Probability, Uncertainty and Quantitative Risk
Vol. 10, No. 4, 2025, 443–470

https://www.aimsciences.org/PUQR
DOI:10.3934/puqr.2025023

Extended adaptive Kalman filter with low noise observations

Yury A. Kutoyants[1,2]

[1]*Le Mans University, Le Mans, France*
[2]*Tomsk State University, Tomsk, Russia*

Email: kutoyants@univ-lemans.fr

Abstract The model of partially observed nonlinear system, called extended Kalman filter (EKF), and depending on some unknown parameters is considered. An approximation of the unobserved component is proposed. This approximation is realized in two steps. First a the method of moments estimator of unknown parameter is constructed and then this estimator is substituted in the equations of extended Kalman filter. The obtained equations describe the adaptive extended Kalman filter. The properties of estimator of the unknown parameter and of the unknown state are described in the asymptotic of small noise in observations.

Keywords Partially observed nonlinear system, Hidden markov process, Parameter estimation, Method of moments estimators, On-line approximation, Adaptive extended Kalman filter

2020 Mathematics Subject Classification 62M05, 60G35, 62M15

1. Introduction

Let us consider the nonlinear partially observed system

$$\mathrm{d}X_t = f\left(\vartheta, t, Y_t\right)\mathrm{d}t + \varepsilon\sigma\left(t\right)\mathrm{d}W_t, \qquad X_0 = 0, \qquad 0 \leqslant t \leqslant T, \tag{1.1}$$

$$\mathrm{d}Y_t = a\left(\vartheta, t, Y_t\right)\mathrm{d}t + b\left(\vartheta, t, Y_t\right)\mathrm{d}V_t, \qquad Y_0 = y_0, \tag{1.2}$$

where the Wiener processes $W_t, V_t, 0 \leqslant t \leqslant T$ are independent, the functions $f\left(\cdot\right), \sigma\left(\cdot\right)$, $a\left(\cdot\right), b\left(\cdot\right)$, initial value y_0 and the parameter $\varepsilon \in (0,1]$ are supposed to be known. The parameter $\vartheta \in \Theta = (\alpha, \beta)$ is unknown and we have to estimate the unobserved component Y_t by the observations $X^t = (X_s, 0 \leqslant s \leqslant t)$. The properties of this estimator has to be described in the asymptotic of *small noise* in observations, i.e., as $\varepsilon \to 0$.

Remind that if the system is linear w.r.t. Y_t,

$$\mathrm{d}X_t = f\left(\vartheta, t\right)Y_t\mathrm{d}t + \varepsilon\sigma\left(t\right)\mathrm{d}W_t, \qquad X_0 = 0, \qquad 0 \leqslant t \leqslant T, \tag{1.3}$$

$$\mathrm{d}Y_t = a\left(\vartheta, t\right)Y_t\mathrm{d}t + b\left(\vartheta, t\right)\mathrm{d}V_t, \qquad Y_0 = y_0, \tag{1.4}$$

Received 25 May 2024; Accepted 14 September 2025; Early access 11 October 2025

and ϑ is known, then the mean squared optimal estimator of Y_t is its conditional expectation $m\left(\vartheta,t\right)=\mathbf{E}_\vartheta\left(Y_t|X^t\right)$. The random process $m\left(\vartheta,t\right),\ 0\leqslant t\leqslant T$ is solution of the well-known Kalman-Bucy equations [11, 20]:

$$\mathrm{d}m\left(\vartheta,t\right)=a\left(\vartheta,t\right)m\left(\vartheta,t\right)\mathrm{d}t+\frac{\gamma\left(\vartheta,t\right)f\left(\vartheta,t\right)}{\varepsilon^2\sigma\left(t\right)^2}\left[\mathrm{d}X_t-f\left(\vartheta,t\right)m\left(\vartheta,t\right)\mathrm{d}t\right] \tag{1.5}$$

with initial value $m\left(\vartheta,0\right)=y_0$. The function $\gamma\left(\vartheta,t\right)=\mathbf{E}_\vartheta\left(Y_t-m\left(\vartheta,t\right)\right)^2$ is solution of Riccati equation

$$\frac{\partial\gamma\left(\vartheta,t\right)}{\partial t}=2a\left(\vartheta,t\right)\gamma\left(\vartheta,t\right)-\frac{\gamma\left(\vartheta,t\right)^2f\left(\vartheta,t\right)^2}{\varepsilon^2\sigma\left(t\right)^2}+b\left(\vartheta,t\right)^2,\quad \gamma\left(\vartheta,0\right)=0. \tag{1.6}$$

If ϑ is unknown then the equations (1.5), (1.6) can not be used for the calculation of $m\left(\vartheta,t\right),\ 0\leqslant t\leqslant T$ and some approximation of these equations is needed. We are interested by such approximation in the situation where it is possible to suppose that the level of the noise ε in the observations is small, i.e., in the asymptotic $\varepsilon\rightarrow 0$. One possibility is to obtain first some estimator, say, $\bar\vartheta_\varepsilon$ of the unknown parameter and then to put it in the underlying equations (1.5), (1.6). This behavior seems to be reasonable because if the estimator has good properties (small error), then we can expect that the corresponding solutions of these equations will be close to the true solutions with known ϑ. There are some technical problems with the realization of such program. First question is the choice of the estimator.

Recall that the MLE $\hat\vartheta_\varepsilon$ is defined by the equation

$$L(\hat\vartheta_\varepsilon,X^T)=\sup_{\vartheta\in\Theta}L\left(\vartheta,X^T\right). \tag{1.7}$$

Here the likelihood ratio function $L\left(\,\cdot\,,X^T\right)$ is

$$L\left(\vartheta,X^T\right)=\exp\left(\int_0^T\frac{f\left(\vartheta,t\right)m\left(\vartheta,t\right)}{\varepsilon^2\sigma\left(t\right)^2}\mathrm{d}X_t-\int_0^T\frac{f\left(\vartheta,t\right)^2m\left(\vartheta,t\right)^2}{2\varepsilon^2\sigma\left(t\right)^2}\mathrm{d}t\right),\qquad \vartheta\in\Theta.$$

It is known (see [14]) that this estimator has good asymptotic properties and, in particular, is asymptotically efficient, but to use it for approximation $m\left(\vartheta,t\right),\ 0\leqslant t\leqslant T$ can be numerically difficult problem. To solve equation (1.7) we need many solutions of the equations (1.5), (1.6) and even if we found $\hat\vartheta_\varepsilon=\hat\vartheta_\varepsilon\left(X^T\right)$ we cannot put it directly in the equation (1.5) because the stochastic integral

$$\int_0^t\frac{\gamma(\hat\vartheta_\varepsilon,s)f(\hat\vartheta_\varepsilon,s)}{\varepsilon^2\sigma\left(s\right)^2}\mathrm{d}X_s$$

is not well defined. That is why it seems reasonable to construct a consistent estimator $\bar\vartheta_{\tau,\varepsilon}$ by observations X^τ, where τ takes a (small) value $\tau\in\left(0,T\right)$ and then to define the estimator $m_t^*,\quad \tau\leqslant t\leqslant T$ as solution of (1.5), (1.6), where ϑ is replaced by $\bar\vartheta_{\tau,\varepsilon}$. Moreover, in this linear case it was possible to introduce the One-step MLE-process $\vartheta^\star_{t,\varepsilon},\ \tau<t\leqslant T$ and to construct an asymptotically optimal in the minimax sens estimator of the conditional expectation $m\left(\vartheta,t\right)$. This program was realized in the works [14, 16] for the linear system (1.3), (1.4).

The goal of the present work is to realize partially a similar program but in the case of nonlinear partially observed system (1.1), (1.2). Of course, there is no more equations like (1.5), (1.6) and we study, so-called, *extended Kalman filter* (EKF), which is introduced as follows:

$$\mathrm{d}m_t(\vartheta)=a(\vartheta,t,m_t(\vartheta))\,\mathrm{d}t+\frac{f'_y(\vartheta,t,m_t(\vartheta))\,\gamma_\varepsilon(\vartheta,t)}{\varepsilon\sigma(t)^2}\left[\mathrm{d}X_t-f(\vartheta,t,m_t(\vartheta))\,\mathrm{d}t\right], \tag{1.8}$$

$$\frac{\partial\gamma_\varepsilon(\vartheta,t)}{\partial t}=2a'_y(\vartheta,t,m_t(\vartheta))\,\gamma_\varepsilon(\vartheta,t)-\frac{f'_y(\vartheta,t,m_t(\vartheta))^2\,\gamma_\varepsilon(\vartheta,t)^2}{\varepsilon^2\sigma(t)^2}+b(\vartheta,t,m_t(\vartheta))^2, \tag{1.9}$$

subject to initial values $m_0(\vartheta)=y_0$, $\gamma_\varepsilon(\vartheta,0)=0$. It is easy to see that if $f(\vartheta,t,y)=f(\vartheta,t)\,y$, $a(\vartheta,t,y)=a(\vartheta,t)\,y$, $b(\vartheta,t,y)=b(\vartheta,t)$, then these equations coincide with the system (1.5), (1.6).

If ϑ is known then the random process $m_t(\vartheta),0\leqslant t\leqslant T$ can be considered as the estimator of Y_t. It is shown below that under mild regularity conditions $m_t(\vartheta)\longrightarrow Y_t$ as $\varepsilon\to 0$. In the case of unknown ϑ this parameter is estimated on the learning interval $[0,\tau]$, where $\tau<T$ with the help of method of moments estimator $\vartheta^*_{\tau,\varepsilon}$ and then this estimator is substituted in the equations (1.8), (1.9). The obtained solution of (1.8) we denote $m^\star_{t,\varepsilon}$. Thus we obtain the adaptive EKF and then we describe in Theorem 1 (main result of this work) the asymptotics of the error $m^\star_{t,\varepsilon}-Y_t$. Therefore the construction of adaptive Kalman filter or adaptive EKF requires at least two steps: estimation of unknown parameters and description of the adaptive filter after substitution of these estimators in the equations of Kalman filtration (in the linear case) or in equations similar to equations of Kalman filtration (for EKF).

The statistical problems with partially observed systems considered in this work belong to a wide class of problems related with hidden Markov processes. The different models with discrete and continuous time observations can be found in the works [4, 6, 20, 21]. The EKF for diffusion processes like (1.1), (1.2) but with small noise in both equations were studied in the works [1, 10, 12, 13, 23, 26–28] and many others. The exhaustive study of EKF with small noise in both equations and in observations equation only can be found in [2]. In all these works it is shown that the EKF are consistent, i.e., the solutions of the corresponding Kalman type equations converge to the value Y_t and if the system depends on some unknown parameter then there exist consistent estimators. Note that in these models the limit $(\varepsilon=0)$ system is deterministic. The studied here system in the limit $(\varepsilon=0)$ is random. The behavior of the EKF of the partially observed systems like (1.1), (1.2) was studied in the work [3], where the upper and lower bounds on the mean squared error of estimation of the initial value y_0 were proposed. The parameter estimation problem by uniformly discretized observations of slightly simplified model (1.1), (1.2) was considered in the works [7, 8] and the rate of convergence of the proposed there estimator was $\varepsilon^{1/2}$.

There exists a large engineering literature on the adaptive Kalman and extended Kalman filters (AEKF). The examples of AKF and AEKF and their applications in different situations can be found in the works [5, 9, 22, 24, 25].

Remark that the properties of parameter estimators for the linear models like (1.5), (1.6) but with small noise in both equations were described in [15, 19] and for the model (1.5), (1.6) in [14]. The asymptotically optimal in minimax sense AKF was proposed in [16]. The similar asymptotically optimal adaptive Kalman filters were studied in the asymptotics of large samples in [17] (homogeneous system, ergodic case) and [18] (hidden auto-regressive time series).

2. EKF and regularity conditions

We have the partially observed system (1.1), (1.2) with the small noise $(\varepsilon\in(0,1),\varepsilon\to 0)$ in observation equation only. The equations of EKF similar to (1.5), (1.6) are

$$\begin{aligned}\mathrm{d}m_t\left(\vartheta\right) = &\left[a\left(\vartheta,t,m_t\left(\vartheta\right)\right) - \frac{f_y'\left(\vartheta,t,m_t\left(\vartheta\right)\right) f\left(\vartheta,t,m_t\left(\vartheta\right)\right)\gamma_\varepsilon\left(\vartheta,t\right)}{\varepsilon^2\sigma\left(t\right)^2}\right]\mathrm{d}t\\ &+\frac{f_y'\left(\vartheta,t,m_t\left(\vartheta\right)\right)\gamma_\varepsilon\left(\vartheta,t\right)}{\varepsilon^2\sigma\left(t\right)^2}\,\mathrm{d}X_t,\quad m_0\left(\vartheta\right)=y_0,\end{aligned}\tag{2.1}$$

$$\frac{\partial\gamma_\varepsilon\left(\vartheta,t\right)}{\partial t} = 2a_y'\left(\vartheta,t,m_t\left(\vartheta\right)\right)\gamma_\varepsilon\left(\vartheta,t\right) - \frac{f_y'\left(\vartheta,t,m_t\left(\vartheta\right)\right)^2\gamma_\varepsilon\left(\vartheta,t\right)^2}{\varepsilon^2\sigma\left(t\right)^2} + b\left(\vartheta,t,m_t\left(\vartheta\right)\right)^2,\tag{2.2}$$

with the initial value $\gamma_\varepsilon\left(\vartheta,0\right)=0$.

Introduce some notations

$$\begin{aligned}&\dot g_\vartheta\left(\vartheta,t,y\right)=\frac{\partial g\left(\vartheta,t,y\right)}{\partial\vartheta},\quad g_t'\left(\vartheta,t,y\right)=\frac{\partial g\left(\vartheta,t,y\right)}{\partial t},\quad g_y'\left(\vartheta,t,y\right)=\frac{\partial g\left(\vartheta,t,y\right)}{\partial y},\\ &\dot g_{\vartheta y}'\left(\vartheta,t,y\right)=\frac{\partial^2 g\left(\vartheta,t,y\right)}{\partial\vartheta\,\partial t},\quad\ldots\end{aligned}$$

$g\left(\vartheta,t,y\right)\in\mathscr{C}_\vartheta^{i,b}$ means that the function $g\left(\cdot\right)$ has i continuous bounded derivatives with respect to ϑ *etc.* If $i=0$ then this condition means that the function is bounded and we denote C_g such constant that $\left|g\left(\vartheta,t,y\right)\right|\leqslant C_g$.

Conditions $\mathscr{C}$

($\mathscr{C}_1$) *The functions* $f\left(\vartheta,t,y\right), a\left(\vartheta,t,y\right),\ b\left(\vartheta,t,y\right)\in\mathcal{C}_t^1$, $f\left(\vartheta,t,y\right)\in\mathcal{C}_y^{2,b}$, *and* $a\left(\vartheta,t,y\right),\ b\left(\vartheta,t,y\right)\in\mathcal{C}_y^{i,b}, i=0,1$.

($\mathscr{C}_2$) *The functions* $f_y'\left(\vartheta,t,y\right),\ \sigma\left(t\right), b\left(\vartheta,t,y\right)$ *are separated from 0, i.e., there exist constants* $c_y>0, c_\sigma>0, c_b>0$ *such that*

$$f_y'\left(\vartheta,t,y\right)>c_y,\quad\sigma\left(t\right)>c_\sigma,\quad b\left(\vartheta,t,y\right)>c_b.$$

($\mathscr{C}_3$) *The functions* $f\left(\vartheta,t,y\right), b\left(\vartheta,t,y\right), f_y'\left(\vartheta,t,y\right)\in\mathcal{C}_\vartheta^{2,b}$. *There exists constants* $c_4>0$, $c_5>0$ *such that*

$$\dot f_{y\vartheta}'\left(\vartheta,t,y\right)>c_4,\quad\dot b\left(\vartheta,t,y\right)>c_5.$$

3. Asymptotics ($\varepsilon\to0$) of EKF

It can be shown (see Lemma 2 in [14]) that the solution $\gamma_\varepsilon\left(\vartheta,t\right)$ of Riccati equation (2.2) for the values $t\in\left[t_0,T\right]$ and any $t_0>0$ has the following limit in probability

$$\lim_{\varepsilon\to0}\sup_{t_0\leqslant t\leqslant T}\left|\frac{f_y'\left(\vartheta,t,m_t\left(\vartheta\right)\right)\gamma_\varepsilon\left(\vartheta,t\right)}{\varepsilon}-b\left(\vartheta,t,m_t\left(\vartheta\right)\right)\sigma\left(t\right)\right|=0.$$

Introduce the notation:

$$\gamma_*\left(\vartheta,t\right)=\frac{b\left(\vartheta,t,m_t\left(\vartheta\right)\right)\sigma\left(t\right)}{f_y'\left(\vartheta,t,m_t\left(\vartheta\right)\right)},\quad\gamma_*^\circ\left(\vartheta,t\right)=\frac{b\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)\sigma\left(t\right)}{f_y'\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)}.$$

Here $m_t^\circ\left(\vartheta\right)$ is the limit of $m_t\left(\vartheta\right)$ as $\varepsilon\to0$. This convergence is described below in the Lemma 1 (see as well (3.7)). Note that the conditions $\mathscr{C}_1$ and $\mathscr{C}_2$ are sufficient for the proof. The similar result see in [26].

For these values of t we replace $\varepsilon^{-1}\gamma_\varepsilon\left(\vartheta,t\right)$ by $\gamma_*\left(\vartheta,t\right)$ in all further study of this model. This will essentially simplify calculations related with the adaptive EKF.

The equation for estimate $m_t\left(\vartheta\right)$ of Y_t became

$$\begin{aligned}\mathrm{d}m_t(\vartheta)&=a(\vartheta,t,m_t(\vartheta))\,\mathrm{d}t+\frac{f'_y(\vartheta,t,m_t(\vartheta))\,\gamma_*(\vartheta,t)}{\varepsilon\sigma(t)^2}\left[\mathrm{d}X_t-f(\vartheta,t,m_t(\vartheta))\,\mathrm{d}t\right]\\&=a(\vartheta,t,m_t(\vartheta))\,\mathrm{d}t+\frac{b(\vartheta,t,m_t(\vartheta))}{\varepsilon\sigma(t)}\left[\mathrm{d}X_t-f(\vartheta,t,m_t(\vartheta))\,\mathrm{d}t\right],\quad t_0\leqslant t\leqslant T.\end{aligned}\tag{3.1}$$

Consider the limits in probability of the random functions $m_t(\vartheta)$, $\dot m_t(\vartheta)=\frac{\partial}{\partial\vartheta}m_t(\vartheta)$, $t_0\leqslant t\leqslant T$, which we denote as $m_t^\circ(\vartheta)$, $\dot m_t^\circ(\vartheta)$, $t_0\leqslant t\leqslant T$.

Lemma 1 *Let the conditions $\mathscr{C}$ be satisfied, then for any $t_0\in(0,T]$ the following relations hold*

$$\begin{aligned}&f(\vartheta,t,m_t^\circ(\vartheta))=f(\vartheta_0,t,Y_t),\quad m_t^\circ(\vartheta_0)=Y_t,\\&\dot m_t^\circ(\vartheta)=-\frac{\dot f_\vartheta(\vartheta,t,m_t^\circ(\vartheta))}{f'_y(\vartheta,t,m_t^\circ(\vartheta))},\quad \dot m_t^\circ(\vartheta_0)=-\frac{\dot f_\vartheta(\vartheta_0,t,Y_t)}{f'_y(\vartheta_0,t,Y_t)}.\end{aligned}\tag{3.2}$$

Proof The limit $m_t^\circ(\vartheta)$, $t_0<t\leqslant T$ can be described as follows. We have

$$\begin{aligned}\mathrm{d}m_t(\vartheta)=a(\vartheta,t,m_t(\vartheta))\,\mathrm{d}t+\frac{b(\vartheta,t,m_t(\vartheta))}{\varepsilon\sigma(t)}\left[f(\vartheta_0,t,Y_t)-f(\vartheta,t,m_t(\vartheta))\right]\mathrm{d}t\\+b(\vartheta,t,m_t(\vartheta))\,\mathrm{d}W_t.\end{aligned}$$

Hence

$$\begin{aligned}&\int_{t_0}^t\frac{b(\vartheta,s,m_s(\vartheta))}{\sigma(s)}\left[f(\vartheta_0,s,Y_s)-f(\vartheta,s,m_s(\vartheta))\right]\mathrm{d}s\\&=\varepsilon\left[m_t(\vartheta)-m_{t_0}(\vartheta)-\int_{t_0}^t a(\vartheta,s,m_s(\vartheta))\,\mathrm{d}s-\int_{t_0}^t b(\vartheta,s,m_s(\vartheta))\,\mathrm{d}W_s\right]\end{aligned}\tag{3.3}$$

and the limit as $\varepsilon\to0$ is: for all $t\in[t_0,T]$,

$$\int_{t_0}^t\frac{b(\vartheta,s,m_s^\circ(\vartheta))}{\sigma(s)}\left[f(\vartheta_0,s,Y_s)-f(\vartheta,s,m_s^\circ(\vartheta))\right]\mathrm{d}s=0\tag{3.4}$$

because the expression $[\ldots]$ in RHS of (3.3) is bounded, the functions $b(\vartheta,s)>c_b,\sigma(t)<C_\sigma$ and the equality (3.4) is possible iff

$$f(\vartheta_0,s,Y_s)=f(\vartheta,s,m_s^\circ(\vartheta)),\quad t_0\leqslant s\leqslant T.\tag{3.5}$$

By the condition $\mathscr{C}_2$ the function $f(\vartheta,t,y)$ is strictly monotone (increasing) in y and therefore the equation $f(\vartheta_0,t,Y_t)=f(\vartheta_0,t,m_t^\circ(\vartheta_0))$ has a unique solution $m_t^\circ(\vartheta_0)=Y_t$.

Recall that the relation similar to (3.2) we had in the linear case (1.1), (1.2) too (see the proof of Lemma 3 in [14])

$$f(\vartheta_0,s)\,Y_s=f(\vartheta,s)\,m_s^\circ(\vartheta),\quad t_0\leqslant s\leqslant T.$$

We suppose that the limit $m_t^\circ(\vartheta)$ of the solution of (3.1) is an Itô process with some stochastic differential

$$\mathrm{d}m_t^\circ(\vartheta)=A_t\mathrm{d}t+D_t\mathrm{d}V_t,\quad m_{t_0}^\circ(\vartheta),$$

then we use the equality (3.5) to define the random processes A_t, D_t, $t_0\leqslant t\leqslant T$. The last step will be to show that $m_t(\vartheta)-m_t^\circ(\vartheta)\to0$. Therefore if this convergence was proved we can say that the assumption that $m_t^\circ(\vartheta)$ is an Itô process was reasonable.

This stochastic differential we write with the help of (3.2) as follows:

$$
\begin{aligned}
& f_t'\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)\mathrm{d}t+f_y'\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)\mathrm{d}m_t^\circ\left(\vartheta\right)+\frac{1}{2}f_{yy}''\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)D_t^2\mathrm{d}t\\
&=f_t'\left(\vartheta_0,t,Y_t\right)\mathrm{d}t+f_y'\left(\vartheta_0,t,Y_t\right)a\left(\vartheta_0,t,Y_t\right)\mathrm{d}t+\frac{1}{2}f_{yy}''\left(\vartheta_0,t,Y_t\right)b\left(\vartheta_0,t,Y_t\right)^2\mathrm{d}t\\
&\quad+f_y'\left(\vartheta_0,t,Y_t\right)b\left(\vartheta_0,t,Y_t\right)\mathrm{d}V_t.
\end{aligned}
$$

Hence A_t and D_t are defined by the equation

$$
\mathrm{d}m_t^\circ\left(\vartheta\right)=A\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)\mathrm{d}t+\frac{f_y'\left(\vartheta_0,t,Y_t\right)b\left(\vartheta_0,t,Y_t\right)}{f_y'\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)}\mathrm{d}V_t,\quad t_0\leqslant t\leqslant T,\quad m_{t_0}^\circ\left(\vartheta\right),
$$

i.e.,

$$
\begin{aligned}
A_t=A\left(\vartheta,t,y\right)&=\frac{f_t'\left(\vartheta_0,t,Y_t\right)-f_t'\left(\vartheta,t,y\right)+f_y'\left(\vartheta_0,t,Y_t\right)a\left(\vartheta_0,t,Y_t\right)}{f_y'\left(\vartheta,t,y\right)}\\
&\quad+\frac{f_{yy}''\left(\vartheta_0,t,Y_t\right)b\left(\vartheta_0,t,Y_t\right)^2-f_{yy}''\left(\vartheta_0,t,y\right)D_t^2}{2f_y'\left(\vartheta,t,y\right)},\\
D_t=\frac{f_y'\left(\vartheta_0,t,Y_t\right)b\left(\vartheta_0,t,Y_t\right)}{f_y'\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)}&.
\end{aligned}
$$

The equation for $\dot m_t\left(\vartheta\right)$ we obtain by the formal differentiation of the equation (3.1)

$$
\begin{aligned}
\mathrm{d}\dot m_t\left(\vartheta\right)&=\left[\dot a_\vartheta\left(\vartheta,t,m_t\left(\vartheta\right)\right)+a_y'\left(\vartheta,t,m_t\left(\vartheta\right)\right)\dot m_t\left(\vartheta\right)\right]\mathrm{d}t+\dot B_\vartheta\left(\vartheta,t\right)\mathrm{d}W_t\\
&\quad+\frac{\dot B_\vartheta\left(\vartheta,t\right)}{\varepsilon\sigma\left(t\right)}\left[f\left(\vartheta_0,t,Y_t\right)-f\left(\vartheta,t,m_t\left(\vartheta\right)\right)\right]\mathrm{d}t\\
&\quad-\frac{B_\vartheta\left(\vartheta,t\right)}{\varepsilon\sigma\left(t\right)}\left[\dot f_\vartheta\left(\vartheta,t,m_t\left(\vartheta\right)\right)+f_y'\left(\vartheta,t,m_t\left(\vartheta\right)\right)\dot m_t\left(\vartheta\right)\right]\mathrm{d}t,
\end{aligned}
$$

where

$$
\dot B_\vartheta\left(\vartheta,t\right)=\dot b_\vartheta\left(\vartheta,t,m_t\left(\vartheta\right)\right)+b_y'\left(\vartheta,t,m_t\left(\vartheta\right)\right)\dot m_t\left(\vartheta\right).
$$

Therefore

$$
\begin{aligned}
&\int_{t_0}^{t}\left[\frac{\dot B_\vartheta\left(\vartheta,s\right)}{\sigma\left(s\right)}\left[f\left(\vartheta_0,s,Y_s\right)-f\left(\vartheta,s,m_s\left(\vartheta\right)\right)\right]\right]\mathrm{d}s\\
&\quad-\int_{t_0}^{t}\frac{b\left(\vartheta,s,m_s\left(\vartheta\right)\right)}{\sigma\left(s\right)}\left[\dot f_\vartheta\left(\vartheta,s,m_s\left(\vartheta\right)\right)+f_y'\left(\vartheta,t,m_t\left(\vartheta\right)\right)\dot m_t\left(\vartheta\right)\right]\mathrm{d}s\\
&=\varepsilon\left[\dot m_t\left(\vartheta\right)-\dot m_{t_0}\left(\vartheta\right)-\int_{t_0}^{t}\left[\dot a_\vartheta\left(\vartheta,s,m_t\left(\vartheta\right)\right)+a_y'\left(\vartheta,t,m_t\left(\vartheta\right)\right)\dot m_t\left(\vartheta\right)\right]\mathrm{d}s\right].
\end{aligned}
$$

The first integral converges to zero (see (3.4)) and we obtain the limit

$$
\int_{t_0}^{t}\frac{b\left(\vartheta,s,m_s^\circ\left(\vartheta\right)\right)}{\sigma\left(s\right)}\left[\dot f_\vartheta\left(\vartheta,s,m_s^\circ\left(\vartheta\right)\right)+f_y'\left(\vartheta,s,m_s^\circ\left(\vartheta\right)\right)\dot m_s^\circ\left(\vartheta\right)\right]\mathrm{d}s=0,\quad t_0<t\leqslant T.
$$

Hence

$$
\dot f_\vartheta\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)+f_y'\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)\dot m_t^\circ\left(\vartheta\right)=0,\quad \dot m_t^\circ\left(\vartheta\right)=-\frac{\dot f_\vartheta\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)}{f_y'\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)}. \tag{3.6}
$$

The last equality can be obtained as well by the differentiation in ϑ of the equation (3.5). If $\vartheta=\vartheta_0$, then

$$
\dot m_t^\circ\left(\vartheta_0\right)=-\frac{\dot f_\vartheta\left(\vartheta_0,t,m_t^\circ\left(\vartheta_0\right)\right)}{f_y'\left(\vartheta_0,t,m_t^\circ\left(\vartheta_0\right)\right)}=-\frac{\dot f_\vartheta\left(\vartheta_0,t,Y_t\right)}{f_y'\left(\vartheta_0,t,Y_t\right)}.
$$

□

Let us study the differences $m_t\left(\vartheta\right)-m_t^\circ\left(\vartheta\right)$ and $\dot m_t\left(\vartheta\right)-\dot m_t^\circ\left(\vartheta\right)$. Denote

$$\Gamma^\circ\left(\vartheta,t\right)=\sqrt{\frac{b\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)\sigma\left(t\right)}{f_y'\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)}},\quad \Gamma^\circ\left(\vartheta_0,t\right)=\sqrt{\frac{b\left(\vartheta_0,t,Y_t\right)\sigma\left(t\right)}{f_y'\left(\vartheta_0,t,Y_t\right)}},$$

$$\Delta^\circ\left(\vartheta,t\right)=\frac{b\left(\vartheta_0,t,Y_t\right)f_y'\left(\vartheta_0,t,Y_t\right)}{b\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)f_y'\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)},\quad \Delta^\circ\left(\vartheta_0,t\right)=1,$$

$$\dot B_\vartheta\left(\vartheta,t\right)=\dot b\left(\vartheta,t,m_t\left(\vartheta\right)\right)+b_y'\left(\vartheta,t,m_t\left(\vartheta\right)\right)\dot m_t\left(\vartheta\right),$$

$$\dot B_\vartheta^\circ\left(\vartheta,t\right)=\dot b\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)-\frac{b_y'\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)\dot f_\vartheta\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)}{f_y'\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)},$$

$$R^\circ\left(\vartheta,t\right)=\frac{f_{yy}''\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)\dot f_\vartheta\left(\vartheta_0,t,m_t^\circ\left(\vartheta\right)\right)}{f_y'\left(\vartheta_0,t,m_t^\circ\left(\vartheta\right)\right)}-\dot f_{\vartheta y}'\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)-\frac{\dot B_\vartheta^\circ\left(\vartheta,t\right)}{b\left(\vartheta,t\right)}f_y'\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right),$$

$$\Pi^\circ\left(\vartheta,t\right)=\frac{R^\circ\left(\vartheta,t\right)}{f_y'\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)}+\frac{\dot B_\vartheta^\circ\left(\vartheta,t\right)^2}{b\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)^2},\quad \Pi^\circ\left(\vartheta_0,t\right)=\frac{R^\circ\left(\vartheta_0,t\right)}{f_y'\left(\vartheta_0,t,Y_t\right)}+\frac{\dot B_\vartheta^\circ\left(\vartheta_0,t\right)^2}{b\left(\vartheta_0,t,Y_t\right)^2},$$

$$\begin{aligned}\Lambda^\circ\left(\vartheta,t\right)=&\frac{b\left(\vartheta_0,t,Y_t\right)f_y'\left(\vartheta_0,t,Y_t\right)}{b\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)f_y'\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)}\left[\frac{\dot B_\vartheta^\circ\left(\vartheta,t\right)\dot f_{y\vartheta}'\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)}{b\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)f_y'\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)}\right.\\ &\left.-\frac{R^\circ\left(\vartheta,t\right)}{f_y'\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)}-\frac{\dot B_\vartheta^\circ\left(\vartheta,t\right)f_{yy}''\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)\dot f_\vartheta\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)}{b\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)f_y'\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)^2}\right],\end{aligned}$$

$$\Lambda^\circ\left(\vartheta_0,t\right)=\frac{\dot B_\vartheta^\circ\left(\vartheta_0,t\right)\dot f_{y\vartheta}'\left(\vartheta_0,t,Y_t\right)}{b\left(\vartheta_0,t,Y_t\right)f_y'\left(\vartheta_0,t,Y_t\right)}-\frac{\dot B_\vartheta^\circ\left(\vartheta_0,t\right)f_{yy}''\left(\vartheta_0,t,Y_t\right)\dot f_\vartheta\left(\vartheta_0,t,Y_t\right)}{b\left(\vartheta_0,t,Y_t\right)f_y'\left(\vartheta_0,t,Y_t\right)^2}-\frac{R^\circ\left(\vartheta_0,t\right)}{f_y'\left(\vartheta_0,t,Y_t\right)}.$$

Below the expressions like $O\left(\varepsilon\right)$ are understood *in probability*, i.e., it means that $\varepsilon^{-1}O\left(\varepsilon\right)$ is bounded in probability.

Proposition 1 *Let the conditions $\mathscr{C}$ be satisfied, then for any $t_0\in(0,T]$ the following relations hold*

$$\frac{m_t\left(\vartheta\right)-m_t^\circ\left(\vartheta\right)}{\Gamma^\circ\left(\vartheta,t\right)\sqrt{\varepsilon}}=\xi_{t,\varepsilon}-\Delta^\circ\left(\vartheta,t\right)\eta_{t,\varepsilon}+O\left(\varepsilon^{\frac{\nu}{2}}\right),\tag{3.7}$$

$$\frac{m_t\left(\vartheta_0\right)-Y_t}{\Gamma^\circ\left(\vartheta_0,t\right)\sqrt{\varepsilon}}=\xi_{t,\varepsilon}-\eta_{t,\varepsilon}+O\left(\varepsilon^{\frac{\nu}{2}}\right),$$

$$\frac{\dot m_t\left(\vartheta\right)-\dot m_t^\circ\left(\vartheta\right)}{\Gamma^\circ\left(\vartheta,t\right)\sqrt{\varepsilon}}\Longrightarrow\Pi^\circ\left(\vartheta,t\right)\,\xi_t+\Lambda^\circ\left(\vartheta,t\right)\,\eta_t.\tag{3.8}$$

Here $\xi_{t,\varepsilon}\Rightarrow\xi_t\sim\mathcal{N}\left(0,1/2\right)$, $\eta_{t,\varepsilon}\Rightarrow\eta_t\sim\mathcal{N}\left(0,1/2\right)$, r.v.'s $\xi_t,\ 0<t\leqslant T$ are independent, r.v.'s $\eta_t,\ 0<t\leqslant T$ are independent and r.v.'s $\xi_t,\ 0<t\leqslant T$ and $\eta_t,\ 0<t\leqslant T$ are independent too.

Proof The equation (3.1) can be written as follows:

$$\begin{aligned}\mathrm{d}m_t\left(\vartheta\right)&=a\left(\vartheta,t,m_t\left(\vartheta\right)\right)\mathrm{d}t+\frac{b\left(\vartheta,t,m_t\left(\vartheta\right)\right)}{\varepsilon\sigma\left(t\right)}\left[f\left(\vartheta_0,t,Y_t\right)-f\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)\right]\mathrm{d}t\\ &\quad+\frac{b\left(\vartheta,t,m_t\left(\vartheta\right)\right)}{\varepsilon\sigma\left(t\right)}\left[f\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)-f\left(\vartheta,t,m_t\left(\vartheta\right)\right)\right]+b\left(\vartheta,t,m_t\left(\vartheta\right)\right)\mathrm{d}W_t\\ &=a\left(\vartheta,t,m_t\left(\vartheta\right)\right)\mathrm{d}t+\frac{b\left(\vartheta,t,m_t\left(\vartheta\right)\right)}{\varepsilon\sigma\left(t\right)}\left[f\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)-f\left(\vartheta,t,m_t\left(\vartheta\right)\right)\right]\mathrm{d}t\\ &\quad+b\left(\vartheta,t,m_t\left(\vartheta\right)\right)\mathrm{d}W_t\\ &=a\left(\vartheta,t,m_t\left(\vartheta\right)\right)\mathrm{d}t-\frac{b\left(\vartheta,t,m_t\left(\vartheta\right)\right)}{\varepsilon\sigma\left(t\right)}f_y'\left(\vartheta,t,\tilde m_t\left(\vartheta\right)\right)\left[m_t\left(\vartheta\right)-m_t^\circ\left(\vartheta\right)\right]\mathrm{d}t\\ &\quad+b\left(\vartheta,t,m_t\left(\vartheta\right)\right)\mathrm{d}W_t,\end{aligned}$$

where $|\tilde m_t(\vartheta)-m_t^\circ(\vartheta)|\leqslant|m_t(\vartheta)-m_t^\circ(\vartheta)|$. Therefore

$$\begin{aligned}\mathrm{d}\left[m_t(\vartheta)-m_t^\circ(\vartheta)\right]&=\left[a(\vartheta,t,m_t(\vartheta))-A(\vartheta,t,m_t^\circ(\vartheta))\right]\mathrm{d}t\\&\quad-\frac{f_y'(\vartheta_0,t,Y_t)\,b(\vartheta_0,t,Y_t)}{f_y'(\vartheta,t,m_t(\vartheta))}\mathrm{d}V_t+b(\vartheta,t,m_t(\vartheta))\,\mathrm{d}W_t\\&\quad-\frac{b(\vartheta,t,m_t(\vartheta))}{\varepsilon\sigma(t)}f_y'(\vartheta,t,\tilde m_t(\vartheta))\left[m_t(\vartheta)-m_t^\circ(\vartheta)\right]\mathrm{d}t.\end{aligned}\tag{3.9}$$

Introduce the random process $Z_t=m_t(\vartheta)-m_t^\circ(\vartheta)-C_t+F_t$, where

$$C_t=\int_{t_0}^t b(\vartheta,s,m_s(\vartheta))\,\mathrm{d}W_s,\quad F_t=\int_{t_0}^t\frac{f_y'(\vartheta_0,s,Y_s)\,b(\vartheta_0,s,Y_s)}{f_y'(\vartheta,s,m_s(\vartheta))}\mathrm{d}V_s.$$

The equation for Z_t is

$$\begin{aligned}\mathrm{d}Z_t&=\left[a(\vartheta,t,m_t(\vartheta))-A(\vartheta,t,m_t^\circ(\vartheta))\right]\ \mathrm{d}t-\frac{b(\vartheta,t,m_t(\vartheta))}{\varepsilon\sigma(t)}f_y'(\vartheta,t,\tilde m_t(\vartheta))\,Z_t\ \mathrm{d}t\\&\quad+\frac{b(\vartheta,t,m_t(\vartheta))}{\varepsilon\sigma(t)}f_y'(\vartheta,t,\tilde m_t(\vartheta))\left[F_t-C_t\right]\ \mathrm{d}t.\end{aligned}$$

Therefore the solution of this equation is

$$\begin{aligned}Z_t&=Z_{t_0}\Psi(t,t_0)+\frac{1}{\varepsilon}\int_{t_0}^t\Psi(t,s)\,\frac{b(\vartheta,s,m_s(\vartheta))}{\sigma(s)}f_y'(\vartheta,s,\tilde m_s(\vartheta))\left[F_s-C_s\right]\mathrm{d}s\\&\quad+\int_{t_0}^t\Psi(t,s)\left[a(\vartheta,s,m_s(\vartheta))-A(\vartheta,s,m_s^\circ(\vartheta))\right]\ \mathrm{d}s\\&=F_t-C_t+\frac{1}{\varepsilon}\int_{t_0}^t\Psi(t,s)\,\frac{b(\vartheta,s,m_s(\vartheta))}{\sigma(s)}f_y'(\vartheta,s,\tilde m_s(\vartheta))\left[F_s-F_t\right]\mathrm{d}s\\&\quad+\int_{t_0}^t\Psi(t,s)\left[a(\vartheta,s,m_s(\vartheta))-A(\vartheta,s,m_s^\circ(\vartheta))\right]\ \mathrm{d}s+O\left(e^{-\frac{c}{\varepsilon}}\right)\\&\quad+\frac{1}{\varepsilon}\int_{t_0}^t\Psi(t,s)\,\frac{b(\vartheta,s,m_s(\vartheta))}{\sigma(s)}f_y'(\vartheta,s,\tilde m_s(\vartheta))\left[C_t-C_s\right]\mathrm{d}s+Z_{t_0}\Psi(t,t_0),\end{aligned}$$

where we used the abbreviations

$$\Psi(t,s)=\exp\left(-\frac{1}{\varepsilon}\int_s^t q(v)\,\mathrm{d}v\right),\quad q(t)=\frac{b(\vartheta,t,m_t(\vartheta))}{\sigma(t)}f_y'(\vartheta,t,\tilde m_t(\vartheta))$$

and the following property of the integral ($t>t_0$).

$$\begin{aligned}&\frac{1}{\varepsilon}\int_{t_0}^t\Psi(t,s)\,\frac{b(\vartheta,s,m_s(\vartheta))}{\sigma(s)}f_y'(\vartheta,s,\tilde m_s(\vartheta))\ \mathrm{d}s\\&=\Psi(t,0)\,\frac{1}{\varepsilon}\int_{t_0}^t\Psi(s,0)^{-1}\,\frac{b(\vartheta,s,m_s(\vartheta))}{\sigma(s)}f_y'(\vartheta,s,\tilde m_s(\vartheta))\ \mathrm{d}s\\&=\Psi(t,0)\int_{t_0}^t\mathrm{d}\Psi(s,0)^{-1}=1-\Psi(t,t_0)\\&=1+O\left(e^{-\frac{c}{\varepsilon}(t-t_0)}\right),\end{aligned}$$

with some $c>0$.

Denote as well

$$q_*(t)=\frac{b(\vartheta,t,m_t(\vartheta))}{\sigma(t)}f_y'(\vartheta,t,m_t(\vartheta)),\quad\Psi_*(t,s)=\exp\left(-\frac{1}{\varepsilon}(t-s)\,q_*(t)\right).$$

We have the estimates $(t > t_0)$

$$Z_{t_0}\Psi\left(t,t_0\right)=O\left(\varepsilon\right),\qquad \int_{t_0}^{t}\Psi\left(t,s\right)\left[a\left(\vartheta,s,m_s\left(\vartheta\right)\right)-A\left(\vartheta,s,m_s^\circ\left(\vartheta\right),Y_s\right)\right]\,\mathrm{d}s=O\left(\varepsilon\right).$$

To obtain these estimates and to study the other integrals below we use the following well known device used, for example, in the proof of Lemma 1 in [14]. Let $G\left(s\right),\ t_0\leqslant s\leqslant T$ be a bounded function and $t_\varepsilon=t-\varepsilon^\nu,\ s>t_\varepsilon$, where $\nu\in(1/2,1)$. We have the estimate

$$q\left(t\right)\geqslant\frac{c_bc_y}{C_\sigma}=\kappa_*>0,$$

where the constants c_b,c_y are from the condition $\mathscr{C}_2$ and $\sigma\left(t\right)\leqslant C_\sigma$. Hence

$$\Psi\left(t,s\right)\leqslant e^{-\frac{\kappa_*}{\varepsilon}\left(t-s\right)}$$

and

$$\int_{t_0}^{t_\varepsilon}\Psi\left(t,s\right)G\left(s\right)\,\mathrm{d}s\leqslant\int_{t_0}^{t_\varepsilon}e^{-\frac{\kappa_*}{\varepsilon}\left(t-s\right)}G\left(s\right)\,\mathrm{d}s\leqslant e^{-\frac{\kappa_*}{\varepsilon^{1-\nu}}}\int_{t_0}^{t_\varepsilon}G\left(s\right)\,\mathrm{d}s=O\left(e^{-\frac{\kappa_*}{\varepsilon^{1-\nu}}}\right).$$

Then

$$\begin{aligned}&\int_{t_0}^{t}\Psi\left(t,s\right)\frac{b\left(\vartheta,s,m_s\left(\vartheta\right)\right)}{\sigma\left(s\right)}f_y'\left(\vartheta,s,\tilde m_s\left(\vartheta\right)\right)G\left(s\right)\,\mathrm{d}s\\&=\int_{t_0}^{t_\varepsilon}\Psi\left(t,s\right)\frac{b\left(\vartheta,s,m_s\left(\vartheta\right)\right)}{\sigma\left(s\right)}f_y'\left(\vartheta,s,\tilde m_s\left(\vartheta\right)\right)G\left(s\right)\,\mathrm{d}s\\&\quad+\int_{t_\varepsilon}^{t}\Psi\left(t,s\right)\frac{b\left(\vartheta,s,m_s\left(\vartheta\right)\right)}{\sigma\left(s\right)}f_y'\left(\vartheta,s,\tilde m_s\left(\vartheta\right)\right)G\left(s\right)\,\mathrm{d}s\\&=\int_{t_\varepsilon}^{t}\Psi\left(t,s\right)\frac{b\left(\vartheta,s,m_s\left(\vartheta\right)\right)}{\sigma\left(s\right)}f_y'\left(\vartheta,s,\tilde m_s\left(\vartheta\right)\right)G\left(s\right)\,\mathrm{d}s+O\left(e^{-\frac{\kappa_*}{\varepsilon^{1-\nu}}}\right).\end{aligned}$$

As the terms like $O\left(e^{-\frac{\kappa_*}{\varepsilon^{1-\nu}}}\right)$ will be added to more large errors we will write $O\left(\varepsilon\right)$ to simplify the exposition.

Lemma 2 *If the function $G\left(\cdot\right)$ is Hölder of order $\mu\in(0,1]$ then*

$$\frac{1}{\varepsilon}\int_{t_0}^{t}\Psi\left(t,s\right)\frac{b\left(\vartheta,s,m_s\left(\vartheta\right)\right)}{\sigma\left(s\right)}f_y'\left(\vartheta,s,\tilde m_s\left(\vartheta\right)\right)G\left(s\right)\,\mathrm{d}s=G\left(t\right)+O\left(\varepsilon^\mu\right).\tag{3.10}$$

Proof Omitting terms of order $o\left(\varepsilon^\mu\right)$ we can write

$$\begin{aligned}&\left|\frac{1}{\varepsilon}\int_{t_\varepsilon}^{t}\Psi\left(t,s\right)\frac{b\left(\vartheta,s,m_s\left(\vartheta\right)\right)}{\sigma\left(s\right)}f_y'\left(\vartheta,s,\tilde m_s\left(\vartheta\right)\right)G\left(s\right)\,\mathrm{d}s-G\left(t\right)\right|\\&\leqslant\frac{1}{\varepsilon}\int_{t_\varepsilon}^{t}\Psi\left(t,s\right)\frac{b\left(\vartheta,s,m_s\left(\vartheta\right)\right)}{\sigma\left(s\right)}f_y'\left(\vartheta,s,\tilde m_s\left(\vartheta\right)\right)\left|G\left(s\right)-G\left(t\right)\right|\,\mathrm{d}s\\&\leqslant\frac{C}{\varepsilon}\int_{t_\varepsilon}^{t}e^{-\frac{\kappa_*}{\varepsilon}\left(t-s\right)}\left|t-s\right|^\mu\,\mathrm{d}s\leqslant C\int_0^{\frac{\kappa_*\left(t-t_\varepsilon\right)}{\varepsilon}}e^{-u}u^\mu\mathrm{d}u\,\varepsilon^\mu\leqslant C\,\varepsilon^\mu.\end{aligned}$$

These estimates allow us to write for such functions $G\left(\cdot\right)$ the representation (3.10). □

We have on the interval $[t_\varepsilon,t]$ the estimates

$$\begin{aligned}\left|q_*\left(t\right)-q\left(s\right)\right|&=\left|\frac{b\left(\vartheta,t,m_t\left(\vartheta\right)\right)}{\sigma\left(t\right)}f_y'\left(\vartheta,t,m_t\left(\vartheta\right)\right)-\frac{b\left(\vartheta,s,m_s\left(\vartheta\right)\right)}{\sigma\left(s\right)}f_y'\left(\vartheta,s,\tilde m_s\left(\vartheta\right)\right)\right|\\&\leqslant C\left|t-s\right|+C\left|f_y'\left(\vartheta,t,m_t\left(\vartheta\right)\right)-f_y'\left(\vartheta,t,\tilde m_s\left(\vartheta\right)\right)\right|\\&\leqslant C\,\varepsilon^\nu+C\left|m_t\left(\vartheta\right)-m_s\left(\vartheta\right)\right|+C\left|m_s\left(\vartheta\right)-m_s^\circ\left(\vartheta\right)\right|.\end{aligned}$$

Hence

$$\begin{aligned}Z_t &= m_t\left(\vartheta\right)-m_t^\circ\left(\vartheta\right)-C_t+F_t\\&=F_t-C_t+\frac{1}{\varepsilon}\int_{t_\varepsilon}^{t}\Psi\left(t,s\right)\frac{b\left(\vartheta,s,m_s\left(\vartheta\right)\right)}{\sigma\left(s\right)}f_y'\left(\vartheta,s,\tilde m_s\left(\vartheta\right)\right)\left[F_s-F_t\right]{\rm d}s\\&\qquad+\frac{1}{\varepsilon}\int_{t_\varepsilon}^{t}\Psi\left(t,s\right)\frac{b\left(\vartheta,s,m_s\left(\vartheta\right)\right)}{\sigma\left(s\right)}f_y'\left(\vartheta,s,\tilde m_s\left(\vartheta\right)\right)\left[C_t-C_s\right]{\rm d}s+O\left(\varepsilon\right)\end{aligned}$$

and we can write

$$\begin{aligned}m_t\left(\vartheta\right)-m_t^\circ\left(\vartheta\right)&=\frac{1}{\varepsilon}\int_{t_\varepsilon}^{t}\Psi\left(t,s\right)\frac{b\left(\vartheta,s,m_s\left(\vartheta\right)\right)}{\sigma\left(s\right)}f_y'\left(\vartheta,s,\tilde m_s\left(\vartheta\right)\right)\left[F_s-F_t\right]{\rm d}s\\&\qquad+\frac{1}{\varepsilon}\int_{t_\varepsilon}^{t}\Psi\left(t,s\right)\frac{b\left(\vartheta,s,m_s\left(\vartheta\right)\right)}{\sigma\left(s\right)}f_y'\left(\vartheta,s,\tilde m_s\left(\vartheta\right)\right)\left[C_t-C_s\right]{\rm d}s+O\left(\varepsilon\right).\end{aligned}$$

We estimate the order of $m_t\left(\vartheta\right)-m_t^\circ\left(\vartheta\right)$ first. We have

$$\begin{aligned}\mathbf{E}_{\vartheta_0}\left|m_t\left(\vartheta\right)-m_t^\circ\left(\vartheta\right)\right|&\leqslant\frac{C}{\varepsilon}\int_{t_\varepsilon}^{t}\Psi_*\left(t,s\right)\mathbf{E}_{\vartheta_0}\left(\left|F_s-F_t\right|+\left|C_s-C_t\right|\right){\rm d}s+O\left(\varepsilon\right)\\&\leqslant\frac{C}{\varepsilon}\int_{t_\varepsilon}^{t}e^{-\frac{\kappa_*}{\varepsilon}\left(t-s\right)}\left|t-s\right|^{1/2}{\rm d}s+O\left(\varepsilon\right)=O\left(\sqrt{\varepsilon}\right).\end{aligned}$$

Therefore $m_t\left(\vartheta\right)-m_t^\circ\left(\vartheta\right)=O\left(\sqrt{\varepsilon}\right)$ and

$$\frac{1}{t-t_\varepsilon}\int_{t_\varepsilon}^{t}\frac{b\left(\vartheta,v,m_v\left(\vartheta\right)\right)}{\sigma\left(v\right)}f_y'\left(\vartheta,v,\tilde m_v\left(\vartheta\right)\right){\rm d}v\longrightarrow\frac{b\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)}{\sigma\left(t\right)}f_y'\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)=q^\circ\left(t\right).$$

We have as well

$$\begin{aligned}&\frac{1}{\varepsilon}\int_{s}^{t}\frac{b\left(\vartheta,r,m_r\left(\vartheta\right)\right)}{\sigma\left(r\right)}f_y'\left(\vartheta,r,\tilde m_r\left(\vartheta\right)\right){\rm d}r-\frac{\left(t-s\right)}{\varepsilon}q\left(t\right)\\&=\frac{\left(t-s\right)}{\varepsilon}\Big[O\left(\varepsilon^{\nu}\right)+O\left(\varepsilon^{\nu/2}\right)+O\left(\sqrt{\varepsilon}\right)\Big]=\frac{\left(t-s\right)}{\varepsilon^{1-\frac{\nu}{2}}}O\left(1\right)=\frac{\left(t-s\right)}{\varepsilon}h_\varepsilon,\\&\qquad h_\varepsilon=O\left(\varepsilon^{\frac{\nu}{2}}\right).\end{aligned}$$

To study these integrals we use the elements of the proof of Lemma 1 [14], which allow us to write

$$\begin{aligned}&\frac{1}{\varepsilon}\int_{t_\varepsilon}^{t}\Psi\left(t,s\right)\frac{b\left(\vartheta,s,m_s\left(\vartheta\right)\right)}{\sigma\left(s\right)}f_y'\left(\vartheta,s,\tilde m_s\left(\vartheta\right)\right)\left[C_t-C_s\right]{\rm d}s\\&=\frac{1}{\varepsilon}\int_{t_\varepsilon}^{t}\Psi\left(t,s\right)\frac{b\left(\vartheta,s,m_s\left(\vartheta\right)\right)}{\sigma\left(s\right)}f_y'\left(\vartheta,s,\tilde m_s\left(\vartheta\right)\right)\int_{s}^{t}b\left(\vartheta,r,m_r\left(\vartheta\right)\right){\rm d}W_r\ {\rm d}s\\&=\frac{1}{\varepsilon}\int_{t_\varepsilon}^{t}e^{-\frac{1}{\varepsilon}\left(t-s\right)\left[q\left(t\right)+h_\varepsilon\right]}\frac{b\left(\vartheta,s,m_s\left(\vartheta\right)\right)}{\sigma\left(s\right)}f_y'\left(\vartheta,s,\tilde m_s\left(\vartheta\right)\right)\int_{s}^{t}b\left(\vartheta,r,m_r\left(\vartheta\right)\right){\rm d}W_r\ {\rm d}s\\&=\frac{b\left(\vartheta,t,m_t\left(\vartheta\right)\right)}{\varepsilon\sigma\left(t\right)}f_y'\left(\vartheta,t,\tilde m_t\left(\vartheta\right)\right)\int_{t_\varepsilon}^{t}e^{-\frac{1}{\varepsilon}\left(t-s\right)q\left(t\right)}\int_{s}^{t}b\left(\vartheta,r,m_r\left(\vartheta\right)\right){\rm d}W_r\ {\rm d}s\left(1+O\left(\varepsilon^{\frac{\nu}{2}}\right)\right)\\&=\frac{b\left(\vartheta,t,m_t\left(\vartheta\right)\right)^2}{\varepsilon\sigma\left(t\right)}f_y'\left(\vartheta,t,\tilde m_t\left(\vartheta\right)\right)\int_{t_\varepsilon}^{t}e^{-\frac{1}{\varepsilon}\left(t-s\right)q\left(t\right)}\left[W_t-W_s\right]\ {\rm d}s\left(1+O\left(\varepsilon^{\frac{\nu}{2}}\right)\right)\\&=\sqrt{\varepsilon}\sqrt{\frac{b\left(\vartheta,t,m_t\left(\vartheta\right)\right)\sigma\left(t\right)}{f_y'\left(\vartheta,t,\tilde m_t\left(\vartheta\right)\right)}}\int_{0}^{\frac{\left(t-t_\varepsilon\right)q\left(t\right)}{\varepsilon}}e^{-u}\,w_{t,\varepsilon}\left(u\right)\,{\rm d}u\left(1+O\left(\varepsilon^{\frac{\nu}{2}}\right)\right)\\&=\sqrt{\varepsilon}\sqrt{\frac{b\left(\vartheta,t,m_t\left(\vartheta\right)\right)\sigma\left(t\right)}{f_y'\left(\vartheta,t,\tilde m_t\left(\vartheta\right)\right)}}\ \xi_{t,\varepsilon}\left(1+O\left(\varepsilon^{\frac{\nu}{2}}\right)\right),\end{aligned}$$

where we changed the variable $s = t - \frac{\varepsilon}{q(t)} u$ and denoted $w_{t,\varepsilon}(u) = \sqrt{\frac{q(t)}{\varepsilon}} \left[W_t - W_{t-\frac{\varepsilon u}{q(t)}} \right]$ the Wiener process.

Note that

$$\int_0^{\frac{(t-t_\varepsilon)q(t)}{\varepsilon}} e^{-u}\, w_{t,\varepsilon}(u)\, \mathrm{d}u = \int_0^{\frac{(t-t_\varepsilon)q(t)}{\varepsilon}} e^{-u}\, \mathrm{d}w_{t,\varepsilon}(u) + O\left(e^{-\frac{c}{\varepsilon}}\right)$$
$$\Longrightarrow \xi_t = \int_0^\infty e^{-u}\, \mathrm{d}w_t(u) \sim \mathcal{N}\left(0, \frac{1}{2}\right).$$

Here $w_t(\cdot)$ is some Wiener process and the Wiener processes $w_{t_1}(\cdot)$ and $w_{t_2}(\cdot)$ for any $t_1 \neq t_2$ are independent. Therefore the random variables $\xi_t,\ t_0 < t \leqslant T$ are independent.

For the first integral we have the similar relations

$$\begin{aligned}
&\frac{1}{\varepsilon} \int_{t_\varepsilon}^t \Psi(t,s) \frac{b(\vartheta, s, m_s(\vartheta))}{\sigma(s)} f'_y(\vartheta, s, \tilde m_s(\vartheta)) \int_s^t \frac{f'_y(\vartheta_0, r, Y_r)\, b(\vartheta_0, r, Y_r)}{f'_y(\vartheta, r, m_r(\vartheta))} \mathrm{d}V_r\, \mathrm{d}s \\
&= \frac{b(\vartheta, t, \bar m_t(\vartheta))\, f'_y(\vartheta, t, \tilde m_t(\vartheta))\, f'_y(\vartheta_0, t, Y_t)\, b(\vartheta_0, t, Y_t)}{\varepsilon \sigma(t)\, f'_y(\vartheta, t, \bar m_t(\vartheta))} \\
&\quad \times \int_{t_\varepsilon}^t e^{-\frac{1}{\varepsilon}(t-s)q(t)} \left[V_t - V_s\right] \mathrm{d}s \left(1 + O\left(\varepsilon^{\frac{\nu}{2}}\right)\right) \\
&= \sqrt{\varepsilon}\, \frac{b(\vartheta_0, t, Y_t)\, f'_y(\vartheta_0, t, Y_t)}{b(\vartheta, t, \bar m_t(\vartheta))\, f'_y(\vartheta, t, m_t(\vartheta))} \sqrt{\frac{b(\vartheta, t, \bar m_t(\vartheta))\, \sigma(t)}{f'_y(\vartheta, t, \tilde m_t(\vartheta))}}\ \eta_{t,\varepsilon} \left(1 + O\left(\varepsilon^{\frac{\nu}{2}}\right)\right) \\
&= \sqrt{\varepsilon}\, \Delta(\vartheta, t) \sqrt{\frac{b(\vartheta, t, \bar m_t(\vartheta))\, \sigma(t)}{f'_y(\vartheta, t, \tilde m_t(\vartheta))}}\ \eta_{t,\varepsilon} \left(1 + O\left(\varepsilon^{\frac{\nu}{2}}\right)\right).
\end{aligned}$$

Here

$$\eta_{t,\varepsilon} = \int_0^{\frac{(t-t_\varepsilon)q(t)}{\varepsilon}} e^{-u}\, v_{t,\varepsilon}(u)\, \mathrm{d}u, \quad v_{t,\varepsilon}(u) = \sqrt{\frac{q(t)}{\varepsilon}} \left[V_t - V_{t-\frac{\varepsilon u}{q(t)}}\right]$$

and

$$\eta_{t,\varepsilon} = \int_0^{\frac{(t-t_\varepsilon)q(t)}{\varepsilon}} e^{-u}\, \mathrm{d}v_{t,\varepsilon}(u) + O\left(e^{-\frac{c}{\varepsilon}}\right) \Longrightarrow \eta_t = \int_0^\infty e^{-u}\, \mathrm{d}v_t(u) \sim \mathcal{N}\left(0, \frac{1}{2}\right).$$

The Wiener processes $w_t(\cdot)$ and $v_t(\cdot)$ are independent and the random variables η_{t_1}, η_{t_2}, $t_0 \leqslant t \leqslant T$ for any $t_1 \neq t_2$ are independent too.

We obtained the representation

$$\begin{aligned}
\frac{m_t(\vartheta) - m_t^\circ(\vartheta)}{\sqrt{\varepsilon}} &= \sqrt{\frac{\sigma(t)\, b(\vartheta, t, \bar m_t(\vartheta))}{f'_y(\vartheta, t, \tilde m_t(\vartheta))}} \left[\xi_{t,\varepsilon} - \frac{b(\vartheta_0, t, Y_t)\, f'_y(\vartheta_0, t, Y_t)}{b(\vartheta, t, \bar m_t(\vartheta)) f'_y(\vartheta, t, m_t(\vartheta))} \eta_{t,\varepsilon}\right] + O\left(\varepsilon^{\frac{\nu}{2}}\right) \\
&= \sqrt{\frac{\sigma(t)\, b(\vartheta, t, m_t^\circ(\vartheta))}{f'_y(\vartheta, t, m_t^\circ(\vartheta))}} \left[\xi_{t,\varepsilon} - \frac{b(\vartheta_0, t, Y_t)\, f'_y(\vartheta_0, t, Y_t)}{b(\vartheta, t, m_t^\circ(\vartheta)) f'_y(\vartheta, t, m_t^\circ(\vartheta))}\ \eta_{t,\varepsilon}\right] + O\left(\varepsilon^{\frac{\nu}{2}}\right) \\
&= \Gamma^\circ(\vartheta, t) \left[\xi_{t,\varepsilon} - \Delta^\circ(\vartheta, t)\, \eta_{t,\varepsilon}\right] + O\left(\varepsilon^{\frac{\nu}{2}}\right).
\end{aligned}$$

Note that as $m_t(\vartheta) - m_t^\circ(\vartheta) \approx \sqrt{\varepsilon}$, we replaced in this expression $m_t(\vartheta)$ and $\tilde m_t(\vartheta)$ by $m_t^\circ(\vartheta)$.

If $\vartheta = \vartheta_0$, then $m_t^\circ(\vartheta_0) = Y_t$ (see (2.4)) and

$$\begin{aligned}
m_t(\vartheta_0) - Y_t &= \Gamma^\circ(\vartheta_0, t) \left[\xi_{t,\varepsilon} - \Delta^\circ(\vartheta_0, t)\, \eta_{t,\varepsilon}\right] \sqrt{\varepsilon} \left(1 + O\left(\varepsilon^{\frac{\nu}{2}}\right)\right) \\
&= \Gamma^\circ(\vartheta_0, t) \left[\xi_{t,\varepsilon} - \eta_{t,\varepsilon}\right]\ \sqrt{\varepsilon} \left(1 + O\left(\varepsilon^{\frac{\nu}{2}}\right)\right).
\end{aligned}$$

Therefore we can formally evaluate the error

$$\begin{aligned}\mathbf{E}_{\vartheta_0}\left(m_t\left(\vartheta_0\right)-Y_t\right)^2&=\sigma\left(t\right)\mathbf{E}_{\vartheta_0}\left[\frac{b\left(\vartheta_0,t,Y_t\right)}{f'_y\left(\vartheta_0,t,Y_t\right)}\right]\ \varepsilon+O\left(\varepsilon^{1+\frac{\nu}{2}}\right)\\&=\mathbf{E}_{\vartheta_0}\gamma_*\left(\vartheta_0,t\right)\ \varepsilon+O\left(\varepsilon^{1+\frac{\nu}{2}}\right).\end{aligned}$$

Hence the function $\mathbf{E}_{\vartheta_0}\gamma_*\left(\vartheta_0,t\right)$ asymptoticaly plays the same role as in the linear case, i.e., describes the mean squared error of estimation Y_t by $m_t\left(\vartheta_0\right)$.

Let us denote

$$P\left(\vartheta,t,y,z\right)=a\left(\vartheta,t,y\right)-A\left(\vartheta,t,z\right)$$

and consider now the difference $\dot m_t\left(\vartheta\right)-\dot m_t^\circ\left(\vartheta\right)$. The derivative in ϑ of (3.9) yields the equation

$$\begin{aligned}&\mathrm{d}\left[\dot m_t\left(\vartheta\right)-\dot m_t^\circ\left(\vartheta\right)\right]\\&=\dot P_\vartheta\left(\vartheta,t,m_t\left(\vartheta\right),m_t^\circ\left(\vartheta\right)\right)\mathrm{d}t+P'_y\left(\vartheta,t,m_t\left(\vartheta\right),m_t^\circ\left(\vartheta\right)\right)\dot m_t\left(\vartheta\right)\mathrm{d}t+P'_z\left(\vartheta,t,m_t\left(\vartheta\right),m_t^\circ\left(\vartheta\right)\right)\dot m_t^\circ\left(\vartheta\right)\mathrm{d}t\\&\quad+\frac{f'_y\left(\vartheta_0,t,Y_t\right)b\left(\vartheta_0,t,Y_-\right)}{f'_y\left(\vartheta,t,m_t\left(\vartheta\right)\right)^2}\left[\dot f'_{y\vartheta}\left(\vartheta,t,m_t\left(\vartheta\right)\right)+f''_{yy}\left(\vartheta,t,m_t\left(\vartheta\right)\right)\dot m_t\left(\vartheta\right)\right]\mathrm{d}V_t\\&\quad-\frac{\left[\dot B_\vartheta\left(\vartheta,t\right)f'_y\left(\vartheta,t,\tilde m_t\left(\vartheta\right)\right)+b\left(\vartheta,t,m_t\left(\vartheta\right)\right)\dot f'_{\vartheta y}\left(\vartheta,t,\tilde m_t\left(\vartheta\right)\right)\right]}{\varepsilon\sigma\left(t\right)}\left[m_t\left(\vartheta\right)-m_t^\circ\left(\vartheta\right)\right]\mathrm{d}t\\&\quad-\frac{b\left(\vartheta,t,m_t\left(\vartheta\right)\right)}{\varepsilon\sigma\left(t\right)}\left[f'_y\left(\vartheta,t,m_t\left(\vartheta\right)\right)\dot m_t\left(\vartheta\right)-f'_y\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)\dot m_t^\circ\left(\vartheta\right)\right]\mathrm{d}t+\dot B\left(\vartheta,t\right)\mathrm{d}W_t\\&=Q\left(\vartheta,t\right)\mathrm{d}t+\dot B_\vartheta\left(\vartheta,t\right)\mathrm{d}W_t-\frac{b\left(\vartheta,t,m_t\left(\vartheta\right)\right)}{\varepsilon\sigma\left(t\right)}f''_{yy}\left(\vartheta,t,\bar m_t\right)\dot m_t\left(\vartheta\right)\left[m_t\left(\vartheta\right)-m_t^\circ\left(\vartheta\right)\right]\mathrm{d}t\\&\quad-\frac{\left[\dot B_\vartheta\left(\vartheta,t\right)f'_y\left(\vartheta,t,\tilde m_t\left(\vartheta\right)\right)+b\left(\vartheta,t,m_t\left(\vartheta\right)\right)\dot f'_{\vartheta y}\left(\vartheta,t,\tilde m_t\left(\vartheta\right)\right)\right]}{\varepsilon\sigma\left(t\right)}\left[m_t\left(\vartheta\right)-m_t^\circ\left(\vartheta\right)\right]\mathrm{d}t\\&\quad+\frac{f'_y\left(\vartheta_0,t,Y_t\right)b\left(\vartheta_0,t,Y_t\right)}{f'_y\left(\vartheta,t,m_t\left(\vartheta\right)\right)^2}\left[\dot f'_{y\vartheta}\left(\vartheta,t,m_t\left(\vartheta\right)\right)+f''_{yy}\left(\vartheta,t,m_t\left(\vartheta\right)\right)\dot m_t\left(\vartheta\right)\right]\mathrm{d}V_t\\&\quad-\frac{b\left(\vartheta,t,m_t\left(\vartheta\right)\right)}{\varepsilon\sigma\left(t\right)}f'_y\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)\left[\dot m_t\left(\vartheta\right)-\dot m_t^\circ\left(\vartheta\right)\right]\mathrm{d}t\left(1+O\left(\sqrt{\varepsilon}\right)\right)\end{aligned}$$

with obvious notation for $Q\left(\vartheta,t\right)$. Here we used the relations

$$\begin{aligned}&f\left(\vartheta_0,t,Y_t\right)-f\left(\vartheta,t,m_t\left(\vartheta\right)\right)=f\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)-f\left(\vartheta,t,m_t\left(\vartheta\right)\right)\\&\quad=f'_y\left(\vartheta,t,\tilde m_t\right)\left(m_t^\circ\left(\vartheta\right)-m_t\left(\vartheta\right)\right)=f'_y\left(\vartheta,t,m_t^\circ\right)\left(m_t^\circ\left(\vartheta\right)-m_t\left(\vartheta\right)\right)\left(1+O\left(\sqrt{\varepsilon}\right)\right).\end{aligned}$$

Introduce the function

$$R\left(\vartheta,t\right)=-f''_{yy}\left(\vartheta,t,\bar m_t\right)\dot m_t\left(\vartheta\right)-\frac{\dot B_\vartheta\left(\vartheta,t\right)}{b\left(\vartheta,t,m_t\left(\vartheta\right)\right)}f'_y\left(\vartheta,t,\tilde m_t\left(\vartheta\right)\right)-\dot f'_{\vartheta y}\left(\vartheta,t,\tilde m_t\left(\vartheta\right)\right).$$

Then we can write

$$\begin{aligned}\mathrm{d}\left[\dot m_t\left(\vartheta\right)-\dot m_t^\circ\left(\vartheta\right)\right]&=Q\left(\vartheta,t\right)\mathrm{d}t+\frac{b\left(\vartheta,t,m_t\left(\vartheta\right)\right)R\left(\vartheta,t\right)}{\varepsilon\sigma\left(t\right)}\left[m_t\left(\vartheta\right)-m_t^\circ\left(\vartheta\right)\right]\mathrm{d}t\\&\quad+\frac{f'_y\left(\vartheta_0,t,Y_t\right)b\left(\vartheta_0,t,Y_t\right)}{f'_y\left(\vartheta,t,m_t\left(\vartheta\right)\right)^2}\left[\dot f'_{y\vartheta}\left(\vartheta,t,m_t\left(\vartheta\right)\right)+f''_{yy}\left(\vartheta,t,m_t\left(\vartheta\right)\right)\dot m_t\left(\vartheta\right)\right]\mathrm{d}V_t\\&\quad-\frac{b\left(\vartheta,t,m_t\left(\vartheta\right)\right)}{\varepsilon\sigma\left(t\right)}f'_y\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)\left[\dot m_t\left(\vartheta\right)-\dot m_t^\circ\left(\vartheta\right)\right]\mathrm{d}t\\&\quad+\dot B_\vartheta\left(\vartheta,t\right)\mathrm{d}W_t\left(1+O\left(\sqrt{\varepsilon}\right)\right).\end{aligned}$$

Introduce the notations

$$\dot{C}_t = \int_{t_0}^{t} \dot{B}_\vartheta\left(\vartheta, s\right) \mathrm{d}W_s,$$

$$\dot{F}_t = \int_{t_0}^{t} \frac{f_y'\left(\vartheta_0, s, Y_s\right) b\left(\vartheta_0, s, Y_s\right)}{f_y'\left(\vartheta, s, m_s\left(\vartheta\right)\right)^2} \left[\dot{f}_{y\vartheta}'\left(\vartheta, s, m_s\left(\vartheta\right)\right) + f_{yy}''\left(\vartheta, s, m_s\left(\vartheta\right)\right) \dot{m}_s\left(\vartheta\right)\right] \mathrm{d}V_s.$$

Then for the process $\dot{Z}_t = \dot{m}_t\left(\vartheta\right) - \dot{m}_t^\circ\left(\vartheta\right) - \dot{C}_t - \dot{F}_t$ we obtain the equation

$$\mathrm{d}\dot{Z}_t = \frac{b\left(\vartheta, t, m_t\left(\vartheta\right)\right) R\left(\vartheta, t\right)}{\varepsilon\sigma\left(t\right)} \left[m_t\left(\vartheta\right) - m_t^\circ\left(\vartheta\right)\right] \mathrm{d}t - \frac{b\left(\vartheta, t, m_t\left(\vartheta\right)\right)}{\varepsilon\sigma\left(t\right)} f_y'\left(\vartheta, t, m_t^\circ\left(\vartheta\right)\right) \dot{Z}_t \mathrm{d}t$$
$$+ Q\left(\vartheta, t\right) \mathrm{d}t - \frac{b\left(\vartheta, t, m_t\left(\vartheta\right)\right)}{\varepsilon\sigma\left(t\right)} f_y'\left(\vartheta, t, m_t^\circ\left(\vartheta\right)\right) \left[\dot{C}_t + \dot{F}_t\right] \mathrm{d}t.$$

Therefore $(t > t_0)$

$$\dot{Z}_t = \int_{t_0}^{t} \Psi\left(t, s\right) Q\left(\vartheta, s\right) \mathrm{d}s + \frac{1}{\varepsilon} \int_{t_0}^{t} \Psi\left(t, s\right) \frac{b\left(\vartheta, s, m_s\left(\vartheta\right)\right) R\left(\vartheta, s\right)}{\sigma\left(s\right)} \left[m_s\left(\vartheta\right) - m_s^\circ\left(\vartheta\right)\right] \mathrm{d}s$$
$$+ \dot{Z}_{t_0} \Psi\left(t, 0\right) - \frac{1}{\varepsilon} \int_{t_0}^{t} \Psi\left(t, s\right) \frac{\dot{B}_\vartheta\left(\vartheta, s\right)}{\sigma\left(s\right)} f_y'\left(\vartheta, s, m_s^\circ\left(\vartheta\right)\right) \left[\dot{C}_s + \dot{F}_s\right] \mathrm{d}s$$
$$= \frac{R\left(\vartheta, t\right)}{f_y'\left(\vartheta, t, m_t^\circ\left(\vartheta\right)\right)} \left[m_t\left(\vartheta\right) - m_t^\circ\left(\vartheta\right)\right] - \dot{C}_t - \dot{F}_t$$
$$+ \frac{1}{\varepsilon} \int_{t_\varepsilon}^{t} \Psi\left(t, s\right) \frac{\dot{B}_\vartheta\left(\vartheta, s\right)}{\sigma\left(s\right)} f_y'\left(\vartheta, s, m_s^\circ\left(\vartheta\right)\right) \left[\dot{C}_t - \dot{C}_s\right] \mathrm{d}s$$
$$+ \frac{1}{\varepsilon} \int_{t_\varepsilon}^{t} \Psi\left(t, s\right) \frac{\dot{B}_\vartheta\left(\vartheta, s\right)}{\sigma\left(s\right)} f_y'\left(\vartheta, s, m_s^\circ\left(\vartheta\right)\right) \left[\dot{F}_t - \dot{F}_s\right] \mathrm{d}s + O\left(\varepsilon\right).$$

Here we used the slightly different notation

$$\Psi\left(t, s\right) = \exp\left(-\frac{1}{\varepsilon} \int_{s}^{t} q\left(r\right) \mathrm{d}r\right), \qquad q\left(t\right) = \frac{b\left(\vartheta, t, m_t^\circ\left(\vartheta\right)\right)}{\sigma\left(t\right)} f_y'(\vartheta, t, m_t^\circ\left(\vartheta\right))$$

and the estimates

$$\dot{Z}_{t_0} \Psi\left(t, 0\right) = O\left(e^{-\frac{c}{\varepsilon}}\right), \qquad \int_{t_\varepsilon}^{t} \Psi\left(t, s\right) Q\left(\vartheta, s\right) \mathrm{d}s = O\left(\varepsilon\right).$$

Further

$$\frac{1}{\varepsilon} \int_{t_0}^{t} \Psi\left(t, s\right) \frac{\dot{B}_\vartheta\left(\vartheta, s\right)}{\sigma\left(s\right)} f_y'\left(\vartheta, s, m_s^\circ\left(\vartheta\right)\right) \int_{s}^{t} \dot{B}\left(\vartheta, r\right) \mathrm{d}W_r \;\mathrm{d}s$$
$$= \frac{\dot{B}_\vartheta\left(\vartheta, t\right)^2}{\varepsilon\sigma\left(t\right)} f_y'\left(\vartheta, t, m_t^\circ\left(\vartheta\right)\right) \int_{t_\varepsilon}^{t} \Psi\left(t, s\right) \left[W_t - W_s\right] \mathrm{d}s \left(1 + O\left(\varepsilon^{\frac{\nu}{2}}\right)\right)$$
$$= \sqrt{\varepsilon}\, \frac{\dot{B}_\vartheta\left(\vartheta, t\right)^2}{\sigma\left(t\right)} f_y'\left(\vartheta, t, m_t^\circ\left(\vartheta\right)\right) q\left(t\right)^{-3/2} \xi_{t,\varepsilon} \left(1 + O\left(\varepsilon^{\frac{\nu}{2}}\right)\right)$$
$$= \sqrt{\varepsilon}\, \frac{\dot{B}_\vartheta\left(\vartheta, t\right)^2}{b\left(\vartheta, t, m_t^\circ\left(\vartheta\right)\right)^2} \sqrt{\frac{b\left(\vartheta, t, m_t^\circ\left(\vartheta\right)\right) \sigma\left(t\right)}{f_y'\left(\vartheta, t, m_t^\circ\left(\vartheta\right)\right)}}\; \xi_{t,\varepsilon} \left(1 + O\left(\varepsilon^{\frac{\nu}{2}}\right)\right),$$

and

$$\begin{aligned}
&\frac{1}{\varepsilon}\int_{t_0}^{t}\Psi\left(t,s\right)\frac{\dot B_\vartheta\left(\vartheta,s\right)}{\sigma\left(s\right)}f_y'\left(\vartheta,s,m_s^\circ\left(\vartheta\right)\right)\left[\dot F_t-\dot F_s\right]\mathrm{d}s\\
&=\sqrt{\varepsilon}\frac{\dot B_\vartheta\left(\vartheta,t\right)f_y'\left(\vartheta_0,t,Y_t\right)b\left(\vartheta_0,t,Y_t\right)}{\sigma\left(t\right)f_y'\left(\vartheta,t,m_t\left(\vartheta\right)\right)q\left(t\right)^{3/2}}\Bigl[\dot f_{y\vartheta}'\left(\vartheta,t,m_t\left(\vartheta\right)\right)\\
&\quad+f_{yy}''\left(\vartheta,t,m_t\left(\vartheta\right)\right)\dot m_t\left(\vartheta\right)\Bigr]\,\eta_{t,\varepsilon}\left(1+O\left(\varepsilon^{\frac{\nu}{2}}\right)\right)\\
&=\sqrt{\varepsilon}\frac{\dot B_\vartheta\left(\vartheta,t\right)f_y'\left(\vartheta_0,t,Y_t\right)b\left(\vartheta_0,t,Y_t\right)\sqrt{\sigma\left(t\right)}}{b\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)^{3/2}f_y'\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)^{7/2}}\Bigl[\dot f_{y\vartheta}'\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)f_y'\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)\\
&\quad-f_{yy}''\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)\dot f_\vartheta\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)\Bigr]\,\eta_{t,\varepsilon}\left(1+O\left(\varepsilon^{\frac{\nu}{2}}\right)\right).
\end{aligned}$$

These relations allow us to write

$$\begin{aligned}
\dot m_t\left(\vartheta\right)-\dot m_t^\circ\left(\vartheta\right)&=\frac{R\left(\vartheta,t\right)\Gamma^\circ\left(\vartheta,t\right)}{f_y'\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)}\left[\xi_{t,\varepsilon}-\frac{b\left(\vartheta_0,t,Y_t\right)f_y'\left(\vartheta_0,t,Y_t\right)}{b\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)f_y'\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)}\,\eta_{t,\varepsilon}\right]\sqrt{\varepsilon}\\
&\quad+\frac{\dot B_\vartheta\left(\vartheta,t\right)^2}{b\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)^2}\sqrt{\frac{b\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)\sigma\left(t\right)}{f_y'\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)}}\,\xi_{t,\varepsilon}\,\sqrt{\varepsilon}\\
&\quad+\frac{\dot B_\vartheta\left(\vartheta,t\right)f_y'\left(\vartheta_0,t,Y_t\right)b\left(\vartheta_0,t,Y_t\right)\sqrt{\sigma\left(t\right)}}{b\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)^{3/2}f_y'\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)^{7/2}}\Bigl[\dot f_{y\vartheta}'\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)f_y'\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)\\
&\quad-f_{yy}''\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)\dot f_\vartheta\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)\Bigr]\,\eta_{t,\varepsilon}\,\sqrt{\varepsilon}\,+O\left(\varepsilon^{\frac{1+\nu}{2}}\right).
\end{aligned}$$

Hence

$$\begin{aligned}
\frac{\dot m_t\left(\vartheta\right)-\dot m_t^\circ\left(\vartheta\right)}{\Gamma^\circ\left(\vartheta,t\right)\,\sqrt{\varepsilon}}&=\frac{R^\circ\left(\vartheta,t\right)}{f_y'\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)}\left[\xi_{t,\varepsilon}-\frac{b\left(\vartheta_0,t,Y_t\right)f_y'\left(\vartheta_0,t,Y_t\right)}{b\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)f_y'\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)}\,\eta_{t,\varepsilon}\right]\\
&\quad+\frac{\dot b_\vartheta\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)f_y'\left(\vartheta_0,t,Y_t\right)b\left(\vartheta_0,t,m_t^\circ\left(\vartheta\right)\right)}{b\left(\vartheta,t\right)^2f_y'\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)}\Biggl[\frac{\dot f_{y\vartheta}'\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)}{f_y'\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)}\\
&\quad-\frac{f_{yy}''\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)\dot f_\vartheta\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)}{f_y'\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)^2}\Biggr]\,\eta_{t,\varepsilon}+\frac{\dot B_\vartheta\left(\vartheta,t\right)^2}{b\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)^2}\,\xi_{t,\varepsilon}+O\left(\varepsilon^{\frac{\nu}{2}}\right)\\
&=\left[\frac{R^\circ\left(\vartheta,t\right)}{f_y'\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)}+\frac{\dot B_\vartheta\left(\vartheta,t\right)^2}{b\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)^2}\right]\,\xi_{t,\varepsilon}\\
&\quad+\frac{b\left(\vartheta_0,t,Y_t\right)f_y'\left(\vartheta_0,t,Y_t\right)}{b\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)f_y'\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)}\Biggl[\frac{\dot B_\vartheta\left(\vartheta,t\right)\dot f_{y\vartheta}'\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)}{b\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)f_y'\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)}\\
&\quad-\frac{R^\circ\left(\vartheta,t\right)}{f_y'\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)}-\frac{\dot B_\vartheta\left(\vartheta,t\right)f_{yy}''\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)\dot f_{y\vartheta}'\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)}{b\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)f_y'\left(\vartheta,t,m_t^\circ\left(\vartheta\right)\right)^2}\Biggr]\,\eta_{t,\varepsilon}+O\left(\varepsilon^{\frac{\nu}{2}}\right)\\
&=\Pi^\circ\left(\vartheta,t\right)\,\xi_{t,\varepsilon}+\Lambda^\circ\left(\vartheta,t\right)\,\eta_{t,\varepsilon}+O\left(\varepsilon^{\frac{\nu}{2}}\right).
\end{aligned}$$

Therefore if $\vartheta=\vartheta_0$, then this expression is simplified and we have

$$
\begin{aligned}
\frac{\dot m_t(\vartheta_0)-\dot m_t^\circ(\vartheta_0)}{\Gamma^\circ(\vartheta_0,t)\ \sqrt{\varepsilon}} \Longrightarrow{}& \frac{R^\circ(\vartheta_0,t)}{f'_y(\vartheta_0,t,Y_t)}\left[\xi_t-\eta_t\right]+\frac{\dot B^\circ_\vartheta(\vartheta_0,t)^2}{b(\vartheta_0,t,Y_t)^2}\xi_t\\
&+\frac{\dot B^\circ_\vartheta(\vartheta_0,t)}{b(\vartheta_0,t,Y_t)}\left[\frac{\dot f'_{y\vartheta}(\vartheta_0,t,Y_t)}{f'_y(\vartheta_0,t,Y_t)}-\frac{f''_{yy}(\vartheta_0,t,Y_t)\,\dot f_\vartheta(\vartheta_0,t,Y_t)}{f'_y(\vartheta_0,t,Y_t)^2}\right]\eta_t\\
={}&\left[\frac{R^\circ(\vartheta_0,t)}{f'_y(\vartheta_0,t,Y_t)}+\frac{\dot B^\circ_\vartheta(\vartheta_0,t)^2}{b(\vartheta_0,t,Y_t)^2}\right]\ \xi_t+\left[\frac{\dot B^\circ_\vartheta(\vartheta_0,t)\,\dot f'_{y\vartheta}(\vartheta_0,t,Y_t)}{b(\vartheta_0,t,Y_t)\,f'_y(\vartheta_0,t,Y_t)}\right.\\
&\left.-\frac{\dot B^\circ_\vartheta(\vartheta_0,t)\,f''_{yy}(\vartheta_0,t,Y_t)\,\dot f_\vartheta(\vartheta_0,t,Y_t)}{b(\vartheta_0,t,Y_t)\,f'_y(\vartheta_0,t,Y_t)^2}-\frac{R^\circ(\vartheta_0,t)}{f'_y(\vartheta_0,t,Y_t)}\right]\eta_t\\
={}&\Pi^\circ(\vartheta_0,t)\,\xi_t+\Lambda^\circ(\vartheta_0,t)\,\eta_t.
\end{aligned}
$$

□

Corollary 1 *The representations (3.7), (3.8) allow us to write*

$$
\begin{aligned}
\frac{1}{\Gamma^\circ(\vartheta,t)\sqrt{\varepsilon}}\frac{\mathrm{d}f(\vartheta,t,m_t(\vartheta))}{\mathrm{d}\vartheta}={}&K(\vartheta,t,m_t^\circ(\vartheta))\left[\xi_{t,\varepsilon}-\Delta^\circ(\vartheta,t)\ \eta_{t,\varepsilon}\right]\\
&+f'_y(\vartheta,t,m_t(\vartheta))\left[\Pi^\circ(\vartheta,t)\ \xi_{t,\varepsilon}+\Lambda^\circ(\vartheta,t)\ \eta_{t,\varepsilon}\right]+O\left(\varepsilon^{\frac{\nu}{2}}\right)\\
={}&\left[K(\vartheta,t,m_t^\circ(\vartheta))+f'_y(\vartheta,t,m_t(\vartheta))\,\Pi^\circ(\vartheta,t)\right]\xi_{t,\varepsilon}+O\left(\varepsilon^{\frac{\nu}{2}}\right)\\
&+\left[f'_y(\vartheta,t,m_t(\vartheta))\,\Lambda^\circ(\vartheta,t)-K(\vartheta,t,m_t^\circ(\vartheta))\,\Delta^\circ(\vartheta,t)\right]\ \eta_{t,\varepsilon}.
\end{aligned}
\tag{3.11}
$$

Here

$$
K(\vartheta,t,y)=\dot f'_{\vartheta y}(\vartheta,t,y)-\frac{f''_{yy}(\vartheta,t,y)}{f'_y(\vartheta,t,y)}.
$$

Indeed we have

$$
\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}\vartheta}f(\vartheta,t,m_t(\vartheta))\\
&=\frac{\mathrm{d}}{\mathrm{d}\vartheta}\left[f(\vartheta,t,m_t(\vartheta))-f(\vartheta_0,t,m_t^\circ(\vartheta_0))\right]\\
&=\frac{\mathrm{d}}{\mathrm{d}\vartheta}\left[f(\vartheta,t,m_t(\vartheta))-f(\vartheta,t,m_t^\circ(\vartheta))\right]\\
&=\dot f_\vartheta(\vartheta,t,m_t(\vartheta))-\dot f_\vartheta(\vartheta,t,m_t^\circ(\vartheta))\\
&\quad+f'_y(\vartheta,t,m_t(\vartheta))\,\dot m_t(\vartheta)-f'_y(\vartheta,t,m_t^\circ(\vartheta))\,\dot m_t^\circ(\vartheta)\\
&=\dot f'_{\vartheta y}(\vartheta,t,\tilde m_t)\left[m_t(\vartheta)-m_t^\circ(\vartheta)\right]+f'_y(\vartheta,t,m_t(\vartheta))\left[\dot m_t(\vartheta)-\dot m_t^\circ(\vartheta)\right]\\
&\quad+\left[f'_y(\vartheta,t,m_t(\vartheta))-f'_y(\vartheta,t,m_t^\circ(\vartheta))\right]\dot m_t(\vartheta)\\
&=\left[\dot f'_{\vartheta y}(\vartheta,t,\tilde m_t)+f''_{yy}(\vartheta,t,\tilde m_t)\,\dot m_t(\vartheta)\right]\left[m_t(\vartheta)-m_t^\circ(\vartheta)\right]\\
&\quad+f'_y(\vartheta,t,m_t(\vartheta))\left[\dot m_t(\vartheta)-\dot m_t^\circ(\vartheta)\right]\\
&=\left[\dot f'_{\vartheta y}(\vartheta,t,m_t^\circ(\vartheta))+f''_{yy}(\vartheta,t,m_t^\circ(\vartheta))\,\dot m_t(\vartheta)\right]\left[m_t^\circ(\vartheta)-m_t^\circ(\vartheta)\right]\\
&\quad+f'_y(\vartheta,t,m_t^\circ(\vartheta))\left[\dot m_t(\vartheta)-\dot m_t^\circ(\vartheta)\right]+O\left(\sqrt{\varepsilon}\right)\\
&=\left[\dot f'_{\vartheta y}(\vartheta,t,m_t^\circ(\vartheta))-\frac{f''_{yy}(\vartheta,t,m_t^\circ(\vartheta))}{f'_y(\vartheta,t,m_t^\circ(\vartheta))}\right]\left[m_t^\circ(\vartheta)-m_t^\circ(\vartheta)\right]\\
&\quad+f'_y(\vartheta,t,m_t^\circ(\vartheta))\left[\dot m_t(\vartheta)-\dot m_t^\circ(\vartheta)\right]+O\left(\sqrt{\varepsilon}\right)\\
&=K(\vartheta,t,m_t^\circ(\vartheta))\left[m_t^\circ(\vartheta)-m_t^\circ(\vartheta)\right]+f'_y(\vartheta,t,m_t^\circ(\vartheta))\left[\dot m_t(\vartheta)-\dot m_t^\circ(\vartheta)\right]+O\left(\sqrt{\varepsilon}\right).
\end{aligned}
$$

The substitution of (3.7), (3.8) here gives (3.11).

4. Empirical quadratic variation of the trend coefficient

Introduce the same statistic as in the linear case [16] ($\delta_\varepsilon=\varepsilon,\ t_{i+1}=t_i+\varphi_\varepsilon,\ \varphi_\varepsilon=\varepsilon^{2/3}$, $N_{\tau,\varepsilon}=\left[\frac{\tau}{\varphi_\varepsilon}\right]$)

$$\Psi_{\tau,\varepsilon}=\sum_{i=0}^{N_{\tau,\varepsilon}-1}\left(\frac{X_{t_{i+1}+\delta_\varepsilon}-X_{t_{i+1}}}{\delta_\varepsilon}-\frac{X_{t_i+\delta_\varepsilon}-X_{t_i}}{\delta_\varepsilon}\right)^2,\quad 0<\tau\leqslant T.$$

Remind that

$$\frac{X_{t_{i+1}+\delta_\varepsilon}-X_{t_{i+1}}}{\delta_\varepsilon}\quad\text{and}\quad\frac{X_{t_i+\delta_\varepsilon}-X_{t_i}}{\delta_\varepsilon}$$

are *asymptotic derivatives* ($\delta_\varepsilon\rightarrow 0$) of X_t at the points t_{i+1} and t_i respectively and the sum $\Psi_{\tau,\varepsilon}$ is the quadratic variation of these "derivatives".

Note that $\Psi_{\tau,\varepsilon}$ converges to the limit of the quadratic variation of the trend coefficient $f\left(\vartheta_0,t,Y_t\right)$. To obtain the expression for this limit we first write the formal expansions and later in the proof of Lemma 1 we give the detailed proof. The quadratic variation of $f\left(\vartheta_0,t,Y_t\right)$ is the limit of the following sum

$$\begin{aligned}&\sum_{i=0}^{N_{\tau,\varepsilon}-1}\left[f\left(\vartheta_0,t_{i+1},Y_{t_{i+1}}\right)-f\left(\vartheta_0,t_i,Y_{t_i}\right)\right]^2\\&=\sum_{i=0}^{N_{\tau,\varepsilon}-1}\left[f\left(\vartheta_0,t_i,Y_{t_{i+1}}\right)-f\left(\vartheta_0,t_i,Y_{t_i}\right)\right]^2+o\left(1\right)\\&=\sum_{i=0}^{N_{\tau,\varepsilon}-1}f'_y(\vartheta_0,t_i,\tilde Y_{t_i})^2\left(Y_{t_{i+1}}-Y_{t_i}\right)^2+o\left(1\right)\\&=\sum_{i=0}^{N_{\tau,\varepsilon}-1}f'_y(\vartheta_0,t_i,Y_{t_i})^2\,b\left(\vartheta_0,t_i,Y_{t_i}\right)^2\left(V_{t_{i+1}}-V_{t_i}\right)^2+o\left(1\right)\\&\longrightarrow\int_0^\tau f'_y(\vartheta_0,t,Y_t)^2\,b\left(\vartheta_0,t,Y_t\right)^2{\rm d}t=\Psi_\tau\left(\vartheta_0\right).\end{aligned}$$

Here the function $\Psi_\tau\left(\vartheta\right),\vartheta\in\Theta$ is defined by the last equality.

Let us denote

$$\pi_1=2\int_0^\tau\sigma\left(t\right)^2{\rm d}t,\quad \pi_2\left(\vartheta_0\right)=\frac{2}{3}\int_0^\tau f'_y\left(\vartheta_0,t,Y_t\right)^2b\left(\vartheta_0,t,Y_t\right)^2{\rm d}t,$$

$$\pi_3\left(\vartheta_0\right)=\int_0^\tau f'_y\left(\vartheta_0,t,Y_t\right)^2b\left(\vartheta_0,t,Y_t\right)^2{\rm d}v_t,\quad \pi\left(\vartheta_0\right)=\pi_1+\pi_2\left(\vartheta_0\right)+\pi_3\left(\vartheta_0\right),$$

where $v_t,0\leqslant t\leqslant\tau$ is some Wiener process.

Details of this convergence are given in the following proposition.

Proposition 2 *Let the condition $\mathscr{C}_1$, be satisfied, then*

$$\Psi_{\tau,\varepsilon}=\Psi_\tau\left(\vartheta_0\right)+\pi_\varepsilon\,\varepsilon^{1/3}+O\left(\varepsilon^{2/3}\right),\quad \varepsilon^{-1/3}\left(\Psi_{\tau,\varepsilon}-\Psi_\tau\left(\vartheta_0\right)\right)\Longrightarrow\pi\left(\vartheta_0\right)$$

and for any $p>0$ there exists a constant $C=C_p>0$ such that

$$\sup_{\vartheta_0\in\Theta}\mathbf{E}_{\vartheta_0}\left|\Psi_{\tau,\varepsilon}-\Psi_\tau\left(\vartheta_0\right)\right|^p\leqslant C\,\varepsilon^{p/3}.$$

The random variable π_ε is defined below in (4.3)–(4.6) and (4.7).

Proof The proof follows the main steps of the proof of the Proposition 1 in [16]. Let us consider first the simplest case just to see how this rate of convergence can be obtained. Put $f\left(\vartheta_0,t,y\right)\equiv y,\ a\left(\vartheta_0,t,y\right)\equiv 0,\ \sigma\left(t\right)\equiv 1,\ b\left(\vartheta_0,t,y\right)\equiv 1,\ \tau=1$. Then $\Psi_\tau=1$ and the observed process is

$$\mathrm{d}X_t=V_t\mathrm{d}t+\varepsilon\mathrm{d}W_t,\quad X_0=0,\quad 0\leqslant t\leqslant 1.$$

Denote

$$\xi_{t_{i+1,\varepsilon}}=\frac{1}{\sqrt{\delta_\varepsilon}}\left[W_{t_{i+1}+\delta_\varepsilon}-W_{t_{i+1}}\right]\sim\mathcal{N}\left(0,1\right),\quad \Delta\xi_{t_{i,\varepsilon}}=\xi_{t_{i+1,\varepsilon}}-\xi_{t_{i,\varepsilon}},$$

$$\eta_{t_{i+1,\varepsilon}}=\frac{1}{\delta_\varepsilon^{3/2}}\int_{t_{i+1}}^{t_{i+1}+\delta_\varepsilon}\left[V_s-V_{t_{i+1}}\right]\,\mathrm{d}s\sim\mathcal{N}\left(0,1/3\right),\quad \Delta\eta_{t_{i,\varepsilon}}=\eta_{t_{i+1,\varepsilon}}-\eta_{t_{i,\varepsilon}}.$$

Then we can write

$$\begin{aligned}\bar\Psi_{\tau,\varepsilon}&=\sum_{i=0}^{N_{\tau,\varepsilon}}\left(\frac{1}{\delta_\varepsilon}\int_{t_{i+1}}^{t_{i+1}+\delta_\varepsilon}V_s\mathrm{d}s+\frac{\varepsilon}{\sqrt{\delta_\varepsilon}}\xi_{t_{i+1,\varepsilon}}-\frac{1}{\delta_\varepsilon}\int_{t_i}^{t_i+\delta_\varepsilon}V_s\mathrm{d}s-\frac{\varepsilon}{\sqrt{\delta_\varepsilon}}\xi_{t_{i,\varepsilon}}\right)^2\\
&=\sum_{i=0}^{N_{\tau,\varepsilon}}\left(V_{t_{i+1}}-V_{t_i}+\sqrt{\delta_\varepsilon}\Delta\eta_{t_{i,\varepsilon}}+\frac{\varepsilon}{\sqrt{\delta_\varepsilon}}\Delta\xi_{t_{i,\varepsilon}}\right)^2\\
&=\sum_{i=0}^{N_{\tau,\varepsilon}}\left[V_{t_{i+1}}-V_{t_i}\right]^2+\delta_\varepsilon\sum_{i=0}^{N_{\tau,\varepsilon}}\Delta\eta_{t_{i,\varepsilon}}^2+\frac{\varepsilon^2}{\delta_\varepsilon}\sum_{i=0}^{N_{\tau,\varepsilon}}\Delta\xi_{t_{i,\varepsilon}}^2+2\sqrt{\delta_\varepsilon}\sum_{i=0}^{N_{\tau,\varepsilon}}\left[V_{t_{i+1}}-V_{t_i}\right]\Delta\eta_{t_{i,\varepsilon}}\\
&\quad+\frac{2\varepsilon}{\sqrt{\delta_\varepsilon}}\sum_{i=0}^{N_{\tau,\varepsilon}}\left[V_{t_{i+1}}-V_{t_i}\right]\Delta\xi_{t_{i,\varepsilon}}+2\varepsilon\sum_{i=0}^{N_{\tau,\varepsilon}}\Delta\eta_{t_{i,\varepsilon}}\Delta\xi_{t_{i,\varepsilon}}.\end{aligned}\tag{4.1}$$

By Itô formula

$$\sum_{i=0}^{N_{\tau,\varepsilon}}\left[V_{t_{i+1}}-V_{t_i}\right]^2=\sum_{i=0}^{N_{\tau,\varepsilon}}\left(t_{i+1}-t_i\right)+2\sum_{i=0}^{N_{\tau,\varepsilon}}\int_{t_i}^{t_{i+1}}\left(V_s-V_{t_i}\right)\mathrm{d}V_s=1+\sqrt{\varphi_\varepsilon}\,O\left(1\right)\tag{4.2}$$

because

$$\mathbf{E}\left(\sum_{i=0}^{N_{\tau,\varepsilon}}\int_{t_i}^{t_{i+1}}\left(V_s-V_{t_i}\right)\mathrm{d}V_s\right)^2=\sum_{i=0}^{N_{\tau,\varepsilon}}\int_{t_i}^{t_{i+1}}\left(s-t_i\right)\mathrm{d}s=\frac{\varphi_\varepsilon}{2}.$$

For the second term we have

$$\varphi_\varepsilon\sum_{i=0}^{N_{\tau,\varepsilon}}\Delta\eta_{t_{i,\varepsilon}}^2\longrightarrow\frac{2}{3},\quad \delta_\varepsilon\sum_{i=0}^{N_{\tau,\varepsilon}}\Delta\eta_{t_{i,\varepsilon}}^2=\frac{\delta_\varepsilon}{\varphi_\varepsilon}\,\varphi_\varepsilon\sum_{i=0}^{N_{\tau,\varepsilon}}\Delta\eta_{t_{i,\varepsilon}}^2=\frac{\delta_\varepsilon}{\varphi_\varepsilon}\,\left(\frac{2}{3}+o\left(1\right)\right).$$

We have as well

$$\varphi_\varepsilon\sum_{i=0}^{N_{\tau,\varepsilon}}\Delta\xi_{t_{i,\varepsilon}}^2\longrightarrow 2,\quad \frac{\varepsilon^2}{\delta_\varepsilon\varphi_\varepsilon}\,\varphi_\varepsilon\sum_{i=0}^{N_{\tau,\varepsilon}}\Delta\xi_{t_{i,\varepsilon}}^2=\frac{\varepsilon^2}{\delta_\varepsilon\varphi_\varepsilon}\,\left(2+o\left(1\right)\right).$$

For the other terms in (4.1) we obtain the relations

$$4\varepsilon^2\mathbf{E}\left(\sum_{i=0}^{N_{\tau,\varepsilon}}\Delta\xi_{t_{i,\varepsilon}}\Delta\eta_{t_{i,\varepsilon}}\right)^2=\frac{16\varepsilon^2}{3\varphi_\varepsilon},\quad \varepsilon\sum_{i=0}^{N_{\tau,\varepsilon}}\Delta\xi_{t_{i,\varepsilon}}\Delta\eta_{t_{i,\varepsilon}}=\frac{\varepsilon}{\sqrt{\varphi_\varepsilon}}\,O\left(1\right),$$

$$\frac{4\varepsilon^2}{\delta_\varepsilon}\mathbf{E}\left(\sum_{i=0}^{N_{\tau,\varepsilon}}\left[V_{t_{i+1}}-V_{t_i}\right]\Delta\xi_{t_{i,\varepsilon}}\right)^2=\frac{8\varepsilon^2}{\delta_\varepsilon},\quad \frac{\varepsilon}{\sqrt{\delta_\varepsilon}}\sum_{i=0}^{N_{\tau,\varepsilon}}\left[V_{t_{i+1}}-V_{t_i}\right]\Delta\xi_{t_{i,\varepsilon}}=\frac{\varepsilon}{\sqrt{\delta_\varepsilon}}\,O\left(1\right).$$

Remark that

$$\mathbf{E}\left[V_{t_{i+1}}-V_{t_i}\right]\Delta\eta_{t_{i,\varepsilon}}=\mathbf{E}\left[V_{t_i}-V_{t_{i+1}}\right]\eta_{t_{i,\varepsilon}}=\frac{1}{\delta_\varepsilon^{3/2}}\mathbf{E}\left[V_{t_i}-V_{t_{i+1}}\right]\int_{t_i}^{t_i+\delta_\varepsilon}\left[V_s-V_{t_i}\right]\mathrm{d}s$$
$$=-\frac{1}{\delta_\varepsilon^{3/2}}\int_{t_i}^{t_i+\delta_\varepsilon}\mathbf{E}\left[V_s-V_{t_i}\right]^2\mathrm{d}s=-\frac{\sqrt{\delta_\varepsilon}}{2}$$

and therefore

$$\sqrt{\delta_\varepsilon}\sum_{i=0}^{N_{\tau,\varepsilon}}\left[V_{t_{i+1}}-V_{t_i}\right]\Delta\eta_{t_{i,\varepsilon}}=-\frac{\delta_\varepsilon}{\varphi_\varepsilon}\left(\frac{1}{2}+o\left(1\right)\right).$$

Therefore the main terms are

$$\Psi_{\tau,\varepsilon}=1+\sqrt{\varphi_\varepsilon}\,O\left(1\right)+\frac{\delta_\varepsilon}{\varphi_\varepsilon}\,O\left(1\right)+\frac{\varepsilon^2}{\delta_\varepsilon\varphi_\varepsilon}\,O\left(1\right).$$

If we write two equations of balance for $\delta_\varepsilon=\varepsilon^q$ and $\varphi_\varepsilon=\varepsilon^l$

$$\sqrt{\varphi_\varepsilon}=\frac{\delta_\varepsilon}{\varphi_\varepsilon},\quad \frac{\delta_\varepsilon}{\varphi_\varepsilon}=\frac{\varepsilon^2}{\delta_\varepsilon\varphi_\varepsilon},$$

then we obtain the values $q=1$ and $l=2/3$, i.e., $\delta_\varepsilon=\varepsilon$ and $\varphi_\varepsilon=\varepsilon^{2/3}$.

Now we repeat the given above calculations but for the model (1.8), (1.9) with description of the corresponding errors. We omit for instant ϑ_0.

Using elementary expansions we write

$$\begin{aligned}\frac{X_{t_i+\delta_\varepsilon}-X_{t_i}}{\delta_\varepsilon}&=\frac{1}{\delta_\varepsilon}\int_{t_i}^{t_i+\delta_\varepsilon}f\left(t,Y_t\right)\,\mathrm{d}t+\frac{\varepsilon}{\delta_\varepsilon}\int_{t_i}^{t_i+\delta_\varepsilon}\sigma\left(t\right)\,\mathrm{d}W_t\\
&=f\left(t_i,Y_{t_i}\right)+\sigma\left(t_i\right)\frac{W_{t_i+\delta_\varepsilon}-W_{t_i}}{\sqrt{\delta_\varepsilon}}\frac{\varepsilon}{\sqrt{\delta_\varepsilon}}+\int_{t_i}^{t_i+\delta_\varepsilon}\left[\sigma\left(t\right)-\sigma\left(t_i\right)\right]\mathrm{d}W_t\frac{\varepsilon}{\delta_\varepsilon}\\
&\quad+\frac{1}{\delta_\varepsilon}\int_{t_i}^{t_i+\delta_\varepsilon}\left[f\left(t,Y_t\right)-f\left(t_i,Y_{t_i}\right)\right]\,\mathrm{d}t\\
&=f\left(t_i,Y_{t_i}\right)+\sigma\left(t_i\right)\xi_{t_i,\varepsilon}\sqrt{\varepsilon}+\frac{1}{\delta_\varepsilon}\int_{t_i}^{t_i+\delta_\varepsilon}f'_y\left(t_i,Y_{t_i}\right)\left[Y_t-Y_{t_i}\right]\,\mathrm{d}t,\\
&\quad+\frac{1}{\delta_\varepsilon}\int_{t_i}^{t_i+\delta_\varepsilon}\left[f\left(t,Y_t\right)-f\left(t_i,Y_{t_i}\right)\right]\,\mathrm{d}t+O\left(\varepsilon^{3/2}\right)\\
&=f\left(t_i,Y_{t_i}\right)+\sigma\left(t_i\right)\xi_{t_i,\varepsilon}\sqrt{\varepsilon}+\frac{1}{\delta_\varepsilon}\int_{t_i}^{t_i+\delta_\varepsilon}f'_y\left(t_i,Y_{t_i}\right)\int_{t_i}^{t}a\left(s,Y_s\right)\,\mathrm{d}s\,\mathrm{d}t\\
&\quad+\frac{1}{\delta_\varepsilon}\int_{t_i}^{t_i+\delta_\varepsilon}f'_y\left(t_i,Y_{t_i}\right)\int_{t_i}^{t}b\left(s,Y_s\right)\,\mathrm{d}V_s\,\mathrm{d}t+O\left(\varepsilon^{3/2}\right),\\
&=f\left(t_i,Y_{t_i}\right)+\sigma\left(t_i\right)\xi_{t_i,\varepsilon}\sqrt{\varepsilon}+f'_y\left(t_i,Y_{t_i}\right)b\left(t_i,Y_{t_i}\right)\eta_{t_i,\varepsilon}\sqrt{\varepsilon}+O\left(\varepsilon\right).\end{aligned}$$

Therefore we have as well the representations

$$\begin{aligned}\frac{X_{t_{i+1}+\delta_\varepsilon}-X_{t_{i+1}}}{\delta_\varepsilon}&=f\left(t_{i+1},Y_{t_{i+1}}\right)+\sigma\left(t_{i+1}\right)\xi_{t_{i+1},\varepsilon}\sqrt{\varepsilon}\\
&\quad+f'_y\left(t_{i+1},Y_{t_{i+1}}\right)b\left(t_{i+1},Y_{t_{i+1}}\right)\eta_{t_{i+1},\varepsilon}\sqrt{\varepsilon}+O\left(\varepsilon\right)\end{aligned}$$

and

$$\begin{aligned}
&\frac{X_{t_{i+1}+\delta_\varepsilon}-X_{t_{i+1}}}{\delta_\varepsilon}-\frac{X_{t_i+\delta_\varepsilon}-X_{t_i}}{\delta_\varepsilon}\\
&=f\left(t_{i+1},Y_{t_{i+1}}\right)-f\left(t_i,Y_{t_i}\right)+\sigma\left(t_{i+1}\right)\xi_{t_{i+1},\varepsilon}\sqrt{\varepsilon}\\
&\quad-\sigma\left(t_i\right)\xi_{t_i,\varepsilon}\sqrt{\varepsilon}+f_y'\left(t_{i+1},Y_{t_{i+1}}\right)b\left(t_{i+1},Y_{t_{i+1}}\right)\eta_{t_{i+1},\varepsilon}\sqrt{\varepsilon}\\
&\quad-f_y'\left(t_i,Y_{t_i}\right)b\left(t_i,Y_{t_i}\right)\eta_{t_i,\varepsilon}\sqrt{\varepsilon}+O\left(\varepsilon\right)\\
&=f\left(t_i,Y_{t_{i+1}}\right)-f\left(t_i,Y_{t_i}\right)+f_t'\left(t_i,Y_{t_i}\right)\varepsilon^{2/3}+\sigma\left(t_i\right)\Delta\xi_{t_i,\varepsilon}\varepsilon^{1/2}\\
&\quad+f_y'\left(t_i,Y_{t_i}\right)b\left(t_i,Y_{t_i}\right)\Delta\eta_{t_i,\varepsilon}\varepsilon^{1/2}+O\left(\varepsilon^{5/6}\right)\\
&=f_y'\left(t_i,Y_{t_i}\right)\left[Y_{t_{i+1}}-Y_{t_i}\right]+\sigma\left(t_i\right)\Delta\xi_{t_i,\varepsilon}\varepsilon^{1/2}+f_y'\left(t_i,Y_{t_i}\right)b\left(t_i,Y_{t_i}\right)\Delta\eta_{t_i,\varepsilon}\varepsilon^{1/2}+O\left(\varepsilon^{2/3}\right)\\
&=f_y'\left(t_i,Y_{t_i}\right)b\left(t_i,Y_{t_i}\right)\left[V_{t_{i+1}}-V_{t_i}\right]+\sigma\left(t_i\right)\Delta\xi_{t_i,\varepsilon}\varepsilon^{1/2}\\
&\quad+f_y'\left(t_i,Y_{t_i}\right)b\left(t_i,Y_{t_i}\right)\Delta\eta_{t_i,\varepsilon}\varepsilon^{1/2}+O\left(\varepsilon^{2/3}\right).
\end{aligned}$$

Hence

$$\begin{aligned}
\Psi_{\tau,\varepsilon}=&\sum_{i=0}^{N_{\tau,\varepsilon}-1}f_y'\left(t_i,Y_{t_i}\right)^2b\left(t_i,Y_{t_i}\right)^2\left[V_{t_{i+1}}-V_{t_i}\right]^2+\sum_{i=0}^{N_{\tau,\varepsilon}-1}\sigma\left(t_i\right)^2\left(\Delta\xi_{t_i,\varepsilon}\right)^2\varepsilon\\
&+\sum_{i=0}^{N_{\tau,\varepsilon}-1}f_y'\left(t_i,Y_{t_i}\right)^2b\left(t_i,Y_{t_i}\right)^2\left(\Delta\eta_{t_i,\varepsilon}\right)^2\varepsilon\\
&+2\sum_{i=0}^{N_{\tau,\varepsilon}-1}f_y'\left(t_i,Y_{t_i}\right)b\left(t_i,Y_{t_i}\right)\sigma\left(t_i\right)\Delta\eta_{t_i,\varepsilon}\Delta\xi_{t_i,\varepsilon}\varepsilon\\
&+2\sum_{i=0}^{N_{\tau,\varepsilon}-1}f_y'\left(t_i,Y_{t_i}\right)^2b\left(t_i,Y_{t_i}\right)^2\left[V_{t_{i+1}}-V_{t_i}\right]\Delta\eta_{t_i,\varepsilon}\,\varepsilon^{1/2}\\
&+2\sum_{i=0}^{N_{\tau,\varepsilon}-1}f_y'\left(t_i,Y_{t_i}\right)b\left(t_i,Y_{t_i}\right)\sigma\left(t_i\right)\left[V_{t_{i+1}}-V_{t_i}\right]\Delta\xi_{t_i,\varepsilon}\,\varepsilon^{1/2}+O\left(\varepsilon^{5/6}\right).
\end{aligned}$$

Note that

$$\pi_{1,\varepsilon}\equiv\varphi_\varepsilon\sum_{i=0}^{N_{\tau,\varepsilon}-1}\sigma\left(t_i\right)^2\left(\Delta\xi_{t_i,\varepsilon}\right)^2\longrightarrow2\int_0^\tau\sigma\left(t\right)^2\mathrm{d}t=\pi_1,\tag{4.3}$$

$$\varepsilon\sum_{i=0}^{N_{\tau,\varepsilon}-1}\sigma\left(t_i\right)^2\left(\Delta\xi_{t_i,\varepsilon}\right)^2=2\int_0^\tau\sigma\left(t\right)^2\mathrm{d}t\left(1+o\left(1\right)\right)\,\varepsilon^{1/3},$$

$$\pi_{2,\varepsilon}\equiv\varphi_\varepsilon\sum_{i=0}^{N_{\tau,\varepsilon}-1}f_y'\left(t_i,Y_{t_i}\right)^2b\left(t_i,Y_{t_i}\right)^2\left(\Delta\eta_{t_i,\varepsilon}\right)^2\longrightarrow\frac{2}{3}\int_0^\tau f_y'\left(t,Y_t\right)^2b\left(t,Y_t\right)^2\mathrm{d}t=\pi_2\left(\vartheta_0\right),\tag{4.4}$$

$$\varepsilon\sum_{i=0}^{N_{\tau,\varepsilon}-1}f_y'\left(t_i,Y_{t_i}\right)^2b\left(t_i\right)^2\left(\Delta\eta_{t_i,\varepsilon}\right)^2=\frac{2}{3}\int_0^\tau f_y'\left(t,Y_t\right)^2b\left(t,Y_t\right)^2\mathrm{d}t\left(1+o\left(1\right)\right)\,\varepsilon^{1/3}$$

and

$$\sum_{i=0}^{N_{\tau,\varepsilon}-1}f_y'\left(t_i,Y_{t_i}\right)b\left(t_i,Y_{t_i}\right)\sigma\left(t_i\right)\Delta\eta_{t_i,\varepsilon}\Delta\xi_{t_i,\varepsilon}\,\varepsilon=O\left(\varepsilon^{2/3}\right).$$

It is easy to verify that

$$\begin{aligned}&\mathbf{E}\left(\sum_{i=0}^{N_{\tau,\varepsilon}-1}f_y'\left(t_i,Y_{t_i}\right)b\left(t_i,Y_{t_i}\right)\sigma\left(t_i\right)\Delta\eta_{t_i,\varepsilon}\Delta\xi_{t_i,\varepsilon}\right)^2\\&\quad=\sum_{i=0}^{N_{\tau,\varepsilon}-1}\mathbf{E}f_y'\left(t_i,Y_{t_i}\right)^2b\left(t_i,Y_{t_i}\right)^2\sigma\left(t_i\right)^2\Delta\eta_{t_i,\varepsilon}^2\Delta\xi_{t_i,\varepsilon}^2\\&\quad\leqslant C\sum_{i=0}^{N_{\tau,\varepsilon}-1}\mathbf{E}\Delta\eta_{t_i,\varepsilon}^2\,\mathbf{E}\Delta\xi_{t_i,\varepsilon}^2\leqslant C\,\varepsilon^{-2/3}.\end{aligned}$$

Further we have

$$\sum_{i=0}^{N_{\tau,\varepsilon}-1}f_y'\left(t_i,Y_{t_i}\right)b\left(t_i,Y_{t_i}\right)\sigma\left(t_i\right)\Delta\xi_{t_i,\varepsilon}\left[V_{t_{i+1}}-V_{t_i}\right]=\int_0^\tau H_\varepsilon\left(t\right){\rm d}V_t,$$

where the step function $H_\varepsilon\left(t\right)$ is

$$H_\varepsilon\left(t\right)=f_y'\left(t_i,Y_{t_i}\right)b\left(t_i,Y_{t_i}\right)\sigma\left(t_i\right)\Delta\xi_{t_i,\varepsilon}\;1\!\!1_{\{t_i\leqslant t<t_{i+1}\}}.$$

Recall that $W_t,\ 0\leqslant t\leqslant T$ and $V_t,\ 0\leqslant t\leqslant T$ are independent Wiener processes. Therefore

$$\begin{aligned}\int_0^\tau H_\varepsilon\left(t\right)^2{\rm d}t&=\sum_{i=0}^{N_{\tau,\varepsilon}-1}f_y'\left(t_i,Y_{t_i}\right)^2b\left(t_i,Y_{t_i}\right)^2\sigma\left(t_i\right)^2\Delta\xi_{t_i,\varepsilon}^2\left(t_{i+1}-t_i\right)\\&\longrightarrow2\int_0^\tau f_y'\left(t,Y_t\right)^2b\left(t,Y_t\right)^2\sigma\left(t\right)^2{\rm d}t\end{aligned}$$

because

$$\begin{aligned}&\mathbf{E}\left(\sum_{i=0}^{N_{\tau,\varepsilon}-1}f_y'\left(t_i,Y_{t_i}\right)^2b\left(t_i,Y_{t_i}\right)^2\sigma\left(t_i\right)^2\left(\Delta\xi_{t_i,\varepsilon}^2-2\right)\left(t_{i+1}-t_i\right)\right)^2\\&=\sum_{i=0}^{N_{\tau,\varepsilon}-1}\mathbf{E}\left(\left(f_y'\left(t_i,Y_{t_i}\right)^4b\left(t_i,Y_{t_i}\right)^4\sigma\left(t_i\right)^4\left(\Delta\xi_{t_i,\varepsilon}^2-2\right)\right)^2\right)\left(t_{i+1}-t_i\right)^2\\&\leqslant C\;\varepsilon^{2/3}.\end{aligned}$$

Therefore

$$\int_0^\tau H_\varepsilon\left(t\right){\rm d}V_t\Longrightarrow\tilde\pi\left(\vartheta_0\right),\quad\mathbf{E}\tilde\pi\left(\vartheta_0\right)^2=2\int_0^\tau f_y'\left(t,Y_t\right)^2b\left(t,Y_t\right)^2\sigma\left(t\right))^2{\rm d}t$$

and we obtained

$$\sum_{i=0}^{N_{\tau,\varepsilon}-1}f_y'\left(t_i,Y_{t_i}\right)b\left(t_i,Y_{t_i}\right)\sigma\left(t_i\right)\Delta\xi_{t_i,\varepsilon}\left[V_{t_{i+1}}-V_{t_i}\right]\sqrt{\varepsilon}=O\left(\sqrt{\varepsilon}\right).$$

For estimation of the next term we denote $G\left(t\right)=f_y'\left(t,Y_t\right)^2b\left(t,Y_t\right)^2$ and write

$$\begin{aligned}&\varepsilon^{3/2}\sum_{i=0}^{N_{\tau,\varepsilon}-1}f_y'\left(t_i,Y_{t_i}\right)^2b\left(t_i,Y_{t_i}\right)^2\left[V_{t_{i+1}}-V_{t_i}\right]\Delta\eta_{t_i,\varepsilon}\\&=\sum_{i=0}^{N_{\tau,\varepsilon}-1}G\left(t_i\right)\left[V_{t_{i+1}}-V_{t_i}\right]\left[\int_{t_{i+1}}^{t_{i+1}+\varepsilon}\left[V_s-V_{t_{i+1}}\right]{\rm d}s-\int_{t_i}^{t_i+\varepsilon}\left[V_s-V_{t_i}\right]{\rm d}s\right]\\&=\sum_{i=0}^{N_{\tau,\varepsilon}-1}G\left(t_i\right)\left[V_{t_{i+1}}-V_{t_i}\right]\int_{t_{i+1}}^{t_{i+1}+\varepsilon}\left[V_s-V_{t_{i+1}}\right]{\rm d}s\\&\quad-\sum_{i=0}^{N_{\tau,\varepsilon}-1}G\left(t_i\right)\left[V_{t_{i+1}}-V_{t_i}\right]\int_{t_i}^{t_i+\varepsilon}\left[V_s-V_{t_i}\right]{\rm d}s\\&\equiv S_{1,\varepsilon}-S_{2,\varepsilon}\end{aligned}$$

with obvious notations. We have

$$\begin{aligned}\mathbf{E}S_{1,\varepsilon}^2&=\mathbf{E}\left(\sum_{i=0}^{N_{\tau,\varepsilon}-1}G\left(t_i\right)\left[V_{t_{i+1}}-V_{t_i}\right]\int_{t_{i+1}}^{t_{i+1}+\varepsilon}\left[V_s-V_{t_{i+1}}\right]\mathrm{d}s\right)^2\\&=\sum_{i=0}^{N_{\tau,\varepsilon}-1}\mathbf{E}G\left(t_i\right)^2\left[V_{t_{i+1}}-V_{t_i}\right]^2\mathbf{E}\left(\int_{t_{i+1}}^{t_{i+1}+\varepsilon}\left[V_s-V_{t_{i+1}}\right]\mathrm{d}s\right)^2\leqslant C\,\varepsilon^3.\end{aligned}$$

For the second sum $S_{2,\varepsilon}$ we first note that $V_{t_{i+1}}-V_{t_i+\varepsilon}$ and $\int_{t_i}^{t_i+\varepsilon}\left[V_s-V_{t_i}\right]\mathrm{d}s$ are independent and can be estimated as the sum $S_{1,\varepsilon}$. Therefore it is sufficient to study

$$S_{3,\varepsilon}=\sum_{i=0}^{N_{\tau,\varepsilon}-1}G\left(t_i\right)\left[V_{t_i+\varepsilon}-V_{t_i}\right]\int_{t_i}^{t_i+\varepsilon}\left[V_s-V_{t_i}\right]\mathrm{d}s.$$

We have

$$\begin{aligned}\mathbf{E}S_{3,\varepsilon}^2&\leqslant C\sum_{i=0}^{N_{\tau,\varepsilon}-1}\mathbf{E}\left(\int_{t_i}^{t_i+\varepsilon}\left[V_{t_i+\varepsilon}-V_{t_i}\right]\left[V_s-V_{t_i}\right]\mathrm{d}s\right)^2\\&\leqslant C\varepsilon\sum_{i=0}^{N_{\tau,\varepsilon}-1}\int_{t_i}^{t_i+\varepsilon}\mathbf{E}\left[V_{t_i+\varepsilon}-V_{t_i}\right]^2\left[V_s-V_{t_i}\right]^2\mathrm{d}s\leqslant C\,\varepsilon^{10/3}.\end{aligned}$$

Finally we obtain

$$\varepsilon^{1/2}\sum_{i=0}^{N_{\tau,\varepsilon}-1}f_y'\left(t_i,Y_{t_i}\right)^2b\left(t_i,Y_{t_i}\right)^2\left[V_{t_{i+1}}-V_{t_i}\right]\Delta\eta_{t_i,\varepsilon}=O\left(\varepsilon^{2/3}\right).$$

Consider now the term (see (4.2))

$$\begin{aligned}\sum_{i=0}^{N_{\tau,\varepsilon}-1}G\left(t_i\right)\left[V_{t_{i+1}}-V_{t_i}\right]^2&=\sum_{i=0}^{N_{\tau,\varepsilon}-1}G\left(t_i\right)\left[t_{i+1}-t_i\right]+2\sum_{i=0}^{N_{\tau,\varepsilon}-1}G\left(t_i\right)\int_{t_i}^{t_{i+1}}\left[V_s-V_{t_i}\right]\mathrm{d}V_s\\&=\sum_{i=0}^{N_{\tau,\varepsilon}-1}G\left(t_i\right)\left[t_{i+1}-t_i\right]+\varepsilon^{1/3}\int_0^\tau R_\varepsilon\left(s\right)\mathrm{d}V_s,\end{aligned}\tag{4.5}$$

where we denoted

$$R_\varepsilon\left(s\right)=\varepsilon^{-1/3}G\left(t_i\right)\left[V_s-V_{t_i}\right]\mathbb{1}_{\{t_i\leqslant s<t_{i+1}\}}.$$

We have

$$\begin{aligned}\int_0^\tau R_\varepsilon\left(s\right)^2\mathrm{d}s&=\varepsilon^{-2/3}\sum_{i=0}^{N_{\tau,\varepsilon}-1}G\left(t_i\right)^2\int_{t_i}^{t_{i+1}}\left[V_s-V_{t_i}\right]^2\mathrm{d}s\\&=\sum_{i=0}^{N_{\tau,\varepsilon}-1}G\left(t_i\right)^2\mu_{i,\varepsilon}\left(t_{i+1}-t_i\right)\left(1+o\left(1\right)\right)\longrightarrow\frac{1}{2}\int_0^\tau G\left(s\right)^2\mathrm{d}s.\end{aligned}$$

Here

$$\mu_{i,\varepsilon}=\varepsilon^{-4/3}\int_{t_i}^{t_{i+1}}\left[V_s-V_{t_i}\right]^2\mathrm{d}s=\int_0^1v_{i,\varepsilon}\left(r\right)^2\mathrm{d}r,\quad i=0,1,\ldots$$

are i.i.d. random variables with mean $\frac{1}{2}$ and $v_{i,\varepsilon}\left(r\right),0\leqslant r\leqslant 1$ are independent Wiener processes.

Therefore

$$\pi_{3,\varepsilon}\equiv\varepsilon^{-1/3}\;2\sum_{i=0}^{N_{\tau,\varepsilon}-1}G\left(t_i\right)\int_{t_i}^{t_{i+1}}\left[V_s-V_{t_i}\right]{\rm d}V_s\Longrightarrow\pi_3\left(\vartheta_0\right),\tag{4.6}$$

$$\mathbf{E}\pi_3\left(\vartheta_0\right)^2=\int_0^\tau f_y'\left(t,Y_t\right)^4b\left(t,Y_t\right)^4{\rm d}t.$$

The random variable $\pi_3\left(\vartheta_0\right)$ can be written in the integral form

$$\pi_3\left(\vartheta_0\right)=\int_0^\tau f_y'\left(t,Y_t\right)^2b\left(t,Y_t\right)^2{\rm d}v_t,$$

where $v_t,\ 0\leqslant t\leqslant\tau$ is some Wiener process.

For the term

$$\sum_{i=0}^{N_{\tau,\varepsilon}-1}G\left(t_i\right)\left(t_{i+1}-t_i\right)=\sum_{i=0}^{N_{\tau,\varepsilon}-1}f_y'(t_i,Y_{t_i})^2b\left(t_i,Y_{t_i}\right)^2\left(t_{i+1}-t_i\right)$$

in (4.5) we write

$$\begin{aligned}
&\sum_{i=0}^{N_{\tau,\varepsilon}-1}f_y'(t_i,Y_{t_i})^2b\left(t_i,Y_{t_i}\right)^2\left(t_{i+1}-t_i\right)-\int_0^\tau f_y'(t,Y_t)^2b\left(t,Y_t\right)^2{\rm d}t\\
&=\sum_{i=0}^{N_{\tau,\varepsilon}-1}\int_{t_i}^{t_{i+1}}\left[f_y'(t,Y_t)^2b\left(t,Y_t\right)^2-f_y'(t_i,Y_{t_i})^2b\left(t_i,Y_{t_i}\right)^2\right]{\rm d}t\\
&=2\sum_{i=0}^{N_{\tau,\varepsilon}-1}\int_{t_i}^{t_{i+1}}f_y'(t,Y_t)b\left(t,Y_t\right)\Big[f_{yy}''(t,\tilde Y_t)b\left(t,Y_t\right)\\
&\quad+f_y'(t,Y_t)b_y'(t,\tilde Y_t)\Big]\int_{t_i}^t b\left(s,Y_s\right){\rm d}V_s\;{\rm d}t+O\left(\varepsilon^{2/3}\right)\\
&=2\sum_{i=0}^{N_{\tau,\varepsilon}-1}\int_{t_i}^{t_{i+1}}f_y'(t,Y_t)b\left(t,Y_t\right)^2\left[f_{yy}''(t,Y_t)b\left(t,Y_t\right)\right.\\
&\quad\left.+f_y'(t,Y_t)b_y'\left(t,Y_t\right)\right]\left[V_t-V_{t_i}\right]\;{\rm d}t+O\left(\varepsilon^{2/3}\right)\\
&=2\varepsilon\sum_{i=0}^{N_{\tau,\varepsilon}-1}f_y'(t_i,Y_{t_i})b\left(t_i,Y_{t_i}\right)^2\left[f_{yy}''(t_i,Y_{t_i})b\left(t_i,Y_{t_i}\right)+f_y'(t_i,Y_{t_i})b_y'\left(t_i,Y_{t_i}\right)\right]\tilde\eta_{i,\varepsilon}+O\left(\varepsilon^{2/3}\right)\\
&=\varepsilon^{1/3}\sum_{i=0}^{N_{\tau,\varepsilon}-1}L_\varepsilon\left(t_i\right)\tilde\eta_{i,\varepsilon}\left(t_{i+1}-t_i\right)+O\left(\varepsilon^{2/3}\right),
\end{aligned}$$

where the two last lines define $L_\varepsilon\left(t_i\right)$ and we denoted

$$\tilde\eta_{i,\varepsilon}=\varepsilon^{-1}\int_{t_i}^{t_{i+1}}\left[V_t-V_{t_i}\right]\;{\rm d}t\sim\mathcal{N}\left(0,1/3\right).$$

We have

$$\mathbf{E}\left(\sum_{i=0}^{N_{\tau,\varepsilon}-1}L_\varepsilon\left(t_i\right)\tilde\eta_{i,\varepsilon}\left(t_{i+1}-t_i\right)\right)^2=\frac{1}{3}\sum_{i=0}^{N_{\tau,\varepsilon}-1}\mathbf{E}L_\varepsilon\left(t_i\right)^2\left(t_{i+1}-t_i\right)^2=O\left(\varepsilon^{2/3}\right).$$

Hence

$$\varepsilon^{1/3}\sum_{i=0}^{N_{\tau,\varepsilon}-1}L_\varepsilon\left(t_i\right)\tilde\eta_{i,\varepsilon}\left(t_{i+1}-t_i\right)+O\left(\varepsilon^{2/3}\right)=O\left(\varepsilon^{2/3}\right).$$

Finally,

$$\begin{gathered}\Psi_{\tau,\varepsilon}=\Psi_\tau\left(\vartheta_0\right)+\pi_\varepsilon\,\varepsilon^{1/3}+O\left(1\right)\ \varepsilon^{2/3},\quad \pi_\varepsilon=\pi_{1,\varepsilon}+\pi_{2,\varepsilon}+\pi_{3,\varepsilon},\\ \varepsilon^{-1/3}\left(\Psi_{\tau,\varepsilon}-\Psi_\tau\left(\vartheta_0\right)\right)\Longrightarrow\pi\left(\vartheta_0\right).\end{gathered}\tag{4.7}$$

The term $O\left(1\right)$ here has finite all polynomial moments. □

5. Method of moments estimator

The method of moments estimator (MME) ϑ_ε^* can be introduced by the equation

$$\Psi_{\tau,\varepsilon}^*=\Psi_\tau^*\left(\vartheta_{\tau,\varepsilon}^*\right),\tag{5.1}$$

where we slightly modified the definitions of $\Psi_{\tau,\varepsilon}$ and $\Psi_\tau\left(\vartheta\right)$

$$\Psi_{\tau,\varepsilon}^*=\sum_{i=i_0}^{N_{\tau,\varepsilon}-1}\left(\frac{X_{t_{i+1}+\varepsilon}-X_{t_{i+1}}}{\varepsilon}-\frac{X_{t_i+\varepsilon}-X_{t_i}}{\varepsilon}\right)^2,\tag{5.2}$$

Here $0<t_0<\tau<T,\ t_{i+1}-t_i=\varepsilon^{2/3},\ i_0=\left[\frac{t_0}{\varepsilon^{2/3}}\right],\ N_{\tau,\varepsilon}=\left[\frac{\tau}{\varepsilon^{2/3}}\right]$, $[A]$ is integer part of A.

The random function

$$\Psi_\tau^*\left(\vartheta\right)=\int_{t_0}^{\tau}f_y'(\vartheta,t,m_t\left(\vartheta\right))^2\,b\left(\vartheta,t,m_t\left(\vartheta\right)\right)^2\mathrm{d}t,\quad \vartheta\in\Theta$$

is continuous with probability 1. The random process $m_t\left(\vartheta\right),\ t_0\leqslant t\leqslant T$ is solution of the equation

$$\mathrm{d}m_t\left(\vartheta\right)=a\left(\vartheta,t,m_t\left(\vartheta\right)\right)\mathrm{d}t+\frac{b\left(\vartheta,t,m_t\left(\vartheta\right)\right)}{\varepsilon\sigma\left(t\right)}\left[\mathrm{d}X_t-f\left(\vartheta,t,m_t\left(\vartheta\right)\right)\mathrm{d}t\right],\quad m_{t_0}\left(\vartheta\right).\tag{5.3}$$

Here the initial value $m_{t_0}\left(\vartheta\right)$ is solution of the system (2.1), (2.2). As the equation (3.1) is simpler to solve than (2.1), (2.2) we start these two statistics $\Psi_\tau^*\left(\vartheta\right)$ and $\Psi_{\tau,\varepsilon}^*$ from the moment $t_0\in\left(0,\tau\right)$.

We have to make several remarks concerning the definition (5.1)–(5.3) of the MME $\vartheta_{\tau,\varepsilon}^*$. For a given ε and observed trajectory $X^\tau=\left(X_t,0\leqslant t\leqslant\tau\right)$ we obtain a random interval $\left[\psi_m,\psi_M\right]$, where

$$\psi_m=\inf_{\vartheta\in\Theta}\Psi_\tau\left(\vartheta\right),\quad \psi_M=\sup_{\vartheta\in\Theta}\Psi_\tau\left(\vartheta\right).$$

If $\Psi_{\tau,\varepsilon}^*\notin\left[\psi_m,\psi_M\right]$, then (5.1) has no solution and has to be replaced by the definition

$$\vartheta_{\tau,\varepsilon}^*=\alpha\mathbb{1}_{\left\{\Psi_{\tau,\varepsilon}<\psi_m\right\}}+\bar\vartheta_{\tau,\varepsilon}\mathbb{1}_{\left\{\psi_m<\Psi_{\tau,\varepsilon}<\psi_M\right\}}+\beta\mathbb{1}_{\left\{\Psi_{\tau,\varepsilon}>\psi_M\right\}},\quad \Psi_{\tau,\varepsilon}=\Psi_\tau\left(\bar\vartheta_{\tau,\varepsilon}\right).$$

Note that in the case of the consistent estimation the probabilities of the events $\Psi_{\tau,\varepsilon}<\psi_m$ and $\Psi_{\tau,\varepsilon}>\psi_M$ are asymptotically negligible.

The equation (5.3) can be simplified as follows. The initial value $m_{t_0}\left(\vartheta\right)$ can be replaced by any other value, say, y_0 and the trend coefficient $a\left(\vartheta,t,m_t\left(\vartheta\right)\right)$ can be omitted. Let us explain why (5.3) can be replaced by

$$\mathrm{d}m_t^*\left(\vartheta\right)=\frac{b\left(\vartheta,t,m_t^*\left(\vartheta\right)\right)}{\varepsilon\sigma\left(t\right)}\left[\mathrm{d}X_t-f\left(\vartheta,t,m_t^*\left(\vartheta\right)\right)\mathrm{d}t\right],\quad m_{t_0}^*\left(\vartheta\right)=y_0,\quad t_0\leqslant t\leqslant\tau.\tag{5.4}$$

Consider the difference $v_t=m_t^*\left(\vartheta\right)-m_t\left(\vartheta\right)$ and the equation for v_t^4

$$\frac{\mathrm{d}v_t^4}{\mathrm{d}t}=-\frac{4b\left(\vartheta,t,m_t^*\left(\vartheta\right)\right)}{\varepsilon\sigma\left(t\right)}v_t^3\left[f\left(\vartheta,t,m_t^*\left(\vartheta\right)\right)-f\left(\vartheta,t,m_t\left(\vartheta\right)\right)\right]-4v_t^3a\left(\vartheta,t,m_t\left(\vartheta\right)\right).$$

We have

$$\begin{aligned}
\mathbf{E}_{\vartheta_0} v_t^4 &= \mathbf{E}_{\vartheta_0}\left(m_{t_0}^*(\vartheta)-y_0\right)^4 - 4\int_{t_0}^{t} \mathbf{E}_{\vartheta_0} v_s^3 a\left(\vartheta,s,m_s(\vartheta)\right){\rm d}s\\
&\quad -\int_{t_0}^{t}\frac{4}{\varepsilon\sigma(s)}\,\mathbf{E}_{\vartheta_0} v_s^3 b\left(\vartheta,s,m_s^*(\vartheta)\right)\left[f\left(\vartheta,s,m_s^*(\vartheta)\right)-f\left(\vartheta,s,m_s(\vartheta)\right)\right]{\rm d}s\\
&= \mathbf{E}_{\vartheta_0}\left(m_{t_0}^*(\vartheta)-y_0\right)^4 - \int_{t_0}^{t}\frac{4}{\varepsilon\sigma(s)}\,\mathbf{E}_{\vartheta_0} v_s^4 b\left(\vartheta,s,m_s(\vartheta)\right)f_y'\left(\vartheta,s,\bar m_s\right){\rm d}s-\\
&\quad -4\int_{t_0}^{t}\mathbf{E}_{\vartheta_0} v_s^3 a\left(\vartheta,s,m_s(\vartheta)\right){\rm d}s\\
&\leqslant C-\frac{4c_bc_y}{\varepsilon C_\sigma}\int_{t_0}^{t}\mathbf{E}_{\vartheta_0} v_s^4{\rm d}s+4\int_{t_0}^{t}\mathbf{E}_{\vartheta_0} v_s^4{\rm d}s+4\int_{t_0}^{t}\mathbf{E}_{\vartheta_0} a\left(\vartheta,s,m_s(\vartheta)\right)^4{\rm d}s\\
&\leqslant C-\frac{2c_bc_y}{\varepsilon C_\sigma}\int_{t_0}^{t}\mathbf{E}_{\vartheta_0} v_s^4{\rm d}s
\end{aligned}$$

for sufficiently small ε. Here following [26], Lemma 2.1 we used the estimate $fg \leqslant f^k+g^m$, where $f>0$, $g>0$ and $k^{-1}+m^{-1}=1$, which allows us to write $\mathbf{E}_{\vartheta_0}\left|v_s^3 a\left(\vartheta,s,m_s(\vartheta)\right)\right|\leqslant \mathbf{E}_{\vartheta_0} v_s^4 + \mathbf{E}_{\vartheta_0} a\left(\vartheta,s,m_s(\vartheta)\right)^4$, i.e., $k=4/3$ and $m=4$.

Therefore

$$\mathbf{E}_{\vartheta_0}\left(m_t^*(\vartheta)-m_t(\vartheta)\right)^4\leqslant Ce^{-\frac{2c_bc_y}{\varepsilon C_\sigma}(t-t_0)}$$

and this proves that the difference between the solutions of (5.3) and (5.4) is asymptotically negligible. This is typical result for such short memory stochastic systems (see Corollary 2.3 and remarks just after in [26]).

Therefore the statistic $\Psi_\tau(\vartheta)$ can be replaced by

$$\Psi_\tau^*(\vartheta)=\int_{t_0}^{\tau} f_y'(\vartheta,t,m_t^*(\vartheta))^2\, b\left(\vartheta,t,m_t^*(\vartheta)\right)^2{\rm d}t,\qquad \vartheta\in\Theta \tag{5.5}$$

with the corresponding definition of the MME $\vartheta_{\tau,\varepsilon}^*$.

We will need the estimator $\vartheta_{\tau,\varepsilon}^*$ in the next section for the construction of the adaptive EKF and the direct substitution of it in the EKF leads to the equation with the stochastic integral

$$\int_{t_0}^{t}\frac{b\left(\vartheta_{\tau,\varepsilon}^*,s,m_s^*\left(\vartheta_{\tau,\varepsilon}^*\right)\right)}{\sigma(s)}{\rm d}X_s.$$

As the estimator $\vartheta_{\tau,\varepsilon}^*$ depends on the whole trajectory X^τ this Itô integral is not well defined. To avoid this problem we consider construction of this estimator for the values $0<t_0<\tau<T$, where t_0 and τ are small and the adaptive filter $m_t^\star$ will be given for the values $t\in[\tau,T]$.

Let us denote

$$\begin{aligned}
\Psi_\tau^\circ(\vartheta)&=\int_{t_0}^{\tau} f_y'(\vartheta,t,m_t^\circ(\vartheta))^2\, b\left(\vartheta,t,m_t^\circ(\vartheta)\right)^2{\rm d}t,\qquad \vartheta\in\Theta,\\
H(\vartheta,t,y,z)&=2\left[\dot f_{\vartheta y}'(\vartheta,t,y)+f_{yy}''(\vartheta,t,y)\,z\right]f_y'(\vartheta,t,y)b\left(\vartheta,t,y\right)^2\\
&\quad+2\left[\dot b_\vartheta\left(\vartheta,t,y\right)+b_y'\left(\vartheta,t,y\right)z\right]f_y'(\vartheta,t,y)^2 b\left(\vartheta,t,y\right),\\
H^\circ(\vartheta,t,y)&=2\left[\dot f_{\vartheta y}'(\vartheta,t,y)f_y'(\vartheta,t,y)-f_{yy}''(\vartheta,t,y)\,\dot f_\vartheta(\vartheta,t,y)\right]b\left(\vartheta,t,y\right)^2\\
&\quad+2\left[\dot b_\vartheta\left(\vartheta,t,y\right)f_y'(\vartheta,t,y)-b_y'\left(\vartheta,t,y\right)\dot f_\vartheta(\vartheta,t,y)\right]f_y'(\vartheta,t,y)b\left(\vartheta,t,y\right).
\end{aligned}$$

The MME is defined by the equations (5.1)–(5.5).

Introduce the conditions $\mathscr{D}$

($\mathscr{D}_1$) *There exists a constant $c_H > 0$, which does not depend on $\vartheta \in \Theta, t \in [t_0, \tau], y \in \mathcal{R}$ such that*

$$H^\circ(\vartheta, t, y) \geqslant c_H.$$

($\mathscr{D}_2$) *The derivatives $\dot f''_{\vartheta yy}(\vartheta, t, y)$ and $f^{(3)}_{yyy}(\vartheta, t, y)$ are bounded.*

Proposition 3 *Let the conditions $\mathscr{C}, \mathscr{D}$ be satisfied, then the MME $\vartheta^*_{\tau,\varepsilon}$ is consistent and*

$$\varepsilon^{-1/3}\left(\vartheta^*_{\tau,\varepsilon} - \vartheta_0\right) = \frac{\pi(\vartheta_0)}{\int_{t_0}^{\tau} H^\circ(\vartheta_0, t, Y_t)\,\mathrm{d}t} + O\left(\varepsilon^{1/6}\right).$$

Proof To prove the consistency we write as usual in such problems the following chain of relations: for any $\nu > 0$

$$\begin{aligned}
&\mathbf{P}_{\vartheta_0}\left(\inf_{|\vartheta-\vartheta_0|<\nu}\left|\vartheta^*_{\tau,\varepsilon} - \vartheta_0\right| > \inf_{|\vartheta-\vartheta_0|\geqslant\nu}\left|\vartheta^*_{\tau,\varepsilon} - \vartheta_0\right|\right)\\
&= \mathbf{P}_{\vartheta_0}\left(\inf_{|\vartheta-\vartheta_0|<\nu}\left|\Psi_{\tau,\varepsilon} - \Psi^*_\tau(\vartheta)\right| > \inf_{|\vartheta-\vartheta_0|\geqslant\nu}\left|\Psi_{\tau,\varepsilon} - \Psi^*_\tau(\vartheta)\right|\right)\\
&\leqslant \mathbf{P}_{\vartheta_0}\left(\inf_{|\vartheta-\vartheta_0|<\nu}\Big(\left|\Psi_{\tau,\varepsilon} - \Psi^*_\tau(\vartheta_0)\right| + \left|\Psi^*_\tau(\vartheta) - \Psi^*_\tau(\vartheta_0)\right|\Big)\right.\\
&\qquad \left.> \inf_{|\vartheta-\vartheta_0|\geqslant\nu}\Big(\left|\Psi^*_\tau(\vartheta) - \Psi^*_\tau(\vartheta_0)\right| - \left|\Psi_{\tau,\varepsilon} - \Psi^*_\tau(\vartheta_0)\right|\Big)\right)\\
&= \mathbf{P}_{\vartheta_0}\left(2\left|\Psi_{\tau,\varepsilon} - \Psi^*_\tau(\vartheta_0)\right| \geqslant \inf_{|\vartheta-\vartheta_0|\geqslant\nu}\left|\Psi^*_\tau(\vartheta) - \Psi^*_\tau(\vartheta_0)\right|\right).
\end{aligned}$$

Recall that $m^*_t(\vartheta) = m^\circ_t(\vartheta) + O(\sqrt{\varepsilon})$, $\dot m^*_t(\vartheta) = \dot m^\circ_t(\vartheta) + O(\sqrt{\varepsilon})$ (see (3.7), (3.8)) and $\dot m^\circ_t(\vartheta) = -\dot f_\vartheta(\vartheta, t, m^\circ_t(\vartheta))\, f'_y(\vartheta, t, m^\circ_t(\vartheta))^{-1}$ (see (3.6)). Note as well that the derivatives of $\Psi^*_\tau(\vartheta)$ and $\Psi^\circ_\tau(\vartheta)$ are

$$\begin{aligned}
\dot\Psi^*_\tau(\vartheta) &= \int_{t_0}^{\tau} H\left(\vartheta, t, m^*_t(\vartheta), \dot m^*_t(\vartheta)\right)\mathrm{d}t, \quad \dot\Psi^\circ_\tau(\vartheta) = \int_{t_0}^{\tau} H^\circ\left(\vartheta, t, m^\circ_t(\vartheta)\right)\mathrm{d}t,\\
\dot\Psi^\circ_\tau(\vartheta_0) &= \int_{t_0}^{\tau} H^\circ\left(\vartheta_0, t, m^\circ_t(\vartheta_0)\right)\mathrm{d}t = \int_{t_0}^{\tau} H^\circ\left(\vartheta_0, t, Y_t\right)\mathrm{d}t.
\end{aligned}$$

Therefore

$$\begin{aligned}
\Psi^*_\tau(\vartheta) - \Psi^*_\tau(\vartheta_0) &= \Psi^\circ_\tau(\vartheta) - \Psi^\circ_\tau(\vartheta_0) + O\left(\sqrt{\varepsilon}\right)\\
&= (\vartheta - \vartheta_0)\,\dot\Psi^\circ_\tau(\tilde\vartheta) + O\left(\sqrt{\varepsilon}\right)\\
&= (\vartheta - \vartheta_0)\int_{t_0}^{\tau} H^\circ(\tilde\vartheta, t, m^\circ_t(\tilde\vartheta))\mathrm{d}t + O(1)\sqrt{\varepsilon}
\end{aligned}$$

and

$$\begin{aligned}
&\mathbf{P}_{\vartheta_0}\left(2\left|\Psi^*_{\tau,\varepsilon} - \Psi^*_\tau(\vartheta_0)\right| \geqslant \inf_{|\vartheta-\vartheta_0|\geqslant\nu}\left|\Psi^*_\tau(\vartheta) - \Psi^*_\tau(\vartheta_0)\right|\right)\\
&\geqslant \mathbf{P}_{\vartheta_0}\left(2\left|\Psi^*_{\tau,\varepsilon} - \Psi^*_\tau(\vartheta_0)\right| \geqslant \inf_{|\vartheta-\vartheta_0|\geqslant\nu}|\vartheta - \vartheta_0|\left|\int_{t_0}^{\tau} H^\circ(\tilde\vartheta, t, m^\circ_t(\tilde\vartheta))\mathrm{d}t\right|\right)\\
&= \mathbf{P}_{\vartheta_0}\Big(\left|\Psi_{\tau,\varepsilon} - \Psi^*_\tau(\vartheta_0)\right| \geqslant \frac{\nu\, c_H}{4}(\tau - t_0)\Big)\\
&= \mathbf{P}_{\vartheta_0}\Big(O\left(\varepsilon^{1/3}\right) \geqslant \frac{\nu\, c_H}{4}\,(\tau - t_0)\Big) \longrightarrow 0.
\end{aligned}$$

The left and right sides of the equation $\Psi^*_{\tau,\varepsilon}=\Psi^*_\tau\left(\vartheta^*_{\tau,\varepsilon}\right)$ can be written as follows

$$\Psi^*_{\tau,\varepsilon}=\Psi_\tau\left(\vartheta_0\right)+\pi\left(\vartheta_0\right)\varepsilon^{1/3}+O\left(\varepsilon^{2/3}\right)=\Psi^\circ_\tau\left(\vartheta_0\right)+\pi\left(\vartheta_0\right)\varepsilon^{1/3}+O\left(\varepsilon^{1/2}\right),$$

$$\Psi^*_\tau\left(\vartheta^*_{\tau,\varepsilon}\right)=\Psi^\circ_\tau\left(\vartheta^*_{\tau,\varepsilon}\right)+O\left(\sqrt{\varepsilon}\right)=\Psi^\circ_\tau\left(\vartheta_0\right)+\left(\vartheta^*_{\tau,\varepsilon}-\vartheta_0\right)\dot\Psi^\circ_\tau(\tilde\vartheta)+O\left(\sqrt{\varepsilon}\right)$$

and

$$\varepsilon^{-1/3}\left(\vartheta^*_{\tau,\varepsilon}-\vartheta_0\right)=\frac{\varepsilon^{-1/3}\left(\Psi^*_{\tau,\varepsilon}-\Psi^\circ_\tau\left(\vartheta_0\right)\right)}{\dot\Psi^\circ_\tau(\vartheta_0)+O\left(\sqrt{\varepsilon}\right)}=\frac{\pi\left(\vartheta_0\right)}{\int_{t_0}^\tau H^\circ\left(\vartheta_0,t,Y_t\right)\mathrm{d}t}+O\left(\varepsilon^{1/6}\right).$$

□

6. Adaptive EKF

Recall that we have the model is

$$\begin{aligned}\mathrm{d}X_t&=f\left(\vartheta,t,Y_t\right)\mathrm{d}t+\varepsilon\sigma\left(t\right)\mathrm{d}W_t,\quad X_0=0,\quad 0\leqslant t\leqslant T,\\ \mathrm{d}Y_t&=a\left(\vartheta,t,Y_t\right)\mathrm{d}t+b\left(\vartheta,t,Y_t\right)\mathrm{d}V_t,\quad Y_0=y_0,\end{aligned}$$

where the observed process is $X^T=\left(X_t,0\leqslant t\leqslant T\right)$ and the diffusion process Y^T is hidden. The parameter $\vartheta\in\Theta=\left(\alpha,\beta\right)$ is unknown and we have to construct an approximation $m^\star_{t,\varepsilon}$ of the unobserved component Y_t. We expect that this approximation has to be consistent $\left(m^\star_{t,\varepsilon}-Y_t\rightarrow 0\right)$ and its calculation has to be relatively simple.

The observation interval $[0,T]$ is subdivided on three intervals: $[0,t_0]$, $[t_0,\tau]$ and $[\tau,T]$. The observations on $[0,t_0]$ are neglected because our problem involves a boundary layer at $t=0$. During this time interval the solution $\varepsilon^{-1}\gamma_\varepsilon\left(\vartheta,t\right)$ of the Riccati equation (2.2) exponentially fast converges to the value $\gamma_*\left(\vartheta,t\right)$. Therefore starting since t_0 we can replace $\varepsilon^{-1}\gamma_\varepsilon\left(\vartheta,t\right)$ in the equations (2.1), (2.2) by $\gamma_*\left(\vartheta,t\right)$. This simplifies the further calculations. By the observations $X_t,\ t_0\leqslant t\leqslant\tau$ we construct the MME $\vartheta^*_{\tau,\varepsilon}$ (see (5.1)–(5.5)).

The adaptive EKF is

$$\mathrm{d}m^\star_{t,\varepsilon}=\frac{b\left(\vartheta^*_{\tau,\varepsilon},t,m^\star_{t,\varepsilon}\right)}{\varepsilon\sigma\left(t\right)}\left[\mathrm{d}X_t-f\left(\vartheta^*_{\tau,\varepsilon},t,m^*_{t,\varepsilon}\right)\mathrm{d}t\right],\quad m^\star_{t_0,\varepsilon},\quad \tau\leqslant t\leqslant T.$$

The initial value $m^\star_{t_0,\varepsilon}$ is arbitrary, since it can be proved that the initial value leads to a boundary layer term.

Theorem 1 *Suppose that the conditions $\mathscr{C},\mathscr{D}$ are satisfied. Then*

$$\varepsilon^{-1/3}\left(m^\star_{t,\varepsilon}-Y_t\right)\Longrightarrow-\frac{\dot f_\vartheta\left(\vartheta_0,t,Y_t\right)}{f'_y\left(\vartheta_0,t,Y_t\right)}\frac{\pi\left(\vartheta_0\right)}{\int_{t_0}^\tau H^\circ\left(\vartheta_0,t,Y_t\right)}.$$

Proof Consider the difference $m^\star_t-Y_t$. We have

$$\begin{aligned}m^\star_{t,\varepsilon}-Y_t&=m_{t,\varepsilon}\left(\vartheta^*_{\tau,\varepsilon}\right)-m_t\left(\vartheta_0\right)+m_t\left(\vartheta_0\right)-Y_t\\&=\dot m_{t,\varepsilon}\left(\bar\vartheta\right)\left(\vartheta^*_{\tau,\varepsilon}-\vartheta_0\right)+m_t\left(\vartheta_0\right)-m^\circ_t\left(\vartheta_0\right)\\&=\dot m_t\left(\vartheta_0\right)\left(\vartheta^*_{\tau,\varepsilon}-\vartheta_0\right)\left(1+O\left(\varepsilon^{1/3}\right)\right)+\Gamma^\circ\left(\vartheta_0,t\right)\left[\xi_{t,\varepsilon}-\eta_{t,\varepsilon}\right]\sqrt{\varepsilon}\left(1+o\left(1\right)\right)\\&=\dot m^\circ_{t,\varepsilon}\left(\vartheta_0\right)\left(\vartheta^*_{\tau,\varepsilon}-\vartheta_0\right)\left(1+O\left(\varepsilon^{1/3}\right)\right)+\Gamma^\circ\left(\vartheta_0,t\right)\left[\xi_{t,\varepsilon}-\eta_{t,\varepsilon}\right]\sqrt{\varepsilon}\left(1+o\left(1\right)\right)\\&=-\frac{\dot f_\vartheta\left(\vartheta_0,t,Y_t\right)}{f'_y\left(\vartheta_0,t,Y_t\right)}\left(\vartheta^*_{\tau,\varepsilon}-\vartheta_0\right)\left(1+O\left(\varepsilon^{1/3}\right)\right)+\Gamma^\circ\left(\vartheta_0,t\right)\left[\xi_{t,\varepsilon}-\eta_{t,\varepsilon}\right]\sqrt{\varepsilon}\left(1+o\left(1\right)\right).\end{aligned}$$

Therefore

$$\varepsilon^{-1/3}\left(m^{\star}_{t,\varepsilon}-Y_t\right)\Longrightarrow-\frac{\dot f_{\vartheta}\left(\vartheta_0,t,Y_t\right)}{f'_y\left(\vartheta_0,t,Y_t\right)}\frac{\pi\left(\vartheta_0\right)}{\int_{t_0}^{\tau}H^{\circ}\left(\vartheta_0,t,Y_t\right){\rm d}t}.$$

□

7. Conclusions

The presented here AEKF is just a consistent approximation of the unknown conditional expectation and the question of the construction of asymptotically efficient approximation is always open. The difficulties are related with the absence of closed form equations for conditional expectation as in the linear case considered in [16]. Remark that if we have approximation of the conditional expectation (see [2]) and of its derivative w.r.t. ϑ, then it will be possible to use the One-step MLE-process which provides the recurrent consistent estimation of the unknown parameter and together with the recurrent extended Kalman filter this couple can provide the asymptotically optimal AEKF as it was already done several times for the linear models [16–18]. The first step, of course, has to be the construction of the preliminary estimator of the unknown parameter with the optimal rate $\varepsilon^{1/2}$ as in [8].

Acknowledgements

I would like to express my gratitude to Referees for helpful comments and suggestions. This research was financially supported by the Russian Science Foundation research project (Grant No. 24-11-00191).

References

[1] Bensoussan, A., On some approximation techniques in non linear filtering, In: Proceedings of June 1986 Workshop, Institute for Mathematics and Its Applications, University of Minnesota, 1986.

[2] Bensoussan, A., Stochastic Control of Partially Observable Systems, Cambridge University Press, Cambridge, 1992.

[3] Bobrovsky, B.-Z., Zakai, M. and Zeitouni, O., Error bounds for the nonlinear filtering of signals with small diffusion coefficients, IEEE Transactions on Information Theory, 1988, 34(4): 710–721.

[4] Cappé, O., Moulines, E. and Rydén, T., Inference in Hidden Markov Models, Springer, New York, 2005.

[5] Chui, C. K. and Chen, G., Kalman Filtering with Real-Time Applications, Springer, Berlin, 1999.

[6] Elliott, R. J., Aggoun, L. and Moor, J. B., Hidden Markov Models, Springer, New York, 1995.

[7] Gloter, A. and Jacod, J., Diffusion with measurement errors. I. Local asymptotic normality, ESAIM: Probability and Statistics, 2001, 5: 225–242.

[8] Gloter, A. and Jacod, J., Diffusion with measurement errors. II. Optimal estimators, ESAIM: Probability and Statistics, 2001, 5: 243–260.

[9] Haykin, S., Adaptive Filter Theory, 5th ed., Pearson Education, Boston, 2014.

[10] James, M. and Le Gland, F., Consistent parameter estimation for partially observed diffusions with small noise, Applied Mathematics and Optimization, 1995, 32: 47–72.

[11] Kalman, R. E. and Bucy, R. S., New results in linear filtering and prediction theory, Journal of Basic Engineering, 1961, 83(1): 95–108.

[12] Katzur, R., Bobrovsky, B. Z. and Schuss, Z., Asymptotic analysis of the optimal filtering problem for one-dimensional diffusions measured in a low noise channel, Part II, SIAM Journal on Applied Mathematics, 1984, 44(6): 1176–1191.

[13] Kutoyants, Y. A., Identification of Dynamical Systems with Small Noise, Kluwer Academic Publishers, Dordrecht,

1994.

[14] Kutoyants, Y. A., On parameter estimation of the hidden Ornstein-Uhlenbeck process, Journal of Multivariate Analysis, 2019, 169: 248−263.

[15] Kutoyants, Y. A., Parameter estimation for continuous time hidden Markov processes, Automation and Remote Control, 2020, 81(3): 446−469.

[16] Kutoyants, Y. A., Volatility estimation of hidden Markov process and adaptive filtration, Stochastic Processes and their Applications, 2024, 173: 104381.

[17] Kutoyants, Y. A., Hidden ergodic Ornstein-Uhlenbeck process and adaptive filter, Statistical Inference for Stochastic Processes, 2025, 28: 4.

[18] Kutoyants, Y. A., Hidden AR process and adaptive Kalman filter, Annals of the Institute of Statistical Mathematics, 2024, 77(1): 61−103.

[19] Kutoyants, Y. A. and Zhou, L., On parameter estimation of the hidden Gaussian process in perturbed SDE, Electronic Journal of Statistics, 2021, 15(1): 211−234.

[20] Liptser, R. S. and Shiryaev, A. N., Statistics of Random Processes II: Applications, 2nd ed., Springer, New York, 2001.

[21] Ljung, L., System Identification: Theory for the User, 2nd ed., Prentice-Hall, New Jersey, 1999.

[22] Marchthaler, R., Adaptive extended Kalman filter (ROSE-Filter) for positioning system, arXiv: 2108.11321, 2021.

[23] Milheiro De Oliveira, P., Approximate filters for a nonlinear discrete time filtering problem with small observation noise, Stochastics and Stochastic Reports, 1994, 46(1–2): 1−24.

[24] Mishra, A. K., Shimjith, S. R. and Tiwari, A. P., Adaptive extended Kalman filtering for reactivity estimation, IFAC-PapersOnLine, 2018, 51(1): 702−707.

[25] Ozbek, L. and Efe, M., An adaptive extended Kalman filter with application to compartment models, Communications in Statistics-Simulation and Computation, 2004, 33(1): 145−158.

[26] Picard, J., Nonlinear filtering of one-dimensional diffusion in the case of a high signal-to-noise ratio, SIAM Journal on Applied Mathematics, 1986, 46(6): 1098−1125.

[27] Picard, J., Efficiency of the extended Kalman filter for nonlinear systems with small noise, SIAM Journal on Applied Mathematics, 1991, 51(4): 843−885.

[28] Skorokhod, A. V. and Khasminskii, R. Z., On parameter estimation by indirect observations, Problems of Information Transmission, 1996, 32(1): 58−68.