Stat Inference Stoch Process (2019) 22:41–75
https://doi.org/10.1007/s11203-018-9174-z

On conditional least squares estimation for affine diffusions based on continuous time observations

Beáta Bolyog[1] · **Gyula Pap**[1]

Received: 26 July 2017 / Accepted: 27 January 2018 / Published online: 5 February 2018

Abstract We study asymptotic properties of conditional least squares estimators for the drift parameters of two-factor affine diffusions based on continuous time observations. We distinguish three cases: subcritical, critical and supercritical. For all the drift parameters, in the subcritical and supercritical cases, asymptotic normality and asymptotic mixed normality is proved, while in the critical case, non-standard asymptotic behavior is described.

Keywords Affine processes · Continuous time observations · Conditional least squares estimators

Mathematics Subject Classification 60J80 · 62F12

1 Introduction

Affine processes are applied in mathematical finance in several models including interest rate models (e.g. the Cox–Ingersoll–Ross, Vasiček or general affine term structure short rate models), option pricing (e.g. the Heston model) and credit risk models, see e.g. Duffie et al. (2003), Filipović (2009), Baldeaux and Platen (2013), and Alfonsi (2015). In this paper we consider two-factor affine processes, i.e. affine processes with state-space $[0, \infty) \times \mathbb{R}$. Dawson and Li (2006) derived a jump-type stochastic differential equation (SDE) for such processes. Specializing this result to the diffusion case, i.e. two-factor affine processes without jumps, we obtain that for every $a \in [0, \infty)$, $b, \alpha, \beta, \gamma \in \mathbb{R}$, $\sigma_1, \sigma_2, \sigma_3 \in [0, \infty)$ and $\varrho \in [-1, 1]$, the SDE

✉ Gyula Pap
papgy@math.u-szeged.hu

Beáta Bolyog
bbeata@math.u-szeged.hu

[1] Bolyai Institute, University of Szeged, Aradi Vértanúk Tere 1, 6720 Szeged, Hungary

$$\begin{cases} \mathrm{d}Y_t = (a - bY_t)\,\mathrm{d}t + \sigma_1\sqrt{Y_t}\,\mathrm{d}W_t, \\ \mathrm{d}X_t = (\alpha - \beta Y_t - \gamma X_t)\,\mathrm{d}t + \sigma_2\sqrt{Y_t}\,(\varrho\,\mathrm{d}W_t + \sqrt{1-\varrho^2}\,\mathrm{d}B_t) + \sigma_3\,\mathrm{d}L_t, \end{cases} \qquad t \in [0,\infty), \tag{1.1}$$

with an arbirary initial value (Y_0, X_0) with $\mathbb{P}(Y_0 \in [0,\infty)) = 1$ and independent of a 3-dimensional standard Wiener process $(W_t, B_t, L_t)_{t\in[0,\infty)}$, has a pathwise unique strong solution being a two-factor affine diffusion process, and conversely, every two-factor affine diffusion process is a pathwise strong solution of a SDE (1.1) with appropriate parameters $a \in [0,\infty)$, $b, \alpha, \beta, \gamma \in \mathbb{R}$, $\sigma_1, \sigma_2, \sigma_3 \in [0,\infty)$ and $\varrho \in [-1,1]$, see Proposition 2.1.

The aim of this paper is to study the asymptotic properties of the conditional least squares estimators (CLSE) $(\widehat{a}_T, \widehat{b}_T, \widehat{\alpha}_T, \widehat{\beta}_T, \widehat{\gamma}_T)$ of the drift parameters $(a, b, \alpha, \beta, \gamma)$ based on continuous time observations $(Y_t, X_t)_{t\in[0,T]}$ with $T > 0$. This estimator is the high frequency limit in probability as $n \to \infty$ of the CLSE based on discrete time observations $(Y_{k/n}, X_{k/n})_{k\in\{0,\ldots,\lfloor nT\rfloor\}}$, $n \in \mathbb{N}$. We do not estimate the parameters $\sigma_1, \sigma_2, \sigma_3$ and ϱ, since for all $T \in (0,\infty)$, they are measurable functions (i.e., statistics) of $(Y_t, X_t)_{t\in[0,T]}$, see Appendix C in the extended arXiv version Bolyog and Pap (2017) of this paper. For the calculation of $(\widehat{a}_T, \widehat{b}_T, \widehat{\alpha}_T, \widehat{\beta}_T, \widehat{\gamma}_T)$ one does not need to know the values of the diffusion coefficients $\sigma_1, \sigma_2, \sigma_3$ and ϱ, see (3.4).

The first coordinate process Y in (1.1) is called a Cox–Ingersoll–Ross (CIR) process (see Cox et al. 1985). In the submodel consisting only of the process Y, (Overbeck and Rydén 1997 Theorems 3.4, 3.5 and 3.6) derived the CLSE of (a, b) based on continuous time observations $(Y_t)_{t\in[0,T]}$ with $T > 0$, i.e., the limit in probability as $n \to \infty$ of the CLSE based on discrete time observations $(Y_{k/n})_{k\in\{0,\ldots,\lfloor nT\rfloor\}}$, $n \in \mathbb{N}$, which turns to be the same as the CLSE $(\widehat{a}_T, \widehat{b}_T)$ of (a, b) based on continuous time observations $(Y_t, X_t)_{t\in[0,T]}$, and they proved strong consistency and asymptotic normality in case of a subcritical CIR process Y, i.e., when $b > 0$ and the initial distribution is the unique stationary distribution of the model.

Barczy et al. (2014) considered a submodel of (1.1) with $a \in (0,\infty)$, $\beta = 0$, $\sigma_1 = 1$, $\sigma_2 = 1$, $\varrho = 0$ and $\sigma_3 = 0$. The estimator of the parameters (α, γ) based on continuous time observations $(X_t)_{t\in[0,T]}$ with $T > 0$ (which they call a least square estimator) is in fact the CLSE, i.e., the limit in probability as $n \to \infty$ of the CLSE based on discrete time observations $(X_{k/n})_{k\in\{0,\ldots,\lfloor nT\rfloor\}}$, $n \in \mathbb{N}$, which can be shown by the method of the proof of Lemma 3.3. They proved strong consistency and asymptotic normality in case of a subcritical process (Y, X), i.e., when $b > 0$ and $\gamma > 0$.

Barczy et al. (2016) considered the so-called Heston model, which is a submodel of (1.1) with $a, \sigma_1, \sigma_2 \in (0,\infty)$, $\gamma = 0$, $\varrho \in (-1, 1)$ and $\sigma_3 = 0$. The estimator of the parameters (a, b, α, β) based on continuous time observations $(Y_t, X_t)_{t\in[0,T]}$ with $T > 0$ (which they call least square estimator) is in fact the CLSE, i.e., the limit in probability as $n \to \infty$ of the CLSE based on discrete time observations $(Y_{k/n}, X_{k/n})_{k\in\{0,\ldots,\lfloor nT\rfloor\}}$, $n \in \mathbb{N}$ which can be shown by the method of the proof of Lemma 3.3. They proved strong consistency and asymptotic normality in case of a subcritical process (Y, X), i.e., when $b > 0$. Note that Barczy and Pap (2016) studied the maximum likelihood estimator (MLE) $(\widetilde{a}_T, \widetilde{b}_T, \widetilde{\alpha}_T, \widetilde{\beta}_T)$ of the parameters (a, b, α, β) in this Heston model under the additional assumption $a \geqslant \frac{\sigma_1^2}{2}$. In the subcritical case, i.e., when $b > 0$, for $(\widetilde{a}_T, \widetilde{b}_T, \widetilde{\alpha}_T, \widetilde{\beta}_T)$, they proved strong consistency and asymptotic normality in case of $a > \frac{\sigma_1^2}{2}$, and weak consistency in case of $a = \frac{\sigma_1^2}{2}$. In the critical case, namely, if $b = 0$, under the additional assumption $a > \frac{\sigma_1^2}{2}$, they showed weak consistency of $(\widetilde{a}_T, \widetilde{b}_T, \widetilde{\alpha}_T, \widetilde{\beta}_T)$, asymptotic normality of $(\widetilde{a}_T, \widetilde{\alpha}_T)$, and determined the asymptotic behavior of $(\widetilde{a}_T, \widetilde{b}_T, \widetilde{\alpha}_T, \widetilde{\beta}_T)$. In the supercritical case, namely, when $b < 0$, they showed that $\widetilde{b}_T$ is strongly consistent, $\widetilde{\beta}_T$ is weakly consistent, $(\widetilde{b}_T, \widetilde{\beta}_T)$ is asymptotically mixed normal, and determined the asymptotic behavior of $(\widetilde{a}_T, \widetilde{b}_T, \widetilde{\alpha}_T, \widetilde{\beta}_T)$. Barczy et al. (2018a, b) studied the asymptotic behavior of maximum likelihood estimators for a jump-type Heston model and for the growth rate of a jump-type CIR process, respectively, based on continuous time observations.

We consider general two-factor affine diffusions (1.1). In the subcritical case, i.e., when $b > 0$ and $\gamma > 0$, we prove strong consistency and asymptotic normality of $(\widehat{a}_T, \widehat{b}_T, \widehat{\alpha}_T, \widehat{\beta}_T, \widehat{\gamma}_T)$ under the additional assumptions $a > 0$, $\sigma_1 > 0$ and $(1-\varrho^2)\sigma_2^2 + \sigma_3^2 > 0$. In a special critical case, namely if $b = 0$ and $\gamma = 0$, we show weak consistency of $(\widehat{b}_T, \widehat{\beta}_T, \widehat{\gamma}_T)$ and determine the asymptotic behavior of $(\widehat{a}_T, \widehat{b}_T, \widehat{\alpha}_T, \widehat{\beta}_T, \widehat{\gamma}_T)$ under the additional assumptions $\beta = 0$ and $(1-\varrho^2)\sigma_2^2 + \sigma_3^2 > 0$. In a special supercritical case, namely, when $\gamma < b < 0$, we show strong consistency of $\widehat{b}_T$, weak consistency of $(\widehat{\beta}_T, \widehat{\gamma}_T)$ and prove asymptotic mixed normality of $(\widehat{a}_T, \widehat{b}_T, \widehat{\alpha}_T, \widehat{\beta}_T, \widehat{\gamma}_T)$ under the additional assumptions $\alpha\beta \leqslant 0$, $\sigma_1 > 0$, and either $\sigma_3 > 0$, or $\big(a - \frac{\sigma_1^2}{2}\big)(1-\varrho^2)\sigma_2^2 > 0$. Note that we decided to deal with the CLSE of $(a, b, \alpha, \beta, \gamma)$, since the MLE of $(a, b, \alpha, \beta, \gamma)$ contains, for example, $\int_0^T \frac{X_t}{(1-\varrho^2)\sigma_2^2 Y_t + \sigma_3^2}\,\mathrm{d}t$, and the question of the asymptotic behavior of this integral as $T \to \infty$ is still open in the critical and supercritical cases. For the sake of brevity of the paper some simple proofs and calculation steps are omitted. However, all these details are included in the extended arXiv version Bolyog and Pap (2017) of this paper.

2 The affine two-factor model

Let $\mathbb{N}$, $\mathbb{Z}_+$, $\mathbb{R}$, $\mathbb{R}_+$, $\mathbb{R}_{++}$, $\mathbb{R}_-$, $\mathbb{R}_{--}$ and $\mathbb{C}$ denote the sets of positive integers, non-negative integers, real numbers, non-negative real numbers, positive real numbers, non-positive real numbers, negative real numbers and complex numbers, respectively. For $x, y \in \mathbb{R}$, we will use the notations $x \wedge y := \min(x, y)$ and $x \vee y := \max(x, y)$. By $C_c^2(\mathbb{R}_+ \times \mathbb{R}, \mathbb{R})$, we denote the set of twice continuously differentiable real-valued functions on $\mathbb{R}_+ \times \mathbb{R}$ with compact support. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space equipped with the augmented filtration $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ corresponding to $(W_t, B_t, L_t)_{t\in\mathbb{R}_+}$ and a given initial value (η_0, ξ_0) being independent of $(W_t, B_t, L_t)_{t\in\mathbb{R}_+}$ such that $\mathbb{P}(\eta_0 \in \mathbb{R}_+) = 1$, constructed as in Karatzas and Shreve (1991, Sect. 5.2). Note that $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ satisfies the usual conditions, i.e., the filtration $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ is right-continuous and $\mathcal{F}_0$ contains all the $\mathbb{P}$-null sets in $\mathcal{F}$. We will denote the convergence in distribution, convergence in probability, almost surely convergence and equality in distribution by $\xrightarrow{\mathcal{D}}$, $\xrightarrow{\mathbb{P}}$, $\xrightarrow{\text{a.s.}}$ and $\overset{\mathcal{D}}{=}$, respectively. By $\|\boldsymbol{x}\|$ and $\|\boldsymbol{A}\|$, we denote the Euclidean norm of a vector $\boldsymbol{x} \in \mathbb{R}^d$ and the spectral norm of a matrix $\boldsymbol{A} \in \mathbb{R}^{d\times d}$, respectively. By $\boldsymbol{I}_d \in \mathbb{R}^{d\times d}$, we denote the $d \times d$ unit matrix. For square matrices $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_k$, $\mathrm{diag}(\boldsymbol{A}_1, \ldots, \boldsymbol{A}_k)$ will denote the square block matrix containing the matrices $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_k$ in its diagonal.

The next proposition is about the existence and uniqueness of a strong solution of the SDE (1.1), see Bolyog and Pap (2016, Proposition 2.2).

Proposition 2.1 *Let* (η_0, ξ_0) *be a random vector independent of the process* $(W_t, B_t, L_t)_{t\in\mathbb{R}_+}$ *satisfying* $\mathbb{P}(\eta_0 \in \mathbb{R}_+) = 1$. *Then for all* $a \in \mathbb{R}_+$, $b, \alpha, \beta, \gamma \in \mathbb{R}$, $\sigma_1, \sigma_2, \sigma_3 \in \mathbb{R}_+$, $\varrho \in [-1, 1]$, *there is a (pathwise) unique strong solution* $(Y_t, X_t)_{t\in\mathbb{R}_+}$ *of the SDE* (1.1) *such that* $\mathbb{P}((Y_0, X_0) = (\eta_0, \xi_0)) = 1$ *and* $\mathbb{P}(Y_t \in \mathbb{R}_+ \text{ for all } t \in \mathbb{R}_+) = 1$. *Further, for all* $s, t \in \mathbb{R}_+$ *with* $s \leqslant t$, *we have*

$$Y_t = \mathrm{e}^{-b(t-s)} Y_s + a \int_s^t \mathrm{e}^{-b(t-u)} \, \mathrm{d}u + \sigma_1 \int_s^t \mathrm{e}^{-b(t-u)} \sqrt{Y_u} \, \mathrm{d}W_u \tag{2.1}$$

and

$$\begin{aligned} X_t = \mathrm{e}^{-\gamma(t-s)} X_s &+ \int_s^t \mathrm{e}^{-\gamma(t-u)} (\alpha - \beta Y_u) \, \mathrm{d}u \\ &+ \sigma_2 \int_s^t \mathrm{e}^{-\gamma(t-u)} \sqrt{Y_u} \, (\varrho \, \mathrm{d}W_u + \sqrt{1-\varrho^2} \, \mathrm{d}B_u) + \sigma_3 \int_s^t \mathrm{e}^{-\gamma(t-u)} \, \mathrm{d}L_u. \end{aligned} \tag{2.2}$$

Moreover, $(Y_t, X_t)_{t \in \mathbb{R}_+}$ *is a two-factor affine process with infinitesimal generator*

$$\begin{aligned} (\mathcal{A}_{(Y,X)} f)(y, x) = (a - by) f_1'(y, x) &+ (\alpha - \beta y - \gamma x) f_2'(y, x) \\ &+ \frac{1}{2} y \big[\sigma_1^2 f_{1,1}''(y, x) + 2\varrho\sigma_1\sigma_2 f_{1,2}''(y, x) + \sigma_2^2 f_{2,2}''(y, x)\big] \\ &+ \frac{1}{2} \sigma_3^2 f_{2,2}''(y, x), \end{aligned} \tag{2.3}$$

where $(y, x) \in \mathbb{R}_+ \times \mathbb{R}$, $f \in \mathcal{C}_c^2(\mathbb{R}_+ \times \mathbb{R}, \mathbb{R})$, *and* f_i', $i \in \{1, 2\}$, *and* $f_{i,j}''$, $i, j \in \{1, 2\}$, *denote the first and second order partial derivatives of* f *with respect to its* i*-th and* i*-th and* j*-th variables.*

Conversely, every two-factor affine diffusion process is a (pathwise) unique strong solution of a SDE (1.1) *with suitable parameters* $a \in \mathbb{R}_+$, $b, \alpha, \beta, \gamma \in \mathbb{R}$, $\sigma_1, \sigma_2, \sigma_3 \in \mathbb{R}_+$ *and* $\varrho \in [-1, 1]$.

The next proposition gives the asymptotic behavior of the first moment of the process $(Y_t, X_t)_{t \in \mathbb{R}_+}$ as $t \to \infty$, see Bolyog and Pap (2016, Prop. 2.3).

Proposition 2.2 *Let us consider the two-factor affine diffusion model* (1.1) *with* $a \in \mathbb{R}_+$, $b, \alpha, \beta, \gamma \in \mathbb{R}$, $\sigma_1, \sigma_2, \sigma_3 \in \mathbb{R}_+$, $\varrho \in [-1, 1]$. *Suppose that* $\mathbb{E}(Y_0 |X_0|) < \infty$.

- (i) *If* $b, \gamma \in \mathbb{R}_{++}$ *then* $\lim_{t\to\infty} \mathbb{E}(Y_t) = \frac{a}{b}$ *and* $\lim_{t\to\infty} \mathbb{E}(X_t) = \frac{\alpha}{\gamma} - \frac{a\beta}{b\gamma}$.
- (ii) *If* $b \in \mathbb{R}_{++}$ *and* $\gamma = 0$ *then* $\lim_{t\to\infty} \mathbb{E}(Y_t) = \frac{a}{b}$ *and* $\lim_{t\to\infty} t^{-1}\mathbb{E}(X_t) = \alpha - \frac{a\beta}{b}$.
- (iii) *If* $b = 0$ *and* $\gamma \in \mathbb{R}_{++}$ *then* $\lim_{t\to\infty} t^{-1}\mathbb{E}(Y_t) = a$ *and* $\lim_{t\to\infty} t^{-1}\mathbb{E}(X_t) = -\frac{a\beta}{\gamma}$.
- (iv) *If* $b = \gamma = 0$ *then* $\lim_{t\to\infty} t^{-1}\mathbb{E}(Y_t) = a$ *and* $\lim_{t\to\infty} t^{-2}\mathbb{E}(X_t) = -\frac{1}{2}a\beta$.
- (v) *Otherwise, there exists* $c \in \mathbb{R}_{++}$ *such that* $\lim_{t\to\infty} \mathrm{e}^{-ct}\mathbb{E}(Y_t) \in \mathbb{R}$ *or* $\lim_{t\to\infty} \mathrm{e}^{-ct}\mathbb{E}(Y_t) \in \mathbb{R}$.

Based on the asymptotic behavior of the first moment of the process $(Y_t, X_t)_{t \in \mathbb{R}_+}$ as $t \to \infty$, we can classify two-factor affine diffusions in the following way.

Definition 2.3 Let $(Y_t, X_t)_{t \in \mathbb{R}_+}$ be the unique strong solution of the SDE (1.1) satisfying $\mathbb{P}(Y_0 \in \mathbb{R}_+) = 1$. We call $(Y_t, X_t)_{t \in \mathbb{R}_+}$ subcritical, critical or supercritical if $b \wedge \gamma \in \mathbb{R}_{++}$, $b \wedge \gamma = 0$ or $b \wedge \gamma \in \mathbb{R}_{--}$, respectively.

3 CLSE based on continuous time observations

Overbeck and Rydén (1997) investigated the CIR process Y, and for each $T \in \mathbb{R}_{++}$, they defined a CLSE $(\widehat{a}_T, \widehat{b}_T)$ of (a, b) based on continuous time observations $(Y_t)_{t \in [0,T]}$ as the limit in probability of the CLSE $(\widehat{a}_{T,n}, \widehat{b}_{T,n})$ of (a, b) based on discrete time observations $(Y_{\frac{iT}{n}})_{i \in \{0,1,\ldots,n\}}$ as $n \to \infty$.

We consider a two-factor affine diffusion process $(Y_t, X_t)_{t\in\mathbb{R}_+}$ given in (1.1) with known $\sigma_1 \in \mathbb{R}_{++}$, $\sigma_2, \sigma_3 \in \mathbb{R}_+$ and $\varrho \in [-1, 1]$, and with a random initial value (η_0, ζ_0) independent of $(W_t, B_t, L_t)_{t\in\mathbb{R}_+}$ satisfying $\mathbb{P}(\eta_0 \in \mathbb{R}_+) = 1$, and we will consider $\boldsymbol{\theta} = (a, b, \alpha, \beta, \gamma)^\top \in \mathbb{R}_+ \times \mathbb{R}^4$ as a parameter. The aim of the following discussion is to construct a CLSE of $\boldsymbol{\theta}$ based on continuous time observations $(Y_t, X_t)_{t\in[0,T]}$ with some $T \in \mathbb{R}_{++}$.

Let us recall the CLSE $\widehat{\boldsymbol{\theta}}_{T,n}$ of $\boldsymbol{\theta}$ based on discrete time observations $(Y_{\frac{i}{n}}, X_{\frac{i}{n}})_{i\in\{0,1,\ldots,\lfloor nT\rfloor\}}$ with some $n \in \mathbb{N}$, which can be obtained by solving the extremum problem

$$\widehat{\boldsymbol{\theta}}_{T,n} := \operatorname*{arg\,min}_{\boldsymbol{\theta}\in\mathbb{R}^5} \sum_{i=1}^{\lfloor nT\rfloor} \left[\left(Y_{\frac{i}{n}} - \mathbb{E}\Big(Y_{\frac{i}{n}} \,\Big|\, \mathcal{F}_{\frac{i-1}{n}}\Big)\right)^2 + \left(X_{\frac{i}{n}} - \mathbb{E}\Big(X_{\frac{i}{n}} \,\Big|\, \mathcal{F}_{\frac{i-1}{n}}\Big)\right)^2\right].$$

By (2.1) and (2.2), together with Proposition 3.2.10 in Karatzas and Shreve (1991), for all $i \in \mathbb{N}$, we obtain

$$\mathbb{E}\Big(Y_{\frac{i}{n}} \,|\, \mathcal{F}_{\frac{i-1}{n}}\Big) = \mathrm{e}^{-\frac{b}{n}} Y_{\frac{i-1}{n}} + a \int_0^{\frac{1}{n}} \mathrm{e}^{-bw}\,\mathrm{d}w$$

and

$$\begin{aligned}\mathbb{E}\Big(X_{\frac{i}{n}} \,|\, \mathcal{F}_{\frac{i-1}{n}}\Big) &= \mathrm{e}^{-\frac{\gamma}{n}} X_{\frac{i-1}{n}} + \alpha \int_0^{\frac{1}{n}} \mathrm{e}^{-\gamma w}\,\mathrm{d}w - \beta Y_{\frac{i-1}{n}} \int_0^{\frac{1}{n}} \mathrm{e}^{(\gamma-b)w-\frac{\gamma}{n}}\,\mathrm{d}w\\ &\quad - a\beta \int_0^{\frac{1}{n}} \mathrm{e}^{\gamma w-\frac{\gamma}{n}} \left(\int_0^w \mathrm{e}^{-b(w-v)}\,\mathrm{d}v\right)\mathrm{d}w.\end{aligned}$$

Consequently,

$$\begin{aligned}\widehat{\boldsymbol{\theta}}_{T,n} = \operatorname*{arg\,min}_{(a,b,\alpha,\beta,\gamma)^\top\in\mathbb{R}^5} \sum_{i=1}^{\lfloor nT\rfloor} \Bigg[&\left(Y_{\frac{i}{n}} - Y_{\frac{i-1}{n}} - \Big(c - dY_{\frac{i-1}{n}}\Big)\right)^2 \\ &+ \left(X_{\frac{i}{n}} - X_{\frac{i-1}{n}} - \Big(\delta - \varepsilon Y_{\frac{i-1}{n}} - \zeta X_{\frac{i-1}{n}}\Big)\right)^2\Bigg], \end{aligned} \tag{3.1}$$

where

$$(c, d, \delta, \varepsilon, \zeta) := (c_n(a, b), d_n(b), \delta_n(a, b, \alpha, \beta, \gamma), \varepsilon_n(b, \beta, \gamma), \zeta_n(\gamma)) := g_n(a, b, \alpha, \beta, \gamma) \tag{3.2}$$

with

$$\begin{aligned} &c := c_n(a, b) := a \int_0^{\frac{1}{n}} \mathrm{e}^{-bw}\,\mathrm{d}w, \qquad d := d_n(b) := 1 - \mathrm{e}^{-\frac{b}{n}},\\ &\delta := \delta_n(a, b, \alpha, \beta, \gamma) := \alpha \int_0^{\frac{1}{n}} \mathrm{e}^{-\gamma w}\,\mathrm{d}w - a\beta \int_0^{\frac{1}{n}} \mathrm{e}^{\gamma w-\frac{\gamma}{n}} \left(\int_0^w \mathrm{e}^{-b(w-v)}\,\mathrm{d}v\right)\mathrm{d}w,\\ &\varepsilon := \varepsilon_n(b, \beta, \gamma) := \beta \int_0^{\frac{1}{n}} \mathrm{e}^{(\gamma-b)w-\frac{\gamma}{n}}\,\mathrm{d}w, \qquad \zeta := \zeta_n(\gamma) := 1 - \mathrm{e}^{-\frac{\gamma}{n}}.\end{aligned}$$

Since the function $g_n : \mathbb{R}^5 \to \mathbb{R} \times (-\infty, 1) \times \mathbb{R}^2 \times (-\infty, 1)$ is bijective, first we determine the CLSE $(\widehat{c}_{T,n}, \widehat{d}_{T,n}, \widehat{\delta}_{T,n}, \widehat{\varepsilon}_{T,n}, \widehat{\zeta}_{T,n})$ of the transformed parameters $(c, d, \delta, \varepsilon, \zeta)$ by minimizing the sum on the right-hand side of (3.1) with respect to $(c, d, \delta, \varepsilon, \zeta)$. We have

$$\left(\widehat{c}_{T,n}, \widehat{d}_{T,n}\right) = \underset{(c,d)^\top \in \mathbb{R}^2}{\arg\min} \sum_{i=1}^{\lfloor nT \rfloor} \left(Y_{\frac{i}{n}} - Y_{\frac{i-1}{n}} - \left(c - dY_{\frac{i-1}{n}}\right)\right)^2,$$

$$\left(\widehat{\delta}_{T,n}, \widehat{\varepsilon}_{T,n}, \widehat{\zeta}_{T,n}\right) = \underset{(\delta,\varepsilon,\zeta)^\top \in \mathbb{R}^3}{\arg\min} \sum_{i=1}^{\lfloor nT \rfloor} \left(X_{\frac{i}{n}} - X_{\frac{i-1}{n}} - \left(\delta - \varepsilon Y_{\frac{i-1}{n}} - \zeta X_{\frac{i-1}{n}}\right)\right)^2,$$

hence, similarly as on page 675 in Barczy et al. (2013), we get

$$\begin{bmatrix} \widehat{c}_{T,n} \\ \widehat{d}_{T,n} \end{bmatrix} = \left(\boldsymbol{\Gamma}_{T,n}^{(1)}\right)^{-1} \boldsymbol{\varphi}_{T,n}^{(1)}, \qquad \begin{bmatrix} \widehat{\delta}_{T,n} \\ \widehat{\varepsilon}_{T,n} \\ \widehat{\zeta}_{T,n} \end{bmatrix} = \left(\boldsymbol{\Gamma}_{T,n}^{(2)}\right)^{-1} \boldsymbol{\varphi}_{T,n}^{(2)} \tag{3.3}$$

with

$$\boldsymbol{\Gamma}_{T,n}^{(1)} := \begin{bmatrix} \lfloor nT \rfloor & -\sum_{i=1}^{\lfloor nT \rfloor} Y_{\frac{i-1}{n}} \\ -\sum_{i=1}^{\lfloor nT \rfloor} Y_{\frac{i-1}{n}} & \sum_{i=1}^{\lfloor nT \rfloor} Y_{\frac{i-1}{n}}^2 \end{bmatrix}, \qquad \boldsymbol{\varphi}_{T,n}^{(1)} := \begin{bmatrix} Y_{\frac{\lfloor nT \rfloor}{n}} - Y_0 \\ -\sum_{i=1}^{\lfloor nT \rfloor} \left(Y_{\frac{i}{n}} - Y_{\frac{i-1}{n}}\right) Y_{\frac{i-1}{n}} \end{bmatrix},$$

$$\boldsymbol{\Gamma}_{T,n}^{(2)} := \begin{bmatrix} \lfloor nT \rfloor & -\sum_{i=1}^{\lfloor nT \rfloor} Y_{\frac{i-1}{n}} & -\sum_{i=1}^{\lfloor nT \rfloor} X_{\frac{i-1}{n}} \\ -\sum_{i=1}^{\lfloor nT \rfloor} Y_{\frac{i-1}{n}} & \sum_{i=1}^{\lfloor nT \rfloor} Y_{\frac{i-1}{n}}^2 & \sum_{i=1}^{\lfloor nT \rfloor} Y_{\frac{i-1}{n}} X_{\frac{i-1}{n}} \\ -\sum_{i=1}^{\lfloor nT \rfloor} X_{\frac{i-1}{n}} & \sum_{i=1}^{\lfloor nT \rfloor} Y_{\frac{i-1}{n}} X_{\frac{i-1}{n}} & \sum_{i=1}^{\lfloor nT \rfloor} X_{\frac{i-1}{n}}^2 \end{bmatrix},$$

$$\boldsymbol{\varphi}_{T,n}^{(2)} := \begin{bmatrix} X_{\frac{\lfloor nT \rfloor}{n}} - X_0 \\ -\sum_{i=1}^{\lfloor nT \rfloor} \left(X_{\frac{i}{n}} - X_{\frac{i-1}{n}}\right) Y_{\frac{i-1}{n}} \\ -\sum_{i=1}^{\lfloor nT \rfloor} \left(X_{\frac{i}{n}} - X_{\frac{i-1}{n}}\right) X_{\frac{i-1}{n}} \end{bmatrix}$$

on the event where the random matrices $\boldsymbol{\Gamma}_{T,n}^{(1)}$ and $\boldsymbol{\Gamma}_{T,n}^{(2)}$ are invertible.

Lemma 3.1 *Let us consider the two-factor affine diffusion model* (1.1) *with* $a \in \mathbb{R}_+$, $b, \alpha, \beta, \gamma \in \mathbb{R}$, $\sigma_1 \in \mathbb{R}_{++}$, $\sigma_2, \sigma_3 \in \mathbb{R}_+$ *and* $\varrho \in [-1, 1]$ *with a random initial value* (η_0, ζ_0) *independent of* $(W_t, B_t, L_t)_{t \in \mathbb{R}_+}$ *satisfying* $\mathbb{P}(\eta_0 \in \mathbb{R}_+) = 1$. *Suppose that* $(1 - \varrho^2)\sigma_2^2 + \sigma_3^2 > 0$. *Then for each* $T \in \mathbb{R}_{++}$ *and* $n \in \mathbb{N}$, *the random matrices* $\boldsymbol{\Gamma}_{T,n}^{(1)}$ *and* $\boldsymbol{\Gamma}_{T,n}^{(2)}$ *are invertible almost surely, and hence there exists a unique CLSE* $\left(\widehat{c}_{T,n}, \widehat{d}_{T,n}, \widehat{\delta}_{T,n}, \widehat{\varepsilon}_{T,n}, \widehat{\zeta}_{T,n}\right)$ *of* $(c, d, \delta, \varepsilon, \zeta)$ *taking the form given in* (3.3).

A proof can be found in the Arxiv version of this paper Bolyog and Pap (2017).

Remark 3.2 The first order Taylor approximation of $g_n(a, b, \alpha, \beta, \gamma)$ at $(0, 0, 0, 0, 0)$ is $\frac{1}{n}(a, b, \alpha, \beta, \gamma)$, hence we obtain the first order Taylor approximations

$$Y_{\frac{i}{n}} - \mathbb{E}\left(Y_{\frac{i}{n}} \mid \mathcal{F}_{\frac{i-1}{n}}\right) \approx Y_{\frac{i}{n}} - Y_{\frac{i-1}{n}} - \frac{1}{n}\left(a - bY_{\frac{i-1}{n}}\right),$$

$$X_{\frac{i}{n}} - \mathbb{E}\left(X_{\frac{i}{n}} \mid \mathcal{F}_{\frac{i-1}{n}}\right) \approx X_{\frac{i}{n}} - X_{\frac{i-1}{n}} - \frac{1}{n}\left(\alpha - \beta Y_{\frac{i-1}{n}} - \gamma X_{\frac{i-1}{n}}\right).$$

Using these approximations, one can define an approximate CLSE $\widehat{\boldsymbol{\theta}}_{T,n}^{\text{approx}}$ of $\boldsymbol{\theta}$ based on discrete time observations $(Y_i, X_i)_{i\in\{0,1,\ldots,\lfloor nT\rfloor\}}$, $n \in \mathbb{N}$, by solving the extremum problem

$$\widehat{\boldsymbol{\theta}}_{T,n}^{\text{approx}} := \operatorname*{arg\,min}_{(a,b,\alpha,\beta,\gamma)^\top\in\mathbb{R}^5} \sum_{i=1}^{\lfloor nT\rfloor} \Bigg[\Big(Y_{\frac{i}{n}} - Y_{\frac{i-1}{n}} - \frac{1}{n}\Big(a - bY_{\frac{i-1}{n}}\Big)\Big)^2 + \Big(X_{\frac{i}{n}} - X_{\frac{i-1}{n}} - \frac{1}{n}\Big(\alpha - \beta Y_{\frac{i-1}{n}} - \gamma X_{\frac{i-1}{n}}\Big)\Big)^2\Bigg],$$

hence $\widehat{\boldsymbol{\theta}}_{T,n}^{\text{approx}} = n\big(\widehat{c}_{T,n}, \widehat{d}_{T,n}, \widehat{\delta}_{T,n}, \widehat{\varepsilon}_{T,n}, \widehat{\zeta}_{T,n}\big)^\top$. This definition of approximate CLSE can be considered as the definition of LSE given in Hu and Long (2009a, formula (1.2)) for generalized Ornstein–Uhlenbeck processes driven by α-stable motions, see also Hu and Long (2009b, formula (3.1)). For a heuristic motivation of the estimator $\widehat{\boldsymbol{\theta}}_{n}^{\text{approx}}$ based on discrete observations, see, e.g., Hu and Long (2007, p. 178) (formulated for Langevin equations). □

We have

$$\frac{1}{n}\boldsymbol{\Gamma}_{T,n}^{(1)} \xrightarrow{\text{a.s.}} \begin{bmatrix} T & -\int_0^T Y_s\,\mathrm{d}s \\ -\int_0^T Y_s\,\mathrm{d}s & \int_0^T Y_s^2\,\mathrm{d}s \end{bmatrix} =: \boldsymbol{G}_T^{(1)},$$

$$\frac{1}{n}\boldsymbol{\Gamma}_{T,n}^{(2)} \xrightarrow{\text{a.s.}} \begin{bmatrix} T & -\int_0^T Y_s\,\mathrm{d}s & -\int_0^T X_s\,\mathrm{d}s \\ -\int_0^T Y_s\,\mathrm{d}s & \int_0^T {Y_s}^2\,\mathrm{d}s & \int_0^T X_sY_s\,\mathrm{d}s \\ -\int_0^T X_s\,\mathrm{d}s & \int_0^T X_sY_s\,\mathrm{d}s & \int_0^T X_s^2\,\mathrm{d}s \end{bmatrix} =: \boldsymbol{G}_T^{(2)}$$

as $n \to \infty$, since $(Y_t, X_t)_{t\in\mathbb{R}_+}$ is almost surely continuous. By Proposition I.4.44 in Jacod and Shiryaev (2003) with the Riemann sequence of deterministic subdivisions $\big(\frac{i}{n} \wedge T\big)_{i\in\mathbb{N}}$, $n \in \mathbb{N}$., we obtain

$$\boldsymbol{\varphi}_{T,n}^{(1)} \xrightarrow{\mathbb{P}} \begin{bmatrix} Y_T - Y_0 \\ -\int_0^T Y_s\,\mathrm{d}Y_s \end{bmatrix} =: \boldsymbol{f}_T^{(1)}, \qquad \boldsymbol{\varphi}_{T,n}^{(2)} \xrightarrow{\mathbb{P}} \begin{bmatrix} X_T - X_0 \\ -\int_0^T Y_s\,\mathrm{d}X_s \\ -\int_0^T X_s\,\mathrm{d}X_s \end{bmatrix} =: \boldsymbol{f}_T^{(2)},$$

as $n \to \infty$. By Slutsky's lemma, using also Lemma 3.1, we conclude

$$\widehat{\boldsymbol{\theta}}_{T,n}^{\text{approx}} = n \begin{bmatrix} \widehat{c}_{T,n} \\ \widehat{d}_{T,n} \\ \widehat{\delta}_{T,n} \\ \widehat{\varepsilon}_{T,n} \\ \widehat{\zeta}_{T,n} \end{bmatrix} \xrightarrow{\mathbb{P}} \begin{bmatrix} (\boldsymbol{G}_T^{(1)})^{-1}\boldsymbol{f}_T^{(1)} \\ (\boldsymbol{G}_T^{(2)})^{-1}\boldsymbol{f}_T^{(2)} \end{bmatrix} =: \begin{bmatrix} \widehat{a}_T \\ \widehat{b}_T \\ \widehat{\alpha}_T \\ \widehat{\beta}_T \\ \widehat{\gamma}_T \end{bmatrix} =: \widehat{\boldsymbol{\theta}}_T \qquad \text{as } n \to \infty, \tag{3.4}$$

whenever the random matrices $\boldsymbol{G}_T^{(1)}$ and $\boldsymbol{G}_T^{(2)}$ are invertible.

Lemma 3.3 *Let us consider the two-factor affine diffusion model (1.1) with* $a \in \mathbb{R}_+$, $b, \alpha, \beta, \gamma \in \mathbb{R}$, $\sigma_1 \in \mathbb{R}_{++}$, $\sigma_2, \sigma_3 \in \mathbb{R}_+$ *and* $\varrho \in [-1, 1]$ *with a random initial value* (η_0, ζ_0) *independent of* $(W_t, B_t, L_t)_{t\in\mathbb{R}_+}$ *satisfying* $\mathbb{P}(\eta_0 \in \mathbb{R}_+) = 1$. *Suppose that* $(1 - \varrho^2)\sigma_2^2 + \sigma_3^2 > 0$. *Then for each* $T \in \mathbb{R}_{++}$, *the random matrices* $\boldsymbol{G}_T^{(1)}$ *and* $\boldsymbol{G}_T^{(2)}$ *are invertible almost surely, and hence* $\widehat{\boldsymbol{\theta}}_T$ *given in (3.4) exists almost surely. Moreover,* $\widehat{\boldsymbol{\theta}}_{T,n} \xrightarrow{\mathbb{P}} \widehat{\boldsymbol{\theta}}_T$ *as* $n \to \infty$.

Proof A proof of the first statement can be found in the Arxiv version of this paper Bolyog and Pap (2017). Next we are going to show $\widehat{\boldsymbol{\theta}}_{T,n} \xrightarrow{\mathbb{P}} \widehat{\boldsymbol{\theta}}_T$ as $n \to \infty$. The function g_n introduced in (3.2) admits an inverse $g_n^{-1} : \mathbb{R} \times (-\infty, 1) \times \mathbb{R}^2 \times (-\infty, 1) \to \mathbb{R}^5$ satisfying

$$g_n^{-1}(c, d, \delta, \varepsilon, \zeta) = (a, b, \alpha, \beta, \gamma)$$

with

$$b = -n \log(1-d), \qquad a = \frac{c}{\int_0^{\frac{1}{n}} \mathrm{e}^{-bw} \, \mathrm{d}w}, \qquad \gamma = -n \log(1-\zeta),$$

$$\beta = \frac{\varepsilon}{\int_0^{\frac{1}{n}} \mathrm{e}^{(\gamma-b)w-\frac{\gamma}{n}} \, \mathrm{d}w}, \qquad \alpha = \frac{\delta + a\beta \int_0^{\frac{1}{n}} \mathrm{e}^{\gamma w - \frac{\gamma}{n}} \big(\int_0^w \mathrm{e}^{-b(w-v)} \, \mathrm{d}v\big) \mathrm{d}w}{\int_0^{\frac{1}{n}} \mathrm{e}^{-\gamma w} \, \mathrm{d}w}.$$

Convergence (3.4) yields $(\widehat{c}_{T,n}, \widehat{d}_{T,n}, \widehat{\delta}_{T,n}, \widehat{\varepsilon}_{T,n}, \widehat{\zeta}_{T,n}) \xrightarrow{\mathbb{P}} \mathbf{0}$ as $n \to \infty$, hence $\widehat{d}_{T,n} \in (-\infty, 1)$ and $\widehat{\zeta}_{T,n} \in (-\infty, 1)$ with probability tending to one as $n \to \infty$. Consequently, $g_n^{-1}(\widehat{c}_{T,n}, \widehat{d}_{T,n}, \widehat{\delta}_{T,n}, \widehat{\varepsilon}_{T,n}, \widehat{\zeta}_{T,n}) = \widehat{\boldsymbol{\theta}}_{T,n}$ with probability tending to one as $n \to \infty$. We have

$$\widehat{b}_{T,n} = -n \log(1 - \widehat{d}_{T,n}) = n\widehat{d}_{T,n} h_1(\widehat{d}_{T,n})$$

with probability tending to one as $n \to \infty$, where the continuous function $h_1 : (-\infty, 1) \to \mathbb{R}$ is given by

$$h_1(x) := \begin{cases} -\frac{1}{x} \log(1-x) & \text{if } x \neq 0, \\ 1 & \text{if } x = 0. \end{cases}$$

By (3.4), we have $n\widehat{d}_{T,n} \xrightarrow{\mathbb{P}} \widehat{b}_T$ and $\widehat{d}_{T,n} \xrightarrow{\mathbb{P}} 0$, thus we obtain $h_1(\widehat{d}_{T,n}) \xrightarrow{\mathbb{P}} h_1(0) = 1$, and hence $\widehat{b}_{T,n} \xrightarrow{\mathbb{P}} \widehat{b}_T$ as $n \to \infty$.

Moreover,

$$\widehat{a}_{T,n} = \frac{\widehat{c}_{T,n}}{\int_0^{\frac{1}{n}} \mathrm{e}^{-\widehat{b}_{T,n} w} \, \mathrm{d}w} = \frac{n\widehat{c}_{T,n}}{n \int_0^{\frac{1}{n}} \mathrm{e}^{-\widehat{b}_{T,n} w} \, \mathrm{d}w} = \frac{n\widehat{c}_{T,n}}{\int_0^1 \exp\big\{-n^{-1}\widehat{b}_{T,n} v\big\} \, \mathrm{d}v} = \frac{n\widehat{c}_{T,n}}{h_2(n^{-1}\widehat{b}_{T,n})}$$

with probability tending to one as $n \to \infty$, where the continuous function $h_2 : \mathbb{R} \to \mathbb{R}$ is given by

$$h_2(x) := \int_0^1 \mathrm{e}^{-xv} \, \mathrm{d}v = \begin{cases} \frac{1-\mathrm{e}^{-x}}{x} & \text{if } x \neq 0, \\ 1 & \text{if } x = 0. \end{cases}$$

We have already showed $\widehat{b}_{T,n} \xrightarrow{\mathbb{P}} \widehat{b}_T$, yielding $n^{-1}\widehat{b}_{T,n} \xrightarrow{\mathbb{P}} 0$, and hence $h_2(n^{-1}\widehat{b}_{T,n}) \xrightarrow{\mathbb{P}} h_2(0) = 1$ as $n \to \infty$. By (3.4), we have $n\widehat{c}_{T,n} \xrightarrow{\mathbb{P}} \widehat{a}_T$, thus we obtain $\widehat{a}_{T,n} \xrightarrow{\mathbb{P}} \widehat{a}_T$ as $n \to \infty$.

In a similar way,

$$\widehat{\gamma}_{T,n} = -n \log(1 - \widehat{\zeta}_{T,n}) = n\widehat{\zeta}_{T,n} h_1(\widehat{\zeta}_{T,n})$$

with probability tending to one as $n \to \infty$. By (3.4), we have $n\widehat{\zeta}_{T,n} \xrightarrow{\mathbb{P}} \widehat{\gamma}_T$ and $\widehat{\zeta}_{T,n} \xrightarrow{\mathbb{P}} 0$, thus we obtain $h_1(\widehat{\zeta}_{T,n}) \xrightarrow{\mathbb{P}} h_1(0) = 1$, and hence $\widehat{\gamma}_{T,n} \xrightarrow{\mathbb{P}} \widehat{\gamma}_T$ as $n \to \infty$.

Further,

$$\widehat{\beta}_{T,n} = \frac{\widehat{\varepsilon}_{T,n}}{\int_0^{\frac{1}{n}} \mathrm{e}^{(\widehat{\gamma}_{T,n} - \widehat{b}_{T,n})w - \frac{\widehat{\gamma}_{T,n}}{n}} \, \mathrm{d}w} = \frac{n\widehat{\varepsilon}_{T,n} \mathrm{e}^{\frac{\widehat{\gamma}_{T,n}}{n}}}{h_2(n^{-1}(\widehat{b}_{T,n} - \widehat{\gamma}_{T,n}))}$$

with probability tending to one as $n \to \infty$. We have already showed $\widehat{b}_{T,n} \stackrel{\mathbb{P}}{\longrightarrow} \widehat{b}_T$ and $\widehat{\gamma}_{T,n} \stackrel{\mathbb{P}}{\longrightarrow} \widehat{\gamma}_T$, yielding $n^{-1}\widehat{b}_{T,n} \stackrel{\mathbb{P}}{\longrightarrow} 0$ and $n^{-1}\widehat{\gamma}_{T,n} \stackrel{\mathbb{P}}{\longrightarrow} 0$, and hence $\mathrm{e}^{\frac{\widehat{\gamma}_{T,n}}{n}} \stackrel{\mathbb{P}}{\longrightarrow} 1$ and $h_2(n^{-1}(\widehat{b}_{T,n} - \widehat{\gamma}_{T,n})) \stackrel{\mathbb{P}}{\longrightarrow} h_2(0) = 1$ as $n \to \infty$. By (3.4), we have $n\widehat{\varepsilon}_{T,n} \stackrel{\mathbb{P}}{\longrightarrow} \widehat{\beta}_T$, thus we obtain $\widehat{\beta}_{T,n} \stackrel{\mathbb{P}}{\longrightarrow} \widehat{\beta}_T$ as $n \to \infty$.

Finally,

$$\widehat{\alpha}_{T,n} = \frac{\widehat{\delta}_{T,n} + \widehat{a}_{T,n}\widehat{\beta}_{T,n} \int_0^{\frac{1}{n}} \mathrm{e}^{\widehat{\gamma}_{T,n}w - \frac{\widehat{\gamma}_{T,n}}{n}} \left(\int_0^w \mathrm{e}^{-b(w-v)} \, \mathrm{d}v\right) \mathrm{d}w}{\int_0^{\frac{1}{n}} \mathrm{e}^{-\widehat{\gamma}_{T,n}w} \, \mathrm{d}w} = \frac{n\widehat{\delta}_{T,n} + \widehat{a}_{T,n}\widehat{\beta}_{T,n}\mathrm{e}^{-\frac{\widehat{\gamma}_{T,n}}{n}} I_{T,n}}{h_2(n^{-1}\widehat{\gamma}_{T,n})}$$

with probability tending to one as $n \to \infty$, where

$$I_{T,n} = n \int_0^{\frac{1}{n}} \mathrm{e}^{\widehat{\gamma}_{T,n}w} \left(\int_0^w \mathrm{e}^{-b(w-v)} \, \mathrm{d}v\right) \mathrm{d}w \leqslant \frac{1}{n} \mathrm{e}^{\frac{|\widehat{\gamma}_{T,n}|}{n}} \mathrm{e}^{\frac{|\widehat{b}_{T,n}|}{n}}.$$

We have already showed $\widehat{a}_{T,n} \stackrel{\mathbb{P}}{\longrightarrow} \widehat{a}_T$, $\widehat{b}_{T,n} \stackrel{\mathbb{P}}{\longrightarrow} \widehat{b}_T$, $\widehat{\beta}_{T,n} \stackrel{\mathbb{P}}{\longrightarrow} \widehat{\beta}_T$ and $\widehat{\gamma}_{T,n} \stackrel{\mathbb{P}}{\longrightarrow} \widehat{\gamma}_T$, yielding $n^{-1}\widehat{b}_{T,n} \stackrel{\mathbb{P}}{\longrightarrow} 0$ and $n^{-1}\widehat{\gamma}_{T,n} \stackrel{\mathbb{P}}{\longrightarrow} 0$, and hence $h_2(n^{-1}\widehat{\gamma}_{T,n}) \stackrel{\mathbb{P}}{\longrightarrow} h_2(0) = 1$, $\mathrm{e}^{\frac{\widehat{\gamma}_{T,n}}{n}} \stackrel{\mathbb{P}}{\longrightarrow} 1$, $\mathrm{e}^{\frac{|\widehat{\gamma}_{T,n}|}{n}} \stackrel{\mathbb{P}}{\longrightarrow} 1$ and $\mathrm{e}^{\frac{|\widehat{b}_{T,n}|}{n}} \stackrel{\mathbb{P}}{\longrightarrow} 1$, implying $I_{T,n} \stackrel{\mathbb{P}}{\longrightarrow} 0$ as $n \to \infty$. By (3.4), we have $n\widehat{\delta}_{T,n} \stackrel{\mathbb{P}}{\longrightarrow} \widehat{\alpha}_T$, thus we obtain $\widehat{\alpha}_{T,n} \stackrel{\mathbb{P}}{\longrightarrow} \widehat{\alpha}_T$ as $n \to \infty$. □

Using the SDE (1.1) and Corollary 3.2.20 in Karatzas and Shreve (1991), one can check that

$$\widehat{\boldsymbol{\theta}}_T - \boldsymbol{\theta} = \begin{bmatrix} \widehat{a}_T - a \\ \widehat{b}_T - b \\ \widehat{\alpha}_T - \alpha \\ \widehat{\beta}_T - \beta \\ \widehat{\gamma}_T - \gamma \end{bmatrix} = \begin{bmatrix} (\boldsymbol{G}_T^{(1)})^{-1}\boldsymbol{h}_T^{(1)} \\ (\boldsymbol{G}_T^{(2)})^{-1}\boldsymbol{h}_T^{(2)} \end{bmatrix} = \boldsymbol{G}_T^{-1}\boldsymbol{h}_T \tag{3.5}$$

on the event where the random matrices $\boldsymbol{G}_T^{(1)}$ and $\boldsymbol{G}_T^{(2)}$ are invertible, where

$$\boldsymbol{G}_T := \begin{bmatrix} \boldsymbol{G}_T^{(1)} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{G}_T^{(2)} \end{bmatrix}, \qquad \boldsymbol{h}_T := \begin{bmatrix} \boldsymbol{h}_T^{(1)} \\ \boldsymbol{h}_T^{(2)} \end{bmatrix},$$

with

$$\boldsymbol{h}_T^{(1)} := \sigma_1 \int_0^T \sqrt{Y_s} \begin{bmatrix} 1 \\ -Y_s \end{bmatrix} \mathrm{d}W_s, \qquad \boldsymbol{h}_T^{(2)} := \int_0^T \begin{bmatrix} 1 \\ -Y_s \\ -X_s \end{bmatrix} (\sigma_2\sqrt{Y_s}\,\mathrm{d}\widetilde{W}_s + \sigma_3\,\mathrm{d}L_s),$$

where

$$\widetilde{W}_s := \varrho W_s + \sqrt{1-\varrho^2} B_s, \qquad s \in \mathbb{R}_+, \tag{3.6}$$

is a standard Wiener process, independent of L. For details see the Arxiv version of this paper Bolyog and Pap (2017).

4 Consistency of CLSE

First we consider the case of subcritical Heston models, i.e., when $b \in \mathbb{R}_{++}$.

Theorem 4.1 *Let us consider the two-factor affine diffusion model* (1.1) *with* $a, b \in \mathbb{R}_{++}$, $\alpha, \beta \in \mathbb{R}$, $\gamma \in \mathbb{R}_{++}$, $\sigma_1 \in \mathbb{R}_{++}$, $\sigma_2, \sigma_3 \in \mathbb{R}_+$ *and* $\varrho \in [-1, 1]$ *with a random initial value* (η_0, ζ_0) *independent of* $(W_t, B_t, L_t)_{t \in \mathbb{R}_+}$ *satisfying* $\mathbb{P}(\eta_0 \in \mathbb{R}_+) = 1$. *Suppose that* $(1-\varrho^2)\sigma_2^2 + \sigma_3^2 > 0$. *Then the CLSE of* $\boldsymbol{\theta} = (a, b, \alpha, \beta, \gamma)^\top$ *is strongly consistent, i.e.,* $\widehat{\boldsymbol{\theta}}_T = \big(\widehat{a}_T, \widehat{b}_T, \widehat{\alpha}_T, \widehat{\beta}_T, \widehat{\gamma}_T\big)^\top \xrightarrow{\text{a.s.}} \boldsymbol{\theta} = (a, b, \alpha, \beta, \gamma)^\top$ *as* $T \to \infty$.

Proof By (3.5), we have

$$\widehat{\boldsymbol{\theta}}_T - \boldsymbol{\theta} = (T^{-1}\boldsymbol{G}_T)^{-1}(T^{-1}\boldsymbol{h}_T) \tag{4.1}$$

on the event, where the random matrix $\boldsymbol{G}_T$ is invertible, which has propapility 1, see Lemma 3.3.

By Theorem A.2, we obtain

$$T^{-1}\boldsymbol{G}_T \xrightarrow{\text{a.s.}} \mathbb{E}(\boldsymbol{G}_\infty) \qquad \text{as } T \to \infty, \tag{4.2}$$

where

$$\boldsymbol{G}_\infty := \begin{bmatrix} \boldsymbol{G}_\infty^{(1)} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{G}_\infty^{(2)} \end{bmatrix} \tag{4.3}$$

with

$$\boldsymbol{G}_\infty^{(1)} := \begin{bmatrix} 1 & -Y_\infty \\ -Y_\infty & Y_\infty^2 \end{bmatrix}, \qquad \boldsymbol{G}_\infty^{(2)} := \begin{bmatrix} 1 & -Y_\infty & -X_\infty \\ -Y_\infty & Y_\infty^2 & Y_\infty X_\infty \\ -X_\infty & Y_\infty X_\infty & X_\infty^2 \end{bmatrix},$$

where the random vector (Y_∞, X_∞) is given by Theorem A.1, since, by Theorem B.2, the entries of $\mathbb{E}(\boldsymbol{G}_\infty)$ exist and finite.

The matrix $\mathbb{E}(\boldsymbol{G}_\infty^{(1)})$ is strictly positive definite, since for all $\boldsymbol{x} \in \mathbb{R}^2 \setminus \{\boldsymbol{0}\}$, we have $\boldsymbol{x}^\top \mathbb{E}(\boldsymbol{G}_\infty^{(1)})\boldsymbol{x} > 0$. Indeed, for all $\boldsymbol{x} = (x_1, x_2)^\top \in \mathbb{R}^2 \setminus \{\boldsymbol{0}\}$,

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}^\top \mathbb{E}(\boldsymbol{G}_\infty^{(1)}) \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbb{E}\big[(x_1 - x_2 Y_\infty)^2\big] > 0,$$

since, by Theorem A.2, the distribution of Y_∞ is absolutely continuous, hence $x_1 - x_2 Y_\infty \neq 0$ with probability 1. In a similar way, the matrix $\mathbb{E}(\boldsymbol{G}_\infty^{(2)})$ is strictly positive definite, since for all $\boldsymbol{x} \in \mathbb{R}^3 \setminus \{\boldsymbol{0}\}$, we have $\boldsymbol{x}^\top \mathbb{E}(\boldsymbol{G}_\infty^{(2)})\boldsymbol{x} > 0$. Indeed, for all $\boldsymbol{x} = (x_1, x_2, x_3)^\top \in \mathbb{R}^3 \setminus \{\boldsymbol{0}\}$,

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}^\top \mathbb{E}(\boldsymbol{G}_\infty^{(2)}) \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \mathbb{E}\big[(x_1 - x_2 Y_\infty - x_3 Z_\infty)^2\big] > 0,$$

since, by Theorem A.2, the distribution of (Y_∞, X_∞) is absolutely continuous, hence $x_1 - x_2 Y_\infty - x_3 X_\infty \neq 0$ with probability 1. Thus the matrices $\mathbb{E}(\boldsymbol{G}_\infty^{(1)})$ and $\mathbb{E}(\boldsymbol{G}_\infty^{(2)})$ are invertible, whence we conclude

$$(T^{-1}\boldsymbol{G}_T)^{-1} \xrightarrow{\text{a.s.}} \begin{bmatrix} [\mathbb{E}(\boldsymbol{G}_\infty^{(1)})]^{-1} & \boldsymbol{0} \\ \boldsymbol{0} & [\mathbb{E}(\boldsymbol{G}_\infty^{(2)})]^{-1} \end{bmatrix} = [\mathbb{E}(\boldsymbol{G}_\infty)]^{-1} \qquad \text{as } T \to \infty. \tag{4.4}$$

The aim of the next discussion is to show convergence

$$T^{-1}\boldsymbol{h}_T \xrightarrow{\text{a.s.}} \boldsymbol{0} \qquad \text{as} \quad T \to \infty. \tag{4.5}$$

We have

$$\frac{1}{T}\int_0^T \sqrt{Y_s}\,\mathrm{d}W_s = \frac{1}{T}\int_0^T Y_s\,\mathrm{d}s \cdot \frac{\int_0^T \sqrt{Y_s}\,\mathrm{d}W_s}{\int_0^T Y_s\,\mathrm{d}s} \xrightarrow{\text{a.s.}} 0 \qquad \text{as} \quad T \to \infty.$$

Indeed, we have already proved

$$\frac{1}{T}\int_0^T Y_s\,\mathrm{d}s \xrightarrow{\text{a.s.}} \mathbb{E}(Y_\infty) = \frac{a}{b} \in \mathbb{R}_{++} \qquad \text{as} \quad T \to \infty,$$

and the strong law of large numbers for continuous local martingales (see, e.g., Theorem C.1) implies

$$\frac{\int_0^T \sqrt{Y_s}\,\mathrm{d}W_s}{\int_0^T Y_s\,\mathrm{d}s} \xrightarrow{\text{a.s.}} 0 \qquad \text{as} \quad T \to \infty,$$

since we have

$$\int_0^T Y_s\,\mathrm{d}s = T \cdot \frac{1}{T}\int_0^T Y_s\,\mathrm{d}s \xrightarrow{\text{a.s.}} \infty \qquad \text{as} \quad T \to \infty.$$

Further,

$$\frac{1}{T}\int_0^T (\sigma_2\sqrt{Y_s}\,\mathrm{d}\widetilde{W}_s + \sigma_3\,\mathrm{d}L_s) = \frac{1}{T}\int_0^T (\sigma_2^2 Y_s + \sigma_3^2)\mathrm{d}s \cdot \frac{\int_0^T (\sigma_2\sqrt{Y_s}\,\mathrm{d}\widetilde{W}_s + \sigma_3\,\mathrm{d}L_s)}{\int_0^T (\sigma_2^2 Y_s + \sigma_3^2)\,\mathrm{d}s} \xrightarrow{\text{a.s.}} 0$$

as $T \to \infty$. Indeed, we have already proved

$$\frac{1}{T}\int_0^T (\sigma_2^2 Y_s + \sigma_3^2)\,\mathrm{d}s \xrightarrow{\text{a.s.}} \mathbb{E}(\sigma_2^2 Y_\infty + \sigma_3^2) = \sigma_2^2\frac{a}{b} + \sigma_3^2 \in \mathbb{R}_{++} \qquad \text{as} \quad T \to \infty,$$

and the strong law of large numbers for continuous local martingales (see, e.g., Theorem C.1) implies

$$\frac{\int_0^T (\sigma_2\sqrt{Y_s}\,\mathrm{d}\widetilde{W}_s + \sigma_3\,\mathrm{d}L_s)}{\int_0^T (\sigma_2^2 Y_s + \sigma_3^2)\,\mathrm{d}s} \xrightarrow{\text{a.s.}} 0 \qquad \text{as} \quad T \to \infty,$$

since we have

$$\int_0^T (\sigma_2^2 Y_s + \sigma_3^2)\,\mathrm{d}s = T \cdot \frac{1}{T}\int_0^T (\sigma_2^2 Y_s + \sigma_3^2)\,\mathrm{d}s \xrightarrow{\text{a.s.}} \infty \qquad \text{as} \quad T \to \infty.$$

One can check

$$\frac{1}{T}\int_0^T Y_s\sqrt{Y_s}\,\mathrm{d}W_s \xrightarrow{\text{a.s.}} 0, \qquad \frac{1}{T}\int_0^T Y_s(\sigma_2\sqrt{Y_s}\,\mathrm{d}\widetilde{W}_s + \sigma_3\,\mathrm{d}L_s) \xrightarrow{\text{a.s.}} 0,$$
$$\frac{1}{T}\int_0^T X_s(\sigma_2\sqrt{Y_s}\,\mathrm{d}\widetilde{W}_s + \sigma_3\,\mathrm{d}L_s) \xrightarrow{\text{a.s.}} 0$$

as $T \to \infty$ in the same way, since

$$\frac{1}{T}\int_0^T Y_s^3\,\mathrm{d}s \xrightarrow{\text{a.s.}} \mathbb{E}(Y_\infty^3) \in \mathbb{R}_{++},$$
$$\frac{1}{T}\int_0^T Y_s^2(\sigma_2^2 Y_s + \sigma_3^2)\,\mathrm{d}s \xrightarrow{\text{a.s.}} \mathbb{E}\big[Y_s^2(\sigma_2^2 Y_s + \sigma_3^2)\big] \in \mathbb{R}_{++},$$
$$\frac{1}{T}\int_0^T X_s^2(\sigma_2^2 Y_s + \sigma_3^2)\,\mathrm{d}s \xrightarrow{\text{a.s.}} \mathbb{E}\big[X_s^2(\sigma_2^2 Y_s + \sigma_3^2)\big] \in \mathbb{R}_{++}$$

as $T \to \infty$. Consequently, we conclude (4.5). Finally, by (4.4) and (4.5), we obtain the statement. □

In order to handle supercritical two-factor affine diffusion models when $b \in \mathbb{R}_{--}$, we need the following integral version of the Toeplitz Lemma, due to Dietz and Kutoyants (1997).

Lemma 4.2 *Let $\{\varphi_T : T \in \mathbb{R}_+\}$ be a family of probability measures on $\mathbb{R}_+$ such that $\varphi_T([0,T]) = 1$ for all $T \in \mathbb{R}_+$, and $\lim_{T\to\infty} \varphi_T([0,K]) = 0$ for all $K \in \mathbb{R}_{++}$. Then for every bounded and measurable function $f : \mathbb{R}_+ \to \mathbb{R}$ for which the limit $f(\infty) := \lim_{t\to\infty} f(t)$ exists, we have*

$$\lim_{T\to\infty} \int_0^\infty f(t)\,\varphi_T(\mathrm{d}t) = f(\infty).$$

As a special case, we have the following integral version of the Kronecker Lemma, see Küchler and Sørensen (1997, Lemma B.3.2).

Lemma 4.3 *Let $a : \mathbb{R}_+ \to \mathbb{R}_+$ be a measurable function. Put $b(T) := \int_0^T a(t)\,\mathrm{d}t$, $T \in \mathbb{R}_+$. Suppose that $\lim_{T\to\infty} b(T) = \infty$. Then for every bounded and measurable function $f : \mathbb{R}_+ \to \mathbb{R}$ for which the limit $f(\infty) := \lim_{t\to\infty} f(t)$ exists, we have*

$$\lim_{T\to\infty} \frac{1}{b(T)} \int_0^T a(t) f(t)\,\mathrm{d}t = f(\infty).$$

Next we present an auxiliary lemma in the supercritical case on the asymptotic behavior of Y_t as $t \to \infty$.

Lemma 4.4 *Let us consider the two-factor affine diffusion model (1.1) with $a \in \mathbb{R}_+$, $b \in \mathbb{R}_{--}$, $\alpha, \beta, \gamma \in \mathbb{R}$, $\sigma_1, \sigma_2, \sigma_3 \in \mathbb{R}_+$ and $\varrho \in [-1, 1]$ with a random initial value (η_0, ζ_0) independent of $(W_t, B_t, L_t)_{t\in\mathbb{R}_+}$ satisfying $\mathbb{P}(\eta_0 \in \mathbb{R}_+) = 1$. Then there exists a random variable V_Y such that*

$$\mathrm{e}^{bt} Y_t \xrightarrow{\text{a.s.}} V_Y \qquad \text{as } t \to \infty \tag{4.6}$$

with $\mathbb{P}(V_Y \neq 0) = 1$, and, for each $k \in \mathbb{N}$,

$$\mathrm{e}^{kbt} \int_0^t Y_u^k\,\mathrm{d}u \xrightarrow{\text{a.s.}} -\frac{V_Y^k}{kb} \qquad \text{as } t \to \infty. \tag{4.7}$$

Proof By (2.1),

$$\mathbb{E}(Y_t \mid \mathcal{F}_s) = \mathbb{E}(Y_t \mid Y_s) = \mathrm{e}^{-b(t-s)} Y_s + a \int_s^t \mathrm{e}^{-b(t-u)}\,\mathrm{d}u$$

for all $s, t \in \mathbb{R}_+$ with $0 \leqslant s \leqslant t$. Thus

$$\mathbb{E}(\mathrm{e}^{bt} Y_t \mid \mathcal{F}_s^Y) = \mathrm{e}^{bs} Y_s + a \int_s^t \mathrm{e}^{bu}\,\mathrm{d}u \geqslant \mathrm{e}^{bs} Y_s$$

for all $s, t \in \mathbb{R}_+$ with $0 \leqslant s \leqslant t$, consequently, the process $(\mathrm{e}^{bt} Y_t)_{t\in\mathbb{R}_+}$ is a non-negative submartingale with respect to the filtration $(\mathcal{F}_t^Y)_{t\in\mathbb{R}_+}$. Moreover, $b \in \mathbb{R}_{--}$ implies

$$\mathbb{E}(\mathrm{e}^{bt} Y_t) = y_0 + a \int_0^t \mathrm{e}^{bu}\,\mathrm{d}u \leqslant y_0 + a \int_0^\infty \mathrm{e}^{bu}\,\mathrm{d}u = y_0 - \frac{a}{b} < \infty, \qquad t \in \mathbb{R}_+,$$

hence, by the submartingale convergence theorem, there exists a non-negative random variable V_Y such that (4.6) holds.

The distribution of V_Y coincides with the distribution of $\widetilde{\mathcal{Y}}_{-1/b}$, where $(\widetilde{\mathcal{Y}}_t)_{t\in\mathbb{R}_+}$ is a CIR process given by the SDE

$$\mathrm{d}\widetilde{\mathcal{Y}}_t = a\mathrm{d}t + \sigma_1\sqrt{\widetilde{\mathcal{Y}}_t}\,\mathrm{d}\mathcal{W}_t, \qquad t \in \mathbb{R}_+,$$

with initial value $\widetilde{\mathcal{Y}}_0 = y_0$, where $(\mathcal{W}_t)_{t\in\mathbb{R}_+}$ is a standard Wiener process, see Ben Alaya and Kebaier (2012, Proposition 3). Consequently, $\mathbb{P}(V_Y \in \mathbb{R}_{++}) = 1$, since $\widetilde{\mathcal{Y}}_t$, $t \in \mathbb{R}_{++}$, are absolutely continuous random variables.

If $\omega \in \Omega$ such that $\mathbb{R}_+ \ni t \mapsto Y_t(\omega)$ is continuous and $\mathrm{e}^{bt} Y_t(\omega) \to V_Y(\omega)$ as $t \to \infty$, then, by the integral Kronecker Lemma 4.3 with $f(t) = \mathrm{e}^{kbt} Y_t(\omega)^k$ and $a(t) = \mathrm{e}^{-kbt}$, $t \in \mathbb{R}_+$, we have

$$\frac{1}{\int_0^t \mathrm{e}^{-kbu}\,\mathrm{d}u} \int_0^t \mathrm{e}^{-kbu}(\mathrm{e}^{kbu} Y_u(\omega)^k)\,\mathrm{d}u \to V_Y(\omega)^k \qquad \text{as } t \to \infty.$$

Here $\int_0^t \mathrm{e}^{-kbu}\,\mathrm{d}u = -\frac{\mathrm{e}^{-kbt}-1}{kb}$, $t \in \mathbb{R}_+$, thus we conclude the second convergence in (4.7). □

The next theorem states strong consistency of the CLSE of b in the supercritical case.

Theorem 4.5 *Let us consider the two-factor affine diffusion model* (1.1) *with* $a \in \mathbb{R}_+$, $b \in \mathbb{R}_{--}$, $\alpha, \beta, \gamma \in \mathbb{R}$, $\sigma_1 \in \mathbb{R}_{++}$, $\sigma_2, \sigma_3 \in \mathbb{R}_+$ *and* $\varrho \in [-1, 1]$ *with a random initial value* (η_0, ζ_0) *independent of* $(W_t, B_t, L_t)_{t\in\mathbb{R}_+}$ *satisfying* $\mathbb{P}(\eta_0 \in \mathbb{R}_+) = 1$. *Then the CLSE of* b *is strongly consistent, i.e.,* $\widehat{b}_T \xrightarrow{\text{a.s.}} b$ *as* $T \to \infty$.

Proof By Lemma 3.3, there exists a unique CLSE $\widehat{b}_T$ of b for all $T \in \mathbb{R}_{++}$ which has the form given in (3.4). By Ito's formula,

$$\int_0^T Y_s\,\mathrm{d}Y_s = \frac{1}{2}(Y_T^2 - Y_0^2) - \frac{1}{2}\sigma_1^2 \int_0^T Y_s\,\mathrm{d}s, \qquad T \in \mathbb{R}_+,$$

hence, by (4.6) and (4.7), we have

$$\begin{aligned}
\widehat{b}_T &= \frac{(Y_T - Y_0)\int_0^T Y_s\,\mathrm{d}s - T\int_0^T Y_s\,\mathrm{d}Y_s}{T\int_0^T Y_s^2\,\mathrm{d}s - \left(\int_0^T Y_s\,\mathrm{d}s\right)^2} \\
&= \frac{(Y_T - Y_0)\int_0^T Y_s\,\mathrm{d}s - \frac{T}{2}(Y_T^2 - Y_0^2) + \frac{T}{2}\sigma_1^2\int_0^T Y_s\,\mathrm{d}s}{T\int_0^T Y_s^2\,\mathrm{d}s - \left(\int_0^T Y_s\,\mathrm{d}s\right)^2} \\
&= \frac{\frac{1}{T}\left(\mathrm{e}^{bT}Y_T - \mathrm{e}^{bT}Y_0\right)\left(\mathrm{e}^{bT}\int_0^T Y_s\,\mathrm{d}s\right) - \frac{1}{2}\left(\mathrm{e}^{2bT}Y_T^2 - \mathrm{e}^{2bT}Y_0^2\right) + \frac{1}{2}\sigma_1^2\mathrm{e}^{bT}\left(\mathrm{e}^{bT}\int_0^T Y_s\,\mathrm{d}s\right)}{\mathrm{e}^{2bT}\int_0^T Y_s^2\,\mathrm{d}s - \frac{1}{T}\left(\mathrm{e}^{bT}\int_0^T Y_s\,\mathrm{d}s\right)^2} \\
&\xrightarrow{\text{a.s.}} \frac{0(V_Y - 0)\left(-\frac{V_Y}{b}\right) - \frac{1}{2}(V_Y^2 - 0) + \frac{1}{2}\sigma_1^2 0\left(-\frac{V_Y}{b}\right)}{-\frac{V_Y^2}{2b} - 0\left(-\frac{V_Y}{b}\right)^2} = b
\end{aligned}$$

as $T \to \infty$. □

Remark 4.6 For critical two-factor affine diffusion models, it will turn out that the CLSE of a and α are not even weakly consistent, but the CLSE of b, β and γ are weakly consistent, see Theorem 6.2. □

Remark 4.7 For supercritical two-factor affine diffusion models, it will turn out that the CLSE of a and α are not even weakly consistent, but the CLSE of β and γ are weakly consistent, see Theorem 7.3. □

5 Asymptotic behavior of CLSE: subcritical case

Theorem 5.1 *Let us consider the two-factor affine diffusion model (1.1) with* $a, b \in \mathbb{R}_{++}$, $\alpha, \beta \in \mathbb{R}$, $\gamma \in \mathbb{R}_{++}$, $\sigma_1 \in \mathbb{R}_{++}$, $\sigma_2, \sigma_3 \in \mathbb{R}_+$ *and* $\varrho \in [-1, 1]$ *with a random initial value* (η_0, ζ_0) *independent of* $(W_t, B_t, L_t)_{t\in\mathbb{R}_+}$ *satisfying* $\mathbb{P}(\eta_0 \in \mathbb{R}_+) = 1$. *Suppose that* $(1-\varrho^2)\sigma_2^2 + \sigma_3^2 > 0$. *Then the CLSE of* $\boldsymbol{\theta} = (a, b, \alpha, \beta, \gamma)^\top$ *is asymptotically normal, namely,*

$$T^{\frac{1}{2}}(\widehat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}) \xrightarrow{\mathcal{D}} \mathcal{N}_5(\boldsymbol{0}, [\mathbb{E}(\boldsymbol{G}_\infty)]^{-1}\mathbb{E}(\widetilde{\boldsymbol{G}}_\infty)[\mathbb{E}(\boldsymbol{G}_\infty)]^{-1}) \quad \text{as } T \to \infty, \tag{5.1}$$

where $\boldsymbol{G}_\infty$ *is given in (4.3) and* $\widetilde{\boldsymbol{G}}_\infty$ *has the form*

$$\begin{bmatrix} \sigma_1^2 Y_\infty & -\sigma_1^2 Y_\infty^2 & \varrho\sigma_1\sigma_2 Y_\infty & -\varrho\sigma_1\sigma_2 Y_\infty^2 & -\varrho\sigma_1\sigma_2 Y_\infty X_\infty \\ -\sigma_1^2 Y_\infty^2 & \sigma_1^2 Y_\infty^3 & -\varrho\sigma_1\sigma_2 Y_\infty^2 & \varrho\sigma_1\sigma_2 Y_\infty^3 & \varrho\sigma_1\sigma_2 Y_\infty^2 X_\infty \\ \varrho\sigma_1\sigma_2 Y_\infty & -\varrho\sigma_1\sigma_2 Y_\infty^2 & \sigma_2^2 Y_\infty + \sigma_3^2 & -(\sigma_2^2 Y_\infty + \sigma_3^2) Y_\infty & -(\sigma_2^2 Y_\infty + \sigma_3^2) X_\infty \\ -\varrho\sigma_1\sigma_2 Y_\infty^2 & \varrho\sigma_1\sigma_2 Y_\infty^3 & -(\sigma_2^2 Y_\infty + \sigma_3^2) Y_\infty & (\sigma_2^2 Y_\infty + \sigma_3^2) Y_\infty^2 & (\sigma_2^2 Y_\infty + \sigma_3^2) Y_\infty X_\infty \\ -\varrho\sigma_1\sigma_2 Y_\infty X_\infty & \varrho\sigma_1\sigma_2 Y_\infty^2 X_\infty & -(\sigma_2^2 Y_\infty + \sigma_3^2) X_\infty & (\sigma_2^2 Y_\infty + \sigma_3^2) Y_\infty X_\infty & (\sigma_2^2 Y_\infty + \sigma_3^2) X_\infty^2 \end{bmatrix},$$

where the random vector (Y_∞, X_∞) *is given by Theorem A.1.*

Proof By (3.5), we have

$$T^{\frac{1}{2}}(\widehat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}) = (T^{-1}\boldsymbol{G}_T)^{-1}(T^{-\frac{1}{2}}\boldsymbol{h}_T) \tag{5.2}$$

on the event where $\boldsymbol{G}_T$ is invertible, which holds almost surely, see Lemma 3.3. We have $(T^{-1}\boldsymbol{G}_T)^{-1} \xrightarrow{\text{a.s.}} [\mathbb{E}(\boldsymbol{G}_\infty)]^{-1}$ as $T \to \infty$ by (4.4). The process $(\boldsymbol{h}_t)_{t\in\mathbb{R}_+}$ is a 5-dimensional continuous local martingale with quadratic variation process $\langle \boldsymbol{h} \rangle_t = \widetilde{\boldsymbol{G}}_t$, $t \in \mathbb{R}_+$, where

$$\widetilde{\boldsymbol{G}}_t := \int_0^t \begin{bmatrix} \sigma_1^2 Y_s & -\sigma_1^2 Y_s^2 & \varrho\sigma_1\sigma_2 Y_s & -\varrho\sigma_1\sigma_2 Y_s^2 & -\varrho\sigma_1\sigma_2 Y_s X_s \\ -\sigma_1^2 Y_s^2 & \sigma_1^2 Y_s^3 & -\varrho\sigma_1\sigma_2 Y_s^2 & \varrho\sigma_1\sigma_2 Y_s^3 & \varrho\sigma_1\sigma_2 Y_s^2 X_s \\ \varrho\sigma_1\sigma_2 Y_s & -\varrho\sigma_1\sigma_2 Y_s^2 & \sigma_2^2 Y_s + \sigma_3^2 & -(\sigma_2^2 Y_s + \sigma_3^2) Y_s & -(\sigma_2^2 Y_s + \sigma_3^2) X_s \\ -\varrho\sigma_1\sigma_2 Y_s^2 & \varrho\sigma_1\sigma_2 Y_s^3 & -(\sigma_2^2 Y_s + \sigma_3^2) Y_s & (\sigma_2^2 Y_s + \sigma_3^2) Y_s^2 & (\sigma_2^2 Y_s + \sigma_3^2) Y_s X_s \\ -\varrho\sigma_1\sigma_2 Y_s X_s & \varrho\sigma_1\sigma_2 Y_s^2 X_s & -(\sigma_2^2 Y_s + \sigma_3^2) X_s & (\sigma_2^2 Y_s + \sigma_3^2) Y_s X_s & (\sigma_2^2 Y_s + \sigma_3^2) X_s^2 \end{bmatrix} \mathrm{d}s.$$

By Theorem A.2, we obtain

$$T^{-1}\widetilde{\boldsymbol{G}}_T \xrightarrow{\text{a.s.}} \mathbb{E}(\widetilde{\boldsymbol{G}}_\infty) \quad \text{as } T \to \infty, \tag{5.3}$$

since, by Theorem B.2, the entries of $\mathbb{E}(\widetilde{\boldsymbol{G}}_\infty)$ exist and finite. Using (5.3), Theorem C.2 yields $T^{-\frac{1}{2}}\boldsymbol{h}_T \xrightarrow{\mathcal{D}} \mathcal{N}_5(\boldsymbol{0}, \mathbb{E}(\widetilde{\boldsymbol{G}}_\infty))$ as $T \to \infty$. Hence, by (5.2) and by Slutsky's lemma,

$$T^{\frac{1}{2}}(\widehat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}) \xrightarrow{\mathcal{D}} [\mathbb{E}(\boldsymbol{G}_\infty)]^{-1}\mathcal{N}_5\big(\boldsymbol{0}, \mathbb{E}(\widetilde{\boldsymbol{G}}_\infty)\big) = \mathcal{N}_5\big(\boldsymbol{0}, [\mathbb{E}(\boldsymbol{G}_\infty)]^{-1}\mathbb{E}(\widetilde{\boldsymbol{G}}_\infty)\big([\mathbb{E}(\boldsymbol{G}_\infty)]^{-1}\big)^\top\big)$$

as $T \to \infty$. □

6 Asymptotic behavior of CLSE: critical case

First we present an auxiliary lemma. A proof can be found in the Arxiv version of this paper Bolyog and Pap (2017).

Lemma 6.1 *If* $(\mathcal{Y}_t, \mathcal{X}_t)_{t\in\mathbb{R}_+}$ *and* $(\widetilde{\mathcal{Y}}_t, \widetilde{\mathcal{X}}_t)_{t\in\mathbb{R}_+}$ *are continuous semimartingales such that* $(\mathcal{Y}_t, \mathcal{X}_t)_{t\in\mathbb{R}_+} \stackrel{\mathcal{D}}{=} (\widetilde{\mathcal{Y}}_t, \widetilde{\mathcal{X}}_t)_{t\in\mathbb{R}_+}$, *then*

$$\begin{aligned}&\left(\mathcal{Y}_1, \mathcal{X}_1, \int_0^1 \mathcal{X}_s \,\mathrm{d}\mathcal{Y}_s, \int_0^1 \mathcal{Y}_s^k \mathcal{X}_s^\ell \,\mathrm{d}s : k, \ell \in \mathbb{Z}_+, k+\ell \leqslant n\right)\\ &\stackrel{\mathcal{D}}{=} \left(\widetilde{\mathcal{Y}}_1, \widetilde{\mathcal{X}}_1, \int_0^1 \widetilde{\mathcal{X}}_s \,\mathrm{d}\widetilde{\mathcal{Y}}_s, \int_0^1 \widetilde{\mathcal{Y}}_s^k \widetilde{\mathcal{X}}_s^\ell \,\mathrm{d}s : k, \ell \in \mathbb{Z}_+, k+\ell \leqslant n\right)\end{aligned}$$

for each $n \in \mathbb{N}$.

Theorem 6.2 *Let us consider the two-factor affine diffusion model* (1.1) *with* $a \in \mathbb{R}_+$, $b = 0$, $\alpha \in \mathbb{R}$, $\beta = 0$, $\gamma = 0$, $\sigma_1, \sigma_2, \sigma_3 \in \mathbb{R}_+$ *and* $\varrho \in [-1, 1]$ *with a random initial value* (η_0, ζ_0) *independent of* $(W_t, B_t, L_t)_{t\in\mathbb{R}_+}$ *satisfying* $\mathbb{P}(\eta_0 \in \mathbb{R}_+) = 1$. *Suppose that* $(1-\varrho^2)\sigma_2^2 + \sigma_3^2 > 0$. *Then*

$$\begin{bmatrix} \widehat{a}_T - a \\ T\widehat{b}_T \\ \widehat{\alpha}_T - \alpha \\ T\widehat{\beta}_T \\ T\widehat{\gamma}_T \end{bmatrix} \stackrel{\mathcal{D}}{\longrightarrow} \begin{bmatrix} \left(\int_0^1 \begin{bmatrix} 1 \\ -\mathcal{Y}_s \end{bmatrix} \begin{bmatrix} 1 \\ -\mathcal{Y}_s \end{bmatrix}^\top \mathrm{d}s\right)^{-1} \begin{bmatrix} \mathcal{Y}_1 - a \\ -\frac{1}{2}\mathcal{Y}_1^2 + \left(a + \frac{\sigma_1^2}{2}\right)\int_0^1 \mathcal{Y}_s \,\mathrm{d}s \end{bmatrix} \\ \left(\int_0^1 \begin{bmatrix} 1 \\ -\mathcal{Y}_s \\ -\mathcal{X}_s \end{bmatrix} \begin{bmatrix} 1 \\ -\mathcal{Y}_s \\ -\mathcal{X}_s \end{bmatrix}^\top \mathrm{d}s\right)^{-1} \begin{bmatrix} \mathcal{X}_1 - \alpha \\ -\mathcal{Y}_1\mathcal{X}_1 + (\alpha + \varrho\sigma_1\sigma_2)\int_0^1 \mathcal{Y}_s \,\mathrm{d}s + \int_0^1 \mathcal{X}_s \,\mathrm{d}\mathcal{Y}_s \\ -\frac{1}{2}\mathcal{X}_1^2 + \alpha\int_0^1 \mathcal{X}_s \,\mathrm{d}s + \frac{\sigma_2^2}{2}\int_0^1 \mathcal{Y}_s \,\mathrm{d}s \end{bmatrix} \end{bmatrix} \tag{6.1}$$

as $T \to \infty$, *where* $(\mathcal{Y}_t, \mathcal{X}_t)_{t\in\mathbb{R}_+}$ *is the unique strong solution of the SDE*

$$\begin{cases} \mathrm{d}\mathcal{Y}_t = a \,\mathrm{d}t + \sigma_1\sqrt{\mathcal{Y}_t}\,\mathrm{d}W_t, \\ \mathrm{d}\mathcal{X}_t = \alpha \,\mathrm{d}t + \sigma_2\sqrt{\mathcal{Y}_t}\,(\varrho \,\mathrm{d}W_t + \sqrt{1-\varrho^2}\,\mathrm{d}B_t), \end{cases} \qquad t \in [0, \infty), \tag{6.2}$$

with initial value $(\mathcal{Y}_0, \mathcal{X}_0) = (0, 0)$.

Proof By (3.5), we have

$$\begin{aligned}\begin{bmatrix} \widehat{a}_T - a \\ T\widehat{b}_T \end{bmatrix} &= \begin{bmatrix} 1 & 0 \\ 0 & T \end{bmatrix} \begin{bmatrix} \widehat{a}_T - a \\ \widehat{b}_T \end{bmatrix} = \operatorname{diag}(1, T)(\boldsymbol{G}_T^{(1)})^{-1}\boldsymbol{h}_T^{(1)} \\ &= \left(\operatorname{diag}(T^{-\frac{1}{2}}, T^{-\frac{3}{2}})\boldsymbol{G}_T^{(1)} \operatorname{diag}(T^{-\frac{1}{2}}, T^{-\frac{3}{2}})\right)^{-1} \operatorname{diag}(T^{-1}, T^{-2})\boldsymbol{h}_T^{(1)} \\ &= \begin{bmatrix} 1 & -\frac{1}{T^2}\int_0^T Y_s \,\mathrm{d}s \\ -\frac{1}{T^2}\int_0^T Y_s \,\mathrm{d}s & \frac{1}{T^3}\int_0^T Y_s^2 \,\mathrm{d}s \end{bmatrix}^{-1} \begin{bmatrix} \frac{\sigma_1}{T}\int_0^T Y_s^{\frac{1}{2}} \,\mathrm{d}W_s \\ -\frac{\sigma_1}{T^2}\int_0^T Y_s^{\frac{3}{2}} \,\mathrm{d}W_s \end{bmatrix}.\end{aligned}$$

In a similar way,

$$\begin{bmatrix} \widehat{\alpha}_T - \alpha \\ T\widehat{\beta}_T \\ T\widehat{\gamma}_T \end{bmatrix} = \begin{bmatrix} 1 & -\frac{1}{T^2}\int_0^T Y_s\,\mathrm{d}s & -\frac{1}{T^2}\int_0^T X_s\,\mathrm{d}s \\ -\frac{1}{T^2}\int_0^T Y_s\,\mathrm{d}s & \frac{1}{T^3}\int_0^T Y_s^2\,\mathrm{d}s & \frac{1}{T^3}\int_0^T Y_s X_s\,\mathrm{d}s \\ -\frac{1}{T^2}\int_0^T X_s\,\mathrm{d}s & \frac{1}{T^3}\int_0^T Y_s X_s\,\mathrm{d}s & \frac{1}{T^3}\int_0^T X_s^2\,\mathrm{d}s \end{bmatrix}^{-1}$$

$$\times \begin{bmatrix} \frac{\sigma_2}{T}\int_0^T Y_s^{\frac{1}{2}}\,\mathrm{d}\widetilde{W}_s + \frac{\sigma_3}{T} L_T \\ -\frac{\sigma_2}{T^2}\int_0^T Y_s^{\frac{3}{2}}\,\mathrm{d}\widetilde{W}_s - \frac{\sigma_3}{T^2}\int_0^T Y_s\,\mathrm{d}L_s \\ -\frac{\sigma_2}{T^2}\int_0^T Y_s^{\frac{1}{2}} X_s\,\mathrm{d}\widetilde{W}_s - \frac{\sigma_3}{T^2}\int_0^T X_s\,\mathrm{d}L_s \end{bmatrix}.$$

The aim of the following discussion is to prove

$$\begin{aligned}&\left(\tfrac{1}{T}Y_T, \tfrac{1}{T}X_T, \tfrac{1}{T^2}\int_0^T X_s\,\mathrm{d}Y_s, \tfrac{1}{T^{k+\ell+1}}\int_0^T Y_s^k X_s^\ell\,\mathrm{d}s : k,\ell\in\mathbb{Z}_+, k+\ell\leqslant 2\right)\\ &\quad\xrightarrow{\mathcal{D}} \left(\mathcal{Y}_1, \mathcal{X}_1, \int_0^1 \mathcal{X}_s\,\mathrm{d}\mathcal{Y}_s, \int_0^1 \mathcal{Y}_s^k \mathcal{X}_s^\ell\,\mathrm{d}s : k,\ell\in\mathbb{Z}_+, k+\ell\leqslant 2\right) \end{aligned} \tag{6.3}$$

as $T\to\infty$. By part (ii) of Remark 2.7 in Barczy et al. (2013), we have

$$(\widetilde{\mathcal{Y}}_t^{(T)}, \widetilde{\mathcal{X}}_t^{(T)})_{t\in\mathbb{R}_+} := \left(\frac{1}{T}\mathcal{Y}_{Tt}, \frac{1}{T}\mathcal{X}_{Tt}\right)_{t\in\mathbb{R}_+} \stackrel{\mathcal{D}}{=} (\mathcal{Y}_t, \mathcal{X}_t)_{t\in\mathbb{R}_+} \qquad \text{for all } T\in\mathbb{R}_{++},$$

since, by Proposition 2.1, $(\mathcal{Y}_t, \mathcal{X}_t)_{t\in\mathbb{R}_+}$ is an affine process with infinitesimal generator

$$\begin{aligned}(\mathcal{A}_{(\mathcal{Y},\mathcal{X})}f)(y,x) &= af_1'(y,x) + \alpha f_2'(y,x)\\ &\quad + \frac{1}{2}y\big[\sigma_1^2 f_{1,1}''(y,x) + 2\varrho\sigma_1\sigma_2 f_{1,2}''(y,x) + \sigma_2^2 f_{2,2}''(y,x)\big].\end{aligned}$$

Hence, by Lemma 6.1, we obtain

$$\begin{aligned}&\left(\mathcal{Y}_1, \mathcal{X}_1, \int_0^1 \mathcal{X}_s\,\mathrm{d}\mathcal{Y}_s, \int_0^1 \mathcal{Y}_s^k\mathcal{X}_s^\ell\,\mathrm{d}s : k,\ell\in\mathbb{Z}_+, k+\ell\leqslant 2\right)\\ &\stackrel{\mathcal{D}}{=} \left(\widetilde{\mathcal{Y}}_1^{(T)}, \widetilde{\mathcal{X}}_1^{(T)}, \int_0^1 \widetilde{\mathcal{X}}_s^{(T)}\,\mathrm{d}\widetilde{\mathcal{Y}}_s^{(T)}, \int_0^1 (\widetilde{\mathcal{Y}}_s^{(T)})^k(\widetilde{\mathcal{X}}_s^{(T)})^\ell\,\mathrm{d}s : k,\ell\in\mathbb{Z}_+, k+\ell\leqslant 2\right)\\ &= \left(\frac{1}{T}\mathcal{Y}_T, \frac{1}{T}\mathcal{X}_T, \frac{1}{T^2}\int_0^T \mathcal{X}_s\,\mathrm{d}\mathcal{Y}_s, \frac{1}{T^{k+\ell+1}}\int_0^T \mathcal{Y}_s^k\mathcal{X}_s^\ell\,\mathrm{d}s : k,\ell\in\mathbb{Z}_+, k+\ell\leqslant 2\right)\end{aligned}$$

for all $T\in\mathbb{R}_{++}$. Then, by Slutsky's lemma, in order to prove (6.3), it suffices to show the convergences

$$\frac{1}{T}(Y_T - \mathcal{Y}_T) \xrightarrow{\mathbb{P}} 0, \qquad \frac{1}{T}(X_T - \mathcal{X}_T) \xrightarrow{\mathbb{P}} 0, \tag{6.4}$$

$$\frac{1}{T^2}\left(\int_0^T X_s\,\mathrm{d}Y_s - \int_0^T \mathcal{X}_s\,\mathrm{d}\mathcal{Y}_s\right) \xrightarrow{\mathbb{P}} 0, \qquad \frac{1}{T^{k+\ell+1}}\int_0^T (Y_s^k X_s^\ell - \mathcal{Y}_s^k\mathcal{X}_s^\ell)\,\mathrm{d}s \xrightarrow{\mathbb{P}} 0 \tag{6.5}$$

as $T\to\infty$ for all $k,\ell\in\mathbb{Z}_+$ with $k+\ell\leqslant 2$. By (3.21) in Barczy et al. (2013), we have

$$\mathbb{E}(|Y_s - \mathcal{Y}_s|) \leqslant \mathbb{E}(Y_0), \qquad s\in\mathbb{R}_+, \tag{6.6}$$

hence

$$\mathbb{E}\left(\left|\frac{1}{T}(Y_T - \mathcal{Y}_T)\right|\right) \leqslant \frac{1}{T}\mathbb{E}(Y_0) \to 0,$$
$$\mathbb{E}\left(\left|\frac{1}{T^2}\int_0^T (Y_s - \mathcal{Y}_s)\,\mathrm{d}s\right|\right) \leqslant \frac{1}{T^2}\int_0^T \mathbb{E}(|Y_s - \mathcal{Y}_s|)\,\mathrm{d}s \leqslant \frac{1}{T}\mathbb{E}(Y_0) \to 0,$$

as $T \to \infty$, implying $\frac{1}{T}(Y_T - \mathcal{Y}_T) \xrightarrow{\mathbb{P}} 0$ and $\frac{1}{T^2}\int_0^T (Y_s - \mathcal{Y}_s)\,\mathrm{d}s \xrightarrow{\mathbb{P}} 0$ as $T \to \infty$, i.e., the first convergence in (6.4) and the second convergence in (6.5) for $(k, \ell) = (1, 0)$.

As in (3.23) in Barczy et al. (2013), we have $\mathbb{E}(|X_s - \mathcal{X}_s|) \leqslant \mathbb{E}(|X_0|) + \sqrt{(\sigma_2^2\mathbb{E}(Y_0) + \sigma_3^2)s}$ for all $s \in \mathbb{R}_+$, hence

$$\sup_{s\in[0,T]} \mathbb{E}(|X_s - \mathcal{X}_s|) = \mathrm{O}(T^{\frac{1}{2}}) \qquad \text{as } T \to \infty, \tag{6.7}$$

thus

$$\mathbb{E}\left(\left|\frac{1}{T}(X_T - \mathcal{X}_T)\right|\right) = \frac{1}{T}\,\mathrm{O}(T^{\frac{1}{2}}) \to 0,$$
$$\mathbb{E}\left(\left|\frac{1}{T^2}\int_0^T (X_s - \mathcal{X}_s)\,\mathrm{d}s\right|\right) \leqslant \frac{1}{T^2}\int_0^T \mathbb{E}(|X_s - \mathcal{X}_s|)\,\mathrm{d}s = \frac{1}{T^2}\int_0^T \mathrm{O}(T^{\frac{1}{2}})\,\mathrm{d}s$$
$$= \frac{1}{T^2}\,\mathrm{O}(T^{\frac{3}{2}}) \to 0,$$

as $T \to \infty$, implying $\frac{1}{T}(X_T - \mathcal{X}_T) \xrightarrow{\mathbb{P}} 0$ and $\frac{1}{T^2}\int_0^T (X_s - \mathcal{X}_s)\,\mathrm{d}s \xrightarrow{\mathbb{P}} 0$ as $T \to \infty$, i.e., the second convergence in (6.4) and the second convergence in (6.5) for $(k, \ell) = (0, 1)$.

As in (3.25) in Barczy et al. (2013), we have $\mathbb{E}[(Y_s - \mathcal{Y}_s)^2] \leqslant 2\mathbb{E}(Y_0^2) + 2s\sigma_1^2\mathbb{E}(Y_0)$ for all $s \in \mathbb{R}_+$, hence

$$\sup_{s\in[0,T]} \mathbb{E}[(Y_s - \mathcal{Y}_s)^2] = \mathrm{O}(T) \qquad \text{as } T \to \infty. \tag{6.8}$$

By Proposition B.1, $\mathbb{E}(Y_s^2) = \mathbb{E}(Y_0^2) + (2a + \sigma_1^2)\big(\mathbb{E}(Y_0)s + a\frac{s^2}{2}\big)$ for all $s \in \mathbb{R}_+$, hence

$$\sup_{s\in[0,T]} \mathbb{E}(Y_s^2) = \mathrm{O}(T^2) \qquad \text{as } T \to \infty, \tag{6.9}$$

and $\sup_{s\in[0,T]} \mathbb{E}(\mathcal{Y}_s^2) = \mathrm{O}(T^2)$ as $T \to \infty$. We have

$$\mathbb{E}(|Y_s^2 - \mathcal{Y}_s^2|) = \mathbb{E}(|(Y_s - \mathcal{Y}_s)(Y_s + \mathcal{Y}_s)|) \leqslant \sqrt{\mathbb{E}[(Y_s - \mathcal{Y}_s)^2]\mathbb{E}[(Y_s + \mathcal{Y}_s)^2]}$$
$$\leqslant \sqrt{2\mathbb{E}[(Y_s - \mathcal{Y}_s)^2](\mathbb{E}(Y_s^2) + \mathbb{E}(\mathcal{Y}_s^2))},$$

yielding

$$\sup_{s\in[0,T]} \mathbb{E}(|Y_s^2 - \mathcal{Y}_s^2|) = \sqrt{2\,\mathrm{O}(T)(\mathrm{O}(T^2) + \mathrm{O}(T^2))} = \mathrm{O}(T^{\frac{3}{2}}) \qquad \text{as } T \to \infty,$$

thus

$$\mathbb{E}\left(\left|\frac{1}{T^3}\int_0^T (Y_s^2 - \mathcal{Y}_s^2)\,\mathrm{d}s\right|\right) \leqslant \frac{1}{T^3}\int_0^T \mathbb{E}(|Y_s^2 - \mathcal{Y}_s^2|)\,\mathrm{d}s = \frac{1}{T^3}\int_0^T \mathrm{O}(T^{\frac{3}{2}})\,\mathrm{d}s$$
$$= \frac{1}{T^3}\,\mathrm{O}(T^{\frac{5}{2}}) \to 0,$$

as $T \to \infty$, implying $\frac{1}{T^3}\int_0^T (Y_s^2 - \mathcal{Y}_s^2)\,\mathrm{d}s \xrightarrow{\mathbb{P}} 0$ as $T \to \infty$., i.e., the second convergence in (6.5) for $(k, \ell) = (2, 0)$.

In a similar way, $\mathbb{E}[(X_s - \mathcal{X}_s)^2] \leqslant 2\mathbb{E}(X_0^2) + 2s(\sigma_2^2\mathbb{E}(Y_0) + \sigma_3^2)$ for all $s \in \mathbb{R}_+$, hence

$$\sup_{s\in[0,T]} \mathbb{E}[(X_s - \mathcal{X}_s)^2] = \mathrm{O}(T) \qquad \text{as } T \to \infty. \tag{6.10}$$

By Proposition B.1, $\mathbb{E}(X_s^2) = \mathbb{E}(X_0^2) + \alpha\big(s\mathbb{E}(X_0) + \alpha\frac{s^2}{2}\big) + \sigma_2^2\big(s\mathbb{E}(Y_0) + a\frac{s^2}{2}\big) + \sigma_3^2 s$, thus $\sup_{s\in[0,T]}\mathbb{E}(X_s^2) = \mathrm{O}(T^2)$ and $\sup_{s\in[0,T]}\mathbb{E}(\mathcal{X}_s^2) = \mathrm{O}(T^2)$ as $T \to \infty$. We have

$$\mathbb{E}(|X_s^2 - \mathcal{X}_s^2|) \leqslant \sqrt{2\mathbb{E}[(X_s - \mathcal{X}_s)^2](\mathbb{E}(X_s^2) + \mathbb{E}(\mathcal{X}_s^2))},$$

yielding

$$\sup_{s\in[0,T]} \mathbb{E}(|X_s^2 - \mathcal{X}_s^2|) = \sqrt{2\,\mathrm{O}(T)(\mathrm{O}(T^2) + \mathrm{O}(T^2))} = \mathrm{O}(T^{\frac{3}{2}}) \qquad \text{as } T \to \infty,$$

thus

$$\begin{aligned}\mathbb{E}\left(\left|\frac{1}{T^3}\int_0^T (X_s^2 - \mathcal{X}_s^2)\,\mathrm{d}s\right|\right) &\leqslant \frac{1}{T^3}\int_0^T \mathbb{E}(|X_s^2 - \mathcal{X}_s^2|)\,\mathrm{d}s = \frac{1}{T^3}\int_0^T \mathrm{O}(T^{\frac{3}{2}})\,\mathrm{d}s\\ &= \frac{1}{T^3}\mathrm{O}(T^{\frac{5}{2}}) \to 0,\end{aligned}$$

as $T \to \infty$, implying $\frac{1}{T^3}\int_0^T (X_s^2 - \mathcal{X}_s^2)\,\mathrm{d}s \xrightarrow{\mathbb{P}} 0$ as $T \to \infty$, i.e., the second convergence in (6.5) for $(k, \ell) = (0, 2)$.

Further,

$$\begin{aligned}\mathbb{E}(|Y_s X_s - \mathcal{Y}_s\mathcal{X}_s|) &\leqslant \mathbb{E}(|Y_s - \mathcal{Y}_s||X_s|) + \mathbb{E}(\mathcal{Y}_s|X_s - \mathcal{X}_s|)\\ &\leqslant \sqrt{\mathbb{E}[(Y_s - \mathcal{Y}_s)^2]\mathbb{E}(X_s^2)} + \sqrt{\mathbb{E}(\mathcal{Y}_s^2)\mathbb{E}[(X_s - \mathcal{X}_s)^2]}\end{aligned}$$

yields

$$\sup_{s\in[0,T]} \mathbb{E}(|Y_s X_s - \mathcal{Y}_s\mathcal{X}_s|) = \sqrt{\mathrm{O}(T)\,\mathrm{O}(T^2)} + \sqrt{\mathrm{O}(T^2)\,\mathrm{O}(T))} = \mathrm{O}(T^{\frac{3}{2}}) \qquad \text{as } T \to \infty,$$

thus

$$\begin{aligned}\mathbb{E}\left(\left|\frac{1}{T^3}\int_0^T (Y_s X_s - \mathcal{Y}_s\mathcal{X}_s)\,\mathrm{d}s\right|\right) &\leqslant \frac{1}{T^3}\int_0^T \mathbb{E}(|Y_s X_s - \mathcal{Y}_s\mathcal{X}_s|)\,\mathrm{d}s = \frac{1}{T^3}\int_0^T \mathrm{O}(T^{\frac{3}{2}})\,\mathrm{d}s\\ &= \frac{1}{T^3}\mathrm{O}(T^{\frac{5}{2}}) \to 0,\end{aligned}$$

as $T \to \infty$, implying $\frac{1}{T^3}\int_0^T (Y_s X_s - \mathcal{Y}_s\mathcal{X}_s)\,\mathrm{d}s \xrightarrow{\mathbb{P}} 0$ as $T \to \infty$, i.e., the second convergence in (6.5) for $(k, \ell) = (1, 1)$.

Using the Cauchy–Schwarz inequality, we obtain

$$\begin{aligned}\mathbb{E}\left(\left|\int_0^T X_s\,\mathrm{d}Y_s - \int_0^T \mathcal{X}_s\,\mathrm{d}\mathcal{Y}_s\right|\right) &\leqslant \mathbb{E}\left(\left|\int_0^T (X_s - \mathcal{X}_s)\,\mathrm{d}Y_s\right|\right) + \mathbb{E}\left(\left|\int_0^T \mathcal{X}_s\,\mathrm{d}(Y_s - \mathcal{Y}_s)\right|\right)\\ &\leqslant \sqrt{E_1(T)} + \sqrt{E_2(T)}\end{aligned}$$

with

$$E_1(T) := \mathbb{E}\left(\left|\int_0^T (X_s - \mathcal{X}_s)\,\mathrm{d}Y_s\right|^2\right), \qquad E_2(T) := \mathbb{E}\left(\left|\int_0^T \mathcal{X}_s\,\mathrm{d}(Y_s - \mathcal{Y}_s)\right|^2\right).$$

Using $\mathrm{d}Y_s = a\,\mathrm{d}s + \sigma_1\sqrt{Y_s}\,\mathrm{d}W_s$, we have

$$E_1(T) = \mathbb{E}\left(\left|a\int_0^T (X_s - \mathcal{X}_s)\,\mathrm{d}s + \sigma_1\int_0^T (X_s - \mathcal{X}_s)\sqrt{Y_s}\,\mathrm{d}W_s\right|^2\right) \leqslant 2a^2 E_{1,1}(T) + 2\sigma_1^2 E_{1,2}(T)$$

with

$$E_{1,1}(T) := \mathbb{E}\left(\left|\int_0^T (X_s - \mathcal{X}_s)\,\mathrm{d}s\right|^2\right), \qquad E_{1,2}(T) := \mathbb{E}\left(\left|\int_0^T (X_s - \mathcal{X}_s)\sqrt{Y_s}\,\mathrm{d}W_s\right|^2\right).$$

Applying (6.10), we obtain

$$\begin{aligned} E_{1,1}(T) &= \mathbb{E}\left(\int_0^T\int_0^T (X_s - \mathcal{X}_s)(X_u - \mathcal{X}_u)\,\mathrm{d}s\,\mathrm{d}u\right) \\ &= \int_0^T\int_0^T \mathbb{E}[(X_s - \mathcal{X}_s)(X_u - \mathcal{X}_u)]\,\mathrm{d}s\,\mathrm{d}u \\ &\leqslant \int_0^T\int_0^T \sqrt{\mathbb{E}[(X_s - \mathcal{X}_s)^2]\mathbb{E}[(X_u - \mathcal{X}_u)^2]}\,\mathrm{d}s\,\mathrm{d}u \\ &= \int_0^T\int_0^T \sqrt{\mathrm{O}(T)\,\mathrm{O}(T)}\,\mathrm{d}s\,\mathrm{d}u = \mathrm{O}(T^3). \end{aligned}$$

Again by the Cauchy–Schwarz inequality, we obtain

$$\begin{aligned} E_{1,2}(T) &= \mathbb{E}\left(\int_0^T (X_s - \mathcal{X}_s)^2 Y_s\,\mathrm{d}s\right) = \int_0^T \mathbb{E}[(X_s - \mathcal{X}_s)^2 Y_s]\,\mathrm{d}s \\ &\leqslant \int_0^T \sqrt{\mathbb{E}[(X_s - \mathcal{X}_s)^4]\mathbb{E}(Y_s^2)}\,\mathrm{d}s. \end{aligned}$$

Using $X_t = X_0 + \sigma_2\int_0^t \sqrt{Y_s}\,\mathrm{d}\widetilde{W}_s + \sigma_3 L_t$ and $\mathcal{X}_t = \sigma_2\int_0^t \sqrt{\mathcal{Y}_s}\,\mathrm{d}\widetilde{W}_s$, we get $X_t - \mathcal{X}_t = X_0 + \sigma_2\int_0^t(\sqrt{Y_s} - \sqrt{\mathcal{Y}_s})\,\mathrm{d}\widetilde{W}_s + \sigma_3 L_t$, and, applying Minkowski inequality and a martingale moment inequality in Karatzas and Shreve (1991, 3.3.25), we obtain

$$\begin{aligned} (\mathbb{E}[(X_t - \mathcal{X}_t)^4])^{\frac14} &\leqslant [\mathbb{E}(X_0^4)]^{\frac14} + \sigma_2\left(\mathbb{E}\left[\left(\int_0^t(\sqrt{Y_s} - \sqrt{\mathcal{Y}_s})\,\mathrm{d}\widetilde{W}_s\right)^4\right]\right)^{\frac14} + \sigma_3[\mathbb{E}(L_t^4)]^{\frac14} \\ &\leqslant [\mathbb{E}(X_0^4)]^{\frac14} + \sigma_2\left((2\cdot 3)^2 t\,\mathbb{E}\left(\int_0^t(\sqrt{Y_s} - \sqrt{\mathcal{Y}_s})^4\,\mathrm{d}s\right)\right)^{\frac14} + \sigma_3\sqrt[4]{3}\sqrt{t} \\ &\leqslant [\mathbb{E}(X_0^4)]^{\frac14} + \sigma_2\left(36t\int_0^t \mathbb{E}[(Y_s - \mathcal{Y}_s)^2]\,\mathrm{d}s\right)^{\frac14} + \sigma_3\sqrt[4]{3}\sqrt{t}. \end{aligned}$$

Applying (6.8), we get

$$\sup_{t\in[0,T]} \mathbb{E}[(X_t - \mathcal{X}_t)^4] = \mathrm{O}(T^3) \qquad \text{as } T\to\infty, \tag{6.11}$$

which, by (6.9), implies $E_{1,2}(T) = \int_0^T \sqrt{\mathrm{O}(T^3)\,\mathrm{O}(T^2)}\,\mathrm{d}s = \mathrm{O}(T^{\frac72})$ as $T\to\infty$. Using $E_{1,1}(T) = \mathrm{O}(T^3)$ as $T\to\infty$, we conclude $E_1(T) = \mathrm{O}(T^3) + \mathrm{O}(T^{\frac72}) = \mathrm{O}(T^{\frac72})$ as $T\to\infty$.

Using $\mathrm{d}Y_s = a\,\mathrm{d}s + \sigma_1\sqrt{Y_s}\,\mathrm{d}W_s$ and $\mathrm{d}\mathcal{Y}_s = a\,\mathrm{d}s + \sigma_1\sqrt{\mathcal{Y}_s}\,\mathrm{d}W_s$, we obtain $\mathrm{d}(Y_t - \mathcal{Y}_t) = \sigma_1(\sqrt{Y_t} - \sqrt{\mathcal{Y}_t})\,\mathrm{d}W_t$, thus

$$\begin{aligned} E_2(T) &= \sigma_1^2\mathbb{E}\left(\int_0^T \mathcal{X}_s^2(\sqrt{Y_s} - \sqrt{\mathcal{Y}_s})^2\,\mathrm{d}s\right) \leqslant \sigma_1^2\int_0^T \mathbb{E}[\mathcal{X}_s^2|Y_s - \mathcal{Y}_s|]\,\mathrm{d}s \\ &\leqslant \sigma_1^2\int_0^T \sqrt{\mathbb{E}(\mathcal{X}_s^4)\mathbb{E}[(Y_s - \mathcal{Y}_s)^2]}\,\mathrm{d}s. \end{aligned}$$

Using $\mathcal{X}_t = \alpha t + \sigma_2\int_0^t \sqrt{\mathcal{Y}_s}\,\mathrm{d}\widetilde{W}_s$, we obtain

$$\begin{aligned} [\mathbb{E}(\mathcal{X}_t^4)]^{\frac{1}{4}} &\leqslant |\alpha|t + \sigma_2\left(\mathbb{E}\left[\left(\int_0^t \sqrt{\mathcal{Y}_s}\,\mathrm{d}\widetilde{W}_s\right)^4\right]\right)^{\frac{1}{4}} \leqslant |\alpha|t + \sigma_2\left((2\cdot 3)^2 t\,\mathbb{E}\left(\int_0^t \mathcal{Y}_s^2\,\mathrm{d}s\right)\right)^{\frac{1}{4}} \\ &= |\alpha|t + \sigma_2\left(36t\int_0^t a\left(a + \frac{\sigma_1^2}{2}\right)s^2\,\mathrm{d}s\right)^{\frac{1}{4}} = \left(|\alpha| + \sigma_2\sqrt[4]{6a(2a + \sigma_1^2)}\right)t, \end{aligned}$$

hence we conclude

$$\sup_{s\in[0,T]}\mathbb{E}(\mathcal{X}_s^4) = \mathrm{O}(T^4) \qquad \text{as } T \to \infty. \tag{6.12}$$

Using (6.8), we obtain $E_2(T) = \int_0^T \sqrt{\mathrm{O}(T^4)\,\mathrm{O}(T)}\,\mathrm{d}s = \mathrm{O}(T^{\frac{7}{2}})$ as $T \to \infty$. Hence

$$\mathbb{E}\left(\left|\frac{1}{T^2}\left(\int_0^T X_s\,\mathrm{d}Y_s - \int_0^T \mathcal{X}_s\,\mathrm{d}\mathcal{Y}_s\right)\right|\right) \leqslant \frac{1}{T^2}\left(\sqrt{E_1(T)} + \sqrt{E_2(T)}\right) = \frac{1}{T^2}\,\mathrm{O}(T^{\frac{7}{4}}) \to 0$$

as $T \to \infty$, implying $\frac{1}{T^2}\left(\int_0^T X_s\,\mathrm{d}Y_s - \int_0^T \mathcal{X}_s\,\mathrm{d}\mathcal{Y}_s\right) \xrightarrow{\mathbb{P}} 0$ as $T \to \infty$, i.e., the first convergence in (6.5). Thus we conclude convergence (6.3).

Applying the first equation of (1.1) and using $b = 0$, we obtain

$$\frac{\sigma_1}{T}\int_0^T Y_s^{\frac{1}{2}}\,\mathrm{d}W_s = \frac{1}{T}(Y_T - Y_0) - a.$$

By Itô's formula and using $b = 0$,

$$\begin{aligned} \mathrm{d}(Y_t^2) &= 2Y_t\,\mathrm{d}Y_t + \sigma_1^2 Y_t\,\mathrm{d}t = 2Y_t(a\,\mathrm{d}t + \sigma_1 Y_t^{\frac{1}{2}}\,\mathrm{d}W_t) + \sigma_1^2 Y_t\,\mathrm{d}t \\ &= (2a + \sigma_1^2)Y_t\,\mathrm{d}t + 2\sigma_1 Y_t^{\frac{3}{2}}\,\mathrm{d}W_t, \end{aligned}$$

hence

$$Y_T^2 = Y_0^2 + (2a + \sigma_1^2)\int_0^T Y_s\,\mathrm{d}s + 2\sigma_1\int_0^T Y_s^{\frac{3}{2}}\,\mathrm{d}W_s.$$

Consequently,

$$-\frac{\sigma_1}{T^2}\int_0^T Y_s^{\frac{3}{2}}\,\mathrm{d}W_s = -\frac{1}{2T^2}(Y_T^2 - Y_0^2) + \frac{2a + \sigma_1^2}{2T^2}\int_0^T Y_s\,\mathrm{d}s.$$

In a similar way, applying the second equation of (1.1) and using $\beta = 0$ and $\gamma = 0$, we obtain

$$\frac{\sigma_2}{T}\int_0^T Y_s^{\frac{1}{2}}\,\mathrm{d}\widetilde{W}_s + \frac{\sigma_3}{T}L_T = \frac{1}{T}(X_T - X_0) - \alpha.$$

By Itô's formula and using $\beta = 0$ and $\gamma = 0$,

$$\begin{aligned}\mathrm{d}(Y_t X_t) =& Y_t\,\mathrm{d}X_t + X_t\,\mathrm{d}Y_t + \varrho\sigma_1\sigma_2 Y_t\,\mathrm{d}t = Y_t(\alpha\,\mathrm{d}t + \sigma_2 Y_t^{\frac{1}{2}}\,\mathrm{d}\widetilde{W}_t + \sigma_3\,\mathrm{d}L_t)\\ &+ X_t\,\mathrm{d}Y_t + \varrho\sigma_1\sigma_2 Y_t\,\mathrm{d}t\\ =& (\alpha + \varrho\sigma_1\sigma_2)Y_t\,\mathrm{d}t + \sigma_2 Y_t^{\frac{3}{2}}\,\mathrm{d}\widetilde{W}_t + X_t\,\mathrm{d}Y_t + \sigma_3 Y_t\,\mathrm{d}L_t,\end{aligned}$$

hence

$$Y_T X_T = Y_0 X_0 + (\alpha + \varrho\sigma_1\sigma_2)\int_0^T Y_s\,\mathrm{d}s + \sigma_2\int_0^T Y_s^{\frac{3}{2}}\,\mathrm{d}\widetilde{W}_s + \int_0^T X_s\,\mathrm{d}Y_s + \sigma_3\int_0^T Y_s\,\mathrm{d}L_s.$$

Consequently,

$$\begin{aligned}&-\frac{\sigma_2}{T^2}\int_0^T Y_s^{\frac{3}{2}}\,\mathrm{d}\widetilde{W}_s - \frac{\sigma_3}{T^2}\int_0^T Y_s\,\mathrm{d}L_s\\ &= -\frac{1}{T^2}(Y_T X_T - Y_0 X_0) + \frac{\alpha + \varrho\sigma_1\sigma_2}{T^2}\int_0^T Y_s\,\mathrm{d}s + \frac{1}{T^2}\int_0^T X_s\,\mathrm{d}Y_s.\end{aligned}$$

Again by Itô's formula and using $\beta = 0$ and $\gamma = 0$,

$$\mathrm{d}(X_t^2) = 2X_t\,\mathrm{d}X_t + (\sigma_2^2 Y_t + \sigma_3^2)\,\mathrm{d}t = 2X_t(\alpha\,\mathrm{d}t + \sigma_2 Y_t^{\frac{1}{2}}\,\mathrm{d}\widetilde{W}_t + \sigma_3\,\mathrm{d}L_t) + (\sigma_2^2 Y_t + \sigma_3^2)\,\mathrm{d}t,$$

hence

$$X_T^2 = X_0^2 + \int_0^T (2\alpha X_s + \sigma_2^2 Y_s + \sigma_3^2)\,\mathrm{d}s + 2\sigma_2\int_0^T Y_s^{\frac{1}{2}} X_s\,\mathrm{d}\widetilde{W}_s + 2\sigma_3\int_0^T X_s\,\mathrm{d}L_s.$$

Consequently,

$$\begin{aligned}&-\frac{\sigma_2}{T^2}\int_0^T Y_s^{\frac{1}{2}} X_s\,\mathrm{d}\widetilde{W}_s - \frac{\sigma_3}{T^2}\int_0^T X_s\,\mathrm{d}L_s\\ &= -\frac{1}{2T^2}(X_T^2 - X_0^2) + \frac{\alpha}{T^2}\int_0^T X_s\,\mathrm{d}s + \frac{\sigma_2^2}{2T^2}\int_0^T Y_s\,\mathrm{d}s + \frac{\sigma_3^2}{2T}.\end{aligned}$$

Applying (6.3) and the continuous mapping theorem, we obtain

$$\begin{bmatrix} 1 & -\frac{1}{T^2}\int_0^T Y_s\,\mathrm{d}s \\ -\frac{1}{T^2}\int_0^T Y_s\,\mathrm{d}s & \frac{1}{T^3}\int_0^T Y_s^2\,\mathrm{d}s \end{bmatrix} \xrightarrow{\mathcal{D}} \begin{bmatrix} 1 & -\int_0^1 \mathcal{Y}_s\,\mathrm{d}s \\ -\int_0^1 \mathcal{Y}_s\,\mathrm{d}s & \int_0^1 \mathcal{Y}_s^2\,\mathrm{d}s \end{bmatrix},$$

$$\begin{bmatrix} \frac{1}{T}(Y_T - Y_0) - a \\ -\frac{1}{2T^2}(Y_T^2 - Y_0^2) + \frac{2a+\sigma_1^2}{2T^2}\int_0^T Y_s\,\mathrm{d}s \end{bmatrix} \xrightarrow{\mathcal{D}} \begin{bmatrix} \mathcal{Y}_1 - a \\ -\frac{1}{2}\mathcal{Y}_1^2 + \frac{2a+\sigma_1^2}{2}\int_0^1 \mathcal{Y}_s\,\mathrm{d}s \end{bmatrix},$$

$$\begin{aligned}&\begin{bmatrix} 1 & -\frac{1}{T^2}\int_0^T Y_s\,\mathrm{d}s & -\frac{1}{T^2}\int_0^T X_s\,\mathrm{d}s \\ -\frac{1}{T^2}\int_0^T Y_s\,\mathrm{d}s & \frac{1}{T^3}\int_0^T Y_s^2\,\mathrm{d}s & \frac{1}{T^3}\int_0^T Y_s X_s\,\mathrm{d}s \\ -\frac{1}{T^2}\int_0^T X_s\,\mathrm{d}s & \frac{1}{T^3}\int_0^T Y_s X_s\,\mathrm{d}s & \frac{1}{T^3}\int_0^T X_s^2\,\mathrm{d}s \end{bmatrix}\\ &\xrightarrow{\mathcal{D}} \begin{bmatrix} 1 & -\int_0^1 \mathcal{Y}_s\,\mathrm{d}s & -\int_0^1 \mathcal{X}_s\,\mathrm{d}s \\ -\int_0^1 \mathcal{Y}_s\,\mathrm{d}s & \int_0^1 \mathcal{Y}_s^2\,\mathrm{d}s & \int_0^1 \mathcal{Y}_s\mathcal{X}_s\,\mathrm{d}s \\ -\int_0^1 \mathcal{X}_s\,\mathrm{d}s & \int_0^1 \mathcal{Y}_s\mathcal{X}_s\,\mathrm{d}s & \int_0^1 \mathcal{X}_s^2\,\mathrm{d}s \end{bmatrix},\end{aligned}$$

$$\begin{bmatrix} \frac{1}{T}(X_T - X_0) - \alpha \\ -\frac{1}{T^2}(Y_T X_T - Y_0 X_0) + \frac{\alpha+\varrho\sigma_1\sigma_2}{T^2}\int_0^T Y_s\,\mathrm{d}s + \frac{1}{T^2}\int_0^T X_s\,\mathrm{d}Y_s \\ -\frac{1}{2T^2}(X_T^2 - X_0^2) + \frac{\alpha}{T^2}\int_0^T X_s\,\mathrm{d}s + \frac{\sigma_2^2}{2T^2}\int_0^T Y_s\,\mathrm{d}s + \frac{\sigma_3^2}{2T} \end{bmatrix}$$
$$\xrightarrow{\mathcal{D}} \begin{bmatrix} \mathcal{X}_1 - \alpha \\ -\mathcal{Y}_1\mathcal{X}_1 + (\alpha+\varrho\sigma_1\sigma_2)\int_0^1 \mathcal{Y}_s\,\mathrm{d}s + \int_0^1 \mathcal{X}_s\,\mathrm{d}\mathcal{Y}_s \\ -\frac{1}{2}\mathcal{X}_1^2 + \alpha\int_0^1 \mathcal{X}_s\,\mathrm{d}s + \frac{\sigma_2^2}{2}\int_0^1 \mathcal{Y}_s\,\mathrm{d}s \end{bmatrix}$$

jointly as $T \to \infty$. Applying again the continuous mapping theorem, we conclude (6.1), since the limiting random matrices in the first and third convergences above are almost surely invertible by Lemma 3.1. □

7 Asymptotic behavior of CLSE: supercritical case

First we present an auxiliary lemma about the asymptotic behavior of $\mathbb{E}(X_t^2)$ as $t \to \infty$.

Lemma 7.1 *Let us consider the two-factor affine diffusion model (1.1) with $a \in \mathbb{R}_+$, $b \in \mathbb{R}_{--}$, $\alpha, \beta \in \mathbb{R}$, $\gamma \in (-\infty, b)$, $\sigma_1 \in \mathbb{R}_{++}$, $\sigma_2, \sigma_3 \in \mathbb{R}_+$ and $\varrho \in [-1, 1]$ with a random initial value (η_0, ζ_0) independent of $(W_t, B_t, L_t)_{t\in\mathbb{R}_+}$ satisfying $\mathbb{P}(\eta_0 \in \mathbb{R}_+) = 1$. Then $\sup_{t\in\mathbb{R}_+} \mathrm{e}^{2\gamma t}\mathbb{E}(X_t^2) < \infty$.*

Proof By Proposition B.1,

$$\sup_{t\in\mathbb{R}_+} \mathrm{e}^{bt}\mathbb{E}(Y_t) = \sup_{t\in\mathbb{R}_+} \left(\mathbb{E}(Y_0) + a\int_0^t \mathrm{e}^{bu}\,\mathrm{d}u\right) = \mathbb{E}(Y_0) + a\int_0^\infty \mathrm{e}^{bu}\,\mathrm{d}u < \infty,$$

since $b < 0$. Moreover,

$$\begin{aligned} \sup_{t\in\mathbb{R}_+} \mathrm{e}^{\gamma t}|\mathbb{E}(X_t)| &= \sup_{t\in\mathbb{R}_+}\left|\mathbb{E}(X_0) + \alpha\int_0^t \mathrm{e}^{\gamma u}\,\mathrm{d}u - \beta\int_0^t \mathrm{e}^{\gamma u}\mathbb{E}(Y_u)\,\mathrm{d}u\right| \\ &\leqslant |\mathbb{E}(X_0)| + |\alpha|\int_0^\infty \mathrm{e}^{\gamma u}\,\mathrm{d}u \\ &\quad + |\beta|\left(\sup_{u\in\mathbb{R}_+} \mathrm{e}^{bu}\mathbb{E}(Y_u)\right)\int_0^\infty \mathrm{e}^{(\gamma-b)u}\,\mathrm{d}u < \infty, \end{aligned}$$

using $\gamma < 0$ and $\gamma - b < 0$. Again by Proposition B.1,

$$\begin{aligned} \sup_{t\in\mathbb{R}_+} \mathrm{e}^{2bt}\mathbb{E}(Y_t^2) &= \sup_{t\in\mathbb{R}_+}\left(\mathbb{E}(Y_0^2) + (2a+\sigma_1^2)\int_0^t \mathrm{e}^{2bu}\mathbb{E}(Y_u)\,\mathrm{d}u\right) \\ &\leqslant \mathbb{E}(Y_0^2) + (2a+\sigma_1^2)\left(\sup_{u\in\mathbb{R}_+} \mathrm{e}^{bu}\mathbb{E}(Y_u)\right)\int_0^\infty \mathrm{e}^{bu}\,\mathrm{d}u < \infty, \end{aligned}$$

using $b < 0$. Hence

$$\begin{aligned} \sup_{t\in\mathbb{R}_+} \mathrm{e}^{(b+\gamma)t}|\mathbb{E}(Y_t X_t)| = \sup_{t\in\mathbb{R}_+}\Big|&\mathbb{E}(Y_0 X_0) + a\int_0^t \mathrm{e}^{(b+\gamma)u}\mathbb{E}(X_u)\,\mathrm{d}u \\ &+ (\alpha+\varrho\sigma_1\sigma_2)\int_0^t \mathrm{e}^{(b+\gamma)u}\mathbb{E}(Y_u)\,\mathrm{d}u - \beta\int_0^t \mathrm{e}^{(b+\gamma)u}\mathbb{E}(Y_u^2)\,\mathrm{d}u\Big| \end{aligned}$$

$$\leqslant |\mathbb{E}(Y_0X_0)| + a\biggl(\sup_{u\in\mathbb{R}_+} \mathrm{e}^{\gamma u}|\mathbb{E}(X_u)|\biggr)\int_0^\infty \mathrm{e}^{bu}\,\mathrm{d}u$$
$$+ (|\alpha| + |\varrho|\sigma_1\sigma_2)\biggl(\sup_{u\in\mathbb{R}_+} \mathrm{e}^{bu}\mathbb{E}(Y_u)\biggr)\int_0^\infty \mathrm{e}^{\gamma u}\,\mathrm{d}u$$
$$+ |\beta|\biggl(\sup_{u\in\mathbb{R}_+} \mathrm{e}^{2bu}\mathbb{E}(Y_u^2)\biggr)\int_0^\infty \mathrm{e}^{(\gamma-b)u}\,\mathrm{d}u < \infty,$$

using $b < 0$, $\gamma < 0$ and $\gamma - b < 0$. Consequently,

$$\sup_{t\in\mathbb{R}_+} \mathrm{e}^{2\gamma t}\mathbb{E}(X_t^2) = \sup_{t\in\mathbb{R}_+}\biggl(\mathbb{E}(X_0^2) + \alpha\int_0^t \mathrm{e}^{2\gamma u}X_u\,\mathrm{d}u - 2\beta\int_0^t \mathrm{e}^{2\gamma u}Y_uX_u\,\mathrm{d}u$$
$$+ \sigma_2^2\int_0^t \mathrm{e}^{2\gamma u}Y_u\,\mathrm{d}u + \sigma_3^2\int_0^t \mathrm{e}^{2\gamma u}\,\mathrm{d}u\biggr)$$
$$\leqslant \mathbb{E}(X_0^2) + |\alpha|\biggl(\sup_{u\in\mathbb{R}_+} \mathrm{e}^{\gamma u}|\mathbb{E}(X_u)|\biggr)\int_0^\infty \mathrm{e}^{\gamma u}\,\mathrm{d}u$$
$$+ 2|\beta|\biggl(\sup_{u\in\mathbb{R}_+} \mathrm{e}^{(b+\gamma)u}|\mathbb{E}(Y_uX_u)|\biggr)\int_0^\infty \mathrm{e}^{(\gamma-b)u}\,\mathrm{d}u$$
$$+ \sigma_2^2\biggl(\sup_{u\in\mathbb{R}_+} \mathrm{e}^{bu}\mathbb{E}(Y_u)\biggr)\int_0^\infty \mathrm{e}^{(2\gamma-b)u}\,\mathrm{d}u + \sigma_3^2\int_0^\infty \mathrm{e}^{2\gamma u}\,\mathrm{d}u < \infty$$

using $\gamma < 0$, $\gamma - b < 0$ and $2\gamma - b < 0$. □

Next we present an auxiliary lemma about the asymptotic behavior of X_t as $t \to \infty$.

Lemma 7.2 *Let us consider the two-factor affine diffusion model (1.1) with* $a \in \mathbb{R}_+$, $b \in \mathbb{R}_{--}$, $\alpha, \beta \in \mathbb{R}$, $\gamma \in (-\infty, b)$, $\sigma_1 \in \mathbb{R}_{++}$, $\sigma_2, \sigma_3 \in \mathbb{R}_+$ *and* $\varrho \in [-1, 1]$ *with a random initial value* (η_0, ζ_0) *independent of* $(W_t, B_t, L_t)_{t\in\mathbb{R}_+}$ *satisfying* $\mathbb{P}(\eta_0 \in \mathbb{R}_+) = 1$. *Suppose that* $\alpha\beta \in \mathbb{R}_-$. *Then there exists a random variable* V_X *such that*

$$\mathrm{e}^{\gamma t}X_t \xrightarrow{\text{a.s.}} V_X \qquad \text{as } t \to \infty \tag{7.1}$$

and, for each $k, \ell \in \mathbb{Z}_+$ *with* $k + \ell > 0$,

$$\mathrm{e}^{(kb+\ell\gamma)t}\int_0^t Y_u^kX_u^\ell\,\mathrm{d}u \xrightarrow{\text{a.s.}} -\frac{V_Y^kV_X^\ell}{kb+\ell\gamma} \qquad \text{as } t \to \infty, \tag{7.2}$$

where V_Y *is given in (4.6). If, in addition,* $\sigma_3 \in \mathbb{R}_{++}$ *or* $\bigl(a - \frac{\sigma_1^2}{2}\bigr)(1-\varrho^2)\sigma_2^2 \in \mathbb{R}_{++}$, *then the distribution of the random variable* V_X *is absolutely continuous. Particularly,* $\mathbb{P}(V_X \neq 0) = 1$.

Proof By (2.2),

$$\mathbb{E}(X_t \mid \mathcal{F}_s) = \mathbb{E}(X_t \mid Y_s, X_s) = \mathrm{e}^{-\gamma(t-s)}X_s + \int_s^t \mathrm{e}^{-\gamma(t-u)}(\alpha - \beta Y_u)\,\mathrm{d}u$$

for all $s, t \in \mathbb{R}_+$ with $0 \leqslant s \leqslant t$. If $\alpha \in \mathbb{R}_+$ and $\beta \in \mathbb{R}_-$, then

$$\mathbb{E}(\mathrm{e}^{\gamma t}X_t \mid \mathcal{F}_s^{Y,X}) = \mathrm{e}^{\gamma s}X_s + \int_s^t \mathrm{e}^{\gamma u}(\alpha - \beta Y_u)\,\mathrm{d}u \geqslant \mathrm{e}^{\gamma s}X_s$$

for all $s, t \in \mathbb{R}_+$ with $0 \leqslant s \leqslant t$, consequently, the process $(e^{\gamma t} X_t)_{t\in\mathbb{R}_+}$ is a submartingale with respect to the filtration $(\mathcal{F}_t^{Y,X})_{t\in\mathbb{R}_+}$. If $\alpha \in \mathbb{R}_-$ and $\beta \in \mathbb{R}_+$, then

$$\mathbb{E}(e^{\gamma t} X_t \mid \mathcal{F}_s^{Y,X}) = e^{\gamma s} X_s + \int_s^t e^{\gamma u}(\alpha - \beta Y_u)\, du \leqslant e^{\gamma s} X_s$$

for all $s, t \in \mathbb{R}_+$ with $0 \leqslant s \leqslant t$, consequently, the process $(e^{\gamma t} X_t)_{t\in\mathbb{R}_+}$ is a supermartingale with respect to the filtration $(\mathcal{F}_t^{Y,X})_{t\in\mathbb{R}_+}$, hence the process $(-e^{\gamma t} X_t)_{t\in\mathbb{R}_+}$ is a submartingale with respect to the filtration $(\mathcal{F}_t^{Y,X})_{t\in\mathbb{R}_+}$. In both cases, $\sup_{t\in\mathbb{R}_+} \mathbb{E}(|e^{\gamma t} X_t|^2) < \infty$, see Lemma 7.1. Hence, by the submartingale convergence theorem, there exists a random variable V_X such that (7.1) holds.

If $\omega \in \Omega$ such that $\mathbb{R}_+ \ni t \mapsto (Y_t(\omega), X_t(\omega))$ is continuous and $(e^{bt} Y_t(\omega), e^{\gamma t} X_t(\omega)) \to (V_Y(\omega), V_X(\omega))$ as $t \to \infty$, then, by the integral Kronecker Lemma 4.3 with $f(t) = e^{(kb+\ell\gamma)t} Y_t(\omega)^k X_t(\omega)^\ell$ and $a(t) = e^{-(kb+\ell\gamma)t}$, $t \in \mathbb{R}_+$, we have

$$\begin{aligned} &\frac{1}{\int_0^t e^{-(kb+\ell\gamma)u}\, du} \int_0^t e^{-(kb+\ell\gamma)u} (e^{(kb+\ell\gamma)u} Y_u(\omega)^k X_u(\omega)^\ell)\, du \\ &\quad \to V_Y(\omega)^k V_X(\omega)^\ell \qquad \text{as } t \to \infty. \end{aligned}$$

Here $\int_0^t e^{-(kb+\ell\gamma)u}\, du = -\frac{e^{-(kb+\ell\gamma)t} - 1}{kb+\ell\gamma}$, $t \in \mathbb{R}_+$, thus we conclude (7.2).

Now suppose that $\sigma_3 \in \mathbb{R}_{++}$ or $\big(a - \frac{\sigma_1^2}{2}\big)(1-\varrho^2)\sigma_2^2 \in \mathbb{R}_{++}$. We are going to show that the random variable V_X is absolutely continuous. Put $Z_t := X_t - rY_t$, $t \in \mathbb{R}_+$ with $r := \frac{\sigma_2 \varrho}{\sigma_1}$. Then the process $(Y_t, Z_t)_{t\in\mathbb{R}_+}$ is an affine process satisfying

$$\begin{cases} dY_t = (a - bY_t)\, dt + \sigma_1 \sqrt{Y_t}\, dW_t, \\ dZ_t = (A - BY_t - \gamma Z_t)\, dt + \Sigma_2 \sqrt{Y_t}\, dB_t + \sigma_3\, dL_t. \end{cases} \qquad t \in \mathbb{R}_+,$$

where $A := \alpha - ra$, $B := \beta - r(b-\gamma)$ and $\Sigma_2 := \sigma_2\sqrt{1-\varrho^2}$, see (Bolyog and Pap 2016, Proposition 2.5). We have

$$\begin{aligned} e^{\gamma t} X_t &= r e^{\gamma t} Y_t + e^{\gamma t} Z_t \\ &= r e^{\gamma t} Y_t + Z_0 + \int_0^t e^{\gamma u}(A - BY_u)\, du + \Sigma_2 \int_0^t e^{\gamma u} \sqrt{Y_u}\, dB_u \\ &\quad + \sigma_3 \int_0^t e^{\gamma u}\, dL_u, \end{aligned}$$

where we used (2.2) with $s = 0$ multiplied both sides by $e^{\gamma t}$. Thus the conditional distribution of $e^{\gamma t} X_t$ given $(Y_u)_{u\in[0,t]}$ and X_0 is a normal distribution with mean $r e^{\gamma t} Y_t + Z_0 + \int_0^t e^{\gamma u}(A - BY_u)\, du$ and with variance $\Sigma_2^2 \int_0^t e^{2\gamma u} Y_u\, du + \sigma_3^2 \int_0^t e^{2\gamma u}\, du$. Hence

$$\begin{aligned} &\mathbb{E}\big(e^{i\lambda e^{\gamma t} X_t} \,\big|\, (Y_u)_{u\in[0,t]}, X_0\big) \\ &= \exp\bigg\{ i\lambda \bigg(r e^{\gamma t} Y_t + Z_0 + \int_0^t e^{\gamma u}(A - BY_u)\, du \bigg) \\ &\quad - \frac{\lambda^2}{2} \bigg(\Sigma_2^2 \int_0^t e^{2\gamma u} Y_u\, du + \sigma_3^2 \int_0^t e^{2\gamma u}\, du \bigg) \bigg\}. \end{aligned}$$

Consequently,

$$\begin{aligned}\big|\mathbb{E}\big(\mathrm{e}^{\mathrm{i}\lambda \mathrm{e}^{\gamma t} X_t}\big)\big| &= \big|\mathbb{E}\big(\mathbb{E}\big(\mathrm{e}^{\mathrm{i}\lambda \mathrm{e}^{\gamma t} X_t} \,\big|\, (Y_u)_{u\in[0,t]}, X_0\big)\big)\big| \\ &= \bigg|\mathbb{E}\bigg(\exp\bigg\{\mathrm{i}\lambda\bigg(r\mathrm{e}^{\gamma t} Y_t + Z_0 + \int_0^t \mathrm{e}^{\gamma u}(A - BY_u)\,\mathrm{d}u\bigg) \\ &\quad - \frac{\lambda^2}{2}\bigg(\varSigma_2^2 \int_0^t \mathrm{e}^{2\gamma u} Y_u\,\mathrm{d}u + \sigma_3^2 \int_0^t \mathrm{e}^{2\gamma u}\,\mathrm{d}u\bigg)\bigg\}\bigg)\bigg| \\ &\leqslant \mathbb{E}\bigg(\bigg|\exp\bigg\{\mathrm{i}\lambda\bigg(r\mathrm{e}^{\gamma t} Y_t + Z_0 + \int_0^t \mathrm{e}^{\gamma u}(A - BY_u)\,\mathrm{d}u\bigg) \\ &\quad - \frac{\lambda^2}{2}\bigg(\varSigma_2^2 \int_0^t \mathrm{e}^{2\gamma u} Y_u\,\mathrm{d}u + \sigma_3^2 \int_0^t \mathrm{e}^{2\gamma u}\,\mathrm{d}u\bigg)\bigg\}\bigg|\bigg) \\ &= \mathbb{E}\bigg(\exp\bigg\{-\frac{\lambda^2}{2}\bigg(\varSigma_2^2 \int_0^t \mathrm{e}^{2\gamma u} Y_u\,\mathrm{d}u + \sigma_3^2 \int_0^t \mathrm{e}^{2\gamma u}\,\mathrm{d}u\bigg)\bigg\}\bigg).\end{aligned}$$

Convergence (7.1) implies $\mathrm{e}^{\gamma t} X_t \xrightarrow{\mathcal{D}} V_X$ as $t \to \infty$, hence, by the continuity theorem and by the monotone convergence theorem,

$$\begin{aligned}\big|\mathbb{E}\big(\mathrm{e}^{\mathrm{i}\lambda V_X}\big)\big| &= \lim_{t\to\infty}\big|\mathbb{E}\big(\mathrm{e}^{\mathrm{i}\lambda \mathrm{e}^{\gamma t} X_t}\big)\big| \\ &\leqslant \lim_{t\to\infty} \mathbb{E}\bigg(\exp\bigg\{-\frac{\lambda^2}{2}\bigg(\varSigma_2^2 \int_0^t \mathrm{e}^{2\gamma u} Y_u\,\mathrm{d}u + \sigma_3^2 \int_0^t \mathrm{e}^{2\gamma u}\,\mathrm{d}u\bigg)\bigg\}\bigg) \\ &= \mathbb{E}\bigg(\exp\bigg\{-\frac{\lambda^2}{2}\bigg(\varSigma_2^2 \int_0^\infty \mathrm{e}^{2\gamma u} Y_u\,\mathrm{d}u + \sigma_3^2 \int_0^\infty \mathrm{e}^{2\gamma u}\,\mathrm{d}u\bigg)\bigg\}\bigg).\end{aligned}$$

for all $\lambda \in \mathbb{R}$. If $\sigma_3 \in \mathbb{R}_{++}$, then we have

$$\big|\mathbb{E}\big(\mathrm{e}^{\mathrm{i}\lambda V_X}\big)\big| \leqslant \exp\left\{-\frac{\sigma_3^2}{4(-\gamma)}\lambda^2\right\}$$

for all $\lambda \in \mathbb{R}$, hence $\int_{-\infty}^{\infty}\big|\mathbb{E}\big(\mathrm{e}^{\mathrm{i}\lambda V_X}\big)\big|\mathrm{d}\lambda < \infty$, implying absolute continuity of the distribution of V_X.

If $\big(a - \frac{\sigma_1^2}{2}\big)(1 - \varrho^2)\sigma_2^2 \in \mathbb{R}_{++}$, then we have

$$\big|\mathbb{E}\big(\mathrm{e}^{\mathrm{i}\lambda V_X}\big)\big| \leqslant \mathbb{E}\bigg(\exp\bigg\{-\frac{\varSigma_2^2}{2}\lambda^2 \int_0^\infty \mathrm{e}^{2\gamma u} Y_u\,\mathrm{d}u\bigg\}\bigg) \leqslant \mathbb{E}\bigg(\exp\bigg\{-\frac{\varSigma_2^2 \mathrm{e}^{4\gamma}}{2}\lambda^2 \int_1^2 Y_u\,\mathrm{d}u\bigg\}\bigg)$$

for all $\lambda \in \mathbb{R}$. Applying the comparison theorem (see, e.g., Karatzas and Shreve 1991, 5.2.18), we obtain $\mathbb{P}(\mathcal{Y}_t \leqslant Y_t \text{ for all } t \in \mathbb{R}_+) = 1$, where $(\mathcal{Y}_t)_{t\in\mathbb{R}_+}$ is the unique strong solution of the SDE

$$\mathrm{d}\mathcal{Y}_t = (a - b\mathcal{Y}_t)\,\mathrm{d}t + \sigma_1\sqrt{\mathcal{Y}_t}\,\mathrm{d}W_t, \qquad t \in [0, \infty),$$

with initial value $\mathcal{Y}_0 = 0$. Consequently, taking into account $\varSigma_2 = \sigma_2\sqrt{1 - \varrho^2} > 0$, we obtain

$$\begin{aligned}\int_{-\infty}^{\infty}\big|\mathbb{E}\big(\mathrm{e}^{\mathrm{i}\lambda V_X}\big)\big|\,\mathrm{d}\lambda &\leqslant \int_{-\infty}^{\infty} \mathbb{E}\bigg(\exp\bigg\{-\frac{\varSigma_2^2 \mathrm{e}^{4\gamma}}{2}\lambda^2 \int_1^2 \mathcal{Y}_u\,\mathrm{d}u\bigg\}\bigg)\mathrm{d}\lambda \\ &= \mathbb{E}\bigg(\int_{-\infty}^{\infty} \exp\bigg\{-\frac{\varSigma_2^2 \mathrm{e}^{4\gamma}}{2}\lambda^2 \int_1^2 \mathcal{Y}_u\,\mathrm{d}u\bigg\}\mathrm{d}\lambda\bigg)\end{aligned}$$

$$= \mathbb{E}\left(\frac{\sqrt{2\pi}}{\Sigma_2 \mathrm{e}^{2\gamma}\sqrt{\int_1^2 \mathcal{Y}_u \, \mathrm{d}u}}\right) = \frac{\sqrt{2\pi}}{\Sigma_2 \mathrm{e}^{2\gamma}} \mathbb{E}\left(\frac{1}{\sqrt{\int_1^2 \mathcal{Y}_u \, \mathrm{d}u}}\right) < \infty,$$

whenever

$$\mathbb{E}\left(\frac{1}{\sqrt{\int_1^2 \mathcal{Y}_u \, \mathrm{d}u}}\right) < \infty. \tag{7.3}$$

By the Cauchy–Schwarz inequality, we have

$$1 = \left(\int_1^2 \sqrt{\mathcal{Y}_u} \cdot \frac{1}{\sqrt{\mathcal{Y}_s}} \, \mathrm{d}u\right)^2 \leqslant \int_1^2 \mathcal{Y}_u \, \mathrm{d}u \int_1^2 \frac{1}{\mathcal{Y}_u} \, \mathrm{d}u,$$

hence

$$\mathbb{E}\left(\frac{1}{\sqrt{\int_1^2 \mathcal{Y}_u \, \mathrm{d}u}}\right) \leqslant \mathbb{E}\left(\sqrt{\int_1^2 \frac{1}{\mathcal{Y}_u} \, \mathrm{d}u}\right) \leqslant \sqrt{\mathbb{E}\left(\int_1^2 \frac{1}{\mathcal{Y}_u} \, \mathrm{d}u\right)} = \sqrt{\int_1^2 \mathbb{E}\Big(\frac{1}{\mathcal{Y}_u}\Big) \mathrm{d}u}.$$

For each $u \in \mathbb{R}_{++}$, we have $\mathcal{Y}_u \stackrel{\mathcal{D}}{=} c(u)\xi$, where the distribution of ξ has a chi-square distribution with degrees of freedom $\frac{4a}{\sigma_1^2}$ and $c(u) := \frac{\sigma_1^2}{4} \int_0^u \mathrm{e}^{-bv} \, \mathrm{d}v = \frac{\sigma_1^2(\mathrm{e}^{-bu}-1)}{4(-b)}$, see Proposition B.1. Hence

$$\mathbb{E}\Big(\frac{1}{\mathcal{Y}_u}\Big) = \frac{1}{c(u)} \mathbb{E}\Big(\frac{1}{\xi}\Big),$$

where $\mathbb{E}\big(\frac{1}{\xi}\big) < \infty$, since the density of ξ has the form

$$\mathbb{R} \ni x \mapsto \frac{1}{2^{\frac{2a}{\sigma_1^2}} \Gamma\big(\frac{2a}{\sigma_1^2}\big)} x^{\frac{2a}{\sigma_1^2}-1} \mathrm{e}^{-\frac{x}{2}} \mathbb{1}_{\mathbb{R}_{++}}(x)$$

and the assumption $a - \frac{\sigma_1^2}{2} > 0$ yields $\frac{2a}{\sigma_1^2} - 1 > 0$. Consequently,

$$\int_1^2 \mathbb{E}\big(\frac{1}{\mathcal{Y}_u}\big) \mathrm{d}u = \mathbb{E}\Big(\frac{1}{\xi}\Big) \int_1^2 \frac{1}{c(u)} \, \mathrm{d}u = \mathbb{E}\Big(\frac{1}{\xi}\Big) \int_1^2 \frac{4(-b)}{\sigma_1^2(\mathrm{e}^{-bu} - 1)} \, \mathrm{d}u < \infty,$$

thus we obtain (7.3), and hence $\int_{-\infty}^{\infty} \big|\mathbb{E}\big(\mathrm{e}^{\mathrm{i}\lambda V_X}\big)\big| \mathrm{d}\lambda < \infty$, and we conclude absolute continuity of the distribution of V_X. □

Theorem 7.3 *Let us consider the two-factor affine diffusion model (1.1) with* $a \in \mathbb{R}_+$, $b \in \mathbb{R}_{--}$, $\alpha, \beta \in \mathbb{R}$, $\gamma \in (-\infty, b)$, $\sigma_1 \in \mathbb{R}_{++}$, $\sigma_2, \sigma_3 \in \mathbb{R}_+$ *and* $\varrho \in [-1, 1]$ *with a random initial value* (η_0, ζ_0) *independent of* $(W_t, B_t, L_t)_{t \in \mathbb{R}_+}$ *satisfying* $\mathbb{P}(\eta_0 \in \mathbb{R}_+) = 1$. *Suppose that* $\alpha\beta \in \mathbb{R}_-$. *Suppose that* $\sigma_3 \in \mathbb{R}_{++}$ *or* $\big(a - \frac{\sigma_1^2}{2}\big)(1 - \varrho^2)\sigma_2^2 \in \mathbb{R}_{++}$. *Then*

$$\begin{bmatrix} T \mathrm{e}^{\frac{bT}{2}} (\widehat{a}_T - a) \\ \mathrm{e}^{-\frac{bT}{2}} (\widehat{b}_T - b) \\ T \mathrm{e}^{\frac{bT}{2}} (\widehat{\alpha}_T - \alpha) \\ \mathrm{e}^{-bT/2} (\widehat{\beta}_T - \beta) \\ \mathrm{e}^{\frac{(b-2\gamma)T}{2}} (\widehat{\gamma}_T - \gamma) \end{bmatrix} \xrightarrow{\mathcal{D}} \boldsymbol{V}^{-1} \boldsymbol{\eta} \boldsymbol{\xi} \tag{7.4}$$

as $T \to \infty$ *with*

$$\boldsymbol{V} := \begin{bmatrix} 1 & \frac{V_Y}{b} & 0 & 0 & 0 \\ 0 & -\frac{V_Y^2}{2b} & 0 & 0 & 0 \\ 0 & 0 & 1 & \frac{V_Y}{b} & \frac{V_X}{\gamma} \\ 0 & 0 & 0 & -\frac{V_Y^2}{2b} & -\frac{V_Y V_X}{b+\gamma} \\ 0 & 0 & 0 & -\frac{V_Y V_X}{b+\gamma} & -\frac{V_X^2}{2\gamma} \end{bmatrix},$$

where V_Y *and* V_X *are given in* (4.6) *and* (7.1)*, respectively,* $\boldsymbol{\eta}$ *is a* 5×5 *random matrix such that*

$$\boldsymbol{\eta}\boldsymbol{\eta}^\top = \begin{bmatrix} -\frac{\sigma_1^2 V_Y}{b} & \frac{\sigma_1^2 V_Y^2}{2b} & -\frac{\varrho\sigma_1\sigma_2 V_Y}{b} & \frac{\varrho\sigma_1\sigma_2 V_Y^2}{2b} & \frac{\varrho\sigma_1\sigma_2 V_Y V_X}{b+\gamma} \\ \frac{\sigma_1^2 V_Y^2}{2b} & -\frac{\sigma_1^2 V_Y^3}{3b} & \frac{\varrho\sigma_1\sigma_2 V_Y^2}{2b} & -\frac{\varrho\sigma_1\sigma_2 V_Y^3}{3b} & -\frac{\varrho\sigma_1\sigma_2 V_Y^2 V_X}{2b+\gamma} \\ -\frac{\varrho\sigma_1\sigma_2 V_Y}{b} & \frac{\varrho\sigma_1\sigma_2 V_Y^2}{2b} & -\frac{\sigma_2^2 V_Y}{b} & \frac{\sigma_2^2 V_Y^2}{2b} & \frac{\sigma_2^2 V_Y V_X}{b+\gamma} \\ \frac{\varrho\sigma_1\sigma_2 V_Y^2}{2b} & -\frac{\varrho\sigma_1\sigma_2 V_Y^3}{3b} & \frac{\sigma_2^2 V_Y^2}{2b} & -\frac{\sigma_2^2 V_Y^3}{3b} & -\frac{\sigma_2^2 V_Y^2 V_X}{2b+\gamma} \\ \frac{\varrho\sigma_1\sigma_2 V_Y V_X}{b+\gamma} & -\frac{\varrho\sigma_1\sigma_2 V_Y^2 V_X}{2b+\gamma} & \frac{\sigma_2^2 V_Y V_X}{b+\gamma} & -\frac{\sigma_2^2 V_Y^2 V_X}{2b+\gamma} & -\frac{\sigma_2^2 V_Y V_X^2}{b+2\gamma} \end{bmatrix},$$

and $\boldsymbol{\xi}$ *is a 5-dimensional standard normally distributed random vector independent of* (V_Y, V_X).

Proof We have

$$\begin{bmatrix} T\mathrm{e}^{\frac{bT}{2}}(\widehat{a}_T - a) \\ \mathrm{e}^{-\frac{bT}{2}}(\widehat{b}_T - b) \\ T\mathrm{e}^{\frac{bT}{2}}(\widehat{\alpha}_T - \alpha) \\ \mathrm{e}^{-\frac{bT}{2}}(\widehat{\beta}_T - \beta) \\ \mathrm{e}^{\frac{(b-2\gamma)T}{2}}(\widehat{\gamma}_T - \gamma) \end{bmatrix} = \operatorname{diag}\Big(T\mathrm{e}^{\frac{bT}{2}}, \mathrm{e}^{-\frac{bT}{2}}, T\mathrm{e}^{\frac{bT}{2}}, \mathrm{e}^{-\frac{bT}{2}}, \mathrm{e}^{\frac{(b-2\gamma)T}{2}}\Big)\big(\widehat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}\big),$$

where, by (3.5),

$$\widehat{\boldsymbol{\theta}}_T - \boldsymbol{\theta} = \boldsymbol{G}_T^{-1}\boldsymbol{h}_T = \begin{bmatrix} \boldsymbol{G}_T^{(1)} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{G}_T^{(2)} \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{h}_T^{(1)} \\ \boldsymbol{h}_T^{(2)} \end{bmatrix}.$$

We are going to apply Theorem C.2 for the continuous local martingale $(\boldsymbol{h}_T)_{T\in\mathbb{R}_+}$ with quadratic variation process $\langle \boldsymbol{h}\rangle_T = \widetilde{\boldsymbol{G}}_T$, $T \in \mathbb{R}_+$ (introduced in the proof of Theorem 5.1). With scaling matrices

$$\boldsymbol{Q}(T) := \operatorname{diag}\Big(\mathrm{e}^{\frac{bT}{2}}, \mathrm{e}^{\frac{3bT}{2}}, \mathrm{e}^{\frac{bT}{2}}, \mathrm{e}^{\frac{3bT}{2}}, \mathrm{e}^{\frac{(b+2\gamma)T}{2}}\Big), \qquad t \in \mathbb{R}_{++},$$

by (7.2), we have

$$\boldsymbol{Q}(T)\langle \boldsymbol{h}\rangle_T \boldsymbol{Q}(T)^\top \xrightarrow{\text{a.s.}} \begin{bmatrix} -\frac{\sigma_1^2 V_Y}{b} & \frac{\sigma_1^2 V_Y^2}{2b} & -\frac{\varrho\sigma_1\sigma_2 V_Y}{b} & \frac{\varrho\sigma_1\sigma_2 V_Y^2}{2b} & \frac{\varrho\sigma_1\sigma_2 V_Y V_X}{b+\gamma} \\ \frac{\sigma_1^2 V_Y^2}{2b} & -\frac{\sigma_1^2 V_Y^3}{3b} & \frac{\varrho\sigma_1\sigma_2 V_Y^2}{2b} & -\frac{\varrho\sigma_1\sigma_2 V_Y^3}{3b} & -\frac{\varrho\sigma_1\sigma_2 V_Y^2 V_X}{2b+\gamma} \\ -\frac{\varrho\sigma_1\sigma_2 V_Y}{b} & \frac{\varrho\sigma_1\sigma_2 V_Y^2}{2b} & -\frac{\sigma_2^2 V_Y}{b} & \frac{\sigma_2^2 V_Y^2}{2b} & \frac{\sigma_2^2 V_Y V_X}{b+\gamma} \\ \frac{\varrho\sigma_1\sigma_2 V_Y^2}{2b} & -\frac{\varrho\sigma_1\sigma_2 V_Y^3}{3b} & \frac{\sigma_2^2 V_Y^2}{2b} & -\frac{\sigma_2^2 V_Y^3}{3b} & -\frac{\sigma_2^2 V_Y^2 V_X}{2b+\gamma} \\ \frac{\varrho\sigma_1\sigma_2 V_Y V_X}{b+\gamma} & -\frac{\varrho\sigma_1\sigma_2 V_Y^2 V_X}{2b+\gamma} & \frac{\sigma_2^2 V_Y V_X}{b+\gamma} & -\frac{\sigma_2^2 V_Y^2 V_X}{2b+\gamma} & -\frac{\sigma_2^2 V_Y V_X^2}{b+2\gamma} \end{bmatrix} = \boldsymbol{\eta}\boldsymbol{\eta}^\top$$

as $T \to \infty$. Hence by Theorem C.2, for each random matrix $\boldsymbol{A}$ defined on $(\Omega, \mathcal{F}, \mathbb{P})$, we obtain

$$(\boldsymbol{Q}(T)\boldsymbol{h}_T, \boldsymbol{A}) \xrightarrow{\mathcal{D}} (\boldsymbol{\eta}\boldsymbol{\xi}, \boldsymbol{A}) \qquad \text{as } T \to \infty, \tag{7.5}$$

where $\boldsymbol{\xi}$ is a 5-dimensional standard normally distributed random vector independent of $(\boldsymbol{\eta}, \boldsymbol{A})$. The aim of the following discussion is to include appropriate scaling matrices for $\boldsymbol{G}_T$. The matrices $\boldsymbol{G}_T^{(1)}$ and $\boldsymbol{G}_T^{(2)}$ can be written in the form

$$\boldsymbol{G}_T^{(1)} = \operatorname{diag}\big(T^{\frac{1}{2}}, \mathrm{e}^{-bT}\big) \begin{bmatrix} 1 & -\frac{\mathrm{e}^{bT}}{\sqrt{T}} \int_0^T Y_s \,\mathrm{d}s \\ -\frac{\mathrm{e}^{bT}}{\sqrt{T}} \int_0^T Y_s \,\mathrm{d}s & \mathrm{e}^{2bT} \int_0^T Y_s^2 \,\mathrm{d}s \end{bmatrix} \operatorname{diag}\big(T^{\frac{1}{2}}, \mathrm{e}^{-bT}\big)$$

and

$$\begin{aligned} \boldsymbol{G}_T^{(2)} &= \operatorname{diag}\big(T^{\frac{1}{2}}, \mathrm{e}^{-bT}, \mathrm{e}^{-\gamma T}\big) \\ &\times \begin{bmatrix} 1 & -\frac{\mathrm{e}^{bT}}{\sqrt{T}} \int_0^T Y_s \,\mathrm{d}s & -\frac{\mathrm{e}^{\gamma T}}{\sqrt{T}} \int_0^T X_s \,\mathrm{d}s \\ -\frac{\mathrm{e}^{bT}}{\sqrt{T}} \int_0^T Y_s \,\mathrm{d}s & \mathrm{e}^{2bT} \int_0^T Y_s^2 \,\mathrm{d}s & \mathrm{e}^{(b+\gamma)T} \int_0^T Y_s X_s \,\mathrm{d}s \\ -\frac{\mathrm{e}^{\gamma T}}{\sqrt{T}} \int_0^T X_s \,\mathrm{d}s & \mathrm{e}^{(b+\gamma)T} \int_0^T Y_s X_s \,\mathrm{d}s & \mathrm{e}^{2\gamma T} \int_0^T X_s^2 \,\mathrm{d}s \end{bmatrix} \\ &\times \operatorname{diag}\big(T^{\frac{1}{2}}, \mathrm{e}^{-bT}, \mathrm{e}^{-\gamma T}\big), \end{aligned}$$

hence the matrices $(\boldsymbol{G}_T^{(1)})^{-1}$ and $(\boldsymbol{G}_T^{(2)})^{-1}$ can be written in the form

$$(\boldsymbol{G}_T^{(1)})^{-1} = \operatorname{diag}\big(T^{-\frac{1}{2}}, \mathrm{e}^{bT}\big) \begin{bmatrix} 1 & -\frac{\mathrm{e}^{bT}}{\sqrt{T}} \int_0^T Y_s \,\mathrm{d}s \\ -\frac{\mathrm{e}^{bT}}{\sqrt{T}} \int_0^T Y_s \,\mathrm{d}s & \mathrm{e}^{2bT} \int_0^T Y_s^2 \,\mathrm{d}s \end{bmatrix}^{-1} \operatorname{diag}\big(T^{-\frac{1}{2}}, \mathrm{e}^{bT}\big)$$

and

$$\begin{aligned} (\boldsymbol{G}_T^{(2)})^{-1} &= \operatorname{diag}\big(T^{-\frac{1}{2}}, \mathrm{e}^{bT}, \mathrm{e}^{\gamma T}\big) \\ &\times \begin{bmatrix} 1 & -\frac{\mathrm{e}^{bT}}{\sqrt{T}} \int_0^T Y_s \,\mathrm{d}s & -\frac{\mathrm{e}^{\gamma T}}{\sqrt{T}} \int_0^T X_s \,\mathrm{d}s \\ -\frac{\mathrm{e}^{bT}}{\sqrt{T}} \int_0^T Y_s \,\mathrm{d}s & \mathrm{e}^{2bT} \int_0^T Y_s^2 \,\mathrm{d}s & \mathrm{e}^{(b+\gamma)T} \int_0^T Y_s X_s \,\mathrm{d}s \\ -\frac{\mathrm{e}^{\gamma T}}{\sqrt{T}} \int_0^T X_s \,\mathrm{d}s & \mathrm{e}^{(b+\gamma)T} \int_0^T Y_s X_s \,\mathrm{d}s & \mathrm{e}^{2\gamma T} \int_0^T X_s^2 \,\mathrm{d}s \end{bmatrix}^{-1} \\ &\times \operatorname{diag}\big(T^{-\frac{1}{2}}, \mathrm{e}^{bT}, \mathrm{e}^{\gamma T}\big). \end{aligned}$$

We have

$$\begin{aligned} &\operatorname{diag}\Big(T\mathrm{e}^{\frac{bT}{2}}, \mathrm{e}^{-\frac{bT}{2}}, T\mathrm{e}^{\frac{bT}{2}}, \mathrm{e}^{-\frac{bT}{2}}, \mathrm{e}^{\frac{(b-2\gamma)T}{2}}\Big) \operatorname{diag}\Big(T^{-\frac{1}{2}}, \mathrm{e}^{bT}, T^{-\frac{1}{2}}, \mathrm{e}^{bT}, \mathrm{e}^{\gamma T}\Big) \\ &\quad = \operatorname{diag}\Big(T^{\frac{1}{2}}\mathrm{e}^{\frac{bT}{2}}, \mathrm{e}^{\frac{bT}{2}}, T^{\frac{1}{2}}\mathrm{e}^{\frac{bT}{2}}, \mathrm{e}^{\frac{bT}{2}}, \mathrm{e}^{\frac{bT}{2}}\Big) \end{aligned}$$

and

$$\begin{aligned} &\operatorname{diag}\Big(T^{-\frac{1}{2}}, \mathrm{e}^{bT}, T^{-\frac{1}{2}}, \mathrm{e}^{bT}, \mathrm{e}^{\gamma T}\Big) \boldsymbol{Q}(T)^{-1} \\ &= \operatorname{diag}\Big(T^{-\frac{1}{2}}, \mathrm{e}^{bT}, T^{-\frac{1}{2}}, \mathrm{e}^{bT}, \mathrm{e}^{\gamma T}\Big) \operatorname{diag}\Big(\mathrm{e}^{-\frac{bT}{2}}, \mathrm{e}^{-\frac{3bT}{2}}, \mathrm{e}^{-\frac{bT}{2}}, \mathrm{e}^{-\frac{3bT}{2}}, \mathrm{e}^{-\frac{(b+2\gamma)T}{2}}\Big) \\ &= \operatorname{diag}\Big(T^{-\frac{1}{2}}\mathrm{e}^{-\frac{bT}{2}}, \mathrm{e}^{-\frac{bT}{2}}, T^{-\frac{1}{2}}\mathrm{e}^{-\frac{bT}{2}}, \mathrm{e}^{-\frac{bT}{2}}, \mathrm{e}^{-\frac{bT}{2}}\Big). \end{aligned}$$

Moreover,

$$\operatorname{diag}\left(T^{\frac{1}{2}}\mathrm{e}^{\frac{bT}{2}}, \mathrm{e}^{\frac{bT}{2}}\right)\begin{bmatrix} 1 & -\frac{\mathrm{e}^{bT}}{\sqrt{T}}\int_0^T Y_s\,\mathrm{d}s \\ -\frac{\mathrm{e}^{bT}}{\sqrt{T}}\int_0^T Y_s\,\mathrm{d}s & \mathrm{e}^{2bT}\int_0^T Y_s^2\,\mathrm{d}s \end{bmatrix}\operatorname{diag}\left(T^{-\frac{1}{2}}\mathrm{e}^{-\frac{bT}{2}}, \mathrm{e}^{-\frac{bT}{2}}\right)$$
$$= \begin{bmatrix} 1 & -\mathrm{e}^{bT}\int_0^T Y_s\,\mathrm{d}s \\ -\frac{\mathrm{e}^{bT}}{T}\int_0^T Y_s\,\mathrm{d}s & \mathrm{e}^{2bT}\int_0^T Y_s^2\,\mathrm{d}s \end{bmatrix} =: \boldsymbol{J}_T^{(1)}$$

and

$$\operatorname{diag}\left(T^{\frac{1}{2}}\mathrm{e}^{\frac{bT}{2}}, \mathrm{e}^{\frac{bT}{2}}, \mathrm{e}^{\frac{bT}{2}}\right)\begin{bmatrix} 1 & -\frac{\mathrm{e}^{bT}}{\sqrt{T}}\int_0^T Y_s\,\mathrm{d}s & -\frac{\mathrm{e}^{\gamma T}}{\sqrt{T}}\int_0^T X_s\,\mathrm{d}s \\ -\frac{\mathrm{e}^{bT}}{\sqrt{T}}\int_0^T Y_s\,\mathrm{d}s & \mathrm{e}^{2bT}\int_0^T Y_s^2\,\mathrm{d}s & \mathrm{e}^{(b+\gamma)T}\int_0^T Y_s X_s\,\mathrm{d}s \\ -\frac{\mathrm{e}^{\gamma T}}{\sqrt{T}}\int_0^T X_s\,\mathrm{d}s & \mathrm{e}^{(b+\gamma)T}\int_0^T Y_s X_s\,\mathrm{d}s & \mathrm{e}^{2\gamma T}\int_0^T X_s^2\,\mathrm{d}s \end{bmatrix}$$
$$\times \operatorname{diag}\left(T^{-\frac{1}{2}}\mathrm{e}^{-\frac{bT}{2}}, \mathrm{e}^{-\frac{bT}{2}}, \mathrm{e}^{-\frac{bT}{2}}\right)$$
$$= \begin{bmatrix} 1 & -\mathrm{e}^{bT}\int_0^T Y_s\,\mathrm{d}s & -\mathrm{e}^{\gamma T}\int_0^T X_s\,\mathrm{d}s \\ -\frac{\mathrm{e}^{bT}}{T}\int_0^T Y_s\,\mathrm{d}s & \mathrm{e}^{2bT}\int_0^T Y_s^2\,\mathrm{d}s & \mathrm{e}^{(b+\gamma)T}\int_0^T Y_s X_s\,\mathrm{d}s \\ -\frac{\mathrm{e}^{\gamma T}}{T}\int_0^T X_s\,\mathrm{d}s & \mathrm{e}^{(b+\gamma)T}\int_0^T Y_s X_s\,\mathrm{d}s & \mathrm{e}^{2\gamma T}\int_0^T X_s^2\,\mathrm{d}s \end{bmatrix} =: \boldsymbol{J}_T^{(2)}$$

Consequently,

$$\begin{bmatrix} T\mathrm{e}^{\frac{bT}{2}}(\widehat{a}_T - a) \\ \mathrm{e}^{-\frac{bT}{2}}(\widehat{b}_T - b) \\ T\mathrm{e}^{\frac{bT}{2}}(\widehat{\alpha}_T - \alpha) \\ \mathrm{e}^{-\frac{bT}{2}}(\widehat{\beta}_T - \beta) \\ \mathrm{e}^{\frac{(b-2\gamma)T}{2}}(\widehat{\gamma}_T - \gamma) \end{bmatrix} = \operatorname{diag}\left(\boldsymbol{J}_T^{(1)}, \boldsymbol{J}_T^{(2)}\right)^{-1}\boldsymbol{Q}(T)\boldsymbol{h}_T,$$

where, by Lemma 7.2,

$$\operatorname{diag}\left(\boldsymbol{J}_T^{(1)}, \boldsymbol{J}_T^{(2)}\right) \xrightarrow{\mathbb{P}} \boldsymbol{V} \qquad \text{as } T \to \infty. \tag{7.6}$$

By (7.5) with $\boldsymbol{A} = \boldsymbol{V}$, by (7.6) and by Theorem 2.7 (iv) of van der Vaart (1998), we obtain

$$\left(\boldsymbol{Q}(T)\boldsymbol{h}_T, \operatorname{diag}\left(\boldsymbol{J}_T^{(1)}, \boldsymbol{J}_T^{(2)}\right)\right) \xrightarrow{\mathcal{D}} (\boldsymbol{\eta}\boldsymbol{\xi}, \boldsymbol{V}) \qquad \text{as } T \to \infty.$$

The random matrix $\boldsymbol{V}$ is invertible almost surely, since

$$\det(\boldsymbol{V}) = -\frac{(b-\gamma)^2 V_Y^4 V_X^2}{8(b+\gamma)^2 b^2 \gamma} > 0$$

almost surely by Lemma 7.2. Consequently, $\operatorname{diag}\left(\boldsymbol{J}_T^{(1)}, \boldsymbol{J}_T^{(2)}\right)^{-1}\boldsymbol{Q}(T)\boldsymbol{h}_T \xrightarrow{\mathcal{D}} \boldsymbol{V}^{-1}\boldsymbol{\eta}\boldsymbol{\xi}$ as $T \to \infty$. □

8 Summary

The following table summarize the results of the present paper on the asymptotic properties of the CLSE $(\widehat{a}_T, \widehat{b}_T, \widehat{\alpha}_T, \widehat{\beta}_T, \widehat{\gamma}_T)$ for the drift parameters $(a, b, \alpha, \beta, \gamma)$ of general two-factor affine diffusions (1.1). We recall that $a \in [0, \infty)$, $b, \alpha, \beta, \gamma \in \mathbb{R}$, $\sigma_1, \sigma_2, \sigma_3 \in [0, \infty)$ and $\varrho \in [-1, 1]$.

$b > 0,\ \gamma > 0$	$a > 0,\ \sigma_1 > 0,$ $(1-\varrho^2)\sigma_2^2 + \sigma_3^2 > 0$	$(\widehat{a}_T, \widehat{b}_T, \widehat{\alpha}_T, \widehat{\beta}_T, \widehat{\gamma}_T)$ is strongly consistent and asymptotically normal
$b = 0,\ \gamma = 0$	$\beta = 0,\ (1-\varrho^2)\sigma_2^2 + \sigma_3^2 > 0$	$(\widehat{b}_T, \widehat{\beta}_T, \widehat{\gamma}_T)$ is weakly consistent
		$\widehat{a}_T$ and $\widehat{\alpha}_T$ are not weakly consistent
		asymptotic behavior of $(\widehat{a}_T, \widehat{b}_T, \widehat{\alpha}_T, \widehat{\beta}_T, \widehat{\gamma}_T)$
$\gamma < \beta < 0$	$\alpha\beta \leqslant 0,\ \sigma_1 > 0,$ $(a - \frac{\sigma_1^2}{2})(1-\varrho^2)\sigma_2^2 + \sigma_3^2 > 0$	$\widehat{b}_T$ is strongly consistent, $(\widehat{\beta}_T, \widehat{\gamma}_T)$ is weakly consistent
		$(\widehat{a}_T, \widehat{b}_T, \widehat{\alpha}_T, \widehat{\beta}_T, \widehat{\gamma}_T)$ is asymptotically mixed normal

For comparison, the following table summarize the results of Barczy and Pap (2016) on the asymptotic properties of the MLE $(\widetilde{a}_T, \widetilde{b}_T, \widetilde{\alpha}_T, \widetilde{\beta}_T)$ for the drift parameters (a, b, α, β) of a Heston model, which is a submodel of (1.1) with $a \geqslant \frac{\sigma_1^2}{2}$, $\sigma_1, \sigma_2 \in (0, \infty)$, $\gamma = 0$, $\varrho \in (-1, 1)$ and $\sigma_3 = 0$.

$b > 0$	$a = \frac{\sigma_1^2}{2}$	$(\widetilde{a}_T, \widetilde{b}_T, \widetilde{\alpha}_T, \widetilde{\beta}_T)$ is weakly consistent
	$a > \frac{\sigma_1^2}{2}$	$(\widetilde{a}_T, \widetilde{b}_T, \widetilde{\alpha}_T, \widetilde{\beta}_T)$ is strongly consistent and asymptotically normal
$b = 0$	$a > \frac{\sigma_1^2}{2}$	$(\widetilde{a}_T, \widetilde{b}_T, \widetilde{\alpha}_T, \widetilde{\beta}_T)$ is weakly consistent
		$(\widetilde{a}_T, \widetilde{\alpha}_T)$ is asymptotically normal
		asymptotic behavior of $(\widetilde{a}_T, \widetilde{b}_T, \widetilde{\alpha}_T, \widetilde{\beta}_T)$
$b < 0$		$\widetilde{b}_T$ is strongly consistent, $\widetilde{\beta}_T$ is weakly consistent
		$\widetilde{a}_T$ and $\widetilde{\alpha}_T$ are not weakly consistent
		$(\widetilde{b}_T, \widetilde{\beta}_T)$ is asymptotically mixed normal
		asymptotic behavior of $(\widetilde{a}_T, \widetilde{b}_T, \widetilde{\alpha}_T, \widetilde{\beta}_T)$

Appendix

A Stationarity and exponential ergodicity

The following result states the existence of a unique stationary distribution of the affine diffusion process given by the SDE (1.1), see Bolyog and Pap (2016, Theorem 3.1). Let $\mathbb{C}_- := \{z \in \mathbb{C} : \mathrm{Re}(z) \leqslant 0\}$.

Theorem A.1 *Let us consider the two-factor affine diffusion model* (1.1) *with* $a \in \mathbb{R}_+$, $b \in \mathbb{R}_{++}$, $\alpha, \beta \in \mathbb{R}$, $\gamma \in \mathbb{R}_{++}$, $\sigma_1, \sigma_2, \sigma_3 \in \mathbb{R}_+$, $\varrho \in [-1, 1]$, *and with a random initial value* (η_0, ζ_0) *independent of* $(W_t, B_t, L_t)_{t \in \mathbb{R}_+}$ *satisfying* $\mathbb{P}(\eta_0 \in \mathbb{R}_+) = 1$. *Then*

(i) $(Y_t, X_t) \xrightarrow{\mathcal{D}} (Y_\infty, X_\infty)$ *as* $t \to \infty$, *and we have*

$$\mathbb{E}\big(\mathrm{e}^{u_1 Y_\infty + \mathrm{i}\lambda_2 X_\infty}\big) = \exp\left\{a \int_0^\infty \kappa_s(u_1, \lambda_2)\,\mathrm{d}s + \mathrm{i}\frac{\alpha}{\gamma}\lambda_2 - \frac{\sigma_3^2}{4\gamma}\lambda_2^2\right\} \tag{A.1}$$

for $(u_1, \lambda_2) \in \mathbb{C}_- \times \mathbb{R}$, *where* $\kappa_t(u_1, \lambda_2)$, $t \in \mathbb{R}_+$, *is the unique solution of the (deterministic) differential equation*

$$\begin{cases} \frac{\partial \kappa_t}{\partial t}(u_1, \lambda_2) = -b\kappa_t(u_1, \lambda_2) - \mathrm{i}\beta \mathrm{e}^{-\gamma t}\lambda_2 + \frac{1}{2}\sigma_1^2 \kappa_t(u_1, \lambda_2)^2 \\ \qquad\qquad + \mathrm{i}\varrho\sigma_1\sigma_2 \mathrm{e}^{-\gamma t}\lambda_2 \kappa_t(u_1, \lambda_2) - \frac{1}{2}\sigma_2^2 \mathrm{e}^{-2\gamma t}\lambda_2^2, \\ \kappa_0(u_1, \lambda_2) = u_1; \end{cases} \tag{A.2}$$

(ii) *supposing that the random initial value* (η_0, ζ_0) *has the same distribution as* (Y_∞, X_∞) *given in part* (i), $(Y_t, X_t)_{t\in\mathbb{R}_+}$ *is strictly stationary.*

In the subcritical case, the following result states the exponential ergodicity and a strong law of large numbers for the process $(Y_t, X_t)_{t\in\mathbb{R}_+}$, see Bolyog and Pap (2016, Theorem 4.1).

Theorem A.2 *Let us consider the two-factor affine diffusion model* (1.1) *with* $a, b \in \mathbb{R}_{++}$, $\alpha, \beta \in \mathbb{R}$, $\gamma \in \mathbb{R}_{++}$, $\sigma_1 \in \mathbb{R}_{++}$, $\sigma_2, \sigma_3 \in \mathbb{R}_+$ *and* $\varrho \in [-1, 1]$ *with a random initial value* (η_0, ζ_0) *independent of* $(W_t, B_t, L_t)_{t\in\mathbb{R}_+}$ *satisfying* $\mathbb{P}(\eta_0 \in \mathbb{R}_+) = 1$. *Suppose that* $(1-\varrho^2)\sigma_2^2 + \sigma_3^2 > 0$. *Then the process* $(Y_t, X_t)_{t\in\mathbb{R}_+}$ *is exponentially ergodic, namely, there exist* $\delta \in \mathbb{R}_{++}$, $B \in \mathbb{R}_{++}$ *and* $\kappa \in \mathbb{R}_{++}$, *such that*

$$\sup_{|g|\leqslant V+1} \left|\mathbb{E}\big(g(Y_t, X_t) \mid (Y_0, X_0) = (y_0, x_0)\big) - \mathbb{E}(g(Y_\infty, X_\infty))\right| \leqslant B(V(y_0, x_0) + 1)\mathrm{e}^{-\delta t} \tag{A.3}$$

for all $t \in \mathbb{R}_+$ *and* $(y_0, x_0) \in \mathbb{R}_+ \times \mathbb{R}$, *where the supremum is running for Borel measurable functions* $g : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$,

$$V(y, x) := y^2 + \kappa x^2, \qquad (y, x) \in \mathbb{R}_+ \times \mathbb{R}, \tag{A.4}$$

and the distribution of (Y_∞, X_∞) *is given by* (A.1) *and* (A.2). *Moreover, for all Borel measurable functions* $f : \mathbb{R}^2 \to \mathbb{R}$ *with* $\mathbb{E}(|f(Y_\infty, X_\infty)|) < \infty$, *we have*

$$\mathbb{P}\left(\lim_{T\to\infty} \frac{1}{T}\int_0^T f(Y_s, X_s)\,\mathrm{d}s = \mathbb{E}(f(Y_\infty, X_\infty))\right) = 1. \tag{A.5}$$

B Moments

The next proposition gives a recursive formula for the moments of the process $(Y_t, X_t)_{t\in\mathbb{R}_+}$.

Proposition B.1 *Let us consider the two-factor affine diffusion model* (1.1) *with* $a \in \mathbb{R}_+$, $b, \alpha, \beta, \gamma \in \mathbb{R}$, $\sigma_1, \sigma_2, \sigma_3 \in \mathbb{R}_+$, $\varrho \in [-1, 1]$. *Suppose that* $\mathbb{E}(Y_0^n |X_0|^p) < \infty$ *for some* $n, p \in \mathbb{Z}_+$. *Then for each* $t \in \mathbb{R}_+$, *we have* $\mathbb{E}(Y_t^k |X_t|^\ell) < \infty$ *for all* $k \in \{0, \dots, n\}$ *and* $\ell \in \{0, \dots, p\}$, *and the recursion*

$$\begin{aligned} \mathbb{E}(Y_t^k X_t^\ell) = {} & \mathrm{e}^{-(kb+\ell\gamma)t}\mathbb{E}(Y_0^k X_0^\ell) + \left(ka + \frac{1}{2}k(k-1)\sigma_1^2\right)\int_0^t \mathrm{e}^{-(kb+\ell\gamma)(t-u)}\mathbb{E}(Y_u^{k-1} X_u^\ell)\,\mathrm{d}u \\ & + (\alpha + k\varrho\sigma_1\sigma_2)\int_0^t \mathrm{e}^{-(kb+\ell\gamma)(t-u)}\mathbb{E}(Y_u^k X_u^{\ell-1})\,\mathrm{d}u \\ & - \ell\beta \int_0^t \mathrm{e}^{-(kb+\ell\gamma)(t-u)}\mathbb{E}(Y_u^{k+1} X_u^{\ell-1})\,\mathrm{d}u \\ & + \frac{1}{2}\ell(\ell-1)\sigma_2^2 \int_0^t \mathrm{e}^{-(kb+\ell\gamma)(t-u)}\mathbb{E}(Y_u^{k+1} X_u^{\ell-2})\,\mathrm{d}u \\ & + \frac{1}{2}\ell(\ell-1)\sigma_3^2 \int_0^t \mathrm{e}^{-(kb+\ell\gamma)(t-u)}\mathbb{E}(Y_u^k X_u^{\ell-2})\,\mathrm{d}u \end{aligned}$$

for all $t \in \mathbb{R}_+$*, where* $\mathbb{E}(Y_t^i X_t^j) := 0$ *if* $i, j \in \mathbb{Z}$ *with* $i < 0$ *or* $j < 0$*. Especially,*

$$\begin{aligned}
\mathbb{E}(Y_t) &= \mathrm{e}^{-bt}\mathbb{E}(Y_0) + a\int_0^t \mathrm{e}^{-b(t-u)}\,\mathrm{d}u,\\
\mathbb{E}(X_t) &= \mathrm{e}^{-\gamma t}\mathbb{E}(X_0) + \alpha\int_0^t \mathrm{e}^{-\gamma(t-u)}\,\mathrm{d}u - \beta\int_0^t \mathrm{e}^{-\gamma(t-u)}\mathbb{E}(Y_u)\,\mathrm{d}u,\\
\mathbb{E}(Y_t^2) &= \mathrm{e}^{-2bt}\mathbb{E}(Y_0^2) + (2a+\sigma_1^2)\int_0^t \mathrm{e}^{-2b(t-u)}\mathbb{E}(Y_u)\,\mathrm{d}u,\\
\mathbb{E}(Y_t X_t) &= \mathrm{e}^{-(b+\gamma)t}\mathbb{E}(Y_0 X_0) + a\int_0^t \mathrm{e}^{-(b+\gamma)(t-u)}\mathbb{E}(X_u)\,\mathrm{d}u\\
&\quad + (\alpha + \varrho\sigma_1\sigma_2)\int_0^t \mathrm{e}^{-(b+\gamma)(t-u)}\mathbb{E}(Y_u)\,\mathrm{d}u - \beta\int_0^t \mathrm{e}^{-(b+\gamma)(t-u)}\mathbb{E}(Y_u^2)\,\mathrm{d}u,\\
\mathbb{E}(X_t^2) &= \mathrm{e}^{-2\gamma t}\mathbb{E}(X_0^2) + \alpha\int_0^t \mathrm{e}^{-2\gamma(t-u)}\mathbb{E}(X_u)\,\mathrm{d}u - 2\beta\int_0^t \mathrm{e}^{-2\gamma(t-u)}\mathbb{E}(Y_u X_u)\,\mathrm{d}u\\
&\quad + \sigma_2^2\int_0^t \mathrm{e}^{-2\gamma(t-u)}\mathbb{E}(Y_u)\,\mathrm{d}u + \sigma_3^2\int_0^t \mathrm{e}^{-2\gamma(t-u)}\,\mathrm{d}u.
\end{aligned}$$

If $\sigma_1 > 0$ *and* $Y_0 = y_0$*, then the Laplace transform of* Y_t*,* $t \in \mathbb{R}_{++}$*, takes the form*

$$\mathbb{E}(\mathrm{e}^{-\lambda Y_t}) = \left(1 + \frac{\sigma_1^2}{2}\lambda\int_0^t \mathrm{e}^{-bu}\,\mathrm{d}u\right)^{-\frac{2a}{\sigma_1^2}} \exp\left\{-\frac{\lambda \mathrm{e}^{-bt} y_0}{1 + \frac{\sigma_1^2}{2}\lambda\int_0^t \mathrm{e}^{-bu}\,\mathrm{d}u}\right\}, \qquad \lambda \in \mathbb{R}_+, \tag{B.1}$$

i.e., Y_t *has a non-centered chi-square distribution up to a multiplicative constant* $\frac{\sigma_1^2}{4}\int_0^t \mathrm{e}^{-bu}\,\mathrm{d}u$*, with degrees of freedom* $\frac{4a}{\sigma_1^2}$ *and with non-centrality parameter* $\frac{4\mathrm{e}^{-bt} y_0}{\sigma_1^2 \int_0^t \mathrm{e}^{-bu}\,\mathrm{d}u}$*.*

If $\sigma_1 > 0$ *and* $(1-\varrho^2)\sigma_2^2 + \sigma_3^2 > 0$*, then for each* $t \in \mathbb{R}_{++}$*, the distribution of* (Y_t, X_t) *is absolutely continuous.*

The next theorem gives a recursive formula for the moments of the stationary distribution of the process $(Y_t, X_t)_{t\in\mathbb{R}_+}$ in the subcritical case, see Bolyog and Pap (2016, Theorem 5.1).

Theorem B.2 *Let us consider the two-factor affine diffusion model (1.1) with* $a \in \mathbb{R}_+$*,* $b \in \mathbb{R}_{++}$*,* $\alpha, \beta \in \mathbb{R}$*,* $\gamma \in \mathbb{R}_{++}$*,* $\sigma_1, \sigma_2, \sigma_3 \in \mathbb{R}_+$*,* $\varrho \in [-1, 1]$*, and the random vector* (Y_∞, X_∞) *given by Theorem A.1. Then all the (mixed) moments of* (Y_∞, X_∞) *of any order are finite, i.e., we have* $\mathbb{E}(Y_\infty^n |X_\infty|^p) < \infty$ *for all* $n, p \in \mathbb{Z}_+$*, and the recursion*

$$\begin{aligned}
\mathbb{E}(Y_\infty^n X_\infty^p) = &\frac{1}{nb + p\gamma}\Bigg[\left(na + \frac{1}{2}n(n-1)\sigma_1^2\right)\mathbb{E}\left(Y_\infty^{n-1} X_\infty^p\right) - p\beta\mathbb{E}\left(Y_\infty^{n+1} X_\infty^{p-1}\right)\\
&+ p(\alpha + n\varrho\sigma_1\sigma_2)\mathbb{E}(Y_\infty^n X_\infty^{p-1})\\
&+ \frac{1}{2}p(p-1)\sigma_2^2\mathbb{E}(Y_\infty^{n+1} X_\infty^{p-2}) + \frac{1}{2}p(p-1)\sigma_3^2\mathbb{E}(Y_\infty^n X_\infty^{p-2})\Bigg],
\end{aligned}$$

holds for all $n, p \in \mathbb{Z}_+$ *with* $n + p \geqslant 1$*, where* $\mathbb{E}(Y_\infty^k X_\infty^\ell) := 0$ *for* $k, \ell \in \mathbb{Z}$ *with* $k < 0$ *or* $\ell < 0$*. Especially,*

$$\mathbb{E}(Y_\infty) = \frac{a}{b}, \qquad \mathbb{E}(Y_\infty^2) = \frac{a(2a+\sigma_1^2)}{2b^2}, \qquad \mathbb{E}(Y_\infty^3) = \frac{a(a+\sigma_1^2)(2a+\sigma_1^2)}{2b^3},$$

$$\mathbb{E}(X_\infty) = \frac{b\alpha - a\beta}{b\gamma}, \qquad \mathbb{E}(Y_\infty X_\infty) = \frac{a\mathbb{E}(X_\infty) - \beta\mathbb{E}(Y_\infty^2) + (\alpha + \varrho\sigma_1\sigma_2)\mathbb{E}(Y_\infty)}{b+\gamma},$$

$$\mathbb{E}(X_\infty^2) = \frac{-2\beta\mathbb{E}(Y_\infty X_\infty) + 2\alpha\mathbb{E}(X_\infty) + \sigma_2^2\mathbb{E}(Y_\infty) + \sigma_3^2}{2\gamma},$$

$$\mathbb{E}(Y_\infty^2 X_\infty) = \frac{(2a+\sigma_1^2)\mathbb{E}(Y_\infty X_\infty) - \beta\mathbb{E}(Y_\infty^3) + (\alpha + 2\varrho\sigma_1\sigma_2)\mathbb{E}(Y_\infty^2)}{2b+\gamma},$$

$$\mathbb{E}(Y_\infty X_\infty^2) = \frac{a\mathbb{E}(X_\infty^2) - 2\beta\mathbb{E}(Y_\infty^2 X_\infty) + 2(\alpha + \varrho\sigma_1\sigma_2)\mathbb{E}(Y_\infty X_\infty) + \sigma_2^2\mathbb{E}(Y_\infty^2) + \sigma_3^2\mathbb{E}(Y_\infty)}{b+2\gamma}.$$

If $\sigma_1 > 0$, then the Laplace transform of Y_∞ takes the form

$$\mathbb{E}(\mathrm{e}^{-\lambda Y_\infty}) = \left(1 + \frac{\sigma_1^2}{2b}\lambda\right)^{-2a/\sigma_1^2}, \qquad \lambda \in \mathbb{R}_+, \tag{B.2}$$

i.e., Y_∞ has gamma distribution with parameters $2a/\sigma_1^2$ and $2b/\sigma_1^2$, hence

$$\mathbb{E}(Y_\infty^\kappa) = \frac{\Gamma\left(\frac{2a}{\sigma_1^2} + \kappa\right)}{\left(\frac{2b}{\sigma_1^2}\right)^\kappa \Gamma\left(\frac{2a}{\sigma_1^2}\right)}, \qquad \kappa \in \left(-\frac{2a}{\sigma_1^2}, \infty\right).$$

If $\sigma_1 > 0$ and $(1-\varrho^2)\sigma_2^2 + \sigma_3^2 > 0$, then the distribution of (Y_∞, X_∞) is absolutely continuous.

C Limit theorems for continuous local martingales

In what follows we recall some limit theorems for continuous local martingales. We use these limit theorems for studying the asymptotic behaviour of the MLE of $\boldsymbol{\theta} = (a, b, \alpha, \beta, \gamma)^\top$. First we recall a strong law of large numbers for continuous local martingales.

Theorem C.1 (Liptser and Shiryaev 2001, Lemma 17.4) *Let $\big(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbb{R}_+}, \mathbb{P}\big)$ be a filtered probability space satisfying the usual conditions. Let $(M_t)_{t\in\mathbb{R}_+}$ be a square-integrable continuous local martingale with respect to the filtration $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ such that $\mathbb{P}(M_0 = 0) = 1$. Let $(\xi_t)_{t\in\mathbb{R}_+}$ be a progressively measurable process such that*

$$\mathbb{P}\left(\int_0^t \xi_u^2 \,\mathrm{d}\langle M\rangle_u < \infty\right) = 1, \quad t \in \mathbb{R}_+, \qquad \text{and} \qquad \int_0^t \xi_u^2 \,\mathrm{d}\langle M\rangle_u \xrightarrow{\text{a.s.}} \infty \quad \text{as } t \to \infty,$$

where $(\langle M\rangle_t)_{t\in\mathbb{R}_+}$ denotes the quadratic variation process of M. Then

$$\frac{\int_0^t \xi_u \,\mathrm{d}M_u}{\int_0^t \xi_u^2 \,\mathrm{d}\langle M\rangle_u} \xrightarrow{\text{a.s.}} 0 \qquad \text{as } t \to \infty. \tag{C.1}$$

If $(M_t)_{t\in\mathbb{R}_+}$ is a standard Wiener process, the progressive measurability of $(\xi_t)_{t\in\mathbb{R}_+}$ can be relaxed to measurability and adaptedness to the filtration $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$.

The next theorem is about the asymptotic behaviour of continuous multivariate local martingales.

Theorem C.2 (van Zanten 2000, Theorem 4.1) *Let $\big(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbb{R}_+}, \mathbb{P}\big)$ be a filtered probability space satisfying the usual conditions. Let $(\boldsymbol{M}_t)_{t\in\mathbb{R}_+}$ be a d-dimensional square-integrable continuous local martingale with respect to the filtration $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ such that*

$\mathbb{P}(\boldsymbol{M}_0 = \boldsymbol{0}) = 1$. *Suppose that there exists a function* $\boldsymbol{Q} : [t_0, \infty) \to \mathbb{R}^{d\times d}$ *with some* $t_0 \in \mathbb{R}_+$ *such that* $\boldsymbol{Q}(t)$ *is an invertible (non-random) matrix for all* $t \in [t_0, \infty)$, $\lim_{t\to\infty} \|\boldsymbol{Q}(t)\| = 0$ *and*

$$\boldsymbol{Q}(t)\langle \boldsymbol{M}\rangle_t\, \boldsymbol{Q}(t)^\top \xrightarrow{\mathbb{P}} \boldsymbol{\eta}\boldsymbol{\eta}^\top \qquad \text{as } t \to \infty,$$

where $\boldsymbol{\eta}$ *is a* $d \times d$ *random matrix. Then, for each* $\mathbb{R}^{k\times\ell}$*-valued random matrix* $\boldsymbol{A}$ *defined on* $(\Omega, \mathcal{F}, \mathbb{P})$, *we have*

$$(\boldsymbol{Q}(t)\boldsymbol{M}_t, \boldsymbol{A}) \xrightarrow{\mathcal{D}} (\boldsymbol{\eta}\boldsymbol{Z}, \boldsymbol{A}) \qquad \text{as } t \to \infty,$$

where $\boldsymbol{Z}$ *is a* d*-dimensional standard normally distributed random vector independent of* $(\boldsymbol{\eta}, \boldsymbol{A})$.

References

Alfonsi A (2015) Affine diffusions and related processes: simulation, theory and applications. Springer, Cham, Bocconi University Press, Milan

Baldeaux J, Platen E (2013) Functionals of multidimensional diffusions with applications to finance. Springer, Cham, Bocconi University Press, Milan

Barczy M, Pap G (2016) Asymptotic properties of maximum-likelihood estimators for Heston models based on continuous time observations. Statistics 50(2):389–417. https://doi.org/10.1080/02331888.2015.1044991

Barczy M, Döring L, Li Z, Pap G (2013) On parameter estimation for critical affine processes. Electron J Stat 7:647–696. https://doi.org/10.1214/13-EJS786

Barczy M, Döring L, Li Z, Pap G (2014) Parameter estimation for a subcritical affine two factor model. J Stat Plann Inference 151–152:37–59. https://doi.org/10.1016/j.jspi.2014.04.001

Barczy M, Nyul B, Pap G (2016) Least squares estimation for the subcritical Heston model based on continuous time observations. Accessed 18 Nov 2015

Barczy M, Ben Alaya M, Kebaier A, Pap G (2018a) Asymptotic behavior of maximum likelihood estimators for a jump-type Heston model. To appear in J Stat Plann Inference. Accessed 15 Dec 2016

Barczy M, Ben Alaya M, Kebaier A, Pap G (2018b) Asymptotic properties of maximum likelihood estimator for the growth rate for a jump-type CIR process based on continuous time observations. To appear in Stochastic Process Appl. Accessed 1 Apr 2017 https://doi.org/10.1016/j.spa.2017.07.004

Ben Alaya M, Kebaier A (2012) Parameter estimation for the square root diffusions: ergodic and nonergodic cases. Stoch Models 28(4):609–634. https://doi.org/10.1080/15326349.2012.726042

Bolyog B, Pap G (2016) Conditions for stationarity and ergodicity of two-factor affine diffusions. Commun Stoch Anal 10(4):587–610

Bolyog B, Pap G (2017) On conditional least squares estimation for affine diffusions based on continuous time observations. Accessed 8 Mar 2017

Cox JC, Ingersoll JE, Ross SA (1985) A theory of the term structure of interest rates. Econometrica 53(2):385–407. https://doi.org/10.2307/1911242

Dawson DA, Li Z (2006) Skew convolution semigroups and affine Markov processes. Ann Probab 34(3):1103–1142. https://doi.org/10.1214/009117905000000747

Dietz HM, Kutoyants YuA (1997) A class of minimum-distance estimators for diffusion processes with ergodic properties. Stat Decisions 15:211–227

Duffie D, Filipović D, Schachermayer W (2003) Affine processes and applications in finance. Ann Appl Probab 13:984–1053. https://doi.org/10.1214/aoap/1060202833

Filipović D (2009) Term-structure models. Springer, Berlin

Hu Y, Long H (2007) Parameter estimation for Ornstein-Uhlenbeck processes driven by α-stable Lévy motions. Commun Stoch Anal 1(2):175–192

Hu Y, Long H (2009a) Least squares estimator for Ornstein-Uhlenbeck processes driven by α-stable motions. Stoch Process Appl 119(8):2465–2480. https://doi.org/10.1016/j.spa.2008.12.006

Hu Y, Long H (2009b) On the singularity of least squares estimator for mean-reverting α-stable motions. Acta Math Sci Ser B Engl Ed 29B(3):599–608. https://doi.org/10.1016/S0252-9602(09)60056-4

Jacod J, Shiryaev AN (2003) Limit theorems for stochastic processes, 2nd edn. Springer, Berlin

Karatzas I, Shreve SE (1991) Brownian motion and stochastic calculus, 2nd edn. Springer, New York

Küchler U, Sørensen M (1997) Exponential families of stochastic processes. Springer, New York
Liptser RS, Shiryaev AN (2001) Statistics of random processes II. Applications, 2nd edn. Springer, Berlin
Overbeck L, Rydén T (1997) Estimation in the Cox–Ingersoll–Ross model. Econom Theory 13(3):430–461. https://doi.org/10.1017/S0266466600005880
van der Vaart AW (1998) Asymptotic statistics. Cambridge University Press, Cambridge
van Zanten H (2000) A multivariate central limit theorem for continuous local martingales. Statist Probab Lett 50(3):229–235. https://doi.org/10.1016/S0167-7152(00)00108-5