover, the boundedness of the functions $\|\chi_1'\|_p$ and γ respectively Lemma 1.5 imply that

$$\begin{aligned} \int_0^t \int_s^t \chi_1'(x_r, m_r)\tilde{D}_s x_r(\tilde{D}_s x_t)^{-1}\, dr\, ds \leq c(p) \int_0^t \int_s^t \zeta_{sr}\zeta_{ts}\, dr\, ds \qquad (28) \\ \leq \quad c(p) \int_0^t \int_s^t \exp\big[\frac{\tilde{a}(p)}{\varepsilon}(r-s)\big] \exp\big[-\frac{\tilde{a}(p)}{\varepsilon}(t-s)\big]\, dr\, ds \\ \leq \quad \varepsilon\, c(p) \int_0^t \big(\exp\big[\frac{\tilde{a}(p)}{\varepsilon}(t-s)\big] - 1\big) \exp\big[\frac{-\tilde{a}(p)}{\varepsilon}(t-s)\big]\, ds \\ = \quad \varepsilon\, t\, c(p). \end{aligned}$$

Hence we have (ii).

Similar calculations imply the expressions (iii)-(v) (let us notice that in expression (iii) there appears an integral with respect to m_t, which is of order $O(\frac{1}{\sqrt{\varepsilon}})$, but $\|\chi_2'\|_p$ is of order $O(\sqrt{\varepsilon})$, so there is no problem to conclude).

This completes the proof of Theorem 3.1. □

References

[1] A. Bensoussan, On some approximation techniques in nonlinear filtering, *Stochastic Differential Systems, Stochastic Control and Applications*, Minneapolis (1986); *IMA*, **10**, Springer, New York (1986).

[2] B.Z. Bobrovsky, M. Zakai, Asomptotic a priori estimates for the error in the nonlinear filtering problem, *IEEE Trans. Inform. Th.*, **28** (1982), 371-376.

[3] A.H. Jazwinski, *Stochastic Filtering Theory*, Applications of Mathematics, Springer, Berlin, Heidelberg, New-York (1980).

[4] G. Kallianpur, *Stochastic Filtering Theory*, Springer, Berlin, Heidelberg, New-York (1980).

[5] I. Karatzas, S.E. Shreve, *Brownian Motion and Stochastic Calculus*, Second Edition, Graduate Texts in Mathematics, Volume **113**, Springer-Verlag, Berlin, New York, Heidelberg (1999).

[6] R. Katzur, B.Z. Bobrovsky, Z. Schuss, Asymptotic analysis of the optimal filtering problem for one-dimensional diffusions measured in a low noise channel, *Siam J. Appl. Math.*, **44** (1984), 591-604 and 1176-1191.

[7] P. Malliavin, *Stochastic Analysis*, Grundlehren der mathematischen Wissenschaften, Volume **313**, Springer, Berlin, New York, Heidelberg (1997).

[8] D. Nualart, *The Malliavin Calculus and Related Topics*, Applications of Mathematics, Springer, New York (1995).

[9] E. Pardoux, Filtrage non linéaire et équations aux dérivées partielles stochastique associées, *Lect. Notes Math.*, **1464**, Springer, Berlin Heidelberg New York (1989), 69-163.

[10] J. Picard, Non linear filtering of one-dimensional diffusions in the case of a high signal to noise ratio, *Siam J. Appl. Math.*, **46** (1986), 1098-1125.

[11] J. Picard, Filtrage de diffusions vectorielles faiblement bruitées, *Lect. Notes Control and Info. Scie.*, **83**, Springer, New York (1986).

[12] J. Picard, Nonlinear filtering and smoothing with high signal-to- noise ratio, In: *Stochastic Processes in Physics and Engineering* (Ed. D. Reidel) (1986).

[13] J.Picard, Estimation of the quadratic variation of nearly observed semimartingales with application to filtering, *Siam J. Control and Optimization*, **31** (1993), 494-517.

[14] M. Zakai, On the optimal filtering of diffusion processes, *Z.F.W.*, **11** (1969), 230-243.