Eur. J. Math. Anal. 3 (2023) 25
doi: 10.28924/ada/ma.3.25

On the Kolmogorov Distance for the Maximum Likelihood Estimator in the Explosive Ornstein-Uhlenbeck Process

Jaya P. N. Bishwal

Department of Mathematics and Statistics, University of North Carolina at Charlotte, 376 Fretwell Bldg, 9201 University City Blvd. Charlotte, NC 28223-0001, USA

Correspondence: J.Bishwal@uncc.edu

Abstract. The paper estimates the Kolmogorov distance between the distribution of the normalized maximum likelihood estimator of the positive drift parameter in the nonergodic Ornstein-Uhlenbeck process and the standard Cauchy distribution and shows exponential error rate for large time limit.

1. Introduction

Estimating the rate in the Kolmogorov distance between two distributions has a long history in probability and statistics. The estimate could be useful in finding confidence interval and in hypothesis testing, see Bishwal [8, 11]. In the i.i.d. case, the Berry-Esseen bound for minimum contrast estimators was obtained in Pfanzagl [28] improving that from Michel and Pfanzagl [24]. Borokov [14] obtained the rate of convergence for the invariance principle in the i.i.d. case. Hall and Heyde [20] obtained rate of convergence in the central limit theorem for martingales using Skorohod embedding. Uniform rate of weak convergence for the minimum contrast estimator in the Ornstein-Uhlenbeck (O-U) process was studied in Bishwal [5]. The rates of convergence of the conditional least squares estimator and an approximate maximum likelihood estimator when the O-U process is observed at discrete time points in $[0, T]$ has been studied in Bishwal and Bose [13](2001) in the ergodic case. In a Bayesian framework, the rates of convergence of the posterior distributions and the Bayes estimators has been studied in Bishwal [6] and Bishwal [10] for the continuous observation and discrete observations respectively in the ergodic case. In finance, asset price may behave in nonergodic manner, i.e., efficient market hypotheses may not hold, possibly be due to social interaction among consumers among other reasons, see Horst and Wenzelburger [21]. We study the nonergodic Ornstein-Uhlenbeck process in this paper and focus on the rate of convergence of the Kolmogorov distance.

Received: 5 Jul 2023.
Key words and phrases. Itô stochastic differential equation, explosive Ornstein-Uhlenbeck process, maximum likelihood estimator, Kolmogorov distance, inefficient market.

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

$$dX_t = \theta X_t dt + dW_t, t \geq 0, \quad X_0 = 0 \tag{1.1}$$

where $\{W_t\}_{t\geq 0}$ is a standard Wiener process with respect to the filtration $\{\mathcal{F}_t\}_{t\geq 0}$ and $\theta > 0$ is the unknown parameter to be estimated on the basis of continuous observation of the process $\{X_t\}_{t\geq 0}$ on the time interval $[0, T]$.

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

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

Maximizing the log-likelihood with respect to θ provides the maximum likelihood estimate (MLE)

$$\theta_T := \frac{\int_0^T X_t dX_t}{\int_0^T X_t^2 dt}. \tag{1.3}$$

In this transient case, we show that this estimator converges to the Cauchy distribution with an error rate $O(e^{-\theta T})$. Note that in the transient case, with random norming, specifically if one normalizes the MLE by the square root of the observed Fisher information, then the MLE converges to the normal distribution, see Feigin [16]. Maximum likelihood estimation in non-recurrent case was studied in Dietz and Kutoyants [15]. Local asymptotic mixed normality for discretely observed non-recurrent Ornstein-Uhlenbeck processes was studied in Shimizu [30].

$$\theta_T - \theta := \frac{\int_0^T X_t dW_t}{\int_0^T X_t^2 dt} = \frac{Z_T}{I_T} \tag{1.4}$$

where

$$Z_T := \int_0^T X_t dW_t \quad \text{and} \quad I_T := \int_0^T X_t^2 dt. \tag{1.5}$$

Hence

$$\frac{e^{\theta T}}{2\theta}(\theta_T - \theta) = \frac{e^{-\theta T} 2\theta Z_T}{e^{-2\theta T} 4\theta^2 I_T} = \frac{\left(e^{-2\theta T} 4\theta^2\right)^{1/2} Z_T}{e^{-2\theta T} 4\theta^2 I_T} \tag{1.6}$$

In (1.6), the numerator of the normalized MLE is a normalized martingale which converges to the standard normal variable and the denominator is its corresponding increasing process which converges to a chi-square random variable as $T \to \infty$ which is independent of the numerator. Hence the ratio converges to a Cauchy distribution with parameters $(0, 1)$.

Let us introduce two Wiener integrals:

$$\xi_t := \int_0^t e^{-\theta s} dW_s, \quad \text{and } \eta_t := \int_0^t e^{\theta s} dW_s, \quad t \geq 0. \tag{1.7}$$

and the respective limits $\xi = \lim_{T\to\infty} \int_0^T e^{-\theta s} dW_s := \int_0^\infty e^{-\theta s} dW_s$ which has $\mathcal{N}(0, \frac{1}{2\theta})$ distribution and $\eta := \lim_{T\to\infty} e^{-\theta T}\eta_T = \lim_{T\to\infty} e^{-\theta T} \int_0^T e^{\theta s} dW_s = \int_0^\infty e^{-\theta s} dW_s$ which has $\mathcal{N}(0, \frac{1}{2\theta})$ distribution.

With these notations

$$Z_T := \int_0^T X_t dW_t = \int_0^T e^{\theta t}\xi_t dW_t \quad \text{and} \quad I_T := \int_0^T X_t^2 dt = \int_0^T e^{2\theta t}\xi_t^2 dt,$$

$$\theta_T - \theta = \frac{\int_0^T e^{\theta t}\xi_t dW_t}{\int_0^T e^{2\theta t}\xi_t^2 dt}, \tag{1.8}$$

$$\frac{e^{\theta T}}{2\theta}(\theta_T - \theta) = \frac{\left(e^{-2\theta T}4\theta^2\right)^{1/2}\int_0^T e^{\theta t}\xi_t dW_t}{e^{-2\theta T}4\theta^2\int_0^T e^{2\theta t}\xi_t^2 dt} = \frac{\left(e^{-2\theta T}4\theta^2\right)^{1/2}\int_0^T e^{\theta t}(\int_0^t e^{-\theta s} dW_s) dW_t}{e^{-2\theta T}4\theta^2\int_0^T e^{2\theta t}(\int_0^t e^{-\theta s} dW_s)^2 dt}$$

$$= \frac{\xi_T\xi}{2\theta e^{-2\theta T}\int_0^T e^{2\theta t}\xi_t^2 dt} \times \frac{e^{-\theta T}\int_0^T e^{\theta t} dW_t}{\xi} = \frac{\xi_T\xi}{2\theta e^{-2\theta T} I_T} \times \frac{e^{-\theta T}\eta_T}{\xi} =: A_T^\theta \times B_T^\theta. \tag{1.9}$$

We have

$$A_T^\theta \to 1 \text{ almost surely as } T\to\infty, \tag{1.10}$$

$$B_T^\theta \xrightarrow{\mathcal{D}} \frac{N}{\sqrt{2\theta}\xi} \text{ as } T\to\infty \tag{1.11}$$

where $\sqrt{2\theta}\xi = N_1$, and N_1 and N are independent standard normal random variables. Since

$$\frac{N}{\sqrt{2\theta}\xi} \stackrel{\mathcal{D}}{=} \mathcal{C}(1) \text{ as } T\to\infty \tag{1.12}$$

where $\mathcal{C}(1)$ is the standard Cauchy distribution, by Slutsky's theorem, we have

$$A_T^\theta \times B_T^\theta \xrightarrow{\mathcal{D}} \mathcal{C}(1) \text{ as } T\to\infty. \tag{1.13}$$

Note that

$$\xi_T \xrightarrow{\mathcal{D}} \xi \text{ as } T\to\infty. \tag{1.14}$$

Using Borel-Cantelli lemma and stochastic Fubini theorem, it can be shown that $\xi_T \to \xi$ almost surely and in $L_2(\Omega)$ as $T\to\infty$. By integration by parts we have

$$e^{-2\theta T} I_T = e^{-2\theta T}\int_0^T X_s^2 ds = e^{-2\theta T}\int_0^T e^{2\theta s}\xi_s^2 ds$$

$$= \frac{\xi_T^2}{2\theta} - \frac{e^{-2\theta T}}{\theta}\int_0^T e^{2\theta s}\xi_s d\xi_s - \frac{Te^{-2\theta T}}{\theta} = \frac{\xi_T^2}{2\theta} - \frac{e^{-2\theta T}}{\theta}\int_0^T e^{\theta s}\xi_s dW_s - \frac{Te^{-2\theta T}}{\theta}. \tag{1.15}$$

This equality together with

$$E\left(\int_0^T e^{\theta s}\xi_s dW_s\right)^2 = \int_0^T e^{2\theta s} E(\xi_s^2) ds = \frac{1}{2\theta}\int_0^T e^{2\theta s}(1 - e^{-2\theta s}) ds = \frac{e^{2\theta T} - 1 - 2\theta T}{4\theta^2} \tag{1.16}$$

by the CLT for stochastic integrals provides

$$e^{-2\theta T}\int_0^T X_s^2 ds \xrightarrow{\mathcal{D}} \frac{\xi^2}{2\theta} \text{ as } T\to\infty \tag{1.17}$$

$$\text{i.e., } e^{-2\theta T} I_T \xrightarrow{\mathcal{D}} \frac{\xi^2}{2\theta} \quad \text{as } T \to \infty.$$

It can be shown that

$$e^{-2\theta T} I_T \to \frac{\xi^2}{2\theta} \text{ almost surely} \quad \text{as } T \to \infty. \tag{1.18}$$

By Itô formula, we have

$$Z_T = \int_0^T X_s dW_s = \int_0^T e^{\theta s} \xi_s dW_s = \int_0^T \xi_s d\eta_s = \xi_T \eta_T - T - \int_0^T \eta_s d\xi_s$$

$$= \xi_T \eta_T - T - \int_0^T \eta_s e^{-\theta s} dW_s. \tag{1.19}$$

Hence

$$e^{-\theta T} Z_T = e^{-\theta T} \xi_T \eta_T - e^{-\theta T} T - e^{-\theta T} \int_0^T \eta_s e^{-\theta s} dW_s. \tag{1.20}$$

Direct calculation gives

$$e^{-\theta T} \eta_T \xrightarrow{\mathcal{D}} \eta \sim \mathcal{N}\left(0, \frac{1}{2\theta}\right) \quad \text{as } T \to \infty \tag{1.21}$$

and

$$E(\xi\eta) = \lim_{T\to\infty} E(\xi_T \eta_T e^{-\theta T}) = \lim_{T\to\infty} T e^{-\theta T} = 0. \tag{1.22}$$

Hence

$$e^{-\theta T} Z_T \xrightarrow{\mathcal{D}} \xi\eta \text{ as } T \to \infty. \tag{1.23}$$

Hence the limit distribution of the pair $(\xi_T, e^{-\theta T}\eta_T)$ is a Gaussian distribution of two independent variables. Thus

$$e^{\theta T} \frac{\xi_T}{\eta_T} \xrightarrow{\mathcal{D}} \zeta \text{ as } T \to \infty \tag{1.24}$$

where ζ is the standard Cauchy variable with probability density function

$$f(x) = \frac{1}{\pi(1+x^2)}, \quad x \in \mathbb{R}. \tag{1.25}$$

and cdf

$$\mathcal{C}(x) = \frac{1}{2} + \frac{1}{\pi} \arctan x, \ \ x \in \mathbb{R} \tag{1.26}$$

and characteristic function

$$\int_{-\infty}^{\infty} e^{i\lambda x} d\mathcal{C}(x) = e^{-|\lambda|}. \tag{1.27}$$

Hence

$$\frac{e^{\theta T}}{2\theta} (\theta_T - \theta) \xrightarrow{\mathcal{D}} \zeta. \tag{1.28}$$

We estimate the rate of convergence in this phenomenon. We need the following lemma in the sequel.

Lemma 1.1 (Esseen's Smoothing Lemma)

Let F be a non-decreasing function and H be a differentiable function of bounded variation on the real line with $F(\pm\infty) = G(\pm\infty)$. Denote the corresponding Fourier-Stieltjes transforms by $\widehat{F}$ and $\widehat{G}$, respectively. Then for all $\Lambda > 0$,

$$\sup_{x\in\mathbb{R}} |F(x) - G(x)| \leq \frac{1}{\pi} \int_{-\Lambda}^{\Lambda} \frac{\left|\widehat{F}(\lambda) - \widehat{G}(\lambda)\right|}{|\lambda|} d\lambda + \frac{24}{\pi\Lambda} \sup_{x\in\mathbb{R}} |G'(x)|.$$

Proof: See Petrov [27] or Feller [18]. □

Let $\Phi(\cdot)$ denote the standard normal distribution function and $\mathcal{C}(\cdot)$ denotes the standard Cauchy distribution function. Throughout the paper C denotes a generic constant (perhaps depending on θ, but not on anything else).

We need the following well known inequality.

Lemma 1.2

$$\frac{1}{\sqrt{2\pi}} \exp(\frac{-x^2}{2})(\frac{1}{x} - \frac{1}{x^3}) \leq 1 - \Phi(x) \leq \frac{1}{\sqrt{2\pi}x} \exp(\frac{-x^2}{2})$$

for $x > 0$. As $x \to \infty$,

$$1 - \Phi(x) \sim \frac{1}{\sqrt{2\pi}x} \exp(\frac{-x^2}{2}).$$

Proof: See Feller ([17], p.166). □

2. Main Results

We start with the Dambis-Dubins-Schwarz (DDS) theorem, see Protter [29]. Since Z_T is a continuous time martingale, due to time change (Skorohod embedding), $Z_T = B_{I_T}$ where B is a Brownian motion independent of W, we have

$$E(\exp(iuZ_T)) = E\exp(-\frac{u^2}{2}I_T), \tag{2.1}$$

$$E(\exp(iue^{-\theta T}\sqrt{2\theta}Z_T)) = E\exp(-\frac{u^2}{2}e^{-2\theta T}2\theta I_T), \tag{2.2}$$

$$\theta_T - \theta = \frac{B_{I_T}}{I_T}, \tag{2.3}$$

$$\frac{e^{\theta T}}{2\theta}(\theta_T - \theta) = \frac{e^{-\theta T}B_{I_T}}{e^{-2\theta T}2\theta I_T} = \frac{\left(e^{-2\theta T}4\theta^2\right)^{1/2}B_{I_T}}{e^{-2\theta T}4\theta^2 I_T} =: Y_T \tag{2.4}$$

where

$$Y_T = \frac{e^{-\theta T}B_{I_T}}{e^{-2\theta T}2\theta I_T}. \tag{2.5}$$

Our main claim in the paper is to show that

$$|E(e^{iuY_T}) - e^{-|u|}| \leq C|u|e^{-|u|/2}e^{-\theta T}. \tag{2.6}$$

This is done through several lemmas. Once it is shown, let

$$F(x) = P(Y_T \leq x), \tag{2.7}$$

$$\mathcal{C}(x) = \frac{1}{2} + \frac{1}{\pi}\arctan x, \quad x \in \mathbb{R}, \tag{2.8}$$

Take $\Lambda = e^{\theta T}$. Then

$$\sup_{x\in\mathbb{R}}|F(x) - \mathcal{C}(x)| \leq \frac{1}{\pi}J + \frac{24}{\pi e^{\theta T}}\sup \mathcal{C}'(x) \tag{2.9}$$

where

$$J := \frac{1}{\pi}\int_{|\lambda|\leq e^{\theta T}} \frac{\left|\widehat{F}(\lambda) - \widehat{\mathcal{C}}(\lambda)\right|}{|\lambda|}d\lambda. \tag{2.10}$$

Clearly

$$\sup \mathcal{C}'(x) < \infty$$

and

$$J \leq \frac{C}{e^{\theta T}}\int_{|\lambda|\leq e^{\theta T}} e^{-|\lambda|/2}d\lambda \leq \frac{C}{e^{\theta T}}\int_{-\infty}^{\infty} e^{-|\lambda|/2}d\lambda = O(e^{-\theta T}). \tag{2.11}$$

which would ultimately give

$$\sup_{x\in\mathbb{R}}|F(x) - \mathcal{C}(x)| = O(e^{-\theta T}). \tag{2.12}$$

First we start with Kolmogorov distance for Wiener chaos and its relative:

Lemma 2.1 We have the following rate of convergence for the double-stochastic integral or the second Wiener chaos $\int_0^T e^{\theta t}\left(\int_0^t e^{-\theta s}dW_s\right)dW_t$:

$$\text{(a)} \sup_{x\in\mathbb{R}}\left|P\left\{e^{\theta T}\left(2\theta e^{-2\theta T}\int_0^T e^{\theta t}\left(\int_0^t e^{-\theta s}dW_s\right)dW_t\right)\le x\right\}-\Phi(x)\right|\le Ce^{-\theta T}.$$

$$\text{(b)} \sup_{x\in\mathbb{R}}\left|P\left\{e^{\theta T}\left(2\theta e^{-2\theta T}\int_0^T e^{2\theta t}\xi_t^2dt-\xi^2\right)\le x\right\}-\Phi(x)\right|\le Ce^{-\theta T}.$$

Proof. Observe that

$$Z_T=\int_0^T X_t dW_t=\int_0^T e^{\theta t}\left(\int_0^t e^{-\theta s}dW_s\right)dW_t$$

The integral

$$\int_0^T e^{\theta t}\left(\int_0^t e^{-\theta s}dW_s\right)dW_t$$

is second Wiener chaos. One can use the Stein-Malliavin method (see Nourdin and Peccati ([25], [26]) and estimate the Kolmogorov distance for Z_T. However, part (a) follows as a consequence of Lemma 2.4(c) below along with Lemma 1.1 above. Part (b) follows as a consequence of Lemma 2.2 below along with Lemma 1.1 above. □

Note that $\xi^2\sim\chi_1^2$. The next theorem gives an exponential estimate on the rate of convergence to the chi-square distribution for energy I_T of the O-U process.

Theorem 2.1

$$\sup_{x\in\mathbb{R}}\left|P\left\{e^{-2\theta T}2\theta I_T\le x\right\}-P\left\{\xi^2\le x\right\}\right|=O(e^{-\theta T}).$$

The above theorem is a consequence of the following lemma and the *Esseen's smoothing lemma 1.1*

Lemma 2.2 For $|u|\le e^{\theta T}\varepsilon$, ε sufficiently small, we have

$$\left|E\exp\left(iue^{-2\theta T}2\theta I_T\right)-\frac{1}{(1-2iu)^{\frac{1}{2}}}\right|\le C(|u|+|u|^3)e^{-\theta T}.$$

Proof. From Liptser and Shiryayev [23], we have

$$E\exp\left(iue^{-2\theta T}2\theta I_T\right)=\exp\left(\frac{\theta T}{2}\right)\left[\frac{2\gamma}{(\gamma-\theta)e^{-\gamma T}+(\gamma+\theta)e^{\gamma T}}\right]^{1/2} \tag{2.12}$$

where

$$\gamma:=\left(\theta^2-2iue^{-2\theta T}2\theta\right)^{1/2}. \tag{2.13}$$

The lemma is an easy consequence of this result. □

Lemma 2.3 For every $\delta>0$,

$$P\left\{\left|e^{-2\theta T}2\theta I_T-\xi^2\right|\ge\delta\right\}\le Ce^{-2\theta T}\delta^{-2}.$$

Proof : It is clear that

$$X_T = \int_0^T e^{-\theta(T-s)} dW_s. \tag{2.14}$$

Further, Itô formula (see Friedman [19]), we have

$$\int_0^T e^{\theta(T-s)} dW_s = W_T - \theta \int_0^T e^{\theta(T-s)} W_s ds,$$

$$\xi_t = \int_0^t e^{-\theta s} dW_s = e^{-\theta t} W_t - \theta \int_0^t e^{-\theta s} dW_s, \ \eta_t = \int_0^t e^{\theta s} dW_s = e^{\theta t} W_t + \theta \int_0^t e^{\theta s} dW_s.$$

Note that

$$E(X_T^2) = \frac{1 - e^{-2\theta T}}{2\theta}, \quad E(X_T^4) = \frac{3(1 - e^{-2\theta T})^2}{4\theta} \text{ and } E(I_T) = \frac{2\theta T - 1 + e^{-2\theta T}}{4\theta^2}. \tag{2.15}$$

By Itô formula, we have

$$I_T = \frac{X_T^2}{2\theta} - \frac{T}{2\theta} - \frac{Z_T}{\theta}. \tag{2.16}$$

By Chebyshev inequality, we have

$$\begin{aligned} P\left\{\left|e^{-2\theta T} 2\theta I_T - \xi^2\right| \geq \delta\right\} &\leq \frac{1}{\delta^2} E\left|e^{-2\theta T} 2\theta I_T - \xi^2\right|^2 \\ = \frac{1}{\delta^2} E\left|e^{-2\theta T} 2\theta \int_0^T e^{2\theta t} \xi_t^2 dt - \xi^2\right|^2 &= \frac{1}{\delta^2} E\left|e^{-2\theta T} 2\theta \int_0^T e^{2\theta t} \xi_t^2 dt - \xi_T^2 + \xi_T^2 - \xi^2\right|^2 \\ \leq \frac{2}{\delta^2}\left[E\left|e^{-2\theta T} 2\theta \int_0^T e^{2\theta t} \xi_t^2 dt - \xi_T^2\right|^2 + E\left|\xi_T^2 - \xi^2\right|^2\right] \\ \leq \frac{2}{\delta^2}\left[E\left|e^{-2\theta T} 2\theta \int_0^T e^{2\theta t} \xi_t^2 dt - \xi_T^2\right|^2 + E|\xi_T - \xi|^2 E|\xi_T + \xi|^2\right] \\ \leq \frac{2}{\delta^2}\left[E\left|e^{-2\theta T} 2\theta \int_0^T e^{2\theta t} \xi_t^2 dt - \xi_T^2\right|^2 + \frac{e^{-2\theta T}}{\sqrt{2\theta}}\right] &\leq C e^{-2\theta T} \delta^{-2} \end{aligned} \tag{2.17}$$

since $E|\xi_T + \xi|^2 \leq 2E|\xi_T|^2 + 2E|\xi|^2 < \infty$.

Since

$$E(\xi_T - \xi)^2 = \int_T^\infty \int_T^\infty e^{-\theta r} e^{-\theta s} |r - s|^{-1} dr ds = \frac{e^{-2\theta T}}{\sqrt{2\theta}} \tag{2.18}$$

hence

$$E(\xi_T - \xi)^2 = \frac{e^{-2\theta T}}{\sqrt{2\theta}} \tag{2.19}$$

gives the L_2 convergence rate. Recall that

$$I_T = \int_0^T e^{2\theta t} \xi_t^2 dt, \tag{2.20}$$

$$E(\xi_T - \xi)^2 \to 0 \text{ as } T \to \infty, \tag{2.21}$$

$$E(\xi_t - \xi_s)^2 \leq C(t - s). \tag{2.22}$$

We have

$$E\left(2\theta e^{-2\theta T} I_T - \xi^2\right)^2 = E\left(2\theta e^{-2\theta T} \int_0^T e^{2\theta t} \xi_t^2 dt - \xi^2\right)^2 = E\left(\frac{2\theta}{e^{2\theta T}} \int_0^T e^{2\theta t} \xi_t^2 dt - \xi^2\right)^2. \tag{2.23}$$

Further, by Toeplitz's lemma

$$\lim_{T\to\infty} 2\theta e^{-2\theta T}\int_0^T e^{2\theta t}\xi_t^2 dt = \lim_{T\to\infty}\xi_T^2 = \xi^2 \text{ almost surely.} \tag{2.24}$$

$E(\xi^2) < \infty$ which implies that $P(\xi = 0) = 0$. We have

$$\lim_{T\to\infty}\left[2\theta e^{-2\theta T}\int_0^T e^{2\theta t}\xi_t^2 dt - \xi_T^2\right] = 0 \quad \text{almost surely.} \tag{2.25}$$

Because of the continuity of ξ_t, for every $t \geq 0$,

$$\int_0^T e^{2\theta t}\xi_t^2 dt \geq \int_{\frac{T}{2}}^T e^{2\theta t}\xi_t^2 dt \geq \frac{T}{2}e^{\theta T}\left(\inf_{\frac{T}{2}<t<T}\xi_t^2\right) \text{ almost surely.} \tag{2.26}$$

Furthermore the continuity of ξ_t, gives

$$\lim_{T\to\infty}\left(\inf_{\frac{T}{2}<t<T}\xi_t^2\right) = \xi^2 \text{ almost surely.} \tag{2.27}$$

$$\lim_{T\to\infty}\int_0^T e^{2\theta t}\xi_t^2 dt = \infty \text{ almost surely.} \tag{2.28}$$

By L'Hopital rule,

$$\lim_{T\to\infty}\frac{\int_0^T e^{2\theta t}\xi_t^2 dt}{e^{2\theta T}} = \lim_{T\to\infty}\frac{\xi_T^2}{2\theta} = \frac{\xi^2}{2\theta} \text{ almost surely.} \tag{2.29}$$

$$\theta_T - \theta = \frac{\int_0^T e^{\theta t}\xi_t dW_t}{\int_0^T e^{2\theta t}\xi_t^2 dt} = \frac{\xi_T^2}{2e^{-2\theta T}\int_0^T e^{2\theta t}\xi_t^2 dt} - \theta. \tag{2.30}$$

$$\theta_T - \theta \to 0 \text{ almost surely.} \tag{2.31}$$

$$\hat{\theta}_T - \theta = \frac{\xi_T^2 - 2\theta e^{-2\theta T}\int_0^T e^{2\theta t}\xi_t^2 dt}{2e^{-2\theta T}\int_0^T e^{2\theta t}\xi_t^2 dt}. \tag{2.32}$$

It is easy to verify that

$$E\left[\xi_T^2 - 2\theta e^{-2\theta T}\int_0^T e^{2\theta t}\xi_t^2 dt\right]^2 \leq Ce^{-2\theta T}. \tag{2.33}$$

This completes the proof of the lemma. □

The following lemma (*Cameron-Martin Type Theorem*) gives the bound on the joint characteristic functions of the sufficient statistics defining the MLE:

Lemma 2.4 (a) Let $\phi_T(z_1, z_2) := E\exp(z_1 I_T + z_2 X_T^2), z_1, z_2 \in \mathbb{C}$. Then $\phi_T(z_1, z_2)$ exists for $|z_i| \leq \delta$, $1 = 1,2$ for some $\delta > 0$ and is given by

$$\phi_T(z_1, z_2) = \exp\left(\frac{\theta T}{2}\right)\left[\frac{2\gamma}{(\gamma - \theta + 2z_2)e^{-\gamma T} + (\gamma + \theta - 2z_2)e^{\gamma T}}\right]^{1/2}$$

where $\gamma = (\theta^2 - 2z_1)^{1/2}$ and we choose the principal branch of the square root.

(b) Let $H_{T,x} := \left(e^{-2\theta T}4\theta^2\right)^{1/2} Z_T - \left(e^{-2\theta T}4\theta^2 I_T - \xi^2\right) x$. Then for $|x| \le 2(\log e^{2\theta T})^{1/2}$ and for $|u| \le \epsilon e^{\theta T}$, where ϵ is sufficiently small

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

(c) For $|u| \le \varepsilon_1 e^{\theta T}$, where ε_1 is sufficiently small, we have as $T \to \infty$,

$$\left|E\exp\left\{iu\left(e^{-\theta T}2\theta\right)Z_T\right\} - \exp(-\frac{u^2}{2})\right| \le C\exp(-\frac{|u|}{2})(|u|+|u|^3)e^{-\theta T}.$$

Part (a) is from Bishwal [5].

We shall prove part (b) in details. Proof of part (c) is very similar to part (b) and will be omitted.

Proof : By Itô formula,

$$Z_T = \theta I_T + \frac{X_T^2}{2} - \frac{T}{2}.$$

Note that

$$\begin{aligned} E\exp(iuH_{T,x}) &= E\exp\left[-iu\left(e^{-2\theta T}4\theta^2\right)^{1/2} Z_T - iu\left(\left(e^{-2\theta T}4\theta^2\right)I_T - \xi^2\right)x\right] \\ &= E\exp\left[-iu\left(e^{-2\theta T}4\theta^2\right)^{1/2}\left\{\theta I_T + \frac{X_T^2}{2} - \frac{T}{2}\right\} - it\left(\left(e^{-2\theta T}4\theta^2\right)I_T - \xi^2\right)x\right] \\ &= E\exp(z_1 I_T + z_2 X_T^2 + z_3) = \exp(z_3)\phi_T(z_1,z_2) \end{aligned}$$

where

$$z_1 = -iu\theta\delta_{T,x}, \quad z_2 = -\frac{iu}{2}\left(e^{-2\theta T}4\theta^2\right)^{1/2}, \quad z_3 = \frac{iuT}{2}\delta_{T,x}, \quad \delta_{T,x} = \left(e^{-2\theta T}4\theta^2\right)^{1/2} + \frac{2x}{T}.$$

Note that (z_1, z_2) satisfies the conditions of (a) by choosing ϵ sufficiently small. Let $\alpha_{1,T}(u), \alpha_{2,T}(u)$,
$\alpha_{3,T}(u)$ and $\alpha_{4,T}(u)$ be functions which are of the orders $O(|u|e^{-\theta T/2}), O(|u|^2 e^{-\theta T/2})$, $O(|u|^3 e^{-3\theta T/2})$ and $O(|u|^3 e^{-\theta T/2})$ respectively. Note that for the given range of values of x and u, the conditions on z_i for part (a) of Lemma are satisfied. Note also that $z_2 = \alpha_{1,T}(u)$. Further, with

$$\beta_T(t) = 1 + iu\frac{\delta_{T,x}}{\theta} + \frac{u^2\delta_{T,x}^2}{2\theta^2},$$

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

Thus $\gamma - \theta = \alpha_{1,T}$, $\gamma + \theta = 2\theta + \alpha_{1,T}$. Hence the above expectation equals

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

where

$$\chi_T(u) := -\theta T\beta_T(u) + \alpha_{4,T}(u) - 2z_3 - \theta T = -2\theta T + \alpha_{1,T}(u) + t^2\alpha_{1,T}(u),$$

$$\psi_T(u) := \theta T\beta_T(u) + \alpha_{4,T}(u) - 2z_3 - \theta e^{\theta T} = \theta T\left[1 + iu\frac{\delta_{T,x}}{\theta} + \frac{u^2\delta_{T,x}^2}{2\theta^2}\right] + \alpha_{4,T}(u) - ite^{\theta T}\delta_{T,x} - \theta e^{\theta T}$$

$$= \frac{u^2 e^{\theta T}}{2\theta}\left[\left(\frac{4\theta^2}{e^{2\theta T}}\right)^{1/2} + \frac{2x}{e^{\theta T}}\right]^2 = u^2 + u^2\alpha_{1,T}(u).$$

Hence, for the given range of values of u, $\chi_T(u) - \psi_T(u) \le -\theta e^{\theta T}$. Hence the above expectation equals

$$\exp(-\frac{t^2}{2})(1+\alpha_{1,T})^{1/2}\left[\alpha_{1,T}\exp\{-2\theta e^{\theta T} + \alpha_{1,T} + u^2\alpha_{1,T}\} + (1+\alpha_{1,T}(u))\exp\{t^2\alpha_{1,T}(u)\}\right]^{-1/2}$$

$$= \exp(-\frac{u^2}{2})\left[1+\alpha_{1,T})(1+\alpha_{1,T}(1+\alpha_{1,T})\exp\{-\theta e^{\theta T} + \alpha_{1,T} + t^2\alpha_{1,T}\}\right]\exp(u^2\alpha_{1,T}(u)).$$

Lemma 2.4 (c) and Lemma 2.2 respectively give the Berry-Esseen rate for Z_T and I_T immediately by using the Esseen's lemma 1.1.

Corollary 2.1

$$\text{(a)} \quad \sup_{x\epsilon\mathbb{R}}\left|P\left\{\left(\frac{4\theta^2}{e^{2\theta T}}\right)^{1/2} Z_T \le x\right\} - \Phi(x)\right| \le Ce^{-\theta T}.$$

$$\text{(b)} \quad \sup_{x\in\mathbb{R}}\left|P\left\{\left(\frac{4\theta^2}{e^{2\theta T}}\right)^{1/2}\left(\theta I_T - \xi^2\frac{e^{\theta T}}{2}\right) \le x\right\} - \Phi(x)\right| \le Ce^{-\theta T}.$$

Remark Though this was basically shown in Lemma 2.1, here we obtain Kolmogorov distance for a martingale and Kolmogorov distance for its quadratic variation through Cameron-Matin type results which are generalization of Levy area formula. In Lemma 2.1, one could go directly to the Stein-Malliavin way through Wiener chaos expansion which does not depend on any martingale characteristics.

Before we prove the results on the Berry-Esseen bound on the Kolmogorov distance for the MLE with random norming we need the following large deviation result for the MLE. This can be obtained as a consequence of Lemma 3.1 of Bercu *et al.* [3] or Bercu and Richou [4] who use the Gartner-Ellis's theorem and the contraction principle. However we give a direct proof using Feller's approach.

Lemma 2.5

$$P\left\{\left(\frac{e^{2\theta T}}{4\theta^2}\right)^{1/2}|\theta_T - \theta| \ge 2(2\theta T)^{1/2}\right\} \le Ce^{-\theta T}.$$

Proof : Observe that

$$P\left\{\left(\frac{e^{2\theta T}}{4\theta^2}\right)^{1/2}|\theta_T - \theta| \ge 2(2\theta T)^{1/2}\right\} = P\left\{\left|\frac{\left(\frac{4\theta^2}{e^{2\theta T}}\right)^{1/2} Z_T}{\left(\frac{2\theta}{e^{2\theta T}}\right)I_T}\right| \ge 2(2\theta T)^{1/2}\right\}$$

$$\begin{aligned}
&\leq P\left\{\left|\left(\frac{4\theta^2}{e^{2\theta T}}\right)^{1/2} Z_T\right| \geq (\theta T)^{1/2}\right\} + P\left\{\left|\frac{2\theta}{e^{2\theta T}} I_T\right| \leq \frac{1}{2}\right\} \\
&\leq \left|P\left\{\left(\frac{4\theta^2}{e^{2\theta T}}\right)^{1/2} |Z_T| \geq (\theta T)^{1/2}\right\} - 2\Phi(-(2\theta T)^{1/2})\right| + 2\Phi(-(2\theta T)^{1/2}) + P\left\{\left|\frac{2\theta}{e^{2\theta T}} I_T - \xi^2\right| \geq \frac{1}{2}\right\} \\
&\leq \sup_{x\in\mathbb{R}}\left|P\left\{\left(\frac{4\theta^2}{e^{2\theta T}}\right)^{1/2} |Z_T| \geq x\right\} - 2\Phi(-x)\right| + 2\Phi(-(2\theta T)^{1/2}) + P\left\{\left|\left(\frac{4\theta^2}{e^{2\theta T}}\right) I_T - \xi^2\right| \geq \frac{1}{2}\right\} \\
&\leq \sup_{x\in\mathbb{R}}\left|P\left\{\left(\frac{4\theta^2}{e^{2\theta T}}\right)^{1/2} |Z_T| \geq x\right\} - 2\Phi(-x)\right| + 2\Phi(-(2\theta T)^{1/2}) + P\left\{\left|\left(\frac{4\theta^2}{e^{2\theta T}}\right) I_T - \xi^2\right| \geq \frac{1}{2}\right\} \\
&\leq Ce^{-\theta T} + C(e^{2\theta T} 2\theta T)^{-1/2} + C(e^{2\theta T})^{-1} \leq Ce^{-\theta T}.
\end{aligned}$$

The bounds for the first and the third terms come from Corollary 2.1 (a) and Lemma 2.3 respectively and that for the middle term comes from Feller ([17], p. 166). □

We are now in a position to obtain the Berry-Esseen bound of the order $O(e^{-\theta T})$ on the Kolmogorov distance for the MLE.

Theorem 2.2

$$\sup_{x\in\mathbb{R}}\left|P\left\{\left(\frac{e^{\theta T}}{2\theta}\right)(\theta_T - \theta) \leq x\right\} - \mathcal{C}(x)\right| = O(e^{-\theta T}).$$

Proof : We shall consider two possibilities: (i) $|x| > 2(\theta T)^{1/2}$ and (ii) $|x| \leq 2(\theta T)^{1/2}$.

(i) We shall give a proof for the case $x > 2(\theta T)^{1/2}$. The proof for the case $x < -2(\theta T)^{1/2}$ runs similarly. Note that

$$\left|P\left\{\left(\frac{e^{\theta T}}{2\theta}\right)(\theta_T - \theta) \leq x\right\} - \mathcal{C}(x)\right| \leq P\left\{\left(\frac{e^{\theta T}}{2\theta}\right)(\theta_T - \theta) \geq x\right\} + \mathcal{C}(-x)$$

But $\mathcal{C}(-x) \leq \mathcal{C}(-2(\theta T)^{1/2}) \leq Ce^{-2\theta T}$. Moreover by Lemma 2.5, we have

$$P\left\{\left(\frac{e^{\theta T}}{2\theta}\right)(\theta_T - \theta) \geq 2(\theta T)^{1/2}\right\} \leq Ce^{-\theta T/2}.$$

Hence

$$\left|P\left\{\left(\frac{e^{\theta T}}{2\theta}\right)^{1/2}(\theta_T - \theta) \leq x\right\} - \mathcal{C}(x)\right| \leq Ce^{-\theta T/2}.$$

(ii)

$$\text{Let } A_T := \left\{\left(\frac{e^{\theta T}}{2\theta}\right)|\theta_T - \theta| \leq 2(\theta T)^{1/2}\right\} \quad \text{and} \quad B_T := \left\{\frac{I_T}{e^{\theta T}} > c_0\right\}$$

where $0 < c_0 < \frac{1}{2\theta}$. By Lemma 2.5, we have

$$P(A_T^c) \leq Ce^{-\theta T}. \tag{2.34}$$

By Lemma 2.3, we have

$$P(B_T^c) = P\left\{\frac{2\theta}{e^{\theta T}} I_T - \xi^2 < 2\theta c_0 - \xi^2\right\} < P\left\{\left|\frac{2\theta}{e^{\theta T}} I_T - \xi^2\right| > \xi^2 - 2\theta c_0\right\} \leq Ce^{-\theta T}. \tag{2.35}$$

Let b_0 be some positive number. For $\omega \in A_T \cap B_T$ and for all $T > T_0$ with $4b_0(2\theta T_0)^{1/2}(\frac{2\theta}{e^{\theta T_0}})^{1/2} \leq c_0$, we have

$$
\begin{aligned}
&\left(\frac{e^{\theta T}}{2\theta}\right)(\theta_T - \theta) \leq x \\
\Rightarrow\quad & I_T + b_0 e^{\theta T}(\theta_T - \theta) < I_T + \left(\frac{e^{\theta T}}{2\theta}\right) 2b_0\theta x \\
\Rightarrow\quad & \left(\frac{e^{\theta T}}{2\theta}\right)(\theta_T - \theta)[I_T + b_0 e^{\theta T}(\theta_T - \theta)] < x[I_T + \left(\frac{e^{\theta T}}{2\theta}\right) 2b_0\theta x] \\
\Rightarrow\quad & (\theta_T - \theta)I_T + b_0 T(\theta_T - \theta)^2 < \left(\frac{2\theta}{e^{\theta T}}\right) I_T x + 2b_0\theta x^2 \\
\Rightarrow\quad & Z_T + (\theta_T - \theta)I_T + b_0 e^{\theta T}(\theta_T - \theta)^2 < Z_T + \left(\frac{2\theta}{e^{\theta T}}\right) I_T x + 2b_0\theta x^2 \\
\Rightarrow\quad & 0 < Z_T + \left(\frac{2\theta}{e^{\theta T}}\right) I_T x + 2b_0\theta x^2
\end{aligned}
$$

since

$$
\begin{aligned}
& I_T + b_0 e^{\theta T}(\theta_T - \theta) > e^{\theta T} c_0 + b_0 e^{\theta T}(\theta_T - \theta) \\
> \quad & 4b_0(\theta T)^{1/2}\left(\frac{2\theta}{e^{\theta T}}\right) - 2b_0(\theta T)^{1/2}(\frac{2\theta}{e^{\theta T}}) = 2b_0(\theta T)^{1/2}\left(\frac{2\theta}{e^{\theta T}}\right) > 0.
\end{aligned}
$$

Hence, for $\omega \in A_T \cap B_T$,

$$
\left(\frac{e^{\theta T}}{2\theta}\right)(\theta_T - \theta) \leq x \Rightarrow Z_T + \left(\frac{2\theta}{e^{\theta T}}\right) I_T x + 2b_0\theta x^2 > 0.
$$

On the other hand, for $\omega \in A_T \cap B_T$ and for all $T > T_0$ with $4b_0(2\theta T_0)^{1/2}(\frac{2\theta}{e^{\theta T_0}}) \leq c_0$, we have

$$
\begin{aligned}
&\left(\frac{e^{\theta T}}{2\theta}\right)(\theta_T - \theta) > x \\
\Rightarrow\quad & I_T - b_0 e^{\theta T}(\theta_T - \theta) < I_T - \left(\frac{e^{\theta T}}{2\theta}\right) 2b_0\theta x \\
\Rightarrow\quad & \left(\frac{e^{\theta T}}{2\theta}\right)(\theta_T - \theta)[I_T - b_0 e^{\theta T}(\theta_T - \theta)] > x[I_T - \left(\frac{e^{\theta T}}{2\theta}\right) 2b_0\theta x] \\
\Rightarrow\quad & (\theta_T - \theta)I_T - b_0 e^{\theta T}(\theta_T - \theta)^2 > \left(\frac{2\theta}{e^{\theta T}}\right) I_T x - 2b_0\theta x^2 \\
\Rightarrow\quad & Z_T + (\theta_T - \theta)I_T - b_0 e^{\theta T}(\theta_T - \theta)^2 > Z_T + \left(\frac{2\theta}{e^{\theta T}}\right) I_T x - 2b_0\theta x^2 \\
\Rightarrow\quad & 0 > Z_T + \left(\frac{2\theta}{e^{\theta T}}\right) I_T x - 2b_0\theta x^2
\end{aligned}
$$

since

$$
\begin{aligned}
& I_T - b_0 e^{\theta T}(\theta_T - \theta) > e^{\theta T} c_0 - b_0 e^{\theta T}(\theta_T - \theta) \\
> \quad & 4b_0(\theta T)^{1/2}\left(\frac{2\theta}{e^{\theta T}}\right) - 2b_0(\theta T)^{1/2}\left(\frac{2\theta}{e^{\theta T}}\right) = 2b_0(\theta T)^{1/2}\left(\frac{2\theta}{e^{\theta T}}\right) > 0.
\end{aligned}
$$

Hence, for $\omega \in A_T \cap B_T$,

$$0 < Z_T + \left(\frac{2\theta}{e^{\theta T}}\right) I_T x - 2b_0\theta x^2 \Rightarrow \left(\frac{e^{\theta T}}{2\theta}\right)(\theta_T - \theta) \le x.$$

We use the squeezing method developed in Pfanzagl [28] for the i.i.d. case instead of the splitting method of Michel and Pfanzagl [28]. Let us introduce the piecewise quadratic random functions involving the martingale and quadratic variation part of $\theta_T - \theta$:

$$g^{\pm}(x) := Z_T + (\frac{2\theta}{e^{\theta T}}) I_T x \pm 2b_0\theta x^2.$$

Let us introduce the events

$$D^{\pm}_{T,x} := \left\{ Z_T + (\frac{2\theta}{e^{\theta T}}) I_T x \pm 2b_0\theta x^2 > 0 \right\}.$$

Thus we have

$$D^{-}_{T,x} \cap A_T \cap B_T \subseteq A_T \cap B_T \cap \left\{ \left(\frac{e^{\theta T}}{2\theta}\right)(\theta_T - \theta) \le x \right\} \subseteq D^{+}_{T,x} \cap A_T \cap B_T. \tag{2.36}$$

This gives

$$P(D^{-}_{T,x} \cap A_T \cap B_T) \le P\left(A_T \cap B_T \cap \left\{ \left(\frac{e^{\theta T}}{2\theta}\right)(\theta_T - \theta) \le x \right\}\right) \le P(D^{+}_{T,x} \cap A_T \cap B_T)$$

so that

$$\begin{aligned} &\left| P\left(A_T \cap B_T \cap \left\{ \left(\frac{e^{\theta T}}{2\theta}\right)(\theta_T - \theta) \le x \right\}\right) - \mathcal{C}(x) \right| \\ &\le \max\left\{ |P(D^{-}_{T,x} \cap A_T \cap B_T) - \mathcal{C}(x)|,\ |P(D^{+}_{T,x} \cap A_T \cap B_T) - \mathcal{C}(x)| \right\} \\ &\le \max\left\{ |P(D^{-}_{T,x}) - \mathcal{C}(x)|,\ |P(D^{+}_{T,x}) - \mathcal{C}(x)| \right\} + P(A_T \cap B_T)^c. \end{aligned}$$

From (2.34) and (2.35),

$$P(A_T \cap B_T)^c \le Ce^{-\theta T}$$

for all $T > T_0$ and $|x| \le 2(\theta T)^{1/2}$. If it is shown that

$$\left| P\left\{ D^{\pm}_{T,x} \right\} - \mathcal{C}(x) \right| \le Ce^{-\theta T} \tag{2.37}$$

for all $T > T_0$ and $|x| \le 2(\theta T)^{1/2}$, then the theorem would follow from (2.34) – (2.37).

We shall prove (2.37) for $D^{+}_{T,x}$. The proof for $D^{-}_{T,x}$ is analogous.

Note that

$$\begin{aligned} \left| P\left\{ D^{+}_{T,x} \right\} - \mathcal{C}(x) \right| &= \left| P\left\{ -(\frac{2\theta}{e^{\theta T}}) Z_T - \left(\frac{2\theta}{e^{\theta T}} I_T - \xi^2\right) x < x + 2\left(\frac{2\theta}{e^{\theta T}}\right) b_0\theta x^2 \right\} - \mathcal{C}(x) \right| \\ \le\ & \sup_{y \in \mathbb{R}} \left| P\left\{ -\left(\frac{2\theta}{e^{\theta T}}\right) Z_T - \left(\frac{2\theta}{e^{\theta T}} I_T - \xi^2\right) x \le y \right\} - \mathcal{C}(y) \right| + \left| \mathcal{C}\left(x + \left(\frac{2\theta}{e^{\theta T}}\right) b_0\theta x^2\right) - \mathcal{C}(x) \right| \\ =:\ & \Delta_1 + \Delta_2. \end{aligned} \tag{2.38}$$

Lemma 2.4 (b) and Esseen's Smoothing Lemma 1.1 immediately yield

$$\Delta_1 \le Ce^{-\theta T}. \tag{2.39}$$

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

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

where

$$|\overline{x} - x| \leq 2\left(\frac{2\theta}{e^{\theta T}}\right)b_0\theta x^2.$$

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

$$\Delta_2 \quad \leq \quad 2\left(\frac{2\theta}{e^{\theta T}}\right)b_0\theta x^2(2\pi)^{-1/2}x^2\exp(-x^2/8) \leq Ce^{-\theta T}. \tag{2.40}$$

From (2.38) - (2.40), we obtain

$$\left|P\left\{D^+_{T,x}\right\} - \mathcal{C}(x)\right| \leq Ce^{-\theta T}.$$

This completes the proof of the theorem. □

Concluding Remarks

(1) The bound in Theorem 2.2 is uniform over compact subsets of the parameter space Θ.

(2) The bound in Theorem 2.2 is optimal and cannot be improved further.

(3) Note that in the critical case, i.e., when $\theta = 0$, the MlE has a distribution concentrated on a half line, precisely the distribution of the ratio of a noncentral chisquare to the to the sum of chisquares. Note that the behaviour of the O-U process depends on both the initial condition $X_0 = X^0$ and the parameter space. Classically it has been assumed that X^0 is either has a normal distribution or a nonzero constant and $\theta < 0$ which makes the process stationary with Gaussian invariant distribution. If $X^0 = 0$ is with $\theta < 0$, then the process is asymptotically stationary and ergodic. In above two cases the model satisfies the LAN (local asymptotic normality) property. With X^0 a nonzero constant and $\theta > 0$ the process is transient and satisfies the LAMN (local asymptotic mixed normality) property. With $\theta = 0$, the process is nonstationary and satisfies the LABF (local asymptotic Brownian functional) property. For all $\theta \in \mathbb{R}$, the model satisfies the LABF property, see Bishwal [9] for the definitions of these LAN, LAMN and LABF properties. Bishwal [9] has shown that sequential sampling based on a stopping rule unifies the three properties and makes them LAN.

(4) It remains to study the Kolmogorov distnace for Bayes estimator from both continuous and siscrete observations and approximate maximum likelihood estimator from discrete observations in the nonergodic case.

(5) Extension to multidimensional process and to multiparameter case remains to be investigated.

(6) It remains to investigate the nonuniform rates of convergence to Cauchy distribution which are more useful.

References

[1] B. Bercu, On large deviations in the Gaussian autoregressive process: stable, unstable and explosive cases, Bernoulli. 7 (2001) 299-316.

[2] B. Bercu, A. Roualt, Sharp large deviation for the Ornstein-Uhlenbeck process, Theory Prob. Appl. 46 (2002) 1-19.

[3] B. Bercu, L. Coutin, N. Savy, Sharp large deviations for the non-stationary Ornstein-Uhlenbeck process, Stoch. Process Appl. 122 (2012) 3393-3424.

[4] B. Bercu, A. Richou, Large deviations for the Ornstein-Uhlenbeck process without tears, Stat. Prob. Lett. 123 (2017) 45-55.

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

[6] 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 (2000b) 51-70.

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

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

[9] J.P.N. Bishwal, Sequential maximum likelihood estimation in nonlinear non-Markov diffusion type processes, Dyn. Syst. Appl. 27 (2018) 107-124.

[10] J.P.N. Bishwal, Berry-Esseen bounds of approximate Bayes estimators for the discretely observed Ornstein-Uhlenbeck process, Asian J. Stat. Sci. 1 (2021) 83-122.

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

[12] J.P.N. Bishwal, A. Bose, Speed of convergence of the maximum likelihood estimator in the Ornstein-Uhlenbeck process, Calcutta 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] H.M. Dietz, Y.A. Kutoyants, Parameter estimation in some non-recurrent solutions of SDE, Stat. Decis. 21 (2003) 29-45.

[16] P.D. Feigin, Maximum likelihood estimation for continuous time stochastic processes, Adv. Appl. Prob. 8 (1976) 712-736.

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

[18] W. Feller, An introduction to probability theory and its applications, Vol. II, Wiley, New York, (1971).

[19] A. Friedman, Stochastic differential equations, Vol. I, Academic Press, New York, (1975).

[20] P. Hall, C.C. Heyde, Martingale limit theory and its applications, Academic Press, New York, (1980).

[21] U. Horst, J. Wenzelburger, On non-ergodic asset prices, Econ. Theory. 34 (2008) 207-234.

[22] A.V. Ivanov, The Berry-Esseen inequality for the distribution of the least squares estimae, Math. Notes. 20 (1976) 721-727.

[23] R.S. Liptser, A.N. Shiryayev, Statistics of random processes I : general theory, Springer-Verlag, Berlin, (1977).

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

[25] I. Nourdin, G. Peccati, Stein's method on Wiener chaos, Prob. Theory Related Fields. 145 (2009) 75-118.

[26] I. Nourdin, G. Peccati, Normal approximation with Malliavin calculus: from Stein's method to universality, Cambridge University Press, Cambridge, (2012).

[27] V.V. Petrov, Limit theorems of probability theory, Oxford University Press, Oxford, (1995).

[28] J. Pfanzagl, The Berry-Esseen bound for minimum contrast estimators, Metrika. 17 (1971) 82-91.

[29] P. Protter, Stochastic integration and differential equations: a new approach, Springer-Verlag, Berlin, (1990).

[30] Y. Shimizu, Local asymptotic mixed normality for discretely observed non-recurrent Ornstein-Uhlenbeck processes, Ann. Inst. Stat. Math. 64 (2012) 193-211.