Hybrid estimators for small diffusion processes based on reduced data

[1]Yusuke Kaino and [1,2]Masayuki Uchida
[1]Graduate School of Engineering Science, Osaka University and
[2]Center for Mathematical Modeling and Data Science (MMDS),
Osaka University, CREST, JST
Toyonaka, Osaka 560-8531, Japan

Asymptotical Statistics of Stochastic Processes XI (SAPS XI)
at "New Peterhof" (Steklov Mathematical Institute), St Petersburg
17-21 July 2017

Plan of today's talk

- Motivation for this talk
- Initial Bayes type estimators of both drift and volatility parameters for small diffusion processes based on reduced data with sample size n_0 $(\leq n)$ in the case when $\epsilon \to 0$ and $n_0 \to \infty$.
- hybrid estimator with the initial Bayes type estimator for small diffusion processes in the case when $\frac{1}{\epsilon\sqrt{n}} = O(1)$.
- Example and simulation results

1. Introduction

We treat a d-dimensional small diffusion process defined by the following stochastic differential equation

$$\begin{cases} dX_t = a(X_t, \alpha)dt + \epsilon b(X_t, \beta)dw_t, \quad t \in [0, T], \quad \epsilon \in (0, 1], \\ X_0 = x_0, \end{cases} \tag{1}$$

where

ϵ and T are known constants, x_0 is a deterministic initial condition,
w is an r-dimensional standard Wiener process,
$\theta = (\alpha, \beta) \in \Theta = \Theta_\alpha \times \Theta_\beta$ with Θ_α and Θ_β being compact convex subsets of $\mathbb{R}^p$ and $\mathbb{R}^q$, respectively,
$a : \mathbb{R}^d \times \Theta_\alpha \to \mathbb{R}^d$ and $b : \mathbb{R}^d \times \Theta_\beta \to \mathbb{R}^d \otimes \mathbb{R}^r$, and $\theta^* = (\alpha^*, \beta^*) \in \mathrm{Int}(\Theta)$ is the true value of θ.
The data are discrete observations $\mathbb{X}_n = (X_{t_i})_{0 \le i \le n}$ with $t_i = ih_n$, $h_n = T/n$.
We will consider the case when $\epsilon \to 0$, $n \to \infty$, $\frac{1}{\epsilon\sqrt{n}} = O(1)$, and
there exists $\gamma \in (0, 1]$ satisfying that $\epsilon(\sqrt{n})^\gamma = O(1)$.

History

A family of small diffusion processes defined by (1) is an important class of dynamical systems with small perturbations. For dynamical systems with small perturbations, see Azencott (1982) and Freidlin and Wentzell (1998).

Statistical inference for continuously observed small diffusion processes is well developed by Kutoyants (1984, 1994), Yoshida (1992a, 2003), Iacus (2000), Iacus and Kutoyants (2001), Uchida and Yoshida (2004a), Brouste et. al. (2014) and references therein.

Furthermore, there are a number of researches on parametric inference for discretely observed small diffusion processes, see Genon-Catalot (1990), Laredo (1990), Sørensen (2000, 2012), Sørensen and Uchida (2003), Uchida (2003, 2004, 2006, 2008), Gloter and Sørensen (2009), Guy et. al. (2014) and Nomura and Uchida (2016).

For applications of small diffusion processes to mathematical finance and mathematical biology, see Yoshida (1992b), Kunitomo and Takahashi (2001), Takahashi and Yoshida (2004), Uchida and Yoshida (2004b), Fuchs (2013), Guy et. al. (2014, 2015) and references therein.

Motivation

In order to explain the goal of this paper, we first review the joint estimation of both drift and volatility parameters for discretely observed small diffusion processes.

Joint estimation

Set $A^{\otimes 2} = AA^\star$ and $C[A] = \mathrm{tr}(CA^\star)$ for matrices A and C of the same size, where $\star$ means the transpose. Let $B(x,\beta) = bb^\star(x,\beta)$, $\Delta X_i = X_{t_i} - X_{t_{i-1}}$, $a_{i-1}(\alpha) = a(X_{t_{i-1}}, \alpha)$ and $B_{i-1}(\beta) = B(X_{t_{i-1}}, \beta)$. The quasi-log likelihood function is defined as

$$U_{\epsilon,n}(\alpha,\beta) = -\frac{1}{2}\sum_{i=1}^{n}\left\{\log\det B_{i-1}(\beta) + (\epsilon^2 h_n)^{-1} B_{i-1}^{-1}(\beta)\left[(\Delta X_i - h_n a_{i-1}(\alpha))^{\otimes 2}\right]\right\}.$$

The joint maximum likelihood (ML) type estimators $\hat{\alpha}_{\epsilon,n}^{(J)}$ and $\hat{\beta}_{\epsilon,n}^{(J)}$ are defined as

$$U_{\epsilon,n}(\hat{\alpha}_{\epsilon,n}^{(J)}, \hat{\beta}_{\epsilon,n}^{(J)}) = \sup_{\alpha\in\Theta_\alpha, \beta\in\Theta_\beta} U_{\epsilon,n}(\alpha,\beta).$$

Sørensen and Uchida (2003) showed that under some regularity conditions, as $\epsilon \to 0$, $n \to \infty$ and $\frac{1}{\epsilon\sqrt{n}} = O(1)$,

$$\left(\epsilon^{-1}(\hat{\alpha}_{\epsilon,n}^{(J)} - \alpha^*), \sqrt{n}(\hat{\beta}_{\epsilon,n}^{(J)} - \beta^*)\right) \xrightarrow{d} (\zeta_1, \zeta_2) \sim N_{p+q}(0, I(\theta^*)^{-1}), \tag{2}$$

where $\xrightarrow{d}$ means convergence in distribution, $N_{p+q}(0, I(\theta^*)^{-1})$ is the normal random variable with mean zero and the covariance matrix $I(\theta^*)^{-1}$ and $I(\theta^*)$ is the asymptotic Fisher information matrix, see Section 2 below.

Adaptive estimation

From the viewpoint of numerical analysis, the joint ML type estimator is unstable when the dimension of Θ is large. For that reason, we consider the adaptive ML type estimators. In the same way as Uchida and Yoshida (2012) for ergodic diffusion models, the quasi-log likelihood functions are defined as

$$\begin{aligned} V^{(1)}_{\epsilon,n}(\beta) &= -\frac{1}{2}\sum_{i=1}^{n}\left\{\log\det B_{i-1}(\beta) + (\epsilon^2 h_n)^{-1} B^{-1}_{i-1}(\beta)\left[(\Delta X_i)^{\otimes 2}\right]\right\}, \\ V^{(2)}_{\epsilon,n}(\alpha,\beta) &= -\frac{1}{2}\sum_{i=1}^{n}(\epsilon^2 h_n)^{-1} B^{-1}_{i-1}(\beta)\left[(\Delta X_i - h_n a_{i-1}(\alpha))^{\otimes 2}\right]. \end{aligned}$$

The adaptive ML type estimators $\hat{\alpha}^{(E)}_{\epsilon,n}$ and $\hat{\beta}^{(E)}_{\epsilon,n}$ are defined as

$$V^{(1)}_{\epsilon,n}(\hat{\beta}^{(E)}_{\epsilon,n}) = \sup_{\beta\in\Theta_\beta} V^{(1)}_{\epsilon,n}(\beta), \tag{3}$$

$$V^{(2)}_{\epsilon,n}(\hat{\alpha}^{(E)}_{\epsilon,n}, \hat{\beta}^{(E)}_{\epsilon,n}) = \sup_{\alpha\in\Theta_\alpha} V^{(2)}_{\epsilon,n}(\alpha, \hat{\beta}^{(E)}_{\epsilon,n}). \tag{4}$$

Then, under some regularity conditions, as $\epsilon \to 0$, $n \to \infty$ and $\frac{1}{\epsilon^2\sqrt{n}} = o(1)$,

$$\left(\epsilon^{-1}(\hat{\alpha}^{(E)}_{\epsilon,n} - \alpha^*), \sqrt{n}(\hat{\beta}^{(E)}_{\epsilon,n} - \beta^*)\right) \xrightarrow{d} (\zeta_1, \zeta_2) \sim N_{p+q}(0, I(\theta^*)^{-1}).$$

In the case of small diffusion process, the adaptive ML type estimators (3) and (4), which are obtained by the same method as the case of the ergodic diffusion processes, are worse than the joint ML type estimators $\hat{\alpha}^{(J)}_{\epsilon,n}$ and $\hat{\beta}^{(J)}_{\epsilon,n}$ since the stronger condition $\frac{1}{\epsilon^2\sqrt{n}} = o(1)$ is needed to get the same asymptotic properties as (2).

Therefore, the aim of this talk is to propose the adaptive ML type estimator which has the same asymptotic normality as (2) under $\frac{1}{\epsilon\sqrt{n}} = O(1)$ from the viewpoint of numerical analysis.

In order to compute the adaptive ML type estimators, it is indispensable to get a suitable initial estimator for optimization of quasi-log likelihood function.

Nomura and Uchida (2016) obtained the initial Bayes type estimator from full data of small diffusion processes. They considered the hybrid estimator with the initial Bayes type estimator and showed that the hybrid estimator has asymptotic normality and convergence of moments.

However, it takes much time to compute the initial Bayes type estimator when the sample size is large.

Kutoyants (2017) considered the multi-step ML type estimation procedure for ergodic diffusion processes from continuous path data on $[0, T]$. He proposed the multi-step estimator with the initial estimator derived from the reduced continuous path data on $[0, T_0]$ for $T_0 \leq T$ and showed asymptotic efficiency of the multi-step ML type estimator as $T_0 \to \infty$.

Uchida and Yoshida (2017) studied the initial Bayes type estimator based on reduced sampled data for a discretely observed ergodic diffusion processes and they showed asymptotic normality and convergence of moments for the adaptive ML type estimator with the initial Bayes type estimator.

In this talk, we consider the initial Bayes type estimator based on reduced sampled data for a discretely observed small diffusion process by applying the initial estimator with reduced data for a ergodic diffusion process in Kutoyants (2017) and Uchida and Yoshida (2017) to the initial Bayes type estimator for a small diffusion model from the viewpoint of numerical analysis.

The adaptive ML type estimator with the initial Bayes type estimator, which is called the hybrid estimator with the initial Bayes type estimator, is proposed for a small diffusion process.

Moreover, it is shown that the proposed hybrid estimator has asymptotic normality and convergence of moments by applying the Ibragimov-Has'minskii program (1972a,b, 1981) and the polynomial type large deviation inequality for statistical random field in Yoshida (2011) to the case of discretely observed small diffusion processes.

Needless to say, the convergence of moments and the polynomial type large deviation inequality of statistical random field play an important part to show the mathematical validity of asymptotic expansions and asymptotic unbiasedness of information criteria for model selection, see Yoshida (1992a, 1992b), Uchida and Yoshida (2001, 2004a, 2004b, 2006) and Uchida (2010).

2. Initial Bayes type estimators with reduced data and hybrid estimators

Although the data are discrete observations $\mathbb{X}_n = (X_{t_i})_{0\le i\le n}$ with $t_i = ih_n$, $h_n = T/n$, from the viewpoint of numerical analysis, we consider initial estimators with reduced data $\mathbb{Y}_{n_0} = (X_{t_i})_{0\le i\le n_0}$, where $n_0 = \left[\frac{n}{c}\right]$ for $c \ge 1$.

For a matrix A, we define $\|A\| = \mathrm{tr}(AA^\star)^{1/2}$ and $|\cdot|$ denotes the Euclidian norm.
Let $\xrightarrow{p}$ and $\xrightarrow{d}$ be the convergence in probability and convergence in distribution, respectively.
Let X_t^0 be the solution of the ordinary differential equation corresponding to $\epsilon = 0$, i.e., $dX_t^0 = a(X_t^0, \alpha^*)dt$, $X_0^0 = x_0$.
Let $C_\uparrow^{k,l}(\mathbb{R}^d \times \Theta; \mathbb{R}^d)$ denote the space of all functions f satisfying the following conditions:

(i) $f(x,\theta)$ is an $\mathbb{R}^d$-valued function on $\mathbb{R}^d \times \Theta$ and is continuously differentiable with respect to x and θ up to order k and l, respectively.

(ii) for $|\boldsymbol{n}| = 0, 1, \ldots, k$ and $|\boldsymbol{\nu}| = 0, 1, \ldots, l$, there exists $C > 0$ such that $\sup\limits_{\theta\in\Theta}|\delta^{\boldsymbol{\nu}}\partial^{\boldsymbol{n}} f| \le C(1 + |x|)^C$ for all x.

Here, $\boldsymbol{n} = (n_1, \ldots, n_d)$ and $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_l)$ are multi-indices, $l = \dim(\Theta)$, $|\boldsymbol{n}| = n_1 + \cdots + n_d$, $|\boldsymbol{\nu}| = \nu_1 + \cdots + \nu_l$, $\partial^{\boldsymbol{n}} = \partial_1^{n_1}\cdots\partial_d^{n_d}$, $\partial_i = \partial/\partial x_i$, $i = 1, \ldots, d$, $\delta^{\boldsymbol{\nu}} = \delta_1^{\nu_1}\cdots\delta_l^{\nu_l}$, $\delta_j = \partial/\partial\theta_j$, $j = 1, \ldots, l$.

In this talk, we make the assumptions as follows.

[A1] (i) There exists $K > 0$ such that for all $x, y \in \mathbb{R}^d$,

$$\sup_{\alpha \in \Theta_\alpha} |a(x,\alpha) - a(y,\alpha)| + \sup_{\beta \in \Theta_\beta} \|b(x,\beta) - b(y,\beta)\| \le K|x - y|.$$

(ii) $\inf_{x,\beta} \det B(x,\beta) > 0$.

[A2] $a(x,\alpha) \in C_{\uparrow}^{6,4}(\mathbb{R}^d \times \Theta_\alpha; \mathbb{R}^d)$, $b(x,\beta) \in C_{\uparrow}^{6,4}(\mathbb{R}^d \times \Theta_\beta; \mathbb{R}^d \otimes \mathbb{R}^r)$.

The quasi log-likelihood functions $U_{\epsilon,n}^{(1)}(\alpha)$ and $U_{\epsilon,n}^{(2)}(\alpha,\beta)$ with reduced data $\mathbb{Y}_{n_0}$, and the quasi log-likelihood functions $U_{\epsilon,n}^{(3)}(\alpha,\beta)$ and $U_{\epsilon,n}^{(4)}(\alpha,\beta)$ with full data $\mathbb{X}_n$ are defined as follows.

$$\begin{aligned}
U_{\epsilon,n_0}^{(1)}(\alpha) &= -\frac{1}{2\epsilon^2 h_n} \sum_{i=1}^{n_0} |\Delta X_i - h_n a_{i-1}(\alpha)|^2, \\
U_{\epsilon,n_0}^{(2)}(\alpha,\beta) &= -\frac{1}{2\epsilon^4 h_n^2} \sum_{i=1}^{n_0} \left\| (\Delta X_i - h_n a_{i-1}(\alpha))^{\otimes 2} - (\epsilon^2 h_n) B_{i-1}(\beta) \right\|^2, \\
U_{\epsilon,n}^{(3)}(\alpha,\beta) &= -\frac{1}{2\epsilon^2 h_n} \sum_{i=1}^{n} B_{i-1}^{-1}(\beta) \left[(\Delta X_i - h_n a_{i-1}(\alpha))^{\otimes 2} \right], \\
U_{\epsilon,n}^{(4)}(\alpha,\beta) &= -\frac{1}{2} \sum_{i=1}^{n} \left\{ \log\det B_{i-1}(\beta) + (\epsilon^2 h_n)^{-1} B_{i-1}^{-1}(\beta) \left[(\Delta X_i - h_n a_{i-1}(\alpha))^{\otimes 2} \right] \right\}.
\end{aligned}$$

Note that under $[A1]$-$[A2]$, as $\frac{1}{\epsilon\sqrt{n}} = O(1)$, uniformly in $\theta \in \Theta$,

$$\epsilon^2 \left\{U_{\epsilon,n_0}^{(1)}(\alpha) - U_{\epsilon,n_0}^{(1)}(\alpha^*)\right\} \xrightarrow{p} \mathbb{Y}^{(1)}(\alpha),$$
$$h_n \left\{U_{\epsilon,n_0}^{(2)}(\alpha^*,\beta) - U_{\epsilon,n_0}^{(2)}(\alpha^*,\beta^*)\right\} \xrightarrow{p} \mathbb{Y}^{(2)}(\beta),$$
$$\epsilon^2 \left\{U_{\epsilon,n}^{(3)}(\alpha,\beta^*) - U_{\epsilon,n}^{(3)}(\alpha^*,\beta^*)\right\} \xrightarrow{p} \mathbb{Y}^{(3)}(\alpha),$$
$$h_n \left\{U_{\epsilon,n}^{(4)}(\alpha^*,\beta) - U_{\epsilon,n}^{(4)}(\alpha^*,\beta^*)\right\} \xrightarrow{p} \mathbb{Y}^{(4)}(\beta),$$

where

$$\begin{aligned}
\mathbb{Y}^{(1)}(\alpha) &= -\frac{1}{2}\frac{c}{T}\int_0^{T/c} \left|a(X_t^0,\alpha) - a(X_t^0,\alpha^*)\right|^2 dt,\\
\mathbb{Y}^{(2)}(\beta) &= -\frac{1}{2}\frac{c}{T}\int_0^{T/c} \left\|B(X_t^0,\beta) - B(X_t^0,\beta^*)\right\|^2 dt,\\
\mathbb{Y}^{(3)}(\alpha) &= -\frac{1}{2}\frac{1}{T}\int_0^{T} B(X_t^0,\beta^*)^{-1}\left[\left(a(X_t^0,\alpha) - a(X_t^0,\alpha^*)\right)^{\otimes 2}\right] dt,\\
\mathbb{Y}^{(4)}(\beta) &= -\frac{1}{2}\frac{1}{T}\int_0^{T} \left\{\mathrm{tr}\left[B(X_t^0,\beta)^{-1}B(X_t^0,\beta^*) - I_d\right] + \log\frac{\det B(X_t^0,\beta)}{\det B(X_t^0,\beta^*)}\right\} dt.
\end{aligned}$$

[A3] There exist positive constants $\chi^{(1)}, \chi^{(2)}$, $\chi^{(3)}$ and $\chi^{(4)}$ such that for all $\alpha \in \Theta_\alpha$ and $\beta \in \Theta_\beta$,

$$
\begin{aligned}
\mathbb{Y}^{(1)}(\alpha) &\leq -\chi^{(1)}|\alpha - \alpha^*|^2, \\
\mathbb{Y}^{(2)}(\beta) &\leq -\chi^{(2)}|\beta - \beta^*|^2, \\
\mathbb{Y}^{(3)}(\alpha) &\leq -\chi^{(3)}|\alpha - \alpha^*|^2, \\
\mathbb{Y}^{(4)}(\beta) &\leq -\chi^{(4)}|\beta - \beta^*|^2.
\end{aligned}
$$

[A4] We assume that $\frac{1}{\epsilon\sqrt{n}} = O(1)$ as $\epsilon \to 0$ and $n \to \infty$, and there exists $\gamma \in (0, 1]$ satisfying that $\epsilon(\sqrt{n})^\gamma = O(1)$. Moreover, $r_2 \leq 2r_1\gamma$ for $r_1, r_2 \in (0, 1]$.

The statistical random fields $\mathbb{H}^{(1)}_{\epsilon,n_0,r_1}(\alpha)$ and $\mathbb{H}^{(2)}_{\epsilon,n_0,r_2}(\alpha,\beta)$ with reduced data $\mathbb{Y}_{n_0}$ are given by

$$\mathbb{H}^{(1)}_{\epsilon,n_0,r_1}(\alpha) = \epsilon^{2-2r_1} U^{(1)}_{\epsilon,n_0}(\alpha),$$
$$\mathbb{H}^{(2)}_{\epsilon,n_0,r_2}(\alpha,\beta) = \frac{1}{(\sqrt{n_0})^{2-2r_2}} U^{(2)}_{\epsilon,n_0}(\alpha,\beta).$$

We assume that the prior densities $\pi_1(\alpha)$ and $\pi_2(\beta)$ are continuous and satisfy that $0 < \inf_{\alpha\in\Theta_\alpha} \pi_1(\alpha) \le \sup_{\alpha\in\Theta_\alpha} \pi_1(\alpha) < \infty$ and $0 < \inf_{\beta\in\Theta_\beta} \pi_2(\beta) \le \sup_{\beta\in\Theta_\beta} \pi_2(\beta) < \infty$.

The initial Bayes type estimators $\tilde{\alpha}^{(1)}_{\epsilon,n_0,r_1}$ and $\tilde{\beta}^{(2)}_{\epsilon,n_0,r_2}$ with reduced data $\mathbb{Y}_{n_0}$ are defined by

$$\tilde{\alpha}^{(1)}_{\epsilon,n_0,r_1} = \frac{\int_{\Theta_\alpha} \alpha \exp\left\{\mathbb{H}^{(1)}_{\epsilon,n_0,r_1}(\alpha))\right\} \pi_1(\alpha) d\alpha}{\int_{\Theta_\alpha} \exp\left\{\mathbb{H}^{(1)}_{\epsilon,n_0,r_1}(\alpha)\right\} \pi_1(\alpha) d\alpha},$$

$$\tilde{\beta}^{(2)}_{\epsilon,n_0,r_2} = \frac{\int_{\Theta_\beta} \beta \exp\left\{\mathbb{H}^{(2)}_{\epsilon,n_0,r_2}(\tilde{\alpha}^{(1)}_{\epsilon,n_0,r_1},\beta)\right\} \pi_2(\beta) d\beta}{\int_{\Theta_\beta} \exp\left\{\mathbb{H}^{(2)}_{\epsilon,n_0,r_2}(\tilde{\alpha}^{(1)}_{\epsilon,n_0,r_1},\beta)\right\} \pi_2(\beta) d\beta}.$$

The hybrid estimators $\hat{\alpha}_{\epsilon,n}$ and $\hat{\beta}_{\epsilon,n}$ with full data $\mathbb{X}_n$ are defined by

$$U^{(3)}_{\epsilon,n}(\hat{\alpha}_{\epsilon,n}, \tilde{\beta}^{(2)}_{\epsilon,n_0,r_2}) = \sup_{\alpha\in\Theta_\alpha} U^{(3)}_{\epsilon,n}(\alpha, \tilde{\beta}^{(2)}_{\epsilon,n_0,r_2}),$$
$$U^{(4)}_{\epsilon,n}(\hat{\alpha}_{\epsilon,n}, \hat{\beta}_{\epsilon,n}) = \sup_{\beta\in\Theta_\beta} U^{(4)}_{\epsilon,n}(\hat{\alpha}_{\epsilon,n}, \beta).$$

Proposition 1

Let $r_1, r_2 \in (0, 1]$. Assume [A1]–[A4]. Then, for all $M > 0$, as $\epsilon \to 0$ and $n \to \infty$,

(i) $\displaystyle \sup_{\epsilon,n} E_{\theta^*}\left[\left|\epsilon^{-r_1}(\tilde{\alpha}^{(1)}_{\epsilon,n_0,r_1} - \alpha^*)\right|^M\right] < \infty$.

(ii) $\displaystyle \sup_{\epsilon,n} E_{\theta^*}\left[\left|(\sqrt{n_0})^{r_2}(\tilde{\beta}^{(2)}_{\epsilon,n_0,r_2} - \beta^*)\right|^M\right] < \infty$.

(iii) $\displaystyle \sup_{\epsilon,n} E_{\theta^*}\left[\left|\epsilon^{-1}(\hat{\alpha}_{\epsilon,n} - \alpha^*)\right|^M\right] < \infty$.

(iv) $\displaystyle \sup_{\epsilon,n} E_{\theta^*}\left[\left|\sqrt{n}(\hat{\beta}_{\epsilon,n} - \beta^*)\right|^M\right] < \infty$.

Remark 1

It follows from Proposition 1 that when $r_1, r_2 \in (0, 1)$, the convergence rates of the initial Bayes type estimators $\tilde{\alpha}^{(1)}_{\epsilon,n_0,r_1}$ and $\tilde{\beta}^{(2)}_{\epsilon,n_0,r_2}$ are ϵ^{r_1} and $\frac{1}{(\sqrt{n_0})^{r_2}}$, respectively, which means that the initial Bayes type estimators do not have optimal rates. However, the hybrid estimators $\hat{\alpha}_{\epsilon,n}$ and $\hat{\beta}_{\epsilon,n}$ have optimal rates, ϵ and $\frac{1}{\sqrt{n}}$, respectively.

Let

$$I(\theta^*) = \begin{pmatrix} \left(I_a^{ij}(\theta^*)\right)_{1\le i,j\le p} & 0 \\ 0 & \left(I_b^{ij}(\beta^*)\right)_{1\le i,j\le q} \end{pmatrix},$$

$$I_a^{ij}(\theta^*) = \int_0^T \left(\partial_{\alpha_i} a(X_t^0, \alpha^*)\right)^\star B(X_t^0, \beta^*) \partial_{\alpha_j} a(X_t^0, \alpha^*) dt,$$

$$I_b^{ij}(\beta^*) = \frac{1}{2}\frac{1}{T}\int_0^T \mathrm{tr}\left\{B^{-1}(\partial_{\beta_i} B) B^{-1}(\partial_{\beta_j} B)(X_t^0, \beta^*)\right\} dt.$$

Theorem 1

Assume [A1]–[A4]. Then, as $\epsilon \to 0$ *and* $n \to \infty$*,*

$$\left(\epsilon^{-1}(\hat{\alpha}_{\epsilon,n} - \alpha^*), \sqrt{n}(\hat{\beta}_{\epsilon,n} - \beta^*)\right) \xrightarrow{d} (\zeta_1, \zeta_2) \sim N_{p+q}(0, I(\theta^*)^{-1})$$

and

$$E_{\theta^*}\left[f\left(\epsilon^{-1}(\hat{\alpha}_{\epsilon,n} - \alpha^*), \sqrt{n}(\hat{\beta}_{\epsilon,n} - \beta^*)\right)\right] \to \mathbb{E}\left[f(\zeta_1, \zeta_2)\right]$$

for all continuous functions f *of at most polynomial growth.*

3. Examples and simulation results

Consider the three-dimensional diffusion process defined by

$$\begin{aligned} dX_t &= a(X_t, \alpha) + \epsilon b(X_t, \beta) dW_t, \quad t \in [0,1], \quad \epsilon \in (0,1], \\ X_0 &= (1,1,1)^\star, \end{aligned}$$

where

$$\begin{aligned} a(X_t, \alpha) &= \begin{pmatrix} 1 - X_{t,1} - 10\sin(\alpha_1 X_{t,2} + \alpha_2 X_{t,2}^2) \\ 2 - \alpha_3 X_{t,2} - 10\sin(\alpha_4 X_{t,3}^2) \\ 3 - \alpha_5 X_{t,3} - 10\sin(\alpha_6 X_{t,1}^2) \end{pmatrix}, \\ b(X_t, \beta) &= \begin{pmatrix} \sqrt{\beta_1(2+\cos(X_{t,3}^2))} & 0.01 & 0 \\ 0.01 & \sqrt{\beta_2(2+\cos(X_{t,1}^2))} & 0 \\ 0 & 0 & \sqrt{\beta_3(2+\cos(X_{t,2}^2))} \end{pmatrix}. \end{aligned}$$

Furthermore, $\alpha = (\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6)$, and $\beta = (\beta_1, \beta_2, \beta_3)$ are unknown parameters, and their true values $(\alpha_1^*, \alpha_2^*, \alpha_3^*, \alpha_4^*, \alpha_5^*, \alpha_6^*, \beta_1^*, \beta_2^*, \beta_3^*) = (3, 7, 5, 2, 4, 6, 1, 2, 3)$. The parameter space is assumed to be $\Theta = [0.1, 50]^9$.

The simulations were done for $T = 1$, $h = 10^{-5}$, which means that $n = 10^5$.
Let $c = 10$ and $n_0 = n/10 = 10^4$. We set $\epsilon = 0.05, 0.01$.
The initial Bayes type estimator $\tilde{\theta}_B = (\tilde{\alpha}^{(1)}_{\epsilon,n_0,r_1}, \tilde{\beta}^{(2)}_{\epsilon,n_0,r_2})$ with reduced data $\mathbb{Y}_{n_0}$ is defined by

$$\tilde{\alpha}^{(1)}_{\epsilon,n_0,r_1} = \frac{\int_{\Theta_\alpha} \alpha \exp\left\{\mathbb{H}^{(1)}_{\epsilon,n_0,r_1}(\alpha))\right\} \pi_1(\alpha) d\alpha}{\int_{\Theta_\alpha} \exp\left\{\mathbb{H}^{(1)}_{\epsilon,n_0,r_1}(\alpha)\right\} \pi_1(\alpha) d\alpha},$$

$$\tilde{\beta}^{(2)}_{\epsilon,n_0,r_2} = \frac{\int_{\Theta_\beta} \beta \exp\left\{\mathbb{H}^{(2)}_{\epsilon,n_0,r_2}(\tilde{\alpha}^{(1)}_{\epsilon,n_0,r_1}, \beta)\right\} \pi_2(\beta) d\beta}{\int_{\Theta_\beta} \exp\left\{\mathbb{H}^{(2)}_{\epsilon,n_0,r_2}(\tilde{\alpha}^{(1)}_{\epsilon,n_0,r_1}, \beta)\right\} \pi_2(\beta) d\beta},$$

where $\mathbb{H}^{(1)}_{\epsilon,n_0,r_1}(\alpha)$ and $\mathbb{H}^{(2)}_{\epsilon,n_0,r_2}(\alpha, \beta)$ with reduced data $\mathbb{Y}_{n_0}$ are given by

$$\mathbb{H}^{(1)}_{\epsilon,n_0,r_1}(\alpha) = \epsilon^{2-2r_1} U^{(1)}_{\epsilon,n_0}(\alpha),$$
$$\mathbb{H}^{(2)}_{\epsilon,n_0,r_2}(\alpha, \beta) = \frac{1}{(\sqrt{n_0})^{2-2r_2}} U^{(2)}_{\epsilon,n_0}(\alpha, \beta).$$

It follows from Proposition 1 that

$$\sup_{\epsilon,n} E_{\theta^*}\left[\left|\epsilon^{-r_1}(\tilde{\alpha}^{(1)}_{\epsilon,n_0,r_1} - \alpha^*)\right|^M\right] < \infty,$$

$$\sup_{\epsilon,n} E_{\theta^*}\left[\left|(\sqrt{n_0})^{r_2}(\tilde{\beta}^{(2)}_{\epsilon,n_0,r_2} - \beta^*)\right|^M\right] < \infty.$$

The Bayes type estimators are calculated with MpCN method proposed by Kamatani (2014). MpCN algorithm is as follows.

- Choose $x \in \mathbf{R}^d$ and $\mu \in \mathbf{R}^d$.
- Generate r from the gamma distribution with the shape parameter $d/2$ and the scale parameter $\parallel x - \mu \parallel^2 /2$.
- Generate $x^* = \mu + \rho^{1/2}(x - \mu) + (1 - \rho)^{1/2} r^{-1/2} \omega$ where w follows the standard normal distribution.
- Accept x^* as x with probability $\min\left\{1, \frac{p(x^*)\|x-\mu\|^2}{p(x)\|x^*-\mu\|^2}\right\}$. Otherwise, discard x^*.

In practice, it is advisable to take the two-stage procedure.

- Choose $x \in \mathbf{R}^d$ and $\mu \in \mathbf{R}^d$. Run MpCN algorithm. Let $(x_1, \ldots, x_M)$ be the output.
- Set $x = x_M$, $\mu = \sum_{m=1}^{M} \frac{x_m}{M}$ and run MpCN algorithm again.

In this paper, we set $\rho = 0.8$.

We used 10^7 and 10^6 Markov chains and 10^6 and 10^5 burn-in iterations for estimation of α and β, respectively.

The adaptive ML type estimator $(\hat{\alpha}^{(3)}_{A,n}, \hat{\beta}^{(4)}_{A,n})$ is defined as

$$
\begin{aligned}
\hat{\alpha}^{(1)}_{A,n} &= \arg\sup_{\alpha\in\Theta_\alpha} U^{(1)}_{\epsilon,n}(\alpha), \\
\hat{\beta}^{(2)}_{A,n} &= \arg\sup_{\beta\in\Theta_\beta} U^{(2)}_{\epsilon,n}(\hat{\alpha}^{(1)}_{A,n}, \beta), \\
\hat{\alpha}^{(3)}_{A,n} &= \arg\sup_{\alpha\in\Theta_\alpha} U^{(3)}_{\epsilon,n}(\alpha, \hat{\beta}^{(2)}_{A,n}), \\
\hat{\beta}^{(4)}_{A,n} &= \arg\sup_{\beta\in\Theta_\beta} U^{(4)}_{\epsilon,n}(\hat{\alpha}^{(3)}_{A,n}, \beta).
\end{aligned}
$$

In order to compute the ML type estimator, we used **optim()** with the "L-BFGS-B" method in the R Language.

The hybrid estimators $\hat{\alpha}_{\epsilon,n}$ and $\hat{\beta}_{\epsilon,n}$ with full data $\mathbb{X}_n$ are computed as follows.

$$\begin{aligned} U^{(3)}_{\epsilon,n}(\hat{\alpha}_{\epsilon,n}, \hat{\beta}^{(2)}_{\epsilon,n}) &= \sup_{\alpha\in\Theta_\alpha} U^{(3)}_{\epsilon,n}(\alpha, \hat{\beta}^{(2)}_{\epsilon,n}), \\ U^{(4)}_{\epsilon,n}(\hat{\alpha}_{\epsilon,n}, \hat{\beta}_{\epsilon,n}) &= \sup_{\beta\in\Theta_\beta} U^{(4)}_{\epsilon,n}(\hat{\alpha}_{\epsilon,n}, \beta), \end{aligned}$$

where $\hat{\alpha}^{(1)}_{\epsilon,n}$ is obtained by using **optim()** for $U^{(1)}_{\epsilon,n}(\alpha)$ with the initial Bayes type estimator $\tilde{\alpha}^{(1)}_{\epsilon,n_0,r_1}$, and $\hat{\beta}^{(2)}_{\epsilon,n}$ is given by using **optim()** for $U^{(2)}_{\epsilon,n}(\hat{\alpha}^{(1)}_{\epsilon,n}, \beta)$ with the initial Bayes type estimator $\tilde{\beta}^{(2)}_{\epsilon,n_0,r_2}$.

For the true model, 100 independent sample paths are generated by the Euler-Maruyama scheme, and the mean and the standard deviation (s.d.) for the estimators are computed.

Tables 1-10 and 11-20 are simulation results for $\epsilon = 0.01$ and 0.05, respectively.

The time in each table is the computation time of estimation for one sample path.

The personal computer with Intel i7-5930K (3.5GHz base clock) was used for simulations.

3.1 In case that $\epsilon = 0.01$

Table 1: adaptive ML type estimator with the initial value being the true value

	$\hat{\alpha}_1(3)$	$\hat{\alpha}_2(7)$	$\hat{\alpha}_3(5)$	$\hat{\alpha}_4(2)$	$\hat{\alpha}_5(4)$	$\hat{\alpha}_6(6)$	$\hat{\beta}_1(1)$	$\hat{\beta}_2(2)$	$\hat{\beta}_3(3)$	time(sec.)
	3.000	7.000	5.000	2.000	4.000	6.000	1.000	2.000	3.000	
true	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.005)	(0.010)	(0.014)	40

Table 2: adaptive ML type estimator with the initial value being the uniform random number on Θ

	$\hat{\alpha}_1(3)$	$\hat{\alpha}_2(7)$	$\hat{\alpha}_3(5)$	$\hat{\alpha}_4(2)$	$\hat{\alpha}_5(4)$	$\hat{\alpha}_6(6)$	$\hat{\beta}_1(1)$	$\hat{\beta}_2(2)$	$\hat{\beta}_3(3)$	time(sec.)
	24.149	25.555	3.078	23.788	3.118	26.158	3.588	4.549	5.628	
unif	(15.155)	(15.678)	(2.877)	(17.492)	(1.005)	(15.668)	(1.247)	(1.439)	(1.189)	40

Table 1 shows the simulation results of the adaptive ML type estimator $(\hat{\alpha}^{(3)}_{A,n}, \hat{\beta}^{(4)}_{A,n})$ when the initial value is the true value. We see from Table 1 that all estimators have good behavior.
Table 2 is the simulation results of the adaptive ML type estimator $(\hat{\alpha}^{(3)}_{A,n}, \hat{\beta}^{(4)}_{A,n})$ with the initial value being the uniform random number on Θ. All estimators have considerable biases, which means that the optimization fails since the initial value may be far from the true value.
As we know very well, it is quite important to choose the initial value for optimization.

Table 3: initial Bayes type estimator $\tilde{\alpha}^{(1)}_{\epsilon,n_0,r_1}$ with $n_0 = 10^4$.

	$\hat{\alpha}_1(3)$	$\hat{\alpha}_2(7)$	$\hat{\alpha}_3(5)$	$\hat{\alpha}_4(2)$	$\hat{\alpha}_5(4)$	$\hat{\alpha}_6(6)$	time(hour)
	3.006	6.989	5.120	2.006	3.986	6.000	
$r_1 = 1.0$	(0.064)	(0.085)	(0.502)	(0.027)	(0.119)	(0.011)	8
	2.994	7.003	5.072	2.003	3.997	6.000	
$r_1 = 0.7$	(0.069)	(0.084)	(0.323)	(0.023)	(0.105)	(0.012)	8
	3.000	6.997	5.112	2.005	3.985	5.998	
$r_1 = 0.5$	(0.052)	(0.069)	(0.247)	(0.016)	(0.090)	(0.010)	8
	2.981	7.011	5.793	2.042	3.997	5.996	
$r_1 = 0.3$	(0.048)	(0.063)	(0.347)	(0.024)	(0.091)	(0.010)	8
	3.437	6.267	10.489	2.914	4.198	5.971	
$r_1 = 0.1$	(0.160)	(0.070)	(0.717)	(0.407)	(0.151)	(0.033)	8

Table 4: initial Bayes type estimator $\tilde{\beta}^{(2)}_{\epsilon,n_0,r_2}$ with $\tilde{\alpha}^{(1)}_{\epsilon,n_0,r_1}$ and $n_0 = 10^4$.

	$\hat{\beta}_1(1)$	$\hat{\beta}_2(2)$	$\hat{\beta}_3(3)$	time(hour)
	0.999	2.000	3.002	
$r_1 = 1.0, r_2 = 1.0$	(0.014)	(0.028)	(0.043)	1.5
	0.999	2.000	3.002	
$r_1 = 0.7, r_2 = 1.0$	(0.014)	(0.028)	(0.043)	1.5
	0.999	1.999	3.002	
$r_1 = 0.5, r_2 = 1.0$	(0.014)	(0.028)	(0.043)	1.5
	0.999	2.000	3.002	
$r_1 = 0.3, r_2 = 0.6$	(0.014)	(0.028)	(0.043)	1.5
	0.999	2.217	3.002	
$r_1 = 0.1, r_2 = 0.2$	(0.014)	(1.035)	(0.043)	1.5

Tables 3-4 show the simulation results of the initial Bayes type estimator $\tilde{\theta}_B = (\tilde{\alpha}^{(1)}_{\epsilon,n_0,r_1}, \tilde{\beta}^{(2)}_{\epsilon,n_0,r_2})$ when the sample size of the reduced data $n_0 = 10^4$ and the tuning parameters $(r_1, r_2) = (1.0, 1.0)$, $(0.7, 1.0)$, $(0.5, 1.0)$, $(0.3, 0.6)$ and $(0.1, 0.2)$.

In Table 3, the initial Bayes type estimators with $r_1 = 1.0, 0.7, 0.5$ have good behavior.

In Tables 4, most of the initial Bayes type estimators have good performance.

Table 5: hybrid estimator $\hat{\alpha}$ with the initial Bayes estimator $\tilde{\alpha}^{(1)}_{\epsilon,n_0,r_1}$ and $\tilde{\beta}^{(2)}_{\epsilon,n_0,r_2}$

	$\hat{\alpha}_1(3)$	$\hat{\alpha}_2(7)$	$\hat{\alpha}_3(5)$	$\hat{\alpha}_4(2)$	$\hat{\alpha}_5(4)$	$\hat{\alpha}_6(6)$	time(sec.)
	2.999	7.000	5.027	2.001	3.988	6.000	
$r_1 = 1.0,\ r_2 = 1.0$	(0.029)	(0.034)	(0.210)	(0.014)	(0.079)	(0.008)	40
	3.000	6.998	5.021	2.001	3.990	5.999	
$r_1 = 0.7,\ r_2 = 1.0$	(0.045)	(0.056)	(0.200)	(0.012)	(0.086)	(0.008)	40
	3.002	6.994	5.051	2.002	3.988	5.999	
$r_1 = 0.5,\ r_2 = 1.0$	(0.031)	(0.041)	(0.178)	(0.010)	(0.076)	(0.008)	40
	2.997	6.999	5.214	2.010	4.004	5.998	
$r_1 = 0.3,\ r_2 = 0.6$	(0.028)	(0.036)	(0.359)	(0.021)	(0.057)	(0.007)	40
	2.998	7.002	4.939	1.961	3.993	5.999	
$r_1 = 0.1,\ r_2 = 0.2$	(0.008)	(0.010)	(0.539)	(0.591)	(0.036)	(0.004)	40

Table 6: hybrid estimator $\hat{\beta}$ with $\tilde{\alpha}^{(1)}_{\epsilon,n_0,r_1}$ and $\tilde{\beta}^{(2)}_{\epsilon,n_0,r_2}$

	$\hat{\beta}_1(1)$	$\hat{\beta}_2(2)$	$\hat{\beta}_3(3)$	time(sec.)
	1.001	2.001	3.000	
$r_1 = 1.0,\ r_2 = 1.0$	(0.005)	(0.010)	(0.014)	40
	1.001	2.001	3.000	
$r_1 = 0.7,\ r_2 = 1.0$	(0.005)	(0.010)	(0.014)	40
	1.001	2.000	3.000	
$r_1 = 0.5,\ r_2 = 1.0$	(0.005)	(0.010)	(0.014)	40
	1.001	2.002	3.000	
$r_1 = 0.3,\ r_2 = 0.6$	(0.005)	(0.011)	(0.014)	40
	1.000	2.124	3.000	
$r_1 = 0.1,\ r_2 = 0.2$	(0.005)	(0.522)	(0.014)	40

Tables 5-6 show the results of the hybrid estimators $\hat{\theta}_n = (\hat{\alpha}_{\epsilon,n}, \hat{\beta}_{\epsilon,n})$ with the initial B. E.s in Tables 3-4.
In Tables 5-6, the hybrid estimators of α with the tuning parameters $(r_1, r_2) = (1.0, 1.0)$, $(0.7, 1.0)$, $(0.5, 1.0)$ have good behavior and
the hybrid estimators of β with the tuning parameters $(r_1, r_2) = (1.0, 1.0)$, $(0.7, 1.0)$, $(0.5, 1.0)$, $(0.3, 0.6)$ are unbiased.

Next, in order to compare with the hybrid estimator $(\hat{\alpha}_n, \hat{\beta}_n)$ based on the initial Bayes type estimator from reduced data, we consider the following two kinds of initial estimators $(\hat{\alpha}^{(1)}_{G,n_0}, \hat{\beta}^{(2)}_{G,n_0})$ and $(\hat{\alpha}^{(1)}_{U,n_0}, \hat{\beta}^{(2)}_{U,n_0})$. Let $n_0 = 10^4$.

Method G (Grid points method). For 18^6 points $\bar{\alpha}_{0,m}$ $(m = 1, \dots, 18^6)$ with 18 equally spaced points on each axis on $[0.1, 50]^6$, the initial estimator $\hat{\alpha}^{(1)}_{G,n_0}$ is defined as

$$U^{(1)}_{\epsilon,n_0}(\hat{\alpha}^{(1)}_{G,n_0}) = \max\left\{U^{(1)}_{\epsilon,n_0}(\bar{\alpha}_{0,1}), U^{(1)}_{\epsilon,n_0}(\bar{\alpha}_{0,2}), \dots, U^{(1)}_{\epsilon,n_0}(\bar{\alpha}_{0,18^6})\right\}.$$

Next, for 130^3 points $\bar{\beta}_{0,m}$ $(m = 1, \dots, 130^3)$ with 130 equally spaced points on each axis on $[0.1, 50]^3$, the initial estimator $\hat{\beta}^{(2)}_{G,n_0}$ is defined as

$$U^{(2)}_{\epsilon,n_0}(\hat{\alpha}^{(1)}_{G,n_0}, \hat{\beta}^{(2)}_{G,n_0}) = \max\left\{U^{(2)}_{\epsilon,n_0}(\hat{\alpha}^{(1)}_{G,n_0}, \bar{\beta}_{0,1}), U^{(2)}_{\epsilon,n_0}(\hat{\alpha}^{(1)}_{G,n_0}, \bar{\beta}_{0,2}), \dots, U^{(2)}_{\epsilon,n_0}(\hat{\alpha}^{(1)}_{G,n_0}, \bar{\beta}_{0,130^3})\right\}$$

Method U (Uniform r.n. + optim() method). Using 7^6 uniform random numbers $\alpha_{0,m}$ $(m = 1, \dots, 7^6)$ on $[0.1, 50]^6$, we compute

$$\hat{\alpha}_m^{(1)} = \arg\sup_{\alpha} U_{\epsilon,n_0}^{(1)}(\alpha)$$

by means of **optim()** with each initial value $\alpha_{0,m}$. The initial estimator $\hat{\alpha}_{U,n_0}^{(1)}$ is defined as

$$U_{\epsilon,n_0}^{(1)}(\hat{\alpha}_{U,n_0}^{(1)}) = \max\left\{U_{\epsilon,n_0}^{(1)}(\hat{\alpha}_1^{(1)}), U_{\epsilon,n_0}^{(1)}(\hat{\alpha}_2^{(1)}), \dots, U_{\epsilon,n_0}^{(1)}(\hat{\alpha}_{7^6}^{(1)})\right\}.$$

Next, using 34^3 uniform random numbers $\beta_{0,m}$ $(m = 1, \dots, 34^3)$ on $[0.1, 50]^3$, we compute

$$\hat{\beta}_m^{(2)} = \arg\sup_{\beta} U_{\epsilon,n_0}^{(2)}(\hat{\alpha}_{U,n_0}^{(1)}, \beta)$$

by means of **optim()** with each initial value $\beta_{0,m}$. The initial estimator $\hat{\beta}_{U,n_0}^{(2)}$ is defined as

$$U_{\epsilon,n_0}^{(2)}(\hat{\alpha}_{U,n_0}^{(1)}, \hat{\beta}_{U,n_0}^{(2)}) = \max\left\{U_{\epsilon,n_0}^{(2)}(\hat{\alpha}_{U,n_0}^{(1)}, \hat{\beta}_1^{(2)}), U_{\epsilon,n_0}^{(2)}(\hat{\alpha}_{U,n_0}^{(1)}, \hat{\beta}_2^{(2)}), \dots, U_{\epsilon,n_0}^{(2)}(\hat{\alpha}_{U,n_0}^{(1)}, \hat{\beta}_{34^3}^{(2)})\right\}.$$

Let $k = G, U$. The hybrid estimator $(\bar{\alpha}_{k,n}^{(3)}, \bar{\beta}_{k,n}^{(4)})$ is computed as follows.

$$\begin{aligned} \bar{\alpha}_{k,n}^{(3)} &= \arg\sup_{\alpha\in\Theta_\alpha} U_{\epsilon,n}^{(3)}(\alpha, \bar{\beta}_{k,n}^{(2)}), \\ \bar{\beta}_{k,n}^{(4)} &= \arg\sup_{\beta\in\Theta_\beta} U_{\epsilon,n}^{(4)}(\bar{\alpha}_{k,n}^{(3)}, \beta), \end{aligned}$$

where $\bar{\alpha}_{k,n}^{(1)}$ is obtained by using **optim()** for $U_{\epsilon,n}^{(1)}(\alpha)$ with the initial estimator $\hat{\alpha}_{k,n_0}^{(1)}$, and $\bar{\beta}_{k,n}^{(2)}$ is given by using **optim()** for $U_{\epsilon,n}^{(2)}(\bar{\alpha}_{k,n}^{(1)}, \beta)$ with the initial estimator $\hat{\beta}_{k,n_0}^{(2)}$.

Let $\hat{\theta}_B = (\hat{\alpha}_n, \hat{\beta}_n)$ with the initial Bayes type estimator $\tilde{\theta}_B$. Let $\hat{\theta}_G = (\bar{\alpha}_{G,n}^{(3)}, \bar{\beta}_{G,n}^{(4)})$ with $\tilde{\theta}_G = (\hat{\alpha}_{G,n_0}^{(1)}, \hat{\beta}_{G,n_0}^{(2)})$ and $\hat{\theta}_U = (\bar{\alpha}_{U,n}^{(3)}, \bar{\beta}_{U,n}^{(4)})$ with $\tilde{\theta}_U = (\hat{\alpha}_{U,n_0}^{(1)}, \hat{\beta}_{U,n_0}^{(2)})$.

Table 7: initial estimators $\tilde{\alpha}_B = \tilde{\alpha}_{\epsilon,n_0,r_1}$ $(r_1 = 0.7)$, $\tilde{\alpha}_G$ $(18^6 + 130^3$ lattice points), $\tilde{\alpha}_U$ $(7^6 + 34^3$ random numbers) with $n_0 = 10^4$.

	$\hat{\alpha}_1(3)$	$\hat{\alpha}_2(7)$	$\hat{\alpha}_3(5)$	$\hat{\alpha}_4(2)$	$\hat{\alpha}_5(4)$	$\hat{\alpha}_6(6)$	time(hour)
$\tilde{\alpha}_B$	2.994 (0.068)	7.004 (0.084)	5.072 (0.326)	2.003 (0.023)	3.997 (0.105)	5.999 (0.012)	8
$\tilde{\alpha}_G$	0.010 (0)	11.121 (0)	16.677 (0)	2.787 (0)	4.149 (1.395)	5.566 (0)	12
$\tilde{\alpha}_U$	3.698 (3.013)	6.919 (2.566)	6.565 (3.316)	2.901 (5.428)	3.850 (0.822)	5.838 (1.266)	15

Table 8: initial estimators $\tilde{\beta}_B = \tilde{\beta}_{\epsilon,n_0,r_1,r_2}$ $(r_1 = 0.7, r_2 = 1.0)$, $\tilde{\beta}_G$ $(18^6 + 130^3$ lattice points), $\tilde{\beta}_U$ $(7^6 + 34^3$ random numbers) with $n_0 = 10^4$.

	$\hat{\beta}_1(1)$	$\hat{\beta}_2(2)$	$\hat{\beta}_3(3)$	time(hour)
$\tilde{\beta}_B$	0.999 (0.014)	2.000 (0.028)	3.002 (0.043)	1.5
$\tilde{\beta}_G$	1.979 (1.464)	1.964 (0.308)	3.087 (0)	1.5
$\tilde{\beta}_U$	1.079 (0.406)	2.342 (1.031)	3.098 (0.441)	1.5

Tables 7-8 show the simulation results of the initial estimators $\tilde{\theta}_B = (\tilde{\alpha}_B, \tilde{\beta}_B)$, $\tilde{\theta}_G = (\tilde{\alpha}_G, \tilde{\beta}_G)$, $\tilde{\theta}_U = (\tilde{\alpha}_U, \tilde{\beta}_U)$.
In Tables 7-8, although $\tilde{\theta}_G$ and $\tilde{\theta}_U$ have considerable biases, $\tilde{\theta}_B$ with $(r_1, r_2) = (0.7, 1.0)$ is unbiased.

Table 9: hybrid estimators $\hat{\alpha}_B$, $\hat{\alpha}_G$ and $\hat{\alpha}_U$ with $\tilde{\alpha}_B$, $\tilde{\alpha}_G$ and $\tilde{\alpha}_U$, respectively.

	$\hat{\alpha}_1(3)$	$\hat{\alpha}_2(7)$	$\hat{\alpha}_3(5)$	$\hat{\alpha}_4(2)$	$\hat{\alpha}_5(4)$	$\hat{\alpha}_6(6)$	time(sec.)
	3.000	6.998	5.021	2.001	3.990	5.999	
$\hat{\alpha}_B$	(0.045)	(0.056)	(0.200)	(0.012)	(0.086)	(0.008)	40
	2.489	10.153	4.965	1.981	4.001	5.999	
$\hat{\alpha}_G$	(1.853)	(4.444)	(0.320)	(0.197)	(0.031)	(0.004)	40
	3.689	6.933	4.878	2.760	3.951	5.832	
$\hat{\alpha}_U$	(3.018)	(2.564)	(1.662)	(5.263)	(0.245)	(1.386)	40

Table 10: hybrid estimators $\hat{\beta}_B$, $\hat{\beta}_G$ and $\hat{\beta}_U$ with $\tilde{\theta}_B$, $\tilde{\theta}_G$ and $\tilde{\theta}_U$, respectively.

	$\hat{\beta}_1(1)$	$\hat{\beta}_2(2)$	$\hat{\beta}_3(3)$	time(sec.)
	1.001	2.001	3.000	
$\hat{\beta}_B$	(0.005)	(0.010)	(0.014)	40
	2.007	2.000	3.000	
$\hat{\beta}_G$	(1.404)	(0.010)	(0.014)	40
	1.115	2.295	3.104	
$\hat{\beta}_U$	(0.504)	(0.817)	(0.472)	40

Tables 9 -10 show the results of the hybrid estimators $\hat{\theta}_B = (\hat{\alpha}_B, \hat{\beta}_B)$, $\hat{\theta}_G = (\hat{\alpha}_G, \hat{\beta}_G)$, $\hat{\theta}_U = (\hat{\alpha}_U, \hat{\beta}_U)$.
In Tables 9-10, $\hat{\theta}_G$ and $\hat{\theta}_U$ have considerable biases.
On the other hand, $\hat{\theta}_B$ with $(r_1, r_2) = (0.7, 1.0)$ in Table 9 has good behavior.

3.2. In case that $\epsilon = 0.05$

Table 11: adaptive ML type estimator with the initial value being the true value

	$\hat{\alpha}_1(3)$	$\hat{\alpha}_2(7)$	$\hat{\alpha}_3(5)$	$\hat{\alpha}_4(2)$	$\hat{\alpha}_5(4)$	$\hat{\alpha}_6(6)$	$\hat{\beta}_1(1)$	$\hat{\beta}_2(2)$	$\hat{\beta}_3(3)$	time(sec.)
	2.999	6.999	5.002	2.000	3.997	6.000	1.000	2.000	3.000	
true	(0.009)	(0.012)	(0.034)	(0.006)	(0.073)	(0.008)	(0.005)	(0.010)	(0.014)	40

Table 12: adaptive ML type estimator with the initial value being the uniform random number on Θ

	$\hat{\alpha}_1(3)$	$\hat{\alpha}_2(7)$	$\hat{\alpha}_3(5)$	$\hat{\alpha}_4(2)$	$\hat{\alpha}_5(4)$	$\hat{\alpha}_6(6)$	$\hat{\beta}_1(1)$	$\hat{\beta}_2(2)$	$\hat{\beta}_3(3)$	time(sec.)
	24.226	24.702	2.890	25.031	3.156	26.583	1.114	2.117	3.104	
unif	(15.068)	(15.596)	(2.618)	(15.876)	(1.042)	(15.534)	(0.049)	(0.050)	(0.055)	40

Tables 11 shows the simulation results of the adaptive ML type estimator $(\hat{\alpha}^{(3)}_{A,n}, \hat{\beta}^{(4)}_{A,n})$ when the initial value is the true value.
We see from Tables 11 that all estimators have good behavior.
Tables 12 shows the simulation results of the adaptive ML type estimator $(\hat{\alpha}^{(3)}_{A,n}, \hat{\beta}^{(4)}_{A,n})$ with the initial value being the uniform random number on Θ.
All estimators have considerable biases, which means that the optimization fails since the initial value may be far from the true value.

Table 13: initial Bayes type estimator $\tilde{\alpha}^{(1)}_{\epsilon,n_0,r_1}$ with $n_0 = 10^4$.

	$\hat{\alpha}_1(3)$	$\hat{\alpha}_2(7)$	$\hat{\alpha}_3(5)$	$\hat{\alpha}_4(2)$	$\hat{\alpha}_5(4)$	$\hat{\alpha}_6(6)$	time(hour)
$r_1 = 1.0$	3.587 (2.638)	6.654 (1.389)	5.317 (1.457)	2.383 (3.751)	3.966 (0.483)	5.984 (0.081)	8
$r_1 = 0.7$	3.388 (2.564)	7.183 (3.408)	5.386 (1.403)	2.443 (4.505)	3.962 (0.483)	5.986 (0.082)	8
$r_1 = 0.5$	4.432 (5.982)	7.235 (4.564)	5.632 (1.379)	2.0279 (0.091)	3.976 (0.497)	5.976 (0.082)	8
$r_1 = 0.3$	3.506 (1.088)	6.607 (0.536)	7.412 (1.956)	2.134 (0.139)	4.020 (0.544)	5.974 (0.142)	8
$r_1 = 0.1$	5.734 (2.384)	6.873 (0.980)	10.529 (2.878)	4.623 (1.311)	4.344 (0.757)	6.369 (0.441)	8

Table 14: initial Bayes type estimator $\tilde{\beta}^{(2)}_{\epsilon,n_0,r_2}$ with $\tilde{\alpha}^{(1)}_{\epsilon,n_0,r_1}$ and $n_0 = 10^4$.

	$\hat{\beta}_1(1)$	$\hat{\beta}_2(2)$	$\hat{\beta}_3(3)$	time(hour)
$r_1 = 1.0, r_2 = 1.0$	1.001 (0.015)	2.002 (0.034)	3.002 (0.043)	1.5
$r_1 = 0.7, r_2 = 1.0$	1.002 (0.024)	2.002 (0.028)	3.002 (0.044)	1.5
$r_1 = 0.5, r_2 = 1.0$	1.003 (0.020)	2.000 (0.028)	3.002 (0.043)	1.5
$r_1 = 0.3, r_2 = 0.6$	1.001 (0.019)	2.000 (0.028)	3.002 (0.043)	1.5
$r_1 = 0.1, r_2 = 0.2$	1.033 (0.063)	2.052 (0.101)	3.006 (0.052)	1.5

Tables 13-14 show the simulation results of the initial Bayes type estimator $\tilde{\theta}_B = (\tilde{\alpha}^{(1)}_{\epsilon,n_0,r_1}, \tilde{\beta}^{(2)}_{\epsilon,n_0,r_2})$ when the sample size of the reduced data $n_0 = 10^4$ and the tuning parameters $(r_1, r_2) = (1.0, 1.0)$, $(0.7, 1.0)$, $(0.5, 1.0)$, $(0.3, 0.6)$ and $(0.1, 0.2)$.

In Table 13, all initial Bayes type estimators have considerable biases.

In Tables 14, most of the initial Bayes estimators have good performance.

Table 15: hybrid estimator $\hat{\alpha}$ with the initial Bayes estimator $\tilde{\alpha}^{(1)}_{\epsilon,n_0,r_1}$ and $\tilde{\beta}^{(2)}_{\epsilon,n_0,r_2}$

	$\hat{\alpha}_1(3)$	$\hat{\alpha}_2(7)$	$\hat{\alpha}_3(5)$	$\hat{\alpha}_4(2)$	$\hat{\alpha}_5(4)$	$\hat{\alpha}_6(6)$	time(sec.)
$r_1 = 1.0, r_2 = 1.0$	3.438 (2.265)	6.762 (1.200)	5.078 (0.819)	2.441 (4.370)	4.012 (0.214)	5.994 (0.026)	40
$r_1 = 0.7, r_2 = 1.0$	3.350 (2.614)	7.325 (4.157)	5.144 (0.790)	2.431 (4.506)	4.039 (0.194)	5.995 (0.024)	40
$r_1 = 0.5, r_2 = 1.0$	4.262 (5.677)	7.346 (4.657)	5.084 (0.418)	2.002 (0.024)	4.039 (0.212)	5.995 (0.025)	40
$r_1 = 0.3, r_2 = 0.6$	3.085 (0.790)	6.972 (0.205)	5.145 (0.617)	2.004 (0.026)	4.027 (0.170)	5.995 (0.024)	40
$r_1 = 0.1, r_2 = 0.2$	4.862 (3.425)	7.034 (1.888)	4.880 (2.015)	3.418 (2.038)	3.925 (0.559)	5.954 (0.653)	40

Table 16: hybrid estimator $\hat{\beta}$ with $\tilde{\alpha}^{(1)}_{\epsilon,n_0,r_1}$ and $\tilde{\beta}^{(2)}_{\epsilon,n_0,r_2}$

	$\hat{\beta}_1(1)$	$\hat{\beta}_2(2)$	$\hat{\beta}_3(3)$	time(sec.)
$r_1 = 1.0, r_2 = 1.0$	1.004 (0.019)	2.002 (0.019)	3.000 (0.014)	40
$r_1 = 0.7, r_2 = 1.0$	1.002 (0.015)	2.001 (0.015)	3.000 (0.014)	40
$r_1 = 0.5, r_2 = 1.0$	1.007 (0.030)	2.000 (0.010)	3.000 (0.014)	40
$r_1 = 0.3, r_2 = 0.6$	1.001 (0.007)	2.000 (0.010)	3.000 (0.014)	40
$r_1 = 0.1, r_2 = 0.2$	1.025 (0.046)	2.039 (0.060)	3.003 (0.021)	40

Tables 15-16 show the results of the hybrid estimators $\hat{\theta}_n = (\hat{\alpha}_{\epsilon,n}, \hat{\beta}_{\epsilon,n})$ with the initial Bayes type estimators in Tables 13-14, respectively.
In Tables 15-16, the hybrid estimator of (α, β) with $(r_1, r_2) = (0.3, 0.6)$ is best among the competing hybrid estimators.

Table 17: initial estimators $\tilde{\alpha}_B = \tilde{\alpha}_{\epsilon,n_0,r_1,r_2}$ $(r_1 = 0.3, r_2 = 0.6)$, $\tilde{\alpha}_G$ $(18^6 + 130^3$ lattice points), $\tilde{\alpha}_U$ $(7^6 + 34^3$ random numbers) with $n_0 = 10^4$.

	$\hat{\alpha}_1(3)$	$\hat{\alpha}_2(7)$	$\hat{\alpha}_3(5)$	$\hat{\alpha}_4(2)$	$\hat{\alpha}_5(4)$	$\hat{\alpha}_6(6)$	time(hour)
	3.506	6.607	7.412	2.134	4.020	5.974	
$\tilde{\alpha}_B$	(1.088)	(0.536)	(1.956)	(0.139)	(0.544)	(0.142)	8
	4.566	7.538	13.566	2.316	4.510	5.566	
$\tilde{\alpha}_G$	(6.878)	(4.872)	(5.615)	(1.120)	(1.617)	(0.000)	12
	4.130	6.387	6.596	2.671	3.864	6.001	
$\tilde{\alpha}_U$	(3.712)	(1.909)	(4.069)	(4.410)	(0.871)	(1.258)	15

Table 18: initial estimators $\tilde{\beta}_B = \tilde{\beta}_{\epsilon,n_0,r_1,r_2}$ $(r_1 = 0.3, r_2 = 0.6)$, $\tilde{\beta}_G$ $(18^6 + 130^3$ lattice points), $\tilde{\beta}_U$ $(7^6 + 34^3$ random numbers) with $n_0 = 10^4$.

	$\hat{\beta}_1(1)$	$\hat{\beta}_2(2)$	$\hat{\beta}_3(3)$	time(hour)
	1.001	2.000	3.002	
$\tilde{\beta}_B$	(0.019)	(0.028)	(0.043)	1.5
	1.160	1.945	3.087	
$\tilde{\beta}_G$	(0.038)	(0.066)	(0.000)	1.5
	1.004	2.010	3.006	
$\tilde{\beta}_U$	(0.023)	(0.041)	(0.049)	1.5

Tables 17-18 show the simulation results of the initial estimators $\tilde{\theta}_B = (\tilde{\alpha}_B, \tilde{\beta}_B)$, $\tilde{\theta}_G = (\tilde{\alpha}_G, \tilde{\beta}_G)$, $\tilde{\theta}_U = (\tilde{\alpha}_U, \tilde{\beta}_U)$.

In Tables 17-18, $\tilde{\beta}_B$, $\tilde{\beta}_G$ and $\tilde{\beta}_U$ are unbiased, but $\tilde{\alpha}_B$ with $(r_1, r_2) = (0.3, 0.6)$, $\tilde{\alpha}_G$ and $\tilde{\alpha}_U$ have considerable biases.

Table 19: hybrid estimators $\hat{\alpha}_B$, $\hat{\alpha}_G$ and $\hat{\alpha}_U$ with $\tilde{\alpha}_B$, $\tilde{\alpha}_G$ and $\tilde{\alpha}_U$, respectively.

	$\hat{\alpha}_1(3)$	$\hat{\alpha}_2(7)$	$\hat{\alpha}_3(5)$	$\hat{\alpha}_4(2)$	$\hat{\alpha}_5(4)$	$\hat{\alpha}_6(6)$	time(sec.)
	3.085	6.972	5.145	2.004	4.027	5.995	
$\hat{\alpha}_B$	(0.790)	(0.205)	(0.617)	(0.026)	(0.170)	(0.024)	40
	5.607	7.205	4.259	1.651	4.017	6.046	
$\hat{\alpha}_G$	(5.801)	(4.904)	(1.768)	(0.817)	(0.163)	(0.509)	40
	3.997	6.561	4.992	2.765	4.037	6.035	
$\hat{\alpha}_U$	(3.607)	(1.488)	(1.064)	(4.424)	(0.263)	(1.142)	40

Table 20: hybrid estimators $\hat{\beta}_B$, $\hat{\beta}_G$ and $\hat{\beta}_U$ with $\tilde{\theta}_B$, $\tilde{\theta}_G$ and $\tilde{\theta}_U$, respectively.

	$\hat{\beta}_1(1)$	$\hat{\beta}_2(2)$	$\hat{\beta}_3(3)$	time(sec.)
	1.001	2.000	3.000	
$\hat{\beta}_B$	(0.007)	(0.010)	(0.014)	40
	1.051	2.012	3.002	
$\hat{\beta}_G$	(0.044)	(0.027)	(0.024)	40
	1.006	2.011	3.003	
$\hat{\beta}_U$	(0.022)	(0.036)	(0.021)	40

Tables 19-20 show the results of the hybrid estimators $\hat{\theta}_B = (\hat{\alpha}_B, \hat{\beta}_B)$, $\hat{\theta}_G = (\hat{\alpha}_G, \hat{\beta}_G)$, $\hat{\theta}_U = (\hat{\alpha}_U, \hat{\beta}_U)$. In Tables 19, $\hat{\alpha}_G$ and $\hat{\alpha}_U$ have considerable biases. On the other hand, $\hat{\alpha}_B$ with $(r_1, r_2) = (0.3, 0.6)$ in Table 19 has good behavior.

Concluding remarks:

In order to calculate the ML type estimator by optim() in R language, it is quite crucial to select a suitable initial value.

It is useful to obtain the Bayes type estimator as an initial estimator since the Bayes type estimator does not strongly depend on the initial value.

We need to develop the theory of computational statistics, which expands high speed and rigorous high frequency data analysis.