Article

Parameter Estimation of Linear Stochastic Differential Equations with Sparse Observations

Yuecai Han 1,2,†, Zhe Yin 1,† and Dingwen Zhang 1,*,†

1 School of Mathematics, Jilin University, Changchun 130012, China
2 Key Laboratory of Symbolic Computation and Knowledge Engineering of Ministry of Education, Jilin University, Changchun 130012, China
* Correspondence: zhangdw20@mails.jlu.edu.cn
† These authors contributed equally to this work.

Abstract: We consider parameter estimation for linear stochastic differential equations with independent experiments observed at infrequent and irregularly spaced follow-up times. The maximum likelihood method is used to obtain an asymptotically consistent estimator. A kernel-weighted score function is proposed for the parameter in drift terms. The strong consistency and the rate of convergence of the estimator are obtained. The numerical results show that the proposed estimator performs well with moderate sample sizes.

Keywords: kernel-weighted estimation; linear stochastic differential equations; geometric Brownian motion; likelihood function

1. Introduction

To simulate the dynamic behavior of a complex system, linear stochastic differential equations (LSDEs) are frequently used. In many real-world applications, it is customary that the parameters that define the system must be estimated from the data. As an example, geometric Brownian motion (GBM) is one of the most popular stochastic processes and undoubtedly an effective instrument in modeling and predicting random changes in stock prices [1,2]. Both deterministic and stochastic components contribute to the pharmacokinetic and pharmacodynamic models: Although there are predictable trends in drug concentrations, it is not always possible to establish the precise concentration at any particular time [3].

In biometrics, a GBM model and an estimation procedure were developed for predicting the height growth of even-aged forest stands as part of a methodology for modeling growth in forest plantations [4].

Due to its growing use in a variety of domains, parameter estimation problems involving stochastic differential equations have received a lot of attention lately. Using the data, one should estimate the parameters that characterize the system. Several methods are proposed to evaluate the parameters, such as the least squares method [5–9], the maximum likelihood method [10–14], and the numerical approximation approach [15]. Several other methods, such as the generalized method of moments procedures [16,17], local linearization method [11,18], and MCMC methods [19], are also proposed.

Assume that \( n \) identical and independently distributed paths are observed. When a number of patients can be watched, for example, this situation can arise in pharmacokinetics. For each patient, a bolus of the medication is given, and the “path” of its diffusion through the body can be observed [3]. Such observations are typically sparse and only observed at infrequent and irregularly spaced follow-up times; the above methods are no longer applicable. In this case, we develop a computationally efficient method to deal with the observations with infrequent and irregularly spaced follow-up times. In this paper, we apply kernel methods to the parameter estimation of LSDEs. At the heart of the proposed
The likelihood function is \( L_n(y^*; \mu) = \prod_{i=1}^{n} p(y_i^*; \mu) \)

\[
L_n(y^*; \mu) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2_i t^*}} \exp\left\{ -\frac{(\log y_i^* - \phi(t^*_i))^2}{2\sigma^2_i t^*} \right\},
\]

where \( y_i^* = y(t^*_i) \)

are independent and identically distributed variables on the same probability space. Then we get the likelihood function

\[
L_n(y^*; \mu) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2_i t^*}} \exp\left\{ -\frac{(\log y_i^* - \phi(t^*_i))^2}{2\sigma^2_i t^*} \right\},
\]

2. Models and Methods

2.1. Description of Models

Consider an LSDE model as follows

\[
\begin{aligned}
\text{d}S(t) &= \mu S(t) \text{d}t + \sigma S(t) \text{d}W(t), \\
S(0) &= s_0,
\end{aligned}
\]

where \( \mu \) is an unknown parameter, \( \sigma \) is a constant, and \( W(t) \) are independent standard Brownian motions. \( W(t) \) is characterized by the following properties: (1) \( W(0) = 0; \) (2) \( W(t) \) has independent increments, which is, for every \( t > 0 \), the future increments \( W(t + u) - W(t), u \geq 0, \) are independent of the past values \( W(s), s \leq t, \) and \( W(t + u) - W(t) \sim N(0, u) \); (3) \( W(t) \) is continuous in \( t \). Under the condition of Lipschitz and linear growth, the LSDE (1) has a unique strong solution \( S(t) \),

\[
S(t) = s_0 \exp\left\{ \left( \mu - \frac{1}{2} \sigma^2 \right) t + \sigma W(t) \right\}.
\]

Let \( \{S_i(t), s_0, \mu_i, \sigma_i; i = 1, \cdots, n \} \) be \( n \) dependent copies of \( \{S(t), s_0, \mu, \sigma \} \). As is known to all (see, e.g., [22]), for given \( t \), \( S_i(t), i = 1, \cdots, n \) follow a lognormal distribution

\[
\log S_i(t) \sim N \left( \log s_0 + \mu t - \frac{1}{2} \sigma^2 t \right).
\]

We are aiming to use the observations to estimate \( \mu \), where the observations consist of \( y = \{y_i(t_{ik}); i = 1, \cdots, n; k = 1, \cdots, d_i \} \), and \( d_i < \infty \). The probability of the \( i \)th subject at time point \( t_{ik} \) is

\[
p(y_i(t_{ik}); \mu) = \frac{1}{\sqrt{2\pi\sigma^2_i t_{ik}}} \exp\left\{ -\frac{(\log y_i(t_{ik}) - \phi(t_{ik})]^2}{2\sigma^2_i t_{ik}} \right\},
\]

where \( \phi(t) = \log s_0 + \mu t - \frac{1}{2} \sigma^2 t \).

If the observations are continuous, consider a time point \( t^* \). We know that \( y_i^* = y(t^*_i) \) are independent and identically distributed variables on the same probability space. Then we get the likelihood function

\[
L_n(y^*; \mu) = \prod_{i=1}^{n} p(y_i^*; \mu)
\]

\[
= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2_i t^*}} \exp\left\{ -\frac{(\log y_i^* - \phi(t^*_i))^2}{2\sigma^2_i t^*} \right\},
\]
where \( y^* = (y^*_1, y^*_2, \ldots, y^*_n) \). The log-likelihood function is

\[
  l_n(y^*; \mu) = \frac{1}{n} \sum_{i=1}^{n} \left[ \frac{(\log y^*_i - \phi(t^*))^2}{2(\Sigma_i(t^*))^2} - \log \sqrt{2\pi} - \log \Sigma_i(t^*) \right],
\]

where \( \Sigma_i(t^*) = \sqrt{\sigma^2_i t^*} \) and the score function is

\[
  B_n(y^*, \mu) = \frac{\partial l_n(y^*; \mu)}{\partial \mu}.
\]

### 2.2. Kernel Estimation with Forward and Lagged Observation

The data \( y_i(t), i = 1, \ldots, n \) are usually not observed continuously, and it is almost impossible for each individual to be observed at \( t^* \). Hence \( l_n(y^*, \theta) \) is not computable from the observations. We propose a method that formalizes the forwarding and lagging strategy, with kernel weighting enabling the use of all available forward and lagged observations. We “smooth” the observations’ contribution to the likelihood based on the distance of their observation time to the time of interest. If data continues to be collected on subjects for which observation has occurred, as in the case of the recurrent event, we use the kernel to impute missing values using both forward and backward-lagged observations. We construct a smoothed log-likelihood function by using kernel estimation

\[
l_n(y^*; \mu) = \frac{1}{n} \sum_{i=1}^{n} \int \left[ \frac{(K_{h_0}(s - t_k) \log y_i(s) - \phi_i(t^*))^2}{2\Sigma_i^2(t^*)} + \log \sqrt{2\pi} + \log \Sigma_i^2(t^*) \right] dN_i(s),
\]

where \( \Sigma_i^2(t^*) < \infty \) is the variance of \( y_i^* \), \( K_{h_0}(t) = K((t - t^*)/h_n)/h_n \), \( h_n \) is the bandwidth, and the kernel function \( K(t) \) is a symmetric probability density with support \([-1, 1]\) and mean 0 that bound the first derivative. In addition, \( E[dN_i(t)] = \lambda(t)dt \), where \( \lambda(t) \) is twice continuously differentiable and strictly positive for \( t \in [0, T] \). The scoring function is given by

\[
  U_n(\mu) = \frac{1}{n} \sum_{i=1}^{n} \int K_{h_n}(t^* - t) \log y_i(t) - \left( \log s_0 + \left( \mu - \frac{1}{2} \Sigma_i^2(t^*) \right) t^* \right) \Sigma_i^2(t^*) \] dN_i(t).
\]

Assume that the following conditions hold:

(A.1) \( \Theta_{\mu_0} \) is an open sets of \( \mathbb{R} \), and \( \Theta_{\mu_0} = \{ \mu : |\mu - \mu_0| < \rho \} \) for some \( \rho > 0 \) and \( \mu_0 \) is the true parameter.

(A.2) \( \lambda^*(t) \) is twice continuously differentiable.

(A.3) \( K(z) \) is a symmetric density function satisfying \( \int_{-\infty}^{\infty} K(z)^2 dz < \infty \). In addition, \( h_n \to 0, nh_n \to \infty, nh_n^5 \to 0 \).

Condition (A.1) is a usual assumption for the proof of consistency, and condition (A.2) ensures the Taylor expansion of the score function to the second order. Our methods depend on a proper choice of bandwidth, which is shown in condition (A.3). The estimator \( \hat{\mu}_n \) is obtained based on solving Equation (5) with a kernel bandwidth selected to obtain the consistency.

**Lemma 1.** Under conditions (A.1)–(A.3), we have

\[
  E[(nh_n)^{\frac{1}{2}} U_n(\mu_0)] = 0,
\]

as \( n \to \infty \).

**Proof of Lemma 1.** From the smoothed likelihood function (5), we have the smoothed scoring function
By taking expectations together with Taylor expansion, \( \int_{-\infty}^{\infty} K(z) \log (t^* - h_n z) - (\log s_0 + (\mu - \frac{1}{2} \Sigma^2(t^*)) t^*) \, dN_i(z) = 0 \). By taking expectations together with Taylor expansion, \( \int_{-\infty}^{\infty} zK(z) \, dz = 1 \) and \( \int_{-\infty}^{\infty} zK(z) \, dz = 0 \), we have

\[
E[I] = h_n^2 n^{-\frac{1}{2}} \sum_{i=1}^{n} \left[ \int K(z) \log (t^* - h_n z) - (\log s_0 + (\mu - \frac{1}{2} \Sigma^2(t^*)) t^*) \right] \frac{\lambda(t^* - h_n z) d(z)}{\Sigma^2(t^*)^2} \\
= h_n^2 n^{-\frac{1}{2}} \frac{1}{\sigma^2} \left( -\frac{\partial F}{\partial \sigma} \bigg|_{\sigma=0} \frac{\partial \lambda}{\partial \sigma} \bigg|_{t=t^* \sigma=0} + \frac{1}{2} \frac{\partial^2 F}{\partial \sigma^2} \bigg|_{\sigma=0} h_n^2 \right) \\
= o \left( n^{-\frac{1}{2}} h_n^2 \right).
\]

From condition (A.3), we have \( E[I] = o(1) \). □

The following theorem shows the consistency of the proposed estimator \( \hat{\mu}_n \) obtained based on solving Equation (5).

**Theorem 1.** Under conditions (A.1)–(A.3), \( \hat{\mu}_n \) admits the consistency as \( n \to \infty \).

**Proof of Theorem 1.** Solving Equation (5), we have

\[
\beta_n t^* = \frac{1}{nh_n} \sum_{i=1}^{n} \frac{K_{h_n}(t - t^*) \log y_i(t) \, dN_i(t)}{\int K_{h_n}(t - t^*) \, dN_i(t)} - \log s_0 + \frac{1}{2} \sigma^2 t^*.
\]

By properties of the kernel function \( K(\cdot) \), we have

\[
|\beta_n t^* - \mu t^*| \leq \frac{1}{nh_n} \sum_{i=1}^{n} \frac{K_{h_n}(t - t^*) (\log y_i(t) - \log s_0 - (\mu - \frac{\sigma^2}{2}) t^*) \, dN_i(t)}{\int K_{h_n}(t - t^*) \, dN_i(t)} \\
\leq \frac{1}{nh_n} \left( \sum_{i=1}^{n} (\log y_i(t) - \log s_0 - (\mu - \frac{\sigma^2}{2}) t^*) \right).
\]

where \( b_i \in [t^* - h_n, t^* + h_n] \) is some constant. By Equation (3), we have

\[
\log y(t) \sim N(\log s_0 + (\mu - \frac{\sigma^2}{2}) t, \sigma^2 t).
\]
Hence, we have
\[
\log y(b_i) - \log s_0 - (\mu - \frac{a^2}{2}) t^* \sim \mathcal{N}(a, \sigma^2 t),
\]
where \(a = O(h_n)\) is some constant. By the Wiener–Khinchin law of large numbers, we have
\[
|\tilde{\mu}_n t^* - \mu t^*| = O(h_n),
\]
which goes to zero, as \(n \to \infty\). \(\square\)

The following theorem shows the asymptotic normality of \(\tilde{\mu}_n\).

**Theorem 2.** Assume conditions (A.1)–(A.3) hold, \(\tilde{\mu}_n\) is consistent, and the asymptotic distribution of \(\tilde{\mu}_n\) satisfies
\[
(nh_n)^{\frac{1}{2}}(\tilde{\mu}_n - \mu) \sim N\left(0, (C_n(\mu_0))^2\Sigma(\mu_0)\right),
\]
as \(n \to \infty\), where
\[
C_n(\mu_0) = -\left(\int_0^1 \frac{\partial U_n}{\partial \mu} (\mu_0 + \lambda(\mu - \mu_0))d\lambda\right)^{-1},
\]
and
\[
\Gamma(\mu_0) = \int (K(z))^2 \frac{1}{\sigma^2} \lambda(t^* - h_n z)dz.
\]

**Proof of Theorem 2.** Let \(\{\mu_n\}\) be a strongly consistent sequence of \(\mu\), i.e., \(\mu_n \xrightarrow{a.s.} \mu_0\), and \(\{\mu_n\} \in \Theta_{\mu_0}\). We can seek a solution \(\tilde{\mu}_n\) of the log-likelihood function \(l_n(y^*_n, \tilde{\mu}_n)\), and \(\tilde{\mu}_n\) is a strongly consistent sequence. Note that
\[
U_n(\mu) = \frac{1}{n} \sum_{i=1}^n \int K_{h_n}(t^* - t)y_i(t) - \left(\log s_0 + (\mu - \frac{1}{2}\Sigma(t^*))^2 t^*\right) \frac{dN_i(t)}{\Sigma^2(t^*)},
\]
we denote \(U_n(\mu) = \sum_{i=1}^n \psi(y^*_i, \mu)\). Expand \(U_n\) as
\[
U_n(\mu) = U_n(\mu_0) + \int_{\mu_0}^\mu \frac{\partial U_n(\mu)}{\partial \mu} d\mu
\]
\[
= U_n(\mu_0) + \left(1 - \int_0^1 \frac{\partial U_n}{\partial \mu} (\mu_0 + \lambda(\mu - \mu_0))d\lambda\right)(\mu - \mu_0).
\]
Let \(\mu = \tilde{\mu}_n\), we have \(U_n(\tilde{\mu}_n) = 0\). Then
\[
\tilde{\mu}_n - \mu = -\left(1 - \int_0^1 \frac{\partial U_n}{\partial \mu} (\mu_0 + \lambda(\mu - \mu_0))d\lambda\right)^{-1} U_n(\mu_0).
\]
Multiply both sides of Equation (7) by \((nh_n)^{\frac{1}{2}}\), we denote
\[
C_n(\mu_0) = -\left(\int_0^1 \frac{\partial U_n}{\partial \mu} (\mu_0 + \lambda(\mu - \mu_0))d\lambda\right)^{-1}.
\]
Then
\[
(nh_n)^{\frac{1}{2}}(\tilde{\mu}_n - \mu) = C_n(\mu_0)(nh_n)^{\frac{1}{2}} U_n(\mu_0).
\]
Hence we give the variance of \((nh_n)^{\frac{1}{2}} U_n(\mu_0)\),
\[
(nh_n)^{\frac{1}{2}} U_n(\mu_0) = \frac{(nh_n)^{\frac{1}{2}}}{n} \sum_{i=1}^n \psi(y^*_i, \mu_0) = \frac{n}{2} \left(1 + \frac{1}{n} \sum_{i=1}^n h_n^2 \psi(y^*_i, \mu_0)\right).
\]
From Lemma 1, we have that $E[h_n^2 \psi(y_i^*, \mu_0)] = 0$, $i = 1, \ldots, n$. By central limit theorem, $\text{var} \left( (nh_n)^{1/2} U_n(\mu_0) \right) = \text{var} \left( h_n^2 \psi(y_i^*, \mu_0) \right)$, and we denote $\phi_i(t) = y_i(t) - (\log s_0 + (\mu_0 - \frac{1}{2} \sigma^2) t^*)$. Then

$$\text{var} \left( h_n^2 \psi(y_i^*, \mu_0) \right) = E \left[ (h_n^2 \psi(y_i^*, \mu_0))^2 \right]$$

$$= h_n^4 E \left[ \left( \sum_{k=1}^{n} K_{h_n} (t^* - t_{ik})(y_i(t_{ik}) - (\log s_0 + (\mu_0 - \frac{1}{2} \Sigma^2(t^*))t^*)) \right)^2 \right]$$

$$= h_n^4 \left( \int_{t_1 \neq t_2} K_{h_n} (t^* - t_1) K_{h_n} (t^* - t_2) E[\phi_i(t_1)\phi_i(t_2)] E[dN(t_1) dN(t_2)] \right) + \int_{t_1 = t_2} (K_{h_n} (t^* - t_1))^2 \lambda(t_1) E[\phi_i(t_1)^2] dt_1.$$  

Assume that for $t_1 \neq t_2$, $pr(dN(t_1) = 1 | N(t_2) - N(t_2^-) = 1) = g(t_1, t_2) dt_1$, where $pr$ means the probability, $g(t_1, t_2)$ is continuous for $t_1 \neq t_2$, and $g(t_1, t_2)$ exists. Then

$$E \left[ \left( h_n^2 \psi(y_i^*, \mu_0) \right)^2 \right]$$

$$= h_n^4 \left( \int_{t_1 \neq t_2} K_{h_n} (t^* - t_1) K_{h_n} (t^* - t_2) E[\phi_i(t_1)\phi_i(t_2)] g(t_1, t_2) E[dN(t_2)] dt_1 \right) + \frac{1}{\sigma^4} \int_{t_1 = t_2} (K_{h_n} (t^* - t_1))^2 \lambda(t_1) E[\phi_i(t_1)^2] dt_1$$

$$= h_n (h_1 + I_2).$$

Using a change of variables, we have $h_n h_1 = O(h_n)$.

With notation $E[(\phi_i(t^* - h_n z))^2] = \sigma^2 t^*$, we have

$$h_n I_2 = h_n \frac{1}{\sigma^4} \int (h_n)^{-2} (K(z))^2 E[(\phi_i(t^* - h_n z))^2] \lambda(t^* - h_n z) h_n dz$$

$$= \int (K(z))^2 \frac{1}{\sigma^2} t^* \lambda(t^* - h_n z) dz.$$  

From the Lyapunov central limit theorem, we have that $(nh_n)^{1/2} U_n(\mu_0)$ converges to a continuous Gaussian process $Z \sim N(0, \Gamma(\mu_0))$. Hence, we have

$$(nh_n)^{1/2} (\hat{\mu}_n - \mu) = C_n(\mu_0) (nh_n)^{1/2} U_n(\mu_0) \sim N \left( 0, (C_n(\mu_0))^2 \Gamma(\mu_0) \right).$$

\[ \square \]

**Remark 1.** When there are several observations $d_i > 1$, one can estimate the drift parameter $\mu$ by a standard maximum likelihood method. Let

$$z_{ik} = \log \left( \frac{y_i(t_{ik+1})}{y_i(t_{ik})} \right),$$

for $k = 1, \ldots, d_i$ and $i = 1, 2, \ldots, n$. Thus, conditional on the observation times,

$$z_{ik} \sim N \left( \left( \mu - \sigma^2 / 2 \right)(t_{ik+1} - t_{ik}), \sigma^2 (t_{ik+1} - t_{ik}) \right).$$
and they are independent. For example, if we reparameterize $\mu$ as $\nu = \mu - \sigma^2/2$, the MLE for $\nu$ would be given by

$$\hat{\nu} = \frac{\sum_{i=1}^{n} \sum_{j=1}^{d_i} z_j^{(i)} / \Delta_k^{(i)}}{\sum_{i=1}^{n} (d_i - 1)}.$$  

(8)

where $\Delta_k^{(i)} = t_{ik+1} - t_{ik}$. Then $\mu$ is estimated by $\hat{\mu} = \hat{\nu} + \sigma^2/2$.

**Remark 2.** When there is only one observation $d_i = 1$, the estimator proposed in Equation (8) is not effective. Our estimator performs reasonably well in this extreme case, and we could give an explicit asymptotic variance for our estimator, which is $\sigma^2$.  

3. Simulation

In this section, utilizing both forward and backward-lagged observation, we examine the kernel estimator. We generate 1000 datasets, and each dataset consists of $n = 100, 400, 900$ subjects with different bandwidths (BD). The process is generated through model (1); we set the initial condition $s_0 = 1, \mu = 2$ and $\sigma = 0.02$. Then the solution is

$$S(t) = s_0 \exp\left\{1.9998t + 0.02w(t)\right\},$$  

(9)

where $w(t)$ is a standard Brownian motion. The number of observation times for each subject is a Poisson distributed with an intensity rate of 5. The time points of each individual’s observation are generated from a uniform distribution, Unif(0, 1). The outcomes from other models’ parameter selections are not mentioned because they are essentially identical. All simulations were performed on a laptop running R 4.2.9 with 8G of RAM.

Based on Theorems 1 and 2, we obtain a kernel estimator with asymptotically negligible bias and employ bandwidths in the range $(n^{-1}, n^{-1/2})$ when calculating $\hat{\mu}$ using the smoothed likelihood score function (6). The kernel function we choose is the Epanechnikov kernel, which is $K(x) = \frac{3}{4}(1 - x^2)_{+}$. The usage of additional kernel functions has little effect on the estimator’s empirical performance, according to additional simulations (not published).

The simulation results show that the estimates for the parameter in the model are accurate. Table 1 summarizes the main findings from over 1000 simulations. We note that the bias diminishes and is minor as the sample size grows. The performance improves the larger sample sizes and smaller bandwidths. The overall parameter estimates are evaluated by the bias and relative bias (RB), which are defined as

$$\text{Bias}(\hat{\mu}) = \mu_0 - \hat{\mu}, \quad \text{RB}(\hat{\mu}) = \frac{|\text{Bias}(\hat{\mu})|}{|\mu_0|},$$

where $\theta_0$ denotes the true parameter.

<table>
<thead>
<tr>
<th>n</th>
<th>BD</th>
<th>$\hat{\mu}$</th>
<th>Bias ($\hat{\mu}$)</th>
<th>RB ($\hat{\mu}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>$n^{-0.25}$</td>
<td>2.095</td>
<td>0.095</td>
<td>0.048</td>
</tr>
<tr>
<td></td>
<td>$n^{-0.3}$</td>
<td>2.065</td>
<td>0.065</td>
<td>0.033</td>
</tr>
<tr>
<td>400</td>
<td>$n^{-0.25}$</td>
<td>2.052</td>
<td>0.052</td>
<td>0.026</td>
</tr>
<tr>
<td></td>
<td>$n^{-0.3}$</td>
<td>2.030</td>
<td>0.030</td>
<td>0.015</td>
</tr>
<tr>
<td>900</td>
<td>$n^{-0.25}$</td>
<td>2.020</td>
<td>0.020</td>
<td>0.010</td>
</tr>
<tr>
<td></td>
<td>$n^{-0.3}$</td>
<td>2.020</td>
<td>0.020</td>
<td>0.010</td>
</tr>
</tbody>
</table>

4. Conclusions

In this paper, we have presented kernel-weighting methods for the estimation of the LSDE model (1) in repeatable experiments when the observation time is a random variable, and the number of observations of each individual is uncertain or even sparse. This is a real
improvement because the past literature usually supposed that observation intervals are equally spaced and could not deal with the sparse observations. We consider the maximum likelihood estimation of the drift parameter. This method has some assumptions, and we give the asymptotic normality of the proposed estimator. In numerical studies, we set the true parameter \( \mu_0 = 2, \sigma_0 = 0.02 \), and the initial condition \( s_0 = 1 \) for each individual with sparse observations (the frequency of observation follows a Poisson distribution with mean 5). Using the smoothed scoring function, we obtain the estimation of the drift parameter.

**Author Contributions:** All authors contributed equally to this work. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was partially supported by NSFC (grants 11871244), and the Fundamental Research Funds for the Central Universities, JLU.

**Data Availability Statement:** There is no data used in this paper.

**Conflicts of Interest:** There are no competing interest to declare that arose during the preparation or publication process of this article.

**References**