Drift parameter estimation in stochastic differential equation with multiplicative stochastic volatility

Meriem Bel Hadj Khlifa\textsuperscript{a}, Yuliya Mishura\textsuperscript{b}, Kostiantyn Ralchenko\textsuperscript{b,∗}, Mounir Zili\textsuperscript{a}

\textsuperscript{a}University of Monastir, Faculty of Sciences of Monastir, Department of Mathematics, Avenue de l’Environnement, 5000, Monastir, Tunisia

\textsuperscript{b}Department of Probability Theory, Statistics and Actuarial Mathematics, Taras Shevchenko National University of Kyiv, 64 Volodymyrska, 01601 Kyiv, Ukraine

meriem.bhk@outlook.fr (M. Bel Hadj Khlifa), myus@univ.kiev.ua (Yu. Mishura), k.ralchenko@gmail.com (K. Ralchenko), Mounir.Zili@fsm.rnu.tn (M. Zili)

Received: 9 November 2016, Revised: 3 December 2016, Accepted: 4 December 2016, Published online: 13 December 2016

Abstract We consider a stochastic differential equation of the form

\[ dX_t = \theta a(t, X_t) \, dt + \sigma_1(t, X_t) \sigma_2(t, Y_t) \, dW_t \]

with multiplicative stochastic volatility, where \( Y \) is some adapted stochastic process. We prove existence–uniqueness results for weak and strong solutions of this equation under various conditions on the process \( Y \) and the coefficients \( a, \sigma_1, \text{ and } \sigma_2 \). Also, we study the strong consistency of the maximum likelihood estimator for the unknown parameter \( \theta \). We suppose that \( Y \) is in turn a solution of some diffusion SDE. Several examples of the main equation and of the process \( Y \) are provided supplying the strong consistency.

Keywords Stochastic differential equation, weak and strong solutions, stochastic volatility, drift parameter estimation, maximum likelihood estimator, strong consistency

2010 MSC 60H10, 62F10, 62F12
1 Introduction

The goal of the paper is to study the stochastic differential equation (SDE), the diffusion coefficient of which includes an additional stochastic process:

\[ dX_t = \theta a(t, X_t) \, dt + \sigma(t, X_t, Y_t) \, dW_t, \tag{1} \]

where \( \sigma(t, x, y) = \sigma_1(t, x)\sigma_2(t, y) \), and to estimate the drift parameter \( \theta \) by the observations of stochastic processes \( X \) and \( Y \). Such equations often arise as models of a financial market in mathematical finance. For example, one of the first models of such a type with \( \sigma(t, x, y) = xy \) was proposed in [8], where \( Y \) was the square root of the geometric Brownian motion process. A similar model was considered by Heston [6], where the volatility was governed by the Ornstein–Uhlenbeck process. Fouque et al. used the model with stochastic volatility driven by the Cox–Ingersoll–Ross process; see [4, 5]. The case where \( \sigma(t, x, y) = x\sigma_2(y) \) and \( Y \) is the Ornstein–Uhlenbeck process was studied in [12, 13].

In the present paper, we investigate the existence and uniqueness of weak and strong solutions to the equation (1). We adapt the approaches of Skorokhod [20], Stroock and Varadhan [21, 22], and Krylov [10, 11] to establish the weak existence and weak uniqueness. Concerning the strong existence and uniqueness, we use the well-known approaches of Yamada and Watanabe [23] (see also [2]) for inhomogeneous coefficients and Lipschitz conditions. In the present paper, we consider only the case of multiplicative stochastic volatility, where, as it was mentioned, the diffusion coefficient is factorized as \( \sigma(t, x, y) = \sigma_1(t, x)\sigma_2(t, y) \). Then we construct the maximum likelihood estimator for the unknown drift parameter and prove its strong consistency. As an example, we consider a linear model with stochastic volatility driven by a solution to some Itô’s SDE. In particular, we study in details an SDE with constant coefficients, the Ornstein–Uhlenbeck process, and the geometric Brownian motion, as the model for volatility (note that process \( Y \) can be interpreted not only as a volatility, but also as an additional source of randomness). Note that the maximum likelihood estimation in the Ornstein–Uhlenbeck model with stochastic volatility was studied in [1]. Similar statistical methods for the case of deterministic volatility can be found in [7, 9, 16, 18].

The paper is organized as follows. In Section 2, we prove the existence of weak and strong solutions under different conditions. In Section 3, we establish the strong consistency of the maximum likelihood estimator of the unknown drift parameter \( \theta \). Section 4 contains the illustrations of our results with some simulations. Auxiliary statements are gathered in Section 5.

2 Existence and uniqueness results for weak and strong solutions

Let \( (\Omega, \mathcal{F}, \mathcal{G}, \mathbb{P}) \) be a complete probability space with filtration \( \mathcal{G} = \{ \mathcal{G}_t, t \geq 0 \} \) satisfying the standard assumptions. We assume that all processes under consideration are adapted to the filtration \( \mathcal{G} \).
2.1 Existence of weak solution in terms of Skorokhod conditions
Consider the following stochastic differential equation:

\[ dX_t = a(t, X_t) \, dt + \sigma_1(t, X_t) \, dY_t, \quad t = 0, 1, \ldots, T, \]

where \( X_{t=0} = X_0 \in \mathbb{R}, \) \( W \) is a Wiener process, and \( Y \) is some adapted stochastic process to be specified later.

**Theorem 1.** Let \( Y \) be a measurable and continuous process, \( a, \sigma_1, \) and \( \sigma_2 \) be continuous w.r.t. \( x \in \mathbb{R}, \) \( y \in \mathbb{R}, \) and \( t \in [0, T], \) \( \sigma_2 \) be bounded, and

\[ |\sigma_1(t, x)|^2 + |a(t, x)|^2 \leq K (1 + |x|^2) \]

for some constant \( K > 0. \) Then Eq. (2) has a weak solution.

**Proof.** Consider a sequence of partitions of \( [0, T]; 0 = t^n_0 < t^n_1 < \ldots < t^n_n = T \) such that \( \lim_{n \to \infty} \max_k (t^n_{k+1} - t^n_k) = 0. \) Define \( \xi^n_k \) by \( \xi^n_0 = X(0) \) and

\[ \xi^n_{k+1} = \xi^n_k + a(t^n_k, \xi^n_k) \Delta t^n_k + \sigma_1(t^n_k, \xi^n_k) \sigma_2(t^n_k, Y(t^n_k)) \Delta W^n_k. \]

It follows from Lemma 1, Lemma 2, and Proposition 1 in Section 5 that it is possible to choose a subsequence \( n' \) and construct processes \( \tilde{\xi}^{n'}_{\cdot}, \tilde{W}^{n'}_{\cdot}, \) and \( \tilde{Y}^{n'}_{\cdot} \) such that the finite-dimensional distributions of \( \tilde{\xi}^{n'}_{\cdot}, \tilde{W}^{n'}_{\cdot}, \) and \( \tilde{Y}^{n'}_{\cdot} \) coincide with those of \( \xi^{n'}_{\cdot}, W_{\cdot}, \) and \( Y_{\cdot} \) and \( \tilde{\xi}^{n'}_{\cdot} \to \tilde{\xi}, \tilde{W}^{n'}_{\cdot} \to \tilde{W} \) and \( \tilde{Y}^{n'}_{\cdot} \to \tilde{Y} \) in probability, where \( \tilde{\xi}, \tilde{W}, \) and \( \tilde{Y} \) are some stochastic processes (evidently, \( \tilde{W} \) is a Wiener process). It suffices to prove that \( \tilde{\xi} \) is a solution of Eq. (2) when \( W \) and \( Y \) are replaced by \( \tilde{W} \) and \( \tilde{Y}. \)

We have that \( \tilde{\xi}^{n'}_{\cdot} \) satisfies the equation

\[ \tilde{\xi}^{n'}_{t} = \tilde{\xi}^{n'}_{0} + \sum_{t^n_k < t} a(t^n_k, \tilde{\xi}^{n'}_{t^n_k}) \Delta t^n_k + \sum_{t^n_k < t} \sigma_1(t^n_k, \tilde{\xi}^{n'}_{t^n_k}) \sigma_2(t^n_k, \tilde{Y}^{n'}_{t^n_k}) \Delta \tilde{W}^{n'}_k. \]

Since \( \sigma_2 \) is bounded and \( \sigma_1 \) is of linear growth, their product is of linear growth:

\[ |\sigma_1(t, x)\sigma_2(t, y)| \leq K_1 (1 + |x|), \]

where \( K_1 > 0 \) is a constant. Therefore,

\[ P\left( \sup_{0 \leq t \leq T} |\sigma_1(t, \tilde{\xi}^{n'}_{t})\sigma_2(t, \tilde{Y}^{n'}_{t})| > C \right) \]

\[ \leq P\left( \sup_{0 \leq t \leq T} K_1 (1 + |\tilde{\xi}^{n'}_{t}|) > C \right) = P\left( \sup_{0 \leq t \leq T} |\tilde{\xi}^{n'}_{t}| > \frac{C}{K_1} - 1 \right) \]

\[ = P\left( \sup_{0 \leq t \leq T} |\tilde{\xi}^{n'}_{t}| > \frac{C}{K_1} - 1 \right). \]

Using Lemma 1, we get that

\[ P\left( \sup_{0 \leq t \leq T} |\sigma_1(t, \tilde{\xi}^{n'}_{t})\sigma_2(t, \tilde{Y}^{n'}_{t})| > C \right) \to 0 \quad \text{as} \quad C \to \infty. \]
Moreover, we have that \( \sigma_1(t, x)\sigma_2(t, y) \) is continuous w.r.t. \( t \in [0, T] \), \( x, y \in \mathbb{R} \).

Then, for any \( \varepsilon > 0 \), there exists \( \delta > 0 \) such that
\[
|\sigma_1(t_1, x_1)\sigma_2(t_1, y_1) - \sigma_1(t_2, x_2)\sigma_2(t_2, y_2)| < \varepsilon
\]
whenever \( |t_1 - t_2| < \delta, |x_1 - x_2| < \delta, |y_1 - y_2| < \delta \).

Therefore,
\[
P(|\sigma_1(t_1, \tilde{\xi}_n(t_1))\sigma_2(t_1, \tilde{Y}_n(t_1)) - \sigma_1(t_2, \tilde{\xi}_n(t_2))\sigma_2(t_2, \tilde{Y}_n(t_2))| > \varepsilon)
\leq P(|\tilde{\xi}_n(t_1) - \tilde{\xi}_n(t_2)| < \delta, |\tilde{Y}_n(t_1) - \tilde{Y}_n(t_2)| < \delta,
|\sigma_1(t_1, \tilde{\xi}_n(t_1))\sigma_2(t_1, \tilde{Y}_n(t_1)) - \sigma_1(t_2, \tilde{\xi}_n(t_2))\sigma_2(t_2, \tilde{Y}_n(t_2))| > \varepsilon)
+ P(|\tilde{\xi}_n(t_1) - \tilde{\xi}_n(t_2)| \geq \delta) + P(|\tilde{Y}_n(t_1) - \tilde{Y}_n(t_2)| \geq \delta)
= P(|\tilde{\xi}_n(t_1) - \tilde{\xi}_n(t_2)| \geq \delta) + P(|Y(t_1) - Y(t_2)| \geq \delta),
\]
and the last relation implies the following one:
\[
\lim_{h \to 0} \lim_{n' \to \infty} \sup_{|t_1 - t_2| \leq h} P(|\sigma_1(t_1, \tilde{\xi}_n'(t_1))\sigma_2(t_1, \tilde{Y}_n(t_1)) - \sigma_1(t_2, \tilde{\xi}_n(t_2))\sigma_2(t_2, \tilde{Y}_n(t_2))| > \varepsilon) = 0.
\]

Applying Lemma 1, we get that
\[
\sum_{t'n'_{k+1} \leq t} \sigma_1(t'_n, \tilde{\xi}_n(t'_n))^\prime)\sigma_2(t'_n, \tilde{Y}_n(t'_n)) \Delta W^n_k \to \int_0^T \sigma_1(s, \tilde{\xi}(s))\sigma_2(s, \tilde{Y}(s)) \, d\tilde{W}(s)
\]
in probability as \( n' \to \infty \), and we also have that
\[
\sum_{t'n'_k \leq t} a(t'_n, \tilde{\xi}_n(t'_n)) \Delta t'_n \to \int_0^T a(s, \tilde{\xi}(s)) \, ds,
\]
whence the proof follows.

2.2 Existence and uniqueness of weak solution in terms of Stroock–Varadhan conditions

In this approach, we assume additionally that the process \( Y \) also is a solution of some diffusion stochastic differential equation. Let \( W^1 \) and \( W^2 \) be two Wiener processes, possibly correlated, so that \( dW^1_t \, dW^2_t = \rho \, dt \) for some \( |\rho| \leq 1 \). In this case, we can represent \( W^2_t = \rho W^1_t + \sqrt{1 - \rho^2} W^3_t \), where \( W^3 \) is a Wiener process independent of \( W^1 \).

**Theorem 2.** Consider the system of stochastic differential equations
\[
\begin{aligned}
\text{d}X_t &= a(t, X_t) \, dt + \sigma_1(t, X_t)\sigma_2(t, Y_t) \, dW^1_t, \\
\text{d}Y_t &= \alpha(t, Y_t) \, dt + \beta(t, Y_t) \, dW^2_t,
\end{aligned}
\tag{3}
\tag{4}
\]
where all coefficients \( a, \alpha, \sigma_1, \sigma_2, \) and \( \beta \) are nonrandom measurable and bounded functions, \( \sigma_1, \sigma_2, \) and \( \beta \) are continuous in all arguments. Let \( |\rho| < 1 \), and let \( \sigma_1(t, x) > 0, \sigma_2(t, y) > 0, \beta(t, y) > 0 \) for all \( t, x, y \). Then the weak existence and uniqueness in law hold for system (3)–(4), and in particular, the weak existence and uniqueness in law hold for Eq. (3) with \( Y \) being a weak solution of Eq. (4).
Theorem 3. Let Yamada–Watanabe conditions for inhomogeneous coefficients from [2]. Now we consider strong existence–uniqueness conditions for Eq. (2), adapting the 2.3 Existence of strong solution in terms of Yamada–Watanabe conditions in the left-hand side of (6) is positive definite since its discriminant

\[ D = \rho^2 (1 - \rho^2) \beta^4 - \beta^2 (1 - \rho^2) (\sigma_1^2 \sigma_2^2 + \beta^2 \rho^2) = -\beta^2 (1 - \rho^2) \sigma_1^2 \sigma_2^2 < 0. \]

The continuity of \( Q(\lambda_1, \lambda_2) \) implies the existence of \( \min_{k=1}^2 Q(\lambda_1, \lambda_2) > 0. \) Then, putting \( \varepsilon = \min_{k=1}^2 Q(\lambda_1, \lambda_2) \) and using homogeneity, we get (6). □

2.3 Existence of strong solution in terms of Yamada–Watanabe conditions

Now we consider strong existence–uniqueness conditions for Eq. (2), adapting the Yamada–Watanabe conditions for inhomogeneous coefficients from [2].

Theorem 3. Let \( a, \alpha_1, \) and \( \alpha_2 \) be nonrandom measurable bounded functions such that

(i) There exists a positive increasing function \( \rho(u), u \in (0, \infty), \) satisfying \( \rho(0) = 0 \) such that

\[ |\sigma_1(t, x) - \sigma_1(t, y)| \leq \rho(|x - y|), \quad t \geq 0, \ x, y \in \mathbb{R}, \]

and \( \int_0^\infty \rho^{-2}(u)du = +\infty. \)

(ii) There exists a positive increasing concave function \( k(u), u \in (0, \infty), \) satisfying \( k(0) = 0 \) such that
\[ |a(t, x) - a(t, y)| \leq k(|x - y|), \quad t \geq 0, \quad x, y \in \mathbb{R}, \]

and \( \int_0^\infty k^{-1}(u) du = +\infty \). Also, let \( Y \) be an adapted continuous stochastic process. Then the pathwise uniqueness of solution holds for Eq. (2), and hence it has a unique strong solution.

**Proof.** Let \( 1 > a_1 > a_2 > \cdots > a_n > \cdots > 0 \) be defined by

\[
\int_{a_1}^{a_2} \rho^{-2}(u) \, du = 1, \quad \int_{a_2}^{a_3} \rho^{-2}(u) \, du = 2, \ldots, \quad \int_{a_{n-1}}^{a_n} \rho^{-2}(u) \, du = n, \ldots.
\]

We have that \( a_n \to 0 \) as \( n \to \infty \). Let \( \phi_n(u), n = 1, 2, \ldots, \) be a continuous function with support contained in \( (a_n, a_{n-1}) \) such that \( 0 \leq \phi_n(u) \leq \frac{2 \rho^{-2}(u)}{n} \) and \( \int_{a_{n-1}}^{a_n} \phi_n(u) \, du = 1 \). Such a function obviously exists. Set

\[
\varphi_n(x) = \int_0^{|x|} \int_0^y \phi_n(u) \, du \, dy, \quad x \in \mathbb{R}.
\]

Clearly, \( \varphi_n \in C^2(\mathbb{R}) \), \( |\varphi_n'(x)| \leq 1 \), and \( \varphi_n(x) \nrightarrow |x| \) as \( n \to \infty \).

Let \( X_1 \) and \( X_2 \) be two solutions of Eq. (2) on the same probability space with the same Wiener process and such that \( X_1(0) = X_2(0) \). Then we can present their difference as

\[
X_1(t) - X_2(t) = \int_0^t \sigma_2(s, Y(s)) (\sigma_1(s, X_1(s)) - \sigma_1(s, X_2(s))) \, dW(s)
+ \int_0^t (a(s, X_1(s)) - a(s, X_2(s))) \, ds.
\]

By the Itô formula,\n
\[
\varphi_n(X_1(t) - X_2(t)) = \int_0^t \varphi_n'(X_1(s) - X_2(s)) \sigma_2(s, Y(s)) (\sigma_1(s, X_1(s)) - \sigma_1(s, X_2(s))) \, dW(s)
+ \int_0^t \varphi_n'(X_1(s) - X_2(s)) (a(s, X_1(s)) - a(s, X_2(s))) \, ds
+ \frac{1}{2} \int_0^t \varphi_n''(X_1(s) - X_2(s)) \sigma_2(s, Y(s))^2 (\sigma_1(s, X_1(s)) - \sigma_1(s, X_2(s)))^2 \, ds
= J_1 + J_2 + J_3.
\]

We have that \( \mathbb{E}(J_1) = 0 \),

\[
|\mathbb{E}(J_2)| \leq \int_0^t \mathbb{E} |a(s, X_1(s)) - a(s, X_2(s))| \, ds
\leq \int_0^t \mathbb{E}(k(|X_1(s) - X_2(s)|)) \, ds \leq \int_0^t k(\mathbb{E}(|X_1(s) - X_2(s)|)) \, ds
\]
by Jensen’s inequality, and
\[
|E(J_3)| \leq \frac{C^2}{2} \int_0^t E\left( \frac{2}{n} \rho^{-2}(|X_1(s) - X_2(s)|) \rho^2(|X_1(s) - X_2(s)|) \right) ds
\]
\[
\leq \frac{t}{n} \to 0 \quad \text{as } n \to \infty.
\]
So by letting \( n \to \infty \) we get
\[
E(|X_1(s) - X_2(s)|) \leq \int_0^t k(E(|X_1(s) - X_2(s)|) ds.
\]
We have that \( \int_0^\infty k^{-1}(u)du = +\infty \). Then we get \( E(|X_1(s) - X_2(s)|) = 0 \), and hence \( X_1(s) = X_2(s) \) a.s.

2.4 Existence and uniqueness for strong solution in terms of Lipschitz conditions

**Theorem 4.** Let \( a, \sigma_1, \) and \( \sigma_2 \) be nonrandom measurable functions, and let \( Y \) be an adapted continuous stochastic process. Consider the following assumptions:

(i) There exists \( K > 0 \) such that, for all \( t \geq 0 \) and \( x \in \mathbb{R} \),
\[
|\sigma_1(t, x)|^2 + |a(t, x)|^2 \leq K^2(1 + |x|^2);
\]
(ii) For any \( n \in \mathbb{N} \), there exists \( K_N > 0 \) such that, for all \( t \geq 0 \) and for all \( (x, y) \) satisfying \( |x| \leq N \) and \( |y| \leq N \),
\[
|a(t, x) - a(t, y)| + |\sigma_1(t, x) - \sigma_1(t, y)| \leq K_N |x - y|;
\]
(iii) \( \sup_{s \geq 0} \sup_{|x| \leq N} |\sigma_2(s, x)| \leq C_N \).

Then Eq. (2) has a unique strong solution.

This result can be proved by using the successive approximation method; see, for example, [19, Thm. 1.2].

3 Drift parameter estimation

3.1 General results
Let \( (\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P}) \) be a complete probability space with filtration \( \mathcal{F} = \{ \mathcal{F}_t, t \geq 0 \} \) satisfying the standard assumptions. We assume that all processes under consideration are adapted to the filtration \( \mathcal{F} \). Consider a parameterized version of Eq. (2)
\[
dX_t = \theta a(t, X_t) dt + \sigma_1(t, X_t)\sigma_2(t, Y_t) dW_t,
\] (7)
where \( W \) is a Wiener process. Assume that Eq. (7) has a unique strong solution \( X = \{ X_t, t \in [0, T] \} \). Our main problem is to estimate the unknown parameter \( \theta \) by continuous observations of \( X \) and \( Y \).
Denote
\[ f(t, x, y) = \frac{a(t, x)}{\sigma_1^2(t, x)\sigma_2^2(t, y)}, \quad g(t, x, y) = \frac{a(t, x)}{\sigma_1(t, x)\sigma_2(t, y)}. \]

Assume that, for all \( t > 0 \),
\begin{align*}
\sigma_1(t, X_t)\sigma_2(t, Y_t) &\neq 0 \quad \text{a.s.,} \quad (8) \\
\int_0^t g^2(s, X_s, Y_s) \, ds &< \infty \quad \text{a.s.,} \quad (9) \\
\int_0^\infty g^2(s, X_s, Y_s) \, ds &= \infty \quad \text{a.s.} \quad (10)
\end{align*}

Then a likelihood function for Eq. (7) has the form
\[ \frac{dP_{\theta}(T)}{dP_0(T)} = \exp \left\{ \theta \int_0^T f(t, X_t, Y_t) \, dX_t - \frac{\theta^2}{2} \int_0^T g^2(t, X_t, Y_t) \, dt \right\}; \]
see [15, Ch. 7]. Hence, the maximum likelihood estimator of parameter \( \theta \) constructed by observations of \( X \) and \( Y \) on the interval \([0, T]\) has the form
\[ \hat{\theta}_T = \frac{\int_0^T f(t, X_t, Y_t) \, dX_t}{\int_0^T g^2(t, X_t, Y_t) \, dt} = \theta + \frac{\int_0^T g(t, X_t, Y_t) \, dW_t}{\int_0^T g^2(t, X_t, Y_t) \, dt}. \quad (11)
\]

**Theorem 5.** Under assumptions (8)–(10), the estimator \( \hat{\theta}_T \) is strongly consistent as \( T \to \infty \).

**Proof.** Note that, under condition (9) the process \( M_t = \int_0^t g(s, X_s, Y_s) \, dW_s \) is a square-integrable local martingale with quadratic variation \( (M)_t = \int_0^t g^2(s, X_s, Y_s) \, ds \). According to the strong law of large numbers for martingales [14, Ch. 2, § 6, Thm. 10, Cor. 1], under the condition \( (M)_T \to \infty \) a.s. as \( T \to \infty \), we have that \( \frac{M_T}{(M)_T} \to 0 \) a.s. as \( T \to \infty \). Therefore, it follows from representation (11) that \( \hat{\theta}_T \) is strongly consistent.

3.2 Linear equation with stochastic volatility

As an example, let us consider the model
\[ dX_t = \theta X_t \, dt + X_t \sigma_2(Y_t) \, dW_t, \quad X_0 = x_0 \in \mathbb{R}, \quad (12)\]
where \( W_t \) is a Wiener process, and \( Y_t \) is a continuous stochastic process with values from an open interval \( J = (l, r) \) (further, in examples, we will consider \( J = \mathbb{R} \) or \( J = (0, +\infty) \)). By Theorem 4, under the assumption

(A1) \( \sigma_2(y) \) is locally bounded on \( J \),

there exists a unique strong solution of (12).
Let $Y$ be a $J$-valued solution of the equation
\begin{equation}
\frac{dY_t}{Y_t} = \alpha(Y_t) \, dt + \beta(Y_t) \, dW^1_t, \quad Y_0 = y_0 \in J,
\end{equation}
where $W^1$ is a Wiener process, possibly correlated with $W$.

By $L_{\text{loc}}^1(J)$ we denote the set of Borel functions $f: J \to [-\infty, \infty]$ that are locally integrable on $J$, that is, integrable on compact subsets of $J$. By $L_{\text{loc}}^1(l+)$ we denote the set of Borel functions $f: J \to [-\infty, \infty]$ such that $\int_l^z |f(y)| \, dy < \infty$ for some $z \in J$. The notation $L_{\text{loc}}^1(r-)$ is introduced similarly.

Assume that coefficients $\alpha$ and $\beta$ satisfy the Engelbert–Schmidt conditions
\begin{enumerate}[(A2)]
\item $\beta(y) \neq 0$ for all $y \in J$, and
\item $\beta^{-2}, \alpha \beta^{-2} \in L_{\text{loc}}^1(J)$.
\end{enumerate}

Let us introduce the following notation:
\begin{align*}
\rho(y) &= \exp \left\{ -2 \int_c^y \frac{\alpha(u)}{\beta^2(u)} \, du \right\}, \quad y \in J, \\
s(y) &= \int_c^y \rho(u) \, du, \quad y \in \bar{J} = [l, r],
\end{align*}
for some $c \in J$. Assume additionally that
\begin{enumerate}[(A4)]
\item $s(r) = \infty$ or $\frac{s(r)-s}{\rho \beta^2} \notin L_{\text{loc}}^1(r-)$,
\item $s(l) = -\infty$ or $\frac{s-l}{\rho \beta^2} \notin L_{\text{loc}}^1(l+)$.
\end{enumerate}

Under (A2)–(A3), the SDE (13) has a weak solution, unique in law, which possibly exits $J$ at some time $\zeta$. Moreover, $\zeta = \infty$ a.s. if and only if conditions (A4)–(A5) are satisfied, see, for example, [17, Prop. 2.6].

Assume also that
\begin{enumerate}[(A6)]
\item $\beta^{-2} \sigma_2^{-2} \in L_{\text{loc}}^1(J)$,
\item one of the following four conditions holds:
\begin{enumerate}[(i)]
\item $s(r) = \infty, s(l) = -\infty$,
\item $s(r) < \infty, s(l) = -\infty, \frac{s(r)-s}{\rho \beta^2} \notin L_{\text{loc}}^1(r-),$
\item $s(r) = \infty, s(l) > -\infty, \frac{s-l}{\rho \beta^2} \notin L_{\text{loc}}^1(l+),$
\item $s(r) < \infty, s(l) > -\infty, \frac{s-r}{\rho \beta^2} \notin L_{\text{loc}}^1(r-), \frac{s-l}{\rho \beta^2} \notin L_{\text{loc}}^1(l+),$
\end{enumerate}
\item $X_t \sigma(Y_t) \neq 0$ a.s., $t \geq 0$.
\end{enumerate}

The maximum likelihood estimator (11) for model (12) equals
\begin{equation}
\hat{\theta}_T = \frac{\int_0^T X_t^{-1} \sigma_2^{-2}(Y_t) \, dX_t}{\int_0^T \sigma_2^{-2}(Y_t) \, dt}.
\end{equation}
Theorem 6. Under assumptions (A1)–(A8), the estimator $\hat{\theta}_T$ is strongly consistent as $T \to \infty$.

Proof. We need to verify conditions (8)–(10) of Theorem 5. For model (12), they read as follows:

$X_t\sigma_2(Y_t) \neq 0, \ t \geq 0, \ a.s., \quad (15)$

$$\int_0^t \sigma_2^{-2}(Y_s) \, ds < \infty, \ t > 0, \ a.s., \quad (16)$$

$$\int_0^\infty \sigma_2^{-2}(Y_s) \, ds = \infty \ a.s. \quad (17)$$

Note that (15) is assumption (A8). By [17, Thm. 2.7] the local integrability condition (A6), together with (A2)–(A5), implies (16). Further, if assumption (A7)(i) holds, then (17) is satisfied by [17, Thm. 2.11]. In the remaining case $s(r) < \infty$ or $s(l) > -\infty$, we have that

$$\Omega = \left\{ \lim_{t \uparrow \infty} Y_t = r \right\} \cup \left\{ \lim_{t \uparrow \infty} Y_t = l \right\};$$

see [17]. Moreover, if $s(r) = \infty$, then $\mathbb{P}(\lim_{t \uparrow \infty} Y_t = r) = 0$ by [17, Prop. 2.4]. If $s(r) < \infty$ and $\frac{s(r) - s}{\rho^2 \sigma_2} \notin L^1_{loc}(r-)$, then $\int_0^\infty \sigma_2^{-2}(Y_s) \, ds = \infty$ a.s. on $\{\lim_{t \uparrow \infty} Y_t = r\}$ by [17, Thm. 2.12]. The similar statements hold for $\{\lim_{t \uparrow \infty} Y_t = l\}$. This implies that (17) is satisfied under each of conditions (ii)–(iv) of assumption (A7).

Now we consider several examples of the process $Y$, namely the Bachelier model, the Ornstein–Uhlenbeck model, the geometric Brownian motion, and the Cox–Ingersoll–Ross model. We concentrate on verification of assumption (A7) for these models, assuming that other conditions of Theorem 6 are satisfied.

Example 1 (Bachelier model). Let $Y$ be a solution of the SDE

$$dY_t = \alpha \, dt + \beta \, dW_t, \quad Y_0 = y_0 \in \mathbb{R},$$

where $\alpha \in \mathbb{R}$ and $\beta \neq 0$ are some constants. Assume that $\sigma_2^{-2}(y) \in L^1_{loc}(\mathbb{R})$ and one of the following assumptions holds:

(i) $\alpha = 0$,

(ii) $\alpha > 0$ and $\sigma_2^{-2}(y) \notin L^1_{loc}(+\infty)$,

(iii) $\alpha < 0$ and $\sigma_2^{-2}(y) \notin L^1_{loc}(-\infty)$.

Then estimator (14) is strongly consistent.

Indeed, in this case, $J = \mathbb{R}$,

$$\rho(y) = \exp\left\{-\frac{2\alpha}{\beta^2} y\right\}, \quad \text{and} \quad s(y) = \int_0^y \exp\left\{-\frac{2\alpha}{\beta^2} u\right\} \, du.$$
\[ s(y) = \frac{\beta^2}{2\alpha} \left( 1 - \exp \left\{ -\frac{2\alpha}{\beta^2} y \right\} \right). \]

If \( \alpha > 0 \), then \( s(+\infty) = \frac{\beta^2}{2\alpha} \), \( s(-\infty) = -\infty \), and

\[ \frac{s(+\infty) - s(y)}{\rho(y)\beta^2\sigma^2_2(y)} = \frac{1}{2\alpha\sigma^2_2(y)} \notin L^1_\text{loc}(+\infty), \]

and hence (A7)(ii) holds. The case \( \alpha < 0 \) is considered similarly.

**Example 2** (Ornstein–Uhlenbeck or Vasicek model). *Let \( Y \) be a solution of the SDE*

\[ dY_t = a(b - Y_t) \, dt + \gamma \, dW_t, \quad Y_0 = y_0 \in \mathbb{R}, \]

*where \( a, b \in \mathbb{R} \), and \( \gamma > 0 \) are some constants. Assume that \( \sigma^{-2}_2 \in L^1_\text{loc}(\mathbb{R}) \) and one of the following assumptions holds:

(i) \( a \geq 0 \),

(ii) \( a < 0 \), \( y^{-1}\sigma^{-2}_2(y) \notin L^1_\text{loc}(+\infty) \cup L^1_\text{loc}(-\infty) \).

*Then estimator (14) is strongly consistent.*

In this case, we also take \( J = \mathbb{R} \). Then

\[ \rho(y) = \exp \left\{ -2 \int_b^y \frac{a(b - u)}{\gamma^2} \, du \right\} = \exp \left\{ \frac{a}{\gamma^2} (y - b)^2 \right\}, \]

\[ s(y) = \int_b^y \exp \left\{ \frac{a}{\gamma^2} (u - b)^2 \right\} \, du. \]

If \( a \geq 0 \), then \( \exp\left\{ \frac{a}{\gamma^2} (u - b)^2 \right\} \geq 1 \), and we get that \( s(+\infty) = \infty \), \( s(-\infty) = -\infty \).

If \( a < 0 \), then

\[ s(y) = \frac{\gamma}{\sqrt{-a}} \int_0^{\frac{\sqrt{-a}}{\gamma}(y-b)} e^{-z^2} \, dz. \]

Therefore, \( s(+\infty) = -s(-\infty) = \frac{\gamma}{\sqrt{2\pi}} \sqrt{-a} < \infty \), and we need to verify (A7)(iv). Since

\[ \int_x^{\infty} e^{-z^2} \, dz \sim \frac{1}{2x} e^{-x^2} \text{ as } x \to \infty, \]

we see that

\[ \frac{s(+\infty) - s(y)}{\rho(y)\gamma^2\sigma^2_2(y)} = \frac{\gamma}{\sqrt{-a}} \int \frac{e^{-z^2} \, dz}{\sqrt{\frac{\sqrt{-a}}{\gamma}(y-b)}} \sim -2a(y-b)\sigma^2_2(y) \]

as \( y \to \infty \). Then \( \frac{s(+\infty) - s(y)}{\rho(y)\gamma^2\sigma^2_2(y)} \notin L^1_\text{loc}(+\infty) \) if \( y^{-1}\sigma^{-2}_2(y) \notin L^1_\text{loc}(+\infty) \). The condition \( \frac{s-s(-\infty)}{\rho\beta^2\sigma^2_2} \notin L^1_\text{loc}(-\infty) \) is considered similarly.

**Example 3** (Geometric Brownian motion). *Let \( Y \) be a solution of the SDE*

\[ dY_t = aY_t \, dt + \beta Y_t \, dW_t, \quad Y_0 = y_0 > 0, \]

*where \( a, b \in \mathbb{R} \), and \( \gamma > 0 \) are some constants.*
where $\alpha \in \mathbb{R}$ and $\beta \neq 0$ are some constants. Assume that $y^{-2} \sigma_2^{-2}(y) \in L^1_{\text{loc}}((0, +\infty))$ and one of the following assumptions holds:

(i) $\beta^2 = 2\alpha^2$,

(ii) $\beta^2 < 2\alpha^2$ and $y^{-1}\sigma_2^{-2}(y) \notin L^1_{\text{loc}}(+\infty)$,

(iii) $\beta^2 > 2\alpha^2$ and $y^{-1}\sigma_2^{-2}(y) \notin L^1_{\text{loc}}(0+)$.

Then estimator (14) is strongly consistent.

In this case, the process $Y$ is positive, and hence $J = (0, \infty)$. We have

$$\rho(y) = \exp \left\{ -2 \int_1^y \frac{\alpha^2}{\beta^2 u} \, du \right\} = y^{-\frac{2\alpha^2}{\beta^2}}.$$

$$s(y) = \int_1^y u \frac{2\alpha^2}{\beta^2} \, du = \begin{cases} \frac{1}{2} - \frac{2\alpha^2}{\beta^2} \frac{1}{1 - \frac{2\alpha^2}{\beta^2}}, & \beta^2 \neq 2\alpha^2, \\
\ln y, & \beta^2 = 2\alpha^2. \end{cases}$$

If $\beta^2 = 2\alpha^2$, then $s(0) = -\infty$ and $s(+\infty) = \infty$. If $\beta^2 < 2\alpha^2$, then $s(0) = -\infty$, $s(+\infty) < \infty$, and

$$\frac{s(+\infty) - s(y)}{\rho(y)\beta^2 y^2 \sigma_2^2(y)} \neq L^1_{\text{loc}}(+\infty).$$

If $\beta^2 > 2\alpha^2$, then $s(0) > -\infty$, $s(+\infty) = \infty$, and

$$\frac{s(y) - s(0)}{\rho(y)\beta^2 y^2 \sigma_2^2(y)} \neq L^1_{\text{loc}}(0+).$$

**Example 4** (Cox–Ingersoll–Ross model). Let $Y$ be a solution of the SDE

$$dY_t = a(b - Y_t) \, dt + \gamma \sqrt{Y_t} \, dW_t^1, \quad Y_0 = y_0 \in \mathbb{R},$$

where $a$, $b$, $\gamma$ are positive constants, and $2ab \geq \gamma^2$. Assume that

$$y^{-1}\sigma_2^{-2}(y) \in L^1_{\text{loc}}((0, +\infty)).$$

Then estimator (14) is strongly consistent.

Under the condition $2ab \geq \gamma^2$, the process $Y$ is positive, and hence $J = (0, \infty)$. Further,

$$\rho(y) = \exp \left\{ -2 \int_1^y \frac{a(b - u)}{\gamma^2 u} \, du \right\} = y^{-\frac{2ab}{\gamma^2}} e^{\frac{2\alpha}{\gamma} (y-1)},$$

$$s(y) = e^{-\frac{2\alpha}{\gamma^2}} \int_1^y u \frac{2ab}{\gamma^2} e^{\frac{2\alpha}{\gamma} u} \, du.$$
Table 1. The means and standard deviations of $\hat{\theta}_T$

<table>
<thead>
<tr>
<th>$\alpha(y)$</th>
<th>$\beta(y)$</th>
<th>$\sigma_2(y)$</th>
<th>$T$</th>
<th>10</th>
<th>50</th>
<th>100</th>
<th>200</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>$</td>
<td>y</td>
<td>^{1/4}$</td>
<td></td>
<td>Mean</td>
<td>1.9455</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Std.dev.</td>
<td>0.4260</td>
<td>0.2576</td>
<td>0.2367</td>
</tr>
<tr>
<td>$y$</td>
<td>2$y$</td>
<td>$\sqrt{y}$</td>
<td></td>
<td>Mean</td>
<td>2.0104</td>
<td>2.0000</td>
<td>2.0000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Std.dev.</td>
<td>0.1225</td>
<td>5.7 \cdot 10^{-5}</td>
<td>4.7 \cdot 10^{-8}</td>
</tr>
<tr>
<td>$y$</td>
<td>$y$</td>
<td>$(1 + y)^{-1}$</td>
<td></td>
<td>Mean</td>
<td>2.0008</td>
<td>2.0001</td>
<td>2.0000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Std.dev.</td>
<td>0.0769</td>
<td>0.0010</td>
<td>2.2 \cdot 10^{-12}</td>
</tr>
<tr>
<td>$y$</td>
<td>1</td>
<td>$2 + \sin y$</td>
<td></td>
<td>Mean</td>
<td>1.9358</td>
<td>1.9819</td>
<td>1.9927</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Std.dev.</td>
<td>0.5436</td>
<td>0.2437</td>
<td>0.1679</td>
</tr>
<tr>
<td>$-y$</td>
<td>1</td>
<td>$2 + \sin y$</td>
<td></td>
<td>Mean</td>
<td>1.9061</td>
<td>1.9684</td>
<td>1.9700</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Std.dev.</td>
<td>0.5994</td>
<td>0.2472</td>
<td>0.1781</td>
</tr>
<tr>
<td>$2 - y$</td>
<td>$\sqrt{y}$</td>
<td>$\sqrt{y}$</td>
<td></td>
<td>Mean</td>
<td>1.9923</td>
<td>2.0039</td>
<td>1.9796</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Std.dev.</td>
<td>0.3540</td>
<td>0.1604</td>
<td>0.1173</td>
</tr>
<tr>
<td>$2 - y$</td>
<td>$\sqrt{y}$</td>
<td>$y$</td>
<td></td>
<td>Mean</td>
<td>2.0830</td>
<td>1.9835</td>
<td>1.9803</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Std.dev.</td>
<td>0.4347</td>
<td>0.1974</td>
<td>0.1205</td>
</tr>
</tbody>
</table>

Since $u = \frac{2ab}{\gamma^2} e^{\frac{2a}{\gamma^2} u} \to \infty$ as $u \to \infty$, we see that $s(+\infty) = \infty$. Moreover, using the inequality $e^{\frac{2a}{\gamma^2} u} \geq 1$, we get

$$s(0) = -e^{-\frac{2a}{\gamma^2}} \int_0^1 u^{\frac{2a}{\gamma^2}} e^{\frac{2a}{\gamma^2} u} du \leq -e^{-\frac{2a}{\gamma^2}} \int_0^1 u^{-\frac{2a}{\gamma^2}} du = -\infty$$

since $\frac{2ab}{\gamma^2} > 1$. Thus, assumption (A7)(i) is satisfied.

4 Simulations

We illustrate the quality of the estimator $\hat{\theta}_T$ in model (12)--(13) by simulation experiments. We simulate the trajectories of the Wiener processes $W$ and $W^1$ at the points $t = 0, h, 2h, 3h, \ldots$ and compute the approximate values of the process $Y$ and $X$ as solutions to SDEs using Euler's approximations. For each set of parameters, we simulate 100 sample paths with step $h = 0.0001$. The initial values of the processes are $x_0 = y_0 = 1$, and the true value of the parameter is $\theta = 2$. The results are reported in Table 1.

5 Appendix

The next two propositions are taken from [20].

**Proposition 1.** Assume that we have $r$ sequences of stochastic processes $\xi_n^{(1)}, \ldots, \xi_n^{(r)}$ such that, for all $i = 1, \ldots, r$,

(i) for every $\delta > 0$, $\lim_{h \to 0} \lim_{n \to \infty} \sup_{|t_1 - t_2| \leq h} \mathbb{P}(|\xi_n^{(i)}(t_1) - \xi_n^{(i)}(t_2)| > \delta) = 0$,
(ii) \( \lim_{C \to \infty} \lim_{n \to \infty} \sup_{0 \leq t \leq T} P(|\xi_n(t)| > C) = 0. \)

Then, for some sequence \( n_k \) we can construct processes \( X_{n_k}^{(1)}, \ldots, X_{n_k}^{(r)} \) on the probability space \( (\Omega', \mathcal{F}', P') \), where \( \Omega' = [0, 1] \), \( \mathcal{F}' = \mathcal{B}([0, 1]) \), and \( P' \) is the Lebesgue measure, such that the finite-dimensional distributions of \( X_{n_k}^{(1)}, \ldots, X_{n_k}^{(r)} \) coincide with those of \( \xi_{n_k}^{(1)}, \ldots, \xi_{n_k}^{(r)} \) and each of the sequences \( X_{n_k}^{(1)}, \ldots, X_{n_k}^{(r)} \) converges in probability to some limit.

**Proposition 2.** Let \( \eta_n(t) \) be a sequence of martingales such that \( \eta_n(t) \to W(t) \) in probability for all \( t \) and \( \mathbb{E}\eta_n(t)^2 \to t \) as \( n \to \infty \). Let \( f_n(t) \) be a sequence such that \( \int_0^T f_n(t) \, d\eta_n(t) \) exists for all \( n \), \( f_n(t) \to f(t) \) in probability for all \( t \), and \( \int_0^T f(t) \, dW(t) \) exists. Suppose that, in addition, the following conditions hold:

(i) for all \( \varepsilon > 0 \), there exists \( C > 0 \) such that, for all \( n \),

\[
P\left( \sup_{0 \leq t \leq T} |f_n(t)| > C \right) \leq \varepsilon,
\]

(ii) for all \( \varepsilon > 0 \),

\[
\lim_{h \to 0} \lim_{n \to \infty} \sup_{|t_1 - t_2| \leq h} P\left( |f_n(t_2) - f_n(t_1)| > \varepsilon \right) = 0.
\]

Then \( \int_0^T f_n(t) \, d\eta_n(t) \to \int_0^T f(t) \, dW(t) \) in probability.

In the next two lemmas, we modify the corresponding auxiliary results from [20] to Eq. (1) with multiplicative diffusion. Consider a sequence of partitions \( 0 = t^n_0 < t^n_1 < \cdots < t^n_n = T \) of \( [0, T] \) such that \( \lim_{n \to \infty} \max_k (t^n_{k+1} - t^n_k) = 0 \). Define \( \xi^n_k \) by \( \xi^n_0 = X(0) \) and

\[
\xi^n_{k+1} = \xi^n_k + a(t^n_k, \xi^n_k) \Delta t^n_k + \sigma_1(t^n_k, \xi^n_k) \sigma_2(t^n_k, Y(t^n_k)) \Delta W^n_k.
\]

**Lemma 1.** The random variables \( \sup_k |\xi^n_k| \) are bounded in probability uniformly w.r.t. \( n \).

**Proof.** Let \( \eta^n_0 = \xi^n_0 \mathbb{1}_{|\xi^n_0| \leq N} \) and

\[
\eta^n_{k+1} = \eta^n_k + a^N(t^n_k, \eta^n_k) \Delta t^n_k + \sigma_1^N(t^n_k, \eta^n_k) \sigma_2(t^n_k, Y(t^n_k)) \Delta W^n_k,
\]

where \( a^N(t, x) = a(t, x) \mathbb{1}_{|x| \leq N} \), \( \sigma_1^N(t, x) = \sigma_1(t, x) \mathbb{1}_{|x| \leq N} \). If \( |\eta^n_k| > N \), then \( \eta^n_{k+1} = \eta^n_k \), and if \( |\eta^n_k| \leq N \), then

\[
|\eta^n_{k+1}| \leq N + |a^N(t^n_k, \eta^n_k)| \Delta t^n_k + |\sigma_1^N(t^n_k, \eta^n_k) \sigma_2(t^n_k, Y(t^n_k))| \Delta W^n_k.
\]

Then, for any \( 1 \leq k \leq n \),

\[
|\eta^n_k| \leq N + \sum_{r=0}^{k-1} |a^N(t^n_r, \eta^n_r)| \Delta t^n_r + \sum_{r=0}^{k-1} |\sigma_1^N(t^n_r, \eta^n_r) \sigma_2(t^n_r, Y(t^n_r))| \Delta W^n_r.
\]
and is square-integrable. Furthermore,

\[\mathbb{E}|\eta_{k+1}^N|^2 = \mathbb{E}|\eta_k^N|^2 + 2\mathbb{E}(\alpha^N(t_k^n, \eta_k^n)\eta_k^n)\Delta t_k^n + \mathbb{E}\alpha^N(t_k^n, \eta_k^n)^2(\Delta t_k^n)^2\]

\[\leq \mathbb{E}|\eta_k^N|^2 + 2\mathbb{E}(\alpha^N(t_k^n, \eta_k^n)\eta_k^n)\Delta t_k^n + \mathbb{E}\alpha^N(t_k^n, \eta_k^n)^2(\Delta t_k^n)^2\]

\[+ C^2\mathbb{E}((\sigma_1^N(t_k^n, \eta_k^n))^2(\Delta t_k^n)^2).\]

Then there exists a constant \(H = H(T, K)\) such that

\[\mathbb{E}|\eta_{k+1}^n|^2 \leq \mathbb{E}|\eta_k^n|^2 + H\Delta t_k^n + H\Delta t_k^n,\]

\[\mathbb{E}|\eta_{k+1}^n|^2 + 1 \leq (\mathbb{E}|\eta_k|^2 + 1)e^{HT} \leq (\mathbb{E}|\eta_0|^2 + 1)e^{HT},\]

\[\mathbb{E}|\eta_{k+1}^n|^2 \leq (\mathbb{E}|\eta_0|^2 + 1)e^{HT} - 1.\]

We have that

\[\sup_k|\eta_k^n| \leq |\eta_0^n| + \sum_{j=0}^{n-1}|a^N(t_j^n, \eta_j^n)|\Delta t_j^n + \sup_k\sum_{j=0}^{k-1}\sigma_1^N(t_j^n, \eta_j^n)\sigma_2(t_j^n, Y(t_j^n))\Delta W_j^n.\]

So, since

\[\mathbb{E}\sup_{0 \leq r \leq n}\left|\sum_{k=0}^r \sigma_1^N(t_k^n, \eta_k^n)\sigma_2(t_k^n, Y(t_k^n))\Delta W_k^n\right|^2 \leq 4\sum_{k=0}^n|\sigma_1^N(t_k^n, \eta_k^n)\sigma_2(t_k^n, Y(t_k^n))|^2\Delta t_k^n \leq 4K\sum_{k=0}^n(\mathbb{E}|\eta_k^n|^2 + 1)\Delta t_k^n\]

and

\[\mathbb{E}\left(\sum_{k=0}^{n-1}|a^N(t_k^n, \eta_k^n)|\Delta t_k^n\right)^2 \leq KT\sum_{k=0}^{n-1}(\mathbb{E}|\eta_k^n|^2 + 1)\Delta t_k^n,\]

we get

\[\mathbb{E}\sup_k|\eta_k^n|^2 \leq A + B|\eta_0^n|^2.\]

We have that \(\eta_0^n\) is bounded uniformly w.r.t. \(n\) and \(N\). Then \(\sup_k|\eta_k^n|^2\) is bounded in \(L_2\) uniformly w.r.t. \(n, N\).

For \(\sup_k|\eta_k^n| < N\), we have \(\sup_k|\eta_k^n| = \sup_k|\xi_k^n|\). Hence, \(\sup_k|\xi_k^n|\) is bounded in probability uniformly w.r.t. \(n\). \(\Box\)

**Remark 1.** Using Lemma 1, we have that, for all \(\varepsilon > 0\), there exists \(N > 0\) such that, for every \(n \geq N\), \(\mathbb{P}(\sup_k|\xi_k^n| > 0) < \varepsilon\).

**Lemma 2.** Let \(\xi_n(t) = \xi_k^n\) for \(t \in [t_k^n, t_{k+1}^n]\). Then, for all \(\delta > 0\),

\[\lim_{h \to 0} \lim_{n \to \infty} \sup_{|t_1 - t_2| \leq h} \mathbb{P}(|\xi_n(t_1) - \xi_n(t_2)| > \delta) = 0.\]
Proof. Let $\eta_n(t) = \eta^n_k$, $t \in [t^n_k, t^n_{k+1})$. Then

$$
\sup_{|t_1 - t_2| \leq h} \mathbb{P}(\|\xi_n(t_1) - \xi_n(t_2)\| > \delta) \leq \sup_{|t_1 - t_2| \leq h} \mathbb{P}(\|\eta_n(t_1) - \eta_n(t_2)\| > \delta) + \mathbb{P}\left(\sup_k |\xi^n_k - \eta^n_k| > 0\right).
$$

From (18) and the boundedness of $a^N$, $\sigma^N_1$, and $\sigma_2$ we have that

$$
\lim_{h \to 0} \lim_{n \to \infty} \sup_{|t_1 - t_2| \leq h} \mathbb{P}(\|\eta_n(t_1) - \eta_n(t_2)\| > \delta) = 0.
$$

Therefore,

$$
\lim_{h \to 0} \lim_{n \to \infty} \sup_{|t_1 - t_2| \leq h} \mathbb{P}(\|\xi_n(t_1) - \xi_n(t_2)\| > \delta) \leq \lim_{n \to \infty} \mathbb{P}\left(\sup_k |\xi^n_k - \eta^n_k| > 0\right).
$$

The proof follows now from Remark 1.

Acknowledgment

The authors are grateful to the anonymous referee for his useful remarks and suggestions, which contributed to a substantial improvement of the text.

References


