Convergence in nonlinear filtering for stochastic delay systems *

Antonella Calzolari † Patrick Florchinger ‡ Giovanna Nappo §

Abstract

We study an approximation scheme for a nonlinear filtering problem when the state process $\boldsymbol{X}$ is the solution of a stochastic delay diffusion equation, and the observation process is a noisy function of $X(s)$ for $s \in [t-\tau, t]$, where τ is a constant. The approximating state is the piecewise linear Euler-Maruyama scheme, and the observation process is a noisy function of the approximating state. The rate of convergence of this scheme is computed.

1 Introduction

Stochastic diffusion processes with delay have been used as models in many applications: in population dynamics (see Goel, Maitra and Montroll (1971) [12]), in respiratory systems (see Longtin, Milton, Bos and Mackey (1990) [23]), in eyes movements control (see Vasilakos and Beuter (1993) [34]), postural control (see Peterka (2000) [29]), transmission delays for neural networks and/or ensemble of coupled neural oscillators (see Niebur, Schuster and Kammen (1991) [27]).

In most of the literature the stochastic process is assumed to be completely observable. However, this cannot always be the case, since measurement errors may occur. This difficulty can be overcome by modelling this situation as a nonlinear filtering problem.

The aim of stochastic nonlinear filtering is to compute the conditional law at time t of a state process, which cannot be directly observed, given an observation process up to time t. This task can be achieved only in a few specific cases, and therefore the problem of the approximation of the conditional law naturally arises.

A classical model of partially observed system extensively studied in the last past years arises when both the state and the observation processes are diffusion processes.

For this model, under suitable hypotheses on the coefficients, the filtering equations are well known (see for example Pardoux (1991) [28] or Kallianpur (1980) [14] and the references therein) and different approximation schemes have been studied (see for example Kushner (1990) [18] or Le Gland (1989) [21] and the references therein).

In this paper we are interested in nonlinear filtering of partially observed delay systems of the following form.

The state process $\boldsymbol{X} = \big(X(t)\big)_{t\in[-\tau,T]}$ satisfies the stochastic delay differential equation on the probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, P)$

$$\begin{cases} X(t) = \eta(t), & -\tau \le t \le 0, \\ X(t) = \eta(0) + \int_0^t a(u, \Pi_u X)du + \int_0^t b(u, \Pi_u X)d\tilde{W}_u, & 0 \le t \le T, \end{cases} \tag{1}$$

*Partially supported by MURST project *"Modelli probabilistici e 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

where τ is a positive constant, $\big(\Pi_t X\big)_{t\in[0,T]}$ is a $C([-\tau,0],\mathbb{R})$ random valued process defined by

$$\Pi_t X(s) = X(t+s) \qquad -\tau \le s \le 0,$$

$\tilde{\boldsymbol{W}} = \big(\tilde{W}(t)\big)_{t\in[0,T]}$ is a standard Brownian motion, and $\boldsymbol{\eta} = \big(\eta(s)\big)_{s\in[-\tau,0]}$ is a $C([-\tau,0],\mathbb{R})$ valued random variable.

The observation process $\boldsymbol{Y} = \big(Y(t)\big)_{t\in[0,T]}$ is given by

$$Y(t) = \int_0^t h(u,\Pi_u X)du + W(t), \qquad 0\le t\le T, \tag{2}$$

where $\boldsymbol{W} = \big(W(t)\big)_{t\in[0,T]}$ is a standard Brownian motion, independent of $\tilde{\boldsymbol{W}}$, and $h : [0,T]\times C([-\tau,0],\mathbb{R}) \to \mathbb{R}$ is a Borel measurable function.

As an example the functions $a(t,\theta)$, $b(t,\theta)$ and $h(t,\theta)$, for $\theta\in C([-\tau,0],\mathbb{R})$ can be taken of the form

$$g\left(t, \max_{u\in[\tau_{i-1},\tau_i]} \theta(u);\; i=1,\cdots,r\right) \tag{3}$$

where $-\tau = \tau_0 < \tau_1 < \cdots < \tau_r = 0$, or

$$g\left(t, \textstyle\int_{-\tau}^0 \psi_i(u,\theta(u))\gamma_i(du);\; i=1,\cdots,r\right), \tag{4}$$

where γ_i are finite measures on $[-\tau,0]$, and g and ψ_i are continuous functions.
By taking in (4) $\psi_i(u,x)=x$ for all i, and $\gamma_i(ds)$ in the set $\{e^{\lambda s}ds,\ \delta_{-\tau}(ds),\ \delta_0(ds)\}$, we recover the stochastic delay differential equation considered in a control framework by Larssen and Risebro (2003) in [20] and by Elsanousi and Larssen (2001) in [8]

$$\begin{aligned} dX(t) &= g_a\big(t, X(t),\, X(t-\tau), \textstyle\int_{t-\tau}^t e^{\lambda(u-t)}X(u)\,du,\, w(t)\big)dt \\ &\quad + g_b\big(t, X(t),\, X(t-\tau), \textstyle\int_{t-\tau}^t e^{\lambda(u-t)}X(u)\,du,\, w(t)\big)d\widetilde{W}_t, \end{aligned}$$

where $w(t)$ is the control. Moreover, when the function $g_a(t,x_1,x_2,x_3,w)$, as well as $g_b(t,x_1,x_2,x_3,w)$, depends only on (x_1,x_2), we recover the fixed time delay model

$$dX(t) = g_a\big(X(t),\, X(t-\tau)\big)dt + g_b\big(X(t),\, X(t-\tau)\big)d\widetilde{W}_t, \tag{5}$$

which is studied in Frank (2002) [10] and in particular, for $g_a(x_1,x_2) = \big(H - K\log(x_2)\big)\,x_1$ and $g_b(x_1,x_2) = x_1$, we recover the stochastic Gompertz model with delay, see Goel, Maitra and Montroll (1971) [12] and Frank and Beek (2001) [11]. The stochastic Gompertz model without delay has been used in population growth (see Ricciardi and al. (1986) [32]) or in biomedical sciences (see Ferrante et al. (2000) [9]).

At first the aim was to obtain a computable approximation for $E\big[\varphi\big(X(t)\big)/\mathcal{F}_t^Y\big]$, for all functions φ belonging to a determining class, i.e. the best estimate of $\varphi\big(X(t)\big)$ given the $\sigma-$algebra of the observations up to time t, $\mathcal{F}_t^Y = \sigma\{Y(s),\, s\le t\}$. In fact, since $\Pi_t X(0) = X(t)$, we shall give a computable approximation for the filter π_t associated with the delay system $\big(\Pi_t X, Y(t)\big)_{t\in[0,T]}$, defined for any measurable and bounded functions ϕ mapping $C([-\tau,0],\mathbb{R})$ into $\mathbb{R}$ by

$$\pi_t(\phi) = E[\phi(\Pi_t X)|\mathcal{F}_t^Y]. \tag{6}$$

In our paper the state is approximated by the piecewise linear Euler-Maryuama scheme (see (12) and (15)), while the observation is approximated by a diffusion (see (13)), which can be considered as a continuous Euler-Maryuama scheme (see Remark 2.1, for the peculiarities due to the delay). The filter process of this approximating system is the first approximation scheme of $\boldsymbol{\pi} = (\pi_t)_{t\in[0,T]}$ we consider (see

(20)). This approximation scheme has a drawback: it depends on the approximating observation process, which is not the actual observation process $\boldsymbol{Y}$. Instead, the other approximation scheme we consider depends on the actual observation process (see (21)).

As the time step converges to zero, the two approximating filters converge in probability to $\boldsymbol{\pi}$ as measure valued processes (see Theorem 2.2).

To our knowledge there are only three papers dealing with nonlinear filtering for delay systems: Kwong and Willsky (1978) [19], Chang (1987) [7], and Kallianpur and Mandal (2002) [15].

In [19] Kwong and Willsky give a characterization of the optimal filter when dealing with nonlinear delay systems with Gaussian noises, i.e. with b depending only on time. A Fujisaki-Kallianpur-Kunita equation for the filter is deduced from a representation result which characterizes conditional moment functionals of nonlinear delay systems. However the uniqueness of the solution of this equation is not guaranteed.

In [7] Chang gives a computable approximation for the optimal filter when dealing with one dimensional nonlinear delay filtering systems with $b = 1$. The original model is approximated by a discrete-time model obtained by applying an Euler discretization scheme. An optimal filter for the approximate system is obtained by an explicit procedure and the weak convergence of the approximating process and the approximating filter to the original ones are verified.

In [15], Kallianpur and Mandal study a nonlinear filtering problem where the state process is solution of the stochastic delay differential equation (1), in the homogeneous case, and the observation process is given by (2). By using some extensions of results obtained by Mohammed (1984) [25] for stochastic delay differential equations they prove that the signal process is the unique solution to an appropriate martingale problem. By taking this fact into account the authors prove that the optimal filter corresponding to the nonlinear filtering problem solves a Zakai-type equation. The uniqueness for the solution of the Zakai equation is deduced from the results of Bhatt, Kallianpur and Karandikar (1995) [2], and a Fujisaki-Kallianpur-Kunita equation for the filter is deduced from the Zakai equation by usual arguments in nonlinear filtering theory.

In addition to the previous references we also quote the paper [3] by Bhatt, Kallianpur and Karandikar (1999), which is the starting point in some of our analysis, and Bhatt and Karandikar (2002) [4]. Though none of these papers is explicitly connected with filtering models involving delays, the results achieved by these authors can be used in the delay context.

This paper is divided into 6 sections and is organized as follows. In Section 2, we introduce the approximation scheme for the system we are dealing with in this paper and we state the main results. The first result concerns the convergence of the approximation schemes for the filter, while the second result deals with the rate of convergence with respect to the bounded Lipschitz metric. In Section 3, we prove the convergence for the filter by making use of a convergence result deduced from the papers by Bhatt, Kallianpur and Karandikar (1999) [3] and Bhatt and Karandikar (2002) [4]. In Section 4, we compute an upper bound for the rate of convergence with respect to the bounded Lipschitz metric for our approximation scheme by combining filter approximation techniques similar to those in Calzolari, Florchinger and Nappo (2005) [6] with a convergence result for the approximation of stochastic delay differential equations. In Section 5, we give the proofs of some technical results that we use in Sections 3 and 4. Moreover we recall a result on the expectation of the modulus of continuity for diffusion due to Słomiǹski (2001) [33], which we use in the proof of the approximation result for stochastic delay differential equations. In Section 6, we conclude by giving a comparison between our result and the one by Chang [7] and we discuss briefly some works on approximation for stochastic delay differential equations, in particular the one by Mao and Sabanis (2003) [24].

2 Approximation scheme and main results

For the partially observed delay system (1) and (2) stated above we assume the following standing hypotheses:

(A1) η is a $\mathcal{F}_0$-measurable $C([-\tau,0],\mathbb{R})$ valued random variable, with

$$E\left(\|\Pi_0 X\|^4\right) = E\Big(\sup_{s\in[-\tau,0]}|\eta(s)|^4\Big) < \infty$$

(A2) The functionals $a(t,\theta)$ and $b(t,\theta)$ on $[0,T]\times C([-\tau,0],\mathbb{R})$ are jointly globally Lipschitz, i.e.

$$|a(t,\theta)-a(t',\bar\theta)|^2 + |b(t,\theta)-b(t',\bar\theta)|^2 \le K\left(|t-t'|^2 + \|\theta-\bar\theta\|^2\right), \tag{7}$$

and satisfy the growth condition

$$|a(t,\theta)|^2 + |b(t,\theta)|^2 \le K\left(1+\|\theta\|^2\right), \tag{8}$$

for some constant $K>0$.

(A3) $h:[0,T]\times C([-\tau,0],\mathbb{R})\to\mathbb{R}$ is jointly continuous and sublinear, i.e.

$$|h(t,\theta)|^2 \le K(1+\|\theta\|^2).$$

Condition **(A2)** assures the existence and the uniqueness of the solution of equation (1) (see Kallianpur and Mandal [15]). Condition **(A1)** together with (8) implies that

$$E\Big[\sup_{u\in[0,T]}\|\Pi_u X\|^2\Big] < \infty, \tag{9}$$

(see [25], Theorem II.2.1 and Lemma III.1.2), which together with the sublinearity of h implies that

$$\int_0^T E[|h(u,\Pi_u X)|^2]du < \infty.$$

The above condition, together with the independence of the noises, is a classical assumption in non-linear filtering theory which guarantees that the filter π_t can be represented via a Kallianpur–Striebel formula

$$\pi_t(\phi) = \frac{\sigma_t(\phi)}{\sigma_t(\mathbf{1})},$$

with

$$\sigma_t(\phi) = E^0\left[\phi(\Pi_t X)\exp\left\{\int_0^t h(s,\Pi_s X)dY_s - \frac{1}{2}\int_0^t |h(s,\Pi_s X)|^2\,ds\right\}\Big|\mathcal{F}_t^Y\right],$$

where E^0 denotes the expectation w.r.t. the reference probability measure P^0, defined by the Radon-Nikodym derivative

$$\frac{dP^0}{dP} = \exp\left\{-\int_0^T h(s,\Pi_s X)dY_s + \frac{1}{2}\int_0^T |h(s,\Pi_s X)|^2\,ds\right\}. \tag{10}$$

This fact will play a fundamental role in the proof of our approximation results. In particular it implies that there exists a deterministic functional, with values in $D_{\mathcal{P}(C([-\tau,0],\mathbb{R}))}([0,T])$, where $\mathcal{P}(C([-\tau,0],\mathbb{R}))$ is the metric space of probability measures on $C([-\tau,0],\mathbb{R})$, endowed with the Prohorov metric,

$$\begin{aligned} U:\ [0,T]\times C([0,T],\mathbb{R}) &\to D_{\mathcal{P}(C([-\tau,0],\mathbb{R}))}([0,T]) \\ (t,\boldsymbol{y}) &\mapsto U\big(t,\boldsymbol{y}\big) \end{aligned}$$

with the property $U\big(t,\boldsymbol{y}\big) = U\big(t, y(\cdot \wedge t)\big)$, and such that

$$\pi_t = U\big(t, \boldsymbol{Y}\big). \tag{11}$$

In this paper we consider the following approximation scheme:

The approximation $\boldsymbol{X^n} = \big(X^n(t)\big)_{t\in[-\tau,T]}$ of the state process $\boldsymbol{X} = \big(X(t)\big)_{t\in[-\tau,T]}$ is the piecewise linear Euler-Maruyama scheme, that is the linear interpolation of the Euler discretization scheme with step $\delta = \delta_n = T/n$, with $\tau = m\delta$ (as in Chang [7], for the sake of simplicity, we assume that T/τ is rational)[1]:

$$\begin{cases} X^n(\ell\delta) = \eta(\ell\delta), & -m \le \ell \le 0, \\ \\ X^n((\ell+1)\delta) = X^n(\ell\delta) + a(\ell\delta, \Pi_{\ell\delta}X^n)\delta & \\ \qquad\qquad\qquad + b(\ell\delta, \Pi_{\ell\delta}X^n)\big[\tilde{W}((\ell+1)\delta) - \tilde{W}(\ell\delta)\big], & 0 \le \ell \le n-1. \end{cases} \tag{12}$$

With this approximation for the state process $\boldsymbol{X}$ we can consider the piecewise-constant $C([-\tau,0],\mathbb{R})$-valued process $\big(\Pi_{\lfloor t/\delta\rfloor\cdot\delta}X^n\big)_{t\in[0,T]}$ as an approximation of the $C([-\tau,0],\mathbb{R})$-valued process $\big(\Pi_t X\big)_{t\in[0,T]}$.

For the approximation of the observation process, we define $\boldsymbol{Y^n} = \big(Y^n(t)\big)_{t\in[0,T]}$ by

$$Y^n(t) = \int_0^t h(\lfloor s/\delta\rfloor\cdot\delta, \Pi_{\lfloor s/\delta\rfloor\cdot\delta}X^n)ds + W(t), \qquad 0 \le t \le T, \tag{13}$$

where $\lfloor x\rfloor$ is the integer part of x.

Remark 2.1. *Note that, unlike in the finite dimensional Euler scheme, the interpolation has to be performed at every step in order to evaluate $\Pi_{\ell\delta}X^n$. Nevertheless it is clear that*

$$\big\{\big(X^n(\ell\delta), X^n((\ell-1)\delta), \cdots, X^n((\ell-m)\delta)\big)\big\}_{0\le\ell\le n} \tag{14}$$

is an $(m+1)$-dimensional Markov chain, and for $t \in [\ell\delta,(\ell+1)\delta]$, $0 \le \ell \le n-1$

$$\begin{aligned} X^n(t) = X^n(\ell\delta) &+ a(\ell\delta, \Pi_{\ell\delta}X^n)(t-\ell\delta) \\ &+ b(\ell\delta, \Pi_{\ell\delta}X^n)\big[\tilde{W}((\ell+1)\delta) - \tilde{W}(\ell\delta)\big](t-\ell\delta)/\delta, \end{aligned} \tag{15}$$

with $X^n(0) = \eta(0)$, and

$$Y^n(t) = Y^n(\ell\delta) + h(\ell\delta, \Pi_{\ell\delta}X^n)(t-\ell\delta) + \big[W(t) - W(\ell\delta)\big], \tag{16}$$

with $Y^n(0) = 0$.

When the state is given by fixed time delay model (5), the linear interpolation, in the above discrete Euler-Maruyama scheme, is not needed in order to compute the sequence $\{X^n(\ell\delta)\}_{0\le\ell\le n}$. Indeed in this case

$$\begin{aligned} X^n((\ell+1)\delta) = X^n(\ell\delta) &+ g_a\big(\ell\delta, X^n(\ell\delta), X^n((\ell-m)\delta)\big)\delta \\ &+ g_b\big(\ell\delta, X^n(\ell\delta), X^n((\ell-m)\delta)\big)\big[\tilde{W}((\ell+1)\delta) - \tilde{W}(\ell\delta)\big], \end{aligned} \tag{17}$$

with $X^n(0) = \eta(0)$, and therefore the computation of the discrete Markov chain (14) is much simpler.

[1]It is clear that, assuming $T = \frac{p}{q}\tau$, we first fix $m = kq$ a multiple of q and then set $\delta = \tau/m$, so that $T = kp\delta$ and $n = kp$.
Or better, we first fix m, then set $\delta = \tau/m$, and finally take the interval $[-\tau, \lfloor T/\delta\rfloor\delta]$, instead of $[-\tau, T]$, so that $n = n(m)$.

The filter π_t^n associated with the approximating delay system $\left(\Pi_{\lfloor t/\delta\rfloor\cdot\delta} X^n, Y^n(t)\right)_{t\in[0,T]}$ is given by Kallianpur Striebel formula:

$$\pi_t^n(\phi) = \frac{\sigma_t^n(\phi)}{\sigma_t^n(\mathbf{1})} = \frac{E^{0,n}\left[\phi(\Pi_{\lfloor t/\delta\rfloor\cdot\delta} X^n)\mathcal{L}_t^n|\mathcal{F}_t^{Y^n}\right]}{E^{0,n}\left[\mathcal{L}_t^n|\mathcal{F}_t^{Y^n}\right]}, \tag{18}$$

where $E^{0,n}$ denotes the expectation w.r.t. the reference probability measure $P^{0,n}$, defined by the Radon-Nikodym derivative

$$\frac{dP^{0,n}}{dP} = \left(\mathcal{L}_T^n\right)^{-1},$$

with

$$\mathcal{L}_t^n = \exp\left\{\int_0^t h(\lfloor s/\delta\rfloor\cdot\delta, \Pi_{\lfloor s/\delta\rfloor\cdot\delta}X^n)dY_s^n - \frac{1}{2}\int_0^t |h(\lfloor s/\delta\rfloor\cdot\delta, \Pi_{\lfloor s/\delta\rfloor\cdot\delta}X^n)|^2\, ds\right\}, \tag{19}$$

which is well defined thanks to the sublinearity of h and **(A1)**.

Taking into account that $s\mapsto h(\lfloor s/\delta\rfloor\cdot\delta, \Pi_{\lfloor s/\delta\rfloor\cdot\delta}X^n)$ is piecewise constant, we have that

$$\mathcal{L}_t^n = L_t^n(X^n(\cdot), Y_0^n, Y_\delta^n, \cdots, Y_{\lfloor t/\delta\rfloor}^n, Y_t^n),$$

where, for $0\le\ell\le n$,

$$\log L_{\ell\delta}^n(x(\cdot), y_0, y_1, \cdots, y_\ell) = \sum_{k=0}^{\ell-1} h(k\delta, \Pi_{k\delta}x(\cdot))(y_{k+1}-y_k) - \frac{1}{2}\sum_{k=0}^{\ell-1}|h(k\delta, \Pi_{k\delta}x(\cdot))|^2\delta,$$

and, for $t\in(\ell\delta, (\ell+1)\delta)$, $0\le\ell\le n-1$,

$$\begin{aligned}\log L_t^n(x(\cdot), y_0, y_1, \cdots, y_\ell, y) &= \log L_{\ell\delta}^n(x(\cdot), y_0, y_1, \cdots, y_\ell)\\ &+ h(\lfloor t/\delta\rfloor\cdot\delta, \Pi_{\lfloor t/\delta\rfloor\cdot\delta}x(\cdot))(y-y_\ell) - \frac{1}{2}|h(\lfloor t/\delta\rfloor\cdot\delta, \Pi_{\lfloor t/\delta\rfloor\cdot\delta}x(\cdot))|^2(t-\ell\delta).\end{aligned}$$

Moreover, under $P^{0,n}$, the processes $\boldsymbol{X^n}$ and $\boldsymbol{Y^n}$ are independent and the law of the approximated state process is invariant under P and $P^{0,n}$, and hence, for $t\in[\ell\delta, (\ell+1)\delta)$, $0\le\ell\le n-1$,

$$\sigma_t^n(\phi) = E\left[\phi(\Pi_{\lfloor t/\delta\rfloor\cdot\delta}X^n)L_t^n(X^n(\cdot), y_0, y_1, \cdots, y_\ell, y)\right]\Big|_{y_0=Y_0^n, y_1=Y_\delta^n, \cdots, y_\ell=Y_{\ell\delta}^n, y=Y_t^n}.$$

Therefore, by taking the above equality into account, one can explicitly obtain a deterministic functional

$$\begin{aligned}U^n:\ [0,T]\times C([0,T],\mathbb{R}) &\to D_{\mathcal{P}(C([-\tau,0],\mathbb{R}))}([0,T])\\ (t,\boldsymbol{y}) &\mapsto U^n\big(t,\boldsymbol{y}\big)\end{aligned}$$

with the property that $U^n\big(t,\boldsymbol{y}\big)$ depends only on $\big(y(k\delta)\big)_{0\le k\le\lfloor\frac{t}{\delta}\rfloor}$ and $y(t)$, such that

$$\pi_t^n = U^n\big(t, \boldsymbol{Y^n}\big). \tag{20}$$

So the filter π_t^n defined above depends explicitly on the approximated observation process $\boldsymbol{Y^n}$, which however is not directly observable. To overcome this difficulty we also consider the following approximation $\tilde\pi_t^n$ for the filter

$$\tilde\pi_t^n = U^n\big(t,\boldsymbol{Y}\big) = \frac{\tilde\sigma_t^n}{\tilde\sigma_t^n(\mathbf{1})}, \tag{21}$$

where

$$\tilde\sigma_t^n(\phi) = E\left[\phi(\Pi_{\lfloor t/\delta\rfloor\cdot\delta}X^n)L_t^n(X^n(\cdot), y_0, y_1, \cdots, y_\ell, y)\right]\Big|_{y_0=Y_0, y_1=Y_\delta, \cdots, y_\ell=Y_{\ell\delta}, y=Y_t}.$$

Then, the following convergence result, which will be proved in the following section, holds.

Theorem 2.2. *Let $\boldsymbol{\pi} = (\pi_t; t \geq 0)$, $\boldsymbol{\pi^n} = (\pi^n_t; t \geq 0)$, and $\tilde{\boldsymbol{\pi}}^{\boldsymbol{n}} = (\pi^n_t; t \geq 0)$, be the càdlàg probability measure-valued processes defined by* (6), (18) *and* (21)*, respectively. Then:*
1. The sequence of filters $\boldsymbol{\pi^n}$ converges in probability (and therefore weakly) to the original filter $\boldsymbol{\pi}$, in $D_{\mathcal{P}(C([-\tau,0],\mathbb{R}))}([0,T])$.
2. The sequence of measure valued processes $\tilde{\boldsymbol{\pi}}^{\boldsymbol{n}}$ converges in probability to the original filter $\boldsymbol{\pi}$.
3. The sequence $\max_{k=1,\cdots,n} d(\tilde{\pi}^n_{k\delta}, \pi_{k\delta})$, where d denotes the Prohorov metric, converges in probability to zero.

In addition, the following result concerning the rate of convergence with respect to the bounded Lipschitz metric of our approximation scheme will be proved in Section 4. For the ease of the reader we recall that, for any metric space S, and probability measures ν_1 and ν_2 on S,

$$d_{BL}(\nu_1, \nu_2) = \sup \left\{ \frac{|\nu_1(\varphi) - \nu_2(\varphi)|}{\|\varphi\| \vee L_\varphi}; \varphi \text{ bounded and Lipschitz} \right\}$$

where $\|\varphi\|$ denotes the sup-norm, and L_φ is the Lipschitz constant of φ.

Theorem 2.3. *Assume further that the functions a and b are bounded, the function h is jointly globally Lipschitz, and that there exists a constant C_η such that the modulus of continuity of the initial condition η satisfies*

$$E\big[\omega^2_\eta\big(\delta; [-\tau, 0]\big)\big] \leq C_\eta\, \delta\, \log(\tfrac{1}{\delta}). \tag{22}$$

Then there exists a constant C such that

$$E\big[d_{BL}(\pi_t, \tilde{\pi}^n_t)\big] \leq C(\tfrac{\log n}{n})^{\frac{1}{2}}, \tag{23}$$

where d_{BL} is the bounded Lipschitz metric on the space $\mathcal{P}(C([-\tau,0],\mathbb{R}))$.

Remark 2.4. *As it will be shown in the example at the end of Section 4, by considering the case where $\eta = 0$, $a = 0$, $b = 1$ and $h = 0$, the upper bound for the rate of convergence given by* (23) *appears to be the best we can obtain in our context.*

Furthermore as it is clear from the proof (see (52)*) if $E\big[\omega^2_\eta\big(\delta; [-\tau, 0]\big)\big]$ converges to zero with an order of convergence lower than $O\big(\delta\, \log(\frac{1}{\delta})\big)$, then*

$$E\big[d_{BL}(\pi_t, \tilde{\pi}^n_t)\big] \leq C\left(E\big[\omega^2_\eta\big(\delta; [-\tau, 0]\big)\big]\right)^{\frac{1}{2}}$$

for a suitable constant C.

To conclude this section, note that in order to evaluate π^n_t and $\tilde{\pi}^n_t$ we need to compute

(a) the transition probability of the $(m+1)$-dimensional Markov chain

$$\big(X^n(\ell\delta), X^n((\ell-1)\delta), \cdots, X^n((\ell-m)\delta)\big),$$

(b) the explicit expression of L^n_t.

Consequently we need the explicit expression of $a(\ell\delta, \Pi_{\ell\delta}X^n)$, $b(\ell\delta, \Pi_{\ell\delta}X^n)$, and $h(\ell\delta, \Pi_{\ell\delta}X^n)$.

When the functionals a, b and h are taken to be of the form (4) with $r = 1$ we need to evaluate expressions like

$$g\big(\ell\delta, \int_{-\tau}^{0} \psi(u, \Pi_{\ell\delta}X^n(u))\, \gamma(du)\big),$$

where

$$\int_{-\tau}^{0} \psi(u, \Pi_{\ell\delta} X^n(u))\gamma(du) = \int_{-\tau}^{0} \psi(u, X^n(\ell\delta + u))\,\gamma(du)$$
$$= \sum_{k=-m}^{-1} \int_{k\delta}^{(k+1)\delta} \psi\big(u, X^n((\ell+k)\delta) + \frac{u-k\delta}{\delta}\big[X^n((\ell+k+1)\delta) - X^n((\ell+k)\delta)\big]\big)\,\gamma(du).$$

3 The convergence result

This section is dedicated to the proof of Theorem 2.2. With this aim, we will make use of a result deduced from the papers by Bhatt, Kallianpur and Karandikar [3] and [4] in the following context.

Consider a signal process $\boldsymbol{\mathcal{X}} = (\mathcal{X}_t)_{t\in[0,T]}$, with values in a complete separable metric space (S, d_S), defined on $(\Omega, \mathcal{F}, P)$, with *càdlàg* paths and continuous in probability, and the observation process $\boldsymbol{Y} = (Y(t))_{t\in[0,T]}$ given by

$$Y(t) = \int_0^t \boldsymbol{h}(\mathcal{X}_s)ds + W(t), \tag{24}$$

where $\boldsymbol{W} = (W(t))_{t\in[0,T]}$ is a standard Brownian motion, defined on $(\Omega, \mathcal{F}, P)$, independent of $\boldsymbol{\mathcal{X}}$, and $\boldsymbol{h}$ is a measurable function on S with values in $\mathbb{R}^k$, such that

$$P\left(\int_0^T |\boldsymbol{h}(\mathcal{X}_s)|^2\, ds < \infty\right) = 1.$$

The approximation signal processes $\boldsymbol{\mathcal{X}^n} = (\mathcal{X}^n_t)_{t\in[0,T]}$ are defined on $(\Omega, \mathcal{F}, P)$, and take values in S as well. The approximation observation processes $\boldsymbol{Y^n} = (Y^n(t))_{t\in[0,T]}$ are defined by

$$Y^n(t) = \int_0^t \boldsymbol{h}^n(\mathcal{X}^n_s)ds + W(t), \tag{25}$$

where $\boldsymbol{W}$ is independent of $\boldsymbol{\mathcal{X}^n}$, and $\boldsymbol{h}^n$ are measurable functions on S with values in $\mathbb{R}^k$, such that

$$P\left(\int_0^T |\boldsymbol{h}^n(\mathcal{X}^n_s)|^2\, ds < \infty\right) = 1.$$

Then the following result is an easy consequence of [3] and Remark 7.4 in [4].

Theorem 3.1. *Assume that*

(B1) $\boldsymbol{h}^n$ *converges to* $\boldsymbol{h}$ *uniformly on compact sets,*

(B2) $\boldsymbol{h}$ *is continuous,*

(B3) $\boldsymbol{\mathcal{X}^n}$ *converges in P-probability (and therefore weakly) to* $\boldsymbol{\mathcal{X}}$ *in* $D_S([0,T])$*,*

(B4) $\lim_{n\to\infty} E\left(\int_0^T |\boldsymbol{h}^n(\mathcal{X}^n_s) - \boldsymbol{h}(\mathcal{X}_s)|^2\, ds\right) = 0.$

Then the filters of the system $(\boldsymbol{\mathcal{X}^n}, \boldsymbol{Y^n})$ *converge in probability (and therefore weakly) to the filter of the system* $(\boldsymbol{\mathcal{X}}, \boldsymbol{Y})$ *as processes with values in* $D_{\mathcal{P}(S)}([0,T])$*, where* $\mathcal{P}(S)$ *is the metric space of probability measures on* S*, endowed with the Prohorov metric.*

Note that the above conditions **(B1)**—**(B4)** are only sufficient conditions, and that in [3] weaker conditions and different frameworks can be found.

In order to be in the framework described above we take $S = [0, T] \times C([-\tau, 0], \mathbb{R})$, endowed with the distance

$$\|(t, \theta) - (t', \theta')\|_S = |t - t'| + \|\theta - \theta'\|,$$

and we consider the limit model $\big(\mathcal{X}_t, Y(t)\big)_{t\in[0,T]}$, where

$$\mathcal{X}_t = (t, \Pi_t X),$$

and $Y(t)$ is given by (2), and the approximating model $\big(\mathcal{X}^n_t, Y^n(t)\big)_{t\in[0,T]}$, where

$$\mathcal{X}^n_t = (\delta \cdot \lfloor t/\delta \rfloor, \Pi_{\delta\cdot\lfloor t/\delta\rfloor} X^n), \tag{26}$$

and $Y^n(t)$ is given by (16).

Then with this choice, the processes $\big(\mathcal{X}_t\big)_{t\in[0,T]}$ and $\big(\mathcal{X}^n_t\big)_{t\in[0,T]}$ have paths in $D_S([0, T])$, and conditions **(B1)** and **(B2)** are obviously satisfied with $\boldsymbol{h}^n = \boldsymbol{h} = h$.

Condition **(B3)** which asserts that $\boldsymbol{\mathcal{X}^n} = \big((\delta \cdot \lfloor t/\delta \rfloor, \Pi_{\delta\cdot\lfloor t/\delta\rfloor} X^n)\big)_{t\in[0,T]}$ converges in probability to $\boldsymbol{\mathcal{X}} = \big((t, \Pi_t X)\big)_{t\in[0,T]}$ in $D_S([0, T]) = D_{[0,T]\times C([-\tau,0],\mathbb{R})}([0, T])$ is also satisfied thanks to the following Proposition which will be proved in Section 5.

Proposition 3.2. *Assume that conditions* **(A1)** *and* **(A3)** *are satisfied. Then*

$$\lim_{n\to\infty} E\left[\sup_{t\in[0,T]} \big\|(\delta \cdot \lfloor t/\delta \rfloor, \Pi_{\delta\cdot\lfloor t/\delta\rfloor} X^n) - (t, \Pi_t X)\big\|^2_S\right] = 0. \tag{27}$$

In our approximation scheme, since the function h appears in the definitions of both $Y(t)$ and $Y^n(t)$, condition **(B4)** is implied by

$$\lim_{n\to\infty} E\left(\int_0^T |h(\delta \cdot \lfloor s/\delta \rfloor, \Pi_{\delta\cdot\lfloor s/\delta\rfloor} X^n)|^2\, ds\right) = E\left(\int_0^T |h(s, \Pi_s X)|^2\, ds\right). \tag{28}$$

Under the continuity and sublinear growth conditions in **(A3)** for the observation function h, we can assume w.l.o.g., since the convergence condition **(B3)** holds, that the integrals inside the expectation in the left hand side of (28) converge to the integral inside the expectation in the right hand side of (28). Therefore we only need a uniform integrability condition. To this end observe that

$$\sup_n E\left[\left(\int_0^T |h(\delta \cdot \lfloor s/\delta \rfloor, \Pi_{\delta\cdot\lfloor s/\delta\rfloor} X^n)|^2\, ds\right)^2\right] \le \sup_n E\left[T\int_0^T |h(\delta \cdot \lfloor s/\delta \rfloor, \Pi_{\delta\cdot\lfloor s/\delta\rfloor} X^n)|^4\, ds\right]$$
$$\le \sup_n T\int_0^T E\left[C'\left(1 + \sup_{t\in[-\tau,T]} |X^n(t)|^4\right)\right] ds < \infty$$

and hence, condition **(B4)** will hold thanks to the following result, which will be proved in Section 5.

Lemma 3.3. *Assume that conditions* **(A1)** *and* **(A2)** *are satisfied, then, for* $k = 1, 2$,

$$\sup_n E\left[\sup_{t\in[0,T]} \|\Pi_t X^n\|^{2k}\right] = \sup_n E\left[\sup_{u\in[-\tau,T]} |X^n(u)|^{2k}\right] < \infty.$$

Therefore, since all the conditions in Theorem 3.1 hold, the filters of the systems $(\boldsymbol{\mathcal{X}^n}, \boldsymbol{Y^n})$ convergence in probability to the filter of the system $(\boldsymbol{\mathcal{X}}, \boldsymbol{Y})$ as random processes with values in $D_{\mathcal{P}(S)}([0,T])$, and this implies the convergence of $\boldsymbol{\pi^n}$ to $\boldsymbol{\pi}$ as random processes with values in $D_{\mathcal{P}(C([-\tau,0],\mathbb{R}))}([0,T])$.

To prove the second assertion of Theorem 2.2, we make use, like in Section 2, of a representation result for the filters $(\boldsymbol{\mathcal{X}^n}, \boldsymbol{Y^n})$ and $(\boldsymbol{\mathcal{X}}, \boldsymbol{Y})$ in the state space S.

In this setting there exist functionals

$$V^n,\, V:\ [0,T] \times D_{\mathbb{R}^k}([0,T]) \mapsto \mathcal{P}(S) \tag{29}$$

with the properties $V^n\big(t, \boldsymbol{y}\big) = V^n\big(t, y(\cdot \wedge t)\big)$ and $V\big(t, \boldsymbol{y}\big) = V\big(t, y(\cdot \wedge t)\big)$, such that the filters of $(\boldsymbol{\mathcal{X}^n}, \boldsymbol{Y^n})$ and $(\boldsymbol{\mathcal{X}}, \boldsymbol{Y})$ are given by $V^n(t, \boldsymbol{Y^n})$ and $V(t, \boldsymbol{Y})$, respectively. This fact is true under very general conditions (see for instance, Kurtz and Ocone (1998) [17]). Note that if one uses the general representation result, then one could only say that the functionals V^n and V are defined almost surely with respect to $P_{\boldsymbol{Y^n}}$, the law of $\boldsymbol{Y^n}$, and with respect to $P_{\boldsymbol{Y}}$, the law of $\boldsymbol{Y}$, respectively. Therefore, the approximation $V^n(t, \boldsymbol{Y})$ of the $(\boldsymbol{\mathcal{X}}, \boldsymbol{Y})$-filter (as the one provided in (21)) could be not well defined. However in this case $P_{\boldsymbol{Y^n}}$ and $P_{\boldsymbol{Y}}$ are equivalent, and this problem does not occur.

In our delay case $S = [0,T] \times C([-\tau,0],\mathbb{R})$ and the functional V^n can be computed starting from the Kallianpur-Striebel formula with the Radon-Nikodym derivative $\mathcal{L}_t^n$ defined by (19). This fact implies also that, for any $(t, \boldsymbol{y})$ in $[0,T] \times C([0,T],\mathbb{R})$ the projection on the space $C([-\tau,0],\mathbb{R})$ of the probability measure $V^n(t, \boldsymbol{y})$ coincides with $U^n(t, \boldsymbol{y})$, defined in Section 2.

Moreover in [3] (see Theorem 3.3-(a)), as a step in the proof of a weak convergence result, the authors prove that for any Wiener process $\boldsymbol{B} = (B(t))_{t\in[0,T]}$, the $\mathcal{P}(S)$-valued processes $(V^n(t, \boldsymbol{B}))_{t\in[0,T]}$ converge in probability to the $\mathcal{P}(S)$-valued process $(V(t, \boldsymbol{B}))_{t\in[0,T]}$. This amounts to say that if P^0 is the reference probability measure defined by the Radon-Nikodym derivative

$$\frac{dP^0}{dP} = \exp\left\{ -\int_0^T \boldsymbol{h}(\mathcal{X}_s) dY_s + \frac{1}{2}\int_0^T |\boldsymbol{h}(\mathcal{X}_s)|^2\, ds \right\}, \tag{30}$$

i.e. the measure under which the process $\boldsymbol{Y}$ is a Wiener process, independent of the state process $\boldsymbol{\mathcal{X}}$, then the $\mathcal{P}(S)$-valued processes $(V^n(t, \boldsymbol{Y}))_{t\in[0,T]}$ converge in P^0-probability to the $\mathcal{P}(S)$-valued process $(V(t, \boldsymbol{Y}))_{t\in[0,T]}$. In addition, since the measure P is also absolutely continuous w.r.t. P^0, the convergence also holds in P-probability. This implies that

$$\tilde{\boldsymbol{\pi}}^{\boldsymbol{n}} \text{ converges in } P\text{-probability to } \boldsymbol{\pi}$$

which is the second assertion in Theorem 2.2.

Since the filter $\boldsymbol{\pi}$ is continuous in time, the last statement of the Theorem is an immediate consequence of the convergence in probability of $\tilde{\boldsymbol{\pi}}^{\boldsymbol{n}}$ to $\boldsymbol{\pi}$.

4 Rate of convergence

The aim of this section is to compute an upper bound for the rate of convergence of our scheme, i.e. to prove Theorem 2.3.

Let $(\boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{X}^n}, \boldsymbol{Y}, \boldsymbol{Y^n})$ be the stochastic processes introduced at the beginning of Section 3, with values in a complete separable metric space (S, d_S), and let P^n be the probability measure defined by

$$\frac{dP^n}{dP^0} = \exp\left\{ \int_0^T \boldsymbol{h}^n(\mathcal{X}_s^n) dY_s - \frac{1}{2}\int_0^T |\boldsymbol{h}^n(\mathcal{X}_s^n)|^2\, ds \right\}, \tag{31}$$

where P^0 is the reference probability measure on $(\Omega, \mathcal{F})$ defined in (30).

Then the law of $(\boldsymbol{\mathcal{X}^n}, \boldsymbol{Y})$ under P^n is the same as the law of $(\boldsymbol{\mathcal{X}^n}, \boldsymbol{Y^n})$ under P, so that the processes $(\boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{X}^n}, \boldsymbol{Y})$ and the probabilities P^0, P and P^n satisfy conditions ***(a)***, ***(a_n)***, ***(b1)*** and ***(b2)*** of Calzolari, Florchinger and Nappo (2005) [6], apart from the fact that we are in a complete separable metric space (S, d_S). Therefore, with slight modifications in the proof of (32) in Theorem 2.3 of [6], we get

$$E\big[d_{BL}^S\big(V(t,\boldsymbol{Y}), V^n(t,\boldsymbol{Y})\big)\big] \leq 2E^0\Big[\big|(dP^n/dP^0)|_{\widetilde{\mathcal{F}}_t} - (dP/dP^0)|_{\widetilde{\mathcal{F}}_t}\big|\Big] + E\big[d_S\big(\mathcal{X}_t, \mathcal{X}_t^n\big)\big], \tag{32}$$

where V and V^n are the functionals defined as in (29), d_{BL}^S is the bounded Lipschitz metric on $\mathcal{P}(S)$ and $\widetilde{\mathcal{F}}_t = \mathcal{F}_t^{\mathcal{X},\mathcal{X}^n,Y}$.

The above inequality is the starting point of the proof of Theorem 2.3 and, as a consequence, we need the estimates for the quantities in the right hand side of (32) stated in the following Proposition.

Proposition 4.1. *For all $t \leq T$, we have*

$$\begin{aligned} &E^0\Big[\big|(dP^n/dP^0)|_{\widetilde{\mathcal{F}}_t} - (dP/dP^0)|_{\widetilde{\mathcal{F}}_t}\big|\Big] \\ &\leq 2\left(E\Big[\int_0^t \big[\boldsymbol{h}^n(\mathcal{X}_s^n) - \boldsymbol{h}(\mathcal{X}_s)\big]^2 ds\Big]\right)^{\frac{1}{2}} + E\Big[\int_0^t \big[\boldsymbol{h}^n(\mathcal{X}_s^n) - \boldsymbol{h}(\mathcal{X}_s)\big]^2 ds\Big]. \end{aligned} \tag{33}$$

In the particular case when $\boldsymbol{h}^n = \boldsymbol{h}$ and $\boldsymbol{h}$ is a globally Lipschitz function then, for all $t \leq T$, we have, for a suitable constant $K(T)$,

$$\begin{aligned} &E^0\Big[\big|(dP^n/dP^0)|_{\widetilde{\mathcal{F}}_t} - (dP/dP^0)|_{\widetilde{\mathcal{F}}_t}\big|\Big] \\ &\leq 2\,K(T)\left\{\sup_{s\in[0,T]}\big(E\big[d_S^2(\mathcal{X}_s^n, \mathcal{X}_s)\big]\big)^{\frac{1}{2}} + \sup_{s\in[0,T]} E\big[d_S^2(\mathcal{X}_s^n, \mathcal{X}_s)\big]\right\}. \end{aligned} \tag{34}$$

Proof Define Λ_t and Λ_t^n by

$$\Lambda_t = \int_0^t h(\mathcal{X}_s)dY_s - \frac{1}{2}\int_0^t |h(\mathcal{X}_s)|^2\, ds \tag{35}$$

and

$$\Lambda_t^n = \int_0^t h^n(\mathcal{X}_s^n)dY_s - \frac{1}{2}\int_0^t |h^n(\mathcal{X}_s^n)|^2\, ds. \tag{36}$$

Then, using the fact that $|e^a - e^b| \leq e^a\,|a-b| + e^b\,|a-b|$, we have

$$\begin{aligned} E^0\Big[\big|(dP^n/dP^0)|_{\widetilde{\mathcal{F}}_t} - (dP/dP^0)|_{\widetilde{\mathcal{F}}_t}\big|\Big] &= E^0\Big[\big|e^{\Lambda_t^n} - e^{\Lambda_t}\big|\Big] \\ &\leq E^0\Big[e^{\Lambda_t^n}\big|\Lambda_t^n - \Lambda_t\big|\Big] + E^0\Big[e^{\Lambda_t}\big|\Lambda_t^n - \Lambda_t\big|\Big] \\ &= E^n\Big[\big|\Lambda_t^n - \Lambda_t\big|\Big] + E\Big[\big|\Lambda_t^n - \Lambda_t\big|\Big], \end{aligned}$$

where E^n is the expectation with respect to P^n.

An easy calculation gives

$$\begin{aligned} \Lambda_t^n - \Lambda_t &= \int_0^t \big[h^n(\mathcal{X}_s^n) - h(\mathcal{X}_s)\big]dY_s - \frac{1}{2}\int_0^t \big[|h^n(\mathcal{X}_s^n)|^2 - |h(\mathcal{X}_s)|^2\big]\, ds \\ &= \int_0^t \big[h^n(\mathcal{X}_s^n) - h(\mathcal{X}_s)\big]\big(dY_s - h^n(\mathcal{X}_s^n)ds\big) + \frac{1}{2}\int_0^t \big[h^n(\mathcal{X}_s^n) - h(\mathcal{X}_s)\big]^2\, ds \\ &= \int_0^t \big[h^n(\mathcal{X}_s^n) - h(\mathcal{X}_s)\big]\big(dY_s - h(\mathcal{X}_s)ds\big) + \frac{1}{2}\int_0^t \big[h^n(\mathcal{X}_s^n) - h(\mathcal{X}_s)\big]^2\, ds. \end{aligned}$$

Therefore we have

$$
\begin{aligned}
&E^0\Big[\big|(dP^n/dP^0)|_{\widetilde{\mathcal{F}}_t}-(dP/dP^0)|_{\widetilde{\mathcal{F}}_t}\big|\Big]\le E^n\Big[\big|\Lambda_t^n-\Lambda_t\big|\Big]+E\Big[\big|\Lambda_t^n-\Lambda_t\big|\Big]\\
&\le E^n\Big[\Big|\int_0^t\big[h^n(\mathcal{X}_s^n)-h(\mathcal{X}_s)\big]\big(dY_s-h^n(\mathcal{X}_s^n)ds\big)\Big|\Big]+\frac{1}{2}E^n\Big[\int_0^t\big[h^n(\mathcal{X}_s^n)-h(\mathcal{X}_s)\big]^2ds\Big]\\
&+E\Big[\Big|\int_0^t\big[h^n(\mathcal{X}_s^n)-h(\mathcal{X}_s)\big]\big(dY_s-h(\mathcal{X}_s)ds\big)\Big|\Big]+\frac{1}{2}E\Big[\int_0^t\big[h^n(\mathcal{X}_s^n)-h(\mathcal{X}_s)\big]^2ds\Big].
\end{aligned}
$$

By recalling that

$$Y_t-\int_0^t h^n(\mathcal{X}_s^n)ds \quad\text{and}\quad Y_t-\int_0^t h(\mathcal{X}_s)ds$$

are Wiener processes under P^n and P respectively, we get, by Cauchy-Schwarz inequality, and the isometry of stochastic integrals,

$$
\begin{aligned}
&E^0\Big[\big|(dP^n/dP^0)|_{\widetilde{\mathcal{F}}_t}-(dP/dP^0)|_{\widetilde{\mathcal{F}}_t}\big|\Big]\\
&\le\left(E^n\Big[\int_0^t\big[h^n(\mathcal{X}_s^n)-h(\mathcal{X}_s)\big]^2ds\Big]\right)^{\frac{1}{2}}+\frac{1}{2}E^n\Big[\int_0^t\big[h^n(\mathcal{X}_s^n)-h(\mathcal{X}_s)\big]^2ds\Big]\\
&+\left(E\Big[\int_0^t\big[h^n(\mathcal{X}_s^n)-h(\mathcal{X}_s)\big]^2ds\Big]\right)^{\frac{1}{2}}+\frac{1}{2}E\Big[\int_0^t\big[h^n(\mathcal{X}_s^n)-h(\mathcal{X}_s)\big]^2ds\Big].
\end{aligned}
$$

As the joint laws of $\boldsymbol{\mathcal{X}}$ and $\boldsymbol{\mathcal{X}^n}$ under P^n and P coincides (with the joint law under P^0) the final upper bound is

$$
\begin{aligned}
&E^0\Big[\big|(dP^n/dP^0)|_{\widetilde{\mathcal{F}}_t}-(dP/dP^0)|_{\widetilde{\mathcal{F}}_t}\big|\Big]\\
&\le 2\left(E\Big[\int_0^T\big[h^n(\mathcal{X}_s^n)-h(\mathcal{X}_s)\big]^2ds\Big]\right)^{\frac{1}{2}}+E\Big[\int_0^T\big[h^n(\mathcal{X}_s^n)-h(\mathcal{X}_s)\big]^2ds\Big],
\end{aligned}
$$

which is inequality (33), and immediately implies inequality (34).

□

As in the previous section, we take $S=[0,T]\times C([-\tau,0],\mathbb{R})$, $\mathcal{X}_t=(t,\Pi_tX)$, $\mathcal{X}_t^n=(\delta\cdot\lfloor t/\delta\rfloor,\Pi_{\delta\cdot\lfloor t/\delta\rfloor}X^n)$ and $\boldsymbol{h}^n=\boldsymbol{h}=h$. Therefore (32) implies that

$$E\big[d_{BL}(\pi_t,\tilde{\pi}_t^n)\big]\le 2E^0\Big[\big|(dP^n/dP^0)|_{\widetilde{\mathcal{F}}_t}-(dP/dP^0)|_{\widetilde{\mathcal{F}}_t}\big|\Big]+E\big[\|\mathcal{X}_t-\mathcal{X}_t^n\|_S\big], \tag{37}$$

and, when h is a jointly globally Lipschitz function, inequality (34) holds. Finally the result of Theorem 2.3 is a direct consequence of the following improvement of the result of Proposition 3.2.

Proposition 4.2. *Assume that conditions* **(A1)** *and* **(A3)** *are satisfied, and furthermore that the initial condition η satisfies* (22)*, and that the functions a and b are bounded. Then there exists a constant C_X such that*

$$E\left[\sup_{t\in[0,T]}\big\|(\delta\cdot\lfloor t/\delta\rfloor,\Pi_{\delta\cdot\lfloor t/\delta\rfloor}X^n)-(t,\Pi_tX)\big\|_S^2\right]\le C_X\frac{\log n}{n}. \tag{38}$$

The proof of this result will be given in Section 5.

Remark 4.3. *Obviously, from* (38) *there exists a constant* $C'_X \leq C_X$ *such that*

$$\sup_{t\in[0,T]} E\left[\left\|(\delta\cdot\lfloor t/\delta\rfloor, \Pi_{\delta\cdot\lfloor t/\delta\rfloor}X^n) - (t, \Pi_t X)\right\|_S^2\right] \leq C'_X \frac{\log n}{n} \tag{39}$$

and, by (34)*, the above inequality is a sufficient condition to get the upper bound* (23) *in Theorem 2.3. From the following example it appears that the rate of convergence cannot be improved neither in the r.h.s. of the previous inequality nor in* (23)*.*

Example Take $\eta = 0$, $a = 0$, $b = 1$ and $h = 0$. In this case $X = \tilde{W}$ and $X^n = \tilde{W}^n$ is the piecewise linear interpolation of $\tilde{W}$. Moreover π_t and $\tilde{\pi}_t^n$ coincide with the law of $\Pi_t \tilde{W}$ and $\Pi_{\delta\cdot\lfloor t/\delta\rfloor}\tilde{W}^n$ respectively, and therefore

$$E\big[d_{BL}(\pi_t, \tilde{\pi}_t^n)\big] = d_{BL}(\pi_t, \tilde{\pi}_t^n) \leq E\left[\left\|\Pi_{\delta\cdot\lfloor t/\delta\rfloor}\tilde{W}^n - \Pi_t\tilde{W})\right\|\right] \leq \left(E\left[\left\|\Pi_{\delta\cdot\lfloor t/\delta\rfloor}\tilde{W}^n - \Pi_t\tilde{W})\right\|^2\right]\right)^{\frac{1}{2}},$$

where the first inequality follows by standard coupling techniques.

Furthermore

$$E\left[\left\|\Pi_{\delta\cdot\lfloor t/\delta\rfloor}\tilde{W}^n - \Pi_t\tilde{W})\right\|^2\right] = O(\tfrac{\log n}{n}),$$

for any $t \in [0, T]$, and uniformly in $[0, T]$. This fact can be shown by using the results established by Pickands (1968) in [30] (see also Nappo (2005) [26]). This result could be expected thanks to Lévy's modulus of continuity, which implies that there exists a finite random variable M such that

$$\sup_{\substack{s,t\in[0,1]\\ |s-t|\leq\delta,}} |\tilde{W}_s - \tilde{W}_t| \leq M\sqrt{\delta\log(1/\delta)}$$

holds a.s. (see the paper by Pinsky (2001) [31] for a simple proof).

5 Technical results

This section is devoted to the proofs of Lemma 3.3, Proposition 3.2 and Proposition 4.2. In order to prove these results we introduce, as a technical tool,

- the operator P^δ which gives the linear interpolation of a function $\big(x(s)\big)_{s\in[-\tau,T]}$, with step δ, so that $\big(P^\delta x(s)\big)_{s\in[-\tau,T]}$ is the linear interpolation of $(\ell\delta, x(\ell\delta))$, for $\ell = -m, \ldots, n$,

and, as another approximation for the state,

- the continuous Euler-Maruyama scheme, i.e. the diffusion processes $\boldsymbol{Z^n} = \big(Z^n(t)\big)_{t\in[0,T]}$ where

$$\begin{cases} Z^n(t) := \eta(t) & -\tau \leq t \leq 0, \\ Z^n(t) := \eta(0) + \displaystyle\int_0^t a(\delta\cdot\lfloor s/\delta\rfloor, \Pi_{\delta\cdot\lfloor s/\delta\rfloor}X^n)ds & \\ \qquad\qquad + \displaystyle\int_0^t b(\delta\cdot\lfloor s/\delta\rfloor, \Pi_{\delta\cdot\lfloor s/\delta\rfloor}X^n)d\tilde{W}_s, & 0 \leq t \leq T, \end{cases} \tag{40}$$

which can be considered as intermediate approximation processes for the state $\mathbf{X}$.

The processes $\boldsymbol{Z^n}$ have the property that

$$P^\delta Z^n(s) = X^n(s), \qquad \text{for } s \in [-\tau, 0] \cup [0, T], \tag{41}$$

since

$$Z^n(\ell\delta) = X^n(\ell\delta), \qquad \text{for } \ell \geq -m.$$

Indeed $Z^n(\ell\delta) = X^n(\ell\delta) = \eta(\ell\delta)$, for $-m \leq \ell \leq 0$. The case $\ell \geq 0$ follows by observing that for $t \in [\ell\delta, (\ell+1)\delta]$

$$\begin{aligned} Z^n(t) &= Z^n(\ell\delta) + \int_{\ell\delta}^t a(\ell\delta, \Pi_{\ell\delta}X^n)ds + \int_{\ell\delta}^t b(\ell\delta, \Pi_{\ell\delta}X^n)d\tilde{W}_s \\ &= Z^n(\ell\delta) + a(\ell\delta, \Pi_{\ell\delta}X^n)(t-\ell\delta) + b(\ell\delta, \Pi_{\ell\delta}X^n)[\tilde{W}(t) - \tilde{W}(\ell\delta)], \end{aligned}$$

so that

$$Z^n((\ell+1)\delta) = Z^n(\ell\delta) + a(\ell\delta, \Pi_{\ell\delta}X^n)\delta + b(\ell\delta, \Pi_{\ell\delta}X^n)[\tilde{W}((\ell+1)\delta) - \tilde{W}(\ell\delta)],$$

and finally comparing the above recursive formula with the definition of $X^n((\ell+1)\delta)$ in (12).

We are now able to prove Lemma 3.3.

Proof of Lemma 3.3 First observe that by (41)

$$\sup_{t\in[-\tau,T]} |X^n(t)| \leq \max\left(\|\eta\|, \sup_{t\in[0,T]} |Z^n(t)|\right),$$

and set

$$M_t^n := \int_0^t b_n(u)d\tilde{W}_u \qquad \text{where } b_n(u) := b(\delta\cdot\lfloor u/\delta\rfloor, \Pi_{\delta\cdot\lfloor u/\delta\rfloor}X^n). \tag{42}$$

For any $\ell \in \{1,2\}$ take $\alpha = 2\ell$ and $\beta = 2\ell/(2\ell-1)$, so that $(1/\alpha)+(1/\beta) = 1$. Then for any stopping time σ, there exists a suitable constant C_ℓ such that

$$\begin{aligned} &\sup_{u\in[-\tau,t\wedge\sigma]} |X^n(u)|^{2\ell} \\ &\leq C_\ell\left\{\|\eta\|^{2\ell} + \left(\int_0^t \sup_{u\in[0,s\wedge\sigma]} |a(\delta\cdot\lfloor u/\delta\rfloor, \Pi_{\delta\cdot\lfloor u/\delta\rfloor}X^n)|ds\right)^{2\ell} + \sup_{s\in[0,t\wedge\sigma]} |M_s^n|^{2\ell}\right\} \\ &\leq C_\ell\left\{\|\eta\|^{2\ell} + t^{2\ell/\beta}\left(\int_0^t \sup_{u\in[0,s\wedge\sigma]} |a(\delta\cdot\lfloor u/\delta\rfloor, \Pi_{\delta\cdot\lfloor u/\delta\rfloor}X^n)|^\alpha ds\right)^{2\ell/\alpha} + \sup_{s\in[0,t\wedge\sigma]} |M_s^n|^{2\ell}\right\} \\ &\leq C_\ell\left\{\|\eta\|^{2\ell} + t^{2\ell-1}\int_0^t \ell K^\ell\left(1 + \sup_{u\in[-\tau,s\wedge\sigma]} |X^n(u)|^{2\ell}\right)ds + \sup_{s\in[0,t]} |M_{s\wedge\sigma}^n|^{2\ell}\right\}. \end{aligned}$$

If $M_{t\wedge\sigma}^n$ is a martingale, setting

$$\phi_{\sigma,\ell}^n(t) := E\left[\sup_{u\in[-\tau,t\wedge\sigma]} |X^n(u)|^{2\ell}\right],$$

and applying Doob's inequality for $p = 2\ell$ to $M_{t\wedge\sigma}^n$ yields

$$\phi_{\sigma,\ell}^n(t) \leq C'_\ell\left\{1 + \|\eta\|^{2\ell} + \int_0^t \phi_{\sigma,\ell}^n(s)ds + E\left[|M_{t\wedge\sigma}^n|^{2\ell}\right]\right\}, \tag{43}$$

for all $t \in [0,T]$, for a suitable constant $C'_\ell = C'_\ell(T)$.
Taking $\sigma = \sigma_N^n := \inf\{s > 0;\ \sup_{u\in[-\tau,s]} |X^n(u)| \geq N\}$, we have

$$\int_0^T E[\mathbf{1}_{s\leq\sigma_N^n} b_n^{2\ell}(s)]ds < \infty, \tag{44}$$

and $M^n_{t\wedge\sigma} = M^n_{t\wedge\sigma^n_N}$ is a martingale. Indeed, by the sublinearity condition (8) on b we have

$$\int_0^t E[\mathbf{1}_{s\le\sigma^n_N} b_n^{2\ell}(s)]ds \le \int_0^t E[\ell K^\ell(1 + \sup_{u\in[-\tau,s\wedge\sigma^n_N]} |X^n(u)|^{2\ell})]ds, \tag{45}$$

and the r.h.s. of the previous inequality is finite since $\sup_{u\in[-\tau,s\wedge\sigma^n_N]} |X^n(u)| \le N$. Then

$$E[|M^n_{t\wedge\sigma^n_N}|^2] = E\left[\left(\int_0^t \mathbf{1}_{s\le\sigma^n_N} b_n(s)d\tilde{W}_s\right)^2\right] = \int_0^t E[\mathbf{1}_{s\le\sigma^n_N} b_n^2(s)]ds,$$

and (see e.g. Lemma 4.12, page 125, in Liptser and Shiryayev [22])

$$E[|M^n_{t\wedge\sigma}|^4] = E\left[\left(\int_0^t \mathbf{1}_{s\le\sigma^n_N} b_n(s)d\tilde{W}_s\right)^4\right] \le 6^2 t \int_0^t E[\mathbf{1}_{s\le\sigma^n_N} b_n^4(s)]ds.$$

Then, taking into account (43) and (45), and invoking Gronwall's inequality we get a bound for $\phi^n_{\sigma,\ell}(T) = \phi^n_{\sigma^n_N,\ell}(T)$, uniform in n and N. Therefore, applying Fatou's Lemma and making use of the fact that $\sigma^n_N \to \infty$ as $N \to \infty$, we get the theses.

□

Remark 5.1. *Note that with the same technique one could prove that under the same assumptions of Lemma 3.3*

$$\sup_n E\left[\sup_{t\in[0,T]} \|\Pi_t Z^n\|^{2\ell}\right] = \sup_n E\left[\sup_{u\in[-\tau,T]} |Z^n(u)|^{2\ell}\right] < \infty, \qquad \textit{for } \ell = 1, 2.$$

In order to prove Proposition 3.2 we need some intermediate results, stated in the following Lemmas, which will be proved at the end of this section.

The first Lemma concerns the behaviour of the modulus of continuity.

Lemma 5.2. *Denoting by*

$$\omega_x(\delta; [-\tau, T]) := \sup_{\substack{s,t\in[-\tau,T] \\ |s-t|\le\delta,}} |x(s) - x(t)|$$

the modulus of continuity of the function $\big(x(s)\big)_{s\in[-\tau,T]}$, *we have*

$$\sup_{t\in[0,T]} \|\Pi_{\delta\cdot\lfloor t/\delta\rfloor} P^\delta x - \Pi_t x\| \le 2\,\omega_x(\delta; [-\tau, T]), \tag{46}$$

and, for $\delta = \delta_n$,

$$\lim_{n\to\infty} E\left[\omega_X^2(\delta, [-\tau, T])\right] = 0, \tag{47}$$

and

$$\lim_{n\to\infty} E\left[\sup_{t\in[0,T]} \|\Pi_{\delta\cdot\lfloor t/\delta\rfloor} P^\delta X - \Pi_t X\|^2\right] = 0. \tag{48}$$

The second Lemma concerns the convergence of the approximation $\boldsymbol{Z^n}$.

Lemma 5.3. *Under the hypotheses of Proposition 3.2*

$$\lim_{n\to\infty} E\left[\sup_{t\in[0,T]} \|\Pi_t Z^n - \Pi_t X\|^2\right] = 0.$$

With the above results the proof of Proposition 3.2 is straightforward.

Proof of Proposition 3.2 First of all we note that

$$\left\|(\delta\cdot\lfloor t/\delta\rfloor, \Pi_{\delta\cdot\lfloor t/\delta\rfloor}X^n) - (t, \Pi_t X)\right\|_S^2 \leq 2\,\delta^2 + 2\,\left\|\Pi_{\delta\cdot\lfloor t/\delta\rfloor}X^n - \Pi_t X\right\|^2. \tag{49}$$

Then, by adding and subtracting $\Pi_{\delta\cdot\lfloor t/\delta\rfloor}P^\delta X$ in the second term on the right hand side of the above expression, it yields

$$\left\|\Pi_{\delta\cdot\lfloor t/\delta\rfloor}X^n - \Pi_t X\right\|^2 \leq 2\,\left\|\Pi_{\delta\cdot\lfloor t/\delta\rfloor}X^n - \Pi_{\delta\cdot\lfloor t/\delta\rfloor}P^\delta X\right\|^2 + 2\,\left\|\Pi_{\delta\cdot\lfloor t/\delta\rfloor}P^\delta X - \Pi_t X\right\|^2. \tag{50}$$

Then, taking into account (48), and that

$$\begin{aligned}
&\sup_{t\in[0,T]}\|\Pi_t X^n - \Pi_t P^\delta X\| = \sup_{k:k\delta\in[-\tau,T]}|X^n(k\delta) - X(k\delta)| \\
&= \sup_{k:k\delta\in[-\tau,T]}|Z^n(k\delta) - X(k\delta)| \leq \sup_{t\in[-\tau,T]}|Z^n(t) - X(t)| = \sup_{t\in[0,T]}\|\Pi_t Z^n - \Pi_t X\|
\end{aligned}$$

the result follows by Lemma 5.3.

□

Proof of Lemma 5.2 Noticing that $\Pi_{\delta\cdot\lfloor t/\delta\rfloor}P^\delta x = P^\delta\Pi_{\delta\cdot\lfloor t/\delta\rfloor}x$, we have

$$\begin{aligned}
\|\Pi_{\delta\cdot\lfloor u/\delta\rfloor}P^\delta x - \Pi_u x\| &\leq \|\Pi_{\delta\cdot\lfloor u/\delta\rfloor}P^\delta x - \Pi_{\delta\cdot\lfloor u/\delta\rfloor}x\| + \|\Pi_{\delta\cdot\lfloor u/\delta\rfloor}x - \Pi_u x\| \\
&= \|P^\delta\Pi_{\delta\cdot\lfloor u/\delta\rfloor}x - \Pi_{\delta\cdot\lfloor u/\delta\rfloor}x\| + \|\Pi_{\delta\cdot\lfloor u/\delta\rfloor}x - \Pi_u x\|.
\end{aligned}$$

Furthermore, since

$$P^\delta\theta(v) = \lambda(v)\,\theta(\delta\cdot\lfloor v/\delta\rfloor + \delta) + (1-\lambda(v))\,\theta(\delta\cdot\lfloor v/\delta\rfloor)$$

with $\lambda(v) = v/\delta - \lfloor v/\delta\rfloor$, and since $\theta(v) = \lambda(v)\,\theta(v) + (1-\lambda(v))\,\theta(v)$, we deduce that

$$\|\Pi_{\delta\cdot\lfloor u/\delta\rfloor}P^\delta x - \Pi_u x\| \leq 2\sup_{\substack{s,t\in[0,T]\\|s-t|\leq\delta,}}\|\Pi_s x - \Pi_t x\| = 2\sup_{\substack{s,t\in[-\tau,T]\\|s-t|\leq\delta,}}|x(s) - x(t)|,$$

which is the first assertion (46) of the Lemma.

Equality (47) follows by the dominated convergence theorem: indeed

$$\omega_x(\delta;[-\tau,T]) \leq 2\sup_{t\in[-\tau,T]}|x(t)|,$$

the integrability condition (9) holds, and finally the modulus of continuity $\omega_X(\delta,[-\tau,T])$ converge to zero as $\delta = \delta_n$ converge to zero, as the paths of X are continuous.
The last assertion (48) is an interesting observation which is a straightforward consequence of (46) and (47).

□

Proof of Lemma 5.3 Noticing that $P^\delta Z^n(s) = X^n(s)$, for $s\in[-\tau,0]\cup[0,T]$, then we can rewrite (40) as

$$Z^n(t) = \eta(0) + \int_0^t a(\delta\cdot\lfloor s/\delta\rfloor, \Pi_{\delta\cdot\lfloor s/\delta\rfloor}P^\delta Z^n)ds + \int_0^t b(\delta\cdot\lfloor s/\delta\rfloor, \Pi_{\delta\cdot\lfloor s/\delta\rfloor}P^\delta Z^n)d\tilde{W}_s.$$

Therefore, taking into account that $Z^n(t) = X(t) = \eta(t)$ for $t \in [-\tau, 0]$,

$$
\begin{aligned}
&\sup_{s\in[-\tau,t]} |Z^n(s) - X(s)|^2 \\
&\leq 2\left(\int_0^t |a(\delta\cdot\lfloor u/\delta\rfloor, \Pi_{\delta\cdot\lfloor u/\delta\rfloor}P^\delta Z^n) - a(u,\Pi_u X)|du\right)^2 \\
&+ 2\sup_{s\in[0,t]}\left(\int_0^s \left[b(\delta\cdot\lfloor u/\delta\rfloor, \Pi_{\delta\cdot\lfloor u/\delta\rfloor}P^\delta Z^n) - b(u,\Pi_u X)\right] d\tilde{W}_u\right)^2 \\
&\leq 2t\int_0^t |a(\delta\cdot\lfloor u/\delta\rfloor, \Pi_{\delta\cdot\lfloor u/\delta\rfloor}P^\delta Z^n) - a(u,\Pi_u X)|^2 du \\
&+ 2\sup_{s\in[0,t]}\left(\int_0^s \left[b(\delta\cdot\lfloor u/\delta\rfloor, \Pi_{\delta\cdot\lfloor u/\delta\rfloor}P^\delta Z^n) - b(u,\Pi_u X)\right] d\tilde{W}_u\right)^2
\end{aligned}
$$

By Lemma 3.3, Remark 5.1 and the sublinearity of b,

$$
\int_0^s \left[b(\delta\cdot\lfloor u/\delta\rfloor, \Pi_{\delta\cdot\lfloor u/\delta\rfloor}P^\delta Z^n) - b(u,\Pi_u X)\right] d\tilde{W}_u
$$

is a martingale.

Then taking the expectations we can apply Doob's inequality, and we get for $t \in [0,T]$

$$
\begin{aligned}
&E\left[\sup_{u\in[0,t]} \|\Pi_u Z^n - \Pi_u X\|^2\right] = E\left[\sup_{s\in[-\tau,t]} |Z^n(s) - X(s)|^2\right] \\
&\leq 2t\int_0^t E\left[|a(\delta\cdot\lfloor u/\delta\rfloor, \Pi_{\delta\cdot\lfloor u/\delta\rfloor}P^\delta Z^n) - a(u,\Pi_u X)|^2\right] du \\
&+ 8\int_0^t E\left[\,|b(\delta\cdot\lfloor u/\delta\rfloor, \Pi_{\delta\cdot\lfloor u/\delta\rfloor}P^\delta Z^n) - b(u,\Pi_u X)|^2\,\right] du \\
&\leq \max(2T, 8)\int_0^t KE\left[\sup_{u\in[0,s]} \|\Pi_{\delta\cdot\lfloor u/\delta\rfloor}P^\delta Z^n - \Pi_u X\|^2 + \delta^2\right] ds \\
&\leq C(T)\int_0^t E\left[\sup_{u\in[0,s]} \left(\|\Pi_{\delta\cdot\lfloor u/\delta\rfloor}P^\delta Z^n - \Pi_{\delta\cdot\lfloor u/\delta\rfloor}P^\delta X\|^2 + \|\Pi_{\delta\cdot\lfloor u/\delta\rfloor}P^\delta X - \Pi_u X\|^2\right) + \delta^2\right] ds \\
&\leq C(T)\int_0^t E\left[\sup_{u\in[0,s]} \|\Pi_u Z^n - \Pi_u X\|^2 + 4\omega_X^2(\delta; [-\tau, T]) + \delta^2\right] ds,
\end{aligned}
$$

where we have used (46). Then Gronwall's inequality gives the upper bound

$$
E\left[\sup_{u\in[0,T]} \|\Pi_u Z^n - \Pi_u X\|^2\right] \leq C_1(T)\left(E\left[\omega_X^2(\delta; [-\tau, T])\right] + \delta^2\right) \tag{51}
$$

and the proof is accomplished, since $\delta = \delta_n$ and (according to (47)) $E\left[\omega_X^2(\delta; [-\tau, T])\right]$ go to zero as n goes to infinity.

□

We conclude this section with the proof of Proposition 4.2.

Proof of Proposition 4.2 First of all we note that the inequalities in the proof of Proposition 3.2 together with (46) imply

$$
\left\|(\delta\cdot\lfloor t/\delta\rfloor, \Pi_{\delta\cdot\lfloor t/\delta\rfloor}X^n) - (t, \Pi_t X)\right\|_S^2 \leq 4\left(\delta^2 + \sup_{t\in[0,T]} \|\Pi_t Z^n - \Pi_t X\|^2 + \omega_X^2(\delta; [-\tau, T])\right).
$$

Then by (51) we get

$$E\left[\left\|(\delta\cdot\lfloor t/\delta\rfloor,\Pi_{\delta\cdot\lfloor t/\delta\rfloor}X^n)-(t,\Pi_tX)\right\|_S^2\right]\le C_2(T)\ \left(\delta^2+E\left[\omega_X^2(\delta;[-\tau,T])\right]\right).$$

The thesis follows by observing that

$$\omega_X(\delta;[-\tau,T])\le\omega_\eta(\delta;[-\tau,0])+\omega_X(\delta;[0,T]) \tag{52}$$

and taking into account the result of Lemma A.4 of Słomiǹski (2001) [33], which is recalled in the sequel for the ease of the reader.

□

Lemma 5.4 (Lemma A.4 of Słomiǹski [33]). *Let H, G be two adapted processes with values in $\mathbb{R}^d\otimes\mathbb{R}^d$ and $\mathbb{R}^d$ respectively, such that $\|H_t\|_{\mathbb{R}^d\otimes\mathbb{R}^d}$, $|G_t|\le L<\infty$, for some constant $L>0$, and let $Y_t=\int_0^t H_s\,dW_s+\int_0^t G_s\,ds$, $t\in\mathbb{R}^+$. Then for every $p\in\mathbb{N}$*

$$E\left[\omega_Y^{2p}(\tfrac{1}{n};[0,T])\right]=O\left(\left(\tfrac{\log n}{n}\right)^p\right). \tag{53}$$

Remark 5.5. *We observe that the thesis holds for any real $p>0$, as can be seen following the lines in the proof by Słomiǹski.*

6 Conclusion

As we have already recalled in the Introduction, Chang in [7] gives a computable approximation for the optimal filter associated with the partially observable delay system (1) and (2), with $b(t,\theta)=1$. The state process is approximated by the linear interpolation of $X^n(\ell\delta)$ as in (12) and the approximation for the observation process is the linear interpolation of $Y^n(\ell\delta)$ defined by (13), while our approximation of the observation process is a continuous time diffusion. However the two approximation processes coincide at times $\ell\delta$.

The author proves weak convergence of the filters under the assumption that there exists a strictly positive constant k such that

$$E[\exp\{k\|\eta\|^2\}]<\infty,$$

and for any partition $-\tau\le\tau_0<\cdots<\tau_n=0$ the $(n+1)-$dimensional random vector $(\eta(\tau_i);0\le i\le n)$ has a density w.r.t. the Lebesgue measure in $\mathbb{R}^{n+1}$. There are other minor differences between our assumptions and the one by Chang, about the diffusion coefficients, which allow to consider coefficient of the form (4) but not of the form (3).

The problem of strong approximation for stochastic delay differential equations has been the subject of research for many authors in the last past years. There is a quite substantial work in this field and in the following we mention only some of them. Küchler and Platen (2000) [16] have proposed a Taylor approximation scheme, besides the Euler scheme, and have proved the strong convergence of their scheme (see also Baker and Buckwar (2000) [1] and Buckwar (2000) [5]). Hu, Mohammed and Yan (2004) [13], have studied strong convergence of Milstein schemes for stochastic delay differential equations with tame coefficient functions g_a and g_b as in (4), with $\gamma_i=\delta_{s_i}$, for $s_i\in[-\tau,0]$. Though these schemes have better performances than Euler schemes the authors do not deal with convergence of the expectation of the uniform norm on $[-\tau,T]$, which we need in order to get our first result. In [24] Mao and Sabanis (2003) have investigated the uniform norm for the continuous Euler scheme instead of the piecewise linear Euler scheme, in the framework of a variable delay, namely when in equation (5) the term $X(t-\tau)$ is replaced by $X(\delta(t))$, where $\delta(t)$ is a Lipschitz function with $-\tau\le\delta(t)\le t$, and when g_a and g_b are locally Lipschitz functions. Moreover Mao and Sabanis get, under suitable assumptions, a rate of convergence of order less or equal to $\sqrt{1/n}$.

References

[1] Baker, C. T. H., and Buckwar, E. Numerical analysis of explicit one-step methods for stochastic delay differential equations. *LMS J. Comput. Math. 3* (2000), 315–335 (electronic).

[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] Bhatt, A. G., Kallianpur, G., and Karandikar, R. L. Robustness of the nonlinear filter. *Stochastic Process. Appl. 81*, 2 (1999), 247–254.

[4] Bhatt, A. G., and Karandikar, R. L. Robustness of the nonlinear filter: the correlated case. *Stochastic Process. Appl. 97*, 1 (2002), 41–58.

[5] Buckwar, E. Introduction to the numerical analysis of stochastic delay differential equations. *J. Comput. Appl. Math. 125*, 1-2 (2000), 297–307. Numerical analysis 2000, Vol. VI, Ordinary differential equations and integral equations.

[6] Calzolari, A., Florchinger, P., and Nappo, G. Approximation of nonlinear filtering for Markov systems with delayed observations. *(to appear in SIAM J. Control Optim.)* (2005).

[7] Chang, M. H. Discrete approximation of nonlinear filtering for stochastic delay equations. *Stochastic Anal. Appl. 5*, 3 (1987), 267–298.

[8] Elsanousi, S. A., and Bjørnar, L. Optimal consumption under partial observations for a stochastic system with delay. Tech. Rep. 9, University of Oslo, 2001.

[9] Ferrante, L., Bompadre, S., Possati, L., and Leone, L. Parameter estimation in a Gompertzian stochastic model for tumor growth. *Biometrics 56*, 4 (2000), 1076–1081.

[10] Frank, T. D. Multivariate Markov processes for stochastic systems with delays: application to the stochastic Gompertz model with delay. *Phys. Rev. E (3) 66*, 1 (2002), 011914, 8.

[11] Frank, T. D., and Beek, P. J. Stationary solutions of linear stochastic delay differential equations: applications to biological system. *Phys. Rev. E (3) 64*, 2 (2001), 021917, 12.

[12] Goel, N. S., Maitra, S. C., and Montroll, E. W. On the Volterra and other nonlinear models of interacting populations. *Rev. Modern Phys. 43* (1971), 231–276.

[13] Hu, Y., Mohammed, S.-E. A., and Yan, F. Discrete-time approximations of stochastic delay equations: the Milstein scheme. *Ann. Probab. 32*, 1A (2004), 265–314.

[14] Kallianpur, G. *Stochastic filtering theory*, vol. 13 of *Applications of Mathematics*. Springer-Verlag, New York, 1980.

[15] Kallianpur, G. Mandal, P. K. Nonlinear filtering with stochastic delay equations. In *Advances on Theoretical and Methodological Aspects of Probability and Statistics; IISA 1998; Volume 2; Editor: Balakrishnan, N.* Gordon and Breach Science Publishers, 2002, pp. 3–36.

[16] Küchler, U., and Platen, E. Strong discrete time approximation of stochastic differential equations with time delay. *Math. Comput. Simulation 54*, 1-3 (2000), 189–205.

[17] Kurtz, T. G., and Ocone, D. L. Unique characterization of conditional distributions in nonlinear filtering. *Ann. Probab. 16*, 1 (1988), 80–107.

[18] Kushner, H. J. *Weak convergence methods and singularly perturbed stochastic control and filtering problems*, vol. 3 of *Systems & Control: Foundations & Applications*. Birkhäuser Boston Inc., Boston, MA, 1990.

[19] Kwong, R. H., and Willsky, A. S. Estimation and filter stability of stochastic delay systems. *SIAM J. Control Optim. 16*, 4 (1978), 660–681.

[20] Larssen, B., and Risebro, N. H. When are HJB-equations in stochastic control of delay systems finite dimensional? *Stochastic Anal. Appl. 21*, 3 (2003), 643–671.

[21] Le Gland, F. Time discretization of nonlinear filtering equations. In *Proceedings of the 28th IEEE Conference on Decision and Control, Vol. 1–3 (Tampa, FL, 1989)* (New York, 1989), IEEE, pp. 2601–2606.

[22] 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.

[23] Longtin, A., Milton, J., Bos, J., and Mackey, M. Noise and critical behavior of the pupil light reflex at oscillation onset. *Phys. Rev. A 41* (1990), 6992–7005.

[24] Mao, X., and Sabanis, S. Numerical solutions of stochastic differential delay equations under local Lipschitz condition. *J. Comput. Appl. Math. 151*, 1 (2003), 215–227.

[25] Mohammed, S. E. A. *Stochastic functional differential equations*, vol. 99 of *Research Notes in Mathematics*. Pitman (Advanced Publishing Program), Boston, MA, 1984.

[26] Nappo, G. The rate of L^p uniform convergence of Euler approximations for stochastic delay differential equations. (in preparation), 2005.

[27] Niebur, E., Schuster, H. G., and Kammen, D. Collective frequencies and metastability in networks of limit cycle oscillators with time delay. *Phys. Rev. Lett. 67*, 20 (1991), 2753–2756.

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

[29] Peterka, R. J. Postural control model interpretation of stabilogram diffusion analysis. *Biol. Cybernet. 82*, 4 (2000), 335–343.

[30] Pickands, III, J. Moment convergence of sample extremes. *Ann. Math. Statist. 39* (1968), 881–889.

[31] Pinsky, M. A. Brownian continuity modulus via series expansions. *J. Theoret. Probab. 14*, 1 (2001), 261–266.

[32] Ricciardi, L. M. Stochastic population theory: diffusion processes. In *Mathematical ecology (Miramare-Trieste, 1982)*, vol. 17 of *Biomathematics*. Springer, Berlin, 1986, pp. 191–238.

[33] Słomiński, L. Euler's approximations of solutions of SDEs with reflecting boundary. *Stochastic Process. Appl. 94*, 2 (2001), 317–337.

[34] Vasilakos, K., and Beuter, A. Effects of noise on a delayed visual feedback system. *J. Theor. Bio.*, 165 (1993), 389–407.