Asymptotic behavior of maximum likelihood estimators for a jump-type Heston model

Mátyás Barczy*,◇, Mohamed Ben Alaya**,

Ahmed Kebaier*** and Gyula Pap****

* MTA-SZTE Analysis and Stochastics Research Group, Bolyai Institute, University of Szeged, Aradi vértanúk tere 1, H–6720 Szeged, Hungary.

** Laboratoire De Mathématiques Raphaël Salem, UMR 6085, Université De Rouen, Avenue de L'Université Technopôle du Madrillet, 76801 Saint-Etienne-Du-Rouvray, France.

*** Université Paris 13, Sorbonne Paris Cité, LAGA, CNRS (UMR 7539), Villetaneuse, France.

**** Bolyai Institute, University of Szeged, Aradi vértanúk tere 1, H–6720 Szeged, Hungary.

e–mails: barczy@math.u-szeged.hu (M. Barczy),
mohamed.ben-alaya@univ-rouen.fr (M. Ben Alaya),
kebaier@math.univ-paris13.fr (A. Kebaier),
papgy@math.u-szeged.hu (G. Pap).

◇ Corresponding author

Abstract

We study asymptotic properties of maximum likelihood estimators of drift parameters for a jump-type Heston model based on continuous time observations, where the jump process can be any purely non-Gaussian Lévy process of not necessarily bounded variation with a Lévy measure concentrated on $(-1, \infty)$. We prove strong consistency and asymptotic normality for all admissible parameter values except one, where we show only weak consistency and mixed normal (but non-normal) asymptotic behavior. It turns out that the volatility of the price process is a measurable function of the price process. We also present some numerical illustrations to confirm our results.

1 Introduction

Parameter estimation, especially studying asymptotic properties of maximum likelihood estimator (MLE) of drift parameters for Cox–Ingersoll–Ross (CIR) and Heston models is an active

---

area of research mainly due to the wide range of applications of these models in financial mathematics.

The present paper gives a new contribution to the theory of asymptotic properties of MLE for jump-type Heston models based on continuous time observations. Concerning related works, due to the vast literature on parameter estimation for Heston models, we will restrict ourselves to mention only papers that investigate the very same types of questions. For a detailed and recent survey on parameter estimation for Heston models in general, see the Introduction of Barczy and Pap [6].

Overbeck [33] studied MLE of the drift parameters of the first coordinate process of a (diffusion type) Heston model (see (1.1)) based on continuous time observations, which is nothing else but a CIR process, also called square root process or Feller process. Ben-Alaya and Kebaier [8], [9] made a progress in MLE for the CIR process, giving explicit forms of joint Laplace transforms of the building blocks of this MLE as well.

Barczy and Pap [6] considered a Heston model

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

where $a, \sigma_1, \sigma_2 \in (0, \infty)$, $b, \alpha, \beta \in \mathbb{R}$, $\varrho \in (-1, 1)$ and $(W_t, B_t)_{t \in [0,\infty)}$ is a 2-dimensional standard Wiener process. Here $(X_t)_{t \in [0,\infty)}$ is the log-price process of an asset, $(Y_t)_{t \in [0,\infty)}$ is its stochastic volatility (or instantaneous variance), $\sigma_1 \in (0, \infty)$ is the so-called volatility of the volatility, and $\varrho \in (-1, 1)$ is the correlation between the driving standard Wiener processes $(W_t)_{t \in [0,\infty)}$ and $(\varrho W_t + \sqrt{1 - \varrho^2}B_t)_{t \in [0,\infty)}$. The MLE of the drift parameters (a, b, α, β) and its asymptotic behavior have been investigated based on continuous time observations $(Y_t, X_t)_{t \in [0,T]}$ with $T \in (0, \infty)$ for all admissible parameter values (according to $b > 0$, $b = 0$, and $b < 0$). It turned out that, for all $t \in [0, T]$, Y_t is a measurable function of $(X_t)_{t \in [0,T]}$, hence, for the calculation of the MLE in question, one does not need the sample $(Y_t)_{t \in [0,T]}$.

The original Heston model (see Heston [18]) takes the form

$$
(1.2) \qquad \begin{cases} \mathrm{d}Y_t = \kappa(\theta - Y_t)\,\mathrm{d}t + \sigma\sqrt{Y_t}\,\mathrm{d}W_t, \\ \mathrm{d}S_t = \mu S_t\,\mathrm{d}t + S_t\sqrt{Y_t}\big(\varrho\,\mathrm{d}W_t + \sqrt{1 - \varrho^2}\,\mathrm{d}B_t\big), \end{cases} \qquad t \in [0, \infty),
$$

where $(S_t)_{t \in [0,\infty)}$ is the price process of an asset, $\mu \in \mathbb{R}$ is the rate of return of the asset, $\theta \in (0, \infty)$ is the so-called long variance (long run average price variance, i.e., the limit of $\mathbb{E}(Y_t)$ as $t \to \infty$), $\kappa \in (0, \infty)$ is the rate at which $(Y_t)_{t \in [0,\infty)}$ reverts to θ, and $\sigma \in (0, \infty)$ is the so-called volatility of the volatility. We call the attention that there are two differences between the models (1.1) and (1.2). Namely, in (1.2) the coefficient κ can only be positive, while in (1.1) the corresponding coefficient b can be an arbitrary real number. In other words, the first coordinate process in (1.1) can be subcritical, critical or supercritical (according to $b > 0$, $b = 0$, and $b < 0$), but in (1.2) it can only be subcritical (since $\kappa > 0$). Moreover, the second coordinate process in (1.2) is the price process, while in (1.1) it is the log-price process.

In this paper we study a jump-type Heston model (also called a stochastic volatility with jumps model, SVJ model)

$$
(1.3) \qquad \begin{cases} \mathrm{d}Y_t = \kappa(\theta - Y_t)\,\mathrm{d}t + \sigma\sqrt{Y_t}\,\mathrm{d}W_t, \\ \mathrm{d}S_t = \mu S_t\,\mathrm{d}t + S_t\sqrt{Y_t}\big(\varrho\,\mathrm{d}W_t + \sqrt{1-\varrho^2}\,\mathrm{d}B_t\big) + S_{t-}\,\mathrm{d}L_t, \end{cases} \qquad t \in [0,\infty),
$$

where $(L_t)_{t\in[0,\infty)}$ is a purely non-Gaussian Lévy process independent of $(W_t, B_t)_{t\in[0,\infty)}$ with Lévy–Khintchine representation

$$
(1.4) \qquad \mathbb{E}(\mathrm{e}^{\mathrm{i}uL_1}) = \exp\left\{ \mathrm{i}\gamma u + \int_{-1}^{\infty} (\mathrm{e}^{\mathrm{i}uz} - 1 - \mathrm{i}uz\mathbb{1}_{(-1,1]}(z))\, m(\mathrm{d}z) \right\}, \qquad u \in \mathbb{R},
$$

where $\gamma \in \mathbb{R}$ and m is a Lévy measure concentrated on $(-1,\infty)$ with $m(\{0\}) = 0$. Here, let us recall that the Lévy process L has finite variation on each interval $[0,t]$, $t \in [0,\infty)$, if and only if $\int_{-1}^{1} |z|\, m(\mathrm{d}z) < \infty$, see, e.g., Sato [35, Theorem 21.9]. We point out that the assumption $\mathbb{P}(Y_0 \in [0,\infty), S_0 \in (0,\infty)) = 1$ and the assumption in question on the support of the Lévy measure m assure that $\mathbb{P}(S_t \in (0,\infty)$ for all $t \in [0,\infty)) = 1$ (see Proposition 2.1), so the process S can be used for modeling prices in a financial market. From the point of view of financial mathematics, a natural question may occur concerning the model (1.3). Namely, is the drift coefficient of the second SDE in (1.3) well-adjusted in the sense that the discounted price process forms a martingale under some suitable equivalent martingale measure? We renounce to consider this question, we just note that one may have to choose the parameter μ in an appropriate way to assure this property. In Lamberton and Lapeyre [27, Section 7] one can find a detailed discussion of the same type of question for a jump-type Black-Scholes model, where the jumps of the log-price process is modeled by a compound Poisson process. They derived a necessary and sufficient condition for the drift coefficient of the underlying SDE in terms of the discounting factor and the parameters of the compound Poisson process in question in order that the discounted price process is a martingale, see [27, page 146]. For a good survey on jump-type Heston models, pricing and hedging in these models, see Runggaldier [34]. In fact, the model (1.3) is quite popular in finance with the special choice of the Lévy process L as a compound Poisson process. Namely, let

$$
(1.5) \qquad L_t := \sum_{i=1}^{\pi_t} (\mathrm{e}^{J_i} - 1), \qquad t \in [0,\infty),
$$

where $(\pi_t)_{t\in[0,\infty)}$ is a Poisson process with intensity $\lambda \in (0,\infty)$, $(J_i)_{i\in\mathbb{N}}$ is a sequence of independent identically distributed random variables having no atom at zero (i.e., $\mathbb{P}(J_1 = 0) = 0$), and being independent of π as well. We also suppose that π, $(J_i)_{i\in\mathbb{N}}$, W and B are independent. One can interpret J as the jump size of the logarithm of the asset price. Then

$$
\mathbb{E}(\mathrm{e}^{\mathrm{i}uL_1}) = \exp\left\{ \lambda \int_{-1}^{\infty} (\mathrm{e}^{\mathrm{i}uz} - 1)\, m(\mathrm{d}z) \right\}
$$

$$
= \exp\left\{ \mathrm{i}u\lambda \int_{-1}^{1} z\, m(\mathrm{d}z) + \lambda \int_{-1}^{\infty} (\mathrm{e}^{\mathrm{i}uz} - 1 - \mathrm{i}uz\mathbb{1}_{[-1,1]}(z))\, m(\mathrm{d}z) \right\}, \qquad u \in \mathbb{R},
$$

3

has the form (1.4) with m being the distribution of $e^{J_1} - 1$ and $\gamma = \lambda \int_{-1}^{1} z\, m(\mathrm{d}z)$. Moreover, S_t takes the form

$$(1.6) \qquad S_t = S_0 \exp\left\{ \int_0^t \left(\mu - \frac{1}{2}Y_u\right) \mathrm{d}u + \int_0^t \sqrt{Y_u}\left(\varrho\, \mathrm{d}W_u + \sqrt{1-\varrho^2}\, \mathrm{d}B_u\right) + \sum_{i=1}^{\pi_t} J_i \right\}$$

for $t \in [0,\infty)$, see (2.1). We note that the SDE (1.3) with the Lévy process L given in (1.5) has been studied, e.g., by Bates [7, equation (1)], Bakshi et al. [3, equations (1) and (2) with $R \equiv 0$], by Broadie and Kaya [12, equations (30)-(31)] (where a factor S_{t-} is missing from the last term of equation (30)), by Runggaldier [34, Remark 3.1 with $\lambda_t \equiv \lambda$] and by Sun et al. [37, equation (1) with $J_v = 0$]. Bates [7], Bakshi et al. [3] and Broadie and Kaya [12] have chosen the common distribution of J as a normal distribution. Bakshi et al. [3] used this model for studying (European style) S&P 500 options, e.g., they derived a practically implementable closed-form pricing formula. Broadie and Kaya [12] gave an exact simulation algorithm for this model, further, they considered the pricing of forward start options in this model. Sun et al. [37] have chosen the common distribution of J as a normal distribution, a one-sided exponential distribution or a two-sided distribution, and they applied the Fourier-cosine series expansion method for pricing vanilla options under these jump-type Heston models.

The aim of this paper is to study the MLE of the parameter $\boldsymbol{\psi} := (\theta, \kappa, \mu)$ for the model (1.3) based on continuous time observations $(Y_t, S_t)_{t \in [0,T]}$ with $T \in (0,\infty)$, starting the process (Y, S) from some deterministic initial value $(y_0, s_0) \in (0,\infty)^2$ supposing that σ, ϱ, γ and the Lévy measure m are known. Here we stress that under these assumptions, the underlying statistical space corresponding to the parameters $(\kappa, \theta, \mu) \in (0,\infty)^2 \times \mathbb{R}$ is identifiable, however it would not be true for the statistical space corresponding to the parameters $(\kappa, \theta, \mu, \gamma) \in (0,\infty)^2 \times \mathbb{R}^2$. We call the attention that the MLE in question contains stochastic integrals with respect to L. We prove that, for all $t \in [0,T]$, L_t is a measurable function (i.e., a statistic) of $(S_t)_{t \in [0,T]}$, by providing a sequence of measurable functions of $(S_t)_{t \in [0,T]}$ converging in probability to L_t, see Remark 2.4 (note that this sequence depends on γ and m as well). Further, it turns out that Y_t for all $t \in [0,T]$, and the parameters σ and ϱ are also measurable functions of $(S_t)_{t \in [0,T]}$, see Remarks 2.5 and 2.6, respectively. Hence, for the calculation of the MLE in question, one needs only the sample $(S_t)_{t \in [0,T]}$, the parameter γ and the Lévy measure m (γ and m are needed for the reconstruction of $(L_t)_{t \in [0,T]}$). Though we do not need to estimate the parameters σ and ϱ, it is worth mentioning that the market microstructure effects may cause serious damage to the approximation of σ and ϱ given in Remark 2.6 and to the MLE of (θ, κ, μ) in case of high-frequency observations as in Zhang et al. [41]. This type of question can be another interesting topic for future research.

The paper is organized as follows. In Section 2, we prove that the SDE (1.3) has a pathwise unique strong solution (under some appropriate conditions), see Proposition 2.1, we recall a result about the existence of a unique stationary distribution and ergodicity for the process $(Y_t)_{t \in [0,\infty)}$ given by the first equation in (1.3), see Theorem 2.2. In Proposition 2.3, we derive a Grigelionis representation for the process $(S_t)_{t \in [0,\infty)}$. Further, we prove that for all $t \in [0,T]$, L_t and Y_t are measurable functions of $(S_t)_{t \in [0,T]}$, and we justify why we do not estimate the

parameters σ and ϱ, see Remarks 2.4, 2.5 and 2.6. Section 3 is devoted to study the existence and uniqueness of the MLE $(\widehat{\theta}_T, \widehat{\kappa}_T, \widehat{\mu}_T)$ of (θ, κ, μ) based on observations $(Y_t, S_t)_{t \in [0,T]}$ with $T \in (0, \infty)$. In Proposition 3.2, under appropriate conditions, we prove the unique existence of $(\widehat{\theta}_T, \widehat{\kappa}_T, \widehat{\mu}_T)$, and we derive an explicit formula for it as well, see (3.11). In Remark 3.5, we describe the connection with the so called score vector due to Sørensen [36] and the estimating equation due to Luschgy [31], [32] leading to the same estimator. In Section 4, we prove that the MLE of (θ, κ, μ) is strongly consistent if $\theta, \kappa \in (0, \infty)$ with $\theta\kappa \in \left(\frac{\sigma^2}{2}, \infty\right)$, and weakly consistent if $\theta, \kappa \in (0, \infty)$ with $\theta\kappa = \frac{\sigma^2}{2}$, see Theorem 4.1 and Remark 4.2, respectively. Section 5 is devoted to investigate the asymptotic behaviour of the MLE of (θ, κ, μ). In Theorem 5.1, provided that $\theta, \kappa \in (0, \infty)$ with $\theta\kappa \in \left(\frac{\sigma^2}{2}, \infty\right)$, we show that the MLE of (θ, κ, μ) is asymptotically normal with a usual square root normalization $(T^{1/2})$, but as usual, the asymptotic covariance matrix depends on the unknown parameters θ and κ, as well. To get around this problem, we also replace the normalization $T^{1/2}$ by a random one (depending only on the sample, but not on the parameters θ, κ and μ) with the advantage that the MLE of (θ, κ, μ) with the random scaling is asymptotically 3-dimensional standard normal. Theorem 5.3 is a counterpart of Theorem 5.1 in some sense. Namely, provided that $\theta, \kappa \in (0, \infty)$ with $\theta\kappa = \frac{\sigma^2}{2}$, we derive two limit theorems for the MLE $(\widehat{\theta}_T, \widehat{\kappa}_T, \widehat{\mu}_T)$ with mixed normal limit distributions. First, we have a non-random scaling, but for $\widehat{\mu}_T$ instead of the usual scaling $T^{1/2}$ we have T; and then we have a random scaling as well. We point out that, surprisingly, the limit distributions in Theorems 5.1 and 5.3 do not depend on L (roughly speaking, they do not depend on the jump part). From a practical point of view, a natural question can occur, namely, how one can decide whether Theorems 5.1 and 5.3 can be applied (if yes, then which one), since one does not know the product $\theta\kappa$ of the unknown parameters θ and κ in advance. To answer this question, one can build up a probe for testing the null hypothesis $\theta\kappa = \frac{\sigma^2}{2}$ against some alternative hypothesis, e.g., $\theta\kappa > \frac{\sigma^2}{2}$. In Section 6 we present some numerical illustrations of our limit theorems. We close the paper with Appendices, where we recall certain sufficient conditions for the absolute continuity of probability measures induced by semimartingales together with a representation of the Radon–Nikodym derivative (Appendix A), some limit theorems for continuous local martingales for studying asymptotic behavior of $(\widehat{\theta}_T, \widehat{\kappa}_T, \widehat{\mu}_T)$ (Appendix B) and a version of the continuous mapping theorem (Appendix C), and we give an explicit formula for the non-normal but mixed normal density function of the limit distribution of $T(\widehat{\mu}_T - \mu)$ as $T \to \infty$ in Theorem 5.3 (Appendix D).

We call the attention that in both cases $\theta\kappa > \frac{\sigma^2}{2}$ and $\theta\kappa = \frac{\sigma^2}{2}$, the CIR process Y has a unique stationary distribution and is ergodic, nevertheless, in case $\theta\kappa > \frac{\sigma^2}{2}$ the asymptotic limit distribution of the MLE of $\boldsymbol{\psi} = (\theta, \kappa, \mu)$ is normal, while in case $\theta\kappa = \frac{\sigma^2}{2}$ it is mixed normal. The interesting point is that we have an ergodic case with an asymptotically mixed normal (but non-normal) limit distribution. The main difference between the two ergodic cases is that $\mathbb{E}\left(\frac{1}{Y_\infty}\right) < \infty$ if $\theta\kappa > \frac{\sigma^2}{2}$, but $\mathbb{E}\left(\frac{1}{Y_\infty}\right) = \infty$ if $\theta\kappa = \frac{\sigma^2}{2}$.

2 Preliminaries

The next proposition is about the existence and uniqueness of a strong solution of the SDE (1.3).

Proposition 2.1 *Let (η_0, ζ_0) be a random vector such that η_0 is independent of $(W_t)_{t \in [0,\infty)}$ satisfying $\mathbb{P}(\eta_0 \in [0,\infty), \zeta_0 \in (0,\infty)) = 1$. Then for all $\theta, \kappa, \sigma \in (0,\infty)$, $\mu \in \mathbb{R}$ and $\varrho \in (-1,1)$, there is a (pathwise) unique strong solution $(Y_t, S_t)_{t \in [0,\infty)}$ of the SDE (1.3) such that $\mathbb{P}((Y_0, S_0) = (\eta_0, \zeta_0)) = 1$ and $\mathbb{P}(Y_t \in [0,\infty)$ and $S_t \in (0,\infty)$ for all $t \in [0,\infty)) = 1$. Further,*

(2.1)
$$S_t = S_0 \exp\left\{ \int_0^t \left(\mu - \frac{1}{2} Y_u \right) du + \int_0^t \sqrt{Y_u} \left(\varrho \, dW_u + \sqrt{1-\varrho^2} \, dB_u \right) + L_t \right\} \prod_{u \in [0,t]} (1 + \Delta L_u) e^{-\Delta L_u}$$

for $t \in [0,\infty)$, where $\Delta L_u := L_u - L_{u-}$, $u \in (0,\infty)$, $\Delta L_0 := 0$, and the (possibly) infinite product is absolutely convergent. If, in addition, $\theta\kappa \in [\frac{\sigma^2}{2}, \infty)$ and $\mathbb{P}(\eta_0 \in (0,\infty)) = 1$, then $\mathbb{P}(Y_t \in (0,\infty)$ for all $t \in [0,\infty)) = 1$.

Note that, due to Sato [35, Theorem 21.3], for each $t \in (0,\infty)$, the product $\prod_{u \in [0,t]} (1 + \Delta L_u) e^{-\Delta L_u}$ in (2.1) contains finitely many terms different from 1 if and only if $m((-1,1]) < \infty$.

Proof of Proposition 2.1. By a theorem due to Yamada and Watanabe (see, e.g., Karatzas and Shreve [25, Proposition 5.2.13]), the strong uniqueness holds for the first equation in (1.3). By Ikeda and Watanabe [19, Example 8.2, page 221], there is a (pathwise) unique non-negative strong solution $(Y_t)_{t \in [0,\infty)}$ of the first equation in (1.3) with any initial value η_0 such that $\mathbb{P}(\eta_0 \in [0,\infty)) = 1$. The second equation in (1.3) can be written in the form

$$dS_t = S_{t-} \, dL_t^*, \qquad t \in [0,\infty),$$

where

(2.2)
$$L_t^* := \mu t + \int_0^t \sqrt{Y_u} \left(\varrho \, dW_u + \sqrt{1-\varrho^2} \, dB_u \right) + L_t, \qquad t \in [0,\infty),$$

is a semimartingale, since the process $(\sqrt{Y_t})_{t \in [0,\infty)}$ has continuous sample paths almost surely and hence locally bounded almost surely yielding that $\left(\int_0^t \sqrt{Y_u} (\varrho \, dW_u + \sqrt{1-\varrho^2} \, dB_u) \right)_{t \in [0,\infty)}$ is a square integrable martingale, and since L is a semimartingale being a Lévy process (see, e.g., Jacod and Shiryaev [22, Corollary II.4.19]). Using $\Delta L_t^* = \Delta L_t$, $t \in [0,\infty)$, and Theorem 1 in Jaschke [23], which is a generalization of the Doléans–Dade exponential formula (see, e.g., Jacod and Shiryaev [22, I.4.61]), we obtain

$$S_t = S_0 \exp\left\{ L_t^* - L_0^* - \frac{1}{2} \langle (L^*)^{\mathrm{cont}} \rangle_t \right\} \prod_{u \in [0,t]} (1 + \Delta L_u^*) e^{-\Delta L_u^*}$$

$$= S_0 \exp\left\{ \mu t + \int_0^t \sqrt{Y_u} \left(\varrho \, dW_u + \sqrt{1-\varrho^2} \, dB_u \right) + L_t - \frac{1}{2} \int_0^t Y_u \, du \right\} \prod_{u \in [0,t]} (1 + \Delta L_u) e^{-\Delta L_u},$$

6

where $(\langle(L^*)^{\mathrm{cont}}\rangle_t)_{t\in[0,\infty)}$ denotes the (predictable) quadratic variation process of the continuous martingale part $(L^*)^{\mathrm{cont}}$ of L^*, and the (possibly) infinite product is absolutely convergent. Here we used that $\langle(L^*)^{\mathrm{cont}}\rangle_t = \int_0^t Y_u\,\mathrm{d}u$, $t\in[0,\infty)$, being a consequence of the fact that

$$(2.3) \qquad (L^*)_t^{\mathrm{cont}} = \int_0^t \sqrt{Y_u}\bigl(\varrho\,\mathrm{d}W_u + \sqrt{1-\varrho^2}\,\mathrm{d}B_u\bigr), \qquad t\in[0,\infty),$$

which can be checked as follows. The Lévy–Itô's representation of L takes the form

$$
\begin{aligned}
(2.4)\qquad L_t &= \lim_{\delta\downarrow 0} \int_{(0,t]}\int_{\{\delta<|z|\leqslant 1\}} z\bigl(\mu^L(\mathrm{d}u,\mathrm{d}z) - \mathrm{d}u\,m(\mathrm{d}z)\bigr) + \int_{(0,t]}\int_{\{1<|z|<\infty\}} z\,\mu^L(\mathrm{d}u,\mathrm{d}z) + \gamma t \\
&=: \int_0^t\int_{\mathbb{R}} h_1(z)\bigl(\mu^L(\mathrm{d}u,\mathrm{d}z) - \mathrm{d}u\,m(\mathrm{d}z)\bigr) + \int_0^t\int_{\mathbb{R}}(z - h_1(z))\,\mu^L(\mathrm{d}u,\mathrm{d}z) + \gamma t
\end{aligned}
$$

for $t\in[0,\infty)$, where $\mu^L(\mathrm{d}u,\mathrm{d}z) := \sum_{v\in[0,\infty)}\mathbb{1}_{\{\Delta L_v\neq 0\}}\varepsilon_{(v,\Delta L_v)}(\mathrm{d}u,\mathrm{d}z)$ is the integer-valued Poisson random measure on $[0,\infty)\times\mathbb{R}$ associated with the jumps of the process L, $\varepsilon_{(v,x)}$ denotes the Dirac measure at the point $(v,x)\in[0,\infty)\times\mathbb{R}$, and

$$(2.5) \qquad h_1(z) := z\mathbb{1}_{[-1,1]}(z), \qquad z\in\mathbb{R},$$

is a truncation function, see, e.g., Sato [35, Theorem 19.2]. The first term in (2.4) is a purely discontinuous local martingale, see, e.g., Jacod and Shiryaev [22, Definitions II.1.27]. The second term in (2.4) can be written as a finite sum (see, e.g., Sato [35, Lemma 20.1])

$$\int_0^t\int_{\mathbb{R}}(z - h_1(z))\,\mu^L(\mathrm{d}u,\mathrm{d}z) = \sum_{u\in[0,t]}\mathbb{1}_{\{|\Delta L_u|>1\}}\Delta L_u, \qquad t\in[0,\infty),$$

which is a compound Poisson process with Lévy-Khintchine representation

$$\mathbb{E}\left(\mathrm{e}^{\mathrm{i}\theta\sum_{u\in[0,1]}\mathbb{1}_{\{|\Delta L_u|>1\}}\Delta L_u}\right) = \exp\left\{\int_1^{\infty}(\mathrm{e}^{\mathrm{i}\theta u} - 1)\,m(\mathrm{d}u)\right\}, \qquad \theta\in\mathbb{R}.$$

Hence it is a process with finite variation over each finite interval $[0,t]$, $t\in[0,\infty)$, see, e.g., Sato [35, Theorem 21.9]. Consequently, we conclude (2.3). An alternative way for deriving (2.3) is as follows. Using (2.4), the process L^* can be written in the form III.2.23 in Jacod and Shiryaev [22], and hence, by Jacod and Shiryaev [22, Remarks III.2.28, part 1)], we get (2.3). Thus the (pathwise) unique strong solution $(S_t)_{t\in[0,\infty)}$ of the second equation in (1.3) is given by (2.1). Further,

$$\mathbb{P}(\Delta L_t \in (-1,\infty) \text{ for all } t\in[0,\infty)) = 1,$$

since the Lévy measure m of L is concentrated on $(-1,0)\cup(0,\infty)$. Using again $\Delta L_t^* = \Delta L_t$, $t\in[0,\infty)$, we obtain $\mathbb{P}(\Delta L_t^* \in (-1,\infty)$ for all $t\in[0,\infty)) = 1$, and hence $\mathbb{P}(S_t\in(0,\infty)$ for all $t\in[0,\infty)) = 1$. Indeed, if $S_0 = 1$, then this follows, e.g., from Theorem I.4.61 (c) in Jacod and Shiryaev [22], hence, in general, this is a consequence of formula (2.1) and $\mathbb{P}(S_0\in(0,\infty)) = 1$.

The proof of the last statement can be found, e.g., in Ikeda and Watanabe [19, Chapter IV, Example 8.2] and in Lamberton and Lapeyre [27, Proposition 6.2.4]. $\qquad\square$

In the sequel $\xrightarrow{\mathbb{P}}$, $\xrightarrow{\mathcal{D}}$ and $\xrightarrow{\text{a.s.}}$ will denote convergence in probability, in distribution and almost surely, respectively.

The following result states the existence of a unique stationary distribution and the ergodicity for the process $(Y_t)_{t\in[0,\infty)}$ given by the first equation in (1.3), see, e.g., Feller [17], Cox et al. [13, Equation (20)], Li and Ma [29, Theorem 2.6] or Theorem 3.1 with $\alpha = 2$ and Theorem 4.1 in Barczy et al. [5].

Theorem 2.2 Let $\theta, \kappa, \sigma \in (0, \infty)$. Let $(Y_t)_{t\in[0,\infty)}$ be the unique strong solution of the first equation of the SDE (1.3) satisfying $\mathbb{P}(Y_0 \in [0, \infty)) = 1$.

(i) Then $Y_t \xrightarrow{\mathcal{D}} Y_\infty$ as $t \to \infty$, and the distribution of Y_∞ is given by

$$(2.6) \qquad \mathbb{E}(e^{-\lambda Y_\infty}) = \left(1 + \frac{\sigma^2}{2\kappa}\lambda\right)^{-\frac{2\theta\kappa}{\sigma^2}}, \qquad \lambda \in [0, \infty),$$

i.e., Y_∞ has Gamma distribution with parameters $2\theta\kappa/\sigma^2$ and $2\kappa/\sigma^2$, hence

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

Especially, $\mathbb{E}(Y_\infty) = \theta$. Further, if $\theta\kappa \in \left(\frac{\sigma^2}{2}, \infty\right)$, then $\mathbb{E}\left(\frac{1}{Y_\infty}\right) = \frac{2\kappa}{2\theta\kappa - \sigma^2}$.

(ii) For all Borel measurable functions $f : \mathbb{R} \to \mathbb{R}$ such that $\mathbb{E}(|f(Y_\infty)|) < \infty$, we have

$$(2.7) \qquad \frac{1}{T}\int_0^T f(Y_u)\,\mathrm{d}u \xrightarrow{\text{a.s.}} \mathbb{E}(f(Y_\infty)) \qquad as \ T \to \infty.$$

Note that, by Proposition 2.1, the process $(Y_t, S_t)_{t\in[0,\infty)}$ is a semimartingale, see, e.g., Jacod and Shiryaev [22, I.4.34]. Now we derive a so-called Grigelionis form for the semimartingale $(S_t)_{t\in[0,\infty)}$, see, e.g., Jacod and Shiryaev [22, III.2.23] or Jacod and Protter [21, Theorem 2.1.2].

Proposition 2.3 Let $\theta, \kappa, \sigma \in (0, \infty)$, $\mu \in \mathbb{R}$, $\varrho \in (-1, 1)$. Let $(Y_t, S_t)_{t\in[0,\infty)}$ be the unique strong solution of the SDE (1.3) satisfying $\mathbb{P}(Y_0 \in [0, \infty), S_0 \in (0, \infty)) = 1$. Then

$$S_t = S_0 + (\mu + \gamma)\int_0^t S_u\,\mathrm{d}u + \int_0^t \left(\int_\mathbb{R}(h_1(S_{u-}z) - S_{u-}h_1(z))\,m(\mathrm{d}z)\right)\mathrm{d}u$$

$$(2.8) \qquad + \int_0^t S_u\sqrt{Y_u}\left(\varrho\,\mathrm{d}W_u + \sqrt{1-\varrho^2}\,\mathrm{d}B_u\right)$$

$$+ \int_0^t\int_\mathbb{R} h_1(S_{u-}z)\left(\mu^L(\mathrm{d}u, \mathrm{d}z) - \mathrm{d}u\,m(\mathrm{d}z)\right) + \int_0^t\int_\mathbb{R}(S_{u-}z - h_1(S_{u-}z))\,\mu^L(\mathrm{d}u, \mathrm{d}z)$$

for $t \in [0, \infty)$, where h_1 is defined in (2.5).

8

Proof. Using (2.4) and Proposition II.1.30 in Jacod and Shiryaev [22], we obtain

$$S_t = S_0 + (\mu + \gamma) \int_0^t S_u \, \mathrm{d}u + \int_0^t S_u \sqrt{Y_u} \left(\varrho \, \mathrm{d}W_u + \sqrt{1 - \varrho^2} \, \mathrm{d}B_u \right)$$

$$+ \int_0^t \int_{\mathbb{R}} S_{u-} h_1(z) \left(\mu^L(\mathrm{d}u, \mathrm{d}z) - \mathrm{d}u \, m(\mathrm{d}z) \right) + \int_0^t \int_{\mathbb{R}} S_{u-} (z - h_1(z)) \, \mu^L(\mathrm{d}u, \mathrm{d}z)$$

for $t \in [0, \infty)$. In order to prove the statement, it is enough to show

$$(2.9) \qquad \int_0^t \int_{\mathbb{R}} S_{u-} h_1(z) \left(\mu^L(\mathrm{d}u, \mathrm{d}z) - \mathrm{d}u \, m(\mathrm{d}z) \right) = I_1 - I_2,$$

$$(2.10) \qquad \int_0^t \int_{\mathbb{R}} S_{u-} (z - h_1(z)) \, \mu^L(\mathrm{d}u, \mathrm{d}z) = I_3 + I_4,$$

with

$$I_1 := \int_0^t \int_{\mathbb{R}} h_1(S_{u-} z) \left(\mu^L(\mathrm{d}u, \mathrm{d}z) - \mathrm{d}u \, m(\mathrm{d}z) \right),$$

$$I_2 := \int_0^t \int_{\mathbb{R}} (h_1(S_{u-} z) - S_{u-} h_1(z)) \left(\mu^L(\mathrm{d}u, \mathrm{d}z) - \mathrm{d}u \, m(\mathrm{d}z) \right),$$

$$I_3 := \int_0^t \int_{\mathbb{R}} (S_{u-} z - h_1(S_{u-} z)) \, \mu^L(\mathrm{d}u, \mathrm{d}z),$$

$$I_4 := \int_0^t \int_{\mathbb{R}} (h_1(S_{u-} z) - S_{u-} h_1(z)) \, \mu^L(\mathrm{d}u, \mathrm{d}z),$$

and the equality

$$(2.11) \qquad I_4 - I_2 = I_5 \qquad \text{with} \qquad I_5 := \int_0^t \left(\int_{\mathbb{R}} (h_1(S_{u-} z) - S_{u-} h_1(z)) \, m(\mathrm{d}z) \right) \mathrm{d}u.$$

For the equations (2.9), (2.10) and (2.11), it suffices to check the existence of I_2, I_3 and I_5.

First note that for every $s \in (0, \infty)$ we have

$$(2.12) \qquad h_1(sz) - s h_1(z) = \begin{cases} sz \mathbb{1}_{\{1 < |z| \leqslant \frac{1}{s}\}} & \text{if } s \in (0, 1), \ z \in \mathbb{R}, \\ 0 & \text{if } s = 1, \ z \in \mathbb{R}, \\ -sz \mathbb{1}_{\{\frac{1}{s} < |z| \leqslant 1\}} & \text{if } s \in (1, \infty), \ z \in \mathbb{R}. \end{cases}$$

The existence of I_2 will be a consequence of $I_2 = I_{2,1} - I_{2,2} - I_{2,3}$ with

$$I_{2,1} := \int_0^t \int_{\mathbb{R}} S_{u-} z \mathbb{1}_{\{1 < |z| \leqslant \frac{1}{S_{u-}}\}} \mathbb{1}_{\{S_{u-} \in (0,1)\}} \, \mu^L(\mathrm{d}u, \mathrm{d}z),$$

$$I_{2,2} := \int_0^t \int_{\mathbb{R}} S_{u-} z \mathbb{1}_{\{1 < |z| \leqslant \frac{1}{S_{u-}}\}} \mathbb{1}_{\{S_{u-} \in (0,1)\}} \, \mathrm{d}u \, m(\mathrm{d}z),$$

$$I_{2,3} := \int_0^t \int_{\mathbb{R}} S_{u-} z \mathbb{1}_{\{\frac{1}{S_{u-}} < |z| \leqslant 1\}} \mathbb{1}_{\{S_{u-} \in (1,\infty)\}} \left(\mu^L(\mathrm{d}u, \mathrm{d}z) - \mathrm{d}u \, m(\mathrm{d}z) \right).$$

9

Here we have

$$|I_{2,1}| \leqslant \int_0^t \int_{\mathbb{R}} |S_{u-} z| \mathbb{1}_{\{1 < |z| \leqslant \frac{1}{S_{u-}}\}} \mathbb{1}_{\{S_{u-} \in (0,1)\}} \, \mu^L(\mathrm{d}u, \mathrm{d}z) \leqslant \int_0^t \int_{\mathbb{R}} \mathbb{1}_{\{1 < |z|\}} \, \mu^L(\mathrm{d}u, \mathrm{d}z) < \infty,$$

see, e.g., Sato [35, Lemma 20.1]. Moreover,

$$|I_{2,2}| \leqslant \int_0^t \int_{\mathbb{R}} |S_{u-} z| \mathbb{1}_{\{1 < |z| \leqslant \frac{1}{S_{u-}}\}} \mathbb{1}_{\{S_{u-} \in (0,1)\}} \, \mathrm{d}u \, m(\mathrm{d}z)$$

$$\leqslant \int_0^t \int_{\mathbb{R}} \mathbb{1}_{\{1 < |z|\}} \, \mathrm{d}u \, m(\mathrm{d}z) = t m(\{z \in \mathbb{R} : |z| > 1\}) < \infty.$$

Further, the function $\Omega \times [0, \infty) \times \mathbb{R} \ni (\omega, t, z) \mapsto h_1(z)$ belongs to $G_{\mathrm{loc}}(\mu^L)$, see Jacod and Shiryaev [22, Definitions II.1.27, Theorem II.2.34]. We have $|z \mathbb{1}_{\{\frac{1}{S_{u-}} < |z| \leqslant 1\}} \mathbb{1}_{\{S_{u-} \in (1, \infty)\}}| \leqslant |h_1(z)|$, hence, by the definition of $G_{\mathrm{loc}}(\mu^L)$, the function $\Omega \times [0, \infty) \times \mathbb{R} \ni (\omega, t, z) \mapsto z \mathbb{1}_{\{\frac{1}{S_{u-}} < |z| \leqslant 1\}} \mathbb{1}_{\{S_{u-} \in (1, \infty)\}}$ also belongs to $G_{\mathrm{loc}}(\mu^L)$. By Jacod and Shiryaev [22, Proposition II.1.30], we conclude that the function $\Omega \times [0, \infty) \times \mathbb{R} \ni (\omega, t, z) \mapsto S_{u-} z \mathbb{1}_{\{\frac{1}{S_{u-}} < |z| \leqslant 1\}} \mathbb{1}_{\{S_{u-} \in (1, \infty)\}}$ also belongs to $G_{\mathrm{loc}}(\mu^L)$, thus the integral $I_{2,3}$ exists, and hence we obtain the existence of I_2, and hence of I_1.

Next observe that we have $\Delta S_t = S_{t-} \Delta L_t$, $t \in [0, \infty)$, see, e.g., Jacod and Shiryaev [22, page 60, formula (5)]. Consequently,

$$I_3 = \int_0^t \int_{\mathbb{R}} S_{u-} z \mathbb{1}_{\{|S_{u-} z| > 1\}} \, \mu^L(\mathrm{d}u, \mathrm{d}z) = \sum_{u \in [0,t]} S_{u-}(\Delta L_u) \mathbb{1}_{\{|S_{u-} \Delta L_u| > 1\}} = \sum_{u \in [0,t]} \Delta S_u \mathbb{1}_{\{|\Delta S_u| > 1\}}$$

is a finite sum, since the process $(S_t)_{t \in [0, \infty)}$ admits càdlàg trajectories, hence there can be at most finitely many points $u \in [0, t]$ at which the jump $|\Delta S_u|$ exceeds 1, see, e.g., Billingsley [11, page 122]. Thus we obtain the existence of I_3, and hence of I_4.

Finally, we have

$$|I_5| \leqslant \int_0^t \left(\int_{\mathbb{R}} |S_{u-} z| \mathbb{1}_{\{1 < |z| \leqslant \frac{1}{S_{u-}}\}} \mathbb{1}_{\{S_{u-} \in (0,1)\}} \, m(\mathrm{d}z) \right) \mathrm{d}u$$

$$+ \int_0^t \left(\int_{\mathbb{R}} |S_{u-} z| \mathbb{1}_{\{\frac{1}{S_{u-}} < |z| \leqslant 1\}} \mathbb{1}_{\{S_{u-} \in (1, \infty)\}} \, m(\mathrm{d}z) \right) \mathrm{d}u$$

$$\leqslant \int_0^t \left(\int_{\mathbb{R}} \mathbb{1}_{\{1 < |z|\}} \, m(\mathrm{d}z) \right) \mathrm{d}u + \left(\int_{\mathbb{R}} |S_{u-} z|^2 \mathbb{1}_{\{|z| \leqslant 1\}} \, m(\mathrm{d}z) \right) \mathrm{d}u$$

$$= t m(\{z \in \mathbb{R} : |z| > 1\}) + \int_0^t S_{u-}^2 \, \mathrm{d}u \int_{-1}^1 |z|^2 \, m(\mathrm{d}z) < \infty,$$

hence we conclude the existence of I_5. $\qquad\square$

In the next remark, we show that, for all $t \in [0, T]$, L_t is a measurable function of $(S_t)_{t \in [0, T]}$ depending on the parameter γ and the Lévy measure m.

10

Remark 2.4 For all $t \in [0, T]$ and $\delta \in (0, 1)$,

$$\int_{(0,t]} \int_{\{\delta < |z| \leqslant 1\}} z \big(\mu^L(\mathrm{d}u, \mathrm{d}z) - \mathrm{d}u\, m(\mathrm{d}z)\big) = \sum_{u \in [0,t]} \mathbb{1}_{\{\delta < |\Delta L_u| \leqslant 1\}} \Delta L_u - \int_{(0,t]} \int_{\{\delta < |z| \leqslant 1\}} z\, \mathrm{d}u\, m(\mathrm{d}z)$$

$$= \sum_{u \in [0,t]} \mathbb{1}_{\{\delta < \frac{|\Delta S_u|}{S_{u-}} \leqslant 1\}} \frac{\Delta S_u}{S_{u-}} - t \int_{\{\delta < |z| \leqslant 1\}} z\, m(\mathrm{d}z),$$

which is a measurable function of $(S_t)_{t \in [0,T]}$. Similarly, for all $t \in [0, T]$,

$$\int_{(0,t]} \int_{\{1 < |z| < \infty\}} z\, \mu^L(\mathrm{d}u, \mathrm{d}z) = \sum_{u \in [0,t]} \mathbb{1}_{\{|\Delta L_u| > 1\}} \Delta L_u = \sum_{u \in [0,t]} \mathbb{1}_{\{\frac{|\Delta S_u|}{S_{u-}} > 1\}} \frac{\Delta S_u}{S_{u-}},$$

which is a measurable function of $(S_t)_{t \in [0,T]}$ as well. Hence, using (2.4), for all $t \in [0, T]$,

$$\sum_{u \in [0,t]} \mathbb{1}_{\{\frac{|\Delta S_u|}{S_{u-}} > \delta\}} \frac{\Delta S_u}{S_{u-}} - t \int_{\{\delta < |z| \leqslant 1\}} z\, m(\mathrm{d}z) + \gamma t \xrightarrow{\mathbb{P}} L_t \quad \text{as } \delta \downarrow 0,$$

yielding that L_t is a measurable function of $(S_t)_{t \in [0,T]}$. In the special case of

$$(2.13) \qquad L_t = \sum_{s \in [0,t]} \Delta L_s, \qquad t \in [0, \infty),$$

the above statement readily follows from $\Delta L_s = \frac{\Delta S_s}{S_{s-}}$, $s \in [0, \infty)$. Condition (2.13) is satisfied if $\int_{-1}^1 |z|\, m(\mathrm{d}z) < \infty$ and $\gamma = \int_{-1}^1 z\, m(\mathrm{d}z)$, since, by (1.4),

$$L_t = \int_0^t \int_{\mathbb{R}} z\, \mu^L(\mathrm{d}u, \mathrm{d}z) + t\left(\gamma - \int_{-1}^1 z\, m(\mathrm{d}z)\right) = \sum_{s \in [0,t]} \Delta L_s + t\left(\gamma - \int_{-1}^1 z\, m(\mathrm{d}z)\right)$$

for $t \in [0, \infty)$, see Sato [35, Theorem 19.3]. Recall that, using (1.4), the Lévy process L has finite variation on each interval $[0, t]$, $t \in [0, \infty)$, if and only if $\int_{-1}^1 |z|\, m(\mathrm{d}z) < \infty$, see, e.g., Sato [35, Theorem 21.9]. For example, it is satisfied for a compound Poisson process given in (1.5), where m is a probability measure. □

In the next remark, we show that, for all $t \in [0, T]$, Y_t is a measurable function of $(S_t)_{t \in [0,T]}$.

Remark 2.5 Let $\theta, \kappa, \sigma \in (0, \infty)$, $\mu \in \mathbb{R}$, $\varrho \in (-1, 1)$. Let $(Y_t, S_t)_{t \in [0,\infty)}$ be the unique strong solution of the SDE (1.3) satisfying $\mathbb{P}(Y_0 \in [0, \infty), S_0 \in (0, \infty)) = 1$. The Grigelionis representation given in Proposition 2.3 implies that the continuous martingale part S^{cont} of S is

$$(2.14) \qquad S_t^{\mathrm{cont}} = \int_0^t S_u \sqrt{Y_u} \big(\varrho\, \mathrm{d}W_u + \sqrt{1 - \varrho^2}\, \mathrm{d}B_u\big), \qquad t \in [0, \infty),$$

11

see Jacod and Shiryaev [22, III.2.28 Remarks, part 1)]. Consequently, the (predictable) quadratic variation process of S^{cont} is $\langle S^{\mathrm{cont}} \rangle_t = \int_0^t S_u^2 Y_u \, \mathrm{d}u$, $t \in [0, \infty)$. Since

$$\mathbb{P}(S_t, \, S_{t-} \in (0, \infty) \ \text{ for all } \ t \in [0, \infty)) = 1$$

with the convention $S_{0-} := S_0$ (due to Proposition 2.1), one can apply Itô's rule to the function $f(x) = \log(x)$, $x \in (0, \infty)$, for which $f'(x) = 1/x$, $f''(x) = -1/x^2$, $x \in (0, \infty)$, and we obtain

(2.15)

$$
\log S_T = \log S_0 + \int_0^T \frac{\mathrm{d}S_u}{S_{u-}} - \frac{1}{2} \int_0^T \frac{1}{S_u^2} \, \mathrm{d}\langle S^{\mathrm{cont}} \rangle_u + \sum_{u \in [0,T]} \left(\log S_u - \log S_{u-} - \frac{1}{S_{u-}} \Delta S_u \right)
$$

$$
= \log S_0 + \mu T + \int_0^T \sqrt{Y_u} (\varrho \, \mathrm{d}W_u + \sqrt{1 - \varrho^2} \, \mathrm{d}B_u) - \frac{1}{2} \int_0^T Y_u \, \mathrm{d}u + L_T
$$

$$
+ \sum_{u \in [0,T]} \left(\log \frac{S_u}{S_{u-}} + 1 - \frac{S_u}{S_{u-}} \right)
$$

for $T \in [0, \infty)$, see, e.g., von Weizsäcker and Winkler [39, Theorem 8.4.1]. All terms in (2.15) are well-defined. In particular, the last term is a process with finite variation over each finite interval $[0, t]$, $t \in [0, \infty)$, see, e.g., Sato [35, Lemma 21.8.(iii)]. Taking into account of the Lévy–Itô's representation (2.4) of L, we conclude that the continuous martingale part $(\log S)^{\mathrm{cont}}$ of $\log S$ is $(\log S)_t^{\mathrm{cont}} = \int_0^t \sqrt{Y_u} (\varrho \, \mathrm{d}W_u + \sqrt{1 - \varrho^2} \, \mathrm{d}B_u)$, $t \in [0, \infty)$. Hence the (predictable) quadratic variation process of $(\log S)^{\mathrm{cont}}$ is

$$
\langle (\log S)^{\mathrm{cont}} \rangle_t = \int_0^t Y_u \, \mathrm{d}u, \qquad t \in [0, \infty).
$$

By Theorem I.4.47 a) in Jacod and Shiryaev [22],

$$
\sum_{i=1}^{\lfloor nt \rfloor} (\log S_{\frac{i}{n}} - \log S_{\frac{i-1}{n}})^2 \xrightarrow{\mathbb{P}} [\log S]_t \qquad \text{as } \ n \to \infty, \ \ t \in [0, \infty),
$$

where $\lfloor x \rfloor$ and $([\log S]_t)_{t \in [0, \infty)}$ denotes the integer part of a real number $x \in \mathbb{R}$, and the quadratic variation process of the semimartingale $\log S$, respectively. By Theorem I.4.52 in Jacod and Shiryaev [22],

$$
[\log S]_t = \langle (\log S)^{\mathrm{cont}} \rangle_t + \sum_{u \in [0,t]} (\log S_u - \log S_{u-})^2, \qquad t \in [0, \infty).
$$

Consequently, for all $t \in [0, \infty)$, we have

$$
\sum_{i=1}^{\lfloor nt \rfloor} (\log S_{\frac{i}{n}} - \log S_{\frac{i-1}{n}})^2 - \sum_{u \in [0,t]} (\log S_u - \log S_{u-})^2 \xrightarrow{\mathbb{P}} \langle (\log S)^{\mathrm{cont}} \rangle_t \qquad \text{as } \ n \to \infty.
$$

Note that this convergence holds almost surely along a suitable subsequence, the members of this sequence are measurable functions of $(S_u)_{u \in [0,t]}$, hence, using Theorems 4.2.2 and 4.2.8 in Dudley [16], we obtain that $\langle (\log S)^{\text{cont}} \rangle_t = \int_0^t Y_u \, du$ is a measurable function (i.e., a statistic) of $(S_u)_{u \in [0,t]}$. Moreover,

$$(2.16) \qquad \frac{\langle (\log S)^{\text{cont}} \rangle_{t+h} - \langle (\log S)^{\text{cont}} \rangle_t}{h} = \frac{1}{h} \int_t^{t+h} Y_s \, ds \xrightarrow{\text{a.s.}} Y_t \qquad \text{as } h \to 0, \quad t \in [0, \infty),$$

since Y has continuous sample paths almost surely. Consequently, for all $t \in [0, T]$, Y_t is a measurable function (i.e., a statistic) of $(S_u)_{u \in [0,T]}$ (where for $t = T$, one may take $h \uparrow 0$), however, we also point out that this measurable function remains inexplicit. $\qquad \square$

Next we give statistics for the parameters σ and ϱ using continuous time observations $(S_t)_{t \in [0,T]}$ with some $T > 0$. Due to this result we do not consider the estimation of these parameters, they are supposed to be known.

Remark 2.6 Let $\theta, \kappa, \sigma \in (0, \infty)$, $\mu \in \mathbb{R}$, $\varrho \in (-1, 1)$, and $\mathbb{P}(Y_0 \in [0, \infty), S_0 \in (0, \infty)) = 1$. Then for all $T > 0$,

$$\begin{bmatrix} \sigma^2 & \varrho\sigma \\ \varrho\sigma & 1 \end{bmatrix} = \frac{1}{\int_0^T Y_s \, ds} \begin{bmatrix} \langle Y \rangle_T & \langle Y, (\log S)^{\text{cont}} \rangle_T \\ \langle Y, (\log S)^{\text{cont}} \rangle_T & \langle (\log S)^{\text{cont}} \rangle_T \end{bmatrix} =: \widehat{\Sigma}_T,$$

where $(\langle Y, (\log S)^{\text{cont}} \rangle_t)_{t \in [0, \infty)}$ denotes the (predictable) quadratic covariation process of Y and $(\log S)^{\text{cont}}$, since, by the SDEs (1.3) and (2.15),

$$\langle Y \rangle_T = \sigma^2 \int_0^T Y_u \, du, \qquad \langle (\log S)^{\text{cont}} \rangle_T = \int_0^T Y_u \, du, \qquad \langle Y, (\log S)^{\text{cont}} \rangle_T = \varrho\sigma \int_0^T Y_u \, du.$$

We point out that $\mathbb{P}\left(\int_0^T Y_u \, du \in (0, \infty)\right) = 1$. Indeed, if $\omega \in \Omega$ is such that $[0, T] \ni s \mapsto Y_s(\omega)$ is continuous and $Y_t(\omega) \in [0, \infty)$ for all $t \in [0, \infty)$, then we have $\int_0^T Y_u(\omega) \, du = 0$ if and only if $Y_u(\omega) = 0$ for all $u \in [0, T]$. Using the method of the proof of Theorem 3.1 in Barczy et. al [4], we get $\mathbb{P}(\int_0^T Y_u \, du = 0) = 0$, as desired. We note that $\widehat{\Sigma}_T$ is a statistic, i.e., there exists a measurable function $\Xi : D([0,T], \mathbb{R}) \to \mathbb{R}^{2 \times 2}$ such that $\widehat{\Sigma}_T = \Xi((S_u)_{u \in [0,T]})$, where $D([0,T], \mathbb{R})$ denotes the space of real-valued càdlàg functions defined on $[0, T]$, since

$$(2.17) \quad \begin{aligned} & \frac{1}{\frac{1}{n}\sum_{i=1}^{\lfloor nT \rfloor} Y_{\frac{i-1}{n}}} \sum_{i=1}^{\lfloor nT \rfloor} \begin{bmatrix} Y_{\frac{i}{n}} - Y_{\frac{i-1}{n}} \\ \log S_{\frac{i}{n}} - \log S_{\frac{i-1}{n}} \end{bmatrix} \begin{bmatrix} Y_{\frac{i}{n}} - Y_{\frac{i-1}{n}} \\ \log S_{\frac{i}{n}} - \log S_{\frac{i-1}{n}} \end{bmatrix}^\top \\ & - \frac{1}{\frac{1}{n}\sum_{i=1}^{\lfloor nT \rfloor} Y_{\frac{i-1}{n}}} \sum_{u \in [0,T]} \begin{bmatrix} 0 & 0 \\ 0 & (\log S_u - \log S_{u-})^2 \end{bmatrix} \xrightarrow{\mathbb{P}} \widehat{\Sigma}_T \qquad \text{as } n \to \infty, \end{aligned}$$

where the convergence in (2.17) holds almost surely along a suitable subsequence, the members of the sequence in (2.17) are measurable functions of $(S_u)_{u \in [0,T]}$ (due to Remark 2.5), and one

13

can use Theorems 4.2.2 and 4.2.8 in Dudley [16]. Next we prove (2.17). By Theorems I.4.47 a) and I.4.52 in Jacod and Shiryaev [22],

$$\sum_{i=1}^{\lfloor nT \rfloor} (Y_{\frac{i}{n}} - Y_{\frac{i-1}{n}})^2 \xrightarrow{\mathbb{P}} [Y]_T = \langle Y \rangle_T,$$

$$\sum_{i=1}^{\lfloor nT \rfloor} (\log S_{\frac{i}{n}} - \log S_{\frac{i-1}{n}})^2 \xrightarrow{\mathbb{P}} [\log S]_T = \langle (\log S)^{\mathrm{cont}} \rangle_T + \sum_{u \in [0,T]} (\log S_u - \log S_{u-})^2,$$

$$\sum_{i=1}^{\lfloor nT \rfloor} (Y_{\frac{i}{n}} - Y_{\frac{i-1}{n}})(\log S_{\frac{i}{n}} - \log S_{\frac{i-1}{n}}) \xrightarrow{\mathbb{P}} [Y, \log S]_T = \langle Y, (\log S)^{\mathrm{cont}} \rangle_T$$

as $n \to \infty$, where $([Y, \log S]_t)_{t \in [0,\infty)}$ denotes the quadratic covariation process of the semimartingales Y and $\log S$. Consequently,

$$\sum_{i=1}^{\lfloor nT \rfloor} \begin{bmatrix} Y_{\frac{i}{n}} - Y_{\frac{i-1}{n}} \\ \log S_{\frac{i}{n}} - \log S_{\frac{i-1}{n}} \end{bmatrix} \begin{bmatrix} Y_{\frac{i}{n}} - Y_{\frac{i-1}{n}} \\ \log S_{\frac{i}{n}} - \log S_{\frac{i-1}{n}} \end{bmatrix}^\top$$

$$- \sum_{u \in [0,T]} \begin{bmatrix} 0 & 0 \\ 0 & (\log S_u - \log S_{u-})^2 \end{bmatrix} \xrightarrow{\mathbb{P}} \left(\int_0^T Y_u \, \mathrm{d}u \right) \widehat{\Sigma}_T \qquad \text{as } n \to \infty,$$

see, e.g., van der Vaart [38, Theorem 2.7, part (vi)]. Moreover,

$$\frac{1}{n} \sum_{i=1}^{\lfloor nT \rfloor} Y_{\frac{i-1}{n}} \xrightarrow{\text{a.s.}} \int_0^T Y_u \, \mathrm{d}u \qquad \text{as } n \to \infty$$

since Y has continuous sample paths almost surely. Hence (2.17) follows by Slutsky's lemma. Finally, we note that the sample size T is fixed above, and it is enough to know any short sample $(S_u)_{u \in [0,T]}$ to carry out the above calculations. $\qquad \square$

3 Existence and uniqueness of MLE

From this section, we will consider the jump-type Heston model (1.3) with known $\sigma \in (0,\infty)$, $\varrho \in (-1,1)$, $\gamma \in \mathbb{R}$, Lévy measure m, and deterministic initial value $(Y_0, S_0) = (y_0, s_0) \in (0,\infty)^2$, and we will consider $\boldsymbol{\psi} := (\theta, \kappa, \mu) \in (0,\infty)^2 \times \mathbb{R} =: \Psi$ as a parameter.

Let $\mathbb{P}_{\boldsymbol{\psi}}$ denote the probability measure induced by $(Y_t, S_t)_{t \in [0,\infty)}$ on the measurable space $(D([0,\infty), \mathbb{R}^2), \mathcal{D}([0,\infty), \mathbb{R}^2))$ of $\mathbb{R}^2$-valued càdlàg functions defined on $[0,\infty)$ endowed with a right continuous filtration $(\mathcal{D}_t([0,\infty), \mathbb{R}^2))_{t \in [0,\infty)}$, see Appendix A. Further, for all $T \in (0,\infty)$, let $\mathbb{P}_{\boldsymbol{\psi},T} := \mathbb{P}_{\boldsymbol{\psi}}|_{\mathcal{D}_T([0,\infty),\mathbb{R}^2)}$ be the restriction of $\mathbb{P}_{\boldsymbol{\psi}}$ to $\mathcal{D}_T([0,\infty), \mathbb{R}^2)$.

Let us write the Heston model (1.3) in the form

$$(3.1) \qquad \begin{bmatrix} \mathrm{d}Y_t \\ \mathrm{d}S_t \end{bmatrix} = A(Y_t, S_t) H(\boldsymbol{\psi}) \, \mathrm{d}t + \Gamma(Y_t, S_t) \begin{bmatrix} \mathrm{d}W_t \\ \mathrm{d}B_t \end{bmatrix} + \begin{bmatrix} 0 \\ S_{t-} \, \mathrm{d}L_t \end{bmatrix}, \qquad t \in [0,\infty),$$

14

where the functions $A : [0, \infty) \times (0, \infty) \to \mathbb{R}^{2 \times 3}$, $\Gamma : [0, \infty) \times (0, \infty) \to \mathbb{R}^{2 \times 2}$ and $H : \mathbb{R}^3 \to \mathbb{R}^3$ are defined by

$$
A(y, s) := \begin{bmatrix} 1 & -y & 0 \\ 0 & 0 & s \end{bmatrix}, \qquad \Gamma(y, s) := \sqrt{y} \begin{bmatrix} \sigma & 0 \\ \varrho s & \sqrt{1 - \varrho^2} s \end{bmatrix}, \qquad H(x_1, x_2, x_3) := \begin{bmatrix} x_1 x_2 \\ x_2 \\ x_3 \end{bmatrix}
$$

for $(y, s) \in [0, \infty) \times (0, \infty)$ and $(x_1, x_2, x_3) \in \mathbb{R}^3$. Note that H is bijective on the set $\mathbb{R} \times (\mathbb{R} \setminus \{0\}) \times \mathbb{R}$ having inverse

$$(3.2) \qquad H^{-1}(y_1, y_2, y_3) = \left(\frac{y_1}{y_2}, y_2, y_3 \right), \qquad (y_1, y_2, y_3) \in \mathbb{R} \times (\mathbb{R} \setminus \{0\}) \times \mathbb{R}.$$

Let us introduce the function $\Sigma : [0, \infty) \times (0, \infty) \to \mathbb{R}^{2 \times 2}$ given by

$$
\Sigma(y, s) := \Gamma(y, s)\Gamma(y, s)^\top = y \begin{bmatrix} \sigma^2 & \varrho \sigma s \\ \varrho \sigma s & s^2 \end{bmatrix}, \qquad (y, s) \in [0, \infty) \times (0, \infty).
$$

If $(y, s) \in (0, \infty)^2$ then $\Sigma(y, s)$ is invertible, namely,

$$(3.3)$$
$$
\Sigma(y, s)^{-1} = (\Gamma(y, s)^\top)^{-1} \Gamma(y, s)^{-1} = \frac{1}{(1 - \varrho^2)\sigma^2 y s^2} \begin{bmatrix} \sqrt{1 - \varrho^2} s & -\varrho s \\ 0 & \sigma \end{bmatrix} \begin{bmatrix} \sqrt{1 - \varrho^2} s & 0 \\ -\varrho s & \sigma \end{bmatrix}
$$
$$
= \frac{1}{(1 - \varrho^2)\sigma^2 y s^2} \begin{bmatrix} s^2 & -\varrho \sigma s \\ -\varrho \sigma s & \sigma^2 \end{bmatrix}.
$$

Further, let

$$
\boldsymbol{G}_t := \int_0^t A(Y_u, S_u)^\top \Sigma(Y_u, S_u)^{-1} A(Y_u, S_u) \, \mathrm{d}u, \qquad t \in [0, \infty),
$$

$$
\boldsymbol{f}_t := \int_0^t A(Y_{u-}, S_{u-})^\top \Sigma(Y_{u-}, S_{u-})^{-1} \begin{bmatrix} \mathrm{d}Y_u \\ \mathrm{d}S_u - S_{u-} \, \mathrm{d}L_u \end{bmatrix}, \qquad t \in [0, \infty),
$$

provided that $\mathbb{P}(Y_t \in (0, \infty) \text{ for all } t \in [0, \infty)) = 1$, which holds if $\theta \kappa \in \left[\frac{\sigma^2}{2}, \infty \right)$. Using (3.3), we obtain

$$
\boldsymbol{G}_t = \int_0^t A(Y_u, S_u)^\top (\Gamma(Y_u, S_u)^\top)^{-1} \Gamma(Y_u, S_u)^{-1} A(Y_u, S_u) \, \mathrm{d}u
$$

$$(3.4)$$
$$
= \int_0^t (\Gamma(Y_u, S_u)^{-1} A(Y_u, S_u))^\top (\Gamma(Y_u, S_u)^{-1} A(Y_u, S_u)) \, \mathrm{d}u
$$

$$
= \frac{1}{(1 - \varrho^2)\sigma^2} \int_0^t \frac{1}{Y_u} \begin{bmatrix} 1 & -Y_u & -\varrho \sigma \\ -Y_u & Y_u^2 & \varrho \sigma Y_u \\ -\varrho \sigma & \varrho \sigma Y_u & \sigma^2 \end{bmatrix} \mathrm{d}u, \qquad t \in [0, \infty),
$$

15

provided that $\mathbb{P}(Y_t \in (0, \infty)$ for all $t \in [0, \infty)) = 1$, since

$$(3.5) \quad \Gamma(y, s)^{-1} A(y, s) = \frac{1}{\sigma\sqrt{(1 - \varrho^2)y}} \begin{bmatrix} \sqrt{1 - \varrho^2} & -y\sqrt{1 - \varrho^2} & 0 \\ -\varrho & \varrho y & \sigma \end{bmatrix}, \quad (y, s) \in (0, \infty)^2.$$

The next lemma is about the form of the Radon–Nikodym derivative $\frac{\mathrm{d}\,\mathbb{P}_{\psi,T}}{\mathrm{d}\,\mathbb{P}_{\widetilde{\psi},T}}$ for certain $\psi, \widetilde{\psi} \in \Psi$.

Lemma 3.1 *Let* $\psi = (\theta, \kappa, \mu) \in \Psi$ *and* $\widetilde{\psi} := (\widetilde{\theta}, \widetilde{\kappa}, \widetilde{\mu}) \in \Psi$ *with* $\theta\kappa, \widetilde{\theta}\widetilde{\kappa} \in \left[\frac{\sigma^2}{2}, \infty\right)$. *Then for all* $T \in (0, \infty)$, *the probability measures* $\mathbb{P}_{\psi,T}$ *and* $\mathbb{P}_{\widetilde{\psi},T}$ *are absolutely continuous with respect to each other, and, under* $\mathbb{P}$,

$$(3.6) \quad \log \frac{\mathrm{d}\,\mathbb{P}_{\psi,T}}{\mathrm{d}\,\mathbb{P}_{\widetilde{\psi},T}}(\widetilde{Y}, \widetilde{S}) = \left(H(\psi) - H(\widetilde{\psi})\right)^\top \widetilde{f}_T - \frac{1}{2}\left(H(\psi) - H(\widetilde{\psi})\right)^\top \widetilde{G}_T\left(H(\psi) + H(\widetilde{\psi})\right),$$

where $\widetilde{Y}$, $\widetilde{S}$, $\widetilde{G}$ *and* $\widetilde{f}$ *are the processes corresponding to the parameter* $\widetilde{\psi}$.

Proof. In what follows, we will apply Theorem III.5.34 in Jacod and Shiryaev [22] (see also Appendix A). We will work on the canonical space $(D([0, \infty), \mathbb{R}^2), \mathcal{D}([0, \infty), \mathbb{R}^2))$. Let $(\eta_t, \zeta_t)_{t \in [0,\infty)}$ denote the canonical process $(\eta_t, \zeta_t)(\omega) := \omega(t)$, $\omega \in D([0, \infty), \mathbb{R}^2)$, $t \in [0, \infty)$. Using (3.1) and (2.4), the Heston model (1.3) can be written in the form

$$\begin{bmatrix} Y_t \\ S_t \end{bmatrix} = \begin{bmatrix} y_0 \\ s_0 \end{bmatrix} + \int_0^t \left(A(Y_u, S_u) H(\psi) + \begin{bmatrix} 0 \\ \gamma S_u \end{bmatrix} \right) \mathrm{d}u$$

$$+ \int_0^t \Gamma(Y_u, S_u) \begin{bmatrix} \mathrm{d}W_u \\ \mathrm{d}B_u \end{bmatrix} + \begin{bmatrix} 0 \\ \int_0^t \int_\mathbb{R} S_{u-} h_1(z)(\mu^L(\mathrm{d}u, \mathrm{d}z) - \mathrm{d}u\, m(\mathrm{d}z)) \end{bmatrix}$$

$$+ \begin{bmatrix} 0 \\ \int_0^t \int_\mathbb{R} S_{u-}(z - h_1(z))\, \mu^L(\mathrm{d}u, \mathrm{d}z) \end{bmatrix}, \quad t \in [0, \infty).$$

Using Proposition 2.3, we obtain

$$\begin{bmatrix} Y_t \\ S_t \end{bmatrix} = \begin{bmatrix} y_0 \\ s_0 \end{bmatrix} + \int_0^t \left(A(Y_u, S_u) H(\psi) + \begin{bmatrix} 0 \\ \gamma S_u + \int_\mathbb{R}(h_1(S_{u-}z) - S_{u-}h_1(z))\, m(\mathrm{d}z) \end{bmatrix} \right) \mathrm{d}u$$

$$+ \int_0^t \Gamma(Y_u, S_u) \begin{bmatrix} \mathrm{d}W_u \\ \mathrm{d}B_u \end{bmatrix} + \begin{bmatrix} 0 \\ \int_0^t \int_\mathbb{R} h_1(S_{u-}z)(\mu^L(\mathrm{d}u, \mathrm{d}z) - \mathrm{d}u\, m(\mathrm{d}z)) \end{bmatrix}$$

$$+ \begin{bmatrix} 0 \\ \int_0^t \int_\mathbb{R}(S_{u-}z - h_1(S_{u-}z))\, \mu^L(\mathrm{d}u, \mathrm{d}z) \end{bmatrix}, \quad t \in [0, \infty),$$

16

which is a special case of III.2.23 in Jacod and Shiryaev [22], since

$$
\begin{bmatrix} Y_t \\ S_t \end{bmatrix} = \begin{bmatrix} y_0 \\ s_0 \end{bmatrix} + \int_0^t \left(A(Y_u, S_u) H(\boldsymbol{\psi}) + \begin{bmatrix} 0 \\ \gamma S_u + \int_{\mathbb{R}} (h_1(S_{u-}z) - S_{u-}h_1(z))\, m(\mathrm{d}z) \end{bmatrix} \right) \mathrm{d}u
$$

$$
(3.7) \qquad + \int_0^t \Gamma(Y_u, S_u) \begin{bmatrix} \mathrm{d}W_u \\ \mathrm{d}B_u \end{bmatrix} + \int_0^t \int_{\mathbb{R}} h_2 \left(\begin{bmatrix} 0 \\ S_{u-}z \end{bmatrix} \right) (\mu^L(\mathrm{d}u, \mathrm{d}z) - \mathrm{d}u\, m(\mathrm{d}z))
$$

$$
+ \int_0^t \int_{\mathbb{R}} \left(\begin{bmatrix} 0 \\ S_{u-}z \end{bmatrix} - h_2 \left(\begin{bmatrix} 0 \\ S_{u-}z \end{bmatrix} \right) \right) \mu^L(\mathrm{d}u, \mathrm{d}z), \qquad t \in [0, \infty),
$$

with the truncation function $h_2(\boldsymbol{z}) := \boldsymbol{z} \mathbb{1}_{[-1,1]^2}(\boldsymbol{z})$, $\boldsymbol{z} \in \mathbb{R}^2$, where we used that

$$
\begin{bmatrix} 0 \\ h_1(z) \end{bmatrix} = h_2 \left(\begin{bmatrix} 0 \\ z \end{bmatrix} \right), \qquad z \in \mathbb{R}.
$$

By Proposition 2.1, the SDE (1.3) has a pathwise unique strong solution (with the given deterministic initial value $(y_0, s_0) \in (0, \infty)^2$), and hence, by Theorem III.2.26 in Jacod and Shiryaev [22], under the probability measure $\mathbb{P}_{\boldsymbol{\psi}}$, the canonical process $(\eta_t, \zeta_t)_{t \in [0, \infty)}$ is a semimartingale with semimartingale characteristics $(B^{(\boldsymbol{\psi})}, C, \nu)$ associated with the truncation function h_2, where

$$
B_t^{(\boldsymbol{\psi})} = \int_0^t \left(A(\eta_u, \zeta_u) H(\boldsymbol{\psi}) + \begin{bmatrix} 0 \\ \gamma \zeta_u + \int_{\mathbb{R}} (h_1(\zeta_u z) - \zeta_u h_1(z))\, m(\mathrm{d}z) \end{bmatrix} \right) \mathrm{d}u,
$$

$$
C_t = \int_0^t \Gamma(\eta_u, \zeta_u) \Gamma(\eta_u, \zeta_u)^\top \, \mathrm{d}u = \int_0^t \Sigma(\eta_u, \zeta_u)\, \mathrm{d}u,
$$

for $t \in [0, \infty)$, and

$$
\nu(\mathrm{d}t, \mathrm{d}y, \mathrm{d}z) = K(\eta_t, \zeta_t, \mathrm{d}y, \mathrm{d}z)\, \mathrm{d}t
$$

with the Borel transition kernel K from $[0, \infty)^2 \times \mathbb{R}^2$ into $\mathbb{R}^2$ given by

$$
K(y, s, R) := \int_{\mathbb{R}} \mathbb{1}_{R \setminus \{\boldsymbol{0}\}}(0, sz)\, m(\mathrm{d}z) \qquad \text{for } (y, s) \in [0, \infty)^2 \text{ and } R \in \mathcal{B}(\mathbb{R}^2).
$$

The aim of the following discussion is to check the set of sufficient conditions presented in Appendix A (of which the notations will be used) in order to have right to apply Theorem III.5.34 in Jacod and Shiryaev [22]. First note that $(C_t)_{t \in [0, \infty)}$ and $\nu(\mathrm{d}t, \mathrm{d}y, \mathrm{d}z)$ do not depend on the unknown parameter $\boldsymbol{\psi}$, and hence $V^{(\boldsymbol{\psi}, \widetilde{\boldsymbol{\psi}})}$ is identically one and then (A.3) and (A.4) readily hold. We also have

$$
\mathbb{P}_{\boldsymbol{\psi}} \left(\nu(\{t\} \times \mathbb{R}^2) = 0 \right) = \mathbb{P}_{\boldsymbol{\psi}} \left(\int_{\{t\}} K(\eta_s, \zeta_s, \mathbb{R}^2)\, \mathrm{d}s = 0 \right) = 1, \qquad t \in [0, \infty), \quad \boldsymbol{\psi} \in \Psi.
$$

Further, $(C_t)_{t \in [0,\infty)}$ can be represented as $C_t = \int_0^t c_u \, \mathrm{d}F_u$, $t \in [0,\infty)$, where the stochastic processes $(c_t)_{t \in [0,\infty)}$ and $(F_t)_{t \in [0,\infty)}$ are given by $c_t := \Sigma(\eta_t, \zeta_t)$, $t \in [0,\infty)$, and $F_t = t$, $t \in [0,\infty)$. Next, note that, under the condition $\theta\kappa \in [\frac{\sigma^2}{2}, \infty)$, we have $\mathbb{P}_{\psi}((\eta_t, \zeta_t) \in (0,\infty)^2$ for all $t \in [0,\infty)) = 1$ (due to Proposition 2.1), hence, by (3.3), for each $t \in [0,\infty)$, the matrix c_t is invertible $\mathbb{P}_{\psi}$-almost surely. Consequently, for all $\boldsymbol{\psi} = (\theta, \kappa, \mu) \in \Psi$ and $\widetilde{\boldsymbol{\psi}} = (\widetilde{\theta}, \widetilde{\kappa}, \widetilde{\mu}) \in \Psi$ with $\theta\kappa, \widetilde{\theta}\widetilde{\kappa} \in [\frac{\sigma^2}{2}, \infty)$,

$$B_t^{(\boldsymbol{\psi})} - B_t^{(\widetilde{\boldsymbol{\psi}})} = \int_0^t A(\eta_u, \zeta_u)(H(\boldsymbol{\psi}) - H(\widetilde{\boldsymbol{\psi}})) \, \mathrm{d}u = \int_0^t c_u \beta_u^{(\widetilde{\boldsymbol{\psi}}, \boldsymbol{\psi})} \, \mathrm{d}F_u, \qquad t \in [0,\infty),$$

where the stochastic process $(\beta_t^{(\widetilde{\boldsymbol{\psi}}, \boldsymbol{\psi})})_{t \in [0,\infty)}$ is given by

$$\beta_t^{(\widetilde{\boldsymbol{\psi}}, \boldsymbol{\psi})} = c_t^{-1} A(\eta_t, \zeta_t)(H(\boldsymbol{\psi}) - H(\widetilde{\boldsymbol{\psi}})) = \Sigma(\eta_t, \zeta_t)^{-1} A(\eta_t, \zeta_t)(H(\boldsymbol{\psi}) - H(\widetilde{\boldsymbol{\psi}})), \qquad t \in [0,\infty),$$

which yields (A.5).

Next we check (A.6), i.e.,

(3.8) $$\mathbb{P}_{\psi}\left(\int_0^t (\beta_u^{(\widetilde{\boldsymbol{\psi}}, \boldsymbol{\psi})})^{\top} c_u \beta_u^{(\widetilde{\boldsymbol{\psi}}, \boldsymbol{\psi})} \, \mathrm{d}F_u < \infty\right) = 1, \qquad t \in [0,\infty).$$

We have

$$\int_0^t (\beta_s^{(\widetilde{\boldsymbol{\psi}}, \boldsymbol{\psi})})^{\top} c_s \beta_s^{(\widetilde{\boldsymbol{\psi}}, \boldsymbol{\psi})} \, \mathrm{d}F_s$$

$$= (H(\boldsymbol{\psi}) - H(\widetilde{\boldsymbol{\psi}}))^{\top} \int_0^t A(\eta_s, \zeta_s)^{\top} (\Sigma(\eta_s, \zeta_s)^{-1})^{\top} A(\eta_s, \zeta_s) \, \mathrm{d}s \, (H(\boldsymbol{\psi}) - H(\widetilde{\boldsymbol{\psi}}))$$

$$= (H(\boldsymbol{\psi}) - H(\widetilde{\boldsymbol{\psi}}))^{\top} G_t (H(\boldsymbol{\psi}) - H(\widetilde{\boldsymbol{\psi}})), \qquad t \in [0,\infty),$$

where G_t, $t \in [0,\infty)$, is understood as the original $\boldsymbol{G}_t$, $t \in [0,\infty)$, replacing (Y, S) by (η, ζ). Since η has continuous sample paths $\mathbb{P}_{\psi}$-almost surely and $\mathbb{P}_{\psi}(\eta_t \in (0,\infty), \forall t \in [0,\infty)) = 1$ (due to $\theta\kappa \in [\frac{\sigma^2}{2}, \infty)$), we have $\mathbb{P}_{\psi}(\inf_{t \in [0,T]} \eta_t \in (0,\infty)) = 1$ for all $T \in [0,\infty)$, which, together with the $\mathbb{P}_{\psi}$-almost sure continuity of η and formula (3.4), yield (3.8).

Next, we check that, under the probability measure $\mathbb{P}_{\psi}$, local uniqueness holds for the martingale problem on the canonical space corresponding to the triplet $(B^{(\boldsymbol{\psi})}, C, \nu)$ with the given initial value (y_0, s_0) with $\mathbb{P}_{\psi}$ as its unique solution. By Proposition 2.1, the SDE (1.3) has a pathwise unique strong solution (with the given deterministic initial value $(y_0, s_0) \in (0,\infty)^2$), and hence Theorem III.2.26 in Jacod and Shiryaev [22] yields that the set of all solutions to the martingale problem on the canonical space corresponding to $(B^{(\boldsymbol{\psi})}, C, \nu)$ has only one element $(\mathbb{P}_{\psi})$ yielding the desired local uniqueness. We also mention that Theorem III.4.29 in Jacod and Shiryaev [22] implies that under the probability measure $\mathbb{P}_{\psi}$, all local martingales have the integral representation property relative to (η, ζ).

By Theorem III.5.34 in Jacod and Shiryaev [22] (see also Appendix A), $\mathbb{P}_{\psi,T}$ and $\mathbb{P}_{\widetilde{\psi},T}$ are equivalent (one can change the roles of $\boldsymbol{\psi}$ and $\widetilde{\boldsymbol{\psi}}$), and under the probability measure

18

$\mathbb{P}_{\widetilde{\psi},T},$ we have

$$\frac{\mathrm{d}\,\mathbb{P}_{\psi,T}}{\mathrm{d}\,\mathbb{P}_{\widetilde{\psi},T}}(\eta,\zeta) = \exp\left\{ \int_0^T \left(\beta_u^{(\widetilde{\psi},\psi)}\right)^\top \begin{bmatrix} \mathrm{d}(\eta^{\mathrm{cont}})_u^{(\widetilde{\psi})} \\ \mathrm{d}(\zeta^{\mathrm{cont}})_u^{(\widetilde{\psi})} \end{bmatrix} - \frac{1}{2}\int_0^T \left(\beta_u^{(\widetilde{\psi},\psi)}\right)^\top c_u \beta_u^{(\widetilde{\psi},\psi)}\,\mathrm{d}u \right\}$$

for $T \in (0,\infty)$, where $((\eta^{\mathrm{cont}})_t^{(\widetilde{\psi})}, (\zeta^{\mathrm{cont}})_t^{(\widetilde{\psi})})_{t\in[0,\infty)}$ denotes the continuous (local) martingale part of $(\eta_t, \zeta_t)_{t\in[0,\infty)}$ under $\mathbb{P}_{\widetilde{\psi}}$. Using part 1) of Remarks III.2.28 in Jacod and Shiryaev [22] and (3.7), the continuous (local) martingale part $(\widetilde{Y}_t^{\mathrm{cont}}, \widetilde{S}_t^{\mathrm{cont}})_{t\in[0,\infty)}$ of $(\widetilde{Y}_t, \widetilde{S}_t)_{t\in[0,\infty)}$ takes the form

$$\begin{bmatrix} \widetilde{Y}_t^{\mathrm{cont}} \\ \widetilde{S}_t^{\mathrm{cont}} \end{bmatrix} = \int_0^t \Gamma(\widetilde{Y}_u, \widetilde{S}_u) \begin{bmatrix} \mathrm{d}W_u \\ \mathrm{d}B_u \end{bmatrix}, \qquad t \in [0,\infty),$$

and, by (3.1), we have

$$(3.9) \qquad \begin{bmatrix} \mathrm{d}\widetilde{Y}_t^{\mathrm{cont}} \\ \mathrm{d}\widetilde{S}_t^{\mathrm{cont}} \end{bmatrix} = \begin{bmatrix} \mathrm{d}\widetilde{Y}_t \\ \mathrm{d}\widetilde{S}_t \end{bmatrix} - A(\widetilde{Y}_t, \widetilde{S}_t)H(\widetilde{\psi})\,\mathrm{d}t - \begin{bmatrix} 0 \\ \widetilde{S}_{t-}\,\mathrm{d}L_t \end{bmatrix}, \qquad t \in [0,\infty).$$

Hence, under $\mathbb{P}$,

$$\log\frac{\mathrm{d}\,\mathbb{P}_{\psi,T}}{\mathrm{d}\,\mathbb{P}_{\widetilde{\psi},T}}(\widetilde{Y},\widetilde{S}) = (H(\psi) - H(\widetilde{\psi}))^\top \int_0^T A(\widetilde{Y}_u, \widetilde{S}_{u-})^\top \Sigma(\widetilde{Y}_u, \widetilde{S}_{u-})^{-1} \begin{bmatrix} \mathrm{d}\widetilde{Y}_u \\ \mathrm{d}\widetilde{S}_u - \widetilde{S}_{u-}\,\mathrm{d}L_u \end{bmatrix}$$

$$- (H(\psi) - H(\widetilde{\psi}))^\top \left(\int_0^T A(\widetilde{Y}_u, \widetilde{S}_u)^\top \Sigma(\widetilde{Y}_u, \widetilde{S}_u)^{-1} A(\widetilde{Y}_u, \widetilde{S}_u)\,\mathrm{d}u \right) H(\widetilde{\psi})$$

$$- \frac{1}{2}(H(\psi) - H(\widetilde{\psi}))^\top \left(\int_0^T A(\widetilde{Y}_u, \widetilde{S}_u)^\top \Sigma(\widetilde{Y}_u, \widetilde{S}_u)^{-1} A(\widetilde{Y}_u, \widetilde{S}_u)\,\mathrm{d}u \right) (H(\psi) - H(\widetilde{\psi}))$$

$$= (H(\psi) - H(\widetilde{\psi}))^\top \widetilde{f}_T - (H(\psi) - H(\widetilde{\psi}))^\top \widetilde{G}_T H(\widetilde{\psi})$$

$$- \frac{1}{2}(H(\psi) - H(\widetilde{\psi}))^\top \widetilde{G}_T (H(\psi) - H(\widetilde{\psi})),$$

which yields the statement. $\qquad\square$

Note that $\widetilde{f}_T$ in Lemma 3.1 contains a stochastic integral with respect to L, but, by Remark 2.4, for all $t \in [0,T]$, L_t is a measurable function of $(\widetilde{S}_t)_{t\in[0,T]}$ (depending on γ and m).

We point out that we use the condition $\theta\kappa \in [\frac{\sigma^2}{2}, \infty)$ in the proof of Lemma 3.1 to assure the invertibility of $(c_t)_{t\in[0,\infty)}$.

Next, using Lemma 3.1, by considering $\mathbb{P}_{\widetilde{\psi},T}$ as a fixed reference measure, we will derive an MLE for the parameter $\psi = (\theta,\kappa,\mu)$ based on the observations $(Y_t, S_t)_{t\in[0,T]}$. Our method for deriving an MLE is one of the known ones in the literature, and it turns out that these lead to the same estimator $\widehat{\psi}_T$, see Remark 3.5. Let us denote the right hand side of (3.6) by

$\Lambda_T(\boldsymbol{\psi}, \widetilde{\boldsymbol{\psi}})$ replacing $(\widetilde{\boldsymbol{f}}_T, \widetilde{\boldsymbol{G}}_T)$ by $(\boldsymbol{f}_T, \boldsymbol{G}_T)$. For convenience, first we calculate an MLE $\widehat{\boldsymbol{\psi}}_T$ of the parameter $\boldsymbol{\psi}$ on the set $\mathbb{R}^3$ based on the observations $(Y_t, S_t)_{t \in [0,T]}$, namely,

$$\widehat{\boldsymbol{\psi}}_T := \arg\max_{\boldsymbol{\psi} \in \mathbb{R}^3} \Lambda_T(\boldsymbol{\psi}, \widetilde{\boldsymbol{\psi}}),$$

which will turn out to be not dependent on $\widetilde{\boldsymbol{\psi}}$. Here the function Λ_T is extended for all $\boldsymbol{\psi} = (\theta, \kappa, \mu) \in \mathbb{R}^3$ in a natural way (note that for the calculation of the random matrices $\boldsymbol{G}_t$, $t \in [0, \infty)$, and the random vectors $\boldsymbol{f}_t$, $t \in [0, \infty)$, one does not need to know the parameters $\boldsymbol{\psi}$ or $\widetilde{\boldsymbol{\psi}}$). In Remark 3.3, we describe the connection between $\widehat{\boldsymbol{\psi}}_T$ and an MLE given by $\arg\max_{\boldsymbol{\psi} \in \Psi} \Lambda_T(\boldsymbol{\psi}, \widetilde{\boldsymbol{\psi}})$ on the set Ψ.

Proposition 3.2 *Let* $\theta, \kappa \in (0, \infty)$ *with* $\theta\kappa \in \left[\frac{\sigma^2}{2}, \infty\right)$, $\mu \in \mathbb{R}$, $\sigma \in (0, \infty)$, $\varrho \in (-1, 1)$, *and* $(Y_0, S_0) = (y_0, s_0) \in (0, \infty)^2$. *Then for all* $T \in (0, \infty)$, *there exists a unique MLE* $\widehat{\boldsymbol{\psi}}_T = (\widehat{\theta}_T, \widehat{\kappa}_T, \widehat{\mu}_T)^\top$ *of* $\boldsymbol{\psi} = (\theta, \kappa, \mu)^\top$ *on the set* $\mathbb{R}^3$ *based on the observations* $(Y_t, S_t)_{t \in [0,T]}$ *taking the form*

$$(3.10) \qquad \widehat{\boldsymbol{\psi}}_T = \begin{bmatrix} \widehat{\theta}_T \\ \widehat{\kappa}_T \\ \widehat{\mu}_T \end{bmatrix} = \begin{bmatrix} \frac{(\boldsymbol{G}_T^{-1} \boldsymbol{f}_T)_1}{(\boldsymbol{G}_T^{-1} \boldsymbol{f}_T)_2} \\ (\boldsymbol{G}_T^{-1} \boldsymbol{f}_T)_2 \\ (\boldsymbol{G}_T^{-1} \boldsymbol{f}_T)_3 \end{bmatrix},$$

provided that $\boldsymbol{G}_T$ *is strictly positive definite and* $(\boldsymbol{G}_T^{-1} \boldsymbol{f}_T)_2 \neq 0$, *which hold almost surely. Further, we have*

$$\widehat{\theta}_T = \frac{\int_0^T Y_u \, \mathrm{d}u \int_0^T \frac{\mathrm{d}Y_u}{Y_u} - T \int_0^T \mathrm{d}Y_u + \varrho\sigma T \int_0^T \frac{\mathrm{d}S_u - S_{u-} \, \mathrm{d}L_u}{S_{u-}} - \varrho\sigma T^2 \frac{\int_0^T \frac{\mathrm{d}S_u - S_{u-} \, \mathrm{d}L_u}{Y_u S_{u-}}}{\int_0^T \frac{\mathrm{d}u}{Y_u}}}{T \int_0^T \frac{\mathrm{d}Y_u}{Y_u} - \int_0^T \frac{\mathrm{d}u}{Y_u} \int_0^T \mathrm{d}Y_u + \varrho\sigma \int_0^T \frac{\mathrm{d}u}{Y_u} \int_0^T \frac{\mathrm{d}S_u - S_{u-} \, \mathrm{d}L_u}{S_{u-}} - \varrho\sigma T \int_0^T \frac{\mathrm{d}S_u - S_{u-} \, \mathrm{d}L_u}{Y_u S_{u-}}},$$

$$(3.11) \qquad \widehat{\kappa}_T = \frac{T \int_0^T \frac{\mathrm{d}Y_u}{Y_u} - \int_0^T \frac{\mathrm{d}u}{Y_u} \int_0^T \mathrm{d}Y_u + \varrho\sigma \int_0^T \frac{\mathrm{d}u}{Y_u} \int_0^T \frac{\mathrm{d}S_u - S_{u-} \, \mathrm{d}L_u}{S_{u-}} - \varrho\sigma T \int_0^T \frac{\mathrm{d}S_u - S_{u-} \, \mathrm{d}L_u}{Y_u S_{u-}}}{\int_0^T Y_u \, \mathrm{d}u \int_0^T \frac{\mathrm{d}u}{Y_u} - T^2},$$

$$\widehat{\mu}_T = \frac{\int_0^T \frac{\mathrm{d}S_u - S_{u-} \, \mathrm{d}L_u}{Y_u S_{u-}}}{\int_0^T \frac{\mathrm{d}u}{Y_u}}.$$

Proof. The function Λ_T can be written in the form

$$\Lambda_T(\boldsymbol{\psi}, \widetilde{\boldsymbol{\psi}}) = -\frac{1}{2} H(\boldsymbol{\psi})^\top \boldsymbol{G}_T H(\boldsymbol{\psi}) + H(\boldsymbol{\psi})^\top \boldsymbol{f}_T - \boldsymbol{c}, \qquad \boldsymbol{\psi} \in \mathbb{R}^3,$$

with

$$\boldsymbol{c} := H(\widetilde{\boldsymbol{\psi}})^\top \boldsymbol{f}_T - \frac{1}{2} H(\widetilde{\boldsymbol{\psi}})^\top \boldsymbol{G}_T H(\widetilde{\boldsymbol{\psi}}),$$

since the symmetry of $\Sigma(y, s)^{-1}$, $(y, s) \in (0, \infty)^2$, implies the symmetry of $\boldsymbol{G}_T$, and hence $H(\widetilde{\boldsymbol{\psi}})^\top \boldsymbol{G}_T H(\boldsymbol{\psi}) = H(\boldsymbol{\psi})^\top \boldsymbol{G}_T H(\widetilde{\boldsymbol{\psi}})$. The symmetric random matrix $\boldsymbol{G}_T$ is almost surely

20

strictly positive definite, since its $k \times k$ minors, $k \in \{1, 2, 3\}$ (see, (3.4)), are almost surely positive, namely,

$$\frac{1}{(1 - \varrho^2)\sigma^2} \int_0^T \frac{\mathrm{d}u}{Y_u} > 0, \qquad \frac{1}{((1 - \varrho^2)\sigma^2)^2} \left(\int_0^T Y_u \, \mathrm{d}u \int_0^T \frac{\mathrm{d}u}{Y_u} - T^2 \right) > 0,$$

$$\frac{1}{((1 - \varrho^2)\sigma^2)^2} \left(\int_0^T Y_u \, \mathrm{d}u \int_0^T \frac{\mathrm{d}u}{Y_u} - T^2 \right) \int_0^T \frac{\mathrm{d}u}{Y_u} > 0$$

almost surely. Indeed, $\int_0^T \frac{\mathrm{d}u}{Y_u} > 0$ a.s. follows from $\mathbb{P}(Y_t \in (0, \infty)$ for all $t \in [0, \infty)) = 1$, which can be found, e.g., in Lamberton and Lapeyre [27, Proposition 6.2.4] (see also Proposition 2.1), and the proof of $\int_0^T Y_u \, \mathrm{d}u \int_0^T \frac{\mathrm{d}u}{Y_u} > T^2$ a.s. is given, e.g., in Barczy and Pap [6, Lemma 3.3]. Thus the matrix $\boldsymbol{G}_T$ is almost surely invertible, namely,

$$\boldsymbol{G}_T^{-1} = \frac{1}{\left(\int_0^T Y_u \, \mathrm{d}u \int_0^T \frac{\mathrm{d}u}{Y_u} - T^2 \right) \int_0^T \frac{\mathrm{d}u}{Y_u}}$$

$$\times \begin{bmatrix} \sigma^2 \int_0^T Y_u \, \mathrm{d}u \int_0^T \frac{\mathrm{d}u}{Y_u} - \varrho^2\sigma^2 T^2 & (1 - \varrho^2)\sigma^2 T \int_0^T \frac{\mathrm{d}u}{Y_u} & \varrho\sigma \int_0^T Y_u \, \mathrm{d}u \int_0^T \frac{\mathrm{d}u}{Y_u} - \varrho\sigma T^2 \\ (1 - \varrho^2)\sigma^2 T \int_0^T \frac{\mathrm{d}u}{Y_u} & (1 - \varrho^2)\sigma^2 \left(\int_0^T \frac{\mathrm{d}u}{Y_u} \right)^2 & 0 \\ \varrho\sigma \int_0^T Y_u \, \mathrm{d}u \int_0^T \frac{\mathrm{d}u}{Y_u} - \varrho\sigma T^2 & 0 & \int_0^T Y_u \, \mathrm{d}u \int_0^T \frac{\mathrm{d}u}{Y_u} - T^2 \end{bmatrix}$$

whenever $\int_0^T Y_u \, \mathrm{d}u \int_0^T \frac{\mathrm{d}u}{Y_u} > T^2$ and $\int_0^T \frac{\mathrm{d}u}{Y_u} > 0$, which hold almost surely. Provided that $\boldsymbol{G}_T$ is strictly positive definite, we have

$$\Lambda_T(\boldsymbol{\psi}, \widetilde{\boldsymbol{\psi}}) = -\frac{1}{2} \left(H(\boldsymbol{\psi}) - \boldsymbol{G}_T^{-1} \boldsymbol{f}_T \right)^\top \boldsymbol{G}_T \left(H(\boldsymbol{\psi}) - \boldsymbol{G}_T^{-1} \boldsymbol{f}_T \right) + \frac{1}{2} \boldsymbol{f}_T^\top \boldsymbol{G}_T^{-1} \boldsymbol{f}_T - \boldsymbol{c} \leqslant \frac{1}{2} \boldsymbol{f}_T^\top \boldsymbol{G}_T^{-1} \boldsymbol{f}_T - \boldsymbol{c},$$

and equality holds if and only if $H(\boldsymbol{\psi}) = \boldsymbol{G}_T^{-1} \boldsymbol{f}_T$. The aim of the following discussion is to show that the inverse mapping H^{-1} given in (3.2) can be applied to $\boldsymbol{G}_T^{-1} \boldsymbol{f}_T =: ((\boldsymbol{G}_T^{-1} \boldsymbol{f}_T)_1, (\boldsymbol{G}_T^{-1} \boldsymbol{f}_T)_2, (\boldsymbol{G}_T^{-1} \boldsymbol{f}_T)_3)$ almost surely, that is, $\mathbb{P}((\boldsymbol{G}_T^{-1} \boldsymbol{f}_T)_2 = 0) = 0$. Applying (3.3), we obtain

$$\boldsymbol{f}_T = \frac{1}{(1 - \varrho^2)\sigma^2} \begin{bmatrix} \int_0^T \frac{\mathrm{d}Y_u}{Y_u} - \varrho\sigma \int_0^T \frac{\mathrm{d}S_u - S_{u-} \, \mathrm{d}L_u}{Y_u S_{u-}} \\ -\int_0^T \mathrm{d}Y_u + \varrho\sigma \int_0^T \frac{\mathrm{d}S_u - S_{u-} \, \mathrm{d}L_u}{S_{u-}} \\ -\varrho\sigma \int_0^T \frac{\mathrm{d}Y_u}{Y_u} + \sigma^2 \int_0^T \frac{\mathrm{d}S_u - S_{u-} \, \mathrm{d}L_u}{Y_u S_{u-}} \end{bmatrix}, \qquad T \in (0, \infty).$$

Using the explicit formula for $\boldsymbol{G}_T^{-1}$, we obtain

$$(\boldsymbol{G}_T^{-1} \boldsymbol{f}_T)_2 = \frac{T \int_0^T \frac{\mathrm{d}Y_u}{Y_u} - \int_0^T \frac{\mathrm{d}u}{Y_u} \int_0^T \mathrm{d}Y_u + \varrho\sigma \int_0^T \frac{\mathrm{d}u}{Y_u} \int_0^T \frac{\mathrm{d}S_u - S_{u-} \, \mathrm{d}L_u}{S_{u-}} - \varrho\sigma T \int_0^T \frac{\mathrm{d}S_u - S_{u-} \, \mathrm{d}L_u}{Y_u S_{u-}}}{\int_0^T Y_u \, \mathrm{d}u \int_0^T \frac{\mathrm{d}u}{Y_u} - T^2}.$$

By the SDE (1.3),

(3.12)
$$\int_0^T \frac{\mathrm{d}S_u - S_{u-} \, \mathrm{d}L_u}{S_{u-}} = \int_0^T \left[\mu \, \mathrm{d}u + \sqrt{Y_u} \left(\varrho \, \mathrm{d}W_u + \sqrt{1 - \varrho^2} \, \mathrm{d}B_u \right) \right],$$

$$\int_0^T \frac{\mathrm{d}S_u - S_{u-} \, \mathrm{d}L_u}{Y_u S_{u-}} = \int_0^T \frac{\mu \, \mathrm{d}u + \sqrt{Y_u} \left(\varrho \, \mathrm{d}W_u + \sqrt{1 - \varrho^2} \, \mathrm{d}B_u \right)}{Y_u}.$$

21

Applying again the SDE (1.3), we have

$$\int_0^T \sqrt{Y_u}\, dW_u = \frac{1}{\sigma}\left(Y_T - y_0 - \int_0^T \kappa(\theta - Y_u)\, du\right), \qquad \int_0^T \frac{dW_u}{\sqrt{Y_u}} = \frac{1}{\sigma}\int_0^T \frac{dY_u - \kappa(\theta - Y_u)\, du}{Y_u},$$

where

$$\int_0^T \frac{dY_u}{Y_u} = \log(Y_T) - \log(y_0) + \frac{\sigma^2}{2}\int_0^T \frac{du}{Y_u}.$$

Indeed, since $\mathbb{P}(Y_t \in (0,\infty)$ for all $t \in [0,\infty)) = 1$, one can apply Itô's rule to the function $f(x) = \log(x)$, $x \in (0,\infty)$, for which $f'(x) = 1/x$, $f''(x) = -1/x^2$, $x \in (0,\infty)$, and we obtain

$$(3.13) \qquad \log(Y_T) = \log(y_0) + \int_0^T \frac{dY_u}{Y_u} - \frac{\sigma^2}{2}\int_0^T \frac{du}{Y_u}, \qquad T \in [0,\infty),$$

see von Weizsäcker and Winkler [39, Theorem 8.1.1]. Hence, using the independence of the processes Y and B, the conditional distribution of $(\boldsymbol{G}_T^{-1}\boldsymbol{f}_T)_2$ given $(Y_u)_{u \in [0,T]}$ is Gaussian and hence absolutely continuous, implying

$$\mathbb{P}((\boldsymbol{G}_T^{-1}\boldsymbol{f}_T)_2 = 0) = \mathbb{E}(\mathbb{P}((\boldsymbol{G}_T^{-1}\boldsymbol{f}_T)_2 = 0 \,|\, (Y_u)_{u \in [0,T]})) = 0.$$

Consequently,

$$H^{-1}(\boldsymbol{G}_T^{-1}\boldsymbol{f}_T) = \arg\max_{\boldsymbol{\psi} \in \mathbb{R}^3} \Lambda_T(\boldsymbol{\psi}, \widetilde{\boldsymbol{\psi}}),$$

provided that $\boldsymbol{G}_T$ is strictly positive definite and $(\boldsymbol{G}_T^{-1}\boldsymbol{f}_T)_2 \neq 0$, which hold almost surely, hence there exists a unique MLE $\widehat{\boldsymbol{\psi}}_T = (\widehat{\theta}_T, \widehat{\kappa}_T, \widehat{\mu}_T)^\top = H^{-1}(\boldsymbol{G}_T^{-1}\boldsymbol{f}_T)$ of $\boldsymbol{\psi} = (\theta, \kappa, \mu)^\top$ on the set $\mathbb{R}^3$ based on the observations $(Y_t, S_t)_{t \in [0,T]}$ yielding (3.10). Using again the explicit formula for $\boldsymbol{G}_T^{-1}$, we obtain (3.11) as well. Note that $\widehat{\boldsymbol{\psi}}_T$ is a measurable function of the observations $(Y_t, S_t)_{t \in [0,T]}$, since all the integrals appearing in $\widehat{\boldsymbol{\psi}}_T$ are measurable functions of this process. Indeed, in Remark 2.4 we showed that for all $t \in [0,T]$, L_t is a measurable function of $(S_u)_{u \in [0,T]}$, and one can use the arguments of Remarks 2.5 and 2.6 together with Proposition I.4.44 in Jacod and Shiryaev [22], and Theorems 4.2.2 and 4.2.8 in Dudley [16]. For example, for all $T \in [0,\infty)$,

$$\sum_{i=1}^{\lfloor nT \rfloor} \frac{(S_{\frac{i}{n}} - S_{\frac{i-1}{n}}) - S_{\frac{i-1}{n}-}(L_{\frac{i}{n}} - L_{\frac{i-1}{n}})}{Y_{\frac{i-1}{n}}S_{\frac{i-1}{n}-}} \xrightarrow{\mathbb{P}} \int_0^T \frac{dS_u - S_{u-}\, dL_u}{Y_u S_{u-}} \qquad \text{as } n \to \infty.$$

$\square$

Remark 3.3 We call the attention that later on it will turn out that $\widehat{\boldsymbol{\psi}}_T$ is a weakly consistent estimator of $\boldsymbol{\psi}$ (see, Theorem 4.1 and Remark 4.2) yielding that $\mathbb{P}(\widehat{\boldsymbol{\psi}}_T \in \Psi) = \mathbb{P}(H^{-1}(\boldsymbol{G}_T^{-1}\boldsymbol{f}_T) \in \Psi) \to 1$ as $T \to \infty$ for each $\boldsymbol{\psi} \in \Psi$, and hence

$$\mathbb{P}\left(H^{-1}(\boldsymbol{G}_T^{-1}\boldsymbol{f}_T) = \arg\max_{\boldsymbol{\psi} \in \Psi} \Lambda_T(\boldsymbol{\psi}, \widetilde{\boldsymbol{\psi}})\right) \to 1 \qquad \text{as } T \to \infty.$$

Consequently, the probability that there exists a unique MLE $\boldsymbol{\psi}_T^*$ of $\boldsymbol{\psi}$ on the set Ψ based on the observations $(Y_t, S_t)_{t \in [0,T]}$ converges to 1 as $T \to \infty$, and $\mathbb{P}(\boldsymbol{\psi}_T^* = \widehat{\boldsymbol{\psi}}_T) \to 1$ as $T \to \infty$. $\square$

Remark 3.4 To make it clear, we point out that the expression for $(\widehat{\theta}_T, \widehat{\kappa}_T)$ in (3.11) is not the same as the MLE of (θ, κ) based only on the continuous time observation $(Y_t)_{t \in [0,T]}$ for the first coordinate process of (1.3), see, e.g., Overbeck [33], because our statistical setup is different. $\qquad \square$

Remark 3.5 In the literature there is another way of deriving an MLE. Sørensen [36] defined an MLE of $\boldsymbol{\psi}$ as a solution of the equation $\dot{\Lambda}_T(\boldsymbol{\psi}) = 0$, where $\dot{\Lambda}_T(\boldsymbol{\psi})$ is the so-called score vector given in formula (3.3) in Sørensen [36]. Luschgy [31], [32] called this equation as an estimating equation. With the notations of the proof of Lemma 3.1, taking into account of the form of $\beta^{(\widetilde{\psi}, \psi)}$ and the fact that $V^{(\psi, \widetilde{\psi})}$ is identically one, we have

$$\dot{\Lambda}_T(\boldsymbol{\psi}) := \int_0^T \dot{H}(\boldsymbol{\psi})^\top A(Y_u, S_u)^\top \Sigma(Y_u, S_u)^{-1} \begin{bmatrix} \mathrm{d}Y_u^{\mathrm{cont}} \\ \mathrm{d}S_u^{\mathrm{cont}} \end{bmatrix}$$

$$= \int_0^T \dot{H}(\boldsymbol{\psi})^\top A(Y_u, S_u)^\top \Sigma(Y_u, S_u)^{-1} \left(\begin{bmatrix} \mathrm{d}Y_u \\ \mathrm{d}S_u \end{bmatrix} - A(Y_u, S_u)H(\boldsymbol{\psi})\,\mathrm{d}u - \begin{bmatrix} 0 \\ S_{u-}\,\mathrm{d}L_u \end{bmatrix} \right)$$

for $\boldsymbol{\psi} \in \mathbb{R}^3$ and $T \in (0, \infty)$ with

$$\dot{H}(\boldsymbol{\psi}) = \begin{bmatrix} \kappa & \theta & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad \boldsymbol{\psi} = \begin{bmatrix} \theta \\ \kappa \\ \mu \end{bmatrix} \in \mathbb{R}^3.$$

Using (3.9) and the definitions of $\boldsymbol{f}_T$ and $\boldsymbol{G}_T$, we obtain

$$\dot{\Lambda}_T(\boldsymbol{\psi}) = \dot{H}(\boldsymbol{\psi})^\top (\boldsymbol{f}_T - \boldsymbol{G}_T H(\boldsymbol{\psi})),$$

hence the estimating equation $\dot{\Lambda}_T(\boldsymbol{\psi}) = 0$, $\boldsymbol{\psi} \in \mathbb{R}^3$, has a unique solution $H^{-1}(\boldsymbol{G}_T^{-1}\boldsymbol{f}_T)$ provided that $\boldsymbol{G}_T$ is strictly positive definite and $(\boldsymbol{G}_T^{-1}\boldsymbol{f}_T)_2 \neq 0$, which hold almost surely. Recall that this unique solution coincides with $\widehat{\boldsymbol{\psi}}_T$, see (3.10). $\qquad \square$

4 Consistency of MLE

Theorem 4.1 If $\theta, \kappa \in (0, \infty)$ with $\theta\kappa \in \left(\frac{\sigma^2}{2}, \infty \right)$, $\mu \in \mathbb{R}$, $\sigma \in (0, \infty)$, $\varrho \in (-1, 1)$, and $(Y_0, S_0) = (y_0, s_0) \in (0, \infty)^2$, then the MLE of $\boldsymbol{\psi} = (\theta, \kappa, \mu)$ is strongly consistent, i.e., $\widehat{\boldsymbol{\psi}}_T = \left(\widehat{\theta}_T, \widehat{\kappa}_T, \widehat{\mu}_T \right) \xrightarrow{\text{a.s.}} \boldsymbol{\psi} = (\theta, \kappa, \mu)$ as $T \to \infty$.

Proof. Obviously, it is enough to show that $\boldsymbol{G}_T^{-1}\boldsymbol{f}_T \xrightarrow{\text{a.s.}} H(\boldsymbol{\psi})$ as $T \to \infty$, since then $\widehat{\boldsymbol{\psi}}_T = H^{-1}(\boldsymbol{G}_T^{-1}\boldsymbol{f}_T) \xrightarrow{\text{a.s.}} H^{-1}(H(\boldsymbol{\psi})) = \boldsymbol{\psi}$ as $T \to \infty$, using the continuity of H^{-1} and that $\mathbb{P}((\boldsymbol{G}_T^{-1}\boldsymbol{f}_T)_2 = 0) = 0$, see Section 3. Using the SDE (3.1), we obtain $\boldsymbol{f}_T = \boldsymbol{G}_T H(\boldsymbol{\psi}) + \boldsymbol{h}_T$, $T \in [0, \infty)$, with

$$\boldsymbol{h}_T := \int_0^T A(Y_u, S_u)^\top \Sigma(Y_u, S_u)^{-1} \Gamma(Y_u, S_u) \begin{bmatrix} \mathrm{d}W_u \\ \mathrm{d}B_u \end{bmatrix}.$$

23

Thus

(4.1) $\qquad \boldsymbol{G}_T^{-1}\boldsymbol{f}_T - H(\boldsymbol{\psi}) = \boldsymbol{G}_T^{-1}\boldsymbol{h}_T = (\boldsymbol{D}_T\boldsymbol{G}_T)^{-1}(\boldsymbol{D}_T\boldsymbol{h}_T), \qquad T \in (0, \infty),$

with

$$\boldsymbol{D}_T := \begin{bmatrix} \frac{1}{\int_0^T \frac{du}{Y_u}} & 0 & 0 \\ 0 & \frac{1}{\int_0^T Y_u\, du} & 0 \\ 0 & 0 & \frac{1}{\int_0^T \frac{du}{Y_u}} \end{bmatrix},$$

provided that $\boldsymbol{G}_T$ is invertible and $\int_0^T \frac{du}{Y_u} > 0$ and $\int_0^T Y_u\, du > 0$, which hold almost surely, see Section 3. We have

$$\boldsymbol{D}_T\boldsymbol{G}_T = \frac{1}{(1-\varrho^2)\sigma^2} \begin{bmatrix} 1 & -\frac{1}{T^{-1}\int_0^T \frac{du}{Y_u}} & -\varrho\sigma \\ -\frac{1}{T^{-1}\int_0^T Y_u\, du} & 1 & \frac{\varrho\sigma}{T^{-1}\int_0^T Y_u\, du} \\ -\varrho\sigma & \frac{\varrho\sigma}{T^{-1}\int_0^T \frac{du}{Y_u}} & \sigma^2 \end{bmatrix}, \qquad T \in (0, \infty).$$

Using that $\Sigma(Y_u, S_u)^{-1} = (\Gamma(Y_u, S_u)^\top)^{-1}\Gamma(Y_u, S_u)^{-1}, \; u \in [0,T],$ and (3.5), we obtain

$$\boldsymbol{h}_T = \int_0^T A(Y_u, S_u)^\top (\Gamma(Y_u, S_u)^\top)^{-1} \begin{bmatrix} dW_u \\ dB_u \end{bmatrix} = \int_0^T (\Gamma(Y_u, S_u)^{-1}A(Y_u, S_u))^\top \begin{bmatrix} dW_u \\ dB_u \end{bmatrix}$$

(4.2)
$$= \frac{1}{\sigma\sqrt{1-\varrho^2}} \int_0^T \frac{1}{\sqrt{Y_u}} \begin{bmatrix} \sqrt{1-\varrho^2}\, dW_u - \varrho\, dB_u \\ -\sqrt{1-\varrho^2}Y_u\, dW_u + \varrho Y_u\, dB_u \\ \sigma\, dB_u \end{bmatrix},$$

and hence

$$\boldsymbol{D}_T\boldsymbol{h}_T = \frac{1}{\sigma\sqrt{1-\varrho^2}} \begin{bmatrix} \overline{h}_T^{(1)} \\ \overline{h}_T^{(2)} \\ \overline{h}_T^{(3)} \end{bmatrix}$$

with

$$\overline{h}_T^{(1)} := \frac{\int_0^T \frac{\sqrt{1-\varrho^2}\, dW_u - \varrho\, dB_u}{\sqrt{Y_u}}}{\int_0^T \frac{du}{Y_u}}, \quad \overline{h}_T^{(2)} := -\frac{\int_0^T \sqrt{Y_u}(\sqrt{1-\varrho^2}\, dW_u - \varrho\, dB_u)}{\int_0^T Y_u\, du}, \quad \overline{h}_T^{(3)} := \sigma\frac{\int_0^T \frac{dB_u}{\sqrt{Y_u}}}{\int_0^T \frac{du}{Y_u}}.$$

By part (i) of Theorem 2.2, $\mathbb{E}(Y_\infty) = \theta$ and $\mathbb{E}\left(\frac{1}{Y_\infty}\right) = \frac{2\kappa}{2\theta\kappa - \sigma^2}$, and hence, part (ii) of Theorem 2.2 implies

(4.3) $\quad \boldsymbol{D}_T\boldsymbol{G}_T \xrightarrow{\text{a.s.}} \frac{1}{(1-\varrho^2)\sigma^2} \begin{bmatrix} 1 & -\frac{1}{\mathbb{E}\left(\frac{1}{Y_\infty}\right)} & -\varrho\sigma \\ -\frac{1}{\mathbb{E}(Y_\infty)} & 1 & \frac{\varrho\sigma}{\mathbb{E}(Y_\infty)} \\ -\varrho\sigma & \frac{\varrho\sigma}{\mathbb{E}\left(\frac{1}{Y_\infty}\right)} & \sigma^2 \end{bmatrix} =: \frac{1}{(1-\varrho^2)\sigma^2}\boldsymbol{S} \qquad$ as $T \to \infty.$

24

We have

$$\det(\boldsymbol{S}) = (1 - \varrho^2)\sigma^2 \left(1 - \frac{1}{\mathbb{E}(Y_\infty)\,\mathbb{E}\left(\frac{1}{Y_\infty}\right)} \right) > 0$$

since $\mathbb{E}(Y_\infty)\,\mathbb{E}\left(\frac{1}{Y_\infty}\right) = \frac{2\theta\kappa}{2\theta\kappa - \sigma^2} > 1$, hence the matrix $\boldsymbol{S}$ is invertible, and we conclude

$$(4.4) \qquad (\boldsymbol{D}_T\boldsymbol{G}_T)^{-1} \xrightarrow{\text{a.s.}} (1 - \varrho^2)\sigma^2\boldsymbol{S}^{-1} \qquad \text{as } T \to \infty.$$

The aim of the following discussion is to show convergence

$$(4.5) \qquad \boldsymbol{D}_T\boldsymbol{h}_T \xrightarrow{\text{a.s.}} \boldsymbol{0} \qquad \text{as } T \to \infty.$$

The strong law of large numbers for continuous local martingales (see, e.g., Theorem B.1) implies $\overline{h}_T^{(1)} \xrightarrow{\text{a.s.}} 0$ as $T \to \infty$, since, by part (ii) of Theorem 2.2,

$$\frac{1}{T}\int_0^T \frac{\mathrm{d}s}{Y_s} \xrightarrow{\text{a.s.}} \mathbb{E}\left(\frac{1}{Y_\infty}\right) = \frac{2\kappa}{2\theta\kappa - \sigma^2} \in (0,\infty) \qquad \text{as } T \to \infty,$$

implying

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

Convergences $\overline{h}_T^{(2)} \xrightarrow{\text{a.s.}} 0$ as $T \to \infty$, and $\overline{h}_T^{(3)} \xrightarrow{\text{a.s.}} 0$ as $T \to \infty$ can be proved in the same way, since, by part (ii) of Theorem 2.2,

$$\frac{1}{T}\int_0^T Y_s\,\mathrm{d}s \xrightarrow{\text{a.s.}} \mathbb{E}(Y_\infty) = \theta \in (0,\infty) \qquad \text{as } T \to \infty,$$

implying

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

Consequently, we conclude (4.5). By (4.1), (4.4) and (4.5), we obtain $\boldsymbol{G}_T^{-1}\boldsymbol{f}_T \xrightarrow{\text{a.s.}} H(\boldsymbol{\psi})$ as $T \to \infty$, hence we conclude the statement. $\qquad\square$

Remark 4.2 For the case $\theta\kappa = \frac{\sigma^2}{2}$, Theorem 5.3 implies weak consistency of the MLE of (θ, κ, μ). $\qquad\square$

5   Asymptotic behaviour of MLE

Theorem 5.1 If $\theta, \kappa \in (0,\infty)$ with $\theta\kappa \in \left(\frac{\sigma^2}{2},\infty\right)$, $\mu \in \mathbb{R}$, $\sigma \in (0,\infty)$, $\varrho \in (-1,1)$, and $(Y_0, S_0) = (y_0, s_0) \in (0,\infty)^2$, then the MLE of $\boldsymbol{\psi} = (\theta, \kappa, \mu)$ is asymptotically normal, namely,

$$(5.1) \qquad T^{1/2}(\widehat{\boldsymbol{\psi}}_T - \boldsymbol{\psi}) \xrightarrow{\mathcal{D}} \mathcal{N}_3(\boldsymbol{0}, \boldsymbol{V}) \qquad \text{as } T \to \infty,$$

where the matrix $\boldsymbol{V}$ *is given by*

$$(5.2) \qquad \frac{1}{2\kappa^3} \begin{bmatrix} \sigma^2(2\theta\kappa - \varrho^2\sigma^2) & -2(1-\varrho^2)\sigma^2\kappa^2 & \varrho\sigma\kappa(2\theta\kappa - \sigma^2) \\ -2(1-\varrho^2)\sigma^2\kappa^2 & 4\kappa^4(1-\varrho^2) & 0 \\ \varrho\sigma\kappa(2\theta\kappa - \sigma^2) & 0 & \kappa^2(2\theta\kappa - \sigma^2) \end{bmatrix}.$$

With a random scaling, we have

$$(5.3) \qquad \boldsymbol{R}_T \boldsymbol{Q}_T (\widehat{\boldsymbol{\psi}}_T - \boldsymbol{\psi}) \xrightarrow{\mathcal{D}} \mathcal{N}_3(\boldsymbol{0}, \boldsymbol{I}_3) \qquad as \ T \to \infty,$$

where $\boldsymbol{I}_3$ *denotes the* 3×3 *identity matrix, and* $\boldsymbol{R}_T$, $T \in (0, \infty)$, *and* $\boldsymbol{Q}_T$, $T \in (0, \infty)$, *are* 3×3 *(not uniquely determined) random matrices with properties* $T^{-1/2}\boldsymbol{R}_T \xrightarrow{\mathbb{P}} \boldsymbol{C}$ *as* $T \to \infty$ *with some* $\boldsymbol{C} \in \mathbb{R}^{3 \times 3}$, $\boldsymbol{R}_T^\top \boldsymbol{R}_T = \boldsymbol{G}_T$, $T \in (0, \infty)$, *and* $\boldsymbol{Q}_T \xrightarrow{\mathbb{P}} \boldsymbol{Q}$ *as* $T \to \infty$, *where*

$$\boldsymbol{Q} := \begin{bmatrix} \kappa & \theta & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

For a possible choice of $\boldsymbol{R}_T$ *and* $\boldsymbol{Q}_T$, $T \in (0, \infty)$, *see Remark 5.2.*

Remark 5.2 Note that the limiting covariance matrix $\boldsymbol{V}$ in (5.1) depends only on the unknown parameters θ and κ, but not on (the unknown) μ. The advantage of the random scaling is that the limiting covariance matrix in (5.3) is the 3×3 identity matrix $\boldsymbol{I}_3$ which does not depend on any of the unknown parameters. Note also that for $\boldsymbol{R}_T$ and $\boldsymbol{Q}_T$ one can choose, for instance,

$$\boldsymbol{R}_T = \frac{1}{\sigma\sqrt{1-\varrho^2}\sqrt{\int_0^T \frac{\mathrm{d}u}{Y_u}}} \begin{bmatrix} \int_0^T \frac{\mathrm{d}u}{Y_u} & -T & -\varrho\sigma\int_0^T \frac{\mathrm{d}u}{Y_u} \\ 0 & \sqrt{\int_0^T Y_u\,\mathrm{d}u \int_0^T \frac{\mathrm{d}u}{Y_u} - T^2} & 0 \\ 0 & 0 & \sigma\sqrt{1-\varrho^2}\int_0^T \frac{\mathrm{d}u}{Y_u} \end{bmatrix}$$

and

$$\boldsymbol{Q}_T = \begin{bmatrix} \frac{\frac{\sigma^2}{2T}\int_0^T \frac{\mathrm{d}u}{Y_u}}{\left(\frac{1}{T}\int_0^T Y_u\,\mathrm{d}u\right)\left(\frac{1}{T}\int_0^T \frac{\mathrm{d}u}{Y_u}\right) - 1} & \frac{1}{T}\int_0^T Y_u\,\mathrm{d}u & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Indeed, we have $\boldsymbol{R}_T^\top \boldsymbol{R}_T = \boldsymbol{G}_T$, $T \in (0, \infty)$ (which is, in fact, the Cholesky factorization of $\boldsymbol{G}_T$),

$$T^{-1/2}\boldsymbol{R}_T = \frac{1}{\sigma\sqrt{1-\varrho^2}\sqrt{\frac{1}{T}\int_0^T \frac{\mathrm{d}u}{Y_u}}} \begin{bmatrix} \frac{1}{T}\int_0^T \frac{\mathrm{d}u}{Y_u} & -1 & -\frac{\varrho\sigma}{T}\int_0^T \frac{\mathrm{d}u}{Y_u} \\ 0 & \sqrt{\left(\frac{1}{T}\int_0^T Y_u\,\mathrm{d}u\right)\left(\frac{1}{T}\int_0^T \frac{\mathrm{d}u}{Y_u}\right) - 1} & 0 \\ 0 & 0 & \frac{\sigma\sqrt{1-\varrho^2}}{T}\int_0^T \frac{\mathrm{d}u}{Y_u} \end{bmatrix}$$

$$\xrightarrow{\text{a.s.}} \frac{1}{\sigma\sqrt{1-\varrho^2}\sqrt{\mathbb{E}\left(\frac{1}{Y_\infty}\right)}} \begin{bmatrix} \mathbb{E}\left(\frac{1}{Y_\infty}\right) & -1 & -\varrho\sigma\,\mathbb{E}\left(\frac{1}{Y_\infty}\right) \\ 0 & \sqrt{\mathbb{E}(Y_\infty)\,\mathbb{E}\left(\frac{1}{Y_\infty}\right) - 1} & 0 \\ 0 & 0 & \sigma\sqrt{1-\varrho^2}\,\mathbb{E}\left(\frac{1}{Y_\infty}\right) \end{bmatrix}$$

as $T \to \infty$, and $\boldsymbol{Q}_T \xrightarrow{\text{a.s.}} \boldsymbol{Q}$ as $T \to \infty$, since

$$\frac{\frac{\sigma^2}{2T} \int_0^T \frac{du}{Y_u}}{\left(\frac{1}{T} \int_0^T Y_u \, du\right)\left(\frac{1}{T} \int_0^T \frac{du}{Y_u}\right) - 1} \xrightarrow{\text{a.s.}} \frac{\frac{\sigma^2}{2} \mathbb{E}\left(\frac{1}{Y_\infty}\right)}{\mathbb{E}(Y_\infty) \mathbb{E}\left(\frac{1}{Y_\infty}\right) - 1} = \kappa, \qquad \frac{1}{T} \int_0^T Y_s \, ds \xrightarrow{\text{a.s.}} \mathbb{E}(Y_\infty) = \theta$$

as $T \to \infty$ by part (i) of Theorem 2.2. Hence then the random scaling factor has the form

$$\boldsymbol{R}_T \boldsymbol{Q}_T = \frac{1}{\sigma\sqrt{1-\varrho^2}\sqrt{\int_0^T \frac{du}{Y_u}}} \begin{bmatrix} \frac{\frac{\sigma^2 T}{2}\left(\int_0^T \frac{du}{Y_u}\right)^2}{\int_0^T Y_u \, du \int_0^T \frac{du}{Y_u} - T^2} & \frac{1}{T}\left(\int_0^T Y_u \, du \int_0^T \frac{du}{Y_u} - T^2\right) & -\varrho\sigma \int_0^T \frac{du}{Y_u} \\ & \sqrt{\int_0^T Y_u \, du \int_0^T \frac{du}{Y_u} - T^2} & \\ 0 & 0 & 0 \\ 0 & 0 & \sigma\sqrt{1-\varrho^2}\int_0^T \frac{du}{Y_u} \end{bmatrix}.$$

$\square$

Proof of Theorem 5.1. For (5.1), it is enough to prove

(5.4) $$T^{1/2}(\boldsymbol{G}_T^{-1}\boldsymbol{f}_T - H(\boldsymbol{\psi})) \xrightarrow{\mathcal{D}} \mathcal{N}_3(\boldsymbol{0}, \boldsymbol{V}_0) \qquad \text{as } T \to \infty,$$

where

$$\boldsymbol{V}_0 := \frac{2\theta\kappa - \sigma^2}{2\kappa} \begin{bmatrix} \sigma^2 + (1-\varrho^2)(2\theta\kappa - \sigma^2) & 2\kappa(1-\varrho^2) & \varrho\sigma \\ 2\kappa(1-\varrho^2) & \frac{4\kappa^2(1-\varrho^2)}{2\theta\kappa - \sigma^2} & 0 \\ \varrho\sigma & 0 & 1 \end{bmatrix}.$$

Indeed, then one can apply Lemma C.1 with $\mathcal{S}_1 = \mathcal{S}_2 = \mathbb{R}^3$, $\mathcal{C} = \mathbb{R}^3$, with a random vector $\boldsymbol{\xi}$ having distribution $\mathcal{N}_3(\boldsymbol{0}, \boldsymbol{V}_0)$, with $\xi_T = T^{1/2}(\boldsymbol{G}_T^{-1}\boldsymbol{f}_T - H(\boldsymbol{\psi}))$, $T \in (0, \infty)$, and with functions $F : \mathbb{R}^3 \to \mathbb{R}^3$ and $F_T : \mathbb{R}^3 \to \mathbb{R}^3$, $T \in (0, \infty)$, given by

$$F(\boldsymbol{x}) := \boldsymbol{Q}^{-1}\boldsymbol{x}, \qquad F_T(\boldsymbol{x}) := \begin{cases} \begin{bmatrix} \frac{x_1 - \theta x_2}{T^{-1/2}x_2 + \kappa} \\ x_2 \\ x_3 \end{bmatrix} & \text{if } x_2 \neq -T^{1/2}\kappa, \\ \boldsymbol{0} & \text{if } x_2 = -T^{1/2}\kappa, \end{cases}$$

for $\boldsymbol{x} = (x_1, x_2, x_3) \in \mathbb{R}^3$ and $T \in (0, \infty)$. We have

$$F_T(T^{1/2}(\boldsymbol{G}_T^{-1}\boldsymbol{f}_T - H(\boldsymbol{\psi}))) = \begin{bmatrix} T^{1/2}\frac{(\boldsymbol{G}_T^{-1}\boldsymbol{f}_T)_1 - \theta\kappa - \theta((\boldsymbol{G}_T^{-1}\boldsymbol{f}_T)_2 - \kappa)}{(\boldsymbol{G}_T^{-1}\boldsymbol{f}_T)_2} \\ T^{1/2}((\boldsymbol{G}_T^{-1}\boldsymbol{f}_T)_2 - \kappa) \\ T^{1/2}((\boldsymbol{G}_T^{-1}\boldsymbol{f}_T)_3 - \mu) \end{bmatrix} = T^{1/2}(\widehat{\boldsymbol{\psi}}_T - \boldsymbol{\psi}),$$

provided that $(\boldsymbol{G}_T^{-1}\boldsymbol{f}_T)_2 \neq 0$, which holds almost surely. Moreover, $F_T(\boldsymbol{x}_T) \to F(\boldsymbol{x})$ as $T \to \infty$ if $\boldsymbol{x}_T \to \boldsymbol{x}$ as $T \to \infty$, since then, for sufficiently large $T \in (0, \infty)$, we have $(\boldsymbol{x}_T)_2 \neq -T^{1/2}\kappa$. Consequently, (5.4) and Lemma C.1 imply

$$T^{1/2}(\widehat{\boldsymbol{\psi}}_T - \boldsymbol{\psi}) = F_T(T^{1/2}(\boldsymbol{G}_T^{-1}\boldsymbol{f}_T - H(\boldsymbol{\psi}))) \xrightarrow{\mathcal{D}} F(\boldsymbol{\xi}) = \boldsymbol{Q}^{-1}\boldsymbol{\xi} \stackrel{\mathcal{D}}{=} \mathcal{N}_3(\boldsymbol{0}, \boldsymbol{Q}^{-1}\boldsymbol{V}_0(\boldsymbol{Q}^{-1})^\top)$$

as $T \to \infty$, where $\boldsymbol{Q}^{-1}\boldsymbol{V}_0(\boldsymbol{Q}^{-1})^\top = \boldsymbol{V}$, hence we obtain (5.1).

By the first equality in (4.1), we have

(5.5)
$$T^{1/2}(\boldsymbol{G}_T^{-1}\boldsymbol{f}_T - H(\boldsymbol{\psi})) = T^{1/2}\boldsymbol{G}_T^{-1}\boldsymbol{h}_T = (T^{-1}\boldsymbol{G}_T)^{-1}(T^{-1/2}\boldsymbol{h}_T),$$

provided that $\boldsymbol{G}_T$ is invertible, which holds almost surely, see Section 3. By part (i) of Theorem 2.2, $\mathbb{E}(Y_\infty) = \theta \in (0, \infty)$ and $\mathbb{E}\left(\frac{1}{Y_\infty}\right) = \frac{2\kappa}{2\theta\kappa - \sigma^2} \in (0, \infty)$, and hence, part (ii) of Theorem 2.2 and (3.4) imply

(5.6)
$$T^{-1}\boldsymbol{G}_T \xrightarrow{\text{a.s.}} \mathbb{E}(\boldsymbol{G}_\infty) \qquad \text{as } T \to \infty,$$

with

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

where Y_∞ has Gamma distribution with parameters $2\theta\kappa/\sigma^2$ and $2\kappa/\sigma^2$. The matrix $\mathbb{E}(\boldsymbol{G}_\infty)$ is invertible, namely,

$$[\mathbb{E}(\boldsymbol{G}_\infty)]^{-1} = \frac{1}{\left(\mathbb{E}(Y_\infty)\,\mathbb{E}\left(\frac{1}{Y_\infty}\right) - 1\right)\mathbb{E}\left(\frac{1}{Y_\infty}\right)}$$

$$\times \begin{bmatrix} \sigma^2\,\mathbb{E}(Y_\infty)\,\mathbb{E}\left(\frac{1}{Y_\infty}\right) - \varrho^2\sigma^2 & (1-\varrho^2)\sigma^2\,\mathbb{E}\left(\frac{1}{Y_\infty}\right) & \varrho\sigma\,\mathbb{E}(Y_\infty)\,\mathbb{E}\left(\frac{1}{Y_\infty}\right) - \varrho\sigma \\ (1-\varrho^2)\sigma^2\,\mathbb{E}\left(\frac{1}{Y_\infty}\right) & (1-\varrho^2)\sigma^2\left(\mathbb{E}\left(\frac{1}{Y_\infty}\right)\right)^2 & 0 \\ \varrho\sigma\,\mathbb{E}(Y_\infty)\,\mathbb{E}\left(\frac{1}{Y_\infty}\right) - \varrho\sigma & 0 & \mathbb{E}(Y_\infty)\,\mathbb{E}\left(\frac{1}{Y_\infty}\right) - 1 \end{bmatrix},$$

since $\mathbb{E}(Y_\infty)\,\mathbb{E}\left(\frac{1}{Y_\infty}\right) = \frac{2\theta\kappa}{2\theta\kappa - \sigma^2} > 1$, which yields $[\mathbb{E}(\boldsymbol{G}_\infty)]^{-1} = \boldsymbol{V}_0$. Whence we conclude

(5.7)
$$(T^{-1}\boldsymbol{G}_T)^{-1} \xrightarrow{\text{a.s.}} \boldsymbol{V}_0 \qquad \text{as } T \to \infty.$$

By (4.2), the process $(\boldsymbol{h}_t)_{t\in[0,\infty)}$ is a 3-dimensional continuous local martingale with (predictable) quadratic variation process $\langle\boldsymbol{h}\rangle_t = \boldsymbol{G}_t$, $t \in [0,\infty)$. Using (5.6), the central limit theorem for multidimensional continuous local martingales, see Theorem B.2, yields $T^{-1/2}\boldsymbol{h}_T \xrightarrow{\mathcal{D}} \mathcal{N}_3(\boldsymbol{0}, \mathbb{E}(\boldsymbol{G}_\infty)) = \mathcal{N}_3(\boldsymbol{0}, \boldsymbol{V}_0^{-1})$ as $T \to \infty$. Hence, by (5.5) and (5.7),

$$T^{1/2}(\boldsymbol{G}_T^{-1}\boldsymbol{f}_T - H(\boldsymbol{\psi})) \xrightarrow{\mathcal{D}} \mathcal{N}_3(\boldsymbol{0}, \boldsymbol{V}_0\boldsymbol{V}_0^{-1}\boldsymbol{V}_0) = \mathcal{N}_3(\boldsymbol{0}, \boldsymbol{V}_0) \qquad \text{as } T \to \infty,$$

thus we obtain (5.4).

With random scaling, by (5.1) and Slutsky's lemma, we obtain

$$\boldsymbol{R}_T\boldsymbol{Q}_T(\widehat{\boldsymbol{\psi}}_T - \boldsymbol{\psi}) = (T^{-1/2}\boldsymbol{R}_T)\boldsymbol{Q}_T\left[T^{1/2}(\widehat{\boldsymbol{\psi}}_T - \boldsymbol{\psi})\right] \xrightarrow{\mathcal{D}} \mathcal{N}_3(\boldsymbol{0}, (\boldsymbol{CQ})\boldsymbol{V}(\boldsymbol{CQ})^\top)$$

as $T \to \infty$. Moreover, by the assumptions on $\boldsymbol{R}_T$, $T \in (0, \infty)$,

$$T^{-1}\boldsymbol{G}_T = (T^{-1/2}\boldsymbol{R}_T)^\top(T^{-1/2}\boldsymbol{R}_T) \xrightarrow{\mathbb{P}} \boldsymbol{C}^\top\boldsymbol{C} \qquad \text{as } T \to \infty.$$

28

Thus, comparing with (5.6), we obtain $C^\top C = \mathbb{E}(G_\infty) = V_0^{-1}$. Using $Q^{-1}V_0(Q^{-1})^\top = V$, we obtain

$$(CQ)V(CQ)^\top = (CQ)Q^{-1}(C^\top C)^{-1}(Q^{-1})^\top(CQ)^\top = I_3,$$

and we conclude (5.3). □

Theorem 5.3 *If* $\theta, \kappa \in (0, \infty)$ *with* $\theta\kappa = \frac{\sigma^2}{2}$, $\mu \in \mathbb{R}$, $\sigma \in (0, \infty)$, $\varrho \in (-1, 1)$, *and* $(Y_0, S_0) = (y_0, s_0) \in (0, \infty)^2$, *then*

$$(5.8) \qquad \begin{bmatrix} T^{1/2}(\widehat{\theta}_T - \theta) \\ T^{1/2}(\widehat{\kappa}_T - \kappa) \\ T(\widehat{\mu}_T - \mu) \end{bmatrix} \xrightarrow{\mathcal{D}} \begin{bmatrix} -\frac{\sigma^2\sqrt{1-\varrho^2}}{\sqrt{2\kappa^3}} Z_1 \\ \sqrt{2(1-\varrho^2)\kappa}\, Z_1 \\ \frac{\varrho\sigma}{\kappa\mathcal{T}} + \frac{\sigma\sqrt{1-\varrho^2}}{\kappa\sqrt{\mathcal{T}}} Z_2 \end{bmatrix} \qquad as \ T \to \infty,$$

where $\mathcal{T} := \inf\{t \in [0, \infty) : \mathcal{W}_t = 1\}$ *with a standard Wiener process* $(\mathcal{W}_t)_{t \in [0,\infty)}$, *and* Z_1 *and* Z_2 *are independent standard normally distributed random variables, independent from* $\mathcal{T}$. *With a random scaling, we have*

$$(5.9) \qquad \begin{bmatrix} \frac{\sigma T^2}{2\sqrt{1-\varrho^2}\left(\int_0^T Y_u\,du\right)^{3/2}}(\widehat{\theta}_T - \theta) \\ \frac{1}{\sigma\sqrt{1-\varrho^2}}\left(\int_0^T Y_u\,du\right)^{1/2}(\widehat{\kappa}_T - \kappa) \\ \frac{\sigma T^2}{2\int_0^T Y_u\,du}(\widehat{\mu}_T - \mu) \end{bmatrix} \xrightarrow{\mathcal{D}} \begin{bmatrix} -Z_1 \\ Z_1 \\ \frac{\varrho}{\mathcal{T}} + \frac{\sqrt{1-\varrho^2}}{\sqrt{\mathcal{T}}} Z_2 \end{bmatrix} \qquad as \ T \to \infty.$$

Note that the limit distribution in Theorem 5.3 (which can be considered as the asymptotic error of the estimator $(\widehat{\theta}_T, \widehat{\kappa}_T, \widehat{\mu}_T)$) is a mixed normal distribution. Moreover, the first and second coordinates of the limit distributions in (5.8) and (5.9) are linearly dependent. In spite of this fact, one can give asymptotic confidence sets for (θ, κ), namely, ellipses together with their interiors and with center $(\widehat{\theta}_T, \widehat{\kappa}_T)$. Indeed, the sum of the squares of the first two coordinates of the left-hand side of (5.9), which one can call a normalized squared error of (θ, κ), converges weakly to $2Z_1^2$, being a chi-squared distribution of degree 1 (multiplied by 2). Surprisingly, the mixed normal limit distributions of the third coordinate in (5.8) and (5.9) are not centered.

In Appendix D we derive an explicit formula for the density function of $\frac{\varrho\sigma}{\kappa\mathcal{T}} + \frac{\sigma\sqrt{1-\varrho^2}}{\kappa\sqrt{\mathcal{T}}} Z_2$, which is the limit distribution of $T(\widehat{\mu}_T - \mu)$ as $T \to \infty$ in Theorem 5.3.

Proof of Theorem 5.3. Since $\mathbb{P}(\mathcal{T} \in (0, \infty)) = 1$, the limit distributions in (5.8) and (5.9) are well defined. We have again $\mathbb{E}(Y_\infty) = \theta \in (0, \infty)$, implying

$$(5.10) \qquad \frac{1}{T}\int_0^T Y_u\,du \xrightarrow{a.s.} \theta \qquad and \qquad \int_0^T Y_u\,du \xrightarrow{a.s.} \infty \qquad as \ T \to \infty.$$

Due to Ben Alaya and Kebaier [8, Proposition 4], we have

$$(5.11) \qquad \frac{1}{T^2}\int_0^T \frac{du}{Y_u} \xrightarrow{\mathcal{D}} \mathcal{T}^* \qquad as \ T \to \infty,$$

where $\mathcal{T}^* := \inf\{t \in [0, \infty) : \mathcal{W}_t^* = \frac{\kappa}{\sigma}\}$ with a standard Wiener process $(\mathcal{W}_t^*)_{t \in [0, \infty)}$. Applying the scaling property of a standard Wiener process, we obtain

$$\mathcal{T}^* = \inf\left\{t \in [0, \infty) : \frac{\sigma}{\kappa}\mathcal{W}_t^* = 1\right\} \stackrel{\mathcal{D}}{=} \inf\left\{t \in [0, \infty) : \mathcal{W}_{\frac{\sigma^2}{\kappa^2}t} = 1\right\}$$

$$= \inf\left\{\frac{\kappa^2}{\sigma^2}s \in [0, \infty) : \mathcal{W}_s = 1\right\} = \frac{\kappa^2}{\sigma^2}\inf\left\{s \in [0, \infty) : \mathcal{W}_s = 1\right\} = \frac{\kappa^2}{\sigma^2}\mathcal{T},$$

where $\stackrel{\mathcal{D}}{=}$ denotes equality in distribution. We may and do suppose that $\mathcal{T}^* = \frac{\kappa^2}{\sigma^2}\mathcal{T}$. Using $\mathbb{P}(\mathcal{T}^* \in (0, \infty)) = 1$, we conclude

$$\frac{1}{\frac{1}{T}\int_0^T \frac{du}{Y_u}} = \frac{1}{T}\frac{1}{\frac{1}{T^2}\int_0^T \frac{du}{Y_u}} \xrightarrow{\mathcal{D}} 0 \cdot \frac{1}{\mathcal{T}^*} = 0 \qquad \text{as } T \to \infty,$$

and hence,

(5.12) $$\frac{1}{\frac{1}{T}\int_0^T \frac{du}{Y_u}} \xrightarrow{\mathbb{P}} 0 \qquad \text{as } T \to \infty,$$

implying also

$$\frac{1}{\int_0^T \frac{du}{Y_u}} = \frac{1}{T}\frac{1}{\frac{1}{T}\int_0^T \frac{du}{Y_u}} \xrightarrow{\mathbb{P}} 0 \qquad \text{as } T \to \infty.$$

Since the function $(0, \infty) \ni T \mapsto \frac{1}{\int_0^T \frac{du}{Y_u}}$ is monotone decreasing, we obtain

$$\frac{1}{\int_0^T \frac{du}{Y_u}} \xrightarrow{\text{a.s.}} 0 \qquad \text{and} \qquad \int_0^T \frac{du}{Y_u} \xrightarrow{\text{a.s.}} \infty \qquad \text{as } T \to \infty.$$

For (5.8), it is enough to prove

(5.13) $$\boldsymbol{C}_T(\boldsymbol{G}_T^{-1}\boldsymbol{f}_T - H(\boldsymbol{\psi})) \xrightarrow{\mathcal{D}} \begin{bmatrix} \frac{\sigma^2}{\kappa\mathcal{T}} \\ Z_1\sqrt{2(1-\varrho^2)\kappa} \\ \frac{\varrho\sigma}{\kappa\mathcal{T}} + \frac{Z_2\sigma\sqrt{1-\varrho^2}}{\kappa\sqrt{\mathcal{T}}} \end{bmatrix} \qquad \text{as } T \to \infty,$$

where

$$\boldsymbol{C}_T := \begin{bmatrix} T & 0 & 0 \\ 0 & T^{1/2} & 0 \\ 0 & 0 & T \end{bmatrix}.$$

Indeed, then one can apply Lemma C.1 with $\mathcal{S}_1 = \mathcal{S}_2 = \mathbb{R}^3$, $\mathcal{C} = \mathbb{R}^3$, with

$$\xi = \begin{bmatrix} \frac{\sigma^2}{\kappa\mathcal{T}} \\ Z_1\sqrt{2(1-\varrho^2)\kappa} \\ \frac{\varrho\sigma}{\kappa\mathcal{T}} + \frac{Z_2\sigma\sqrt{1-\varrho^2}}{\kappa\sqrt{\mathcal{T}}} \end{bmatrix},$$

30

with $\xi_T = \boldsymbol{C}_T(\boldsymbol{G}_T^{-1}\boldsymbol{f}_T - H(\boldsymbol{\psi}))$, $T \in (0, \infty)$, and with functions $F : \mathbb{R}^3 \to \mathbb{R}^3$ and $F_T : \mathbb{R}^3 \to \mathbb{R}^3$, $T \in (0, \infty)$, given by

$$F(\boldsymbol{x}) := \boldsymbol{B}\boldsymbol{x}, \qquad F_T(\boldsymbol{x}) := \begin{cases} \begin{bmatrix} \frac{T^{-1/2}x_1 - \theta x_2}{T^{-1/2}x_2 + \kappa} \\ x_2 \\ x_3 \end{bmatrix} & \text{if } x_2 \neq -T^{1/2}\kappa, \\ \boldsymbol{0} & \text{if } x_2 = -T^{1/2}\kappa, \end{cases}$$

for $\boldsymbol{x} = (x_1, x_2, x_3) \in \mathbb{R}^3$ and $T \in (0, \infty)$, where

$$\boldsymbol{B} := \begin{bmatrix} 0 & -\frac{\sigma^2}{2\kappa^2} & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

We have

$$F_T(\boldsymbol{C}_T(\boldsymbol{G}_T^{-1}\boldsymbol{f}_T - H(\boldsymbol{\psi}))) = \begin{bmatrix} \frac{T^{-1/2}T((\boldsymbol{G}_T^{-1}\boldsymbol{f}_T)_1 - \theta\kappa) - \theta T^{1/2}((\boldsymbol{G}_T^{-1}\boldsymbol{f}_T)_2 - \kappa)}{(\boldsymbol{G}_T^{-1}\boldsymbol{f}_T)_2} \\ T^{1/2}((\boldsymbol{G}_T^{-1}\boldsymbol{f}_T)_2 - \kappa) \\ T((\boldsymbol{G}_T^{-1}\boldsymbol{f}_T)_3 - \mu) \end{bmatrix} = \widetilde{\boldsymbol{C}}_T(\widehat{\boldsymbol{\psi}}_T - \boldsymbol{\psi}),$$

provided that $(\boldsymbol{G}_T^{-1}\boldsymbol{f}_T)_2 \neq 0$, which holds almost surely, where

$$\widetilde{\boldsymbol{C}}_T := \begin{bmatrix} T^{1/2} & 0 & 0 \\ 0 & T^{1/2} & 0 \\ 0 & 0 & T \end{bmatrix}.$$

Moreover, $F_T(\boldsymbol{x}_T) \to F(\boldsymbol{x})$ as $T \to \infty$ if $\boldsymbol{x}_T \to \boldsymbol{x}$ as $T \to \infty$, since then, for sufficiently large $T \in (0, \infty)$, we have $(\boldsymbol{x}_T)_2 \neq -T^{1/2}\kappa$. Consequently, (5.13) and Lemma C.1 imply

$$\widetilde{\boldsymbol{C}}_T(\widehat{\boldsymbol{\psi}}_T - \boldsymbol{\psi}) = F_T(\boldsymbol{C}_T(\boldsymbol{G}_T^{-1}\boldsymbol{f}_T - H(\boldsymbol{\psi}))) \xrightarrow{\mathcal{D}} F(\boldsymbol{\xi}) = \boldsymbol{B}\boldsymbol{\xi} \qquad \text{as } T \to \infty,$$

hence we obtain (5.8).

Now we turn to prove (5.13). By the first equality in (4.1), we have

(5.14)
$$\boldsymbol{C}_T(\boldsymbol{G}_T^{-1}\boldsymbol{f}_T - H(\boldsymbol{\psi})) = \boldsymbol{C}_T\boldsymbol{G}_T^{-1}\boldsymbol{h}_T = (\boldsymbol{C}_T\boldsymbol{G}_T^{-1}\boldsymbol{C}_T)(\boldsymbol{C}_T^{-1}\boldsymbol{h}_T)$$
$$= (\boldsymbol{C}_T^{-1}\boldsymbol{G}_T\boldsymbol{C}_T^{-1})^{-1}(\boldsymbol{C}_T^{-1}\boldsymbol{h}_T),$$

provided that $\boldsymbol{G}_T$ is invertible, which holds almost surely, see Section 3. We have

$$\boldsymbol{C}_T^{-1}\boldsymbol{G}_T\boldsymbol{C}_T^{-1} = \frac{1}{(1 - \varrho^2)\sigma^2} \begin{bmatrix} T^{-2}\int_0^T \frac{du}{Y_u} & -T^{-1/2} & -\varrho\sigma T^{-2}\int_0^T \frac{du}{Y_u} \\ -T^{-1/2} & T^{-1}\int_0^T Y_u \, du & \varrho\sigma T^{-1/2} \\ -\varrho\sigma T^{-2}\int_0^T \frac{du}{Y_u} & \varrho\sigma T^{-1/2} & \sigma^2 T^{-2}\int_0^T \frac{du}{Y_u} \end{bmatrix},$$

31

and, by (4.2),

$$\boldsymbol{C}_T^{-1}\boldsymbol{h}_T = \frac{1}{\sigma\sqrt{1-\varrho^2}}\begin{bmatrix}\widetilde{h}_T^{(1)}\\\widetilde{h}_T^{(2)}\\\widetilde{h}_T^{(3)}\end{bmatrix}$$

with

$$\widetilde{h}_T^{(1)} := \frac{1}{T}\int_0^T \frac{\sqrt{1-\varrho^2}\,\mathrm{d}W_u - \varrho\,\mathrm{d}B_u}{\sqrt{Y_u}}, \qquad \widetilde{h}_T^{(2)} := -\frac{1}{\sqrt{T}}\int_0^T \sqrt{Y_u}(\sqrt{1-\varrho^2}\,\mathrm{d}W_u - \varrho\,\mathrm{d}B_u),$$

$$\widetilde{h}_T^{(3)} := \frac{\sigma}{T}\int_0^T \frac{\mathrm{d}B_u}{\sqrt{Y_u}}.$$

By (5.10), $\frac{1}{T}\int_0^T Y_u\,\mathrm{d}u \xrightarrow{\text{a.s.}} \theta$ as $T \to \infty$. Ben Alaya and Kebaier [9, proof of Theorem 7] proved $\frac{\log(Y_T)}{T} \xrightarrow{\mathbb{P}} 0$ as $T \to \infty$. Using the SDE (3.1), (3.13) and $\theta\kappa = \frac{\sigma^2}{2}$,

$$(5.15) \qquad \int_0^T \frac{\sigma\,\mathrm{d}W_u}{\sqrt{Y_u}} = \int_0^T \frac{\mathrm{d}Y_u}{Y_u} - \theta\kappa\int_0^T \frac{\mathrm{d}u}{Y_u} + \kappa T = \log(Y_T) - \log(y_0) + \kappa T,$$

thus

$$(5.16) \qquad \frac{1}{T}\int_0^T \frac{\mathrm{d}W_u}{\sqrt{Y_u}} \xrightarrow{\mathbb{P}} \frac{\kappa}{\sigma} \qquad \text{as } T \to \infty.$$

Consequently, (5.13) will follow from

$$(5.17) \qquad (\boldsymbol{C}_T^{-1}\boldsymbol{G}_T\boldsymbol{C}_T^{-1}, \boldsymbol{C}_T^{-1}\boldsymbol{h}_T) \xrightarrow{\mathcal{D}} (\widetilde{\boldsymbol{G}}_\infty, \widetilde{\boldsymbol{h}}_\infty) \qquad \text{as } T \to \infty$$

with

$$\widetilde{\boldsymbol{G}}_\infty := \frac{1}{(1-\varrho^2)\sigma^2}\begin{bmatrix}\mathcal{T}^* & 0 & -\varrho\sigma\mathcal{T}^*\\ 0 & \theta & 0\\ -\varrho\sigma\mathcal{T}^* & 0 & \sigma^2\mathcal{T}^*\end{bmatrix}, \quad \widetilde{\boldsymbol{h}}_\infty := \frac{1}{\sigma\sqrt{1-\varrho^2}}\begin{bmatrix}\frac{\kappa\sqrt{1-\varrho^2}}{\sigma} - \varrho Z_2\sqrt{\mathcal{T}^*}\\ -Z_3\sqrt{\theta(1-\varrho^2)} + \varrho Z_4\sqrt{\theta}\\ \sigma Z_2\sqrt{\mathcal{T}^*}\end{bmatrix},$$

where Z_3 and Z_4 are independent standard normally distributed random variables, independent from Z_2 and $\mathcal{T}^*$. Indeed, provided that (5.17) holds, by the continuous mapping theorem,

$$(\boldsymbol{C}_T\boldsymbol{G}_T^{-1}\boldsymbol{C}_T, \boldsymbol{C}_T^{-1}\boldsymbol{h}_T) \xrightarrow{\mathcal{D}} (\widetilde{\boldsymbol{G}}_\infty^{-1}, \widetilde{\boldsymbol{h}}_\infty) \qquad \text{as } T \to \infty,$$

since $\widetilde{\boldsymbol{G}}_\infty$ is invertible almost surely with inverse

$$\widetilde{\boldsymbol{G}}_\infty^{-1} = \frac{1}{\mathcal{T}^*}\begin{bmatrix}\sigma^2 & 0 & \varrho\sigma\\ 0 & \frac{1}{\theta}(1-\varrho^2)\sigma^2\mathcal{T}^* & 0\\ \varrho\sigma & 0 & 1\end{bmatrix},$$

and hence, by (5.14) and the continuous mapping theorem,

$$\boldsymbol{C}_T(\boldsymbol{G}_T^{-1}\boldsymbol{f}_T - H(\boldsymbol{\psi})) \xrightarrow{\mathcal{D}} \widetilde{\boldsymbol{G}}_\infty^{-1}\widetilde{\boldsymbol{h}}_\infty \qquad \text{as } T \to \infty,$$

32

where, with $Z_1 := -\sqrt{1-\varrho^2}\,Z_3 + \varrho Z_4$,

$$
\widetilde{\boldsymbol{G}}_\infty^{-1}\widetilde{\boldsymbol{h}}_\infty = \frac{1}{\sigma\mathcal{T}^*\sqrt{1-\varrho^2}}
\begin{bmatrix}
\sigma^2\left(\frac{\kappa\sqrt{1-\varrho^2}}{\sigma} - \varrho Z_2\sqrt{\mathcal{T}^*}\right) + \varrho\sigma^2 Z_2\sqrt{\mathcal{T}^*} \\
\frac{1}{\sqrt{\theta}}(1-\varrho^2)\sigma^2\mathcal{T}^*(-Z_3\sqrt{1-\varrho^2} + \varrho Z_4) \\
\varrho\sigma\left(\frac{\kappa\sqrt{1-\varrho^2}}{\sigma} - \varrho Z_2\sqrt{\mathcal{T}^*}\right) + \sigma Z_2\sqrt{\mathcal{T}^*}
\end{bmatrix}
$$

$$
= \frac{1}{\sigma\mathcal{T}^*\sqrt{1-\varrho^2}}
\begin{bmatrix}
\sigma\kappa\sqrt{1-\varrho^2} \\
\frac{1}{\sqrt{\theta}}(1-\varrho^2)\sigma^2\mathcal{T}^* Z_1 \\
\varrho\kappa\sqrt{1-\varrho^2} + \sigma(1-\varrho^2)Z_2\sqrt{\mathcal{T}^*}
\end{bmatrix}
= \begin{bmatrix}
\frac{\kappa}{\mathcal{T}^*} \\
Z_1\sqrt{2(1-\varrho^2)\kappa} \\
\frac{\varrho\kappa}{\sigma\mathcal{T}^*} + \frac{Z_2\sqrt{1-\varrho^2}}{\sqrt{\mathcal{T}^*}}
\end{bmatrix},
$$

thus we obtain (5.13) using that $\mathcal{T}^* = \frac{\kappa^2}{\sigma^2}\mathcal{T}$.

Now we turn to prove (5.17). It will follow from Slutsky's lemma, continuous mapping theorem and from

$$
\left(\frac{1}{\sqrt{T}}\int_0^T \sqrt{Y_u}\,\mathrm{d}W_u, \frac{1}{T}\int_0^T \frac{\mathrm{d}W_u}{\sqrt{Y_u}}, \frac{1}{\sqrt{T}}\int_0^T \sqrt{Y_u}\,\mathrm{d}B_u, \frac{1}{T}\int_0^T \frac{\mathrm{d}B_u}{\sqrt{Y_u}}, \frac{1}{T}\int_0^T Y_u\,\mathrm{d}u, \frac{1}{T^2}\int_0^T \frac{\mathrm{d}u}{Y_u}\right)
$$

$$
\overset{\mathcal{D}}{\longrightarrow} \left(\sqrt{\theta}\,Z_3, \frac{\kappa}{\sigma}, \sqrt{\theta}Z_4, Z_2\sqrt{\mathcal{T}^*}, \theta, \mathcal{T}^*\right) \qquad \text{as } T \to \infty,
$$

which will be a consequence of (5.10), (5.11), (5.16), Slutsky's lemma (or part (v) of Theorem 2.7 in van der Vaart [38]), and

(5.18)
$$
\left(\frac{1}{\sqrt{T}}\int_0^T \sqrt{Y_s}\,\mathrm{d}W_s, \frac{1}{\sqrt{T}}\int_0^T \sqrt{Y_s}\,\mathrm{d}B_s, \frac{1}{T}\int_0^T \frac{\mathrm{d}B_s}{\sqrt{Y_s}}, \frac{1}{T^2}\int_0^T \frac{\mathrm{d}s}{Y_s}\right)
$$
$$
\overset{\mathcal{D}}{\longrightarrow} \left(\sqrt{\theta}\,Z_3, \sqrt{\theta}\,Z_4, Z_2\sqrt{\mathcal{T}^*}, \mathcal{T}^*\right) \qquad \text{as } T \to \infty.
$$

Using the SDE (1.3),

(5.19)
$$
\sigma\int_0^T \sqrt{Y_s}\,\mathrm{d}W_s = Y_T - y_0 - \theta\kappa T + \kappa\int_0^T Y_s\,\mathrm{d}s,
$$

consequently, $\int_0^T \sqrt{Y_s}\,\mathrm{d}W_s$ is measurable with respect to the σ-algebra $\sigma((Y_s)_{s\in[0,T]})$. For all $(u_1, u_2, u_3, v) \in \mathbb{R}^4$ and $T \in (0,\infty)$, we have

$$
\mathbb{E}\left(\exp\left\{\frac{iu_1}{\sqrt{T}}\int_0^T \sqrt{Y_s}\,\mathrm{d}W_s + \frac{iu_2}{\sqrt{T}}\int_0^T \sqrt{Y_s}\,\mathrm{d}B_s + \frac{iu_3}{T}\int_0^T \frac{\mathrm{d}B_s}{\sqrt{Y_s}} + \frac{iv}{T^2}\int_0^T \frac{\mathrm{d}s}{Y_s}\right\}\right)
$$

$$
= \mathbb{E}\left[\mathbb{E}\left(\exp\left\{\frac{iu_1}{\sqrt{T}}\int_0^T \sqrt{Y_s}\,\mathrm{d}W_s + \frac{iu_2}{\sqrt{T}}\int_0^T \sqrt{Y_s}\,\mathrm{d}B_s + \frac{iu_3}{T}\int_0^T \frac{\mathrm{d}B_s}{\sqrt{Y_s}} + \frac{iv}{T^2}\int_0^T \frac{\mathrm{d}s}{Y_s}\right\}\,\middle|\,(Y_s)_{s\in[0,T]}\right)\right]
$$

$$= \mathbb{E}\left[\exp\left\{\frac{iu_1}{\sqrt{T}}\int_0^T \sqrt{Y_s}\, dW_s + \frac{iv}{T^2}\int_0^T \frac{ds}{Y_s}\right\}\right.$$

$$\left. \times \mathbb{E}\left(\exp\left\{i\int_0^T\left(\frac{u_2\sqrt{Y_s}}{\sqrt{T}} + \frac{u_3}{T\sqrt{Y_s}}\right)dB_s\right\}\middle|(Y_s)_{s\in[0,T]}\right)\right]$$

$$= \mathbb{E}\left(\exp\left\{\frac{iu_1}{\sqrt{T}}\int_0^T \sqrt{Y_s}\, dW_s + \frac{iv}{T^2}\int_0^T \frac{ds}{Y_s}\right\}\exp\left\{-\frac{1}{2}\int_0^T\left(\frac{u_2^2 Y_s}{T} + \frac{u_3^2}{T^2 Y_s} + \frac{2u_2 u_3}{T^{3/2}}\right)ds\right\}\right)$$

$$= \exp\left\{-\frac{u_2 u_3}{\sqrt{T}}\right\}\mathbb{E}\left(\exp\left\{\frac{iu_1}{\sqrt{T}}\int_0^T \sqrt{Y_s}\, dW_s - \frac{u_2^2}{2T}\int_0^T Y_s\, ds - \frac{u_3^2}{2T^2}\int_0^T \frac{ds}{Y_s} + \frac{iv}{T^2}\int_0^T \frac{ds}{Y_s}\right\}\right),$$

where we used the independence of Y and B yielding that the conditional distribution of $\int_0^T\left(\frac{u_2\sqrt{Y_s}}{\sqrt{T}} + \frac{u_3}{T\sqrt{Y_s}}\right)dB_s$ given $(Y_s)_{s\in[0,T]}$ is normal. We may and do suppose that Z_2 and Z_4 are independent also from $(Y_s)_{s\in[0,T]}$. Then, in a similar way, for all $(u_1, u_2, u_3, v) \in \mathbb{R}^4$ and $T \in (0, \infty)$, we have

$$\mathbb{E}\left(\exp\left\{\frac{iu_1}{\sqrt{T}}\int_0^T \sqrt{Y_s}\, dW_s + \frac{iu_2 Z_4}{\sqrt{T}}\left(\int_0^T Y_s\, ds\right)^{1/2} + \frac{iu_3 Z_2}{T}\left(\int_0^T \frac{ds}{Y_s}\right)^{1/2} + \frac{iv}{T^2}\int_0^T \frac{ds}{Y_s}\right\}\right)$$

$$= \mathbb{E}\left[\mathbb{E}\left(\exp\left\{\frac{iu_1}{\sqrt{T}}\int_0^T \sqrt{Y_s}\, dW_s + \frac{iu_2 Z_4}{\sqrt{T}}\left(\int_0^T Y_s\, ds\right)^{1/2} + \frac{iu_3 Z_2}{T}\left(\int_0^T \frac{ds}{Y_s}\right)^{1/2}\right.\right.\right.$$

$$\left.\left.\left. + \frac{iv}{T^2}\int_0^T \frac{ds}{Y_s}\right\}\middle|(Y_s)_{s\in[0,T]}\right)\right]$$

$$= \mathbb{E}\left[\exp\left\{\frac{iu_1}{\sqrt{T}}\int_0^T \sqrt{Y_s}\, dW_s + \frac{iv}{T^2}\int_0^T \frac{ds}{Y_s}\right\}\right.$$

$$\left. \times \mathbb{E}\left(\exp\left\{\frac{iu_2 Z_4}{\sqrt{T}}\left(\int_0^T Y_s\, ds\right)^{1/2} + \frac{iu_3 Z_2}{T}\left(\int_0^T \frac{ds}{Y_s}\right)^{1/2}\right\}\middle|(Y_s)_{s\in[0,T]}\right)\right]$$

$$= \mathbb{E}\left(\exp\left\{\frac{iu_1}{\sqrt{T}}\int_0^T \sqrt{Y_s}\, dW_s + \frac{iv}{T^2}\int_0^T \frac{ds}{Y_s}\right\}\exp\left\{-\frac{u_2^2}{2T}\int_0^T Y_s\, ds - \frac{u_3^2}{2T^2}\int_0^T \frac{ds}{Y_s}\right\}\right),$$

which is the same as the previous expectation except the factor $\exp\{-\frac{u_2 u_3}{\sqrt{T}}\}$. Ben Alaya and Kebaier [9, proof of Theorem 7] proved

$$\left(\frac{Y_T - \theta\kappa T + \kappa\int_0^T Y_s\, ds}{\sqrt{T}}, \frac{1}{T^2}\int_0^T \frac{ds}{Y_s}\right) \xrightarrow{\mathcal{D}} \left(\frac{\sigma^2}{\sqrt{2\kappa}}Z_3, \mathcal{T}^*\right) \qquad \text{as } T \to \infty,$$

hence, by (5.19),

$$\left(\frac{1}{\sqrt{T}}\int_0^T \sqrt{Y_s}\, dW_s, \frac{1}{T^2}\int_0^T \frac{ds}{Y_s}\right) \xrightarrow{\mathcal{D}} \left(\sqrt{\theta}\, Z_3, \mathcal{T}^*\right) \qquad \text{as } T \to \infty.$$

By Slutsky's lemma, we obtain

$$(5.20) \qquad \left(\frac{1}{\sqrt{T}} \int_0^T \sqrt{Y_s}\, dW_s, Z_2, Z_4, \frac{1}{T} \int_0^T Y_s\, ds, \frac{1}{T^2} \int_0^T \frac{ds}{Y_s} \right) \xrightarrow{\mathcal{D}} \left(\sqrt{\theta}\, Z_3, Z_2, Z_4, \theta, \mathcal{T}^* \right)$$

as $T \to \infty$. Using the continuity theorem, we obtain

$$\mathbb{E}\left(\exp\left\{ \frac{iu_1}{\sqrt{T}} \int_0^T \sqrt{Y_s}\, dW_s + \frac{iu_2}{\sqrt{T}} \int_0^T \sqrt{Y_s}\, dB_s + \frac{iu_3}{T} \int_0^T \frac{dB_s}{\sqrt{Y_s}} + \frac{iv}{T^2} \int_0^T \frac{ds}{Y_s} \right\} \right)$$

$$= \exp\left\{ -\frac{u_2 u_3}{\sqrt{T}} \right\} \mathbb{E}\left(\exp\left\{ \frac{iu_1}{\sqrt{T}} \int_0^T \sqrt{Y_s}\, dW_s - \frac{u_2^2}{2T} \int_0^T Y_s\, ds - \frac{u_3^2}{2T^2} \int_0^T \frac{ds}{Y_s} + \frac{iv}{T^2} \int_0^T \frac{ds}{Y_s} \right\} \right)$$

$$= \exp\left\{ -\frac{u_2 u_3}{\sqrt{T}} \right\}$$

$$\times \mathbb{E}\left(\exp\left\{ \frac{iu_1}{\sqrt{T}} \int_0^T \sqrt{Y_s}\, dW_s + \frac{iu_2 Z_4}{\sqrt{T}} \left(\int_0^T Y_s\, ds \right)^{1/2} + \frac{iu_3 Z_2}{T} \left(\int_0^T \frac{ds}{Y_s} \right)^{1/2} + \frac{iv}{T^2} \int_0^T \frac{ds}{Y_s} \right\} \right)$$

$$\to \mathbb{E}\left(\exp\left\{ iu_1 \sqrt{\theta}\, Z_3 + iu_2 \sqrt{\theta}\, Z_4 + iu_3 Z_2 \sqrt{\mathcal{T}^*} + iv\mathcal{T}^* \right\} \right) \qquad \text{as } T \to \infty$$

for all $(u_1, u_2, u_3, v) \in \mathbb{R}^4$. By the continuity theorem, we obtain (5.18).

With a random scaling, we have

$$\begin{bmatrix} \frac{\sigma T^2}{2\sqrt{1-\varrho^2}\, (\int_0^T Y_u\, du)^{3/2}} (\widehat{\theta}_T - \theta) \\[2mm] \frac{1}{\sigma\sqrt{1-\varrho^2}} \left(\int_0^T Y_u\, du \right)^{1/2} (\widehat{\kappa}_T - \kappa) \\[2mm] \frac{\sigma T^2}{2 \int_0^T Y_u\, du} (\widehat{\mu}_T - \mu) \end{bmatrix} = \widetilde{\widetilde{C}}_T(\widehat{\psi}_T - \psi), \qquad T \in (0, \infty),$$

with

$$\widetilde{\widetilde{C}}_T := \begin{bmatrix} \frac{\sigma T^2}{2\sqrt{1-\varrho^2}\,(\int_0^T Y_u\, du)^{3/2}} & 0 & 0 \\[2mm] 0 & \frac{1}{\sigma\sqrt{1-\varrho^2}} \left(\int_0^T Y_u\, du \right)^{1/2} & 0 \\[2mm] 0 & 0 & \frac{\sigma T^2}{2 \int_0^T Y_u\, du} \end{bmatrix}.$$

We have

$$\widetilde{\widetilde{C}}_T(\widehat{\psi}_T - \psi) = (\widetilde{\widetilde{C}}_T \widetilde{C}_T^{-1}) \widetilde{C}_T(\widehat{\psi}_T - \psi),$$

where

$$\widetilde{\widetilde{C}}_T \widetilde{C}_T^{-1} = \begin{bmatrix} \frac{\sigma}{2\sqrt{1-\varrho^2}\,(\frac{1}{T}\int_0^T Y_u\, du)^{3/2}} & 0 & 0 \\[2mm] 0 & \frac{1}{\sigma\sqrt{1-\varrho^2}} \left(\frac{1}{T}\int_0^T Y_u\, du \right)^{1/2} & 0 \\[2mm] 0 & 0 & \frac{\sigma}{\frac{2}{T} \int_0^T Y_u\, du} \end{bmatrix}.$$

Applying (5.10), (5.11), (5.16), (5.18), Slutsky's lemma (or part (v) of Theorem 2.7 in van der Vaart [38]) and the continuous mapping theorem, we obtain

$$
(\widetilde{\widehat{C}}_T \widetilde{C}_T^{-1}, \widetilde{C}_T(\widehat{\psi}_T - \psi)) \xrightarrow{\mathcal{D}} \left(\begin{bmatrix} \frac{\sqrt{2\kappa^3}}{\sigma^2\sqrt{1-\varrho^2}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{2\kappa(1-\varrho^2)}} & 0 \\ 0 & 0 & \frac{\kappa}{\sigma} \end{bmatrix}, \begin{bmatrix} -\frac{\sigma^2 Z_1}{\sqrt{2\kappa^3}}\sqrt{1-\varrho^2} \\ Z_1\sqrt{2(1-\varrho^2)\kappa} \\ \frac{\varrho\sigma}{\kappa\mathcal{T}} + \frac{Z_2\sigma\sqrt{1-\varrho^2}}{\kappa\sqrt{\mathcal{T}}} \end{bmatrix} \right) \qquad \text{as } T \to \infty.
$$

Using again the continuous mapping theorem, we obtain (5.9). □

Remark 5.4 Putting formally $\theta\kappa = \frac{\sigma^2}{2}$ into the formula of V given in (5.2) of Theorem 5.1, one can observe that the joint limit distribution of the first two coordinates in (5.1) of Theorem 5.1 and in (5.8) of Theorem 5.3 coincide. □

Remark 5.5 According to Theorem 7 in Ben Alaya and Kebaier [9], if $a = \frac{\sigma_1^2}{2}$ and $b \in (0, \infty)$, then, based on continuous time observations $(Y_t)_{t \in [0,T]}$, $T \in (0, \infty)$, for the MLE $(\widehat{a}_T, \widehat{b}_T)$ of (a, b) for the first coordinate process of the SDE (1.1), we have

$$
\text{(5.21)} \qquad \begin{bmatrix} T(\widehat{a}_T - a) \\ T^{1/2}(\widehat{b}_T - b) \end{bmatrix} \xrightarrow{\mathcal{D}} \begin{bmatrix} \frac{\sigma_1^2}{b\mathcal{T}} \\ \sqrt{2b}Z_1 \end{bmatrix} \qquad \text{as } T \to \infty,
$$

where Z_1 is a standard normally distributed random variable independent of $\mathcal{T}$ introduced in Theorem 5.3. Hence, using Slutsky's lemma and that $\widehat{b}_T$ converges in probability to b as $T \to \infty$ (following from (5.21)), we get

$$
T^{1/2}\left(\frac{\widehat{a}_T}{\widehat{b}_T} - \frac{a}{b}\right) = T^{1/2}\frac{b\widehat{a}_T - a\widehat{b}_T}{b\widehat{b}_T} = T^{1/2}\frac{b(\widehat{a}_T - a) - a(\widehat{b}_T - b)}{b\widehat{b}_T}
$$

$$
= \frac{T^{-1/2}bT(\widehat{a}_T - a) - aT^{1/2}(\widehat{b}_T - b)}{b\widehat{b}_T} \xrightarrow{\mathcal{D}} -\frac{a}{b^2}\sqrt{2b}Z_1 = -\frac{\sqrt{2}a}{\sqrt{b^3}}Z_1
$$

as $T \to \infty$. Let us observe that in the special case of $\varrho = 0$, we have $\frac{\widehat{a}_T}{\widehat{b}_T} = \widehat{\theta}_T$ and $\widehat{b}_T = \kappa_T$, $T > 0$ (for the explicit formulae for $\widehat{a}_T$ and $\widehat{b}_T$, see Ben Alaya and Kebaier [9, Section 3.1]). Moreover, in case of $a = \theta\kappa = \frac{\sigma^2}{2}$ and $b = \kappa$, we have $-\frac{\sqrt{2}a}{\sqrt{b^3}} = -\frac{\sigma^2}{\sqrt{2\kappa^3}}$. Hence, under the conditions of Theorem 5.3 together with $\varrho = 0$, the joint (weak) convergence of the first two coordinates of (5.8) follows from Theorem 7 in Ben Alaya and Kebaier [9]. □

6 Numerical illustrations

We present some numerical illustrations in order to confirm our limit theorems given in Sections 4 and 5. We call the attention to the fact that our numerical illustrations using synthetic data can not be considered as simulations or a receipt for handling real data set of (Y, S), since, as it will turn out, we use the standard Wiener processes $(W_t)_{t \in [0,\infty)}$ and $(B_t)_{t \in [0,\infty)}$ appearing

36

in (1.3) that can not be observed. Hence the main aim of this section is to confirm the scaling factors and the limit distributions of the derived MLE in Theorems 5.1 and 5.3. In order to approximate the estimator $\widehat{\psi}_T$ given in (3.10), one could generate sample paths of the model (1.3), and then one could approximate the estimator $\widehat{\psi}_T$ given in (3.10) based on the generated sample paths. For this, it would be sufficient to simulate, for a large time $T > 0$, the random variables

$$Y_T, \qquad I_{1,T} := \int_0^T Y_u \, \mathrm{d}u, \qquad I_{2,T} := \int_0^T \frac{\mathrm{d}u}{Y_u}, \qquad I_{3,T} := \int_0^T \frac{\mathrm{d}Y_u}{Y_u},$$

$$I_{4,T} := \int_0^T \frac{\mathrm{d}S_u - S_{u-} \, \mathrm{d}L_u}{S_{u-}}, \qquad I_{5,T} := \int_0^T \frac{\mathrm{d}S_u - S_{u-} \, \mathrm{d}L_u}{Y_u S_{u-}}.$$

It is well known that the random variable Y_T has a non-central chi-squared distribution (see, e.g., Alfonsi [2, Proposition 1.2.11]) that can be simulated exactly. Further, Broadie and Kaya [12, Section 3.2] proposed an exact simulation method of $(Y_T, I_{1,T})$, and more recently, Ben Alaya and Kebaier [8, Sections 4.1 and 4.2] developed an analogous method to simulate $(Y_T, I_{2,T})$. In the context of our current study, it would be possible to compute the Laplace transform of the couple $(I_{1,T}, I_{2,T})$ conditionally on Y_T, and using relation (3.13), we could derive an exact simulation method for the random vector $(Y_T, I_{1,T}, I_{2,T}, I_{3,T})$. However, due to the lack of an exact simulation method for the couple $(I_{4,T}, I_{5,T})$, we choose to approximate the quantities $(Y_T, I_{1,T}, I_{2,T}, I_{3,T}, I_{4,T}, I_{5,T})$ using discretization schemes, like the famous Euler one (see, e.g., Kloeden and Platen [26] or Alfonsi [2, Chapter 2]). Nevertheless, it is important to note that the discretization of the CIR process presents some troubles because of the square root in the diffusion coefficient. Several papers deal with this problem, see for example Alfonsi [1] and Berkaoui et al. [10].

For a given time step $\frac{T}{n}$ with $n \in \mathbb{N}$, we use the drift implicit Euler scheme introduced by Alfonsi [1] to approximate the process $(Y_t)_{t \in [0,T]}$ at times $t_i^n = i\frac{T}{n}$, $i \in \{0, \dots, n\}$, by the following non-linear recursion, $Y_0^n = y_0 \in (0, \infty)$ and

$$Y_{t_{i+1}^n}^n = \left(\frac{\frac{\sigma}{2}(W_{t_{i+1}^n} - W_{t_i^n}) + \sqrt{Y_{t_i^n}^n} + \sqrt{(\frac{\sigma}{2}(W_{t_{i+1}^n} - W_{t_i^n}) + \sqrt{Y_{t_i^n}^n})^2 + (1 + \frac{\kappa T}{2n})(2\theta\kappa - \frac{\sigma^2}{2})\frac{T}{n}}}{2 + \frac{\kappa T}{n}} \right)^2,$$

for $i \in \{0, \dots, n-1\}$. Note that, due to Alfonsi [1], this scheme is well defined for $\theta, \kappa \in (0, \infty)$ and $\theta\kappa \in (\frac{\sigma^2}{4}, \infty)$ covering the case $\theta\kappa \in [\frac{\sigma^2}{2}, \infty)$ as well, which ensures the unique existence of a MLE of (θ, κ, μ), see Proposition 3.2. Moreover, the strong convergence rate of this approximation is of order 1 in case of $\theta\kappa \in (\frac{\sigma^2}{2}, \infty)$, see Alfonsi [1] for more details. Then, we can easily approximate $I_{1,T}$, $I_{2,T}$ and $I_{3,T}$ respectively, by

$$I_{1,T}^n := \frac{T}{n} \sum_{i=0}^{n-1} Y_{t_i^n}^n, \qquad I_{2,T}^n := \frac{T}{n} \sum_{i=0}^{n-1} \frac{1}{Y_{t_i^n}^n}, \qquad I_{3,T}^n := \sum_{i=0}^{n-1} \frac{Y_{t_{i+1}^n}^n - Y_{t_i^n}^n}{Y_{t_i^n}^n}.$$

Alternatively, using the relation (3.13), one can also use the approximation

$$\widetilde{I}_{3,T}^n := \log(Y_{t_n^n}^n) - \log(y_0) + \frac{\sigma^2}{2} I_{2,T}^n.$$

Since we just would like to present some numerical illustrations of our limit theorems and not to provide simulations, we will not approximate the processes $(L_t)_{t\in[0,\infty)}$ and $(S_t)_{t\in[0,\infty)}$, instead, applying the equations (3.12), we can use

$$I_{4,T}^n := \mu T + \sum_{i=0}^{n-1} \sqrt{Y_{t_i^n}^n}(\varrho(W_{t_{i+1}^n} - W_{t_i^n}) + \sqrt{1-\varrho^2}(B_{t_{i+1}^n} - B_{t_i^n})),$$

$$I_{5,T}^n := \mu I_{2,T}^n + \sum_{i=0}^{n-1} \frac{\varrho(W_{t_{i+1}^n} - W_{t_i^n}) + \sqrt{1-\varrho^2}(B_{t_{i+1}^n} - B_{t_i^n})}{\sqrt{Y_{t_i^n}^n}}.$$

Here, we point out that $I_{4,T}^n$ and $I_{5,T}^n$ use the standard Wiener processes $(W_t)_{t\in[0,\infty)}$ and $(B_t)_{t\in[0,\infty)}$ appearing in (1.3) that can not be observed, so $I_{4,T}^n$ and $I_{5,T}^n$ can not be used for approximating $I_{4,T}$ and $I_{5,T}$, respectively, given a real dataset of (Y,S). However, the main advantage of this procedure is that it allows us to handle numerical illustrations involving any arbitrary purely non-Gaussian Lévy process $(L_t)_{t\in\mathbb{R}_+}$ with Lévy–Khintchine representation given in (1.4). Hence, by (3.11), we approximate $\widehat{\theta}_T$, $\widehat{\kappa}_T$ and $\widehat{\mu}_T$ by

$$\widehat{\theta}_T^n := \frac{I_{1,T}^n I_{2,T}^n I_{3,T}^n - TI_{2,T}^n(Y_T^n - y_0) + \varrho\sigma TI_{2,T}^n I_{4,T}^n - \varrho\sigma T^2 I_{5,T}^n}{TI_{2,T}^n I_{3,T}^n - (I_{2,T}^n)^2(Y_T^n - y_0) + \varrho\sigma(I_{2,T}^n)^2 I_{4,T}^n - \varrho\sigma TI_{2,T}^n I_{5,T}^n},$$

$$\widehat{\kappa}_T^n := \frac{TI_{3,T}^n - I_{2,T}^n(Y_T^n - y_0) + \varrho\sigma I_{2,T}^n I_{4,T}^n - \varrho\sigma TI_{5,T}^n}{(I_{1,T}^n I_{2,T}^n - T^2)},$$

$$\widehat{\mu}_T^n := \frac{I_{5,T}^n}{I_{2,T}^n}.$$

For the numerical implementation, we consider two case studies, one with $\theta\kappa > \sigma^2/2$, and another with $\theta\kappa = \sigma^2/2$.

First we take $\theta = 2$, $\kappa = 0.5$, $\mu = 1 - \sqrt{e}$, $\sigma = 0.2$, $\varrho = 0.5$, $y_0 = 1$, $s_0 = 100$, $\frac{T}{n} = 0.01$, and we simulate $M = 4000$ independent trajectories of the normalized error $T^{1/2}(\widehat{\psi}_T - \psi)$. Note that $\theta\kappa > \frac{\sigma^2}{2}$ with this choice of parameters. In Table 1 we give the relative errors for $T \in \{10, 100, 300\}$. Note that, when T increases we need of course a suitable number of time steps n to guarantee a good approximation. The obtained relative errors confirm the strong

Relative error	$T = 10$	$T = 100$	$T = 300$
$|\widehat{\theta}_T^n - \theta|/\theta$	0.0010578	0.0002387	0.0000658
$|\widehat{\kappa}_T^n - \kappa|/\kappa$	0.2803024	0.0532183	0.0214441
$|\widehat{\mu}_T^n - \mu|/\mu$	0.0380512	0.0060456	0.0034771

Table 1: Relative errors.

consistency of the estimator $\widehat{\psi}_T$ stated in Theorem 4.1. In Figure 1 we illustrate the law of each suitably scaled coordinate of the MLE $\widehat{\psi}_T = (\widehat{\theta}_T, \widehat{\kappa}_T, \widehat{\mu}_T)$ for $T = 300$. As a consequence
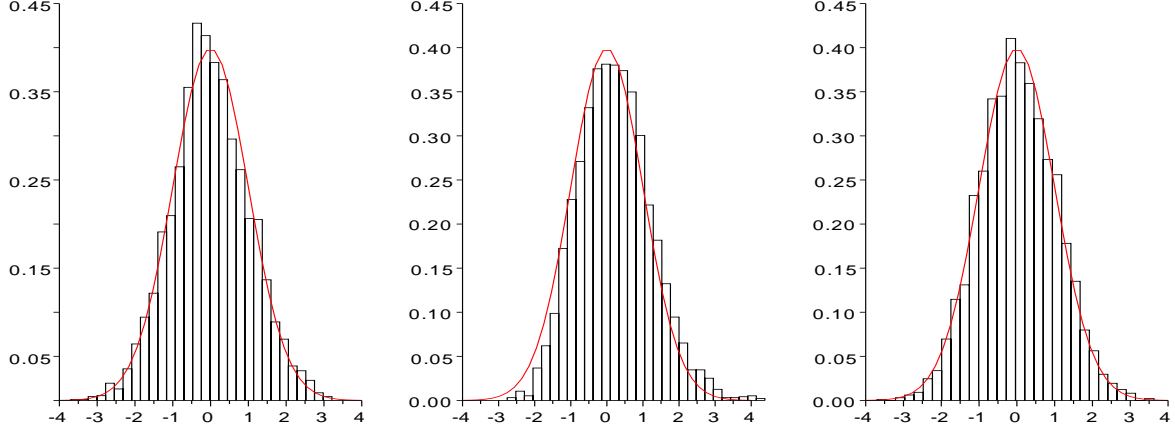
Figure 1: From the left to the right, the density histograms of the suitably scaled errors given in (6.1), (6.2) and (6.3). In each case, the red line denotes the density function of the standard normal distribution.

of Theorem 5.1, we have

$$\text{(6.1)} \qquad \sqrt{\frac{2\kappa^3 T}{\sigma^2(2\theta\kappa - \varrho^2\sigma^2)}} \left(\widehat{\theta}_T - \theta\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1) \qquad \text{as } T \to \infty,$$

$$\text{(6.2)} \qquad \sqrt{\frac{T}{2\kappa(1 - \varrho^2)}} \left(\widehat{\kappa}_T - \kappa\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1) \qquad \text{as } T \to \infty,$$

$$\text{(6.3)} \qquad \sqrt{\frac{2\kappa T}{2\theta\kappa - \sigma^2}} \left(\widehat{\mu}_T - \mu\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1) \qquad \text{as } T \to \infty.$$

The obtained density histograms in Figure 1 confirm our results in Theorem 5.1.

Next we take $\theta = 2$, $\kappa = 0.5$, $\mu = 1 - \sqrt{e}$, $\sigma = \sqrt{2}$, $\varrho = 0.5$, $y_0 = 1$, $s_0 = 100$, $n = 30000$, and we simulate $M = 4000$ independent trajectories of the appropriately normalized error $\widehat{\psi}_T - \psi$. Note that $\theta\kappa = \frac{\sigma^2}{2}$ with this choice of parameters. In Figure 2 we illustrate the law of each suitably scaled coordinate of the MLE $\widehat{\psi}_T = (\widehat{\theta}_T, \widehat{\kappa}_T, \widehat{\mu}_T)$ for $T = 300$. As a consequence of Theorem 5.3, we have

$$\text{(6.4)} \qquad \sqrt{\frac{2\kappa^3 T}{\sigma^4(1 - \varrho^2)}} \left(\widehat{\theta}_T - \theta\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1) \qquad \text{as } T \to \infty,$$

$$\text{(6.5)} \qquad \sqrt{\frac{T}{2\kappa(1 - \varrho^2)}} \left(\widehat{\kappa}_T - \kappa\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1) \qquad \text{as } T \to \infty,$$

$$\text{(6.6)} \qquad T(\widehat{\mu}_T - \mu) \xrightarrow{\mathcal{D}} \frac{\varrho\sigma}{\kappa\mathcal{T}} + \frac{\sigma\sqrt{1 - \varrho^2}}{\kappa\sqrt{\mathcal{T}}} Z_2 \qquad \text{as } T \to \infty.$$
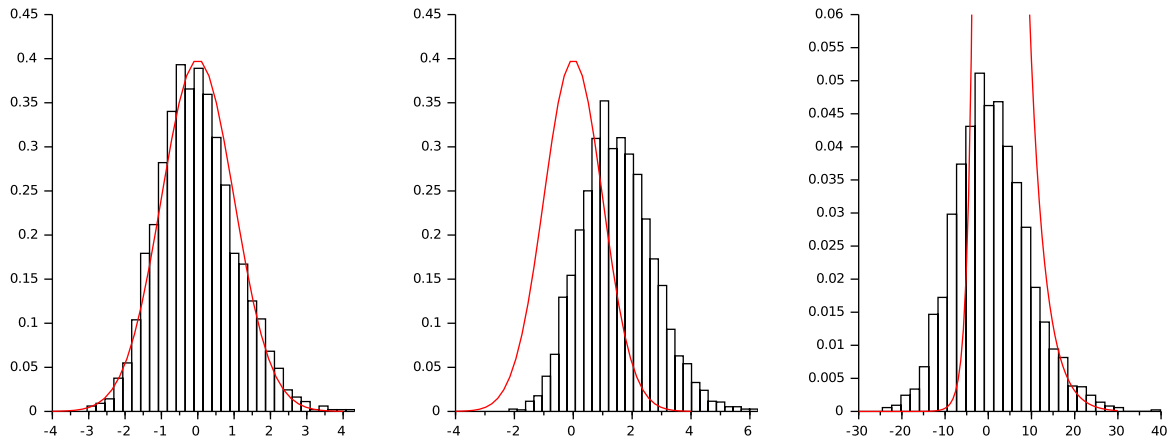
39

Figure 2: From the left to the right, the density histograms of the suitably scaled errors given in (6.4), (6.5) and (6.6). The red line denotes the density functions of corresponding limit distributions.

We plot the density function of the limit distribution in (6.6) using its explicit form given in Appendix D. Note that it is cutted at the level $0.06,$ since it tends to infinity at $0.$ In case of the parameter $\kappa,$ one can see a bias in Figure 2, which, in our opinion, is caused by the bad performance of the applied discretization scheme together with the approximation method of the integrals in question, when $\theta\kappa = \frac{\sigma^2}{2}.$ We have not been able to find any discretization scheme to explain the bias (we tried the truncated Euler scheme, see, e.g., Deelstra and Delbaen [14], and the symmetrized Euler scheme, see, e.g., Diop [15] or Berkaoui et al. [10]). Eventually, this bad performance can also be observed whenever the ratio $\frac{2\theta\kappa}{\sigma^2}$ is close to 1. And to top it all, one can observe the same phenomena already in case of the MLE $(\widehat{a}_T, \widehat{b}_T)$ of (a, b) of the first coordinate process of the SDE (1.1) based on continuous time observations $(Y_t)_{t\in[0,T]},$ $T \in (0, \infty),$ for both $\widehat{a}_T$ and $\widehat{b}_T$ (for an expression of $(\widehat{a}_T, \widehat{b}_T),$ see Overbeck [33]). So we conclude that the bias for κ seen in Figure 2 is not related to the fact that the model (1.3) contains a jump part.

As we mentioned in the Introduction, the model (1.3) with L as a compound Poisson process given in (1.5) is quite popular in finance. In this special case, one can use another illustration method without applying the equations (3.12), but still using the standard Wiener processes $(W_t)_{t\in[0,\infty)}$ and $(B_t)_{t\in[0,\infty)}.$ Namely, for all $0 \leqslant s < t,$ by (1.6),

$$S_t = S_s \exp\left\{ \int_s^t \left(\mu - \frac{1}{2} Y_u \right) \mathrm{d}u + \int_s^t \sqrt{Y_u} \left(\varrho \, \mathrm{d}W_u + \sqrt{1 - \varrho^2} \, \mathrm{d}B_u \right) + \sum_{k=\pi_s+1}^{\pi_t} J_k \right\},$$

hence we can approximate the price process $(S_t)_{t\in[0,T]}$ by the recursion $S_0^n = s_0 \in (0, \infty)$ and

$$\frac{S_{t_{i+1}^n}^n}{S_{t_i^n}^n} = \exp\left\{ \frac{\mu T}{n} - \frac{T}{2n} Y_{t_i^n}^n + \sqrt{Y_{t_i^n}^n} \left(\varrho (W_{t_{i+1}^n} - W_{t_i^n}) + \sqrt{1 - \varrho^2} (B_{t_{i+1}^n} - B_{t_i^n}) \right) + \sum_{k=\pi_{t_i^n}+1}^{\pi_{t_{i+1}^n}} J_k \right\}$$

40

for $i \in \{0, \ldots, n-1\}$. Note that the process $(\pi_t)_{t \in [0,\infty)}$ is a Poisson process with intensity λ being independent of $(W_t, B_t)_{t \in [0,\infty)}$, and it can be easily simulated. Therefore, given independently an i.i.d. sequence of random variables $(J_k)_{k \in \mathbb{N}}$, one can simulate at the same time the term $\sum_{k=\pi_{t_i^n}+1}^{\pi_{t_{i+1}^n}} J_k$ together with the increments $L_{t_{i+1}^n} - L_{t_i^n} = \sum_{k=\pi_{t_i^n}+1}^{\pi_{t_{i+1}^n}} (e^{J_k} - 1)$, $i \in \{0, \ldots, n-1\}$. Further, one can approximate $I_{4,T}$ and $I_{5,T}$ respectively by

$$\widetilde{I}_{4,T}^n := \sum_{i=0}^{n-1} \left(\frac{S_{t_{i+1}^n}^n}{S_{t_i^n}^n} - 1 \right) - L_T \quad \text{and} \quad \widetilde{I}_{5,T}^n := \sum_{i=0}^{n-1} \frac{1}{Y_{t_i^n}^n} \left(\frac{S_{t_{i+1}^n}^n}{S_{t_i^n}^n} - 1 \right) - \sum_{i=0}^{n-1} \frac{L_{t_{i+1}^n} - L_{t_i^n}}{Y_{t_i^n}^n}.$$

We remark that $S_{t_i^n}^n \in (0, \infty)$, $n \in \mathbb{N}$, $i \in \{0, 1, \ldots, n-1\}$, so $\widetilde{I}_{4,T}^n$ and $\widetilde{I}_{5,T}^n$ are well-defined. Here we take again $\theta = 2$, $\kappa = 0.5$, $\mu = 1 - \sqrt{e}$, $\sigma = 0.2$, $\varrho = 0.5$, $y_0 = 1$, $s_0 = 100$, $n = 30000$ and additionally, $\lambda = 1$ and a random variable J_1 with standard normal distribution. We simulate $M = 2000$ independent trajectories of the normalized error $T^{1/2}(\widehat{\boldsymbol{\psi}}_T - \boldsymbol{\psi})$. Note that $\theta\kappa > \frac{\sigma^2}{2}$ with this choice of parameters. In Figure 3 we illustrate the law of each suitably scaled coordinate of the MLE $\widehat{\boldsymbol{\psi}}_T = (\widehat{\theta}_T, \widehat{\kappa}_T, \widehat{\mu}_T)$ which confirms our results in Theorem 5.1.
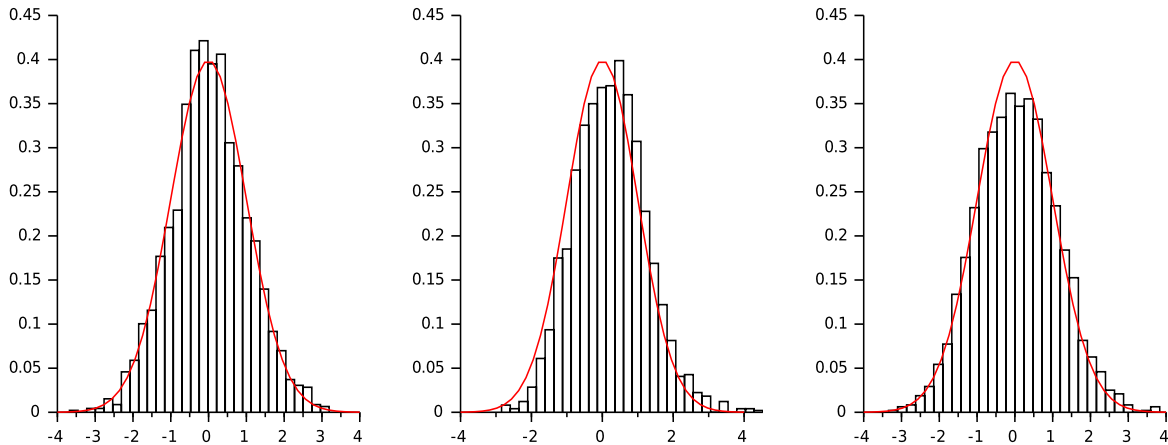

Figure 3: From the left to the right, the density histograms of the suitably scaled errors given in (6.1), (6.2) and (6.3). In each case, the red line denotes the density function of the standard normal distribution.

Finally, we note that we used the open source software Scilab for making the simulations.

Appendix

A   Likelihood-ratio process

Based on Jacod and Shiryaev [22], see also Jacod and Mémin [20], Sørensen [36] and Luschgy [32], we recall certain sufficient conditions for the absolute continuity of probability measures in-

duced by semimartingales together with a representation of the corresponding Radon–Nikodym derivative (likelihood-ratio process).

Let $D([0, \infty), \mathbb{R}^d)$ denote the space of $\mathbb{R}^d$-valued càdlàg functions defined on $[0, \infty)$. Let $(\eta_t)_{t \in [0, \infty)}$ denote the canonical process $\eta_t(\omega) := \omega(t)$, $\omega \in D([0, \infty), \mathbb{R}^d)$, $t \in [0, \infty)$. Put $\mathcal{F}_t^\eta := \sigma(\eta_s, s \in [0, t])$, $t \in [0, \infty)$, and

$$\mathcal{D}_t([0, \infty), \mathbb{R}^d) := \bigcap_{\varepsilon > 0} \mathcal{F}_{t+\varepsilon}^\eta, \quad t \in [0, \infty), \qquad \mathcal{D}([0, \infty), \mathbb{R}^d) := \sigma \left(\bigcup_{t \in [0, \infty)} \mathcal{F}_t^\eta \right).$$

Let $\Psi \subset \mathbb{R}^k$ be an arbitrary non-empty set, and let $\mathbb{P}_\psi$, $\psi \in \Psi$, are probability measures on the canonical space $(D([0, \infty), \mathbb{R}^d), \mathcal{D}([0, \infty), \mathbb{R}^d))$. Suppose that for each $\psi \in \Psi$, under $\mathbb{P}_\psi$, the canonical process $(\eta_t)_{t \in [0, \infty)}$ is a semimartingale with semimartingale characteristics $(B^{(\psi)}, C, \nu^{(\psi)})$ associated with a fixed Borel measurable truncation function $h : \mathbb{R}^d \to \mathbb{R}^d$, see Jacod and Shiryaev [22, Definition II.2.6 and Remark II.2.8]. Namely, $C_t := \langle (\eta^{\mathrm{cont}})^{(\psi)} \rangle_t$, $t \in [0, \infty)$, where $(\langle (\eta^{\mathrm{cont}})^{(\psi)} \rangle_t)_{t \in [0, \infty)}$ denotes the (predictable) quadratic variation process (with values in $\mathbb{R}^{d \times d}$) of the continuous martingale part $(\eta^{\mathrm{cont}})^{(\psi)}$ of η under $\mathbb{P}_\psi$, $\nu^{(\psi)}$ is the compensator of the integer-valued random measure μ^η on $[0, \infty) \times \mathbb{R}^d$ associated with the jumps of η under $\mathbb{P}_\psi$ given by

$$(\mathrm{A.1}) \qquad \mu^\eta(\omega, \mathrm{d}t, \mathrm{d}\boldsymbol{x}) := \sum_{s \in [0, \infty)} \mathbb{1}_{\{\Delta \eta_s(\omega) \neq \boldsymbol{0}\}} \varepsilon_{(s, \Delta \eta_s(\omega))}(\mathrm{d}t, \mathrm{d}\boldsymbol{x}), \qquad \omega \in D([0, \infty), \mathbb{R}^d),$$

where $\varepsilon_{(t, \boldsymbol{x})}$ denotes the Dirac measure at the point $(t, \boldsymbol{x}) \in [0, \infty) \times \mathbb{R}^d$, and $\Delta \eta_t := \eta_t - \eta_{t-}$, $t \in (0, \infty)$, $\Delta \eta_0 := \boldsymbol{0}$, and $B^{(\psi)}$ is the predictable process (with values in $\mathbb{R}^d$ having finite variation over each finite interval $[0, t]$, $t \in [0, \infty)$) appearing in the canonical decomposition

$$\widetilde{\eta}_t = \eta_0 + M_t^{(\psi)} + B_t^{(\psi)}, \qquad t \in [0, \infty),$$

of the special semimartingale $(\widetilde{\eta}_t)_{t \in [0, \infty)}$ under $\mathbb{P}_\psi$ given by

$$\widetilde{\eta}_t := \eta_t - \sum_{s \in [0, t]} (\eta_s - h(\Delta \eta_s)), \qquad t \in [0, \infty),$$

where $(M_t^{(\psi)})_{t \in [0, \infty)}$ is a local martingale with $M_0^{(\psi)} = \boldsymbol{0}$. We call the attention that, by our assumption, the process $C = \langle (\eta^{\mathrm{cont}})^{(\psi)} \rangle$ does not depend on ψ, although $(\eta^{\mathrm{cont}})^{(\psi)}$ might depend on ψ. In addition, assume that $\mathbb{P}_\psi(\nu^{(\psi)}(\{t\} \times \mathbb{R}^d) = 0) = 1$ for every $\psi \in \Psi$, $t \in [0, \infty)$, and $\mathbb{P}_\psi(\eta_0 = \boldsymbol{x}_0) = 1$ with some $\boldsymbol{x}_0 \in \mathbb{R}^d$ for every $\psi \in \Psi$. Note that we have the semimartingale representation

$$\eta_t = \boldsymbol{x}_0 + B_t^{(\psi)} + (\eta^{\mathrm{cont}})_t^{(\psi)} + \int_0^t \int_{\mathbb{R}^d} h(\boldsymbol{x}) (\mu^\eta - \nu^{(\psi)})(\mathrm{d}s, \mathrm{d}\boldsymbol{x})$$

$$(\mathrm{A.2})$$

$$+ \int_0^t \int_{\mathbb{R}^d} (\boldsymbol{x} - h(\boldsymbol{x})) \mu^\eta(\mathrm{d}s, \mathrm{d}\boldsymbol{x}), \qquad t \in [0, \infty),$$

42

of η under $\mathbb{P}_{\psi}$, see Jacod and Shiryaev [22, Theorem II.2.34]. Moreover, for each $\psi \in \Psi$, let us choose a nondecreasing, continuous, adapted process $(F_t^{(\psi)})_{t\in[0,\infty)}$ with $F_0^{(\psi)} = 0$ and a predictable process $(c_t^{(\psi)})_{t\in[0,\infty)}$ with values in the set of all symmetric positive semidefinite $d \times d$ matrices such that

$$C_t = \int_0^t c_s^{(\psi)} \, \mathrm{d}F_s^{(\psi)}$$

$\mathbb{P}_{\psi}$-almost sure for every $t \in [0,\infty)$. Due to the assumption $\mathbb{P}_{\psi}(\nu^{(\psi)}(\{t\} \times \mathbb{R}^d) = 0) = 1$ for every $\psi \in \Psi$, $t \in [0,\infty)$, such choices of $(F_t^{(\psi)})_{t\in[0,\infty)}$ and $(c_t^{(\psi)})_{t\in[0,\infty)}$ are possible, see Jacod and Shiryaev [22, Proposition II.2.9 and Corollary II.1.19]. Let $\mathcal{P}$ denote the predictable σ-algebra on $D([0,\infty),\mathbb{R}^d) \times [0,\infty)$. Assume also that for every $\psi, \widetilde{\psi} \in \Psi$, there exist a $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^d)$-measurable function $V^{(\widetilde{\psi},\psi)} : D([0,\infty),\mathbb{R}^d) \times [0,\infty) \times \mathbb{R}^d \to (0,\infty)$ and a predictable $\mathbb{R}^d$-valued process $\beta^{(\widetilde{\psi},\psi)}$ satisfying

(A.3) $$\nu^{(\psi)}(\mathrm{d}t, \mathrm{d}\boldsymbol{x}) = V^{(\widetilde{\psi},\psi)}(t, \boldsymbol{x})\nu^{(\widetilde{\psi})}(\mathrm{d}t, \mathrm{d}\boldsymbol{x}),$$

(A.4) $$\int_0^t \int_{\mathbb{R}^d} \left(\sqrt{V^{(\widetilde{\psi},\psi)}(s, \boldsymbol{x})} - 1\right)^2 \nu^{(\widetilde{\psi})}(\mathrm{d}s, \mathrm{d}\boldsymbol{x}) < \infty,$$

(A.5) $$B_t^{(\psi)} = B_t^{(\widetilde{\psi})} + \int_0^t c_s^{(\psi)}\beta_s^{(\widetilde{\psi},\psi)} \, \mathrm{d}F_s^{(\psi)} + \int_0^t \int_{\mathbb{R}^d} (V^{(\widetilde{\psi},\psi)}(s, \boldsymbol{x}) - 1)h(\boldsymbol{x})\, \nu^{(\widetilde{\psi})}(\mathrm{d}s, \mathrm{d}\boldsymbol{x}),$$

(A.6) $$\int_0^t (\beta_s^{(\widetilde{\psi},\psi)})^\top c_s^{(\psi)}\beta_s^{(\widetilde{\psi},\psi)} \, \mathrm{d}F_s^{(\psi)} < \infty,$$

$\mathbb{P}_{\psi}$-almost sure for every $t \in [0,\infty)$. Further, assume that for each $\psi \in \Psi$, local uniqueness holds for the martingale problem on the canonical space corresponding to the triplet $(B^{(\psi)}, C, \nu^{(\psi)})$ with the given initial value $\boldsymbol{x}_0$ with $\mathbb{P}_{\psi}$ as its unique solution. Then for each $T \in [0,\infty)$, $\mathbb{P}_{\psi,T}$ is absolutely continuous with respect to $\mathbb{P}_{\widetilde{\psi},T}$, where $\mathbb{P}_{\psi,T} := \mathbb{P}_{\psi}|_{\mathcal{D}_T([0,\infty),\mathbb{R}^d)}$ denotes the restriction of $\mathbb{P}_{\psi}$ to $\mathcal{D}_T([0,\infty),\mathbb{R}^d)$ (similarly for $\mathbb{P}_{\widetilde{\psi},T}$), and, under $\mathbb{P}_{\widetilde{\psi},T}$, the corresponding likelihood-ratio process takes the form

$$\log \frac{\mathrm{d}\mathbb{P}_{\psi,T}}{\mathrm{d}\mathbb{P}_{\widetilde{\psi},T}}(\eta) = \int_0^T (\beta_s^{(\widetilde{\psi},\psi)})^\top \, \mathrm{d}(\eta^{\mathrm{cont}})_s^{(\widetilde{\psi})} - \frac{1}{2}\int_0^T (\beta_s^{(\widetilde{\psi},\psi)})^\top c_s^{(\psi)}\beta_s^{(\widetilde{\psi},\psi)} \, \mathrm{d}F_s^{(\psi)}$$

(A.7) $$+ \int_0^T \int_{\mathbb{R}^d} (V^{(\widetilde{\psi},\psi)}(s, \boldsymbol{x}) - 1)\, (\mu^\eta - \nu^{(\widetilde{\psi})})(\mathrm{d}s, \mathrm{d}\boldsymbol{x})$$

$$+ \int_0^T \int_{\mathbb{R}^d} (\log(V^{(\widetilde{\psi},\psi)}(s, \boldsymbol{x})) - V^{(\widetilde{\psi},\psi)}(s, \boldsymbol{x}) + 1)\, \mu^\eta(\mathrm{d}s, \mathrm{d}\boldsymbol{x})$$

for all $T \in (0,\infty)$, see Jacod and Shiryaev [22, Theorem III.5.34].

In what follows we give a proof for (A.7) using Jacod and Shiryaev [22], since in the literature we could not find a detailed proof. Using the notations of Jacod and Shiryaev [22], under $\mathbb{P}_{\psi}$ the triplets $(B^{(\psi)}, C, \nu^{(\psi)})$ and $(B^{(\widetilde{\psi})}, C, \nu^{(\widetilde{\psi})})$ satisfy III.5.5 in Jacod and Shiryaev [22] with $Y = V^{(\widetilde{\psi},\psi)}$, $A = F^{(\psi)}$ and $c = c^{(\psi)}$, and the filtration $\mathcal{D}_t([0,\infty),\mathbb{R}^d)$, $t \in [0,\infty)$, is generated by $(\eta_t)_{t\in[0,\infty)}$. Moreover, $a_t = \nu^{(\psi)}(\{t\} \times \mathbb{R}^d) = 0$ and $\widehat{Y}_t =$

$\int_{\mathbb{R}^d} V^{(\widetilde{\psi},\psi)}(t,\boldsymbol{x})\,\nu^{(\psi)}(\{t\}\times\mathrm{d}\boldsymbol{x}) = 0$ $\mathbb{P}_\psi$-almost surely for all $t\in[0,\infty)$, hence $\sigma = \inf\{t\in [0,\infty) : \text{either } \widehat{Y}_t > 1, \text{ or } a_t = 1 \text{ and } \widehat{Y}_t < 1\} = \infty$ $\mathbb{P}_\psi$-almost surely. Then, by (A.4) and (A.6), we have

$$H_t = \int_0^t (\beta_s^{(\widetilde{\psi},\psi)})^\top c_s^{(\psi)}\beta_s^{(\widetilde{\psi},\psi)}\,\mathrm{d}F_s^{(\psi)} + \int_0^t\int_{\mathbb{R}^d}\left(\sqrt{V^{(\widetilde{\psi},\psi)}(s,\boldsymbol{x})}-1\right)^2\nu^{(\widetilde{\psi})}(\mathrm{d}s,\mathrm{d}\boldsymbol{x}) < \infty$$

$\mathbb{P}_\psi$-almost sure for every $t\in[0,\infty)$. Consequently, $T_n = \inf\{t\in[0,\infty) : H_t \geq n\}\to\infty$ as $n\to\infty$ $\mathbb{P}_\psi$-almost sure (due to the fact that $(H_t)_{t\in[0,\infty)}$ is a nondecreasing process), and hence $\mathbb{P}_\psi(\Delta=[0,\infty))=1$, and the Hypothesis III.5.29 in Jacod and Shiryaev [22] holds. Thus, by Theorem III.5.34 in Jacod and Shiryaev [22], $\mathbb{P}_{\psi,T}$ is absolutely continuous with respect to $\mathbb{P}_{\widetilde{\psi},T}$ for all $T\in[0,\infty)$, and under $\mathbb{P}_{\widetilde{\psi},T}$, the density process (likelihood ratio process) $(Z_T)_{T\in[0,\infty)}$ takes the form

$$Z_T = \frac{\mathrm{d}\,\mathbb{P}_{\psi,T}}{\mathrm{d}\,\mathbb{P}_{\widetilde{\psi},T}}(\eta) = \exp\left\{N_T - \frac{1}{2}\int_0^T (\beta_s^{(\widetilde{\psi},\psi)})^\top c_s^{(\psi)}\beta_s^{(\widetilde{\psi},\psi)}\,\mathrm{d}F_s^{(\psi)}\right\}\prod_{s\in[0,T]}(1+\Delta N_s)\mathrm{e}^{-\Delta N_s}$$

for $T\in[0,\infty)$ with

$$N_T := \int_0^T (\beta_s^{(\widetilde{\psi},\psi)})^\top\,\mathrm{d}(\eta^{\mathrm{cont}})_s^{(\widetilde{\psi})} + \int_0^T\int_{\mathbb{R}^d}(V^{(\widetilde{\psi},\psi)}(s,\boldsymbol{x})-1)\,(\mu^\eta - \nu^{(\widetilde{\psi})})(\mathrm{d}s,\mathrm{d}\boldsymbol{x})$$

for $T\in[0,\infty)$. Further, the density process $(Z_T)_{T\in[0,\infty)}$ satisfies

$$Z_T = 1 + \int_0^T Z_{s-}\,\mathrm{d}N_s \qquad \text{under } \mathbb{P}_{\widetilde{\psi},T} \text{ for each } T\in[0,\infty),$$

see Jacod and Shiryaev [22, III.5.20]. Taking into account the fact that $\mathbb{P}_\psi(\nu^{(\psi)}(\{s\}\times\mathbb{R}^d) = 0) = 1$ for every $\psi\in\Psi$, $s\in[0,\infty)$, and the definition of the stochastic integral with respect to the random measure $\mu^\eta - \nu^{(\widetilde{\psi})}$ (see Jacod and Shiryaev [22, Definition II.1.27]), we obtain

$$\Delta N_s(\omega) = (V^{(\widetilde{\psi},\psi)}(s,\Delta\eta_s(\omega))(\omega) - 1)\mathbb{1}_{\{\Delta\eta_s(\omega)\neq 0\}}, \qquad \omega\in D([0,\infty),\mathbb{R}^d), \quad s\in[0,\infty).$$

Hence, using that $V^{(\widetilde{\psi},\psi)}$ is positive, we have $\mathbb{P}_{\widetilde{\psi}}(\inf\{t\in[0,\infty) : \Delta N_t = -1\} = \infty) = 1$, and consequently, $\mathbb{P}_{\widetilde{\psi}}(Z_T\in(0,\infty)) = 1$ for all $T\in(0,\infty)$, see Jacod and Shiryaev [22, Theorem I.4.61]. Further,

$$\prod_{s\in[0,T]}(1+\Delta N_s)\mathrm{e}^{-\Delta N_s} = \prod_{s\in[0,T]}(V^{(\widetilde{\psi},\psi)}(s,\Delta\eta_s)\mathbb{1}_{\{\Delta\eta_s\neq 0\}} + \mathbb{1}_{\{\Delta\eta_s=0\}})\mathrm{e}^{-(V^{(\widetilde{\psi},\psi)}(s,\Delta\eta_s)-1)\mathbb{1}_{\{\Delta\eta_s\neq 0\}}}$$

$$= \exp\left\{\sum_{s\in[0,T]}\left(\log(V^{(\widetilde{\psi},\psi)}(s,\Delta\eta_s)) - V^{(\widetilde{\psi},\psi)}(s,\Delta\eta_s) + 1\right)\mathbb{1}_{\{\Delta\eta_s\neq 0\}}\right\}$$

$$= \exp\left\{\int_0^T\int_{\mathbb{R}^d}\left(\log(V^{(\widetilde{\psi},\psi)}(s,\boldsymbol{x})) - V^{(\widetilde{\psi},\psi)}(s,\boldsymbol{x}) + 1\right)\mu^\eta(\mathrm{d}s,\mathrm{d}\boldsymbol{x})\right\},$$

yielding (A.7), where the existence of the integral in the exponent above follows from the facts that $\mathbb{P}_{\widetilde{\psi}}(Z_T\in(0,\infty)) = 1$, assumption (A.6) and $\mathbb{P}_{\widetilde{\psi}}(N_T\in\mathbb{R}) = 1$ for all $T\in\mathbb{R}$ (due to Jacod and Shiryaev [22, Proposition III.5.10 and III.5.12]).

44

B   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 $\boldsymbol{\psi} = (\theta, \kappa, \mu)$. First we recall a strong law of large numbers for continuous local martingales.

Theorem B.1 (Liptser and Shiryaev [30, Lemma 17.4]) *Let* $\big(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0, \infty)}, \mathbb{P}\big)$ *be a filtered probability space satisfying the usual conditions. Let* $(M_t)_{t \in [0, \infty)}$ *be a square-integrable continuous local martingale with respect to the filtration* $(\mathcal{F}_t)_{t \in [0, \infty)}$ *such that* $\mathbb{P}(M_0 = 0) = 1$. *Let* $(\xi_t)_{t \in [0, \infty)}$ *be a progressively measurable process such that*

$$\mathbb{P}\left(\int_0^t \xi_u^2 \, \mathrm{d}\langle M \rangle_u < \infty \right) = 1, \qquad t \in [0, \infty),$$

and

$$(B.1) \qquad \int_0^t \xi_u^2 \, \mathrm{d}\langle M \rangle_u \xrightarrow{\text{a.s.}} \infty \qquad \text{as } t \to \infty,$$

where $(\langle M \rangle_t)_{t \in [0, \infty)}$ *denotes the (predictable) quadratic variation process of* M. *Then*

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

If $(M_t)_{t \in [0, \infty)}$ *is a standard Wiener process, the progressive measurability of* $(\xi_t)_{t \in [0, \infty)}$ *can be relaxed to measurability and adaptedness to the filtration* $(\mathcal{F}_t)_{t \in [0, \infty)}$.

The next theorem is a special case of the central limit theorem for multidimensional square-integrable continuous local martingales, see, e.g., Jacod and Shiryaev [22, Corollary VIII.3.24] or van Zanten [40, Theorem 4.1].

Theorem B.2 *Let* $\big(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0, \infty)}, \mathbb{P}\big)$ *be a filtered probability space satisfying the usual conditions. Let* $(\boldsymbol{M}_t)_{t \in [0, \infty)}$ *be a d-dimensional square-integrable continuous local martingale with respect to the filtration* $(\mathcal{F}_t)_{t \in [0, \infty)}$ *such that* $\mathbb{P}(\boldsymbol{M}_0 = \boldsymbol{0}) = 1$ *and*

$$t^{-1} \langle \boldsymbol{M} \rangle_t \xrightarrow{\mathbb{P}} \boldsymbol{U} \qquad \text{as } t \to \infty,$$

where $\boldsymbol{U} \in \mathbb{R}^{d \times d}$. *Then*

$$t^{-1/2} \boldsymbol{M}_t \xrightarrow{\mathcal{D}} \mathcal{N}_d(\boldsymbol{0}, \boldsymbol{U}) \qquad \text{as } t \to \infty.$$

C   A version of the continuous mapping theorem

The following version of continuous mapping theorem can be found for example in Kallenberg [24, Theorem 3.27].

Lemma C.1 *Let* $(\mathcal{S}_1, d_{\mathcal{S}_1})$ *and* $(\mathcal{S}_2, d_{\mathcal{S}_2})$ *be metric spaces and* $(\xi_n)_{n\in\mathbb{N}},$ ξ *be random elements with values in* $\mathcal{S}_1$ *such that* $\xi_n \xrightarrow{\mathcal{D}} \xi$ *as* $n \to \infty$. *Let* $F : \mathcal{S}_1 \to \mathcal{S}_2$ *and* $F_n : \mathcal{S}_1 \to \mathcal{S}_2,$ $n \in \mathbb{N}$, *be measurable mappings and* $\mathcal{C} \in \mathcal{B}(\mathcal{S}_1)$ *such that* $\mathbb{P}(\xi \in \mathcal{C}) = 1$ *and* $\lim_{n\to\infty} d_{\mathcal{S}_2}(F_n(s_n), F(s)) = 0$ *if* $\lim_{n\to\infty} d_{\mathcal{S}_1}(s_n, s) = 0$ *and* $s \in \mathcal{C}$. *Then* $F_n(\xi_n) \xrightarrow{\mathcal{D}} F(\xi)$ *as* $n \to \infty$.

D Explicit formula for a density function

We show that the mixed normal but non-normal density function of $\frac{\varrho\sigma}{\kappa\mathcal{T}} + \frac{\sigma\sqrt{1-\varrho^2}}{\kappa\sqrt{\mathcal{T}}}Z_2$, which is the limit distribution of $T(\widehat{\mu}_T - \mu)$ as $T \to \infty$ in Theorem 5.3, has the form

$$(\text{D.1}) \qquad f(x) = \frac{\kappa}{2\pi\sigma\sqrt{1-\varrho^2}} \int_0^\infty \frac{1}{t} \exp\left\{-\frac{1}{2t} - \frac{(\kappa x t - \varrho\sigma)^2}{2\sigma^2(1-\varrho^2)t}\right\} \mathrm{d}t, \qquad x \in \mathbb{R} \setminus \{0\},$$

and

$$(\text{D.2}) \qquad \lim_{x\to 0} f(x) = \infty.$$

It is known that the density function of $\mathcal{T}$ takes the form $f_{\mathcal{T}}(t) = (2\pi t^3)^{-1/2}\mathrm{e}^{-1/(2t)}\mathbb{1}_{\mathbb{R}_{++}}(t)$, $t \in \mathbb{R}$. Using the independence of $\mathcal{T}$ and Z_2, we have

$$\mathbb{P}\left(\frac{\varrho\sigma}{\kappa\mathcal{T}} + \frac{\sigma\sqrt{1-\varrho^2}}{\kappa\sqrt{\mathcal{T}}}Z_2 \leqslant x \,\bigg|\, \mathcal{T} = t\right) = \mathbb{P}\left(Z_2 \leqslant \frac{(x - \frac{\varrho\sigma}{\kappa\mathcal{T}})\kappa\sqrt{\mathcal{T}}}{\sigma\sqrt{1-\varrho^2}} \,\bigg|\, \mathcal{T} = t\right)$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{\kappa x t - \varrho\sigma}{\sigma\sqrt{(1-\varrho^2)t}}} \mathrm{e}^{-u^2/2} \,\mathrm{d}u, \qquad x \in \mathbb{R}, \quad t \in \mathbb{R}_{++}.$$

By the law of total expectation, we obtain

$$\mathbb{P}\left(\frac{\varrho\sigma}{\kappa\mathcal{T}} + \frac{\sigma\sqrt{1-\varrho^2}}{\kappa\sqrt{\mathcal{T}}}Z_2 \leqslant x\right) = \int_0^\infty g(x,t)\,\mathrm{d}t, \qquad x \in \mathbb{R},$$

with

$$g(x,t) := \frac{1}{\sqrt{2\pi t^3}}\mathrm{e}^{-\frac{1}{2t}}\left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{\kappa x t - \varrho\sigma}{\sigma\sqrt{(1-\varrho^2)t}}} \mathrm{e}^{-u^2/2}\,\mathrm{d}u\right), \qquad x \in \mathbb{R}, \quad t \in \mathbb{R}_{++}.$$

The aim of the following discussion is to show that, by the dominated convergence theorem,

$$(\text{D.3}) \qquad \frac{\mathrm{d}}{\mathrm{d}x} \int_0^\infty g(x,t)\,\mathrm{d}t = \lim_{h\to 0} \int_0^\infty \frac{g(x+h,t) - g(x,t)}{h}\,\mathrm{d}t = f(x), \qquad x \in \mathbb{R} \setminus \{0\}.$$

For all $x \in \mathbb{R}$ and $t \in \mathbb{R}_{++}$, we have

$$(\text{D.4}) \qquad \lim_{h\to 0} \frac{g(x+h,t) - g(x,t)}{h} = \partial_1 g(x,t) = \frac{\kappa}{2\pi\sigma\sqrt{1-\varrho^2}t} \exp\left\{-\frac{1}{2t} - \frac{(\kappa x t - \varrho\sigma)^2}{2\sigma^2(1-\varrho^2)t}\right\}.$$

Moreover, for all $x \in \mathbb{R} \setminus \{0\}$, $t \in \mathbb{R}_{++}$ and $h \in \left[-\frac{|x|}{2}, \frac{|x|}{2}\right] \setminus \{0\}$, we have

$$\left|\frac{g(x+h,t) - g(x,t)}{h}\right| \leqslant \sup_{\xi \in [-1,1]} |\partial_1 g(x+\xi h, t)| \leqslant \sup_{\xi \in [-1,1]} \left|\partial_1 g\left(x + \xi\frac{x}{2}, t\right)\right|.$$

For $x \in \mathbb{R} \setminus \{0\}$, $\xi \in [-1,1]$ and $t \in \mathbb{R}_{++}$, we have

$$\left|\partial_1 g\left(x + \xi\frac{x}{2}, t\right)\right| \leqslant \frac{\kappa}{2\pi\sigma\sqrt{1-\varrho^2}\,t} e^{-\frac{1}{2t}} \leqslant \frac{\kappa}{e\pi\sigma\sqrt{1-\varrho^2}},$$

since $\sup_{t \in \mathbb{R}_{++}} t^{-1} e^{-\frac{1}{2t}} = 2e^{-1}$. Further, for $x \in \mathbb{R} \setminus \{0\}$, $\xi \in [-1,1]$ and $t \in \mathbb{R}_{++}$, we have

$$\left|\kappa\left(x+\xi\frac{x}{2}\right)t - \varrho\sigma\right| \geqslant \left|\kappa\left(x+\xi\frac{x}{2}\right)t\right| - |\varrho\sigma| \geqslant \frac{1}{2}\left|\kappa\left(x+\xi\frac{x}{2}\right)t\right|$$

whenever

$$\frac{1}{2}\left|\kappa\left(x+\xi\frac{x}{2}\right)t\right| \geqslant |\varrho\sigma|,$$

equivalently, whenever

$$t \geqslant \frac{2|\varrho\sigma|}{\left|\kappa\left(x+\xi\frac{x}{2}\right)\right|} = \frac{2|\varrho|\sigma}{\kappa\left(1+\frac{1}{2}\xi\right)|x|},$$

which holds if

$$t \geqslant \frac{4|\varrho|\sigma}{\kappa|x|} + 1 =: T_0(x) \in \mathbb{R}_{++}.$$

Consequently, for $x \in \mathbb{R} \setminus \{0\}$, $\xi \in [-1,1]$ and $t \in (T_0(x), \infty)$, we have

$$\left|\kappa\left(x+\xi\frac{x}{2}\right)t - \varrho\sigma\right| \geqslant \frac{1}{2}\left|\kappa\left(x+\xi\frac{x}{2}\right)t\right| = \frac{1}{2}\kappa\left(1+\frac{\xi}{2}\right)|x|t \geqslant \frac{\kappa|x|t}{4},$$

and hence, by the second equality in (D.4),

$$\left|\partial_1 g\left(x + \xi\frac{x}{2}, t\right)\right| \leqslant \frac{\kappa}{2\pi\sigma\sqrt{1-\varrho^2}\,T_0(x)} \exp\left\{-\frac{\kappa^2 x^2}{32\sigma^2(1-\varrho^2)}t\right\}.$$

We conclude that for $x \in \mathbb{R} \setminus \{0\}$, $\xi \in [-1,1]$ and $t \in \mathbb{R}_{++}$, we have

$$\left|\partial_1 g\left(x + \xi\frac{x}{2}, t\right)\right| \leqslant G(t,x)$$

with

$$G(t,x) := \frac{\kappa}{e\pi\sigma\sqrt{1-\varrho^2}} \mathbb{1}_{[0, T_0(x)]}(t) + \frac{\kappa}{2\pi\sigma\sqrt{1-\varrho^2}\,T_0(x)} \exp\left\{-\frac{\kappa^2 x^2}{32\sigma^2(1-\varrho^2)}t\right\} \mathbb{1}_{(T_0(x), \infty)}(t)$$

for $t \in \mathbb{R}_{++}$ and $x \in \mathbb{R} \setminus \{0\}$, and the function $\mathbb{R}_+ \ni t \mapsto G(t,x)$ is integrable on $\mathbb{R}_+$, hence the dominated convergence theorem can be used, and we obtain (D.1).

One can derive (D.3) in another way. Since $g(x,t) \leqslant \frac{1}{\sqrt{2\pi t^3}} e^{-\frac{1}{2t}}$, $x \in \mathbb{R}$, $t \in \mathbb{R}_{++}$, the improper integral $\int_0^\infty g(x,t)\, \mathrm{d}t$ is uniformly convergent for $x \in \mathbb{R}$. Further, for any $t, a, A \in \mathbb{R}_{++}$ and $x \in \mathbb{R}$ with $a < |x| < A$, we have

$$G(t,x) \leqslant \frac{\kappa}{e\pi\sigma\sqrt{1-\varrho^2}} \mathbb{1}_{[0, T_0(a)]}(t) + \frac{\kappa}{2\pi\sigma\sqrt{1-\varrho^2}\,T_0(A)} \exp\left\{-\frac{\kappa^2 a^2}{32\sigma^2(1-\varrho^2)}t\right\} \mathbb{1}_{(T_0(A), \infty)}(t)$$

47

showing that the improper integral $\int_0^\infty \partial_1 g(x,t)\,\mathrm{d}t$ is uniformly convergent for $a < |x| < A$ with $a < A$, $a, A \in \mathbb{R}_{++}$. This together with the continuity of the functions $\mathbb{R} \times \mathbb{R}_{++} \ni (x,t) \mapsto g(x,t) \in \mathbb{R}$ and $\mathbb{R} \times \mathbb{R}_{++} \ni (x,t) \mapsto \partial_1 g(x,t) \in \mathbb{R}$ yield (D.3), see, e.g., Lang [28, pages 337-339].

Moreover, for all $x \in \mathbb{R} \setminus \{0\}$, we have

$$f(x) \geqslant \frac{\kappa}{2\pi\sigma\sqrt{1-\varrho^2}} \exp\left\{-\frac{1}{2} + \frac{2\varrho\sigma\kappa x}{2\sigma^2(1-\varrho^2)} - \frac{\varrho^2\sigma^2}{2\sigma^2(1-\varrho^2)}\right\} \int_1^\infty \frac{1}{t} \exp\left\{-\frac{\kappa^2 x^2 t}{2\sigma^2(1-\varrho^2)}\right\} \mathrm{d}t,$$

where

$$\lim_{x \to 0} \exp\left\{-\frac{1}{2} + \frac{2\varrho\sigma\kappa x}{2\sigma^2(1-\varrho^2)} - \frac{\varrho^2\sigma^2}{2\sigma^2(1-\varrho^2)}\right\} = \exp\left\{-\frac{1}{2} - \frac{\varrho^2\sigma^2}{2\sigma^2(1-\varrho^2)}\right\}$$

and

$$\int_1^\infty \frac{1}{t} \exp\left\{-\frac{\kappa^2 x^2 t}{2\sigma^2(1-\varrho^2)}\right\} \mathrm{d}t = \int_{x^2}^\infty \frac{1}{u} \exp\left\{-\frac{\kappa^2 u}{2\sigma^2(1-\varrho^2)}\right\} \mathrm{d}u \to \infty \qquad \text{as } x \to 0,$$

hence we obtain (D.2).

Acknowledgements

References

[1] Alfonsi, A., (2005). On the discretization schemes for the CIR (and Bessel squared) processes. *Monte Carlo Methods and Applications* **11(4)** 355–384.

[2] Alfonsi, A. (2015). *Affine Diffusions and Related Processes: Simulation, Theory and Applications.* Springer, Cham, Bocconi University Press, Milan.

[3] Bakshi, G., Cao, C. and Chen, Z. (1997). Empirical performance of alternative option pricing models. *The Journal of Finance* **LII(5)** 2003–2049.

[4] Barczy, M., Döring, L., Li, Z. and Pap, G. (2013). On parameter estimation for critical affine processes. *Electronic Journal of Statistics* **7** 647–696.

[5] Barczy, M., Döring, L., Li, Z. and Pap, G. (2014). Stationarity and ergodicity for an affine two factor model. *Advances in Applied Probability* **46(3)** 878–898.

[6] Barczy, M. and Pap, G. (2016). Asymptotic properties of maximum likelihood estimators for Heston models based on continuous time observations. *Statistics* **50(2)** 389–417.

[7] Bates, D. S. (1996). Jumps and stochastic volatility: exchange rate processes implicit in Deutsche mark options. *The Review of Financial Studies* **9(1)** 69–107.

[8] Ben Alaya, M. and Kebaier, A. (2012). Parameter estimation for the square root diffusions: ergodic and nonergodic cases. *Stochastic Models* **28(4)** 609–634.

[9] Ben Alaya, M. and Kebaier, A. (2013). Asymptotic behavior of the maximum likelihood estimator for ergodic and nonergodic square-root diffusions. *Stochastic Analysis and Applications* **31(4)** 552–573.

[10] Berkaoui, A., Bossy, M., and Diop, A. (2008). Euler scheme for SDEs with non-Lipschitz diffusion coefficient: strong convergence. *ESAIM. Probability and Statistics* **12** 1–11 (electronic).

[11] Billingsley, P. (1999). *Convergence of probability measures*, 2nd ed. John Wiley & Sons, Inc., New York.

[12] Broadie, M. and Kaya, Ö. (2006). Exact simulation of stochastic volatility and other affine jump diffusion processes. *Operations Research* **54(2)** 217–231.

[13] Cox, J. C., Ingersoll, J. E. and Ross, S. A. (1985). A theory of the term structure of interest rates. *Econometrica* **53(2)** 385–407.

[14] Deelstra, G. and Delbaen, F. (1998). Convergence of discretized stochastic (interest rate) processes with stochastic drift term. *Applied Stochastic Models and Data Analysis* **14(1)** 77–84.

[15] Diop, A. (2003). *Sur la discretisation et le comportement a petit bruit d'EDS multidimensionnelles dont les coefficients sont a derivees singulieres.* Ph.D Thesis, INRIA, France.

[16] Dudley, R. M. (1989). *Real Analysis and Probability.* Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, California.

[17] Feller, W. (1951). Two Singular Diffusion Problems. *Annals of Mathematics* **54(1)** 173–182.

[18] Heston, S. (1993). A closed-form solution for options with stochastic volatilities with applications to bond and currency options. *The Review of Financial Studies* **6** 327–343.

[19] Ikeda, N. and Watanabe, S. (1981). *Stochastic Differential Equations and Diffusion Processes.* North-Holland Publishing Company.

[20] Jacod, J. and Mémin, J. (1976). Caractéristiques locales et conditions de continuité absolue pour les semi-martingales. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete,* **35** 1–37.

[21] Jacod, J. and Protter, P. (2012). *Discretization of processes,* Springer, Heidelberg.

[22] JACOD, J. and SHIRYAEV, A. N. (2003). *Limit Theorems for Stochastic Processes*, 2nd ed. Springer-Verlag, Berlin.

[23] JASCHKE, S. (2003). A note on the inhomogeneous linear stochastic differential equation. *Insurance: Mathematics & Economics* **32(3)** 461–464.

[24] KALLENBERG, O. (1997). *Foundations of Modern Probability.* Springer, New York, Berlin, Heidelberg.

[25] KARATZAS, I. and SHREVE, S. E. (1991). *Brownian Motion and Stochastic Calculus,* 2nd ed. Springer-Verlag, New York.

[26] KLOEDEN, P. E. and PLATEN, E. (1995). *Numerical Methods for Stochastic Differential Equations.* CRC, Boca Raton, FL.

[27] LAMBERTON, D. and LAPEYRE, B. (1996). *Introduction to Stochastic Calculus Applied to Finance.* Chapman and Hall CRC Press.

[28] LANG, S. (1997). *Undergraduate Analysis,* 2nd ed. Springer, New York.

[29] LI, Z. and MA, C. (2015). Asymptotic properties of estimators in a stable Cox–Ingersoll–Ross model. *Stochastic Processes and their Applications* **125(8)** 3196–3233.

[30] LIPTSER, R. S. and SHIRYAEV, A. N. (2001). *Statistics of Random Processes II. Applications*, 2nd edition. Springer-Verlag, Berlin, Heidelberg.

[31] LUSCHGY, H. (1992). Local asymptotic mixed normality for semimartingale experiments. *Probability Theory and Related Fields* **92** 151–176.

[32] LUSCHGY, H. (1994). Asymptotic inference for semimartingale models with singular parameter points. *Journal of Statistical Planning and Inference* **39** 155–186.

[33] OVERBECK, L. (1998). Estimation for continuous branching processes. *Scandinavian Journal of Statistics* **25(1)** 111–126.

[34] RUNGGALDIER, W. J. (2003). *Jump-diffusion models.* In: Handbook of Heavy Tailed Distributions in Finance, Volume 1: Handbooks in Finance edited by S. T. Rachev, Elsevier Science.

[35] SATO, K.-I. (1999). *Lévy Processes and Infinitely Divisible Distributions.* Cambridge University Press, Cambridge.

[36] SØRENSEN, M. (1991). Likelihood methods for diffusions with jumps. In: N. U. Prabhu and I.V. Basawa, Eds., *Statistical Inference in Stochastic Processes,* Marcel Dekker, New York, 67–105.

[37] SUN, Y., LIU, C. and GUO, S. (2017). Stochastic volatility double-jump-diffusions model: The importance of distribution type of jump amplitude. *International Journal of Computer Mathematics* **94(5)** 989–1014.

[38] VAN DER VAART, A. W. (1998). *Asymptotic Statistics*, Cambridge University Press.

[39] VON WEIZSÄCKER, H. and WINKLER, G. (1990). *Stochastic integrals. An introduction*, Advanced Lectures in Mathematics. Friedr. Vieweg & Sohn, Braunschweig.

[40] VAN ZANTEN, H. (2000). A multivariate central limit theorem for continuous local martingales. *Statistics & Probability Letters* **50(3)** 229–235.

[41] ZHANG, L., MYKLAND, P. A. and AÏT-SAHALIA, Y. (2005). A tale of two time scales: Determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association* **100(472)** 1394–1411.