DOI: 10.28919/ejma.2023.3.22

Eur. J. Math. Appl. (2023)3:22
URL: http://ejma.euap.org

RATE OF CONVERGENCE IN THE KOLMOGOROV DISTANCE FOR THE MINIMUM CONTRAST ESTIMATOR IN THE HESTON MODEL

JAYA P. N. BISHWAL

ABSTRACT. We develop a new explicit estimator of the mean reversion parameter in the Heston model by using the minimum contrast method. We obtain a bound on the Kolmogorov distance for the distribution of the approximate minimum contrast estimator and the normal distribution for high frequency data.

1. Introduction

Due to the availability of high frequency market price data of stocks, currencies, and other financial instruments, statistics of high frequency data has seen a revolution recently. One of the fundamental problems is the estimation of integrated volatility in the statistics of high frequency data. Hence realized volatility which is a measure of the integrated volatility has received considerable interest in recent days empirical finance. The realized volatility is defined as the sum of squared increments of returns, which is basically quadratic variation of log-prices. In Heston stochastic volatility model (see Heston [1]), the unknown parameters are present in the unobserved volatility process. We focus on the estimation of the mean reversion parameter based on the high frequency log-price data. Barndorff-Nielsen and Shephard [2–4] also studied stochastic volatility models where the unobserved stochastic volatility process is Ornstein-Uhlenbeck type driven by positive Levy processes. Woerner [5] studied estimation of integrated volatility in stochastic volatility models. Jacod [6] studied asymptotic properties of realized power variations for semimartingales. Jacod and Reiss [7] studied the rates of convergence for integrated volatility estimation in the presence of semimartingales. Parameter estimation in stochastic differential equations from direct observations is studied in Bishwal [8]. Parameter estimation in stochastic volatility models from partial observations is extensively studied in see Bishwal [9].

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, P)$ be a stochastic basis on which the Cox-Ingersoll-Ross (CIR) process $\{X_t\}$ is defined satisfying the Itô stochastic differential equation

$$dX_t = (1 - 2\theta X_t)dt + 2\sqrt{X_t}dW_t,\ t \geq 0 \tag{1.1}$$

Department of Mathematics and Statistics, University of North Carolina at Charlotte, 376 Fretwell Bldg., 9201 University City Blvd., Charlotte, NC 28223, USA

E-mail address: J.Bishwal@uncc.edu.

Key words and phrases. Itô stochastic differential equation; Heston model; discrete observations; approximate minimum contrast estimators; robustness; efficiency; Fourier method; Berry-Esseen bound; Kolmogorov distance.

Received 11/09/2023.

where $\{W_t, t \geq 0\}$ is a standard Wiener process with the filtration $\{\mathcal{F}_t\}_{t\geq 0}$ and consider the classical direct estimation problem where $\theta > 0$ is the unknown parameter to be estimated on the basis of discrete observations of the process $\{X_t\}$ at times $0 = t_0 < t_1 < \cdots t_n = T$ with $t_i - t_{i-1} = \frac{T}{n}$, $i = 1, 2 \cdots, n$. For our asymptotic framework, we assume two types of high frequency data with long observation time: 1) $T \to \infty, n \to \infty, \frac{T}{\sqrt{n}} \to 0$, 2) $T \to \infty, n \to \infty, \frac{T}{n^{2/3}} \to 0$. Recall that in the standard definition of CIR process, there is a mean reversion level parameter α, mean reversion speed parameter β and a volatility of volatility parameter σ. However, first we consider the simplified model (1.1) in the beginning in order to keep the presentation simple.

For the moment assume that a continuous realization $\{X_t, 0 \leq t \leq T\}$ be denoted by X_0^T. Let P_θ^T be the measure generated on the space (C_T, B_T) of continuous functions on $[0, T]$ with the associated Borel σ-algebra B_T generated under the supremum norm by the process X_0^T and let P_0^T be the standard Wiener measure. It is well known that when θ is the true value of the parameter P_θ^T is absolutely continuous with respect to P_0^T and the Radon-Nikodym derivative (likelihood) of P_θ^T with respect to P_0^T based on X_0^T is given by

$$L_T(\theta) := \frac{dP_\theta^T}{dP_0^T}(X_0^T) = \exp\left\{-\theta \int_0^T dX_t - \frac{\theta^2}{2}\int_0^T X_t dt\right\}. \tag{1.2}$$

Consider the score function, the derivative of the log-likelihood function, which is given by

$$\gamma_T(\theta) := -\int_0^T dX_t - \theta \int_0^T X_t dt. \tag{1.3}$$

A solution of the estimating equation $\gamma_T(\theta) = 0$ provides the maximum likelihood estimate (MLE)

$$\hat{\theta}_T := \frac{-X_T + X_0 + T/2}{\int_0^T X_t dt}. \tag{1.4}$$

Minimum contrast estimator (MCE) is an alternative to the maximum likelihood estimator which does not involve the stochastic integral and hence is easier for simulation. It preserves similar asymptotic properties of the MLE. The popular M-estimator is reduced to the minimum contrast estimator, see Bishwal [8]. As far as we know, rate of normal approximation in the Kolmogorov distance for the minimum contrast estimator has not been studied earlier. Our aim in this paper is to bridge this gap. Consider the minimum contrast estimate (MCE)

$$\theta_T := \frac{T/2}{\int_0^T X_t dt}. \tag{1.5}$$

Note that the volatility which is given by the CIR process is not observed. In the following section we obtain nonparametric estimator of the minimum contrast estimator of the mean reversion parameter in the Heston model using approximations to θ_T defined in (1.5).

2. Approximate Minimum Contrast Estimator

Consider the Heston stochastic volatility model

$$dS_t = \mu S_t dt + \sqrt{X_t} S_t dW_t, \tag{2.1}$$

$$dX_t = (1 - 2\theta X_t)\, dt + 2\sigma\sqrt{X_t} dZ_t, \tag{2.2}$$

where $\{W_t\}$, a standard Brownian motion, is independent of another standard Brownian motion $\{Z_t\}$ and $\theta > 0$. We assume that $X_0 > 0$. The process X is strictly positive and never hits zero. Here θ corresponds to the speed of adjustment, $1/\theta$ is called the mean and σ is called the volatility of volatility. It is well known that the X process is ergodic and has a stationary distribution. The distribution of its future value given the current is non-central chi-square and the distribution of the limit value is gamma. The integrated volatility is given by

$$I_T := \int_0^T X_t dt. \tag{2.3}$$

The process I_T which is the integrated volatility (energy) of the CIR process which plays a important role in clustering time or activity persistence in stochastic volatility modeling.

The estimator of $\theta_{n,T}$ based on discrete observations of the process $\{S_t\}$ at times $0 = t_0 < t_1 < \cdots t_n = T$ is approximate minimum contrast estimator given by

$$\theta_{n,T} = \frac{T}{2\widehat{I}_T} \tag{2.4}$$

where $\widehat{I}_T$ is nonparametric estimator of I_T based on discrete observations of the process $\{S_t\}$ at times $0 = t_0 < t_1 < \cdots t_n = T$.

Integrated volatility has to be estimated on the basis of discrete observations of the process $\{S_t\}$ at times $0 = t_0 < t_1 < \cdots t_n = T$ with $t_i - t_{i-1} = \frac{T}{n}$, $i = 1, 2 \cdots, n$. Denote

$$\Delta S_{t_{i-1}} := S_{t_i} - S_{t_{i-1}}. \tag{2.5}$$

The realized volatility or "approximate quadratic variation" is defined as

$$R_{n,T} := \sum_{i=1}^{n} (\Delta S_{t_{i-1}})^2. \tag{2.6}$$

It is well known from Barndorff-Nielsen and Shephard [2] that

$$\operatorname*{P\text{-}lim}_{n\to\infty} R_{n,T} = I_T \tag{2.7}$$

where $P - lim$ stands for convergence in probability. and the stable convergence in law at the rate $\sqrt{n}$ holds, see Barndorff-Nielsen and Shephard [3]:

$$\sqrt{n}(R_{n,T} - I_T) \to^{\mathcal{D}-s} \sqrt{2} \int_0^T X_s dW'_s \quad \text{as} \quad n \to \infty \tag{2.8}$$

where $\mathcal{D} - s$ stands for convergence in distribution stably.

where W' is a Brownian motion defined on an extension of the probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, P)$, and which is independent of the σ-field $\mathcal{F}$. Now we define the approximate minimum contrast estimator (AMCE) of θ as

$$\theta_{n,T} := \frac{T}{2R_{n,T}} = \frac{T}{2\sum_{i=1}^{n}(S_{t_i} - S_{t_{i-1}})^2}. \tag{2.9}$$

3. Rates of Weak Convergence in the Kolmogorov Distance

For our asymptotic framework, we assume two types of high frequency data with long observation time: 1) $T \to \infty, n \to \infty, \frac{T}{\sqrt{n}} \to 0$, 2) $T \to \infty, n \to \infty, \frac{T}{n^{2/3}} \to 0$.

Observe that

$$\left(\frac{T}{2\theta}\right)^{1/2}(\tilde{\theta}_T-\theta)=\frac{\left(\frac{2\theta}{T}\right)^{1/2}M_T}{\left(\frac{2\theta}{T}\right)I_T} \tag{3.1}$$

where

$$M_T:=\frac{T}{2}-\theta I_T \quad \text{and} \quad I_T:=\int_0^T X_t dt. \tag{3.2}$$

We have by Itô's formula, $\sqrt{X_t}$ is an Ornstein-Uhlenbeck process satisfying

$$d(\sqrt{X_t})=\theta\sqrt{X_t}dt+dW_t \tag{3.3}$$

since

$$\begin{aligned} dX_t &= d(\sqrt{X_t})^2=2\sqrt{X_t}d\sqrt{X_t}+d(\sqrt{X_t})^2=2\theta X_t dt+2X_t dW_t+dt\\ &=2\theta(\frac{1}{2\theta}-X_t)dt+2\sqrt{X_t}dW_t=(1-2\theta X_t dt)+2\sqrt{X_t}dW_t. \end{aligned}$$

Denote $U_t:=\sqrt{X_t}$. Thus

$$dU_t=\theta U_t dt+dW_t,\ t\geq 0. \tag{3.4}$$

Hence

$$U_t=\int_0^t e^{\theta(t-u)}dW_u,\ t\geq 0.$$

Introduce the following notations :

$$\begin{aligned} &Y_{n,T}:=\sum_{i=1}^n U_{t_{i-1}}\left[W_{t_i}-W(t_{i-1})\right],\ Y_T:=\int_0^T U_t dW_t,\\ &Z_{n,T}:=\sum_{i=1}^n U_{t_{i-1}}\left[U_{t_i}-U(t_{i-1})\right],\ Z_T:=\int_0^T U_t dU_t,\\ &I_{n,T}:=\sum_{i=1}^n X_{t_{i-1}}(t_i-t_{i-1}),\ V_{n,T}:=\sum_{i=1}^n\int_{t_{i-1}}^{t_i}U_{t_{i-1}}\left[U_t-U_{t_{i-1}}\right]dt. \end{aligned}$$

In this paper, $O_P(\delta_n)$ are random variables which are bounded in probability of the order δ_n, also called stochastically bounded of the order δ_n.

We need the following lemma in the sequel.

Lemma 3.2 (a) Let

$$\Xi_{T,x}:=\left(\frac{2\theta}{T}\right)^{1/2}M_T-\left(\frac{2\theta}{T}I_T-1\right)x.$$

Then for $|x|\leq 2(\log T)^{1/2}$ and for $|u|\leq\epsilon T^{1/2}$, where ϵ is sufficiently small

$$\left|E\exp(iu\Xi_{T,x})-\exp(\frac{-u^2}{2})\right|\leq C\exp(\frac{-u^2}{4})(|u|+|u|^3)T^{-1/2}.$$

(b) $\sup_{x\in\mathbb{R}}\left|P\left\{\left(\frac{2\theta}{T}\right)^{1/2}M_T\leq x\right\}-\Phi(x)\right|\leq CT^{-1/2}.$

Proof : From Pitman and Yor [15], the characteristic function of I_T, closely associated with Levy's stochastic area formula, is given by

$$E\exp(iuI_T)=\exp\left(x\frac{2iu}{\theta+\gamma\coth(\frac{\gamma T}{2})}\right)\frac{\exp(\frac{\theta T}{2})}{\cosh(\frac{\gamma T}{2})+\frac{\theta}{\gamma}\sinh(\frac{\gamma T}{2})}$$

where $u \in \mathbb{R}$, $\gamma := \sqrt{\theta^2 - 4iu}$ and $X_0 = x$. Alternatively, this can be written as

$$E \exp(iuI_T) = \exp\left(\frac{2iu(e^{\frac{\gamma T}{2}} - e^{\frac{-\gamma T}{2}})}{e^{\frac{\gamma T}{2}}(\gamma+\theta) + e^{-\frac{\gamma T}{2}}(\gamma-\theta)}\right) 2\left[e^{\frac{\gamma T}{2}}(\gamma+\theta) + e^{-\frac{\gamma T}{2}}(\gamma-\theta)\right]^{-1}.$$

Now consider

$$\begin{aligned} & E\exp(iu\Xi_{T,x}) \\ = \; & E\exp\left[iu\left(\tfrac{2\theta}{T}\right)^{1/2} M_T - iu\left(\tfrac{2\theta}{T}I_T - 1\right)x\right] \\ = \; & E\exp\left[-iu\left(\tfrac{2\theta}{T}\right)^{1/2}\{\theta I_T - 1\} - iu\left(\tfrac{2\theta}{T}I_T - 1\right)x\right] \\ = \; & E\exp(z_1 I_T + z_3) \\ = \; & \exp(z_3)\phi_T(z_1) \end{aligned} \tag{3.5}$$

where $z_1 := -iu\theta\delta_{T,x}$, and $z_3 := \frac{iuT}{2}\delta_{T,x}$ with $\delta_{T,x} := \left(\frac{2\theta}{T}\right)^{1/2} + \frac{2x}{T}$. Note that $\phi_T(z_1)$ satisfies the conditions of (a) by choosing ϵ sufficiently small. Let $\omega_{1,T}(u)$, $\omega_{2,T}(u)$, $\omega_{3,T}(u)$ and $\omega_{4,T}(u)$ be functions which are $O(|u|T^{-1/2}), O(|u|^2T^{-1/2})$, $O(|u|^3T^{-3/2})$ and $O(|u|^3T^{-1/2})$ respectively. Note that for the given range of values of x and u, the conditions on z_1 of the Lemma are satisfied. Further, with

$$\varpi_T(u) := 1 + iu\frac{\delta_{T,x}}{\beta} + \frac{u^2\delta_{T,x}^2}{2\beta^2},$$

we obtain

$$\begin{aligned} \gamma &= (\beta^2 - 2z_1)^{1/2} \\ &= \theta\left[1 - \frac{z_1}{\theta^2} - \frac{z_1^2}{2\theta^4} + \frac{z_1^3}{2\theta^8} + \cdots\right] \\ &= \theta\left[1 + iu\frac{\delta_{T,x}}{\theta} + \frac{u^2\delta_{T,x}^2}{2\theta^2} + \frac{iu^3\delta_{T,x}^3}{2\theta^3} + \cdots\right] \\ &= \beta[1 + \omega_{1,T}(u) + \omega_{2,T}(u) + \omega_{3,T}(u)] \\ &= \beta\varpi_T(u) + \omega_{3,T}(u) \\ &= \beta[1 + \omega_{1,T}(u)]. \end{aligned} \tag{3.6}$$

Thus

$$\gamma - \theta = \omega_{1,T}, \;\; \gamma + \theta = 2\beta + \omega_{1,T}. \tag{3.7}$$

Hence the above expectation equals

$$\begin{aligned} & \exp\left(z_3 + \frac{\theta T}{2}\right)\left[\frac{2\theta\varpi_T(u) + \omega_{3,T}(u)}{\omega_{1,T}\exp\{-\theta T\varpi_T(u) + \omega_{4,T}(u)\} + (2\theta + \omega_{1,T}(u))\exp\{\theta T\varpi_T(u) + \omega_{4,T}(u)\}}\right]^{1/2} \\ = & \left[\frac{1 + \omega_{1,T}(u)}{\omega_{1,T}\exp(\chi_T(u)) + (1 + \omega_{1,T}(u))\exp(\psi_T(t))}\right]^{1/2} \end{aligned} \tag{3.8}$$

where

$$\begin{aligned} \chi_T(u) &:= -\theta T\beta_T(u) + \alpha_{4,T}(u) - 2z_3 - \theta T \\ &= -2\theta T + \omega_{1,T}(u) + t^2\omega_{1,T}(u). \end{aligned} \tag{3.9}$$

and

$$\begin{aligned} \psi_T(u) &:= \theta T\varpi_T(u) + \omega_{4,T}(u) - 2z_3 - \theta T \\ &= \theta T\left[1 + iu\frac{\delta_{T,x}}{\theta} + \frac{u^2\delta_{T,x}^2}{2\theta^2}\right] + \alpha_{4,T}(u) - iuT\delta_{T,x} - \theta T \\ &= \frac{u^2T}{2\theta}\left[\left(\frac{2\theta}{T}\right)^{1/2} + \frac{2x}{T}\right]^2 \\ &= u^2 + u^2\omega_{1,T}(u). \end{aligned} \tag{3.10}$$

Hence, for the given range of values of u, $\chi_T(u) - \psi_T(u) \leq -\theta T$.

Hence the above expectation equals

$$\begin{aligned} & \exp(-\frac{u^2}{2})(1+\omega_{1,T})^{1/2} \\ & \times\left[\omega_{1,T}\exp\{-2\theta T+\omega_{1,T}+u^2\omega_{1,T}\}+(1+\omega_{1,T}(u))\exp\{u^2\omega_{1,T}(u)\}\right]^{-1/2} \\ = \; & \exp(-\tfrac{u^2}{2})\left[1+\omega_{1,T})(1+\omega_{1,T}(1+\omega_{1,T})\exp\{-\theta T+\omega_{1,T}+t^2\omega_{1,T}\}\right]\exp(u^2\omega_{1,T}(u)). \end{aligned} \tag{3.11}$$

Applying part (a) along with Esseen's smoothing lemma (see Petov [10] or Feller [11]) yields part (b). This completes the proof of the lemma. □

Lemma 3.3 We have

$$\text{(a)}\quad E\left|Y_{n,T}-Y_T\right|^2=O\left(\frac{T^2}{n}\right),$$

$$\text{(b)}\quad E\left|Z_{n,T}-Z_T\right|^2=O\left(\frac{T^2}{n}\right),$$

$$\text{(c)}\quad E\left|I_{n,T}-I_T\right|^2=O\left(\frac{T^4}{n^2}\right),$$

$$\text{(d)}\quad E\left|R_{n,T}-I_T\right|^2=O\left(\frac{T^4}{n^2}\right).$$

Proof: See Appendix.

The following theorem gives the bound on the error of normal approximation of the AMCEs. Note that part (a) uses parameter dependent nonrandom norming. While this is useful for testing hypotheses about θ, it may not necessarily give a confidence interval. The normings in parts (b) and (c) are sample dependent which can be used for obtaining a confidence interval. Following theorem shows that asymptotic normality of the AMCEs need $T\to\infty$ and $\frac{T}{\sqrt{n}}\to 0$.

Theorem 3.1 Let $\delta_{n,T}=T^{-1/2}(\log T)^{1/2}\bigvee\frac{T^2}{n}(\log T)^{-1}$. We have,

$$\text{(a)}\ \sup_{x\in\mathbb{R}}\left|P\left\{\left(\frac{T}{-2\theta}\right)^{1/2}(\theta_{n,T}-\theta)\le x\right\}-\Phi(x)\right|=O(\delta_{n,T}).$$

$$\text{(b)}\ \sup_{x\in\mathbb{R}}\left|P\left\{R_{n,T}^{1/2}(\theta_{n,T}-\theta)\le x\right\}-\Phi(x)\right|=O(\delta_{n,T}).$$

$$\text{(c)}\ \sup_{x\in\mathbb{R}}\left|P\left\{\left(\frac{T}{|2\theta_{n,T}|}\right)^{1/2}(\theta_{n,T}-\theta)\le x\right\}-\Phi(x)\right|=O(\delta_{n,T}).$$

Proof (a) From (2.9) and (3.2), we have

$$R_{n,T}\theta_{n,T}=\frac{T}{2}=M_T+\theta I_T. \tag{3.12}$$

Hence using (3.1)

$$\begin{aligned} \left(\frac{T}{2\theta}\right)^{1/2}(\theta_{n,T}-\theta) &= \frac{\left(-\frac{T}{2\theta}\right)^{1/2}M_T+\theta\left(\frac{T}{2\theta}\right)^{1/2}(I_T-R_{n,T})}{R_{n,T}} \\ = \; & \frac{\left(\frac{2\theta}{T}\right)^{1/2}M_T+\left(-\frac{2\theta}{T}\right)^{1/2}(I_T-R_{n,T})}{\left(\frac{2\theta}{T}\right)R_{n,T}}. \end{aligned} \tag{3.13}$$

Further, using Lemma 2.2 (a) of Bishwal and Bose [14] and Lemma 3.3 (d), we have

$$\begin{aligned}
& P\left\{\left|\left(\frac{2\theta}{T}\right)(R_{n,T}-1)\right|>\epsilon\right\} = \left\{\left|\left(\frac{2\theta}{T}\right)(R_{n,T}-I_T+I_T)-1\right|>\epsilon\right\} \\
\leq\ & P\left\{\left|\left(\frac{2\theta}{T}\right)I_T-1\right|>\frac{\epsilon}{2}\right\}+P\left\{\left(\frac{2\theta}{T}\right)|R_{n,T}-I_T|>\frac{\epsilon}{2}\right\} \\
\leq\ & C\exp\left(\frac{-T\theta}{16}\epsilon^2\right)+\frac{16\theta^2}{T^2}\frac{E|R_{n,T}-I_T|^2}{\epsilon^2} \\
\leq\ & C\exp\left(\frac{-T\theta}{16}\epsilon^2\right)+C\frac{T^2/n^2}{\epsilon^2}. \qquad (3.14)
\end{aligned}$$

Next, observe that

$$\begin{aligned}
& \sup_{x\in\mathbb{R}}\left|P\left\{\left(\frac{T}{2\theta}\right)^{1/2}(\theta_{n,T}-\theta)\leq x\right\}-\Phi(x)\right| \\
=\ & \sup_{x\in\mathbb{R}}\left|P\left\{\frac{\left(\frac{2\theta}{T}\right)^{1/2}M_T+\left(\frac{2\theta}{T}\right)^{1/2}(I_T-R_{n,T})}{\left(\frac{2\theta}{T}\right)R_{n,T}}\leq x\right\}-\Phi(x)\right| \\
\leq\ & \sup_{x\in\mathbb{R}}\left|P\left\{\left(\frac{2\theta}{T}\right)^{1/2}M_T\leq x\right\}-\Phi(x)\right|+P\left\{\left|\theta\left(\frac{2\theta}{T}\right)^{1/2}(R_{n,T}-I_T)\right|>\epsilon\right\} \\
& +P\left\{\left|\left(\frac{2\theta}{T}\right)R_{n,T}-1\right|>\epsilon\right\}+2\epsilon \\
\leq\ & CT^{-1/2}+\theta^2\frac{\left(\frac{2\theta}{T}\right)E|R_{n,T}-I_T|^2}{\epsilon^2}+C\exp\left(-\frac{T\theta}{4}\epsilon^2\right)+C\frac{T^2}{n^2\epsilon^2}+2\epsilon, \qquad (3.15)
\end{aligned}$$

(the bound for the first term in the right hand side of (3.15) comes from Lemma 3.2(b) and that for the 3rd term is obtained from (3.14))

$$\leq\ CT^{-1/2}+C\frac{T^2}{n^2\epsilon^2}+C\exp\left(-\frac{T\theta}{4}\epsilon^2\right)+C\frac{T}{n^2\epsilon^2}+\epsilon \qquad (3.16)$$

(by Lemma 3.3(d)).

Choosing $\epsilon = CT^{-1/2}(\log T)^{1/2}$, the terms in the right hand side of (3.16) are of the order $O(\max(T^{-1/2}(\log T)^{1/2}, (\frac{T^4}{n^2})(\log T)^{-1}))$.

(b) Using (3.1), we have

$$R_{n,T}^{1/2}(\theta_{n,T}-\theta)=\frac{M_T+\theta(I_T-R_{n,T})}{R_{n,T}^{1/2}}.$$

Then,

$$\begin{aligned}
& \sup_{x\in\mathbb{R}}\left|P\left\{R_{n,T}^{1/2}(\theta_{n,T}-\theta)\leq x\right\}-\Phi(x)\right| = \sup_{x\in\mathbb{R}}\left|P\left\{\frac{M_T}{R_{n,T}^{1/2}}+\theta\frac{I_T-R_{n,T}}{R_{n,T}^{1/2}}\leq x\right\}-\Phi(x)\right| \\
\leq\ & \sup_{x\in\mathbb{R}}\left|P\left\{\frac{M_T}{I_{n,T}^{1/2}}\leq x\right\}-\Phi(x)\right|+P\left\{\left|\frac{\theta(I_T-R_{n,T})}{R_{n,T}^{1/2}}\right|>\epsilon\right\}+\epsilon \\
=:\ & U_1+U_2+\epsilon. \qquad (3.17)
\end{aligned}$$

We have from (3.14),

$$U_1\leq CT^{-1/2}+C\exp\left(\frac{T\theta}{16}\epsilon^2\right)+C\frac{T^2}{n^2\epsilon^2}+\epsilon. \qquad (3.18)$$

Now,

$$\begin{aligned} U_2 &= P\left\{|\theta|\left|\frac{R_{n,T}-I_T}{R_{n,T}^{1/2}}\right| > \epsilon\right\} = P\left\{|\theta|\frac{\left|\left(-\frac{2\theta}{T}\right)^{1/2}(R_{n,T}-I_T)\right|}{\left|\left(-\frac{2\theta}{T}\right)^{1/2}R_{n,T}^{1/2}\right|} > \epsilon\right\} \\ &\le P\left\{\left|\left(-\frac{2\theta}{T}\right)^{1/2}\right||R_{n,T}-I_T| > \delta\right\} + P\left\{\left|\left(-\frac{2\theta}{T}\right)^{1/2}R_{n,T}^{1/2}-1\right| > \delta_1\right\} \qquad (3.19) \\ &\quad \text{(where } \delta = \epsilon - C\epsilon^2 \text{ and } \delta_1 = (\epsilon-\delta)/\epsilon > 0) \\ &\le \left(-\frac{2\theta}{T}\right)\frac{E|R_{n,T}-I_T|^2}{\delta^2} + P\left\{\left|\left(-\frac{2\theta}{T}\right)R_{n,T}-1\right| > \delta_1\right\} \\ &\le C\frac{T^3}{n^2\delta^2} + C\exp\left(\frac{T\theta}{16}\delta_1^2\right) + C\frac{T^2}{n^2\delta_1^2}. \qquad (3.20) \end{aligned}$$

Here, the bound for the first term in the right hand side of (3.20) comes from Lemma 3.4(d) and that for the second term is obtained from (3.14).

Now, using the bounds (3.18) and (3.20) in (3.17) with $\epsilon = CT^{-1/2}(\log T)^{1/2}$, we obtain that the terms in (3.17) are of the order $O(\max(T^{-1/2}(\log T)^{1/2}, (\frac{T^4}{n^2})(\log T)^{-1}))$.

(c) Let $G_T := \{|\theta_{n,T}-\theta| \le d_T\}$, and $d_T := CT^{-1/2}(\log T)^{1/2}$.

On the set G_T, expanding $(2|\theta_{n,T}|)^{1/2}$, we obtain

$$(2\theta_{n,T})^{-1/2} = (2\theta)^{1/2}\left[1-\frac{\theta-\theta_{n,T}}{\theta}\right]^{-1/2} = (2\theta)^{1/2}\left[1+\frac{1}{2}\left(\frac{\theta-\theta_{n,T}}{\theta}\right)\right] + O(d_T^2).$$

Then,

$$\begin{aligned} &\sup_{x\in\mathbb{R}}\left|P\left\{\left(\frac{T}{2|\theta_{n,T}|}\right)^{1/2}(\theta_{n,T}-\theta) \le x\right\} - \Phi(x)\right| \\ &\le \sup_{x\in\mathbb{R}}\left\{P\left(\frac{T}{2|\theta_{n,T}|}\right)^{1/2}(\theta_{n,T}-\theta) \le x, G_T\right\} + P(G_T^c). \end{aligned}$$

Further,

$$\begin{aligned} P(G_T^c) &= P\left\{|\theta_{n,T}-\theta| > CT^{-1/2}(\log T)^{1/2}\right\} \\ &= P\left\{\left(\frac{T}{2\theta}\right)^{1/2}|\theta_{n,T}-\theta| > C(\log T)^{1/2}(2\theta)^{-1/2}\right\} \\ &\le C\max\left(T^{-1/2}(\log T)^{1/2}, \frac{T^3}{n^2}(\log T)^{-1}\right) + 2(1-\Phi\log T^{1/2}(2\theta)^{-1/2}) \\ &\quad \text{(by Theorem 3.1(a))} \\ &\le C\max\left(T^{-1/2}(\log T)^{1/2}, \frac{T^3}{n^2}(\log T)^{-1}\right). \end{aligned}$$

On the set G_T,

$$\left|\left(\frac{\theta_{n,T}}{\theta}\right)^{1/2} - 1\right| \le CT^{-1/2}(\log T)^{1/2}.$$

Hence, upon choosing $\epsilon = CT^{-1/2}(\log T)^{1/2}, C$ large, we obtain

$$\left|P\left\{\left(\frac{T}{2\theta_{n,T}}\right)^{1/2}(\theta_{n,T}-\theta) \le x, G_T\right\} - \Phi(x)\right|$$

$$\leq \left| P\left\{ \left(\frac{T}{2\theta}\right)^{1/2} (\theta_{n,T} - \theta) \leq x, G_T \right\} \right| + P\left\{ \left| \left(\frac{\theta_{n,T}}{\theta}\right)^{1/2} - 1 \right| > \epsilon, G_T \right\} + \epsilon$$

(by Lemma 1.1 (b) in Bishwal and Bose [13])

$$\leq C \max \left(T^{-1/2} (\log T)^{1/2}, \frac{T^4}{n^2} (\log T)^{-1} \right) \quad \text{(by Theorem 3.1(a))}.$$

This completes the proof of the theorem. □

In the following theorem, we improve the bound on the error of normal approximation using a mixture of random and nonrandom normings. Thus asymptotic normality of the AMCE needs $T \to \infty$ and $\frac{T}{n^{2/3}} \to 0$ which are sharper than the bound in Theorem 3.1. Using this norming, we do not need the rapidly increasing experimental design condition $T \to \infty$ and $\frac{T}{n^{1/2}} \to 0$ as in Theorem 3.1.

Theorem 3.2

$$\sup_{x\in\mathbb{R}} \left| P\left\{ R_{n,T} \left(-\frac{2\theta}{T}\right)^{1/2} (\theta_{n,T} - \theta) \leq x \right\} - \Phi(x) \right| = O\left(T^{-1/2} \bigvee \left(\frac{T^3}{n^2}\right)^{1/3} \right).$$

Proof : Let $b_{n,T} := R_{n,T} - I_T$. By Lemma 3.3 (d), we have

$$E|b_{n,T}|^2 = O\left(\frac{T^4}{n^2}\right). \tag{3.21}$$

From (3.12), we have

$$R_{n,T}(\theta_{n,T} - \theta) = M_T + \theta(R_{n,T} - I_T) = M_T + \theta b_{n,T}.$$

Combining Lemma 3.3 (c) and (d), we have

$$E\left|R_{n,T} - I_{n,T}\right|^2 = O\left(\frac{T^4}{n^2}\right).$$

Thus

$$\begin{aligned}
& \sup_{x\in\mathbb{R}} \left| P\left\{ R_{n,T} \left(\frac{2\theta}{T}\right)^{1/2} (\theta_{n,T} - \theta) \leq x \right\} - \Phi(x) \right| \\
= & \sup_{x\in\mathbb{R}} \left| P\left\{ \left(\frac{2\theta}{T}\right)^{1/2} [M_T + \theta b_{n,T}] \leq x \right\} - \Phi(x) \right| \\
\leq & \sup_{x\in\mathbb{R}} \left| P\left\{ \left(\frac{2\theta}{T}\right)^{1/2} Y_T \leq x \right\} - \Phi(x) \right| + P\left\{ \left| \left(\frac{2\theta}{T}\right)^{1/2} [\theta b_{n,T}] \right| > \epsilon \right\} + \epsilon \\
\leq & \; CT^{-1/2} + \left(\frac{2\theta}{T}\right) \frac{E|\theta b_{n,T}|^2}{\epsilon^2} + \epsilon
\end{aligned}$$

(by Lemma 1 in Michel and Pfanzagl [12] and Lemma 1.2 in Bishwal and Bose [13])

$$\leq CT^{-1/2} + C\frac{T^3}{n^2\epsilon^2} + \epsilon \text{ (by (3.21))}. \tag{3.22}$$

Choosing $\epsilon = (\frac{T^3}{n^2})^{1/3}$, the rate is $O\left(T^{-1/2} \bigvee \left(\frac{T^3}{n^2}\right)^{1/3}\right)$. □

The following theorem gives stochastic bound on the error of approximation of the continuous AMCE by the AMCE.

Theorem 3.3

$$|\theta_{n,T} - \theta_T| = O_P\left(\frac{T^2}{n}\right).$$

Proof : Note that $\theta_{n,T} - \theta_T = \frac{T}{2R_{n,T}} - \frac{T}{2I_T}$. From Lemma 3.3 it follows that $|R_{n,T} - I_T| = O_P(T^4/n^2)^{1/2}$. Now the theorem follows easily from the from the Lemma 1.2 in Bishwal and Bose [13]. □

We propose three theoretical ratio estimators of the drift. The first ratio estimator of θ is defined as

$$\hat{\theta}_n := -\log\left[\frac{\sigma^2_{n\Delta}}{\sigma^2_{(n-1)\Delta}}\right]. \tag{3.23}$$

The second ratio estimator of θ is defined as

$$\tilde{\theta}_n := -\log\left[\frac{\sum_{i=1}^n \sigma^2_{i\Delta}}{\sum_{i=1}^n \sigma^2_{(i-1)\Delta}}\right]. \tag{3.24}$$

The third ratio estimator of θ is defined as

$$\check{\theta}_n := -\log\left[\min_{1\le i\le n} \frac{\sigma^2_{i\Delta}}{\sigma^2_{(i-1)\Delta}}\right]. \tag{3.25}$$

We propose three observable ratio estimators of the drift. The first observable ratio estimator of θ is defined as

$$\hat{\theta}_{n,T} := -\log\left[\frac{R_{n,T}}{R_{n-1,T}}\right]. \tag{3.26}$$

The second observable ratio estimator of θ is defined as

$$\tilde{\theta}_{n,T} := -\log\left[\frac{\sum_{i=1}^n R_{n,i\Delta}}{\sum_{i=1}^n R_{n,(i-1)\Delta}}\right]. \tag{3.27}$$

The third observable ratio estimator of θ is defined as

$$\check{\theta}_{n,T} := -\log\left[\min_{1\le i\le n} \frac{R_{n,i\Delta}}{R_{n,(i-1)\Delta}}\right]. \tag{3.28}$$

It would be interesting to study Kolmogorov distance of these estimators.

4. Estimation of Mean Reversion Level and Speed

We consider the general case with two parameters. Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\ge0}, P)$ be a stochastic basis on which the Cox-Ingersoll-Ross process $\{X_t\}$ is defined satisfying the Itô stochastic differential equation

$$dX_t = (\alpha + \beta X_t)\,dt + \ 2\sqrt{X_t}\,dW_t,\ \ t \ge 0, \quad X_0 = 1 \tag{4.1}$$

where $\{W_t\}$ is a standard Wiener process with the filtration $\{\mathcal{F}_t\}_{t\ge0}$, $\alpha > 0$ and $\beta < 0$ are the unknown parameters to be estimated on the basis of observations of the process $\{X_t\}$.

The true transition density which is the fundamental solution to the PDE

$$u_t = 2xu_{xx} + \alpha u_x - \left(\frac{\mu}{x} + \lambda x\right) u \tag{4.2}$$

is given by

$$q(t,x,y,\alpha,\beta) := -2\beta \left(\frac{y}{x}\right)^{\alpha-\frac{1}{2}} \frac{e^{(\frac{1}{2}-\alpha)\beta t}}{1-e^{\beta t}} \exp\left[\frac{2\beta(x+y)}{e^{-\beta t}-1}\right] I_\nu\left(\frac{-2\beta\sqrt{xy}}{\sinh(-\frac{1}{2}\beta t)}\right) \tag{4.3}$$

where I_ν is the modified Bessel function of first kind with index ν, which is noncentral chi-square density. The invariant density is gamma as $t \to \infty$.

Let the continuous realization $\{X_t, 0 \le t \le T\}$ be denoted by X_0^T. Let P_β^T be the measure generated on the space (C_T, B_T) of continuous functions on $[0,T]$ with the associated Borel σ-algebra B_T generated under the supremum norm by the process X_0^T and let P_0^T be the standard Wiener measure. It is well known that when β is the true value of the parameter $P_{\beta,\alpha}^T$ is absolutely continuous with respect to P_0^T and the Radon-Nikodym derivative (likelihood) of P_β^T with respect to P_0^T based on X_0^T is given by

$$L_T(\beta,\alpha) := \frac{dP_{\beta,\alpha}^T}{dP_0^T}(X_0^T) = \exp\left\{\int_0^T \frac{\alpha+\beta X_t}{4X_t} dX_t - \int_0^T \frac{(\alpha+\beta X_t)^2}{8X_t} dt\right\}. \tag{4.4}$$

Consider the score function, the derivative of the log-likelihood function, which is given by

$$\gamma_T(\beta,\alpha) := \left\{\int_0^T \frac{\alpha+\beta X_t}{4X_t} dX_t - \int_0^T \frac{(\alpha+\beta X_t)^2}{8X_t} dt\right\}. \tag{4.5}$$

We estimate α and β. A solution of the estimating equation $\gamma_T(\beta,\alpha) = 0$ provides the maximum likelihood estimates (MLEs)

$$\hat{\beta}_T := \frac{X_0 - X_T + \alpha T}{\int_0^T X_t dt}, \quad \hat{\alpha}_T := \frac{\int_0^T X_t^{-1} dX_t + \beta T}{\int_0^T X_t^{-1} dt} = \frac{\log X_T - \log X_0 + \int_0^T X_t^{-1} dt + \beta T}{\int_0^T X_t^{-1} dt}.$$

It is important to note that if $\beta > 0$ and $\alpha \ge 2$, the MLE $\hat{\alpha}_T$ is *inconsistent*. It remains an open problem to find a consistent estimator in this case.

Consider the minimum contrast estimates (MCE)

$$\widetilde{\beta}_T := \frac{\alpha T}{\int_0^T X_t dt} = \frac{\alpha}{\overline{X_T}} \quad \text{where} \quad \overline{X_T} = \frac{1}{T}\int_0^T X_t dt$$

and

$$\widetilde{\alpha}_T := \frac{\beta T}{\int_0^T X_t^{-1} dt} = \frac{\beta}{\overline{X_T^{-1}}} \quad \text{where} \quad \overline{X_T^{-1}} = \frac{1}{T}\int_0^T X_t^{-1} dt.$$

Note that using the Skorohod embedding of martingale which has been the one of the basic tools for normal approximation of martingales, will not give a rate better than $O(T^{-1/4})$. To obtain the rate of normal approximation of the order $O(T^{-1/2})$, we adopt the Fourier method followed by the squeezing technique of Bishwal [8].

Observe that

$$\left(\frac{T\alpha}{-4\beta}\right)^{1/2} (\widetilde{\beta}_T - \beta) = \frac{\left(\frac{-4\beta}{T\alpha}\right)^{1/2} N_T}{\left(\frac{-4\beta}{T\alpha}\right) I_T} \tag{4.6}$$

and

$$\left(\frac{T\beta}{-4(\alpha-2)}\right)^{1/2} (\widetilde{\alpha}_T - \alpha) = \frac{\left(\frac{-4(\alpha-2)}{T\beta}\right)^{1/2} M_T}{\left(\frac{-4(\alpha-2)}{T\beta}\right) J_T}. \tag{4.7}$$

where

$$N_T := \alpha T - \beta I_T, \quad M_T := \beta T - \alpha J_T, \quad I_T := \int_0^T X_t dt, \quad \text{and} \quad J_T := \int_0^T X_t^{-1} dt.$$

The process I_T which is energy of the CIR process which plays a important role in clustering time or activity persistence in stochastic volatility modeling.

Based on continuous time observation $\{X_t, 0 \leq t \leq T\}$ the continuous conditional least squares estimators of β and α are respectively given by

$$\beta_T := \frac{\int_0^T X_s dX_s - (X_T - X_0)\widetilde{X_T}}{\int_0^T (X_t - \widetilde{X_T})^2 dt}, \tag{4.8}$$

$$\alpha_T := -\widetilde{X_T}\beta_T + T^{-1}(X_T - X_0) \tag{4.9}$$

where

$$\widetilde{X_T} := \int_0^T X_t dt. \tag{4.10}$$

Note that by Itô formula

$$X_T^2 - X_0^2 = 2\int_0^T X_s dX_s + \int_0^T X_s ds. \tag{2.11}$$

Hence

$$\begin{aligned} \beta_T &= \frac{T\widetilde{X_T}}{\int_0^T (X_t - \widetilde{X_T})^2 dt} + o_P(T^{-1/2}) \\ &= \frac{T\widetilde{X_T}}{2(\widetilde{X_T^2} - (\widetilde{X_T})^2)} + o_P(T^{-1/2}) \end{aligned} \tag{4.12}$$

$$\alpha_T = \frac{\widetilde{X_T^2}}{2(\widetilde{X_T^2} - (\widetilde{X_T})^2)} + o_P(T^{-1/2}) \tag{4.13}$$

where

$$\widetilde{X_T^2} := \int_0^T X_t^2 dt \tag{4.14}$$

We define the minimum contrast estimators as

$$\breve{\beta}_T := \frac{T\widetilde{X_T}}{2(\widetilde{X_T^2} - (\widetilde{X_T})^2)}, \tag{4.15}$$

$$\breve{\alpha}_T := \frac{\widetilde{X_T^2}}{2(\widetilde{X_T^2} - (\widetilde{X_T})^2)}. \tag{4.16}$$

Consider the Heston model under the risk-neutral measure

$$dS_t = rS_t dt + \sqrt{X_t} S_t dZ_t$$

$$dX_t = (\alpha + \beta X_t)\, dt + \; 2\sqrt{X_t}\, dW_t, \; t \geq 0, \quad X_0 = 1 \tag{4.17}$$

where Z_t and W_t are correlated Brownian motions with correlation parameter ρ. For instance,

$$Z_t = \rho W_t + \sqrt{1-\rho^2} V_t$$

where W_t and V_t are two independent Brownian motions.

We define the minimum contrast estimators as

$$\hat{\beta}_T := \frac{T\widetilde{\hat{X}_T}}{2(\widetilde{\hat{X}_T^2} - (\widetilde{\hat{X}_T})^2)}, \tag{4.18}$$

$$\hat{\alpha}_T := \frac{\widetilde{\hat{X}_T^2}}{2(\widetilde{\hat{X}_T^2} - (\widetilde{\hat{X}_T})^2)}. \tag{4.19}$$

where

$$\widetilde{\hat{X}_T} = \int_0^T \hat{X}_t dt, \quad \widetilde{\hat{X}_T^2} = \int_0^T \hat{X}_t^2 dt \tag{4.20}$$

and $\hat{X}_t = E(X_t|\mathcal{S}_t)$ where $\mathcal{S}_t = \sigma(S_u, 0 \leq u \leq t)$ which can be estimated as follows:

Using Itô formula to $Y_t = \log S_t$, we have the observation process

$$dY_t = (r - \frac{1}{2}\sigma^2)dt + \sqrt{X_t}dZ_t, \ t \geq 0. \tag{4.21}$$

We follow conditional least squares (CLS) estimation method as in ARCH(1) model. Observe that

$$(dY_t)^2 = X_t(dZ_t)^2, \quad E[(dY_t)^2|\mathcal{F}_t] = X_t. \tag{4.22}$$

Denote $\hat{X}_{t_i} = (Y_{t_i} - Y_{t_{i-1}})^2$. We have

$$\sum_{i=1}^n (Y_{t_i} - Y_{t_{i-1}})^2 \to^P \int_0^T X_t dt, \quad \sum_{i=1}^n (Y_{t_i}^2 - Y_{t_{i-1}}^2)^2 \to^P \int_0^T X_t^2 dt \tag{4.23}$$

We use these estimates for (4.18) and (4.19) and obtain

$$\hat{\beta}_{n,T} := \frac{n\Delta \sum_{i=1}^n (Y_{t_i} - Y_{t_{i-1}})^2}{2(\sum_{i=1}^n (Y_{t_i}^2 - Y_{t_{i-1}}^2)^2 - (\sum_{i=1}^n (Y_{t_i} - Y_{t_{i-1}})^2)^2)}, \tag{4.24}$$

$$\hat{\alpha}_{n,T} := \frac{\sum_{i=1}^n (Y_{t_i}^2 - Y_{t_{i-1}}^2)^2}{2(\sum_{i=1}^n (Y_{t_i}^2 - Y_{t_{i-1}}^2)^2 - (\sum_{i=1}^n (Y_{t_i} - Y_{t_{i-1}})^2)^2)}. \tag{4.25}$$

By following the methods in the previous section, we have

Theorem 4.1

a) $$\sup_{x \in \mathbb{R}} \left| P\left\{ R_{n,T} \left(-\frac{4(\alpha-2)}{T\hat{\beta}_{n,T}} \right)^{1/2} (\alpha_{n,T} - \alpha) \leq x \right\} - \Phi(x) \right| = O\left(T^{-1/2} \bigvee \left(\frac{T^3}{n^2} \right)^{1/3} \right).$$

b) $$\sup_{x \in \mathbb{R}} \left| P\left\{ R_{n,T} \left(-\frac{4\beta}{T\hat{\alpha}_{n,T}} \right)^{1/2} (\beta_{n,T} - \beta) \leq x \right\} - \Phi(x) \right| = O\left(T^{-1/2} \bigvee \left(\frac{T^3}{n^2} \right)^{1/3} \right).$$

5. Estimation of Correlation

Next we consider the SDEs with five parameters

$$dY_t = (\mu + \beta X_t)dt + \sqrt{X_t}dW_t + \rho dZ_{\theta t}, \tag{5.1}$$

$$dX_t = -\theta X_t dt + dZ_{\theta t}, \tag{5.2}$$

where Z_t is a inverse Gaussian (IG) Levy process independent of X_0 with $\mathcal{L}(X_0) = IG(\delta, \gamma)$. We suppose that the parameters δ and γ are known. Here $\theta > 0$ and $\rho < 0$. When the process Z is inverse-Gaussian, the model is called the IGOU process. In the IGOU model, calculation of conditional cummulants of the integrated volatility conditioned on the initial value is enough to be able to compute European style options. We consider fixed time interval $t_i - t_{i-1} = \Delta,\ i = 1, 2, \ldots, n.$

The process Z is the sum of two independent Levy processes $Z = Z^{(1)} + Z^{(2)}$ where $\mathcal{L}(X_0) = IG(\delta/2, \gamma)$ and $Z^{(2)}$ is a compound Poisson process given by $Z^{(2)} = \gamma^{-2}\sum_{j=1}^{N_t} u_j$ with N being a Poisson process with intensity $\delta\gamma/2$ and u_j is a sequence of independent and identically χ_1^2-distributed random variables independent of N, see Barndorff-Neilsen and Shephard [2].

The processes Z and X have infinitely many jumps in any finite time interval, hence they are infinite activity processes. Note that the cumulative process or the integrated process $I_t = \int_0^t X_u du$ has long range dependence or long memory, see Barndorff-Neilsen and Shephard [2].

The cumulant functions of IGOU process are given by

$$k(u) = \log E[e^{-uZ(1)}] = -u\delta\gamma^{-1}(1 + 2u\gamma^{-2})^{-1/2}, \tag{5.3}$$

$$k'(u) = \log E(e^{-uX_t}) = \delta\gamma - \delta\gamma(1 + 2u\gamma^{-2})^{1/2}, \quad u \in \mathbb{R}. \tag{5.4}$$

In order to construct the estimating functions, we use the first and second cumulants which are given respectively by

$$\kappa_{y_1}^{(1)} = \theta\rho\Delta\kappa_{IG}^{(1)}, \quad \kappa_{y_1}^{(2)} = \Delta\kappa_{IG}^{(1)} + 2\theta\rho^2\Delta\kappa_{IG}^{(2)} \tag{5.5}$$

where $y_j := Y_{j\Delta} - Y_{(j-1)\Delta}, j = 1, 2, \ldots, n$. Inverting these cumulants and replacing the cumulants by their sample quantities, we obtain the explicit the moment estimators of ρ and θ.

The moment estimators of ρ and θ are given by

$$\hat{\rho}_n := \frac{\gamma(\gamma s_y^2 - \Delta\delta)}{2\bar{y}}, \quad \hat{\theta}_n := \frac{\gamma\bar{y}}{\Delta\delta\hat{\rho}_n}$$

where

$$s_y^2 := \frac{1}{n}\sum_{j=1}^{n}(y_j - \bar{y})^2 = \frac{1}{n}\sum_{j=1}^{n} y_j^2 - (\bar{y})^2,$$

$$\bar{y} := \frac{1}{n}\sum_{j=1}^{n} y_j, \quad y_j := Y_{j\Delta} - Y_{(j-1)\Delta}$$

Let $\vartheta = (\rho, \theta)$. and $\hat{\vartheta}_n = (\hat{\rho}_n, \hat{\theta}_n)$. We have the following properties of the estimators:

Proposition 5.1 For fixed $\Delta > 0$ as $n \to \infty$,

$$(a) \quad \hat{\vartheta}_n \to \vartheta_0 \ \text{ a.s.} \quad \text{as } n \to \infty.$$

$$(b) \quad \sqrt{n}(\hat{\vartheta}_n - \vartheta_0) \to^{\mathcal{D}} \mathcal{N}_2(0, (2\theta\rho^2\Delta^2\delta^2\gamma^{-4})^{-2}V(\vartheta_0)) \text{ as } n \to \infty.$$

where $V(\vartheta_0)$ is the limiting covariance matrix.

Next we consider the Heston model with correlation under the risk neutral measure:

$$dS_t = rS_t dt + \rho\sqrt{X_t}S_t dW_t$$

which gives

$$dY_t = (r - \frac{1}{2}\rho^2)dt + \rho\sqrt{X_t}dW_t, \tag{5.6}$$

$$dX_t = (1 - 2\theta X_t)dt + dW_t. \tag{5.7}$$

where $Y_t = \log S_t$. Invariant distribution of X_t is gamma with parameters a and b. We assume a and b known. The moment estimators of θ and ρ are given by

$$\hat{\theta}_n := \frac{\frac{1}{n^2}\left[\sum_{i=1}^n (Y_{i\Delta} - Y_{(i-1)\Delta})\right]^2}{\frac{1}{n^2}\sum_{i=1}^n (Y_{i\Delta} - Y_{(i-1)\Delta})^2 - \frac{\Delta}{n}\left[\sum_{i=1}^n (Y_{i\Delta} - Y_{(i-1)\Delta})\right]} \frac{2a^3(a+1)}{b^4\Delta},$$

$$\hat{\rho}_n := \frac{\frac{1}{n^2}\sum_{i=1}^n (Y_{i\Delta} - Y_{(i-1)\Delta})^2 - \frac{\Delta}{n}\left[\sum_{i=1}^n (Y_{i\Delta} - Y_{(i-1)\Delta})\right]}{\frac{1}{n^2}\left[\sum_{i=1}^n (Y_{i\Delta} - Y_{(i-1)\Delta})\right]} \frac{b^3\Delta}{2a^2(a+1)}.$$

By using Theorem 4.1 Van der Vaart [20], we obtain the strong consistency and asymptotic normality of the method of moments (MM) estimators:

Proposition 5.2

a) $\hat{\theta}_n \to \theta_0$ a.s. as $n \to \infty$.
b) $\sqrt{n}(\hat{\theta}_n - \theta_0) \to^{\mathcal{D}} \mathcal{N}(0,\ L^{-1}(\theta_0))$ as $n \to \infty$.
c) $\hat{\rho}_n \to \rho_0$ a.s. as $n \to \infty$.
d) $\sqrt{n}(\hat{\rho}_n - \rho_0) \to^{\mathcal{D}} \mathcal{N}(0,\ K^{-1}(\rho_0))$ as $n \to \infty$
where $L(\theta_0)$ and $K(\rho_0)$ are the corresponding Fisher-information.

Next we show that the Berry-Esseen bounds in part d) is of the order $O(n^{-1/2})$. Let

$$dY_t = b_t dt + \sqrt{\rho}\sqrt{X_t} dW_t \tag{5.8}$$

where b_t is an Itô process and X_t is a positive Itô processes satisfying the SDEs

$$db_t = b_t^{[0]} dt + b_t^{[1]} dW_t,\ d\sqrt{X}_t = \sigma_t^{[0]} dt + \sigma_t^{[1]} dW_t \tag{5.9}$$

$$db_t^{[0]} = b_t^{[0,0]} dt + b_t^{[0,1]} dW_t,\ d\sigma_t^{[0]} = \sigma_t^{[0,0]} dt + \sigma_t^{[0,1]} dW_t, \tag{5.10}$$

$$db_t^{[1]} = b_t^{[1,0]} dt + b_t^{[1,1]} dW_t,\ d\sigma_t^{[1]} = \sigma_t^{[1,0]} dt + \sigma_t^{[1,1]} dW_t \tag{5.11}$$

We assume that the drift and the volatility coefficients of the hidden diffusions are smooth, i.e., $\sup_{t\in[0,1]} \|f_t\|_{p,3} < \infty$ for $f_t = b_t, b_t^{[0]}, b_t^{[1]}, b_t^{[1,0]}, b_t^{[1,1]}, \sigma_t, \sigma_t^{[0]}, \sigma_t^{[1]}, \sigma_t^{[0,0]}, \sigma_t^{[0,1]}, \sigma_t^{[1,0]}, \sigma_t^{[1,1]}$ for $p > 1$. We also assume that $\sup_{t\in[0,1]} \|1/\sigma_t\|_{L^p} < \infty$ for each for $p > 1$. Our aim is to estimate ρ based on observations of Y at $0 = t_0 < t_1 < t_2 \ldots < t_n = 1$. Thus the observations are $Y_{t_0}, Y_{t_1}, Y_{t_2}, \ldots, Y_{t_n}$. Let $\Delta W_i := W_{t_i} - W_{t_{i-1}}$.

The quasi maximum likelihood estimator (QMLE) of ρ is given by

$$\hat{\rho}_n = \sum_{i=1}^n \left(\frac{Y_{t_i} - Y_{t_{i-1}}}{\sqrt{X}_{t_{i-1}}}\right)^2 = \sum_{i=1}^n \frac{(Y_{t_i} - Y_{t_{i-1}})^2}{X_{t_{i-1}}}.$$

We have the following optimal rate of convergence on the Kolmogorov distance:

Theorem 5.3

$$\sup_{x\in\mathbb{R}} \left| \frac{\sqrt{n}}{\sqrt{2}\rho}(\hat{\rho}_n - \rho_0) - \Phi(x) \right| = O(n^{-1/2}).$$

Proof Using Burkholder's inequality for martingales and Itô's formula, we show that

$$\frac{\sqrt{n}}{\sqrt{2}\rho}(\hat{\rho}_n - \rho) = M_n + \frac{1}{\sqrt{n}} N_n \tag{5.12}$$

where

$$M_n := \sum_{i=1}^{n} \frac{1}{\sqrt{2n}}[(\sqrt{n}\Delta W_i)^2 - 1], \tag{5.13}$$

and

$$N_n := \sum_{i=1}^{n} \frac{1}{\sqrt{6n}} \frac{\sqrt{3}\sigma^{[1]}_{t_{i-1}}}{\sqrt{X_{t_{i-1}}}}[(\sqrt{n}\Delta W_i)^3 - 3\sqrt{n}\Delta W_i]$$

$$+ \sum_{i=1}^{n} \frac{1}{\sqrt{n}} \frac{\sqrt{2}\sigma^{[1]}_{t_{i-1}}}{\sqrt{X_{t_{i-1}}}} + \frac{\sqrt{2}b_{t_{i-1}}}{\sqrt{\theta}\sqrt{X_{t_{i-1}}}})(\sqrt{n}\Delta W_i) + \sum_{i=1}^{n} \frac{1}{n} F_{t_{i-1}} + O_P(n^{-1/2}) \tag{5.14}$$

and

$$F_t := \frac{\sqrt{2}\sigma^{[0]}_t}{2\sqrt{X_t}} + \frac{(\sigma^{[1]}_t)^2}{2\sqrt{2}X_t} + \frac{\sqrt{2}b^{[1]}_t}{2\sqrt{\rho}\sqrt{X_t}} + \frac{\sqrt{2}b_t^2}{2\rho X_t}. \tag{5.15}$$

Further,

$$\hat{\rho}_n - \int_0^T \left(\frac{\sqrt{X_t}}{\sqrt{X_{t_{i-1}}}}\right)^2 dt = \sum_{i=1}^{n} \left(\frac{Y_{t_i} - Y_{t_{i-1}}}{\sqrt{X_{t_{i-1}}}}\right)^2 - \int_0^T \left(\frac{\sqrt{X_t}}{\sqrt{X_{t_{i-1}}}}\right)^2 dt \tag{5.16}$$

$$= \sum_{i=1}^{n} \left(\frac{Y_{t_i} - Y_{t_{i-1}}}{\sqrt{X_{t_{i-1}}}}\right)^2 - \sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} \left(\frac{\sqrt{X_t}}{\sqrt{X_{t_{i-1}}}}\right)^2 dt. \tag{5.17}$$

Denote

$$\mathcal{E}_n := \frac{\sqrt{n}}{\sqrt{2}\rho}(\hat{\rho}_n - \rho). \tag{5.18}$$

We decompose this as the sum of four terms:

$$\mathcal{E}_n = \Psi + \Phi_2 + \Phi_3 + \Phi_4 \tag{5.19}$$

where

$$\Psi = \sum_{i=1}^{n} \frac{\sqrt{n}}{\sqrt{2}} \left\{ \frac{1}{X_{t_{i-1}}} \left(\int_{t_{i-1}}^{t_i} \sqrt{X_t} dW_t \right)^2 - \frac{1}{n} \right\}, \quad \Phi_2 = \sum_{i=1}^{n} \frac{\sqrt{2n}}{\sqrt{\rho}X_{t_{i-1}}} \int_{t_{i-1}}^{t_i} \int_{t_{i-1}}^{t} b_s ds \sqrt{X_t} dW_t, \tag{5.20}$$

$$\Phi_3 = \sum_{i=1}^{n} \frac{\sqrt{2n}}{\sqrt{\rho}X_{t_{i-1}}} \int_{t_{i-1}}^{t_i} \int_{t_{i-1}}^{t} \sqrt{X_s} dW_s b_t dt, \quad \Phi_4 = \sum_{i=1}^{n} \frac{\sqrt{2n}}{\rho X_{t_{i-1}}} \int_{t_{i-1}}^{t_i} \int_{t_{i-1}}^{t} b_s ds b_t dt. \tag{5.21}$$

$$\sqrt{n}\Phi_3 = \sum_{i=1}^{n} \frac{\sqrt{2n}}{\sqrt{\rho}X_{t_{i-1}}} \int_{t_{i-1}}^{t_i} \sqrt{X_{t_{i-1}}} b_{t_{i-1}} \int_{t_{i-1}}^{t} dW_s dt + \int_{t_{i-1}}^{t_i} b_{t_{i-1}} \left(\int_{t_{i-1}}^{s} \sqrt{X}^{[1]}_s dW_u \right) dW_s dt$$

$$+ \int_{t_{i-1}}^{t_i} \sigma_{t_{i-1}} \int_{t_{i-1}}^{t} b^{[1]}_s ds dt + \int_{t_{i-1}}^{t_i} \sigma_{t_{i-1}} \left[\int_{t_{i-1}}^{t} \int_{t_{i-1}}^{v} b^{[1]}_s dW_s dW_v + \int_{t_{i-1}}^{t} \int_{t_{i-1}}^{v} dW_s \cdot b^{[1]}_v dW_v \right] dt + O_P(n^{-0.5})$$

$$= \sum_{i=1}^{n} \frac{\sqrt{2n}}{\sqrt{\rho}X_{t_{i-1}}} \left\{ b_{t_{i-1}} \int_{t_{i-1}}^{t_i} \int_{t_{i-1}}^{t} dW_s dt + b^{[1]}_{t_{i-1}} \int_{t_{i-1}}^{t_i} \int_{t_{i-1}}^{t} ds dt \right\} + O_P(n^{-0.5}). \tag{5.22}$$

$$\sqrt{n}\Phi_2 = \sum_{i=1}^{n} \frac{\sqrt{2}nb_{t_{i-1}}}{\sqrt{\rho}\sigma^2_{t_{i-1}}} \int_{t_{i-1}}^{t_i} \int_{t_{i-1}}^{t} ds dW_t + O_P(n^{-0.5}). \tag{5.23}$$

$$\sqrt{n}\Phi_2 + \sqrt{n}\Phi_3 = \sum_{i=1}^{n} \frac{\sqrt{2}nb_{t_{i-1}}}{\sqrt{\rho}X_{t_{i-1}}}(W_{t_i} - W_{t_{i-1}}) + \sum_{i=1}^{n} \frac{\sqrt{2}nb_{t_{i-1}}}{2\sqrt{\rho}X_{t_{i-1}}} \cdot \frac{1}{n} + O_P(n^{-0.5}). \tag{5.24}$$

$$\sqrt{n}\Phi_4 = \sum_{i=1}^{n} \frac{\sqrt{2}b^2_{t_{i-1}}}{2\rho X_{t_{i-1}}} \cdot \frac{1}{n} + O_P(n^{-0.5}). \tag{5.25}$$

On the other hand, Ψ is decomposed as the sum of three terms:

$$\Psi = \Psi_1 + \Psi_2 + \Psi_3 \tag{5.26}$$

where

$$\Psi_1 = \sum_{i=1}^{n} \frac{\sqrt{n}}{\sqrt{2}} \left((\Delta W_i)^2 - \frac{1}{n} \right), \ \Psi_2 = \sum_{i=1}^{n} \frac{\sqrt{2n}}{\sqrt{X}_{t_{i-1}}} \Delta W_i \int_{t_{i-1}}^{t_i} (\sqrt{X}_t - \sqrt{X}_{t_{i-1}}) dW_t, \tag{5.27}$$

$$\Psi_3 = \sum_{i=1}^{n} \frac{\sqrt{n}}{\sqrt{2} X_{t_{i-1}}} \left(\int_{t_{i-1}}^{t_i} (\sqrt{X}_t - \sqrt{X}_{t_{i-1}}) dW_t \right)^2. \tag{5.28}$$

$$\sqrt{n}\Psi_1 = \sum_{i=1}^{n} \frac{n}{\sqrt{2}} \left((\Delta W_i)^2 - \frac{1}{n} \right), \tag{5.29}$$

$$\sqrt{n}\Psi_3 = \sum_{i=1}^{n} \frac{n}{\sqrt{2} X_{t_{i-1}}} \left(\int_{t_{i-1}}^{t_i} \int_{t_{i-1}}^{t} (\sigma_s^{[1]} dW_s dW_t \right)^2 + O_P(n^{-0.5}) \tag{5.30}$$

$$= \sum_{i=1}^{n} \frac{n}{\sqrt{2}\sqrt{X}_{t_{i-1}}} \int_{t_{i-1}}^{t_i} \left(\int_{t_{i-1}}^{t} \sigma_s^{[1]} dW_s \right)^2 dt + O_P(n^{-0.5}) \tag{5.31}$$

$$= \sum_{i=1}^{n} \frac{n}{\sqrt{2}\sqrt{X}_{t_{i-1}}} \int_{t_{i-1}}^{t_i} \int_{t_{i-1}}^{t} \left(\sigma_s^{[1]} \right)^2 ds dt + O_P(n^{-0.5}) \tag{5.32}$$

$$= \sum_{i=1}^{n} \frac{1}{2\sqrt{2}} \frac{\left(\sigma_{t_{i-1}}^{[1]} \right)^2}{X_{t_{i-1}}} \cdot \frac{1}{n} + O_P(n^{-0.5}), \tag{5.33}$$

$$\sqrt{n}\Psi_2 = \sum_{i=1}^{n} \frac{\sqrt{2}}{\sqrt{X}_{t_{i-1}}} n \Delta W_i \int_{t_{i-1}}^{t_i} \int_{t_{i-1}}^{t} dW_s dW_t + + \sum_{i=1}^{n} \frac{\sqrt{2}\sigma_{t_{i-1}}^{[0]}}{2\sqrt{X}_{t_{i-1}}} \cdot \frac{1}{n} + O_P(n^{-0.5}). \tag{5.34}$$

Due to uniform non-degeneracy of the Malliavian covariance of the scaled stochastic integral of the diffusion functional with respect to Brownian motion, we have

$$\sup_{x \in \mathbb{R}} |P(M_n \le x) - \Phi(x)| \le C n^{-1/2} \tag{5.35}$$

and

$$P(\frac{1}{\sqrt{n}} N_n > \epsilon) \le C n^{-1}. \tag{5.36}$$

Hence

$$\sup_{x \in \mathbb{R}} |P(\frac{\sqrt{n}}{\sqrt{2}\rho}(\hat{\rho}_n - \rho) \le x) - \Phi(x)| \le C n^{-1/2}. \tag{5.37}$$

With $\rho = 1$,

$$\sup_{x \in \mathbb{R}} |P(\frac{\sqrt{n}}{\sqrt{2}}(\hat{\rho}_n - 1) \le x) - \Phi(x)| \le C n^{-1/2}. \tag{5.38}$$

When the process X is unobserved, we replace X_t by the conditional expectation $\hat{X}_t = E(X_t|\mathcal{Y}_t)$ where $\mathcal{Y}_t = \sigma(Y_s, 0 \le s \le t)$, which can be evaluated by Kitagawa [18] algorithm or by Monte Carlo methods. □

6. Estimation of Stochastic Correlation and Stochastic Elasticity of Volatility

In general, the correlation should be stochastic. It can be modeled by a Jacobi process, see Veraart and Veraart [17]. A real asset price model should be of the following hybrid type with

14 parameters. We consider the hybrid stochastic volatility, stochastic interest rate, stochastic leverage and stochastic elasticity model under the risk neutral measure which is given by

$$dS_t = X_t dt + \sqrt{V_{t-}} S_t dW_t + \rho_{\lambda t} dL_{\tau_{\lambda t}}, \tag{6.1}$$

$$dV_t = -\lambda V_t dt + \upsilon_{\lambda t-} dL_{\tau_{\lambda t}}, \tag{6.2}$$

$$dX_t = \alpha(\beta - X_t)dt + \sigma X_t^{\gamma_t} dW_t^H, \tag{6.3}$$

$$d\rho_t = ((2\zeta - \eta) - \eta\rho_t)dt + \theta\sqrt{(1+\rho_t)(1-\rho_t)}dZ_t, \tag{6.4}$$

$$d\xi_t = \kappa(\mu - \xi_t)dt + \varsigma\sqrt{\xi_t}dB_t, \tag{6.5}$$

$$d\gamma_t = \varpi(\psi - \delta))dt + \sqrt{\chi}dM_t, \tag{6.6}$$

$$d\tau_t = \xi_{t-}dt, \; t \geq 0 \tag{6.7}$$

where $(L_t, t \geq 0)$ is a Levy process, $(W^H, t \geq 0)$ is a subfractional Brownian motion, $(B_t, t \geq 0)$, $(Z_t, t \geq 0)$ and $(M_t, t \geq 0)$ are standard Brownian motions. Here $(S_t, t \geq 0)$ is the asset price which a geometric jump-diffusion, $(V_t, t \geq 0)$ is the stochastic volatility which is a Levy O-U process, $(X_t, t \geq 0)$ is the stochastic interest rate which is a sub-fractional Chan-Karolyi-Longstaff-Sanders (CKLS) process, $(\rho_t, t \geq 0)$ is the stochastic leverage Jacobi (Beta) process, $(\xi_t, t \geq 0)$ is a volatility modulation (stochastic time change or stochastic clock process) of the driving Levy subordinator which is a Cox-Ingersoll-Ross (CIR) process, γ_t is the stochastic elasticity models which is another CIR process, and all the 14 parameters $\lambda, \alpha, \beta, \sigma, \xi, \eta, \theta, \kappa, \mu, \varsigma, \varpi, \psi, \delta, \chi$ are positive. See Bishwal [19].

We first discuss estimation of elasticity when it is a constant parameter by the test function estimation method. Generalized method of moments (GMM), which is a generalization of weighted least squares method with the random weight being the inverse of the covariance matrix is a popular estimation method in financial econometrics where likelihood may not be available, that is maximum likelihood estimation is not feasible. Also one may not need the distribution of the error term in the model. GMM estimators are in general consistent, asymptotic normal and asymptotically efficient. Conley *et al.* [21] proposed to estimate the elasticity parameter γ by minimizing a generalized method of moments (GMM) criterion function. The criterion function is based on a combined set of moment conditions constructed from the level and difference test functions, whereas the elasticity is treated as an unknown parameter to be estimated along with the drift parameters. To facilitate the interpretation of the GMM test statistics, we can estimate the elasticity parameter γ by the two-step GMM estimation procedure proposed in Conley *et al.* [21]. In the first step, we use our estimators of the drift parameters of the previous section as a function of the variance elasticity γ and plug them into the moment conditions formed from test functions of the first differences to estimate γ. Then we estimate γ by GMM method. Based on empirical data fitting to the model, the value of the elasticity parameter is known to be near 0.75.

Consider a test function φ in the domain of the generator $\mathcal{G}$ of the diffusion process X satisfying the SDE

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t. \tag{6.8}$$

Since $E[\varphi(X_t)]$ is constant over time, it has zero derivative. We have

$$E[\mathcal{G}\varphi(X_t)] = E[\mu(X_t)\varphi'(X_t) + \frac{1}{2}\sigma^2(X_t)\varphi''(X_t)] = 0. \tag{6.9}$$

An efficient test function would be l_T, the derivative of the log-likelihood. The resulting test function estimator using $E[\mathcal{G}l_T] = 0$ will be efficient. It will be more efficient than the quasi-maximum likelihood estimator (QMLE) that uses the moment condition $E[l_T] = 0$. In effect, the application of the generator to the score function adjusts the moment conditions optimally for the presence of temporal dependence. One can use localized test functions by multiplying the first derivative by a smooth kernel K.

Finally, we propose a fractional stochastic elasticity of volatility model

$$dX_t = \alpha(\beta - X_t)dt + \sigma X_t^{\gamma_t} dW_{1,t}^H, \tag{6.10}$$

$$d\gamma_t = (\alpha_1 + \beta_1\gamma_t)dt + \sigma_1\sqrt{\gamma_t}(\rho W_{1,t}^H + \sqrt{1-\rho^2}B_{1,t}^H) \tag{6.11}$$

where $W_{1,t}^H$ and $B_{1,t}^H$ are two correlated fractional Levy processes and ρ is the correlation between the interest rate and elasticity processes. In order to estimate γ_t and its parameters based on the interest rate data, one can use the stochastic filtering method. It would be very interesting to estimate other parameters in the hybrid model which is a very complex problem. We postpone this work to a future publication.

Concluding Remarks: In order to test hypotheses and obtain confidence intervals of unknown model parameters, rate of normal approximation in the Kolmogorov distance is needed. Heston model is a popular stochastic volatility model which is an improvement of the Black-Scholes model with constant volatility. The option pricing formula involves the parameters in the volatility process which must be estimated from the return data. Rate of normal approximation in the Kolmogorov distance for the minimum contrast estimator in the Heston model had not been studied earlier in the literature. Our aim in this paper was to bridge this gap. First we studied this problem in Heston model without correlation. Then we generalized the problem to Heston model with correlation. Then we studied models with Levy type noise with jumps. Finally we considered rough volatility models with stochastic correlation having long memory and jumps.

References

[1] S.L. Heston, A closed-form solution for options with stochastic volatility, Rev. Financ. Stud. 6 (1993) 327-343.

[2] O.E. Barndorff-Nielsen, N. Shephard, Non-Gaussian Ornstein–Uhlenbeck-based models and some of their uses in financial economics (with discussion), J. R. Stat. Soc. Ser. B. 63 (2001) 167-241.

[3] O.E. Barndorff-Nielsen, N. Shephard, Econometric analysis of realised covariation: high frequency based covariance, regression and correlation in financial economics, Econometrica. 72 (2004a) 885-925.

[4] O.E. Barndorff-Nielsen, N. Shephard, Power and bipower variation with stochastic volatility and jumps (with discussion), J. Financ. Econ. 2 (2004b) 1-48.

[5] J.H.C. Woerner, Estimation of integrated volatility in stochastic volatility models. Appl. Stoch. Models Bus. Ind. 21 (2005) 27-44.

[6] J. Jacod, Asymptotic properties of realized power variations and related functionals of semimartingales. Stoch. Process. Appl. 118 (2008) 517-559.

[7] J. Jacod, M. Reiss, A remark on the rates of convergence for integrated volatility estimation in the presence of semimartingales, Ann. Stat. 42 (2014) 1134-1144.

[8] J.P.N. Bishwal, Parameter estimation in stochastic differential equations, Lecture Notes in Mathematics, Vol. 1923, Berlin: Springer-Verlag, (2008).

[9] J.P.N. Bishwal, Parameter estimation in stochastic volatility models, Cham, Switzerland: Springer Nature, (2022).
[10] V.V. Petrov, Limit theorems of probability theory, Oxford: Oxford University Press, (1995).
[11] W. Feller, An introduction to probability theory and its applications, Vol. I. New York: Wiley, (1957).
[12] R. Michel, J. Pfanzagl, The accuracy of the normal approximation for minimum contrast estimate, Zeit. Wahr. Verw. Gebiete. 18 (1971) 73-84.
[13] J.P.N. Bishwal, A. Bose, Rates of convergence of approximate maximum likelihood estimators in the Ornstein-Uhlenbeck process, Comp. Math. Appl. 42 (2001) 23-38.
[14] J.P.N. Bishwal, A. Bose, Speed of convergence of the maximum likelihood estimator in the Ornstein-Uhlenbeck process, Calc. Stat. Assoc. Bull. 45 (1995) 245-251.
[15] J. Pitman, M. Yor, A decomposition of Bessel bridge, Zeit. Wahr. Verw. Gebiete. 59 (1982) 425-457.
[16] I.I. Gikhman, A.V. Skorohod, Stochastic Differential Equations, Berlin: Springer-Verlag, (1972).
[17] A.E.D. Veraart, L.A.M. Veraart, Stochastic volatility and stochastic leverage, Ann. Finance. 8 (2012) 205-233.
[18] G. Kitagawa, Non-Gaussian state space modelling of nonstationary time series (with discussion), J. Amer. Stat. Assoc. 82 (1987) 1032-1063.
[19] J.P.N. Bishwal, A new algorithm for approximate maximum likelihood estimation in sub-fractional Chan-Karolyi-Longstaff-Sanders model, Asian J. Prob. Stat. 13 (2021) 62-88.
[20] A.W. Van der Vaart, Asymptotic statistics, Cambridge University Press, Cambridge, (2000).
[21] T.G. Conley, L.P. Hansen, E.G.J. Luttmer, J.A. Scheinkman, Short term interest rates as subordinated diffusions, Rev. Financ. Stud. 10 (1997) 525-578.

Appendix

Proof of Lemma 3.3 Let $g_i(t) := U_{t_{i-1}} - U_t$ for $t_{i-1} \le t < t_i, \quad i = 1, 2, \ldots, n$. By Theorem 4 of Gikhman and Skorohod [16, p. 48], there exists $C > 0$ such that

$$E|U_{t_{i-1}} - U_t|^{2k} \le C(t_{i-1} - t)^k, k = 1, 2, \ldots, \tag{A.1}$$

hence

$$\begin{aligned}
& E|Y_{n,T} - Y_T|^2 \\
= \;& E|\sum_{i=1}^{n} U_{t_{i-1}}[W_{t_i} - W_{t_{i-1}}] - \int_0^T U_t dW_t|^2 \\
= \;& E|\int_0^T g_i(t) dW_t|^2 \\
= \;& \int_0^T E(g_i^2(t))dt \\
\le \;& C \sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} |t_{i-1} - t| dt \\
= \;& Cn\frac{(t_i - t_{i-1})^2}{2} = C\frac{T^2}{n}.
\end{aligned}$$

This completes the proof of (a).

Next we prove (b). Using (2.1) and the fact that

$$U_{t_i} - U_{t_{i-1}} = \int_{t_{i-1}}^{t_i} \theta U_t dt + W_{t_i} - W_{t_{i-1}}$$

we obtain

$$\begin{aligned}
& E|Z_{n,T} - Z_T|^2 \\
= \; & E|\sum_{i=1}^{n} U_{t_{i-1}}[U_{t_i} - U_{t_{i-1}}] - \int_0^T U_t dW_t|^2 \\
= \; & E|\sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} \theta U_t U_{t_{i-1}} dt + \sum_{i=1}^{n} U_{t_{i-1}}[W_{t_i} - W_{t_{i-1}}] \\
& - \int_0^T \theta U_t^2 dt - \int_0^T U_t dW_t|^2 \\
\leq \; & 2E|\sum_{i=1}^{n} U_{t_{i-1}}[W_{t_i} - W_{t_{i-1}}] - \int_0^T U_t dW_t|^2 \\
& + 2\theta^2 E|\sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} U_t[U_{t_{i-1}} - X_t]dt|^2. \\
=: \; & N_1 + N_2.
\end{aligned}$$

N_1 is $O(\frac{T^2}{n})$ by Lemma 3.2(a) in Bishwal and Bose [13].

To estimate N_2 let $\psi_i(t) := U_t[U_{t_{i-1}} - U_t]$ for $t_{i-1} \leq t < t_i, \; i = 1, 2, \ldots, n$. Then

$$\begin{aligned}
& E|\sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} \psi_i(t)dt|^2 \\
= \; & \sum_{i=1}^{n} E|\int_{t_{i-1}}^{t_i} \psi_i(t)dt|^2 + 2 \sum_{i,j=1, i<j}^{n} E\left[\int_{t_{i-1}}^{t_i} \psi_i(t)dt \int_{t_{i-1}}^{t_i} \psi_j(s)ds\right] \\
=: \; & D_1 + D_2.
\end{aligned}$$

By the boundedness of $E(U_t^4)$ and (A.1) we have

$$\begin{aligned}
& E(\psi_i^2(t)) \\
= \; & E\{U_t^2[U_{t_{i-1}} - U_t]^2\} \\
\leq \; & \{E(U_t^4)\}^{1/2}\{E[U_{t_{i-1}} - U_t]^4\}^{1/2} \\
\leq \; & C(t_{i-1} - t).
\end{aligned}$$

Note that

$$\begin{aligned}
D_1 \; = \; & \sum_{i=1}^{n} E|\int_{t_{i-1}}^{t_i} \psi_i(t)dt|^2 \\
\leq \; & \sum_{i=1}^{n} (t_i - t_{i-1}) \int_{t_{i-1}}^{t_i} E(\psi_i^2(t))dt \\
\leq \; & C\frac{T}{n} \sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} (t - t_{i-1})dt \\
\leq \; & C\frac{T}{n} \sum_{i=1}^{n} (t_i - t_{i-1})^2 = C\frac{T^3}{n^2}
\end{aligned}$$

and

$$\begin{aligned}
D_2 \; = \; & 2 \sum_{i,j=1, \; i<j}^{n} E \int_{t_{i-1}}^{t_i} \int_{t_{j-1}}^{t_i} [\psi_i(t)\psi_j(s)]dtds \\
= \; & 2 \sum_{i,j=1, \; i<j}^{n} \int_{t_{i-1}}^{t_i} \int_{t_{j-1}}^{t_j} E[\psi_i(t)\psi_j(s)]dtds.
\end{aligned}$$

By Wick's lemma, we have

$$E[\psi_i(t)\psi_j(s)]$$

$$\begin{aligned}
&= E[U_t(U_{t_{i-1}} - U_t)U_s(X_{t_{j-1}} - U_s)] \\
&= E[U_t(U_{t_{i-1}} - U_t)]E[U_s(U_{t_{j-1}} - U_s)] \\
&+ E[U_tU_s]E[(U_{t_{i-1}} - U_t)(U_{t_{j-1}} - U_s)] \\
&+ E[U_t(U(t_{j-1}) - U_s)]E[U_s(U_{t_{i-1}} - U_t)] \\
&=: A_1 + A_2 + A_3.
\end{aligned}$$

Note that

$$U_t = \int_0^t e^{\theta(t-u)} dW_u, t \geq 0.$$

Let $a := e^\theta$. For $s \geq t$, we have

$$\begin{aligned}
& E(U_tU_s) \\
&= E\left(\int_0^t e^{\theta(t-u)} dW_u\right)\left(\int_0^s e^{\theta(s-u)} dW_u\right) \\
&= \int_0^t e^{\theta(t+s-2u)} du \\
&= \frac{1}{2\theta}[a^{s+t} - a^{s-t}].
\end{aligned}$$

Observe that

$$\begin{aligned}
& E(U_t - U_{t_{i-1}})(U_s - U_{t_{j-1}}) \\
&= E(U_tU_s) - E(U_tU_{t_{j-1}}) - E(U_{t_{i-1}}U_s) + E(U_{t_{i-1}}U_{t_{j-1}}) \\
&= \frac{1}{2\theta}(a^s - a^{t_{j-1}})[(a^t - a^{t_{i-1}}) + (a^{-t_{i-1}} - a^{-t})] \\
&= \frac{1}{2\theta}(s - t_{j-1})a^{t^*}[(t - t_{i-1}a^{t^{**}} + (t - t_{i-1})a^{-t^{***}}] \\
& (\text{where } t_{j-1} < t^* < s, t_{i-1} < t^{**}, t^{***} < t) \\
&\leq \frac{1}{2\theta}(s - t_{j-1})a^t(t - t_{i-1})a^{t_{i-1}} + (s - t_{j-1})a^t(t - t_{i-1})a^{-t}] \\
&\leq C(s - t_{j-1})(t - t_{i-1}).
\end{aligned}$$

Thus

$$A_2 \leq C(s - t_{j-1})(t - t_{i-1})$$

since $|E(U_tU_s)|$ is bounded.

Next

$$\begin{aligned}
& |E[U_t(U_{t_{i-1}} - U_t)]| \\
&= \frac{1}{2|\theta|}[a^{t+t_{i-1}} - a^{t-t_{i-1}} - a^{2t} + 1] \\
&= \frac{1}{2|\theta|}a^t[a^{t_{i-1}} - a^{-t_{i-1}} - a^t + a^{-t}] \\
&\leq \frac{1}{2|\theta|}a^t(t - t_{i-1})[a^{t_{i-1}} + a^{-t}] \\
&\leq C(t - t_{i-1})
\end{aligned}$$

and

$$|E[U_s(U_s - U_{t_{j-1}}]|$$

$$\begin{aligned}
&= \frac{1}{2|\theta|}[a^{2s} - 1 - a^{s+t_{j-1}} + a^{s+t_{j-1}}] \\
&= \frac{1}{2|\theta|}a^s[a^s - a^{-s} - a^{t_{j-1}} + a^{-t_{j-1}}] \\
&\leq \frac{1}{2|\theta|}a^s(s - t_{j-1})[a^{t_{j-1}} + a^{-s}] \\
&\leq C(s - t_{j-1}).
\end{aligned}$$

Thus

$$A_1 \leq C(s - t_{j-1})(t - t_{i-1}).$$

Next

$$\begin{aligned}
&|E[U_t(U_s - U_{t_{j-1}}]| \\
&= \frac{1}{2|\theta|}[a^{s+t} - a^{s-t} - a^{t+t_{j-1}} + a^{t_{j-1}-t}] \\
&= \frac{1}{2|\theta|}a^t(a^s - a^{t_{j-1}}) \\
&\leq \frac{1}{2|\theta|}a^t(1 - a^{-2t})(s - t_{j-1})a^t \\
&\leq (a^{2t} - 1)(s - t_{j-1}) \\
&\leq C(s - t_{j-1})
\end{aligned}$$

and

$$\begin{aligned}
&|E[U_s(U_t - U_{t_{i-1}}]| \\
&= \frac{1}{2|\theta|}[a^{t+s} - a^{s-t} - a^{s+t_{i-1}} + a^{s-t_{i-1}}] \\
&= \frac{1}{2|\theta|}a^s[a^t - a^{-t} - a^{t_{i-1}} + a^{-t_{i-1}}] \\
&\leq \frac{1}{2|\theta|}a^s(t - t_{i-1})[a^{t_{i-1}} + a^{-t}] \\
&\leq C(t - t_{i-1}).
\end{aligned}$$

Thus

$$A_3 \leq C(s - t_{j-1})(t - t_{i-1}).$$

Hence

$$E[\psi_i(t)\psi_j(s)] \leq C(s - t_{j-1})(t - t_{i-1}).$$

Thus

$$\begin{aligned}
D_2 &= 2\sum_{i,j=1,i<j}^{n} \int_{t_{i-1}}^{t_i}\int_{t_{j-1}}^{t_j} E[\psi_i(t)\psi_j(s)]dtds \\
&\leq C\sum_{i,j=1,i<j}^{n} \int_{t_{i-1}}^{t_i}\int_{t_{j-1}}^{t_j} (t - t_{i-1})(s - t_{j-1})dtds \\
&= C\sum_{i,j=1,i<j}^{n} (t_{i-1} - t_i)^2(t_{j-1} - t_i)^2
\end{aligned}$$

$$= \quad Cn^2\left(\frac{T}{n}\right)^4 = C\frac{T^4}{n^2}.$$

Hence, N_2 is $O(\frac{T^3}{n^2})$. Combining N_1 and N_2 completes the proof of (b).

Next we prove (c). Let $\chi_i(t) := U_{t_{i-1}}^2 - U_t^2, t_{i-1} \le t < t_i,\ i = 1, 2, \ldots, n.$ Then

$$\begin{aligned}
& E|I_{n,T} - I_t|^2 \\
= \quad & E|\sum_{i=1}^n U_{t_{i-1}}^2(t_i - t_{i-1}) - \int_0^T U_t^2 dt|^2 \\
= \quad & E|\sum_{i=1}^n \int_{t_{i-1}}^{t_i} [U_{t_{i-1}}^2 - U_t^2] dt|^2 \\
= \quad & E|\sum_{i=1}^n \int_{t_{i-1}}^{t_i} \chi_i(t) dt|^2 \\
= \quad & \sum_{i=1}^n E|\int_{t_{i-1}}^{t_i} \chi_i(t) dt|^2 + 2\sum_{i,j=1,i<j}^n E\int_{t_{i-1}}^{t_i}\int_{t_{j-1}}^{t_j} \chi_i(t)\chi_j(s) dt ds \\
=: \quad & B_1 + B_2.
\end{aligned}$$

Thus

$$\begin{aligned}
E\chi_i^2(t) \quad &= \quad E[U_{t_{i-1}}^2 - U_t^2]^2 \\
&= \quad E[U_{t_{i-1}} - U_t]^2[U_{t_{i-1}} + U_t]^2 \\
&\le \quad \{E[U_{t_{i-1}} - U_t]^4\}^{1/2}\{\{E[U_{t_{i-1}} + U_t]^4\}^{1/2} \\
&\le \quad C(t - t_{i-1})
\end{aligned}$$

(by (A.1) and the boundedness of the second term)

$$\begin{aligned}
B_1 \quad &= \quad \sum_{i=1}^n E|\int_{t_{i-1}}^{t_i} \chi_i(t) dt|^2 \le \sum_{i=1}^n (t_i - t_{i-1}) \int_{t_{i-1}}^{t_i} E(\chi_i^2(t)) dt \\
&\le \quad C\frac{T}{n}\sum_{i=1}^n \int_{t_{i-1}}^{t_i} (t - t_{i-1}) dt = C\frac{T^3}{n^2}.
\end{aligned}$$

Note that

$$\begin{aligned}
& E[\chi_i(t)\chi_j(s)] \\
= \quad & E(U_{t_{i-1}}^2 - U_t^2)(U_{t_{j-1}}^2 - U_s^2) \\
= \quad & E(U_{t_{i-1}} - U_t)(U_{t_{i-1}} + U_t)(U_{t_{j-1}} - U_s)(U_{t_{j-1}} + U_s).
\end{aligned}$$

Now using Wick's lemma and proceeding similar to the estimation of M_2 it is easy to see that

$$B_2 \le C\frac{T^4}{n^2}.$$

Combining B_1 and B_2, (c) follows.

Next we prove (d). Let $h_i(t) := U_{t_{i-1}} - U_t$ for $t_{i-1} \le t < t_i, \quad i = 1, 2, \dots, n$. Observe that

$$\begin{aligned}
& E|R_{n,T} - I_T|^2 \\
= \;& E|R_{n,T} - I_{n,T} + I_{n,T} - I_T|^2 \\
\le \;& 2[E|I_{n,T} - R_{n,T}|^2 + 2E|I_{n,T} - I_T|^2] \\
=: \;& 2G_{11} + 2E\left|\sum_{i=1}^{n} U_{t_{i-1}}(t_i - t_{i-1}) - \int_0^T U_t\, dt\right|^2 \\
= \;& 2G_{11} + E\left|\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i} [U_{t_{i-1}} - U_t]\, dt\right|^2 \\
= \;& 2G_{11} + 2E\left|\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i} h_i(t)\, dt\right|^2 \\
= \;& 2G_{11} + 2\sum_{i=1}^{n} 2E\left|\int_{t_{i-1}}^{t_i} h_i(t)\, dt\right|^2 + 4\sum_{i,j=1,\, i<j} E\int_{t_{i-1}}^{t_i}\int_{t_{j-1}}^{t_j} h_i(t)h_j(s)\, dt\, ds \\
=: \;& 2G_{11} + 2B_{11} + 2B_{12}.
\end{aligned}$$

We first estimate $G_{11} = E|R_{n,T} - I_{n,T}|^2$. We have

$$S_{t_i} - S_{t_{i-1}} = \int_{t_{i-1}}^{t_i} \mu S_t dt + 2\int_{t_{i-1}}^{t_i} \sqrt{X_t} S_t dW_t.$$

Hence

$$(S_{t_i} - S_{t_{i-1}})^2 = (S_{t_i} - S_{t_{i-1}})\left(\int_{t_{i-1}}^{t_i} \mu S_t dt + 2\int_{t_{i-1}}^{t_i} \sqrt{X_t} S_t dW_t\right).$$

This gives

$$\sum_{i=1}^{n}(S_{t_i} - S_{t_{i-1}})^2 = S_T - \sum_{i=1}^{n} S_{t_{i-1}}(S_{t_i} - S_{t_{i-1}}) = S_T - \sum_{i=1}^{n} S_{t_{i-1}}\int_{t_{i-1}}^{t_i} dS_t.$$

Observe that by Itô formula,

$$\sum_{i=1}^{n}(S_{t_i} - S_{t_{i-1}})^2 - \int_0^T S_t X_t dt = \sum_{i=1}^{n}(S_{t_i} - S_{t_{i-1}})\int_{t_{i-1}}^{t_i} dS_t - \int_0^T S_t X_t dt = \sum_{i=1}^{n}\int_{t_{i-1}}^{t_i}(S_t - S_{t_{i-1}})dS_t$$

$$= \sum_{i=1}^{n}(S_{t_i} - S_{t_{i-1}})\int_{t_{i-1}}^{t_i}(\mu S_t dt + \sqrt{X}_t S_t dW_t) - \int_0^T S_t X_t dt$$

$$= \sum_{i=1}^{n} S_{t_i}\int_{t_{i-1}}^{t_i} dS_t + \sum_{i=1}^{n}\int_{t_{i-1}}^{t_i}(\mu S_t - S_{t_{i-1}})dt + \sum_{i=1}^{n}\int_{t_{i-1}}^{t_i}\sqrt{X}_t(S_t - S_{t_{i-1}})dW_t - \int_0^T S_t X_t dt.$$

$$\sum_{i=1}^{n}(S_{t_i} - S_{t_{i-1}})^2 - I_T = 2\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i}\mu(S_t - S_{t_{i-1}})dt + 2\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i}(S_t - S_{t_{i-1}})\sqrt{X_t} dW_t.$$

Hence

$$R_{n,T} - I_{n,T} = \sum_{i=1}^{n}\int_{t_{i-1}}^{t_i}(U_{t_{i-1}} - U_t)dW_t.$$

Observe that

$$S_{t_{i-1}} - S_t = \int_t^{t_{i-1}} \mu S_r dr + 2\int_t^{t_{i-1}} \sqrt{X_r} S_r dW_r.$$

Thus

$$G_{11} = E|R_{n,T} - I_T|^2 = 4E\left|\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i}(S_{t_{i-1}} - S_t)\sqrt{X_t} dW_t\right|^2$$

$$= 4E\left|\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i}\sqrt{X_t}\left(\int_t^{t_{i-1}}\mu S_r dr + 2\int_t^{t_{i-1}}\sqrt{X_r}S_r dW_r\right)dW_t\right|^2$$

$$\leq C\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i}(t-t_{i-1})^2 dt \leq C\frac{T^4}{n^2}.$$

Since

$$\begin{aligned} & Eh_i^2(t) = E[U(t_{i-1}) - U_t]^2 = E[U(t_{i-1}) - U_t][U(t_{i-1}) + U_t] \\ \leq & \left\{E[U(t_{i-1}) - U_t]^2\right\}^{1/2}\left\{E[U(t_{i-1}) + U_t]^2\right\}^{1/2} \\ \leq & C(t - t_{i-1}) \end{aligned}$$

(by (A.1) and the boundedness of the second term).

$$\begin{aligned} B_{11} &= \sum_{i=1}^{n} E\left|\int_{t_{i-1}}^{t_i} h_i(t)\, dt\right|^2 \leq \sum_{i=1}^{n}(t_i - t_{i-1})\int_{t_{i-1}}^{t_i} E(h_i^2(t))\, dt \\ &\leq C\frac{T}{n}\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i}(t - t_{i-1})\, dt = C\frac{T^3}{n^2}. \end{aligned}$$

Further,

$$\begin{aligned} & E[h_i(t)h_j(s)] \\ = & E(U_{t_{i-1}} - U_t)(U_{t_{j-1}} - U_s) \\ = & E(U_{t_{i-1}} - U_t)(U_{t_{i-1}} + U_t)(U_{t_{j-1}} - U_s)(U_{t_{j-1}} + U_s). \end{aligned}$$

Now, using Wick's lemma and proceeding similar to the estimation of D_2 it is easy to see that

$$B_{12} \leq C\frac{T^4}{n^2}.$$

Combining bounds for B_{11}, B_{12} and G_{11}, part (d) of the lemma follows. This completes the proof of the lemma. □