On Approximation of the Backward Stochastic Differential Equation

Yury A. Kutoyants, Li Zhou
Laboratoire de Statistique et Processus, Université du Maine.
72085 Le Mans Cedex, France

Abstract

We consider the problem of approximation of the solution of the backward stochastic differential equation in the Markovian case. We suppose that the trend coefficient of the diffusion process depends on some unknown parameter and the diffusion coefficient of this equation is small. We propose an approximation of this solution based on the one-step MLE of the unknown parameter and we show that this approximation is asymptotically efficient in the asymptotics of "small noise".

Keywords: Backward SDE, approximation of the solution, small noise asymptotics.

1 Introduction

We consider the following problem. We are given a stochastic differential equation (called *forward*)

$$\mathrm{d}X_t = b(t, X_t)\ \mathrm{d}t + a(t, X_t)\ \mathrm{d}W_t, \quad X_0 = x_0,\ 0 \leq t \leq T, \tag{1}$$

and two functions $f(t, x, y, z)$ and $\Phi(x)$. We have to construct a couple of processes (Y_t, Z_t) such that the solution of the equation

$$\mathrm{d}Y_t = -f(t, X_t, Y_t, Z_t)\ \mathrm{d}t + Z_t\ \mathrm{d}W_t,\ 0 \leq t \leq T, \tag{2}$$

(called *backward*) has the final value $Y_T = \Phi(X_T)$.

The existence and uniqueness of the solution of *backward stochastic differential equation* (BSDE) in essentially more general situations was studied by Pardoux and Peng [9]. The problem (1)-(2) considered here was introduced as forward-backward stochastic differential equations (FBSDE) in El Karoui & al. [2]. The solution of this FBSDEs is presented as a triple-process $(X_t, Y_t, Z_t)_{t\geq 0}$. It is shown that the solution $(X_t, Y_t, Z_t)_{t\geq 0}$ exists and is unique under the condition that all coefficients are Lipschitzian and so on (see [2] for details). The solution of the problem (1)-(2) proposed in [2] is the following. Suppose that $u(t,x)$ satisfies the equation

$$\frac{\partial u}{\partial t}+b(t,x)\frac{\partial u}{\partial x}+\frac{1}{2}a(t,x)^2\frac{\partial^2 u}{\partial x^2}=-f\left(t,x,u,a(t,x)\frac{\partial u}{\partial x}\right),\quad u(T,x)=\Phi(x) \tag{3}$$

and put $Y_t = u(t,X_t), Z_t = a(t,X_t)\,u'_x(t,X_t)$. Then by Itô's formula the process Y_t has the stochastic differential

$$\begin{aligned}\mathrm{d}Y_t &= \left[\frac{\partial u}{\partial t}(t,X_t)+b(t,X_t)\frac{\partial u}{\partial x}(t,X_t)+\frac{1}{2}a(t,x)^2\frac{\partial^2 u}{\partial x^2}(t,X_t)\right]\mathrm{d}t\\ &\qquad + a(t,X_t)\frac{\partial u}{\partial x}(t,X_t)\,\mathrm{d}W_t\\ &= -f(t,X_t,Y_t,Z_t)\,\mathrm{d}t+Z_t\,\mathrm{d}W_t,\qquad Y_0=u(0,X_0).\end{aligned}$$

The final value $Y_T = u(T,X_T) = \Phi(X_T)$. Therefore the problem is solved and the couple (Y_t, Z_t) provides the desired solution. More details can be found, e.g., in El Karoui & Mazliak [1] and Ma & Yong [8].

In the present work we consider the similar statement but in the situation when the trend coefficient $b(t,x)$ of the diffusion process (1) depends on the unknown parameter $\vartheta \in \Theta$, i.e., $b(t,x) = S(\vartheta,t,x)$. In this case the function $u(t,x) = u(t,x,\vartheta)$ satisfying the equation (3) depends on unknown parameter ϑ and we can not put $Y_t = u(t,X_t,\vartheta)$ because we do not know ϑ. We consider the problem of the construction of the couple $(\hat{Y}_t, \hat{Z}_t)$, where $\hat{Y}_t$ and $\hat{Z}_t$ are some approximations of (Y_t, Z_t). This approximation is done with the help of the one-step maximum likelihood estimator $\tilde{\vartheta}$ as $\hat{Y}_t = u(t,X_t,\tilde{\vartheta})$ and $\hat{Z}_t = a(t,X_t)\,u'_x(t,X_t,\tilde{\vartheta})$. We are interested by a situation when the error of this approximation is small. One of the possibilities to have a small error of approximations is in some sense equivalent to the situation with the small error of estimation of the parameter ϑ, then from the continuity of the function $u(t,x,\vartheta)$ w.r.t. ϑ, we obtain $\hat{Y}_T \sim Y_T = \Phi(X_T)$. The small error of estimation we can have, besides others, in the situations when $T\to\infty$ or when $a(\cdot)\to 0$ (see, e.g., Kutoyants [6] and [5]). We propose to study this model in the asymptotics of *small noise*, i.e. the diffusion coefficient $a(t,x)^2$ tends to 0. This allows us to keep the final time T fixed and, what is as well important, this asymptotics is easier to treat. We show (under regularity conditions) that the proposed $\hat{Y}_t$ is close to Y_t for the small values of ε.

We believe that the presented results can be valid (generalized) for essentially more general, say, nonlinear models and the conditions of regularity can be weakened.

2 Main result

We consider the following model. The observed diffusion process $X^T = (X_t, 0 \leq t \leq T)$ is

$$\mathrm{d}X_t = S\left(\vartheta, t, X_t\right) \mathrm{d}t + \varepsilon \sigma\left(t, X_t\right)\, \mathrm{d}W_t, \quad X_0,\ 0 \leq t \leq T \tag{4}$$

where $\vartheta \in \Theta = (\alpha, \beta)$ and

$$\begin{aligned} &\left|S\left(\vartheta, t, x_2\right) - S\left(\vartheta, t, x_1\right)\right| + \left|\sigma(t, x_2) - \sigma(t, x_2)\right| \leq L\ \left|x_2 - x_1\right|, \\ &\left|S\left(\vartheta, t, x\right)\right| + \left|\sigma(t, x)\right| \leq L\ \left(1 + |x|\right). \end{aligned} \tag{5}$$

We are given two functions $f\left(t, x, y, z\right)$, $\Phi\left(x\right)$ and we have to find a couple of stochastic processes $(X_t, Z_t, 0 \leq t \leq T)$ such that the solution of the equation (*backward SDE*)

$$\mathrm{d}Y_t = -f\left(t, X_t, Y_t, Z_t\right)\, \mathrm{d}t + Z_t\, \mathrm{d}W_t, \qquad Y_0, \quad 0 \leq t \leq T \tag{6}$$

at point $t = T$ satisfies the condition $Y_T = \Phi\left(X_T\right)$.

As the solution of (6) is entirely defined by the initial value Y_0 and by the process $Z^T = (Z_t, 0 \leq t \leq T)$ we can seek Y_0, Z^T, which provide the equality $Y_T = \Phi\left(X_T\right)$.

Let us introduce a family of functions

$$\mathcal{U} = \left\{ \left(u(t, x, \vartheta), t \in [0, T]\,, x \in \mathbb{R}\right), \vartheta \in \Theta \right\}$$

such that for all $\vartheta \in \Theta$ the function $u(t, x, \vartheta)$ satisfies the equation

$$\frac{\partial u}{\partial t} + S(\vartheta, t, x)\frac{\partial u}{\partial x} + \frac{\varepsilon^2 \sigma(t, x)^2}{2}\frac{\partial^2 u}{\partial x^2} = -f\left(t, x, u, \varepsilon\sigma(x)\frac{\partial u}{\partial x}\right)$$

and condition $u(T, x, \vartheta) = \Phi\left(x\right)$. If we put $Y_t = u\left(t, X_t, \vartheta\right)$, then by Itô's formula we obtain (6) with $Z_t = \varepsilon\sigma\left(t, X_t\right) u'_x\left(t, X_t, \vartheta\right)$.

We suppose that the true value ϑ_0 of ϑ is unknown. Therefore we can not put $Y_t = u\left(t, X_t, \vartheta_0\right)$ and our goal is to approximate Y_t and Z_t. We would like to study this problem in the situation where the error of approximation can be small.

Of course, the *natural approximation* is first to estimate ϑ_0 and then to substitute it in the function $u\left(\cdot\right)$. The small error we can have, besides others, in the case when we have a large volume of observations ($T \to \infty$) or when the *noise* $\varepsilon\sigma\left(t, X_t\right)$ is small. At the present work we propose an approximation of Y_t in the case of *small noise*, as $\varepsilon \to 0$. We suppose to treat *large samples case* later.

Remind that the stochastic process X_t of the equation (4) under condition (5) converges to the deterministic function $x_s = x_s\left(\vartheta_0\right)$, where $x_s\left(\vartheta\right)$ is solution of the ordinary differential equation

$$\frac{\mathrm{d}x_s}{\mathrm{d}s} = S\left(\vartheta, s, x_s\right), \qquad x_0, \quad 0 \leq s \leq T, \tag{7}$$

and this convergence is uniform in $s \in [0, T]$ (see, e.g., [3] or [5]). The corresponding PDE with $\varepsilon = 0$ is

$$\frac{\partial u^0}{\partial t} + S(\vartheta, t, x)\frac{\partial u^0}{\partial x} = -f\left(t, x, u^0, 0\right), \qquad u^0\left(T, x, \vartheta\right) = \Phi\left(x\right).$$

and the *limit BSDE*

$$\frac{\mathrm{d}y_t}{\mathrm{d}t} = -f\left(t, x_t, y_t, 0\right), \qquad y_T = \Phi\left(x_T\right)$$

we obtain by putting $y_t = u^0\left(t, x_t, \vartheta\right)$.

To estimate ϑ we can use any *good* estimator. For example, let us denote the likelihood ratio

$$L\left(\vartheta, X^T\right) = \exp\left\{\int_0^T \frac{S\left(\vartheta, t, X_t\right)}{\varepsilon^2\,\sigma\left(t, X_t\right)^2}\,\mathrm{d}X_t - \int_0^T \frac{S\left(\vartheta, t, X_t\right)^2}{2\,\varepsilon^2\,\sigma\left(t, X_t\right)^2}\,\mathrm{d}t\right\}$$

and define the maximum likelihood estimator $\hat\vartheta_\varepsilon$ by the relation

$$L\left(\hat\vartheta_\varepsilon, X^T\right) = \sup_{\vartheta\in\Theta} L\left(\vartheta, X^T\right).$$

Note that we can not use this MLE $\hat\vartheta_\varepsilon$ and to write $\hat Y_t = u\left(t, X_t, \hat\vartheta_\varepsilon\right)$ because $\hat\vartheta_\varepsilon$ depends on all observations and at the moment t the observations $X_s, t < s \leq T$ are not available. If we decide to use just the observations up to instant t and to define the MLE as follows

$$L\left(\hat\vartheta_{t,\varepsilon}, X^t\right) = \sup_{\vartheta\in\Theta} L\left(\vartheta, X^t\right), \tag{8}$$

then we obtain mathematically correct approximation $\hat Y_t = u\left(t, X_t, \hat\vartheta_{t,\varepsilon}\right)$ and the properties of $\hat Y_t$ are described in [11]. Remind that under regularity conditions the estimator $\hat\vartheta_{t,\varepsilon}$ is consistent, asymptotically normal

$$\frac{\hat\vartheta_{t,\varepsilon} - \vartheta_0}{\varepsilon} \Longrightarrow \mathcal{N}\left(0, \mathrm{I}\left(\vartheta, x^t\right)^{-1}\right), \qquad \mathrm{I}\left(\vartheta, x^t\right) = \int_0^t \frac{\dot S\left(\vartheta, s, x_s\right)^2}{\sigma\left(s, x_s\right)^2}\,\mathrm{d}s$$

and asymptotically efficient (see [5]). Here and in the sequel dot means derivative w.r.t. ϑ and $\mathrm{I}\left(\vartheta, x^t\right)$ is the Fisher information. The approximation $\hat{Y}_t = u\left(t, X_t, \hat{\vartheta}_{t,\varepsilon}\right)$ is difficult to realize because to solve the equation (8) for all $t \in (0, T]$ is computationally a quite complicate problem.

We need some regularity conditions. Let us denote $\mathcal{P}$ a class of functions of x having polynomial majorants. For example a function $g\left(t, x, \vartheta, \varepsilon\right) \in \mathcal{P}$ means that there exist constants $C > 0$ and $p > 0$ which do not depend on $t \in [0, T], \vartheta \in \Theta, \varepsilon \in [0, 1]$ such that

$$\left|g\left(t, x, \vartheta, \varepsilon\right)\right| \leq C\left(1 + |x|^p\right). \tag{9}$$

We suppose that the functions $S\left(\vartheta, t, x\right)$ and $u(t, x, \vartheta)$ have two continuous derivatives w.r.t. ϑ and the following derivatives belong to $\mathcal{P}$

$$\dot{S}\left(\vartheta, t, x\right), \ddot{S}\left(\vartheta, t, x\right), \dot{S}_x'\left(\vartheta, t, x\right), \sigma_x'\left(t, x\right), \dot{u}(t, x, \vartheta), \ddot{u}(t, x, \vartheta), \dot{u}_x'(t, x, \vartheta).$$

The function $\sigma\left(t, x\right)^2 \geq \kappa > 0$ and we have the uniform in ϑ convergence (as $\varepsilon \to 0$)

$$u\left(t, x, \vartheta\right) \to u^0\left(t, x, \vartheta\right), \qquad u_x'\left(t, x, \vartheta\right) \to \left(u^0\right)_x'\left(t, x, \vartheta\right). \tag{10}$$

We propose the following solution. Fix some (small) $\delta > 0$ and introduce the *minimum distance estimator* (MDE) $\vartheta_{\delta,\varepsilon}^*$ by the relation

$$\left\|X - x\left(\vartheta_{\delta,\varepsilon}^*\right)\right\|^2 = \inf_{\vartheta \in \Theta}\left\|X - x\left(\vartheta\right)\right\|^2 = \inf_{\vartheta \in \Theta}\int_0^\delta \left[X_t - x_t\left(\vartheta\right)\right]^2 \mathrm{d}t.$$

This estimator is consistent and asymptotically normal

$$\varepsilon^{-1}\left(\vartheta_{\delta,\varepsilon}^* - \vartheta_0\right) \Longrightarrow \mathcal{N}\left(0, D_\delta\left(\vartheta_0\right)^2\right)$$

where $D_\delta\left(\vartheta_0\right)^2 > 0$ (see Theorem 7.5 [5]. The required regularity conditions are : the function $S\left(\vartheta, t, x\right)$ has two continuous derivatives w.r.t. ϑ having polynomial majorants (see (9)) and the following identifiability condition is fulfilled: for any $\nu > 0$

$$\inf_{|\vartheta - \vartheta_0| > \nu}\left\|x\left(\vartheta\right) - x\left(\vartheta_0\right)\right\| > 0.$$

Let us introduce the *one-step MLE*

$$\tilde{\vartheta}_{t,\varepsilon} = \vartheta_{\delta,\varepsilon}^* + \frac{\Delta_t\left(\vartheta_{\delta,\varepsilon}^*, X_\delta^t\right) + \Delta_\delta\left(\vartheta_{\delta,\varepsilon}^*, X^\delta\right)}{\mathrm{I}\left(\vartheta_{\delta,\varepsilon}^*, x^t\left(\vartheta_{\delta,\varepsilon}^*\right)\right)}, \tag{11}$$

where

$$\Delta_t\left(\vartheta, X_\delta^t\right) = \int_\delta^t \frac{\dot S\left(\vartheta, s, X_s\right)}{\sigma\left(s, X_s\right)^2}\left[{\rm d}X_s - S\left(\vartheta, s, X_s\right)\,{\rm d}s\right], \quad t\in[\delta,T],$$

$$\Delta_\delta\left(\vartheta, X^\delta\right) = A\left(\vartheta,\delta,X_\delta\right) - \int_0^\delta A_s'\left(\vartheta,s,X_s\right)\,{\rm d}s$$

$$-\frac{\varepsilon^2}{2}\int_0^\delta B_x'\left(\vartheta,s,X_s\right)\sigma\left(s,X_s\right)^2{\rm d}s - \int_0^\delta \frac{\dot S\left(\vartheta,s,X_s\right)S\left(\vartheta,s,X_s\right)}{\sigma\left(s,X_s\right)^2}{\rm d}s,$$

$$B\left(\vartheta,s,x\right) = \frac{\dot S\left(\vartheta,s,x\right)}{\sigma\left(s,x\right)^2}, \qquad A\left(\vartheta,s,x\right) = \int_{x_0}^x B\left(\vartheta,s,z\right)\,{\rm d}z,$$

$${\rm I}\left(\vartheta, x^t\left(\vartheta\right)\right) = \int_0^t \frac{\dot S\left(\vartheta,s,x_s\left(\vartheta\right)\right)^2}{\sigma\left(s,x_s\left(\vartheta\right)\right)^2}\,{\rm d}s.$$

The approximation of the solution of BSDE is given in the following theorem.

Theorem 1 *Let the conditions of regularity be fulfilled then the processes*

$$\hat Y_t = u\left(t, X_t, \tilde\vartheta_{t,\varepsilon}\right), \qquad \hat Z_t = \varepsilon\sigma\left(t,X_t\right)u_x'\left(t,X_t,\tilde\vartheta_{t,\varepsilon}\right)$$

for the values $t\in[\delta,T]$ *have the representation*

$$\hat Y_t = Y_t + \varepsilon \dot u\left(t,X_t,\vartheta_0\right)\,\xi_t\left(\vartheta_0\right) + o\left(\varepsilon\right), \tag{12}$$

$$\hat Z_t = Z_t + \varepsilon^2\sigma\left(t,X_t\right)\dot u_x'\left(t,X_t,\vartheta_0\right)\,\xi_t\left(\vartheta_0\right) + o\left(\varepsilon^2\right), \tag{13}$$

where

$$\xi_t\left(\vartheta_0\right) = {\rm I}\left(\vartheta_0, x^t\left(\vartheta_0\right)\right)^{-1}\int_0^t \frac{\dot S\left(\vartheta_0,s,x_s\left(\vartheta_0\right)\right)}{\sigma\left(s,x_s\left(\vartheta_0\right)\right)}\,{\rm d}W_s.$$

Proof. Suppose that we already proved that

$$\varepsilon^{-1}\left(\tilde\vartheta_{t,\varepsilon} - \vartheta_0\right) = \xi_t\left(\vartheta_0\right) + o\left(1\right), \tag{14}$$

then the representations (12), (13) we obtain by Taylor's formula

$$\hat Y_t = u\left(t,X_t,\vartheta_0\right) + \left(\tilde\vartheta_{t,\varepsilon}-\vartheta_0\right)\dot u\left(t,X_t,\vartheta_0\right) + o\left(\varepsilon\right)$$
$$= Y_t + \varepsilon\dot u\left(t,X_t,\vartheta_0\right)\xi_t\left(\vartheta_0\right) + o\left(\varepsilon\right),$$

and

$$\hat Z_t = \varepsilon\sigma\left(t,X_t\right)u_x'\left(t,X_t,\vartheta_0\right) + \left(\tilde\vartheta_{t,\varepsilon}-\vartheta_0\right)\varepsilon\sigma\left(t,X_t\right)\dot u_x'\left(t,X_t,\vartheta_0\right) + o\left(\varepsilon\right)$$
$$= Z_t + \varepsilon^2\sigma\left(t,X_t\right)\dot u_x'\left(t,X_t,\vartheta_0\right)\xi_t\left(\vartheta_0\right) + o\left(\varepsilon^2\right).$$

Let us verify (14). Remind that for any $p > 0$

$$
\begin{aligned}
&\sup_{0\leq t\leq T} \left|X_t - x_t\left(\vartheta_0\right)\right| \leq C\varepsilon \sup_{0\leq t\leq T} \left|W_t\right|,\\
&\mathbf{E}_{\vartheta_0} \left|X_t - x_t\left(\vartheta_0\right)\right|^p \leq C\,\varepsilon^p \\
&\mathbf{E}_{\vartheta_0} \left|\vartheta^*_{\delta,\varepsilon} - \vartheta_0\right|^p \leq C\,\varepsilon^p
\end{aligned}
\tag{15}
$$

(see, e.g.; Lemma 1.13, and Theorem 7.5 in [5]). Below we denoted $x_s = x_s\left(\vartheta_0\right)$, use the convergence (15), consistency of the MDE $\vartheta^*_{\delta,\varepsilon}$ and smoothness of $S\left(\vartheta, s, x\right)$ and $\sigma\left(s, x\right)$

$$
\begin{aligned}
\Delta_t\left(\vartheta^*_{\delta,\varepsilon}, X^t_\delta\right) &= \int_\delta^t \frac{\dot S\left(\vartheta^*_{\delta,\varepsilon}, s, X_s\right)}{\sigma\left(s, X_s\right)^2}\,\left[{\rm d}X_s - S\left(\vartheta^*_{\delta,\varepsilon}, s, X_s\right)\,{\rm d}s\right]\\
&= \varepsilon\int_\delta^t \frac{\dot S\left(\vartheta^*_{\delta,\varepsilon}, s, X_s\right)}{\sigma\left(s, X_s\right)}{\rm d}W_s\\
&\qquad + \int_\delta^t \frac{\dot S\left(\vartheta^*_{\delta,\varepsilon}, s, X_s\right)}{\sigma\left(s, X_s\right)^2}\,\left[S\left(\vartheta_0, s, X_s\right) - S\left(\vartheta^*_{\delta,\varepsilon}, s, X_s\right)\right]\,{\rm d}s\\
&= \varepsilon\int_\delta^t \frac{\dot S\left(\vartheta_0, s, x_s\right)}{\sigma\left(s, x_s\right)}{\rm d}W_s - \left(\vartheta^*_{\delta,\varepsilon} - \vartheta_0\right)\int_\delta^t \frac{\dot S\left(\vartheta_0, s, x_s\right)^2}{\sigma\left(s, x_s\right)^2}\,{\rm d}s + o\left(\varepsilon\right).
\end{aligned}
$$

Further, using the same arguments we write

$$
\begin{aligned}
\Delta_\delta&\left(\vartheta^*_{\delta,\varepsilon}, X^\delta\right) = A\left(\vartheta^*_{\delta,\varepsilon}, \delta, X_\delta\right) - \int_0^\delta A'_s\left(\vartheta^*_{\delta,\varepsilon}, s, X_s\right)\,{\rm d}s\\
&- \frac{\varepsilon^2}{2}\int_0^\delta B'_x\left(\vartheta^*_{\delta,\varepsilon}, s, X_s\right)\sigma\left(s, X_s\right)^2{\rm d}s - \int_0^\delta \frac{\dot S\left(\vartheta^*_{\delta,\varepsilon}, s, X_s\right) S\left(\vartheta^*_{\delta,\varepsilon}, s, X_s\right)}{\sigma\left(s, X_s\right)^2}{\rm d}s\\
&\qquad = A\left(\vartheta_0, \delta, X_\delta\right) - \int_0^\delta A'_s\left(\vartheta_0, s, X_s\right)\,{\rm d}s\\
&- \frac{\varepsilon^2}{2}\int_0^\delta B'_x\left(\vartheta_0, s, X_s\right)\sigma\left(s, X_s\right)^2{\rm d}s - \int_0^\delta \frac{\dot S\left(\vartheta_0, s, X_s\right) S\left(\vartheta_0, s, X_s\right)}{\sigma\left(s, X_s\right)^2}{\rm d}s\\
&- \left(\vartheta^*_{\delta,\varepsilon} - \vartheta_0\right)\int_0^\delta \frac{\dot S\left(\vartheta_0, s, X_s\right)^2}{\sigma\left(s, X_s\right)^2}{\rm d}s + \left(\vartheta^*_{\delta,\varepsilon} - \vartheta^*_0\right)\left[\dot A\left(\vartheta_0, \delta, X_\delta\right)\right.\\
&- \int_0^\delta \dot A'_s\left(\vartheta_0, s, X_s\right)\,{\rm d}s - \frac{\varepsilon^2}{2}\int_0^\delta \dot B'_x\left(\vartheta_0, s, X_s\right)\sigma\left(s, X_s\right)^2{\rm d}s\\
&\left.- \int_0^\delta \frac{\ddot S\left(\vartheta_0, s, X_s\right) S\left(\vartheta_0, s, X_s\right)}{\sigma\left(s, X_s\right)^2}{\rm d}s\right] + o\left(\varepsilon\right).
\end{aligned}
$$

Note that by Itô's formula

$$
\begin{aligned}
&A\left(\vartheta_0,\delta,X_\delta\right)-\int_0^\delta A_s'\left(\vartheta_0,s,X_s\right)\,\mathrm{d}s-\frac{\varepsilon^2}{2}\int_0^\delta B_x'\left(\vartheta_0,s,X_s\right)\sigma\left(s,X_s\right)^2\mathrm{d}s\\
&\quad-\int_0^\delta\frac{\dot S\left(\vartheta_0,s,X_s\right)S\left(\vartheta_0,s,X_s\right)}{\sigma\left(s,X_s\right)^2}\mathrm{d}s=\varepsilon\int_0^\delta\frac{\dot S\left(\vartheta_0,s,X_s\right)}{\sigma\left(s,X_s\right)}\mathrm{d}W_s
\end{aligned}
$$

and

$$
\begin{aligned}
&\dot A\left(\vartheta_0,\delta,X_\delta\right)-\int_0^\delta \dot A_s'\left(\vartheta_0,s,X_s\right)\,\mathrm{d}s-\frac{\varepsilon^2}{2}\int_0^\delta \dot B_x'\left(\vartheta_0,s,X_s\right)\sigma\left(s,X_s\right)^2\mathrm{d}s\\
&\qquad-\int_0^\delta\frac{\ddot S\left(\vartheta_0,s,X_s\right)S\left(\vartheta_0,s,X_s\right)}{\sigma\left(s,X_s\right)^2}\mathrm{d}s=\varepsilon\int_0^\delta\frac{\ddot S\left(\vartheta_0,s,X_s\right)}{\sigma\left(s,X_s\right)}\mathrm{d}W_s.
\end{aligned}
$$

Therefore

$$
\begin{aligned}
\Delta_t\left(\vartheta_{\delta,\varepsilon}^*,X_\delta^t\right)+\Delta_\delta\left(\vartheta_{\delta,\varepsilon}^*,X^\delta\right)&=\varepsilon\int_0^t\frac{\dot S\left(\vartheta_0,s,X_s\right)}{\sigma\left(s,X_s\right)}\mathrm{d}W_s\\
&\quad-\left(\vartheta_{\delta,\varepsilon}^*-\vartheta_0\right)\int_0^t\frac{\dot S\left(\vartheta_0,s,X_s\right)^2}{\sigma\left(s,X_s\right)^2}\mathrm{d}s+o\left(\varepsilon\right)\\
&=\varepsilon\int_0^t\frac{\dot S\left(\vartheta_0,s,x_s\right)}{\sigma\left(s,x_s\right)}\mathrm{d}W_s-\left(\vartheta_{\delta,\varepsilon}^*-\vartheta_0\right)\int_0^t\frac{\dot S\left(\vartheta_0,s,x_s\right)^2}{\sigma\left(s,x_s\right)^2}\mathrm{d}s+o\left(\varepsilon\right).
\end{aligned}
$$

For the one-step MLE this allows us to write

$$
\begin{aligned}
\varepsilon^{-1}\left(\tilde\vartheta_{t,\varepsilon}-\vartheta_0\right)&=\frac{\vartheta_{\delta,\varepsilon}^*-\vartheta_0}{\varepsilon}+\mathrm{I}\left(\vartheta_{\delta,\varepsilon}^*,x^t\left(\vartheta_{\delta,\varepsilon}^*\right)\right)^{-1}\left[\int_0^t\frac{\dot S\left(\vartheta_0,s,x_s\right)}{\sigma\left(s,x_s\right)}\mathrm{d}W_s\right.\\
&\qquad\left.-\frac{\left(\vartheta_{\delta,\varepsilon}^*-\vartheta_0\right)}{\varepsilon}\int_0^t\frac{\dot S\left(\vartheta_0,s,x_s\right)^2}{\sigma\left(s,x_s\right)^2}\mathrm{d}s\right]+o\left(1\right)\\
&=\mathrm{I}\left(\vartheta_0,x^t\left(\vartheta_0\right)\right)^{-1}\int_0^t\frac{\dot S\left(\vartheta_0,s,x_s\right)}{\sigma\left(s,x_s\right)}\mathrm{d}W_s+o\left(1\right)=\xi_t\left(\vartheta_0\right)+o\left(1\right).
\end{aligned}
$$

This proves Theorem 1. We used just the continuity of the derivatives $\dot u$

Let us show that the proposed approximation is asymptotically efficient. This means, that the means-quare errors

$$
\mathbf{E}_\vartheta\left|Y_t-\hat Y_t\right|^2,\qquad \mathbf{E}_\vartheta\left|Z_t-\hat Z_t\right|^2,
$$

of estimation Y_t and Z_t can not be improved. This will be done in two steps. First we establish a low bound on the risks of all estimators and then show that the proposed estimators attaint this bound.

Theorem 2 *For all estimators $\bar Y_t$ and $\bar Z_t$ and all $t \in [\delta, T]$ we have the relations*

$$\varliminf_{\nu\to 0}\varliminf_{\varepsilon\to 0}\sup_{|\vartheta-\vartheta_0|\leq\nu}\varepsilon^{-2}\mathbf{E}_\vartheta\left|\bar Y_t - Y_t\right|^2 \geq \frac{\dot u^0\left(t, x_t\left(\vartheta_0\right), \vartheta_0\right)^2}{\mathrm{I}\left(\vartheta_0, x^t\left(\vartheta_0\right)\right)}, \tag{16}$$

$$\varliminf_{\nu\to 0}\varliminf_{\varepsilon\to 0}\sup_{|\vartheta-\vartheta_0|\leq\nu}\varepsilon^{-4}\mathbf{E}_\vartheta\left|\bar Z_t - Z_t\right|^2 \geq \frac{\left(\dot u^0\right)_x'\left(t, x_t\left(\vartheta_0\right), \vartheta_0\right)^2\sigma\left(t, x_t\left(\vartheta_0\right)\right)^2}{\mathrm{I}\left(\vartheta_0, x^t\left(\vartheta_0\right)\right)} \tag{17}$$

Proof. We follow the usual proof of the van Trees and minimax bounds. Let us fix $\nu > 0$ and introduce a probability density $p(\theta), \theta \in [\vartheta_0 - \nu, \vartheta_0 + \nu]$ such that $p(\vartheta_0 - \nu) = 0$, $p(\vartheta_0 + \nu) = 0$ and

$$\mathrm{I}_p = \int_{\vartheta_0-\nu}^{\vartheta_0+\nu}\frac{\dot p(\theta)^2}{p(\theta)}\,\mathrm{d}\theta < \infty.$$

Denote $L_{\theta_0}(\theta, X^t) = L(\theta_0, X^t)^{-1} L(\theta, X^t)$. Integrating by parts we obtain

$$\begin{aligned}\int_{\vartheta_0-\nu}^{\vartheta_0+\nu}&\left[\bar Y_t - u(t, X_t, \vartheta)\right]\frac{\partial}{\partial\theta}\left[L_{\theta_0}\left(\theta, X^t\right)p(\theta)\right]\mathrm{d}\theta\\ &= L_{\theta_0}\left(\theta, X^t\right)p(\theta)\left[\bar Y_t - u(t, X_t, \vartheta)\right]\Big|_{\vartheta_0-\nu}^{\vartheta_0+\nu}\\ &\qquad+\int_{\vartheta_0-\nu}^{\vartheta_0+\nu}\dot u(t, X_t, \vartheta)\,L_{\theta_0}\left(\theta, X^t\right)p(\theta)\,\mathrm{d}\theta\\ &= \int_{\vartheta_0-\nu}^{\vartheta_0+\nu}\dot u(t, X_t, \vartheta)\,L_{\theta_0}\left(\theta, X^t\right)p(\theta)\,\mathrm{d}\theta.\end{aligned}$$

We have

$$\begin{aligned}\mathbf{E}_{\vartheta_0}&\int_{\vartheta_0-\nu}^{\vartheta_0+\nu}\left[\bar Y_t - u(t, X_t, \vartheta)\right]\frac{\partial}{\partial\theta}\left[L_{\theta_0}\left(\theta, X^t\right)p(\theta)\right]\mathrm{d}\theta\\ &= \mathbf{E}_{\vartheta_0}\int_{\vartheta_0-\nu}^{\vartheta_0+\nu}\dot u(t, X_t, \vartheta)\,L_{\theta_0}\left(\theta, X^t\right)p(\theta)\,\mathrm{d}\theta\\ &= \int_{\vartheta_0-\nu}^{\vartheta_0+\nu}\mathbf{E}_\vartheta\dot u(t, X_t, \vartheta)\,p(\theta)\,\mathrm{d}\theta = \int_{\vartheta_0-\nu}^{\vartheta_0+\nu}\dot u(t, x_t, \vartheta)\,p(\theta)\,\mathrm{d}\theta + O(\varepsilon).\end{aligned}$$

We need the equality

$$
\begin{aligned}
&\mathbf{E}_{\vartheta_0}\int_{\vartheta_0-\nu}^{\vartheta_0+\nu}\left(\frac{\partial}{\partial\theta}\ln\left[L_{\theta_0}\left(\theta,X^t\right)p\left(\theta\right)\right]\right)^2 L_{\theta_0}\left(\theta,X^t\right)p\left(\theta\right)\mathrm{d}\theta\\
&\quad=\int_{\vartheta_0-\nu}^{\vartheta_0+\nu}\mathbf{E}_{\vartheta}\left(\int_0^t\frac{\dot S\left(\vartheta,s,X_s\right)}{\varepsilon\,\sigma\left(s,X_s\right)}\mathrm{d}W_s+\frac{\dot p\left(\theta\right)}{p\left(\theta\right)}\right)^2 p\left(\theta\right)\mathrm{d}\theta\\
&\quad=\int_{\vartheta_0-\nu}^{\vartheta_0+\nu}\mathbf{E}_{\theta}\int_0^t\frac{\dot S\left(\theta,s,X_s\right)^2}{\varepsilon^2\sigma\left(s,X_s\right)^2}\mathrm{d}s\;p\left(\theta\right)\;\mathrm{d}\theta+\mathrm{I}_p\\
&\quad=\frac{1}{\varepsilon^2}\int_{\vartheta_0-\nu}^{\vartheta_0+\nu}\mathrm{I}\left(\theta,x^t\left(\theta\right)\right)\;p\left(\theta\right)\;\mathrm{d}\theta+\mathrm{I}_p+o\left(\frac{1}{\varepsilon}\right),
\end{aligned}
$$

where we used the chain rule ($\mathbf{E}_{\vartheta_0}L_{\theta_0}\left(\theta,X^t\right)=\mathbf{E}_{\vartheta}$). Below we apply Cauchy-Shwartz inequality

$$
\begin{aligned}
&\left(\mathbf{E}_{\vartheta_0}\int_{\vartheta_0-\nu}^{\vartheta_0+\nu}\left[\bar Y_t-u\left(t,X_t,\vartheta\right)\right]\frac{\partial}{\partial\theta}\left[L_{\theta_0}\left(\theta,X^t\right)p\left(\theta\right)\right]\mathrm{d}\theta\right)^2\\
&\quad=\left(\mathbf{E}_{\vartheta_0}\int_{\vartheta_0-\nu}^{\vartheta_0+\nu}\left[\bar Y_t-u\left(t,X_t,\vartheta\right)\right]\right.\\
&\qquad\qquad\left.\frac{\partial}{\partial\theta}\ln\left[L_{\theta_0}\left(\theta,X^t\right)p\left(\theta\right)\right]L_{\theta_0}\left(\theta,X^t\right)p\left(\theta\right)\mathrm{d}\theta\right)^2\\
&\quad\leq\mathbf{E}_{\vartheta_0}\int_{\vartheta_0-\nu}^{\vartheta_0+\nu}\left[\bar Y_t-u\left(t,X_t,\vartheta\right)\right]^2 L_{\theta_0}\left(\theta,X^t\right)p\left(\theta\right)\mathrm{d}\theta\\
&\qquad\mathbf{E}_{\vartheta_0}\int_{\vartheta_0-\nu}^{\vartheta_0+\nu}\left(\frac{\partial}{\partial\theta}\ln\left[L_{\theta_0}\left(\theta,X^t\right)p\left(\theta\right)\right]\right)^2 L_{\theta_0}\left(\theta,X^t\right)p\left(\theta\right)\mathrm{d}\theta\\
&\quad=\int_{\vartheta_0-\nu}^{\vartheta_0+\nu}\mathbf{E}_{\vartheta}\left[\bar Y_t-u\left(t,X_t,\vartheta\right)\right]^2 p\left(\theta\right)\mathrm{d}\theta\\
&\qquad\left[\frac{1}{\varepsilon^2}\int_{\vartheta_0-\nu}^{\vartheta_0+\nu}\mathrm{I}\left(\theta,x^t\left(\theta\right)\right)\;p\left(\theta\right)\;\mathrm{d}\theta+\mathrm{I}_p+o\left(\frac{1}{\varepsilon}\right)\right].
\end{aligned}
$$

Therefore we obtained the van Trees inequality

$$
\begin{aligned}
&\int_{\vartheta_0-\nu}^{\vartheta_0+\nu}\mathbf{E}_{\vartheta}\left[\bar Y_t-u\left(t,X_t,\vartheta\right)\right]^2 p\left(\theta\right)\mathrm{d}\theta\\
&\qquad\geq\frac{\left(\int_{\vartheta_0-\nu}^{\vartheta_0+\nu}\dot u\left(t,x_t,\vartheta\right)p\left(\theta\right)\mathrm{d}\theta+O\left(\varepsilon\right)\right)^2}{\frac{1}{\varepsilon^2}\int_{\vartheta_0-\nu}^{\vartheta_0+\nu}\mathrm{I}\left(\theta,x^t\left(\theta\right)\right)\;p\left(\theta\right)\;\mathrm{d}\theta+\mathrm{I}_p+o\left(\frac{1}{\varepsilon}\right)}
\end{aligned}
$$

Further

$$
\begin{aligned}
&\varliminf_{\varepsilon\to0}\sup_{|\vartheta-\vartheta_0|\leq\nu}\varepsilon^{-2}\mathbf{E}_\vartheta\left|\bar Y_t-Y_t\right|^2=\varliminf_{\varepsilon\to0}\sup_{|\vartheta-\vartheta_0|\leq\nu}\varepsilon^{-2}\mathbf{E}_\vartheta\left|\bar Y_t-u\left(t,X_t,\vartheta\right)\right|^2\\
&\quad\geq\varliminf_{\varepsilon\to0}\varepsilon^{-2}\int_{\vartheta_0-\nu}^{\vartheta_0+\nu}\mathbf{E}_\vartheta\left[\bar Y_t-u\left(t,X_t,\vartheta\right)\right]^2p\left(\theta\right)\mathrm{d}\theta\\
&\quad\geq\frac{\left(\int_{\vartheta_0-\nu}^{\vartheta_0+\nu}\dot u^0\left(t,x_t,\vartheta\right)p\left(\theta\right)\mathrm{d}\theta\right)^2}{\int_{\vartheta_0-\nu}^{\vartheta_0+\nu}\mathrm{I}\left(\theta,x^t\left(\theta\right)\right)\;p\left(\theta\right)\;\mathrm{d}\theta}\longrightarrow\frac{\dot u^0\left(t,x_t\left(\vartheta_0\right),\vartheta_0\right)^2}{\mathrm{I}\left(\vartheta_0,x^t\left(\vartheta_0\right)\right)}.
\end{aligned}
$$

The last limit corresponds to $\nu\to0$ and used the continuity of the underlying functions. Therefore the inequality (16) is proved. The proof of (17) is quite similar.

We call an approximation $Y_t^\star$ asymptotically efficient if for all $\vartheta_0\in\Theta$ we have the equality

$$
\lim_{\nu\to0}\lim_{\varepsilon\to0}\sup_{|\vartheta-\vartheta_0|\leq\nu}\varepsilon^{-2}\mathbf{E}_\vartheta\left|Y_t^\star-Y_t\right|^2=\frac{\dot u^0\left(t,x_t\left(\vartheta_0\right),\vartheta_0\right)^2}{\mathrm{I}\left(\vartheta_0,x^t\left(\vartheta_0\right)\right)}\tag{18}
$$

and of course the similar definition is valid in the case of the bound (17).

Theorem 3 *The approximations*

$$
\hat Y_t=u\left(t,X_t,\tilde\vartheta_{t,\varepsilon}\right)\qquad\text{and}\qquad\hat Z_t=\varepsilon\sigma\left(t,X_t\right)u_x'\left(t,X_t,\tilde\vartheta_{t,\varepsilon}\right)
$$

are asymptotically efficient, i.e.,

$$
\begin{aligned}
\lim_{\nu\to0}\lim_{\varepsilon\to0}\sup_{|\vartheta-\vartheta_0|\leq\nu}\varepsilon^{-2}\mathbf{E}_\vartheta\left|\hat Y_t-Y_t\right|^2&=\frac{\dot u^0\left(t,x_t\left(\vartheta_0\right),\vartheta_0\right)^2}{\mathrm{I}\left(\vartheta_0,x^t\left(\vartheta_0\right)\right)},\\
\lim_{\nu\to0}\lim_{\varepsilon\to0}\sup_{|\vartheta-\vartheta_0|\leq\nu}\varepsilon^{-4}\mathbf{E}_\vartheta\left|\hat Z_t-Z_t\right|^2&=\frac{\sigma\left(t,x_t\left(\vartheta_0\right)\right)^2\left(\dot u^0\right)_x'\left(t,x_t\left(\vartheta_0\right),\vartheta_0\right)^2}{\mathrm{I}\left(\vartheta_0,x^t\left(\vartheta_0\right)\right)}
\end{aligned}
$$

Proof. The proof of these equalities follows from (12)-(13) because using standard arguments we can show that the convergence $o\left(1\right)$ in these representations is uniform in ϑ and the moments converge too.

3 Discussion

Uniform approximation. The representations (12), (13) are valid for each $t\in[\delta,T]$. It is possible to show that these equalities are true uniformly in

t. More precisely, let us put $\nu = \varepsilon^\kappa, \kappa > 0$. Then for sufficiently small κ we have the convergence

$$\mathbf{P}_{\vartheta_0}\left\{\sup_{\delta\leq t\leq T}\left|\hat{Y}_t - Y_t\right| > \nu\right\} \longrightarrow 0.$$

Indeed, as the derivatives $\dot{u}(t,x,\vartheta)$ and $\dot{u}'_x(t,x,\vartheta)$ have polynomial majorants, we can write

$$\begin{aligned}\mathbf{P}_{\vartheta_0}&\left\{\sup_{\delta\leq t\leq T}\left|\hat{Y}_t - Y_t\right| > \nu\right\} = \mathbf{P}_{\vartheta_0}\left\{\sup_{\delta\leq t\leq T}\left|\dot{u}\left(t,X_t,\bar{\vartheta}\right)\right|\,\left|\tilde{\vartheta}_{t,\varepsilon}-\vartheta_0\right| > \nu\right\}\\ &\leq \mathbf{P}_{\vartheta_0}\left\{\sup_{\delta\leq t\leq T}\left|\dot{u}\left(t,X_t,\bar{\vartheta}\right)\right| > \nu^{-\frac12}\right\} + \mathbf{P}_{\vartheta_0}\left\{\sup_{\delta\leq t\leq T}\left|\tilde{\vartheta}_{t,\varepsilon}-\vartheta_0\right| > \nu^{\frac32}\right\}\\ &\leq \mathbf{P}_{\vartheta_0}\left\{C\sup_{\delta\leq t\leq T}\left|X_t\right|^p > \nu^{-\frac12}\right\} + \mathbf{P}_{\vartheta_0}\left\{\sup_{\delta\leq t\leq T}\left|\tilde{\vartheta}_{t,\varepsilon}-\vartheta_0\right| > \nu^{\frac32}\right\}.\end{aligned}$$

The estimates (15) allow us to prove the convergence

$$\mathbf{P}_{\vartheta_0}\left\{\sup_{\delta\leq t\leq T}\left|X_t\right|^p > \nu^{-\frac12}\right\} \rightarrow 0.$$

The verification

$$\mathbf{P}_{\vartheta_0}\left\{\sup_{\delta\leq t\leq T}\left|\tilde{\vartheta}_{t,\varepsilon}-\vartheta_0\right| > \nu^{\frac32}\right\}$$

with $3\kappa < 2$ is more complicate, but direct, because we have the uniform convergence

$$\begin{aligned}&\sup_{\delta\leq t\leq T}\left|\mathrm{I}\left(\vartheta^*_{\delta,\varepsilon},x^t\left(\vartheta^*_{\delta,\varepsilon}\right)\right)-\mathrm{I}\left(\vartheta_0,x^t\left(\vartheta_0\right)\right)\right|\rightarrow 0,\\ &\sup_{\delta\leq t\leq T}\left|\mathrm{I}\left(\vartheta^*_{\delta,\varepsilon},X^t\right)-\mathrm{I}\left(\vartheta_0,x^t\left(\vartheta_0\right)\right)\right|\rightarrow 0.\end{aligned}$$

Further $\mathrm{I}\left(\vartheta^*_{\delta,\varepsilon},x^t\left(\vartheta^*_{\delta,\varepsilon}\right)\right) \geq \mathrm{I}\left(\vartheta^*_{\delta,\varepsilon},x^\delta\left(\vartheta^*_{\delta,\varepsilon}\right)\right) \geq \inf_\vartheta \mathrm{I}\left(\vartheta,x^\delta\left(\vartheta\right)\right) > 0$ and

$$\begin{aligned}\mathbf{P}_{\vartheta_0}&\left\{\sup_{\delta\leq t\leq T}\left|\int_\delta^t\left[\frac{\dot{S}\left(\vartheta^*_{\delta,\varepsilon},s,X_s\right)}{\sigma\left(s,X_s\right)}-\frac{\dot{S}\left(\vartheta_0,s,x_s\right)}{\sigma\left(s,x_s\right)}\right]\mathrm{d}W_t\right| > \nu\right\}\\ &\leq C\nu^{-2}\mathbf{E}_{\vartheta_0}\int_\delta^T\left[\frac{\dot{S}\left(\vartheta^*_{\delta,\varepsilon},s,X_s\right)}{\sigma\left(s,X_s\right)}-\frac{\dot{S}\left(\vartheta_0,s,x_s\right)}{\sigma\left(s,x_s\right)}\right]^2\mathrm{d}s \leq C\nu^{-2}\varepsilon^2.\end{aligned}$$

The proof of the last estimate is direct.

Case $\delta\rightarrow 0$. The representations (12), (13) are valid for each $t\in[\delta,T]$ with fixed $\delta > 0$. It is possible to show that $\hat{Y}_t \rightarrow Y_t$ and $\varepsilon^{-1}\hat{Z}_t \rightarrow \varepsilon^{-1}Z_t$ as

$\varepsilon \to 0$ in the situation, where $\delta = \delta_\varepsilon \to 0$ but *slowly.* What we need for the consistency of the estimator $\vartheta^*_{\delta_\varepsilon,\varepsilon}$ is the condition: for any $\nu > 0$

$$\lim_{\varepsilon\to 0} \varepsilon^{-2} \inf_{|\vartheta-\vartheta_0|>\nu} \int_0^{\delta_\varepsilon} \left[x_t\left(\vartheta\right) - x_t\left(\vartheta_0\right)\right]^2 \mathrm{d}t \to \infty.$$

For example, if the derivative $\left|\dot S\left(\vartheta_0, 0, x_0\right)\right| \geq \gamma > 0$, then for the function

$$\dot x_t\left(\vartheta\right) = \int_0^t \exp\left\{\int_s^t S'_x\left(\vartheta, v, x_v\right) \mathrm{d}v\right\} \dot S\left(\vartheta, s, x_s\right)\,\mathrm{d}s$$

and small δ_ε we have

$$\varepsilon^{-2} \inf_{|\vartheta-\vartheta_0|>\nu} \int_0^{\delta_\varepsilon} \left[x_t\left(\vartheta\right) - x_t\left(\vartheta_0\right)\right]^2 \mathrm{d}t \geq \frac{\gamma^2\,\delta_\varepsilon^3\,\nu^2}{6\,\varepsilon^2}.$$

Let us consider the linear case

$$\mathrm{d}X_t = \vartheta X_t \mathrm{d}t + \varepsilon \mathrm{d}W_t, \quad X_0 = x_0 > 0, \quad 0 \leq t \leq T.$$

Then the MLE can be written explicitly

$$\hat\vartheta_{t,\varepsilon} = \frac{\int_0^t X_s \mathrm{d}X_s}{\int_0^t X_s^2 \mathrm{d}s} = \vartheta + \varepsilon \frac{\int_0^t X_s \mathrm{d}W_s}{\int_0^t X_s^2 \mathrm{d}s}$$

and

$$\hat\vartheta_{\delta_\varepsilon,\varepsilon} - \vartheta = \varepsilon \frac{\int_0^{\delta_\varepsilon} X_s \mathrm{d}W_s}{\int_0^{\delta_\varepsilon} X_s^2 \mathrm{d}s} \sim \frac{\varepsilon\, W_{\delta_\varepsilon}}{x_0\ \delta_\varepsilon} \sim \frac{\varepsilon\, W_1}{x_0\ \delta_\varepsilon^{1/2}}.$$

Therefore, if $\varepsilon\delta_\varepsilon^{-1/2} \to 0$ (for example, $\delta_\varepsilon = \varepsilon^2 \ln \frac{1}{\varepsilon}$) then $\hat Y_t \to Y_t$ for all $t \in [\delta_\varepsilon, T]$.

Approximation of the BSDE. Note that $\hat Y_t$ is approximation of the solution of the BSDE (6), but the stochastic process $\hat Y_t$ itself satisfies another stochastic differential equation. To simplify the notations let us put

$$\mathrm{I}_t = \mathrm{I}\left(\vartheta^*_{\delta,\varepsilon}, x^t\left(\vartheta^*_{\delta,\varepsilon}\right)\right), \qquad \Delta_t = \Delta_t\left(\vartheta^*_{\delta,\varepsilon}, x^t_\delta\right) + \Delta_\delta\left(\vartheta^*_{\delta,\varepsilon}, x^\delta\right),$$

$$b_t\left(x\right) = \frac{\dot S\left(\vartheta^*_{\delta,\varepsilon}, t, x\right)}{\sigma\left(t, x\right)}, \qquad c_t\left(x\right) = \frac{S\left(\vartheta_0, t, x\right) - S\left(\vartheta^*_{\delta,\varepsilon}, t, x\right)}{\sigma\left(t, x\right)}.$$

Then we can write the stochastic differential of the one-step MLE $\tilde{\vartheta}_{t,\varepsilon}$ as follows

$$\mathrm{d}\tilde{\vartheta}_{t,\varepsilon} = \mathrm{I}_t^{-1} b_t\left(X_t\right)\left[c_t\left(X_t\right) - \mathrm{I}_t^{-1} b_t\left(x_t\right)\Delta_t\right]\mathrm{d}t + \varepsilon \mathrm{I}_t^{-1} b_t\left(X_t\right)\mathrm{d}W_t, \quad \tilde{\vartheta}_{\delta,\varepsilon},$$

where $t \in [\delta, T]$.

The approximation of the BSDE is the following equation (below $u = u\left(t, X_t, \tilde{\vartheta}_{t,\varepsilon}\right)$ and $\delta \leq t \leq T$)

$$\begin{aligned}
\mathrm{d}\hat{Y}_t &= u_t'\,\mathrm{d}t + u_x'\,S\left(\vartheta_0, t, X_t\right)\mathrm{d}t + \frac{\varepsilon^2}{2} u_{x,x}''\,\sigma\left(t, X_t\right)^2\mathrm{d}t + \varepsilon u_x'\,\sigma\left(t, X_t\right)\mathrm{d}W_t \\
&\quad + \dot{u}\,\mathrm{I}_t^{-1} b_t\left(X_t\right)\left[c_t\left(X_t\right) - \mathrm{I}_t^{-1} b_t\left(x_t\right)\Delta_t\right]\mathrm{d}t + \frac{\varepsilon^2}{2}\ddot{u}\,\mathrm{I}_t^{-2} b_t\left(X_t\right)^2\mathrm{d}t \\
&\quad + \varepsilon\dot{u}\,\mathrm{I}_t^{-1} b_t\left(X_t\right)\mathrm{d}W_t + \frac{\varepsilon^2}{2}\dot{u}_x'\,\mathrm{I}_t^{-1} b_t\left(X_t\right)\sigma\left(t, X_t\right)\mathrm{d}t, \qquad \hat{Y}_\delta.
\end{aligned}$$

It can be written as follows

$$\begin{aligned}
\mathrm{d}\hat{Y}_t &= -f\left(t, X_t, \hat{Y}_t, \hat{Z}_t\right)\mathrm{d}t + \hat{Z}_t \mathrm{d}W_t \\
&\quad + u_x'\,S\left(\vartheta_0, t, X_t\right)\mathrm{d}t + \dot{u}\,\mathrm{I}_t^{-1} b_t\left(X_t\right)\left[c_t\left(X_t\right) - \mathrm{I}_t^{-1} b_t\left(x_t\right)\Delta_t\right]\mathrm{d}t \\
&\quad + \frac{\varepsilon^2}{2}\ddot{u}\,\mathrm{I}_t^{-2} b_t\left(X_t\right)^2\mathrm{d}t + \varepsilon\dot{u}\,\mathrm{I}_t^{-1} b_t\left(X_t\right)\mathrm{d}W_t \\
&\quad + \frac{\varepsilon^2}{2}\dot{u}_x'\,\mathrm{I}_t^{-1} b_t\left(X_t\right)\sigma\left(t, X_t\right)\mathrm{d}t, \qquad \hat{Y}_\delta, \quad \delta \leq t \leq T.
\end{aligned}$$

Other estimators of Y_t. There are many possibilities to construct estimators $\bar{Y}_t$ of the random function Y_t such that $\varepsilon^{-1}\left(\bar{Y}_t - Y_t\right) \Rightarrow \mathcal{N}$. We can put $\bar{Y}_t = u\left(t, X_t, \vartheta_{\delta,\varepsilon}^*\right), t \in [\delta, T]$. Then

$$\varepsilon^{-1}\left(\bar{Y}_t - Y_t\right) \Longrightarrow \mathcal{N}\left(0, \dot{u}\left(t, x_t, \vartheta_0\right)^2 D_\delta\left(\vartheta_0\right)^2\right),$$

where

$$D_\delta\left(\vartheta_0\right)^2 = \int_0^\delta \frac{\sigma\left(s, x_s\right)^2}{\psi\left(s, \vartheta_0\right)}\left(\int_s^\delta \psi\left(t, \vartheta_0\right)\,\dot{x}_t\,\mathrm{d}t\right)^2\mathrm{d}s$$

and

$$\psi\left(t, \vartheta_0\right) = \exp\left\{\int_0^t S_x'\left(s, x_s\right)\,\mathrm{d}s\right\}$$

(see Theorem 7.5, [5]). It is known that for $t \geq \delta$

$$D_\delta\left(\vartheta_0\right)^2 \geq \mathrm{I}\left(\vartheta_0, x^\delta\left(\vartheta_0\right)\right)^{-1} \geq \mathrm{I}\left(\vartheta_0, x^t\left(\vartheta_0\right)\right)^{-1}.$$

Therefore this approximation is not asymptotically efficient.

Another possibility is to use the limit equation ($\varepsilon = 0$)

$$\frac{\partial u^0}{\partial t} + S(\vartheta, x)\frac{\partial u^0}{\partial x} = -f\left(x, u^0, 0\right), \qquad u^0(T, x, \vartheta) = \Phi(x)$$

and to put $\bar{Y}_t = u^0\left(t, X_t, \vartheta^*_{\delta,\varepsilon}\right)$. Under regularity conditions we have the convergence

$$\sup_{\vartheta}\left|u(t, X_t, \vartheta) - u(t, x_t, \vartheta)\right| \to 0, \qquad u(t, x, \vartheta) \longrightarrow u^0(t, x, \vartheta).$$

Therefore $\bar{Y}_t - Y_t \to 0$. This means that the both random functions have the same (deterministic) limit. Therefore such solutions are not asymptotically efficient and are essentially less interesting.

Linear case. Let us consider one example. Suppose that

$$\mathrm{d}X_t = \vartheta \mathrm{d}t + \varepsilon\sigma \mathrm{d}W_t, \quad X_0 = x_0, \quad 0 \le t \le T,$$

where $\vartheta \in \Theta = (a, b)$ and we are given two functions $f(x, z) = \beta y + \gamma z$ and $\Phi(x)$. The variables σ, β, γ are known constants and ϑ is unknown parameter. The function $\Phi(x)$ has two continuous derivatives with polynomial majorants. We have to construct the BSDE

$$\mathrm{d}Y_t = -\left(\beta Y_t + \gamma Z_t\right)\mathrm{d}t + Z_t \mathrm{d}W_t, \qquad Y_T = \Phi(X_T). \tag{19}$$

The corresponding PDE is

$$\begin{cases} \dfrac{\partial u}{\partial t} + \dfrac{1}{2}\varepsilon^2\sigma^2\dfrac{\partial^2 u}{\partial x^2} + (\vartheta + \varepsilon\sigma\gamma)\dfrac{\partial u}{\partial x} + \beta u = 0, \ 0 \le t \le T, \\ u(T, x, \vartheta) = \Phi(x), \ x \in \mathbb{R}. \end{cases} \tag{20}$$

with the solution

$$u(t, x, \vartheta) = e^{\beta(T-t)}\, G(t, x, \vartheta),$$

where we denoted

$$G(t, x, \vartheta) = \int_{-\infty}^{\infty} e^{-\frac{z^2}{2\varepsilon^2\sigma^2(T-t)}}\ \frac{\Phi(x + (\vartheta + \varepsilon\sigma\gamma)(T-t) - z)}{\sqrt{2\pi\varepsilon^2\sigma^2(T-t)}}\,\mathrm{d}z$$

Then we can put

$$\begin{aligned} Y_t &= u(t, X_t, \vartheta) = e^{\beta(T-t)}\, G(t, X_t, \vartheta), \\ Z_t &= \varepsilon\ \sigma\ u'(t, X_t, \vartheta) = \varepsilon\ \sigma\ e^{\beta(T-t)}\, G'_x(t, X_t, \vartheta), \end{aligned}$$

and obtain the BSDE (19).

Note that

$$G_x'(t,x,\vartheta)=\int_{-\infty}^{\infty}e^{-\frac{z^2}{2\varepsilon^2\sigma^2(T-t)}}\;\frac{\Phi'(x+(\vartheta+\varepsilon\sigma\gamma)(T-t)-z)}{\sqrt{2\pi\varepsilon^2\sigma^2(T-t)}}\;{\rm d}z.$$

We have

$$\begin{aligned}\dot u_\vartheta(t,x,\vartheta)&=(T-t)e^{\beta(T-t)}\int_{-\infty}^{\infty}e^{-\frac{z^2}{2\sigma^2(T-t)}}\;\frac{\Phi'(x+(\vartheta+\varepsilon\sigma\gamma)(T-t)-z)}{\sqrt{2\pi\varepsilon^2\sigma^2(T-t)}}\;{\rm d}z\\&=(T-t)e^{\beta(T-t)}G_x'(t,x,\vartheta),\end{aligned}$$

and

$$\begin{aligned}\dot u'(t,x,\vartheta)&=(T-t)e^{\beta(T-t)}G''(t,x,\vartheta),\\ \ddot u(t,x,\vartheta)&=(T-t)^2e^{\beta(T-t)}G''(t,x,\vartheta).\end{aligned}$$

In this model the MLE $\hat\vartheta_{t,\varepsilon}$ can be explicitly written

$$\hat\vartheta_{t,\varepsilon}=\frac{X_t}{t}=\vartheta_0+\varepsilon\sigma\frac{W_t}{t}\sim\mathcal{N}\left(\vartheta_0,\frac{\varepsilon^2\sigma^2}{t}\right)$$

and for all $t\in(0,T]$ is consistent. Therefore we can put

$$\begin{aligned}\hat Y_t&=e^{\beta(T-t)}\;G(t,X_t,\hat\vartheta_{t,\varepsilon}),\qquad t\in(0,T]\\ \hat Z_t&=\varepsilon\;\sigma\;e^{\beta(T-t)}\;G_x'(t,X_t,\hat\vartheta_{t,\varepsilon}),\qquad t\in(0,T]\end{aligned}$$

and according to the Theorem 3 this approximation is asymptotically efficient.

The limit solution of PDE is

$$u^0\left(t,x,\vartheta\right)=e^{\beta(T-t)}\,\Phi\left(x+\vartheta\left(T-t\right)\right).$$

Hence

$$\lim_{\varepsilon\to0}\varepsilon^{-2}\mathbf{E}_\vartheta\left(\hat Y_t-Y_t\right)^2=\left(T-t\right)^2t^{-1}\,\sigma^2e^{2\beta(T-t)}\Phi'\left(x_0+\vartheta T\right)^2.$$

The last expression is also the rhs in the lower bound (16).

On regularity condition (10). The most difficult to verify are the conditions imposed on $u\left(t,x,\vartheta\right)$ (regularity w.r.t. ϑ, convergence to $u^0\left(t,x,\vartheta\right)$ and majorations in x of $u\left(\cdot\right)$ and its derivatives). Sufficient conditions for the convergence of the solution $u(t,x,\vartheta)$ in homogeneous case and a particular case of the linear $f(\cdot)$ are given in the Theorem 1.3.1 in [3].

Some generalizations. The parameter ϑ in our work is supposed to be one-dimensional and the loss function $\ell(y)$ in the theorems 2 and 3 quadratic, $\ell(y) = y^2$. The case of multidimensional parameter and more general class of loss functions can be treated by a similar way, but in this case we have to use Hajek-Le Cam lower bound as follows. Suppose that $\vartheta \in \Theta$ where Θ is an open bounded subset of $\mathbb{R}^d$ and the corresponding regularity conditions are fulfilled. Then the family of measures $\left\{\mathbf{P}_\vartheta^{(t)}, \vartheta \in \Theta\right\}$ (induced in the space of realizations by the stochastic processes $X^t = (X_s, 0 \le s \le t))$ is *locally asymptotically normal* (LAN), i.e., the normalized likelihood ratio function $Z_{t,\varepsilon}(v) = L(\vartheta_0 + \varepsilon v, \vartheta_0, X^t)$ admits the representation

$$Z_\varepsilon(v) = \exp\left\{\langle v, \Delta_t\left(\vartheta_0, X_0^t\right)\rangle - \frac{1}{2} v^* \mathrm{I}\left(\vartheta_0, x^t(\vartheta_0)\right) v + r_\varepsilon\right\}, \qquad v \in \mathbb{R}^d,$$

where $\Delta_t(\vartheta_0, X_0^t) \Rightarrow \mathcal{N}(0, \mathrm{I}(\vartheta_0, x^t(\vartheta_0)))$ and $r_\varepsilon \to 0$. Suppose that the loss function $\ell(y)$ is nonnegative, continuous at point 0 and $\ell(0) = 0$, but is not identically 0, is symmetric $\ell(y) = \ell(|y|)$ and $\ell(y), y \ge 0$ is nondecreasing. Then we have the following Hajek-Le Cam lower bound: for all $\vartheta_0 \in \Theta$ and for all estimators $\bar{Y}_t, t \in [\delta, T]$

$$\varliminf_{\nu \to 0} \varliminf_{\varepsilon \to 0} \sup_{|\vartheta - \vartheta_0| \le \nu} \mathbf{E}_\vartheta \ell\left(\varepsilon^{-1}\left(\bar{Y}_t - Y_t\right)\right) \ge \mathbf{E}_{\vartheta_0} \ell\left(\langle \dot{u}^0(t, x_t(\vartheta_0), \vartheta_0), \xi_t(\vartheta_0)\rangle\right).$$

For the proof of LAN see Lemma 2.2 in [5] and for the lower bound see Theorem 2.12.1 in [4] (we have to modify slightly the proof, because we estimate not ϑ but some random function of ϑ).

The next step, of course, is to prove the asymptotic efficiency of the one-step MLE in the sense of this bound.

References

[1] El Karoui N., Mazliak L. (Eds). (1997) *Backward Stochastic Differential Equations.* Pitman Research Notes in Mathematics Series 364. Longman, Harlow.

[2] El Karoui N., Peng S. and Quenez M. (1997) Backward stochastic differential equations in finance, *Math. Finance*, 7, 1-71.

[3] Freidlin M. and Wentzell A. (1998) *Random Perturbations of Dynamical Systems.* Springer, New York.

[4] Ibragimov I.A. and Khasminskii R. (1981) *Statistical Estimation - Asymptotic Theorey.* Springer-Verlag, New York.

[5] Kutoyants Y.A. (1994) *Identification of Dynamical Systems with Small Noise,* Kluwer Academic Publisher, Dordrecht.

[6] Kutoyants Y.A. (2004) *Statiatical Inference for Ergodic Diffusion Processes.* Springer, London.

[7] Lipster R.S. and Shiryaev A.N. (1978) *Statistics of Random Process I II.* Springer New York.

[8] Ma J., Yong J. (1999) *Forward-Backward Stochastic Differential Equations and their Applications.* Lecture Notes in Mathematics. Springer, Berlin.

[9] Pardoux E. and Peng S. (1990) Adapted solution of a backward stochastic differential equation. *System Control Letter*, 14, 55-61.

[10] Revuz D. and Yor M. (1991) *Continuous Martingles and Brownian Motion.* Spring-Verlag, New-York.

[11] Zhou, L. (2013) *Problèmes Statistiques pour des EDS et les EDS Rétrogrades*, PhD Thesis, March 2013, Université du Maine, Le Mans.