Drift parameter estimation in stochastic differential equation with multiplicative stochastic volatility

Meriem Bel Hadj Khlifa[a], Yuliya Mishura[b], Kostiantyn Ralchenko[b,*],

Mounir Zili[a]

[a]*University of Monastir, Faculty of Sciences of Monastir, Department of Mathematics, Avenue de l'Environnement, 5000, Monastir, Tunisia*
[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)

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, σ_1, and σ_2. Also, we study the strong consistency of the maximum likelihood estimator for the unknown parameter θ. 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

---

*Corresponding author.

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 θ 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 θ. 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 $(\Omega, \mathcal{F}, \overline{\mathcal{F}}, \mathbb{P})$ be a complete probability space with filtration $\overline{\mathcal{F}} = \{\mathcal{F}_t, t \geq 0\}$ satisfying the standard assumptions. We assume that all processes under consideration are adapted to the filtration $\overline{\mathcal{F}}$.

2.1 Existence of weak solution in terms of Skorokhod conditions

Consider the following stochastic differential equation:

$$dX_t = a(t, X_t)\, dt + \sigma_1(t, X_t)\sigma_2(t, Y_t)\, dW_t, \tag{2}$$

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, σ_1, and σ_2 be continuous w.r.t. $x \in \mathbb{R}$, $y \in \mathbb{R}$, and $t \in [0, T]$, σ_2 be bounded, and*

$$\left|\sigma_1(t, x)\right|^2 + \left|a(t, x)\right|^2 \le K\left(1 + |x|^2\right)$$

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

Proof. Consider a sequence of partitions of $[0, T]$: $0 = t_0^n < t_1^n < \cdots < t_n^n = T$ such that $\lim_{n\to\infty} \max_k (t_{k+1}^n - t_k^n) = 0$. Define ξ_k^n by $\xi_0^n = X(0)$ and

$$\xi_{k+1}^n = \xi_k^n + a\left(t_k^n, \xi_k^n\right)\Delta t_k^n + \sigma_1\left(t_k^n, \xi_k^n\right)\sigma_2\left(t_k^n, Y\left(t_k^n\right)\right)\Delta W_k^n.$$

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 $\widetilde{\xi}_{n'}$, $\widetilde{W}_{n'}$, and $\widetilde{Y}_{n'}$ such that the finite-dimensional distributions of $\widetilde{\xi}_{n'}$, $\widetilde{W}_{n'}$, and $\widetilde{Y}_{n'}$ coincide with those of $\xi^{n'}$, W, and Y and $\widetilde{\xi}_{n'} \to \widetilde{\xi}$, $\widetilde{W}_{n'} \to \widetilde{W}$ and $\widetilde{Y}_{n'} \to \widetilde{Y}$ in probability, where $\widetilde{\xi}$, $\widetilde{W}$, and $\widetilde{Y}$ are some stochastic processes (evidently, $\widetilde{W}$ is a Wiener process). It suffices to prove that $\widetilde{\xi}$ is a solution of Eq. (2) when W and Y are replaced by $\widetilde{W}$ and $\widetilde{Y}$.

We have that $\widetilde{\xi}_{n'}$ satisfies the equation

$$\widetilde{\xi}_{n'}(t) = \widetilde{\xi}_{n'}(0) + \sum_{t_{k+1}^{n'} \le t} a\left(t_k^{n'}, \widetilde{\xi}_{n'}\left(t_k^{n'}\right)\right)\Delta t_k^{n'}$$

$$+ \sum_{t_{k+1}^{n'} \le t} \sigma_1\left(t_k^{n'}, \widetilde{\xi}_{n'}\left(t_k^{n'}\right)\right)\sigma_2\left(t_k^{n'}, \widetilde{Y}_{n'}\left(t_k^{n'}\right)\right)\Delta \widetilde{W}_k^{n'}.$$

Since σ_2 is bounded and σ_1 is of linear growth, their product is of linear growth:

$$\left|\sigma_1(t, x)\sigma_2(t, y)\right| \le K_1(1 + |x|),$$

where $K_1 > 0$ is a constant. Therefore,

$$\mathbb{P}\left(\sup_{0\le t\le T} \left|\sigma_1\left(t, \widetilde{\xi}_{n'}(t)\right)\sigma_2\left(t, \widetilde{Y}_{n'}(t)\right)\right| > C\right)$$

$$\le \mathbb{P}\left(\sup_{0\le t\le T} K_1\left(1 + \left|\widetilde{\xi}_{n'}(t)\right|\right) > C\right) = \mathbb{P}\left(\sup_{0\le t\le T} \left|\widetilde{\xi}_{n'}(t)\right| > \frac{C}{K_1} - 1\right)$$

$$= \mathbb{P}\left(\sup_{0\le t\le T} \left|\xi^{n'}(t)\right| > \frac{C}{K_1} - 1\right).$$

Using Lemma 1, we get that

$$\mathbb{P}\left(\sup_{0\le t\le T} \left|\sigma_1\left(t, \widetilde{\xi}_{n'}(t)\right)\sigma_2\left(t, \widetilde{Y}_{n'}(t)\right)\right| > C\right) \to 0 \quad \text{as } C \to \infty.$$

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

$$\left| \sigma_1(t_1, x_1)\sigma_2(t_1, y_1) - \sigma_1(t_2, x_2)\sigma_2(t_2, y_2) \right| < \varepsilon$$

whenever $|t_1 - t_2| < \delta$, $|x_1 - x_2| < \delta$, $|y_1 - y_2| < \delta$. Therefore,

$$\begin{aligned}
\mathbb{P}\big(\big| &\sigma_1\big(t_1, \widetilde{\xi}_{n'}(t_1)\big)\sigma_2\big(t_1, \widetilde{Y}_{n'}(t_1)\big) - \sigma_1\big(t_2, \widetilde{\xi}_{n'}(t_2)\big)\sigma_2\big(t_2, \widetilde{Y}_{n'}(t_2)\big)\big| > \varepsilon\big) \\
&\leq \mathbb{P}\big(\big|\widetilde{\xi}_{n'}(t_1) - \widetilde{\xi}_{n'}(t_2)\big| < \delta, \big|\widetilde{Y}_{n'}(t_1) - \widetilde{Y}_{n'}(t_2)\big| < \delta, \\
&\qquad \big|\sigma_1\big(t_1, \widetilde{\xi}_{n'}(t_1)\big)\sigma_2\big(t_1, \widetilde{Y}_{n'}(t_1)\big) - \sigma_1\big(t_2, \widetilde{\xi}_{n'}(t_2)\big)\sigma_2\big(t_2, \widetilde{Y}_{n'}(t_2)\big)\big| > \varepsilon\big) \\
&\quad + \mathbb{P}\big(\big|\widetilde{\xi}_{n'}(t_1) - \widetilde{\xi}_{n'}(t_2)\big| \geq \delta\big) + \mathbb{P}\big(\big|\widetilde{Y}_{n'}(t_1) - \widetilde{Y}_{n'}(t_2)\big| \geq \delta\big) \\
&= \mathbb{P}\big(\big|\xi_{n'}(t_1) - \xi_{n'}(t_2)\big| \geq \delta\big) + \mathbb{P}\big(|Y(t_1) - Y(t_2)| \geq \delta\big),
\end{aligned}$$

and the last relation implies the following one:

$$\lim_{h \to 0} \lim_{n' \to \infty} \sup_{|t_1 - t_2| \leq h} \mathbb{P}\big(\big| \sigma_1\big(t_1, \widetilde{\xi}_{n'}(t_1)\big)\sigma_2\big(t_1, \widetilde{Y}_{n'}(t_1)\big)$$
$$- \sigma_1\big(t_2, \widetilde{\xi}_{n'}(t_2)\big)\sigma_2\big(t_2, \widetilde{Y}_{n'}(t_2)\big)\big| > \varepsilon\big) = 0.$$

Applying Lemma 1, we get that

$$\sum_{t_{k+1}^{n'} \leq t} \sigma_1\big(t_k^{n'}, \widetilde{\xi}_{n'}\big(t_k^{n'}\big)\big)\sigma_2\big(t_k^{n'}, \widetilde{Y}_{n'}\big(t_k^{n'}\big)\big)\Delta W_k^{n'} \to \int_0^T \sigma_1\big(s, \widetilde{\xi}(s)\big)\sigma_2\big(s, \widetilde{Y}(s)\big)\, d\widetilde{W}(s)$$

in probability as $n' \to \infty$, and we also have that

$$\sum_{t_{k+1}^{n'} \leq t} a\big(t_k^{n'}, \widetilde{\xi}_{n'}\big(t_k^{n'}\big)\big)\Delta t_k^{n'} \to \int_0^T a\big(s, \widetilde{\xi}(s)\big)\, ds,$$

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 $dW_t^1 dW_t^2 = \rho dt$ for some $|\rho| \leq 1$. In this case, we can represent $W_t^2 = \rho W_t^1 + \sqrt{1 - \rho^2} W_t^3$, where W^3 is a Wiener process independent of W^1.

Theorem 2. *Consider the system of stochastic differential equations*

$$\begin{cases} dX_t = a(t, X_t)\, dt + \sigma_1(t, X_t)\sigma_2(t, Y_t)\, dW_t^1, & (3) \\ dY_t = \alpha(t, Y_t)\, dt + \beta(t, Y_t)\, dW_t^2, & (4) \end{cases}$$

where all coefficients a, α, σ_1, σ_2, and β are nonrandom measurable and bounded functions, σ_1, σ_2, and β 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).

Proof. Equations (3) and (4) are equivalent to the two-dimensional stochastic differential equation

$$dZ(t) = A(t, Z_t)\, dt + B(t, Z_t)\, dW(t),$$

where $Z(t) = \binom{X(t)}{Y(t)}$, $W(t) = \binom{W^1(t)}{W^3(t)}$ is a two dimensional Wiener process,

$$A(t, x, y) = \begin{pmatrix} a(t, x) \\ \alpha(t, y) \end{pmatrix}, \quad \text{and} \quad B(t, x, y) = \begin{pmatrix} \sigma_1(t, x)\sigma_2(t, y) & 0 \\ \rho\beta(t, y) & \sqrt{1 - \rho^2}\beta(t, y) \end{pmatrix}.$$

It follows from the measurability and boundedness of a and α and from the continuity and boundedness of σ_1, σ_2, and β that the coefficients of matrices A and B are nonrandom, measurable, and bounded, and additionally the coefficients of B are continuous in all arguments. Then we can apply Theorems 4.2 and 5.6 from [21] (see also Prop. 1.14 in [3]) and deduce that we have to prove the following relation: for any $(t, x, y) \in \mathbb{R}^+ \times \mathbb{R}^2$, there exists $\varepsilon(t, x, y) > 0$ such that, for all $\lambda \in \mathbb{R}^2$,

$$\left\| B(t, x, y)\lambda \right\| \geq \varepsilon(t, x, y)\|\lambda\|. \tag{5}$$

Relation (5) is equivalent to the following one (we omit arguments):

$$\sigma_1^2\sigma_2^2\lambda_1^2 + \beta^2\left(\rho\lambda_1 + \sqrt{1 - \rho^2}\lambda_2\right)^2 \geq \varepsilon^2\left(\lambda_1^2 + \lambda_2^2\right)$$

or

$$\left(\sigma_1^2\sigma_2^2 + \beta^2\rho^2\right)\lambda_1^2 + \beta^2\left(1 - \rho^2\right)\lambda_2^2 + 2\rho\sqrt{1 - \rho^2}\beta^2\lambda_1\lambda_2 \geq \varepsilon^2\left(\lambda_1^2 + \lambda_2^2\right). \tag{6}$$

The quadratic form

$$Q(\lambda_1, \lambda_2) = \left(\sigma_1^2\sigma_2^2 + \beta^2\rho^2\right)\lambda_1^2 + \beta^2\left(1 - \rho^2\right)\lambda_2^2 + 2\rho\sqrt{1 - \rho^2}\beta^2\lambda_1\lambda_2$$

in the left-hand side of (6) is positive definite since its discriminant

$$D = \rho^2\left(1 - \rho^2\right)\beta^4 - \beta^2\left(1 - \rho^2\right)\left(\sigma_1^2\sigma_2^2 + \beta^2\rho^2\right) = -\beta^2\left(1 - \rho^2\right)\sigma_1^2\sigma_2^2 < 0.$$

The continuity of $Q(\lambda_1, \lambda_2)$ implies the existence of $\min_{\lambda_1^2 + \lambda_2^2 = 1} Q(\lambda_1, \lambda_2) > 0$. Then, putting $\varepsilon = \min_{\lambda_1^2 + \lambda_2^2 = 1} Q(\lambda_1, \lambda_2)$ 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, σ_1, and σ_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*

$$\left|\sigma_1(t, x) - \sigma_1(t, y)\right| \leq \rho(|x - y|), \quad t \geq 0, \; x, y \in \mathbb{R},$$

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*

$$\left|a(t, x) - a(t, y)\right| \le k(|x - y|), \quad t \ge 0, \ x, y \in \mathbb{R},$$

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

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

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 $0 \le \phi_n(u) \le \frac{2\rho^{-2}(u)}{n}$ and $\int_{a_n}^{a_{n-1}} \phi_n(u)\, du = 1$. 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, $\varphi_n \in C^2(\mathbb{R})$, $\left|\varphi_n'(x)\right| \le 1$, and $\varphi_n(x) \nearrow |x|$ 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,

$$\varphi_n(X_1(t) - X_2(t))$$
$$= \int_0^t \varphi_n'(X_1(s) - X_2(s))\sigma_2(s, Y(s))(\sigma_1(s, X_1(s)) - \sigma_1(s, X_2(s)))\, dW(s)$$
$$+ \int_0^t \varphi_n'(X_1(s) - X_2(s))(a(s, X_1(s)) - a(s, X_2(s)))\, ds$$
$$+ \frac{1}{2} \int_0^t \varphi_n''(X_1(s) - X_2(s))\sigma_2(s, Y(s))^2(\sigma_1(s, X_1(s)) - \sigma_1(s, X_2(s)))^2\, ds$$
$$= J_1 + J_2 + J_3.$$

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

$$|\mathbb{E}(J_2)| \le \int_0^t \mathbb{E}\left|a(s, X_1(s)) - a(s, X_2(s))\right|\, ds$$
$$\le \int_0^t \mathbb{E}(k(|X_1(s) - X_2(s)|)\, ds \le \int_0^t k(\mathbb{E}(|X_1(s) - X_2(s)|)\, ds$$

by Jensen's inequality, and

$$\left|\mathbb{E}(J_3)\right| \leq \frac{C^2}{2} \int_0^t \mathbb{E}\left(\frac{2}{n}\rho^{-2}\left(|X_1(s) - X_2(s)|\right)\rho^2\left(|X_1(s) - X_2(s)|\right)\right)ds$$

$$\leq \frac{t}{n} \to 0 \quad \text{as } n \to \infty.$$

So by letting $n \to \infty$ we get

$$\mathbb{E}\left(|X_1(s) - X_2(s)|\right) \leq \int_0^t k\left(\mathbb{E}\left(|X_1(s) - X_2(s)|\right)\right) ds.$$

We have that $\int_0^\infty k^{-1}(u)du = +\infty$. Then we get $\mathbb{E}(|X_1(s) - X_2(s)|) = 0$, and hence $X_1(s) = X_2(s)$ a.s. $\qquad\square$

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

Theorem 4. *Let a, σ_1, and σ_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 $x \in \mathbb{R}$,*

$$\left|\sigma_1(t, x)\right|^2 + \left|a(t, x)\right|^2 \leq K^2\left(1 + |x|^2\right);$$

(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 $|y| \leq N$,*

$$\left|a(t, x) - a(t, y)\right| + \left|\sigma_1(t, x) - \sigma_1(t, y)\right| \leq K_N |x - y|;$$

(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 $(\Omega, \mathcal{F}, \overline{\mathcal{F}}, \mathbb{P})$ be a complete probability space with filtration $\overline{\mathcal{F}} = \{\mathcal{F}_t, t \geq 0\}$ satisfying the standard assumptions. We assume that all processes under consideration are adapted to the filtration $\overline{\mathcal{F}}$. Consider a parameterized version of Eq. (2)

$$dX_t = \theta a(t, X_t) dt + \sigma_1(t, X_t)\sigma_2(t, Y_t) dW_t, \tag{7}$$

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 θ by continuous observations of X and Y.

Denote

$$f(t, x, y) = \frac{a(t, x)}{\sigma_1^2(t, x)\sigma_2^2(t, y)}, \qquad g(t, x, y) = \frac{a(t, x)}{\sigma_1(t, x)\sigma_2(t, y)}.$$

Assume that, for all $t > 0$,

$$\sigma_1(t, X_t)\sigma_2(t, Y_t) \neq 0 \quad \text{a.s.,} \tag{8}$$

$$\int_0^t g^2(s, X_s, Y_s)\, ds < \infty \quad \text{a.s.,} \tag{9}$$

$$\int_0^\infty g^2(s, X_s, Y_s)\, ds = \infty \quad \text{a.s.} \tag{10}$$

Then a likelihood function for Eq. (7) has the form

$$\frac{d P_\theta(T)}{d P_0(T)} = \exp\left\{\theta \int_0^T f(t, X_t, Y_t)\, dX_t - \frac{\theta^2}{2} \int_0^T g^2(t, X_t, Y_t)\, dt\right\};$$

see [15, Ch. 7]. Hence, the maximum likelihood estimator of parameter θ constructed by observations of X and Y on the interval $[0, T]$ has the form

$$\hat{\theta}_T = \frac{\int_0^T f(t, X_t, Y_t)\, dX_t}{\int_0^T g^2(t, X_t, Y_t)\, dt} = \theta + \frac{\int_0^T g(t, X_t, Y_t)\, dW_t}{\int_0^T g^2(t, X_t, Y_t)\, dt}. \tag{11}$$

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 $\langle M \rangle_t = \int_0^t g^2(s, X_s, Y_s)\, ds$. According to the strong law of large numbers for martingales [14, Ch. 2, § 6, Thm. 10, Cor. 1], under the condition $\langle M \rangle_T \to \infty$ a.s. as $T \to \infty$, we have that $\frac{M_T}{\langle M \rangle_T} \to 0$ a.s. as $T \to \infty$. Therefore, it follows from representation (11) that $\hat{\theta}_T$ is strongly consistent. $\qquad\square$

3.2   Linear equation with stochastic volatility

As an example, let us consider the model

$$dX_t = \theta X_t\, dt + X_t \sigma_2(Y_t)\, dW_t, \quad X_0 = x_0 \in \mathbb{R}, \tag{12}$$

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

$$dY_t = \alpha(Y_t)\,dt + \beta(Y_t)\,dW_t^1, \quad Y_0 = y_0 \in J, \tag{13}$$

where W^1 is a Wiener process, possibly correlated with W.

By $L_{\text{loc}}^1(J)$ we denote the set of Borel functions $J \to [-\infty, \infty]$ 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 α and β satisfy the Engelbert–Schmidt conditions

(A2) $\beta(y) \neq 0$ for all $y \in J$, and

(A3) $\beta^{-2}, \alpha\beta^{-2} \in L_{\text{loc}}^1(J)$.

Let us introduce the following notation:

$$\rho(y) = \exp\left\{-2\int_c^y \frac{\alpha(u)}{\beta^2(u)}\,du\right\}, \quad y \in J,$$

$$s(y) = \int_c^y \rho(u)\,du, \quad y \in \bar{J} = [l, r],$$

for some $c \in J$. Assume additionally that

(A4) $s(r) = \infty$ or $\frac{s(r)-s}{\rho\beta^2} \notin L_{\text{loc}}^1(r-)$,

(A5) $s(l) = -\infty$ or $\frac{s-s(l)}{\rho\beta^2} \notin L_{\text{loc}}^1(l+)$.

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

Assume also that

(A6) $\beta^{-2}\sigma_2^{-2} \in L_{\text{loc}}^1(J)$,

(A7) one of the following four conditions holds:

    (i) $s(r) = \infty$, $s(l) = -\infty$,

    (ii) $s(r) < \infty$, $s(l) = -\infty$, $\frac{s(r)-s}{\rho\beta^2\sigma_2^2} \notin L_{\text{loc}}^1(r-)$,

    (iii) $s(r) = \infty$, $s(l) > -\infty$, $\frac{s-s(l)}{\rho\beta^2\sigma_2^2} \notin L_{\text{loc}}^1(l+)$,

    (iv) $s(r) < \infty$, $s(l) > -\infty$, $\frac{s(r)-s}{\rho\beta^2\sigma_2^2} \notin L_{\text{loc}}^1(r-)$, $\frac{s-s(l)}{\rho\beta^2\sigma_2^2} \notin L_{\text{loc}}^1(l+)$,

(A8) $X_t\sigma(Y_t) \neq 0$ a.s., $t \geq 0$.

The maximum likelihood estimator (11) for model (12) equals

$$\hat{\theta}_T = \frac{\int_0^T X_t^{-1}\sigma_2^{-2}(Y_t)\,dX_t}{\int_0^T \sigma_2^{-2}(Y_t)\,dt}. \tag{14}$$

Theorem 6. *Under assumptions* (A1)–(A8)*, the estimator* $\hat{\theta}_T$ *is strongly consistent as* $T \rightarrow \infty$.

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

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

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

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

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 $\frac{s(r)-s}{\rho \beta^2 \sigma_2^2} \notin L_{\text{loc}}^1(r-)$, 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*

$$dY_t = \alpha \, dt + \beta \, dW_t^1, \quad Y_0 = y_0 \in \mathbb{R},$$

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

(i) $\alpha = 0$,

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

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

Then estimator (14) *is strongly consistent.*

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

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

If $\alpha = 0$, then $s(y) = y$, $s(+\infty) = \infty$, $s(-\infty) = -\infty$, and assumption (A7)(i) is satisfied. Otherwise, we have

$$s(y) = \frac{\beta^2}{2\alpha}\left(1 - \exp\left\{-\frac{2\alpha}{\beta^2}y\right\}\right).$$

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*

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

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 $\exp\{\frac{a}{\gamma^2}(u-b)^2\} \geq 1$, 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, $s(+\infty) = -s(-\infty) = \frac{\gamma\sqrt{\pi}}{2\sqrt{-a}} < \infty$, and we need to verify (A7)(iv). Since $\int_x^\infty e^{-z^2}\,dz \sim \frac{1}{2x}e^{-x^2}$ as $x \to \infty$, we see that

$$\frac{s(+\infty) - s(y)}{\rho(y)\gamma^2\sigma_2^2(y)} = \frac{\frac{\gamma}{\sqrt{-a}}\int_{\frac{\sqrt{-a}}{\gamma}(y-b)}^{\infty} e^{-z^2}\,dz}{\exp\left\{\frac{a}{\gamma^2}(y-b)^2\right\}\gamma^2\sigma_2^2(y)} \sim \frac{1}{-2a(y-b)\sigma_2^2(y)}$$

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*

$$dY_t = \alpha Y_t\,dt + \beta Y_t\,dW_t^1, \quad Y_0 = y_0 > 0,$$

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

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

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

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

Then estimator (14) *is strongly consistent.*

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

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

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

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

$$\frac{s(+\infty) - s(y)}{\rho(y)\beta^2 y^2 \sigma_2^2(y)} = \frac{1}{(2\alpha^2 - \beta^2)y\sigma_2^2(y)} \notin L^1_{\mathrm{loc}}(+\infty).$$

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

$$\frac{s(y) - s(0)}{\rho(y)\beta^2 y^2 \sigma_2^2(y)} = \frac{1}{(\beta^2 - 2\alpha^2)y\sigma_2^2(y)} \notin L^1_{\mathrm{loc}}(0+).$$

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, γ are positive constants, and $2ab \geq \gamma^2$. Assume that

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

Then estimator (14) *is strongly consistent.*

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

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

$$s(y) = e^{-\frac{2a}{\gamma^2}}\int_1^y u^{-\frac{2ab}{\gamma^2}} e^{\frac{2a}{\gamma^2}u}\, du.$$

Table 1. The means and standard deviations of $\hat{\theta}_T$

					T		
$\alpha(y)$	$\beta(y)$	$\sigma_2(y)$		10	50	100	200
1	1	$|y|^{1/4}$	Mean	1.9455	1.9431	1.9711	1.9762
			Std.dev.	0.4260	0.2576	0.2367	0.2022
y	$2y$	$\sqrt{y}$	Mean	2.0104	2.0000	2.0000	2.0000
			Std.dev.	0.1225	$5.7 \cdot 10^{-5}$	$4.7 \cdot 10^{-8}$	$1.6 \cdot 10^{-14}$
y	y	$(1+y)^{-1}$	Mean	2.0008	2.0001	2.0000	2.0000
			Std.dev.	0.0769	0.0010	$2.2 \cdot 10^{-12}$	$1.4 \cdot 10^{-14}$
y	1	$2 + \sin y$	Mean	1.9358	1.9819	1.9927	1.9939
			Std.dev.	0.5436	0.2437	0.1679	0.1077
$-y$	1	$2 + \sin y$	Mean	1.9061	1.9684	1.9700	1.9786
			Std.dev.	0.5994	0.2472	0.1781	0.1254
$2-y$	$\sqrt{y}$	$\sqrt{y}$	Mean	1.9923	2.0039	1.9796	1.9872
			Std.dev.	0.3540	0.1604	0.1173	0.0782
$2-y$	$\sqrt{y}$	y	Mean	2.0830	1.9835	1.9803	1.9886
			Std.dev.	0.4347	0.1974	0.1205	0.0840

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

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

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 $\delta > 0$, $\lim_{h \to 0} \overline{\lim}_{n \to \infty} \sup_{|t_1 - t_2| \leq h} \mathbb{P}(|\xi_n^{(i)}(t_1) - \xi_n^{(i)}(t_2)| > \delta) = 0$,*

(ii) $\lim_{C \to \infty} \lim_{n \to \infty} \sup_{0 \le t \le T} \mathbb{P}(|\xi_n^{(i)}(t)| > C) = 0$.

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,*

$$\mathbb{P}\left(\sup_{0 \le t \le T} |f_n(t)| > C \right) \le \varepsilon,$$

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

$$\lim_{h \to 0} \lim_{n \to \infty} \sup_{|t_1 - t_2| \le h} \mathbb{P}(|f_n(t_2) - f_n(t_1)| > \varepsilon) = 0.$$

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 $0 = t_0^n < t_1^n < \cdots < t_n^n = T$ of $[0, T]$ such that $\lim_{n \to \infty} \max_k (t_{k+1}^n - t_k^n) = 0$. Define ξ_k^n by $\xi_0^n = X(0)$ and

$$\xi_{k+1}^n = \xi_k^n + a(t_k^n, \xi_k^n)\Delta t_k^n + \sigma_1(t_k^n, \xi_k^n)\sigma_2(t_k^n, Y(t_k^n))\Delta W_k^n.$$

Lemma 1. *The random variables $\sup_k |\xi_k^n|$ are bounded in probability uniformly w.r.t. n.*

Proof. Let $\eta_0^n = \xi_0^n \mathbb{1}_{|\xi_0^n| \le N}$ and

$$\eta_{k+1}^n = \eta_k^n + a^N(t_k^n, \eta_k^n)\Delta t_k^n + \sigma_1^N(t_k^n, \eta_k^n)\sigma_2(t_k^n, Y(t_k^n))\Delta W_k^n, \qquad (18)$$

where $a^N(t, x) = a(t, x)\mathbb{1}_{|x| \le N}$, $\sigma_1^N(t, x) = \sigma_1(t, x)\mathbb{1}_{|x| \le N}$. If $|\eta_k^n| > N$, then $\eta_{k+1}^n = \eta_k^n$, and if $|\eta_k^n| \le N$, then

$$|\eta_{k+1}^n| \le N + |a^N(t_k^n, \eta_k^n)|\Delta t_k^n + |\sigma_1^N(t_k^n, \eta_k^n)\sigma_2(t_k^n, Y(t_k^n))| |\Delta W_k^n|.$$

Then, for any $1 \le k \le n$,

$$|\eta_k^n| \le N + \sum_{r=0}^{k-1} |a^N(t_r^n, \eta_r^n)|\Delta t_r^n + \sum_{r=0}^{k-1} |\sigma_1^N(t_r^n, \eta_r^n)\sigma_2(t_r^n, Y(t_r^n))| |\Delta W_r^n|$$

and is square-integrable. Furthermore,

$$
\begin{aligned}
\mathbb{E}\left|\eta_{k+1}^n\right|^2 &= \mathbb{E}\left|\eta_k^n\right|^2 + 2\mathbb{E}\big(a^N\big(t_k^n, \eta_k^n\big)\eta_k^n \Delta t_k^n\big) + \mathbb{E}a^N\big(t_k^n, \eta_k^n\big)^2\big(\Delta t_k^n\big)^2 \\
&\quad + \mathbb{E}\big(\big(\sigma_1^N\big(t_k^n, \eta_k^n\big)\big)^2\big(\sigma_2\big(t_k^n, Y\big(t_k^n\big)\big)\big)^2\big(\Delta W_k^n\big)^2\big) \\
&\leq \mathbb{E}\left|\eta_k^n\right|^2 + 2\mathbb{E}\big(a^N\big(t_k^n, \eta_k^n\big)\eta_k^n \Delta t_k^n\big) + \mathbb{E}a^N\big(t_k^n, \eta_k^n\big)^2\big(\Delta t_k^n\big)^2 \\
&\quad + C^2\mathbb{E}\big(\big(\sigma_1^N\big(t_k^n, \eta_k^n\big)\big)^2 \Delta t_k^n\big).
\end{aligned}
$$

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

$$
\begin{aligned}
\mathbb{E}\left|\eta_{k+1}^n\right|^2 &\leq \mathbb{E}\left|\eta_k^n\right|^2\big(1 + H\Delta t_k^n\big) + H\Delta t_k^n \leq \mathbb{E}\left|\eta_k^n\right|^2 e^{H\Delta t_k^n} + H\Delta t_k^n, \\
\mathbb{E}\left|\eta_{k+1}^n\right|^2 + 1 &\leq \big(\mathbb{E}\left|\eta_k^n\right|^2 + 1\big)e^{H\Delta t_k^n} \leq \big(\mathbb{E}\left|\eta_0^n\right|^2 + 1\big)e^{HT}, \\
\mathbb{E}\left|\eta_{k+1}^n\right|^2 &\leq \big(\mathbb{E}\left|\eta_0^n\right|^2 + 1\big)e^{HT} - 1.
\end{aligned}
$$

We have that

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

So, since

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

and

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

we get

$$
\mathbb{E}\sup_k \left|\eta_k^n\right|^2 \leq A + B\left|\eta_0^n\right|^2.
$$

We have that η_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. $\qquad\square$

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

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

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

Proof. Let $\eta_n(t) = \eta_k^n$, $t \in [t_k^n, t_{k+1}^n)$. Then

$$\sup_{|t_1-t_2|\leq h} \mathbb{P}\big(\big|\xi_n(t_1) - \xi_n(t_2)\big| > \delta\big) \leq \sup_{|t_1-t_2|\leq h} \mathbb{P}\big(\big|\eta_n(t_1) - \eta_n(t_2)\big| > \delta\big)$$
$$+ \mathbb{P}\Big(\sup_k \big|\xi_k^n - \eta_k^n\big| > 0\Big).$$

From (18) and the boundedness of a^N, σ_1^N, and σ_2 we have that

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

Therefore,

$$\lim_{h\to 0} \overline{\lim_{n\to\infty}} \sup_{|t_1-t_2|\leq h} \mathbb{P}\big(\big|\xi_n(t_1) - \xi_n(t_2)\big| > \delta\big) \leq \overline{\lim_{n\to\infty}} \mathbb{P}\Big(\sup_k \big|\xi_k^n - \eta_k^n\big| > 0\Big).$$

The proof follows now from Remark 1. □

Acknowledgment

References

[1] Aït-Sahalia, Y., Kimmel, R.: Maximum likelihood estimation of stochastic volatility models. J. Financ. Econ. **83**, 413–452 (2007)

[2] Altay, S., Schmock, U.: Lecture notes on the Yamada–Watanabe condition for the pathwise uniqueness of solutions of certain stochastic differential equations. fam.tuwien.ac.at/~schmock/notes/Yamada-Watanabe.pdf. Accessed 25 April 2013

[3] Cherny, A.S., Engelbert, H.-J.: Singular Stochastic Differential Equations. Lecture Notes in Mathematics, vol. 1858, Springer (2005) MR2112227. doi:10.1007/b104187

[4] Fouque, J.-P., Papanicolaou, G., Sircar, K.R.: Derivatives in Financial Markets with Stochastic Volatility. Cambridge University Press, Cambridge (2000) MR1768877

[5] Fouque, J.-P., Papanicolaou, G., Sircar, K.R.: Mean-reverting stochastic volatility. Int. J. Theor. Appl. Finance **3**(1), 101–142 (2000)

[6] Heston, S.: A closed-form solution of options with stochastic volatility with applications to bond and currency options. Rev. Financ. Stud. **6**(2), 327–343 (1993)

[7] Heyde, C.C.: Quasi-likelihood and Its Application. Springer Series in Statistics. Springer (1997) MR1461808. doi:10.1007/b98823

[8] Hull, J., White, A.: The pricing of options on assets with stochastic volatilities. J. Finance **42**, 281–300 (1987)

[9] Jankunas, A., Khasminskii, R.Z.: Estimation of parameters of linear homogenous stochastic differential equations. Stochastic Process. Appl. **72**(2), 205–219 (1997) MR1486553. doi:10.1016/S0304-4149(97)00083-5

[10] Krylov, N.V.: On Itô's stochastic integral equations. Theory Probab. Appl. **14**(2), 330–336 (1969) MR0270462

[11] Krylov, N.V.: Controlled Diffusion Processes. Applications of Mathematics, vol. 14, Springer (1980) MR0601776

[12] Kuchuk-Iatsenko, S., Mishura, Y.: Option pricing in the model with stochastic volatility driven by Ornstein–Uhlenbeck process. Simulation. Mod. Stoch. Theory Appl. **2**(4), 355–369 (2015) MR3456143. doi:10.15559/15-VMSTA43

[13] Kuchuk-Iatsenko, S., Mishura, Y.: Pricing the European call option in the model with stochastic volatility driven by Ornstein–Uhlenbeck process. Exact formulas. Mod. Stoch. Theory Appl. **2**(3), 233–249 (2015) MR3407504. doi:10.15559/15-VMSTA36CNF

[14] Liptser, R.S., Shiryayev, A.N.: Theory of Martingales. Mathematics and its Applications (Soviet Series), vol. 49. Kluwer Academic Publishers Group, Dordrecht (1989) MR1022664. doi:10.1007/978-94-009-2438-3

[15] Liptser, R.S., Shiryaev, A.N.: Statistics of Random Processes. I. Applications of Mathematics (New York), vol. 5, Springer (2001) MR1800858

[16] Liptser, R.S., Shiryaev, A.N.: Statistics of Random Processes. II. Applications of Mathematics (New York), vol. 6, Springer (2001) MR1800858

[17] Mijatović, A., Urusov, M.: Convergence of integral functionals of one-dimensional diffusions. Electron. Commun. Probab. **17**, 61–13 (2012) MR3005734. doi:10.1214/ECP.v17-1825

[18] Mishura, Y.: Standard maximum likelihood drift parameter estimator in the homogeneous diffusion model is always strongly consistent. Statist. Probab. Lett. **86**, 24–29 (2014) MR3162713. doi:10.1016/j.spl.2013.12.004

[19] Nisio, M.: Stochastic Control Theory. Dynamic Programming Principle, 2nd edn. Probability Theory and Stochastic Modelling, vol. 72, Springer (2015) MR3290231. doi:10.1007/978-4-431-55123-2

[20] Skorokhod, A.V.: Studies in the Theory of Random Processes. Translated from the Russian by Scripta Technica, Inc. Addison-Wesley Publishing Co., Inc., Reading, Mass. (1965) MR0185620

[21] Stroock, D.W., Varadhan, S.R.S.: Diffusion processes with continuous coefficients. I. Comm. Pure Appl. Math. **22**, 345–400 (1969) MR0253426

[22] Stroock, D.W., Varadhan, S.R.S.: Diffusion processes with continuous coefficients. II. Comm. Pure Appl. Math. **22**, 479–530 (1969) MR0254923

[23] Yamada, T., Watanabe, S.: On the uniqueness of solutions of stochastic differential equations. J. Math. Kyoto Univ. **11**, 155–167 (1971) MR0278420