DOI: 10.28919/ejma.2024.4.22

Eur. J. Math. Appl. (2024)4:22
URL: http://ejma.euap.org

ON THE KOLMOGOROV DISTANCE FOR THE ESTIMATORS IN THE COX-INGERSOLL-ROSS MODEL

JAYA P. N. BISHWAL

Abstract. We study the bounds on the Kolmogorov distance of some new estimators of the Cox-Ingersoll-Ross model. First we obtain the rate of weak convergence of the distribution of the normalized minimum contrast estimator of the drift parameter based on continuous observation which are of theoretical interest. Then we obtain the rates of normal approximation of the normalized approximate minimum contrast estimators when the process is densely observed at discrete time points which are of practical interest in finance and biology. The approximation, which could have independent interest, is based on Hausdorff moment problem.

1. Introduction

There are close connections between some models in biology and finance. Feller [20] reached at the square-root process as the weak limit of Galton-Watson branching process with immigration while studying a problem in genetics. Using the Feller's square-root process, Cox *et al.* [17] studied the theory of term structure of interest rates and the model is now known as the Cox-Ingersoll-Ross model. Overbeck and Ryden [30] studied asymptotics of conditional least squares estimators of Cox-Ingersoll-Ross process from discrete observations using an auto-regressive type representation of the model with non-Gaussian error. Dehtiar *et al.* [18] studied strong consistency for the maximum likelihood method and an alternative method of estimation of the drift parameters of the Cox-Ingersoll-Ross process based on continuous observations. Mishura and Yurchenko-Tytarenko [29] studied hitting probability of fractional Cox-Ingersoll-Ross model which involves long memory. Mackevicius [27] used stochastic Verhulst model as an alternative to CIR model for modeling interest rate as both processes have similar behavior. Mackevicius [26] studied weak approximation of CIR equation by discrete random variables. Lenkasas and Mackevicius [24] obtained a second order weak approximation of Heston model by discrete random variables. Lileika and Mackevicius [25] studied weak approximation of CKLS and CEV process (cf. Cox [16]) by discrete random variables. The Cox-Ingersoll-Ross (CIR) model is extensively used as a short rate mean reverting model in term structure of interest rates and a stochastic volatility process in the Heston model, see Bishwal [9]. In view of this, it becomes necessary to estimate the unknown parameters in the model from discrete data. See

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, Cox-Ingersoll-Ross process, minimum contrast estimator, approximate minimum contrast estimators, uniform rate of weak convergence, Kolmogorov distance, Fourier method.

Received 24/06/2024.

Bishwal [8] for asymptotic results on approximate likelihood asymptotics and approximate Bayes asymptotics for drift estimation of discretely observed diffusions based on high frequency data.

In this paper, we assume constant volatility and without loss of generality assume it to be one. To estimate the drift parameter, we adopt minimum contrast method and study the accuracy of distributional approximation by estimating the Kolmogorov distance of the resulting estimators both from continuous data and high frequency discrete data.

2. Continuous Observation

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

$$dX_t = (\alpha + \beta X_t)\, dt + \ 2\sqrt{X_t}\, dW_t,\ t \geq 0, \quad X_0 = 1 \tag{2.1}$$

where $\{W_t\}$ is a standard Wiener process with the filtration $\{\mathcal{F}_t\}_{t\geq 0}$, $\alpha > 0$ 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

$$u_t = 2xu_{xx} + \alpha u_x - \left(\frac{\mu}{x} + \lambda x\right) u \tag{2.2}$$

is given by

$$q(t, x, y, \alpha, \beta) := -2\beta \left(\frac{y}{x}\right)^{\alpha - \frac{1}{2}} \frac{e^{(\frac{1}{2}-\alpha)\beta t}}{1 - e^{\beta t}} \exp\left[\frac{2\beta(x+y)}{e^{-\beta t} - 1}\right] I_\nu \left(\frac{-2\beta\sqrt{xy}}{\sinh(-\frac{1}{2}\beta t)}\right) \tag{2.3}$$

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

Let the continuous realization $\{X_t, 0 \leq t \leq T\}$ 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

$$L_T(\beta, \alpha) := \frac{dP_{\beta,\alpha}^T}{dP_0^T}(X_0^T) = \exp\left\{\int_0^T \frac{\alpha + \beta X_t}{4X_t} dX_t - \int_0^T \frac{(\alpha + \beta X_t)^2}{8X_t} dt\right\}. \tag{2.4}$$

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

$$\gamma_T(\beta, \alpha) := \left\{\int_0^T \frac{\alpha + \beta X_t}{4X_t} dX_t - \int_0^T \frac{(\alpha + \beta X_t)^2}{8X_t} dt\right\}. \tag{2.5}$$

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 $\alpha \geq 2$, the MLE $\hat{\alpha}_T$ is inconsistent. It remains open to find a consistent estimator in this case.

As far as we know, the rate of normal approximation for the minimum contrast estimator has not been studied earlier. Our aim in this paper is to bridge this gap.

Consider the minimum contrast estimates (MCE)

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

and

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

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 $O(T^{-1/4})$ (see Borokov [14]). 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 Pfanzagl [28].

Observe that

$$\left(\frac{T\alpha}{-4\beta}\right)^{1/2}(\widetilde{\beta}_T - \beta) = \frac{\left(\frac{-4\beta}{T\alpha}\right)^{1/2} N_T}{\left(\frac{-4\beta}{T\alpha}\right) I_T} \tag{2.6}$$

and

$$\left(\frac{T\beta}{-4(\alpha-2)}\right)^{1/2}(\widetilde{\alpha}_T - \alpha) = \frac{\left(\frac{-4(\alpha-2)}{T\beta}\right)^{1/2} M_T}{\left(\frac{-4(\alpha-2)}{T\beta}\right) J_T}. \tag{2.7}$$

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 $\{X_t, 0 \le t \le T\}$ the continuous conditional least squares estimators of β and α are respectively given by

$$\beta_T := \frac{\int_0^T X_s dX_s - (X_T - X_0)\widetilde{X_T}}{\int_0^T (X_t - \widetilde{X_T})^2 dt}, \tag{2.8}$$

$$\alpha_T := -\widetilde{X_T}\beta_T + T^{-1}(X_T - X_0) \tag{2.9}$$

where

$$\widetilde{X_T} := \int_0^T X_t dt. \tag{2.10}$$

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

$$\begin{aligned} \beta_T &= \frac{T\widetilde{X_T}}{\int_0^T (X_t - \widetilde{X_T})^2 dt} + o_P(T^{-1/2}) \\ &= \frac{T\widetilde{X_T}}{2(\widetilde{X_T^2} - (\widetilde{X_T})^2)} + o_P(T^{-1/2}) \end{aligned} \tag{2.12}$$

$$\alpha_T = \frac{\widetilde{X_T^2}}{2(\widetilde{X_T^2} - (\widetilde{X_T})^2)} + o_P(T^{-1/2}) \tag{2.13}$$

where

$$\widetilde{X_T^2} := \int_0^T X_t^2 dt \tag{2.14}$$

We define the minimum contrast estimators as

$$\check{\beta}_T := \frac{T\widetilde{X_T}}{2(\widetilde{X_T^2} - (\widetilde{X_T})^2)}, \tag{2.15}$$

$$\check{\alpha}_T := \frac{\widetilde{X_T^2}}{2(\widetilde{X_T^2} - (\widetilde{X_T})^2)}. \tag{2.16}$$

We digress a bit and provide some closely related results to our problem.

Laplace and Fourier Transforms, and Cameron-Martin Theorems

The case $\beta = 0$.

Recall that the Kummer's confluent hypergeometric function is given by

$${}_1F_1(r, s, z) = \frac{\Gamma(s)}{\Gamma(r)\Gamma(s-r)} z^{1-s} \int_0^z e^u u^{r-1}(z-u)^{s-r-1} du. \tag{2.17}$$

and the Whittaker function of first kind is given by

$$M_{s,r}(z) = z^{r+\frac{1}{2}} e^{-z/2} \ {}_1F_1(r - s + \frac{1}{2}, 2r+1, z).$$

The following three propositions are from Ben Alaya and Kebaier [1].

Proposition 2.1 Let$\nu := \frac{1}{\sigma}\sqrt{(\alpha-\sigma)^2 + 4u}$. Recall that $J_T := \int_0^T X_t^{-1} dt$. We have

$$E \exp(-uJ_T) = \frac{\Gamma(k + \frac{\nu}{2} + \frac{1}{2})}{\Gamma(\nu+1)} (\frac{x}{\sigma t})^{\frac{\nu}{2}+\frac{1}{2}-k} \exp(-\frac{x}{\sigma t}) \ {}_1F_1(k + \frac{\nu}{2} + \frac{1}{2}, \frac{x}{\sigma t}).$$

Proposition 2.2 We have

$$E \exp(-uI_T - vJ_T) = \frac{\Gamma(k + \frac{\nu}{2} + \frac{1}{2})}{\Gamma(\nu+1)} \left(\frac{\sqrt{\sigma u}x \coth(\sqrt{\sigma u}t)}{\sigma}\right)^{-k} \exp\left(\frac{\sqrt{\sigma u}x}{\sigma} \coth(\sqrt{\sigma u}t)\right)$$
$$\times \exp\left(\frac{\sqrt{\sigma u}x}{2\sigma \sinh(\sqrt{\sigma u}t) \cosh\sqrt{\sigma u}t}\right) M_{-k,\frac{\nu}{2}}\left(\frac{\sqrt{\sigma u}x}{\sigma \sinh(\sqrt{\sigma u}t)\cosh\sqrt{\sigma u}t}\right).$$

Proposition 2.3 Let $\rho := 2\sqrt{\sigma u}$. We have

$$E \exp\left(-uX_T - vI_T\right) = \left(\frac{2\sigma u}{\rho} \sinh(\frac{\rho t}{2}) + \cosh(\frac{\rho t}{2})\right)^{-a/\sigma} \frac{\frac{2v}{\rho}\sinh(\frac{\rho t}{2}) + u\cosh(\frac{\rho t}{2})}{\frac{2\sigma u}{\rho}\sinh(\frac{\rho t}{2}) + \cosh(\frac{\rho t}{2})}.$$

The case $\beta \neq 0$.

The following characteristic function of I_T is closely associated with *Levy's stochastic area formula* and is well known from Brownian motion literature and also from the work of Cox, Ingersoll and Ross [17].

Proposition 2.4 a) Let $\phi_T(u) := E\exp(iuI_T), u \in \mathbb{R}$. Then

$$E\exp(iuI_T) = \exp\left(\frac{2iu}{\beta + \gamma\coth\frac{\gamma T}{2}}\right)\left[\cosh\frac{\gamma T}{2} + \frac{\beta}{\gamma}\sinh\frac{\gamma T}{2}\right]^{-1}$$

where $\gamma := (\beta^2 - 2iu)^{1/2}$.

b) We have

$$\lim_{T\to\infty} E\left(e^{-uI_T}\right) = \lim_{T\to\infty}\exp\left(-\frac{ab}{2\sigma}T\left(\sqrt{1+\frac{4u\sigma}{b^2}}-1\right)\right).$$

Proposition 2.5

$$E\exp(-uX_T - vI_T) = \left(\frac{2\rho e^{(b-\rho)T/2}}{2\sigma u(1-e^{-\rho T}) + (\rho - b)e^{-\rho T} + (\rho + b)}\right)^{a/\sigma}$$
$$\times\exp\left(\frac{u((\rho+b)e^{-\rho T} + (\rho - b) + 2v(1-e^{-\rho T})}{2\sigma u(1-e^{-\rho T}) + (\rho - b)e^{-\rho T} + (\rho + b)}\right)$$

where $\rho =: \sqrt{b^2 + 4\sigma v}$.

Proposition 2.6

$$E\exp(-uJ_T) = \frac{\Gamma(k+\frac{\nu}{2}+\frac{1}{2})}{\Gamma(\nu+1)}(\frac{x}{\alpha})^{-k}\beta^{\frac{\nu}{2}+\frac{1}{2}}\exp(\frac{b}{2\sigma}[at - \frac{2x}{e^{bt}-1}])_1F_1(k+\frac{\nu}{2}+\frac{1}{2}, \nu+1, \beta)$$

where

$$k =: \frac{a}{2\sigma},\ \alpha =: \frac{be^{bt}}{\sigma(e^{bt}-1)},\ \nu =: \frac{1}{\sigma}\sqrt{(\alpha-\sigma)^2 + 4u\sigma}$$

and ${}_1F_1$ is Kummer's confluent hypergeometric function.

Bond and Option Pricing

Here we give the bond price formula since it is closely connected to the Laplace transform to the integrated interest rate and our mail tool for obtaining the bound on the Kolmogorov distance is the characteristic function of the integrated interest rate. Further we also give the bond price formula for the Vasicek model.

Proposition 2.7 For the CIR model

$$dX_t = a(b - X_t)dt + \sigma\sqrt{X_t}dW_t$$

the price at time t of a zero-coupon bond that pays \$1 at time T is given by

$$P(t,T) = E_Q\left(e^{-\int_t^T X_t dt}\right) = A(t,T)e^{-B(t,T)X_t}$$

where

$$B(t,T) = \frac{2(e^{\gamma(T-t)-1})}{(\gamma+a)(e^{\gamma(T-t)-1}) + 2\gamma},\quad A(t,T) = \left(\frac{2\gamma e^{(a+\gamma)(T-t)/2}}{(\gamma+a)(e^{\gamma(T-t)-1}) + 2\gamma}\right)^{2ab/\sigma^2}$$

with $\gamma = \sqrt{a^2 + 2\sigma^2}$ and Q is the risk-neutral measure.

Proposition 2.8 The European call option is given by

$$C_t = P(t,s)\chi^2\left(2\frac{\log(A(t,s)/K)}{B(T,s)}[\phi+\psi+B(T,s)]; \frac{4ab}{\sigma^2}, \frac{2\phi^2 X_t e^{\gamma(T-t)}}{\phi+\psi+B(T,s)}\right)$$
$$-KP(t,T)\chi^2\left(2\frac{\log(A(t,s)/K)}{B(T,s)}[\phi+\psi]; \frac{4ab}{\sigma^2}, \frac{2\phi^2 X_t e^{\gamma(T-t)}}{\phi+\psi}\right)$$

where

$$\phi =: \frac{2\gamma}{\sigma^2(e^{\gamma(T-t)-1})}, \quad \psi =: \frac{a+\gamma}{\sigma^2}$$

and $\chi^2(x; d, \lambda)$ is the noncentral chi-square distribution with d degrees of freedom and noncentrality parameter λ.

Proposition 2.9 For the Vasicek model for short rate

$$dr_t = a(b - r_t)dt + \sigma dW_t$$

the price of a zero coupon bond at time t maturing at time T is given by

$$P(t,T) = A(t,T)e^{-B(t,T)r(t)}$$

where

$$B(t,T) = \frac{1-e^{-a(T-t)}}{a}, \quad A(t,T) = \exp\left(\frac{(B(t,T)-T+t)(a^2b-\sigma^2/2)}{a^2} - \frac{\sigma^2 B(t,T)^2}{4a}\right).$$

Let $\Phi(\cdot)$ denote the standard normal distribution function. Throughout the paper, C denotes a generic constant (which does not depend on T and x).

Proposition 2.10 The interest rate derivative, European call option is given by

$$C = LP(0,s)\Phi(h) - KP(0,T)\Phi(h-\sigma_p)$$

where L is the bond principal, s is the bond maturity, T is the option maturity, K is the strike price,

$$h = \frac{1}{\sigma_P}\ln\frac{LP(0,s)}{P(0,T)K} + \frac{\sigma_P}{2}, \quad \sigma_P =: \frac{\sigma}{a}(1-e^{-a(s-T)})\sqrt{\frac{1-e^{-2aT}}{2a}}$$

When $a = 0$, $\sigma_P = \sigma(s-T)\sqrt{T}$.

See Brigo and Mercurio [15] and Hull [23].

Rate of Convergence and Esseen's Lemma

The first lemma gives the rate in an ergodic theorem and we omit the proof.

Lemma 2.1 a) For every $\delta > 0$,

$$P\left\{\left|\left(\frac{-4\beta}{T\alpha}\right)I_T - 1\right| \geq \delta\right\} \leq CT^{-1}\delta^{-2}.$$

b) For every $\delta > 0$,

$$P\left\{\left|\left(\frac{-4(\alpha-2)}{T\beta}\right)J_T - 1\right| \geq \delta\right\} \leq CT^{-1}\delta^{-2}.$$

We omit the details of the proof of the next lemma which uses the bounds on the characteristic functions given later in Lemma 2.3 and Lemma 2.7 along with Esseen's smoothing lemma.

Lemma 2.2

$$a)\quad \sup_{x\in\mathbb{R}}\left|P\left\{\left(\frac{-4\beta}{T\alpha}\right)^{1/2} N_T \le x\right\} - \Phi(x)\right| \le C\, T^{-1/2}.$$

$$b)\quad \sup_{x\in\mathbb{R}}\left|P\left\{\left(\frac{-4(\alpha-2)}{T\beta}\right)^{1/2} M_T \le x\right\} - \Phi(x)\right| \le C\, T^{-1/2}.$$

Lemma 2.3 Let

$$H_{T,x} := \left(\frac{-4\beta}{T\alpha}\right)^{1/2} N_T - \left(\frac{-4\beta}{T\alpha} I_T - 1\right) x.$$

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

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

Proof : Observe that

$$E\exp(iuI_T) = \exp\left(\frac{2iu(e^{\frac{\gamma T}{2}} - e^{\frac{-\gamma T}{2}})}{e^{\frac{\gamma T}{2}}(\gamma+\beta) + e^{-\frac{\gamma T}{2}}(\gamma-\beta)}\right) \times 2\left[e^{\frac{\gamma T}{2}}(\gamma+\beta) + e^{-\frac{\gamma T}{2}}(\gamma-\beta)\right]^{-1}. \quad (2.18)$$

Now consider

$$\begin{aligned} E\exp(iuH_{T,x}) &= E\exp\left[-iu\left(\tfrac{-4\beta}{T\alpha}\right)^{1/2} N_T - iu\left(\tfrac{-4\beta}{T\alpha}I_T - 1\right)x\right] \\ &= E\exp\left[-iu\left(\tfrac{-4\beta}{T\alpha}\right)^{1/2}\{\beta I_T - \alpha T\} - iu\left(\tfrac{-4\beta}{T\alpha}I_T - 1\right)x\right] \\ &= E\exp(z_1 I_T + z_3) =: \exp(z_3)\phi_T(z_1) \end{aligned} \quad (2.19)$$

where $z_1 := -iu\beta\delta_{T,x}$, $z_3 := \frac{iuT}{2}\delta_{T,x}$ with $\delta_{T,x} := \left(\frac{-4\beta}{T\alpha}\right)^{1/2} + \frac{2x}{T}$. Note that $\phi_T(z_1) = E(\exp z_1 I_T)$ satisfies (2.18) 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 for the Lemma are satisfied. Further, with $\varpi_T(u) := 1 + iu\frac{\delta_{T,x}}{\beta} + \frac{u^2\delta_{T,x}^2}{2\beta^2}$,

$$\gamma = (\beta^2 - 2z_1)^{1/2} = \beta\left[1 - \frac{z_1}{\beta^2} - \frac{z_1^2}{2\beta^4} + \frac{z_1^3}{2\beta^8} + \cdots\right] = \beta\left[1 + iu\frac{\delta_{T,x}}{\beta} + \frac{u^2\delta_{T,x}^2}{2\beta^2} + \frac{iu^3\delta_{T,x}^3}{2\beta^3} + \cdots\right]$$

$$= \beta[1 + \omega_{1,T}(u) + \omega_{2,T}(u) + \omega_{3,T}(u)] = \beta\varpi_T(u) + \omega_{3,T}(u) = \beta[1 + \omega_{1,T}(u)]. \quad (2.20)$$

Thus

$$\gamma - \beta = \omega_{1,T},\ \ \gamma + \beta = 2\beta + \omega_{1,T}. \quad (2.21)$$

Hence the above expectation equals

$$\exp\left(z_3 + \tfrac{\beta T}{2}\right) \times \left[\frac{2\beta\varpi_T(u) + \omega_{3,T}(u)}{\omega_{1,T}\exp\{-\beta T\varpi_T(u) + \omega_{4,T}(u)\} + (2\beta + \omega_{1,T}(u))\exp\{\beta T\varpi_T(u) + \omega_{4,T}(u)\}}\right]^{1/2}$$

$$= \left[\frac{1 + \omega_{1,T}(u)}{\omega_{1,T}\exp(\chi_T(u)) + (1 + \omega_{1,T}(u))\exp(\psi_T(t))}\right]^{1/2} \quad (2.22)$$

where

$$\chi_T(u) \;\; := \;\; -\beta T\beta_T(u) + \alpha_{4,T}(u) - 2z_3 - \beta T = -2\beta T + \omega_{1,T}(u) + t^2\omega_{1,T}(u), \quad (2.23)$$

$$\begin{aligned}\psi_T(u) &:= \beta T\varpi_T(u)+\omega_{4,T}(u)-2z_3-\beta T=\beta T\left[1+iu\frac{\delta_{T,x}}{\beta}+\frac{u^2\delta_{T,x}^2}{2\beta^2}\right]+\alpha_{4,T}(u)-iuT\delta_{T,x}-\beta T\\ &= \frac{u^2T\alpha}{-4\beta}\left[\left(\frac{-4\beta}{T\alpha}\right)^{1/2}+\frac{2x}{T}\right]^2=u^2+u^2\omega_{1,T}(u). \qquad (2.24)\end{aligned}$$

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

Hence the above expectation equals

$$\begin{aligned}&\exp(-\tfrac{u^2}{2})(1+\omega_{1,T})^{1/2}\left[\omega_{1,T}\exp\{-2\beta T+\omega_{1,T}+u^2\omega_{1,T}\}+(1+\omega_{1,T}(u))\exp\{u^2\omega_{1,T}(u)\}\right]^{-1/2}\\ &=\exp(-\tfrac{u^2}{2})\left[1+\omega_{1,T})(1+\omega_{1,T}(1+\omega_{1,T})\exp\{-\beta T+\omega_{1,T}+t^2\omega_{1,T}\}\right]\exp(u^2\omega_{1,T}(u)). \qquad (2.25)\end{aligned}$$

This completes the proof of the lemma. □

To obtain the rate of normal approximation for the MCE, we need the following estimate on the tail behavior of the MCE.

Lemma 2.4

$$P\left\{(\frac{T\alpha}{-4\beta})^{1/2}|\widetilde{\beta}_T-\beta|\ge 2(\log T)^{1/2}\right\}\le CT^{-1/2}.$$

Proof : Observe that

$$\begin{aligned}&P\left\{(\frac{T\alpha}{-4\beta})^{1/2}|\widetilde{\beta}_T-\beta|\ge 2(\log T)^{1/2}\right\}=P\left\{\left|\frac{(\frac{-4\beta}{T\alpha})^{1/2}N_T}{(\frac{-4\beta}{T\alpha})I_T}\right|\ge 2(\log T)^{1/2}\right\}\\ \le\;& P\left\{\left|(\frac{-4\beta}{T\alpha})^{1/2}N_T\right|\ge(\log T)^{1/2}\right\}+P\left\{\left|\frac{-4\beta}{T\alpha}I_T\right|\le\frac12\right\}\\ \le\;& \left|P\left\{(\frac{-4\beta}{T\alpha})^{1/2}|N_T|\ge(\log T)^{1/2}\right\}-2\Phi(-(\log T)^{1/2})\right|+2\Phi(-(\log T)^{1/2})+P\left\{\left|\frac{-4\beta}{T\alpha}I_T-1\right|\ge\frac12\right\}\\ \le\;& \sup_{x\in\mathbb{R}}\left|P\left\{(\frac{-4\beta}{T\alpha})^{1/2}|N_T|\ge x\right\}-2\Phi(-x)\right|\\ \le\;& \sup_{x\in\mathbb{R}}\left|P\left\{(\frac{-4\beta}{T\alpha})^{1/2}|N_T|\ge x\right\}-2\Phi(-x)\right|+2\Phi(-(\log T)^{1/2})+P\left\{\left|(\frac{-4\beta}{T\alpha})I_T-1\right|\ge\frac12\right\}\\ \le\;& CT^{-1/2}+C(T\log T)^{-1/2}+CT^{-1}\le CT^{-1/2}.\end{aligned}$$

The bounds for the first and the third terms come from Lemma 2.2 and Lemma 2.1 respectively and that for the middle term comes from Feller ([21], p. 166). □

Lemma 2.5

$$P\left\{\left(\frac{T\beta}{4(\alpha-2}\right)^{1/2}|\widetilde{\alpha}_T-\alpha|\ge 2(\log T)^{1/2}\right\}\le CT^{-1/2}.$$

Proof : Observe that

$$\begin{aligned}&P\left\{(\frac{T\beta}{4(\alpha-2)})^{1/2}|\widetilde{\alpha}_T-\alpha|\ge 2(\log T)^{1/2}\right\}=P\left\{\left|\frac{(\frac{2\alpha}{T})^{1/2}N_T}{(\frac{4(\alpha-2)}{T\beta})I_T}\right|\ge 2(\log T)^{1/2}\right\}\\ \le\;& P\left\{\left|(\frac{4(\alpha-2)}{T\beta})^{1/2}N_T\right|\ge(\log T)^{1/2}\right\}+P\left\{\left|\frac{4(\alpha-2)}{T\beta}I_T\right|\le\frac12\right\}\\ \le\;& \left|P\left\{(\frac{4(\alpha-2)}{T\beta})^{1/2}|N_T|\ge(\log T)^{1/2}\right\}-2\Phi(-(\log T)^{1/2})\right|+2\Phi(-(\log T)^{1/2})\end{aligned}$$

$$
\begin{aligned}
&+P\left\{\left|\frac{4(\alpha-2)}{T\beta}I_T-1\right|\geq\frac{1}{2}\right\}\leq\sup_{x\in\mathbb{R}}\left|P\left\{(\frac{4(\alpha-2)}{T\beta})^{1/2}|N_T|\geq x\right\}-2\Phi(-x)\right|\\
\leq\;&\sup_{x\in\mathbb{R}}\left|P\left\{(\frac{4(\alpha-2)}{T\beta})^{1/2}|N_T|\geq x\right\}-2\Phi(-x)\right|+2\Phi(-(\log T)^{1/2})+P\left\{\left|(\frac{4(\alpha-2)}{T\beta})I_T-1\right|\geq\frac{1}{2}\right\}\\
\leq\;&CT^{-1/2}+C(T\log T)^{-1/2}+CT^{-1}\leq CT^{-1/2}.
\end{aligned}
$$

The bounds for the first and the third terms come from Lemma 2.2 and Lemma 2.1 respectively and that for the middle term comes from Feller ([21], p. 166). □

Now we are ready to obtain the uniform rate of normal approximation of the distribution of the MCEs. We obtain the bound for the estimator of β assuming α is known.

Theorem 2.6

$$\sup_{x\in\mathbb{R}}\left|P\left\{\left(\frac{T\alpha}{-4\beta}\right)^{1/2}(\widetilde{\beta}_T-\beta)\leq x\right\}-\Phi(x)\right|\leq C_\beta T^{-1/2}.$$

Proof : We shall consider two possibilities (i) $|x|>2(\log T)^{1/2}$ and (ii) $|x|\leq 2(\log T)^{1/2}$.
(i) We shall give a proof for the case $x>2(\log T)^{1/2}$. The proof for the case $x<-2(\log T)^{1/2}$ runs similarly. Note that

$$\left|P\left\{(\frac{T\alpha}{-4\beta})^{1/2}(\widetilde{\beta}_T-\beta)\leq x\right\}-\Phi(x)\right|\leq P\left\{(\frac{T\alpha}{-4\beta})^{1/2}(\widetilde{\beta}_T-\beta)\geq x\right\}+\Phi(-x).$$

But $\Phi(-x)\leq\Phi(-2(\log T)^{1/2})\leq CT^{-2}$. See Feller (1957, p. 166).
Moreover by Lemma 2.4, we have

$$P\left\{(\frac{T\alpha}{-4\beta})^{1/2}(\widetilde{\beta}_T-\beta)\geq 2(\log T)^{1/2}\right\}\leq CT^{-1/2}.$$

Hence

$$\left|P\left\{(\frac{T\alpha}{-4\beta})^{1/2}(\widetilde{\beta}_T-\beta)\leq x\right\}-\Phi(x)\right|\leq CT^{-1/2}.$$

(ii) Let

$$A_T:=\left\{(\frac{T\alpha}{-4\beta})^{1/2}|\widetilde{\beta}_T-\beta|\leq 2(\log T)^{1/2}\right\}\text{ and }B_T:=\left\{\frac{I_T}{T}>c_0\right\}$$

where $0<c_0<\frac{1}{-4\beta}$. By Lemma 2.4, we have

$$P(A_T^c)\leq CT^{-1/2}. \tag{2.26}$$

By Lemma 2.1, we have

$$P(B_T^c)=P\left\{\tfrac{-4\beta}{T\alpha}I_T-1<2\beta c_0-1\right\}<P\left\{|\tfrac{-4\beta}{T\alpha}I_T-1|>1-2\beta c_0\right\}\leq CT^{-1}. \tag{2.27}$$

Let b_0 be some positive number. On the set $A_T\cap B_T$ for all $T>T_0$ with $4b_0(\log T_0)^{1/2}(\frac{-4\beta}{T_0\alpha})^{1/2}\leq c_0$, we have

$$
\begin{aligned}
&(\tfrac{T\alpha}{-4\beta})^{1/2}(\widetilde{\beta}_T-\beta)\leq x\\
\Rightarrow\;& I_T+b_0T(\widetilde{\beta}_T-\beta)<I_T+(\tfrac{T\alpha}{-4\beta})^{1/2}2b_0\beta x\\
\Rightarrow\;& (\tfrac{T\alpha}{-4\beta})^{1/2}(\widetilde{\beta}_T-\beta)[I_T+b_0T(\beta_T-\beta)]<x[I_T+(\tfrac{T\alpha}{-4\beta})^{1/2}2b_0\beta x]\\
\Rightarrow\;& (\widetilde{\beta}_T-\beta)I_T+b_0T(\beta_T-\beta)^2<(\tfrac{-4\beta}{T\alpha})^{1/2}I_Tx+2b_0\beta x^2\\
\Rightarrow\;& -N_T+(\widetilde{\beta}_T-\beta)I_T+b_0T(\widetilde{\beta}_T-\beta)^2<-N_T+(\tfrac{-4\beta}{T\alpha})^{1/2}I_Tx+2b_0\beta x^2\\
\Rightarrow\;& 0<-N_T+(\tfrac{-4\beta}{T\alpha})^{1/2}I_Tx+2b_0\beta x^2
\end{aligned}
\tag{2.28}
$$

since

$$\begin{aligned} & I_T + b_0T(\widetilde{\beta}_T - \beta) > Tc_0 + b_0T(\widetilde{\beta}_T - \beta) \\ > \ & 4b_0(\log T)^{1/2}(\tfrac{-4\beta}{T\alpha})^{1/2} - 2b_0(\log T)^{1/2}(\tfrac{-4\beta}{T\alpha})^{1/2} = 2b_0(\log T)^{1/2}(\tfrac{-4\beta}{T\alpha})^{1/2} > 0. \end{aligned} \tag{2.29}$$

On the other hand, on the set $A_T \cap B_T$ for all $T > T_0$ with $4b_0(\log T_0)^{1/2}(\frac{-4\beta}{T_0\alpha})^{1/2} \leq c_0$, we have

$$\begin{aligned} & (\tfrac{T\alpha}{-4\beta})^{1/2}(\widetilde{\beta}_T - \beta) > x \\ \Rightarrow \ & I_T - b_0T(\widetilde{\beta}_T - \beta) < I_T - (\tfrac{T\alpha}{-4\beta})^{1/2}2b_0\beta x \\ \Rightarrow \ & (\tfrac{T\alpha}{-4\beta})^{1/2}(\widetilde{\beta}_T - \beta)[I_T - b_0T(\widetilde{\beta}_T - \beta)] > x[I_T - (\tfrac{T\alpha}{-4\beta})^{1/2}2b_0\beta x] \\ \Rightarrow \ & (\widetilde{\beta}_T - \beta)I_T - b_0T(\widetilde{\beta}_T - \beta)^2 > (\tfrac{-4\beta}{T\alpha})^{1/2}I_T x - 2b_0\beta x^2 \\ \Rightarrow \ & -N_T + (\widetilde{\beta}_T - \beta)I_T - b_0T(\beta_T - \beta)^2 > -N_T + (\tfrac{-4\beta}{T\alpha})^{1/2}I_T x - 2b_0\beta x^2 \\ \Rightarrow \ & 0 > -N_T + (\tfrac{-4\beta}{T\alpha})^{1/2}I_T x - 2b_0\beta x^2 \end{aligned} \tag{2.30}$$

since

$$\begin{aligned} & I_T - b_0T(\widetilde{\beta}_T - \beta) > Tc_0 - b_0T(\widetilde{\beta}_T - \beta) \\ > \ & 4b_0(\log T)^{1/2}(\tfrac{-4\beta}{T\alpha})^{1/2} - 2b_0(\log T)^{1/2}(\tfrac{-4\beta}{T\alpha})^{1/2} = 2b_0(\log T)^{1/2}(\tfrac{-4\beta}{T\alpha})^{1/2} > 0. \end{aligned} \tag{2.31}$$

Hence

$$0 < -N_T + (\frac{-4\beta}{T\alpha})^{1/2}I_T x - 2b_0\beta x^2 \Rightarrow (\frac{T\alpha}{-4\beta})^{1/2}(\beta_T - \beta) \leq x.$$

Letting $D^{\pm}_{T,x} := \left\{-N_T + (\frac{-4\beta}{T})^{1/2}I_T x \pm 2b_0\beta x^2 > 0\right\}$, we obtain

$$D^{-}_{T,x} \cap A_T \cap B_T \subseteq A_T \cap B_T \cap \{(\frac{T\alpha}{-4\beta})^{1/2}(\beta_T - \beta) \leq x\} \subseteq D^{+}_{T,x} \cap A_T \cap B_T. \tag{2.32}$$

This gives

$$P(D^{-}_{T,x} \cap A_T \cap B_T) \leq P(A_T \cap B_T \cap \{(\frac{T\alpha}{-4\beta})^{1/2}(\beta_T - \beta) \leq x\}) \leq P(D^{+}_{T,x} \cap A_T \cap B_T)$$

so that

$$\begin{aligned} & \left| P\left(A_T \cap B_T \cap \left\{\left(\frac{T\alpha}{-4\beta}\right)^{1/2}(\beta_T - \beta) \leq x\right\}\right) - \Phi(x)\right| \\ & \leq \max\left\{|P(D^{-}_{T,x} \cap A_T \cap B_T) - \Phi(x)|,\ |P(D^{+}_{T,x} \cap A_T \cap B_T) - \Phi(x)|\right\} \\ & \leq \max\left\{|P(D^{-}_{T,x}) - \Phi(x)|,\ |P(D^{+}_{T,x}) - \Phi(x)|\right\} + P(A_T \cap B_T)^c. \end{aligned}$$

If it is shown that

$$\left|P\left\{D^{\pm}_{T,x}\right\} - \Phi(x)\right| \leq CT^{-1/2} \tag{2.33}$$

for all $T > T_0$ and $|x| \leq 2(\log T)^{1/2}$, then the theorem would follow from (2.31) - (2.33). We shall prove (2.33) for $D^{+}_{T,x}$. The proof for $D^{-}_{T,x}$ is analogous.

Observe that

$$\begin{aligned} & \left|P\left\{D^{+}_{T,x}\right\} - \Phi(x)\right| = \left|P\left\{(\frac{-4\beta}{T\alpha})^{1/2}N_T - \left(\frac{-4\beta}{T}I_T - 1\right)x < x + 2(\frac{-4\beta}{T})^{1/2}b_0\beta x^2\right\} - \Phi(x)\right| \\ \leq \ & \sup_{y\in\mathbb{R}}\left|P\left\{(\frac{-4\beta}{T\alpha})^{1/2}N_T - \left(\frac{-4\beta}{T\alpha}I_T - 1\right)x \leq y\right\} - \Phi(y)\right| + \left|\Phi\left(x + (\frac{-4\beta}{T\alpha})^{1/2}b_0\beta x^2\right) - \Phi(x)\right| \\ =: \ & \Delta_1 + \Delta_2. \end{aligned} \tag{2.34}$$

Lemma 2.3 and Esseen's smoothing lemma (see Feller [21]) immediately yield

$$\Delta_1 \leq CT^{-1/2}. \tag{2.35}$$

On the other hand, for all $T > T_0$,

$$\Delta_2 \leq 2(\frac{-4\beta}{T\alpha})^{1/2} b_0 \beta x^2 (2\pi)^{-1/2} \exp(-\overline{x}^2/2)$$

where

$$|\overline{x} - x| \leq 2(\frac{-4\beta}{T\alpha})^{1/2} b_0 \beta x^2.$$

Since $|x| \leq 2(\log T)^{1/2}$, it follows that $|\overline{x}| > |x|/2$ for all $T > T_0$ and consequently

$$\Delta_2 \;\leq\; 2(\frac{-4\beta}{T\alpha})^{1/2} b_0 \beta x^2 (2\pi)^{-1/2} x^2 \exp(-x^2/8) \leq CT^{-1/2}. \tag{2.36}$$

From (2.34) - (2.36), we obtain

$$\left| P\left\{ D^+_{T,x} \right\} - \Phi(x) \right| \leq CT^{-1/2}. \tag{2.37}$$

This completes the proof of the theorem. □

Lemma 2.7 Let

$$G_{T,x} := \left(\frac{-4(\alpha-2)}{T\beta} \right)^{1/2} M_T - \left(\frac{-4(\alpha-2)}{T\beta} J_T - 1 \right) x.$$

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

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

Proof : Observe that

$$E\exp(-uJ_T) = \frac{\Gamma(k + \frac{\nu}{2} + \frac{1}{2})}{\Gamma(\nu+1)} (\frac{x}{\sigma t})^{\frac{\nu}{2}+\frac{1}{2}-k} \exp(\frac{x}{\sigma t}) \;\; {}_1F_1(k + \frac{\nu}{2} + \frac{1}{2}, \frac{x}{\sigma t}) \tag{2.38}$$

where $\nu = \frac{1}{\sigma}\sqrt{(\alpha-\sigma)^2 + 4u}$. Now consider

$$\begin{aligned} & E\exp(iuG_{T,x}) = E\exp\left[-iu \left(\tfrac{-4(\alpha-2)}{T\beta} \right)^{1/2} M_T - iu \left(\tfrac{-4(\alpha-2)}{T\beta} J_T - 1 \right) x \right] \\ = \;& E\exp\left[-iu \left(\tfrac{-4\beta}{T\alpha} \right)^{1/2} \{\alpha J_T - \beta T\} - iu \left(\tfrac{-4(\alpha-2)}{T\beta} J_T - 1 \right) x \right] = E\exp(z_1 J_T + z_3) \\ =: \;& \exp(z_3)\tilde{\phi}_T(z_1) \end{aligned} \tag{2.39}$$

where $z_1 := -iu\beta\delta_{T,x}$, and $z_3 := \frac{iuT}{2}\delta_{T,x}$ with $\delta_{T,x} := \left(\frac{-4\beta}{T\alpha}\right)^{1/2} + \frac{2x}{T}$. Note that $\tilde{\phi}_T(z_1) = E(\exp z_1 J_t)$ satisfies (2.38) 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 for (2.38) of Lemma are satisfied. Further, with $\varpi_T(u) := 1 + iu\frac{\delta_{T,x}}{\beta} + \frac{u^2\delta^2_{T,x}}{2\beta^2}$, we have

$$\begin{aligned} & \gamma = (\beta^2 - 2z_1)^{1/2} = \beta\left[1 - \frac{z_1}{\beta^2} - \frac{z_1^2}{2\beta^4} + \frac{z_1^3}{2\beta^8} + \cdots \right] = \beta\left[1 + iu\frac{\delta_{T,x}}{\beta} + \frac{u^2\delta^2_{T,x}}{2\beta^2} + \frac{iu^3\delta^3_{T,x}}{2\beta^3} + \cdots \right] \\ & = \beta[1 + \omega_{1,T}(u) + \omega_{2,T}(u) + \omega_{3,T}(u)] = \beta\varpi_T(u) + \omega_{3,T}(u) = \beta[1 + \omega_{1,T}(u)]. \end{aligned} \tag{2.40}$$

Thus

$$\gamma - \beta = \omega_{1,T}, \;\; \gamma + \beta = 2\beta + \omega_{1,T}. \tag{2.41}$$

Hence the above expectation equals

$$\exp\left(z_3 + \frac{\beta T}{2} \right) \left[\frac{2\beta\varpi_T(u) + \omega_{3,T}(u)}{\omega_{1,T}\exp\{-\beta T\varpi_T(u) + \omega_{4,T}(u)\} + (2\beta + \omega_{1,T}(u))\exp\{\beta T\varpi_T(u) + \omega_{4,T}(u)\}} \right]^{1/2}$$

$$= \left[\frac{1+\omega_{1,T}(u)}{\omega_{1,T}\exp(\chi_T(u)) + (1+\omega_{1,T}(u))\exp(\psi_T(t))}\right]^{1/2} \tag{2.42}$$

where

$$\chi_T(u) \ := \ -\beta T\beta_T(u) + \alpha_{4,T}(u) - 2z_3 - \beta T = -2\beta T + \omega_{1,T}(u) + t^2\omega_{1,T}(u). \tag{2.43}$$

and

$$\begin{aligned}\psi_T(u) \ &:= \ \beta T\varpi_T(u) + \omega_{4,T}(u) - 2z_3 - \beta T = \beta T\left[1 + iu\frac{\delta_{T,x}}{\beta} + \frac{u^2\delta^2_{T,x}}{2\beta^2}\right] + \alpha_{4,T}(u) - iuT\delta_{T,x} - \beta T \\ &= \ \frac{u^2 T\beta}{-4(\alpha-2)}\left[\left(\frac{-4(\alpha-2)}{T\beta}\right)^{1/2} + \frac{2x}{T}\right]^2 = u^2 + u^2\omega_{1,T}(u).\end{aligned} \tag{2.44}$$

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

Hence the above expectation equals

$$\begin{aligned}&\exp(-\frac{u^2}{2})(1+\omega_{1,T})^{1/2} \times \left[\omega_{1,T}\exp\{-2\beta T + \omega_{1,T} + u^2\omega_{1,T}\} + (1+\omega_{1,T}(u))\exp\{u^2\omega_{1,T}(u)\}\right]^{-1/2} \\ &= \exp(-\frac{u^2}{2})\left[1+\omega_{1,T})(1+\omega_{1,T}(1+\omega_{1,T})\exp\{-\beta T + \omega_{1,T} + t^2\omega_{1,T}\}\right] \times \exp(u^2\omega_{1,T}(u)).\end{aligned} \tag{2.45}$$

This completes the proof of the lemma. □

Next we obtain the bound for the estimator of α assuming β is known.

Theorem 2.8

$$\sup_{x\in\mathbb{R}}\left|P\left\{\left(\frac{T\beta}{-4(\alpha-2)}\right)^{1/2}(\widetilde{\alpha}_T - \alpha) \leq x\right\} - \Phi(x)\right| \leq C_\alpha T^{-1/2}.$$

Proof : We shall consider two possibilities (i) $|x| > 2(\log T)^{1/2}$ and (ii) $|x| \leq 2(\log T)^{1/2}$.
(i) We shall give a proof for the case $x > 2(\log T)^{1/2}$. The proof for the case $x < -2(\log T)^{1/2}$ runs similarly. Note that

$$\left|P\left\{(\frac{T\beta}{-4(\alpha-2)})^{1/2}(\widetilde{\alpha}_T - \alpha) \leq x\right\} - \Phi(x)\right| \leq P\left\{(\frac{T\beta}{-4(\alpha-2)})^{1/2}(\widetilde{\alpha}_T - \alpha) \geq x\right\} + \Phi(-x).$$

But $\Phi(-x) \leq \Phi(-2(\log T)^{1/2}) \leq CT^{-2}$. See Feller ([21], p. 166).
Moreover by Lemma 2.4, we have

$$P\left\{(\frac{T\beta}{-4(\alpha-2)})^{1/2}(\widetilde{\alpha}_T - \alpha) \geq 2(\log T)^{1/2}\right\} \leq CT^{-1/2}.$$

Hence

$$\left|P\left\{(\frac{T\beta}{-4(\alpha-2)})^{1/2}(\widetilde{\alpha}_T - \alpha) \leq x\right\} - \Phi(x)\right| \leq CT^{-1/2}.$$

(ii) Let

$$A_T := \left\{(\frac{T\beta}{-4(\alpha-2)})^{1/2}|\widetilde{\alpha}_T - \alpha| \leq 2(\log T)^{1/2}\right\} \text{ and } B_T := \left\{\frac{J_T}{T} > c_0\right\}$$

where $0 < c_0 < \frac{\beta}{-4(\alpha-2)}$. By Lemma 2.5, we have

$$P(A_T^c) \leq CT^{-1/2}. \tag{2.46}$$

By Lemma 2.1, we have

$$\begin{array}{l} P(B_T^c) = P\left\{\frac{-4(\alpha-2)}{T\beta}J_T - 1 < -4(\alpha-2)\beta^{-1}c_0 - 1\right\} \\ < P\left\{|\frac{-4(\alpha-2)}{T\beta}J_T - 1| > 1 + 4(\alpha-2)\beta^{-1}c_0\right\} \le CT^{-1}. \end{array} \tag{2.47}$$

Let b_0 be some positive number. On the set $A_T \cap B_T$ for all $T > T_0$ with $4b_0(\log T_0)^{1/2}(\frac{-4(\alpha-2)}{T_0\beta})^{1/2} \le c_0$, we have

$$\begin{array}{rl} & (\frac{T}{-4(\alpha-2)})^{1/2}(\widetilde{\alpha}_T - \alpha) \le x \\ \Rightarrow & J_T + b_0T(\widetilde{\alpha}_T - \alpha) < J_T + (\frac{T\beta}{-4(\alpha-2)})^{1/2}2b_0\alpha x \\ \Rightarrow & (\frac{T\beta}{-4(\alpha-2)})^{1/2}(\widetilde{\alpha}_T - \alpha)[J_T + b_0T(\alpha_T - \beta)] < x[J_T + (\frac{T\beta}{-4(\alpha-2)})^{1/2}2b_0\alpha x] \\ \Rightarrow & (\widetilde{\alpha}_T - \alpha)J_T + b_0T(\alpha_T - \alpha)^2 < (\frac{-4(\alpha-2)}{T\beta})^{1/2}J_Tx + 2b_0\alpha x^2 \\ \Rightarrow & -N_T + (\widetilde{\alpha}_T - \alpha)J_T + b_0T(\widetilde{\alpha}_T - \alpha)^2 < -N_T + (\frac{-4(\alpha-2)}{T\beta})^{1/2}J_Tx + 2b_0\alpha x^2 \\ \Rightarrow & 0 < -N_T + (\frac{-4(\alpha-2)}{T\beta})^{1/2}J_Tx + 2b_0\alpha x^2 \end{array} \tag{2.48}$$

since

$$\begin{array}{rl} & J_T + b_0T(\widetilde{\alpha}_T - \alpha) > Tc_0 + b_0T(\widetilde{\alpha}_T - \alpha) \\ > & 4b_0(\log T)^{1/2}(\frac{-4(\alpha-2)}{T\beta})^{1/2} - 2b_0(\log T)^{1/2}(\frac{-4(\alpha-2)}{T\beta})^{1/2} = 2b_0(\log T)^{1/2}(\frac{-4(\alpha-2)}{T\beta})^{1/2} > 0. \end{array} \tag{2.49}$$

On the other hand, on the set $A_T \cap B_T$ for all $T > T_0$ with $4b_0(\log T_0)^{1/2}(\frac{-4(\alpha-2)}{T_0\beta})^{1/2} \le c_0$, we have

$$\begin{array}{rl} & (\frac{T}{-4(\alpha-2)})^{1/2}(\widetilde{\alpha}_T - \alpha) > x \\ \Rightarrow & J_T - b_0T(\widetilde{\alpha}_T - \alpha) < J_T - (\frac{T\beta}{-4(\alpha-2)})^{1/2}2b_0\beta x \\ \Rightarrow & (\frac{T\beta}{-4(\alpha-2)})^{1/2}(\widetilde{\alpha}_T - \beta)[J_T - b_0T(\widetilde{\alpha}_T - \beta)] > x[J_T - (\frac{T\beta}{-4(\alpha-2)})^{1/2}2b_0\alpha x] \\ \Rightarrow & (\widetilde{\alpha}_T - \alpha)J_T - b_0T(\widetilde{\alpha}_T - \alpha)^2 > (\frac{-4(\alpha-2)}{T\beta})^{1/2}J_Tx - 2b_0\alpha x^2 \\ \Rightarrow & -N_T + (\widetilde{\alpha}_T - \alpha)J_T - b_0T(\alpha_T - \alpha)^2 > -N_T + (\frac{-4(\alpha-2)}{T\beta})^{1/2}J_Tx - 2b_0\alpha x^2 \\ \Rightarrow & 0 > -N_T + (\frac{-4(\alpha-2)}{T\beta})^{1/2}J_Tx - 2b_0\alpha x^2 \end{array} \tag{2.50}$$

since

$$\begin{array}{rl} & J_T - b_0T(\widetilde{\alpha}_T - \alpha) > Tc_0 - b_0T(\widetilde{\alpha}_T - \alpha) \\ > & 4b_0(\log T)^{1/2}(\frac{-4(\alpha-2)}{T\beta})^{1/2} - 2b_0(\log T)^{1/2}(\frac{-4(\alpha-2)}{T\beta})^{1/2} \\ = & 2b_0(\log T)^{1/2}(\frac{-4(\alpha-2)}{T\beta})^{1/2} > 0. \end{array} \tag{2.51}$$

Hence

$$0 < -N_T + (\frac{-4(\alpha-2)}{T})^{1/2}J_Tx - 2b_0\alpha x^2 \Rightarrow (\frac{T\beta}{-4(\alpha-2)})^{1/2}(\alpha_T - \alpha) \le x.$$

Letting $D_{T,x}^{\pm} := \left\{-N_T + (\frac{-4(\alpha-2)}{T\beta})^{1/2}I_Tx \pm 2b_0\alpha x^2 > 0\right\}$, we obtain

$$D_{T,x}^{-} \cap A_T \cap B_T \subseteq A_T \cap B_T \cap \{(\frac{T\beta}{-4(\alpha-2)})^{1/2}(\alpha_T - \alpha) \le x\} \subseteq D_{T,x}^{+} \cap A_T \cap B_T. \tag{2.52}$$

If it is shown that

$$\left|P\left\{D_{T,x}^{\pm}\right\} - \Phi(x)\right| \le CT^{-1/2} \tag{2.53}$$

for all $T > T_0$ and $|x| \le 2(\log T)^{1/2}$, then the theorem would follow from (2.51) – (2.53). We shall prove (2.53) for $D_{T,x}^{+}$. The proof for $D_{T,x}^{-}$ is analogous.

Observe that

$$\begin{aligned}&\left|P\left\{D^{+}_{T,x}\right\}-\Phi(x)\right|=\left|P\left\{(\tfrac{-4(\alpha-2)}{T\beta})^{1/2}N_T-\left(\tfrac{-4(\alpha-2)}{T\beta}J_T-1\right)x\ <x+2(\tfrac{-4(\alpha-2)}{T\beta})^{1/2}b_0\alpha x^2\right\}-\Phi(x)\right|\\&\leq \sup_{y\in\mathbb{R}}\left|P\left\{(\tfrac{-4(\alpha-2)}{T\beta})^{1/2}N_T-\left(\tfrac{-4(\alpha-2)}{T\beta}J_T-1\right)x\leq y\right\}-\Phi(y)\right|+\left|\Phi\left(x+(\tfrac{-4(\alpha-2)}{T\beta})^{1/2}b_0\alpha x^2\right)-\Phi(x)\right|\\&=:\Delta_1+\Delta_2.\end{aligned} \tag{2.54}$$

Lemma 2.3 and Esseen's lemma (see Feller [21]) immediately yield

$$\Delta_1\leq CT^{-1/2}. \tag{2.55}$$

On the other hand, for all $T>T_0$,

$$\Delta_2\leq 2(\frac{-4(\alpha-2)}{T\beta})^{1/2}b_0\alpha x^2(2\pi)^{-1/2}\exp(-\overline{x}^2/2)$$

where

$$|\overline{x}-x|\leq 2(\frac{-4(\alpha-2)}{T\beta})^{1/2}b_0\alpha x^2.$$

Since $|x|\leq 2(\log T)^{1/2}$, it follows that $|\overline{x}|>|x|/2$ for all $T>T_0$ and consequently

$$\Delta_2\ \leq\ 2(\frac{-4(\alpha-2)}{T\beta})^{1/2}b_0\alpha x^2(2\pi)^{-1/2}x^2\exp(-x^2/8)\leq CT^{-1/2}. \tag{2.56}$$

From (2.54) - (2.56), we obtain

$$\left|P\left\{D^{+}_{T,x}\right\}-\Phi(x)\right|\leq CT^{-1/2}. \tag{2.57}$$

This completes the proof of the theorem. □

3. Discrete Observations

In finance, we have discrete observations of the interest rate process $\{X_t\}$ at times $0=t_0<t_1<\cdots t_n=T$ with $t_i-t_{i-1}=\frac{T}{n}, i=1,2\cdots,n$. 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) $T\to\infty, n\to\infty, \frac{T}{n^{2/3}}\to 0$.

The approximate minimum contrast estimator based on discrete observations is the following:

$$\tilde{\beta}_{n,T}:=\frac{\alpha T}{\sum_{i=1}^{n}(X_{t_i}-X_{t_{i-1}})^2}. \tag{3.1}$$

Let

$$I_{n,T}:=\sum_{i=1}^{n}(X_{t_i}-X_{t_{i-1}})^2. \tag{3.2}$$

The conditional least squares estimators (CLSEs) of β and α are respectively given by

$$\hat{\beta}_n:=\frac{1}{\Delta}\log\frac{\sum_{i=1}^{n}(X_{t_i}-\overline{X}_n)(X_{t_{i-1}}-\overline{X'_n})}{\sum_{i=1}^{n}(X_{t_i}-\overline{X}_n)^2} \tag{3.3}$$

and

$$\hat{\alpha}_n:=\frac{\overline{X}_n-e^{\hat{\beta}_n\Delta}\overline{X'_n}}{e^{\hat{\beta}_n\Delta}-1}\hat{\beta}_n \tag{3.4}$$

where

$$\overline{X}_n:=\frac{1}{n}\sum_{i=1}^{n}X_{t_i},\quad \overline{X'_n}:=\frac{1}{n}\sum_{i=1}^{n}X_{t_{i-1}}. \tag{3.5}$$

Approximate maximum likelihood estimators (AMLEs) based on approximation of the continuous Girsanov likelihood are given by

$$\tilde{\beta}_n := \frac{\sum_{i=1}^n X_{t_{i-1}}^{-1}(X_{t_i} - X_{t_{i-1}}) - T^{-1}X_T \sum_{i=1}^n X_{t_{i-1}}^{-1}(t_i - t_{i-1})}{T - \overline{X}_n \sum_{i=1}^n X_{t_{i-1}}^{-1}(t_i - t_{i-1})}. \tag{3.6}$$

and

$$\tilde{\alpha}_n := \frac{X_T - \overline{X}_n \sum_{i=1}^n X_{t_{i-1}}^{-1}(t_i - t_{i-1})}{T - \overline{X}_n \sum_{i=1}^n X_{t_{i-1}}^{-1}(t_i - t_{i-1})}. \tag{3.7}$$

The approximate transition density based on Dacunha-Castle and Florens-Zmirou [19] is given by

$$\begin{aligned} p(t,x,y,t) \;:=\; & -\tfrac{1}{2}\log(2\pi t y) - \tfrac{2(\sqrt{y}-\sqrt{x})^2}{t} + \beta(y-x) + (\alpha - \tfrac{1}{4})\log(\tfrac{y}{x}) \\ & -\tfrac{t}{2}\left[\tfrac{1}{3}\left\{\beta\sqrt{x} + (\alpha-\tfrac{1}{4})\tfrac{1}{\sqrt{x}}\right\}^2 + \tfrac{1}{2}\left\{\tfrac{\beta}{2} - \tfrac{1}{2}(\alpha-\tfrac{1}{4})\tfrac{1}{x}\right\}\right. \\ & +\tfrac{1}{3}\left\{\beta\sqrt{y} + (\alpha-\tfrac{1}{4})\tfrac{1}{\sqrt{y}}\right\}^2 + \tfrac{1}{2}\left\{\tfrac{\beta}{2} - \tfrac{1}{2}(\alpha-\tfrac{1}{4})\tfrac{1}{y}\right\} \\ & \left. +\tfrac{1}{3}\left\{\beta\sqrt{x} + (\alpha-\tfrac{1}{4})\tfrac{1}{\sqrt{x}}\right\}\left\{\beta\sqrt{y} + (\alpha-\tfrac{1}{4})\tfrac{1}{\sqrt{y}}\right\}\right]. \end{aligned} \tag{3.8}$$

The AMLEs of β and α are given respectively by

$$\beta_{n,T} := \frac{u_{1,n}w_{2,n} - u_{2,n}w_{1,n}}{v_{1,n}u_{2,n} - v_{2,n}u_{1,n}}, \tag{3.9}$$

$$\alpha_{n,T} := \frac{v_{2,n}w_{1,n} - v_{1,n}w_{2,n}}{v_{1,n}u_{2,n} - v_{2,n}u_{1,n}} \tag{3.10}$$

where

$$\begin{aligned} u_{1,n} &:= \tfrac{2n\Delta}{3} + \tfrac{\Delta}{6}\sum_{i=1}^n \tfrac{X_{t_{i-1}}+X_{t_i}}{\sqrt{X_{t_{i-1}}X_{t_i}}} \\ v_{1,n} &:= \tfrac{\Delta}{3}\sum_{i=1}^n \left(X_{t_{i-1}} + \sqrt{X_{t_{i-1}}X_{t_i}} + X_{t_i}\right) \\ w_{1,n} &:= \tfrac{n\Delta}{12} - \sum_{i=1}^n (X_{t_i} - X_{t_{i-1}}) - \tfrac{\Delta}{24}\sum_{i=1}^n \tfrac{X_{t_i}+X_{t_{i-1}}}{\sqrt{X_{t_{i-1}}X_{t_i}}} \\ u_{2,n} &:= \tfrac{\Delta}{3}\sum_{i=1}^n \left(\tfrac{1}{X_{t_i}} + \tfrac{1}{X_{t_{i-1}}} + \tfrac{1}{\sqrt{X_{t_{i-1}}X_{t_i}}}\right) \\ v_{2,n} &:= \tfrac{\Delta}{6}\left(4n + \sum_{i=1}^n \tfrac{X_{t_{i-1}}+X_{t_i}}{\sqrt{X_{t_{i-1}}X_{t_i}}}\right) \\ w_{2,n} &:= -\sum_{i=1}^n (\log X_{t_i} - \log X_{t_{i-1}}) - \tfrac{\Delta}{8}\sum_{i=1}^n \left(\tfrac{1}{X_{t_i}} + \tfrac{1}{X_{t_{i-1}}}\right) \\ &\quad -\tfrac{\Delta}{12}\sum_{i=1}^n \left(\tfrac{1}{X_{t_i}} + \tfrac{1}{X_{t_{i-1}}} + \tfrac{1}{\sqrt{X_{t_{i-1}}X_{t_i}}}\right). \end{aligned} \tag{3.11}$$

Based on discrete observations we define the approximate minimum contrast estimators (AMCEs) as follows:

In order to define the approximate minimum contrast estimators (AMCEs), we use various discrete approximations of the integrals in the definition (2.15) and (2.16) of MCEs.

An Euler type discrete approximation of (2.15) and (2.16) gives

$$\check{\beta}_n = \frac{T\overline{X'_n}}{2\sum_{i=1}^n (X_{t_{i-1}} - \overline{X}'_n)^2}, \tag{3.12}$$

$$\check{\alpha}_n = \frac{(\overline{X'_n})^2}{2\sum_{i=1}^n (X_{t_{i-1}} - \overline{X'_n})^2}. \tag{3.13}$$

We will next consider weighted AMCES. Define a weighted sum of squares

$$M_{n,T} := \frac{T}{n}\left\{\sum_{i=1}^n w_{t_i}X_{t_{i-1}}^2 + \sum_{i=2}^{n+1} w_{t_i}X_{t_{i-1}}^2\right\}. \tag{3.14}$$

where $w_{t_i} \geq 0$ is a weight function.

Denote the discrete increasing functions

$$I_{n,T} := \frac{T}{n}\sum_{i=1}^{n} X_{t_{i-1}}^2, \tag{3.15}$$

$$J_{n,T} := \frac{T}{n}\sum_{i=2}^{n+1} X_{t_{i-1}}^2 = \frac{T}{n}\sum_{i=1}^{n} X_{t_i}^2. \tag{3.16}$$

General weighted AMCE is defined as

$$\tilde{\beta}_{n,T} := -\left\{\frac{2}{n}M_{n,T}\right\}^{-1}. \tag{3.17}$$

With $w_{t_i} = 1$, we obtain the forward AMCE as

$$\tilde{\beta}_{n,T,F} := -\left\{\frac{2}{n}I_{n,T}\right\}^{-1}. \tag{3.18}$$

With $w_{t_i} = 0$, we obtain the backward AMCE as

$$\tilde{\beta}_{n,T,B} := -\left\{\frac{2}{n}J_{n,T}\right\}^{-1}. \tag{3.19}$$

Analogous to the estimators for the discrete AR (1) model, we define the simple symmetric and weighted symmetric estimators (see Fuller (1996)):

With $w_{t_i} = 0.5$, the simple symmetric AMCE is defined as

$$\tilde{\beta}_{n,T,z} := -\left\{\frac{1}{n}[I_{n,T} + J_{n,T}]\right\}^{-1} = -\left\{\frac{2}{n}\sum_{i=2}^{n} X_{t_{i-1}}^2 + 0.5(X_{t_0}^2 + X_{t_n}^2)\right\}^{-1}. \tag{3.20}$$

With the weight function

$$w_{t_i} = \begin{cases} 0 & : \quad i = 1 \\ \frac{i-1}{n} & : \quad i = 2, 3, \cdots, n \\ 1 & : \quad i = n+1 \end{cases}$$

the weighted symmetric AMCE is defined as

$$\tilde{\beta}_{n,T,w} := -\left\{\frac{2}{n}\sum_{i=2}^{n} X_{t_{i-1}}^2 + \frac{1}{n}\sum_{i=1}^{n} X_{t_{i-1}}^2\right\}^{-1}. \tag{3.21}$$

Note that estimator (3.21) is analogous to the trapezoidal rule in numerical analysis. One can instead use the midpoint rule to define another estimator

$$\tilde{\beta}_{n,T,A} := -\left\{\frac{2}{n}\sum_{i=1}^{n}\left(\frac{X_{t_{i-1}} + X_{t_i}}{2}\right)^2\right\}^{-1}. \tag{3.22}$$

One can use the Simpson's rule to define another estimator where the denominator is a convex combination of the denominators in (3.21) and (3.22)

$$\tilde{\beta}_{n,T,S} := -\left\{\frac{1}{3n}\sum_{i=1}^{n}\left\{X_{t_{i-1}}^2 + 4\left(\frac{X_{t_{i-1}} + X_{t_i}}{2}\right)^2 + X_{t_i}^2\right\}\right\}^{-1}. \tag{3.23}$$

In general, one can generalize Simpson's rule as

$$\tilde{\beta}_{n,T,GS} := -\left\{\frac{2}{n}\sum_{i=1}^{n}\left\{p\frac{X_{t_{i-1}}^2 + X_{t_i}^2}{2} + (1-p)\left(\frac{X_{t_{i-1}} + X_{t_i}}{2}\right)^2\right\}\right\}^{-1} \tag{3.24}$$

for any $0 \leq p \leq 1$. The case $p = 0$ produces the estimator (3.22). The case $p = 1$ produces the estimator (3.21). The case $p = \frac{1}{3}$ produces the estimator (3.23).

I propose a very general form of the quadrature based estimator as

$$\tilde{\beta}_{n,T,w} := -\left\{\frac{2}{n}\sum_{i=1}^{n}\sum_{j=1}^{m}\left[(1-s_j)X_{t_{i-1}} + s_j X_{t_i}\right]^2 p_j\right\}^{-1} \tag{3.25}$$

where $p_j, \quad j \in \{1, 2, \cdots, m\}$ is a probability mass function of a discrete random variable S on $0 \leq s_1 < s_2 < \cdots < s_m \leq 1$ with $P(S = s_j) := p_j,\ j \in \{1, 2, \cdots, m\}$. Denote the k-th moment of the random variable S as $\mu_k := \sum_{j=1}^{m} s_j^k p_j,\ k = 1, 2, \cdots$.

If one chooses the probability distribution as uniform distribution for which the moments are a harmonic sequence $(\mu_1, \mu_2, \mu_3, \mu_4, \mu_5, \mu_6, \cdots) = (\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \frac{1}{6}, \frac{1}{7}, \cdots)$ then there is no change in rate of convergence than second order. If one can construct a probability distribution for which the harmonic sequence is truncated at a point, then there is an improvement in the rate of convergence at the point of truncation.

Given a positive integer m, construct a probability mass function $p_j, \quad j \in \{1, 2, \cdots, m\}$ on $0 \leq s_1 < s_2 < \cdots < s_m \leq 1$ such that

$$\sum_{j=1}^{m} s_j^r p_j = \frac{1}{r+1},\ r \in \{0, \cdots, m-2\} \tag{3.26}$$

$$\sum_{j=1}^{m} s_j^{m-1} p_j \neq \frac{1}{m}. \tag{3.27}$$

Neither the probabilities p_j nor the atoms, s_j, of the distribution are specified in advance.

This problem is related to the truncated Hausdorff moment problem. I obtain examples of such probability distributions and use them to get higher order accurate (up to sixth order) AMCEs.

The order of approximation error (rate of convergence) of an estimator is $n^{-\nu}$ where

$$\nu := \inf\left\{k : \mu_k \neq \frac{1}{1+k},\ \mu_j = \frac{1}{1+j}, j = 1, 2, \cdots, k-1\right\}. \tag{3.28}$$

We construct probability distributions satisfying these moment conditions and obtain estimators of the rate of convergence up to order 6.

Probability $p_1 = 1$ at the point $s_1 = 0$ gives the estimator (3.18) for which $\mu_1 = 0$. Note that $\mu_1 \neq \frac{1}{2}$. Thus $\nu = 1$ Probability $p_1 = 1$ at the point $s_1 = 1$ gives the estimator (3.19) for which $\mu_1 = 1$. Note that $\mu_1 \neq \frac{1}{2}$. Thus $\nu = 1$. Probabilities $(p_1, p_2) = (\frac{1}{2}, \frac{1}{2})$ at the respective points $(s_1, s_2) = (0, 1)$ produces the estimator $\tilde{\beta}_{n,T,Z}$ for which $(\mu_1, \mu_2) = (\frac{1}{2}, \frac{1}{4})$. Thus $\nu = 2$. Probability $p_j = 1$ at the point $s_j = \frac{1}{2}$ produce the estimator $\tilde{\beta}_{n,T,A}$ for which $(\mu_1, \mu_2) = (\frac{1}{2}, \frac{1}{2})$. Thus $\nu = 2$. Probabilities $(p_1, p_2) = (\frac{1}{4}, \frac{3}{4})$ at the respective points $(s_1, s_2) = (0, \frac{2}{3})$ produce the asymmetric estimator

$$\tilde{\beta}_{n,T,3} := -\left\{\frac{2}{n}\frac{1}{4}\sum_{i=1}^{n}\left[(X_{t_{i-1}})^2 + 3(\frac{X_{t_{i-1}}+2X_{t_i}}{3})^2\right]\right\}^{-1} \tag{3.29}$$

for which $(\mu_1, \mu_2, \mu_3) = (\frac{1}{2}, \frac{1}{3}, \frac{2}{9})$. Thus $\nu = 3$. Probabilities $(p_1, p_2) = (\frac{3}{4}, \frac{1}{4})$ at the respective points $(s_1, s_2) = (\frac{1}{3}, 1)$ produce asymmetric estimator

$$\tilde{\beta}_{n,T,4} := -\left\{\frac{2}{n}\frac{1}{4}\sum_{i=1}^{n}\left[3(\frac{2X_{t_{i-1}}+X_{t_i}}{3})^2 + (X_{t_i})^2\right]\right\}^{-1} \tag{3.30}$$

for which $(\mu_1, \mu_2, \mu_3) = (\frac{1}{2}, \frac{1}{3}, \frac{10}{36})$. Thus $\nu = 3$. Probabilities $(p_1, p_2, p_3) = (\frac{1}{6}, \frac{2}{3}, \frac{1}{6})$ at the respective points $(s_1, s_2, s_3) = (0, \frac{1}{2}, 1)$ produce the estimator $\tilde{\beta}_{n,T,5}$ for which $(\mu_1, \mu_2, \mu_3, \mu_4) = (\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{5}{25})$. Thus $\nu = 4$. Probabilities $(p_1, p_2, p_3, p_4) = (\frac{1}{8}, \frac{3}{8}, \frac{3}{8}, \frac{1}{8})$ at the respective points $(s_1, s_2, s_3, s_4) = (0, \frac{1}{3}, \frac{2}{3}, 1)$ produce the symmetric estimator

$$\tilde{\beta}_{n,T,5} := -\left\{\frac{2}{n}\frac{1}{8}\sum_{i=1}^{n}\left[(X_{t_{i-1}})^2 + 3(\tfrac{2X_{t_{i-1}}+X_{t_i}}{3})^2 + 3(\tfrac{X_{t_{i-1}}+2X_{t_i}}{3})^2 + (X_{t_i})^2\right]\right\}^{-1} \tag{3.31}$$

for which $(\mu_1, \mu_2, \mu_3, \mu_4) = (\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{11}{54})$. Thus $\nu = 4$. Probabilities $(p_1, p_2, p_3, p_4, p_5) = (\frac{1471}{24192}, \frac{6925}{24192}, \frac{1475}{12096}, \frac{2725}{12096}, \frac{5675}{24192}, \frac{1721}{24192})$ at the respective points $(s_1, s_2, s_3, s_4, s_5) = (0, \frac{1}{5}, \frac{2}{5}, \frac{3}{5}, \frac{4}{5}, 1)$ produce the asymmetric estimator

$$\begin{aligned}\tilde{\beta}_{n,T,7} := -\Big\{\frac{2}{n}\frac{1}{24192}\sum_{i=1}^{n}\Big[&1471(X_{t_{i-1}})^2 + 6925(\tfrac{X_{t_{i-1}}+X_{t_i}}{5})^2 + 2950(\tfrac{2X_{t_{i-1}}+2X_{t_i}}{5})^2 \\ &+5450(\tfrac{3X_{t_{i-1}}+3X_{t_i}}{5})^2 + 5675(\tfrac{4X_{t_{i-1}}+4X_{t_i}}{5})^2 + 1721(X_{t_i})^2\Big]\Big\}^{-1}\end{aligned} \tag{3.32}$$

for which $(\mu_1, \mu_2, \mu_3, \mu_4, \mu_5, \mu_6) = (\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \frac{841}{5040})$. Thus $\nu = 5$. Probabilities $(p_1, p_2, p_3, p_4, p_5) = (\frac{7}{90}, \frac{16}{45}, \frac{2}{15}, \frac{16}{45}, \frac{7}{90})$ at the respective points
$(s_1, s_2, s_3, s_4, s_5) = (0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1)$ produce the symmetric estimator $\tilde{\beta}_{n,T,8}$ given by

$$\begin{aligned}\tilde{\beta}_{n,T,8} := -\Big\{\frac{2}{n}\frac{1}{90}\sum_{i=1}^{n}\Big[&7(X_{t_{i-1}})^2 + 32(\tfrac{3X_{t_{i-1}}+X_{t_i}}{4})^2 + 12(\tfrac{X_{t_{i-1}}+X_{t_i}}{2})^2 + 32(\tfrac{X_{t_{i-1}}+3X_{t_i}}{4})^2 \\ &+7(t_i, X_{t_i})^2]\}^{-1}\end{aligned} \tag{3.33}$$

for which $(\mu_1, \mu_2, \mu_3, \mu_4, \mu_5, \mu_6) = (\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \frac{1}{6}, \frac{110}{768})$. Thus $\nu = 6$. Probabilities $(p_1, p_2, p_3, p_4, p_5) = (\frac{19}{288}, \frac{75}{288}, \frac{50}{288}, \frac{50}{288}, \frac{75}{288}, \frac{19}{288})$ at the respective points
$(s_1, s_2, s_3, s_4, s_5) = (0, \frac{1}{5}, \frac{2}{5}, \frac{3}{5}, \frac{4}{5}, 1)$ produce symmetric estimator

$$\begin{aligned}\tilde{\beta}_{n,T,9} := -\Big\{\frac{2}{n}\frac{1}{288}\sum_{i=1}^{n}\Big[&19(X_{t_{i-1}})^2 + 75(\tfrac{4X_{t_{i-1}}+X_{t_i}}{5})^2 + 50(\tfrac{3X_{t_{i-1}}+2X_{t_i}}{5})^2 \\ &+50(\tfrac{2X_{t_{i-1}}+3X_{t_i}}{5})^2 + 75(\tfrac{X_{t_{i-1}}+4X_{t_i}}{5})^2 + 19(X_{t_i})^2\Big]\Big\}^{-1}\end{aligned} \tag{3.34}$$

for which $(\mu_1, \mu_2, \mu_3, \mu_4, \mu_5, \mu_6) = (\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \frac{1}{6}, \frac{3219}{22500})$. Thus $\nu = 6$.

The estimator $\tilde{\beta}_{n,T,S}$ is based on the arithmetic mean of $I_{n,T}$ and $J_{n,T}$. One can use geometric mean and harmonic mean instead. The geometric mean based symmetric AMCE (which is based on the ideas of partial autocorrelation) is defined as

$$\tilde{\beta}_{n,T,G} := \frac{-T/2}{\sqrt{I_{n,T}\ J_{n,T}}} \tag{3.35}$$

The harmonic mean based symmetric AMCE is defined as

$$\tilde{\beta}_{n,T,H} := \frac{-T}{\frac{1}{I_{n,T}} + \frac{1}{J_{n,T}}} \tag{3.36}$$

Note that

$$\tilde{\beta}_{n,T,H} \ \le\ \tilde{\beta}_{n,T,G} \ \le\ \tilde{\beta}_{n,T,S}. \tag{3.37}$$

Note that for the Simpson's estimator we have

$$\frac{1}{3}\sum_{i=1}^{n}\left[X_{t_{i-1}} + \sqrt{X_{t_i}X_{t_{i-1}}} + X_{t_i}\right] = \frac{1}{6}\sum_{i=1}^{n}\left[X_{t_{i-1}} + 4\left(\frac{\sqrt{X_{t_{i-1}}} + \sqrt{X_{t_i}}}{2}\right)^2 + X_{t_i}\right]. \tag{3.38}$$

Now we define AMCEs for α.

Define a weighted sum of squares

$$M_{n,T} := \frac{T}{n}\left\{\sum_{i=1}^{n} w_{t_i} X_{t_{i-1}}^2 + \sum_{i=2}^{n+1} w_{t_i} X_{t_{i-1}}^2\right\}. \tag{3.39}$$

where $w_{t_i} \geq 0$ is a weight function.

Denote the discrete realized variance functions as

$$I_{n,T} := \frac{T}{n}\sum_{i=1}^{n} X_{t_{i-1}}^2, \tag{3.41}$$

$$J_{n,T} := \frac{T}{n}\sum_{i=2}^{n+1} X_{t_{i-1}}^2 = \frac{T}{n}\sum_{i=1}^{n} X_{t_i}^2. \tag{3.40}$$

General weighted AMCE is defined as

$$\tilde{\alpha}_{n,T} := -\left\{\frac{2}{n} M_{n,T}\right\}^{-1}. \tag{3.41}$$

With $w_{t_i} = 1$, we obtain the forward AMCE as

$$\tilde{\alpha}_{n,T,F} := -\left\{\frac{2}{n} I_{n,T}\right\}^{-1}. \tag{3.42}$$

With $w_{t_i} = 0$, we obtain the backward AMCE as

$$\tilde{\alpha}_{n,T,B} := -\left\{\frac{2}{n} J_{n,T}\right\}^{-1}. \tag{3.43}$$

Analogous to the estimators for the discrete AR (1) model, we define the simple symmetric and weighted symmetric estimators (see Fuller (1996)):

With $w_{t_i} = 0.5$, the simple symmetric AMCE is defined as

$$\tilde{\alpha}_{n,T,z} := -\left\{\frac{1}{n}[I_{n,T} + J_{n,T}]\right\}^{-1} = -\left\{\frac{2}{n}\sum_{i=2}^{n} X_{t_{i-1}}^2 + 0.5(X_{t_0}^2 + X_{t_n}^2)\right\}^{-1}. \tag{3.44}$$

With the weight function

$$w_{t_i} = \begin{cases} 0 & : \quad i = 1 \\ \frac{i-1}{n} & : \quad i = 2, 3, \cdots, n \\ 1 & : \quad i = n+1 \end{cases} \tag{3.45}$$

the weighted symmetric AMCE is defined as

$$\tilde{\alpha}_{n,T,w} := -\left\{\frac{2}{n}\sum_{i=2}^{n} X_{t_{i-1}}^2 + \frac{1}{n}\sum_{i=1}^{n} X_{t_{i-1}}^2\right\}^{-1}. \tag{3.46}$$

Note that estimator (3.46) is analogous to the trapezoidal rule in numerical analysis. One can instead use the midpoint rule to define another estimator

$$\tilde{\alpha}_{n,T,A} := -\left\{\frac{2}{n}\sum_{i=1}^{n}\left(\frac{X_{t_{i-1}} + X_{t_i}}{2}\right)^2\right\}^{-1}. \tag{3.47}$$

One can use the Simpson's rule to define another estimator where the denominator is a convex combination of the denominators in (3.46) and (3.47)

$$\tilde{\alpha}_{n,T,S} := -\left\{\frac{1}{3n}\sum_{i=1}^{n}\left\{X_{t_{i-1}}^2 + 4\left(\frac{X_{t_{i-1}} + X_{t_i}}{2}\right)^2 + X_{t_i}^2\right\}\right\}^{-1}. \tag{3.48}$$

In general, one can generalize Simpson's rule as

$$\tilde{\alpha}_{n,T,GS} := -\left\{\frac{2}{n}\sum_{i=1}^{n}\left\{a\frac{X_{t_{i-1}}^2+X_{t_i}^2}{2}+(1-a)\left(\frac{X_{t_{i-1}}+X_{t_i}}{2}\right)^2\right\}\right\}^{-1} \tag{3.49}$$

for any $0 \leq a \leq 1$.

The case $a = 0$ produces the estimator (3.47). The case $1 = 1$ produces the estimator (3.44). The case $a = \frac{1}{3}$ produces the estimator (3.48).

I propose a very general form of the quadrature based estimator as

$$\tilde{\alpha}_{n,T,w} := -\left\{\frac{2}{n}\sum_{i=1}^{n}\sum_{j=1}^{m}\left[(1-s_j)X_{t_{i-1}}+s_jX_{t_i}\right]^2 p_j\right\}^{-1} \tag{3.50}$$

where $p_j,\ \ j \in \{1,2,\cdots,m\}$ is a probability mass function of a discrete random variable S on $0 \leq s_1 < s_2 < \cdots < s_m \leq 1$ with $P(S = s_j) := p_j,\ j \in \{1,2,\cdots,m\}$. Denote the k-th moment of the random variable S as $\mu_k := \sum_{j=1}^{m} s_j^k p_j,\ k = 1,2,\cdots$.

If one chooses the probability distribution as uniform distribution for which the moments are a harmonic sequence $(\mu_1,\mu_2,\mu_3,\mu_4,\mu_5,\mu_6,\cdots) = (\frac{1}{2},\frac{1}{3},\frac{1}{4},\frac{1}{5},\frac{1}{6},\frac{1}{7},\cdots)$ then there is no change in rate of convergence than second order. If one can construct a probability distribution for which the harmonic sequence is truncated at a point, then there is a rate of convergence improvement at the point of truncation.

The order of approximation error (rate of convergence) of an estimator is $n^{-\nu}$ where

$$\nu := \inf\left\{k : \mu_k \neq \frac{1}{1+k},\ \mu_j = \frac{1}{1+j}, j = 1,2,\cdots,k-1\right\}. \tag{3.51}$$

We construct probability distributions satisfying these moment conditions and obtain estimators of the rate of convergence up to order 6.

Probability $p_1 = 1$ at the point $s_1 = 0$ gives the estimator (3.42) for which $\mu_1 = 0$. Note that $\mu_1 \neq \frac{1}{2}$. Thus $\nu = 1$ Probability $p_1 = 1$ at the point $s_1 = 1$ gives the estimator (3.43) for which $\mu_1 = 1$. Note that $\mu_1 \neq \frac{1}{2}$. Thus $\nu = 1$. Probabilities $(p_1,p_2) = (\frac{1}{2},\frac{1}{2})$ at the respective points $(s_1,s_2) = (0,1)$ produces the estimator $\tilde{\alpha}_{n,T,Z}$ for which $(\mu_1,\mu_2) = (\frac{1}{2},\frac{1}{4})$. Thus $\nu = 2$. Probability $p_j = 1$ at the point $s_j = \frac{1}{2}$ produce the estimator $\tilde{\alpha}_{n,T,A}$ for which $(\mu_1,\mu_2) = (\frac{1}{2},\frac{1}{2})$. Thus $\nu = 2$. Probabilities $(p_1,p_2) = (\frac{1}{4},\frac{3}{4})$ at the respective points $(s_1,s_2) = (0,\frac{2}{3})$ produce the asymmetric estimator

$$\tilde{\alpha}_{n,T,3} := -\left\{\tfrac{2}{n}\tfrac{1}{4}\textstyle\sum_{i=1}^{n}\left[(X_{t_{i-1}})^2+3(\tfrac{X_{t_{i-1}}+2X_{t_i}}{3})^2\right]\right\}^{-1} \tag{3.52}$$

for which $(\mu_1,\mu_2,\mu_3) = (\frac{1}{2},\frac{1}{3},\frac{2}{9})$. Thus $\nu = 3$. Probabilities $(p_1,p_2) = \frac{3}{4},\frac{1}{4}$ at the respective points $(s_1,s_2) = (\frac{1}{3},1)$ produce asymmetric estimator

$$\tilde{\alpha}_{n,T,4} := -\left\{\tfrac{2}{n}\tfrac{1}{4}\textstyle\sum_{i=1}^{n}\left[3(\tfrac{2X_{t_{i-1}}+X_{t_i}}{3})^2+(X_{t_i})^2\right]\right\}^{-1} \tag{3.53}$$

for which $(\mu_1,\mu_2,\mu_3) = (\frac{1}{2},\frac{1}{3},\frac{10}{36})$. Thus $\nu = 3$. Probabilities $(p_1,p_2,p_3) = (\frac{1}{6},\frac{2}{3},\frac{1}{6})$ at the respective points $(s_1,s_2,s_3) = (0,\frac{1}{2},1)$ produce the estimator $\tilde{\alpha}_{n,T,5}$ for which $(\mu_1,\mu_2,\mu_3,\mu_4) = (\frac{1}{2},\frac{1}{3},\frac{1}{4},\frac{5}{25})$. Thus $\nu = 4$. Probabilities $(p_1,p_2,p_3,p_4) = (\frac{1}{8},\frac{3}{8},\frac{3}{8},\frac{1}{8})$ at the respective points $(s_1,s_2,s_3,s_4) = (0,\frac{1}{3},\frac{2}{3},1)$ produce the symmetric estimator

$$\tilde{\alpha}_{n,T,5} := -\left\{\tfrac{2}{n}\tfrac{1}{8}\textstyle\sum_{i=1}^{n}\left[(X_{t_{i-1}})^2+3(\tfrac{2X_{t_{i-1}}+X_{t_i}}{3})^2+3(\tfrac{X_{t_{i-1}}+2X_{t_i}}{3})^2+(X_{t_i})^2\right]\right\}^{-1} \tag{3.54}$$

for which $(\mu_1, \mu_2, \mu_3, \mu_4) = (\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{11}{54})$. Thus $\nu = 4$. Probabilities $(p_1, p_2, p_3, p_4, p_5) = (\frac{1471}{24192}, \frac{6925}{24192}, \frac{1475}{12096}, \frac{2725}{12096}, \frac{5675}{24192}, \frac{1721}{24192})$ at the respective points $(s_1, s_2, s_3, s_4, s_5) = (0, \frac{1}{5}, \frac{2}{5}, \frac{3}{5}, \frac{4}{5}, 1)$ produce the asymmetric estimator

$$\begin{aligned}\tilde{\alpha}_{n,T,7} := -\Big\{\tfrac{2}{n}\tfrac{1}{24192}\textstyle\sum_{i=1}^{n}\Big[1471(X_{t_{i-1}})^2 + 6925(\tfrac{X_{t_{i-1}}+X_{t_i}}{5})^2 + 2950(\tfrac{2X_{t_{i-1}}+2X_{t_i}}{5})^2 \\ +5450(\tfrac{3X_{t_{i-1}}+3X_{t_i}}{5})^2 + 5675(\tfrac{4X_{t_{i-1}}+4X_{t_i}}{5})^2 + 1721(X_{t_i})^2\Big]\Big\}^{-1} \qquad (3.55)\end{aligned}$$

for which $(\mu_1, \mu_2, \mu_3, \mu_4, \mu_5, \mu_6) = (\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \frac{841}{5040})$. Thus $\nu = 5$. Probabilities $(p_1, p_2, p_3, p_4, p_5) = (\frac{7}{90}, \frac{16}{45}, \frac{2}{15}, \frac{16}{45}, \frac{7}{90})$ at the respective points $(s_1, s_2, s_3, s_4, s_5) = (0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1)$ produce the symmetric estimator $\tilde{\alpha}_{n,T,8}$ given by

$$\begin{aligned}\tilde{\alpha}_{n,T,8} := -\Big\{\tfrac{2}{n}\tfrac{1}{90}\textstyle\sum_{i=1}^{n}\Big[7(X_{t_{i-1}})^2 + 32(\tfrac{3X_{t_{i-1}}+X_{t_i}}{4})^2 + 12(\tfrac{X_{t_{i-1}}+X_{t_i}}{2})^2 + 32(\tfrac{X_{t_{i-1}}+3X_{t_i}}{4})^2 \\ +7(t_i, X_{t_i})^2]\}^{-1} \qquad (3.56)\end{aligned}$$

for which $(\mu_1, \mu_2, \mu_3, \mu_4, \mu_5, \mu_6) = (\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \frac{1}{6}, \frac{110}{768})$. Thus $\nu = 6$. Probabilities $(p_1, p_2, p_3, p_4, p_5) = (\frac{19}{288}, \frac{75}{288}, \frac{50}{288}, \frac{50}{288}, \frac{75}{288}, \frac{19}{288})$ at the respective points
$(s_1, s_2, s_3, s_4, s_5) = (0, \frac{1}{5}, \frac{2}{5}, \frac{3}{5}, \frac{4}{5}, 1)$ produce symmetric estimator

$$\begin{aligned}\tilde{\alpha}_{n,T,9} := -\Big\{\tfrac{2}{n}\tfrac{1}{288}\textstyle\sum_{i=1}^{n}\Big[19(X_{t_{i-1}})^2 + 75(\tfrac{4X_{t_{i-1}}+X_{t_i}}{5})^2 + 50(\tfrac{3X_{t_{i-1}}+2X_{t_i}}{5})^2 \\ +50(\tfrac{2X_{t_{i-1}}+3X_{t_i}}{5})^2 + 75(\tfrac{X_{t_{i-1}}+4X_{t_i}}{5})^2 + 19(X_{t_i})^2\Big]\Big\}^{-1} \qquad (3.57)\end{aligned}$$

for which $(\mu_1, \mu_2, \mu_3, \mu_4, \mu_5, \mu_6) = (\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \frac{1}{6}, \frac{3219}{22500})$. Thus $\nu = 6$.

The estimator $\tilde{\alpha}_{n,T,S}$ is based on the arithmetic mean of $I_{n,T}$ and $J_{n,T}$. One can use geometric mean and harmonic mean instead. The geometric mean based symmetric AMCE (which is based on the ideas of partial autocorrelation) is defined as

$$\tilde{\alpha}_{n,T,G} := \frac{-T/2}{\sqrt{I_{n,T}\ J_{n,T}}} \qquad (3.58)$$

The harmonic mean based symmetric AMCE is defined as

$$\tilde{\alpha}_{n,T,H} := \frac{-T}{\frac{1}{I_{n,T}} + \frac{1}{J_{n,T}}} \qquad (3.59)$$

Note that

$$\tilde{\alpha}_{n,T,H} \ \leq\ \tilde{\alpha}_{n,T,G} \ \leq\ \tilde{\alpha}_{n,T,S}. \qquad (3.60)$$

We need the following lemma in the sequel.

Lemma 3.1 Let ξ, η and ζ be any three random variables on a probability space $(\Omega, \mathcal{F}, P)$ with $P(\zeta > 0) = 1$. Then, for any $\epsilon > 0$, we have

(a) $\sup_{x\in\mathbb{R}} |P\{\xi + \eta \leq x\} - \Phi(x)| \leq \sup_{x\in\mathbb{R}} |P\{\xi \leq x\} - \Phi(x)| + P(|\eta| > \epsilon) + \epsilon.$

(b) $\sup_{x\in\mathbb{R}} |P\{\frac{\xi}{\zeta} \leq x\} - \Phi(x)| \leq \sup_{x\in\mathbb{R}} |P\{\xi \leq x\} - \Phi(x)| + P\{|\zeta - 1| > \epsilon\} + \epsilon.$

(c) Let $\varrho_n, \tau_n, \varrho$ and τ be random variables on the same probability space $(\Omega, \mathcal{F}, P)$ with $P(\tau_n > 0) = 1$ and $P(\tau > 0) = 1$. Suppose $|\varrho_n - \varrho| = O_P(\delta_{1n})$ and $|\tau_n - \tau| = O_P(\delta_{2n})$ where $\delta_{1n},\ \delta_{2n} \to 0$ as $n \to \infty$. Then,

$$\left|\frac{\varrho_n}{\tau_n} - \frac{\varrho}{\tau}\right| = O_P(\max(\delta_{1n}, \delta_{2n})).$$

(d) (Wick's Lemma or Feynman Diagram Formula)
Let $(\xi_1, \xi_2, \xi_3, \xi_4)$ be a Gaussian random vector with zero mean. Then

$$E(\xi_1\xi_2\xi_3\xi_4) = E(\xi_1\xi_2)E(\xi_3\xi_4) + E(\xi_1\xi_3)E(\xi_2\xi_4) + E(\xi_1\xi_4)E(\xi_2\xi_3).$$

Lemma 3.1 (a) is from Michel and Pfanzagl (1971) and Lemma 3.1 (b) and (c) are from Bishwal and Bose (2001).

Lemma 3.2 (a) $E|X_t|^\eta \sim \left(\frac{2}{\beta}\right)^\eta \frac{\Gamma(\alpha/2+\eta)}{\Gamma(\alpha/2)}$ *as* $t \to \infty$. *Hence* $\sup_{t\geq 0} E(X_t)^\eta < \infty$.

(b) *For* $q \geq 1$, *with* $0 \leq s < t$ *such that* $0 < t - s < 1$,

$$E|X_t - X_s|^q \leq C(t-s)^{q/2}.$$

(c) *For* $\alpha > 4$,

$$E|X_t^{-1} - X_s^{-1}| \leq C(t-s)^{1/2}.$$

$$\text{(d)} \quad E\left|I_{n,T} - I_T\right|^2 = O\left(\frac{T^4}{n^2}\right), \qquad \text{(e)} \quad E\left|\frac{1}{2}(I_{n,T} + J_{n,T}) - I_T\right|^2 = O\left(\frac{T^4}{n^4}\right).$$

Proof: Parts (a)–(c) of the Lemma are proved in Propositions 3-4 of Ben Alaya and Kebaier [2]. See also Gikhman and Skorohod ([22], p.48) for general ergodic diffusions. We give a proof for part (d).

Let $h_i(t) := X_{t_{i-1}} - X_t$. Observe that

$$\begin{aligned}
&E|I_{n,T} - I_t|^2 = E\left|\sum_{i=1}^n X_{t_{i-1}}(t_i - t_{i-1}) - \int_0^T X_t\,dt\right|^2 = E\left|\textstyle\sum_{i=1}^n \int_{t_{i-1}}^{t_i} [X_{t_{i-1}} - X_t]\,dt\right|^2 \\
=\;& E\left|\sum_{i=1}^n \int_{t_{i-1}}^{t_i} h_i(t)\,dt\right|^2 = \sum_{i=1}^n E\left|\int_{t_{i-1}}^{t_i} h_i(t)\,dt\right|^2 + 2\sum_{i,j=1,\, i<j}^n E\int_{t_{i-1}}^{t_i}\int_{t_{j-1}}^{t_j} h_i(t)h_j(s)\,dt\,ds \\
=:\;& B_1 + B_2.
\end{aligned} \tag{3.61}$$

Note that

$$\begin{aligned}
&Eh_i^2(t) = E[X_{t_{i-1}} - X_t]^2 = E[\sqrt{X_{t_{i-1}}} - \sqrt{X_t}]^2[\sqrt{X}_{t_{i-1}} + \sqrt{X_t}]^2 \\
\leq\;& \{E[\sqrt{X_{t_{i-1}}} - \sqrt{X_t}]^4\}^{1/2}\{E[\sqrt{X_{t_{i-1}}} + \sqrt{X_t}]^4\}^{1/2} \leq C(t - t_{i-1})
\end{aligned} \tag{3.62}$$

(by (2.1) and the boundedness of the second term.)
Now

$$\begin{aligned}
&B_1 = \sum_{i=1}^n E\left|\int_{t_{i-1}}^{t_i} h_i(t)\,dt\right|^2 \leq \sum_{i=1}^n (t_i - t_{i-1})\int_{t_{i-1}}^{t_i} E(h_i^2(t))\,dt \\
\leq\;& C\tfrac{T}{n}\sum_{i=1}^n \int_{t_{i-1}}^{t_i} (t - t_{i-1})\,dt = C\frac{T^3}{n^2}.
\end{aligned} \tag{3.63}$$

Note, that

$$\begin{aligned}
&E[h_i(t)h_j(s)] = E(X_{t_{i-1}} - X_t)(X_{t_{j-1}} - X_s) \\
=\;& E(\sqrt{X_{t_{i-1}}} - \sqrt{X_t})(\sqrt{X_{t_{i-1}}} + \sqrt{X_t})(\sqrt{X_{t_{j-1}}} - \sqrt{X_s})(\sqrt{X_{t_{j-1}}} + \sqrt{X_s}).
\end{aligned} \tag{3.64}$$

Now, using Wick's lemma it is easy to see that $B_2 \leq C\frac{T^4}{n^2}$. Combining B_1 and B_2, the lemma follows.

Proof of part (e) is similar. We omit the details. □

Theorem 3.3 Let $r_{n,T} = T^{-1/2}(\log T)^{1/2} \bigvee \frac{T^2}{n}(\log T)^{-1}$. We have,

(a) $\sup_{x\in\mathbb{R}} \left| P\left\{ \left(\frac{T\alpha}{-4\beta}\right)^{1/2} (\beta_{n,T} - \beta) \le x \right\} - \Phi(x) \right| = O(r_{n,T})$.

(b) $\sup_{x\in\mathbb{R}} \left| P\left\{ I_{n,T}^{1/2}(\beta_{n,T} - \beta) \le x \right\} - \Phi(x) \right| = O(r_{n,T})$.

(c) $\sup_{x\in\mathbb{R}} \left| P\left\{ \left(\frac{T\alpha}{|4\beta_{n,T}|}\right)^{1/2} (\beta_{n,T} - \beta) \le x \right\} - \Phi(x) \right| = O(r_{n,T})$.

Proof : (a) It is easy to see that

$$\beta_{n,T} - \beta = \frac{Y_{n,T}}{I_{n,T}} + \beta\frac{V_{n,T}}{I_{n,T}} \tag{3.65}$$

Hence

$$\begin{aligned}
&\sup_{x\in\mathbb{R}} \left| P\left\{ \left(\frac{T\alpha}{-4\beta}\right)^{1/2} (\beta_{n,T} - \beta) \le x \right\} - \Phi(x) \right| \\
= &\sup_{x\in\mathbb{R}} \left| P\left\{ \left(\frac{T\alpha}{-4\beta}\right)^{1/2} \frac{Y_{n,T}}{I_{n,T}} + \left(\frac{T}{-2\beta}\right)^{1/2} \beta\frac{V_{n,T}}{I_{n,T}} \le x \right\} - \Phi(x) \right| \\
\le &\sup_{x\in\mathbb{R}} \left| P\left\{ \left(\frac{T\alpha}{-4\beta}\right)^{1/2} \frac{Y_{n,T}}{I_{n,T}} \le x \right\} - \Phi(x) \right| + P\left\{ \left| \left(\frac{T\alpha}{-4\beta}\right)^{1/2} \frac{V_{n,T}}{I_{n,T}} \right| > \epsilon \right\} + \epsilon. \\
=: &\ K_1 + K_2 + \epsilon.
\end{aligned} \tag{3.66}$$

Note that by Lemma 3.1 (b)

$$\begin{aligned}
K_1 &= \sup_{x\in\mathbb{R}} \left| P\left\{ \left(\frac{T\alpha}{-4\beta}\right)^{1/2} \frac{Y_{n,T}}{I_{n,T}} \le x \right\} - \Phi(x) \right| = \sup_{x\in\mathbb{R}} \left| P\left\{ \frac{\left(-\frac{4\beta}{T\alpha}\right)^{1/2} Y_{n,T}}{\left(\frac{-4\beta}{T\alpha}\right) I_{n,T}} \le x \right\} - \Phi(x) \right| \\
&\le \sup_{x\in\mathbb{R}} \left| P\left\{ \left(\frac{-4\beta}{T\alpha}\right)^{1/2} Y_{n,T} \le x \right\} - \Phi(x) \right| + P\left\{ \left(\frac{-4\beta}{T\alpha}\right) I_{n,T} - 1 > \epsilon \right\} + \epsilon \\
&=: J_1 + J_2 + \epsilon.
\end{aligned} \tag{3.67}$$

$$\begin{aligned}
J_1 &= \sup_{x\in\mathbb{R}} \left| P\left\{ \left(\frac{-4\beta}{T\alpha}\right)^{1/2} (Y_{n,T} - Y_T + Y_T) \le x \right\} - \Phi(x) \right| \\
&\le \sup_{x\in\mathbb{R}} \left| P\left\{ \left(\frac{-4\beta}{T\alpha}\right)^{1/2} Y_T \le x \right\} - \Phi(x) \right| + P\left\{ \left(\frac{-4\beta}{T\alpha}\right)^{1/2} |Y_{n,T} - Y_T| > \epsilon \right\} + \epsilon \\
&\le CT^{-1/2} + \left(\frac{-4\beta}{T\alpha}\right) \frac{E|Y_{n,T} - Y_T|^2}{\epsilon^2} + \epsilon \le CT^{-1/2} + C\frac{T/n}{\epsilon^2} + \epsilon \\
&\quad \text{(by Lemma 3.2(a).)}
\end{aligned} \tag{3.68}$$

$$\begin{aligned}
J_2 &= P\left\{ \left| \left(\frac{-4\beta}{T\alpha}\right) (I_{n,T} - I_T + I_T) - 1 \right| > \epsilon \right\} \\
&\le P\left\{ \left| \left(\frac{-4\beta}{T\alpha}\right) I_T - 1 \right| > \frac{\epsilon}{2} \right\} + P\left\{ \left(-\frac{4\beta}{2}\right) |I_{n,T} - I_T| > \frac{\epsilon}{2} \right\} \\
&\le C\exp\left(\frac{T\beta}{16}\epsilon^2\right) + \frac{16\beta^2}{T^2}\frac{E|I_{n,T} - I_T|^2}{\epsilon^2} \le C\exp\left(\frac{T\beta}{16}\epsilon^2\right) + C\frac{T^2/n^2}{\epsilon^2}.
\end{aligned} \tag{3.69}$$

Here the bound for the first term in (3.8) comes from Lemma 2.4(a) and that for the second term from Lemma 3.2(c). From the proof of Lemma 3.2(b) we have

$$E|V_{n,T}|^2 \leq C\tfrac{T^4}{n^2} \tag{3.70}$$

$$\begin{aligned}
K_2 &= P\left\{\left|\left(\frac{T\alpha}{-4\beta}\right)^{1/2}\beta\frac{V_{n,T}}{I_{n,T}}\right| > \epsilon\right\} = P\left\{\left|\frac{\left(\frac{-4\beta}{T\alpha}\right)^{1/2}\beta V_{n,T}}{\left(-\frac{4\beta}{T\alpha}\right)I_{n,T}}\right| > \epsilon\right\}\\
&= P\left\{\left|\left(\frac{-4\beta}{T\alpha}\right)^{1/2}\beta V_{n,T}\right| > \delta\right\} + P\left\{\left(-\frac{4\beta}{T\alpha}\right)I_{n,T} < \frac{\delta}{\epsilon}\right\}\\
&\quad\text{(where we choose } \delta = \epsilon - C\epsilon^2)\\
&\leq P\left\{\left|\left(\frac{-4\beta}{T\alpha}\right)^{1/2}\beta V_{n,T}\right| > \delta\right\} + P\left\{\left|\left(-\frac{4\beta}{T\alpha}\right)I_{n,T} - 1\right| > \delta_1\right\}\\
&\quad\text{(where } \delta_1 = \frac{\epsilon-\delta}{\delta} = C\epsilon)\\
&\leq -\frac{4\beta}{T\alpha}\beta^2\frac{E|V_{n,T}|^2}{\delta^2} + C\exp\left(\frac{T\beta}{16}\delta_1^2\right) + C\frac{T^2/n^2}{\delta_1^2}\\
&\leq C\frac{T^3/n^2}{\delta^2} + C\exp\left(\frac{T\beta}{16}\delta_1^2\right) + C\frac{T^2/n^2}{\delta_1^2} \quad \text{(by (3.9) and (3.8)).}
\end{aligned} \tag{3.71}$$

Now from (3.4) - (3.10), since $T/n \to 0$

$$\begin{aligned}
&\sup_{x\in\mathbb{R}}\left|P\left\{\left(\frac{T\alpha}{-4\beta}\right)^{1/2}(\beta_{n,T}-\beta) \leq x\right\} - \Phi(x)\right|\\
\leq\ & CT^{-1/2} + C\exp\left(\frac{T\beta}{16}\epsilon^2\right) + C\frac{T/n}{\epsilon^2} + C\frac{T^2/n^2}{\epsilon^2} + C\frac{T^3/n^2}{\delta^2} + C\exp\left(\frac{T\beta}{16}\delta_1^2\right) + C\left(\frac{T^2/n^2}{\delta_1^2}\right) + \epsilon.
\end{aligned} \tag{3.72}$$

Choosing $\epsilon = T^{-1/2}(\log T)^{1/2}$, the terms of (3.11) are of the order $O(T^{-1/2}(\log T)^{1/2} \bigvee \frac{T^2}{n}(\log T)^{-1})$. This proves (a).

(b) Using the expression (3.3)

$$\begin{aligned}
&\sup_{x\in\mathbb{R}}\left|P\left\{I_{n,T}^{1/2}(\beta_{n,T}-\beta) \leq x\right\} - \Phi(x)\right| = \sup_{x\in\mathbb{R}}\left|P\left\{\frac{Y_{n,T}}{I_{n,T}^{1/2}} + \beta\frac{V_{n,T}}{I_{n,T}^{1/2}} \leq x\right\} - \Phi(x)\right|\\
\leq\ & \sup_{x\in\mathbb{R}}\left|P\left\{\frac{Y_{n,T}}{I_{n,T}^{1/2}} \leq x\right\} - \Phi(x)\right| + P\left\{\left|\beta\frac{V_{n,T}}{I_{n,T}^{1/2}}\right| > \epsilon\right\} + \epsilon =: H_1 + H_2 + \epsilon.
\end{aligned} \tag{3.73}$$

Note that

$$\begin{aligned}
H_1 &= \sup_{x\in\mathbb{R}}\left|P\left\{\frac{Y_{n,T}-Y_T+Y_T}{I_{n,T}^{1/2}} \leq x\right\} - \Phi(x)\right|\\
&= \sup_{x\in\mathbb{R}}\left|P\left\{\frac{Y_T}{I_{n,T}^{1/2}} \leq x\right\} - \Phi(x)\right| + P\left\{\frac{|Y_{n,T}-Y_T|}{I_{n,T}^{1/2}} > \epsilon\right\} + \epsilon =: F_1 + F_2 + \epsilon.
\end{aligned} \tag{3.74}$$

Now

$$
\begin{aligned}
F_1 &= \sup_{x\in\mathbb{R}} \left| P\left\{ \frac{Y_T}{I_{n,T}^{1/2}} \le x \right\} - \Phi(x) \right| \le \sup_{x\in\mathbb{R}} \left| P\left\{ \left(\frac{-4\beta}{T\alpha}\right)^{1/2} Y_T \le x \right\} - \Phi(x) \right| \\
&\quad + P\left\{ \left| \left(\frac{-4\beta}{T\alpha}\right)^{1/2} I_{n,T}^{1/2} - 1 \right| > \epsilon \right\} + \epsilon \text{ (by Lemma 3.1(b))} \\
&\le C_T^{-1/2} + P\left\{ \left| \left(\frac{-4\beta}{T\alpha}\right) I_{n,T} - 1 \right| > \epsilon \right\} + \epsilon \\
&\le CT^{-1/2} + C\exp\left(-\frac{T\beta}{16}\epsilon^2\right) + C\frac{T^2/n^2}{\epsilon^2} + \epsilon \quad \text{(by (3.8))}.
\end{aligned}
\tag{3.75}
$$

On the other hand,

$$
\begin{aligned}
F_2 &= P\left\{ \frac{|Y_{n,T} - Y_T|}{I_{n,T}^{1/2}} > \epsilon \right\} \\
&\le P\left\{ \left(\frac{-4\beta}{T\alpha}\right)^{1/2} |Y_{n,T} - Y_T| > \delta \right\} + P\left\{ \left| \left(\frac{-4\beta}{T\alpha}\right)^{1/2} I_{n,T}^{1/2} - 1 \right| > \delta_1 \right\} \\
&\quad \text{(where } \delta = \epsilon - C\epsilon^2 \textit{ and } \delta_1 = (\epsilon - \delta)/\epsilon > 0) \\
&\le \frac{\left(\frac{-4\beta}{T\alpha}\right) E|Y_{n,T} - Y_T|^2}{\delta^2} + P\left\{ \left| \left(\frac{-2\beta}{T}\right) I_{n,T} - 1 \right| > \delta_1 \right\} \\
&\le C\frac{T/n}{\delta^2} + C\exp\left(\frac{T\beta}{16}\delta_1^2\right) + C\frac{T^2/n^2}{\delta_1^2} \text{ (from Lemma 3.2(a) and (3.8).)}
\end{aligned}
\tag{3.76}
$$

Using (3.15) and (3.14) in (3.13), we obtain

$$
\begin{aligned}
H_1 &= \sup_{x\in\mathbb{R}} \left| P\left\{ \frac{Y_T}{I_{n,T}^{1/2}} \le x \right\} - \Phi(x) \right| \\
&\le CT^{-1/2} + C\exp\left(\frac{T\beta}{16}\epsilon^2\right) + C\frac{T/n}{\delta^2} + C\frac{T^2/n^2}{\delta_1^2} + C\exp(\frac{T\beta}{16}\delta_1^2) + C\frac{T^2/n^2}{\epsilon^2} + \epsilon.
\end{aligned}
\tag{3.77}
$$

$$
\begin{aligned}
H_2 &= P\left\{ \left| \beta \frac{V_{n,T}}{I_{n,T}^{1/2}} \right| > \epsilon \right\} = P\left\{ \frac{\left| \left(\frac{-4\beta}{T\alpha}\right)^{1/2} \beta V_{n,T} \right|}{\left| \left(\frac{-4\beta}{T\alpha}\right)^{1/2} I_{n,T}^{1/2} \right|} > \epsilon \right\} \\
&\le P\left\{ \left| \left(\frac{-4\beta}{T\alpha}\right)^{1/2} \beta V_{n,T} \right| > \delta \right\} + P\left\{ \left| \left(\frac{-4\beta}{T\alpha}\right)^{1/2} I_{n,T}^{1/2} \right| < \delta/\epsilon \right\} \\
&\le \left(-\frac{4\beta}{T\alpha}\right) \beta^2 \frac{E|V_{n,T}|^2}{\delta^2} + P\left\{ \left| \left(-\frac{2\beta}{T}\right) I_{n,T} - 1 \right| > \delta_1 \right\} \\
&\quad \text{(where } 0 < \delta < \epsilon \text{ and } \delta_1 = (\epsilon - \delta)/\epsilon = C\epsilon > 0) \\
&\le C\frac{T^2/n^2}{\delta^2} + C\exp\left(\frac{T\beta}{16}\delta_1^2\right) + C\frac{T^2/n^2}{\delta_1^2}. \text{ (from (3.9) and (3.8))}
\end{aligned}
\tag{3.78}
$$

Using (3.17) and (3.16) in (3.12) and choosing $\epsilon = T^{-1/2}(\log T)^{1/2}$ the terms of (3.12) are of the order $O(T^{-1/2}(\log T)^{1/2} \bigvee \frac{T^2}{n}(\log T)^{-1})$. This proves (b).

(c) Let $D_T = \{|\beta_{n,T} - \beta| \le d_T\}$ and $d_T = CT^{-1/2}(\log T)^{1/2}$. On the set D_T, expanding $(2|\beta_{n,T}|)^{-1/2}$, we obtain

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

Then

$$\begin{aligned} &\sup_{x\in\mathbb{R}}\left|P\left\{\left(\frac{T\alpha}{4|\beta_{n,T}|}\right)^{1/2}(\beta_{n,T}-\beta)\le x\right\}-\Phi(x)\right| \\ \le\ &\sup_{x\in\mathbb{R}}\left|P\left\{\left(\frac{T\alpha}{4|\beta_{n,T}|}\right)^{1/2}(\beta_{n,T}-\beta)\le x, D_T\right\}-\Phi(x)\right|+P(D_T^c). \end{aligned} \tag{3.79}$$

$$\begin{aligned} &P(D_T^c)=P\left\{|\beta_{n,T}-\beta|>CT^{-1/2}(\log T)^{1/2}\right\} \\ =\ &P\left\{\left(\frac{T\alpha}{-4\beta}\right)^{1/2}|\beta_{n,T}-\beta|>C(\log T)^{1/2}(-4\beta)^{-1/2}\right\} \\ \le\ &C(T^{-1/2}(\log T)^{1/2}\bigvee\frac{T^2}{n}(\log T)^{-1})+2(1-\Phi\left((\log T)^{1/2}(-2\beta)^{-1/2}\right) \\ &\text{(by Theorem 3.3(a))} \\ \le\ &C(T^{-1/2}(\log T)^{1/2}\bigvee\frac{T^2}{n}(\log T)^{-1}). \end{aligned}$$

On the set D_T,

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

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

$$\begin{aligned} &\left|P\left\{\left(\frac{T\alpha}{-4\beta_{n,T}}\right)^{1/2}(\beta_{n,T}-\beta)\le x, D_T\right\}-\Phi(x)\right| \\ \le\ &\left|P\left\{\left(\frac{T\alpha}{-4\beta}\right)^{1/2}(\beta_{n,T}-\beta)\le x, D_T\right\}-\Phi(x)\right|+P\left\{\left|\left(\frac{\beta_{n,T}}{\beta}\right)^{1/2}-1\right|>\epsilon, D_T\right\}+\epsilon \\ &\text{(by Lemma 3.1(b))} \\ \le\ &C(T^{-1/2}(\log T)^{1/2}\bigvee\frac{T^2}{n}(\log T)^{-1}) \\ &\text{(by Theorem 3.3(a)).} \end{aligned} \tag{3.80}$$

(c) follows from (3.18) - (3.20). □

Theorem 3.4

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

Proof : Let $a_{n,T}:=Z_{n,T}-Z_T$, $b_{n,T}:=I_{n,T}-I_T$. By Lemma 3.2, we have

$$E|a_{n,T}|^2=O\left(\frac{T^2}{n}\right)\text{ and }E|b_{n,T}|^2=O\left(\frac{T^4}{n^2}\right). \tag{3.81}$$

From (3.5)

$$I_{n,T}\beta_{n,T}=\sum_{i=1}^{n}X_{t_{i-1}}\left[X_{t_i}-X_{t_{i-1}}\right]=\int_0^T X_tdX_t+a_{n,T}=\int_0^T X_tdW_t+\beta\int_0^T X_t^2dt+a_{n,T}.$$

Hence $I_{n,T}(\beta_{n,T}-\beta)=-\beta b_{n,T}+a_{n,T}$.

Thus

$$\begin{aligned}
&\sup_{x\in\mathbb{R}}\left|P\left\{I_{n,T}\left(-\frac{4\beta}{T\alpha}\right)^{1/2}(\beta_{n,T}-\beta)\le x\right\}-\Phi(x)\right| \\
=&\sup_{x\in\mathbb{R}}\left|P\left\{\left(-\frac{4\beta}{T\alpha}\right)^{1/2}[Y_T-\beta b_{n,T}+a_{n,T}]\le x\right\}-\Phi(x)\right| \\
\le&\sup_{x\in\mathbb{R}}\left|P\left\{\left(-\frac{4\beta}{T\alpha}\right)^{1/2}Y_T\le x\right\}-\Phi(x)\right|+P\left\{\left|\left(-\frac{4\beta}{T\alpha}\right)^{1/2}[-\beta b_{n,T}+a_{n,T}]\right|>\epsilon\right\}+\epsilon \\
\le& CT^{-1/2}+\left(-\frac{4\beta}{T\alpha}\right)\frac{E|-\beta b_{n,T}+a_{n,T}|^2}{\epsilon^2}+\epsilon \\
\le& CT^{-1/2}+C\frac{T/n}{\epsilon^2}+\epsilon \text{ (by (3.21))}.
\end{aligned}$$

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

Theorem 3.5

$$|\beta_{n,T}-\beta_T|=O_P(\frac{T^2}{n})^{1/2}.$$

Proof : Note that $\beta_{n,T}-\beta_T=\frac{Z_{n,T}}{I_{n,T}}-\frac{Z_T}{I_T}$. From Lemma 3.2 it follows that $|Z_{n,T}-Z_T|=O_P(\frac{T^2}{n})^{1/2}$ and $|I_{n,T}-I_T|=O_P(\frac{T^4}{n^2})^{1/2}$. Now the theorem follows easily from the from the Lemma 3.1. □

Bessel Process: If $\beta=0$, the model is 2-dimensional Bessel process

$$dX_t=\alpha dt+2\sqrt{X_t}dW_t$$

the MLE is

$$\hat{\alpha}_T=\frac{\int_0^T X_t^{-1}dX_t}{\int_0^T X_t^{-2}dt}=\frac{\log X_T-\log X_0+2\int_0^T X_t^{-1}dt}{\int_0^T X_t^{-1}dt}.$$

In this case we have a different rate of convergence.

Theorem 3.6 Denote $b_{n,T}:=O(\max((\log T)^{-1/2},(\frac{T^4}{n^2})(\log T)^{-2}))$. If $\beta=0$ and $\alpha>2$, then

$$\text{(a)}\quad \sup_{x\in\mathbb{R}}\left|P\left\{\left(\frac{\log T}{4(\alpha-2)}\right)^{1/2}(\tilde{\alpha}_{n,T,F}-\alpha)\le x\right\}-\Phi(x)\right|=O(b_{n,T}),$$

$$\text{(b)}\quad \sup_{x\in\mathbb{R}}\left|P\left\{I_{n,T}^{1/2}(\tilde{\alpha}_{n,T,F}-\alpha)\le x\right\}-\Phi(x)\right|=O(b_{n,T}),$$

$$\text{(c)}\quad \sup_{x\in\mathbb{R}}\left|P\left\{\left(\frac{\log T}{4|\tilde{\alpha}_{n,T,F}-2|}\right)^{1/2}(\tilde{\alpha}_{n,T,F}-\alpha)\le x\right\}-\Phi(x)\right|=O(b_{n,T}).$$

Theorem 3.7 Denote $b_{n,T}:=O(\max(T^{-1/2}(\log T)^{1/2},(\frac{T^4}{n^2})(\log T)^{-1}))$. If $\beta<0$ and $\alpha>2$, then

$$\text{(a)}\quad \sup_{x\in\mathbb{R}}\left|P\left\{\left(-\frac{T\beta}{4(\alpha-2)}\right)^{1/2}(\tilde{\alpha}_{n,T,F}-\alpha)\le x\right\}-\Phi(x)\right|=O(b_{n,T}),$$

$$\text{(b)}\quad \sup_{x\in\mathbb{R}}\left|P\left\{I_{n,T}^{1/2}(\tilde{\alpha}_{n,T,F}-\alpha)\le x\right\}-\Phi(x)\right|=O(b_{n,T}),$$

$$\text{(c)}\quad \sup_{x\in\mathbb{R}}\left|P\left\{\left(\frac{T\beta}{4|\widetilde{\alpha}_{n,T,F}-2|}\right)^{1/2}(\widetilde{\alpha}_{n,T,F}-\alpha)\le x\right\}-\Phi(x)\right| = O(b_{n,T}).$$

Proof (a) Observe that

$$\left(-\frac{T\beta}{4(\alpha-2)}\right)^{1/2}(\widetilde{\alpha}_T-\alpha)=\frac{\left(-\frac{4(\alpha-2)}{T\beta}\right)^{1/2}Y_T}{\left(-\frac{4(\alpha-2)}{T\beta}\right)I_T} \tag{3.82}$$

where

$$Y_T := -(\alpha-2)I_T-\frac{T\beta}{4}\quad\text{and}\quad I_T:=\int_0^T X_t^2dt.$$

Thus, we have $I_{n,T}\widetilde{\alpha}_{n,T,F}=Y_T+\alpha I_T$. Hence,

$$\begin{aligned}\left(-\frac{T\beta}{4(\alpha-2)}\right)^{1/2}(\widetilde{\alpha}_{n,T,F}-\alpha) &= \frac{\left(-\frac{T\beta}{2\alpha}\right)^{1/2}Y_T+\alpha\left(-\frac{T\beta}{2\alpha}\right)^{1/2}(I_T-I_{n,T})}{I_{n,T}}\\ =\ &\frac{\left(-\frac{4(\alpha-2)}{T\beta}\right)^{1/2}Y_T+\left(-\frac{4(\alpha-2)}{T\beta}\right)^{1/2}(I_T-I_{n,T})}{\left(-\frac{4(\alpha-2)}{T\beta}\right)I_{n,T}}.\end{aligned} \tag{3.83}$$

Further,

$$\begin{aligned}&P\left\{\left|\left(\frac{-4(\alpha-2)}{T\beta}\right)(I_{n,T}-1)\right|>\epsilon\right\}=P\left\{\left|\left(\frac{-4(\alpha-2)}{T\beta}\right)(I_{n,T}-I_T+I_T)-1\right|>\epsilon\right\}\\ \le\ &P\left\{\left|\left(\frac{-4(\alpha-2)}{T\beta}\right)I_T-1\right|>\frac{\epsilon}{2}\right\}+P\left\{\left(-\frac{4(\alpha-2)}{T\beta}\right)|I_{n,T}-I_T|>\frac{\epsilon}{2}\right\}\\ \le\ &C\exp\left(\frac{T\alpha}{16}\epsilon^2\right)+\frac{16\alpha^2}{T^2}\frac{E|I_{n,T}-I_T|^2}{\epsilon^2}\le C\exp\left(\frac{T\alpha}{16}\epsilon^2\right)+C\frac{T^2/n^2}{\epsilon^2}.\end{aligned} \tag{3.84}$$

Next, observe that

$$\begin{aligned}&\sup_{x\in\mathbb{R}}\left|P\left\{\left(-\frac{T}{2\alpha}\right)^{1/2}(\widetilde{\alpha}_{n,T,F}-\alpha)\le x\right\}-\Phi(x)\right|\\ =\ &\sup_{x\in\mathbb{R}}\left|P\left\{\frac{\left(-\frac{4(\alpha-2)}{T\beta}\right)^{1/2}Y_T+\left(-\frac{4(\alpha-2)}{T}\right)^{1/2}(I_T-I_{n,T})}{\left(-\frac{2\alpha}{T}\right)I_{n,T}}\le x\right\}-\Phi(x)\right|\\ \le\ &\sup_{x\in\mathbb{R}}\left|P\left\{\left(-\frac{4(\alpha-2)}{T\beta}\right)^{1/2}Y_T\le x\right\}-\Phi(x)\right|\\ &+P\left\{\left|\alpha\left(-\frac{4(\alpha-2)}{T\beta}\right)^{1/2}(I_{n,T}-I_T)\right|>\epsilon\right\}+P\left\{\left|\left(-\frac{4(\alpha-2)}{T\beta}\right)I_{n,T}-1\right|>\epsilon\right\}+2\epsilon\\ \le\ &CT^{-1/2}+\alpha^2\frac{\left(-\frac{4(\alpha-2)}{T\beta}\right)E|I_{n,T}-I_T|^2}{\epsilon^2}+C\exp\left(\frac{T\alpha}{4}\epsilon^2\right)+C\frac{T^2}{n^2\epsilon^2}+2\epsilon,\end{aligned} \tag{3.85}$$

(the bound for the 3rd term in the right hand side of (2.4)is obtained from (2.3))

$$\le\ CT^{-1/2}+C\frac{T^2}{n^2\epsilon^2}+C\exp\left(\frac{T\alpha}{4}\epsilon^2\right)+C\frac{T}{n^2\epsilon^2}+\epsilon \tag{3.86}$$

(by Lemma 2.3(a)).

Choosing $\epsilon = CT^{-1/2}(\log T)^{1/2}$, the terms in the right hand side of (2.5) are of the order $O(\max(T^{-1/2}(\log T)^{1/2}, (\frac{T^4}{n^2})(\log T)^{-1}))$. □

(b) From (2.1), we have

$$I_{n,T}^{1/2}(\widetilde{\alpha}_{n,T,F} - \alpha) = \frac{Y_T + \alpha(I_T - I_{n,T})}{I_{n,T}^{1/2}}.$$

Then,

$$\begin{aligned}\sup_{x\in\mathbb{R}} \left| P\left\{ I_{n,T}^{1/2}(\widetilde{\alpha}_{n,T,F} - \alpha) \le x \right\} - \Phi(x) \right| &= \sup_{x\in\mathbb{R}} \left| P\left\{ \frac{Y_T}{I_{n,T}^{1/2}} + \alpha \frac{I_T - I_{n,T}}{I_{n,T}^{1/2}} \le x \right\} - \Phi(x) \right| \\ \le \sup_{x\in\mathbb{R}} \left| P\left\{ \frac{Y_T}{I_{n,T}^{1/2}} \le x \right\} - \Phi(x) \right| + P\left\{ \left| \frac{\alpha(I_T - I_{n,T})}{I_{n,T}^{1/2}} \right| > \epsilon \right\} + \epsilon &=: U_1 + U_2 + \epsilon.\end{aligned}$$

We have from (2.3),

$$U_1 \le CT^{-1/2} + C\exp\left(\frac{T\alpha}{16}\epsilon^2\right) + C\frac{T^2}{n^2\epsilon^2} + \epsilon. \tag{3.88}$$

Now,

$$\begin{aligned} U_2 &= P\left\{ |\alpha| \left| \frac{I_{n,T} - I_T}{I_{n,T}^{1/2}} \right| > \epsilon \right\} = P\left\{ |\alpha| \frac{\left| \left(-\frac{4(\alpha-2)}{T}\right)^{1/2} (I_{n,T} - I_T) \right|}{\left| \left(-\frac{2\alpha}{T}\right)^{1/2} I_{n,T}^{1/2} \right|} > \epsilon \right\} \\ &\le P\left\{ \left| \left(-\frac{4(\alpha-2)}{T}\right)^{1/2} \right| |I_{n,T} - I_T| > \delta \right\} + P\left\{ \left| \left(-\frac{4(\alpha-2)}{T}\right)^{1/2} I_{n,T}^{1/2} - 1 \right| > \delta_1 \right\} \quad (3.89) \\ &\quad \text{(where } \delta = \epsilon - C\epsilon^2 \text{ and } \delta_1 = (\epsilon - \delta)/\epsilon > 0) \\ &\le \left(-\frac{4(\alpha-2)}{T}\right) \frac{E|I_{n,T} - I_T|^2}{\delta^2} + P\left\{ \left| \left(-\frac{4(\alpha-2)}{T}\right) I_{n,T} - 1 \right| > \delta_1 \right\} \\ &\le C\frac{T^3}{n^2\delta^2} + C\exp\left(\frac{T\alpha}{16}\delta_1^2\right) + C\frac{T^2}{n^2\delta_1^2}. \quad (3.90) \end{aligned}$$

Here, the bound for the first term in the right hand side of (3.89) comes from Lemma 2.2(c) and that for the second term is obtained from (2.3).

Now, using the bounds (3.88) and (3.90) in (3.87) with $\epsilon = CT^{-1/2}(\log T)^{1/2}$, we obtain that the terms in (3.87) are of the order $O(\max(T^{-1/2}(\log T)^{1/2}, (\frac{T^4}{n^2})(\log T)^{-1}))$. □

(c) Let $G_T := \{|\widetilde{\alpha}_{n,T,F} - \alpha| \le d_T\}$, and $d_T := CT^{-1/2}(\log T)^{1/2}$. On the set G_T, expanding $(2|\widetilde{\alpha}_{n,T,F}|)^{1/2}$, we obtain

$$(-2\widetilde{\alpha}_{n,T})^{-1/2} = (-4(\alpha-2))^{1/2}\left[1 - \frac{\alpha - \widetilde{\alpha}_{n,T,F}}{\alpha}\right]^{-1/2} = (-2\alpha)^{1/2}\left[1 + \frac{1}{2}\left(\frac{\alpha - \widetilde{\alpha}_{n,T,F}}{\alpha}\right)\right] + O(d_T^2).$$

Then,

$$\begin{aligned} &\sup_{x\in\mathbb{R}} \left| P\left\{ \left(\frac{T\beta}{4|\widetilde{\alpha}_{n,T,F-2}|}\right)^{1/2} (\widetilde{\alpha}_{n,T,F} - \alpha) \le x \right\} - \Phi(x) \right| \\ \le &\sup_{x\in\mathbb{R}} \left\{ P\left(\frac{T\beta}{2|\widetilde{\alpha}_{n,T,F-2}|}\right)^{1/2} (\widetilde{\alpha}_{n,T,F} - \alpha) \le x, G_T \right\} + P(G_T^c). \end{aligned}$$

Now,

$$\begin{aligned} P(G_T^c) &= P\left\{|\widetilde{\alpha}_{n,T,F}-\alpha| > CT^{-1/2}(\log T)^{1/2}\right\} \\ &= P\left\{\left(-\frac{T\beta}{4(\alpha-2)}\right)^{1/2}|\widetilde{\alpha}_{n,T,F}-\alpha| > C(\log T)^{1/2}(-4(\alpha-2))^{-1/2}\right\} \\ &\leq C\max\left(T^{-1/2}(\log T)^{1/2}, \frac{T^3}{n^2}(\log T)^{-1}\right) + 2(1-\Phi\log T^{1/2}(-4(\alpha-2))^{-1/2}) \\ &\quad \text{(by Theorem 2.1(a))} \\ &\leq C\max\left(T^{-1/2}(\log T)^{1/2}, \frac{T^3}{n^2}(\log T)^{-1}\right). \end{aligned}$$

On the set G_T,

$$\left|\left(\frac{\widetilde{\alpha}_{n,T,F}}{\alpha}\right)^{1/2} - 1\right| \leq CT^{-1/2}(\log T)^{1/2}.$$

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

$$\begin{aligned} &\left|P\left\{\left(\frac{T\beta}{-4(\widetilde{\alpha}_{n,T,F}-2)}\right)^{1/2}(\widetilde{\alpha}_{n,T,F}-\alpha) \leq x, G_T\right\} - \Phi(x)\right| \\ \leq &\left|P\left\{\left(\frac{T\beta}{-4(\alpha-2)}\right)^{1/2}(\widetilde{\alpha}_{n,T,F}-\alpha) \leq x, G_T\right\}\right| + P\left\{\left|\left(\frac{\widetilde{\alpha}_{n,T,F}}{\alpha}\right)^{1/2} - 1\right| > \epsilon, G_T\right\} + \epsilon \\ &\text{(by Lemma 3.1(b))} \\ \leq &\ C\max\left(T^{-1/2}(\log T)^{1/2}, \frac{T^4}{n^2}(\log T)^{-1}\right) \\ &\text{(by Theorem 2.1(a)).} \end{aligned}$$

□

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 AMCEs need $T \to \infty$ and $\frac{T}{n^{2/3}} \to 0$ which are sharper than the bound in Theorem 2.1.

Theorem 3.8

$$\sup_{x\in\mathbb{R}}\left|P\left\{I_{n,T}\left(-\tfrac{4(\alpha-2)}{T\beta}\right)^{1/2}(\widetilde{\alpha}_{n,T,F}-\alpha) \leq x\right\} - \Phi(x)\right| = O\left(\max\left(T^{-1/2}, \left(\tfrac{T^3}{n^2}\right)^{1/3}\right)\right).$$

Proof. From (2.2), we have

$$I_{n,T}\left(-\tfrac{4(\alpha-2)}{T}\right)^{1/2}(\widetilde{\alpha}_{n,T,F}-\alpha) = \left(-\tfrac{4(\alpha-2)}{T\beta}\right)^{1/2}Y_T + \alpha\left(-\tfrac{4(\alpha-2)}{T\beta}\right)^{1/2}(I_T - I_{n,T}).$$

Hence, by Lemma 2.1–2.3

$$\begin{aligned} &\sup_{x\in\mathbb{R}}\left|P\left\{I_{n,T}\left(-\tfrac{4(\alpha-2)}{T\beta}\right)^{1/2}(\widetilde{\alpha}_{n,T,F}-\alpha) \leq x\right\} - \Phi(x)\right| \\ = &\sup_{x\in\mathbb{R}}\left|P\left\{\left(-\tfrac{4(\alpha-2)}{T\beta}\right)^{1/2}Y_T + \alpha\left(-\tfrac{4(\alpha-2)}{T\beta}\right)^{1/2}(I_T - I_{n,T}) \leq x\right\} - \Phi(x)\right| \\ \leq &\sup_{x\in R}\left|P\left\{\left(-\tfrac{4(\alpha-2)}{T\beta}\right)^{1/2}Y_T \leq x\right\} - \Phi(x)\right| + P\left\{\left|\alpha\left(-\tfrac{4(\alpha-2)}{T\beta}\right)^{1/2}(I_T - I_{n,T})\right| > \epsilon\right\} + \epsilon \\ \leq &\ CT^{-1/2} + C\tfrac{E|I_T - I_{n,T}|^2}{T\epsilon^2} + \epsilon \leq CT^{-1/2} + C\tfrac{T^3}{n^2\epsilon^2} + \epsilon. \end{aligned}$$

Choosing $\epsilon = (\frac{T^3}{n^2})^{1/3}$, the theorem follows. □

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

Theorem 3.9 (a) $|\widetilde{\alpha}_{n,T} - \widetilde{\alpha}_T| = O_P\left(\frac{T}{n}\right)^{1/2}$, (b) $|\widetilde{\alpha}_{n,T,z} - \widetilde{\alpha}_T| = O_P\left(\frac{T^2}{n^2}\right)^{1/2}$.

Proof From (1.9) and (1.14), we have $\widetilde{\alpha}_T = -\frac{T}{2I_T}$, $\widetilde{\alpha}_{n,T} = -\frac{T}{2I_{n,T}}$. Hence, applying Lemma 3.1 (c) with the aid of Lemma 3.2 (d) and noting that $|\frac{I_{n,T}}{T}| = O_P(1)$ and $|\frac{I_T}{T}| = O_P(1)$ the part (a) of theorem follows. From (1.9) and (1.16), we have $\widetilde{\alpha}_T = -\frac{T}{2I_T}$, $\widetilde{\alpha}_{n,T,z} = -\frac{T}{I_{n,t}+J_{n,T}}$. Next applying Lemma 3.1 (c) with the aid of Lemma 3.2 (e) and noting that $|\frac{J_{n,T}}{T}| = O_P(1)$ and $|\frac{I_T}{T}| = O_P(1)$ the part (b) of theorem follows. □

References

[1] M. Ben Alaya, A. Kebaier, Parameter estimation for the square-root diffusions: Ergodic and nonergodic cases, Stoch. Models. 28 (2012), 609-634.

[2] M. Ben Alaya, A. Kebaier, Asymptotic behavior of the maximum likelihood estimator for ergodic and nonergodic square-root diffusions, Stoch. Anal. Appl. 31 (2013), 552-573.

[3] J.P.N. Bishwal, Sharp Berry-Esseen bound for the maximum likelihood estimator in the Ornstein-Uhlenbeck process, Sankhyā, Ser. A. 62 (2000), 1-10.

[4] J.P.N. Bishwal, Rates of convergence of the posterior distributions and the Bayes estimators in the Ornstein-Uhlenbeck process, Rand. Oper. Stoch. Equ. 8 (2000), 51-70.

[5] J.P.N. Bishwal, Accuracy of normal approximation for the maximum likelihood and Bayes estimators in the Ornstein-Uhlenbeck process using random normings, Stat. Prob. Lett. 52 (2001), 427-439.

[6] J.P.N. Bishwal, Rate of weak convergence of the approximate minimum contrast estimators for discretely observed Ornstein-Uhlenbeck process, Stat. Prob. Lett. 76 (2006), 1397-1409.

[7] J.P.N. Bishwal, A new estimating function for discretely sampled diffusions, Rand. Oper. Stoch. Equ. 15 (2007), 65-88.

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

[9] J.P.N. Bishwal, Parameter estimation in stochastic volatility models, Springer Nature, Cham. (2022).

[10] J.P.N. Bishwal, Berry-Esseen inequalities for discretely observed diffusions, Monte Carlo Meth. Appl. 15 (2009), 229-239.

[11] J.P.N. Bishwal, Uniform rate of weak convergence of the minimum contrast estimator in the Ornstein-Uhlenbeck process, Meth. Comp. Appl. Prob. 12 (2010), 323-334.

[12] 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.

[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] A.A. Borokov, On the rate of convergence for the invariance principle, Theory Prob. Appl. 18 (1973), 217-234.

[15] D. Brigo, F. Mercurio, Interest rate models: Theory and practice (with smile, inflation and credit), Second Edition, Springer, Hiedelberg, (2001).

[16] J.C. Cox, Notes on option pricing I: Constant elasticity of diffusions, J. Portf. Manage. 23 (1996), 15-17.

[17] J.C. Cox, J.E. Ingersoll, S.A. Ross, A theory of term structure of interest rates, Econometrica. 53 (1985), 385-407.

[18] O. Dehtiar, Y. Mishura, K. Ralchenko, Two methods of estimation of the drift parameters of the Cox-Ingersoll-Ross process; Continuous observations, Comm. Stat.-Theory Meth. 50 (2021), 1-16.

[19] D. Dacunha-Castelle, D. Florens-Zmirou, Estimation of the coefficients of a diffusion from discrete observations, Stochastics. 19 (1986), 263-284.

[20] W. Feller, Two singular diffusion problems, Ann. Math. 54 (1951), 173-182.

[21] W. Feller, An introduction to probability theory and its applications, Vol. I, Wiley, New York, (1957).

[22] I.I. Gikhman, A.V. Skorohod, Stochastic differential equations, Springer-Verlag, Berlin, (1972).

[23] J.C. Hull, Options, Futures and other derivatives, Ninth Edition, Prentice Hall, (2015).

[24] A. Lenkasas, V. Mackevicius, A second order weak approximation of Heston model by discrete random variables, Lithuanian Math. J. 55 (2015), 555-572.

[25] G. Lileika, V. Mackevicius, Weak approximation of CKLS and CEV process by discrete random variables, Lithuanian Math. J. 60 (2020), 208-224.

[26] V. Mackevicius, Weak approximation of CIR equation by discrete random variables, Lithuanian Math. J. 51 (2011), 385-401.

[27] V. Mackevicius, Verhulst versus CIR, Lithuanian Math. J. 55 (2015), 119-133.

[28] R. Michel, J. Pfanzagl, The accuracy of the normal approximation for minimum contrast estimate, Zeit Wahr. Verw. Gebiete 18 (1971), 73-84.

[29] Y. Mishura, A. Yurchenko-Tytarenko, Fractional Cox-Ingersoll-Ross model with non-zero mean, Mod. Stoch.: Theory Meth. 5 (2018), 99-111.

[30] L. Overbeck, T. Ryden, Estimation in the Cox-Ingersoll-Ross model, Econ. Theory 13 (1997), 430-461.