

9-2017

Time varying parameter estimation scheme for a linear stochastic differential equation

Olusegun Michael Otunuga
Marshall University, otunuga@marshall.edu

Follow this and additional works at: https://mds.marshall.edu/mathematics_faculty



Part of the [Dynamical Systems Commons](#)

Recommended Citation

Otunuga, Osegun, M. (2017). Time varying parameter estimation scheme for a linear stochastic differential equation. *International Journal of Statistics and Probability*, 6(5), 84-100. <https://doi.org/10.5539/ijsp.v6n5p84>

This Article is brought to you for free and open access by the Mathematics at Marshall Digital Scholar. It has been accepted for inclusion in Mathematics Faculty Research by an authorized administrator of Marshall Digital Scholar. For more information, please contact zhangj@marshall.edu, beachgr@marshall.edu.

Time Varying Parameter Estimation Scheme for a Linear Stochastic Differential Equation

Olusegun Michael Otunuga¹

¹ Department of Mathematics, Marshall University, Huntington, WV, USA

Correspondence: Olusegun Michael Otunuga, Department of Mathematics, One John Marshall Dr, Huntington, WV, 25755, USA. E-mail: otunuga@marshall.edu

Received: July 17, 2017 Accepted: July 31, 2017 Online Published: August 11, 2017

doi:10.5539/ijsp.v6n5p84 URL: <https://doi.org/10.5539/ijsp.v6n5p84>

Abstract

In this work, an attempt is made to estimate time varying parameters in a linear stochastic differential equation. By defining m_k as the local admissible sample/data observation size at time t_k , parameters and state at time t_k are estimated using past data on interval $[t_{k-m_k+1}, t_k]$. We show that the parameter estimates at each time t_k converge in probability to the true value of the parameters being estimated. A numerical simulation is presented by applying the local lagged adapted generalized method of moments (LLGMM) method to the stochastic differential models governing prices of energy commodities and stock price processes.

Keywords: Stochastic, Generalized Method of Moments, Maximum Likelihood, Simulation, Local lagged adapted

1. Introduction

In this work, we estimate the time varying parameters in a linear stochastic differential equation (SDE) in a systemic and unified way. We studied two special linear stochastic differential equations, namely the geometric brownian motion and the Ornstein-Uhlenbeck time varying stochastic differential equation. By drawing/recording real world time varying data at discrete-time on a time interval $[t_0, T]$, a stochastic numerical approximation scheme (Kloeden & Platen, 1992) is constructed to study the data. The future states of a continuous time dynamic process is assumed to be influenced by the history of past state. We also assume that response/reaction time delay process influence the present states (Ladde, 1979; Otunuga & Ladde, 2017). Utilizing Monte-Carlo method and the Euler-type (Kloeden & Platen, 1992) stochastic discretization scheme, we construct systems of local moments/observation equations based on the number of parameters present. In addition, using the method of moments (Casella & Berger, 2002) in the context of lagged adaptive expectation process (Paothong & Ladde, 2013), we describe theoretical parametric estimation procedure for the SDE.

The paper is organized as follows.

In Section 2, we describe the general form of linear stochastic differential equation used and give its closed form solution. In Section 3, we describe the theoretical parametric estimation procedure for a geometric stochastic differential equation with time varying parameters. We later show that the estimated parameters converge in probability to the true value of the parameter being estimated. In Section 4, we describe the theoretical parametric estimation procedure for a linear mean reverting stochastic differential equation with time varying parameters and also show the convergence in probability of the estimated parameters. In Section 5, numerical simulation is presented by applying the LLGMM method (Otunuga & Ladde, 2017) to stochastic model governing the stock prices for APPLE Inc., International Businesses Machines (IBM), JPMorgan Chase and Co. for the same period 01/03/2000 – 01/27/2017, (Apple; JPMorgan; IBM) and the energy commodities: Henry Hub natural gas price (dollars/million Btu), crude oil price (dollars/barrel), coal price for the period 01/04/2000–01/09/2017, 01/04/2000–01/09/2017 and 01/03/2000–01/11/2016, respectively, (Coal; Crude; Natural gas). We give a summary of the paper in Section 6.

2. Model Derivation

In this section, we describe the general linear SDE with constant and time-varying parameters. We consider the general linear SDE with time varying parameters of the form:

$$dx = (f(t)x + p(t))dt + (\sigma(t)x + q(t))dW(t), \quad x(t_0) = x_0, \quad (2.1)$$

where x is a state variable; f, p, σ, q are continuous functions defined on a given interval $[t_0, T]$, and $W(t)$ is a standard Wiener process defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$; the filtration $(\mathcal{F}_t)_{t \geq 0}$ is right-continuous. We will assume for the rest of this work that the solution process $x(t)$ is adapted and non-anticipating with respect to $(\mathcal{F}_t)_{t \geq 0}$.

Under these conditions, the closed form solution (interested readers are advised to see Ladde et.al (Ladde & Ladde, 2013)

for details on solution of stochastic differential equations) of (2.1) is given by

$$x(t) = \Phi(t) \left(x_0 + \int_{t_0}^t \Phi^{-1}(s) [p(s) - \sigma(s)q(s)] ds + \int_{t_0}^t \Phi^{-1}(s)q(s) dW(s) \right), \tag{2.2}$$

where $\Phi(t) = \exp \left[\int_{t_0}^t \left(f(s) - \frac{1}{2} \sigma^2(s) \right) ds + \int_{t_0}^t \sigma(s) dW(s) \right]$.

The general linear stochastic differential equation with constant parameters equivalent to (2.1) is described by:

$$dx = (\mu x + \gamma) dt + (\sigma x + \beta) dW(t), \quad x(t_0) = x_0, \tag{2.3}$$

where μ, γ, σ and $\beta \in \mathbb{R}$. The general solution of (2.3) is given by

$$x(t) = \Phi(t) \left(x_0 + [\gamma - \sigma\beta] \int_{t_0}^t \Phi^{-1}(s) ds + \beta \int_{t_0}^t \Phi^{-1}(s) dW(s) \right), \tag{2.4}$$

where $\Phi(t) = \exp \left[\left(\mu - \frac{1}{2} \sigma^2 \right) (t - t_0) + \sigma (W(t) - W(t_0)) \right]$.

We study two special cases of (2.1). For the first case, we set $f(t) = a(t)$, $p(t) = 0$ and $q(t) = 0$. For the second case, we set $f(t) = -\alpha(t)$, $p(t) = \alpha(t)\mu(t)$, $\sigma(t) = 0$ and $q(t) = \sigma(t)$. We note here that the first case is the generalization of the geometric stochastic differential equation with time varying parameters. The second case is the generalization of the Ornstein-Uhlenbeck stochastic differential equation with time varying parameters. From (2.2), it follows that x is normally distributed if $\sigma(t) = 0$ and lognormally distributed if $p(t) = q(t) = 0$.

3. Theoretical Parametric Estimation Procedure: Case 1

In this section, we study the case of (2.1) of the form

$$dx = a(t)x dt + \sigma(t)x dW(t), \quad x(t_0) = x_0, \tag{3.1}$$

where the parameters $a(t)$ and $\sigma(t)$ are unknown for each time t , and are not assumed constant over t . We further assume that for a fixed time T , the functions $a : [t_0, T] \rightarrow \mathbb{R}$ and $\sigma : [t_0, T] \rightarrow \mathbb{R}_+$ are continuous and bounded, \mathbb{R}_+ is the set of positive real numbers. As noted earlier, (3.1) is the generalization of the geometric stochastic differential equation with time varying parameters. It's equivalent version with constant parameters is given by

$$dx = ax dt + \sigma x dW(t), \quad x(t_0) = x_0, \tag{3.2}$$

where $\alpha \in \mathbb{R}$ and $\sigma > 0$.

For any given time $\mathcal{T} \in (t_0, T]$, let $t_0 < t_1 < \dots < t_n = \mathcal{T}$ be a non-random partition P of the interval $[t_0, \mathcal{T}]$ such that $t_i = t_0 + i\Delta t$, $i = 0, 1, \dots, n$, and $\Delta t = \frac{\mathcal{T}-t_0}{n}$.

By the classical existence theorem, the solutions $\mathbf{x}(t)$ and $x(t)$ (see (2.2) and (2.4) for the solutions) satisfying (3.1) and (3.2), respectively, exist and are continuous.

3.1 LLGMM Parameter Estimation Procedure for (3.1)

In the following, we outline the procedure for estimating the time varying parameters $a(t)$ and $\sigma(t)$ in (3.1).

3.1.1 Transformation of Stochastic Differential Equations

Due to the fact that the drift coefficient of (3.1) has only one parameter $a(t)$, we consider the process $V(t, x) = \ln x$. The Ito-Doob stochastic differential for V satisfy

$$d(\ln x) = \left(a(t) - \frac{\sigma^2(t)}{2} \right) dt + \sigma(t) dW(t), \quad \ln x(t_0) = \ln x_0. \tag{3.3}$$

The solution of (3.3) in the interval $[t_{n-1}, t_n] = [\mathcal{T} - \Delta t, \mathcal{T}]$ satisfies

$$\ln \left[\frac{x(t_n)}{x(t_{n-1})} \right] = \int_{t_{n-1}}^{t_n} \left(a(s) - \frac{\sigma^2(s)}{2} \right) ds + \int_{t_{n-1}}^{t_n} \sigma(s) dW(s). \tag{3.4}$$

3.1.2 Generalized Moment Equations

Define $\mathcal{F}_{t_{n-1}} \equiv \mathcal{F}_{n-1}$ as the filtration process up to time t_{n-1} . Applying conditional expectations to $\ln \left(\frac{x(t_n)}{x(t_{n-1})} \right)$ in (3.4) with respect to \mathcal{F}_{n-1} , we obtain

$$\begin{aligned} \mathbb{E} \left[\ln \left(\frac{x(t_n)}{x(t_{n-1})} \right) | \mathcal{F}_{n-1} \right] &= \int_{t_{n-1}}^{t_n} \left(a(s) - \frac{\sigma^2(s)}{2} \right) ds, \\ \mathbb{E} \left[\left(\ln \left(\frac{x(t_n)}{x(t_{n-1})} \right) - \mathbb{E} \left[\ln \left(\frac{x(t_n)}{x(t_{n-1})} \right) | \mathcal{F}_{n-1} \right] \right)^2 | \mathcal{F}_{n-1} \right] &= \int_{t_{n-1}}^{t_n} \sigma^2(s) ds. \end{aligned} \tag{3.5}$$

3.1.3 Basis for Lagged Adaptive Discrete-time Expectation Process

From (3.5), (3.4) reduces to

$$\ln\left(\frac{x(t_n)}{x(t_{n-1})}\right) = \mathbb{E}\left[\ln\left(\frac{x(t_n)}{x(t_{n-1})}\right) \middle| \mathcal{F}_{n-1}\right] + \int_{t_{n-1}}^{t_n} \sigma(s) dW(s). \tag{3.6}$$

Equation (3.6) provides the basis for the development of lagged adaptive expectation process (Otunuga & Ladde, 2017; Paothong & Ladde, 2013) with respect to the stochastic dynamic systems (3.3).

3.1.4 Parameter Estimates at Time $t_n = \mathcal{T}$

Using the fact that $\lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \int_{t_{n-1}}^{t_n} (\mathbf{a}(s) - \frac{\sigma^2(s)}{2}) ds = \mathbf{a}(\mathcal{T}) - \frac{\sigma^2(\mathcal{T})}{2}$ and $\lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \int_{t_{n-1}}^{t_n} \sigma^2(s) ds = \sigma^2(\mathcal{T})$, we have

$$\begin{aligned} \mathbf{a}(t_n) &= \mathbf{a}(\mathcal{T}) = \lim_{\Delta t \rightarrow 0} \left(\frac{1}{\Delta t} \left[\mathbb{E}\left(\ln\left[\frac{x(t_n)}{x(t_{n-1})}\right) \middle| \mathcal{F}_{n-1}\right) + \frac{1}{2} \mathbb{E}\left(\left[\ln\left[\frac{x(t_n)}{x(t_{n-1})}\right] - \mathbb{E}\left(\ln\left[\frac{x(t_n)}{x(t_{n-1})}\right) \middle| \mathcal{F}_{n-1}\right]\right)^2 \middle| \mathcal{F}_{n-1}\right) \right] \right), \\ \sigma^2(t_n) &= \sigma^2(\mathcal{T}) = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \mathbb{E}\left(\left[\ln\left[\frac{x(t_n)}{x(t_{n-1})}\right] - \mathbb{E}\left(\ln\left[\frac{x(t_n)}{x(t_{n-1})}\right) \middle| \mathcal{F}_{n-1}\right]\right)^2 \middle| \mathcal{F}_{n-1}\right). \end{aligned} \tag{3.7}$$

To get a discretize estimate for $\mathbf{a}(t_n)$ and $\sigma^2(t_n)$ in (3.7), we use Monte-Carlo method and Euler scheme to discretize (3.3) on the interval $[t_0, \mathcal{T}]$ as follows:

$$\ln\left[\frac{x_{i,j}}{x_{i-1,j}}\right] = \left(\mathbf{a}(t_{i-1}) - \frac{\sigma^2(t_{i-1})}{2}\right) \Delta t + \sigma(t_{i-1}) \Delta W_{i,j}, \quad i = 1, 2, \dots, n, \quad j = 1, 2, \dots, N, \tag{3.8}$$

where $(x_{i,j})_{j \in \mathbb{N}}$ denotes sequence of independent copies of $x_i \equiv x(t_i)$ at time $t_i, i = 1, 2, \dots, n, \Delta W_{i,j} \sim N(0, \Delta t)$. Let $\hat{\mathbf{a}}_{n,N}$ and $\hat{\sigma}_{n,N}^2$ be the estimates of $\mathbf{a}(\mathcal{T})$ and $\sigma^2(\mathcal{T})$, respectively, at time $t_n = \mathcal{T}$. We derive $\hat{\mathbf{a}}_{n,N}$ and $\hat{\sigma}_{n,N}^2$ from (3.7) as follows:

$$\begin{cases} \hat{\mathbf{a}}_{n,N} &= \frac{1}{N\Delta t} \left[\sum_{j=1}^N \ln\left(\frac{x_{n,j}}{x_{n-1,j}}\right) + \frac{1}{2} \sum_{j=1}^N \left(\ln\left(\frac{x_{n,j}}{x_{n-1,j}}\right) - \frac{1}{N} \sum_{j=1}^N \ln\left(\frac{x_{n,j}}{x_{n-1,j}}\right) \right)^2 \right], \\ \hat{\sigma}_{n,N}^2 &= \frac{1}{N\Delta t} \sum_{j=1}^N \left(\ln\left(\frac{x_{n,j}}{x_{n-1,j}}\right) - \frac{1}{N} \sum_{j=1}^N \ln\left(\frac{x_{n,j}}{x_{n-1,j}}\right) \right)^2. \end{cases} \tag{3.9}$$

3.2 Consistency

We shall show that $\hat{\mathbf{a}}_{n,N} \xrightarrow{P} \mathbf{a}(\mathcal{T})$ and $\hat{\sigma}_{n,N}^2 \xrightarrow{P} \sigma^2(\mathcal{T})$ as $n \rightarrow \infty, N \rightarrow \infty$.

Lemma 1. The random variable $\frac{1}{N\Delta t} \sum_{j=1}^N \Delta W_{n,j} \xrightarrow{P} 0$, and $\frac{1}{N\Delta t} \sum_{j=1}^N \Delta W_{n,j}^2 \xrightarrow{P} 1$ as $n \rightarrow \infty, N \rightarrow \infty$.

Proof. For any $\epsilon > 0$,

$$\begin{aligned} \mathbb{P}\left[\left|\frac{1}{N\Delta t} \sum_{j=1}^N \Delta W_{n,j}\right| \geq \epsilon\right] &< \frac{1}{N^2 \Delta t^2 \epsilon^2} \mathbb{E}\left[\left(\sum_{j=1}^N \Delta W_{n,j}\right)^2\right] = \frac{1}{N\Delta t \epsilon^2} \rightarrow 0 \text{ as } N \rightarrow \infty, \\ \mathbb{P}\left[\left|\frac{1}{N\Delta t} \sum_{j=1}^N \Delta W_{n,j}^2 - 1\right| \geq \epsilon\right] &< \frac{1}{\epsilon^2} \mathbb{E}\left[\left(\frac{1}{N\Delta t} \sum_{j=1}^N \Delta W_{n,j}^2 - 1\right)^2\right] = \frac{2}{N\epsilon^2} \rightarrow 0 \text{ as } N \rightarrow \infty, \end{aligned}$$

where $N\Delta t \rightarrow \infty$ as $N \rightarrow \infty$. ■

Theorem 2. Assume that $\mathbf{a}(t)$ and $\sigma(t)$ are continuous and bounded functions. Then $\hat{\mathbf{a}}_{n,N} \xrightarrow{P} \mathbf{a}(\mathcal{T})$ and $\hat{\sigma}_{n,N}^2 \xrightarrow{P} \sigma^2(\mathcal{T})$ as $N \rightarrow \infty$.

Proof. From (3.8), (3.9) and Lemma 1, it follows directly that

$$\begin{cases} \hat{\mathbf{a}}_{n,N} &\xrightarrow{P} \mathbf{a}(\mathcal{T}) - \frac{\sigma^2(\mathcal{T})}{2} + \frac{\sigma^2(\mathcal{T})}{2} = \mathbf{a}(\mathcal{T}), \\ \hat{\sigma}_{n,N}^2 &\xrightarrow{P} \sigma^2(\mathcal{T}), \end{cases} \tag{3.10}$$

as $n \rightarrow \infty, N \rightarrow \infty$. ■

3.3 LLGMM Parameter Estimation Scheme

Define $\bar{x}_i = \frac{1}{N} \sum_{j=1}^N x_{i,j}$, $i = 1, 2, \dots, n$ and $I_a^b = \{i \in \mathbb{Z}_{++} : a \leq i \leq b\}$, where \mathbb{Z}_{++} is the set of non negative integers. Consider a continuous time stochastic dynamic process defined on the interval $[-\tau, T]$ into \mathbb{R}_{++} , where τ characterizes the influence of past history of the state of the dynamics. We define the discrete version of time delay τ as $r = \left\lfloor \frac{\tau}{\Delta t} \right\rfloor$ and also define m_n as the local admissible sample/data observation size at time t_n . We note here that $m_n \in I_2^n$. For each sample observation size $m_n \in I_2^n$, a partition P_n of closed interval $[t_{n-m_n+1}, t_n]$ of length m_n is called local at time t_n and defined by

$$P_n := t_{n-m_n+1} < t_{n-m_n+2} < \dots < t_n. \tag{3.11}$$

Using (3.5 – 3.7), we formulate a local observation process at time t_n as an algebraic functions of m_n - local functions by restricting the overall finite sample sequence $\{x_i\}_{i=-r}^n$ to a subpartition P_n in (3.11).

Using the LLGMM method by Otunuga et.al (Otunuga & Ladde, 2017), let $\hat{\mathbf{a}}_{n,m_n}$ and $\hat{\sigma}_{n,m_n}^2$ be the moving estimates of $\mathbf{a}(t_n)$ and $\sigma^2(t_n)$, respectively, at time t_n using observation size m_n . We derive $\hat{\mathbf{a}}_{n,m_n}$ and $\hat{\sigma}_{n,m_n}^2$ from (3.7) as follows:

$$\begin{cases} \hat{\mathbf{a}}_{n,m_n} &= \frac{1}{m_n \Delta t} \left[\sum_{i=n-m_n+1}^n \ln \left(\frac{\bar{x}_i}{\bar{x}_{i-1}} \right) + \frac{1}{2} \sum_{i=n-m_n+1}^n \left(\ln \left(\frac{\bar{x}_i}{\bar{x}_{i-1}} \right) - \frac{1}{m_n} \sum_{i=n-m_n+1}^n \ln \left(\frac{\bar{x}_i}{\bar{x}_{i-1}} \right) \right)^2 \right], \\ \hat{\sigma}_{n,m_n}^2 &= \frac{1}{m_n \Delta t} \sum_{i=n-m_n+1}^n \left(\ln \left(\frac{\bar{x}_i}{\bar{x}_{i-1}} \right) - \frac{1}{m_n} \sum_{i=n-m_n+1}^n \ln \left(\frac{\bar{x}_i}{\bar{x}_{i-1}} \right) \right)^2. \end{cases} \tag{3.12}$$

Remark 1. We note here that at each time $t = t_k$, the local admissible observation size m_k is not constant. For a given $\epsilon > 0$, the value of m_k that gives the least simulated error at each time t_k is recorded as the ϵ -best sub-optimal sample size \hat{m}_k and the parameters $\hat{\mathbf{a}}_{k,m_k}$ and $\hat{\sigma}_{k,m_k}^2$ recorded as \mathbf{a}_{k,\hat{m}_k} and σ_{k,\hat{m}_k}^2 , respectively. The method of estimating \hat{m}_k is discussed in Section 5.

Remark 2. For random variables x_1, x_2, \dots, x_n , we compare the maximum likelihood estimates (MLE) \hat{a}_n and $\hat{\sigma}_n$ of the parameters a and σ in model with constant parameters (3.2) with the parameter estimates $\hat{\mathbf{a}}_{n,N}$ and $\hat{\sigma}_{n,N}$ in (3.9). The MLE estimates \hat{a}_N and $\hat{\sigma}_N$ are given by

$$\begin{aligned} \hat{a}_n &= \frac{1}{n \Delta t} \left[\sum_{i=1}^n \ln \left(\frac{x_i}{x_{i-1}} \right) + \frac{1}{2} \sum_{i=1}^n \left(\ln \left(\frac{x_i}{x_{i-1}} \right) - \frac{1}{n} \sum_{i=1}^n \ln \left(\frac{x_i}{x_{i-1}} \right) \right)^2 \right], \\ \hat{\sigma}_n^2 &= \frac{1}{n \Delta t} \sum_{i=1}^n \left(\ln \left(\frac{x_i}{x_{i-1}} \right) - \frac{1}{n} \sum_{i=1}^n \ln \left(\frac{x_i}{x_{i-1}} \right) \right)^2. \end{aligned} \tag{3.13}$$

We note here that the estimates in (3.12) and (3.13) are similar when $m_n = n$. Also, the difference in (3.9) and (3.13) is that (3.13) contains summation over functions of the random variables x_1, x_2, \dots, x_n while (3.9) contains summation over sequence of copies $x_{n,1}, x_{n,2}, \dots, x_{n,N}$ of x_n .

4. Theoretical Parametric Estimation Procedure: Case 2

In this section, we study the case of (2.1) of the form

$$dx = \alpha(t) (\mu(t) - x) dt + \sigma(t) dW(t), \quad x(t_0) = x_0, \tag{4.1}$$

where the parameters $\alpha(t)$, $\mu(t)$, and $\sigma(t)$ are unknown for each time t , and are not assumed constant over t . We further assume that for a fixed time T , the functions $\alpha : [t_0, T] \rightarrow \mathbb{R}_+$, $\mu : [t_0, T] \rightarrow \mathbb{R}_{++}$ and $\sigma : [t_0, T] \rightarrow \mathbb{R}_+$ are continuous and bounded.

We note here that (4.1) is the generalization of the Ornstein-Uhlenbeck stochastic differential equation with time varying parameters. It's equivalent version with constant parameters is

$$dx = \alpha (\mu - x) dt + \sigma dW(t), \quad x(t_0) = x_0, \tag{4.2}$$

where $\mu \geq 0$, $\alpha > 0$, and $\sigma > 0$.

The solution $\mathbf{x}(t)$ satisfying (4.1) is given by

$$x(t) = x_0 e^{-\int_0^t \alpha(s) ds} + \int_0^t e^{-\int_s^t \alpha(u) du} \alpha(s) \mu(s) ds + \int_0^t e^{-\int_s^t \alpha(u) du} \sigma(s) dW_s. \tag{4.3}$$

It follows that $\mathbb{E}[x(t)] = \mathbb{E}[x_0] e^{-\int_0^t \alpha(s) ds} + \int_0^t e^{-\int_s^t \alpha(u) du} \alpha(s) \mu(s) ds$.

4.1 Parameter Estimation Procedure for (4.1)

We outline the procedure for estimating time varying parameters $\alpha(t)$, $\mu(t)$ and $\sigma(t)$ in (4.1) at any given time $\mathcal{T} \in (t_0, T]$.

4.1.1 Transformation of Stochastic Differential Equations

Due to the fact that the drift coefficient of (4.1) has two parameters $\alpha(t)$ and $\mu(t)$, we consider the processes $V_1(t, x) = x$ and $V_2(t, x) = x^2$. The Ito-Doob stochastic differential equations for V_1 and V_2 satisfy

$$\begin{cases} dx &= \alpha(t)(\mu(t) - x) dt + \sigma(t) dW(t), \quad x(t_0) = x_0, \\ d(x^2) &= (2\alpha(t)(\mu(t) - x)x + \sigma^2(t)) dt + 2\sigma(t)x dW(t), \quad x^2(t_0) = x_0^2. \end{cases} \quad (4.4)$$

4.1.2 Euler-type Discretization Scheme for (4.4)

For any given time $\mathcal{T} \in (t_0, T]$, let $t_0 < t_1 < \dots < t_n = \mathcal{T}$ be a non-random partition P of the interval $[t_0, \mathcal{T}]$ such that $t_i = t_0 + i\Delta t$, $i = 0, 1, \dots, n$, where $\Delta t = \frac{\mathcal{T}-t_0}{n}$, and n is the sample size. The solution of (4.4) on the interval $[t_{n-1}, t_n]$ satisfies

$$\begin{aligned} \Delta x_n &= \int_{t_{n-1}}^{t_n} \alpha(t)(\mu(t) - x(t)) dt + \int_{t_{n-1}}^{t_n} \sigma(t) dW(t), \\ \Delta(x_n^2) &= \int_{t_{n-1}}^{t_n} (2\alpha(t)(\mu(t) - x(t))x(t) + \sigma^2(t)) dt + 2 \int_{t_{n-1}}^{t_n} \sigma(t)x(t) dW(t), \end{aligned} \quad (4.5)$$

where $\Delta x_n = x(t_n) - x(t_{n-1})$. Note that $\Delta(x_n^2) = \Delta x_n(x_{n-1} + x_n) = 2(\Delta x_n)x_{n-1} + (\Delta x_n)^2$. From (4.5) and Ito-Doob scheme, it follows that $\Delta(x_n^2) = \int_{t_{n-1}}^{t_n} \sigma^2(t) dt$. Using this and the fact that $\Delta(x_n^2) = 2(\Delta x_n)x_{n-1} + (\Delta x_n)^2$, we reduce (4.5) to the following:

$$\begin{cases} \Delta x_n &= \int_{t_{n-1}}^{t_n} \alpha(t)(\mu(t) - x(t)) dt + \int_{t_{n-1}}^{t_n} \sigma(t) dW(t), \\ (\Delta x_n)x_{n-1} &= \int_{t_{n-1}}^{t_n} \alpha(t)x(t)(\mu(t) - x(t)) dt + \int_{t_{n-1}}^{t_n} \sigma(t)x(t) dW(t). \end{cases} \quad (4.6)$$

4.1.3 Generalized Moment Equations

Applying expectations to (4.6), we obtain

$$\begin{cases} \mathbb{E}[\Delta x_n] &= \int_{t_{n-1}}^{t_n} \alpha(t)(\mu(t) - \mathbb{E}[x(t)]) dt, \\ \mathbb{E}[(\Delta x_n)x_{n-1}] &= \int_{t_{n-1}}^{t_n} \alpha(t)\mu(t)\mathbb{E}[x(t)] dt - \int_{t_{n-1}}^{t_n} \alpha(t)\mathbb{E}[x(t)^2] dt \\ \mathbb{E}[(\Delta x_n - \mathbb{E}[\Delta x_n])^2] &= \int_{t_{n-1}}^{t_n} \sigma^2(t) dt. \end{cases} \quad (4.7)$$

4.1.4 Parameter Estimates at Time $t_n = \mathcal{T}$

By dividing (4.7) by Δt and taking the limit as $\Delta t \rightarrow 0$, we have

$$\begin{cases} \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \mathbb{E}[\Delta x_n] &= \alpha(t_n)(\mu(t_n) - \mathbb{E}[x(t_n)]), \\ \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \mathbb{E}[(\Delta x_n)x_{n-1}] &= \alpha(t_n)\mu(t_n)\mathbb{E}[x(t_n)] - \alpha(t_n)\mathbb{E}[x(t_n)^2], \\ \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \mathbb{E}[(\Delta x_n - \mathbb{E}[\Delta x_n])^2] &= \sigma^2(t_n). \end{cases} \quad (4.8)$$

The solutions $\alpha(t_n)$, $\mu(t_n)$ and $\sigma^2(t_n)$ satisfying (4.8) are given by

$$\begin{cases} \alpha(t_n) &= \frac{\lim_{\Delta t \rightarrow 0} (\frac{1}{\Delta t} \mathbb{E}[\Delta x_n])\mathbb{E}[x_n] - \lim_{\Delta t \rightarrow 0} (\frac{1}{\Delta t} \mathbb{E}[(\Delta x_n)x_{n-1}])}{\mathbb{E}[x_n - \mathbb{E}[x_n]]^2}, \\ \mu(t_n) &= \mathbb{E}[x(t_n)] + \frac{1}{\alpha(t_n)} \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \mathbb{E}[\Delta x_n], \\ \sigma^2(t_n) &= \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \mathbb{E}[(\Delta x_n - \mathbb{E}[\Delta x_n])^2]. \end{cases} \quad (4.9)$$

To get a discretize estimate for $\alpha(t_n)$, $\mu(t_n)$ and $\sigma^2(t_n)$ in (4.9), we use Monte-Carlo method and Euler scheme to discretize (4.1) on the interval $[t_0, \mathcal{T}]$ as follows:

$$\Delta x_{i,j} = \alpha_{i-1}(\mu_{i-1} - x_{i-1,j})\Delta t + \sigma_{i-1}\Delta W_{i,j}, \quad i = 1, 2, \dots, n, \quad j = 1, 2, \dots, N, \quad (4.10)$$

where $(x_{i,j})_{j \in \mathbb{N}}$ denotes sequence of independent copies of $x_i \equiv x(t_i)$ at time t_i , $i = 1, 2, \dots, n$, $\alpha_i = \alpha(t_i)$, $\mu_i = \mu(t_i)$ and $\sigma_i = \sigma(t_i)$.

Let $\hat{\alpha}_{n,N}$, $\hat{\mu}_{n,N}$ and $\hat{\sigma}_{n,N}^2$ be the estimates of $\alpha(\mathcal{T})$, $\mu(\mathcal{T})$ and $\sigma^2(\mathcal{T})$, respectively, at time $t_n = \mathcal{T}$. We derive $\hat{\alpha}_{n,N}$, $\hat{\mu}_{n,N}$ and $\hat{\sigma}_{n,N}^2$ from (4.9) as follows:

$$\left\{ \begin{aligned} \hat{\alpha}_{n,N} &= \frac{\frac{1}{N} \sum_{j=1}^N \Delta x_{n,j} \sum_{j=1}^N x_{n-1,j} - \sum_{j=1}^N (\Delta x_{n,j}) x_{n-1,j}}{\left[\sum_{j=1}^N x_{n-1,j}^2 - \frac{1}{N} \left(\sum_{j=1}^N x_{n-1,j} \right)^2 \right] \Delta t}, \\ \hat{\mu}_{n,N} &= \frac{1}{N} \sum_{j=1}^N x_{n-1,j} + \frac{\frac{1}{N} \sum_{j=1}^N \Delta x_{n,j}}{\hat{\alpha}_{n,N} \Delta t}, \\ \hat{\sigma}_{n,N}^2 &= \frac{1}{\Delta t} \cdot \frac{1}{N} \sum_{j=1}^N \left(\Delta x_{n,j} - \frac{1}{N} \sum_{j=1}^N \Delta x_{n,j} \right)^2, \end{aligned} \right. \tag{4.11}$$

where $\Delta x_{n,j} = x_{n,j} - x_{n-1,j}$.

Remark 3. The solution $x_{n,j}$, $j = 1, 2, \dots, N$, satisfying (4.10) is given by

$$\begin{aligned} x_{n,j} &= \prod_{k=0}^{n-1} (1 - \alpha_k \Delta t) x_0 + \sum_{i=0}^{n-2} \left(\prod_{k=i+1}^{n-1} (1 - \alpha_k \Delta t) \right) \alpha_i \mu_i \Delta t + \alpha_{n-1} \mu_{n-1} \Delta t \\ &+ \sum_{i=0}^{n-2} \left(\prod_{k=i+1}^{n-1} (1 - \alpha_k \Delta t) \right) \sigma_i \Delta W_{i+1,j} + \sigma_{n-1} \Delta W_{n,j}. \end{aligned} \tag{4.12}$$

By writing $\prod_{k=0}^{n-1} (1 - \alpha_k \Delta t) = \exp\left(-\sum_{k=0}^{n-1} \sum_{m=1}^{\infty} \frac{(\alpha_k \Delta t)^m}{m}\right) = \exp\left(-\sum_{k=0}^{n-1} \alpha_k \Delta t\right) \exp\left(-\sum_{k=0}^{n-1} \sum_{m=2}^{\infty} \frac{(\alpha_k \Delta t)^m}{m}\right)$, it follows that as $n \rightarrow \infty$, $\prod_{k=0}^{n-1} (1 - \alpha_k \Delta t) \rightarrow e^{-\int_0^{\mathcal{T}} \alpha(s) ds}$ and $x_{n,j} \rightarrow x_{\mathcal{T},j}$, where

$$x_{\mathcal{T},j} = e^{-\int_0^{\mathcal{T}} \alpha(s) ds} x_0 + \int_{t_0}^{\mathcal{T}} e^{-\int_s^{\mathcal{T}} \alpha(u) du} \alpha(s) \mu(s) ds + \int_{t_0}^{\mathcal{T}} e^{-\int_s^{\mathcal{T}} \alpha(u) du} \sigma(s) dW_{s,j}. \tag{4.13}$$

4.2 Consistency

Lemma 3. Assume that $\mathbb{E}[x_0] < \infty$. Define $A_{1,\mathcal{T}} = e^{-\int_0^{\mathcal{T}} \alpha(s) ds} \mathbb{E}(x_0)$; $A_{2,\mathcal{T}} = \int_{t_0}^{\mathcal{T}} e^{-\int_s^{\mathcal{T}} \alpha(u) du} \alpha(s) \mu(s) ds$ and $V_{\mathcal{T}} = \int_{t_0}^{\mathcal{T}} e^{-2\int_s^{\mathcal{T}} \alpha(u) du} \sigma^2(s) ds$. Then

1. $\frac{1}{N} \sum_{j=1}^N x_{n,j} \xrightarrow{P} A_{1,\mathcal{T}} + A_{2,\mathcal{T}} = \mathbb{E}[x(\mathcal{T})]$,
2. $\frac{1}{N} \sum_{j=1}^N x_{n,j}^2 \xrightarrow{P} (A_{1,\mathcal{T}} + A_{2,\mathcal{T}})^2 + V_{\mathcal{T}} = \mathbb{E}[x^2(\mathcal{T})]$,
3. $\frac{1}{N\Delta t} \sum_{j=1}^N \Delta x_{n,j} \xrightarrow{P} \alpha(\mathcal{T}) (\mu(\mathcal{T}) - \mathbb{E}[x(\mathcal{T})])$,
4. $\frac{1}{N\Delta t} \sum_{j=1}^N \Delta x_{n,j} x_{n-1,j} \xrightarrow{P} \alpha(\mathcal{T}) (\mu(\mathcal{T}) \mathbb{E}[x(\mathcal{T})] - \mathbb{E}[x^2(\mathcal{T})])$, as $n \rightarrow \infty$ and $N \rightarrow \infty$.

Proof. It follows directly from (4.13) that $\lim_{n \rightarrow \infty} \mathbb{E} \left[\frac{1}{N} \sum_{j=1}^N x_{n-1,j} \right] = \mathbb{E}[x(\mathcal{T})]$ and $\lim_{n \rightarrow \infty} \mathbb{E} \left[\frac{1}{N\Delta t} \sum_{j=1}^N \Delta(x_{n,j}) \right] = \alpha(\mathcal{T}) (\mu(\mathcal{T}) - \mathbb{E}[x(\mathcal{T})])$.

Define $U_{n,j} = \sum_{i=0}^{n-2} \left(\prod_{k=i+1}^{n-1} (1 - \alpha_k \Delta t) \right) \sigma_i \Delta W_{i+1,j} + \sigma_{n-1} \Delta W_{n,j}$, $j = 1, 2, \dots, N$. For $\epsilon > 0$,

$$\begin{aligned} \mathbb{P} \left[\left| \frac{1}{N} \sum_{j=1}^N x_{n,j} - \mathbb{E}[x(\mathcal{T})] \right| \geq \epsilon \right] &< \frac{1}{\epsilon^2} \left[2\mathbb{E}(B_{1,n} - A_{1,\mathcal{T}})^2 + 2\mathbb{E}(B_{2,n} - A_{2,\mathcal{T}})^2 + \mathbb{E}(Z_{1,N}^2) \right], \\ \mathbb{P} \left[\left| \frac{1}{N\Delta t} \sum_{j=1}^N \Delta x_{n,j} - \alpha(\mathcal{T})(\mu(\mathcal{T}) - \mathbb{E}[x(\mathcal{T})]) \right| \geq \epsilon \right] &< \frac{1}{\epsilon^2} \left[\mathbb{E}(C_{1,n} + C_{2,n} - \alpha(\mathcal{T})(\mu(\mathcal{T}) - \mathbb{E}[x(\mathcal{T})]))^2 + \mathbb{E}(Z_{2,N}^2) \right], \\ \mathbb{P} \left[\left| \frac{1}{N} \sum_{j=1}^N U_{n,j}^2 - V_{\mathcal{T}} \right| \geq \epsilon \right] &< \frac{1}{\epsilon^2} \mathbb{E} \left[\frac{1}{N} \sum_{j=1}^N \left(\sum_{i=0}^{n-2} \prod_{k=i+1}^{n-1} (1 - \alpha_k \Delta t) \sigma_i \Delta W_{i+1,j} \right)^2 - V_{\mathcal{T}} \right. \\ &+ \frac{2}{N} \sum_{j=1}^N \sigma_{n-1} \sum_{i=0}^{n-2} \prod_{k=i+1}^{n-1} (1 - \alpha_k \Delta t) \sigma_i \Delta W_{i+1,j} \Delta W_{n,j} \\ &\left. + \frac{1}{N} \sum_{j=1}^N \sigma_{n-1}^2 \Delta W_{n,j}^2 \right], \end{aligned}$$

where

$$\begin{aligned} B_{1,n} &= \prod_{k=0}^{n-1} (1 - \alpha_k \Delta t) x_0; & B_{2,n} &= \sum_{i=0}^{n-2} \left(\prod_{k=i+1}^{n-1} (1 - \alpha_k \Delta t) \right) \alpha_i \mu_i \Delta t + \alpha_{n-1} \mu_{n-1} \Delta t; \\ C_{1,n} &= -\alpha_{n-1} \prod_{k=0}^{n-2} (1 - \alpha_k \Delta t) x_0; & C_{2,n} &= -\alpha_{n-1} \sum_{i=0}^{n-3} \left(\prod_{k=i+1}^{n-2} (1 - \alpha_k \Delta t) \right) \alpha_i \mu_i \Delta t - \alpha_{n-1} \alpha_{n-2} \mu_{n-1} \Delta t + \alpha_{n-1} \mu_{n-1}; \\ D_{1,n} &= \sum_{i=0}^{n-2} \left(\prod_{k=i+1}^{n-1} (1 - \alpha_k \Delta t)^2 \right) \sigma_i^2 \Delta t; & Z_{1,N} &= \frac{1}{N} \sum_{j=1}^N U_{n,j}; & Z_{2,N} &= -\frac{1}{N} \sum_{j=1}^N [U_{n-1,j} \alpha_{n-1} - \sigma_{n-1} \Delta W_{n,j} / \Delta t]. \end{aligned}$$

It follows that $B_{1,n} \xrightarrow{p} A_{1,\mathcal{T}}$, $B_{2,n} \xrightarrow{p} A_{2,\mathcal{T}}$ and $D_{1,n} \xrightarrow{p} V_{\mathcal{T}}$ as $n \rightarrow \infty$. Also,

$$C_{1,n} + C_{2,n} \xrightarrow{p} \alpha(\mathcal{T})\mu(\mathcal{T}) - \alpha(\mathcal{T}) \left[e^{-\int_0^{\mathcal{T}} \alpha(s) ds} x_0 + \int_0^{\mathcal{T}} e^{-\int_s^{\mathcal{T}} \alpha_u du} \alpha(s)\mu(s) ds \right] = \alpha(\mathcal{T})\mu(\mathcal{T}) - \alpha(\mathcal{T})\mathbb{E}[x(\mathcal{T})] \text{ as } n \rightarrow \infty.$$

Furthermore, $\mathbb{E}[Z_{1,N}^2] = \frac{1}{N^2} \sum_{j=1}^N \sum_{i=0}^{n-2} \left(\prod_{k=i+1}^{n-1} (1 - \alpha_k \Delta t)^2 \right) \sigma_i^2 \Delta t + \frac{\sigma_{n-1}^2}{N}$ and $\mathbb{E}[Z_{2,N}^2] = \frac{1}{N^2} \sum_{j=1}^N \sum_{i=0}^{n-3} \left(\prod_{k=i+1}^{n-2} (1 - \alpha_k \Delta t)^2 \right) \sigma_i^2 \alpha_{n-1}^2 \Delta t + \alpha_{n-1}^2 \sigma_{n-2}^2 \frac{\Delta t}{N} + \sigma_{n-1}^2 / (N\Delta t)$ both tend to zero as $n \rightarrow \infty$, $N \rightarrow \infty$ and $N\Delta t \rightarrow \infty$. Hence, $\mathbb{P} \left[\left| \frac{1}{N} \sum_{j=1}^N x_{n,j} - \mathbb{E}[x(\mathcal{T})] \right| \geq \epsilon \right] \rightarrow 0$,

$$\mathbb{P} \left[\left| \frac{1}{N\Delta t} \sum_{j=1}^N \Delta x_{n,j} - \alpha(\mathcal{T})(\mu(\mathcal{T}) - \mathbb{E}[x(\mathcal{T})]) \right| \geq \epsilon \right] \rightarrow 0 \text{ and}$$

$$\begin{aligned} \mathbb{P} \left[\left| \frac{1}{N} \sum_{j=1}^N U_{n,j}^2 - V_{\mathcal{T}} \right| \geq \epsilon \right] &< \frac{1}{\epsilon^2 N^2} \left[D_{1,n} (3N\Delta t + (N^2 - N)D_{1,n}\Delta t - 2V_{\mathcal{T}} + 2\Delta t + 4\sigma_{n-1}^2 \Delta t / N) + V_{\mathcal{T}}^2 - 2\sigma_{n-1}^2 V_{\mathcal{T}} \Delta t \right. \\ &\left. + \left(1 + \frac{2}{N} \right) \sigma_{n-1}^2 \Delta t^2 \right] \end{aligned}$$

tends to zero as $n \rightarrow \infty$, $N \rightarrow \infty$. Finally,

$$\begin{aligned} \frac{1}{N} \sum_{j=1}^N x_{n,j}^2 &= (B_{1,n} + B_{2,n})^2 + \frac{2}{N} \sum_{j=1}^N (B_{1,n} + B_{2,n}) U_{n,j} + \frac{1}{N} \sum_{j=1}^N U_{n,j}^2 \\ &\xrightarrow{p} (A_{1,\mathcal{T}} + A_{2,\mathcal{T}})^2 + V_{\mathcal{T}}, \end{aligned}$$

as $n \rightarrow \infty$, $N \rightarrow \infty$ and

$$\begin{aligned} \frac{1}{N\Delta t} \sum_{j=1}^N \Delta x_{n,j} x_{n-1,j} &= \alpha_{n-1} \left(\frac{\mu_{n-1}}{N} \sum_{j=1}^N x_{n-1,j} - \frac{1}{N} \sum_{j=1}^N x_{n-1,j}^2 \right) + \frac{\sigma_{n-1}}{N\Delta t} \sum_{j=1}^N \Delta W_{n,j} x_{n-1,j} \\ &\xrightarrow{P} \alpha(\mathcal{T}) \left(\mu(\mathcal{T}) \mathbb{E}[x(\mathcal{T})] - \mathbb{E}[x^2(\mathcal{T})] \right), \end{aligned}$$

as $n \rightarrow \infty, N \rightarrow \infty$.

Theorem 4. Assume that $\mathbb{E}[x_0] < \infty$. Then $\hat{\alpha}_{n,N} \xrightarrow{P} \alpha(\mathcal{T})$, $\hat{\mu}_{n,N} \xrightarrow{P} \mu(\mathcal{T})$, and $\hat{\sigma}_{n,N} \xrightarrow{P} \sigma(\mathcal{T})$ as $n \rightarrow \infty, N \rightarrow \infty$.

Proof. From (4.10),

$$\begin{aligned} \frac{1}{N\Delta t} \sum_{j=1}^N \Delta x_{n,j}^2 &= \alpha_{n-1}^2 \Delta t \left(\mu_{n-1}^2 - \frac{2\mu_{n-1}}{N} \sum_{j=1}^N x_{n-1,j} + \frac{1}{N} \sum_{j=1}^N x_{n-1,j}^2 \right) + \frac{2\alpha_{n-1}\mu_{n-1}\sigma_{n-1}}{N} \sum_{j=1}^N \Delta W_{n,j} \\ &\quad - \frac{2\alpha_{n-1}\sigma_{n-1}}{N} \sum_{j=1}^N x_{n-1,j} \Delta W_{n,j} + \frac{\sigma_{n-1}^2}{N\Delta t} \sum_{j=1}^N \Delta W_{n,j}^2 \xrightarrow{P} \sigma^2(\mathcal{T}), \end{aligned}$$

as $n \rightarrow \infty, N \rightarrow \infty$. It follows from (4.11) that

$$\begin{cases} \hat{\alpha}_{n,N} &\xrightarrow{P} \frac{\alpha(\mathcal{T})(\mu(\mathcal{T}) - \mathbb{E}[x(\mathcal{T})])\mathbb{E}[x(\mathcal{T})] - \alpha(\mathcal{T})\mu(\mathcal{T})\mathbb{E}[x(\mathcal{T})] - \mathbb{E}[x^2(\mathcal{T})]}{V_{\mathcal{T}}} = \alpha(\mathcal{T}), \\ \hat{\mu}_{n,N} &\xrightarrow{P} \mathbb{E}[x(\mathcal{T})] + \frac{\alpha(\mathcal{T})(\mu(\mathcal{T}) - \mathbb{E}[x(\mathcal{T})])}{\alpha(\mathcal{T})} = \mu(\mathcal{T}), \\ \hat{\sigma}_{n,N}^2 &\xrightarrow{P} \sigma^2(\mathcal{T}), \end{cases}$$

as $n \rightarrow \infty, N \rightarrow \infty$.

4.2.1 LLGMM Parameter Estimation Scheme

Define $\bar{x}_i = \frac{1}{N} \sum_{j=1}^N x_{i,j}, i = 1, 2, \dots, n$. For each sample observation size $m_n \in I_2^n$, a partition P_n of closed interval $[t_{n-m_n+1}, t_n]$ of length m_n is called local at time t_n and defined by (3.11). Using (4.9), we formulate a local observation process at time t_n as an algebraic functions of m_n - local functions of restriction of the overall finite sample sequence $\{x_i\}_{i=-r}^n$ to a subpartition P_n .

Let $\hat{\alpha}_{n,m_n}, \hat{\mu}_{n,m_n}$, and $\hat{\sigma}_{n,m_n}^2$ be the moving estimates of $\alpha(\mathcal{T}), \mu(\mathcal{T})$, and $\sigma^2(\mathcal{T})$, respectively, at time t_n and observation size m_n . We derive $\hat{\alpha}_{n,m_n}, \hat{\mu}_{n,m_n}$, and $\hat{\sigma}_{n,m_n}^2$ as follows:

$$\begin{cases} \hat{\alpha}_{n,m_n} &= \frac{\frac{1}{m_n} \sum_{i=n-m_n+1}^n \Delta \bar{x}_i \sum_{i=n-m_n+1}^n \bar{x}_{i-1} - \sum_{i=n-m_n+1}^n (\Delta \bar{x}_i) \bar{x}_{i-1}}{\left[\sum_{i=n-m_n+1}^n \bar{x}_{i-1}^2 - \frac{1}{m_n} \left(\sum_{i=n-m_n+1}^n \bar{x}_{i-1} \right)^2 \right] \Delta t}, \\ \hat{\mu}_{n,m_n} &= \frac{1}{m_n} \sum_{i=n-m_n+1}^n \bar{x}_{i-1} + \frac{\frac{1}{m_n} \sum_{i=n-m_n+1}^n \Delta \bar{x}_i}{\hat{\alpha}_{n,m_n} \Delta t}, \\ \hat{\sigma}_{n,m_n}^2 &= \frac{1}{\Delta t} \cdot \frac{1}{m_n} \sum_{i=n-m_n+1}^n \left(\Delta \bar{x}_i - \frac{1}{m_n} \sum_{i=n-m_n+1}^n \Delta \bar{x}_i \right)^2. \end{cases} \tag{4.14}$$

Remark 4. As discussed in Remark 1, the local admissible observation size m_k is not constant at each time t_k . For a given $\epsilon > 0$, the value of m_k that gives the least simulated error at each time t_k is recorded as the ϵ -best sub-optimal sample size \hat{m}_k and the parameters $\hat{\alpha}_{k,m_k}, \hat{\mu}_{k,m_k}$ and $\hat{\sigma}_{k,m_k}^2$ recorded as $\alpha_{k,\hat{m}_k}, \mu_{k,\hat{m}_k}$ and σ_{k,\hat{m}_k}^2 , respectively.

Remark 5. We present the maximum likelihood estimate of the parameters α, μ and σ in (4.2) as follows:

For random variables x_1, x_2, \dots, x_n satisfying (4.2) such that $x_i = x(t_i)$, we have

$$x_i = \mu + (x_0 - \mu) e^{-i\alpha\Delta t} + \xi_i, \quad i = 1, 2, 3, \dots, n, \tag{4.15}$$

where $\xi_i = \sigma \int_0^{t_i} e^{-\alpha(t_i-s)} dW(s), i = 1, 2, \dots, n$, are identically normally distributed with mean 0 and variance $\frac{\sigma^2}{2\alpha} (1 - e^{-2\alpha\Delta t})$. The joint density function of the random vector $\mathbf{x} = [x_1, x_2, \dots, x_n]'$ is given by

$$f_{\mathbf{x}}(x_1, x_2, \dots, x_n) = (2\pi)^{-\frac{1}{2}n} |\Sigma|^{-\frac{1}{2}} e^{-\frac{1}{2}(\mathbf{x}-\mathbf{m})' \Sigma^{-1}(\mathbf{x}-\mathbf{m})}, \tag{4.16}$$

where $m_i = \mu + (x_0 - \mu)e^{-i\alpha\Delta t}$, $\Sigma_{i,j} = \Sigma_{j,i}$ and $\Sigma_{i,j} = \frac{\sigma^2}{2\alpha} (e^{-(i-j)\alpha\Delta t} - e^{-(i+j)\alpha\Delta t})$, $1 \leq j \leq i \leq n$. Σ is positive definite with determinant $|\Sigma| = \left(\frac{\sigma^2}{2\alpha}\right)^n (1 - e^{-2\alpha\Delta t})^n$ and inverse defined by

$$\Sigma_{i,j}^{-1} = \begin{cases} \frac{2\alpha}{\sigma^2} \frac{e^{2\alpha\Delta t}}{e^{2\alpha\Delta t} - 1}, & \text{if } i = j = n, \\ \frac{2\alpha}{\sigma^2} \frac{e^{2\alpha\Delta t} + 1}{e^{2\alpha\Delta t} - 1}, & \text{if } i = j \neq n, \\ \frac{2\alpha}{\sigma^2} \frac{e^{\alpha\Delta t}}{1 - e^{-2\alpha\Delta t}}, & \text{if } |i - j| = 1, \\ 0, & \text{if } |i - j| \geq 2. \end{cases}$$

It can be shown that $(\mathbf{x} - \mathbf{m})' \Sigma^{-1} (\mathbf{x} - \mathbf{m}) = \sum_{i=1}^n \frac{(x_i - \mu - (x_{i-1} - \mu)e^{-\alpha\Delta t})^2}{\frac{\sigma^2}{2\alpha} [1 - e^{-2\alpha\Delta t}]}$. The log-likelihood function $\mathcal{L}_n(\alpha, \mu, \sigma : x_1, \dots, x_n)$ is

$$\mathcal{L}_n(\alpha, \mu, \sigma : \mathbf{x}) = -\frac{n}{2} \left[\log\left(\frac{\sigma^2}{2\alpha}\right) + \log(1 - e^{-2\alpha\Delta t}) + \log(2\pi) \right] - \frac{\alpha}{\sigma^2} \sum_{i=1}^n \frac{(x_i - \mu - (x_{i-1} - \mu)e^{-\alpha\Delta t})^2}{[1 - e^{-2\alpha\Delta t}]}. \tag{4.17}$$

Estimating the maximum likelihood estimates $\hat{\alpha}$, $\hat{\mu}$ and $\hat{\sigma}$ of α , μ and σ , respectively, reduces to solving the nonlinear system of equations

$$\begin{cases} \frac{\partial \mathcal{L}_n(\alpha, \mu, \sigma : \mathbf{x})}{\partial \mu} = 0; & \frac{\partial \mathcal{L}_n(\alpha, \mu, \sigma : \mathbf{x})}{\partial \alpha} = 0; & \frac{\partial \mathcal{L}_n(\alpha, \mu, \sigma : \mathbf{x})}{\partial \sigma} = 0, \end{cases} \tag{4.18}$$

for μ , α and σ . Let $\text{var}(\mathbf{x}) = \frac{1}{n} \sum_{i=1}^n \left(x_{i-1} - \frac{1}{n} \sum_{j=1}^n x_{j-1}\right)^2$, $\text{cov}(\mathbf{x}, \mathbf{y}) = \frac{1}{n} \sum_{i=1}^n \left(x_{i-1} - \frac{1}{n} \sum_{j=1}^n x_{j-1}\right) \left(x_i - \frac{1}{n} \sum_{j=1}^n x_j\right)$, where $\mathbf{y} = [x_0, \dots, x_{n-1}]'$. The MLE estimates $\hat{\alpha}_n$, $\hat{\mu}_n$ and $\hat{\sigma}_n$ satisfying (4.18) are given by

$$\begin{aligned} \hat{\alpha}_n &= \frac{1}{\Delta t} \log\left(\frac{\text{var}(\mathbf{x})}{\text{cov}(\mathbf{x}, \mathbf{y})}\right), \\ \hat{\mu}_n &= \frac{1}{n} \sum_{i=1}^n x_i + \frac{e^{-\hat{\alpha}\Delta t}}{1 - e^{-\hat{\alpha}\Delta t}} \frac{1}{n} \sum_{i=1}^n \Delta x_i, \\ \hat{\sigma}_n^2 &= \frac{2\hat{\alpha}}{1 - e^{-2\hat{\alpha}\Delta t}} \cdot \frac{1}{n} \sum_{i=1}^n \left(x_i - \hat{\mu} - (x_{i-1} - \hat{\mu})e^{-\hat{\alpha}\Delta t}\right)^2. \end{aligned} \tag{4.19}$$

Using the Taylor series expansion of $\ln(x)$ around $x = 1$, we can approximate $\ln(x) \approx 1 - \frac{1}{x}$. Also, $\frac{1 - e^{-x}}{e^{-x}} \approx x$ for considerably small x . Thus, we can compare (4.19) with (4.14) for large observation size $m_n = n$ as follows:

$$\begin{aligned} \hat{\alpha}_n &= \frac{1}{\Delta t} \log\left(\frac{\text{var}(\mathbf{x})}{\text{cov}(\mathbf{x}, \mathbf{y})}\right) \approx \frac{\frac{1}{n} \sum_{i=1}^n \Delta x_i \sum_{i=1}^n x_{i-1} - \sum_{i=1}^n (\Delta x_i) x_{i-1}}{\left[\sum_{i=1}^n x_{i-1}^2 - \frac{1}{n} \left(\sum_{i=1}^n x_{i-1}\right)^2\right] \Delta t}, \\ \hat{\mu}_n &= \frac{1}{n} \sum_{i=1}^n x_i + \frac{e^{-\hat{\alpha}\Delta t}}{1 - e^{-\hat{\alpha}\Delta t}} \frac{1}{n} \sum_{i=1}^n \Delta x_i \approx \frac{1}{n} \sum_{i=1}^n x_{i-1} + \frac{\frac{1}{n} \sum_{i=1}^n \Delta x_i}{\hat{\alpha}_n \Delta t}, \\ \hat{\sigma}_n^2 &= \frac{2\hat{\alpha}}{1 - e^{-2\hat{\alpha}\Delta t}} \cdot \frac{1}{n} \sum_{i=1}^n \left(x_i - \hat{\mu} - (x_{i-1} - \hat{\mu})e^{-\hat{\alpha}\Delta t}\right)^2 \approx \frac{1}{n\Delta t} \sum_{i=1}^n \left(\Delta x_i - \frac{1}{n} \sum_{i=1}^n \Delta x_i\right)^2. \end{aligned} \tag{4.20}$$

5. Numerical Simulation

In this section, we simulate stock prices and energy commodities' prices using discretized stochastic differential equations (3.1) and (4.1), together with their estimated parameters derived in (3.12) and (4.14), respectively.

5.1 Numerical Simulation of Stock Price Using LLGMM Method

Let y_{k,m_k}^s be a simulated value of the real stock price satisfying the Euler- discretized version of (3.1) at time t_k . y_{k,m_k}^s corresponds to a local admissible lagged-adapted finite sequences of conditional sample/data observation $\{\mathbb{E}[y_i | \mathcal{F}_{i-1}]\}_{i=k-m_k+1}^k$ of size m_k at time t_k derived from the discretized Euler scheme

$$y_{k,m_k}^s = y_{k-1,m_{k-1}}^s + \hat{\alpha}_{k-1,m_{k-1}} y_{k-1,m_{k-1}}^s \Delta t + \hat{\sigma}_{k-1,m_{k-1}} y_{k-1,m_{k-1}}^s \Delta W_{k,m_k}. \tag{5.1}$$

We define the quadratic square error of $\mathbb{E}[y_k | \mathcal{F}_{k-1}]$ relative to each simulated values y_{k,m_k}^s as

$$\Xi_{m_k,k,y_k} = \left(\mathbb{E}[y_k | \mathcal{F}_{k-1}] - y_{m_k,k}^s\right)^2. \tag{5.2}$$

For any arbitrary small positive number ϵ and for each time t_k , we define the ϵ -sub-optimal admissible subset \mathcal{M}_k of set of m_k -size local admissible lagged sample size m_k at t_k as:

$$\mathcal{M}_k = \{m_k : \Xi_{m_k,k,y_k} < \epsilon\}. \tag{5.3}$$

We record the value $m_k \in \mathcal{M}_k$ that gives the minimum error as \hat{m}_k . The ϵ -best sub-optimal estimates of the parameters \hat{a}_{k,m_k} and $\hat{\sigma}_{k,m_k}^2$ at the ϵ -best sub-optimal sample size \hat{m}_k are also recorded as a_{k,\hat{m}_k} and σ_{k,\hat{m}_k}^2 , respectively.

Finally, the simulated value at time t_k with \hat{m}_k is now recorded as the ϵ -best sub-optimal state estimate y_{k,\hat{m}_k}^s for $\mathbb{E}[y_k|\mathcal{F}_{k-1}]$ at time t_k

Now, we apply the LLGMM conceptual computational algorithm to simulate real stock price data namely International Businesses Machines (IBM), JPMorgan Chase and Co and Apple Inc. for the same period 01/03/2000 – 01/27/2017 (Apple; JPMorgan; IBMI). For simulation purpose, we pick time delay $r = 30$, $\Delta t = 1$ and the quadratic square error of the simulation is chosen to be at most $\epsilon = 0.001$. We note here that r can be chosen as large as the sample size \mathcal{N} . We note that as r approaches \mathcal{N} , the quadratic square error reduces significantly. This is demonstrated by showing side by side simulations for $r = 30$ and $r = 50$ in Figures 3 and 4, respectively.

The graphs of the ϵ -best sub-optimal estimates a_{k,\hat{m}_k} for the three stock prices International Businesses Machines (IBM), JPMorgan Chase and Co and Apple Inc. are shown in Figure 1 below.

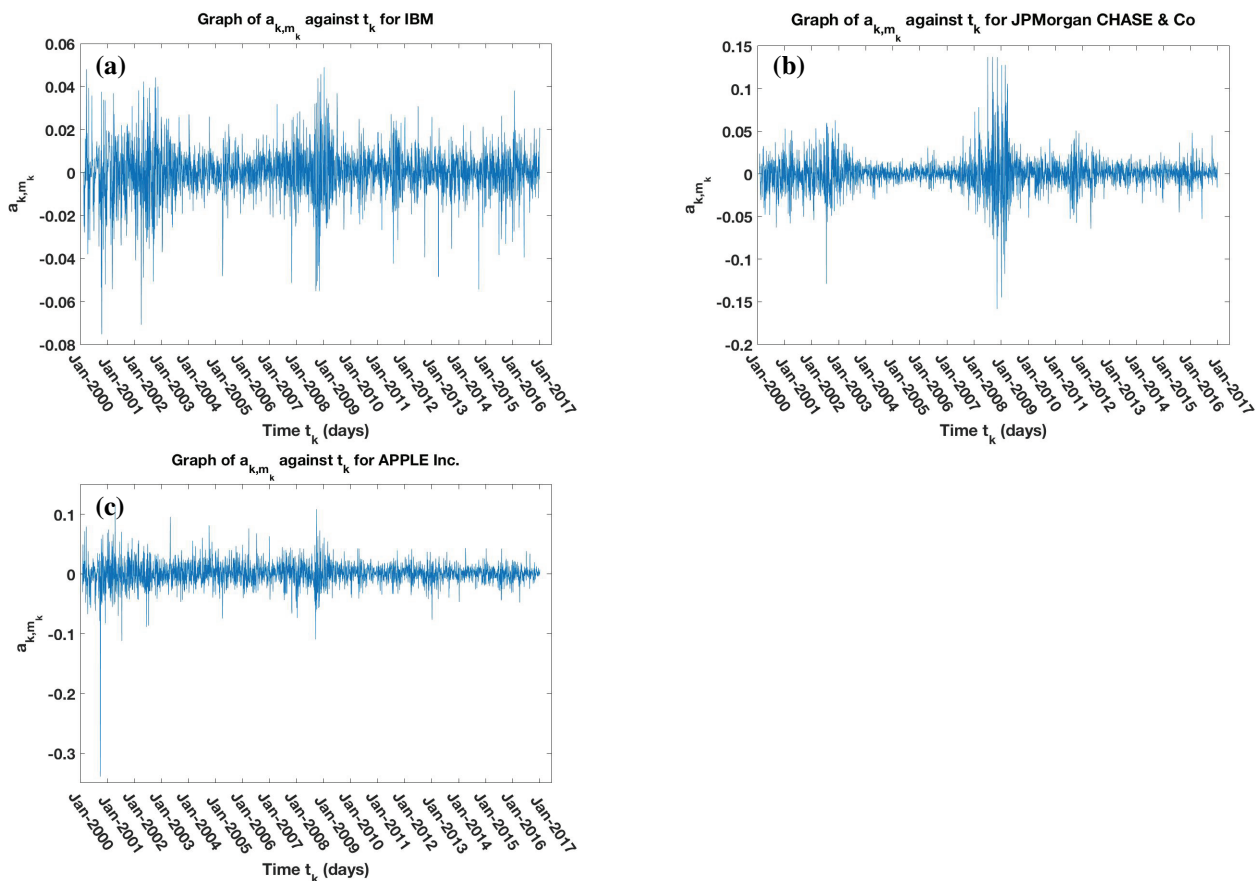


Figure 1. The graphs of parameter a_{k,\hat{m}_k} with time t_k for IBM, JPMorgan CHASE & Co and APPLE Inc stock price

The graphs of the ϵ -best sub-optimal estimates σ_{k,\hat{m}_k}^2 for the three stock prices International Businesses Machines (IBM), JPMorgan Chase and Co and Apple Inc. are shown in Figure 2 below.

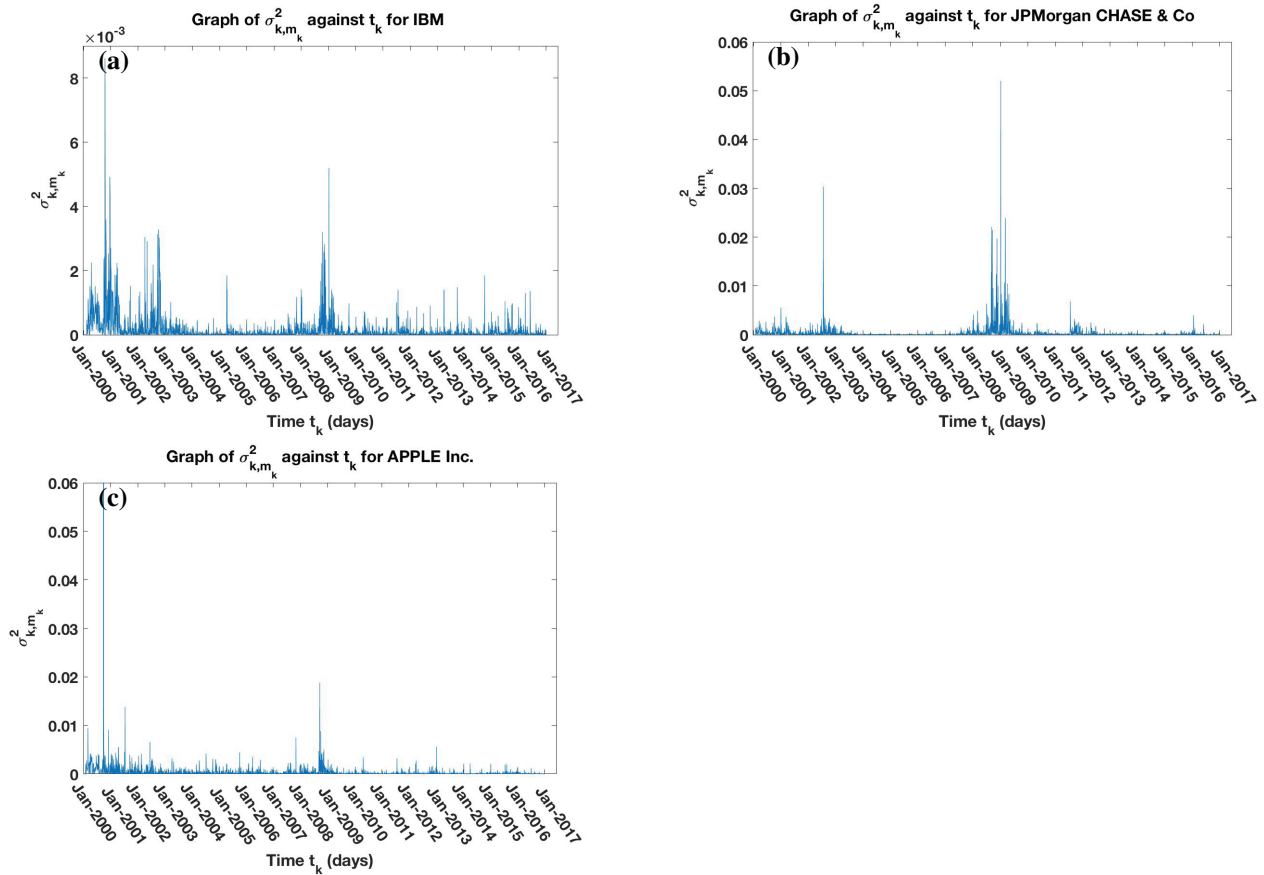
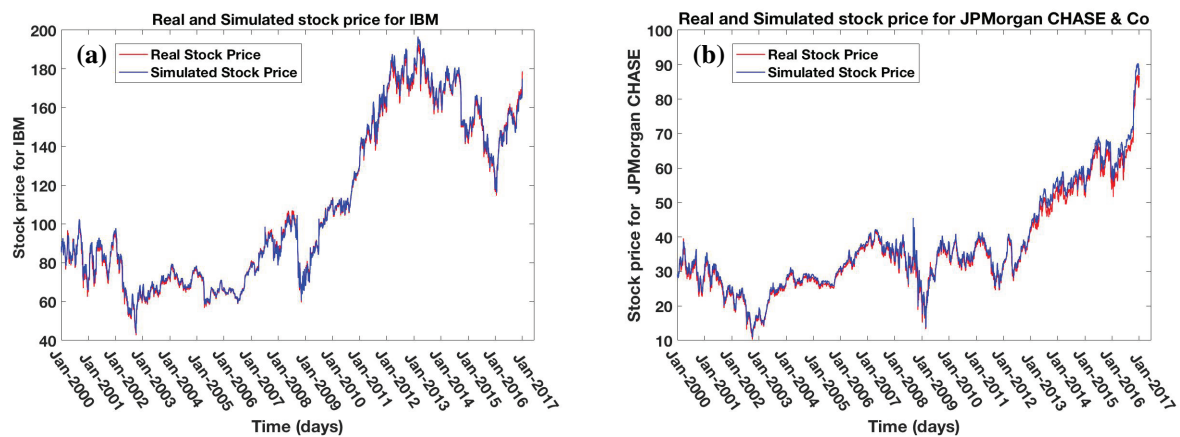


Figure 2. The graphs of parameter σ_{k,m_k}^2 with time t_k for IBM, JPMorgan CHASE & Co and APPLE Inc stock price

The graphs of the ϵ - best sub-optimal simulated stock price estimates for the three stock prices International Businesses Machines (IBM), JPMorgan Chase and Co and Apple Inc. are shown in Figure 3 below.



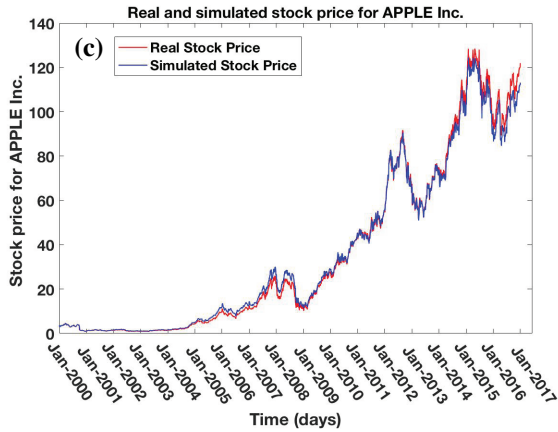


Figure 3. The graphs of real and simulated stock price for IBM, JPMorgan CHASE & Co and APPLE Inc. $r = 30$

In order to show the effect of increasing the delay constant r , we also include the simulation result using $r = 50$. It is obvious from comparison of Figures 3 and 4 that as r approaches \mathcal{N} , the root mean square value approaches zero.

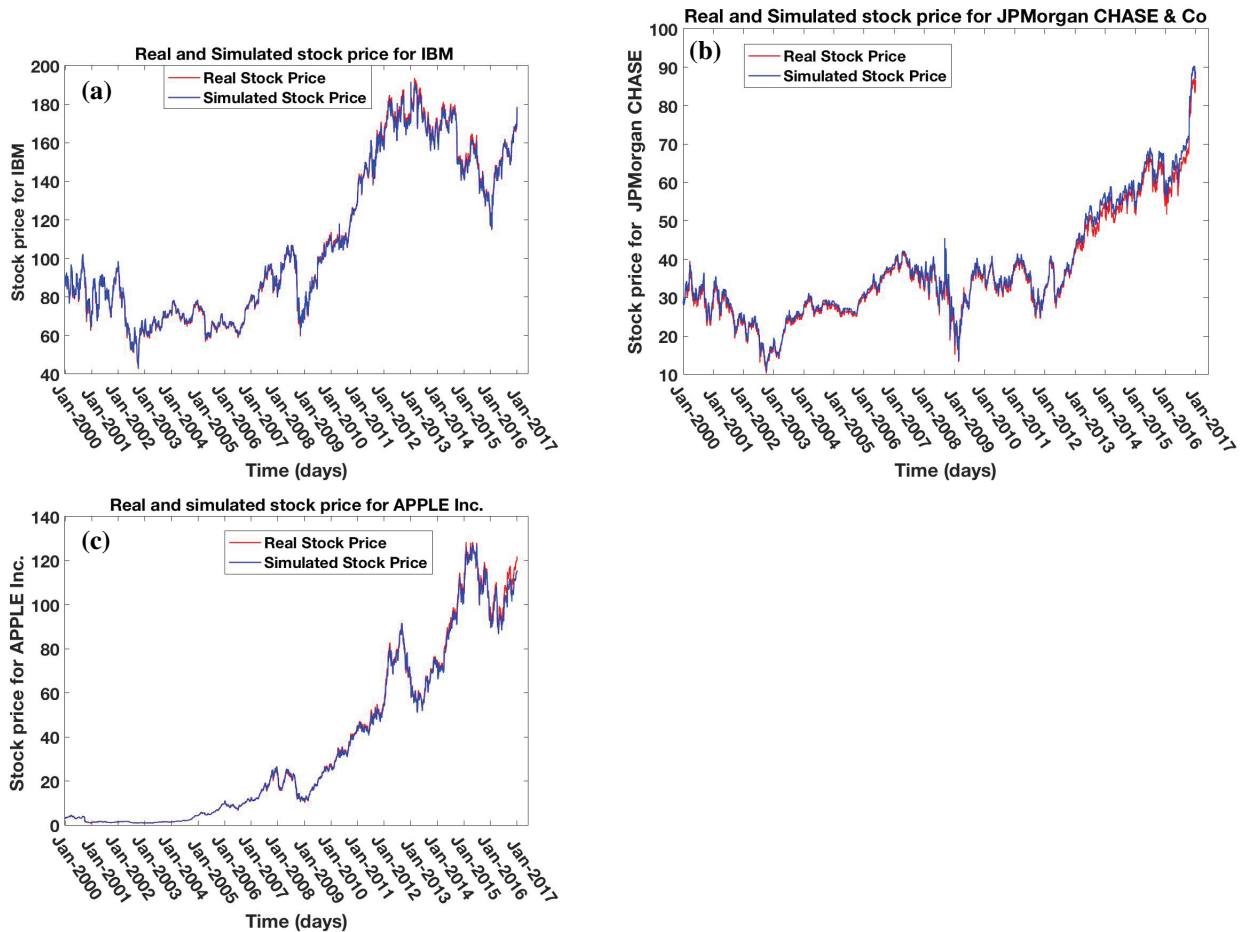


Figure 4. The graphs of real and simulated stock price for IBM, JPMorgan CHASE & Co and APPLE Inc. for $r = 50$

5.2 Numerical Simulation of energy commodity using LLGMM method

Using discretized Euler scheme for (4.1), we apply the above conceptual computational algorithm for the real time data sets namely; daily Henry Hub Natural gas (dollars/million Btu) price data for the period 01/04/2000 – 01/09/2017, crude oil price (dollars/barrel) data for the period 01/04/2000 – 01/09/2017 and coal price data for the period 01/03/2000 –

01/11/2016 (Coal; Crude; Natural gas). For simulation purposes, we pick time delay $r = 30$, $\Delta t = 1$ and the quadratic square error of the simulation is chosen to be at most $\epsilon = 0.001$. We also note here that r can be chosen as large as the sample size, \mathcal{N} . As r approaches \mathcal{N} , the quadratic square error reduces significantly. We demonstrated this by showing side by side simulations for $r = 30$ and $r = 50$ in Figures 8 and 9, respectively.

The graphs of the ϵ - best sub-optimal estimates μ_{k,m_k} for the three energy commodities: natural gas, crude oil and coal price are shown in Figure 5 below.

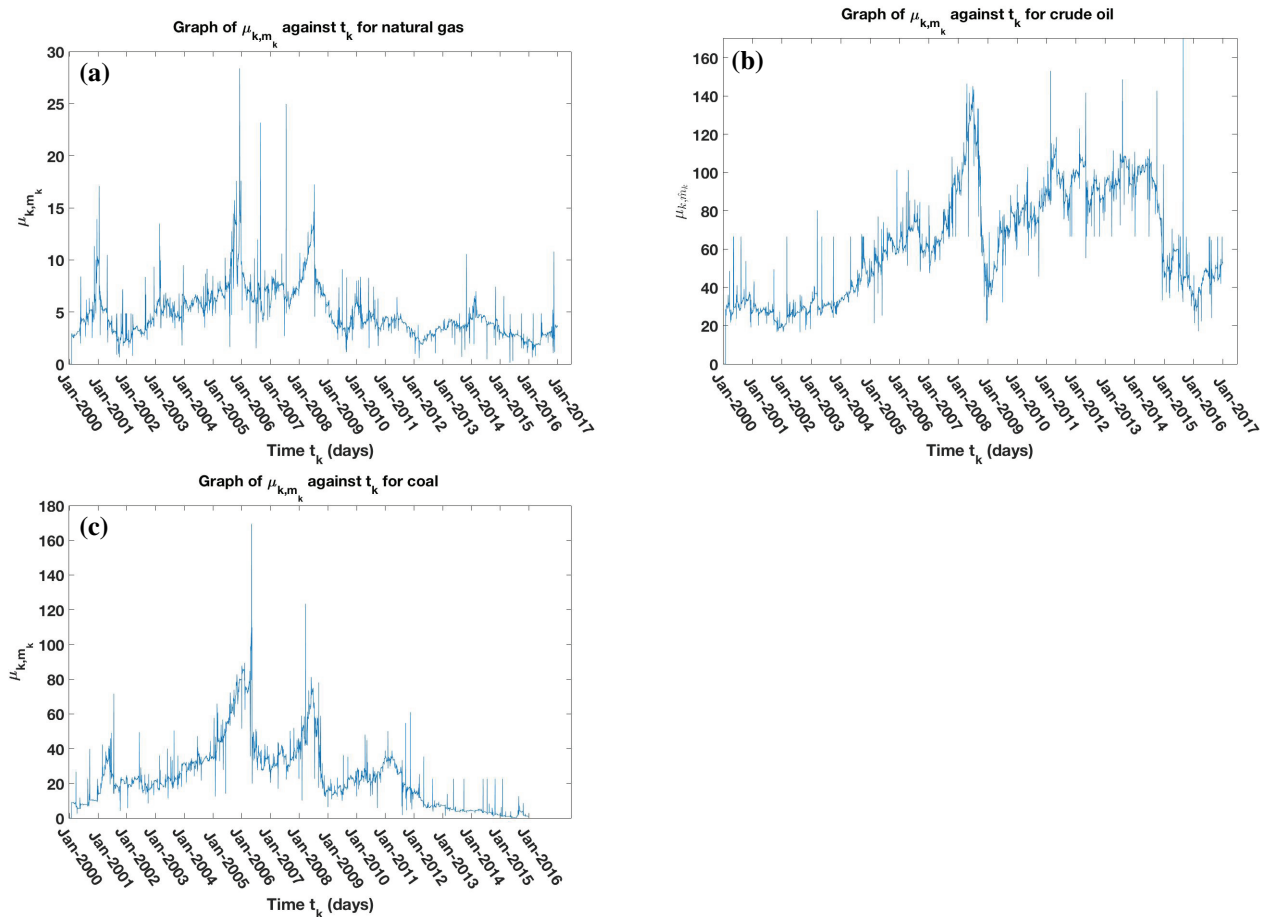


Figure 5. The graphs of parameter μ_{k,m_k} with time t_k for natural gas, crude oil and coal

The graphs of the ϵ - best sub-optimal estimates α_{k,\hat{m}_k} for the three energy commodities: natural gas, crude oil and coal price are shown in Figure 6 below.

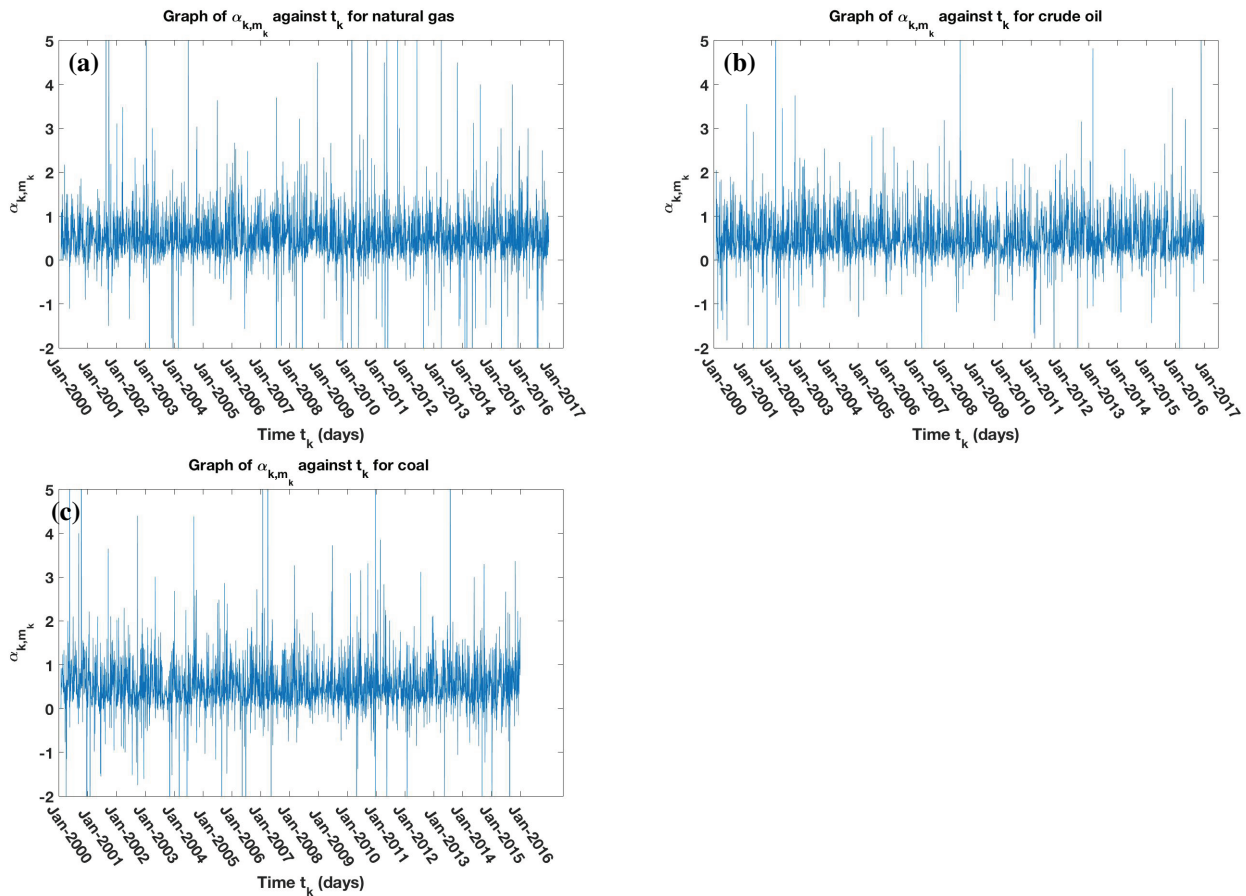
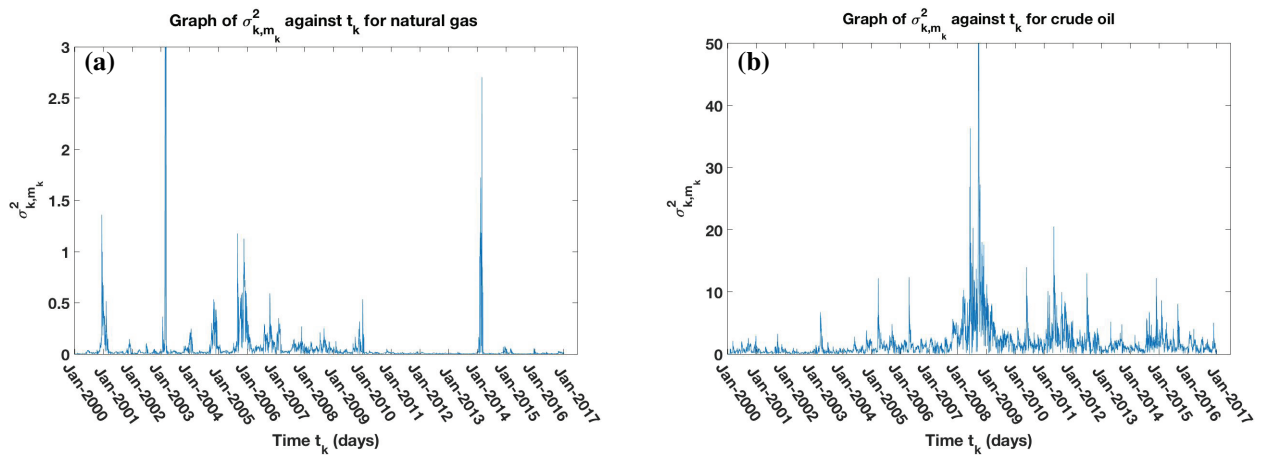


Figure 6. The graphs of parameter α_{k,m_k} with time t_k for natural gas, crude oil and coal

The graphs of the ϵ - best sub-optimal estimates σ_{k,m_k}^2 for the three energy commodities Natural gas, crude oil and coal price are shown in Figure 7 below.



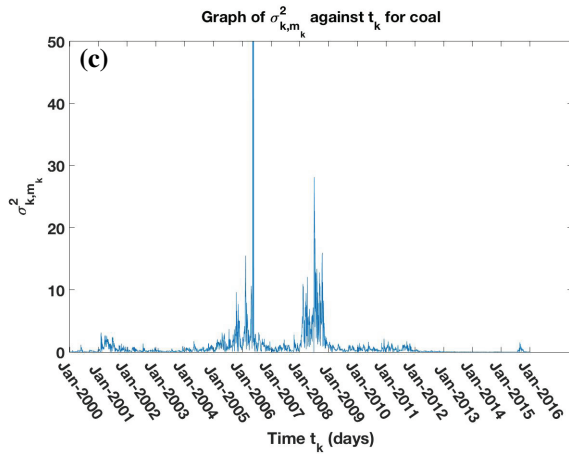


Figure 7. The graphs of parameter σ^2_{k,m_k} with time t_k for natural gas, crude oil and coal

In the following figures, we show the effect of increasing the delay constant r by showing a simulation result for the cases where $r = 30$ and $r = 50$.

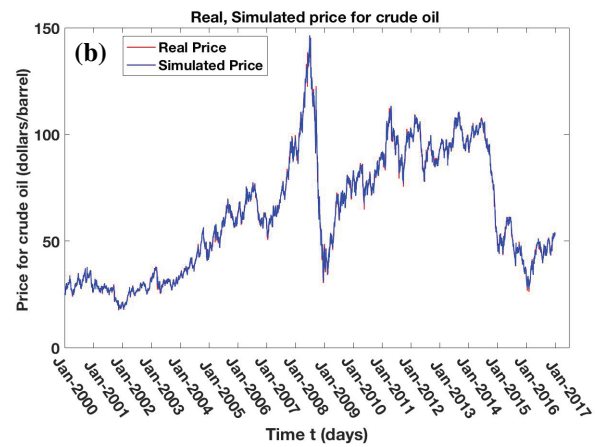
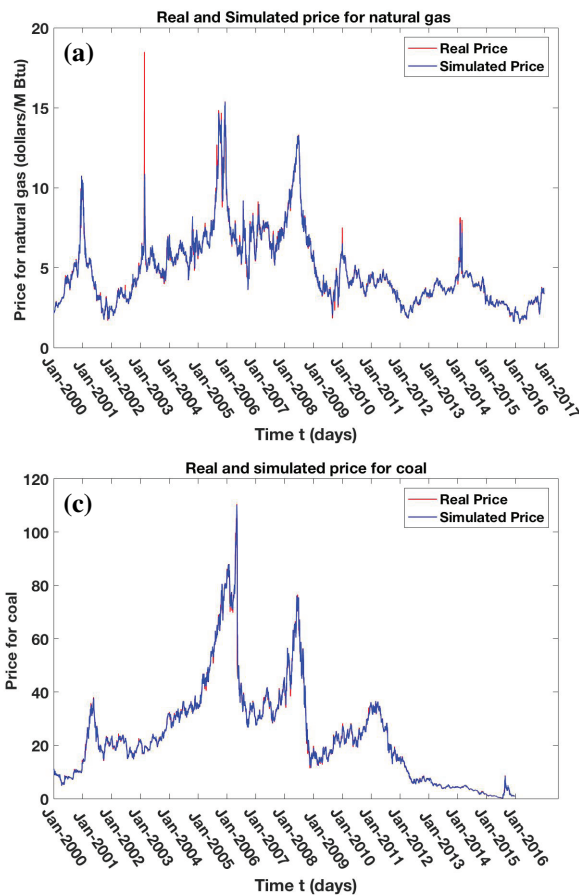


Figure 8. The graphs of real and simulated daily prices of natural gas, crude oil and coal for $r = 30$.

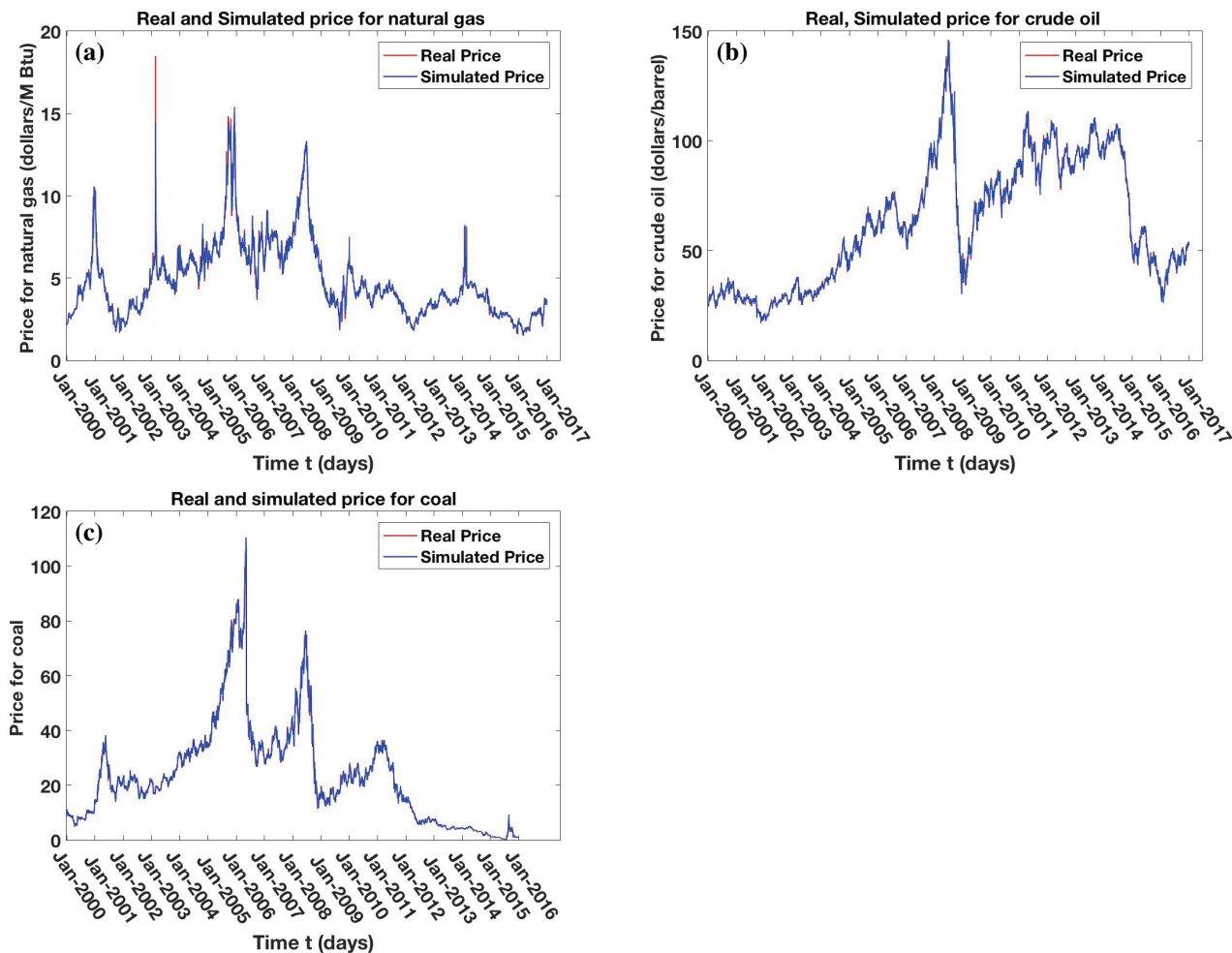


Figure 9. The graphs of real and simulated daily prices for natural gas, crude oil and coal for $r = 50$.

6. Summary

The purpose of this paper is to describe the LLGMM method of parameter estimations in a linear stochastic differential equation with time dependent parameters. It is shown that the estimated parameters converge in probability to the true parameters of the stochastic model. Numerical simulation is carried out in the context of three stock prices and three energy commodity prices. We also show graphically that as the delay constant r increases, the quadratic square error tends to zero.

References

Andrews, Donald, W. K. (2002). *Generalized method of moments estimation when a parameter is on a boundary*. *Journal of Business Economic Statistics*, 20(4), 530-544. <https://doi.org/10.1198/073500102288618667>

Apostol, Tom, M. (1974). *Mathematical Analysis*, Addison-Wesley Publishing Company, Inc.

Bibbona, E., Panfilo, G., & Tavella, P. (2008). The Ornstein-Uhlenbeck process as a model of a low pass filtered white noise. *Metrologia*, 45(6), S117-S126. <https://doi.org/10.1088/0026-1394/45/6/S17>

Cramer, H. (1946). *Mathematical methods of statistics* (Princeton University Press, Princeton, NJ).

Casella, George, & Berger, Roger, L. (2002). *Statistical Inference*, Second Edition, Duxbury Advanced Series, Pacific Grove, California.

Daily Coal data for the period 01/03/2000 – 01/11/2016, *Arch Coal Incorporation*, URL <http://finance.yahoo.com/q/hp?s=ACI>.

Daily Crude Oil data for the period 01/03/2000 – 01/09/2017, *U. S. Energy Information Administration Website*, URL <http://www.eia.gov/>.

- Daily Henry Hub Natural gas data for the period 01/03/2000-01/09/2017, *U. S. Energy Information Administration Website*, URL <http://www.eia.gov/>.
- Daily stock price for Apple Incorporation for the period 01/03/2000 – 01/27/2017, <http://finance.yahoo.com/quote/AAPL/history?p=AAPL>.
- Daily stock price for JPMorgan Chase & Co. for the period 01/03/2000 – 01/27/2017, <http://finance.yahoo.com/quote/JPM/history?p=JPM>.
- Daily stock price for International Businesses Machines Corporation (IBM) for the period 01/03/2000 – 01/27/2017, <http://finance.yahoo.com/quote/IBM/history?p=IBM>.
- Donnet, Sophie, & Samson, Adeline. (2013). A Review on Estimation of Stochastic Differential Equations for Pharmacokinetic/Pharmacodynamic Models. *Advanced Drug Delivery Reviews*, 65(7), 929-939. <https://doi.org/10.1016/j.addr.2013.03.005>
- Hackl, Peter. (1989). Statistical Analysis and Forecasting of Economic Structural Change. *Springer-Verlag Berlin Heidelberg*, 253-264.
- Hansen, Lars, Peter. (1982). Large Sample Properties of Generalized Method of Moments Estimators. *Econometrica*, 50(4), 1029-1054. <https://doi.org/10.2307/1912775>
- Hansen, Lars, Peter, & Scheinkman, Jose Alexandre. (1995). Back to the Future: Generating Moment Implications for Continuous-Time Markov Processes. *Econometrica*, 63(4), 767-804. <https://doi.org/10.2307/2171800>
- Kloeden, P. E., & Platen, E. (1992). *Numerical Solution of Stochastic Differential Equations*. New York: Springer-Verlag. <https://doi.org/10.1007/978-3-662-12616-5>
- Ladde, Anil, G., & Ladde, G. S. (2013). *An Introduction to Differential Equations: Stochastic Modelling Methods and Analysis*. Volume 2, World Scientific Publishing Company, Singapore.
- Ladde, G. S. (1979). Stability and Oscillations in single-species process with past memory. *Int. J. Systems Sci.*, 10(6), 639-647.
- Otunuga, Olusegun, M., Ladde, Gangaram, & Ladde, Nathan. (2017). Local lagged adapted generalized method of moments and applications. *Journal of stochastic analysis and applications*, 35(1), 110-143. <http://dx.doi.org/10.1080/07362994.2016.1213640>
- Ozaki T. (1992). A Bridge between Nonlinear Time Series Models and Nonlinear Stochastic Dynamical Systems: A local Linearization Approach. *Stat. Sin.*, 2, 115-135.
- Paothong, Arnut, & Ladde, G. S. (2013). Adaptive Expectations and Dynamic Models for Network Goods. *Economic Analysis and Policy*, 43, 353-373. [https://doi.org/10.1016/S0313-5926\(13\)50036-1](https://doi.org/10.1016/S0313-5926(13)50036-1)
- Rao, Radhakrishna. (1973). *Linear statistical inference and its applications*, 2nd ed. Wiley, New York.
- Ronchetti, E., & Trojani, F. (2001). Robust inference with GMM estimators. *Journal of Econometrics*, 101, 37-69.

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).