Proceedings of the 28th Conference
on Decision and Control
Tampa, Florida • December 1989

TIME DISCRETIZATION OF NONLINEAR FILTERING EQUATIONS*

François LE GLAND

INRIA Sophia–Antipolis
Route des Lucioles
F–06565 VALBONNE Cédex

Abstract *Some computable approximate expressions are provided for the conditional law of diffusion processes observed in continuous time. The numerical schemes are derived through an approximation of the original filtering problem. Given a partition of the time interval, this procedure consists in sampling the available observation sample path and approximating the a priori law of the diffusion process. This results in approximation schemes for the Zakai equation, for which rate of convergence are provided.*

1 Introduction

The purpose of this paper is to give computable and accurate approximate expressions for the conditional law of a diffusion process observed in continuous time. Since this conditional law depends on both

- the a priori information, provided by the semi–group $\{P_t\,,\,t\geq 0\}$ or equivalently the infinitesimal generator L,
- the available observation sample–path $\{Y_t\,,\,t\geq 0\}$,

the approximation problem under consideration should reduce in some sense to

- approximate the a priori law of the original diffusion process, e.g. by the more simple a priori law of some other process,
- extract the most useful information from the available continuous time measurements $\{Y_t\,,\,t\geq 0\}$.

The general situation of filtering a signal process from noisy continuous measurements will be considered. At each step of the approximation procedure, the general formulas will be applied to the particular case of diffusion processes, in order to check whether or not some computable expression has been obtained. Note that only time discretization is considered here: the discretization with respect to the space variable, e.g. the approximation of the partial differential operator L by finite differences is not taken into consideration.

*Research partially supported by Systems Research Center, University of Maryland, and by USACCE under Contract DAJA45–87–M–0296

2 The filtering problem

On a measurable space $(\Omega,\mathcal{F})$ are given a probability measure P, and a pair of stochastic processes $\{X_t\,,\,t\geq 0\}$ and $\{Y_t\,,\,t\geq 0\}$ taking values in $\mathbf{R}^m$ and $\mathbf{R}^d$ respectively, such that under P

$$dY_t = h(X_t)\,dt + dV_t\;, \tag{1}$$

where $\{V_t\,,\,t\geq 0\}$ is a standard Wiener process, independent of $\{X_t\,,\,t\geq 0\}$.

Note that the a priori law of the signal $\{X_t\,,\,t\geq 0\}$ is not specified at this point. The observation function satisfy the following hypothesis

> *h is a measurable and bounded function from $\mathbf{R}^m$ to $\mathbf{R}^d$.*

Remark 2.1 As usual, (1) is the mathematical way of expressing that some measurement

$$z_t = h(X_t) + \eta_t\;, \tag{2}$$

is available at time t, where $\{\eta_t\,,\,t\geq 0\}$ is a Gaussian white–noise process, independent of $\{X_t\,,\,t\geq 0\}$.

Introduce the σ–algebras

$$\mathcal{F}_t \triangleq \sigma(X_s\,,\,0\leq s\leq t)\;,$$

$$\mathcal{Y}_t^s \triangleq \sigma(Y_\tau - Y_s\,,\,s\leq\tau\leq t)\;,\qquad \mathcal{Y}_t \triangleq \mathcal{Y}_t^0\;.$$

The problem is to estimate X_t from $\mathcal{Y}_t$, i.e. to compute the conditional (a posteriori) law of X_t given $\mathcal{Y}_t$, defined by $\mathbf{E}(\phi(X_t)\mid\mathcal{Y}_t)$. Introducing

$$Z_t^s \triangleq \exp\left\{\int_s^t h^*(X_\tau)\,dY_\tau - \tfrac{1}{2}\int_s^t |h(X_\tau)|^2\,d\tau\right\}\;,$$

and $Z_t \triangleq Z_t^0$, it is standard that, for all $T>0$ the original probability measure P is equivalent on $[0,T]$ to the *reference probability* measure $P^\dagger$ with Radon–Nikodym derivative Z_T, so that under $P^\dagger$ $\{Y_t\,,\,t\geq 0\}$ is a standard Wiener process, independent of $\{X_t\,,\,t\geq 0\}$.

By the Bayes formula

$$\mathbf{E}(\phi(X_t)\mid\mathcal{Y}_t) = \frac{\mathbf{E}^\dagger(\phi(X_t)Z_t\mid\mathcal{Y}_t)}{\mathbf{E}^\dagger(Z_t\mid\mathcal{Y}_t)}\;,$$

so that it is enough to compute $\{p_t\,,\,t\geq 0\}$ defined by

$$(p_t,\phi) \triangleq \mathbf{E}^\dagger(\phi(X_t)Z_t \mid \mathcal{Y}_t)\ .$$

In the particular case where the signal $\{X_t\,,\,t\geq 0\}$ is a diffusion process with infinitesimal generator L, $\{p_t\,,\,t\geq 0\}$ is the unique solution of the Zakai equation

$$dp_t = L^* p_t\,dt + h^* p_t\,dY_t\ . \tag{3}$$

It is readily seen on this equation that p_t depends on the a priori law of $\{X_t\,,\,t\geq 0\}$ represented by the partial differential operator L, and on the observation sample–path $\{Y_t\,,\,t\geq 0\}$. However, equation (3) is not computable and should be approximated. The approach presented here, is to rather approximate the original filtering problem by a simpler problem, and to consider the resulting equation for the conditional law in this new filtering problem as an approximation to equation (3). In Section 5, the rate of convergence for such approximations will be provided, by direct numerical analysis of equation (3).

The presentation adopted follows Korezlioglu–Mazziotto [2]. There is indeed three successive steps in the global approximation procedure. In the first step, sampling and data compression of the observation sample–path $\{Y_t\,,\,t\geq 0\}$ is performed. Then, the signal $\{X_t\,,\,t\geq 0\}$ is approximated by some piecewise constant process $\{\overline{X}_t\,,\,t\geq 0\}$. In the last step, the a priori law of the process $\{\overline{X}_t\,,\,t\geq 0\}$ is approximated. Only the first two steps will be considered here.

3 Sampling of the observation sample–path

Throughout the paper, an infinite partition

$$0 = t_0 < t_1 < \cdots < t_n < \cdots$$

of $[0,+\infty)$ is introduced, to be denoted by π, with time increments $\delta_i \triangleq t_{i+1} - t_i$.

Sampling and data compression is the pre–processing procedure by which the new information contained in the continuous measurements received in the time interval $t_i \leq t \leq t_{i+1}$ and represented by $\mathcal{Y}^{t_i}_{t_{i+1}}$, is summarized into a finite number of random variables. This is formalized in the following

Definition 3.1 *An admissible sampling procedure relative to the partition π is a family $\{\overline{\mathcal{Y}}^{t_i}_{t_{i+1}}\,,\,i\geq 0\}$ of σ–algebras which satisfy, for all $i\geq 0$*

(i) $\overline{\mathcal{Y}}^{t_i}_{t_{i+1}}$ is generated by a finite number of random variables,

(ii) $\overline{\mathcal{Y}}^{t_i}_{t_{i+1}} \subset \mathcal{Y}^{t_i}_{t_{i+1}}$.

In addition, the following notations are used

$$\overline{\mathcal{Y}}^{t_l}_{t_m} \triangleq \bigvee_{i=l}^{m-1} \overline{\mathcal{Y}}^{t_i}_{t_{i+1}}\ , \qquad \overline{\mathcal{Y}}_{t_m} \triangleq \overline{\mathcal{Y}}^{0}_{t_m}\ .$$

Here are two examples of admissible sampling procedures, to be considered throughout the paper.

Example 1. Define

$$\xi_i \triangleq Y_{t_{i+1}} - Y_{t_i} = \int_{t_i}^{t_{i+1}} z_s\,ds\ , \tag{4}$$

which is the mean value of the actual measurements (2) on the time interval $t_i \leq s \leq t_{i+1}$. In this example, $\overline{\mathcal{Y}}^{t_i}_{t_{i+1}}$ is generated by the random variable ξ_i. Note that, under the reference probability measure $P^\dagger$, $\{\xi_i\,,\,i\geq 0\}$ are mutually independent d–dimensional Gaussian random variables with zero mean and covariance matrix $\delta_i I$.

Example 2. Define

$$\xi^\sharp_i \triangleq \frac{1}{\delta_i}\int_{t_i}^{t_{i+1}} (s-t_i)\,dY_s = \frac{1}{\delta_i}\int_{t_i}^{t_{i+1}}\int_s^{t_{i+1}} y_\tau\,d\tau\,ds\ ,$$

$$\xi^\flat_i \triangleq \frac{1}{\delta_i}\int_{t_i}^{t_{i+1}} (t_{i+1}-s)\,dY_s = \frac{1}{\delta_i}\int_{t_i}^{t_{i+1}}\int_{t_i}^{s} y_\tau\,d\tau\,ds\ ,$$

which are two other different ways of computing some mean value of the actual measurements (2) on the time interval $t_i \leq s \leq t_{i+1}$. In this example, $\overline{\mathcal{Y}}^{t_i}_{t_{i+1}}$ is generated by the random variables $\xi^\sharp_i$ and $\xi^\flat_i$. Note that $\xi^\sharp_i + \xi^\flat_i = \xi_i$ and that, under the reference probability measure $P^\dagger$, $\{(\xi^\sharp_i,\xi^\flat_i)\,,\,i\geq 0\}$ are mutually independent $2d$–dimensional Gaussian random variables with zero mean and covariance matrix $\delta_i\Sigma$, where

$$\Sigma = \begin{pmatrix} \frac{1}{3}I & \frac{1}{6}I \\ \frac{1}{6}I & \frac{1}{3}I \end{pmatrix}\ .$$

In particular, the characteristic function of $(\xi^\sharp_i,\xi^\flat_i)$ satisfies

$$\begin{aligned}\chi(a,b) &\triangleq \mathbf{E}^\dagger(\exp\left\{a^*\xi^\sharp_i + b^*\xi^\flat_i\right\}) \\ &= \exp\left\{\tfrac{1}{6}(|a|^2 + a^*b + |b|^2)\delta_i\right\}\ .\end{aligned} \tag{5}$$

The problem is now to estimate X_{t_i} from $\overline{\mathcal{Y}}_{t_i}$, i.e. to compute the conditional law of X_{t_i} given $\overline{\mathcal{Y}}_{t_i}$. By the Bayes formula

$$\mathbf{E}(\phi(X_{t_i}) \mid \overline{\mathcal{Y}}_{t_i}) = \frac{\mathbf{E}^\dagger(\phi(X_{t_i})Z_{t_i} \mid \overline{\mathcal{Y}}_{t_i})}{\mathbf{E}^\dagger(Z_{t_i} \mid \overline{\mathcal{Y}}_{t_i})}\ ,$$

so that it is enough to compute $\{p_i\,,\,i\geq 0\}$ defined by

$$(p_i,\phi) \triangleq \mathbf{E}^\dagger(\phi(X_{t_i})Z_{t_i} \mid \overline{\mathcal{Y}}_{t_i})\ .$$

The first step is provided by the following

Proposition 3.2 *Introduce*

$$\Xi^{t_i}_{t_{i+1}} \triangleq \mathbf{E}^\dagger(Z^{t_i}_{t_{i+1}} \mid \mathcal{F}_{t_{i+1}} \vee \overline{\mathcal{Y}}^{t_i}_{t_{i+1}})\ ,$$

$$U_{i+1}\phi \triangleq \mathbf{E}^\dagger(\phi(X_{t_{i+1}})\Xi^{t_i}_{t_{i+1}} \mid \mathcal{F}_{t_i} \vee \overline{\mathcal{Y}}^{t_i}_{t_{i+1}})\ .$$

Then

$$(p_{i+1},\phi) = \mathbf{E}^\dagger([U_{i+1}\phi]Z_{t_i} \mid \overline{\mathcal{Y}}_{t_{i+1}})\ . \tag{6}$$

PROOF.

$$
\begin{aligned}
&(p_{i+1}, \phi) \\
&= \mathbf{E}^\dagger(\phi(X_{t_{i+1}}) Z_{t_i} Z^{t_i}_{t_{i+1}} \mid \overline{\mathcal{Y}}_{t_{i+1}}) \\
&= \mathbf{E}^\dagger(\phi(X_{t_{i+1}}) Z_{t_i} \\
&\qquad \cdot \mathbf{E}^\dagger(Z^{t_i}_{t_{i+1}} \mid \mathcal{F}_{t_{i+1}} \vee \mathcal{Y}_{t_i} \vee \overline{\mathcal{Y}}^{t_i}_{t_{i+1}}) \mid \overline{\mathcal{Y}}_{t_{i+1}}) \\
&= \mathbf{E}^\dagger(\phi(X_{t_{i+1}}) Z_{t_i} \Xi^{t_i}_{t_{i+1}} \mid \overline{\mathcal{Y}}_{t_{i+1}}) \\
&= \mathbf{E}^\dagger(Z_{t_i}\, \mathbf{E}^\dagger(\phi(X_{t_{i+1}}) \Xi^{t_i}_{t_{i+1}} \mid \mathcal{F}_{t_i} \vee \mathcal{Y}_{t_i} \vee \overline{\mathcal{Y}}^{t_i}_{t_{i+1}}) \mid \overline{\mathcal{Y}}_{t_{i+1}}) \\
&= \mathbf{E}^\dagger([U_{i+1}\phi] Z_{t_i} \mid \overline{\mathcal{Y}}_{t_{i+1}}) \ . \qquad \square
\end{aligned}
$$

Going back to the examples introduced above, the expression for $\Xi^{t_i}_{t_{i+1}}$ will be derived, and it will be checked whether or not the additional hypothesis that the signal $\{X_t\,,\,t \geq 0\}$ is a diffusion process can lead to computable expressions.

Example 1 (Continued). For the sampling procedure defined by ξ_i, it is proved in [2] that

$$\Xi^{t_i}_{t_{i+1}} = \exp\left\{h_i^* \xi_i - \tfrac{1}{2}|h_i|^2 \delta_i\right\} \ , \tag{7}$$

where

$$h_i \triangleq \frac{1}{\delta_i} \int_{t_i}^{t_{i+1}} h(X_s)\, ds \ .$$

However, replacing this expression into (6) does not provide a computable expression, even if the additional hypothesis that the signal $\{X_t\,,\,t \geq 0\}$ is a diffusion process is introduced.

Example 2 (Continued). For the sampling procedure defined by $(\xi_i^\sharp, \xi_i^\flat)$, it can be proved that

$$
\begin{aligned}
\Xi^{t_i}_{t_{i+1}} &= \exp\left\{[h_i^\sharp]^* \xi_i^\sharp + [h_i^\flat]^* \xi_i^\flat \right. \\
&\qquad \left. -\tfrac{1}{6}(|h_i^\sharp|^2 + [h_i^\sharp]^* h_i^\flat + |h_i^\flat|^2)\delta_i\right\} \\
&= \exp\left\{[h_i^\sharp]^* \xi_i^\sharp - \tfrac{1}{4}|h_i^\sharp|^2 \delta_i\right\} \\
&\quad \cdot \exp\left\{[h_i^\flat]^* \xi_i^\flat - \tfrac{1}{4}|h_i^\flat|^2 \delta_i\right\} \cdot \exp\left\{\tfrac{1}{12}|h_i^\sharp - h_i^\flat|^2 \delta_i\right\}
\end{aligned} \tag{8}
$$

where

$$h_i^\sharp \triangleq \frac{1}{\delta_i} \int_{t_i}^{t_{i+1}} \overline{w}\Big(\frac{s - t_i}{t_{i+1} - t_i}\Big)\, h(X_s)\, ds \ ,$$

$$h_i^\flat \triangleq \frac{1}{\delta_i} \int_{t_i}^{t_{i+1}} \overline{w}\Big(\frac{t_{i+1} - s}{t_{i+1} - t_i}\Big)\, h(X_s)\, ds \ ,$$

and the weight function $\overline{w}$ is defined for all $0 \leq \theta \leq 1$ by $\overline{w}(\theta) \triangleq 6\theta - 2$.

Here again, replacing this expression into (6) does not provide a computable expression, even if the additional hypothesis that the signal $\{X_t\,,\,t \geq 0\}$ is a diffusion process is introduced.

4 Piecewise constant approximation of the signal process

The purpose of this section is to investigate the effect of replacing the signal process $\{X_t\,,\,t \geq 0\}$ by a piecewise constant process $\{\overline{X}_t\,,\,t \geq 0\}$ whose values on "pieces" are related in some way to the values taken by the original signal process at some particular instants. This is formalized in the following

Definition 4.1 *A process $\{\overline{X}_t\,,\,t \geq 0\}$ is* subordinate *to the process $\{X_t\,,\,t \geq 0\}$ relatively to the partition π if, for all $i \geq 0$*

$\overline{X}_t$ is $\mathcal{F}_{t_{i+1}}$-measurable, $t_i \leq t \leq t_{i+1}$.

The following example provide a particular class of subordinate process, to be used throughout the paper.

Example. For all $i \geq 0$ are given

- a partition $\{A_i^j\,,\,1 \leq j \leq k(i)\}$ of the time interval $[t_i, t_{i+1})$,
- an increasing sequence

$$t_i \leq \tau_i^1 < \cdots < \tau_i^j < \cdots < \tau_i^{k(i)} \leq t_{i+1} \ .$$

Then the piecewise constant process $\{\overline{X}_t\,,\,t \geq 0\}$ defined by

$$\overline{X}_t = X_{\tau_i^j} \ , \qquad \text{if } t \in A_i^j \ ,$$

is subordinate to $\{X_t\,,\,t \geq 0\}$ relatively to the partition π. There is a similar class of subordinate processes, where the time interval to be partitioned is rather $(t_i, t_{i+1}]$.

The problem is to chose $\{\overline{X}_t\,,\,t \geq 0\}$ in such a way that the conditional law of $\overline{X}_{t_i}$ given $\overline{\mathcal{Y}}_{t_i}$ is more simple to handle than the conditional law of X_{t_i} given $\overline{\mathcal{Y}}_{t_i}$, and is even computable in the particular case where the signal $\{X_t\,,\,t \geq 0\}$ is a diffusion process.

Introduce

$$\overline{Z}_t^s \triangleq \exp\left\{\int_s^t h^*(\overline{X}_\tau)\, dY_\tau - \tfrac{1}{2}\int_s^t |h(\overline{X}_\tau)|^2\, d\tau\right\} \tag{9}$$

and $\overline{Z}_t \triangleq \overline{Z}_t^0$. Under the reference probability measure $P^\dagger$, the processes $\{\overline{X}_t\,,\,t \geq 0\}$ and $\{Y_t\,,\,t \geq 0\}$ are independent, so that the stochastic integral in (9) is well defined, although $\{\overline{X}_t\,,\,t \geq 0\}$ is not necessarily adapted. Therefore, it is possible for all $T > 0$ to define a new probability measure $\overline{P}$ equivalent on $[0,T]$ to $P^\dagger$ with Radon–Nikodym derivative $\overline{Z}_T$, so that under $\overline{P}$

$$dY_t = h(\overline{X}_t)\, dt + d\overline{V}_t \ ,$$

where $\{\overline{V}_t\,,\,t \geq 0\}$ is a standard Wiener process, independent of $\{\overline{X}_t\,,\,t \geq 0\}$.

The problem is now to estimate $\overline{X}_{t_i}$ from $\overline{\mathcal{Y}}_{t_i}$, i.e. to compute the conditional law of $\overline{X}_{t_i}$ given $\overline{\mathcal{Y}}_{t_i}$. By the Bayes formula

$$\overline{\mathbf{E}}(\phi(\overline{X}_{t_i}) \mid \overline{\mathcal{Y}}_{t_i}) = \frac{\mathbf{E}^\dagger(\phi(\overline{X}_{t_i})\overline{Z}_{t_i} \mid \overline{\mathcal{Y}}_{t_i})}{\mathbf{E}^\dagger(\overline{Z}_{t_i} \mid \overline{\mathcal{Y}}_{t_i})} \ ,$$

so that it is enough to compute $\{\overline{p}_i\,,\,i\geq 0\}$ defined by

$$(\overline{p}_i,\phi)\triangleq \mathbf{E}^{\dagger}(\phi(\overline{X}_{t_i})\overline{Z}_{t_i}\mid \overline{\mathcal{Y}}_{t_i})\ .$$

It follows from the proof of Proposition 3.2 that

$$(\overline{p}_{i+1},\phi)=\mathbf{E}^{\dagger}([\overline{U}_{i+1}\phi]\overline{Z}_{t_i}\mid \overline{\mathcal{Y}}_{t_{i+1}})\ , \tag{10}$$

where, for all $i\geq 0$

$$\overline{\Xi}^{t_i}_{t_{i+1}}\triangleq \mathbf{E}^{\dagger}(\overline{Z}^{t_i}_{t_{i+1}}\mid \mathcal{F}_{t_{i+1}}\vee\overline{\mathcal{Y}}^{t_i}_{t_{i+1}})\ ,$$

$$\overline{U}_{i+1}\phi\triangleq \mathbf{E}^{\dagger}(\phi(\overline{X}_{t_{i+1}})\overline{\Xi}^{t_i}_{t_{i+1}}\mid \mathcal{F}_{t_i}\vee\overline{\mathcal{Y}}^{t_i}_{t_{i+1}})\ .$$

Going back to the examples introduced above, some particular piecewise constant subordinate processes will be considered, the corresponding expression for $\overline{\Xi}^{t_i}_{t_{i+1}}$ will be derived, and it will be checked whether or not the additional hypothesis that the signal $\{X_t\,,\,t\geq 0\}$ is a diffusion process can lead to computable expressions.

Example 1 (Continued). For the sampling procedure defined by ξ_i, $\overline{\Xi}^{t_i}_{t_{i+1}}$ has the same form than (7) where now

$$h_i\triangleq \frac{1}{\delta_i}\int_{t_i}^{t_{i+1}} h(\overline{X}_s)\,ds\ .$$

Two different piecewise constant subordinate processes will be considered.

$\boxed{\text{1a}}$ Define

$$\overline{X}_t=X_{t_i}\ ,\qquad \text{if } t_i\leq t<t_{i+1}\ .$$

Then $h_i=h(X_{t_i})$ and $\overline{\Xi}^{t_i}_{t_{i+1}}=\Psi_i(X_{t_i})$, where for all $x\in\mathbf{R}^m$

$$\Psi_i(x)\triangleq \exp\left\{h^*(x)\xi_i-\tfrac{1}{2}|h(x)|^2\delta_i\right\}\ . \tag{11}$$

Therefore

$$\overline{U}_{i+1}\phi=\Psi_i(X_{t_i})\ \mathbf{E}^{\dagger}(\phi(X_{t_{i+1}})\mid\mathcal{F}_{t_i})\ .$$

Under the additional hypothesis that the signal $\{X_t\,,\,t\geq 0\}$ is a diffusion process with semi–group $\{P_t\,,\,t\geq 0\}$

$$\overline{U}_{i+1}\phi=\Psi_i(X_{t_i})\,[P_{\delta_i}\phi](X_{t_i})\ ,$$

and

$$\begin{aligned}(\overline{p}_{i+1},\phi)\ &=\ \mathbf{E}^{\dagger}(\Psi_i(X_{t_i})\,[P_{\delta_i}\phi](X_{t_i})\overline{Z}_{t_i}\mid\overline{\mathcal{Y}}_{t_{i+1}})\\ &=\ (\overline{p}_i,\Psi_i[P_{\delta_i}\phi])\ ,\end{aligned}$$

so that $\{\overline{p}_i\,,\,i\geq 0\}$ satisfies the following recurrence

$$\overline{p}_{i+1}=P^*_{\delta_i}[\Psi_i\overline{p}_i]\ , \tag{12}$$

which is a computable expression, and can be considered as a time discretization scheme for the Zakai equation (3). The rate of convergence of this approximation will be considered in Section 5.

$\boxed{\text{1b}}$ Define

$$\overline{X}_t=X_{t_{i+1}}\ ,\qquad \text{if } t_i<t\leq t_{i+1}\ .$$

Then $h_i=h(X_{t_{i+1}})$ and $\overline{\Xi}^{t_i}_{t_{i+1}}=\Psi_i(X_{t_{i+1}})$. Therefore

$$\overline{U}_{i+1}\phi=\mathbf{E}^{\dagger}(\phi(X_{t_{i+1}})\ \Psi_i(X_{t_{i+1}})\mid\mathcal{F}_{t_i}\vee\overline{\mathcal{Y}}^{t_i}_{t_{i+1}})\ .$$

Under the additional hypothesis that the signal $\{X_t\,,\,t\geq 0\}$ is a diffusion process with semi–group $\{P_t\,,\,t\geq 0\}$

$$\overline{U}_{i+1}\phi=P_{\delta_i}[\Psi_i\phi](X_{t_i})\ ,$$

and

$$\begin{aligned}(\overline{p}_{i+1},\phi)\ &=\ \mathbf{E}^{\dagger}(P_{\delta_i}[\Psi_i\phi](X_{t_i})\overline{Z}_{t_i}\mid\overline{\mathcal{Y}}_{t_{i+1}})\\ &=\ (\overline{p}_i,P_{\delta_i}[\Psi_i\phi])\ .\end{aligned}$$

This results in the following recurrence

$$\overline{p}_{i+1}=\Psi_i[P^*_{\delta_i}\overline{p}_i]\ , \tag{13}$$

which gives another time discretization scheme for the Zakai equation (3).

Remark 4.2 In the numerical scheme (12) (resp. (13)) the transition from $\overline{p}_i$ to $\overline{p}_{i+1}$ reflects the following situation: A new measurement ξ_i is available, which is a compression of the information provided by $\{z_t\,,\,t_i\leq t\leq t_{i+1}\}$ according to (4). This measurement is interpreted as a noisy nonlinear observation of X_{t_i} (resp. $X_{t_{i+1}}$), and is combined with the current estimate $\overline{p}_i$ of X_{t_i} to produce an estimate $\overline{p}_{i+1}$ of $X_{t_{i+1}}$.

Example 2 (Continued). For the sampling procedure defined by $(\xi^\sharp_i,\xi^\flat_i)$, $\overline{\Xi}^{t_i}_{t_{i+1}}$ has the same form than (8) where now

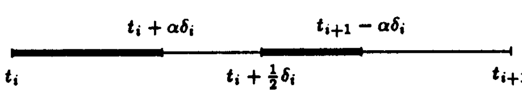
$$h^\sharp_i\triangleq\frac{1}{\delta_i}\int_{t_i}^{t_{i+1}}\varpi\Big(\frac{s-t_i}{t_{i+1}-t_i}\Big)\,h(\overline{X}_s)\,ds\ ,$$

$$h^\flat_i\triangleq\frac{1}{\delta_i}\int_{t_i}^{t_{i+1}}\varpi\Big(\frac{t_{i+1}-s}{t_{i+1}-t_i}\Big)\,h(\overline{X}_s)\,ds\ .$$

The following family, parametrized by $0<\alpha<\frac{1}{2}$, of piecewise constant subordinate processes will be considered

$$\overline{X}_t\triangleq\begin{cases}X_{t_i} & \text{if } t\in A^\alpha_i\\[1ex] X_{t_{i+1}} & \text{if } t\in[t_i,t_{i+1})\setminus A^\alpha_i\end{cases}$$

where for all $i\geq 0$, A^α_i denotes the following subset of the time interval $[t_i,t_{i+1})$



It is then possible to find a particular value α_0 for which

$$h^\sharp_i=h(X_{t_{i+1}})\ ,\qquad h^\flat_i=h(X_{t_i})\ .$$

Therefore (8) becomes

$$\begin{aligned}\overline{\Xi}^{t_i}_{t_{i+1}}\ &=\ \Psi^\sharp_i(X_{t_{i+1}})\ \Psi^\flat_i(X_{t_i})\\ &\qquad\cdot\exp\left\{\tfrac{1}{12}|h(X_{t_{i+1}})-h(X_{t_i})|^2\delta_i\right\}\end{aligned}$$

where for all $x \in \mathbf{R}^m$

$$\Psi_i^\sharp(x) \triangleq \exp\left\{h^*(x)\xi_i^\sharp - \tfrac{1}{4}|h(x)|^2\delta_i\right\} ,$$

$$\Psi_i^\flat(x) \triangleq \exp\left\{h^*(x)\xi_i^\flat - \tfrac{1}{4}|h(x)|^2\delta_i\right\} ,$$

and

$$\overline{U}_{i+1}\phi = \Psi_i^\flat(X_{t_i})\,\mathbf{E}^\dagger(\phi(X_{t_{i+1}})\,\Psi_i^\sharp(X_{t_{i+1}}) \cdot \exp\left\{\tfrac{1}{12}|h(X_{t_{i+1}}) - h(X_{t_i})|^2\delta_i\right\} \mid \mathcal{F}_{t_i} \vee \overline{\mathcal{Y}}_{t_{i+1}}^{t_i}) .$$

Under the additional hypothesis that the signal $\{X_t\,,\,t \geq 0\}$ is a diffusion process with semi-group $\{P_t\,,\,t \geq 0\}$

$$\overline{U}_{i+1}\phi = \Psi_i^\flat(X_{t_i})\,Q_{\delta_i}[\Psi_i^\sharp\phi](X_{t_i}) ,$$

and

$$\begin{aligned}(\overline{p}_{i+1},\phi) &= \mathbf{E}^\dagger(\Psi_i^\flat(X_{t_i})\,Q_{\delta_i}[\Psi_i^\sharp\phi](X_{t_i})\overline{Z}_{t_i} \mid \overline{\mathcal{Y}}_{t_{i+1}}) \\ &= (\overline{p}_i, \Psi_i^\flat Q_{\delta_i}[\Psi_i^\sharp\phi]) ,\end{aligned}$$

so that $\{\overline{p}_i\,,\,i \geq 0\}$ satisfies the following recurrence

$$\overline{p}_{i+1} = \Psi_i^\sharp Q_{\delta_i}^*[\Psi_i^\flat\overline{p}_i] , \tag{14}$$

where the family of operators $\{Q_\delta\,,\,\delta \geq 0\}$ is defined by

$$Q_\delta\phi \triangleq \mathbf{E}^\dagger(\phi(X_{t+\delta})\exp\left\{\tfrac{1}{12}|h(X_{t+\delta}) - h(X_t)|^2\delta\right\} \mid \mathcal{F}_t) .$$

Note that $\Psi_i^\sharp(x)\ \Psi_i^\flat(x) = \Psi_i(x)$, and that the operator Q_{δ_i} can be seen as a perturbation of the semi-group P_{δ_i}. However, it is not obvious that (14) is a computable expression and can be considered as a time discretization of the Zakai equation (3). The relevant analysis and the rate of convergence of this approximation will be considered elsewhere.

5 A product formula and its rate of convergence

The purpose of this section is to study, from the point of view of numerical analysis, the following recurrence

$$\overline{p}_{i+1} = P_{\delta_i}^*[\Psi_i\overline{p}_i] , \tag{15}$$

derived in the previous section, as a time discretization scheme for the Zakai equation

$$dp_t = L^*p_t\,dt + h^*p_t\,dY_t . \tag{16}$$

Recall that

$$(p_{t_i},\phi) = \mathbf{E}^\dagger(\phi(X_{t_i})Z_{t_i} \mid \mathcal{Y}_{t_i}) ,$$

$$(\overline{p}_i,\phi) = \mathbf{E}^\dagger(\phi(\overline{X}_{t_i})\overline{Z}_{t_i} \mid \overline{\mathcal{Y}}_{t_i}) ,$$

so that $\overline{p}_i$ should be "close" to p_{t_i}. Indeed it will be proved below that

$$\left\{\mathbf{E}^\dagger|\overline{p}_i - p_{t_i}|^2\right\}^{1/2} \leq C\delta ,$$

where δ is the mesh of the partition π up to time t_i, and $|\cdot|$ denotes the norm in the Sobolev space $L^2(\mathbf{R}^m)$.

Remark 5.1 Similar rate of convergence has already been obtained for approximation of nonlinear filtering problems, in Picard [6] and Newton [4]. The proof in [6] uses only probabilistic arguments and does not consider the Zakai equation, but rather the underlying nonlinear filtering problem. In [4], the Zakai equation is considered for pure-jump Markov processes rather than diffusion processes, and the approximation procedure relies on the stochastic Taylor formula of Wagner–Platen [7,8].

Define, for all $x \in \mathbf{R}^m$

$$\Psi_t^s(x) \triangleq \exp\left\{h^*(x)\,(Y_t - Y_s) - \tfrac{1}{2}|h(x)|^2\,(t-s)\right\} .$$

Note that two operators are involved in (16)

- the unbounded operator L^* which generates the adjoint semi-group $\{P_t^*\,,\,t \geq 0\}$,
- the multiplication operator B which generates the two-parameter stochastic semi-group $\{\Psi_t^s\,,\,0 \leq s \leq t\}$,

so that the time discretization scheme (15) is a Trotter-like product formula for the Zakai equation (16). See Bensoussan–Glowinski–Rascanu [1] for a related work in this direction.

The main assumption of this section is that the signal $\{X_t\,,\,t \geq 0\}$ is a diffusion process

$$dX_t = b(X_t)\,dt + \sigma(X_t)\,dW_t , \qquad X_0 \sim p_0(x)\,dx$$

observed in continuous time through measurements

$$dY_t = h(X_t)\,dt + dV_t .$$

Define $a \triangleq \sigma\sigma^*$ and $\overline{a}^i \triangleq \sum_{j=1}^m \frac{\partial a^{i,j}}{\partial x_j}$. The coefficients satisfy the following hypotheses

(i) p_0 *is a density on* $\mathbf{R}^m$,

(ii) σ *is a continuous and bounded function on* $\mathbf{R}^m$ *and* a *is a uniformly elliptic* $m \times m$ *matrix, i.e.* $a(x) \geq \alpha I$,

(iii) b *and* $\overline{a}$ *are bounded and measurable functions from* $\mathbf{R}^m$ *to* $\mathbf{R}^m$,

(iv) h *is a measurable and bounded function from* $\mathbf{R}^m$ *to* $\mathbf{R}^d$.

The infinitesimal generator of the semi-group $\{P_t\,,\,t \geq 0\}$ is defined by

$$L \triangleq \tfrac{1}{2}\sum_{i,j=1}^m a^{i,j}\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^m b^i\frac{\partial}{\partial x_i} ,$$

and satisfies, under the hypotheses, the following coercivity property: for all $u \in H^1(\mathbf{R}^m)$

$$2\langle Lu,u\rangle + \mu\|u\|^2 \leq \lambda|u|^2 , \tag{17}$$

where $\|\cdot\|$ denotes the norm in the Sobolev space $H^1(\mathbf{R}^m)$. Existence and uniqueness of a solution to equation (16) is proved in Pardoux [5] and Krylov–Rozovskii [3].

Theorem 5.2 *Suppose that, in addition to (i)–(iv)*

(v) a, b and $\overline{a}$ have bounded first derivative,

(vi) h has bounded derivatives up to order 2 .

Then, if $p_0 \in H^1(\mathbf{R}^m)$

$$\max_{0 \leq k \leq i} \left\{ \mathbf{E}^\dagger |p_{t_k} - \overline{p}_k|^2 \right\}^{1/2} \leq C\delta \; . \tag{18}$$

PROOF. Under the hypotheses, it follows from Theorem 2.1 of [5] that $p \in L^2(\Omega\,;\, C([0,T]\,;\, H^1(\mathbf{R}^m)))$. Also, for all $i \geq 0$, $\overline{p}_i \in L^2(\Omega\,;\, H^1(\mathbf{R}^m))$ and in addition

$$\max_{0 \leq k \leq i} \mathbf{E}^\dagger \|\overline{p}_k\|^2 \leq C \; .$$

For $t \geq t_k$, define $v_t \triangleq P^*_{t-t_k}[\Psi^{t_k}_t \overline{p}_k]$, so that $\overline{p}_k = v_{t_k}$ and $\overline{p}_{k+1} = v_{t_{k+1}}$. Differentiating with respect to t gives

$$\begin{aligned} dv_t \;&=\; L^* v_t \, dt + \left\{ P^*_{t-t_k} [B \Psi^{t_k}_t \overline{p}_k] \right\}^* \, dY_t \\ &=\; L^* v_t \, dt + [B p_t]^* \, dY_t + \beta^*_t \, dY_t \; , \end{aligned}$$

where the perturbation term is defined by

$$\beta_t \triangleq [P^*_{t-t_k} B - B P^*_{t-t_k}] \, [\Psi^{t_k}_t \overline{p}_k] \; .$$

Note that $\beta \in L^2(\Omega\,;\, C([t_k, T]\,;\, H^1(\mathbf{R}^m)))$. The identity of energy of [5] applied to the difference $\epsilon = v - p$, and the coercivity property (17) give

$$\mathbf{E}^\dagger |\epsilon_t|^2 \leq \mathbf{E}^\dagger |\epsilon_{t_k}|^2 + C \int_{t_k}^t \mathbf{E}^\dagger |\epsilon_s|^2 \, ds + C' \int_{t_k}^t \mathbf{E}^\dagger |\beta_s|^2 \, ds \; .$$

Assume the following estimate holds

$$\mathbf{E}^\dagger |\beta_s|^2 \leq C |s - t_k|^2 \, \exp\left\{ C(s - t_k) \right\} \, \mathbf{E}^\dagger \|\overline{p}_k\|^2 \; . \tag{19}$$

Applying Gronwall's lemma and setting $t = t_{k+1}$, yields

$$\begin{aligned} &\mathbf{E}^\dagger |\overline{p}_{k+1} - p_{t_{k+1}}|^2 \\ &\leq \left[\mathbf{E}^\dagger |\overline{p}_k - p_{t_k}|^2 + C |t_{k+1} - t_k|^3 \right] \, \exp\left\{ C(t_{k+1} - t_k) \right\} \; , \end{aligned}$$

and the rate of convergence (18) follows from the discrete Gronwall lemma. The end of the proof is devoted to proving estimate (19).

First, the following perturbation result holds

$$\begin{aligned} &[P^*_{t-t_k} B - B P^*_{t-t_k}] \, u \\ &= \int_{t_k}^t P^*_{t-s} \, [L^* B - B L^*] \, P^*_{s-t_k} u \, ds \; , \end{aligned}$$

provided u is smooth enough. Under the hypotheses, $[L^* B - B L^*]$ is a bounded operator from $H^1(\mathbf{R}^m)$ to $L^2(\mathbf{R}^m)$, so that it is enough that $u \in H^1(\mathbf{R}^m)$ for (20) to hold. Now, $[\Psi^{t_k}_t \overline{p}_k] \in H^1(\mathbf{R}^m)$ a.s. so that

$$\beta_t = \int_{t_k}^t P^*_{t-s} \, [L^* B - B L^*] \, P^*_{s-t_k} \, [\Psi^{t_k}_t \overline{p}_k] \, ds \; .$$

Therefore

$$|\beta_t|^2 \leq C |t - t_k|^2 \; \|\Psi^{t_k}_t \overline{p}_k\|^2 \; ,$$

and

$$\begin{aligned} \mathbf{E}^\dagger |\beta_t|^2 \;&\leq\; C |t - t_k|^2 \; \mathbf{E}^\dagger \|\Psi^{t_k}_t \overline{p}_k\|^2 \\ &\leq\; C |t - t_k|^2 \; \exp\left\{ C(t - t_k) \right\} \; \mathbf{E}^\dagger \|\overline{p}_k\|^2 \; , \end{aligned}$$

which proves (19). □

Remark 5.3 The same rate of convergence holds for the approximation scheme (13).

The next step is to approximate the adjoint semi-group $\{P^*_t \, , \, t \geq 0\}$ itself, i.e. to approximate the associated Fokker–Planck equation. For instance, using an implicit Euler scheme results in the following approximation scheme

$$(I - \delta_i L^*) \overline{p}_{i+1} = \Psi_i \overline{p}_i \; .$$

References

[1] A. BENSOUSSAN, R. GLOWINSKI and A. RASCANU, Approximation of Zakai equation by the splitting–up method, *preprint (September 1988)*.

[2] H. KOREZLIOGLU and G. MAZZIOTTO, Approximations of the nonlinear filter by periodic sampling and quantization, in: *Analysis and Optimization of Systems*, 553–567, Springer–Verlag (LNCIS #62) (1984).

[3] N.V. KRYLOV and B.L. ROZOVSKII, On the Cauchy problem for linear stochastic partial differential equations, *Math. USSR Izvestija* **11**,6,1267–1284 (1977).

[4] N.J. NEWTON, Discrete approximations for Markov chain filters, *Ph.D Thesis, Imperial College* (1983).

[5] E. PARDOUX, Stochastic partial differential equations and filtering of diffusion processes, *Stochastics* **3**,2,127–167 (1979).

[6] J. PICARD, Approximation of nonlinear filtering problems and order of convergence, in: *Filtering and Control of Random Processes*, 219–236, Springer–Verlag (LNCIS #61) (1984).

[7] W. WAGNER and E. PLATEN, Approximation of Itô integral equations, *report ZIMM, Akademie der Wissenschaften der DDR* (1978).

[8] E. PLATEN and W. WAGNER, On a Taylor formula for a class of Itô processes, *Probability and Mathematical Statistics* **3**,1,37–51 (1982).