Estimation for Partially Observed Mean-reversion Type Stochastic Systems

Chao Wei

Abstract—**This paper is concerned with the parameter estimation problem for partially observed mean-reversion type stochastic systems. The Girsanov transformation is used to simplify the equation because of the expression of the drift coefficient. The suboptimal estimation of the state is obtained by constructing the extended Kalman filtering equation. The likelihood function is provided based on state estimation equation. The strong consistency of the estimator is proved by applying maximal inequality for martingales, Borel-Cantelli lemma and uniform ergodic theorem. An example is given to verify the effectiveness of the estimation methods.**

Index Terms—**Mean-reversion type stochastic systems, partially observed stochastic systems, Girsanov transformation, state estimation equation, parameter estimation, strong consistency.**

I. Introduction

Stochastic systems have been widely used in many application areas such as social, physics, physical, engineering and life sciences.([3], [4]). Recently, stochastic systems have been applied to describe the dynamics of a financial asset, asset portfolio and term structure of interest rates, such as Black-Scholes option pricing model ([5]), Vasicek and Cox-Ingersoll-Ross mean-reversion type models ([7], [8], [25]), Chan-Karloyi-Longstaff-Sanders model ([9]) and Ait-Sahalia model([1]). Some parameters in stochastic models describe the related assets dynamic, however, these parameters are always unknown. In the past few decades, some authors studied the parameter estimation problem for economic models. For example, Yu and Phillips([30]) used Gaussian approach to study the parameter estimation for continuous-time short-term interest rates model, Overback and Rydén([20]), Rossi([23]), Wei et al.([27]) investigated the parameter estimation problem for Cox-Ingersoll-Ross model by applying the maximum likelihood method, least-square method and Gaussian method respectively. Moreover, some methods have been used to estimate the parameters in general nonlinear stochastic differential equation. For instance, Bayes estimation([6], [14], [21]), maximum likelihood estimation([2], [28], [29]), least-square estimation([17]), minimum contrast estimation([13]) and M-estimation([24]).

In practice, the state of stochastic systems can not be observed completely. Some authors studied the state estimation problem for stochastic systems by using Kalman filtering or extend Kalman filtering([12], [16], [18], [26]). Furthermore, sometimes the parameters and states of a stochastic system are unknown at the same time. Therefore, the parameter estimation and state estimation needed to be solved simultaneously. In recent years, some authors investigated the parameter estimation problem for partially observed linear stochastic systems. For example, Deck and Theting([10]) used Kalman filtering and Bayes method to study the linear homogenous stochastic systems. Kan et al.([11]) discussed the linear nonhomogenous stochastic systems based on the methods used in ([10]). Mbalawata et al. [19] applied Kalman filtering and maximum likelihood estimation to investigate the parameter and state estimation for linear stochastic systems. However, the asymptotic property of the parameter estimator has not been discussed in ([19]), and in ([10], [11]), only drift parameter estimation has been studied.

In this paper, the parameter estimation problem for partially observed mean-reversion type stochastic systems is investigated. This topic has not been studied in the past literatures. Firstly, The Girsanov theorem has been used to simplify the drift coefficient of the stochastic systems and a new family of probability measures has been indexed. Then, the suboptimal estimation of the state has been obtained by constructing the extended Kalman filtering equation, both drift and diffusion item of the state estimation equation have the unknown parameter. The likelihood function has been provided based on state estimation equation. Finally, The strong consistency of the estimator has been proved by applying maximal inequality for martingales, Borel-Cantelli lemma and uniform ergodic theorem.

This paper is organized as follows. In Section 2, the drift coefficient of the stochastic systems is simplified and a new family of probability measures is indexed, the suboptimal estimation of the state is obtained. In Section 3, the likelihood function is given and the strong consistency of the estimator is proved. An example is given in Section 4. The conclusion is given in Section 5.

II. Problem Formulation and Preliminaries

In this paper, the parameter estimation problem for partially observed mean-reversion type stochastic system is investigated. The stochastic system is described as follows:

$$\begin{cases} dX_t =(\alpha+\mu(X_t))f(X_t,\theta)dt+g(X_t)dW_t \quad X_0\sim u_\theta, \\ dY_t =X_t dt+dV_t \quad Y_0=Y_0, \end{cases} \tag{1}$$

where θ is an unknown parameter, α is a constant, $\mu(x,\varepsilon)\in\mathbb{R}$ is twice differentiable with respect to x, $(W_t, t\geq 0)$ and $(V_t, t\geq 0)$ are independent Wiener processes defined on a complete probability space $(\Omega,\mathcal{F},\mathrm{P})$, X_t is ergodic, u_θ is the invariant measure, $\{Y_t\}$ is observable while $\{X_t\}$ is unobservable.

Because of the complexity of Equation 1, the Girsanov transformation will be used to simplify the drift coefficient. From now on the work is under the assumptions below.

This work was supported in part by the National Natural Science Foundation of China (Grant no. 61403248)

Chao Wei is with the School of Mathematics and Statistics, Anyang Normal University, Anyang 455000, China.(Email: chaowei0806@aliyun.com..

Assumption 1: $|f(x,\theta) - f(y,\theta)| + |g(x) - g(y)| \leq K(\theta)|x-y|, \quad \sup\{K(\theta)\} < \infty, \theta \in \Theta, x, y \in \mathbb{R}$.

Assumption 2: $|\mu(x,\varepsilon)f(x,\theta) - \mu(y,\varepsilon)f(y,\theta)| \leq K_3|x-y|$, K_3 is a positive constant, $\varepsilon \in (0,1]$.

Assumption 3: $|f(x,s)| \leq h(x), \mathbb{E}[h(X_0)]^2 < \infty, \mathbb{E}[g(X_0)]^2 < \infty, s \in I(\theta)$.

Assumption 4: $\mathbb{E}[f(X_0,\theta)(f(X_0,\theta_0) - \frac{1}{2}f(X_0,\theta))]$ has the unique maximal value at $\theta = \theta_0$, where θ_0 is the true parameter and m is a constant.

Firstly, we introduce the Girsanov theorem below.

Lemma 1: [15] Let $Y(t)$ be an Itô process of the form

$$dY(t) = a(t,\omega)dt + dB(t); t \leq T,$$

where $T \leq \infty$ is a given constant and $B(t)$ is Brownian motion. Put

$$M_t = \exp(-\int_0^t a(s,\omega)dB_s - \frac{1}{2}\int_0^t a^2(s,\omega)ds); t \leq T.$$

Assume that $a(s,\omega)$ satisfies Novikov's condition

$$\mathbb{E}[\exp(\frac{1}{2}\int_0^T a^2(s,\omega)ds)] < \infty,$$

where $\mathbb{E}$ is the expectation with respect to P. Define the measure Q on $(\Omega, \mathcal{F}_T)$ by

$$dQ(\omega) = M_T dP(\omega).$$

Then $Y(t)$ is a Brownian motion with respect to the probability law Q, for $t \leq T$.

According to the Girsanov theorem, Equation 1 could be written as

$$dX_t = ((\alpha + \mu(X_t))f(X_t,\theta) + g(X_t)m(X_t))dt + g(X_t)d\widetilde{W}_t, \tag{2}$$

where $m(X_t)$ satisfies the Novikov condition, $\widetilde{W}_t = W_t - \int_0^t m(X_s)ds$ is an $\{\mathcal{F}_t\}_{t\in[0,1]}$-Brownian motion under the probability Q_t.

Let

$$\mu(x)f(x,\theta) + g(x)m(x) = 0,$$

hence

$$\mu(X_t)f(X_{t,\theta}) + g(X_t)m(X_t) = 0.$$

Therefore, Equation 2 becomes

$$dX_t = \alpha f(X_t,\theta)dt + g(X_t)d\widetilde{W}_t. \tag{3}$$

Now we consider the parameter estimation for the following stochastic system:

$$\begin{cases} dX_t = \alpha f(X_t,\theta)dt + g(X_t)d\widetilde{W}_t \quad X_0 \sim u_\theta, \\ dY_t = X_t dt + dV_t \quad Y_0 = Y_0. \end{cases} \tag{4}$$

The likelihood function can not be given directly due to the unobservability of $\{X_t\}$. Therefore, we should estimate $\{X_t\}$ firstly.

The state estimator is designed as follows:

$$\begin{cases} d\widehat{X_t} = \alpha f(\widehat{X_t},\theta)dt + K_t(dY_t - \widehat{X_t}dt) \\ \widehat{X_0} = X_0. \end{cases} \tag{5}$$

According to (1) and (5), one has

$$\begin{aligned} & d(X_t - \widehat{X_t}) \\ = \;& (\alpha f(X_t,\theta) - \alpha f(\widehat{X_t},\theta) - K_t(X_t - \widehat{X_t}))dt \\ + \;& g(X_t)d\widetilde{W}_t - K_t dV_t. \end{aligned} \tag{6}$$

From Itô lemma and (6), it can be checked that

$$\begin{aligned} & d(X_t - \widehat{X_t})^2 \\ = \;& 2(X_t - \widehat{X_t})(\alpha f(X_t,\theta) - \alpha f(\widehat{X_t},\theta) \\ & -K_t(X_t - \widehat{X_t}))dt \\ + \;& 2(X_t - \widehat{X_t})(g(X_t)d\widetilde{W}_t - K_t dV_t) \\ & +(g^2(X_t) + K_t^2)dt \\ = \;& [2(X_t - \widehat{X_t})(\alpha f(X_t,\theta) - \alpha f(\widehat{X_t},\theta)) \\ & -2K_t(X_t - \widehat{X_t})^2 + g^2(X_t) + K_t^2]dt \\ + \;& 2(X_t - \widehat{X_t})(g(X_t)d\widetilde{W}_t - K_t dV_t). \end{aligned} \tag{7}$$

Taking expectation from both sides of (7), we obtain that

$$\begin{aligned} & d\mathbb{E}(X_t - \widehat{X_t})^2 \\ = \;& [2\mathbb{E}(X_t - \widehat{X_t})(\alpha f(X_t,\theta) - \alpha f(\widehat{X_t},\theta)) \\ & -2K_t\mathbb{E}(X_t - \widehat{X_t})^2 + \mathbb{E}(g^2(X_t)) + K_t^2]dt \\ = \;& [K_t^2 - 2K_t\mathbb{E}(X_t - \widehat{X_t})^2 + \mathbb{E}(g^2(X_t)) \\ & +2\mathbb{E}(X_t - \widehat{X_t})(\alpha f(X_t,\theta) - \alpha f(\widehat{X_t},\theta))]dt \\ = \;& [(K_t - \mathbb{E}(X_t - \widehat{X_t})^2)^2 - (\mathbb{E}(X_t - \widehat{X_t})^2)^2 \\ & +\mathbb{E}(g^2(X_t))]dt \\ + \;& 2\mathbb{E}(X_t - \widehat{X_t})(\alpha f(X_t,\theta) - \alpha f(\widehat{X_t},\theta))dt \\ \geq \;& [-(\mathbb{E}(X_t - \widehat{X_t})^2)^2 + \mathbb{E}(g^2(X_0)) + 2\mathbb{E}(X_t - \widehat{X_t}) \\ & (\alpha f(X_t,\theta) - \alpha f(\widehat{X_t},\theta))]dt. \end{aligned} \tag{8}$$

Therefore, when $K_t = \mathbb{E}(X_t - \widehat{X_t})^2$, (8) has the minimum value

$$\begin{aligned} & d\mathbb{E}(X_t - \widehat{X_t})^2 \\ & = [-(\mathbb{E}(X_t - \widehat{X_t})^2)^2 + \mathbb{E}(g^2(X_0)) \\ & +2\mathbb{E}(X_t - \widehat{X_t})(\alpha f(X_t,\theta) - \alpha f(\widehat{X_t},\theta))]dt. \end{aligned}$$

From the Assumption (1), one has

$$\begin{aligned} & [-(\mathbb{E}(X_t - \widehat{X_t})^2)^2 \\ & +\mathbb{E}(g^2(X_0)) + 2\mathbb{E}(X_t - \widehat{X_t}) \\ & (\alpha f(X_t,\theta) - \alpha f(\widehat{X_t},\theta))]dt \\ \leq \;& [-(\mathbb{E}(X_t - \widehat{X_t})^2)^2 + \mathbb{E}(g^2(X_0)) \\ & +2K(\theta)\alpha\mathbb{E}(X_t - \widehat{X_t})^2]dt. \end{aligned}$$

Since $f(X_t,\theta)$ is nonlinear, we can not obtain the optimal state estimation of X_t, the suboptimal state estimation is considered.

Considering the equation

$$\begin{aligned} & d\mathbb{E}(X_t - \widehat{X_t})^2 \\ & = (2K(\theta)\alpha\mathbb{E}(X_t - \widehat{X_t})^2 - (\mathbb{E}(X_t - \widehat{X_t})^2)^2 \\ & +\mathbb{E}(g^2(X_0)))dt. \end{aligned}$$

Let

$$\mathbb{E}(X_t - \widehat{X_t})^2 = \gamma_t, \quad \mathbb{E}(g^2(X_0))) = m, \tag{9}$$

one has

$$d\gamma_t = (2K(\theta)\alpha\gamma_t - \gamma_t^2 + m)dt. \tag{10}$$

It is easy to check that

$$\gamma_t = \frac{(\sqrt{K^2(\theta)\alpha^2 + m} + K(\theta)\alpha)(m - e^{-2t\sqrt{K^2(\theta)\alpha^2+m}})}{m + (\frac{K(\theta)\alpha + \sqrt{K^2(\theta)\alpha^2+m}}{\sqrt{K^2(\theta)\alpha^2+m} - K(\theta)\alpha})e^{-2t\sqrt{K^2(\theta)\alpha^2+m}}}. \tag{11}$$

Then, we obtain that

$$\gamma_t \to K(\theta)\alpha + \sqrt{K^2(\theta)\alpha^2 + m} = \gamma(\theta). \tag{12}$$

Therefore, it is obvious that

$$\begin{cases} d\widehat{X_t} = \alpha f(\widehat{X_t}, \theta)dt + \gamma_t(dY_t - \widehat{X_t}dt), \\ \widehat{X_0} = X_0. \end{cases} \tag{13}$$

Let

$$dV_t^* = dY_t - \widehat{X_t}dt, \tag{14}$$

where $(V_t^*, t \geq 0)$ is assumed as standard Wiener processes defined on complete probability space $(\Omega, \mathscr{F}, \mathrm{P})$.

Hence,

$$\begin{cases} d\widehat{X_t} = \alpha f(\widehat{X_t}, \theta)dt + \gamma_t dV_t^* \\ \widehat{X_0} = X_0. \end{cases} \tag{15}$$

It is assumed that the system (15) reaches the steady state, which means that

$$\begin{cases} d\widehat{X_t} = \alpha f(\widehat{X_t}, \theta)dt + \gamma(\theta) dV_t^* \\ \widehat{X_0} = X_0. \end{cases} \tag{16}$$

In summary, the suboptimal state estimation equation of X_t is (16).

Remark 1: System (15) reaches the steady state means that the Riccatti equation satisfies $\frac{d\gamma_t}{dt} = 0$. Hence, we obtain γ_t=$\gamma(\theta)$.

Remark 2: In (16), both drift item and diffusion item have the parameter. Thus, it is difficult to discuss the asymptotic property of the estimator. In the next section, the problem is solved.

III. Main Results and Proofs

In the following theorem, the strong consistency of the maximum likelihood estimator is proved by using maximal inequality for martingales, Borel-Cantelli lemma and uniform ergodic theorem.

Theorem 1: Under Assumptions 1-4, the maximum likelihood estimator $\widehat{\theta}_t$ is strong consistent, namely

$$\widehat{\theta}_t \overset{a.s.}{\to} \theta_0.$$

The likelihood function has the following expression

$$\ell_t(\theta) = \int_0^t \frac{\alpha f(\widehat{X_s}, \theta)}{\gamma^2(\theta)} d\widehat{X_s} - \frac{1}{2}\int_0^t \frac{\alpha^2 f^2(\widehat{X_s}, \theta)}{\gamma^2(\theta)} ds. \tag{17}$$

Since θ_0 is the true parameter, from (16), (17) can be written as

$$\begin{aligned} & \ell_t(\theta) \\ = & \int_0^t \frac{\alpha f(\widehat{X_s}, \theta)}{\gamma^2(\theta)} (\alpha f(\widehat{X_s}, \theta_0)ds + \gamma(\theta_0)dV_s^*) \\ & -\frac{1}{2}\int_0^t \frac{\alpha^2 f^2(\widehat{X_s}, \theta)}{\gamma^2(\theta)} ds \\ = & \int_0^t \frac{\alpha^2 f(\widehat{X_s}, \theta) f(\widehat{X_s}, \theta_0)}{\gamma^2(\theta)} ds \\ & + \int_0^t \frac{\alpha f(\widehat{X_s}, \theta)\gamma(\theta_0)}{\gamma^2(\theta)} dV_s^* - \frac{1}{2}\int_0^t \frac{\alpha^2 f^2(\widehat{X_s}, \theta)}{\gamma^2(\theta)} ds \\ = & \alpha^2 \int_0^t \frac{f(\widehat{X_s}, \theta)[f(\widehat{X_s}, \theta_0) - \frac{1}{2}f(\widehat{X_s}, \theta)]}{\gamma^2(\theta)} ds \\ & + \int_0^t \frac{\alpha f(\widehat{X_s}, \theta)\gamma(\theta_0)}{\gamma^2(\theta)} dV_s^*. \end{aligned}$$

By applying the uniform ergodic theorem, one has

$$\begin{aligned} & \frac{1}{t}\int_0^t \frac{f(\widehat{X_s}, \theta)[f(\widehat{X_s}, \theta_0) - \frac{1}{2}f(\widehat{X_s}, \theta)]}{\gamma^2(\theta)} ds \\ \overset{a.s.}{\to} & \mathbb{E}[\frac{f(\widehat{X_0}, \theta)[f(\widehat{X_0}, \theta_0) - \frac{1}{2}f(\widehat{X_0}, \theta)]}{\gamma^2(\theta)}] \\ = & \mathbb{E}[\frac{f(X_0, \theta)[f(X_0, \theta_0) - \frac{1}{2}f(X_0, \theta)]}{\gamma^2(\theta)}]. \end{aligned}$$

By using the maximal inequality for martingales, it can be checked that

$$\begin{aligned} & \mathbb{P}_\theta(\sup_{0<t\leq t_0} |\int_0^t f(\widehat{X_s}, \theta)dV_s^*| > \varepsilon) \\ \leq & \frac{\mathbb{E}_\theta(\int_0^{t_0} f(\widehat{X_s}, \theta)dV_s^*)^2}{\varepsilon^2} \\ = & \frac{t_0\mathbb{E}_\theta(f(\widehat{X_0}, \theta))^2}{\varepsilon^2} \\ = & \frac{t_0\mathbb{E}_\theta(f(X_0, \theta))^2}{\varepsilon^2}. \end{aligned}$$

Let

$$\mathcal{B}_n = \{\sup_{2^{n-1}<t<2^n} \sup_\theta |\int_0^t f(\widehat{X_s}, \theta)dV_s^*| > 2^{\frac{n}{2}} n^\alpha\}, \tag{18}$$

where $n \geq 1$, $\alpha > \frac{1}{2}$.

Then,

$$\begin{aligned} & \mathbb{P}_\theta(\mathcal{B}_n) \\ = & \mathbb{P}_\theta(\sup_{0<t<2^{n-1}} \sup_\theta |\int_0^t f(\widehat{X_s}, \theta)dV_s^*| > 2^{\frac{n}{2}} n^\alpha) \\ \leq & \frac{2^{n-1}\mathbb{E}_\theta(f(\widehat{X_s}, \theta))^2}{2^n n^{2\alpha}} \\ = & \frac{\mathbb{E}_\theta(f(\widehat{X_0}, \theta))^2}{2}\frac{1}{n^{2\alpha}} \\ = & \frac{\mathbb{E}_\theta(f(X_0, \theta))^2}{2}\frac{1}{n^{2\alpha}}. \end{aligned}$$

Thus, it is easy to check that

$$\Sigma_{n=1}^\infty \mathbb{P}_\theta(\mathcal{B}_n) < \infty. \tag{19}$$

According to Borel-Cantelli lemma, one has

$$\mathbb{P}_\theta(\limsup_{n\to\infty} \mathcal{B}_n) = 0. \tag{20}$$

As $2^{n-1} < t < 2^n$, it follows that

$$2^{\frac{n}{2}} < 2^{\frac{1}{2}} t^{\frac{1}{2}}, \tag{21}$$

and

$$n^\alpha < (\ln t)^\alpha (\frac{1}{\ln t} + \frac{1}{\ln 2})^\alpha. \tag{22}$$

Therefore, when t is large enough,

$$\limsup_{t\to\infty} \sup_\theta \frac{|\int_0^t f(\widehat{X_s}, \theta)dV_s^*|}{t^{\frac{1}{2}}(\ln t)^\alpha} < 2^{\frac{1}{2}}(\frac{1}{\ln 2})^\alpha \quad a.s. \tag{23}$$

$$\sup_\theta |\frac{1}{t}\int_0^t f(\widehat{X_s}, \theta)dV_s^*| < \frac{(\ln t)^\alpha}{t^{\frac{1}{2}}} 2^{\frac{1}{2}}(\frac{1}{\ln 2})^\alpha, \tag{24}$$

with probability one.

Since

$$\frac{(\ln t)^\alpha}{t^{\frac{1}{2}}} \to 0, \tag{25}$$

we obtain that

$$\sup_{\theta} |\frac{1}{t}\int_0^t f(\widehat{X_s},\theta)dV_s^*| \xrightarrow{a.s.} 0. \quad (26)$$

As

$$|\frac{1}{t}\int_0^t f(\widehat{X_s},\theta)dV_s^*| \le \sup_{\theta} |\frac{1}{t}\int_0^t f(\widehat{X_s},\theta)dV_s^*|, \quad (27)$$

one has

$$\frac{1}{t}\int_0^t f(\widehat{X_s},\theta)dV_s^* \xrightarrow{a.s.} 0. \quad (28)$$

Form the above results, it follows that

$$\frac{1}{t}\ell_t(\theta) \xrightarrow{a.s.} \alpha^2 \mathbb{E}[\frac{f(X_0,\theta)[f(X_0,\theta_0)-\frac{1}{2}f(X_0,\theta)]}{\gamma^2(\theta)}]. \quad (29)$$

Therefore, it can be checked that

$$\widehat{\theta}_t \xrightarrow{a.s.} \theta_0. \quad (30)$$

Remark 3: When the stochastic system is observed discretely, the approximate likelihood function can be written as

$$\ell_n(\theta) = \sum_{i=1}^n \frac{f(\widehat{X_{t_{i-1}}},\theta)}{\gamma^2(\theta)}(\widehat{X_{t_i}}-\widehat{X_{t_{i-1}}}) - \frac{\Delta}{2}\sum_{i=1}^n \frac{f^2(\widehat{X_{t_{i-1}}},\theta)}{\gamma^2(\theta)}. \quad (31)$$

The following lemmas are useful to derive our results.

Lemma 2: Assume that $\{\widehat{X_t}\}$ is a solution of the stochastic differential equation (15). Then, for any integer $n \ge 1$ and $0 \le s \le t$,

$$\mathbb{E}|\widehat{X_t}-\widehat{X_s}|^{2p} = O(|t-s|^p).$$

Proof: Suppose θ_0 is the true parameter value, by applying Holder's inequality, it follows that

$$\begin{aligned}
& |\widehat{X_t}-\widehat{X_s}|^{2p} \\
= \ & |\int_s^t f(\widehat{X_u},\theta_0)du + \gamma(\theta_0)\int_s^t dV_u^*|^{2p} \\
\le \ & 2^{2p-1}(|\int_s^t f(\widehat{X_u},\theta_0)du|^{2p} + (\gamma(\theta_0))^{2p}|\int_s^t dV_u^*|^{2p}) \\
\le \ & 2^{2p-1}((t-s)^{2p-1} \\
& \int_s^t |f(\widehat{X_u},\theta_0)|^{2p}du + (\gamma(\theta_0))^{2p}|\int_s^t dV_u^*|^{2p})
\end{aligned}$$

Since

$$|f(\widehat{X_u},\theta_0)|^{2p} \le K_1(\theta_0)^p 2^{p-1}(1+|\widehat{X_u}|^{2p}), \quad (32)$$

from Assumption 4 together with the stationarity of the process, one has

$$\mathbb{E}[\int_s^t |f(\widehat{X_u},\theta_0)|^{2p}du] = O(|t-s|). \quad (33)$$

From Burkholder-Davis-Gundy inequality, it can be checked that

$$\mathbb{E}[|\int_s^t dV_u^*|^{2p}] \le C_p \mathbb{E}|\int_s^t du|^p = C_p(t-s)^{2p}, \quad (34)$$

where C_p is a positive constant depending only on p.

Then, we have

$$\mathbb{E}[|\int_s^t dV_u^*|^{2p}] = O(|t-s|^p). \quad (35)$$

From the above analysis, it follows that

$$\mathbb{E}|\widehat{X_t}-\widehat{X_s}|^{2p} = O(|t-s|^p). \quad (36)$$

The proof is complete. ■

Lemma 3: When $\Delta \to 0$, one has

$$\begin{aligned}
& \mathbb{E}|\sum_{i=1}^n \int_{t_{i-1}}^{t_i} \frac{f(\widehat{X_{t_{i-1}}},\theta)f(\widehat{X_s},\theta_0)}{\gamma^2(\theta)}ds \\
& -\sum_{i=1}^n \frac{f(\widehat{X_{t_{i-1}}},\theta)f(\widehat{X_{t_{i-1}}},\theta_0)}{\gamma^2(\theta)}\Delta| \to 0.
\end{aligned}$$

Proof: By applying Holder's inequality, it can be checked that

$$\begin{aligned}
& \mathbb{E}|\sum_{i=1}^n \int_{t_{i-1}}^{t_i} \frac{f(\widehat{X_{t_{i-1}}},\theta)f(\widehat{X_s},\theta_0)}{\gamma^2(\theta)}ds \\
& -\sum_{i=1}^n \frac{f(\widehat{X_{t_{i-1}}},\theta)f(\widehat{X_{t_{i-1}}},\theta_0)}{\gamma^2(\theta)}\Delta| \\
= \ & \mathbb{E}|\sum_{i=1}^n \int_{t_{i-1}}^{t_i} \frac{f(\widehat{X_{t_{i-1}}},\theta)(f(\widehat{X_s},\theta_0)-f(\widehat{X_{t_{i-1}}},\theta_0))}{\gamma^2(\theta)}ds| \\
\le \ & \sum_{i=1}^n \int_{t_{i-1}}^{t_i} \mathbb{E}|\frac{f(\widehat{X_{t_{i-1}}},\theta)(f(\widehat{X_s},\theta_0)-f(\widehat{X_{t_{i-1}}},\theta_0))}{\gamma^2(\theta)}|ds \\
\le \ & \sum_{i=1}^n \int_{t_{i-1}}^{t_i} (\mathbb{E}[\frac{f(\widehat{X_{t_{i-1}}},\theta)}{\gamma^2(\theta)}]^2)^{\frac{1}{2}} \\
& (\mathbb{E}[f(\widehat{X_s},\theta_0)-f(\widehat{X_{t_{i-1}}},\theta_0)]^2)^{\frac{1}{2}}ds.
\end{aligned}$$

Then, we have

$$\mathbb{E}[f(\widehat{X_s},\theta_0)-f(\widehat{X_{t_{i-1}}},\theta_0)]^2 = O(\Delta), \quad (37)$$

and $\mathbb{E}[\frac{f(\widehat{X_{t_{i-1}}},\theta)}{\gamma^2(\theta)}]^2$ is bounded.

From the above analysis, it follows that

$$\begin{aligned}
& \mathbb{E}|\sum_{i=1}^n \int_{t_{i-1}}^{t_i} \frac{f(\widehat{X_{t_{i-1}}},\theta)f(\widehat{X_s},\theta_0)}{\gamma^2(\theta)}ds \\
& -\sum_{i=1}^n \frac{f(\widehat{X_{t_{i-1}}},\theta)f(\widehat{X_{t_{i-1}}},\theta_0)}{\gamma^2(\theta)}\Delta| \to 0,
\end{aligned}$$

as $\Delta \to 0$.

The proof is complete. ■

Theorem 2: When $\Delta \to 0$, $n \to \infty$ and $n\Delta \to \infty$,

$$\widehat{\theta}_0 \xrightarrow{P} \theta_0.$$

Proof: According to the expression of the approximate likelihood function and equation (1), it follows that

$$
\begin{aligned}
&\ell_n(\theta)\\
=& \sum_{i=1}^{n} \frac{f(\widehat{X_{t_{i-1}}},\theta)}{\gamma^2(\theta)}(\int_{t_{i-1}}^{t_i} f(\widehat{X_s},\theta_0)ds\\
&+\gamma(\theta_0)\int_{t_{i-1}}^{t_i} dV_s^*)\\
&-\frac{\Delta}{2}\sum_{i=1}^{n}\frac{f^2(\widehat{X_{t_{i-1}}},\theta)}{\gamma^2(\theta)}\\
=& \sum_{i=1}^{n}\int_{t_{i-1}}^{t_i}\frac{f(\widehat{X_{t_{i-1}}},\theta)f(\widehat{X_s},\theta_0)}{\gamma^2(\theta)}ds\\
&+\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i}\frac{f(\widehat{X_{t_{i-1}}},\theta)\gamma(\theta_0)}{\gamma^2(\theta)}dV_s^*\\
&-\frac{1}{2}\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i}\frac{f^2(\widehat{X_{t_{i-1}}},\theta)}{\gamma^2(\theta)}ds\\
=& \sum_{i=1}^{n}\int_{t_{i-1}}^{t_i}\frac{f(\widehat{X_{t_{i-1}}},\theta)(f(\widehat{X_s},\theta_0)-\frac{1}{2}f(\widehat{X_{t_{i-1}}},\theta))}{\gamma^2(\theta)}ds\\
+& \sum_{i=1}^{n}\int_{t_{i-1}}^{t_i}\frac{f(\widehat{X_{t_{i-1}}},\theta)\gamma(\theta_0)}{\gamma^2(\theta)}dV_s^*.
\end{aligned}
$$

Then, we have

$$
\begin{aligned}
&\frac{1}{n\Delta}\ell_n(\theta) &(38)\\
&=\frac{1}{n\Delta}\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i}\\
&\frac{f(\widehat{X_{t_{i-1}}},\theta)(f(\widehat{X_s},\theta_0)-\frac{1}{2}f(\widehat{X_{t_{i-1}}},\theta))}{\gamma^2(\theta)}ds &(39)\\
&+\frac{1}{n\Delta}\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i}\frac{f(\widehat{X_{t_{i-1}}},\theta)\gamma(\theta_0)}{\gamma^2(\theta)}dV_s^*.
\end{aligned}
$$

From the martingale moment inequality, it can be checked that

$$
\begin{aligned}
&\mathbb{E}|\frac{1}{n\Delta}\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i}\frac{f(\widehat{X_{t_{i-1}}},\theta)\gamma(\theta_0)}{\gamma^2(\theta)}dV_s^*|^2\\
\leq\ & \frac{1}{(n\Delta)^2}C\frac{\gamma^2(\theta_0)}{\gamma^4(\theta)}\mathbb{E}\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i}(f(\widehat{X_{t_{i-1}}},\theta))^2ds\\
\leq\ & \frac{1}{n\Delta}C_1\\
\to\ & 0,
\end{aligned}
$$

where C and C_1 are constants.

By applying Chebyshev inequality, it can be obtained that

$$
\frac{1}{n\Delta}\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i}\frac{f(\widehat{X_{t_{i-1}}},\theta)\gamma(\theta_0)}{\gamma^2(\theta)}dV_s^*\xrightarrow{P}0, \tag{40}
$$

when $\Delta\to 0$, $n\to\infty$ and $n\Delta\to\infty$.

By the uniform ergodic theorem (see e.g. [21]), one has

$$
\begin{aligned}
&\frac{1}{n\Delta}\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i}\frac{f(\widehat{X_{t_{i-1}}},\theta)(f(\widehat{X_s},\theta_0)-\frac{1}{2}f(\widehat{X_{t_{i-1}}},\theta))}{\gamma^2(\theta)}ds\\
&\xrightarrow{P}\mathbb{E}[\frac{f(x_0,\theta)(f(x_0,\theta_0)-\frac{1}{2}f(x_0,\theta)}{\gamma^2(\theta)}],
\end{aligned}
$$

when $\Delta\to 0$, $n\to\infty$ and $n\Delta\to\infty$.

Hence, it leads to the relation that

$$
\frac{1}{n\Delta}\ell_n(\theta)\xrightarrow{P}\mathbb{E}[\frac{f(x_0,\theta)(f(x_0,\theta_0)-\frac{1}{2}f(x_0,\theta)}{\gamma^2(\theta)}], \tag{41}
$$

when $\Delta\to 0$, $n\to\infty$ and $n\Delta\to\infty$.

It is easy to check that

$$
\widehat{\theta_0}\xrightarrow{P}\theta_0, \tag{42}
$$

when $\Delta\to 0$, $n\to\infty$ and $n\Delta\to\infty$.

The proof is complete. ■

Theorem 3: When $\Delta\to 0$, $n^{\frac{1}{2}}\Delta\to 0$ and $n\Delta\to\infty$ as $n\to\infty$,

$$
\begin{aligned}
&\sqrt{n\Delta}(\theta_0-\widehat{\theta_0})\\
&\xrightarrow{d}N(0,\frac{\gamma^2(\theta_0)}{\mathbb{E}[f'(x_0,\theta_0)\gamma(\theta_0)-2f(x_0,\theta_0)\gamma'(\theta_0)]^2}).
\end{aligned}
$$

Proof: Expanding $\ell_n'(\theta_0)$ about $\widehat{\theta_0}$, it follows that

$$
\ell_n'(\theta_0)=\ell_n'(\widehat{\theta_0})+\ell_n''(\widetilde{\theta})(\theta_0-\widehat{\theta_0}), \tag{43}
$$

where $\widetilde{\theta}$ is between $\widehat{\theta_0}$ and θ_0.

In view of Theorem 1, it is known that $\ell_n'(\widehat{\theta_0})=0$, then

$$
\ell_n'(\theta_0)=\ell_n''(\widetilde{\theta})(\theta_0-\widehat{\theta_0}). \tag{44}
$$

From the same method used in Theorem 2, it is easy to check that

$$
\begin{aligned}
&\frac{1}{n\Delta}\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i}\frac{f''(\widehat{X_{t_{i-1}}},\theta_0)\gamma(\theta_0)-4\gamma'(\theta_0)f'(\widehat{X_{t_{i-1}}},\theta_0)}{\gamma^2(\theta_0)}\\
&dV_s^*\xrightarrow{P}0,
\end{aligned}
$$

and

$$
\begin{aligned}
&\frac{1}{n\Delta}\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i}\\
&\frac{6f(\widehat{X_{t_{i-1}}},\theta_0)(\gamma'(\theta_0))^2-2\gamma(\theta_0)\gamma''(\theta_0)f(\widehat{X_{t_{i-1}}},\theta_0)}{\gamma^3(\theta_0)}\\
&dV_s^*\xrightarrow{P}0.
\end{aligned}
$$

By applying the results of Lemmas 3 and the uniform ergodic theorem, it follows that

$$
\begin{aligned}
&\frac{1}{n\Delta}\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i}\frac{f''(\widehat{X_{t_{i-1}}},\theta_0)(f(\widehat{X_s},\theta_0)-f(\widehat{X_{t_{i-1}}},\theta_0))}{\gamma^2(\theta_0)}\\
&ds\xrightarrow{P}0,
\end{aligned}
$$

$$
\frac{1}{n\Delta}\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i}\frac{(f'(\widehat{X_{t_{i-1}}},\theta_0))^2}{\gamma^2(\theta_0)}ds\xrightarrow{P}\frac{1}{\gamma^2(\theta_0)}\mathbb{E}[f'(x_0,\theta_0)]^2. \tag{45}
$$

Therefore, we have

$$
\begin{aligned}
&\frac{1}{n\Delta}\ell_n''(\theta_0)\xrightarrow{P}\\
&\mathbb{E}[\frac{3(\gamma'(\theta_0))^2-\gamma^2(\theta_0)(\gamma''(\theta_0))^2}{\gamma^4(\theta_0)}\\
&(f(x_0,\theta_0))^2+\frac{1}{\gamma^2(\theta_0)}(f'(x_0,\theta_0))^2].
\end{aligned}
$$

According to the expression of $\ell_n''(\theta)$ and by employing the martingale moment inequality, Chebyshev inequality, the uniform ergodic theorem and the dominated convergence theorem, it follows that

$$\frac{1}{n\Delta}(\ell_n''(\widetilde{\theta}) - \ell_n''(\theta_0)) \xrightarrow{P} 0. \tag{46}$$

Hence, it can be obtained that

$$\frac{1}{n\Delta}\ell_n''(\widetilde{\theta}) \xrightarrow{P} \mathbb{E}[\frac{3(\gamma'(\theta_0))^2 - \gamma^2(\theta_0)(\gamma''(\theta_0))^2}{\gamma^4(\theta_0)} (f(x_0,\theta_0))^2 + \frac{1}{\gamma^2(\theta_0)}(f'(x_0,\theta_0))^2].$$

Since

$$\begin{aligned}
&\ell_n'(\theta)\\
=& \sum_{i=1}^{n}\frac{f'(\widehat{X_{t_{i-1}}},\theta)\gamma(\theta) - 2f(\widehat{X_{t_{i-1}}},\theta)\gamma'(\theta)}{\gamma^3(\theta)} (\widehat{X_{t_i}} - \widehat{X_{t_{i-1}}})\\
-& \Delta\sum_{i=1}^{n} \frac{f(\widehat{X_{t_{i-1}}},\theta)f'(\widehat{X_{t_{i-1}}},\theta)\gamma(\theta) - f^2(\widehat{X_{t_{i-1}}},\theta)\gamma'(\theta)}{\gamma^3(\theta)}\\
=& \sum_{i=1}^{n}\int_{t_{i-1}}^{t_i}\frac{f'(\widehat{X_{t_{i-1}}},\theta)(f(\widehat{X_s},\theta_0) - f(X_{t_{i-1}},\theta))}{\gamma^2(\theta)}ds\\
+& \sum_{i=1}^{n}\int_{t_{i-1}}^{t_i} \frac{f(\widehat{X_{t_{i-1}}},\theta)\gamma'(\theta)(f(X_{t_{i-1}},\theta) - 2f(\widehat{X_s},\theta_0))}{\gamma^3(\theta)}ds\\
+& \sum_{i=1}^{n}\int_{t_{i-1}}^{t_i} \gamma(\theta_0)\frac{f'(\widehat{X_{t_{i-1}}},\theta)\gamma(\theta) - 2f(\widehat{X_{t_{i-1}}},\theta)\gamma'(\theta)}{\gamma^3(\theta)}dV_s^*,
\end{aligned}$$

it follows that

$$\begin{aligned}
&\frac{1}{\sqrt{n\Delta}}\ell_n'(\theta_0)\\
=& \frac{1}{\sqrt{n\Delta}}\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i} \frac{f'(\widehat{X_{t_{i-1}}},\theta_0)(f(\widehat{X_s},\theta_0) - f(X_{t_{i-1}},\theta_0))}{\gamma^2(\theta_0)}ds\\
+& \frac{1}{\sqrt{n\Delta}}\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i} \frac{f(\widehat{X_{t_{i-1}}},\theta_0)\gamma'(\theta_0)(f(X_{t_{i-1}},\theta_0) - 2f(\widehat{X_s},\theta_0))}{\gamma^3(\theta_0)}ds\\
+& \frac{1}{\sqrt{n\Delta}}\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i} \frac{f'(\widehat{X_{t_{i-1}}},\theta_0)\gamma(\theta_0) - 2f(\widehat{X_{t_{i-1}}},\theta_0)\gamma'(\theta_0)}{\gamma^2(\theta_0)}dV_s^*.
\end{aligned}$$

From the Lemma 1 and Assumption 1, we have

$$\mathbb{E}[f(\widehat{X_s},\theta_0) - f(\widehat{X_{t_{i-1}}},\theta_0)]^2 = O(\Delta). \tag{47}$$

Then, it follows that

$$\mathbb{E}|\frac{1}{\sqrt{n\Delta}}\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i} \frac{f'(\widehat{X_{t_{i-1}}},\theta_0)(f(\widehat{X_s},\theta_0) - f(\widehat{X_{t_{i-1}}},\theta_0))}{\gamma^2(\theta_0)}ds| \to 0,$$

when $\Delta \to 0$, $n^{\frac{1}{2}}\Delta \to 0$ and $n\Delta \to \infty$ as $n \to \infty$.

By applying the Chebyshev inequality, it can be obtained that

$$\frac{1}{\sqrt{n\Delta}}\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i} \frac{f'(\widehat{X_{t_{i-1}}},\theta_0)(f(\widehat{X_s},\theta_0) - f(\widehat{X_{t_{i-1}}},\theta_0))}{\gamma^2(\theta_0)}ds \xrightarrow{P} 0.$$

By applying the same methods, we have

$$\frac{1}{\sqrt{n\Delta}}\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i} \frac{f(\widehat{X_{t_{i-1}}},\theta)\gamma'(\theta)(f(X_{t_{i-1}},\theta) - 2f(\widehat{X_s},\theta_0))}{\gamma^3(\theta)}ds \xrightarrow{P} 0.$$

It is obviously that

$$\begin{aligned}
&\frac{1}{\sqrt{n\Delta}}\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i} \frac{f'(\widehat{X_{t_{i-1}}},\theta_0)\gamma(\theta_0) - 2f(\widehat{X_{t_{i-1}}},\theta_0)\gamma'(\theta_0)}{\gamma^2(\theta_0)}dV_s^* \xrightarrow{d}\\
&N(0, \frac{1}{\gamma^2(\theta_0)}\mathbb{E}[f'(x_0,\theta_0)\gamma(\theta_0) - 2f(x_0,\theta_0)\gamma'(\theta_0)]^2).
\end{aligned}$$

Hence, we have

$$\begin{aligned}
&\frac{1}{\sqrt{n\Delta}}\ell_n'(\theta_0) \xrightarrow{d}\\
&N(0, \frac{1}{\gamma^2(\theta_0)}\mathbb{E}[f'(x_0,\theta_0)\gamma(\theta_0) - 2f(x_0,\theta_0)\gamma'(\theta_0)]^2),
\end{aligned}$$

when $\Delta \to 0$, $n^{\frac{1}{2}}\Delta \to 0$ and $n\Delta \to \infty$ as $n \to \infty$.

From the above analysis, it can be checked that

$$\begin{aligned}
&\sqrt{n\Delta}(\theta_0 - \widehat{\theta}_0) \xrightarrow{d}\\
&N(0, \frac{\gamma^2(\theta_0)}{\mathbb{E}[f'(x_0,\theta_0)\gamma(\theta_0) - 2f(x_0,\theta_0)\gamma'(\theta_0)]^2}),
\end{aligned}$$

when $\Delta \to 0$, $n^{\frac{1}{2}}\Delta \to 0$ and $n\Delta \to \infty$ as $n \to \infty$.

The proof is complete. ■

IV. Example

Consider the following partially observed stochastic systems:

$$\begin{cases} dX_t = (\alpha + \varepsilon X_t^{-1})\theta\dfrac{X_t}{\sqrt{1+X_t^2}}dt + \varepsilon dW_t \quad X_0 \sim u_\theta,\\ dY_t = X_t dt + dV_t \quad Y_0 = Y_0, \end{cases}$$

where θ is an unknown parameter, α is a constant, $(W_t, t \geq 0)$ and $(V_t, t \geq 0)$ are independent Wiener processes defined on a complete probability space $(\Omega, \mathcal{F}, \mathrm{P})$, X_t is ergodic, u_θ is the invariant measure, $\{Y_t\}$ is observable while $\{X_t\}$ is unobservable.

Since

$$|\theta\frac{x}{\sqrt{1+x^2}}-\theta\frac{y}{\sqrt{1+y^2}}|\leq 2\theta|x-y|,$$

$$|\varepsilon x^{-1}\theta\frac{x}{\sqrt{1+x^2}}-\varepsilon y^{-1}\theta\frac{y}{\sqrt{1+y^2}}|\leq \varepsilon\theta|x-y|,$$

$$|s\frac{x}{\sqrt{1+x^2}}|\leq sx,\quad \mathbb{E}[X_0]^2<\infty,$$

and $\mathbb{E}[\theta\frac{X_0}{\sqrt{1+X_0^2}}(\theta_0\frac{X_0}{\sqrt{1+X_0^2}}-\frac{1}{2}\theta\frac{X_0}{\sqrt{1+X_0^2}})]$ has the unique maximal value at $\theta=\theta_0$.

It can be derived that coefficients of the stochastic system satisfy Assumptions 1-4.

Therefore, it is easy to check that

$$\widehat{\theta}_t\overset{a.s.}{\to}\theta_0.$$

V. Conclusion

The aim of this paper is to estimate the parameter for partially observed mean-reversion type stochastic system. The Girsanov transformation has been used to simplify the equation because of the expression of the drift coefficient. The suboptimal state estimation equation has been obtained by constructing the extended Kalman filtering equation. The likelihood function has been given and the strong consistency of the maximum likelihood estimator has been proved by using maximal inequality for martingales, Borel-Cantelli lemma and uniform ergodic theorem. Further research topics will include the parameter estimation for partially observed nonlinear stochastic systems driven by lévy process.

References

[1] Y. Ait-Sahalia. "Testing continuous-time models of the spot interest rate", *Review of Financial Studies*, vol. 9, no. 2, pp. 385-426, 1996.

[2] M. Barczy, G. Pap. "Asymptotic behavior of maximum likelihood estimator for time inhomogeneous diffusion processes", *Journal of Statistical Planning and Inference*, vol. 140, no. 1, pp. 1576-1593, 2010.

[3] B. Øksendal. *Stochastic Differential Equations: an Introduction with Applications*, Springer-Verlag, Berlin, 2003.

[4] J. P. N. Bishwal. *Parameter estimation in stochastic differential equations*, Springer-Verlag, London, 2008.

[5] F. Black, M. Scholes, "The pricing of options and corporate liabilities", *Journal of Political Economy*, vol. 81, no. 1, pp. 637-654, 1973.

[6] J. W. Cai, P. Chen, X. Mei, "Range-based threshold spot volatility estimation for jump diffusion models", *IAENG International Journal of Applied Mathematics*, vol. 47, no. 1, pp. 43-48, 2017.

[7] Cox. J.,Ingersoll. J., S. Ross. "An intertemporal general equilibrium model of asset prices", *Econometrica*, vol. 53, no. 1, pp. 363-384. 1985.

[8] Cox. J.,Ingersoll. J., S. Ross. "A theory of the term structure of interest rates", *Econometrica*. vol. 53, no. 1, pp. 385-408. 1985.

[9] K. C. Chan. "An empirical comparision of alternative models of the short-term interest rate", *The Journal of Finance*, vol. 47, no. 3, pp. 1209-1227, 1992.

[10] T. Deck, T. G. Theting. "Robust parameter estimation for stochastic differential equations", *Acta Applicandae Mathematicae*, vol. 84, no. 1, pp. 279-314, 2004.

[11] X. Kan, H. S. Shu, Y. Che, "Asymptotic parameter estimation for a class of linear stochastic systems using Kalman-Bucy filtering", *Mathematical Problems in Engineering*, Article ID: 342705, DOI: 10.1155/2012/342705, 2012.

[12] R. Kannan. "Orientation estimation based on LKF using differential state equation", *IEEE Sensors Journal*, vol. 15, no. 2, pp. 6156-6163, 2015.

[13] M. Kessler. "Estimation of an ergodic diffusion from discrete observations", *Journal of Statistics*, vol. 24, no. 1, pp. 211-229, 1997.

[14] Y. A. Kutoyants. *Statistical Inference for Ergodie Diffusion Processes*, Springer-Verlag, London, 2004.

[15] R. S. Liptser, A. N. Shiryayev. *Statistics of Rondom Processes*, Springer, New York, 1977.

[16] C. P. Li, H. B. Hao, "Likelihood and Bayesian estimation in stress strength model from generalized exponential distribution containing outliers," *IAENG International Journal of Applied Mathematics*, vol. 46, no. 2, pp. 155-159, 2016.

[17] H. W. Long, Y. Shimizu, W. Sun, "Least squares estimators for discretely observed stochastic processes driven by small Lévy noises", *Journal of Multivariate Analysis*. vol. 116, no. 1, pp. 422-439, 2013.

[18] X. Lu, L. Xie, H. Zhang, W. Wang. "Robust Kalman filtering for discrete-time systems with measurement delay", *IEEE Transactions on Circuits and Systems-II: Express Briefs*, vol. 54, no. 3, pp. 522-526, 2007.

[19] I. S. Mbalawata, S. Särkkä, H. Haario. "Parameter estimation in stochastic differential equations with Markov chain Monte Carlo and non-linear Kalman filtering", *Computational and Applied Statistics*, vol. 28, no. 3, pp. 1195-1223, 2013.

[20] L. Overback, T. Rydén. "Estimation in Cox-Ingersoll-Ross model", *Econometric Theory*, vol. 13, no. 1, pp. 430-461, 1997.

[21] B. L. S. Prakasa Rao. *Statistical Inference for Diffusion Type Processes*, Arnold Publishers, London, 1999.

[22] P. E. Protter. *Stochastic Integration and Differential Equations: Stochastic Modelling and Applied Probability*, 2nd ed, Applications of Mathematics(New York) 21, Springer, Berlin, 2004.

[23] G. D. Rossi. "Maximum likelihood estimation of the Cox-Ingersoll-Ross model using particle filters", *Computational Economics*, vol. 36, no. 1, pp. 1-16, 2010.

[24] Y. Shimizu. "Estimation of parameters for discretely observed diffusion processes with a variety of rates for information", *Annals of the Institute of Statistical Mathematics*, vol. 64, no. 2, pp. 545-575, 2012.

[25] O. Vasicek, "An equilibrium characterization of the term structure", *Journal of Financial Economics*. vol. 5, no. 2, pp. 177-186, 1977.

[26] Z. Wang, J. Lam, X. H. Liu. "Filtering for a class of nonlinear discrete-time stochastic systems with state delays", *Journal of Computational and Applied Mathematics*, vol. 201, no. 1, pp. 153-163, 2007.

[27] C. Wei, H. S. Shu, "Maximum likelihood estimation for the drift parameter in diffusion processes", *Stochastics: An International Journal of Probability and Stochastic Processes*, vol. 88, no. 5, pp. 699-710, 2016.

[28] C. Wei, H. S. Shu, Y. R. Liu, "Gaussian estimation for discretely observed Cox-Ingersoll-Ross model", *International Journal of General Systems*, vol. 45, no. 5, pp. 561-574, 2016.

[29] J. H. Wen, X. J. Wang, S. H. Mao, X. P. Xiao, "Maximum likelihood estimation of McKeanCVlasov stochastic differential equation and its application", *Applied Mathematics and Computation*, vol. 274, no. 1, pp. 237-246, 2015.

[30] J. Yu, P. C. B. Phillips. "A Gaussian approach for continuous time models of the short-term interest rate", *Econometrics Journal*, vol. 4, no. 1, pp. 210-224, 2001.