Strong consistency of the Bayesian estimator for the Ornstein-Uhlenbeck process

Arturo Kohatsu-Higa, Nicolas Vayatis and Kazuhiro Yasuda

Abstract In the accompanying paper Kohatsu-Higa et al. [9], we have introduced a theoretical study of the consistency of a computational intensive parameter estimation method for Markovian models. This method could be considered as an approximate Bayesian estimator method or a filtering problem approximated using particle methods. We showed in [9] that under certain conditions which explicitly relate the number of data, the amount of simulations and the size of the kernel window, one obtains the rate of convergence of the method. In this first study, the conditions do not seem easy to verify and for this reason, we show in this paper how to verify these conditions in the toy example of the Ornstein-Uhlenbeck processes. We hope that this article will help the reader understand the theoretical background of our previous studies and how to interpret the required hypotheses.

1 Introduction

One method to estimate parameters in a Markovian model is to use a filtering method (also known as the Bayesian method). In such a framework the estimation is carried out using a least-square principle which leads to the calculation of the conditional expectation of the unknownn density given the available data.

Arturo Kohatsu-Higa
Osaka University and Japan Science and Technology Agency, Machikaneyama cho 1-3, Osaka, 560-8531, Japan, e-mail: kohatsu@sigmath.es.osaka-u.ac.jp

Nicolas Vayatis
Ecole Normale Superieure de Cachan, 61, avenue du president Wilson 94 235 Cachan cedex, France, e-mail: vayatis@cmla.ens-cachan.fr

Kazuhiro Yasuda
Hosei University, 3-7-2, Kajino-cho, Koganei-shi, Tokyo, 184-8584, Japan, e-mail: yasuda@hosei.ac.jp

This expression is somehow theoretical, so one option is to use simulation to approximate the value of the unknownn transition density if some theoretical model is proposed. This simulation procedure requires the choice of a variety of parameters. The procedure of choosing these parameters “correctly” is called tuning.

Recently, many computational statisticians have successfully proposed and studied several algorithms related with this idea, for example, using the Markov Chain Monte-Carlo method (Roberts et al. [10]) between others. Many papers have confirmed the rate of convergence of the proposed method to the desired value using numerical experiments, but usually no mathematical proof is provided. In an accompanying paper [9], we adopt a particle method (details and other comments about this method can be found in Bain et al. [2]) to approximate the conditional expectation and study theoretically the rate of convergence and the proper tuning needed. This kind of filtering problem under discrete observations were studied by Del Moral et al. [4]. They prove weak consistency and L^2-convergence. After that, Cano et al. [3] studied the convergence of an approximated posterior distribution, which used the Euler-Maruyama approximation for stochastic differential equations (SDE). Kohatsu-Higa et al. [9], gives the rate of convergence of the approximated Bayesian estimator. In that set-up, the transition density function of an observation process is usually unknown, so that one approximates it by using the kernel density estimation method (KDE). As mentioned before, we remark that there are several new algorithms, which may work well on applications, but our objective was to provide a theoretical mathematical framework therefore we choose the most basic method available within particle methods. Our method of analysis uses the Laplace method to obtain the rate of convergence $1/\sqrt{N}$, where N is a number of data under a strong hypothesis of convergence rate for the approximating average of likelihoods (see Assumption (A) (6)-(a)).

In the second part of Kohatsu-Higa et al. [9], we gave an explicit relationship between number of data and approximation parameters, as to ensure that Assumption (A) (6)-(a) is satisfied. Here we have three approximation parameters: (i). the first one is used to approximate the theoretical stochastic processes, (ii). the second one is to express the number of the Monte-Carlo simulations used for the approximating process, (iii). the last one is a bandwidth size of the KDE. We connect these three approximation parameters and the number of data. We believe that our study is the first that provides an explicit theoretical relationship between these parameters in order to achieve a certain rate of convergence. It also shows why a bad choice of tunning parameters may lead to unreliable estimation results.

Assumption (A) below states the needed hypothesis in order to achieve the rate of convergence announced previously. These hypotheses are not necessarily easy to understand/interpret. The objective of the present article is to consider an easy toy example where the reader may see how these conditions could be verified and most importantly what do they mean. In this paper, we consider the following Ornstein-Uhlenbeck process (OU process) as the parametrized observation process:

$$dX_t = -\theta X_t dt + dW_t,$$

where W_t is a Brownian motion and θ is a parameter, which we want to estimate. Then, we check the assumptions, that give the strong consistency and the convergence rate. Clearly this is a toy example as many elements can be directly computed and there is no need to use simulations. Furthermore in that setting many other competing statistical methods exist (see e.g. [1], [6], [7], [8], [11]).

We would like to emphasize again, that the main objective here is to show that the general theory is applicable to a basic example. Clearly, there are still open problems to be considered and in particular, how to apply the results in other examples. We hope that with this article, the reader may understand when a model satisfies the assumptions although verifying them may still require a long procedure.

This paper is constructed as follows: In Section 2, we give the general theorem and the assumptions of Kohatsu-Higa et al. [9]. In Section 3, we check the assumptions with respect to the OU process. and the Euler-Maruyama approximation of the OU process. Finally in Section 4, we give some properties the mean and variance of the OU process and its Euler-Maruyama approximation.

2 Framework and General Theorem

2.1 Framework

In this article, we consider the following problem: Let $\theta_0 \in \Theta := [\theta^l, \theta^u]$, $(\theta^l < \theta^u)$ be a parameter that we want to estimate $\theta_0 \in \dot{\Theta}$, where $\dot{\Theta}$ denotes the interior of the set Θ and $\Theta_0 = \Theta - \{\theta_0\}$. Let $(\Omega, \mathscr{F}, P_{\theta_0})$, $(\bar{\Omega}, \bar{\mathscr{F}}, \bar{P})$ and $(\hat{\Omega}, \hat{\mathscr{F}}, \hat{P})$ be three probability spaces, where the probability measure P_{θ_0} is parametrized by θ_0. $\Delta > 0$ is a fixed parameter that represents the time between observations. The observed Markov chain is defined on the probability space $(\Omega, \mathscr{F}, P_{\theta_0})$. The theoretical Markov chain (with law P_θ) and its approximation are defined on $(\bar{\Omega}, \bar{\mathscr{F}}, \bar{P})$. Finally, simulations are defined on $(\hat{\Omega}, \hat{\mathscr{F}}, \hat{P})$, which will be used in estimating the transition density of the theoretical Markov chain.

- **(i). (Observation process)** Let $\{Y_{i\Delta}\}_{i=0,1,\cdots,N}$ be a sequence of $N+1$-observations of a Markov chain having transition density $p_{\theta_0}(y,z)$, $y, z \in \mathbf{R}$ and invariant measure μ_{θ_0}. This sequence is defined on the probability space $(\Omega, \mathscr{F}, P_{\theta_0})$. We write $Y_i := Y_{i\Delta}$ for $i = 0, 1, \cdots, N$.
- **(ii). (Model process)** Denote by $X^y(\theta)$ a random variable defined on the probability space $(\bar{\Omega}, \bar{\mathscr{F}}, \bar{P})$ such that its law is given by $p_\theta(y,z)$.
- **(iii).** Denote by $(\hat{\Omega}, \hat{\mathscr{F}}, \hat{P})$ the probability space where one generates the simulation of the approximation to the process X^y.
- **(iv). (Approximating process)** Denote by $X^y_{(m)}(\theta)$ the approximation to $X^y(\theta)$, which is defined on $(\bar{\Omega}, \bar{\mathscr{F}}, \bar{P})$. $m = m(N)$ is the parameter that determines the quality of the approximation. Denote by $\tilde{p}^N_\theta(y,z) = \tilde{p}^N_\theta(y,z; m(N))$ the transition density for the process $X^y_{(m)}(\theta)$.

(v). (Approximated transition density) Let $K \in C^2(\mathbf{R};\mathbf{R}_+)$ (usually called kernel), which satisfies $\int K(x)dx = 1$. Denote by $\hat{p}^N_\theta(y,z)$, the kernel density estimate of $\tilde{p}^N_\theta(y,z)$ based on $n \equiv n(N)$ simulated i.i.d. copies of $X^y_{(m)}(\theta)$ which are defined on $(\hat{\Omega}, \hat{\mathscr{F}}, \hat{P})$ and denoted by $X^{y,(k)}_{(m)}(\theta,\cdot)$, $k = 1,\cdots,n$; for $h \equiv h(N) > 0$,

$$\hat{p}^N_\theta(y,z) := \frac{1}{n(N)h(N)} \sum_{k=1}^{n(N)} K\left(\frac{X^{y,(k)}_{(m(N))}(\theta,\hat{\omega}) - z}{h(N)}\right).$$

(vi). For given m, we introduce the “average” approximated transition density over all trajectories with respect to the kernel K;

$$\overline{p}^N_\theta(y,z) := \overline{p}^N_\theta(y,z;m(N),h(N)) := \hat{E}\left[\hat{p}^N_\theta(y,z)\right],$$

where $\hat{E}$ means the expectation with respect to $\hat{P}$.

As it can be deduced from the above set-up, we have preferred to state our problem in abstract terms without explicitly defining the dynamics that generate $X^y(\theta)$ or how the approximation $X^y_{(m)}(\theta)$ is defined. All the properties that will be required for p_θ and $\tilde{p}^N_\theta$ will be satisfied for an appropriate subclass of diffusion processes. Our objective in this article is to show that OU processes are in this class.

Remark 1. Without loss of generality, we can consider the product of the above three probability spaces so that all random variables are defined on the same probability space. We do this without any further mentioning.

Our purpose is to estimate the posterior expectation for some function $f \in C^1(\Theta)$ given the data;

$$E_N[f] := E_\theta[f|Y_0,\cdots,Y_N] := \frac{\int f(\theta)\phi_\theta(Y^N_0)\pi(\theta)d\theta}{\int \phi_\theta(Y^N_0)\pi(\theta)d\theta},$$

where $\phi_\theta(Y^N_0) = \phi_\theta(Y_0,\cdots,Y_N) = \mu_\theta(Y_0)\prod_{j=1}^N p_\theta(Y_{j-1},Y_j)$ is the joint density of $(Y_0,Y_1,\cdots,Y_N)$.

We propose to estimate this quantity on the basis of simulated instances of the process;

$$\hat{E}^n_{N,m}[f] := \frac{\int f(\theta)\hat{\phi}^N_\theta(Y^N_0)\pi(\theta)d\theta}{\int \hat{\phi}^N_\theta(Y^N_0)\pi(\theta)d\theta},$$

where $\hat{\phi}^N_\theta(Y^N_0) := \mu_\theta(Y_0)\prod_{j=1}^N \hat{p}^N_\theta(Y_{j-1},Y_j)$.

2.2 General Theorem of Kohatsu-Higa et al. [9]

Assumption (A): We assume the following

(1). (Observation process) $\{Y_i\}_{i=0,1,\cdots,N}$ is an α-mixing process with $\alpha_n = O(n^{-5})$.

(2). (The prior distribution) The prior distribution $\pi \in C(\Theta)$, and for all $\theta \in \Theta$, $\pi(\theta) > 0$.

(3). (Density regularity) The transition densities $p, \overline{p}^N \in C^{2,0,0}(\Theta \times \mathbf{R}^2; \mathbf{R}_+)$, and for all $\theta \in \Theta$, $y, z \in \mathbf{R}$, we have that $\min\left\{p_\theta(y,z), \overline{p}_\theta^N(y,z)\right\} > 0$. And p_θ admits an invariant measure $\mu \in C_b^{0,0}(\Theta \times \mathbf{R}; \mathbf{R}_+)$, and for all $\theta \in \Theta$, $\mu_\theta(y) > 0$ for every $y \in \mathbf{R}$.

(4). (Identifiability) Assume that there exist c_1, $c_2 : \mathbf{R} \to (0, \infty)$ such that for all $\theta \in \Theta$,

$$\inf_N \int |q_\theta^i(y,z) - q_{\theta_0}^i(y,z)| dz \geq c_i(y) |\theta - \theta_0|,$$

and $C_i(\theta_0) := \int c_i(y)^2 \mu_{\theta_0}(y) dy \in (0, +\infty)$ for $i = 1, 2$ and $q_\theta^1 = p_\theta$ and $q_\theta^2 = \overline{p}_\theta^N$.

(5). (Regularity of the log-density) We assume for $q_\theta = p_\theta, \overline{p}_\theta^N$,

$$\sup_N \sup_{\theta \in \Theta} \int\int \left(\frac{\partial^i}{\partial \theta^i} \ln q_\theta(y,z) \right)^{12} p_{\theta_0}(y,z) \mu_{\theta_0}(y) dy dz < +\infty, \ for\ i = 0,1,2, \quad (1)$$

$$\sup_N \sup_{\theta \in \Theta} \left| \frac{\partial^2}{\partial \theta^2} \int\int (\ln q_\theta(y,z))\, \overline{p}_{\theta_0}^N(y,z) \mu_{\theta_0}(y) dy dz \right| < +\infty, \quad (2)$$

$$\sup_N \sup_{\theta \in \Theta} \int\int \left| \frac{\partial^i}{\partial \theta^i} \ln q_\theta(y,z) \right| \overline{p}_{\theta_0}^N(y,z) \mu_{\theta_0}(y) dy dz < +\infty, \ for\ i = 0,1, \quad (3)$$

where $\frac{\partial^0}{\partial \theta^0} q_\theta = q_\theta$.

(6). (Parameter tuning)

(a). We assume the following boundedness;

$$\sup_N \sup_{\theta \in \Theta} \left| \frac{1}{\sqrt{N}} \sum_{i=0}^{N-1} \left(\frac{\partial}{\partial \theta} \ln \hat{p}_\theta^N(Y_i, Y_{i+1}) - \frac{\partial}{\partial \theta} \ln \overline{p}_\theta^N(Y_i, Y_{i+1}) \right) \right| < +\infty, \ a.s. \quad (4)$$

(b). Assume that for each $y, z \in \mathbf{R}$, there exist functions $C_1^N(y,z)$ and $c_1(y,z)$ such that $|p_{\theta_0}(y,z) - \overline{p}_{\theta_0}^N(y,z)| \leq C_1^N(y,z) a_1(N)$, where $\sup_N C_1^N(y,z) < +\infty$ and $a_1(N) \to 0$ as $N \to \infty$, and $C_1^N(y,z) a_1(N) \sqrt{N} < c_1(y,z)$, where c_1 satisfies; $\sup_N \sup_{\theta \in \Theta} \int\int |\frac{\partial}{\partial \theta} \ln \overline{p}_\theta^N(y,z)| c_1(y,z) \mu_{\theta_0}(y) dy dz < +\infty$.

(c). There exist some function $g^N : \mathbf{R}^2 \to \mathbf{R}$ and constant $a_2(N)$, which depends on N, such that for all $y, z \in \mathbf{R}$,

$$\sup_{\theta \in \Theta} \left| \frac{\partial}{\partial \theta} \ln \overline{p}_\theta^N(y,z) - \frac{\partial}{\partial \theta} \ln p_\theta(y,z) \right| \leq |g^N(y,z)| a_2(N),$$

where $\sup_N E_{\theta_0}[|g^N(Y_0, Y_1)|^4] < +\infty$ and $a_2(N) \to 0$ as $N \to \infty$.

Now we state the main result of the paper.

Theorem 1. *(Kohatsu-Higa et al. [9]) Under Assumption (A), there exists some positive finite random variables Ξ_1 and Ξ_2 such that*

$$|E_N[f]-f(\theta_0)| \le \frac{\Xi_1}{\sqrt{N}} \ a.s., \quad and \quad \left|\hat{E}^n_{N,m}[f]-f(\theta_0)\right| \le \frac{\Xi_2}{\sqrt{N}} \ a.s.,$$

and therefore $|E_N[f]-\hat{E}^n_{N,m}[f]| \le \frac{\Xi_1+\Xi_2}{\sqrt{N}}$ *a.s.*

2.3 Parameter Tuning for Assumption (A) (6)-(a)

All the conditions in Assumption (A) will be directly verified with the exception of Assumption (A) (6)-(a) which requires a special treatment. This section is devoted to show that Assumption (A) (6)-(a) is satisfied under sufficient smoothness hypothesis on the random variables and processes that appear in the problem as well as a certain parameter tuning. We recall that the objective is to find conditions that assure that Assumption (A) (6)-(a) in Section 2.2 is satisfied.

Now $m \equiv m(N)$, $n \equiv n(N)$ and $h \equiv h(N)$ are parameters that depend on N. n is the number of Monte Carlo simulations used in order to estimate the density and m is the generated random numbers used in the simulation of $X^{y,(1)}_{(m)}(\theta,\cdot)$ and h is the window associated to the kernel density estimation method. In this sense we will always think of hypotheses in terms of N although we will drop them from the notation and just use m,n and h. The goal of this section is to prove that under certain hypotheses, there is a choice of m,n and h that ensures that condition (4) is satisfied.

We work in this section under the following hypotheses:

(H1). Assume that there exist some positive constants φ_1,φ_2, where φ_1 is independent of N and φ_2 is independent of N and Δ, such that the following holds; $\inf_{(\mathbf{x},\theta)\in B^N} \overline{p}^N_\theta(x,y) \ge \varphi_1 \exp(-\frac{\varphi_2 a_N^2}{\Delta})$, where sequences of a number a_N and a set B^N are defined by **(ii)** below.

(H2). Assume that the kernel K is the Gaussian kernel; $K(z) := \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}z^2\right)$.

(H5). Assume that there exists some positive constant $C_5 > 0$ such that

$$|\partial_x \overline{p}^N_\theta(x,y)|,\ |\partial_y \overline{p}^N_\theta(x,y)|,\ |\partial_\theta \overline{p}^N_\theta(x,y)| \le C_5 < +\infty,$$

for all $x,y \in \mathbf{R},\ m \in \mathbf{N}$ and $\theta \in \Theta$.

(H5'). Assume that there exists some positive constant $\dot{C}_5 > 0$ such that

$$|\partial_x\partial_\theta \overline{p}^N_\theta(x,y)|,\ |\partial_y\partial_\theta \overline{p}^N_\theta(x,y)|,\ |\partial^2_\theta \overline{p}^N_\theta(x,y)| \le \dot{C}_5 < +\infty,$$

for all $x,y \in \mathbf{R},\ m \in \mathbf{N}$ and $\theta \in \Theta$.

Remark 2. We use the same hypothesis numbering as in Kohatsu-Higa et al. [9] for easy reference. Some of the intermediate hypotheses do not appear here. For the detailed explanations we refer to Kohatsu-Higa et al. [9].

We need to find now a sequence of values for n and h such that all the hypothesis in the previous Theorem are satisfied and that the upper bound is uniformly bounded in N. Now, we rewrite the needed conditions that are related to the parameters n and h. We assume stronger hypothesis that may help us understand better the existence of the right choice of parameters n and h.

The proof of Assumption (A) (6)-(a) uses a series of Borel-Cantelli lemmas for which we need the following hypotheses. We will assume the existence of some sequences of strictly positive numbers which we assume wlog that are bigger than 1.

(ii). (Borel-Cantelli for Y_i)
$m_{c_1} := E[e^{c_1|Y_1|^2}] < +\infty$ for some constant $c_1 > 0$ and $\{a_N\}_{N\in\mathbf{N}} \subset [\theta^u - \theta^l, \infty)$ is a sequence such that for the same c_1,

$$\sum_{N=1}^{\infty} \frac{N}{\exp(c_1 a_N^2)} < +\infty.$$

And set $B^N = \{(\mathbf{x},\theta) \in \mathbf{R}^2 \times \Theta;\ \|\mathbf{x}\| < a_N\}$, where $\|\cdot\|$ denotes the max-norm.

(iii). (Borel-Cantelli for $Z^{(k)}_{3,N}(\omega)$)

For some $r_3 > 0$ and some sequence $b_{3,N} \geq 1$, $N \in \mathbf{N}$, $\sum_{N=1}^{\infty} \frac{n a_N^{2r_3}}{(h^2 b_{3,N})^{r_3}} < +\infty$ and $\sup_{N\in\mathbf{N}} E[|Z_{3,N}(\cdot)|^{r_3}] < +\infty$ for each fixed $m \in \mathbf{N}$, where

$$Z^{(k)}_{3,N}(\omega) := a_N^{-2}\left(\sup_{(\mathbf{x},\theta)\in B^N} |X^{x,(k)}_{(m)}(\theta,\omega)| + 1\right) \sup_{(\mathbf{x},\theta)\in B^N} |\partial_\theta X^{x,(k)}_{(m)}(\theta,\omega)|.$$

(iv). (Borel-Cantelli for $Z^{(k)}_{4,N}(\omega)$)
For some $r_4 > 0$ and some sequence $b_{4,N} \geq 1$, $N \in \mathbf{N}$, $\sum_{N=1}^{\infty} \frac{n}{(b_{4,N})^{r_4}} < +\infty$ and $\sup_{N\in\mathbf{N}} E[|Z_{4,N}(\cdot)|^{r_4}] < +\infty$ for each fixed $m \in \mathbf{N}$, where

$$Z^{(k)}_{4,N}(\omega) := a_N^{-1}\left(\sup_{(\mathbf{x},\theta)\in B^N} |\partial_x X^{x,(k)}_{(m)}(\theta;\omega)| + \sup_{(\mathbf{x},\theta)\in B^N} |\partial_\theta X^{x,(k)}_{(m)}(\theta;\omega)|\right).$$

(vi). (Borel-Cantelli for $\dot{Z}^{(k)}_{4,N}(\omega)$)
For some $\dot{r}_4 > 0$ and some sequence $\dot{b}_{4,N} \geq 1$, $N \in \mathbf{N}$, $\sum_{N=1}^{\infty} \frac{n}{(\dot{b}_{4,N})^{\dot{r}_4}} < +\infty$ and $\sup_{N\in\mathbf{N}} E[|\dot{Z}^{(k)}_{4,N}(\cdot)|^{\dot{r}_4}] < +\infty$ for each fixed $m \in \mathbf{N}$, where

$$\dot{Z}^{(k)}_{4,N}(\omega) := a_N^{-1}\left(h \sup_{(\mathbf{x},\theta)\in B^N} |\partial_x \partial_\theta X^{x,(k)}_{(m)}(\theta;\omega)| + h \sup_{(\mathbf{x},\theta)\in B^N} |\partial_\theta \partial_\theta X^{x,(k)}_{(m)}(\theta;\omega)|\right.$$

$$+(\dot{Z}^{(k)}_{4,N}+1)\sup_{(\mathbf{x},\theta)\in B^N}|\partial_\theta X^{x,(k)}_{(m)}(\theta;\omega)|\Bigg).$$

(viii). (Borel-Cantelli for $\dot{Z}^{(k)}_{6,N}(\omega)$)
For some $\dot{r}_6>0$ and some sequence of positive numbers $\dot{b}_{6,N}$, $\sum_{N=1}^\infty \frac{n}{(\dot{b}_{6,N})^{\dot{r}_6}}<+\infty$ and $\sup_N E[|\dot{Z}_{6,N}(\cdot)|^{\dot{r}_6}]<+\infty$ for each fixed $m\in\mathbf{N}$, where

$$\dot{Z}^{(k)}_{6,N}(\omega):=a_N^{-1}\sup_{(\mathbf{x},\theta)\in B^N}\left\{|\partial_\theta X^{x,(k)}_{(m)}(\theta;\omega)|+E[|\partial_\theta X^{x,(1)}_{(m)}(\theta;\cdot)|]\right\}.$$

(ix). For some $\dot{\alpha}_6>0$, $\dot{q}_6>1$ and $\dot{C}_6>0$, and some positive sequence η_N,

$$\left(\frac{\eta_N h^2}{(\|K'\|_\infty \dot{b}_{6,N})^2 a_N}\exp\left(-\frac{(\eta_N)^2}{2(\frac{\|K'\|_\infty}{h^2}\dot{b}_{6,N}a_N)^2}\right)\right)^{\dot{q}_6}\le\frac{\dot{C}_6}{n^{1+\dot{\alpha}_6}}$$

and $\sup_{N\in\mathbf{N}}E[|\dot{Z}_{6,N}(\cdot)|^{\dot{q}_6}]<+\infty$.

Set $a_N:=\sqrt{c_2\ln N}$ for some positive constant c_2. Set $n=C_1N^{\alpha_1}$ for $\alpha_1,C_1>0$ and $h=C_2N^{-\alpha_2}$ for $\alpha_2,C_2>0$. And also set $b_{3,N}=\frac{C_3(N^{\gamma_3}n)^{\frac{1}{r_3}}c_2\ln N}{h^2}$ for $\gamma_3>1$, and $\dot{b}_{6,N}=(\dot{C}_6 nN^{\dot{\gamma}_6})^{\frac{1}{\dot{r}_6}}$. Then we obtain the following result.

Theorem 2. *(Kohatsu-Higa et al. [9]) Assume that the constants are chosen so as to satisfy $c_1>\frac{2}{c_2}$,*

$$\left(4\alpha_2+2\frac{\alpha_1+\dot{\gamma}_6}{\dot{r}_6}+\frac{\varphi_2c_2}{\Delta}+\frac{1}{2}+\frac{\gamma_3}{r_3}+\frac{\alpha_1}{r_3}\right)\dot{q}_6>\alpha_1, \tag{5}$$

$$\alpha_1\left(1-\frac{2}{r_3}-\frac{2}{\dot{r}_6}\right)>8\alpha_2+1+\frac{2\varphi_2c_2}{\Delta}+\frac{2\gamma_3}{r_3}+2\frac{\dot{\gamma}_6}{\dot{r}_6}. \tag{6}$$

Furthermore, assume that the moment conditions stated in **(ii)**, **(iii)**, **(iv)**, **(vi)**, **(viii)** *and* **(ix)** *above are satisfied. If additionally, we assume* **(H1)**, **(H2)**, **(H5)**, **(H5')**, *then Assumption (A) (6)-***(a)** *is satisfied.*

Furthermore if all other conditions on Assumption (A) are satisfied then there exist some positive finite random variables Ξ_1 and Ξ_2 such that

$$|E_N[f]-f(\theta_0)|\le\frac{\Xi_1}{\sqrt{N}}\ a.s.\quad and\quad \left|E^n_{N,m}[f]-f(\theta_0)\right|\le\frac{\Xi_2}{\sqrt{N}}\ a.s.,$$

and therefore $|E_N[f]-E^n_{N,m}[f]|\le\frac{\Xi_1+\Xi_2}{\sqrt{N}}$ a.s.

*Remark 3.***(i).** In (6), r_3 and $\dot{r}_6$ represent moment conditions on the derivatives of $X^x_{(m)}(\theta)$, φ_2^{-1} represents the variance of $X^x_{(m)}(\theta)$, Δ represents the length of the time interval between observations. Finally $c_2>2c_1^{-1}$ expresses a moment condition on Y_i. In (5), recall that $\dot{q}_6$ determines a moment condition on $X^x_{(m)}(\theta)$.

(ii). Roughly speaking, if $r_3, \dot{r}_6$ and $\dot{q}_6$ are big enough (which implies a conditions on n), and we choose $\alpha_1 > 8\alpha_2 + 1 + \frac{2\varphi_2 c_2}{\Delta}$, $m = \sqrt{N}$, $h = C_2 N^{-\alpha_2}$ and $n = C_1 N^{\alpha_1}$, then Assumption (A) (6)-**(a)** and (6)-**(b)** are satisfied. Then conditions contain the main tuning requirements. (See Proposition 10)

3 The Ornstein-Uhlenbeck Process

We consider the following Ornstein-Uhlenbeck process; without loss of generality for $\theta \in [\alpha, \beta]$, where $0 < \alpha < \beta < 2$,

$$dX_t = -\theta X_t dt + dW_t, \quad X_0 = x, \tag{7}$$

where W_t is a 1-dimensional Brownian motion. Then we can write the solution explicitly as $X_t = X_s e^{-\theta(t-s)} + \int_s^t e^{-\theta(t-u)} dW_u$. As it is well known, the OU Process has the following expectation, variance and covariance: For $s < t$,

$$\mu(X_s, t-s, \theta) := X_s \mu(t-s, \theta) := E[X_t | X_s] = X_s e^{-\theta(t-s)},$$
$$\sigma^2_{t-s}(\theta) := Var(X_t | X_s) = \frac{1}{2\theta} - \frac{1}{2\theta} e^{-2\theta(t-s)},$$
$$Cov(X_t, X_s) := \frac{1}{2\theta} e^{-\theta(t-s)} - \frac{1}{2\theta} e^{-\theta(t+s)}.$$

From moment results for the normal distribution, moments of the OU Process can also be bounded as follows

$$E\left[\left|X_t - X_s e^{-\theta(t-s)}\right|^{2k}\right] = E\left[\left|\int_s^t e^{-\theta(t-u)} dW_u\right|^{2k}\right] = \frac{(2k)!}{2^{2k} k!} \left(\frac{1 - e^{-2\theta(t-s)}}{\theta}\right)^k.$$

Therefore in particular for $s = 0$, by using the Minkowski's inequality,

$$E\left[X_t^{2k}\right] \le C_k \left(\left(\frac{1 - e^{-2\theta t}}{\theta}\right)^k + E\left[X_0^{2k}\right]\right). \tag{8}$$

The conditional density of X_t given X_s is given by

$$p_\theta(X_s, x; s, t) := q\left(X_s, x; \mu(t-s, \theta), \sigma^2_{t-s}(\theta)\right), \tag{9}$$

where $q(y, z; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(z - y\mu)^2}{2\sigma^2}}$. Note that $p_\theta(X_s, y) = p_\theta(X_s, y; s, s + \Delta)$.

3.1 The Euler-Maruyama Approximation of the OU process

For $m \in \mathbf{N}$ and $i = 1, \cdots, m$, $X^x_{i,m}(\theta) := X^x_{i-1,m}(\theta) - \theta X^x_{i-1,m}(\theta)\Delta t + \Delta_{i-1} W$, where $X^x_0(\theta) = x$, $\Delta t = t_i - t_{i-1} = \frac{\Delta}{m}$ and $\Delta_i W = W_{t_{i+1}} - W_{t_i}$. Set $X^x_{(m)}(\theta) = X^x_{m,m}(\theta)$. We will find an explicit expression for this approximation by induction. In fact,

$$\bar{X}_{t_1} = x(1-\theta\Delta t) + W_{t_1},$$

for $\Delta t = \frac{\Delta}{m}$. Similarly for $\Delta_i W = W(t_{i+1}) - W(t_i)$,

$$\bar{X}_{t_2} = (x(1-\theta\Delta t) + \Delta_0 W)(1-\theta\Delta t) + \Delta_1 W.$$

Therefore in general, we have that

$$X^x_{(m)}(\theta) = \bar{X}_{t_m} = x(1-\theta\Delta t)^m + \sum_{i=0}^{m-1} \Delta_i W (1-\theta\Delta t)^{m-1-i}. \tag{10}$$

From the above expression, we can easily find that $X^x_{(m)}(\theta)$ follows the normal distribution with mean $\mu(x,m,\theta)$ and variance $\sigma^2(m,\theta)$;

$$\mu(x,m,\theta) = x\mu(m,\theta) = x(1-\frac{\theta\Delta}{m})^m,$$
$$\sigma^2(m,\theta) = \frac{(1-\frac{\theta\Delta}{m})^{2m} - 1}{\theta(\frac{\theta}{m} - 2)},$$

where we exclude $\frac{\theta}{m} = 2$. For example, if we take $\beta < 2$, then since $m \in \mathbf{N}$, we always have $\frac{\theta}{m} < 2$ for $\theta \in [\alpha, \beta]$, where $0 < \alpha < \beta < 2$. Then the transition density $\tilde{p}^{(m)}_\theta(x,y) \equiv \tilde{p}^N_\theta(x,y)$ is given as follows;

$$\tilde{p}^{(m)}_\theta(x,y) = q\left(x,y,\mu(m,\theta),\sigma^2(m,\theta)\right).$$

Next, we can write $\bar{p}^N_\theta(x,y)$ as follows; set $w = \frac{z-y}{h}$,

$$\bar{p}^N_\theta(x,y) = E\left[\tilde{p}^{(m)}_\theta(x, hX + y)\right] = \frac{d}{dy} P\left(X^{x,(1)}_{(m)}(\theta,\cdot) - hX \le y\right),$$

where $X \sim N(0,1)$ is a random variable with the standard normal distribution. Now $X^{x,(1)}_{(m)}(\theta,\cdot) \sim N(x(1-\frac{\theta\Delta}{m})^m, \frac{(1-\frac{\theta\Delta}{m})^{2m}-1}{\theta(\frac{\theta}{m}-2)})$ and is independent of X. Then $X^{x,(1)}_{(m)}(\theta,\cdot) - hX \sim N(x(1-\frac{\theta\Delta}{m})^m, \frac{(1-\frac{\theta\Delta}{m})^{2m}-1}{\theta(\frac{\theta}{m}-2)} + h^2)$. Therefore

$$\bar{p}^N_\theta(x,y) = q\left(x,y,\mu(m,\theta),\sigma^2(m,\theta,h)\right), \tag{11}$$

where $\sigma^2(m,\theta,h) = \frac{(1-\frac{\theta\Delta}{m})^{2m}-1}{\theta(\frac{\theta}{m}-2)} + h^2$.

Proposition 1. *(Density conditions for $\overline{p}_\theta^N(x,y)$) $\overline{p}_\theta^N(x,y)$ satisfies the hypotheses* **(H1)**, **(H5)** *and* **(H5')**.

The proof follows directly from Lemma 10 in Appendix. In fact, in the OU process case, we can take $\varphi_2 = \frac{6\beta}{\alpha}$ if $0 < 2\alpha\Delta \le \ln 2$, and $\varphi_1 = \frac{1}{\sqrt{2\pi(C(0,\Delta,\alpha)+1)}}$, where $C(0,\Delta,\alpha)$ is given in Lemma 10.

3.2 About Assumption (A) (1)∼(5)

In this section, we will consider Assumption (A) (1)∼(5) for the OU process and its Euler-Maruyama approximation. Assumption (A) (6) will be discussed in the following sections.

Proposition 2. *The OU process satisfies Assumption (A) (1).*

Proof. From Proposition 3 in pp.115 of Doukhan [5], we obtain that the OU process has the geometrically strong mixing property. The OU process satisfies Assumption (A) (1). □

Once we take a prior distribution $\pi(\theta)$ as the uniform distribution on Θ. It satisfies Assumption (A) (2).

Set $\mu_\theta(x) := \sqrt{\frac{\theta}{\pi}} \exp(-\theta x^2)$.

Lemma 1. *$\mu_\theta(x)$ is the probability density function of the invariant measure for the OU process (7).*

Proposition 3. *The OU process and its Euler-Maruyama approximation satisfy Assumption (A) (3).*

Proof. From the expression (9) of the transition density

$$p_\theta(y,z) = p_\theta(y,z;s,s+\Delta)$$

of the OU process, and also, from the assumption of the kernel K, $p_\theta(y,z)$ and $\overline{p}_\theta^N(x,y)$ clearly satisfy Assumption (A) (3), that is, it is continuous in x,y and twice continuously differentiable in θ [1] . And from Lemma 1, the OU process satisfies Assumption (A) (3). □

Now we consider the identifiability condition for p in Assumption (A) (4).

Proposition 4. *The OU process satisfies Assumption (A) (4) for p.*

[1] Note that the solution $X_{(m)}^x(\theta)$ is twice continuously differentiable in θ, since from the definition of the Euler-Maruyama approximation, the OU process is polynomial in θ and the kernel $K(x)$ is infinitely differentiable in x.

Proof. First note that the identifiability condition for p is equivalent to

$$\infty > \int \left(\inf_{\theta\in\Theta} \int \frac{|p_\theta(x,y) - p_{\theta_0}(x,y)|}{|\theta-\theta_0|} dy \right)^2 \mu_{\theta_0}(x)dx \geq \int c(x)^2 \mu_{\theta_0}(x)dx > 0.$$

By using the fundamental theorem of calculus and changing variables; set

$$\beta = \alpha\theta + (1-\alpha)\theta_0,$$

we have

$$\begin{aligned}\infty &> \int \left(\inf_{\theta\in\Theta} \int \left| \int_0^1 \partial_\theta p_{\alpha\theta+(1-\alpha)\theta_0}(x,y)d\alpha \right| dy \right)^2 \mu_{\theta_0}(x)dx \\ &\geq \int c(x)^2 \mu_{\theta_0}(x)dx > 0.\end{aligned}$$

The integrability (upper estimation) is easily obtain as p_θ is a normal density function. That is, set $\theta' = \mathrm{argmax}_{\theta\in\Theta}|\partial_\theta p_\theta(x,y)|$, then from (8) and using the inequalities $(a+b)^2 \leq 2(a^2+b^2)$ and $2|ab| \leq (a^2+b^2)$,

$$\begin{aligned}&\int \left(\inf_{\theta\in\Theta} \int \left| \int_0^1 \partial_\theta p_{\alpha\theta+(1-\alpha)\theta_0}(x,y)d\alpha \right| dy \right)^2 \mu_{\theta_0}(x)dx \\ &\leq 2 \int\int\int_0^1 \frac{p^2_{\alpha\theta+(1-\alpha)\theta_0}(x,y)}{m_0^2} \\ &\qquad \times \left(M_1^2 + 16(t-s)^2(|y|^4+|x|^4) + \frac{32\left(y^4+x^4\right)}{m_0^2} M_1^2 \right) d\alpha dy \mu_{\theta_0}(x)dx < \infty.\end{aligned}$$

Here $\theta' = \alpha\theta + (1-\alpha)\theta_0$ and $E_{\theta_0}\left[X_0^{2k}\right] = \frac{k!}{(4\theta)^k}$. Therefore the above is finite.

Now $\mu_{\theta_0}(x) > 0$ for all $x \in \mathbf{R}$. Therefore it is enough to prove that

$$\inf_{\theta\in\Theta} \int |\int_0^1 \partial_\theta p_{\alpha\theta+(1-\alpha)\theta_0}(x,y)d\alpha|dy > 0,$$

for all $x \in \mathbf{R}$. We use proof by contradiction. We assume that

$$\inf_{\theta\in\Theta} \int |\int_0^1 \partial_\theta p_{\alpha\theta+(1-\alpha)\theta_0}(x,y)d\alpha|dy = 0.$$

This is equivalent that for all $x \in \mathbf{R}$, there exists some $\theta^* = \theta^*(x)$ such that

$$\int \left| \int_0^1 \partial_\theta p_{\alpha\theta^*+(1-\alpha)\theta_0}(x,y)d\alpha \right| dy = 0.$$

Then for all $x \in \mathbf{R}$, there exists some $\theta^* = \theta^*(x)$ such that for all $y \in \mathbf{R}$,

$$\left|\int_0^1 \partial_\theta p_{\alpha\theta^*+(1-\alpha)\theta_0}(x,y)d\alpha\right| = 0.$$

This means that for all $x \in \mathbf{R}$, there exists some $\theta^* = \theta^*(x)$ such that for all $y \in \mathbf{R}$, $p_{\theta^*}(x,y) = p_{\theta_0}(x,y)$. As both density functions are Gaussian then the point where the maximum is taken has to be the same. Therefore the mean values are equal. Similarly if we take y equal to the common mean we obtain that the variances have to be equal. Then analyzing the variance function, we have that it is decreasing in θ, therefore $\theta^* = \theta_0$. □

And by using the similar argument, we obtain the identifiability condition for $\overline{p}^N$.

Proposition 5. *The Euler-Maruyama approximation of the OU process satisfies Assumption (A) (4) for* $\overline{p}^N$.

Proof. Set $B := \int \{\inf_\theta \inf_N \int \frac{|\overline{p}^N_\theta(x,y)-\overline{p}^N_{\theta_0}(x,y)|}{|\theta-\theta_0|}dy\}^2 \mu_{\theta_0}(x)dx \in (0,+\infty)$. As before, it is easy to prove $B < +\infty$.

Here we also use proof by contradiction. If $B = 0$, then from the assumption of $\mathrm{supp}\mu_\theta(x) = \mathbf{R}$, we have, for all $x \in \mathbf{R}$, $\inf_\theta \inf_N \int \frac{|\overline{p}^N_\theta(x,y)-\overline{p}^N_{\theta_0}(x,y)|}{|\theta-\theta_0|}dy = 0$. Then for all $x \in \mathbf{R}$, there exists some sequence $\theta_n = \theta_n(x)$ such that

$$\lim_{n\to\infty} \inf_N \int \frac{|\overline{p}^N_{\theta_n}(x,y) - \overline{p}^N_{\theta_0}(x,y)|}{|\theta_n - \theta_0|}dy = 0.$$

And also, for all $x \in \mathbf{R}$, there exists some sequence $\theta_n = \theta_n(x)$ such that there exists some sequence $N_n = N_n(x,\theta_n)$ such that $\lim_{n\to\infty} \int \frac{|\overline{p}^{N_n}_{\theta_n}(x,y)-\overline{p}^{N_n}_{\theta_0}(x,y)|}{|\theta_n-\theta_0|}dy = 0$.

By using the mean-value theorem, we consider the following;

$$\lim_{n\to\infty} \int \left|\int_0^1 \partial_\theta \overline{p}^{N_n}_{\alpha\theta_n+(1-\alpha)\theta_0}(x,y)d\alpha\right| dy = 0.$$

Then we obtain our conclusion. □

Note that from Lemma 1, we have $E[X_0^{2k}] = \frac{(2k)!}{(4\theta)^k k!}$, and from (8), we have

$$E\left[X_t^{2k}\right] \le C_k \left(\left(\frac{1-e^{-2\theta t}}{\theta}\right)^k + \theta^{-k}\right). \tag{12}$$

Proposition 6. *For the OU process and its Euler-Maruyama approximation, the first regularity conditions (1) of Assumption (A) (5) hold.*

Proof. Using (12), we obtain

$$\begin{aligned}
&\sup_{\theta\in\Theta} \int\int (\ln p_\theta(y,z))^{12} p_{\theta_0}(y,z)\mu_{\theta_0}(y)dydz \\
&= \sup_{\theta\in\Theta} E\left[\left(-\frac{1}{2}\log\left(2\pi\sigma^2_\Delta(\theta)\right) - \frac{\left(X_\Delta(\theta_0) - X_0(\theta_0)e^{-\theta\Delta}\right)^2}{2\sigma^2_\Delta(\theta)}\right)^{12}\right]
\end{aligned}$$

$$\leq C\left\{\sup_{\theta\in\Theta}\log^{12}\left(2\pi\sigma_\Delta^2(\theta)\right)+\sup_{\theta\in\Theta}\sigma_\Delta^{-24}(\theta)E\left[X_\Delta(\theta_0)^{24}+X_0(\theta_0)^{24}e^{-24\theta\Delta}\right]\right\}<\infty. \tag{13}$$

Now $\sigma_\Delta^2(\theta)=\frac{1}{2\theta}(1-e^{-2\theta\Delta})$. Note that $\sigma_\Delta^2(\theta)\geq\frac{1}{2\beta}(1-e^{-2\alpha\Delta})>0$ and also $\sigma_\Delta^2(\cdot)\in C_b^\infty([\alpha,\beta])$. Furthermore let $m(\theta)=e^{-\theta\Delta}$. Note that $m(\cdot)\in C_b^\infty([\alpha,\beta])$. Then by using similar arguments as in the above calculations, we obtain (1), for $q_\theta=p_\theta$ and $i=1,2$. We can also obtain our integrabilities for $q_\theta=p_\theta$.

Therefore, as $\bar{X}_{(m)}^{x,(1)}(\theta,\cdot)$ has the density $\bar{p}_\theta^N(x,y)$ at y (see (11)) then

$$E\left[|\bar{X}_{(m)}^{x,(1)}(\theta,\cdot)|^{2k}\right]\leq C_k\left(\left(\frac{(1-\frac{\theta\Delta}{m})^{2m}-1}{\theta(\frac{\theta}{m}-2)}+h^2\right)^k+x^{2k}\left(1-\frac{\theta\Delta}{m}\right)^{2km}\right).$$

From Lemma 10 in Section 4.1, (12) and as (13), we have

$$\begin{aligned}&\sup_N\sup_{\theta\in\Theta}\int\int\left(\log\bar{p}_\theta^N(y,z)\right)^{12}p_{\theta_0}(y,z)\mu_{\theta_0}(y)dydz\\&\leq C\sup_N\sup_{\theta\in\Theta}\left\{\left(\log\left(\frac{1}{\sqrt{2\pi\sigma^2(m,\theta,h)}}\right)\right)^{12}\right.\\&\qquad\left.+C'\frac{E[X_\Delta(\theta_0)^{24}]+E[X_0(\theta_0)^{24}]\mu(m,\theta)^{24}}{2^{12}\sigma^{24}(m,\theta,h)}\right\}<+\infty.\end{aligned}$$

And for $i=1,2$, as in the above, we obtain (1) for $q_\theta=\bar{p}_\theta^N$. Then we obtain our conclusions. □

Now we check the second condition of the regularity of the log-density (Assumption (A) (5)).

Proposition 7. *For the OU process and its Euler-Maruyama approximation, the second regularity conditions (2) of Assumption (A) (5) hold.*

Proof. For $q_\theta=p_\theta$, we have

$$\begin{aligned}&\int\int\left(-\frac{1}{2}\log\left(2\pi\sigma_\Delta^2(\theta)\right)-\frac{\left(z-ye^{-\theta\Delta}\right)^2}{2\sigma_\Delta^2(\theta)}\right)\bar{p}_{\theta_0}^N(y,z)\mu_{\theta_0}(y)dydz\\&=-\frac{1}{2}\log\left(2\pi\sigma_\Delta^2(\theta)\right)-\frac{\sigma^2(m,\theta_0,h)+(2\theta_0)^{-1}\left(\mu(m,\theta_0)-e^{-\theta\Delta}\right)^2}{2\sigma_\Delta^2(\theta)}.\end{aligned}$$

Therefore the result follows because $\sigma_\Delta^2(\theta)$ is twice continuously differentiable and the above quantities are uniformly bounded in m.

Next we will check equation (2) for $q_\theta=\bar{p}_\theta^N$, Then as before,

$$\int\int\left\{\log\left(\frac{1}{\sqrt{2\pi\sigma^2(m,\theta,h)}}\right)-\frac{(y-x\mu(m,\theta))^2}{2\sigma^2(m,\theta,h)}\right\}\bar{p}_{\theta_0}^N(y,z)\mu_{\theta_0}(y)dydz$$

$$= \log\left(\frac{1}{\sqrt{2\pi\sigma^2(m,\theta,h)}}\right) - \frac{1}{2}.$$

Therefore the property follows as in the previous case. □

Next we consider the third regularity condition of Assumption (A) (5).

Proposition 8. *For the OU process and its Euler-Maruyama approximation, the third regularity conditions (3) of Assumption (A) (5) hold.*

Proof. For $i = 0, 1$, we have

$$\int\int \left(-\frac{1}{2}\frac{\partial^i}{\partial\theta^i}\log\left(2\pi\sigma_\Delta^2(\theta)\right) - \frac{\partial^i}{\partial\theta^i}\frac{\left(z - ye^{-\theta\Delta}\right)^2}{2\sigma_\Delta^2(\theta)}\right)\overline{p}_{\theta_0}^N(y,z)\mu_{\theta_0}(y)dydz$$
$$= -\frac{1}{2}\frac{\partial^i}{\partial\theta^i}\log\left(2\pi\sigma_\Delta^2(\theta)\right) - \left(\frac{\partial^i}{\partial\theta^i}\frac{1}{2\sigma_\Delta^2(\theta)}\right)E\left[\left(\overline{X}_{(m)}^{X_0,(1)}(\theta_0,\cdot) - X_0e^{-\theta\Delta}\right)^2\right]$$
$$- \frac{E\left[\frac{\partial^i}{\partial\theta^i}(\overline{X}_{(m)}^{X_0,(1)}(\theta_0,\cdot) - X_0e^{-\theta\Delta})^2\right]}{2\sigma_\Delta^2(\theta)}.$$

If we expand the last expectation in the above expression, it is clear that

$$\frac{\partial^i}{\partial\theta^i}E\left[\left(\overline{X}_{(m)}^{X_0,(1)}(\theta_0,\cdot) - X_0e^{-\theta\Delta}\right)^2\right] = E\left[\frac{\partial^i}{\partial\theta^i}\left(\overline{X}_{(m)}^{X_0,(1)}(\theta_0,\cdot) - X_0e^{-\theta\Delta}\right)^2\right].$$

And therefore the last property of Assumption (A) (5) follows for $q_\theta = p_\theta$. A similar proof also applies in the case $q_\theta = \overline{p}_\theta^N$. □

3.3 Assumption (A) (6)

3.3.1 Parameter Tuning of Assumption (A) (6)-(a)

If we choose $0 < c_1 < \alpha$, the moment hypothesis of **(ii)** in Section 2.3,

$$E[e^{c_1|Y_1|^2}] < \infty,$$

is satisfied since Y_1 has a normal distribution. Furthermore as $a_N = \sqrt{c_2 \ln N}$ with $c_1 > \frac{2}{c_2}$, then condition **(ii)** is satisfied.

From the explicit expression (10) of the OU process, we have the following derivatives of the Euler-Maruyama approximation of the OU process.

$$\partial_x X_{(m)}^x(\theta) = (1-\theta\Delta t)^m, \tag{14}$$
$$\partial_\theta X_{(m)}^x(\theta) = -mx\Delta t(1-\theta\Delta t)^{m-1} - \Delta t\sum_{i=0}^{m-2}(m-1-i)\Delta_i W\,(1-\theta\Delta t)^{m-2-i},$$

$$\partial_\theta \partial_x X^x_{(m)}(\theta) = -m\Delta t(1-\theta\Delta t)^{m-1},$$
$$\partial^2_\theta X^x_{(m)}(\theta) = m(m-1)\Delta t^2(1-\theta\Delta t)^{m-2} + \Delta t^2 \sum_{i=0}^{m-3}(m-1-i)(m-2-i)\Delta_i W\,(1-\theta\Delta t)^{m-3-i}.$$

Lemma 2. *For any* $j \in \mathbf{N}$,

$$\sup_{m\geq j\vee(\theta\Delta)} \sup_{(\mathbf{x},\theta)\in B^N} a_N^{-1}\left| x\frac{\partial^j}{\partial\theta^j}(1-\theta\Delta t)^m\right| < +\infty.$$

Proof. From the definition of B^N, it is clear that $\sup_{(\mathbf{x},\theta)\in B^N} a_N^{-1}|x| \leq 1$. Next, we have

$$\left|\frac{\partial^j}{\partial\theta^j}(1-\theta\Delta t)^m\right| = \left|(-1)^j 1(1-\frac{1}{m})\cdots(1-\frac{j-1}{m})\Delta^j(1-\theta\Delta t)^{m-j}\right| \leq \Delta^j\left(1-\frac{\theta\Delta}{m}\right)^{m-j}.$$

Set $y = -\frac{m}{\theta\Delta}$. For any m and j so that $m \geq j\vee(\theta\Delta)$, we have

$$\left(1-\frac{\theta\Delta}{m}\right)^{m-j} = \left\{\left(1+\frac{1}{y}\right)^y\right\}^{-\frac{\theta\Delta(m-j)}{m}} \leq e^{-\frac{\theta\Delta(m-j)}{m}} \leq 1, \qquad (15)$$

where we use Lemma 9. Hence we obtain our conclusion. □

For some differentiable function $h(\theta,t)$, set $U(\theta) := \int_0^\Delta h(\theta,s)dW_s$. Then

$$U'(\theta) = \int_0^\Delta \frac{\partial}{\partial\theta}h(\theta,s)dW_s.$$

Lemma 3. *We assume that there exists some positive constant* $C(\Delta)$, *which depends on* Δ, *such that* $\sum_{j=0}^1 \sup_{\substack{\theta\in[\alpha,\beta]\\ t\in[0,\Delta]}} |\frac{\partial^j}{\partial\theta^j}h(\theta,t)| \leq C(\Delta)$. *Then, for* $p\in\mathbf{N}$, *we have*

$$E\left[\sup_{\theta\in[\alpha,\beta]}|U(\theta)|^{2p}\right] \leq C(\Delta)^{2p}\Delta^p(1+(\beta-\alpha)^{2p})\frac{(2p)!}{p!}.$$

Proof. Note that $U(\theta) = U(\alpha) + \int_\alpha^\theta U'(\rho)d\rho$ a.s. From the Hölder's inequality and Fubini's theorem, we have

$$E\left[\sup_{\theta\in[\alpha,\beta]}|U(\theta)|^{2p}\right] \leq 2^p\left(E\left[|U(\alpha)|^{2p}\right] + (\beta-\alpha)^{2p-1}\int_\alpha^\beta E\left[|U'(\rho)|^{2p}\right]d\rho\right). \qquad (16)$$

Now note that $U(\theta)$ and $U'(\theta)$ have normal distribution with mean 0 and variance $\int_0^\Delta h(\theta,s)^2 ds$ and $\int_0^\Delta (\frac{\partial}{\partial\theta} h(\theta,s))^2 ds$ each other. Then from moment properties of Gaussian distribution. We have that

$$E\left[|U(\alpha)|^{2p}\right] = \left(\int_0^\Delta h(\alpha,s)^2 ds\right)^p \frac{(2p)!}{2^p p!} \leq \left(C(\Delta)^2\Delta\right)^p \frac{(2p)!}{2^p p!},$$
$$E\left[|U'(\rho)|^{2p}\right] = \left(\int_0^\Delta \left(\frac{\partial}{\partial\theta} h(\rho,s)\right)^2 ds\right)^p \frac{(2p)!}{2^p p!} \leq \left(C(\Delta)^2\Delta\right)^p \frac{(2p)!}{2^p p!}.$$

Finally we have $(16) \leq \left(C(\Delta)^2\Delta\right)^p \frac{(2p)!}{p!}\left(1+(\beta-\alpha)^{2p}\right)$. Hence we obtain our conclusion. □

We note that $\sum_{i=0}^{m-1} \Delta_i W (1-\theta\Delta t)^{m-1-i} = \int_0^\Delta h_m(\theta,s) dW_s$, where

$$h_m(\theta,t) = (1-\theta\Delta t)^{m-1-i},$$

for $t \in [t_i, t_{i+1})$ and $i = 0,1,\cdots,m-1$. Also we have

$$\frac{\partial}{\partial\theta} h_m(\theta,t) = (m-1-i)(-\Delta t)(1-\theta\Delta t)^{m-2-i},$$

for $t \in [t_i, t_{i+1})$, $i = 0,1,\cdots,m-2$ and $=0$ for $t \in [t_{m-1}, t_m]$, $i = m-1$. Note that from (15), we have, for $m \geq \theta\Delta$,

$$|h_m(\theta,t)| \leq 1 \quad \text{and} \quad \left|\frac{\partial}{\partial\theta} h_m(\theta,t)\right| \leq \Delta.$$

Next we consider that

$$\sum_{i=0}^{m-2} \Delta t(m-1-i)\Delta_i W \left(1-\theta\Delta t\right)^{m-2-i} = \int_0^\Delta h_m^{(1)}(\theta,s) dW_s,$$

where

$$h_m^{(1)}(\theta,t) = \begin{cases} \Delta t(m-1-i)(1-\theta\Delta t)^{m-2-i}, & \text{for } t \in [t_i, t_{i+1}),\ i = 0,1,\cdots,m-2, \\ = 0, & \text{for } t \in [t_{m-1}, t_m]. \end{cases}$$

And also we have

$$\frac{\partial}{\partial\theta} h_m^{(1)}(\theta,t) = \begin{cases} -\Delta t^2(m-1-i)(m-2-i)(1-\theta\Delta t)^{m-3-i}, & \\ & \text{for } t \in [t_i, t_{i+1}),\ i = 0,1,\cdots,m-3, \\ = 0, & \text{for } t \in [t_{m-2}, t_m]. \end{cases}$$

Then as before, from (15), we have, for $m \geq \theta\Delta$,

$$|h_m^{(1)}(\theta,t)| \leq \Delta \quad \text{and} \quad \left|\frac{\partial}{\partial\theta} h_m^{(1)}(\theta,t)\right| \leq \Delta^2.$$

As above, we consider

$$\sum_{i=0}^{m-3} \Delta t^2 (m-1-i)(m-2-i)\Delta_i W(1-\theta\Delta t)^{m-3-i} = \int_0^{\Delta} h_m^{(2)}(\theta,s)dW_s,$$

where

$$h_m^{(2)}(\theta,t) = \begin{cases} \Delta t^2(m-1-i)(m-2-i)(1-\theta\Delta t)^{m-3-i}, & \\ & \text{for } t\in[t_i,t_{i+1}),\ i=0,1,\cdots,m-3, \\ =0, & \text{for } t\in[t_{m-2},t_m]. \end{cases}$$

And also we have

$$\frac{\partial}{\partial\theta}h_m^{(2)}(\theta,t) = \begin{cases} -\Delta t^3(m-1-i)(m-2-i)(m-3-i)(1-\theta\Delta t)^{m-4-i}, & \\ & \text{for } t\in[t_i,t_{i+1}),\ i=0,1,\cdots,m-4, \\ =0, & \text{for } t\in[t_{m-3},t_m]. \end{cases}$$

Then as before, from (15), we have, for $m \geq \theta\Delta$,

$$|h_m^{(2)}(\theta,t)| \leq \Delta^2 \quad \text{and} \quad \left|\frac{\partial}{\partial\theta}h_m^{(2)}(\theta,t)\right| \leq \Delta^3.$$

Lemma 4. *For $H_m(\theta,t) = h_m(\theta,t),\ h_m^{(1)}(\theta,t),\ h_m^{(2)}(\theta,t)$, we have, for $p\in\mathbf{N}$,*

$$\sup_{m\in\mathbf{N}} E\left[\sup_{\theta\in[\alpha,\beta]}\left|\int_0^{\Delta} H_m(\theta,s)dW_s\right|^{2p}\right] < +\infty.$$

Proof. From the calculations before the lemma, we have found that H_m satisfies the assumption of Lemma 3 as we take $C(\Delta) = 1\vee\Delta^3$. Then we apply Lemma 3.

$$E\left[\sup_{\theta\in[\alpha,\beta]}\left|\int_0^{\Delta} H_m(\theta,s)dW_s\right|^{2p}\right] \leq C(\Delta)^{2p}\Delta^p(1+(\beta-\alpha)^{2p})\frac{(2p)!}{p!}.$$

The right hand side does not depend on m. Then we take sup with respect to $m\in\mathbf{N}$ for the left hand side and we have our conclusion. □

From the above lemmas and the explicit formulas (14), we obtain the following two results.

Lemma 5. *For all $p \geq 1$ and $k\in\mathbf{N}$, we have*

$$\sup_{N\in\mathbf{N}} E\left[\left(a_N^{-1}\sup_{(\mathbf{x},\theta)\in B^N}\left|V_{(m)}^{x,(k)}(\theta;\cdot)\right|\right)^p\right] < +\infty,$$

for $V_{(m)}^{x,(k)}(\theta;\omega) = X_{(m)}^{x,(k)}(\theta;\omega), \partial_x X_{(m)}^{x,(k)}(\theta;\omega), \partial_\theta X_{(m)}^{x,(k)}(\theta;\omega), \partial_\theta\partial_x X_{(m)}^{x,(k)}(\theta;\omega)$, $\partial_\theta^2 X_{(m)}^{x,(k)}(\theta;\omega)$.

Proposition 9. *(Moment conditions of* **(iii)**, **(iv)**, **(vi)**, **(viii)** *and* **(ix)** *in Section 2.3) For all* $p \geq 1$, *we have* $\sup_{N\in\mathbf{N}} E[|T_N(\cdot)|^p] < +\infty$, *for* $T_N(\omega) = Z_{3,N}(\omega),\ Z_{4,N}(\omega), \dot{Z}_{4,N}(\omega), \dot{Z}_{6,N}(\omega)$.

From the above result, we obtain enough integrability for $Z_{3,N}(\omega),\ Z_{4,N}(\omega), \dot{Z}_{4,N}(\omega),\ \dot{Z}_{6,N}(\omega)$. Therefore, we can take $r_3, \dot{r}_6, \dot{q}_6$ big enough so as (5) and (6) are satisfied.

Proposition 10. *(Parameter conditions of (5) and (6)) If* $\alpha_1 > 8\alpha_2 + 1 + \frac{2\varphi_2 c_2}{\Delta}$, *then there exist some* $r_3, \dot{r}_6, \dot{q}_6, \gamma_3, \dot{\gamma}_6$ *such that (5) and (6) are satisfied.*

3.3.2 Parameter Tuning of Assumption (A) (6)-(b)

In this section, we consider the parameter tuning **(b)** of Assumption (A) (6). Set $q(y,z;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(z-y\mu)^2}{2\sigma^2}}$. Then we can denote $p_{\theta_0}(y,z)$ and $\overline{p}^N_{\theta_0}(y,z)$ as $q(y,z,\mu_\Delta(\theta_0),\sigma^2_\Delta(\theta_0))$ and $q(y,z,\mu(m,\theta_0),\sigma^2(m,\theta_0,h))$ each other, where

$$\mu_\Delta(\theta_0) = e^{-\theta_0\Delta}.$$

Then applying the mean value theorem and Lemma 8, we obtain

$$\begin{aligned}
&|p_{\theta_0}(y,z) - \overline{p}^N_{\theta_0}(y,z)| \\
&\leq |\mu_\Delta(\theta_0) - \mu(m,\theta_0)| \int_0^1 \left|\partial_\mu q\left(y,z,\gamma\mu_\Delta(\theta_0) + (1-\gamma)\mu(m,\theta_0), \sigma^2_\Delta(\theta_0)\right)\right| d\gamma \\
&\quad + \left|\sigma^2_\Delta(\theta_0) - \sigma^2(m,\theta_0,h)\right| \\
&\qquad\qquad \times \int_0^1 \left|\partial_{\sigma^2} q\left(y,z,\mu(m,\theta_0), \gamma\sigma^2_\Delta(\theta_0) + (1-\gamma)\sigma^2(m,\theta_0,h)\right)\right| d\gamma \\
&\leq C(\alpha,\beta,\Delta)\frac{1}{m}\int_0^1 \left|\partial_\mu q\left(y,z,\gamma\mu_\Delta(\theta_0) + (1-\gamma)\mu(m,\theta_0), \sigma^2_\Delta(\theta_0)\right)\right| d\gamma \qquad (17) \\
&\quad + C(\alpha,\beta,\Delta)\left(\frac{1}{m} + h^2\right) \\
&\qquad\qquad \times \int_0^1 \left|\partial_{\sigma^2} q\left(y,z,\mu(m,\theta_0), \gamma\sigma^2_\Delta(\theta_0) + (1-\gamma)\sigma^2(m,\theta_0,h)\right)\right| d\gamma.
\end{aligned}$$

Next we consider derivatives of q with respect to μ, σ^2. Assume that

$$0 < \mu_{min} \leq \mu \leq \mu_{max} \quad \text{and} \quad 0 < \sigma^2_{min} \leq \sigma^2 \leq \sigma^2_{max}.$$

From Lemma 6, we have, for $c > 1$,

$$\left|\partial_\mu q(y,z;\mu,\sigma^2)\right| \leq \frac{|yz| + y^2\mu_{max}}{\sigma^2_{min}} \frac{1}{\sqrt{2\pi\sigma^2_{min}}} \exp\left(-\frac{\frac{c-1}{c}z^2 - (c-1)(\mu_{max}y)^2}{2\sigma^2_{max}}\right)$$

and

$$\left|\partial_{\sigma^2} q(y,z;\mu,\sigma^2)\right| \le \left\{ \frac{1}{2\sigma_{min}^2}\frac{1}{\sqrt{2\pi\sigma_{min}^2}} + \frac{4(z^2+y^2\mu_{max}^2)}{2(\sigma_{min}^2)^2}\frac{1}{\sqrt{2\pi\sigma_{min}^2}} \right\} \times \exp\left(-\frac{\frac{c-1}{c}z^2-(c-1)(\mu_{max}y)^2}{2\sigma_{max}^2} \right).$$

Next for $0<\gamma<1$, we have

$$\gamma\mu_\Delta(\theta_0)+(1-\gamma)\mu(m,\theta_0) \le \gamma e^{-\theta_0\Delta}+(1-\gamma)e^{-\theta_0\Delta} \le e^{-\theta_0\Delta}$$

and

$$\begin{aligned}\frac{1}{2\beta}(1-e^{-2\alpha\Delta}) &\le \gamma\sigma_\Delta^2(\theta_0)+(1-\gamma)\sigma^2(m,\theta_0,h)\\ &\le \frac{1}{2\alpha}(1-e^{-2\beta\Delta})+C(k,\Delta,\alpha)+1,\end{aligned}$$

where the constant $C(k,\Delta,\alpha)$ is a constant which is defined in (18). Therefore we take as follows;

$$\begin{aligned}&\mu_{max}=e^{-\theta_0\Delta},\quad \sigma_{min}^2=\frac{1}{2\beta}\left(1-e^{-2\alpha\Delta}\right),\\ &\sigma_{max}^2=\frac{1}{2\alpha}\left(1-e^{-2\beta\Delta}\right)+C(k,\Delta,\alpha)+1.\end{aligned}$$

Then we have

$$(17) \le C(\alpha,\beta,\Delta)\frac{1}{\sqrt{2\pi\sigma_{min}^2}}\left\{ \frac{|yz|+y^2\mu_{max}}{\sigma_{min}^2}+\frac{1}{2\sigma_{min}^2}+\frac{4(z^2+y^2\mu_{max}^2)}{2(\sigma_{min}^2)^2} \right\} \times \exp\left(-\frac{\frac{c-1}{c}z^2-(c-1)(\mu_{max}y)^2}{2\sigma_{max}^2} \right)\left(\frac{1}{m}+h^2\right).$$

Then we need the following parameter tuning condition; $\left(\frac{1}{m}+h^2\right)\sqrt{N}\le C$, where C is a constant. Note that $h=C_2N^{-\alpha_2}$ therefore we require that $\alpha_2\ge\frac{1}{2}$. Furthermore, $m\ge\sqrt{N}$. Finally we check the following integrability condition.

$$\sup_N \sup_{\theta\in[\alpha,\beta]} \int\int \left|\frac{\partial}{\partial\theta}\ln \bar{p}_\theta^N(y,z)\right| \left\{ \frac{|yz|+y^2\mu_{max}}{\sigma_{min}^2}+\frac{1}{2\sigma_{min}^2}+\frac{4(z^2+y^2\mu_{max}^2)}{2(\sigma_{min}^2)^2} \right\} \times \exp\left(-\frac{\frac{c-1}{c}z^2-(c-1)(\mu_{max}y)^2}{2\sigma_{max}^2} \right)\mu_{\theta_0}(y)dydz < +\infty.$$

Note that $\mu_\theta(y)$ is the density of $N(0, \frac{1}{2\theta})$ law and that we have the explicit expression for $\frac{\partial}{\partial\theta}\ln \overline{p}_\theta^N(y,z)$, which is a second degree polynomial in y,z. As the parameters, $\sigma^2(m,\theta,h)$ and $\mu(m,\theta)$ satisfy Lemma 10. The above integrability condition is satisfied.

Finally from the above calculations, we obtain

Proposition 11. *In the OU process and its Euler-Maruyama approximation case, for $\alpha_2 \geq \frac{1}{2}$ and $m(N) \geq \sqrt{N}$, Assumption (A) (6)-(b) holds.*

3.3.3 Parameter Tuning of Assumption (A) (6)-(c)

Now we consider the parameter tuning **(c)** of Assumption (A) (6). Note that in order to verify this condition, we can concretely calculate $|\frac{\partial}{\partial\theta}\ln \overline{p}_\theta^N(y,z) - \frac{\partial}{\partial\theta}\ln p_\theta(y,z)|$. From this difference we need to analyze separately each term and use Lemma 8 together with Lemma 10. Then we have some polynomial function $g^N(y,z) = g(y,z)$ with respect to y,z, Assumption (A) (6)-**(c)** is satisfied. In particular, if Y_0 and Y_1 have a normal distribution, it is clear that the integrability condition $E[|g(Y_0,Y_1)|^4] < +\infty$ is satisfied. Then we have

Proposition 12. *In the OU process and its Euler-Maruyama approximation case, Assumption (A) (6)-(c) holds.*

Recall $\Theta = [\alpha,\beta]$ $(0<\alpha<\beta<2)$, $n = C_1 N^{\alpha_1}$, $h = C_2 N^{-\alpha_2}$ and

$$\inf_{(\mathbf{x},\theta)\in B^N} \overline{p}_\theta^N(x,y) \geq (c\sqrt{2c+\Delta}N^{\frac{4cc_2}{\Delta}})^{-1},$$

where $B^N = \{(x,y,\theta);\ |(x,y)| \leq \sqrt{c_2 \ln N}\}$. And we take a prior density function so that $\pi(\theta) > 0$ on Θ and a kernel function K as the Gaussian kernel. Finally, we obtain the following theorem for the OU process and its Euler-Maruyama approximation.

Theorem 3. *Assume $\alpha_1 > 8\alpha_2 + 1 + \frac{4\varphi_2 c_2}{\Delta}$, $\alpha_2 \geq \frac{1}{2}$ and $m \geq \sqrt{N}$. Then there exist some positive finite random variables Ξ_1 and Ξ_2 such that for $f \in C^1(\Theta)$, we have*

$$|E_N[f] - f(\theta_0)| \leq \frac{\Xi_1}{\sqrt{N}} \quad a.s. \quad \text{and} \quad \left|\hat{E}^n_{N,m}[f] - f(\theta_0)\right| \leq \frac{\Xi_1}{\sqrt{N}} \quad a.s.,$$

and therefore $\left|E_N[f] - E^n_{N,m}[f]\right| \leq \frac{\Xi_1+\Xi_2}{\sqrt{N}}$ a.s.

4 Appendix

4.1 Study of the function $(1-\frac{\theta\Delta}{m})^m$ and $\frac{(1-\frac{\theta\Delta}{m})^{2m}-1}{\theta(\frac{\theta}{m}-2)}$

Here we give some lemmas which are used in the parameter tuning sections.

Lemma 6. *For $c>1$, we have*

(i). $(x+y)^2 \le \frac{c}{c-1}x^2 + cy^2$,
(ii). $\frac{c-1}{c}x^2 + (c-1)y^2 \le (x-y)^2$.

The proofs are based on Young's lemma.

Lemma 7. *For $m \ge 2\beta\Delta$, we have* $|(1-\frac{\theta\Delta}{m})^m - e^{-\theta\Delta}| \le e^{-\alpha\Delta}(\beta\Delta)^2\frac{1}{m}$.

After simple calculations, we can obtain this lemma. From this lemma, we obtain

Lemma 8. *For $k=0,1$ and $m \ge 2\beta\Delta$, we have the following estimations;*

(i). $|\frac{\partial^k}{\partial\theta^k}(\sigma_\Delta^2(\theta) - \sigma^2(m,\theta,h))| \le C(\alpha,\beta,\Delta)\{\frac{1}{m} + h^2\mathbf{1}(k=0)\}$,
(ii). $|\frac{\partial^k}{\partial\theta^k}(e^{-\theta\Delta} - \mu(m,\theta))| \le C(\alpha,\beta,\Delta)\frac{1}{m}$,

where $C(\alpha,\beta,\Delta)$ is some positive constant.

Lemma 9. *For $m > \beta\Delta$, we have* $\left(1-\frac{\theta\Delta}{m}\right)^m \le e^{-\theta\Delta}$.

Proof. Set $f(x) = (1+\frac{1}{x})^x$. Then $f(x)$ is an increasing function for $-\infty < x < -1$ and $\lim_{x\to-\infty} f(x) = e$. The proof follows. □

Lemma 10. *We have the following: For $k \in \mathbf{N}\cup\{0\}$,*

(i).
$$\sup_{m\ge\max(\frac{k}{2},\beta\Delta)}\sup_{\theta\in\Theta}\left|\frac{\partial^k}{\partial\theta^k}\mu(m,\theta)\right| = \sup_{m\ge\max(\frac{k}{2},\beta\Delta)}\sup_{\theta\in\Theta}\left|\frac{\partial^k}{\partial\theta^k}(1-\frac{\theta\Delta}{m})^m\right| \le (2\Delta)^k 3^{2\beta\Delta} < +\infty,$$

(ii).
$$\sup_{0\le h\le 1}\sup_{m\ge\max(\frac{k}{2},\beta\Delta)}\sup_{\theta\in\Theta}\left|\frac{\partial^k}{\partial\theta^k}\sigma^2(m,\theta,h)\right|$$
$$= \sup_{0\le h\le 1}\sup_{m\ge\max(\frac{k}{2},\beta\Delta)}\sup_{\theta\in\Theta}\left|\frac{\partial^k}{\partial\theta^k}\frac{(1-\frac{\theta\Delta}{m})^{2m}-1}{\theta(\frac{\theta}{m}-2)} + h^2\mathbf{1}_{\{k=0\}}\right|$$
$$\le C(k,\Delta,\alpha) + 1 < +\infty,$$

(iii).
$$\inf_{0\le h\le 1}\inf_{m\ge\max(\frac{k}{2},\beta\Delta)}\inf_{\theta\in\Theta}\left|\sigma^2(m,\theta,h)\right|$$
$$= \inf_{0\le h\le 1}\inf_{m\ge\max(\frac{k}{2},\beta\Delta)}\inf_{\theta\in\Theta}\left|\frac{(1-\frac{\theta\Delta}{m})^{2m}-1}{\theta(\frac{\theta}{m}-2)} + h^2\right| \ge \frac{2(1-e^{-2\alpha\Delta})}{3\beta} > 0,$$

where some positive constant $C(k,\Delta,\alpha)$ is defined in the proof.

Proof. Now $\mu(m,\theta) = (1-\frac{\theta\Delta}{m})^m$ and set $D_\theta^k = \frac{\partial^k}{\partial\theta^k}$. Note that from Lemma 9, we have $0 \le \mu(m,\theta) \le e^{-\theta\Delta} \le \sup_\theta e^{-\theta\Delta} = e^{-\alpha\Delta}$. Note that

$$D_\theta^k \mu(m,\theta) = (2m)(2m-1)\cdots(2m-(k-1))\left(1-\frac{\theta\Delta}{m}\right)^{2m-k}\left(-\frac{\Delta}{m}\right)^k.$$

Then

$$\begin{aligned}&D_\theta^{k+1}\mu(m,\theta)\\&=(2m)(2m-1)\cdots(2m-(k-1))(2m-k)\left(1-\frac{\theta\Delta}{m}\right)^{2m-(k+1)}\left(-\frac{\Delta}{m}\right)^{k+1}.\end{aligned}$$

And for $2m \ge k$, we have

$$\sup_m \sup_\theta |D_\theta^k \mu(m,\theta)| \le \sup_m \left\{\frac{(2m\Delta)^k}{m^k}(1+\frac{\beta\Delta}{m})^{2m-k}\right\} \le (2\Delta)^k 3^{2\beta\Delta}.$$

Hence we obtain **(i)**.

Remember that $\sigma^2(m,\theta) = \frac{\mu(m,\theta)^2-1}{\theta(\frac{\theta}{m}-2)}$. From the Leibnitz's formula, we have

$$\begin{aligned}&|D_\theta^k\sigma^2(m,\theta)|\\&\le \sum_{i=0}^k C_{k,i}\sup_m\sup_\theta |D_\theta^i(\mu(m,\theta)^2)|\sup_m\sup_\theta\left|D_\theta^{k-i}\frac{1}{\theta(\frac{\theta}{m}-2)}\right| + \sup_m\sup_\theta\left|D_\theta^k\frac{1}{\theta(\frac{\theta}{m}-2)}\right|.\end{aligned}$$

From the above and the Leibnitz's formula, we have for $i = 0,1,\cdots,k$, from the binomial theorem,

$$\begin{aligned}\sup_m\sup_\theta |D_\theta^i(\mu(m,\theta)^2)| &\le \sup_m\sup_\theta\left|\Delta^i(1-\frac{\theta\Delta}{m})^{2m-i}\sum_{j=0}^i\binom{i}{j}\right|\\&\le \Delta^i e^{-\alpha\Delta}\sum_{j=0}^i\binom{i}{j} < +\infty.\end{aligned}$$

And for all $i = 0,1,\cdots,k$, we have, from the binomial theorem,

$$\sup_m\sup_\theta\left|D_\theta^i\frac{1}{\theta(\frac{\theta}{m}-2)}\right| \le \sum_{j=0}^i C_{i,j}\frac{j!}{\alpha^{j+1}}\frac{(i-j)!}{2^{i-j}}.$$

Then we have

$$\begin{aligned}&\sup_m\sup_{\theta\in[\alpha,\beta]}\left|D_\theta^k\sigma^2(m,\theta)\right|\\&\le \sum_{i=0}^k C_{k,i}\left\{\Delta^i e^{-\alpha\Delta}\sum_{j=0}^i\binom{i}{j}\right\}\left\{\sum_{j=0}^{k-i}C_{k-i,j}\frac{j!}{\alpha^{j+1}}\frac{(k-i-j)!}{2^{k-i-j}}\right\}\end{aligned}$$

$$+\sum_{j=0}^{k} C_{k,j} \frac{j!}{\alpha^{j+1}} \frac{(k-j)!}{2^{k-j}} =: C(k,\Delta,\alpha) < +\infty. \tag{18}$$

Therefore we obtain **(ii)**.

Finally for $m \geq \beta\Delta$, we have

$$\sigma^2(m,\theta) \geq \frac{1-e^{-2\theta\Delta}}{\theta\left(2-\frac{\theta}{2\theta}\right)} = \frac{2}{3\theta}\left(1-e^{-2\theta\Delta}\right) \geq \frac{2}{3\beta}\left(1-e^{-2\alpha\Delta}\right) > 0.$$

We obtain **(ii)**.

Here for $m \geq \beta\Delta$, $0 \leq \left(1-\frac{\theta\Delta}{m}\right)^m \leq e^{-\theta\Delta} \leq e^{-\alpha\Delta}$. And for $m \geq \beta\Delta$,

$$\frac{2(1-e^{-2\alpha\Delta})}{3\beta} \leq \frac{(1-\frac{\theta\Delta}{m})^{2m}-1}{\theta(\frac{\theta}{m}-2)} \leq \frac{1}{\alpha(2-\beta)}.$$

Therefore we obtain **(iv)** and finish the proof. □

References

1. Y. Ait-Sahalia, P.A. Mykland, *Estimators of diffusions with randomly spaced discrete observations: a general theory*, Ann. Statist., 32, no. 5, pp.2186–2222, 2004.
2. A. Bain, D. Crisan, *Fundamentals of Stochastic Filtering*, Springer New York, 2009.
3. J.A. Cano, M. Kessler, D. Salmeron, *Approximation of the posterior density for diffusion processes*, Statistics & Probability Letters, 76, no. 1, pp.39–44, 2006.
4. P. Del Moral, J. Jacod, P. Protter, *The Monte-Carlo method for filtering with discrete-time observations*, Probab. Theory and Related Fields, 120, pp.346–368, 2001.
5. P. Doukhan, *Mixing; Properties and Examples*, Lecture Notes in Statistics 85, Springer-Verlag, 1994
6. J. Jacod, *Parametric inference for discretely observed non-ergodic diffusions*, Bernoulli, 12, no. 3, pp.383–401, 2006.
7. L. Kelly, E. Platen, M. Sorensen, *Estimation for discretely observed diffusions using transform functions. Stochastic methods and their applications*, J. Appl. Probab., 41A, pp.99–118, 2004.
8. M. Kessler, *Estimation of an ergodic diffusion from discrete observations*, Scand. J. Statist., 24, no. 2, pp.211–229, 1997.
9. A. Kohatsu-Higa, N. Vayatis, K. Yasuda, *Tuning of a Bayesian Estimator under Discrete Time Observations and Unknown Transition Density*, submitted.
10. G.O. Roberts, O. Stramer, *On inference for partially observed nonlinear diffusion models using the Metropolis-Hastings algorithm*, Biometrika 88, pp.603–621, 2001.
11. N. Yoshida, *Estimation for diffusion processes from discrete observation*, J. Multivariate Anal., 41, no. 2, pp.220–242, 1992.