Parameter estimation for stochastic differential equations with small Lévy noise

小分散レヴィノイズを伴う確率微分方程式におけるパラメータ推定

July, 2023

Mitsuki KOBAYASHI
小林　光木

Parameter estimation for stochastic differential equations with small Lévy noise

小分散レヴィノイズを伴う確率微分方程式におけるパラメータ推定

July, 2023

Waseda University Graduate School of Fundamental Science and Engineering

Department of Pure and Applied Mathematics, Research on Stochastic and Statistical Analysis

Mitsuki KOBAYASHI
小林　光木

Acknowledgements

I would like to express my deepest gratitude to Professor Yasutaka Shimizu and Professor Yoshio Yamada for their invaluable patience and guidance. They told me how to navigate an uncharted water and how interest it was.

I could not have undertaken this journey without my cohort members, especially Jun Yoshida and Shohei Nakajima. Without Yoshida's encouragement I probably did not challenge myself to new fields, and Nakajima always gave me an incentive to hard work. I am also grateful to the disciples of Prof. Shimizu for their taking part in a lot of private seminars with me including late-night sessions, and for their moral support.

Lastly, I would be remiss in not mentioning my familiy members, especially my parents Kiyoki and Michiko, and my sister Yuri and Mami. Their belief in me has kept my spirits and motivation high throughout this process. I would also like to thank to my friends for all the entertainment and emotional support.

Contents

Chapter 1

Introduction

In this thesis, we study two types of parametric estimations for unknown parameters in the coefficient functions of stochastic differential equations (SDEs) with small Lévy noise, and we establish our estimators from a discrete sample path derived from the SDE with true parameters in the coefficient functions, called discretely observed case. Problems of parametric estimation for discretely observed stochastic processes with small diffusion have been studied by various authors (*e.g.*, Genon-Catalot [10], Laredo [21], Sørensen and Uchida [33] and so on) and problems of ones with small Lévy noise have been studied by Long *et al.* [22], Long *et al.* [23] and references therein.

In the first half of this thesis, we focus on the parametric estimation for drift parameter only, and we propose a new type of least square estimator (LSE) based on the Adams method, which is well-known as a numerical computation method for calculating numerical solutions to ordinary differential equations (ODEs). Then, we prove the consistency and the asymptotic normality of of the proposed estimator, and we show that the LSEs based on the Adams method can be better than the usual LSE based on the Euler method in the finite sample performance.

In order to say more precisely about our estimators in the first half of this thesis, let us introduce the well-known Adams method in numerical analysis for ODEs (see, *e.g.*, Butcher [7], Hairer *et al.* [14], Hairer and Wanner [15] and Iserles [16]), which is the combinations of two methods as predictor-corrector pair, says, the Adams-Bashforth and the Adams-Moulton formulae. Here, we give an ODE

$$\frac{\mathrm{d}x_t}{\mathrm{d}t} = a(x_t) \tag{1.1}$$

with the initial condition x_0, where a is a Lipschitz function. For instance, to compute an approximate value $\hat{x}_{t_k}$ of the solution of (1.1) at $t = t_k$, we firstly prepare a predictor $x^*_{t_k}$ given by Adams-Bashforth method with $\ell = 1, 2, \ldots$ as

$$x^*_{t_k} = \hat{x}_{t_{k-1}} + \frac{1}{n}\sum_{\nu=1}^{\ell} \gamma_{\ell\nu} a(\hat{x}_{t_{k-\nu}}), \quad \gamma_{\ell\nu} := \frac{(-1)^{\nu-1}}{(\nu-1)!(\ell-\nu)!}\int_0^1 \prod_{\substack{j=1\\ j\neq\nu}}^{\ell} (u + j - 1)\,\mathrm{d}u\,, \tag{1.2}$$

by using the past approximate values $\hat{x}_{t_{k-1}}, \ldots, \hat{x}_{t_{k-\ell}}$ with $\hat{x}_{t_0} = x_0$, and we secondly modify the value $x^*_{t_k}$ to a corrector $\hat{x}_{t_k}$ given by Adams-Moulton method as

$$\hat{x}_{t_k} = \hat{x}_{t_{k-1}} + \frac{1}{n}\beta_{\ell 0} a(x^*_{t_k}) + \frac{1}{n}\sum_{\nu=1}^{\ell} \beta_{\ell\nu} a(\hat{x}_{t_{k-\nu}}), \quad \beta_{\ell\nu} := \frac{(-1)^\nu}{\nu!(\ell-\nu)!}\int_0^1 \prod_{\substack{j=0\\ j\neq\nu}}^{\ell} (u+j-1)\,\mathrm{d}u\,. \tag{1.3}$$

Both formulae follows by the same argument as in Section 2.1 in Iserles [16], and the predictor-corrector scheme is written in Hairer and Wanner [15]. Some of the values of the coefficients $\gamma_{\ell\nu}$, $\beta_{\ell\nu}$ can be seen in Table 244 in Butcher [7]. Here, we remark that for any $g : \mathbb{R} \to \mathbb{R}$, the coefficients $\gamma_{\ell\nu}$ and $\beta_{\ell\nu}$ satisfy

$$\begin{aligned} \int_{t_{k-1}}^{t_k} P(s; g, t_{k-1}, \dots, t_{k-\ell})\,\mathrm{d}s &= \frac{1}{n}\sum_{\nu=1}^{\ell} \gamma_{\ell\nu} g(x_{t_{k-\nu}}), \\ \int_{t_{k-1}}^{t_k} P(s; g, t_k, \dots, t_{k-\ell})\,\mathrm{d}s &= \frac{1}{n}\sum_{\nu=0}^{\ell} \beta_{\ell\nu} g(x_{t_{k-\nu}}), \end{aligned} \tag{1.4}$$

where $s \mapsto P(s; g, t_k, \dots, t_{k-\ell})$ is the Lagrange interpolating polynomial through the points $(s, g(s))$, $s = t_k, \dots, t_{k-\ell}$ (see, *e.g.*, Section III.1 in Hairer *et al.* [14]). In particular, substituting $g \equiv 1$, we have

$$\sum_{\nu=1}^{\ell} \gamma_{\ell\nu} = \sum_{\nu=0}^{\ell} \beta_{\ell\nu} = 1.$$

Here, we suppose that we have a discrete sample $\{X^\varepsilon_{t_k}\}_{k=0,\dots,n}$ derived by

$$\mathrm{d}X^\varepsilon_t = a(X^\varepsilon_t, \theta_0)\,\mathrm{d}t + \varepsilon\,\mathrm{d}L_t\,, \quad X^\varepsilon_0 = x_0 \in \mathbb{R}^d,$$

where Θ_0 is a smooth bounded open convex set in $\mathbb{R}^p$ with $p \in \mathbb{N}$, Θ denotes the closure of Θ_0, $\theta \in \Theta$, $\varepsilon > 0$, a is a function from $\mathbb{R}^d \times \Theta$ to $\mathbb{R}^d$, and $L = (L_t)_{t\leq 0}$ is a d-dimensional Lévy process. As in the usual literature (*e.g.*, Long et al. [22]), we use the following contrast function for LSE based on Euler method:

$$\Psi_{n,\varepsilon}(\theta) := \frac{n}{\varepsilon^2}\sum_{k=1}^{n}\left| X^\varepsilon_{t_k} - X^\varepsilon_{t_{k-1}} - \frac{1}{n}a(X^\varepsilon_{t_{k-1}}, \theta)\right|^2,$$

though the Euler method sometimes fails to approximate the solution of ODEs (*e.g.*, $a(x,\theta) = -\theta x$ for $x, \theta > 0$ and $\theta/n \notin (0,2)$, in Section 4.2 in Iserles [16]) and is less accurate than the Runge-Kutta method, the Adams method, etc. Of course, these numerical approximation methods except for the Euler method do not work to calculate numerical solutions to SDEs, while they are available for the ODE given in the limit $\varepsilon \to 0$. Thus, we employ the Adams method instead of the Euler method and define the *Adams-Bashforth type* contrast function $\Psi_{n,\varepsilon,\ell}(\theta)$ as

$$\tilde{\Psi}_{n,\varepsilon,\ell}(\theta) := \frac{n}{\varepsilon^2}\sum_{k=\ell\vee 1}^{n}\left| X^\varepsilon_{t_k} - X^\varepsilon_{t_{k-1}} - \frac{1}{n}\sum_{\nu=1}^{\ell}\gamma_{\ell\nu} a(X^\varepsilon_{t_{k-\nu}}, \theta)\right|^2.$$

We can also define the *Adams-Moulton type* contrast function as we shall see in Chapter 2, and will discuss only the *Adams-Moulton type* contrast function, since the proof for *Adams-Moulton type* is analogous.

In the second half, we expand our scope to the joint estimation of the parameters in the drift, diffusion and jump terms, while we restrict Lévy noise to compound Poisson process. In the ergodic case, such joint estimation for SDEs with Lévy noise is proposed in Shimizu and Yoshida [32], and has been considered so far by various researchers (see, *e.g.*, [1, 12, 25, 31], and other references are given in Amorino and Gloter [2]). On the other hand, in the small noise case, no one has succeeded in giving a proof for such joint threshold estimation of the parameter relative to drift, diffusion and jumps. So, the aim of the second half of this thesis is to give a framework and a proof for the threshold estimation in the small noise case.

As an essential part of our framework for estimation, we suppose not only $n \to \infty$ and $\varepsilon \to 0$ but $\lambda_\varepsilon \to \infty$, while the intensity λ_ε is fixed, $n \to \infty$ and $\varepsilon \to 0$ in the previous works of estimations for SDEs with small noise (see, *e.g.*, [13, 20, 22, 33], and references are given in [23]). The asymptotics with $\lambda_\varepsilon \to \infty$ would be the first and new attempt in many works of literature, and enables us to deal with the joint estimation of the parameter (μ, σ, α) relative to drift, diffusion and jumps, while the papers above deal with only the estimation of drift and diffustion parameters (or in some papers drift parameter only). Indeed, one can immediately notice that if the intensity λ_ε is constant, then the number of large jumps never goes to infinity in probability, and so we would never establish a consistent estimator of jump size density. Therefore, we suppose that $\lambda_\varepsilon \to \infty$ as $\varepsilon \downarrow 0$ (λ_ε is not necessary to depend on ε as in Remark 3.2.4). Also, the assumption $\lambda_\varepsilon \to \infty$ seems natural when we deal with data obtained in the long term with the pitch of observations shortened, which is familiar in both cases of ergodic and small noise. Thus, one can agree with our proposal.

Another attempt in the second half of this thesis is to give a proof by using localization argument (as in, *e.g.*, Remark 1 in Sørensen and Uchida [33]) in the entire context, though the argument is usually omitted, or instead, Proposition 1 in Gloter and Sørensen [13] is just referred. As to the proof, we prepare the localization assumptions for jump size densities, *i.e.*, Assumptions 3.2.9 to 3.2.12, together with usual localization assumptions for coefficient functions in (3.1), *i.e.*, Assumptions 3.2.5 and 3.2.6. Owing to prepare Assumptions 3.2.9 to 3.2.12, this thesis has more examples of jump size densities than the papers [25, 32] (see Section 3.5 in this thesis, and see, *e.g.*, Ogihara and Yoshida [25, Example]). On the other hand, Assumptions 3.2.9 to 3.2.12 are too complicated for us to omit the localization argument. Thus, we show our main results under the localization argument in the entirety of our proof, which is one of the novelties.

A further attempt of the second half of this thesis is to simplify the contrast functions used in earlier works [25, 32] by removing φ_n defined in [25, 32] from their contrast functions. As we mentioned above, the class of jump size densities is wide and includes unbounded densities (*e.g.*, log-normal distribution) which are not included in [25, 32]. Note that the class of jump size densities in Shimizu [27] is also wide ([27] does not assume the twice differentiability of jump size densities, while conversely this thesis does not assume $\int |z|^p \frac{\partial}{\partial \alpha_j} f_\alpha(z)\,\mathrm{d}z$ ($p \geq 1$) as in the assumption A5 in [27]), but [27] is concerned with moment estimators in the ergodic case.

In order to see the behavior of our estimator in numerical experiments, we give Table 3.6.1 under the assumption that λ_ε is known. Of course, this assumption is impractical when we deal with only observations, and how to choose threshold v_{nk}/n^ρ in filters $1_{C_k^{n,\varepsilon,\rho}}$ and $1_{D_k^{n,\varepsilon,\rho}}$ defined in Notation 3.2.7 is one of the crucial points for estimation with jumps, but it is not within the scope of this thesis (see, *e.g.*, Shimizu [29, 30] for the readers who are interested in the techniques of the way to choose such threshold, and then Lemma 3.4.8 may also help you estimate the intensity λ_ε). Instead of this discussion, we give another experiment as in Table 3.6.2 to see what will occur by using different thresholds.

In Section 2.3, we set up some notations, assumptions and preliminary propositions for the proof of main results. In Section 2.2, we state our main result and give their proof. In Section 2.4, we give a simulation by numerical computation to compare our estimators with well-known least squares estimators for an Ornstein-Uhlenbeck process.

In Section 3.2, we set up some assumptions and notations. In Section 3.3, we state our main results, *i.e.*, the consistency and the asymptotic normality of our estimator. In Section 3.4, we give a proof of our main results. In Section 3.5, we give some examples of the jump size density for compound Poisson processes in our model. In Section 3.6, we give two numerical experiments to see the finite sample performance of our estimator. In Section 3.7, we state and prove some slightly different versions of well-known results.

Chapter 2

LSE based on Adams method

In this thesis, we first concerned with the following $\mathbb{R}^d$-valued stochastic differential equation

$$\mathrm{d}X_t^\varepsilon = a(X_t^\varepsilon, \theta_0)\,\mathrm{d}t + \varepsilon\,\mathrm{d}L_t\,, \quad X_0^\varepsilon = x_0 \in \mathbb{R}^d, \tag{2.1}$$

where Θ_0 is a smooth bounded open convex set in $\mathbb{R}^p$ with $p \in \mathbb{N}$, Θ denotes the closure of Θ_0, $\theta \in \Theta$, $\varepsilon > 0$, a is a function from $\mathbb{R}^d \times \Theta$ to $\mathbb{R}^d$, and $L = (L_t)_{t \le 0}$ is a d-dimensional Lévy process given by

$$L_t = \sigma B_t + \int_0^t \int_{|z|\le 1} z\,\tilde{N}(\mathrm{d}s\,, \mathrm{d}z) + \int_0^t \int_{|z|>1} z\,N(\mathrm{d}s\,, \mathrm{d}z)$$

with a $d \times r$ real-valued matrix σ, an r-dimensional standard Brownian motion B_t, an independent Poisson random measure $N(\mathrm{d}s\,, \mathrm{d}z)$ with characteristic measure $\mathrm{d}t\;\nu(\mathrm{d}z)$, and a martingale measure $\tilde{N}(\mathrm{d}s\,, \mathrm{d}z) = N(\mathrm{d}s\,, \mathrm{d}z) - \nu(\mathrm{d}z)\,\mathrm{d}s$. Here, we assume that $\nu(\mathrm{d}z)$ is a Lévy measure on $\mathbb{R}^d \setminus \{0\}$ and $\int_{|z|>0} |z|\,\nu(\mathrm{d}z) < \infty$. Suppose that we have discrete data $X_{t_0}^\varepsilon, \ldots, X_{t_n}^\varepsilon$ from (2.1) under $\theta = \theta_0 \in \Theta_0$ with $X_t^\varepsilon := X_t^{\varepsilon,\theta_0}$, and that $0 = t_0 < \cdots < t_n = 1$ and $t_i - t_{i-1} = 1/n$. We consider the problem of estimating the true $\theta_0 \in \Theta_0$ under $n \to \infty$ and $\varepsilon \to 0$ at the same time. We also define x_t as the solution of the corresponding deterministic differential equation

$$\frac{\mathrm{d}x_t}{\mathrm{d}t} = a(x_t, \theta_0) \tag{2.2}$$

with the initial condition x_0.

2.1 Assumtions and notations

Notation 2.1.1. *The following notations will be needed throughout this chapter:*

- $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$, $B_M \subset \mathbb{R}^d$ *is a closed ball centered at the origin with radius* $M > 0$.
- $C^{\infty,0}(\mathbb{R}^d \times \Theta; \mathbb{R}^d) := \left\{ f : \mathbb{R}^d \times \Theta \to \mathbb{R}^d \,\middle|\, \begin{array}{l} f \text{ is smooth with respet to } x \in \mathbb{R}^d, \text{ and for} \\ \text{all } k \in \mathbb{N}, \text{ the } k\text{-th derivatives of } f \text{ with} \\ \text{respect to } x \in \mathbb{R}^d \text{ are continuous on } \mathbb{R}^d \times \Theta \end{array} \right\}$.
- $D_x^\alpha := \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \dots \partial x_d^{\alpha_d}}$ *with* $\alpha \in \mathbb{N}_0^d$, $|\alpha| = \alpha_1 + \cdots + \alpha_d$.
- $\|f\|_{C^{\infty,0}(B_M \times \Theta)} := \sup_{\alpha \in \mathbb{N}_0^d} \|D_x^\alpha f\|_{C(B_M \times \Theta)} = \sup_{\alpha \in \mathbb{N}_0^d} \sup_{(x,\theta) \in B_M \times \Theta} |D_x^\alpha f(x, \theta)|$,

  $\|f(\theta_0)\|_{C^\infty(B_M)} := \sup_{\alpha \in \mathbb{N}_0^d} \|D_x^\alpha f(\cdot, \theta_0)\|_{C(B_M)} = \sup_{\alpha \in \mathbb{N}_0^d} \sup_{x \in B_M} |D_x^\alpha f(x, \theta_0)|$,

  where $f \in C^{\infty,0}(\mathbb{R}^d \times \Theta; \mathbb{R}^d)$, $M > 0$.
- $⨍_{t_{k-1}}^{t_k} f(t)\,\mathrm{d}t$ *denotes the average integral* $\frac{1}{|t_k - t_{k-1}|} \int_{t_{k-1}}^{t_k} f(t)\,\mathrm{d}t$.
- $Y_t^{n,\varepsilon} := X_{\lceil nt \rceil/n}^\varepsilon$ *for* $t \in (-1/n, 1]$, *where* $\lceil \cdot \rceil$ *is the ceiling function.*
- $\|\sigma\|_F^2 := \mathrm{tr}(\sigma^T \sigma) = \sum_{ij} \sigma_{ij}^2$, *where* $\sigma = (\sigma_{ij})$ *is a* $d \times r$ *matrix.*

Assumption 2.1.1. *We will make the following assumptions:*

(A1) *The family* $\{a(\cdot, \theta)\}_{\theta \in \Theta}$ *is equi-Lipschitz continuous, i.e., there is a postive constant* C *called a common Lipschitz constant such that*

$$|a(x, \theta) - a(y, \theta)| \le C|x - y| \quad (x, y \in \mathbb{R}^d,\ \theta \in \Theta).$$

(A2) *The function* a *belongs to* $C^{\infty,0}(\mathbb{R}^d \times \Theta; \mathbb{R}^d)$, *and* $\|b\|_{C^{\infty,0}(B_M \times \Theta)} < \infty$ *for all* $M > 0$.

(A3) *The function* a *is differentiable with respect to* $\theta \in \Theta_0$, *and the families* $\{\frac{\partial b}{\partial \theta_j}(\cdot, \theta)\}_{\theta \in \Theta_0}$ $(j = 1, \dots, p)$ *are equi-Lipschitz continuous.*

(A4) *If* $\theta \neq \theta_0$, *then* $a(x_t, \theta) \neq a(x_t, \theta_0)$ *for some* $t \in [0, 1]$.

(A5) *For any* $\theta \in \Theta$, *a* $p \times p$ *symmetric matrix* $I(\theta)$ *with the following* (i, j)*-th entry is positive definite:*

$$I^{ij}(\theta) := \int_0^1 \frac{\partial a}{\partial \theta_i}(x_t, \theta) \cdot \frac{\partial a}{\partial \theta_j}(x_t, \theta)\,\mathrm{d}t\,.$$

2.2 Main result

We define the *Adams-Moulton type* contrast function $\Psi_{n,\varepsilon,\ell}(\theta)$ as

$$\Psi_{n,\varepsilon,\ell}(\theta) := \sum_{k=\ell \vee 1}^{n} \frac{\left| X_{t_k}^\varepsilon - X_{t_{k-1}}^\varepsilon - \frac{1}{n} A_\ell a(\boldsymbol{X}_{t_k:t_{k-\ell}}^\varepsilon, \theta) \right|^2}{\varepsilon^2/n},$$

where $\boldsymbol{X}^{\varepsilon}_{t_k:t_{k-\ell}} := (X^{\varepsilon}_{t_k}, \ldots, X^{\varepsilon}_{t_{k-\ell}})$ and A_ℓ is the operator from $C^{\ell+1}(\mathbb{R}^d; \mathbb{R}^d)$ to $C^{\ell+1}(\mathbb{R}^{d\times\ell}; \mathbb{R}^d)$ of the form

$$A_\ell f(\boldsymbol{x}) := \sum_{\nu=0}^{\ell} \beta_{\ell\nu} f(x_\nu) \qquad \text{for } \boldsymbol{x} = (x_0, \ldots, x_\ell) \in \mathbb{R}^{d\times\ell},\ \theta \in \Theta,$$

in particular,

$$A_\ell a(\boldsymbol{X}^{\varepsilon}_{t_k:t_{k-\ell}}, \theta) = \sum_{\nu=0}^{\ell} \beta_{\ell\nu} a(X^{\varepsilon}_{t_{k-\nu}}, \theta). \tag{2.3}$$

For simplicity of discussion, it is useful to use the following form for the contrast function

$$\Phi_{n,\varepsilon,\ell}(\theta) := \varepsilon^2(\Psi_{n,\varepsilon,\ell}(\theta) - \Psi_{n,\varepsilon,\ell}(\theta_0)).$$

Then the LSE is given by

$$\hat{\theta}_{n,\varepsilon,\ell} := \operatorname*{argmin}_{\theta\in\Theta} \Psi_{n,\varepsilon,\ell}(\theta) = \operatorname*{argmin}_{\theta\in\Theta} \Phi_{n,\varepsilon,\ell}(\theta). \tag{2.4}$$

Similarly, we denote by $\tilde{\Psi}_{n,\varepsilon,\ell}$ the *Adams-Bashforth type* contrast function

$$\tilde{\Psi}_{n,\varepsilon,\ell}(\theta) := \sum_{k=\ell}^{n} \frac{\left| X^{\varepsilon}_{t_k} - X^{\varepsilon}_{t_{k-1}} - \frac{1}{n}\tilde{A}_\ell a(\boldsymbol{X}^{\varepsilon}_{t_{k-1}:t_{k-\ell}}, \theta) \right|^2}{\varepsilon^2/n},$$

where

$$\tilde{A}_\ell a(\boldsymbol{X}^{\varepsilon}_{t_{k-1}:t_{k-\ell}}, \theta) = \sum_{\nu=1}^{\ell} \gamma_{\ell\nu} a(X^{\varepsilon}_{t_{k-\nu}}, \theta). \tag{2.5}$$

Then the LSE $\tilde{\theta}_{n,\varepsilon,\ell}$ is given by

$$\tilde{\theta}_{n,\varepsilon,\ell} := \operatorname*{argmin}_{\theta\in\Theta} \tilde{\Psi}_{n,\varepsilon,\ell}(\theta). \tag{2.6}$$

We call $\hat{\theta}_{n,\varepsilon,\ell}$ and $\tilde{\theta}_{n,\varepsilon,\ell}$ the *Adams-Moulton type* LSE and the *Adams-Bashforth type* LSE, respectively.

Theorem 2.2.1 (Consistency). *Let $\ell = \ell_n \in \mathbb{N}$ depend on n. Under conditions* (A1)-(A4), *the least squares estimator $\hat{\theta}_{n,\varepsilon,\ell}$ given in* (2.4) *is consistent to θ_0, i.e., if $2^\ell\varepsilon \to 0$ and $\ell 2^{2\ell}/n$ is bounded as $n \to \infty$ and $\varepsilon \to 0$, then*

$$\hat{\theta}_{n,\varepsilon,\ell} \xrightarrow{P_{\theta_0}} \theta_0$$

as $n \to \infty$ and $\varepsilon \to 0$.

Theorem 2.2.2 (Asymptotic distribution). *Let $\ell = \ell_n \in \mathbb{N}$ depend on n. Under conditions* (A1)-(A5), *if $\ell 2^{4\ell}/n \to 0$, $2^\ell\varepsilon \to 0$ and $\ell 2^{2\ell}/n\varepsilon \to 0$ as $n \to \infty$ and $\varepsilon \to 0$, then*

$$\varepsilon^{-1}\left(\hat{\theta}_{n,\varepsilon,\ell} - \theta_0\right) \xrightarrow{P_{\theta_0}} I(\theta_0)^{-1} S(\theta_0)$$

as $n \to \infty$ and $\varepsilon \to 0$, where $I(\theta)$ is the $p \times p$ matrix given in the assumption (A5)*, and $S(\theta)$ is a p-dimensional vector with the i-th entry*

$$S_i(\theta) := \int_0^1 \frac{\partial a}{\partial \theta_i}(x_t, \theta) \cdot \mathrm{d}L_t$$

for $\theta \in \Theta$, respectively.

Remark 2.2.1. *Recall that there is another LSE $\tilde{\theta}_{n,\varepsilon,\ell}$ given by* (2.6) *based on Adams-Bashforth method (the LSE $\hat{\theta}_{n,\varepsilon,\ell}$ given by* (2.4) *is based on Adams-Moulton method). The consistency of $\tilde{\theta}_{n,\varepsilon,\ell}$ also holds if $\ell 2^{\ell}\varepsilon \to 0$ and $\ell^3 2^{2\ell}/n$ is bounded as $n \to \infty$ and $\varepsilon \to 0$. In Theorem 2.2.2, the corresponding convergence for $\tilde{\theta}_{n,\varepsilon,\ell}$ holds if $\ell^5 2^{4\ell}/n \to 0$, $\ell 2^{\ell}\varepsilon \to 0$ and $\ell^3 2^{2\ell}/n\varepsilon \to 0$ as $n \to \infty$ and $\varepsilon \to 0$.*

Remark 2.2.2. *We shall see in Section 2.4 that the performances for numerical experiments become better when ℓ depends on n. It is natural to ask what assumption on (n, ε, ℓ) we need to assume for consistency and asymptotic normality of LSEs. Theorem 2.2.1 and 2.2.2 answewer the question, and in particular the assumption for (n, ε, ℓ) is simply written as $n\varepsilon \to \infty$ when ℓ is bounded. This simple assumption is same as in various research so far (see,* e.g.*, Theorem 2.2 in Long et al. [22], the assumption (B1) in Sørensen and Uchida [33], etc.) If someone wants to consider ℓ unbounded, one should carefully take ℓ go infinity since the larger ℓ we take, the exponentailly larger data we need.*

2.3 Proof

In this section, we begin with setting up some notations and assumptions, and give some useful inequalities and several convergence theorems.

Proposition 2.3.1. *Suppose the assumption* (A1)*.*

(i) *It holds that*

$$\sup_{\substack{\nu=0,\dots,\ell \\ t \in (t_{(\ell-1)\vee 0}, 1]}} \left| Y^{n,\varepsilon}_{t-t_\nu} - x_t \right| \leq C \left(\varepsilon \sup_{s \in [0,1]} |L_s| + \frac{\ell+1}{n} \right),$$

where C is a positive constant, and $Y^{n,\varepsilon}_t := X^{\varepsilon}_{\lceil nt \rceil/n}$ with the ceiling function $\lceil \cdot \rceil$.

(ii) *Let $\ell = \ell_n \in \mathbb{N}$ depend on n. If $\ell/n = O(1)$ as $n, \ell \to \infty$, then*

$$\sup_{\substack{0<\varepsilon<1 \\ n \in \mathbb{N}}} \sup_{\substack{\nu=0,\dots,\ell \\ t \in (t_{(\ell-1)\vee 0}, 1]}} \left| Y^{n,\varepsilon}_{t-t_\nu} \right| < \infty \quad a.s.,$$

and

$$\tau^{n,\varepsilon,\ell}_m := \inf \left\{ t > 0 \,\middle|\, |x_t| \geq m, \min_{\nu=0,\dots,\ell} \left| Y^{n,\varepsilon}_{t-t_\nu} \right| \geq m \right\} \xrightarrow{a.s.} \infty$$

as $m \to \infty$, uniformly in n, $0 < \varepsilon < 1$ and $\ell \in \mathbb{N}_0$.

Proof. It follows by Gronwall's inequality that

$$\sup_{t\in[0,1]} |X_t^\varepsilon - x_t| \le e^C \varepsilon \sup_{t\in[0,1]} |L_t|,$$

where C is the common Lipschitz constant from (A1). Since $|\lceil n(t-t_\nu)\rceil/n - t| \le \frac{\ell+1}{n}$ for all $t \in (t_{(\ell-1)\vee 0}, 1]$, we have

$$\left|Y_{t-t_\nu}^{n,\varepsilon} - x_t\right| \le e^C\varepsilon \sup_{s\in[0,1]} |L_s| + \sup_{\substack{|s-u|\le(\ell+1)/n \\ s,u\in[0,1]}} |x_s - x_u|$$

for all $t \in \left(t_{(\ell-1)\vee 0}, 1\right]$. This implies (i). Moreover, (ii) is immediate from the inequality

$$\left|Y_{t-t_\nu}^{n,\varepsilon}\right| \le \sup_{s\in[0,1]} |x_s| + e^C\varepsilon \sup_{s\in[0,1]} |L_s| + \sup_{\substack{|s-u|\le(\ell+1)/n \\ s,u\in[0,1]}} |x_s - x_u|$$

for all $t \in \left(t_{(\ell-1)\vee 0}, 1\right]$. □

2.3.1 Inequalities for deterministic convergence

We prepare the following inequalities for the solution of (2.2).

Lemma 2.3.2. *Let f be a function in $C^{\infty,0}(\mathbb{R}^d \times \Theta; \mathbb{R}^d)$ such that $\|f\|_{C^{\infty,0}(B_M\times\Theta)} < \infty$ for all $M > 0$, and suppose the assumption* (A2)*. Then,*

$$\sup_{t\in[0,1]} \left| \frac{\mathrm{d}^\ell}{\mathrm{d}t^\ell}\big(f(x_t,\theta)\big)\right| \le \ell! d^\ell \|a(\theta_0)\|_{C^\infty(B_M)}^\ell \|f\|_{C^{\infty,0}(B_M\times\Theta)}$$

for all $\ell \in \mathbb{N}$.

Proof. It is shown by induction that

$$\frac{\mathrm{d}^\ell}{\mathrm{d}t^\ell}\big(f(x_t,\theta)\big) = \sum_{j_1=1}^d \cdots \sum_{j_\ell=1}^d \sum_{|\alpha|+|\nu|=\ell} c_{\alpha,\nu} (D_x^{\alpha_1} b_{j_1} \cdots D_x^{\alpha_\ell} b_{j_\ell} D_x^\nu f_\theta)_{x=x_t}, \qquad \sum_{|\alpha|+|\nu|=\ell} c_{\alpha,\nu} = \ell!,$$

where $\alpha = (\alpha_1, \dots, \alpha_\ell)$ for $\alpha_j \in \mathbb{N}_0^d$, $\nu \in \mathbb{N}_0^d$, $c_{\alpha,\nu} \in \mathbb{N}_0$. We write $b_i(x,\theta_0)$ and $f(x,\theta)$ simply as b_i and f_θ, respectively. Indeed, the derivative of each term with respect to t is

$$c_{\alpha,\nu} \sum_{j_{\ell+1}=1}^d \left(b_{j_{\ell+1}} D_x^{e_{j_{\ell+1}}} (D_x^{\alpha_1} b_{j_1} \cdots D_x^{\alpha_\ell} b_{j_\ell} D_x^\nu f_\theta)\right)_{x=x_t},$$

where e_j denotes d-dimensional multi-index with entry 1 at the jth coordinate, and entry zero elsewhere. □

Lemma 2.3.3. *Let $\ell \in \mathbb{N}$. Let f be a function as in Lemma 2.3.2. Under the assumption* (A2)*, it follows that*

$$\left| A_\ell f(\boldsymbol{x}_{t_k:t_{k-\ell}}, \theta) - \fint_{t_{k-1}}^{t_k} f(x_s,\theta)\,\mathrm{d}s \right| \le \ell! n^{-(\ell+1)} d^{\ell+1} \|a(\theta_0)\|_{C^\infty(B_M)}^{\ell+1} \|f\|_{C^{\infty,0}(B_M\times\Theta)},$$

where $k = \ell \vee 1, \dots, n$, and $M = \sup_{t\in[0,1]} |x_t|$.

Remark 2.3.1. *When we employ $\tilde{A}_\ell f(\boldsymbol{x}_{t_{k-1}:t_{k-\ell}}, \theta)$ given by* (2.5) *with* (1.2) *as the version of the Adams-Bashforth method instead of $A_\ell f(\boldsymbol{x}_{t_k:t_{k-\ell}}, \theta)$, we obtain the following inequality:*

$$\left| \tilde{A}_\ell f(\boldsymbol{x}_{t_{k-1}:t_{k-\ell}}, \theta) - \int_{t_{k-1}}^{t_k} f(x_s, \theta)\, \mathrm{d}s \right| \le \ell! n^{-\ell} d^\ell \|a(\theta_0)\|_{C^\infty(B_M)}^\ell \|f\|_{C^{\infty,0}(B_M \times \Theta)}.$$

Proof. It follows from (1.4) and (2.3) that

$$\left| A_\ell f(\boldsymbol{x}_{t_k:t_{k-\ell}}, \theta) - \int_{t_{k-1}}^{t_k} f(x_s, \theta)\, \mathrm{d}s \right| = \left| \int_{t_{k-1}}^{t_k} \left(P(s; f(x_\cdot, \theta), t_k, \ldots, t_{k-\ell}) - f(x_s, \theta) \right) \mathrm{d}s \right|,$$

where $k = \ell \vee 1, \ldots, n$, and $s \mapsto P(s; f(x_\cdot, \theta), t_k, \ldots, t_{k-\ell})$ is the Lagrange interpolating polynomial through the points $(s, f(x_s, \theta))$, $s = t_k, \ldots, t_{k-\ell}$. It holds from Theorem 3.1.1 in Davis [8] that for each $s \in [t_{k-1}, t_k]$ there exists $\xi_s \in (t_{k-\ell}, t_k)$ such that

$$P(s; f(x_\cdot, \theta), t_k, \ldots, t_{k-\ell}) - f(x_s, \theta) = \frac{1}{(\ell+1)!} \left(\frac{\mathrm{d}^{\ell+1}}{\mathrm{d}t^{\ell+1}} (f(x_t, \theta)) \right)_{t=\xi_s} \prod_{\nu=0}^{\ell} (s - t_{k-\nu}),$$

and that

$$\begin{aligned} \int_{t_{k-1}}^{t_k} \left| \left(\frac{\mathrm{d}^{\ell+1}}{\mathrm{d}t^{\ell+1}} (f(x_t, \theta)) \right)_{t=\xi_s} \prod_{\nu=0}^{\ell} (s - t_{k-\nu}) \right| \mathrm{d}s &\le \sup_{t \in [t_{k-\ell}, t_k]} \left| \frac{\mathrm{d}^{\ell+1}}{\mathrm{d}t^{\ell+1}} (f(x_t, \theta)) \right| \int_{t_{k-1}}^{t_k} \prod_{\nu=0}^{\ell} |s - t_{k-\nu}|\, \mathrm{d}s \\ &\le \frac{\ell!}{n^{\ell+1}} \sup_{t \in [0,1]} \left| \frac{\mathrm{d}^{\ell+1}}{\mathrm{d}t^{\ell+1}} (f(x_t, \theta)) \right|. \end{aligned}$$

This yields the consequence. □

2.3.2 Convergence theorems

Proposition 2.3.4. *Let $\ell = \ell_n \in \mathbb{N}$ depend on n, and let f be a function as in Lemma 2.3.2. Suppose the assumptions* (A1) *and* (A2)*, and that the family $\{f(\cdot, \theta)\}_{\theta \in \Theta}$ is equi-Lipschitz continuous. If $\ell/n \to 0$ and $2^\ell \varepsilon \to 0$ as $n \to \infty$. Then, for all $q \ge 1$*

$$\frac{1}{n} \sum_{k=\ell \vee 1}^{n} \left| A_\ell f(\boldsymbol{X}_{t_k:t_{k-\ell}}^\varepsilon, \theta) \right|^q \xrightarrow{a.s.} \int_0^1 |f(x_t, \theta)|^q\, \mathrm{d}t$$

as $n \to \infty$ and $\varepsilon \to 0$, uniformly in $\theta \in \Theta$.

Proof. We use the triangle inequality to obtain that

$$\begin{aligned}
&\left| \left(\frac{1}{n} \sum_{k=\ell\vee 1}^{n} \left| A_\ell f(\boldsymbol{X}^\varepsilon_{t_k:t_{k-\ell}}, \theta) \right|^q \right)^{1/q} - \left(\int_0^1 |f(x_t,\theta)|^q \,\mathrm{d}t \right)^{1/q} \right| \\
&\qquad \le \left(\frac{1}{n} \sum_{k=\ell\vee 1}^{n} \left| A_\ell f(\boldsymbol{X}^\varepsilon_{t_k:t_{k-\ell}}, \theta) - A_\ell f(\boldsymbol{x}_{t_k:t_{k-\ell}}, \theta) \right|^q \right)^{1/q} \\
&\qquad + \left(\frac{1}{n} \sum_{k=\ell\vee 1}^{n} \left| A_\ell f(\boldsymbol{x}_{t_k:t_{k-\ell}}, \theta) - \fint_{t_{k-1}}^{t_k} f(x_t,\theta)\,\mathrm{d}t \right|^q \right)^{1/q} \\
&\qquad\quad + \left| \left(\frac{1}{n} \sum_{k=\ell\vee 1}^{n} \left| \fint_{t_{k-1}}^{t_k} f(x_t,\theta)\,\mathrm{d}t \right|^q \right)^{1/q} - \left(\int_0^1 |f(x_t,\theta)|^q \,\mathrm{d}t \right)^{1/q} \right|.
\end{aligned}$$

The second and the third term in the right-hand side converge to zero as $n \to \infty$ and $\ell/n \to 0$, uniformly in $\theta \in \Theta$, by Lemma 2.3.3 and Lemma 2.5.2. From Lemma 2.5.1, the first term is estimated from above by

$$\left(\sum_{\nu=0}^{\ell} |\beta_{\ell\nu}| \right) \sup_{s\in[0,1]} |f(X^\varepsilon_s,\theta) - f(x_s,\theta)| \le C 2^\ell \sup_{s\in[0,1]} |X^\varepsilon_s - x_s| \,,$$

where C is the common Lipschitz constant for f. This converges almost surely to zero as $2^\ell \varepsilon \to 0$, uniformly in $\theta \in \Theta$, as we saw in the proof of Proposition 2.3.1. □

Remark 2.3.2. *If we employ* $\tilde{A}_\ell f(\boldsymbol{X}^\varepsilon_{t_{k-1}:t_{k-\ell}}, \theta)$ *instead of* $A_\ell f(\boldsymbol{x}^\varepsilon_{t_k:t_{k-\ell}}, \theta)$*, the convergence in Proposition 2.3.4 holds under* $\ell 2^\ell \varepsilon \to 0$.

Remark 2.3.3. *It is easy to check that, for* f *and* g *satisfying the same assumptions as in Proposition 2.3.4,*

$$\frac{1}{n} \sum_{k=\ell\vee 1}^{n} A_\ell f(\boldsymbol{X}^\varepsilon_{t_k:t_{k-\ell}}, \theta) \cdot A_\ell g(\boldsymbol{X}^\varepsilon_{t_k:t_{k-\ell}}, \theta) \xrightarrow{a.s.} \int_0^1 f(x_t,\theta) \cdot g(x_t,\theta)\,\mathrm{d}t \,.$$

Indeed, we replace f *by* $f \pm g$ *in Proposition 2.3.4 with* $q = 2$*, and apply the relationship between inner product and norm:*

$$\left| \frac{a+b}{2} \right|^2 - \left| \frac{a-b}{2} \right|^2 = 2ab \quad (a, b \in \mathbb{R}).$$

This convergence will appear in the proof of Proposition 2.3.8 (ii).

Lemma 2.3.5. *Let* $\ell = \ell_n \in \mathbb{N}$ *depend on* n*, and let* f *be a function as in Proposition 2.3.4. Suppose the assumptions* (A1) *and* (A2)*, and that* f *is differentiable with respect to* $\theta \in \Theta_0$*, and the families* $\left\{ \frac{\partial f}{\partial \theta_j}(\cdot,\theta) \right\}_{\theta\in\Theta_0}$ $(j = 1, \ldots, p)$ *are equi-Lipschitz continuous. If* $\ell 2^\ell / n \to 0$ *and* $2^\ell \varepsilon \to 0$ *as* $n \to \infty$ *and* $\varepsilon \to 0$*, then it holds that*

$$\sum_{k=\ell\vee 1}^{n} A_\ell f(\boldsymbol{X}^\varepsilon_{t_k:t_{k-\ell}}, \theta) \cdot \left(L_{t_k} - L_{t_{k-1}} \right) \xrightarrow{P_{\theta_0}} \int_0^1 f(x_t,\theta) \cdot \mathrm{d}L_t$$

as $n \to \infty$ *and* $\varepsilon \to 0$*, uniformly in* $\theta \in \Theta$.

Proof. Since

$$\sum_{k=\ell\vee 1}^{n} A_\ell f(\boldsymbol{X}^\varepsilon_{t_k:t_{k-\ell}},\theta)\cdot\big(L_{t_k}-L_{t_{k-1}}\big)=\sum_{\nu=0}^{\ell}\beta_{\ell\nu}\int_{t_{(\ell-1)\vee 0}}^{1} f(Y^{n,\varepsilon}_{t-t_\nu},\theta)\cdot \mathrm{d}L_t$$

and $\sum_{\nu=0}^{\ell}\beta_{\ell\nu}=1$, we have

$$\begin{aligned}\sum_{k=\ell\vee 1}^{n} A_\ell f(\boldsymbol{X}^\varepsilon_{t_k:t_{k-\ell}},\theta)\cdot\big(L_{t_k}-L_{t_{k-1}}\big)-\int_0^1 f(x_t,\theta)\cdot\mathrm{d}L_t\\ =\sum_{\nu=0}^{\ell}\beta_{\ell\nu}\int_{t_{(\ell-1)\vee 0}}^{1}\big(f(Y^{n,\varepsilon}_{t-t_\nu},\theta)-f(x_t,\theta)\big)\cdot\mathrm{d}L_t-\int_0^{t_{(\ell-1)\vee 0}} f(x_t,\theta)\cdot\mathrm{d}L_t\,.\end{aligned}$$

The last term converges almost surely to zero as $n\to\infty$ and $\ell/n\to 0$, uniformly in $\theta\in\Theta$. Let us denote

$$\tilde L_t=\sigma B_t+\int_0^t\int_{|z|\le 1} z\,\tilde N(\mathrm{d}s\,,\mathrm{d}z),$$

then $L_t=at+\tilde L_t+\int_0^t\int_{|z|>1} z\,N(\mathrm{d}s\,,\mathrm{d}z)$. We have

$$\begin{aligned}&\left|\sum_{\nu=0}^{\ell}\beta_{\ell\nu}\int_{t_{(\ell-1)\vee 0}}^{1}\int_{|z|>1}\big(f(Y^{n,\varepsilon}_{t-t_\nu},\theta)-f(x_t,\theta)\big)\cdot z\,N(\mathrm{d}t\,,\mathrm{d}z)\right|\\ &\le\left(\sum_{\nu=0}^{\ell}|\beta_{\ell\nu}|\right)\sup_{\nu=0,\dots,\ell}\int_{t_{(\ell-1)\vee 0}}^{1}\int_{|z|>1}\big|f(Y^{n,\varepsilon}_{t-t_\nu},\theta)-f(x_t,\theta)\big|\,|z|\,N(\mathrm{d}t\,,\mathrm{d}z)\\ &\le C2^\ell\sup_{\substack{\nu=0,\dots,\ell\\ t\in[t_{(\ell-1)\vee 0},1]}}\big|Y^{n,\varepsilon}_{t-t_\nu}-x_t\big|\int_0^1\int_{|z|>1}|z|\,N(\mathrm{d}t\,,\mathrm{d}z),\end{aligned}$$

which converges almost surely to zero as $n\to\infty$, $\varepsilon\to 0$, $\ell 2^\ell/n\to 0$ and $2^\ell\varepsilon\to 0$, uniformly in $\theta\in\Theta$, by Proposition 2.3.1. Analogously, we obtain

$$\left|\sum_{\nu=0}^{\ell}\beta_{\ell\nu}\int_{t_{(\ell-1)\vee 0}}^{1}\big(f(Y^{n,\varepsilon}_{t-t_\nu},\theta)-f(x_t,\theta)\big)\cdot a\,\mathrm{d}t\right|\xrightarrow{a.s.}0$$

as $n\to\infty$, $\varepsilon\to 0$, $\ell 2^\ell/n\to 0$ and $2^\ell\varepsilon\to 0$, uniformly in $\theta\in\Theta$.

Analogous to the proof of Lemma 4 in Ogihara and Yoshida [25], it follows from Markov's inequality and Morrey's inequality (see, *e.g.*, Theorem 5 in Evans [9, Section 5.6]) that for any $q\in(p,\infty]$ and $\eta>0$

$$\begin{aligned}&P\left(\sup_{\theta\in\Theta}\left|\sum_{\nu=0}^{\ell}\beta_{\ell\nu}\int_{t_{(\ell-1)\vee 0}}^{1}\mathbf{1}_{\{t\le\tau_m^{n,\varepsilon,\ell}\}}\big(f(Y^{n,\varepsilon}_{t-t_\nu},\theta)-f(x_t,\theta)\big)\cdot\mathrm{d}\tilde L_t\right|>\eta\right)\\ &\le\frac{1}{\eta}E\left[\sup_{\theta\in\Theta}\left|\int_{t_{(\ell-1)\vee 0}}^{1}\mathbf{1}_{\{t\le\tau_m^{n,\varepsilon,\ell}\}}\sum_{\nu=0}^{\ell}\beta_{\ell\nu}\big(f(Y^{n,\varepsilon}_{t-t_\nu},\theta)-f(x_t,\theta)\big)\cdot\mathrm{d}\tilde L_t\right|\right]\\ &\le\frac{C}{\eta}E\left[\left\|\int_{t_{(\ell-1)\vee 0}}^{1}\mathbf{1}_{\{t\le\tau_m^{n,\varepsilon,\ell}\}}\sum_{\nu=0}^{\ell}\beta_{\ell\nu}\big(f(Y^{n,\varepsilon}_{t-t_\nu},\cdot\,)-f(x_t,\cdot\,)\big)\cdot\mathrm{d}\tilde L_t\right\|_{W^{1,q}(\Theta)}\right],\end{aligned}\tag{2.7}$$

where C is a constant depending only on p, q and Θ. It follows from Hölder's inequality and Fubini's theorem that

$$
\begin{aligned}
&P\left(\sup_{\theta\in\Theta}\left|\sum_{\nu=0}^{\ell}\beta_{\ell\nu}\int_{t_{(\ell-1)\vee 0}}^{1}\mathbf{1}_{\{t\le\tau_m^{n,\varepsilon,\ell}\}}\left(f(Y_{t-t_\nu}^{n,\varepsilon},\theta)-f(x_t,\theta)\right)\cdot \mathrm{d}\tilde{L}_t\right|>\eta\right)\\
&\quad\le \frac{C}{\eta}\left(\int_\Theta E\left[\left|\int_{t_{(\ell-1)\vee 0}}^{1}\mathbf{1}_{\{t\le\tau_m^{n,\varepsilon,\ell}\}}\sum_{\nu=0}^{\ell}\beta_{\ell\nu}\left(f(Y_{t-t_\nu}^{n,\varepsilon},\theta)-f(x_t,\theta)\right)\cdot \mathrm{d}\tilde{L}_t\right|^q\right]\mathrm{d}\theta\right)^{1/q}\\
&\quad+\frac{C}{\eta}\left(\int_\Theta E\left[\left|\int_{t_{(\ell-1)\vee 0}}^{1}1_{\{t\le\tau_m^{n,\varepsilon,\ell}\}}\sum_{\nu=0}^{\ell}\beta_{\ell\nu}\left(\frac{\partial f}{\partial\theta_j}(Y_{t-t_\nu}^{n,\varepsilon},\theta)-\frac{\partial f}{\partial\theta_j}(x_t,\theta)\right)\cdot \mathrm{d}\tilde{L}_t\right|^q\right]\mathrm{d}\theta\right)^{1/q}
\end{aligned}
$$

for $j=1,\dots,p$. By the moment inequality for stochastic integrals (see, *e.g.*, Theorem 7.1 in Chapter 1 in Mao [24]), for $q\ge 2$ we obtain

$$
\begin{aligned}
&\int_\Theta E\left[\left|\int_{t_{(\ell-1)\vee 0}}^{1}\mathbf{1}_{\{t\le\tau_m^{n,\varepsilon,\ell}\}}\sum_{\nu=0}^{\ell}\beta_{\ell\nu}\left(f(Y_{t-t_\nu}^{n,\varepsilon},\theta)-f(x_t,\theta)\right)\cdot \mathrm{d}B_t\right|^q\right]\mathrm{d}\theta\\
&\quad\le\left(\frac{q(q-1)}{2}\right)^{q/2}\int_\Theta E\left[\int_{t_{(\ell-1)\vee 0}}^{1}\mathbf{1}_{\{t\le\tau_m^{n,\varepsilon,\ell}\}}\left|\sum_{\nu=0}^{\ell}\beta_{\ell\nu}\left(f(Y_{t-t_\nu}^{n,\varepsilon},\theta)-f(x_t,\theta)\right)\right|^q\mathrm{d}t\right]\mathrm{d}\theta\\
&\quad\le\left(\frac{q(q-1)}{2}\right)^{q/2}C^q\int_\Theta E\left[\int_{t_{(\ell-1)\vee 0}}^{1}\mathbf{1}_{\{t\le\tau_m^{n,\varepsilon,\ell}\}}2^{q\ell}\sup_{\nu=0,\dots,\ell}\left|Y_{t-t_\nu}^{n,\varepsilon}-x_t\right|^q\mathrm{d}t\right]\mathrm{d}\theta\,,
\end{aligned}
$$

and by Kunita's inequality (see, *e.g.*, Theorem 4.4.23 in Applebaum [3]), for $q\ge 2$, there exists $D(q)>0$ such that

$$
\begin{aligned}
&\int_\Theta E\left[\left|\int_{t_{(\ell-1)\vee 0}}^{1}\int_{0<|z|\le 1}\mathbf{1}_{\{t\le\tau_m^{n,\varepsilon,\ell}\}}\sum_{\nu=0}^{\ell}\beta_{\ell\nu}\left(f(Y_{t-t_\nu}^{n,\varepsilon},\theta)-f(x_t,\theta)\right)\cdot z\,\tilde{N}(\mathrm{d}s\,,\mathrm{d}z)\right|^q\right]\mathrm{d}\theta\\
&\le D(q)\int_\Theta\left\{E\left[\left(\int_{t_{(\ell-1)\vee 0}}^{1}\int_{0<|z|\le 1}\mathbf{1}_{\{t\le\tau_m^{n,\varepsilon,\ell}\}}\left|\sum_{\nu=0}^{\ell}\beta_{\ell\nu}\left(f(Y_{t-t_\nu}^{n,\varepsilon},\theta)-f(x_t,\theta)\right)\cdot z\right|^2\nu(\mathrm{d}z)\,\mathrm{d}t\right)^{q/2}\right]\right.\\
&\quad\left.+E\left[\int_{t_{(\ell-1)\vee 0}}^{1}\int_{0<|z|\le 1}\mathbf{1}_{\{t\le\tau_m^{n,\varepsilon,\ell}\}}\left|\sum_{\nu=0}^{\ell}\beta_{\ell\nu}\left(f(Y_{t-t_\nu}^{n,\varepsilon},\theta)-f(x_t,\theta)\right)\cdot z\right|^q\nu(\mathrm{d}z)\,\mathrm{d}t\right]\right\}\mathrm{d}\theta\\
&\le D(q)C^q\left\{\left(\int_{0<|z|\le 1}|z|^2\,\nu(\mathrm{d}z)\right)^{q/2}E\left[\left(\int_{t_{(\ell-1)\vee 0}}^{1}\mathbf{1}_{\{t\le\tau_m^{n,\varepsilon,\ell}\}}\sup_{\nu=0,\dots,\ell}2^{2\ell}\left|Y_{t-t_\nu}^{n,\varepsilon}-x_t\right|^2\mathrm{d}t\right)^{q/2}\right]\right.\\
&\quad\left.+\left(\int_{0<|z|\le 1}|z|^q\,\nu(\mathrm{d}z)\right)E\left[\int_{t_{(\ell-1)\vee 0}}^{1}\mathbf{1}_{\{t\le\tau_m^{n,\varepsilon,\ell}\}}\sup_{\nu=0,\dots,\ell}2^{q\ell}\left|Y_{t-t_\nu}^{n,\varepsilon}-x_t\right|^q\mathrm{d}t\right]\right\}.
\end{aligned}
$$

C is the constant in (2.7). Both converge to zero as $\ell 2^\ell/n\to 0$ and $2^\ell\varepsilon\to 0$, by dominated convergence theorem, and so does (2.7). □

Proposition 2.3.6. *Let $\ell = \ell_n \in \mathbb{N}$ depend on n, and let f be a function as in Lemma 2.3.5. Under the assumptions* (A1) *and* (A2)*, if $\ell 2^{4\ell}/n \to 0$, $2^\ell \varepsilon \to 0$ and $\ell 2^{2\ell}/n\varepsilon \to 0$ as $n \to \infty$ and $\varepsilon \to 0$, then it holds that*

$$\frac{1}{\varepsilon}\sum_{k=\ell\vee 1}^{n} A_\ell f(\boldsymbol{X}^\varepsilon_{t_k:t_{k-\ell}},\theta)\cdot\left(X^\varepsilon_{t_k} - X^\varepsilon_{t_{k-1}} - \frac{1}{n}A_\ell a(\boldsymbol{X}^\varepsilon_{t_k:t_{k-\ell}},\theta_0)\right) \xrightarrow{P_{\theta_0}} \int_0^1 f(x_t,\theta)\cdot \mathrm{d}L_t$$

as $n \to \infty$ and $\varepsilon \to 0$, uniformly in $\theta \in \Theta$.

Remark 2.3.4. *This lemma will be essentially used for the case $\theta = \theta_0$.*

Proof. It follows that

$$\begin{aligned}
&\frac{1}{\varepsilon}\sum_{k=\ell\vee 1}^{n} A_\ell f(\boldsymbol{X}^\varepsilon_{t_k:t_{k-\ell}},\theta)\cdot\left(X^\varepsilon_{t_k} - X^\varepsilon_{t_{k-1}} - \frac{1}{n}A_\ell a(\boldsymbol{X}^\varepsilon_{t_k:t_{k-\ell}},\theta_0)\right)\\
&\quad= \frac{1}{\varepsilon}\sum_{k=\ell\vee 1}^{n} A_\ell f(\boldsymbol{X}^\varepsilon_{t_k:t_{k-\ell}},\theta)\cdot\left(\int_{t_{k-1}}^{t_k} a(X^\varepsilon_t,\theta_0)\,\mathrm{d}t - \frac{1}{n}A_\ell a(\boldsymbol{X}^\varepsilon_{t_k:t_{k-\ell}},\theta_0)\right)\\
&\qquad+ \sum_{k=\ell\vee 1}^{n} A_\ell f(\boldsymbol{X}^\varepsilon_{t_k:t_{k-\ell}},\theta)\cdot\left(L_{t_k} - L_{t_{k-1}}\right)\\
&\quad=: J_1 + J_2.
\end{aligned}$$

From Lemma 2.3.5, J_2 converges to $\int_0^1 f(x_t,\theta)\cdot\mathrm{d}L_t$ in P_{θ_0} as $n \to \infty$, $\varepsilon \to 0$, $\ell 2^\ell/n \to 0$ and $2^\ell\varepsilon \to 0$, uniformly in $\theta \in \Theta$, and

$$\begin{aligned}
|J_1| &= \frac{1}{\varepsilon}\left|\sum_{\nu,\mu=0}^{\ell}\beta_{\ell\nu}\beta_{\ell\mu}\sum_{k=\ell\vee 1}^{n}\int_{t_{k-1}}^{t_k} f(X^\varepsilon_{t_{k-\mu}},\theta)\cdot\left(a(X^\varepsilon_t,\theta_0) - a(X^\varepsilon_{t_{k-\nu}},\theta_0)\right)\mathrm{d}t\right|\\
&\le \left(\sum_{\nu=0}^{\ell}|\beta_{\ell\nu}|\right)^2 \frac{1}{n\varepsilon}\sum_{k=\ell\vee 1}^{n}\sup_{\substack{\nu,\mu=0,\dots,\ell\\ t\in[t_{k-1},t_k]}}\left|f(X^\varepsilon_{t_{k-\mu}},\theta)\cdot\left(a(X^\varepsilon_t,\theta_0) - a(X^\varepsilon_{t_{k-\nu}},\theta_0)\right)\right|\\
&\qquad\qquad\le \frac{C2^{2\ell}}{n\varepsilon}\sum_{k=\ell\vee 1}^{n}\sup_{\substack{\nu,\mu=0,\dots,\ell\\ t\in[t_{k-1},t_k]}}\left|f(X^\varepsilon_{t_{k-\mu}},\theta)\right|\left|X^\varepsilon_t - X^\varepsilon_{t_{k-\nu}}\right|,
\end{aligned}$$

where C is a Lipschitz constant in (A1). For $t \in [t_{(\ell-1)\vee 0}, 1)$,

$$\begin{aligned}
\left|X^\varepsilon_t - X^\varepsilon_{t_{k-\nu}}\right| &= \left|\int_{t_{k-\nu}}^{t} a(X^\varepsilon_s,\theta_0)\,\mathrm{d}s + \varepsilon(L_t - L_{t_{k-\nu}})\right|\\
&\le C\int_{t_{k-\nu}}^{t}\left|X^\varepsilon_s - X^\varepsilon_{t_{k-\nu}}\right|\mathrm{d}s + \frac{\ell}{n}\left|a(X^\varepsilon_{t_{k-\nu}},\theta_0)\right| + \varepsilon\sup_{s\in[t_{k-1},t_k]}\left|L_s - L_{t_{k-\nu}}\right|
\end{aligned}$$

and by Gronwall's inequality, we obtain

$$\left|X^\varepsilon_t - X^\varepsilon_{t_{k-\nu}}\right| \le e^{C(t-t_{k-\nu})}\left(\frac{\ell}{n}\left|a(X^\varepsilon_{t_{k-\nu}},\theta_0)\right| + \varepsilon\sup_{s\in[t_{k-1},t_k]}\left|L_s - L_{t_{k-\nu}}\right|\right).$$

Thus,

$$|J_1| \le Ce^{C\ell/n}\left(\frac{\ell 2^{2\ell}}{n\varepsilon}\sup_{s,t\in[0,1]}|a(X_s^\varepsilon,\theta_0)f(X_t^\varepsilon,\theta)| + \frac{2^{2\ell}}{n}\sum_{k=\ell\vee 1}^{n}\sup_{\substack{\nu,\mu=0,\dots,\ell\\ s\in[t_{k-1},t_k]}}\left|f(X_{t_{k-\mu}}^\varepsilon,\theta)\right|\left|L_s - L_{t_{k-\nu}}\right|\right).$$

The next to the last term converges almost surely to zero as $\varepsilon \to 0$ and $\ell 2^{2\ell}/n\varepsilon \to 0$, uniformly in $\theta \in \Theta$. We remain to prove that

$$\frac{2^{2\ell}}{n}\sum_{k=\ell\vee 1}^{n}\sup_{\substack{\nu=0,\dots,\ell\\ s\in[t_{k-1},t_k]}}\left|L_s - L_{t_{k-\nu}}\right| \xrightarrow{P_{\theta_0}} 0.$$

This follows from the fact that

$$\sup_{\substack{\nu=0,\dots,\ell\\ s\in[t_{k-1},t_k]}}\left|L_s - L_{t_{k-\nu}}\right| \le \sup_{\substack{\nu=0,\dots,\ell\\ s\in[t_{k-1},t_k]}}\left|\tilde{L}_s - \tilde{L}_{t_{k-\nu}}\right| + (t_k - t_{k-\ell\vee 1}) + \int_{t_{k-\ell\vee 1}}^{t_k}\int_{|z|>1}|z|\,N(\mathrm{d}s\,,\mathrm{d}z),$$

where

$$\frac{2^{2\ell}}{n}\sum_{k=\ell\vee 1}^{n}(t_k - t_{k-\ell\vee 1}) = \frac{2^{2\ell}}{n}\sum_{\nu=0}^{(\ell-1)\vee 0}(t_{n-\nu} - t_\nu) = \frac{2^{2\ell}}{n}\sum_{\nu=0}^{(\ell-1)\vee 0}\left(1 - \frac{2\nu}{n}\right) \le \frac{\ell 2^{2\ell}}{n} \to 0,$$

$$\frac{2^{2\ell}}{n}\sum_{k=\ell\vee 1}^{n}\int_{t_{k-\ell\vee 1}}^{t_k}\int_{|z|>1}|z|\,N(\mathrm{d}s\,,\mathrm{d}z) \le \frac{\ell 2^{2\ell}}{n}\int_0^1\int_{|z|>1}|z|\,N(\mathrm{d}s\,,\mathrm{d}z) \xrightarrow{a.s.} 0 \qquad \text{as} \quad \frac{\ell 2^{2\ell}}{n} \to 0,$$

and by Doob's martingale inequality (see, *e.g.*, Theorem 2.1.5 in Applebaum [3])

$$\begin{aligned}
\frac{2^{2\ell}}{n}\sum_{k=\ell\vee 1}^{n}E\left[\sup_{\substack{\nu=0,\dots,\ell\\ s\in[t_{k-1},t_k]}}\left|\tilde{L}_s - \tilde{L}_{t_{k-\nu}}\right|\right] &\le 2^{2\ell}\left(\frac{1}{n}\sum_{k=\ell\vee 1}^{n}E\left[\sup_{\substack{\nu=0,\dots,\ell\\ s\in[t_{k-1},t_k]}}\left|\tilde{L}_s - \tilde{L}_{t_{k-\nu}}\right|^2\right]\right)^{1/2}\\
&\le C\left(\frac{2^{4\ell}}{n}\sum_{k=\ell\vee 1}^{n}E\left[\left|\tilde{L}_{t_k} - \tilde{L}_{t_{k-\ell}}\right|^2\right]\right)^{1/2}\\
&\le C\left(\frac{\ell 2^{4\ell}}{n}\left(\|\sigma\|_F^2 + \int_{|z|\le 1}|z|^2\,\nu(\mathrm{d}z)\right)\right)^{1/2} \to 0 \quad \text{as} \quad \frac{\ell 2^{4\ell}}{n} \to 0
\end{aligned}$$

with some positive constant C independent of n, ε, ℓ. Thus, for any $\eta > 0$,

$$\begin{aligned}
&P\left(\sup_{\theta\in\Theta}\frac{2^{2\ell}}{n}\sum_{k=\ell\vee 1}^{n}\sup_{\substack{\nu,\mu=0,\dots,\ell\\ t\in[t_{k-1},t_k]}}\left|f(X_{t_{k-\mu}}^\varepsilon,\theta)\right|\left|L_t - L_{t_{k-\nu}}\right| > \eta\right)\\
&\qquad\qquad \le P(1 > \tau_m^{n,\varepsilon,\ell}) + P\left(\|f\|_{C(B_m\times\Theta)}\frac{2^{2\ell}}{n}\sum_{k=\ell\vee 1}^{n}\sup_{\substack{\nu=0,\dots,\ell\\ t\in[t_{k-1},t_k]}}\left|L_t - L_{t_{k-\nu}}\right| > \eta\right)
\end{aligned}$$

converges in P_{θ_0} to zero as $n \to \infty$, $\varepsilon \to 0$ and $\ell 2^{4\ell}/n \to 0$. □

Analogously, we obtain the following proposition.

Proposition 2.3.7. *Let $\ell = \ell_n \in \mathbb{N}$ depend on n, and let f be a function as in Lemma 2.3.5. Under the assumptions* (A1) *and* (A2)*, if $2^\ell \varepsilon \to 0$ and $\ell 2^{2\ell}/n$ is bounded as $n \to \infty$ and $\varepsilon \to 0$, then it holds that*

$$\sum_{k=\ell\vee 1}^{n} A_\ell f(\boldsymbol{X}^\varepsilon_{t_k:t_{k-\ell}}, \theta) \cdot \left(X^\varepsilon_{t_k} - X^\varepsilon_{t_{k-1}} - \frac{1}{n} A_\ell a(\boldsymbol{X}^\varepsilon_{t_k:t_{k-\ell}}, \theta_0) \right) \xrightarrow{P_{\theta_0}} 0$$

as $n \to \infty$ and $\varepsilon \to 0$, uniformly in $\theta \in \Theta$.

2.3.3 Proof of main resutlts

Proof of Theorem 2.2.1. Let $f(x, \theta) = a(x, \theta_0) - a(x, \theta)$. Since

$$\begin{aligned}\Phi_{n,\varepsilon,\ell}(\theta) = 2 \sum_{k=\ell\vee 1}^{n} &\left(X^\varepsilon_{t_k} - X^\varepsilon_{t_{k-1}} - \frac{1}{n} A_\ell a(\boldsymbol{X}^\varepsilon_{t_k:t_{k-\ell}}, \theta_0) \right) \cdot A_\ell f(\boldsymbol{X}^\varepsilon_{t_k:t_{k-\ell}}, \theta) \\ &+ \frac{1}{n} \sum_{k=\ell\vee 1}^{n} \left| A_\ell f(\boldsymbol{X}^\varepsilon_{t_k:t_{k-\ell}}, \theta) \right|^2,\end{aligned}$$

it follows from Proposition 2.3.4 and 2.3.7 that for any $\eta > 0$, if $2^\ell \varepsilon \to 0$ and $\ell 2^{2\ell}/n$ is bounded, then

$$P\left(\sup_{\theta\in\Theta} \left| \Phi_{n,\varepsilon,\ell}(\theta) - \int_0^1 |f(x_s, \theta)|^2 \, \mathrm{d}s \right| > \eta \right) \to 0$$

as $n \to \infty$, $\varepsilon \to 0$. Also, (A4) implies that for any $\delta > 0$

$$\inf_{|\theta-\theta_0|>\delta} \int_0^1 |f(x_s, \theta)|^2 \, \mathrm{d}s > \int_0^1 |f(x_s, \theta_0)|^2 \, \mathrm{d}s = 0.$$

Thus, it follows from Theorem 5.9 in van der Vaart [35] that $\hat{\theta}_{n,\varepsilon,\ell}$ is consistent to θ_0. □

To prove Theorem 2.2.2, we prepare the following proposition.

Proposition 2.3.8. *Let $\ell = \ell_n \in \mathbb{N}$ depend on n. Assume the conditions* (A1)-(A5).

(i) *If $\ell 2^{4\ell}/n \to 0$, $2^\ell \varepsilon \to 0$ and $\ell 2^{2\ell}/n\varepsilon \to 0$ as $n \to \infty$ and $\varepsilon \to 0$, then*

$$\varepsilon^{-1} \frac{\partial \Phi_{n,\varepsilon,\ell}}{\partial \theta_i}(\theta_0) \xrightarrow{P_{\theta_0}} -2S_i(\theta_0)$$

as $n \to \infty$ and $\varepsilon \to 0$.

(ii) *If $2^\ell \varepsilon \to 0$ and $\ell 2^{2\ell}/n$ is bounded as $n \to \infty$ and $\varepsilon \to 0$, then*

$$\frac{\partial^2 \Phi_{n,\varepsilon,\ell}}{\partial \theta_i \partial \theta_j}(\theta) \xrightarrow{P_{\theta_0}} 2I^{ij}(\theta)$$

as $n \to \infty$ and $\varepsilon \to 0$, uniformly in $\theta \in \Theta$.

Proof. (i) Since we have

$$\frac{\partial \Phi_{n,\varepsilon,\ell}}{\partial \theta_i}(\theta) = -2 \sum_{k=\ell\vee 1}^{n} A_\ell \frac{\partial a}{\partial \theta_i}(\boldsymbol{X}^\varepsilon_{t_k:t_{k-\ell}}, \theta) \cdot \left(X^\varepsilon_{t_k} - X^\varepsilon_{t_{k-1}} - \frac{1}{n} A_\ell a(\boldsymbol{X}^\varepsilon_{t_k:t_{k-\ell}}, \theta_0) \right),$$

the consequence follows by Proposition 2.3.6 with $f(x,\theta) = -2\frac{\partial a}{\partial \theta_i}(x,\theta)$.

(ii) We have

$$\begin{aligned}\frac{\partial^2 \Phi_{n,\varepsilon,\ell}}{\partial \theta_i \partial \theta_j}(\theta) = {} & \frac{2}{n} \sum_{k=\ell\vee 1}^{n} A_\ell \frac{\partial a}{\partial \theta_i}(\boldsymbol{X}^\varepsilon_{t_k:t_{k-\ell}}, \theta) \cdot A_\ell \frac{\partial a}{\partial \theta_j}(\boldsymbol{X}^\varepsilon_{t_k:t_{k-\ell}}, \theta) \\ & - 2 \sum_{k=\ell\vee 1}^{n} A_\ell \frac{\partial^2 a}{\partial \theta_i \partial \theta_j}(\boldsymbol{X}^\varepsilon_{t_k:t_{k-\ell}}, \theta) \cdot \left(X^\varepsilon_{t_k} - X^\varepsilon_{t_{k-1}} - \frac{1}{n} A_\ell a(\boldsymbol{X}^\varepsilon_{t_k:t_{k-\ell}}, \theta_0) \right).\end{aligned}$$

By Proposition 2.3.7, the second term in the right-hand side converges in $P_{\theta_{\theta_0}}$ to zero uniformly in $\theta \in \Theta$. Also, by using Proposition 2.3.4 and the relationship between inner product and norm with $f(x,\theta) = \frac{\partial a}{\partial \theta_i}(x,\theta) \pm \frac{\partial a}{\partial \theta_j}(x,\theta)$ and with $q = 2$, the first term converges almost surely to $2I^{ij}(\theta)$ uniformly in $\theta \in \Theta$. □

Proof of Theorem 2.2.2. It follows from the mean value theorem that

$$\varepsilon^{-1} \left(\frac{\partial \Phi_{n,\varepsilon,\ell}}{\partial \theta_i}(\hat\theta_{n,\varepsilon,\ell}) - \frac{\partial \Phi_{n,\varepsilon,\ell}}{\partial \theta_i}(\theta_0) \right) = \varepsilon^{-1}(\hat\theta_{n,\varepsilon,\ell} - \theta_0) \cdot \int_0^1 \nabla_\theta \frac{\partial \Phi_{n,\varepsilon,\ell}}{\partial \theta_i} \left(\theta_0 + u(\hat\theta_{n,\varepsilon,\ell} - \theta_0) \right) \mathrm{d}u \,.$$

By the consistency of $\hat\theta_{n,\varepsilon,\ell}$ and Proposition 2.3.8 (i), the left-hand side converges to $2S_i(\theta_0)$ in P_{θ_0} as $n \to \infty$ and $\varepsilon \to 0$ if $\ell 2^{4\ell}/n \to 0$, $2^\ell \varepsilon \to 0$ and $\ell 2^{2\ell}/n\varepsilon \to 0$.

For an arbitrary convex neighborhood U of $\theta_0 \in \Theta_0$, we have

$$\begin{aligned} & \left| \frac{\partial^2 \Phi_{n,\varepsilon,\ell}}{\partial \theta_i \partial \theta_j}(\theta_0) - \int_0^1 \frac{\partial^2 \Phi_{n,\varepsilon,\ell}}{\partial \theta_i \partial \theta_j} \left(\theta_0 + u(\hat\theta_{n,\varepsilon,\ell} - \theta_0) \right) \mathrm{d}u \right| 1_{\{\hat\theta_{n,\varepsilon,\ell} \in U\}} \\ & \qquad \le \sup_{\theta \in U} \left| \frac{\partial^2 \Phi_{n,\varepsilon,\ell}}{\partial \theta_i \partial \theta_j}(\theta_0) - \frac{\partial^2 \Phi_{n,\varepsilon,\ell}}{\partial \theta_i \partial \theta_j}(\theta) \right| \\ & \qquad \le 2 \sup_{\theta \in U} \left| \frac{\partial^2 \Phi_{n,\varepsilon,\ell}}{\partial \theta_i \partial \theta_j}(\theta) - 2I^{ij}(\theta) \right| + 2 \sup_{\theta \in U} \left| I^{ij}(\theta) - I^{ij}(\theta_0) \right|. \end{aligned}$$

It follows from the consistency of $\hat\theta_{n,\varepsilon,\ell}$, Proposition 2.3.8 (ii) and the continuity of $\theta \mapsto I(\theta)$ that if $2^\ell \varepsilon \to 0$ and $\ell 2^{2\ell}/n$ is bounded, then

$$2I^{ij}_{n,\varepsilon,\ell} := \int_0^1 \frac{\partial^2 \Phi_{n,\varepsilon,\ell}}{\partial \theta_i \partial \theta_j} \left(\theta_0 + u(\hat\theta_{n,\varepsilon,\ell} - \theta_0) \right) \mathrm{d}u \xrightarrow{P_{\theta_0}} 2I^{ij}(\theta_0)$$

as $n \to \infty$ and $\varepsilon \to 0$. By Lemma 2.5.3, the proof is complete. □

2.4 Numerical experiment

In this section, we give a simulation by numerical computation to compare our estimators with well-known least squares estimators for an Ornstein-Uhlenbeck process given by

$$\mathrm{d}X_t = -\theta_0 X_t \,\mathrm{d}t + \varepsilon \,\mathrm{d}L_t\,, \qquad X_0 = x_0, \tag{2.8}$$

where L is one of the following processes:

- L is the standard Brownian motion.
- L is the sum of the standard Brownian motion and the compound Poisson process with intensity 10 and with standard exponentially distributed jumps.

For simplicity, we set $\theta_0 = 1$ and $x_0 = 1$ with $\varepsilon = 0.1, 0.5, 1.0$ and $n = 10, 100, 1000$. As we mentioned in the remark of Theorem 2.2.1 and 2.2.1, we take $\ell = 1, \dots, 6$, and we shall compare our Adams-Moulton type estimators

$$\hat{\theta}_{n,\varepsilon,\ell} := \operatorname*{argmin}_{\theta\in\Theta} \sum_{k=\ell}^{n} \left| X^\varepsilon_{t_k} - X^\varepsilon_{t_{k-1}} + \frac{1}{n}\theta A_\ell a(\boldsymbol{X}^\varepsilon_{t_k:t_{k-\ell}}) \right|^2 \quad (\ell = 1, \dots, 6)$$

to the usual 'Euler-type' estimator

$$\hat{\theta}_{n,\varepsilon} := \operatorname*{argmin}_{\theta\in\Theta} \sum_{k=1}^{n} \left| X^\varepsilon_{t_k} - X^\varepsilon_{t_{k-1}} + \frac{1}{n}\theta X^\varepsilon_{t_{k-1}} \right|^2,$$

where $A_\ell a(\boldsymbol{X}^\varepsilon_{t_k:t_{k-\ell}})$ with $a(x) = x$ are given by

$$A_\ell a(\boldsymbol{X}^\varepsilon_{t_k:t_{k-\ell}}) = \begin{cases} \frac{1}{2}X^\varepsilon_{t_k} + \frac{1}{2}X^\varepsilon_{t_{k-1}} & \text{if } \ell = 1, \\ \frac{5}{12}X^\varepsilon_{t_k} + \frac{2}{3}X^\varepsilon_{t_{k-1}} - \frac{1}{12}X^\varepsilon_{t_{k-2}} & \text{if } \ell = 2, \\ \frac{3}{8}X^\varepsilon_{t_k} + \frac{19}{24}X^\varepsilon_{t_{k-1}} - \frac{5}{24}X^\varepsilon_{t_{k-2}} + \frac{1}{24}X^\varepsilon_{t_{k-3}} & \text{if } \ell = 3, \\ \frac{251}{720}X^\varepsilon_{t_k} + \frac{323}{360}X^\varepsilon_{t_{k-1}} - \frac{11}{30}X^\varepsilon_{t_{k-2}} + \frac{53}{360}X^\varepsilon_{t_{k-3}} - \frac{19}{720}X^\varepsilon_{t_{k-4}} & \text{if } \ell = 4, \\ \frac{95}{288}X^\varepsilon_{t_k} + \frac{1427}{1440}X^\varepsilon_{t_{k-1}} - \frac{133}{240}X^\varepsilon_{t_{k-2}} + \frac{241}{720}X^\varepsilon_{t_{k-3}} - \frac{173}{1440}X^\varepsilon_{t_{k-4}} + \frac{3}{160}X^\varepsilon_{t_{k-5}} & \text{if } \ell = 5, \\ \frac{19087}{60480}X^\varepsilon_{t_k} + \frac{2713}{2520}X^\varepsilon_{t_{k-1}} - \frac{15487}{20160}X^\varepsilon_{t_{k-2}} + \frac{586}{945}X^\varepsilon_{t_{k-3}} & \\ \qquad - \frac{6737}{20160}X^\varepsilon_{t_{k-4}} + \frac{263}{2520}X^\varepsilon_{t_{k-5}} - \frac{863}{60480}X^\varepsilon_{t_{k-6}} & \text{if } \ell = 6. \end{cases}$$

Such coefficients from the Adams-Moulton method can be seen, *e.g.*, in Table 244 in Butcher [7]. Note that the Euler-type LSE $\hat{\theta}_{n,\varepsilon}$ is slightly different from the Adams-Moulton type LSE $\hat{\theta}_{n,\varepsilon,0}$, *i.e.*, 'backward Euler-type', but for the sake of both similarity, we omit to consider $\hat{\theta}_{n,\varepsilon,0}$. Here, we compute

$$\hat{\theta}_{n,\varepsilon} = -\frac{\sum_{k=1}^{n}(X^\varepsilon_{t_k} - X^\varepsilon_{t_{k-1}})X^\varepsilon_{t_{k-1}}}{\frac{1}{n}\sum_{k=1}^{n}|X^\varepsilon_{t_{k-1}}|^2},$$

and

$$\hat{\theta}_{n,\varepsilon,\ell} = -\frac{\sum_{k=\ell}^{n}(X^\varepsilon_{t_k} - X^\varepsilon_{t_{k-1}})A_\ell a(\boldsymbol{X}^\varepsilon_{t_k:t_{k-\ell}})}{\frac{1}{n}\sum_{k=\ell}^{n}|A_\ell a(\boldsymbol{X}^\varepsilon_{t_k:t_{k-\ell}})|^2} \quad (\ell = 1, \dots, 6).$$

In Table 2.4.1, we make a sample path $\{X^{\varepsilon}_{t_k}\}_{k=0}^{n}$ by

$$X^{\varepsilon}_{t_0} = x_0, \quad X_{t_k} = e^{-\theta_0 \Delta t} X_{t_{k-1}} + \varepsilon \sqrt{\frac{1 - e^{-2\theta_0 \Delta t}}{2\theta_0}} N(0,1), \quad \Delta t = t_k - t_{k-1} = \frac{1}{n}$$

as a well-known way of constructing an exact numerical solution of (2.8), where $N(0,1)$ is the standard normal variable. In Table 2.4.2, we make a sample path by using Euler Maruyama method with $100 \times 1,000$ number of subdivisions for the interval $[0,1]$. We iterate this computation 10,000 times and show their sample means and standard deviations in Table 2.4.1. We also plot the sample means and 95% confidence intervals of $\hat{\theta}_{n,\varepsilon,\ell}$ through iterations in Figure 2.4.1.

From Tables 2.4.1 and 2.4.2, and Figure 2.4.1 and 2.4.2, we see that the simulation performances depend on ℓ and become better than one from Euler's method when we employ an appropriate ℓ.

Table 2.4.1:
Sample mean (with standard deviation in parentheses) of LSEs, based on 10,000 sample paths from the OU process (2.8) with $(\theta_0, x_0) = (1.0, 1.0)$ and with standard Brownian motion. We emphasize the best average of LSEs for each (ε, n) using a bold font.

$\varepsilon = 1.0$	$n = 10$	$n = 100$	$n = 1000$
Euler	1.663489 (1.471654)	1.931070 (1.821037)	1.966545 (1.873523)
AM1	0.951550 (1.283734)	0.802641 (1.038229)	0.790167 (1.010930)
AM2	1.028716 (1.564615)	**0.987162** (1.187243)	**0.986894** (1.152948)
AM3	1.074215 (1.833185)	1.078157 (1.261705)	1.084002 (1.225335)
AM4	1.087510 (2.127400)	1.131508 (1.309929)	1.146718 (1.273696)
AM5	**1.026782** (2.333814)	1.167348 (1.354262)	1.190314 (1.307555)
AM6	0.884837 (2.779585)	1.192481 (1.387997)	1.224354 (1.336056)

$\varepsilon = 0.1$	$n = 10$	$n = 100$	$n = 1000$
Euler	0.964199 (0.138219)	1.010592 (0.151375)	1.015443 (0.152811)
AM1	**1.004400** (0.154055)	**1.004306** (0.152030)	**1.004406** (0.151787)
AM2	1.008660 (0.174779)	1.006182 (0.153983)	1.006442 (0.152153)
AM3	1.012103 (0.198794)	1.007291 (0.156226)	1.007306 (0.152466)
AM4	1.013956 (0.230109)	1.007880 (0.157855)	1.007970 (0.152656)
AM5	1.015775 (0.269280)	1.008047 (0.159778)	1.008404 (0.152829)
AM6	1.019984 (0.321273)	1.008658 (0.162266)	1.008764 (0.153157)

$\varepsilon = 0.01$	$n = 10$	$n = 100$	$n = 1000$
Euler	0.951791 (0.013711)	0.995199 (0.014975)	0.999686 (0.015110)
AM1	0.999246 (0.015337)	1.000049 (0.015147)	**1.000074** (0.015126)
AM2	1.000177 (0.017310)	1.000061 (0.015320)	1.000102 (0.015141)
AM3	1.000232 (0.019645)	1.000070 (0.015533)	1.000100 (0.015160)
AM4	1.000138 (0.022662)	1.000062 (0.015683)	1.000110 (0.015169)
AM5	**1.000017** (0.026460)	**1.000026** (0.015867)	1.000113 (0.015182)
AM6	1.000139 (0.031419)	1.000041 (0.016101)	1.000117 (0.015209)

AMℓ: LSE via the Adams-Moulton method with order ℓ ($\ell = 1, \ldots, 6$).

Table 2.4.2:
Sample mean (with standard deviation in parentheses) of LSEs, based on 10,000 sample paths from the OU process (2.8) with $(\theta_0, x_0) = (1.0, 1.0)$ and with Lévy noise which is the sum of the standard Brownian motion and the compound Poisson process. We emphasize the best average of LSEs for each (ε, n) using a bold font.

$\varepsilon = 1.0$	$n = 10$	$n = 100$	$n = 1000$
Euler	2.007064 (2.384299)	2.404387 (2.936047)	2.452525 (3.013250)
AM1	-0.715378 (1.566790)	-0.639030 (1.750689)	-0.629191 (1.514785)
AM2	-0.014319 (2.051735)	-0.104887 (1.768940)	-0.112236 (1.493505)
AM3	0.286297 (2.305605)	0.161202 (1.833517)	0.145195 (1.544060)
AM4	0.477244 (2.574459)	0.331748 (1.895879)	0.308179 (1.595550)
AM5	**0.511744** (2.868428)	0.445835 (1.947106)	0.423175 (1.637366)
AM6	0.439842 (3.366565)	**0.521836** (1.970925)	**0.510284** (1.669400)

$\varepsilon = 0.1$	$n = 10$	$n = 100$	$n = 1000$
Euler	1.166634 (0.741016)	1.280757 (0.936255)	1.294340 (0.977801)
AM1	**1.060503** (0.790647)	**1.034306** (0.747335)	**1.031139** (0.744317)
AM2	1.130144 (0.993074)	1.077497 (0.771562)	1.073988 (0.771562)
AM3	1.167730 (1.108686)	1.100365 (0.788456)	1.096388 (0.792289)
AM4	1.214359 (1.338724)	1.115808 (0.820980)	1.108822 (0.776632)
AM5	1.223887 (1.515457)	1.127240 (0.831186)	1.118895 (0.787421)
AM6	1.176586 (1.749853)	1.135689 (0.842277)	1.126413 (0.794018)

$\varepsilon = 0.01$	$n = 10$	$n = 100$	$n = 1000$
Euler	0.952623 (0.062801)	0.996526 (0.068634)	1.001071 (0.069261)
AM1	0.998390 (0.070618)	0.999100 (0.069312)	0.999127 (0.069208)
AM2	**1.000092** (0.079777)	0.999481 (0.070127)	0.999517 (0.069260)
AM3	1.000229 (0.090495)	0.999731 (0.070769)	0.999721 (0.069369)
AM4	1.001112 (0.103840)	0.999782 (0.071581)	0.999834 (0.069484)
AM5	1.001624 (0.120812)	0.999925 (0.072499)	0.999920 (0.069517)
AM6	1.001718 (0.143539)	**1.000004** (0.073293)	**1.000001** (0.069566)

2.5 Appendix

Lemma 2.5.1. *Let $\ell \in \mathbb{N}$. Let $\gamma_{\ell\nu}$ and $\beta_{\ell\nu}$ be given by* (1.2) *and* (1.3)*. Then,*

$$\sum_{\nu=1}^{\ell} |\gamma_{\ell\nu}| \le \ell 2^{\ell-1} \quad (\ell = 1, 2, \dots), \qquad \sum_{\nu=0}^{\ell} |\beta_{\ell\nu}| \le 2^{\ell} \quad (\ell = 0, 1, \dots).$$

Proof. The conclusion is obtained from

$$\sum_{\nu=1}^{\ell} |\gamma_{\ell\nu}| = \sum_{\nu=1}^{\ell} \frac{1}{(\nu-1)!(\ell-\nu)!} \int_0^1 \prod_{\substack{j=1 \\ j\neq\nu}}^{\ell} (u+j-1)\,\mathrm{d}u \le \sum_{\nu=1}^{\ell} \frac{\ell!}{(\nu-1)!(\ell-\nu)!} = \ell 2^{\ell-1}$$

for $\ell = 1, 2 \dots$, and

$$\sum_{\nu=0}^{\ell} |\beta_{\ell\nu}| = \sum_{\nu=0}^{\ell} \frac{1}{\nu!(\ell-\nu)!} \int_0^1 \prod_{\substack{j=0 \\ j\neq\nu}}^{\ell} (u+j-1)\,\mathrm{d}u \le \sum_{\nu=0}^{\ell} \frac{\ell!}{\nu!(\ell-\nu)!} = 2^{\ell}$$

for $\ell = 0, 1, \dots$. □

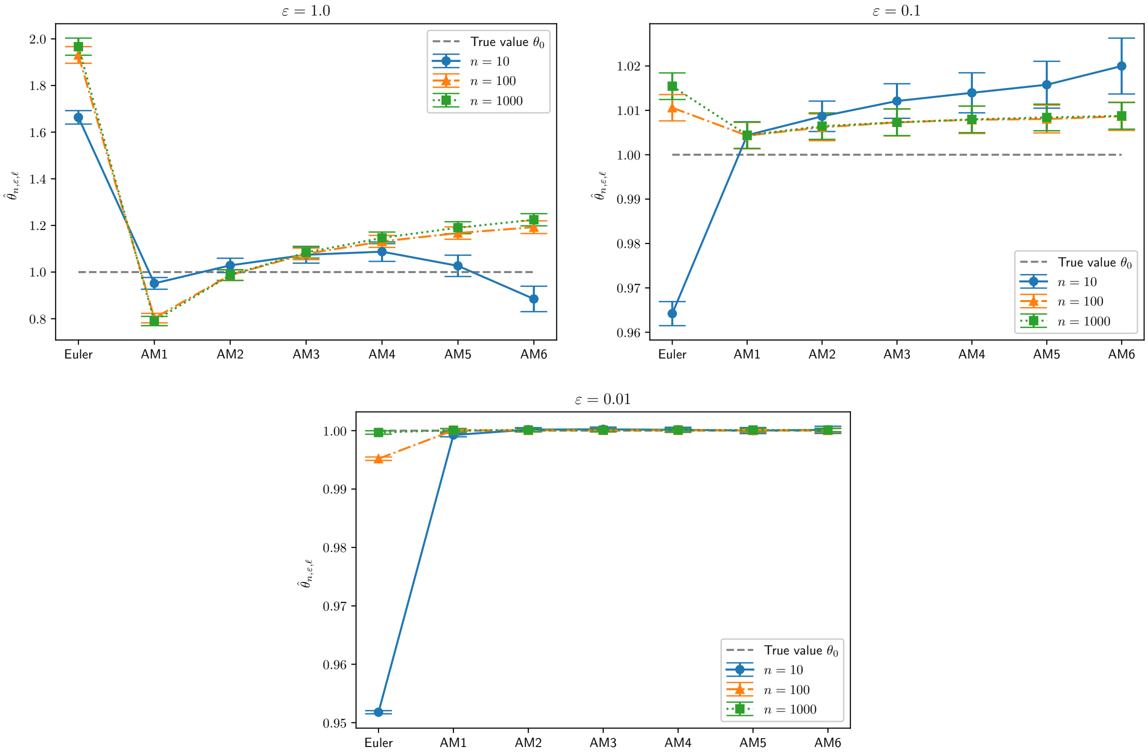

Figure 2.4.1:
The means and 95% confidence intervals through 10,000 iteration for $\hat{\theta}_{n,\varepsilon}$ (Euler) and $\hat{\theta}_{n,\varepsilon,\ell}$ (AMℓ, $\ell = 1, \dots, 6$), with the same data in Figure 2.4.1.

Lemma 2.5.2. *Let g be a continuous function on $\mathbb{R}^d$, let $t \mapsto y_t$ be an $\mathbb{R}^d$-valued continuous function on $[0,1]$, and let $\{f(\cdot,\theta)\}_{\theta\in\Theta}$ be a pointwise equicontinuous family of functions from $\mathbb{R}^d$ to $\mathbb{R}^d$. If $\ell/n \to 0$ as $n \to \infty$, then*

$$\frac{1}{n}\sum_{k=\ell\vee 1}^{n} g\left(⨍_{t_{k-1}}^{t_k} f(y_t,\theta)\,\mathrm{d}t\right) \to \int_0^1 g\circ f(y_t,\theta)\,\mathrm{d}t$$

as $n \to \infty$, uniformly in $\theta \in \Theta$.

Proof. Since $\{f(y_\cdot,\theta)\}_{\theta\in\Theta}$ is uniformly equicontinuous on $[0,1]$, for any $\eta > 0$ there exists $N \in \mathbb{N}$ such that $\theta \in \Theta$, $|s-t| \le 1/N \Rightarrow |f(y_s,\theta) - f(y_t,\theta)| < \eta$. Then, for all $n \ge N$, $t \in [0,1)$ and $\theta \in \Theta$

$$\left|\sum_{k=1}^{n} \mathbf{1}_{[t_{k-1},t_k)}(t) ⨍_{t_{k-1}}^{t_k} f(y_s,\theta)\,\mathrm{d}s - f(y_t,\theta)\right| \le \sum_{k=1}^{n} \mathbf{1}_{[t_{k-1},t_k)}(t) ⨍_{t_{k-1}}^{t_k} |f(y_s,\theta) - f(y_t,\theta)|\,\mathrm{d}s < \eta,$$

and we have

$$\sum_{k=1}^{n} \mathbf{1}_{[t_{k-1},t_k)}(t) ⨍_{t_{k-1}}^{t_k} f(y_s,\theta)\,\mathrm{d}s \to f(y_t,\theta)$$

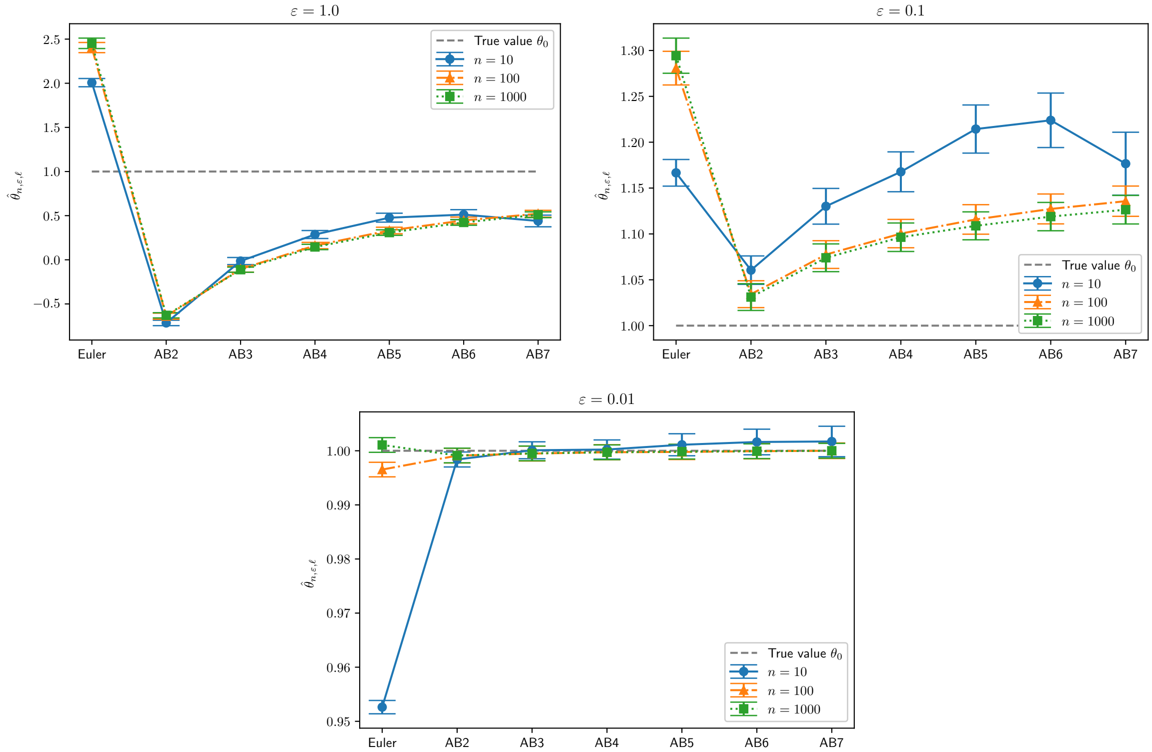

Figure 2.4.2:
The means and 95% confidence intervals through 10,000 iteration for $\hat{\theta}_{n,\varepsilon}$ (Euler) and $\hat{\theta}_{n,\varepsilon,\ell}$ (AMℓ, $\ell = 1, \dots, 6$), with the same data in Figure 2.4.2.

uniformly in $(t, \theta) \in [0, 1) \times \Theta$. By the continuity of g, we obtain

$$\begin{aligned}\frac{1}{n}\sum_{k=1}^{n} g\left(\fint_{t_{k-1}}^{t_k} f(y_t, \theta)\,\mathrm{d}t\right) &= \int_0^1 \sum_{k=1}^{n} \mathbf{1}_{[t_{k-1}, t_k)}(t)\, g\left(\fint_{t_{k-1}}^{t_k} f(y_s, \theta)\,\mathrm{d}s\right)\mathrm{d}t \\ &= \int_0^1 g\left(\sum_{k=1}^{n} \mathbf{1}_{[t_{k-1}, t_k)}(t) \fint_{t_{k-1}}^{t_k} f(y_s, \theta)\,\mathrm{d}s\right)\mathrm{d}t \to \int_0^1 g \circ f(y_t, \theta)\,\mathrm{d}t\end{aligned}$$

as $n \to \infty$, uniformly in $\theta \in \Theta$. Since $\{g \circ f(y_\cdot, \theta)\}_{\theta \in \Theta}$ is equicontinuous at $t = 0$, for $\ell \geq 2$,

$$\frac{1}{n}\sum_{k=1}^{\ell-1} g\left(\fint_{t_{k-1}}^{t_k} f(y_t, \theta)\,\mathrm{d}t\right) \to 0$$

as $\ell/n \to 0$, uniformly in $\theta \in \Theta$. □

Let $(\Omega, P, \mathscr{F})$ be a probability space, and let $\mathrm{Sym}_p(\mathbb{R})$ denote the set of all $p \times p$ symmetric matrix with real entries and with the Frobenius norm $\|\cdot\|_F$.

Lemma 2.5.3. *Suppose that* $v_n \xrightarrow{p} v$ *in* $\mathbb{R}^p$ *and* $M_n \xrightarrow{p} M$ *in* $\mathrm{Sym}_p(\mathbb{R})$ *as* $n \to \infty$*,* w_n *satisfies* $v_n = M_n w_n$*. If* M *is positive definite,* $w_n \xrightarrow{p} M^{-1}v$*.*

Proof. Let η be an arbitrary positive number less than the smallest eigenvalue of M. If $\|M_n - M\|_F < \eta$, then $0 \prec M - \eta\mathbb{I}_{p\times p} \prec M_n \prec M + \eta\mathbb{I}_{p\times p}$, where $\mathbb{I}_{p\times p}$ is the identity matrix of size p and $\prec$ is the Loewner order. This implies that M_n is invetible and

$$(M + \eta\mathbb{I}_{p\times p})^{-1} \prec M_n^{-1} \prec (M - \eta\mathbb{I}_{p\times p})^{-1}.$$

Since $(M \pm \eta\mathbb{I}_{p\times p})^{-1} \to M^{-1}$ in $\mathrm{Sym}_p(\mathbb{R})$ as $\eta \to 0$, there exists a positive number $\tilde{\eta}$ depending only on M, p and η such that $\|M_n^{-1} - M^{-1}\|_F < \tilde{\eta}$ and $\tilde{\eta} \to 0$ as $\eta \to 0$.

Set $\mathscr{D}_n := \{\omega \in \Omega \,|\, M_n(\omega) \text{ is invertible}\}$. Then, if an arbitrary positive number $\tilde{\eta}$ is sufficiently small, for some $\eta > 0$ we have

$$P(\mathscr{D}_n^{\mathrm{C}}) + P(\mathbf{1}_{\mathscr{D}_n}\|M_n^{-1} - M^{-1}\|_F > \tilde{\eta}) \leq 2P(\|M_n^{-1} - M^{-1}\|_F > \eta) \to 0,$$

where $\mathbf{1}_A$ is the indicator function on a set $A \subset \Omega$. Hence, we obtain

$$w_n = M_n^{-1}v_n\mathbf{1}_{\mathscr{D}_n} + w_n\mathbf{1}_{\mathscr{D}_n^{\mathrm{C}}} \xrightarrow{p} M^{-1}v$$

as $n \to \infty$. □

Chapter 3

QMLE type threshold estimation

3.1 Introduction

In this chapter, we are concerned with the following stochastic differential equation (SDE):

$$\mathrm{d}X_t^\varepsilon = a(X_t^\varepsilon, \mu_0)\,\mathrm{d}t + \varepsilon b(X_{t-}^\varepsilon, \sigma_0)\,\mathrm{d}W_t + \varepsilon c(X_{t-}^\varepsilon, \alpha_0)\,\mathrm{d}Z_t^{\lambda_\varepsilon}, \quad X_0^\varepsilon = x_0 \in \mathbb{R}, \tag{3.1}$$

where $\varepsilon > 0$, and Θ_i $(i = 1, 2, 3)$ are smooth bounded open convex sets in $\mathbb{R}^{d_i}$ with $d_i \in \mathbb{N}$ $(i = 1, 2, 3)$, respectively, and $\theta_0 = (\mu_0, \sigma_0, \alpha_0) \in \Theta_0 := \Theta_1 \times \Theta_2 \times \Theta_3 \subset \mathbb{R}^d$ with $d := d_1 + d_2 + d_3$ with $\Theta := \bar{\Theta}_0$, and each domain of a, b, c is $\mathbb{R} \times \bar{\Theta}_i$ $(i = 1, 2, 3)$, respectively. Also, $Z^{\lambda_\varepsilon} = (Z_t^{\lambda_\varepsilon})_{t\ge 0}$ is a compound Poisson process given by

$$Z_t^{\lambda_\varepsilon} = \sum_{i=1}^{N_t^{\lambda_\varepsilon}} V_i, \quad Z_0^{\lambda_\varepsilon} = 0,$$

where $N^{\lambda_\varepsilon} = (N_t^{\lambda_\varepsilon})_{t\ge 0}$ is a Poisson process with intensity $\lambda_\varepsilon > 0$, and V_i's are i.i.d. random variables with common probability density function f_{α_0}, and are independent of N^{λ_ε} (*cf.* Example 1.3.10 in Applebaum [3]). $W = (W_t)_{t\ge 0}$ is a Wiener process. Here, we denote the filtered probability space by $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\ge 0}, P)$. Suppose that we have discrete data $X_{t_0}^\varepsilon, \dots, X_{t_n}^\varepsilon$ from (3.1) for $0 = t_0 < \cdots < t_n = 1$ with $t_i - t_{i-1} = 1/n$. We consider the problem of estimating the true $\theta_0 \in \Theta_0$ under $n \to \infty$ and $\varepsilon \to 0$. We also define x_t as the solution of the corresponding deterministic differential equation

$$\frac{\mathrm{d}x_t}{\mathrm{d}t} = a(x_t, \mu_0)$$

with the initial condition x_0.

3.2 Assumptions and notations

This section is devoted to prepare some notations and assumptions. Before going to see our assumptions, we begin by setting up the following two notations:

Notation 3.2.1. *Let I_{x_0} be the image of $t \mapsto x_t$ on $[0,1]$, and set*

$$I_{x_0}^{\delta} := \left\{ y \in \mathbb{R} \,\middle|\, \mathrm{dist}(y, I_{x_0}) = \inf_{x \in I_{x_0}} |x-y| < \delta \right\}.$$

Notation 3.2.2. *A function ψ on $\mathbb{R} \times \mathbb{R} \times \Theta_3$ is of the form*

$$\psi(x,y,\alpha) := \begin{cases} \log \left| \dfrac{1}{c(x,\alpha)} f_{\alpha}\left(\dfrac{y}{c(x,\alpha)} \right) \right| & \text{if } c(x,\alpha) \neq 0 \text{ and } f_{\alpha}\left(\dfrac{y}{c(x,\alpha)} \right) > 0, \\ 0 & \text{otherwise.} \end{cases}$$

Then, we prepare the following assumptions:

Assumption 3.2.1. *$a(\cdot,\mu_0)$, $b(\cdot,\sigma_0)$ and $c(\cdot,\alpha_0)$ are Lipschitz continuous on $\mathbb{R}$.*

Assumption 3.2.2. *The functions a, b, c are differentiable with respect to θ on $I_{x_0}^{\delta} \times \Theta$ for some $\delta > 0$, and the families $\left\{ \frac{\partial a}{\partial \theta_j}(\cdot,\mu) \right\}_{\mu \in \Theta_1}$, $\left\{ \frac{\partial b}{\partial \theta_j}(\cdot,\sigma) \right\}_{\sigma \in \Theta_2}$, $\left\{ \frac{\partial c}{\partial \theta_j}(\cdot,\alpha) \right\}_{\alpha \in \Theta_3}$ $(j = 1, \ldots, d)$ are equi-Lipschitz continuous on $I_{x_0}^{\delta}$.*

Assumption 3.2.3. *For any $p \geq 0$, let $f_{\alpha_0} : \mathbb{R} \to \mathbb{R}$ satisfy*

$$\int_{\mathbb{R}} |z|^p f_{\alpha_0}(z) \, \mathrm{d}z < \infty.$$

Assumption 3.2.4. *The family $\{f_{\alpha}\}_{\alpha \in \bar{\Theta}_3}$ satisfies either of the following condtions:*

(i) f_{α}, $\alpha \in \bar{\Theta}_3$ are positive and continuous on $\mathbb{R}$.

(ii) f_{α}, $\alpha \in \bar{\Theta}_3$ are positive and continuous on $\mathbb{R}_+ (= (0,\infty))$, and are zero on $(-\infty, 0]$.

Assumption 3.2.5. *The family $\{b(\cdot,\sigma)\}_{\sigma \in \bar{\Theta}_2}$ satisfies*

$$\inf_{(x,\sigma) \in I_{x_0} \times \Theta_2} |b(x_t,\sigma)| > 0.$$

Assumption 3.2.6. *The familiy $\{c(\cdot,\alpha)\}_{\sigma \in \bar{\Theta}_3}$ satisfies*

$$0 < c_1 \leq |c(x,\alpha)| \leq c_2 \quad \text{for } (x,\alpha) \in I_{x_0} \times \Theta_3$$

with some positve constants c_1 and c_2. In this thesis, without loss of generality, we may assume

$$c(x_t,\alpha) > c_1 \quad \text{for } (x,\alpha) \in I_{x_0} \times \Theta_3.$$

Assumption 3.2.7. *If $\mu \neq \mu_0$, $\sigma \neq \sigma_0$ or $\alpha \neq \alpha_0$, then*

$$a(y,\mu) \not\equiv a(y,\mu_0), \quad b(y,\sigma) \not\equiv b(y,\sigma_0) \quad \text{or} \quad \psi(y, c(y,z,\alpha) \neq \psi(y,z,\alpha_0), \quad \text{respectively}$$

for some $y \in I_{x_0}^{\delta}$ with some $\delta > 0$, and for some $z \in \mathbb{R}$.

Assumption 3.2.8. *$v_{n1}, \dots, v_{nn}$ are random variables such that v_{nk} is $\mathcal{F}_{t_{k-1}}$-measurable (or measurable with respect to $\{X_{t_j}; j < k\}$), and they satisfy*

$$0 < v_1 \le v_{nk} \le v_2$$

for some constants v_1 and v_2.

Assumption 3.2.9. *There exists $\delta > 0$ such that for $(x, y, \alpha) \in I_{x_0}^{\delta} \times \mathbb{R} \times \Theta$ with $\psi \neq 0$, ψ is differentiable with respect to α_i $(i = 1, \dots, d_3)$. For $\alpha \in \Theta_3$*

$$x \mapsto \int \psi(x, c(x, \alpha_0)z, \alpha) f_{\alpha_0}(z)\,\mathrm{d}z\,, \quad x \mapsto \int |\psi(x, c(x, \alpha_0)z, \alpha)|^2 f_{\alpha_0}(z)\,\mathrm{d}z$$

are continuous at every points in I_{x_0}, and there exist $\delta > 0$ and $C > 0$ such that

$$\int \left\{ \sup_{(x,\alpha) \in I_{x_0}^{\delta} \times \Theta_3} |\psi(x, c(x, \alpha_0)z, \alpha)| + \sum_{j=1}^{d_3} \sup_{(x,\alpha) \in I_{x_0}^{\delta} \times \Theta_3} \left| \frac{\partial \psi}{\partial \alpha_j}(x, c(x, \alpha_0)z, \alpha) \right| \right\} f_{\alpha_0}(z)\,\mathrm{d}z < \infty.$$

Assumption 3.2.10. *Relative to the choice (i) or (ii) in Assumption 3.2.4, we assume either of the following conditions (i) or (ii), respectively:*

(i) Under Assumption 3.2.4 (i), there exist constants $C > 0$, $q \ge 1$ and $\delta > 0$ such that

$$\sup_{(x,\alpha) \in I_{x_0}^{\delta} \times \Theta_3} \left| \frac{\partial \psi}{\partial y}(x, y, \alpha) \right| \le C(1 + |y|^q) \quad (y \in \mathbb{R}).$$

(ii) Under Assumption 3.2.4 (ii), we assume the following three conditions:

(ii.a) There exists $\delta > 0$ and $L > 0$ such that if $0 < y_1 \le y \le y_2$, then

$$\left| \frac{\partial \psi}{\partial y}(x, y, \alpha) \right| \le \left| \frac{\partial \psi}{\partial y}(x, y_1, \alpha) \right| + \left| \frac{\partial \psi}{\partial y}(x, y_2, \alpha) \right| + L \quad \text{for all } (x, \alpha) \in I_{x_0}^{\delta} \times \Theta_3.$$

(ii.b) There exist constants $q \ge 0$ and $\delta > 0$ such that

$$\sup_{(x,\alpha) \in I_{x_0}^{\delta} \times \Theta_3} \left| \frac{\partial \psi}{\partial y}(x, y, \alpha) \right| \le O\left(\frac{1}{|y|^q} \right) \quad \text{as } |y| \to 0.$$

(ii.c) There exists $\delta > 0$ such that for any $C_1 > 0$ and $C_2 \ge 0$ the map

$$x \mapsto \int \sup_{\alpha \in \Theta_3} \left| \frac{\partial \psi}{\partial y}(x, C_1 y + C_2, \alpha) \right| f_{\alpha_0}(y)\,\mathrm{d}y$$

takes values in $\mathbb{R}$ from $I_{x_0}^{\delta}$, and is continous on $I_{x_0}^{\delta}$.

Assumption 3.2.11. *For $(x, y, \alpha) \in I_{x_0}^{\delta} \times \mathbb{R} \times \Theta$ with $\psi \neq 0$, ψ is differentiable with respect to $\alpha \in \Theta_3$, and*

$$x \mapsto \int \frac{\partial \psi}{\partial \alpha_i} \frac{\partial \psi}{\partial \alpha_j}(x, c(x, \alpha_0)z, \alpha_0) f_{\alpha_0}(z)\,\mathrm{d}z \quad (i, j = 1, \dots, d_3)$$

is continuous at every point $x \in I_{x_0}$.

Assumption 3.2.12. *The functions a, b, c are twice differentiable with respect to θ on $I_{x_0}^{\delta} \times \Theta$ for some δ, and the families $\left\{\frac{\partial^2 a}{\partial\theta_i\partial\theta_j}(\cdot,\mu)\right\}_{\mu\in\Theta_1}$, $\left\{\frac{\partial^2 b}{\partial\theta_i\partial\theta_j}(\cdot,\sigma)\right\}_{\sigma\in\Theta_2}$, $\left\{\frac{\partial^2 c}{\partial\theta_i\partial\theta_j}(\cdot,\alpha)\right\}_{\alpha\in\Theta_3}$ $(i, j = 1,\dots,d)$ are equi-Lipschitz continuous on $I_{x_0}^{\delta}$. There exists $\delta > 0$ such that for $(x, y, \alpha) \in I_{x_0}^{\delta} \times \mathbb{R} \times \Theta$ with $\psi \neq 0$, ψ is twice differentiable with respect to α_i $(i = 1,\dots,d_3)$. For $\alpha \in \Theta$, $i = 1,\dots,d_3$*

$$x \mapsto \int \frac{\partial\psi}{\partial\alpha_i}\big(x, c(x,\alpha_0)z, \alpha\big) f_{\alpha_0}(z)\,\mathrm{d}z\,, \quad x \mapsto \int \left|\frac{\partial\psi}{\partial\alpha_i}\big(x, c(x,\alpha_0)z, \alpha\big)\right|^2 f_{\alpha_0}(z)\,\mathrm{d}z$$

are continuous at every points $x \in I_{x_0}$, and there exist $\delta > 0$ such that

$$\int \sum_{i,j=1}^{d_3} \sup_{(x,\theta)\in I_{x_0}^{\delta}\times\Theta} \left|\frac{\partial^2\psi}{\partial\alpha_i\partial\alpha_j}\big(x, c(x,\alpha_0)z, \theta\big)\right| f_{\alpha_0}(z)\,\mathrm{d}z < \infty.$$

Relative to the choice (i) or (ii) in Assumption 3.2.4, we assume either of the following conditions (i) or (ii), respectively:

(i) Under Assumption 3.2.4 (i), there exist constants $C > 0$, $q \geq 1$ and $\delta > 0$ such that

$$\sup_{(x,\alpha)\in I_{x_0}^{\delta}\times\Theta_3} \left|\frac{\partial^2\psi}{\partial y\partial\alpha_i}(x,y,\alpha)\right| \leq C(1 + |y|^q) \quad (y \in \mathbb{R}).$$

(ii) Under Assumption 3.2.4 (ii), we assume the following three conditions:

(ii.a) There exists $\delta > 0$ and $L > 0$ such that if $0 < y_1 \leq y \leq y_2$, then

$$\left|\frac{\partial^2\psi}{\partial y\partial\alpha_i}(x,y,\alpha)\right| \leq \left|\frac{\partial^2\psi}{\partial y\partial\alpha_i}(x,y_1,\alpha)\right| + \left|\frac{\partial^2\psi}{\partial y\partial\alpha_i}(x,y_2,\alpha)\right| + L \quad \textit{for all } (x,\alpha) \in I_{x_0}^{\delta}\times\Theta_3.$$

(ii.b) There exist constants $q \geq 0$ and $\delta > 0$ such that

$$\sup_{(x,\alpha)\in I_{x_0}^{\delta}\times\Theta_3} \left|\frac{\partial^2\psi}{\partial y\partial\alpha_i}(x,y,\alpha)\right| \leq O\left(\frac{1}{|y|^q}\right) \quad \textit{as } |y| \to 0.$$

(ii.c) There exists $\delta > 0$ such that for any $C_1 > 0$ and $C_2 \geq 0$ the map

$$x \mapsto \int \sup_{\alpha\in\Theta_3} \left|\frac{\partial^2\psi}{\partial y\partial\alpha_i}(x, C_1 y + C_2, \alpha)\right| f_{\alpha_0}(y)\,\mathrm{d}y$$

takes values in $\mathbb{R}$ from $I_{x_0}^{\delta}$, and is continous on $I_{x_0}^{\delta}$.

Remark 3.2.1. *Instead of Assumptions 3.2.5 and 3.2.6, the following stronger assumptions are often used:*

$$\inf_{(x,\sigma)\in\mathbb{R}\times\bar{\Theta}_2} |b(x,\sigma)| > 0, \qquad \inf_{(x,\alpha)\in\mathbb{R}\times\bar{\Theta}_3} |c(x,\alpha)| > 0.$$

(see, e.g., *Remark 1 in Sørensen and Uchida [33]). However, the 'classical' localization argument mentioned in [33] is hard to apply for our purpose. Thus, we employ our milder assumptions and show how it works well.*

Remark 3.2.2. *Under Assumption 3.2.9,*

$$\int \frac{\partial \psi}{\partial \alpha_i}(x, c(x,\alpha_0)z, \alpha_0) f_{\alpha_0}(z)\,\mathrm{d}z = \frac{\partial}{\partial \alpha_i}\left(\int \psi\big(x, c(x,\alpha_0)z, \alpha\big) f_{\alpha_0}(z)\,\mathrm{d}z\right)_{\alpha=\alpha_0},$$

at every $x \in I^\delta_{x_0}$.

Remark 3.2.3. *Assumption 3.2.12 is given by replacing* a, b, c, ψ *with* $\frac{\partial a}{\partial \mu_i}$, $\frac{\partial b}{\partial \sigma_i}$, $\frac{\partial c}{\partial \alpha_i}$, $\frac{\partial \psi}{\partial \alpha_i}$, *respectively, in Assumptions 3.2.2, 3.2.9 and 3.2.10, and is needed for obtaining the convergence* (3.20) *of the matrix containing the second derivatives of the contrast function.*

Furthermore, we introduce the following notations:

Notation 3.2.3. *Denote*

$$\Delta X^\varepsilon_t := X^\varepsilon_t - X^\varepsilon_{t-} \quad \text{for } t > 0,$$

where $\varepsilon > 0$.

Notation 3.2.4. *Denote*

$$\Delta^n_k X^\varepsilon := X^\varepsilon_{t_k} - X^\varepsilon_{t_{k-1}}, \quad \Delta^n_k N^{\lambda_\varepsilon} := N^{\lambda_\varepsilon}_{t_k} - N^{\lambda_\varepsilon}_{t_{k-1}} \quad \text{for } k = 1, \ldots, n,$$

where $n \in \mathbb{N}$, $\varepsilon > 0$.

Notation 3.2.5. *Define random times*

$$\tau_k := \inf\{t \in [t_{k-1}, t_k] \,|\, \Delta X^\varepsilon_t \neq 0 \ \text{or}\ t = t_k\},$$
$$\eta_k := \sup\{t \in [t_{k-1}, t_k] \,|\, \Delta X^\varepsilon_t \neq 0 \ \text{or}\ t = t_{k-1}\}.$$

Notation 3.2.6. *Define events* $J^{n,\varepsilon}_{k,i}$ $(k = 1, \ldots, n,\ i = 0, 1, 2)$ *by*

$$J^{n,\varepsilon}_{k,0} := \left\{\Delta^n_k N^{\lambda_\varepsilon} = 0\right\}, \quad J^{n,\varepsilon}_{k,1} := \left\{\Delta^n_k N^{\lambda_\varepsilon} = 1\right\}, \quad J^{n,\varepsilon}_{k,2} := \left\{\Delta^n_k N^{\lambda_\varepsilon} \geq 2\right\}$$

where $n \in \mathbb{N}$, $\varepsilon > 0$.

Notation 3.2.7. *Under Assumption 3.2.8, set events* $C^{n,\varepsilon,\rho}_k$ *and* $D^{n,\varepsilon,\rho}_k$ $(k = 1, \ldots, n)$ *by*

$$C^{n,\varepsilon,\rho}_k := \begin{cases} \left\{|\Delta^n_k X^\varepsilon| \leq \frac{v_{nk}}{n^\rho}\right\} & \text{under Assumption 3.2.4 (i)}, \\ \left\{\Delta^n_k X^\varepsilon \leq \frac{v_{nk}}{n^\rho}\right\} & \text{under Assumption 3.2.4 (ii)}, \end{cases}$$
$$D^{n,\varepsilon,\rho}_k := \begin{cases} \left\{|\Delta^n_k X^\varepsilon| > \frac{v_{nk}}{n^\rho}\right\} & \text{under Assumption 3.2.4 (i)}, \\ \left\{\Delta^n_k X^\varepsilon > \frac{v_{nk}}{n^\rho}\right\} & \text{under Assumption 3.2.4 (ii)}, \end{cases}$$

where $n \in \mathbb{N}$, $\varepsilon > 0$, $\rho \in (0, 1/2)$. *Then, put*

$$C^{n,\varepsilon,\rho}_{k,i} := C^{n,\varepsilon,\rho}_k \cap J^{n,\varepsilon}_{k,i}, \quad D^{n,\varepsilon,\rho}_{k,i} := D^{n,\varepsilon,\rho}_k \cap J^{n,\varepsilon}_{k,i} \quad \text{for } k = 1, \ldots, n,\ i = 0, 1, 2,$$

where $n \in \mathbb{N}$, $\varepsilon > 0$, $\rho \in (0, 1/2)$. *Furthermore, for sufficiently small* $\delta > 0$, *we may put*

$$\tilde{C}^{n,\varepsilon,\rho}_{k,i} := C^{n,\varepsilon,\rho}_{k,i} \cap \{X^\varepsilon_t \in I^\delta_{x_0} \ \text{for all}\ t \in [0,1]\},$$
$$\tilde{D}^{n,\varepsilon,\rho}_{k,i} := D^{n,\varepsilon,\rho}_{k,i} \cap \{X^\varepsilon_t \in I^\delta_{x_0} \ \text{for all}\ t \in [0,1]\}$$

for $k = 1, \ldots, n$, $i = 0, 1, 2$.

Remark 3.2.4. *We treat (n, ε) as a directed set with a suitable order according to a convergence. For examples, when we say that $n \to \infty$, $\varepsilon \to 0$ and $\lambda_\varepsilon \to \infty$, we mean that the index set (n, ε) is a directed set in $\mathbb{N} \times (0, \infty)$ with order $\prec_1$ defined by*

$$(n_1, \varepsilon_1) \prec_1 (n_2, \varepsilon) \quad \text{if } n_1 < n_2,\ \varepsilon_1 > \varepsilon_2 \text{ and } \lambda_{\varepsilon_1} < \lambda_{\varepsilon_2},$$

and when we say that $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$ and $\lambda_\varepsilon \int_{|z| \le C/n^\rho} f_{\alpha_0}(z)\,\mathrm{d}z \to 0$ with some constants $C, \rho > 0$, we mean that the index set (n, ε) is a directed set in $\mathbb{N} \times (0, \infty)$ with order $\prec_2$ defined by

$$(n_1, \varepsilon_1) \prec_2 (n_2, \varepsilon) \quad \text{if } n_1 < n_2,\ \varepsilon_1 > \varepsilon_2,\ \lambda_{\varepsilon_1} < \lambda_{\varepsilon_2}$$
$$\text{and } \lambda_{\varepsilon_1} \int_{|z| \le \frac{C}{n_1^\rho}} f_{\alpha_0}(z)\,\mathrm{d}z > \lambda_{\varepsilon_2} \int_{|z| \le \frac{C}{n_2^\rho}} f_{\alpha_0}(z)\,\mathrm{d}z\,.$$

Needless to say, the identity map Id *from $(\{(n, \varepsilon)\}, \prec_2)$ to $(\{(n, \varepsilon)\}, \prec_1)$ is monotone, and* $\mathrm{Id}(\{(n, \varepsilon)\})$ *is cofinal in $(\{(n, \varepsilon)\}, \prec_1)$.*

Remark 3.2.5. *We can assume λ_ε does not depend on ε. In this case, we treat $\{(n, \varepsilon, \lambda)\}$ instead of $\{(n, \varepsilon)\}$ as a driected set, and we must write $X^{\varepsilon,\lambda}$, Z^λ, $\Psi_{n,\varepsilon,\lambda}$,* etc., *instead of X^ε, Z^{λ_ε}, $\Psi_{n,\varepsilon}$,* etc., *respectively. But, for simplicity, we assume λ_ε depends on ε.*

3.3 Main results

We define the following contrast function $\Psi_{n,\varepsilon}(\theta)$ after the quasi-log likelihood proposed in Shimizu and Yoshida [31]:

$$\Psi_{n,\varepsilon}(\theta) := \Psi^{(1)}_{n,\varepsilon}(\mu, \sigma) + \Psi^{(2)}_{n,\varepsilon}(\alpha) \quad \text{for } \theta = (\mu, \sigma, \alpha) \in \Theta,$$

where for $\rho \in (0, 1/2)$, $\Psi^{(1)}_{n,\varepsilon}(\mu, \sigma)$ and $\Psi^{(2)}_{n,\varepsilon}(\alpha)$ are given by using Notations 3.2.4 and 3.2.7 as the following:

$$\begin{aligned} \Psi^{(1)}_{n,\varepsilon}(\mu, \sigma) &:= -\frac{1}{n} \sum_{k=1}^{n} \left\{ \frac{\left|\Delta^n_k X^\varepsilon - \frac{1}{n} a(X^\varepsilon_{t_{k-1}}, \mu)\right|^2}{2\frac{1}{n}\left|\varepsilon b(X^\varepsilon_{t_{k-1}}, \sigma)\right|^2} + \frac{1}{2} \log |b(X^\varepsilon_{t_{k-1}}, \sigma)|^2 \right\} 1_{C^{n,\varepsilon,\rho}_k}, \\ \Psi^{(2)}_{n,\varepsilon}(\alpha) &:= \frac{1}{\lambda_\varepsilon} \sum_{k=1}^{n} \psi\left(X^\varepsilon_{t_{k-1}}, \frac{\Delta^n_k X^\varepsilon}{\varepsilon}, \alpha \right) 1_{D^{n,\varepsilon,\rho}_k} \end{aligned} \tag{3.2}$$

with

$$\psi(x, y, \alpha) := \begin{cases} \log \left| \dfrac{1}{c(x, \alpha)} f_\alpha\!\left(\dfrac{y}{c(x, \alpha)} \right) \right| & \text{if } c(x, \alpha) \neq 0 \text{ and } f_\alpha\!\left(\dfrac{y}{c(x, \alpha)} \right) > 0, \\ 0 & \text{otherwise.} \end{cases}$$

Then, the quasi-maximum likelihood estimator is given by

$$\hat{\theta}_{n,\varepsilon} := \operatorname*{argmax}_{\theta \in \Theta} \Psi_{n,\varepsilon}(\theta). \tag{3.3}$$

The goal is to show the asymptotic normality of $\hat{\theta}_{n,\varepsilon}$ when $n \to \infty$ and $\varepsilon \to 0$ at the sametime. In the sequel, we will also assume that $\lambda_\varepsilon \to \infty$ as $\varepsilon \downarrow 0$ for consistency of $\hat{\theta}_{n,\varepsilon}$. Our interest is in a situation where the number of jumps is large and the Lévy noise is small. In practice, λ_ε, the intensity of jumps, should be estimated, and it is possible by Lemma 3.4.8:

$$\lambda_\varepsilon \overset{p}{\sim} \sum_{k=1}^{n} 1_{D_k^{n,\varepsilon,\rho}} \quad \text{as } \varepsilon \downarrow 0.$$

Theorem 3.3.1. *Under Assumptions 3.2.1 to 3.2.10, take ρ as either of the following:*

(i) Under Assumption 3.2.4 (i), take $\rho \in (0, 1/2)$.

(ii) Under Assumption 3.2.4 (ii), take $\rho \in (0, \min\{1/2, 1/4q\})$, where q is the constant in Assumption 3.2.10 (ii.b).

Then,

$$\hat{\theta}_{n,\varepsilon} \xrightarrow{p} \theta_0$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$, $\lambda_\varepsilon^2/n \to 0$, $\varepsilon\lambda_\varepsilon \to 0$ and $\lambda_\varepsilon \int_{|z| \le 4v_2/c_1 n^\rho} f_{\alpha_0}(z)\,\mathrm{d}z \to 0$ with $\lim(\varepsilon^2 n)^{-1} < \infty$. Here, the constants c_1 and v_2 are taken as in Assumptions 3.2.6 and 3.2.8, respectively.

Theorem 3.3.2. *Under Assumptions 3.2.1 to 3.2.12, take ρ as either of the following:*

(i) Under Assumption 3.2.4 (i), take $\rho \in (0, 1/2)$.

(ii) Under Assumption 3.2.4 (ii), take $\rho \in (0, \min\{1/2, 1/4q\})$, where q is the constant in Assumption 3.2.10 (ii.b) and Assumption 3.2.12 (ii.b).

If $\theta_0 \in \Theta$ and I_{θ_0} is positive definite, then

$$\begin{pmatrix} \varepsilon^{-1}(\hat{\mu}_{n,\varepsilon} - \mu_0) \\ \sqrt{n}(\hat{\sigma}_{n,\varepsilon} - \sigma_0) \\ \sqrt{\lambda_\varepsilon}(\hat{\alpha}_{n,\varepsilon} - \alpha_0) \end{pmatrix} \xrightarrow{d} \mathcal{N}\big(0, I_{\theta_0}^{-1}\big)$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$, $\lambda_\varepsilon^2/n \to 0$, $\varepsilon\lambda_\varepsilon \to 0$ and $\lambda_\varepsilon \int_{|z| \le 4v_2/c_1 n^\rho} f_{\alpha_0}(z)\,\mathrm{d}z \to 0$ with $\lim(\varepsilon^2 n)^{-1} < \infty$, where

$$I_{\theta_0} := \begin{pmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_2 \end{pmatrix}$$

and

$$\begin{aligned} I_1^{ij} &:= \int_0^1 \frac{\frac{\partial a}{\partial \mu_i}\frac{\partial a}{\partial \mu_j}(x_t, \mu_0)}{|b(x_t, \mu_0)|^2}\,\mathrm{d}t && (i, j = 1, \dots, d_1), \\ I_2^{ij} &:= 2\int_0^1 \frac{\frac{\partial b}{\partial \sigma_i}\frac{\partial b}{\partial \sigma_j}(x_t, \sigma_0)}{|b(x_t, \sigma_0)|^2}\,\mathrm{d}t && (i, j = 1, \dots, d_2), \\ I_3^{ij} &:= \int_0^1 \int \frac{\partial \psi}{\partial \alpha_i}\frac{\partial \psi}{\partial \alpha_j}\big(x_t, c(x_t, \alpha_0)z, \alpha_0\big) f_{\alpha_0}(z)\,\mathrm{d}z\,\mathrm{d}t && (i, j = 1, \dots, d_4). \end{aligned} \tag{3.4}$$

Remark 3.3.1. *If $\{f_\alpha\}_{\alpha\in\Theta_3}$ is given by the class of the densities of normal distributions as in Example 3.5.1, then the range of ρ in Theorems 3.3.1 and 3.3.2 is same as in Shimizu and Yoshida [32] and Ogihara and Yoshida [25]. However, if $\{f_\alpha\}_{\alpha\in\Theta_3}$ is given by the class of the densities of gamma distributions as in Example 3.5.2, then the range of ρ is $(0,1/4)$ which is different from the range $(3/8+b,1/2)$ of ρ in Ogihara and Yoshida [25], where b is the constant defined in the equation (1) in Ogihara and Yoshida [25].*

3.4 Proofs

3.4.1 Inequalities

Lemma 3.4.1. *Under Assumptions 3.2.1 and 3.2.3, suppose that $0<\varepsilon\le 1$, $\lambda_\varepsilon\ge 1$, $\varepsilon\lambda_\varepsilon\le 1$ and $0\le s<t\le 1$. Then, for $p\ge 2$,*

$$E\left[\sup_{u\in[s,t]}|X_u^\varepsilon-X_s^\varepsilon|^p\,\middle|\,\mathcal{F}_s\right]$$
$$\le C\left\{(t-s)^p+\varepsilon^p\left((t-s)^{p/2}+\lambda_\varepsilon(t-s)+\lambda_\varepsilon^{p/2}(t-s)^{p/2}+\lambda_\varepsilon^p(t-s)^p\right)\right\}(1+|X_s^\varepsilon|^p),$$

where C depends only on p,a,b,c and f_{α_0}. In particular, when $\lambda_\varepsilon/n\le 1$ and $\lambda_\varepsilon\ge 1$, it holds for $p\ge 2$ and $k=1,\dots,n$ that

$$E\left[\sup_{t\in[t_{k-1},t_k]}\frac{\left|X_t^\varepsilon-X_{t_{k-1}}^\varepsilon\right|^p}{\varepsilon^p}\,\middle|\,\mathcal{F}_{t_{k-1}}\right]\le C\left\{\frac{1}{\varepsilon^p n^p}+\frac{1}{n^{p/2}}+\frac{\lambda_\varepsilon}{n}\right\}(1+|X_s^\varepsilon|^p),$$
$$E\left[\sup_{t\in[0,1]}|X_t^\varepsilon-x_0|^p\,\middle|\,\mathcal{F}_{t_0}\right]\le C\left\{1+\varepsilon^p\lambda_\varepsilon^p\right\}(1+|x_0|^p),$$

where C depends only on p,a,b,c and f_{α_0}.

Proof. For any $p\ge 2$, we have

$$\begin{aligned}
&\left(E\left[\sup_{u\in[s,t]}|X_u^\varepsilon-X_s^\varepsilon|^p\,\middle|\,\mathcal{F}_s\right]\right)^{1/p}\\
&\le\left(E\left[\left|\int_s^t|a(X_u^\varepsilon,\mu_0)-a(X_s^\varepsilon,\mu_0)|\,\mathrm{d}u\right|^p\,\middle|\,\mathcal{F}_s\right]\right)^{1/p}\\
&\quad+\varepsilon\left(E\left[\sup_{u\in[s,t]}\left|\int_s^u\{b(X_v^\varepsilon,\sigma_0)-b(X_s^\varepsilon,\sigma_0)\}\,\mathrm{d}W_v\right|^p\,\middle|\,\mathcal{F}_s\right]\right)^{1/p}\\
&\quad+\varepsilon\left(E\left[\sup_{u\in[s,t]}\left|\int_s^u\{c(X_v^\varepsilon,\alpha_0)-c(X_s^\varepsilon,\alpha_0)\}\,\mathrm{d}Z_v^{\lambda_\varepsilon}\right|^p\,\middle|\,\mathcal{F}_s\right]\right)^{1/p}\\
&\quad+(t-s)\,|a(X_s^\varepsilon,\mu_0)|+C\varepsilon\sqrt{t-s}\,|b(X_s^\varepsilon,\sigma_0)|+\varepsilon\,|c(X_s^\varepsilon,\alpha_0)|\left(E\left[\sup_{u\in[s,t]}\left|\int_s^u\mathrm{d}Z_v^{\lambda_\varepsilon}\right|^p\right]\right)^{1/p},
\end{aligned}\tag{3.5}$$

where C depends only on p. Then, it follows from the Lipschitz continuity of $a(\cdot,\mu_0)$ that

$$E\left[\left(\int_s^t |a(X_u^\varepsilon,\mu_0)-a(X_s^\varepsilon,\mu_0)|\,\mathrm{d}u\right)^p \,\middle|\, \mathcal{F}_s\right] \le CE\left[\left(\int_s^t |X_u^\varepsilon - X_s^\varepsilon|\,\mathrm{d}u\right)^p \,\middle|\, \mathcal{F}_s\right]$$
$$\le C(t-s)^{p-1}\int_s^t E\left[|X_u^\varepsilon - X_s^\varepsilon|^p \,|\, \mathcal{F}_s\right]\mathrm{d}u\,, \quad (3.6)$$

where C depends only on a, and it follows from the Lipschitz continuity of $b(\cdot,\sigma_0)$ and Burkholder's inequality (see, *e.g.*, Theorem 4.4.21 in Applebaum [3]) that

$$E\left[\sup_{u\in[s,t]}\left|\int_s^u \{b(X_v^\varepsilon,\sigma_0)-b(X_s^\varepsilon,\sigma_0)\}\,\mathrm{d}W_v\right|^p \,\middle|\, \mathcal{F}_s\right] \le CE\left[\left|\int_s^t |X_u^\varepsilon - X_s^\varepsilon|^2\,\mathrm{d}u\right|^{p/2} \,\middle|\, \mathcal{F}_s\right]$$
$$\le C(t-s)^{p/2-1}\int_s^t E\left[|X_u^\varepsilon - X_s^\varepsilon|^p \,|\, \mathcal{F}_s\right]\mathrm{d}u\,, \quad (3.7)$$

where C depends only on p and b, and from the Lipschitz continuity of $c(\cdot,\alpha_0)$, it is analogous to the proof of Theorem 4.4.23 in Applebaum [3] that

$$E\left[\sup_{u\in[s,t]}\left|\int_s^u \{c(X_v^\varepsilon,\alpha_0)-c(X_s^\varepsilon,\alpha_0)\}\,\mathrm{d}Z_v^{\lambda_\varepsilon}\right|^p \,\middle|\, \mathcal{F}_s\right]$$
$$\le C\left\{E\left[\left(\int_s^t\int_{\mathbb{R}} |X_u^\varepsilon - X_s^\varepsilon|^2|z|^2\lambda_\varepsilon f_{\alpha_0}(z)\,\mathrm{d}z\,\mathrm{d}u\right)^{p/2} \,\middle|\, \mathcal{F}_s\right]\right.$$
$$+ E\left[\int_s^t\int_{\mathbb{R}} |X_u^\varepsilon - X_s^\varepsilon|^p|z|^p\lambda_\varepsilon f_{\alpha_0}(z)\,\mathrm{d}z\,\mathrm{d}u \,\middle|\, \mathcal{F}_s\right]$$
$$\left. + E\left[\left(\int_s^t\int_{\mathbb{R}} |X_u^\varepsilon - X_s^\varepsilon|\,|z|\,\lambda_\varepsilon f_{\alpha_0}(z)\,\mathrm{d}z\,\mathrm{d}u\right)^{p} \,\middle|\, \mathcal{F}_s\right]\right\},$$

where C depends only on p and c. Here, we have

$$E\left[\left(\int_s^t\int_{\mathbb{R}} |X_u^\varepsilon - X_s^\varepsilon|^2\,|z|^2\,\lambda_\varepsilon f_{\alpha_0}(z)\,\mathrm{d}z\,\mathrm{d}u\right)^{p/2} \,\middle|\, \mathcal{F}_s\right]$$
$$\le C\lambda_\varepsilon^{p/2}\left(\int_s^t \left(E\left[|X_u^\varepsilon - X_s^\varepsilon|^p \,|\, \mathcal{F}_s\right]\right)^{2/p}\mathrm{d}u\right)^{p/2},$$
$$\le C\lambda_\varepsilon^{p/2}(t-s)^{p/2-1}\int_s^t E\left[|X_u^\varepsilon - X_s^\varepsilon|^p \,|\, \mathcal{F}_s\right]\mathrm{d}u\,,$$

where C depends only on p and f_{α_0}, and

$$E\left[\left(\int_s^t\int_{\mathbb{R}} |X_u^\varepsilon - X_s^\varepsilon|\,|z|\,\lambda_\varepsilon f_{\alpha_0}(z)\,\mathrm{d}z\,\mathrm{d}u\right)^{p} \,\middle|\, \mathcal{F}_s\right]$$
$$\le C\lambda_\varepsilon^{p}\left(\int_s^t \left(E\left[|X_u^\varepsilon - X_s^\varepsilon|^p \,|\, \mathcal{F}_s\right]\right)^{1/p}\mathrm{d}u\right)^{p}$$
$$\le C\lambda_\varepsilon^{p}(t-s)^{p-1}\int_s^t E\left[|X_u^\varepsilon - X_s^\varepsilon|^p \,|\, \mathcal{F}_s\right]\mathrm{d}u\,,$$

where C depends only on p and f_{α_0}. Thus,

$$E\left[\sup_{u\in[s,t]}\left|\int_s^u (c(X_v^\varepsilon,\alpha_0)-c(X_s^\varepsilon,\alpha_0))\,\mathrm{d}Z_v^{\lambda_\varepsilon}\right|^p \,\middle|\, \mathcal{F}_s\right] \leq C\left(\lambda_\varepsilon^{p/2}(t-s)^{p/2-1}+\lambda_\varepsilon+\lambda_\varepsilon^p(t-s)^{p-1}\right)\int_s^t E\left[|X_u^\varepsilon-X_s^\varepsilon|^p \,|\, \mathcal{F}_s\right]\mathrm{d}u\,, \quad (3.8)$$

where C depends only on p, c and f_{α_0}. By using the Burkholder-Davis-Gundy inequality,

$$E\left[\sup_{u\in[s,t]}\left|\int_s^u \mathrm{d}Z_v^{\lambda_\varepsilon}\right|^p\right] \leq C\left(\lambda_\varepsilon^{p/2}(t-s)^{p/2}+\lambda_\varepsilon(t-s)+\lambda_\varepsilon^p(t-s)^p\right), \quad (3.9)$$

where C depends only on p and f_{α_0}. From (3.5), (3.6), (3.7), (3.8) and (3.9),

$$\begin{aligned}
&E\left[\sup_{u\in[s,t]}|X_u^\varepsilon-X_s^\varepsilon|^p \,\middle|\, \mathcal{F}_s\right] \\
&\quad\leq C\Bigg\{\left((t-s)^{p-1}+\varepsilon^p(t-s)^{\frac{p-2}{2}}+\varepsilon^p\left(\lambda_\varepsilon+\lambda_\varepsilon^{\frac{p}{2}}(t-s)^{\frac{p-2}{2}}+\lambda_\varepsilon^p(t-s)^{p-1}\right)\right) \\
&\qquad\times\int_s^t E\left[|X_u^\varepsilon-X_s^\varepsilon|^p \,|\, \mathcal{F}_s\right]\mathrm{d}u \\
&\qquad+(t-s)^p\,|a(X_s^\varepsilon,\mu_0)|^p+\varepsilon^p(t-s)^{p/2}\,|b(X_s^\varepsilon,\sigma_0)|^p \\
&\qquad+\varepsilon^p\left(\lambda_\varepsilon(t-s)+\lambda_\varepsilon^{p/2}(t-s)^{p/2}+\lambda_\varepsilon^p(t-s)^p\right)|c(X_s^\varepsilon,\alpha_0)|^p\Bigg\}\,,
\end{aligned}$$

where C depends only on p, a, b, c and f_{α_0}. By Gronwall's inequality,

$$\begin{aligned}
&E\left[\sup_{u\in[s,t]}|X_u^\varepsilon-X_s^\varepsilon|^p \,\middle|\, \mathcal{F}_s\right] \\
&\leq C\left\{(t-s)^p\,|a(X_s^\varepsilon,\mu_0)|^p+\varepsilon^p(t-s)^{p/2}\,|b(X_s^\varepsilon,\sigma_0)|^p\right. \\
&\qquad\left.+\varepsilon^p\left(\lambda_\varepsilon(t-s)+\lambda_\varepsilon^{p/2}(t-s)^{p/2}+\lambda_\varepsilon^p(t-s)^p\right)|c(X_s^\varepsilon,\alpha_0)|^p\right\} \\
&\quad\times\exp\left(C\left\{(t-s)^p+\varepsilon^p(t-s)^{p/2}+\varepsilon^p\lambda_\varepsilon(t-s)+\varepsilon^p\lambda_\varepsilon^{p/2}(t-s)^{p/2}+\varepsilon^p\lambda_\varepsilon^p(t-s)^p\right\}\right).
\end{aligned}$$

This implies the conclusion. □

Lemma 3.4.2. *Under Assumptions 3.2.1 and 3.2.3, suppose that* $0<\varepsilon\leq 1$, $\lambda_\varepsilon\geq 1$, $\varepsilon\lambda_\varepsilon\leq 1$ *and* $0\leq s<t\leq 1$. *Then, for* $p\geq 2$

$$E\left[\sup_{u\in[s,t]}|X_u^\varepsilon-x_u|^p \,\middle|\, \mathcal{F}_s\right] \leq C\varepsilon^p\left((t-s)^{p/2}+\lambda_\varepsilon(t-s)+\lambda_\varepsilon^{p/2}(t-s)^{p/2}+\lambda_\varepsilon^p(t-s)^p\right),$$

where C *depends only on* p, a *and* b.

Proof. Same as the proof of Lemma 3.4.1, for any $p \geq 2$, we obtain

$$\begin{aligned} &E\left[\sup_{u\in[s,t]} |X_u^\varepsilon - x_u|^p \,\middle|\, \mathcal{F}_s\right] \\ &\qquad \leq C\varepsilon^p \left((t-s)^{p/2} + \lambda_\varepsilon(t-s) + \lambda_\varepsilon^{p/2}(t-s)^{p/2} + \lambda_\varepsilon^p(t-s)^p\right) \\ &\qquad\quad \times \exp\left(C\left\{(t-s)^p + \varepsilon^p\left((t-s)^{p/2} + \lambda_\varepsilon(t-s) + \lambda_\varepsilon^{p/2}(t-s)^{p/2} + \lambda_\varepsilon^p(t-s)^p\right)\right\}\right), \end{aligned}$$

where C depends only on p, a, b, c and f_{α_0}. □

Lemma 3.4.3. *Under Assumptions 3.2.1 and 3.2.3, for* $p \geq 1$

$$\|X_\cdot^\varepsilon - x_\cdot\|_{L^p(\Omega;L^\infty([0,1]))} = O(\varepsilon\lambda_\varepsilon)$$

as $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$ *and* $\varepsilon\lambda_\varepsilon \to 0$, *and*

$$\left\| \sup_{\substack{0\leq u,s\leq 1 \\ |u-s|\leq 1/n}} |X_u^\varepsilon - x_s| \right\|_{L^p(\Omega)} = O(1/n + \varepsilon\lambda_\varepsilon)$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$ *and* $\varepsilon\lambda_\varepsilon \to 0$.

Proof. Both rates of convergence are obtained immediately from Lemma 3.4.2. □

Lemma 3.4.4. *Under Assumptions 3.2.1 and 3.2.3, suppose that a family* $\{g(\cdot,\theta)\}_{\theta\in\Theta}$ *of functions from* $\mathbb{R}$ *to* $\mathbb{R}$ *is equicontinuous at every points in* I_{x_0}. *Then,*

$$\frac{1}{n}\sum_{k=1}^n g\Big(X_{t_{k-1}}^\varepsilon, \theta\Big) \xrightarrow{p} \int_0^1 g(x_t,\theta)\,\mathrm{d}t$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$ *and* $\varepsilon\lambda_\varepsilon \to 0$, *uniformly in* $\theta \in \Theta$.

Proof. This follows from Lemmas 3.4.3 and 3.7.2. □

Lemma 3.4.5. *Under Assumptions 3.2.1 and 3.2.3 with Notation 3.2.5, suppose that* $0 < \varepsilon \leq 1$, $\lambda_\varepsilon \geq 1$ *and* $\varepsilon\lambda_\varepsilon \leq 1$. *Then, for any* $p \in [1,\infty)$,

$$E\left[\sup_{t\in[t_{k-1},\tau_k)} \left|X_t^\varepsilon - X_{t_{k-1}}^\varepsilon\right|^p \,\middle|\, \mathcal{F}_{t_{k-1}}\right] \leq C\left(\frac{1}{n^p} + \frac{\varepsilon^p}{n^{p/2}}\right)\left(1 + |X_{t_{k-1}}^\varepsilon|^p\right),$$

where C *depends only on* p, a *and* b, *and*

$$E\left[\sup_{t\in[\eta_k,t_k]} \left|X_t^\varepsilon - X_{t_k}^\varepsilon\right|^p \,\middle|\, \mathcal{F}_{t_{k-1}}\right] \leq C\left(\frac{1}{n^p} + \frac{\varepsilon^p}{n^{p/2}}\right)\left(1 + |X_{t_{k-1}}^\varepsilon|^p\right),$$

where C *depends only on* p, a, b, c *and* f_{α_0}.

Proof. For $t \in [t_{k-1}, \tau_k)$ and $p \geq 2$,

$$\left(E \left[\sup_{s \in [t_{k-1}, t)} \left| X_s^\varepsilon - X_{t_{k-1}}^\varepsilon \right|^p \middle| \mathcal{F}_{t_{k-1}} \right] \right)^{1/p}$$
$$\leq C \int_{t_{k-1}}^t \left(E \left[\left| X_s^\varepsilon - X_{t_{k-1}}^\varepsilon \right|^p \middle| \mathcal{F}_{t_{k-1}} \right] \right)^{1/p} \mathrm{d}s + \frac{1}{n} \left| a(X_{t_{k-1}}^\varepsilon, \mu_0) \right|$$
$$+ C\varepsilon \left(\int_{t_{k-1}}^t \left(E \left[\left| X_s^\varepsilon - X_{t_{k-1}}^\varepsilon \right|^p \middle| \mathcal{F}_{t_{k-1}} \right] \right)^{2/p} \mathrm{d}s \right)^{1/2} + \frac{\varepsilon}{\sqrt{n}} \left| b(X_{t_{k-1}}^\varepsilon, \sigma_0) \right|,$$

where C depnds only on p, a and b. By using Gronwall's inequality, we obtain

$$\left(E \left[\sup_{s \in [t_{k-1}, t)} \left| X_s^\varepsilon - X_{t_{k-1}}^\varepsilon \right|^p \middle| \mathcal{F}_{t_{k-1}} \right] \right)^{2/p} \leq C e^{C(1/n + \varepsilon^2)t} \left(\frac{1}{n^2} + \frac{\varepsilon^2}{n} \right) (1 + |X_{t_{k-1}}^\varepsilon|^2),$$

where C depnds only on p, a and b. Similarly,

$$\left(E \left[\sup_{s \in [\eta_k, t_k]} |X_s^\varepsilon - X_{t_k}^\varepsilon|^p \middle| \mathcal{F}_{t_{k-1}} \right] \right)^{2/p} \leq C \left(\frac{1}{n^2} + \frac{\varepsilon^2}{n} \right) \left(1 + E \left[|X_{t_k}^\varepsilon|^2 \,\middle|\, \mathcal{F}_{t_{k-1}} \right] \right),$$

where C depnds only on p, a and b. From Lemma 3.4.1, we have

$$E \left[\sup_{u, s \in [t_{k-1}, t_k]} |X_u^\varepsilon - X_s^\varepsilon|^p \middle| \mathcal{F}_{t_{k-1}} \right] \leq C \left(\frac{1}{n^p} + \varepsilon^p \frac{\lambda_\varepsilon}{n} \right) \left(1 + |X_{t_{k-1}}^\varepsilon|^p \right),$$

where C depnds only on p, a, b, c and f_{α_0}. We can easily extend this result to the case $p \in [1, 2)$ by using Hölder inequality. □

Lemma 3.4.6. *Under Assumptions 3.2.1 and 3.2.3, suppose that* $0 < \varepsilon \leq 1$, $\lambda_\varepsilon \geq 1$, $\varepsilon \lambda_\varepsilon \leq 1$. *Let*

$$Y_k^\varepsilon := \sup_{t \in [t_{k-1}, \tau_k)} \frac{|X_t^\varepsilon - X_{t_{k-1}}^\varepsilon|}{\varepsilon} + \sup_{t \in [\eta_k, t_k]} \frac{|X_t^\varepsilon - X_{t_k}^\varepsilon|}{\varepsilon}.$$

Then, for any $p \in (2, \infty)$,

$$\sup_{k=1,\ldots,n} Y_k^\varepsilon = O_p \left(\frac{1}{\varepsilon n^{1-1/p}} + \frac{1}{n^{1/2-1/p}} \right)$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$ *and* $\varepsilon \lambda_\varepsilon \to 0$.

Proof. By using Lemmas 3.4.4 and 3.4.5, we have

$$\sum_{k=1}^n E \left[|Y_k^\varepsilon|^p \,\middle|\, \mathcal{F}_{t_{k-1}} \right] \leq C \left(\frac{n}{(\varepsilon n)^p} + \frac{n}{n^{p/2}} \right) \frac{1}{n} \sum_{k=1}^n \left(1 + \left| X_{t_{k-1}}^\varepsilon \right|^p \right) = O_p \left(\frac{n}{(\varepsilon n)^p} + \frac{n}{n^{p/2}} \right)$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$ and $\varepsilon \lambda_\varepsilon \to 0$. It follows from Lemma 3.7.3 that

$$\sup_{k=1,\ldots,n} |Y_k^\varepsilon| \leq \left(\sum_{k=1}^n |Y_k^\varepsilon|^p \right)^{1/p} = O_p \left(\frac{1}{\varepsilon n^{1-1/p}} + \frac{1}{n^{1/2-1/p}} \right).$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$ and $\varepsilon \lambda_\varepsilon \to 0$. □

3.4.2 Limit theorems

We make a version of Lemma 2.2 in Shimizu and Yoshida [31] in the sequel.

Lemma 3.4.7. *Under Assumptions 3.2.1, 3.2.3, 3.2.6 and 3.2.8 with Notations 3.2.3 to 3.2.5 and 3.2.7, suppose that* $0 < \varepsilon \le 1$*,* $\lambda_\varepsilon \ge 1$ *and* $\varepsilon\lambda_\varepsilon \le 1$*. Then, for* $p \ge 2$ *and* $\rho \in (0, 1/2)$

$$\begin{aligned}
&P\left[C_{k,0}^{n,\varepsilon,\rho} \,\middle|\, \mathcal{F}_{t_{k-1}}\right] \ge e^{-\lambda_\varepsilon/n}\left\{1 - C\left(\frac{1}{n^{p(1-\rho)}} + \frac{\varepsilon^p}{n^{p(1/2-\rho)}}\right)\left(1 + |X_{t_{k-1}}^\varepsilon|^p\right)\right\}\\
&P\left[D_{k,0}^{n,\varepsilon,\rho} \,\middle|\, \mathcal{F}_{t_{k-1}}\right] \le C\left(\frac{1}{n^{p(1-\rho)}} + \frac{\varepsilon^p}{n^{p(1/2-\rho)}}\right)\left(1 + |X_{t_{k-1}}^\varepsilon|^p\right)\\
&P\left[C_{k,1}^{n,\varepsilon,\rho} \,\middle|\, \mathcal{F}_{t_{k-1}}\right]\\
&\quad \le \frac{\lambda_\varepsilon}{n}\left\{C\left(\frac{1}{n^{p(1-\rho)}} + \frac{\varepsilon^p}{n^{p(1/2-\rho)}} + \frac{\varepsilon^p\lambda_\varepsilon}{n}\right)\left(1 + |X_{t_{k-1}}^\varepsilon|^p\right) + \int_{|z|\le 4v_2/c_1 n^\rho} f_{\alpha_0}(z)\,\mathrm{d}z\right\},\\
&P\left[D_{k,1}^{n,\varepsilon,\rho} \,\middle|\, \mathcal{F}_{t_{k-1}}\right] \le \frac{\lambda_\varepsilon}{n}\left\{C\left(\frac{1}{n^{p(1-\rho)}} + \frac{\varepsilon^p}{n^{p(1/2-\rho)}} + \frac{\varepsilon^p\lambda_\varepsilon}{n}\right)\left(1 + |X_{t_{k-1}}^\varepsilon|^p\right) + 1\right\},\\
&P\left[C_{k,2}^{n,\varepsilon,\rho} \,\middle|\, \mathcal{F}_{t_{k-1}}\right] \le \frac{\lambda_\varepsilon^2}{n^2}, \qquad P\left[D_{k,2}^{n,\varepsilon,\rho} \,\middle|\, \mathcal{F}_{t_{k-1}}\right] \le \frac{\lambda_\varepsilon^2}{n^2},
\end{aligned}$$

where $c_1 := \inf_{t\in[0,1]}|c(x_t,\alpha_0)| > 0$*,* $c_2 := \sup_{t\in[0,1]}|c(x,\alpha_0)|$*, and* C *depends only on* p*,* a*,* b*,* c*,* f_{α_0} *and* v_1*.*

Proof. We only give a proof for the case (i) in Assumption 3.2.4, because the same argument still works under the case (ii) in Assumption 3.2.4-. Same as in the proof of Lemma 2.2 in Shimizu and Yoshida [32, Section 4.2], it follows that

$$P\left[C_{k,2}^{n,\varepsilon,\rho} \,\middle|\, \mathcal{F}_{t_{k-1}}\right],\ P\left[D_{k,2}^{n,\varepsilon,\rho} \,\middle|\, \mathcal{F}_{t_{k-1}}\right] \le \frac{\lambda_\varepsilon^2}{n^2}.$$

Also, it follows from

$$\begin{aligned}
&P\left[\left|X_{t_k}^\varepsilon - X_{\tau_k}^\varepsilon + \Delta X_{\tau_k}^\varepsilon + X_{\tau_k-}^\varepsilon - X_{t_{k-1}}^\varepsilon\right| \le \frac{v_{nk}}{n^\rho} \,\middle|\, \mathcal{F}_{t_{k-1}},\ \Delta_k^n N^{\lambda_\varepsilon} = 1\right]\\
&\quad \le P\left[\left|X_{t_k}^\varepsilon - X_{\tau_k}^\varepsilon\right| + \sup_{t\in[t_{k-1},\tau_k)}\left|X_t^\varepsilon - X_{t_{k-1}}^\varepsilon\right| > \frac{v_{nk}}{n^\rho} \,\middle|\, \mathcal{F}_{t_{k-1}},\ \Delta_k^n N^{\lambda_\varepsilon} = 1\right]\\
&\qquad + P\left[\left|\Delta Z_{\tau_k}^{\lambda_\varepsilon}\right| \le \frac{4v_{nk}}{c_1 n^\rho} \text{ or } \sup_{t\in[t_k,t_{k-1}]}|c(X_t^\varepsilon,\alpha_0) - c(x_t,\alpha_0)| > \frac{c_1}{2} \,\middle|\, \mathcal{F}_{t_{k-1}},\ \Delta_k^n N^{\lambda_\varepsilon} = 1\right],\\
&P\left[\left|X_{t_k}^\varepsilon - X_{\tau_k}^\varepsilon + \Delta X_{\tau_k}^\varepsilon + X_{\tau_k-}^\varepsilon - X_{t_{k-1}}^\varepsilon\right| > \frac{v_{nk}}{n^\rho} \,\middle|\, \mathcal{F}_{t_{k-1}},\ \Delta_k^n N^{\lambda_\varepsilon} = 1\right]\\
&\quad \le P\left[\left|X_{t_k}^\varepsilon - X_{\tau_k}^\varepsilon\right| + \sup_{t\in[t_{k-1},\tau_k)}\left|X_t^\varepsilon - X_{t_{k-1}}^\varepsilon\right| > \frac{v_{nk}}{2n^\rho} \,\middle|\, \mathcal{F}_{t_{k-1}},\ \Delta_k^n N^{\lambda_\varepsilon} = 1\right]\\
&\qquad + P\left[\left|\Delta Z_{\tau_k}^{\lambda_\varepsilon}\right| > \frac{v_{nk}}{4c_2 n^\rho} \text{ or } \sup_{t\in[t_k,t_{k-1}]}|c(X_t^\varepsilon,\alpha_0) - c(x_t,\alpha_0)| > c_2 \,\middle|\, \mathcal{F}_{t_{k-1}},\ \Delta_k^n N^{\lambda_\varepsilon} = 1\right]
\end{aligned}$$

and Lemmas 3.4.2, 3.4.5 and 3.7.2 that

$$P\left[C_{k,1}^{n,\varepsilon,\rho} \,\middle|\, \mathcal{F}_{t_{k-1}}\right] \leq \frac{\lambda_\varepsilon}{n} e^{-\lambda_\varepsilon/n} \left\{ C \left(\frac{1}{n^{p(1-\rho)}} + \frac{\varepsilon^p}{n^{p(1/2-\rho)}} \right) \left(1 + |X_{t_{k-1}}^\varepsilon|^p\right) + \int_{|z|\leq 4v_{nk}/c_1 n^\rho} f_{\alpha_0}(z)\,\mathrm{d}z + C\frac{\varepsilon^p \lambda_\varepsilon}{n} \right\},$$

$$P\left[D_{k,1}^{n,\varepsilon,\rho} \,\middle|\, \mathcal{F}_{t_{k-1}}\right] \leq \frac{\lambda_\varepsilon}{n} e^{-\lambda_\varepsilon/n} \left\{ C \left(\frac{1}{n^{p(1-\rho)}} + \frac{\varepsilon^p}{n^{p(1/2-\rho)}} \right) \left(1 + |X_{t_{k-1}}^\varepsilon|^p\right) + \int_{|z|> v_{nk}/4c_2 n^\rho} f_{\alpha_0}(z)\,\mathrm{d}z + C\frac{\varepsilon^p \lambda_\varepsilon}{n} \right\},$$

where C depends only on p, a, b, c, f_{α_0} and v_1. The other inequalities follow from Lemma 3.4.5. □

In the proof of Proposition 3.3 (ii) in Shimizu and Yoshida [31], the intensity of the Poisson process driving on the background is constant, although we assume the intensity λ_ε goes to infinity. So, we prepare the following lemma.

Lemma 3.4.8. *Under Assumptions 3.2.1, 3.2.3, 3.2.4, 3.2.6 and 3.2.8, for $\rho \in (0, 1/2)$*

$$\frac{1}{\lambda_\varepsilon} \sum_{k=1}^n 1_{D_k^{n,\varepsilon,\rho}} \xrightarrow{p} 1,$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$, $\lambda_\varepsilon/n \to 0$ and $\varepsilon\lambda_\varepsilon \to 0$. More precisely, for $\rho \in (0, 1/2)$ and $p \in [2/(1-2\rho), \infty)$

$$\frac{1}{\lambda_\varepsilon} \sum_{k=1}^n 1_{D_{k,0}^{n,\varepsilon,\rho}} = O_p\left(\frac{1}{\lambda_\varepsilon n^{p(1-\rho)-1}} + \frac{\varepsilon^p}{\lambda_\varepsilon n^{p(1/2-\rho)-1}} \right),$$
$$\frac{1}{\lambda_\varepsilon} \sum_{k=1}^n 1_{D_{k,1}^{n,\varepsilon,\rho}} = 1 + O_p\left(\frac{\lambda_\varepsilon}{n} + \frac{1}{n^{p(1-\rho)}} + \frac{\varepsilon^p}{n^{p(1/2-\rho)}} + \int_{|z|\leq 4v_{nk}/c_1 n^\rho} f_{\alpha_0}(z)\,\mathrm{d}z \right),$$
$$\frac{1}{\lambda_\varepsilon} \sum_{k=1}^n 1_{D_{k,2}^{n,\varepsilon,\rho}} = O_p\left(\frac{\lambda_\varepsilon}{n}\right)$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$ and $\varepsilon\lambda_\varepsilon \to 0$.

Proof. Since

$$\left| \frac{\lambda_\varepsilon}{n} - P\left[D_{k,1}^{n,\varepsilon,\rho} \,\middle|\, \mathcal{F}_{t_{k-1}}\right] \right| \leq \frac{\lambda_\varepsilon}{n} - \frac{\lambda_\varepsilon}{n} e^{-\lambda_\varepsilon/n} + \left| \frac{\lambda_\varepsilon}{n} e^{-\lambda_\varepsilon/n} - P\left[D_{k,1}^{n,\varepsilon,\rho} \,\middle|\, \mathcal{F}_{t_{k-1}}\right] \right|$$
$$\leq \left(\frac{\lambda_\varepsilon}{n}\right)^2 + P\left[C_{k,1}^{n,\varepsilon,\rho} \,\middle|\, \mathcal{F}_{t_{k-1}}\right],$$

it follows from Lemmas 3.4.4 and 3.4.7 that for $p \geq 2$ and $\rho \in (0, 1/2)$

$$\sum_{k=1}^{n} E\left[\left|\frac{1}{\lambda_\varepsilon} 1_{D_{k,1}^{n,\varepsilon,\rho}} - \frac{1}{n}\right| \,\middle|\, \mathcal{F}_{t_{k-1}}\right] \leq \frac{\lambda_\varepsilon}{n} + \frac{1}{\lambda_\varepsilon} \sum_{k=1}^{n} P\left[C_{k,1}^{n,\varepsilon,\rho} \,\middle|\, \mathcal{F}_{t_{k-1}}\right]$$
$$\leq \frac{\lambda_\varepsilon}{n} + C\left(\frac{1}{n^{p(1-\rho)}} + \frac{\varepsilon^p}{n^{p(1/2-\rho)}} + \frac{\varepsilon^p \lambda_\varepsilon}{n}\right) \frac{1}{n} \sum_{k=1}^{n} \left(1 + |X_{t_{k-1}}^\varepsilon|^p\right) + \int_{|z| \leq 4v_2/c_1 n^\rho} f_{\alpha_0}(z)\, \mathrm{d}z \xrightarrow{p} 0$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$, $\lambda_\varepsilon/n \to 0$ and $\varepsilon\lambda_\varepsilon \to 0$. Similarly, we obtain

$$\sum_{k=1}^{n} E\left[\frac{1}{\lambda_\varepsilon} 1_{D_{k,0}^{n,\varepsilon,\rho}} \,\middle|\, \mathcal{F}_{t_{k-1}}\right] \leq C \frac{1}{\lambda_\varepsilon} \left(\frac{1}{n^{p(1-\rho)}} + \frac{\varepsilon^p}{n^{p(1/2-\rho)}}\right) \sum_{k=1}^{n} \left(1 + |X_{t_{k-1}}^\varepsilon|^p\right),$$
$$\sum_{k=1}^{n} E\left[\frac{1}{\lambda_\varepsilon} 1_{D_{k,2}^{n,\varepsilon,\rho}} \,\middle|\, \mathcal{F}_{t_{k-1}}\right] \leq \frac{\lambda_\varepsilon}{n}.$$

Hence, the conclusion follows from Lemma 3.7.3. □

Lemma 3.4.9. *Under Assumptions 3.2.1, 3.2.3, 3.2.4, 3.2.6 and 3.2.8, for* $\rho \in (0, 1/2)$

$$\frac{1}{n} \sum_{k=1}^{n} 1_{C_k^{n,\varepsilon,\rho}} \xrightarrow{p} 1,$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$, $\lambda_\varepsilon/n \to 0$ *and* $\varepsilon\lambda_\varepsilon \to 0$. *More precisely, for* $\rho \in (0, 1/2)$ *and* $p \in [2, \infty)$

$$\frac{1}{n} \sum_{k=1}^{n} 1_{C_{k,0}^{n,\varepsilon,\rho}} = 1 + O_p\left(\frac{\lambda_\varepsilon}{n}\right),$$
$$\frac{1}{n} \sum_{k=1}^{n} 1_{C_{k,1}^{n,\varepsilon,\rho}} = O_p\left(\frac{\lambda_\varepsilon}{n^{p(1-\rho)+1}} + \frac{\varepsilon^p \lambda_\varepsilon}{n^{p(1/2-\rho)+1}} + \frac{\lambda_\varepsilon}{n} \int_{|z| \leq 4v_{nk}/c_1 n^\rho} f_{\alpha_0}(z)\, \mathrm{d}z\right),$$
$$\frac{1}{n} \sum_{k=1}^{n} 1_{C_{k,2}^{n,\varepsilon,\rho}} = O_p\left(\frac{\lambda_\varepsilon^2}{n^2}\right)$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$, $\lambda_\varepsilon/n \to 0$ *and* $\varepsilon\lambda_\varepsilon \to 0$.

Proof. From Lemma 3.4.8 we have

$$\frac{1}{n} \sum_{k=1}^{n} 1_{C_k^{n,\varepsilon,\rho}} - 1 = \frac{\lambda_\varepsilon}{n} \frac{1}{\lambda_\varepsilon} \sum_{k=1}^{n} 1_{D_k^{n,\varepsilon,\rho}} = O_p\left(\frac{\lambda_\varepsilon}{n}\right)$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$, $\lambda_\varepsilon/n \to 0$ and $\varepsilon\lambda_\varepsilon \to 0$. It follows from Lemmas 3.4.4 and 3.4.7

that for any $p \in [0,\infty)$

$$\sum_{k=1}^n E\left[\left|\frac{1}{n}1_{C_{k,1}^{n,\varepsilon,\rho}}\right|\,\middle|\,\mathcal{F}_{t_{k-1}}\right] = \frac{1}{n}\sum_{k=1}^n P\left[C_{k,1}^{n,\varepsilon,\rho}\,\middle|\,\mathcal{F}_{t_{k-1}}\right]$$
$$\le C\frac{\lambda_\varepsilon}{n}\left(\frac{1}{n^{p(1-\rho)}} + \frac{\varepsilon^p}{n^{p(1/2-\rho)}} + \frac{\varepsilon^p\lambda_\varepsilon}{n}\right)\frac{1}{n}\sum_{k=1}^n\left(1+|X_{t_{k-1}}^\varepsilon|^p\right) + \frac{\lambda_\varepsilon}{n}\int_{|z|\le 4v_{nk}/c_1n^\rho} f_{\alpha_0}(z)\,\mathrm{d}z\,,$$
$$\sum_{k=1}^n E\left[\left|\frac{1}{n}1_{C_{k,2}^{n,\varepsilon,\rho}}\right|\,\middle|\,\mathcal{F}_{t_{k-1}}\right] = \frac{1}{n}\sum_{k=1}^n P\left[C_{k,2}^{n,\varepsilon,\rho}\,\middle|\,\mathcal{F}_{t_{k-1}}\right] \le \frac{\lambda_\varepsilon^2}{n^2}.$$

The conclusion follows from Lemma 3.7.3. □

Remark 3.4.1. *From this lemma, under Assumptions 3.2.1, 3.2.3, 3.2.4, 3.2.6 and 3.2.8, for $\rho \in (0,1/2)$ and for any random variables $\xi_{k,\theta}^{n,\varepsilon}$ $(k=1,\dots,n,\ n\in\mathbb{N},\ \varepsilon>0,\ \theta\in\bar\Theta)$, when*

$$\lambda_\varepsilon\to\infty,\quad \varepsilon\lambda_\varepsilon\to 0,\quad \frac{\lambda_\varepsilon^2}{n}\to 0,\quad \lambda_\varepsilon\int_{|z|\le 4v_2/c_1n^\rho} f_{\alpha_0}(z)\,\mathrm{d}z\to 0$$

as $\varepsilon\to 0$,

$$\sum_{k=1}^n \xi_{k,\theta}^{n,\varepsilon}\left\{1_{C_k^{n,\varepsilon,\rho}} - 1_{C_{k,0}^{n,\varepsilon,\rho}}\right\} = o_p(1)$$

as $n\to\infty$ and $\varepsilon\to 0$, uniformly in $\theta\in\Theta$, since for any $\eta>0$

$$P\left(\sup_{\theta\in\Theta}\left|\sum_{k=1}^n \xi_{k,\theta}^{n,\varepsilon}1_{C_{k,j}^{n,\varepsilon,\rho}}\right| > \eta\right) \le P\left(\left|\sum_{k=1}^n 1_{C_{k,j}^{n,\varepsilon,\rho}}\right| > 1/2\right) \quad \text{for } j=1,2.$$

Similarly, from Lemma 3.4.8, when

$$\lambda_\varepsilon\to\infty,\quad \varepsilon\lambda_\varepsilon\to 0,\quad \frac{\lambda_\varepsilon^2}{n}\to 0$$

as $\varepsilon\to 0$,

$$\sum_{k=1}^n \xi_{k,\theta}^{n,\varepsilon}\left\{1_{D_k^{n,\varepsilon,\rho}} - 1_{D_{k,1}^{n,\varepsilon,\rho}}\right\} = o_p(1)$$

as $n\to\infty$ and $\varepsilon\to 0$, uniformly in $\theta\in\Theta$,

Lemma 3.4.10. *Under Assumptions 3.2.1, 3.2.3, 3.2.4, 3.2.6 and 3.2.8, let $\rho\in(0,1/2)$, $\delta>0$ and $\tilde D_{k,1}^{n,\varepsilon,\rho}$ be an event defined by*

$$\tilde D_{k,1}^{n,\varepsilon,\rho} := D_{k,1}^{n,\varepsilon,\rho}\cap\{X_t^\varepsilon\in I_{x_0}^\delta \text{ for all } t\in[0,1]\},$$

and let $\xi_{k,\theta}^{n,\varepsilon}$ $(k=1,\dots,n,\ n\in\mathbb{N},\ \varepsilon>0,\ \theta\in\bar\Theta)$ be random variables. If

$$\lambda_\varepsilon\to\infty,\quad \varepsilon\lambda_\varepsilon\to 0,\quad \frac{\lambda_\varepsilon^2}{n}\to 0,\quad \lambda_\varepsilon\int_{|z|\le 4v_2/c_1n^\rho} f_{\alpha_0}(z)\,\mathrm{d}z\to 0 \tag{3.10}$$

as $\varepsilon \to 0$, then

$$\sum_{k=1}^{n} \xi_{k,\theta}^{n,\varepsilon} \left\{ 1_{D_k^{n,\varepsilon,\rho}} - 1_{\tilde{D}_{k,1}^{n,\varepsilon,\rho}} \right\} = o_p(1), \quad \sum_{k=1}^{n} \xi_{k,\theta}^{n,\varepsilon} \left\{ 1_{\tilde{D}_{k,1}^{n,\varepsilon,\rho}} - 1_{J_{k,1}^{n,\varepsilon}} \right\} = o_p(1)$$

as $n \to \infty$ and $\varepsilon \to 0$, uniformly in $\theta \in \Theta$.

Proof. Since from Lemma 3.4.3

$$P\left(X_t^\varepsilon \in I_{x_0}^\delta \text{ for all } t \in [0,1]\right) \geq P\left(\sup_{t\in[0,1]} |X_t^\varepsilon - x_t| \leq \delta\right) \to 1$$

as $n \to \infty$, $\varepsilon \to 0$ and $\varepsilon\lambda_\varepsilon \to 0$, for any $\eta > 0$

$$\begin{aligned} &P\left(\left|\sum_{k=1}^{n} \xi_{k,\theta}^{n,\varepsilon} \left\{ 1_{D_k^{n,\varepsilon,\rho}} - 1_{\tilde{D}_{k,1}^{n,\varepsilon,\rho}} \right\}\right| > \eta\right) \\ &\quad \leq P\left(\left|\sum_{k=1}^{n} \xi_{k,\theta}^{n,\varepsilon} \left\{ 1_{D_k^{n,\varepsilon,\rho}} - 1_{D_{k,1}^{n,\varepsilon,\rho}} \right\}\right| > \eta/2\right) + P\left(\{X_t^\varepsilon \notin I_{x_0}^\delta,\ \exists t \in [0,1]\}\right), \\ &P\left(\left|\sum_{k=1}^{n} \xi_{k,\theta}^{n,\varepsilon} \left\{ 1_{\tilde{D}_{k,1}^{n,\varepsilon,\rho}} - 1_{J_{k,1}^{n,\varepsilon}} \right\}\right| > \eta\right) \\ &\quad \leq P\left(\{X_t^\varepsilon \notin I_{x_0}^\delta,\ \exists t \in [0,1]\}\right) + P\left(\left|\sum_{k=1}^{n} \xi_{k,\theta}^{n,\varepsilon} 1_{C_{k,1}^{n,\varepsilon,\rho}}\right| > \eta/2\right). \end{aligned}$$

Take sufficiently large $p \in [2/(1-2\rho), \infty)$. Thus, we obtain from Remark 3.4.1 the conclusion. □

Remark 3.4.2. *In this lemma, if $\left\{\xi_{k,\theta}^{n,\varepsilon}\right\}_{n,\varepsilon,k,\theta}$ is bounded in probability, we can replace the condition (3.10) with a milder condition*

$$\lambda_\varepsilon \to \infty, \quad \varepsilon\lambda_\varepsilon \to 0, \quad \frac{\lambda_\varepsilon}{n} \to 0.$$

But, we will never use this fact in this thesis.

Lemma 3.4.11. *Under Assumptions 3.2.1, 3.2.3, 3.2.4, 3.2.6 and 3.2.8, let $\rho \in (0, 1/2)$, and suppose that a family $\{g(\cdot, \theta)\}_{\theta\in\Theta}$ of functions from $\mathbb{R}$ to $\mathbb{R}$ is equicontinuous at every points in I_{x_0}. Then,*

$$\frac{1}{n}\sum_{k=1}^{n} g(X_{t_{k-1}}^\varepsilon, \theta) 1_{C_k^{n,\varepsilon,\rho}} \xrightarrow{p} \int_0^1 g(x_t, \theta)\,\mathrm{d}t$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$, $\lambda_\varepsilon/n \to 0$ and $\varepsilon\lambda_\varepsilon \to 0$, uniformly in $\theta \in \Theta$. Also, for $p \in [2, \infty)$

$$\begin{aligned} &\frac{1}{n}\sum_{k=1}^{n} g\Big(X_{t_{k-1}}^\varepsilon, \theta\Big) 1_{C_{k,0}^{n,\varepsilon,\rho}} \xrightarrow{p} \int_0^1 g(x_t, \theta)\,\mathrm{d}t, \\ &\frac{1}{n}\sum_{k=1}^{n} g\Big(X_{t_{k-1}}^\varepsilon, \theta\Big) 1_{C_{k,1}^{n,\varepsilon,\rho}} = O_p\left(\frac{\lambda_\varepsilon}{n^{p(1-\rho)+1}} + \frac{\varepsilon^p \lambda_\varepsilon}{n^{p(1/2-\rho)+1}} + \frac{\lambda_\varepsilon}{n}\int_{|z|\leq 4v_{nk}/c_1 n^\rho} f_{\alpha_0}(z)\,\mathrm{d}z\right), \\ &\frac{1}{n}\sum_{k=1}^{n} g\Big(X_{t_{k-1}}^\varepsilon, \theta\Big) 1_{C_{k,2}^{n,\varepsilon,\rho}} = O_p\left(\frac{\lambda_\varepsilon^2}{n^2}\right) \end{aligned}$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$, $\lambda_\varepsilon/n \to 0$ and $\varepsilon\lambda_\varepsilon \to 0$, uniformly in $\theta \in \Theta$.

Proof of Lemma 3.4.11. Since $\{g(\cdot,\theta)\}_{\theta\in\Theta}$ is equicontinuous at every points in I_{x_0}, there exists $\delta > 0$ such that

$$\sup_{(x,\theta)\in I_{x_0}^\delta\times\Theta} |g(x,\theta)| < \infty.$$

For any $\eta > 0$

$$P\left(\left|\frac{1}{n}\sum_{k=1}^n g\Big(X_{t_{k-1}}^\varepsilon,\theta\Big)1_{D_k^{n,\varepsilon,\rho}}\right| > \eta\right)$$
$$\le P\left(\sup_{k=0,\dots,n-1}|X_{t_k}^\varepsilon - x_{t_k}| \ge \delta\right) + P\left(\sup_{(x,\theta)\in I_{x_0}^\delta\times\Theta}|g(x,\theta)|\frac{\lambda_\varepsilon}{n}\frac{1}{\lambda_\varepsilon}\sum_{k=1}^n 1_{D_k^{n,\varepsilon,\rho}} > \eta\right),$$
$$P\left(\left|\frac{1}{n}\sum_{k=1}^n g\Big(X_{t_{k-1}}^\varepsilon,\theta\Big)1_{C_{k,j}^{n,\varepsilon,\rho}}\right| > \eta\right)$$
$$\le P\left(\sup_{k=0,\dots,n-1}|X_{t_k}^\varepsilon - x_{t_k}| \ge \delta\right) + P\left(\sup_{(x,\theta)\in I_{x_0}^\delta\times\Theta}|g(x,\theta)|\frac{1}{n}\sum_{k=1}^n 1_{C_{k,j}^{n,\varepsilon,\rho}} > \eta\right) \quad \text{for } j=1,2.$$

It follows from Lemmas 3.4.3, 3.4.4 and 3.4.9 that

$$\left|\frac{1}{n}\sum_{k=1}^n g\Big(X_{t_{k-1}}^\varepsilon,\theta\Big)1_{C_k^{n,\varepsilon,\rho}} - \int_0^1 g(x_t,\theta)\,\mathrm{d}t\right|$$
$$\le \left|\frac{1}{n}\sum_{k=1}^n g\Big(X_{t_{k-1}}^\varepsilon,\theta\Big)1_{D_k^{n,\varepsilon,\rho}}\right| + \left|\frac{1}{n}\sum_{k=1}^n g\Big(X_{t_{k-1}}^\varepsilon,\theta\Big) - \int_0^1 g(x_t,\theta)\,\mathrm{d}t\right| \xrightarrow{p} 0,$$
$$\frac{1}{n}\sum_{k=1}^n g\Big(X_{t_{k-1}}^\varepsilon,\theta\Big)1_{C_{k,1}^{n,\varepsilon,\rho}} = O_p\left(\frac{\lambda_\varepsilon}{n^{p(1-\rho)+1}} + \frac{\varepsilon^p\lambda_\varepsilon}{n^{p(1/2-\rho)+1}} + \frac{\lambda_\varepsilon}{n}\int_{|z|\le 4v_{nk}/c_1n^\rho} f_{\alpha_0}(z)\,\mathrm{d}z\right),$$
$$\frac{1}{n}\sum_{k=1}^n g\Big(X_{t_{k-1}}^\varepsilon,\theta\Big)1_{C_{k,2}^{n,\varepsilon,\rho}} = O_p\left(\frac{\lambda_\varepsilon^2}{n^2}\right)$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$, $\lambda_\varepsilon/n \to 0$ and $\varepsilon\lambda_\varepsilon \to 0$, uniformly in $\theta \in \Theta$. □

Lemma 3.4.12. *Under Assumptions 3.2.1, 3.2.3, 3.2.4, 3.2.6 and 3.2.8, let $\rho \in (0,1/2)$. We assume either of the following conditions (i) or (ii):*

(i) Under Assumption 3.2.4 (i), we assume the following four conditions:

(i.a) There exists $\delta > 0$ such that for every $(x,\theta) \in I_{x_0}^\delta \times \bar{\Theta}$, $g(x,y,\theta)$ is continuously differentiable with respect to $y \in \mathbb{R}$.

(i.b) There exist constants $C > 0$, $q \ge 1$ and $\delta > 0$ such that

$$\sup_{(x,\theta)\in I_{x_0}^\delta\times\bar{\Theta}}\left|\frac{\partial g}{\partial y}(x,y,\theta)\right| \le C(1+|y|^q) \quad (y \in \mathbb{R}).$$

(i.c) *There exists a sufficiently large* $p \geq 2$ *such that*

$$\lambda_\varepsilon \to \infty, \quad \frac{\lambda_\varepsilon^2}{n} \to 0, \quad \varepsilon\lambda_\varepsilon \to 0, \quad \varepsilon n^{1-1/p} \to \infty, \quad \lambda_\varepsilon \int_{|z|\leq 4v_2/c_1 n^\rho} f_{\alpha_0}(z)\,\mathrm{d}z \to 0$$

as $n \to \infty$ *and* $\varepsilon \to 0$.

(i.d) *Let* p *be taken as in the condition (i.c). Put* $r_{n,\varepsilon}$ *by*

$$r_{n,\varepsilon} := \frac{1}{\varepsilon n^{1-1/p}} + \frac{1}{n^{1/2-1/p}}.$$

(ii) *Under Assumption 3.2.4 (ii), we assume the following six conditions:*

(ii.a) *There exists* $\delta > 0$ *such that for every* $(x, \theta) \in I_{x_0}^\delta \times \bar{\Theta}$, $g(x, y, \theta)$ *is continuously differentiable with respect to* $y \in (0, \infty)$.

(ii.b) *There exists* $\delta > 0$ *and* $L > 0$ *such that if* $0 < y_1 \leq y \leq y_2$, *then*

$$\left|\frac{\partial g}{\partial y}(x, y, \theta)\right| \leq \left|\frac{g}{y}(x, y_1, \theta)\right| + \left|\frac{\partial g}{\partial y}(x, y_2, \theta)\right| + L \quad \textit{for all } (x, \theta) \in I_{x_0}^\delta \times \bar{\Theta}.$$

(ii.c) *There exist* $q \geq 0$ *and* $\delta > 0$ *such that*

$$\sup_{(x,\theta)\in I_{x_0}^\delta \times \bar{\Theta}} \left|\frac{\partial g}{\partial y}(x, y, \theta)\right| \leq O\left(\frac{1}{|y|^q}\right) \quad \textit{as } |y| \to 0.$$

(ii.d) *There exists* $\delta > 0$ *such that for any* $C_1 > 0$ *and* $C_2 \geq 0$ *the map*

$$x \mapsto \int \sup_\theta \left|\frac{\partial g}{\partial y}(x, C_1 y + C_2, \theta)\right| f_{\alpha_0}(y)\,\mathrm{d}y$$

takes values in $\mathbb{R}$ *from* $I_{x_0}^\delta$, *and is continous on* $I_{x_0}^\delta$.

(ii.e) *Let* q *be the constant in the condition (ii.c), and let* $\rho < 1/4q$. *For any large* $p \geq 2/(1 - 2q\rho)$,

$$\lambda_\varepsilon \to \infty, \quad \frac{\lambda_\varepsilon^2}{n} \to 0, \quad \varepsilon\lambda_\varepsilon \to 0, \quad \varepsilon n^{1-q\rho-1/p} \to \infty, \quad \lambda_\varepsilon \int_{|z|\leq 4v_2/c_1 n^\rho} f_{\alpha_0}(z)\,\mathrm{d}z \to 0$$

as $n \to \infty$, $\varepsilon \to 0$.

(ii.f) *Let* p *and* q *be the constants in the condition (ii.e). Put* $r_{n,\varepsilon}$ *by*

$$r_{n,\varepsilon} := \frac{1}{\varepsilon n^{1-1/p-q\rho}} + \frac{1}{n^{1/2-1/p-q\rho}}.$$

Then,

$$\left|\frac{1}{\lambda_\varepsilon}\sum_{k=1}^n g\left(X_{t_{k-1}}^\varepsilon, \frac{\Delta_k^n X^\varepsilon}{\varepsilon}, \theta\right) 1_{D_k^{n,\varepsilon,\rho}} - \frac{1}{\lambda_\varepsilon}\sum_{k=1}^n g\left(X_{t_{k-1}}^\varepsilon, c(X_{t_{k-1}}^\varepsilon, \alpha_0) V_{N_{\tau_k}^{\lambda_\varepsilon}}, \theta\right) 1_{J_{k,1}^{n,\varepsilon}}\right| = O_p(r_{n,\varepsilon})$$

as $n \to \infty$ *and* $\varepsilon \to 0$, *uniformly in* $\theta \in \Theta$.

Remark 3.4.3. *Assumption 3.2.4 is used only for defining* $D_k^{n,\varepsilon,\rho}$ *in Lemmas 3.4.7, 3.4.8, 3.4.10 and 3.4.11, while it is essentially used in Lemma 3.4.12.*

Remark 3.4.4. *The assumptions (i.c) and (ii.e) in Lemma 3.4.12 are ensured if*

$$\lambda_\varepsilon \to \infty, \quad \varepsilon\lambda_\varepsilon \to 0, \quad (\varepsilon\sqrt{n})^{-1} < \infty \quad \textit{and} \quad \lambda_\varepsilon \int_{|z|\le 4v_2/c_1 n^\rho} f_{\alpha_0}(z)\,\mathrm{d}z \to 0$$

as $n \to \infty$ *and* $\varepsilon \to 0$*. This condition seems to be natural when we consider the asymptotic normality for our estimator (see,* e.g.*, the condition (B2) in Sørensen and Uchida [33]).*

Proof of Lemma 3.4.12. Let $\delta > 0$ be a sufficiently small number satisfying the conditions of the statement and

$$\frac{c_1}{2} \le c(x, \alpha_0) \le 2c_2 \quad \text{for } x \in I_{x_0}^\delta,$$

where c_1 and c_2 are the constants from Assumption 3.2.6. In this proof, we may simply write the maps

$$(y,\theta) \mapsto g(X_{t_{k-1}}^\varepsilon, y, \theta) =: g_k(y,\theta) \quad \text{and} \quad (y,\theta) \mapsto \frac{\partial g}{\partial y}(x,y,\theta)\bigg|_{x = X_{t_{k-1}}^\varepsilon} =: \frac{\partial g_k}{\partial y}(y,\theta),$$

and we denote the following event by $\tilde{D}_{k,1}^{n,\varepsilon,\rho}$

$$\tilde{D}_{k,1}^{n,\varepsilon,\rho} := D_{k,1}^{n,\varepsilon,\rho} \cap \{X_t^\varepsilon \in I_{x_0}^\delta \text{ for all } t \in [0,1]\}.$$

Since

$$\frac{\lambda_\varepsilon^2}{n} \to 0, \quad \lambda_\varepsilon \int_{|z|\le 4v_2/c_1 n^\rho} f_{\alpha_0}(z)\,\mathrm{d}z \to 0,$$

under either of the assumptions (i.c) or (ii.e), we obtain from Lemma 3.4.10 that for any non-random $r'_{n,\varepsilon} > 0$ $(n \in \mathbb{N}, \varepsilon > 0)$,

$$\frac{1}{\lambda_\varepsilon}\sum_{k=1}^n g_k\left(\frac{\Delta_k^n X^\varepsilon}{\varepsilon}, \theta\right)\left\{1_{D_k^{n,\varepsilon,\rho}} - 1_{\tilde{D}_{k,1}^{n,\varepsilon,\rho}}\right\} = o_p(r'_{n,\varepsilon}),$$

$$\frac{1}{\lambda_\varepsilon}\sum_{k=1}^n g_k\left(c(X_{t_{k-1}}^\varepsilon, \alpha_0)V_{N_{\tau_k}^{\lambda_\varepsilon}}, \theta\right)\left\{1_{\tilde{D}_{k,1}^{n,\varepsilon,\rho}} - 1_{J_{k,1}^{n,\varepsilon}}\right\} = o_p(r'_{n,\varepsilon})$$

as $n \to \infty$, $\varepsilon \to 0$ $\lambda_\varepsilon \to \infty$ and $\varepsilon\lambda_\varepsilon \to 0$, uniformly in $\theta \in \Theta$. Thus, it is sufficient to show that

$$\frac{1}{\lambda_\varepsilon}\sum_{k=1}^n \left\{g_k\left(\frac{\Delta_k^n X^\varepsilon}{\varepsilon}, \theta\right) - g_k\left(\frac{\Delta X_{\tau_k}^\varepsilon}{\varepsilon}, \theta\right)\right\} 1_{\tilde{D}_{k,1}^{n,\varepsilon,\rho}} = O_p(r_{n,\varepsilon}), \tag{3.11}$$

$$\frac{1}{\lambda_\varepsilon}\sum_{k=1}^n \left\{g_k\left(\frac{\Delta X_{\tau_k}^\varepsilon}{\varepsilon}, \theta\right) - g_k\left(c(X_{t_{k-1}}^\varepsilon, \alpha_0)V_{N_{\tau_k}^{\lambda_\varepsilon}}, \theta\right)\right\} 1_{\tilde{D}_{k,1}^{n,\varepsilon,\rho}} = O_p\left(\frac{1}{\varepsilon n^{1-1/p}} + \frac{1}{n^{1/2-1/p}}\right) \tag{3.12}$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$ and $\varepsilon\lambda_\varepsilon \to 0$, uniformly in $\theta \in \Theta$.

Put

$$Y_k^\varepsilon := \frac{X_{t_k}^\varepsilon - X_{\eta_k}^\varepsilon}{\varepsilon} + \frac{X_{\tau_k-}^\varepsilon - X_{t_{k-1}}^\varepsilon}{\varepsilon} \left(= \frac{\Delta_k^n X^\varepsilon}{\varepsilon} - \frac{\Delta X_{\tau_k}^\varepsilon}{\varepsilon} \quad \text{on } D_{k,1}^{n,\varepsilon,\rho} \right).$$

By using Taylor's theorem under either of the assumptions (i.a) or (ii.a), we have

$$g_k\left(\frac{\Delta_k^n X^\varepsilon}{\varepsilon}, \theta\right) - g_k\left(\frac{\Delta X_{\tau_k}^\varepsilon}{\varepsilon}, \theta\right) = \int_0^1 \frac{\partial g_k}{\partial y}\left(\frac{\Delta X_{\tau_k}^\varepsilon}{\varepsilon} + \zeta Y_k^\varepsilon, \theta\right) Y_k^\varepsilon \, \mathrm{d}\zeta \quad \text{on } \tilde{D}_{k,1}^{n,\varepsilon,\rho}.$$

Here, we remark that $\Delta_k^n X^\varepsilon$ and $\Delta X_{\tau_k}^\varepsilon$ are almost surely positive on $\tilde{D}_{k,1}^{n,\varepsilon,\rho}$ under Assumption 3.2.4 (ii). To see (3.11), it is sufficient to show that

$$\sup_{\theta\in\Theta} \left| \frac{1}{\lambda_\varepsilon} \sum_{k=1}^n \int_0^1 \frac{\partial g_k}{\partial y}\left(\frac{\Delta X_{\tau_k}^\varepsilon}{\varepsilon} + \zeta Y_k^\varepsilon, \theta\right) Y_k^\varepsilon \, \mathrm{d}\zeta \, 1_{\tilde{D}_{k,1}^{n,\varepsilon,\rho} \cap \{|Y_k^\varepsilon| \le 1\}} \right| = O_p(r_{n,\varepsilon}) \tag{3.13}$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$ and $\varepsilon\lambda_\varepsilon \to 0$. Indeed, for any $M > 0$

$$\begin{aligned} &P\left(\sup_{\theta\in\Theta} \left| \frac{1}{\lambda_\varepsilon} \sum_{k=1}^n \left\{ g_k\left(\frac{\Delta_k^n X^\varepsilon}{\varepsilon}, \theta\right) - g_k\left(\frac{\Delta X_{\tau_k}^\varepsilon}{\varepsilon}, \theta\right) \right\} 1_{\tilde{D}_{k,1}^{n,\varepsilon,\rho}} \right| > M r_{n,\varepsilon} \right) \\ &\quad \le P\left(\sup_{k=1,\dots,n} |Y_k^\varepsilon| > 1 \right) \\ &\qquad + P\left(\sup_{\theta\in\Theta} \left| \frac{1}{\lambda_\varepsilon} \sum_{k=1}^n \int_0^1 \frac{\partial g_k}{\partial y}\left(\frac{\Delta X_{\tau_k}^\varepsilon}{\varepsilon} + \zeta Y_k^\varepsilon, \theta\right) Y_k^\varepsilon \, \mathrm{d}\zeta \, 1_{\tilde{D}_{k,1}^{n,\varepsilon,\rho} \cap \{|Y_k^\varepsilon| \le 1\}} \right| > M r_{n,\varepsilon} \right), \end{aligned}$$

and from Lemma 3.4.6 the first term converges to zero as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$ and $\varepsilon\lambda_\varepsilon \to 0$, since from either of the assumptions (i.c) or (ii.e) we have $\varepsilon n^{1-1/p} \to \infty$ or $\varepsilon n^{1-q\rho-1/p} \to \infty$, respectively.

We first consider the case (ii) in Assumption 3.2.4. Since for $\zeta \in [0, 1]$ we have

$$\frac{\Delta X_{\tau_k}^\varepsilon}{\varepsilon} + \zeta Y_k^\varepsilon \ge (1-\zeta)\, c(X_{\tau_k-}^\varepsilon, \alpha_0)\, V_{N_{\tau_k}^{\lambda_\varepsilon}} + \zeta \frac{v_{nk}}{n^\rho} \ge \min\left\{ \frac{c_1}{2} V_{N_{\tau_k}^{\lambda_\varepsilon}}, \frac{v_1}{n^\rho} \right\} \quad \text{on } \tilde{D}_{k,1}^{n,\varepsilon,\rho},$$

we obtain from the assumption (ii.b) that

$$\begin{aligned} &\int_0^1 \left| \frac{\partial g_k}{\partial y}\left(\frac{\Delta X_{\tau_k}^\varepsilon}{\varepsilon} + \zeta Y_k^\varepsilon, \theta\right) \right| \mathrm{d}\zeta \, 1_{\tilde{D}_{k,1}^{n,\varepsilon,\rho} \cap \{|Y_k^\varepsilon| \le 1\}} \\ &\quad \le \left\{ \left| \frac{\partial g_k}{\partial y}\left(\frac{c_1}{2} V_{N_{\tau_k}^{\lambda_\varepsilon}}, \theta\right) \right| + \left| \frac{\partial g_k}{\partial y}\left(\frac{v_1}{n^\rho}, \theta\right) \right| + \left| \frac{\partial g_k}{\partial y}\left(2c_2 V_{N_{\tau_k}^{\lambda_\varepsilon}} + 1, \theta\right) \right| + L \right\} 1_{\tilde{D}_{k,1}^{n,\varepsilon,\rho} \cap \{|Y_k^\varepsilon| \le 1\}} \\ &\quad \le \left\{ \left| \frac{\partial g_k}{\partial y}\left(\frac{c_1}{2} V_{N_{\tau_k}^{\lambda_\varepsilon}}, \theta\right) \right| + \left| \frac{\partial g_k}{\partial y}\left(\frac{v_1}{n^\rho}, \theta\right) \right| + \left| \frac{\partial g_k}{\partial y}\left(2c_2 V_{N_{\tau_k}^{\lambda_\varepsilon}} + 1, \theta\right) \right| + L \right\} 1_{J_{k,1}^{n,\varepsilon}} \end{aligned}$$

Since

$$\begin{aligned} &\frac{1}{\lambda_\varepsilon} \sum_{k=1}^n E\left[\sup_\theta \left| \frac{\partial g_k}{\partial y}\left(\frac{c_1}{2} V_{N_{\tau_k}^{\lambda_\varepsilon}}, \theta\right) \right| 1_{J_{k,1}^{n,\varepsilon}} \middle| \mathcal{F}_{t_{k-1}} \right] \\ &\le \frac{1}{\lambda_\varepsilon} \sum_{k=1}^n \int \sup_\theta \left| \frac{\partial g_k}{\partial y}\left(\frac{c_1}{2} z, \theta\right) \right| f_{\alpha_0}(z)\, \mathrm{d}z \cdot P\left(J_{k,1}^{n,\varepsilon}\right) \le \frac{1}{n} \sum_{k=1}^n \int \sup_\theta \left| \frac{\partial g_k}{\partial y}\left(\frac{c_1}{2} z, \theta\right) \right| f_{\alpha_0}(z)\, \mathrm{d}z\,, \end{aligned}$$

it follows from Lemma 3.7.3 (ii), Lemmas 3.4.4 and 3.4.6 and the assumption (ii.d) that

$$\frac{1}{\lambda_\varepsilon}\sum_{k=1}^n \left|\frac{\partial g_k}{\partial y}\Big(\frac{c_1}{2}V_{N^{\lambda_\varepsilon}_{\tau_k}},\theta\Big)\right| 1_{J^{n,\varepsilon}_{k,1}} \sup_{k=1,\dots,n}|Y^\varepsilon_k| = O_p\left(\frac{1}{\varepsilon n^{1-1/p}}+\frac{1}{n^{1/2-1/p}}\right)$$

as $n\to\infty$, $\varepsilon\to 0$, $\lambda_\varepsilon\to\infty$ and $\varepsilon\lambda_\varepsilon\to 0$, uniformly in $\theta\in\Theta$, where p is given in the assumption (ii.e). Similarly, it follows from Lemma 3.7.3 (ii), Lemmas 3.4.4 and 3.4.6 and the assumption (ii.d) that

$$\frac{1}{\lambda_\varepsilon}\sum_{k=1}^n \left|\frac{\partial g_k}{\partial y}\Big(2c_2V_{N^{\lambda_\varepsilon}_{\tau_k}}+1,\theta\Big)\right| 1_{J^{n,\varepsilon}_{k,1}} \sup_{k=1,\dots,n}|Y^\varepsilon_k| = O_p\left(\frac{1}{\varepsilon n^{1-1/p}}+\frac{1}{n^{1/2-1/p}}\right)$$

as $n\to\infty$, $\varepsilon\to 0$, $\lambda_\varepsilon\to\infty$ and $\varepsilon\lambda_\varepsilon\to 0$, uniformly in $\theta\in\Theta$, and it follows from Lemma 3.7.3 (ii), Lemma 3.4.6 and the assumption (ii.c) that

$$\frac{1}{\lambda_\varepsilon}\sum_{k=1}^n \left|\frac{\partial g_k}{\partial y}\Big(\frac{v_1}{n^\rho},\theta\Big)\right| 1_{J^{n,\varepsilon}_{k,1}} \sup_{k=1,\dots,n}|Y^\varepsilon_k| = O_p\left(\frac{1}{\varepsilon n^{1-1/p-q\rho}}+\frac{1}{n^{1/2-1/p-q\rho}}\right)$$

as $n\to\infty$, $\varepsilon\to 0$, $\lambda_\varepsilon\to\infty$ and $\varepsilon\lambda_\varepsilon\to 0$, uniformly in $\theta\in\Theta$. Thus, we obtain (3.13).

Under the case (i) in Assumption 3.2.4, as in the same argument above, we have

$$\begin{aligned}
&\frac{1}{\lambda_\varepsilon}\sum_{k=1}^n \sup_{\theta\in\bar\Theta}\left|\int_0^1 \frac{\partial g_k}{\partial y}\left(\frac{\Delta X^\varepsilon_{\tau_k}}{\varepsilon}+\zeta Y^\varepsilon_k,\theta\right)Y^\varepsilon_k\,\mathrm{d}\zeta\right| 1_{\tilde D^{n,\varepsilon,\rho}_{k,1}\cap\{|Y^\varepsilon_k|\le 1\}}\\
&\le \frac{C}{\lambda_\varepsilon}\sum_{k=1}^n\Big(2+\Big|2c_2V_{N^{\lambda_\varepsilon}_{\tau_k}}\Big|^p\Big)1_{J^{n,\varepsilon}_{k,1}}\sup_{k=1,\dots,n}|Y^\varepsilon_k| = O_p\left(\frac{1}{\varepsilon n^{1-1/p}}+\frac{1}{n^{1/2-1/p}}\right)
\end{aligned}$$

as $n\to\infty$, $\varepsilon\to 0$, $\lambda_\varepsilon\to\infty$ and $\varepsilon\lambda_\varepsilon\to 0$. Thus, we obtain (3.13).

Analogously, it follows that for $\zeta\in[0,1]$

$$(1-\zeta)\frac{\Delta X^\varepsilon_{\tau_k}}{\varepsilon}+\zeta c(X^\varepsilon_{t_{k-1}},\alpha_0)V_{N^{\lambda_\varepsilon}_{\tau_k}} \ge \frac{c_1}{2}V_{N^{\lambda_\varepsilon}_{\tau_k}} \quad \text{on } \tilde D^{n,\varepsilon,\rho}_{k,1},$$

and that on $\tilde D^{n,\varepsilon,\rho}_{k,1}$

$$\begin{aligned}
&\int_0^1\left|\frac{\partial g_k}{\partial y}\left((1-\zeta)\frac{\Delta X^\varepsilon_{\tau_k}}{\varepsilon}+\zeta c(X^\varepsilon_{t_{k-1}},\alpha_0)V_{N^{\lambda_\varepsilon}_{\tau_k}},\theta\right)\right|\mathrm{d}\zeta\\
&\le\begin{cases} C\left(1+\Big|2c_2V_{N^{\lambda_\varepsilon}_{\tau_k}}\Big|^p\right) & \text{in the case (i)},\\ \left|\frac{\partial g_k}{\partial y}\Big(\frac{c_1}{2}V_{N^{\lambda_\varepsilon}_{\tau_k}},\theta\Big)\right|+\left|\frac{\partial g_k}{\partial y}\Big(2c_2V_{N^{\lambda_\varepsilon}_{\tau_k}},\theta\Big)\right| & \text{in the case (ii)},\end{cases}
\end{aligned}$$

so that, (3.12) holds. □

Lemma 3.4.13. *Let $\rho\in(0,1/2)$. Under Assumptions 3.2.1, 3.2.3, 3.2.4, 3.2.6 and 3.2.8, suppose that for $\theta\in\Theta$*

$$x\mapsto\int g(x,c(x,\alpha_0)z,\theta)f_{\alpha_0}(z)\,\mathrm{d}z\,,\quad x\mapsto\int |g(x,c(x,\alpha_0)z,\theta)|^2 f_{\alpha_0}(z)\,\mathrm{d}z \tag{3.14}$$

are continuous at every points in I_{x_0}, and that there exist $\delta > 0$, $C > 0$ and $q \geq 0$ such that

$$\int \left\{ \sup_{(x,\theta)\in I_{x_0}^\delta \times \Theta} |g(x, c(x, \alpha_0)z, \theta)| + \sum_{j=1}^d \sup_{(x,\theta)\in I_{x_0}^\delta \times \Theta} \left| \frac{\partial g}{\partial \theta_j}(x, c(x, \alpha_0)z, \theta) \right| \right\} f_{\alpha_0}(z)\,\mathrm{d}z < \infty. \tag{3.15}$$

Then,

$$\frac{1}{\lambda_\varepsilon} \sum_{k=1}^n g\Big(X_{t_{k-1}}^\varepsilon, c(X_{t_{k-1}}^\varepsilon, \alpha_0) V_{N_{\tau_k}^{\lambda_\varepsilon}}, \theta\Big) 1_{D_{k,1}^{n,\varepsilon,\rho}} \xrightarrow{p} \int_0^1 \int g(x_t, c(x_t, \alpha_0)z, \theta) f_{\alpha_0}(z)\,\mathrm{d}z\,\mathrm{d}t$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$ and $\varepsilon\lambda_\varepsilon \to 0$, uniformly in $\theta \in \Theta$.

Proof. It follows from Lemma 3.4.4 and the assumption (3.14) that for each $\theta \in \Theta$

$$\begin{aligned}
&\sum_{k=1}^n E\left[\frac{1}{\lambda_\varepsilon} g\Big(X_{t_{k-1}}^\varepsilon, c(X_{t_{k-1}}^\varepsilon, \alpha_0) V_{N_{\tau_k}^{\lambda_\varepsilon}}, \theta\Big) 1_{J_{k,1}^{n,\varepsilon}} \,\middle|\, \mathcal{F}_{t_{k-1}} \right] \\
&\quad = \frac{1}{n} \sum_{k=1}^n \int g\Big(X_{t_{k-1}}^\varepsilon, c(X_{t_{k-1}}^\varepsilon, \alpha_0)z, \theta\Big) f_{\alpha_0}(z)\,\mathrm{d}z \\
&\quad \xrightarrow{p} \int_0^1 \int g(x_t, c(x_t, \alpha_0)z, \theta) f_{\alpha_0}(z)\,\mathrm{d}z\,\mathrm{d}t
\end{aligned}$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$ and $\varepsilon\lambda_\varepsilon \to 0$, and that

$$\sum_{k=1}^n E\left[\frac{1}{\lambda_\varepsilon^2} \left| g\Big(X_{t_{k-1}}^\varepsilon, c(X_{t_{k-1}}^\varepsilon, \alpha_0) V_{N_{\tau_k}^{\lambda_\varepsilon}}, \theta\Big) \right|^2 1_{J_{k,1}^{n,\varepsilon}} \,\middle|\, \mathcal{F}_{t_{k-1}} \right] \xrightarrow{p} 0$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$ and $\varepsilon\lambda_\varepsilon \to 0$. Thus, Lemma 9 in Genon-Catalot and Jacod [11] shows us that for each $\theta \in \Theta$

$$\frac{1}{\lambda_\varepsilon} \sum_{k=1}^n g\Big(X_{t_{k-1}}^\varepsilon, c(X_{t_{k-1}}^\varepsilon, \alpha_0) V_{N_{\tau_k}^{\lambda_\varepsilon}}, \theta\Big) 1_{J_{k,1}^{n,\varepsilon}} \xrightarrow{p} \int_0^1 \int g(x_t, c(x_t, \alpha_0)z, \theta) f_{\alpha_0}(z)\,\mathrm{d}z\,\mathrm{d}t$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$ and $\varepsilon\lambda_\varepsilon \to 0$. Put

$$\tilde{J}_{k,1}^{n,\varepsilon} := J_{k,1}^{n,\varepsilon} \cap \{X_t^\varepsilon \in I_{x_0}^\delta \text{ for all } t \in [0,1]\}.$$

Then, by the same argument in the proof of Lemma 3.4.10, it follows from Lemma 3.4.3 that

$$\frac{1}{\lambda_\varepsilon} \sum_{k=1}^n g\Big(X_{t_{k-1}}^\varepsilon, c(X_{t_{k-1}}^\varepsilon, \alpha_0) V_{N_{\tau_k}^{\lambda_\varepsilon}}, \theta\Big) \left\{ 1_{J_{k,1}^{n,\varepsilon}} - 1_{\tilde{J}_{k,1}^{n,\varepsilon}} \right\} \xrightarrow{p} 0$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$ and $\varepsilon\lambda_\varepsilon \to 0$, uniformly in $\theta \in \Theta$.

Now, we have for each $\theta \in \Theta$

$$\frac{1}{\lambda_\varepsilon} \sum_{k=1}^n g\Big(X_{t_{k-1}}^\varepsilon, c(X_{t_{k-1}}^\varepsilon, \alpha_0) V_{N_{\tau_k}^{\lambda_\varepsilon}}, \theta\Big) 1_{\tilde{J}_{k,1}^{n,\varepsilon}} \xrightarrow{p} \int_0^1 \int g(x_t, c(x_t, \alpha_0)z, \theta) f_{\alpha_0}(z)\,\mathrm{d}z\,\mathrm{d}t$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$ and $\varepsilon\lambda_\varepsilon \to 0$. To say the uniformity of this convergence in $\theta \in \Theta$, put

$$\chi^{n,\varepsilon}(\theta) := \frac{1}{\lambda_\varepsilon}\sum_{k=1}^n g\Big(X^\varepsilon_{t_{k-1}}, c(X^\varepsilon_{t_{k-1}},\alpha_0)V_{N^{\lambda_\varepsilon}_{\tau_k}},\theta\Big)1_{\tilde{J}^{n,\varepsilon}_{k,1}} - \int_0^1\int g(x_t, c(x_t,\alpha_0)z,\theta)f_{\alpha_0}(z)\,\mathrm{d}z\,\mathrm{d}t$$

and we shall use Theorem 5.1 in Billingsley [4] with the state space $C(\Theta)$, same as in the proofs of Propositions 3.3 and 3.6 in Shimizu and Yoshida [32] [1]. From the assumption (3.15), we obtain

$$\begin{aligned}
&E\left[\sup_{\theta\in\Theta}\left|\frac{1}{\lambda_\varepsilon}\sum_{k=1}^n g\left(X^\varepsilon_{t_{k-1}}, c(X^\varepsilon_{t_{k-1}},\alpha_0)V_{N^{\lambda_\varepsilon}_{\tau_k}},\theta\right)1_{\tilde{J}^{n,\varepsilon}_{k,1}}\right|\right]\\
&\quad\le \frac{1}{\lambda_\varepsilon}\sum_{k=1}^n E\left[\sup_{(x,\theta)\in I^\delta_{x_0}\times\Theta}\left|g(x, c(x,\alpha_0)V_{N^{\lambda_\varepsilon}_{\tau_k}},\theta)\right|1_{J^{n,\varepsilon}_{k,1}}\right]\\
&\quad= \int \sup_{(x,\theta)\in I^\delta_{x_0}\times\Theta}|g(x,c(x,\alpha_0)z,\theta)|\,f_{\alpha_0}(z)\,\mathrm{d}z\ (<\infty)
\end{aligned}$$

and

$$\begin{aligned}
&E\left[\sup_{\theta\in\Theta}\left|\frac{1}{\lambda_\varepsilon}\sum_{k=1}^n \frac{\partial g}{\partial\theta_j}\Big(X^\varepsilon_{t_{k-1}}, c(X^\varepsilon_{t_{k-1}},\alpha_0)V_{N^{\lambda_\varepsilon}_{\tau_k}},\theta\Big)1_{\tilde{J}^{n,\varepsilon}_{k,1}}\right|\right]\\
&\quad\le \frac{1}{\lambda_\varepsilon}\sum_{k=1}^n E\left[\sup_{(x,\theta)\in I^\delta_{x_0}\times\Theta}\left|\frac{\partial g}{\partial\theta_j}\Big(x, c(x,\alpha_0)V_{N^{\lambda_\varepsilon}_{\tau_k}},\theta\Big)\right|1_{J^{n,\varepsilon}_{k,1}}\right]\\
&\quad= \int \sup_{(x,\theta)\in I^\delta_{x_0}\times\Theta}\left|\frac{\partial g}{\partial\theta_j}(x,c(x,\alpha_0)z,\theta)\right|f_{\alpha_0}(z)\,\mathrm{d}z\ (<\infty) \qquad \text{for } j=1,\dots,d.
\end{aligned}$$

The above equalities hold from the fact that $V_{N^{\lambda_\varepsilon}_{\tau_k}}$ and $1_{J^{n,\varepsilon}_{k,1}}$ are independent. Hence, for any closed ball B_M of radius $M>0$ centered at zero in the Sobolev space $W^{1,\infty}(\Theta)$, we obtain from Markov's inequality that

$$\sup_{n,\varepsilon} P\left(\chi^{n,\varepsilon}\notin B_M\right) = P\left(\|\chi^{n,\varepsilon}\|_{W^{1,\infty}(\Theta)} \ge M\right) \le \frac{2C}{M},$$

where C is defined as (3.15) and for $q \ge 1$

$$\|u\|_{W^{1,q}(\Theta)} := \|u\|_{L^q(\Theta)} + \sum_{j=1}^d \left\|\frac{\partial u}{\partial\theta_j}\right\|_{L^q(\Theta)} \qquad \text{for } u\in W^{1,q}(\Theta).$$

From Rellich-Kondrachov's theorem (see, *e.g.*, Theorem 9.16 in Brezis [5]), it follows that the balls B_M, $M>0$ are compact in $C(\Theta)$, and so from Theorem 5.1 in Billingsley [4] that $\{\chi^{n,\varepsilon}\}$

[1] We cannot use Theorem 20 in Ibragimov and Has'minskii [17, Appendix I] (or Lemma 3.1 in Yoshida [36]), as in the proof of Lemma 2 in Sørensen and Uchida [33]. In fact, we fail to say that $\{\chi^{n,\varepsilon}\}$ satisfies (1) and (2) in Lemma 3.1 in Yoshida [36].

is relatively compact in distribution sense as in the Billingsley's book. Since for each $\theta \in \Theta$ $\{\chi^{n,\varepsilon}(\theta)\}$ converges to zero in probability, all convergent subsequences of $\{\chi^{n,\varepsilon}\}$ converges to zero in probability. Analogously, all subnet of $\{\chi^{n,\varepsilon}\}$ has a subsequence convergent in probability to zero, and so $\{\chi^{n,\varepsilon}\}$ converges to zero in probability as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$ and $\varepsilon\lambda_\varepsilon \to 0$. □

Lemma 3.4.14. *Under Assumptions 3.2.1, 3.2.3, 3.2.4, 3.2.6 and 3.2.8, let $\rho \in (0, 1/2)$, and let $g : \mathbb{R} \times \Theta \to \mathbb{R}$ satisfy that $\left\{\frac{\partial g}{\partial \theta_j}(\cdot, \theta)\right\}_{\theta\in\Theta}$, $j = 1, \ldots, d$ are equi-Lipschitz continuous on $I_{x_0}^\delta$ for some small $\delta > 0$. Then,*

$$\frac{1}{\varepsilon}\sum_{k=1}^n g(X_{t_{k-1}}^\varepsilon, \theta)\left\{\Delta_k^n X^\varepsilon - \frac{1}{n}a(X_{t_{k-1}}^\varepsilon, \mu_0)\right\} 1_{C_k^{n,\varepsilon,\rho}} \xrightarrow{p} \int_0^1 g(x_t, \theta)b(x_t, \theta)\,\mathrm{d}W_t$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$, $\lambda_\varepsilon^2/n \to 0$, $\varepsilon\lambda_\varepsilon \to 0$ and $\lambda_\varepsilon \int_{|z|\le 4v_2/c_1 n^\rho} f_{\alpha_0}(z)\,\mathrm{d}z \to 0$, uniformly in $\theta \in \Theta$.

Proof. At first, we can easily check that

$$\frac{1}{\varepsilon}\sum_{k=1}^n g(X_{t_{k-1}}^\varepsilon, \theta)\left\{\int_{t_{k-1}}^{t_k} a(X_t^\varepsilon, \mu_0)\,\mathrm{d}t - \frac{1}{n}a(X_{t_{k-1}}^\varepsilon, \mu_0)\right\} 1_{C_k^{n,\varepsilon,\rho}} \xrightarrow{p} 0 \qquad (3.16)$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$, $\varepsilon n \to \infty$, and $\varepsilon\lambda_\varepsilon \to 0$, uniformly in $\theta \in \Theta$. Indeed, this follows from Lemmas 3.4.3, 3.7.2 and 3.7.3 with the equicontinuity of g on I_{x_0} and the following estimate:

$$\begin{aligned}
&\frac{1}{\varepsilon}\sum_{k=1}^n E\left[\sup_{\theta\in\Theta}\left| g(X_{t_{k-1}}^\varepsilon, \theta)\left\{\int_{t_{k-1}}^{t_k} a(X_t^\varepsilon, \mu_0)\,\mathrm{d}t - \frac{1}{n}a(X_{t_{k-1}}^\varepsilon, \mu_0)\right\} 1_{C_k^{n,\varepsilon,\rho}}\right| \,\middle|\, \mathcal{F}_{t_{k-1}}\right] \\
&\quad \le C\left(\frac{1}{n}\sum_{k=1}^n E\left[\sup_{\theta\in\Theta}\left| g(X_{t_{k-1}}^\varepsilon, \theta)\sup_{t\in[t_{k-1},t_k]}\frac{|X_t^\varepsilon - X_{t_{k-1}}^\varepsilon|}{\varepsilon}\right|^2 \,\middle|\, \mathcal{F}_{t_{k-1}}\right]\right)^{1/2} \\
&\qquad\qquad\qquad\qquad (\because \text{Schwartz's inequality and 3.2.1}) \\
&\quad = O_p\left(\frac{1}{\varepsilon n} + \frac{1}{\sqrt{n}} + \frac{\lambda_\varepsilon}{n}\right) \quad (\because \text{Lemmas 3.4.1 and 3.4.4})
\end{aligned}$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$ and $\varepsilon\lambda_\varepsilon \to 0$.

At second, we show that

$$\begin{aligned}
&\sum_{k=1}^n g(X_{t_{k-1}}^\varepsilon, \theta)\int_{t_{k-1}}^{t_k} b(X_{t-}^\varepsilon, \sigma_0)\,\mathrm{d}W_t\, 1_{C_k^{n,\varepsilon,\rho}} - \int_0^1 g(X_{t-}^\varepsilon, \theta)b(X_{t-}^\varepsilon, \sigma_0)\,\mathrm{d}W_t \\
&\qquad = \sum_{k=1}^n \int_{t_{k-1}}^{t_k}\left\{g(X_{t_{k-1}}^\varepsilon, \theta) - g(X_{t-}^\varepsilon, \theta)\right\} b(X_{t-}^\varepsilon, \sigma_0)\,\mathrm{d}W_t\, 1_{C_k^{n,\varepsilon,\rho}} \\
&\qquad\qquad - \sum_{k=1}^n \int_{t_{k-1}}^{t_k} g(X_{t-}^\varepsilon, \theta)b(X_{t-}^\varepsilon, \sigma_0)\,\mathrm{d}W_t\, 1_{D_k^{n,\varepsilon,\rho}} \xrightarrow{p} 0 \qquad (3.17)
\end{aligned}$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$, $\lambda_\varepsilon^2/n \to 0$, $\varepsilon\lambda_\varepsilon \to 0$ and $\lambda_\varepsilon \int_{|z|\le 4v_2/c_1 n^\rho} f_{\alpha_0}(z)\,\mathrm{d}z \to 0$, uniformly in $\theta \in \Theta$. When we put

$$\tilde{C}_k^{n,\varepsilon,\rho} := C_k^{n,\varepsilon,\rho} \cap \left\{ \sup_{t\in[0,1]} |X_t^\varepsilon - x_t| < \delta \right\},$$

it holds from Morrey's inequality (see, *e.g.*, Theorem 5 in Evans [9, Section 5.6]) that for $q \in (d,\infty)$

$$\begin{aligned}
&\sum_{k=1}^n E\left[\sup_{\theta\in\Theta} \left| \int_{t_{k-1}}^{t_k} \left\{ g(X_{t_{k-1}}^\varepsilon,\theta) - g(X_{t-}^\varepsilon,\theta)\right\} b(X_{t-}^\varepsilon,\sigma_0)\,\mathrm{d}W_t\, 1_{\tilde{C}_k^{n,\varepsilon,\rho}} \right| \middle| \mathcal{F}_{t_{k-1}} \right] \\
&\quad \le C_1 \sum_{k=1}^n E\left[\left\| \int_{t_{k-1}}^{t_k} \left\{ g(X_{t_{k-1}}^\varepsilon,\theta) - g(X_{t-}^\varepsilon,\theta)\right\} b(X_{t-}^\varepsilon,\sigma_0)\,\mathrm{d}W_t\, 1_{\tilde{C}_k^{n,\varepsilon,\rho}} \right\|_{W^{1,q}(\Theta)} \middle| \mathcal{F}_{t_{k-1}} \right],
\end{aligned}$$

where the constant C_1 depends only on d, q and Θ. Then, it follows that

$$\begin{aligned}
&\sum_{k=1}^n E\left[\left\| \int_{t_{k-1}}^{t_k} \left\{ g(X_{t_{k-1}}^\varepsilon,\theta) - g(X_{t-}^\varepsilon,\theta)\right\} b(X_{t-}^\varepsilon,\sigma_0)\,\mathrm{d}W_t\, 1_{\tilde{C}_k^{n,\varepsilon,\rho}} \right\|_{L^q(\Theta)} \middle| \mathcal{F}_{t_{k-1}} \right] \\
&\le C_2 \sum_{k=1}^n \left(\int_\Theta E\left[\int_{t_{k-1}}^{t_k} \left| \left\{ g(X_{t_{k-1}}^\varepsilon,\theta) - g(X_t^\varepsilon,\theta)\right\} 1_{\tilde{C}_k^{n,\varepsilon,\rho}} b(X_t^\varepsilon,\sigma_0)\right|^2 \mathrm{d}t \,\middle|\, \mathcal{F}_{t_{k-1}} \right]^{q/2} \mathrm{d}\theta \right)^{1/q} \\
&\hspace{20em} (\because \text{Hölder's and Burkholder's inequalities}) \\
&\le \frac{C_3}{\sqrt{n}} \sum_{k=1}^n \left(E\left[\sup_{t\in[t_{k-1},t_k]} |X_{t_{k-1}}^\varepsilon - X_t^\varepsilon|^2 \left(1 + \sup_{t\in[t_{k-1},t_k]} |X_{t_{k-1}}^\varepsilon - X_t^\varepsilon|^2 + |X_{t_{k-1}}^\varepsilon|^2 \right) \middle| \mathcal{F}_{t_{k-1}} \right]^{q/2} \right)^{1/q} \\
&\hspace{8em} (\because \text{Hölder's inequality and the equi-Lipschitz continuity of } g \text{ and } b \text{ on } I_{x_0}^\delta) \\
&= O_p\left(\frac{1}{\sqrt{n}} + \varepsilon + \varepsilon\sqrt{\lambda_\varepsilon} \right) \hspace{14em} (\because \text{Lemmas 3.4.1 and 3.4.4})
\end{aligned}$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$ and $\varepsilon\lambda_\varepsilon \to 0$, where C_2 depends only on q, and C_3 depends only on q, b, g and Θ. By the same argument with Theorem B.4 in Prakasa Rao [26], it follows that

$$\begin{aligned}
&\sum_{k=1}^n E\left[\left\| \int_{t_{k-1}}^{t_k} \left\{ \frac{\partial g}{\partial\theta_j}\left(X_{t_{k-1}}^\varepsilon,\theta\right) - \frac{\partial g}{\partial\theta_j}\left(X_{t-}^\varepsilon,\theta\right)\right\} b(X_{t-}^\varepsilon,\sigma_0)\,\mathrm{d}W_t\, 1_{\tilde{C}_k^{n,\varepsilon,\rho}} \right\|_{L^q(\Theta)} \middle| \mathcal{F}_{t_{k-1}} \right] \\
&\quad = O_p\left(\frac{1}{\sqrt{n}} + \varepsilon + \varepsilon\sqrt{\lambda_\varepsilon} \right)
\end{aligned}$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$ and $\varepsilon\lambda_\varepsilon \to 0$. Thus, it follows from Lemma 3.7.3 that

$$\sum_{k=1}^n \int_{t_{k-1}}^{t_k} \left\{ g(X_{t_{k-1}}^\varepsilon,\theta) - g(X_{t-}^\varepsilon,\theta)\right\} b(X_{t-}^\varepsilon,\sigma_0)\,\mathrm{d}W_t\, 1_{\tilde{C}_k^{n,\varepsilon,\rho}} = O_p\left(\frac{1}{\sqrt{n}} + \varepsilon + \varepsilon\sqrt{\lambda_\varepsilon} \right)$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$ and $\varepsilon\lambda_\varepsilon \to 0$, uniformly in $\theta \in \Theta$, and therefore, from Lemma 3.4.3 we obtain the convergence of the first term in the left-hand side of (3.17). To obtain (3.17), we remain to prove

$$\sum_{k=1}^{n} \int_{t_{k-1}}^{t_k} g(X_{t-}^\varepsilon, \theta) b(X_{t-}^\varepsilon, \sigma_0)\, \mathrm{d}W_t\, 1_{D_k^{n,\varepsilon,\rho}} \xrightarrow{p} 0 \tag{3.18}$$

as $\lambda_\varepsilon \to \infty$, $\varepsilon\lambda_\varepsilon \to 0$, $\frac{\lambda_\varepsilon^2}{n} \to 0$, $\lambda_\varepsilon \int_{|z| \le 4v_2/c_1 n^\rho} f_{\alpha_0}(z)\,\mathrm{d}z \to 0$, uniformly in $\theta \in \Theta$. Put $\tilde{D}_{k,1}^{n,\varepsilon,\rho} := D_{k,1}^{n,\varepsilon,\rho} \cap \{X_t^\varepsilon \in I_{x_0}^\delta \text{ for all } t \in [0,1]\}$. We begin with showing that for any $p \in (2, \infty)$ and $q' \in (1, d/(d-1))$

$$\sum_{k=1}^{n} \int_{t_{k-1}}^{t_k} g(X_{t-}^\varepsilon, \theta) b(X_{t-}^\varepsilon, \sigma_0)\, \mathrm{d}W_t\, 1_{\tilde{D}_{k,1}^{n,\varepsilon,\rho}} = O_p\left(\frac{1}{\sqrt{n}} \left(\frac{\varepsilon^p \lambda_\varepsilon^2}{n} + \lambda_\varepsilon \right)^{1/2+1/q'} \right) \tag{3.19}$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$ and $\varepsilon\lambda_\varepsilon \to 0$, uniformly in $\theta \in \Theta$. It follows from Morrey's inequality (see, *e.g.*, Theorem 5 in Evans [9, Section 5.6]) that for $q \in (d, \infty)$

$$\begin{aligned}
&\sum_{k=1}^{n} E\left[\sup_{\theta \in \Theta} \left| \int_{t_{k-1}}^{t_k} g(X_{t-}^\varepsilon, \theta) b(X_{t-}^\varepsilon, \sigma_0)\, \mathrm{d}W_t\, 1_{\tilde{D}_{k,1}^{n,\varepsilon,\rho}} \right| \middle| \mathcal{F}_{t_{k-1}} \right] \\
&\qquad \le C_1 \sum_{k=1}^{n} E\left[\left\| \int_{t_{k-1}}^{t_k} g(X_{t-}^\varepsilon, \theta) b(X_{t-}^\varepsilon, \sigma_0)\, \mathrm{d}W_t \right\|_{W^{1,q}(\Theta)} 1_{\tilde{D}_{k,1}^{n,\varepsilon,\rho}} \middle| \mathcal{F}_{t_{k-1}} \right],
\end{aligned}$$

where the constant C_1 depends only on d, q and Θ. If we put $q' = q/(q-1)$, then it follows from Hölder's inequality, Burkholder's inequality (see, *e.g.*, Theorem 4.4.21 in Applebaum [3]), the equicontinuity of g and Assumption 3.2.1 that

$$\begin{aligned}
&\sum_{k=1}^{n} E\left[\left\| \int_{t_{k-1}}^{t_k} g(X_{t-}^\varepsilon, \theta) b(X_{t-}^\varepsilon, \sigma_0)\, \mathrm{d}W_t \right\|_{L^q(\Theta)} 1_{\tilde{D}_{k,1}^{n,\varepsilon,\rho}} \middle| \mathcal{F}_{t_{k-1}} \right] \\
&\quad \le \sum_{k=1}^{n} \left(\int_\Theta E\left[\left| \int_{t_{k-1}}^{t_k} g(X_{t-}^\varepsilon, \theta) b(X_{t-}^\varepsilon, \sigma_0) 1_{\tilde{D}_{k,1}^{n,\varepsilon,\rho}}\, \mathrm{d}W_t \right|^q \middle| \mathcal{F}_{t_{k-1}} \right] \mathrm{d}\theta \right)^{1/q} P\left(\tilde{D}_{k,1}^{n,\varepsilon,\rho} \middle| \mathcal{F}_{t_{k-1}} \right)^{1/q'} \\
&\quad \le C_2 \sum_{k=1}^{n} \left(\int_\Theta E\left[\int_{t_{k-1}}^{t_k} \left| g(X_t^\varepsilon, \theta) b(X_t^\varepsilon, \sigma_0) 1_{\tilde{D}_{k,1}^{n,\varepsilon,\rho}} \right|^2 \mathrm{d}t \middle| \mathcal{F}_{t_{k-1}} \right]^{q/2} \mathrm{d}\theta \right)^{1/q} P\left(\tilde{D}_{k,1}^{n,\varepsilon,\rho} \middle| \mathcal{F}_{t_{k-1}} \right)^{1/q'} \\
&\quad \le C_2 \sup_{(x,\theta) \in I_{x_0}^\delta \times \Theta} |g(x,\theta) b(x,\sigma_0)| \frac{|\Theta|^{1/q}}{n^{1/2}} \sum_{k=1}^{n} P\left(\tilde{D}_{k,1}^{n,\varepsilon,\rho} \middle| \mathcal{F}_{t_{k-1}} \right)^{1/2+1/q'},
\end{aligned}$$

where C_2 depends only on q. By using Lemmas 3.4.4 and 3.4.7, for any $p > 2$ we obtain

$$\begin{aligned}
&\sum_{k=1}^{n} E\left[\left\| \int_{t_{k-1}}^{t_k} g(X_{t-}^\varepsilon, \theta) b(X_{t-}^\varepsilon, \sigma_0)\, \mathrm{d}W_t \right\|_{L^q(\Theta)} 1_{\tilde{D}_{k,1}^{n,\varepsilon,\rho}} \middle| \mathcal{F}_{t_{k-1}} \right] \\
&\qquad = O_p\left(\sqrt{n} \left\{ \frac{\lambda_\varepsilon}{n} \left(\frac{1}{n^{p(1-\rho)}} + \frac{\varepsilon^p}{n^{p(1/2-\rho)}} + \frac{\varepsilon^p \lambda_\varepsilon}{n} \right) + \frac{\lambda_\varepsilon}{n} \right\}^{1/2+1/q'} \right)
\end{aligned}$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$ and $\varepsilon\lambda_\varepsilon \to 0$. Similarly, by using Theorem B.4 in Prakasa Rao [26], we obtain for $j = 1, \ldots, d$

$$\sum_{k=1}^{n} E\left[\left\|\int_{t_{k-1}}^{t_k} \frac{\partial g}{\partial\theta_j}(X_{t-}^\varepsilon, \theta) b(X_{t-}^\varepsilon, \sigma_0)\,\mathrm{d}W_t\right\|_{L^q(\Theta)} 1_{\tilde{D}_{k,1}^{n,\varepsilon,\rho}} \middle| \mathcal{F}_{t_{k-1}}\right]$$
$$= O_p\left(\sqrt{n}\left\{\frac{\lambda_\varepsilon}{n}\left(\frac{1}{n^{p(1-\rho)}} + \frac{\varepsilon^p}{n^{p(1/2-\rho)}} + \frac{\varepsilon^p\lambda_\varepsilon}{n}\right) + \frac{\lambda_\varepsilon}{n}\right\}^{1/2+1/q'}\right)$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$ and $\varepsilon\lambda_\varepsilon \to 0$. Since we can take $q' < 2$ small enough, we obtain (3.19) from Remark 3.7.3. Hence, (3.18) holds from (3.19) and Lemma 3.4.10.

At last, it is an immediate consequence from Lemma 3.4.9 that

$$\sum_{k=1}^{n} g(X_{t_{k-1}}^\varepsilon, \theta)\int_{t_{k-1}}^{t_k} c(X_{t-}^\varepsilon, \alpha_0)\,\mathrm{d}Z_t^{\lambda_\varepsilon}\, 1_{C_k^{n,\varepsilon,\rho}} \xrightarrow{p} 0$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$, $\lambda_\varepsilon^2/n \to 0$, $\varepsilon\lambda_\varepsilon \to 0$ and $\lambda_\varepsilon \int_{|z|\le 4v_2/c_1 n^\rho} f_{\alpha_0}(z)\,\mathrm{d}z \to 0$, uniformly in $\theta \in \Theta$. □

Lemma 3.4.15. *Under Assumptions 3.2.1, 3.2.3, 3.2.4, 3.2.6 and 3.2.8, let $\rho \in (0, 1/2)$. and let $g : \mathbb{R} \times \Theta \to \mathbb{R}$ satisfy that $\left\{\frac{\partial g}{\partial\theta_i}(\cdot, \theta)\right\}_{\theta\in\Theta}$ ($i = 1, \ldots, d$) are equicontinuous on $I_{x_0}^\delta$ for some small $\delta > 0$. Then,*

$$\frac{1}{\varepsilon^2}\sum_{k=1}^{n} g(X_{t_{k-1}}^\varepsilon, \theta)\left|\Delta_k^n X^\varepsilon - \frac{1}{n}a(X_{t_{k-1}}^\varepsilon, \mu_0)\right|^2 1_{C_k^{n,\varepsilon,\rho}} \xrightarrow{p} \int_0^1 g(x_t, \theta)|b(x_t, \sigma_0)|^2\,\mathrm{d}t$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$, $\varepsilon\lambda_\varepsilon \to 0$, $\lambda_\varepsilon^2/n \to 0$ and $\lambda_\varepsilon \int_{|z|\le 4v_2/c_1 n^\rho} f_{\alpha_0}(z)\,\mathrm{d}z \to 0$, uniformly in $\theta \in \Theta$.

Proof. From Lemma 3.4.9, it is sufficient to show that

$$\frac{1}{\varepsilon^2}\sum_{k=1}^{n} g(X_{t_{k-1}}^\varepsilon, \theta)\left|\Delta_k^n X^\varepsilon - \frac{1}{n}a(X_{t_{k-1}}^\varepsilon, \mu_0)\right|^2 1_{C_{k,0}^{n,\varepsilon,\rho}} \xrightarrow{p} \int_0^1 g(x_t, \theta)|b(x_t, \sigma_0)|^2\,\mathrm{d}t$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$, $\varepsilon\lambda_\varepsilon \to 0$, $\lambda_\varepsilon^2/n \to 0$ and $\lambda_\varepsilon \int_{|z|\le 4v_2/c_1 n^\rho} f_{\alpha_0}(z)\,\mathrm{d}z \to 0$, uniformly in $\theta \in \Theta$, and we note that

$$\left|\Delta_k^n X^\varepsilon - \frac{1}{n}a(X_{t_{k-1}}^\varepsilon, \mu_0)\right|^2 1_{J_{k,0}^{n,\varepsilon}} = \left\{\left|\int_{t_{k-1}}^{t_k}\left\{a(X_t^\varepsilon, \mu_0) - a(X_{t_{k-1}}^\varepsilon, \mu_0)\right\}\mathrm{d}t\right|^2 + \left|\int_{t_{k-1}}^{t_k} b(X_{t-}^\varepsilon, \sigma_0)\,\mathrm{d}W_t\right|^2\right.$$
$$\left. +2\int_{t_{k-1}}^{t_k}\left\{a(X_t^\varepsilon, \mu_0) - a(X_{t_{k-1}}^\varepsilon, \mu_0)\right\}\mathrm{d}t\int_{t_{k-1}}^{t_k} b(X_{t-}^\varepsilon, \sigma_0)\,\mathrm{d}W_t\right\} 1_{J_{k,0}^{n,\varepsilon}}.$$

Similarly to the proof of (3.16), it follows that

$$\sup_{\theta\in\Theta}\left|\frac{1}{\varepsilon^2}\sum_{k=1}^{n} g(X_{t_{k-1}}^\varepsilon, \theta)\left|\int_{t_{k-1}}^{t_k}\left\{a(X_t^\varepsilon, \mu_0) - a(X_{t_{k-1}}^\varepsilon, \mu_0)\right\}\mathrm{d}t\right|^2 1_{C_{k,0}^{n,\varepsilon,\rho}}\right| = O_p\left(\frac{1}{\varepsilon^2 n^3} + \frac{1}{n^2}\right)$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$, $\varepsilon n \to \infty$, and $\varepsilon\lambda_\varepsilon \to 0$. Also, it holds that

$$\sup_{\theta\in\Theta}\left|\frac{2}{\varepsilon}\sum_{k=1}^n g(X^\varepsilon_{t_{k-1}},\theta)\int_{t_{k-1}}^{t_k}\left\{a(X^\varepsilon_t,\mu_0)-a(X^\varepsilon_{t_{k-1}},\mu_0)\right\}\mathrm{d}t \times\int_{t_{k-1}}^{t_k} b(X^\varepsilon_{t-},\sigma_0)\,\mathrm{d}W_t\ 1_{C^{n,\varepsilon,\rho}_{k,0}}\right| = O_p\left(\frac{1}{\varepsilon n^{3/2}}+\frac{1}{n}\right)$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$ and $\varepsilon\lambda_\varepsilon \to 0$, uniformly in $\theta \in \Theta$. Indeed, by using Assumption 3.2.1, Hölder's inequality and Burkholder's inequality, we obtain

$$\begin{aligned}
&\frac{2}{\varepsilon}\sum_{k=1}^n E\left[\sup_{\theta\in\Theta}\left|g(X^\varepsilon_{t_{k-1}},\theta)\int_{t_{k-1}}^{t_k}\left\{a(X^\varepsilon_t,\mu_0)-a(X^\varepsilon_{t_{k-1}},\mu_0)\right\}\mathrm{d}t\int_{t_{k-1}}^{t_k} b(X^\varepsilon_{t-},\sigma_0)\,\mathrm{d}W_t\ 1_{C^{n,\varepsilon,\rho}_{k,0}}\right|\,\middle|\,\mathcal{F}_{t_{k-1}}\right]\\
&\quad\le \frac{2C}{n}\sum_{k=1}^n \sup_{\theta\in\Theta}|g(X^\varepsilon_{t_{k-1}},\theta)|\left(E\left[\frac{1}{\varepsilon^2}\sup_{t\in[t_{k-1},\tau_k]}|X^\varepsilon_t - X^\varepsilon_{t_{k-1}}|^2\,\middle|\,\mathcal{F}_{t_{k-1}}\right]\right)^{1/2}\\
&\qquad\times\left(\frac{1}{n}E\left[\sup_{t\in[t_{k-1},\tau_k]}|X^\varepsilon_t - X^\varepsilon_{t_{k-1}}|^2 + |b(X^\varepsilon_{t_{k-1}},\sigma_0)|^2\,\middle|\,\mathcal{F}_{t_{k-1}}\right]\right)^{1/2}
\end{aligned}$$

where C depends only on a, b. By applying Lemmas 3.4.3 to 3.4.5 and 3.7.3 and the boundedness of g on $I^\delta_{x_0}\times\Theta$ for some small $\delta > 0$, we obtain the above convergence.

From Lemma 3.4.11, we remain to prove that

$$\sum_{k=1}^n g(X^\varepsilon_{t_{k-1}},\theta)\left|\int_{t_{k-1}}^{t_k} b(X^\varepsilon_{t-},\sigma_0)\,\mathrm{d}W_t\right|^2 1_{C^{n,\varepsilon,\rho}_k}\xrightarrow{p}\int_0^1 g(x_t,\theta)|b(x_t,\sigma_0)|^2\,\mathrm{d}t$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$ and $\varepsilon\lambda_\varepsilon \to 0$, uniformly in $\theta \in \Theta$. At first, by using Lemma 3.4.4, we have

$$\begin{aligned}
\sum_{k=1}^n E\left[g(X^\varepsilon_{t_{k-1}},\theta)\left|\int_{t_{k-1}}^{t_k} b(X^\varepsilon_{t_{k-1}},\sigma_0)\,\mathrm{d}W_t\right|^2\,\middle|\,\mathcal{F}_{t_{k-1}}\right] &= \frac{1}{n}\sum_{k=1}^n g(X^\varepsilon_{t_{k-1}},\theta)|b(X^\varepsilon_{t_{k-1}},\sigma_0)|^2\\
&\xrightarrow{p}\int_0^1 g(x_t,\theta)|b(x_t,\sigma_0)|^2\,\mathrm{d}t
\end{aligned}$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$ and $\varepsilon\lambda_\varepsilon \to 0$, and

$$\sum_{k=1}^n E\left[\left|g(X^\varepsilon_{t_{k-1}},\theta)\left|\int_{t_{k-1}}^{t_k} b(X^\varepsilon_{t_{k-1}},\sigma_0)\,\mathrm{d}W_t\right|^2\right|^2\,\middle|\,\mathcal{F}_{t_{k-1}}\right]\xrightarrow{p} 0$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$ and $\varepsilon\lambda_\varepsilon \to 0$. Thus, by Lemma 9 in Genon-Catalot and Jacod [11], we obtain

$$\sum_{k=1}^n g(X^\varepsilon_{t_{k-1}},\theta)\left|\int_{t_{k-1}}^{t_k} b(X^\varepsilon_{t_{k-1}},\sigma_0)\,\mathrm{d}W_t\right|^2\xrightarrow{p}\int_0^1 g(x_t,\theta)|b(x_t,\theta)|^2\,\mathrm{d}t$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$ and $\varepsilon\lambda_\varepsilon \to 0$. From the equidifferentiablities of g on $I_{x_0}^\delta$ for some $\delta > 0$, the uniform tightness is shown by the same argument in the proof of Lemma 3.4.13. At second, we shall see

$$\sum_{k=1}^n g(X_{t_{k-1}}^\varepsilon, \theta) \left\{ \left| \int_{t_{k-1}}^{t_k} b(X_{t-}^\varepsilon, \sigma_0)\, \mathrm{d}W_t \right|^2 - \left| \int_{t_{k-1}}^{t_k} b(X_{t_{k-1}}^\varepsilon, \sigma_0)\, \mathrm{d}W_t \right|^2 \right\} 1_{C_{k,0}^{n,\varepsilon,\rho}} \xrightarrow{p} 0$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$, $\varepsilon\lambda_\varepsilon \to 0$ and $\lambda_\varepsilon/n \to 0$, uniformly in $\theta \in \Theta$. This convergence is obtained from Lemma 3.7.3 and the following estimate:

$$\begin{aligned}
&\sum_{k=1}^n E\left[\sup_{\theta\in\Theta} \left| g(X_{t_{k-1}}^\varepsilon, \theta) \left\{ \left| \int_{t_{k-1}}^{t_k} b(X_{t-}^\varepsilon, \sigma_0)\, \mathrm{d}W_t \right|^2 - \left| \int_{t_{k-1}}^{t_k} b(X_{t_{k-1}}^\varepsilon, \sigma_0)\, \mathrm{d}W_t \right|^2 \right\} 1_{J_{k,0}^{n,\varepsilon}} \right| \,\middle|\, \mathcal{F}_{t_{k-1}} \right] \\
&\le \sum_{k=1}^n \sup_{\theta\in\Theta} |g(X_{t_{k-1}}^\varepsilon, \theta)| \left(E\left[\left| \int_{t_{k-1}}^{t_k} \{ b(X_{t-}^\varepsilon, \sigma_0) + b(X_{t_{k-1}}^\varepsilon, \sigma_0) \}\, \mathrm{d}W_t \right|^2 1_{J_{k,0}^{n,\varepsilon}} \,\middle|\, \mathcal{F}_{t_{k-1}} \right] \right)^{1/2} \\
&\quad \times \left(E\left[\left| \int_{t_{k-1}}^{t_k} \{ b(X_{t-}^\varepsilon, \sigma_0) - b(X_{t_{k-1}}^\varepsilon, \sigma_0) \}\, \mathrm{d}W_t \right|^2 1_{J_{k,0}^{n,\varepsilon}} \,\middle|\, \mathcal{F}_{t_{k-1}} \right] \right)^{1/2} \quad (\because \text{Hölder's inequality}) \\
&\le \sum_{k=1}^n \sup_{\theta\in\Theta} |g(X_{t_{k-1}}^\varepsilon, \theta)| \left(E\left[\int_{t_{k-1}}^{t_k} |b(X_t^\varepsilon, \sigma_0) + b(X_{t_{k-1}}^\varepsilon, \sigma_0)|^2 1_{J_{k,0}^{n,\varepsilon}}\, \mathrm{d}t \,\middle|\, \mathcal{F}_{t_{k-1}} \right] \right)^{1/2} \\
&\quad \times \left(E\left[\int_{t_{k-1}}^{t_k} |b(X_t^\varepsilon, \sigma_0) - b(X_{t_{k-1}}^\varepsilon, \sigma_0)|^2 1_{J_{k,0}^{n,\varepsilon}}\, \mathrm{d}t \,\middle|\, \mathcal{F}_{t_{k-1}} \right] \right)^{1/2} \quad (\because \text{Burkholder's inequality}) \\
&\le C\frac{1}{n} \sum_{k=1}^n \sup_{\theta\in\Theta} |g(X_{t_{k-1}}^\varepsilon, \theta)| \left(E\left[\sup_{t\in[t_{k-1},\tau_k]} (1 + |X_t^\varepsilon - X_{t_{k-1}}^\varepsilon|^2 + |X_{t_{k-1}}^\varepsilon|^2) \,\middle|\, \mathcal{F}_{t_{k-1}} \right] \right)^{1/2} \\
&\quad \times \left(E\left[\sup_{t\in[t_{k-1},\tau_k]} |X_t^\varepsilon - X_{t_{k-1}}^\varepsilon|^2 \,\middle|\, \mathcal{F}_{t_{k-1}} \right] \right)^{1/2} \quad (\because b \text{ is Lipschitz}) \\
&= O_p\left(\frac{1}{n} + \frac{\varepsilon}{\sqrt{n}} \right) \quad (\because \text{Lemmas 3.4.4 and 3.4.5})
\end{aligned}$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$ and $\varepsilon\lambda_\varepsilon \to 0$.

At last, since

$$\sup_{k=1,\dots,n} n \left| \int_{t_{k-1}}^{t_k} \mathrm{d}W_t \right|^2$$

is bounded in probability, it follows from Lemmas 3.4.1, 3.4.8 and 3.4.9 and the linearity of b that

$$\sum_{k=1}^n g(X_{t_{k-1}}^\varepsilon, \theta) \left| \int_{t_{k-1}}^{t_k} b(X_{t_{k-1}}^\varepsilon, \sigma_0)\, \mathrm{d}W_t \right|^2 1_{D_k^{n,\varepsilon,\rho} \cup C_{k,1}^{n,\varepsilon,\rho} \cup C_{k,2}^{n,\varepsilon,\rho}} \xrightarrow{p} 0$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$, $\varepsilon\lambda_\varepsilon \to 0$ and $\lambda_\varepsilon/n \to 0$, uniformly in $\theta \in \Theta$. □

3.4.3 Proof of main results

Proof of Theorem 3.3.1

Proof of Theorem 3.3.1. It follows from Lemmas 3.4.11 and 3.4.14 that

$$\begin{aligned}\Phi^{(1)}_{n,\varepsilon}(\mu,\sigma) &:= n\varepsilon^2\big(\Psi^{(1)}_{n,\varepsilon}(\mu,\sigma)-\Psi^{(1)}_{n,\varepsilon}(\mu_0,\sigma)\big)\\ &= \sum_{k=1}^n \frac{\Big(\Delta^n_k X^\varepsilon - a(X^\varepsilon_{t_{k-1}},\mu_0)/n\Big)\Big(a(X^\varepsilon_{t_{k-1}},\mu)-a(X^\varepsilon_{t_{k-1}},\mu_0)\Big)}{\big|b(X^\varepsilon_{t_{k-1}},\sigma)\big|^2} 1_{C^{n,\varepsilon,\rho}_k}\\ &\quad - \frac{1}{2n}\sum_{k=1}^n \frac{\Big|a(X^\varepsilon_{t_{k-1}},\mu)-a(X^\varepsilon_{t_{k-1}},\mu_0)\Big|^2}{\big|b(X^\varepsilon_{t_{k-1}},\sigma)\big|^2} 1_{C^{n,\varepsilon,\rho}_k} \xrightarrow{p} -\frac{1}{2}\int_0^1 \frac{|a(x_t,\mu)-a(x_t,\mu_0)|^2}{|b(x_t,\sigma)|^2}\,\mathrm{d}t\end{aligned}$$

as $n\to\infty$, $\varepsilon\to 0$, $\lambda_\varepsilon\to\infty$, $\lambda_\varepsilon^2/n\to 0$, $\varepsilon\lambda_\varepsilon\to 0$ and $\lambda_\varepsilon\int_{|z|\le 4v_2/c_1 n^\rho} f_{\alpha_0}(z)\,\mathrm{d}z\to 0$, uniformly in $(\mu,\sigma)\in\bar\Theta_1\times\bar\Theta_2$, and from Lemmas 3.4.11, 3.4.14 and 3.4.15 that

$$\begin{aligned}&\Psi^{(1)}_{n,\varepsilon}(\mu,\sigma)\\ &= \frac{1}{\varepsilon^2 n}\Phi^{(1)}_{n,\varepsilon}(\mu,\sigma)+\Psi^{(1)}_{n,\varepsilon}(\mu_0,\sigma)\\ &= \frac{1}{\varepsilon^2 n}\Phi^{(1)}_{n,\varepsilon}(\mu,\sigma) - \frac{1}{n}\sum_{k=1}^n\left\{\frac{\Big|\Delta^n_k X^\varepsilon - a(X^\varepsilon_{t_{k-1}},\mu_0)/n\Big|^2}{2\frac{1}{n}\Big|\varepsilon b(X^\varepsilon_{t_{k-1}},\sigma)\Big|^2}+\frac{1}{2}\log|b(X^\varepsilon_{t_{k-1}},\sigma)|^2\right\}1_{C^{n,\varepsilon,\rho}_k}\\ &\xrightarrow{p} -\left(\lim_{\substack{n\to\infty\\ \varepsilon\to 0}}\frac{1}{\varepsilon^2 n}\right)\int_0^1 \frac{|a(x_t,\mu)-a(x_t,\mu_0)|^2}{2\,|b(x_t,\sigma)|^2}\,\mathrm{d}t - \frac{1}{2}\int_0^1\left|\frac{b(x_t,\sigma_0)}{b(x_t,\sigma)}\right|^2\mathrm{d}t - \frac{1}{2}\int_0^1 \log|b(x_t,\sigma)|^2\,\mathrm{d}t\end{aligned}$$

as $n\to\infty$, $\varepsilon\to 0$, $\lambda_\varepsilon\to\infty$, $\lambda_\varepsilon^2/n\to 0$, $\varepsilon\lambda_\varepsilon\to 0$ and $\lambda_\varepsilon\int_{|z|\le 4v_2/c_1 n^\rho} f_{\alpha_0}(z)\,\mathrm{d}z\to 0$, uniformly in $(\mu,\sigma)\in\bar\Theta_1\times\bar\Theta_2$, Also, it follows from Lemmas 3.4.12 and 3.4.13 that

$$\begin{aligned}\Psi^{(2)}_{n,\varepsilon}(\alpha) &= \frac{1}{\lambda_\varepsilon}\sum_{k=1}^n \psi\left(X^\varepsilon_{t_{k-1}},\frac{\Delta^n_k X^\varepsilon}{\varepsilon},\alpha\right)1_{D^{n,\varepsilon,\rho}_k}\\ &\xrightarrow{p}\int_0^1\int_{-\infty}^\infty \frac{1}{c(x_t,\alpha_0)}f_{\alpha_0}\left(\frac{y}{c(x_t,\alpha_0)}\right)\log\left\{\frac{1}{c(x_t,\alpha)}f_\alpha\left(\frac{y}{c(x_t,\alpha)}\right)\right\}\mathrm{d}y\,\mathrm{d}t\end{aligned}$$

as $n\to\infty$, $\varepsilon\to 0$, $\lambda_\varepsilon\to\infty$, $\lambda_\varepsilon^2/n\to 0$, $\varepsilon\lambda_\varepsilon\to 0$ and $\lambda_\varepsilon\int_{|z|\le 4v_2/c_1 n^\rho} f_{\alpha_0}(z)\,\mathrm{d}z\to 0$, uniformly in $\alpha\in\bar\Theta_3$. Thus, by using usual argument (see, *e.g.*, the proof of Theorem 1 in Sørensen and Uchida [33]), the consistency of $\hat\theta_{n,\varepsilon}$ holds under Assumption 3.2.7. □

Proof of Theorem 3.3.2

To establish the proof of this theorem, we set up random variables $\xi_{\ell k}^i$, $\tilde{\xi}_{\ell k}^i$ ($\ell = 1, \dots, 3$, $i = 1, \dots, d_\ell$, $k = 1, \dots, n$) as the followings:

$$\begin{aligned}
\sqrt{\varepsilon^{-2}} \frac{\partial \Phi_{n,\varepsilon}^{(1)}}{\partial \mu_i}(\mu, \sigma)\bigg|_{\theta=\theta_0} &= -\frac{1}{\varepsilon} \sum_{k=1}^n \frac{\left\{\Delta_k^n X^\varepsilon - \frac{1}{n} a(X_{t_{k-1}}^\varepsilon, \mu)\right\} \frac{\partial a}{\partial \mu_i}\left(X_{t_{k-1}}^\varepsilon, \mu\right)}{\left|b(X_{t_{k-1}}^\varepsilon, \sigma)\right|^2} 1_{C_k^{n,\varepsilon,\rho}} \Bigg|_{\theta=\theta_0} \\
&=: \sum_{k=1}^n \xi_{1,k}^i \left(\xrightarrow{p} \int_0^1 \frac{\frac{\partial a}{\partial \mu_i}(x_t, \mu_0)}{b(x_t, \sigma_0)} \, \mathrm{d}W_t \right), \quad (\because \text{Lemma 3.4.14}) \\
\sqrt{n}\, \frac{\partial \Psi_{n,\varepsilon}^{(1)}}{\partial \sigma_i}(\mu, \sigma)\bigg|_{\theta=\theta_0} &= -\frac{1}{\sqrt{n}} \sum_{k=1}^n \left\{ -\frac{\left|\Delta_k^n X^\varepsilon - \frac{1}{n} a(X_{t_{k-1}}^\varepsilon, \mu)\right|^2}{\frac{1}{n}\left|\varepsilon b(X_{t_{k-1}}^\varepsilon, \sigma)\right|^2} + 1 \right\} \frac{\frac{\partial b}{\partial \sigma_i}\left(X_{t_{k-1}}^\varepsilon, \sigma\right)}{b(X_{t_{k-1}}^\varepsilon, \sigma)} 1_{C_k^{n,\varepsilon,\rho}} \Bigg|_{\theta=\theta_0} \\
&=: \sum_{k=1}^n \xi_{2,k}^i, \\
\sqrt{\lambda_\varepsilon} \frac{\partial \Psi_{n,\varepsilon}^{(2)}}{\partial \alpha_i}(\alpha)\bigg|_{\alpha=\alpha_0} &= \frac{1}{\sqrt{\lambda_\varepsilon}} \sum_{k=1}^n \frac{\partial \psi}{\partial \alpha_i}\left(X_{t_{k-1}}^\varepsilon, \frac{\Delta_k^n X^\varepsilon}{\varepsilon}, \alpha_0\right) 1_{D_k^{n,\varepsilon,\rho}} =: \sum_{k=1}^n \xi_{3,k}^i,
\end{aligned}$$

and

$$\begin{aligned}
\sum_{k=1}^n \tilde{\xi}_{1,k}^i &:= \sum_{k=1}^n \frac{\frac{\partial a}{\partial \mu_i}\left(X_{t_{k-1}}^\varepsilon, \mu_0\right)}{b(X_{t_{k-1}}^\varepsilon, \sigma_0)} \int_{t_{k-1}}^{t_k} \mathrm{d}W_t \, 1_{C_{k,0}^{n,\varepsilon,\rho}}, \\
\sum_{k=1}^n \tilde{\xi}_{2,k}^i &:= -\sqrt{n} \sum_{k=1}^n \left\{ -\left| \int_{t_{k-1}}^{t_k} \mathrm{d}W_t \right|^2 + \frac{1}{n} \right\} \frac{\frac{\partial b}{\partial \sigma_i}\left(X_{t_{k-1}}^\varepsilon, \sigma_0\right)}{b(X_{t_{k-1}}^\varepsilon, \sigma_0)} 1_{C_{k,0}^{n,\varepsilon,\rho}}, \\
\sum_{k=1}^n \tilde{\xi}_{3,k}^i &:= \sum_{k=1}^n \frac{1}{\sqrt{\lambda_\varepsilon}} \frac{\partial \psi}{\partial \alpha_i}\left(X_{t_{k-1}}^\varepsilon, c(X_{t_{k-1}}^\varepsilon, \alpha_0) V_{N_{\tau_k}^{\lambda_\varepsilon}}, \alpha_0\right) 1_{J_{k,1}^{n,\varepsilon}}.
\end{aligned}$$

Lemma 3.4.16. *Under Assumptions 3.2.1 to 3.2.6, 3.2.8 and 3.2.10, the following convergences are holds.*

For $\ell = 1, 2$

$$\sum_{k=1}^n \xi_{\ell k}^i - \sum_{k=1}^n \tilde{\xi}_{\ell k}^i \xrightarrow{p} 0 \quad (i = 1, \dots, d_\ell)$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$, $\varepsilon n \to \infty$, $\varepsilon \lambda_\varepsilon \to 0$, $\lambda_\varepsilon^2 / n \to 0$ *and* $\lambda_\varepsilon \int_{|z| \le 4v_2/c_1 n^\rho} f_{\alpha_0}(z) \, \mathrm{d}z \to 0$.

For $\ell = 3$, *take* ρ *as either of the following:*

(i) Under Assumption 3.2.4 (i), take $\rho \in (0, 1/2)$.

(ii) Under Assumption 3.2.4 (ii), take $\rho \in (0, \min\{1/2, 1/4q\})$, *where* q *is the constant given in Assumption 3.2.10 (ii.b).*

Then,

$$\sum_{k=1}^n \xi_{\ell k}^i - \sum_{k=1}^n \tilde{\xi}_{\ell k}^i \xrightarrow{p} 0 \quad (i = 1, \ldots, d_\ell)$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$, $\varepsilon\lambda_\varepsilon \to 0$, $\lambda_\varepsilon^2/n \to 0$ *and* $\lambda_\varepsilon \int_{|z|\le 4v_2/c_1 n^\rho} f_{\alpha_0}(z)\,\mathrm{d}z \to 0$ *with* $\lim(\varepsilon^2 n)^{-1} < \infty$.

Proof. For $\ell = 1, 2$, from Lemmas 3.4.9 and 3.7.3, it is suffcient to show that for $\rho \in (0, 1/2)$

$$\sum_{k=1}^n E\left[\left|\xi_{\ell k}^i 1_{J_{k,0}^{n,\varepsilon}} - \tilde{\xi}_{\ell k}^i\right| \,\middle|\, \mathcal{F}_{t_{k-1}}\right] \xrightarrow{p} 0$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$, $\varepsilon n \to \infty$, $\varepsilon\lambda_\varepsilon \to 0$, $\lambda_\varepsilon^2/n \to 0$ and $\lambda_\varepsilon \int_{|z|\le 4v_2/c_1 n^\rho} f_{\alpha_0}(z)\,\mathrm{d}z \to 0$.

For $\ell = 1$, let $i \in \{1, \ldots, d_1\}$, and put $g(x) = \frac{\partial a}{\partial \mu_i}(x, \mu_0)/|b(x, \sigma_0)|^2$. Then,

$$\begin{aligned}\xi_{1,k}^i 1_{J_{k,0}^{n,\varepsilon}} - \tilde{\xi}_{1,k}^i = g\Big(X_{t_{k-1}}^\varepsilon\Big) \Bigg\{ &\int_{t_{k-1}}^{t_k} a(X_t^\varepsilon, \mu_0)\,\mathrm{d}t - \frac{1}{n} a(X_{t_{k-1}}^\varepsilon, \mu_0) \\ &+ \varepsilon \int_{t_{k-1}}^{t_k} \{b(X_t^\varepsilon, \sigma_0) - b(X_{t_{k-1}}^\varepsilon, \sigma_0)\}\,\mathrm{d}W_t \Bigg\} 1_{C_{k,0}^{n,\varepsilon,\rho}}.\end{aligned}$$

As in the same argument in Lemma 3.4.14, it holds from Assumptions 3.2.1 to 3.2.3 and Lemmas 3.4.4 and 3.4.5 that

$$\frac{1}{\varepsilon}\sum_{k=1}^n E\left[|g(X_{t_{k-1}}^\varepsilon)| \left|\int_{t_{k-1}}^{t_k} a(X_t^\varepsilon, \mu_0)\,\mathrm{d}t - \frac{1}{n} a(X_{t_{k-1}}^\varepsilon, \mu_0)\right| 1_{C_{k,0}^{n,\varepsilon,\rho}} \,\middle|\, \mathcal{F}_{t_{k-1}}\right] = O_p\left(\frac{1}{\varepsilon n} + \frac{1}{\sqrt{n}}\right)$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$ and $\varepsilon\lambda_\varepsilon \to 0$, and from Assumption 3.2.1, Burkholder's inequality, Lemmas 3.4.4 and 3.4.5 that

$$\begin{aligned}&\sum_{k=1}^n E\left[|g(X_{t_{k-1}}^\varepsilon)| \left|\int_{t_{k-1}}^{t_k} \{b(X_t^\varepsilon, \sigma_0) - b(X_{t_{k-1}}^\varepsilon, \sigma_0)\}\,\mathrm{d}W_t\right| 1_{C_{k,0}^{n,\varepsilon,\rho}} \,\middle|\, \mathcal{F}_{t_{k-1}}\right] \\ &\quad \le \frac{C}{\sqrt{n}} \sum_{k=1}^n |g(X_{t_{k-1}}^\varepsilon)| \left(E\left[\sup_{t\in[t_{k-1},\tau_k]} |X_t^\varepsilon - X_{t_{k-1}}^\varepsilon|^2 \,\middle|\, \mathcal{F}_{t_{k-1}}\right]\right)^{1/2} = O_p\left(\frac{1}{\sqrt{n}} + \varepsilon\right)\end{aligned}$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$ and $\varepsilon\lambda_\varepsilon \to 0$.

For $\ell = 2$, let $i \in \{1, \ldots, d_2\}$, and put $g(x) = -\frac{1}{|b|^3}\frac{\partial b}{\partial \sigma_i}(x, \sigma_0)$. Then, we have

$$\begin{aligned}\xi_{2,k}^i 1_{J_{k,0}^{n,\varepsilon}} - \tilde{\xi}_{2,k}^i = g\Big(X_{t_{k-1}}^\varepsilon\Big) \Bigg\{ &\left|\int_{t_{k-1}}^{t_k} \Big\{a(X_t^\varepsilon, \mu_0) - a(X_{t_{k-1}}^\varepsilon, \mu_0)\Big\}\,\mathrm{d}t\right|^2 \\ &+ 2\varepsilon \int_{t_{k-1}}^{t_k} \Big\{a(X_t^\varepsilon, \mu_0) - a(X_{t_{k-1}}^\varepsilon, \mu_0)\Big\}\,\mathrm{d}t \int_{t_{k-1}}^{t_k} b(X_{t-}^\varepsilon, \sigma_0)\,\mathrm{d}W_t \\ &+ \left|\varepsilon \int_{t_{k-1}}^{t_k} b(X_{t-}^\varepsilon, \sigma_0)\,\mathrm{d}W_t\right|^2 - \left|\varepsilon \int_{t_{k-1}}^{t_k} b(X_{t_{k-1}}^\varepsilon, \sigma_0)\,\mathrm{d}W_t\right|^2 \Bigg\} 1_{C_{k,0}^{n,\varepsilon,\rho}},\end{aligned}$$

and by the same argument as in the proof of Lemma 3.4.15, we obtain

$$\begin{aligned}
&\frac{\sqrt{n}}{\varepsilon^2}\sum_{k=1}^n E\left[|g(X^\varepsilon_{t_{k-1}})|\left|\int_{t_{k-1}}^{t_k}\left\{a(X^\varepsilon_t,\mu_0)-a(X^\varepsilon_{t_{k-1}},\mu_0)\right\}\mathrm{d}t\right|^2 1_{C^{n,\varepsilon,\rho}_{k,0}}\,\middle|\,\mathcal{F}_{t_{k-1}}\right]\\
&\quad= O_p\left(\frac{1}{\varepsilon^2 n^{5/2}}+\frac{1}{n^{3/2}}\right),\\
&\frac{\sqrt{n}}{\varepsilon}\sum_{k=1}^n E\left[\left|g(X^\varepsilon_{t_{k-1}})\int_{t_{k-1}}^{t_k}\left\{a(X^\varepsilon_t,\mu_0)-a(X^\varepsilon_{t_{k-1}},\mu_0)\right\}\mathrm{d}t\int_{t_{k-1}}^{t_k} b(X^\varepsilon_{t-},\sigma_0)\,\mathrm{d}W_t\right| 1_{C^{n,\varepsilon,\rho}_{k,0}}\,\middle|\,\mathcal{F}_{t_{k-1}}\right]\\
&\quad= O_p\left(\frac{1}{\varepsilon n}+\frac{1}{\sqrt{n}}\right),\\
&\sqrt{n}\sum_{k=1}^n E\left[\left|g(X^\varepsilon_{t_{k-1}})\left|\int_{t_{k-1}}^{t_k} b(X^\varepsilon_{t-},\sigma_0)\,\mathrm{d}W_t\right|^2-\left|\int_{t_{k-1}}^{t_k} b(X^\varepsilon_{t_{k-1}},\sigma_0)\,\mathrm{d}W_t\right|^2\right| 1_{C^{n,\varepsilon,\rho}_{k,0}}\,\middle|\,\mathcal{F}_{t_{k-1}}\right]\\
&\quad= O_p\left(\frac{1}{\sqrt{n}}+\varepsilon\right)
\end{aligned}$$

as $n\to\infty$, $\varepsilon\to 0$, $\lambda_\varepsilon\to\infty$, $\varepsilon n\to\infty$ and $\varepsilon\lambda_\varepsilon\to 0$.

For $\ell=3$, let $r_{n,\varepsilon}$ be defined as either of the following:

(i) Under Assumption 3.2.4 (i), $r_{n,\varepsilon}=\frac{1}{\varepsilon n^{1-1/p}}+\frac{1}{n^{1/2-1/p}}$ with sufficiently large $p>1$.

(ii) Under Assumption 3.2.4 (ii), $r_{n,\varepsilon}=\frac{1}{\varepsilon n^{1-1/p-q\rho}}+\frac{1}{n^{1/2-1/p-q\rho}}$ with sufficiently large $p>1$.

Then, it follows from Lemmas 3.4.10, 3.4.12 and 3.7.3 that

$$\sum_{k=1}^n E\left[\left|\xi^i_{3,k}1_{J^{n,\varepsilon}_{k,1}}-\tilde{\xi}^i_{3,k}\right|\,\middle|\,\mathcal{F}_{t_{k-1}}\right]=O_p\left(\sqrt{\lambda_\varepsilon}r_{n,\varepsilon}\right)$$

as $n\to\infty$, $\varepsilon\to 0$, $\lambda_\varepsilon\to\infty$, $\varepsilon n\to\infty$, $\varepsilon\lambda_\varepsilon\to 0$, $\lambda_\varepsilon^2/n\to 0$ and $\lambda_\varepsilon\int_{|z|\le 4v_2/c_1n^\rho} f_{\alpha_0}(z)\,\mathrm{d}z\to 0$. □

Lemma 3.4.17. *Under Assumptions 3.2.1 to 3.2.3, 3.2.5, 3.2.6, 3.2.8 and 3.2.9,*

$$\sum_{k=1}^n E\left[\tilde{\xi}^i_{\ell k}\,\middle|\,\mathcal{F}_{t_{k-1}}\right]\xrightarrow{p}0\quad(\ell=1,2,3)$$

as $n\to\infty$, $\varepsilon\to 0$, $\lambda_\varepsilon\to\infty$, $\varepsilon\lambda_\varepsilon\to 0$ and $\lambda_\varepsilon/n\to 0$.

Proof. For $\ell=1$, let $i\in\{1,\dots,d_1\}$, and put $g(x)=\frac{\partial a}{\partial\mu_i}(x,\mu_0)/b(x,\sigma_0)$. Since

$$E\left[\int_{t_{k-1}}^{t_k}\mathrm{d}W_t\,\middle|\,\mathcal{F}_{t_{k-1}}\right]=0,\quad\text{and}\quad\int_{t_{k-1}}^{t_k}\mathrm{d}W_t\ \text{and}\ 1_{J^{n,\varepsilon}_{k,i}}\ (i=1,2)\ \text{are independent},$$

it holds from Lemmas 3.4.4 and 3.4.7 that for any $p \geq 1$

$$\begin{aligned}\left|\sum_{k=1}^n E\left[\tilde{\xi}^i_{1,k} \,\middle|\, \mathcal{F}_{t_{k-1}}\right]\right| &= \left|\sum_{k=1}^n g(X^\varepsilon_{t_{k-1}}) E\left[\int_{t_{k-1}}^{t_k} \mathrm{d}W_t\, 1_{C^{n,\varepsilon,\rho}_{k,0}} \,\middle|\, \mathcal{F}_{t_{k-1}}\right]\right| \\ &= \left|\sum_{k=1}^n g(X^\varepsilon_{t_{k-1}}) E\left[\int_{t_{k-1}}^{t_k} \mathrm{d}W_t\, 1_{D^{n,\varepsilon,\rho}_{k,0} \cup J^{n,\varepsilon}_{k,1} \cup J^{n,\varepsilon}_{k,2}} \,\middle|\, \mathcal{F}_{t_{k-1}}\right]\right| \\ &= O_p\left(\frac{1}{n^{p(1-\rho)-1/2}} + \frac{\varepsilon^p}{n^{p(1/2-\rho)-1/2}}\right) + O_p\left(\frac{\lambda_\varepsilon}{n}\right) + O_p\left(\frac{\lambda_\varepsilon^2}{n^2}\right)\end{aligned}$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$ and $\varepsilon\lambda_\varepsilon \to 0$.

For $\ell = 2$, let $i \in \{1, \ldots, d_2\}$, and put $g(x) = -\frac{1}{b}\frac{\partial b}{\partial \sigma_i}(x, \sigma_0)$. Since

$$E\left[\left|\int_{t_{k-1}}^{t_k} \mathrm{d}W_t\right|^2 \,\middle|\, \mathcal{F}_{t_{k-1}}\right] = \frac{1}{n}, \quad \text{and } \left|\int_{t_{k-1}}^{t_k} \mathrm{d}W_t\right|^2 \text{ and } 1_{J^{n,\varepsilon}_{k,i}} \ (i = 1, 2) \text{ are independent,}$$

it follows from Lemmas 3.4.4 and 3.4.7 that for any $p \geq 1$

$$\begin{aligned}\left|\sum_{k=1}^n E\left[\tilde{\xi}^i_{2,k} \,\middle|\, \mathcal{F}_{t_{k-1}}\right]\right| &= \left|\sqrt{n}\sum_{k=1}^n g(X^\varepsilon_{t_{k-1}}) E\left[\left\{\left|\int_{t_{k-1}}^{t_k} \mathrm{d}W_t\right|^2 - \frac{1}{n}\right\} 1_{C^{n,\varepsilon,\rho}_{k,0}} \,\middle|\, \mathcal{F}_{t_{k-1}}\right]\right| \\ &= \left|\sqrt{n}\sum_{k=1}^n g(X^\varepsilon_{t_{k-1}}) E\left[\left\{\left|\int_{t_{k-1}}^{t_k} \mathrm{d}W_t\right|^2 - \frac{1}{n}\right\} 1_{D^{n,\varepsilon,\rho}_{k,0}} \,\middle|\, \mathcal{F}_{t_{k-1}}\right]\right| \\ &= O_p\left(\frac{1}{n^{p(1-\rho)-1/2}} + \frac{\varepsilon^p}{n^{p(1/2-\rho)-1/2}}\right)\end{aligned}$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$ and $\varepsilon\lambda_\varepsilon \to 0$.

For $\ell = 3$, we may assume $\sup_t |X^\varepsilon_t - x_t| < \delta$ for some enough small $\delta > 0$. From Assumption 3.2.9, we obtain

$$\begin{aligned}\sum_{k=1}^n E\left[\tilde{\xi}^i_{3,k} \,\middle|\, \mathcal{F}_{t_{k-1}}\right] &= \frac{\sqrt{\lambda_\varepsilon}}{n}\sum_{k=1}^n \int \frac{\partial \psi}{\partial \alpha_i}\Big(X^\varepsilon_{t_{k-1}}, c(X^\varepsilon_{t_{k-1}}, \alpha_0)z, \alpha_0\Big) f_{\alpha_0}(z)\,\mathrm{d}z \\ &= \frac{\sqrt{\lambda_\varepsilon}}{n}\sum_{k=1}^n \frac{\partial}{\partial \alpha_i}\left(\int \psi\Big(X^\varepsilon_{t_{k-1}}, c(X^\varepsilon_{t_{k-1}}, \alpha_0)z, \alpha\Big) f_{\alpha_0}(z)\,\mathrm{d}z\right)_{\alpha=\alpha_0} = 0.\end{aligned}$$

The last equality holds from the fact that

$$\alpha \mapsto \int \psi\Big(X^\varepsilon_{t_{k-1}}, c(X^\varepsilon_{t_{k-1}}, \alpha_0)z, \alpha\Big) f_{\alpha_0}(z)\,\mathrm{d}z$$

behaves like the Kullback Leibler divergence from $p_{\alpha,x}$ to $p_{\alpha_0,x}$ at $x = X^\varepsilon_{t_{k-1}}$, where $p_{\alpha,x}(y) = f_\alpha(y/c(x,\alpha))/c(x,\alpha)$. □

Lemma 3.4.18. *Under Assumptions 3.2.1 to 3.2.6, 3.2.8, 3.2.9 and 3.2.11,*

$$\sum_{k=1}^{n} E\left[\tilde{\xi}_{\ell k}^{i_1}\tilde{\xi}_{\ell k}^{i_2}\,\middle|\,\mathcal{F}_{t_{k-1}}\right] \xrightarrow{p} I_\ell^{i_1 i_2} \quad (\ell = 1,2,3,\ i_1,i_2 = 1,\dots,d_\ell),$$

$$\sum_{k=1}^{n} E\left[\tilde{\xi}_{\ell_1 k}^{i_1}\tilde{\xi}_{\ell_2 k}^{i_2}\,\middle|\,\mathcal{F}_{t_{k-1}}\right] \xrightarrow{p} 0 \quad (\ell_1,\ell_2 = 1,2,3,\ \ell_1 \neq \ell_2,\ i_j = 1,\dots,d_{\ell_j},\ j = 1,2)$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$ and $\varepsilon\lambda_\varepsilon \to 0$, where $I_1,\dots,I_3$ are the matrices defined as (3.4)*.*

Proof. For $\ell = 1$, $i,j \in \{1,\dots,d_1\}$, put $g(x) = \frac{\partial a}{\partial \mu_i}\frac{\partial a}{\partial \mu_j}(x,\mu_0)/b(x,\sigma_0)^2$. Since from Lemmas 3.4.4 and 3.4.7 for any $p > 1$ we have

$$\sum_{k=1}^{n} g(X_{t_{k-1}}^\varepsilon) E\left[\left|\int_{t_{k-1}}^{t_k} \mathrm{d}W_t\right|^2 1_{D_{k,0}^{n,\varepsilon,\rho}\cup J_{k,1}^{n,\varepsilon}\cup J_{k,2}^{n,\varepsilon}}\,\middle|\,\mathcal{F}_{t_{k-1}}\right] = O_p\left(\frac{1}{n^{p(1-\rho)}} + \frac{\varepsilon^p}{n^{p(1/2-\rho)}} + \frac{\lambda_\varepsilon}{n}\right)$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$ and $\varepsilon\lambda_\varepsilon \to 0$, we obtain

$$\begin{aligned}\sum_{k=1}^{n} E\left[\tilde{\xi}_{1,k}^{i}\tilde{\xi}_{1,k}^{j}\,\middle|\,\mathcal{F}_{t_{k-1}}\right] &= \sum_{k=1}^{n} g(X_{t_{k-1}}^\varepsilon) E\left[\left|\int_{t_{k-1}}^{t_k} \mathrm{d}W_t\right|^2 1_{C_{k,0}^{n,\varepsilon,\rho}}\,\middle|\,\mathcal{F}_{t_{k-1}}\right]\\ &= \frac{1}{n}\sum_{k=1}^{n} g(X_{t_{k-1}}^\varepsilon) + O_p\left(\frac{1}{n^{p(1-\rho)}} + \frac{\varepsilon^p}{n^{p(1/2-\rho)}} + \frac{\lambda_\varepsilon}{n}\right) \xrightarrow{p} \int_0^1 g(x_t)\,\mathrm{d}t\end{aligned}$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$, $\varepsilon\lambda_\varepsilon \to 0$, and $\lambda_\varepsilon/n \to 0$.

For $\ell = 2$, $i,j \in \{1,\dots,d_2\}$, put $g(x) = \frac{1}{b^2}\frac{\partial b}{\partial \sigma_i}\frac{\partial b}{\partial \sigma_j}(x,\sigma_0)$. Since similarly to the proof of Lemma 3.4.17, it follows from Lemmas 3.4.4 and 3.4.7 that for any $p > 1$

$$\begin{aligned}&n\sum_{k=1}^{n} g(X_{t_{k-1}}^\varepsilon) E\left[\left|\left|\int_{t_{k-1}}^{t_k} \mathrm{d}W_t\right|^2 - \frac{1}{n}\right|^2 1_{D_{k,0}^{n,\varepsilon,\rho}\cup J_{k,1}^{n,\varepsilon}\cup J_{k,2}^{n,\varepsilon}}\,\middle|\,\mathcal{F}_{t_{k-1}}\right]\\ &= O_p\left(\frac{1}{n^{p(1-\rho)}} + \frac{\varepsilon^p}{n^{p(1/2-\rho)}}\right) + O_p\left(\frac{\lambda_\varepsilon}{n}\right) + O_p\left(\frac{\lambda_\varepsilon^2}{n^2}\right)\end{aligned}$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$ and $\varepsilon\lambda_\varepsilon \to 0$, we obtain from Lemma 3.4.4 that

$$\begin{aligned}\sum_{k=1}^{n} E\left[\tilde{\xi}_{2,k}^{i}\tilde{\xi}_{2,k}^{j}\,\middle|\,\mathcal{F}_{t_{k-1}}\right] &= n\sum_{k=1}^{n} g(X_{t_{k-1}}^\varepsilon) E\left[\left|\left|\int_{t_{k-1}}^{t_k} \mathrm{d}W_t\right|^2 - \frac{1}{n}\right|^2 1_{C_{k,0}^{n,\varepsilon,\rho}}\,\middle|\,\mathcal{F}_{t_{k-1}}\right]\\ &= n\sum_{k=1}^{n} g(X_{t_{k-1}}^\varepsilon) E\left[\left|\left|\int_{t_{k-1}}^{t_k} \mathrm{d}W_t\right|^2 - \frac{1}{n}\right|^2\,\middle|\,\mathcal{F}_{t_{k-1}}\right] + O_p\left(\frac{1}{n^{p(1-\rho)}} + \frac{\varepsilon^p}{n^{p(1/2-\rho)}} + \frac{\lambda_\varepsilon}{n}\right)\\ &\xrightarrow{p} 2\int_0^1 g(x_t)\,\mathrm{d}t\end{aligned}$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$, $\varepsilon\lambda_\varepsilon \to 0$ and $\lambda_\varepsilon/n \to 0$.

For $\ell = 3$, $i, j \in \{1, \ldots, d_3\}$, put $g(x, y) = \frac{\partial\psi}{\partial\alpha_i}\frac{\partial\psi}{\partial\alpha_j}(x, y, \alpha_0)$. Then, it follows from Lemma 3.4.4 and Assumption 3.2.11 that

$$\begin{aligned}\sum_{k=1}^n E\left[\tilde{\xi}^i_{3,k}\tilde{\xi}^j_{3,k} \,\middle|\, \mathcal{F}_{t_{k-1}}\right] &= \frac{1}{\lambda_\varepsilon}\sum_{k=1}^n E\left[g\Big(X^\varepsilon_{t_{k-1}}, c(X^\varepsilon_{t_{k-1}}, \alpha_0)V_{N^{\lambda_\varepsilon}_{\tau_k}}\Big)1_{J^{n,\varepsilon}_{k,1}} \,\middle|\, \mathcal{F}_{t_{k-1}}\right] \\ &= \frac{1}{n}\sum_{k=1}^n \int g\Big(X^\varepsilon_{t_{k-1}}, c(X^\varepsilon_{t_{k-1}}, \alpha_0)z\Big) f_{\alpha_0}(z)\,\mathrm{d}z \\ &\xrightarrow{p} \int_0^1 \int g(x_t, c(x_t, \alpha_0)z) f_{\alpha_0}(z)\,\mathrm{d}z\,\mathrm{d}t\end{aligned}$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$ and $\varepsilon\lambda_\varepsilon \to 0$. The second equality holds from the fact that $V_{N^{\lambda_\varepsilon}_{\tau_k}}$ and $1_{J^{n,\varepsilon}_{k,1}}$ are independent.

For $\ell_j = j$, $i_j = 1, \ldots, d_j$ $(j = 1, 2)$, put $g(x) = -\frac{\partial a}{\partial\mu_{i_1}}(x, \mu_0)\frac{1}{b^2}\frac{\partial b}{\partial\sigma_{i_2}}(x, \sigma_0)$. Since

$$E\left[\left(\int_{t_{k-1}}^{t_k}\mathrm{d}W_t\right)^i \,\middle|\, \mathcal{F}_{t_{k-1}}\right] = 0, \text{ and } \left(\int_{t_{k-1}}^{t_k}\mathrm{d}W_t\right)^i \text{ and } 1_{J^{n,\varepsilon}_{k,j}} \text{ are independent } (i = 1, 3,\ j = 1, 2),$$

it follows from Lemmas 3.4.4 and 3.4.7 that for any $p \geq 1$

$$\begin{aligned}\sum_{k=1}^n E\left[\tilde{\xi}^{i_1}_{1,k}\tilde{\xi}^{i_2}_{2,k} \,\middle|\, \mathcal{F}_{t_{k-1}}\right] &= \sqrt{n}\sum_{k=1}^n g(X^\varepsilon_{t_{k-1}})E\left[\left\{-\left|\int_{t_{k-1}}^{t_k}\mathrm{d}W_t\right|^2 + \frac{1}{n}\right\}\int_{t_{k-1}}^{t_k}\mathrm{d}W_t\,1_{C^{n,\varepsilon,\rho}_{k,0}} \,\middle|\, \mathcal{F}_{t_{k-1}}\right] \\ &= \sqrt{n}\sum_{k=1}^n g(X^\varepsilon_{t_{k-1}})E\left[\left\{-\left|\int_{t_{k-1}}^{t_k}\mathrm{d}W_t\right|^2 + \frac{1}{n}\right\}\int_{t_{k-1}}^{t_k}\mathrm{d}W_t\,1_{D^{n,\varepsilon,\rho}_{k,0}} \,\middle|\, \mathcal{F}_{t_{k-1}}\right] \\ &= O_p\left(\frac{1}{n^{p(1-\rho)}} + \frac{\varepsilon^p}{n^{p(1/2-\rho)}}\right)\end{aligned}$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$ and $\varepsilon\lambda_\varepsilon \to 0$. □

Lemma 3.4.19. *Under Assumptions 3.2.1 to 3.2.3, 3.2.5, 3.2.6, 3.2.8 and 3.2.9,*

$$\sum_{k=1}^n \left|E\left[\tilde{\xi}^i_{\ell k} \,\middle|\, \mathcal{F}_{t_{k-1}}\right]\right|^2 \xrightarrow{p} 0 \quad (\ell = 1, 2, 3)$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$, $\varepsilon\lambda_\varepsilon \to 0$ *and* $\lambda_\varepsilon/n \to 0$.

Proof. This follows from the same argument as in the proof of Lemma 3.4.17. □

Lemma 3.4.20. *Under Assumptions 3.2.1 to 3.2.3, 3.2.5, 3.2.6, 3.2.8 and 3.2.11,*

$$\sum_{k=1}^n E\left[\left|\tilde{\xi}^i_{\ell k}\right|^4 \,\middle|\, \mathcal{F}_{t_{k-1}}\right] \xrightarrow{p} 0 \quad (\ell = 1, 2, 3)$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$ *and* $\varepsilon\lambda_\varepsilon \to 0$.

Proof. For $\ell = 1$, let $i \in \{1, \dots, d_1\}$, and put $g(x) = |\frac{\partial a}{\partial \mu_i}(x, \mu_0)/b(x, \sigma_0)|^4$. Then, it holds from Lemma 3.4.4 that

$$\sum_{k=1}^n E\left[\left|\tilde{\xi}^i_{1,k}\right|^4 \middle| \mathcal{F}_{t_{k-1}}\right] \leq \sum_{k=1}^n g\Big(X^\varepsilon_{t_{k-1}}\Big) E\left[\left|\int_{t_{k-1}}^{t_k} \mathrm{d}W_t\right|^4 \middle| \mathcal{F}_{t_{k-1}}\right] \xrightarrow{p} 0$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$ and $\varepsilon\lambda_\varepsilon \to 0$.

For $\ell = 2$, let $i \in \{1, \dots, d_2\}$, and put $g(x) = \left|\frac{1}{b}\frac{\partial b}{\partial \sigma_i}(x, \sigma_0)\right|^4$. Then, it follows from Lemma 3.4.4 that

$$\sum_{k=1}^n E\left[\left|\tilde{\xi}^i_{2,k}\right|^4 \middle| \mathcal{F}_{t_{k-1}}\right] \leq n^2 \sum_{k=1}^n g(X^\varepsilon_{t_{k-1}}) E\left[\left|\left|\int_{t_{k-1}}^{t_k} \mathrm{d}W_t\right|^2 - \frac{1}{n}\right|^4 \middle| \mathcal{F}_{t_{k-1}}\right] \xrightarrow{p} 0$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$ and $\varepsilon\lambda_\varepsilon \to 0$.

For $\ell = 3$, $i \in \{1, \dots, d_3\}$, put $g(x, y) = \left|\frac{\partial \psi}{\partial \alpha_i}(x, y, \alpha_0)\right|^4$. Then, similarly to the proof of Lemma 3.4.18, it follows from Lemma 3.4.4 and Assumption 3.2.11 that

$$\sum_{k=1}^n E\left[\left|\tilde{\xi}^i_{3,k}\right|^4 \middle| \mathcal{F}_{t_{k-1}}\right] = \frac{1}{\lambda_\varepsilon^2} \sum_{k=1}^n E\left[g\Big(X^\varepsilon_{t_{k-1}}, c(X^\varepsilon_{t_{k-1}}, \alpha_0) V_{N^{\lambda_\varepsilon}_{\tau_k}}\Big) 1_{J^{n,\varepsilon}_{k,1}} \middle| \mathcal{F}_{t_{k-1}}\right] \xrightarrow{p} 0$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$ and $\varepsilon\lambda_\varepsilon \to 0$. □

Proof of Theorem 3.3.2. From Theorem A.3 in Shimizu [28] and Lemmas 3.4.16 to 3.4.20,

$$\Lambda_{n,\varepsilon} := \sum_{k=1}^n \left(\xi^1_{1,k}, \dots, \xi^{d_1}_{1,k}, \xi^1_{2,k}, \dots, \xi^{d_2}_{2,k}, \xi^1_{3,k}, \dots, \xi^{d_3}_{3,k}\right)^T \xrightarrow{d} \mathcal{N}\left(0, I_{\theta_0}\right)$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$, $\varepsilon\lambda_\varepsilon \to 0$, $\lambda_\varepsilon^2/n \to 0$ and $\lambda_\varepsilon \int_{|z| \leq 4v_2/c_1 n^\rho} f_{\alpha_0}(z)\,\mathrm{d}z \to 0$ with $\lim(\varepsilon^2 n)^{-1} < \infty$. Also, it follows from Lemmas 3.4.11 to 3.4.15 under Assumption 3.2.12 that

$$C_{\varepsilon,n}(\theta) := \begin{pmatrix} \varepsilon^2 n\left(\frac{\partial^2 \Psi_{n,\varepsilon}}{\partial \mu_i \partial \mu_j}(\theta)\right)_{i,j} & \varepsilon^2 n\left(\frac{\partial^2 \Psi_{n,\varepsilon}}{\partial \mu_i \partial \sigma_j}(\theta)\right)_{i,j} & 0 \\ \left(\frac{\partial^2 \Psi_{n,\varepsilon}}{\partial \sigma_i \partial \mu_j}(\theta)\right)_{i,j} & \left(\frac{\partial^2 \Psi_{n,\varepsilon}}{\partial \sigma_i \partial \sigma_j}(\theta)\right)_{i,j} & 0 \\ 0 & 0 & \left(\frac{\partial^2 \Psi_{n,\varepsilon}}{\partial \alpha_i \partial \alpha_j}(\theta)\right)_{i,j} \end{pmatrix} \xrightarrow{p} -I_{\theta_0} \tag{3.20}$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$, $\varepsilon\lambda_\varepsilon \to 0$, $\lambda_\varepsilon^2/n \to 0$ and $\lambda_\varepsilon \int_{|z| \leq 4v_2/c_1 n^\rho} f_{\alpha_0}(z)\,\mathrm{d}z \to 0$ with

$\lim(\varepsilon^2 n)^{-1} < \infty$, uniformly in $\theta \in \Theta$. Indeed,

$$\varepsilon^2 n \frac{\partial^2 \Psi_{n,\varepsilon}}{\partial\mu_i\partial\mu_j}(\theta) = \sum_{k=1}^n \left\{\Delta_k^n X^\varepsilon - \frac{1}{n}a(X_{t_{k-1}}^\varepsilon,\mu)\right\} \frac{\frac{\partial^2 a}{\partial\mu_i\partial\mu_j}\left(X_{t_{k-1}}^\varepsilon,\mu\right)}{\left|b(X_{t_{k-1}}^\varepsilon,\sigma)\right|^2} 1_{C_k^{n,\varepsilon,\rho}}$$

$$- \frac{1}{n}\sum_{k=1}^n \frac{\frac{\partial a}{\partial\mu_i}\frac{\partial a}{\partial\mu_j}\left(X_{t_{k-1}}^\varepsilon,\mu\right)}{\left|b(X_{t_{k-1}}^\varepsilon,\sigma)\right|^2} 1_{C_k^{n,\varepsilon,\rho}} \xrightarrow{p} -I_1^{ij},$$

$$\varepsilon^2 n \frac{\partial^2 \Psi_{n,\varepsilon}}{\partial\mu_i\partial\sigma_j}(\theta) = -2\sum_{k=1}^n \left\{\Delta_k^n X^\varepsilon - \frac{1}{n}a(X_{t_{k-1}}^\varepsilon,\mu)\right\} \frac{\frac{\partial a}{\partial\mu_i}\left(X_{t_{k-1}}^\varepsilon,\mu\right)\frac{\partial b}{\partial\sigma_j}\left(X_{t_{k-1}}^\varepsilon,\sigma\right)}{\left|b(X_{t_{k-1}}^\varepsilon,\sigma)\right|^3} 1_{C_k^{n,\varepsilon,\rho}} \xrightarrow{p} 0,$$

$$\frac{\partial^2 \Psi_{n,\varepsilon}}{\partial\sigma_i\partial\sigma_j}(\theta) = -\frac{1}{n}\sum_{k=1}^n \left\{ -\frac{\left|\Delta_k^n X^\varepsilon - \frac{1}{n}a(X_{t_{k-1}}^\varepsilon,\mu)\right|^2}{\frac{1}{n}\left|\varepsilon b(X_{t_{k-1}}^\varepsilon,\sigma)\right|^2} + 1\right\} \frac{\partial\left(\frac{1}{b}\frac{\partial b}{\partial\sigma_i}\right)}{\partial\sigma_j}\left(X_{t_{k-1}}^\varepsilon,\sigma\right) 1_{C_k^{n,\varepsilon,\rho}}$$

$$- \frac{2}{\varepsilon^2}\sum_{k=1}^n \left|\Delta_k^n X^\varepsilon - \frac{1}{n}a(X_{t_{k-1}}^\varepsilon,\mu)\right|^2 \frac{\frac{\partial b}{\partial\sigma_i}\frac{\partial b}{\partial\sigma_j}\left(X_{t_{k-1}}^\varepsilon,\sigma\right)}{|b(X_{t_{k-1}}^\varepsilon,\sigma)|^4} 1_{C_k^{n,\varepsilon,\rho}} \xrightarrow{p} -I_2^{i_1 i_2},$$

$$\frac{\partial^2 \Psi_{n,\varepsilon}}{\partial\alpha_i\partial\alpha_j}(\theta) = \frac{1}{\lambda_\varepsilon}\sum_{k=1}^n \frac{1}{\varphi}\frac{\partial^2\varphi}{\partial\alpha_i\partial\alpha_j}\left(X_{t_{k-1}}^\varepsilon, \frac{\Delta_k^n X^\varepsilon}{\varepsilon},\alpha\right) 1_{D_k^{n,\varepsilon,\rho}}$$

$$- \frac{1}{\lambda_\varepsilon}\sum_{k=1}^n \frac{1}{|\varphi|^2}\frac{\partial\varphi}{\partial\alpha_i}\frac{\partial\varphi}{\partial\alpha_j}\left(X_{t_{k-1}}^\varepsilon, \frac{\Delta_k^n X^\varepsilon}{\varepsilon},\alpha\right) 1_{D_k^{n,\varepsilon,\rho}} \xrightarrow{p} -I_3^{i_1 i_2},$$

where $\varphi(x,y,\alpha) := \exp\psi(x,y,\alpha)$. Since

$$D_{n,\varepsilon}\begin{pmatrix} \varepsilon^{-1}(\hat\mu_{n,\varepsilon}-\mu_0) \\ \sqrt{n}(\hat\sigma_{n,\varepsilon}-\sigma_0) \\ \sqrt{\lambda_\varepsilon}(\hat\alpha_{n,\varepsilon}-\alpha_0) \end{pmatrix} = \Lambda_{n,\varepsilon},$$

where

$$D_{n,\varepsilon} := \int_0^1 C_{n,\varepsilon}(\theta_0 + u(\hat\theta_{n,\varepsilon}-\theta_0)),$$

the conclusion follows by the same argument in the proof of Theorem 1 in Sørensen and Uchida [33]. □

3.5 Examples

This section is devoted to give some examples of densities which satisfy Assumptions 3.2.9 to 3.2.12. For simplicity, suppose that $c(x,\alpha)$ is an enough smooth postive function on $I_{x_0}^\delta \times \Theta_3$, and derivatives of c are uniformly continuous. Let D_+ is the interior of the common support of $\{f_\alpha\}_{\alpha\in\Theta_3}$, *i.e.*,

$$f_\alpha(z)\begin{cases} > 0 & \text{for } z \in D_+, \\ = 0 & \text{otherwise.}\end{cases}$$

Note that $y \in D_+(=\mathbb{R}$ or $\mathbb{R}_+)$ if and only if $y/c(x,\alpha) \in D_+$ for $(x,\alpha) \in I_{x_0}^\delta \times \Theta_3$ owing to Assumption 3.2.4. If $(x,y,\alpha) \in I_{x_0}^\delta \times D_+ \times \Theta_3$,

$$\frac{\partial\psi}{\partial y}(x,y,\alpha) = \frac{1}{c(x,\alpha)} \frac{f_\alpha'\left(\frac{y}{c(x,\alpha)}\right)}{f_\alpha\left(\frac{y}{c(x,\alpha)}\right)},$$

$$\frac{\partial\psi}{\partial\alpha_j}(x,y,\alpha) = -\frac{\partial(\log c)}{\partial\alpha_j}(x,\alpha)\left(1 + y\frac{\partial\psi}{\partial y}(x,y,\alpha)\right) + \frac{\frac{\partial f_\alpha}{\partial\alpha_j}\left(\frac{y}{c(x,\alpha)}\right)}{c(x,\alpha) f_\alpha\left(\frac{y}{c(x,\alpha)}\right)}$$

for $(x,\alpha) \in I_{x_0}^\delta \times \Theta_3$. The values of these functions may be undefined if $(x,y,\alpha) \in I_{x_0}^\delta \times \partial D_+ \times \Theta_3$. Otherwise their values are equal to zero.

First, we show an example such that the class of jump size densities satisfies Assumption 3.2.4 (i).

Example 3.5.1 (Normal distribution)**.** *Let Θ_3 be a smooth open convex set which is compactly contained in $\mathbb{R} \times \mathbb{R}_+ \times \mathbb{R}^{d_3-2}$, and let f_α be of the form*

$$f_\alpha(z) = \frac{1}{\sqrt{2\pi\alpha_2^2}} \exp\left(-\frac{|z-\alpha_1|^2}{2\alpha_2^2}\right) \quad \text{for } \alpha = (\alpha_1, \alpha_2) \in \Theta_3.$$

Then,

$$\psi(x,y,\alpha) = -\log c(x,\alpha) - \frac{1}{2}\log\left(2\pi\alpha_2^2\right) - \frac{|\frac{y}{c(x,\alpha)} - \alpha_1|^2}{2\alpha_2^2} \quad \text{on } I_{x_0}^\delta \times \mathbb{R} \times \Theta_3.$$

Since

$$f_\alpha'(z) = -\frac{z-\alpha_1}{\alpha_2^2} f_\alpha(z), \quad \frac{\partial f_\alpha}{\partial\alpha_1}(z) = \frac{z-\alpha_1}{\alpha_2} f_\alpha(z), \quad \frac{\partial f_\alpha}{\partial\alpha_2}(z) = \left\{-\frac{1}{\alpha_2} + \frac{(z-\alpha_1)^2}{\alpha_2^3}\right\} f_\alpha(z),$$

we have

$$\begin{aligned}
\frac{\partial\psi}{\partial y}(x,y,\alpha) &= -\frac{1}{c(x,\alpha)}\frac{1}{\alpha_2^2}\left(\frac{y}{c(x,\alpha)} - \alpha_1\right), \\
\frac{\partial\psi}{\partial\alpha_1}(x,y,\alpha) &= -\frac{\partial(\log c)}{\partial\alpha_1}(x,\alpha)\left(1 + y\frac{\partial\psi}{\partial y}(x,y,\alpha)\right) - \frac{\frac{y}{c(x,\alpha)} - \alpha_1}{\alpha_2 c(x,\alpha)}, \\
\frac{\partial\psi}{\partial\alpha_2}(x,y,\alpha) &= -\frac{\partial(\log c)}{\partial\alpha_2}(x,\alpha)\left(1 + y\frac{\partial\psi}{\partial y}(x,y,\alpha)\right) + \frac{1}{c(x,\alpha)}\left\{-\frac{1}{\alpha_2} + \frac{|\frac{y}{c(x,\alpha)} - \alpha_1|^2}{\alpha_2^3}\right\}, \\
\frac{\partial\psi}{\partial\alpha_2}(x,y,\alpha) &= -\frac{\partial(\log c)}{\partial\alpha_j}(x,\alpha)\left(1 + y\frac{\partial\psi}{\partial y}(x,y,\alpha)\right)
\end{aligned}$$

for $(x,y,\alpha) \in I_{x_0}^\delta \times \mathbb{R}_+ \times \Theta_3$ and $j = 3, \ldots, d_3$. Furthermore, the derivatives of c and $\log c$ with respect to α are bounded on $I_{x_0}^\delta \times \Theta_3$, and so

$$\left|\frac{\partial^2\psi}{\partial\alpha_j\partial y}(x,y,\alpha)\right| \le C(1+|y|), \quad \left|\frac{\partial^2\psi}{\partial\alpha_i\partial\alpha_j}(x,y,\alpha)\right| \le C(1+|y|^2) \quad \text{for } (x,y,\alpha) \in I_{x_0}^\delta \times \mathbb{R} \times \Theta_3,$$

where C is a constant not depending on (x,y,α). Thus, Assumptions 3.2.9 to 3.2.12 are satisfied.

Next, we show examples such that the class of jump size densities satisfies Assumption 3.2.4 (ii).

Example 3.5.2 (Gamma distribution)**.** *Let Θ_3 be an open interval compactly contained in $\mathbb{R}_+ \times (1,\infty) \times \mathbb{R}^{d_3-2}$, and let f_α be of the form*

$$f_\alpha(z) = \begin{cases} \dfrac{1}{\Gamma(\alpha_2)\alpha_1^{\alpha_2}} z^{\alpha_2-1}e^{-z/\alpha_1} & (z>0), \\ 0 & (z \le 0) \end{cases}$$

for $\alpha \in \Theta_3$. Then,

$$\psi(x,y,\alpha) = -\log c(x,\alpha) - \log\Gamma(\alpha_2) - \alpha_2\log\alpha_1 + (\alpha_2-1)\log z - \frac{z}{\alpha_1} \quad \text{on } I_{x_0}^\delta \times \mathbb{R}_+ \times \Theta_3.$$

Since

$$f'_\alpha(z) = \left(\frac{\alpha_2-1}{z} - \frac{1}{\alpha_1}\right) f_\alpha(z),$$
$$\frac{\partial f_\alpha}{\partial\alpha_1}(z) = \left(-\frac{\alpha_2}{\alpha_1} + \frac{z}{\alpha_1^2}\right) f_\alpha(z), \quad \frac{\partial f_\alpha}{\partial\alpha_2}(z) = \left\{-\frac{\Gamma'(\alpha_2)}{\Gamma(\alpha_2)} - \log\alpha_1 + \log z\right\} f_\alpha(z),$$

for $z>0$ and $\alpha \in \Theta_3$, we have

$$\begin{aligned}
\frac{\partial\psi}{\partial y}(x,y,\alpha) &= \frac{\alpha_2-1}{y} - \frac{1}{\alpha_1 c(x,\alpha)}, \\
\frac{\partial\psi}{\partial\alpha_1}(x,y,\alpha) &= -\frac{\partial(\log c)}{\partial\alpha_1}(x,\alpha)\left(1 + y\frac{\partial\psi}{\partial y}(x,y,\alpha)\right) + \frac{1}{c(x,\alpha)}\left\{-\frac{\alpha_2}{\alpha_1} + \frac{y}{\alpha_1^2 c(x,\alpha)}\right\}, \\
\frac{\partial\psi}{\partial\alpha_2}(x,y,\alpha) &= -\frac{\partial(\log c)}{\partial\alpha_2}(x,\alpha)\left(1 + y\frac{\partial\psi}{\partial y}(x,y,\alpha)\right) + \frac{1}{c(x,\alpha)}\left\{-\frac{\Gamma'(\alpha_2)}{\Gamma(\alpha_2)} - \log\alpha_1 + \log\frac{y}{c(x,\alpha)}\right\}, \\
\frac{\partial\psi}{\partial\alpha_j}(x,y,\alpha) &= -\frac{\partial(\log c)}{\partial\alpha_j}(x,\alpha)\left(1 + y\frac{\partial\psi}{\partial y}(x,y,\alpha)\right)
\end{aligned}$$

for $(x,y,\alpha) \in I_{x_0}^\delta \times \mathbb{R}_+ \times \Theta_3$ and $j = 3,\dots,d_3$. Furthermore, the derivatives of c and $\log c$ with respect to α are bounded on $I_{x_0}^\delta \times \Theta_3$, and so

$$\left|\frac{\partial^2\psi}{\partial\alpha_j\partial y}(x,y,\alpha)\right| \le C, \quad \left|\frac{\partial^2\psi}{\partial\alpha_i\partial\alpha_j}(x,y,\alpha)\right| \le C(1+|y|) \quad \text{for } (x,y,\alpha) \in I_{x_0}^\delta \times \mathbb{R} \times \Theta_3,$$

where C is a constant not depending on (x,y,α). Thus, Assumptions 3.2.9 to 3.2.12 are satisfied, and ρ in Theorems 3.3.1 and 3.3.2 can be taken as $\rho \in (0,1/4)$. Here, we remark that

$$\int \frac{1}{z} f_\alpha(z)\,\mathrm{d}z < \infty \quad \text{if and only if} \quad \alpha_2 > 1.$$

Example 3.5.3 (Inverse Gaussian distribution)**.** *Let Θ_3 be smooth, open, convex and compactly contained in $\mathbb{R}_+^2 \times \mathbb{R}^{d_3-2}$, and let f_α be of the form*

$$f_\alpha(z) = \begin{cases} \sqrt{\dfrac{\alpha_2}{2\pi z^3}} e^{-\alpha_2(z-\alpha_1)^2/2\alpha_1^2 z} & (z>0), \\ 0 & (z \le 0) \end{cases}$$

for $\alpha \in \Theta_3$. *Then,*

$$\psi(x, y, \alpha) = \frac{1}{2c(x, \alpha)} \left\{ \log \frac{\alpha_2}{2\pi} - 3 \log \frac{y}{c(x, \alpha)} \right\} - \frac{\alpha_2 \left| \frac{y}{c(x,\alpha)} - \alpha_1 \right|^2}{2\alpha_1^2 y} \quad \text{on } I_{x_0}^\delta \times \mathbb{R}_+ \times \Theta_3.$$

Since

$$f'_\alpha(z) = \left\{ -\frac{3}{2z} - \frac{\alpha_2(z - \alpha_1)}{\alpha_1^2 z} - \frac{\alpha_2(z - \alpha_1)^2}{2\alpha_1 z^2} \right\} f_\alpha(z),$$

$$\frac{\partial f_\alpha}{\partial \alpha_1}(z) = \frac{\alpha_2(z - \alpha_1)}{\alpha_1^2} f_\alpha(z), \quad \frac{\partial f_\alpha}{\partial \alpha_2}(z) = \left\{ \frac{1}{2\alpha_2} - \frac{|z - \alpha_1|^2}{2\alpha_1^2 z} \right\} f_\alpha(z)$$

for $z > 0$ *and* $\alpha \in \Theta_3$, *we have*

$$\begin{aligned}
\frac{\partial \psi}{\partial y}(x, y, \alpha) &= -\frac{3}{2y} - \frac{\alpha_2 \left\{ \frac{y}{c(x,\alpha)} - \alpha_1 \right\}}{\alpha_1^2 y} - \frac{\alpha_2 \left| \frac{y}{c(x,\alpha)} - \alpha_1 \right|^2}{2\alpha_1 \frac{y^2}{c(x,\alpha)}}, \\
\frac{\partial \psi}{\partial \alpha_1}(x, y, \alpha) &= -\frac{\partial (\log c)}{\partial \alpha_1}(x, \alpha) \left(1 + y \frac{\partial \psi}{\partial y}(x, y, \alpha) \right) + \frac{\alpha_2 (\frac{y}{c(x,\alpha)} - \alpha_1)}{\alpha_1^2 c(x, \alpha)}, \\
\frac{\partial \psi}{\partial \alpha_2}(x, y, \alpha) &= -\frac{\partial (\log c)}{\partial \alpha_2}(x, \alpha) \left(1 + y \frac{\partial \psi}{\partial y}(x, y, \alpha) \right) + \frac{1}{2\alpha_2 c(x, \alpha)} - \frac{|\frac{y}{c(x,\alpha)} - \alpha_1|^2}{2\alpha_1^2 y}, \\
\frac{\partial \psi}{\partial \alpha_j}(x, y, \alpha) &= -\frac{\partial (\log c)}{\partial \alpha_j}(x, \alpha) \left(1 + y \frac{\partial \psi}{\partial y}(x, y, \alpha) \right)
\end{aligned}$$

for $(x, y, \alpha) \in I_{x_0}^\delta \times \mathbb{R}_+ \times \Theta_3$ *and* $j = 3, \ldots, d_3$. *Furthermore, the derivatives of* c *and* $\log c$ *with respect to* α *are bounded on* $I_{x_0}^\delta \times \Theta_3$, *and so*

$$\sup_{(x,\alpha) \in I_{x_0}^\delta \times \Theta_3} \left| \frac{\partial^2 \psi}{\partial \alpha_j \partial y}(x, y, \alpha) \right| \leq O\left(\frac{1}{|y|^2} \right) \quad \text{as } y \to 0,$$

for $(x, y, \alpha) \in I_{x_0}^\delta \times \mathbb{R}_+ \times \Theta_3$ *with* $y/c(x, \alpha) \neq \alpha_1$. *Thus, Assumptions 3.2.9 to 3.2.12 are satisfied, and* ρ *in Theorems 3.3.1 and 3.3.2 can be taken as* $\rho \in (0, 1/8)$.

Example 3.5.4 (Weibull distribution)**.** *Let* Θ_3 *be smooth, open, convex and compactly contained in* $\mathbb{R}_+ \times (1, \infty) \times \mathbb{R}^{d_3 - 2}$, *and let* f_α *be of the form*

$$f_\alpha(z) = \begin{cases} \frac{\alpha_2}{\alpha_1} \left(\frac{z}{\alpha_1} \right)^{\alpha_2 - 1} e^{-(z/\alpha_1)^{\alpha_2}} & (z > 0), \\ 0 & (z \leq 0) \end{cases}$$

for $\alpha \in \Theta_3$. *Then,*

$$\psi(x, y, \alpha) = \frac{1}{c(x, \alpha)} \left\{ \log \alpha_2 - \alpha_2 \log \alpha_1 - (\alpha_2 - 1) \log \frac{y}{c(x, \alpha)} \right\} \quad \text{on } I_{x_0}^\delta \times \mathbb{R}_+ \times \Theta_3.$$

Since

$$f'_\alpha(z) = \left(\frac{\alpha_2 - 1}{z} - \alpha_2 \left(\frac{z}{\alpha_1}\right)^{\alpha_2 - 1}\right) f_\alpha(z) \quad (z \neq 0),$$

$$\frac{\partial f_\alpha}{\partial \alpha_1}(z) = -\frac{\alpha_2}{\alpha_1}\left\{1 + \left(\frac{z}{\alpha_1}\right)^{\alpha_2}\right\} f_\alpha(z), \quad \frac{\partial f_\alpha}{\partial \alpha_2}(z) = \left\{\frac{1}{\alpha_2} + \log\frac{z}{\alpha_1} - \left(\frac{z}{\alpha_1}\right)^{\alpha_2}\log\frac{z}{\alpha_1}\right\} f_\alpha(z)$$

for $z > 0$ *and* $\alpha \in \Theta_3$*, we have*

$$\begin{aligned}
\frac{\partial \psi}{\partial y}(x, y, \alpha) &= \frac{(\alpha_2 - 1)}{y} - \frac{\alpha_2}{c(x, \alpha)}\left(\frac{y}{\alpha_1 c(x, \alpha)}\right)^{\alpha_2 - 1} \\
\frac{\partial \psi}{\partial \alpha_1}(x, y, \alpha) &= -\frac{\partial(\log c)}{\partial \alpha_1}(x, \alpha)\left(1 + y\frac{\partial \psi}{\partial y}(x, y, \alpha)\right) - \frac{\alpha_2}{\alpha_1 c(x, \alpha)}\left\{1 + \left(\frac{y}{\alpha_1 c(x, \alpha)}\right)^{\alpha_2}\right\}, \\
\frac{\partial \psi}{\partial \alpha_2}(x, y, \alpha) &= -\frac{\partial(\log c)}{\partial \alpha_2}(x, \alpha)\left(1 + y\frac{\partial \psi}{\partial y}(x, y, \alpha)\right) \\
&\quad + \frac{1}{c(x, \alpha)}\left\{\frac{1}{\alpha_2} + \log\frac{y}{\alpha_1 c(x, \alpha)} - \left(\frac{y}{\alpha_1 c(x, \alpha)}\right)^{\alpha_2}\log\frac{y}{\alpha_1 c(x, \alpha)}\right\}, \\
\frac{\partial \psi}{\partial \alpha_j}(x, y, \alpha) &= -\frac{\partial(\log c)}{\partial \alpha_j}(x, \alpha)\left(1 + y\frac{\partial \psi}{\partial y}(x, y, \alpha)\right)
\end{aligned}$$

for $(x, y, \alpha) \in I_{x_0}^\delta \times \mathbb{R}_+ \times \Theta_3$ *and* $j = 3, \ldots, d_3$*. Furthermore, the derivatives of* c *and* $\log c$ *with respect to* α *are bounded on* $I_{x_0}^\delta \times \Theta_3$*, and so*

$$\sup_{(x,\alpha) \in I_{x_0}^\delta \times \Theta_3}\left|\frac{\partial^2 \psi}{\partial \alpha_j \partial y}(x, y, \alpha)\right| \leq O\left(\frac{1}{y}\right) \quad \text{as } y \to 0,$$

for $(x, y, \alpha) \in I_{x_0}^\delta \times \mathbb{R} \times \Theta_3$ *with* $y/c(x, \alpha) \neq \alpha_1$*, where* C *is a constant not depending on* (x, y, α)*. Here, we remark that*

$$\int \frac{1}{y} f_\alpha(y)\,\mathrm{d}y < \infty \quad \text{if and only if} \quad \alpha_2 > 1$$

and that there exists a constant $C > 0$ *such that*

$$|y^{\alpha_2 - 1}\log y| \leq |y_1^{\alpha_2 - 1}\log y_1| + |y_2^{\alpha_2 - 1}\log y_2| + C \quad \text{for } y_1 \leq y \leq y_2.$$

Thus, Assumptions 3.2.9 to 3.2.12 are satisfied, and ρ *in Theorems 3.3.1 and 3.3.2 can be taken as* $\rho \in (0, 1/4)$*.*

Example 3.5.5 (Log-normal distribution)**.** *Let* Θ_3 *be smooth, open, convex and compactly contained in* $\mathbb{R} \times [0, \infty)$*, and let* f_α *be of the form*

$$f_\alpha(z) = \begin{cases} \dfrac{1}{\sqrt{2\pi}\alpha_2 z} e^{-(\log z - \alpha_1)^2/2\alpha_2^2} & (z > 0), \\ 0 & (z \leq 0) \end{cases}$$

for $\alpha \in \Theta_3$. *Then,*

$$\psi(x,y,\alpha) = \frac{1}{c(x,\alpha)}\left\{-\log\frac{\sqrt{2\pi}\alpha_2 y}{c(x,\alpha)} - \frac{1}{2\alpha_2}\left|\log\frac{y}{c(x,\alpha)} - \alpha_1\right|^2\right\} \quad \text{on } I_{x_0}^\delta \times \mathbb{R}_+ \times \Theta_3.$$

Since

$$f'_\alpha(z) = \left\{-\frac{1}{z} - \frac{\log z - \alpha_1}{\alpha_2^2 z}\right\} f_\alpha(z),$$
$$\frac{\partial f_\alpha}{\partial \alpha_1}(z) = \frac{\log z - \alpha_1}{\alpha_2^2} f_\alpha(z), \quad \frac{\partial f_\alpha}{\partial \alpha_2}(z) = \left\{-\frac{1}{\alpha_2} + \frac{|\log z - \alpha_1|^2}{\alpha_2^3}\right\} f_\alpha(z)$$

for $z > 0$ *and* $\alpha \in \Theta_3$, *we have*

$$\frac{\partial\psi}{\partial y}(x,y,\alpha) = -\frac{1}{\alpha_2^2 y}\left(\alpha_1 + \alpha_2^2 + \log\frac{y}{c(x,\alpha)}\right)$$
$$\frac{\partial\psi}{\partial\alpha_1}(x,y,\alpha) = -\frac{\partial(\log c)}{\partial\alpha_1}(x,\alpha)\left(1 + y\frac{\partial\psi}{\partial y}(x,y,\alpha)\right) + \frac{\log\frac{y}{c(x,\alpha)} - \alpha_1}{\alpha_2^2 c(x,\alpha)},$$
$$\frac{\partial\psi}{\partial\alpha_2}(x,y,\alpha) = -\frac{\partial(\log c)}{\partial\alpha_2}(x,\alpha)\left(1 + y\frac{\partial\psi}{\partial y}(x,y,\alpha)\right) + \frac{1}{c(x,\alpha)}\left\{-\frac{1}{\alpha_2} + \frac{|\log\frac{y}{c(x,\alpha)} - \alpha_1|^2}{\alpha_2^3}\right\},$$
$$\frac{\partial\psi}{\partial\alpha_j}(x,y,\alpha) = -\frac{\partial(\log c)}{\partial\alpha_j}(x,\alpha)\left(1 + y\frac{\partial\psi}{\partial y}(x,y,\alpha)\right)$$

for $(x,y,\alpha) \in I_{x_0}^\delta \times \mathbb{R}_+ \times \Theta_3$ *and* $j = 3,\dots,d_3$. *Furthermore, the derivatives of* c *and* $\log c$ *with respect to* α *are bounded on* $I_{x_0}^\delta \times \Theta_3$, *and so*

$$\sup_{(x,\alpha)\in I_{x_0}^\delta \times \Theta_3}\left|\frac{\partial^2\psi}{\partial\alpha_j\partial y}(x,y,\alpha)\right| \le O\left(\frac{1}{y} + \frac{1}{y}\log y\right) \quad \text{as } y \to 0,$$

for $(x,y,\alpha) \in I_{x_0}^\delta \times \mathbb{R} \times \Theta_3$ *with* $y/c(x,\alpha) \neq \alpha_1$, *where* C *is a constant not depending on* (x,y,α). *Here, we remark that*

$$\int\left(\frac{1}{y} + \frac{\log y}{y}\right) f_\alpha(y)\,\mathrm{d}y < \infty$$

and that there exists a constant $C > 0$ *such that*

$$\left|\frac{1}{y}\log y\right| \le \left|\frac{1}{y_1}\log y_1\right| + \left|\frac{1}{y_2}\log y_2\right| + C \quad \text{for } y_1 \le y \le y_2.$$

Thus, Assumptions 3.2.9 to 3.2.12 are satisfied, and ρ *in Theorems 3.3.1 and 3.3.2 can be taken as* $\rho \in (0, 1/4)$.

Remark 3.5.1. *As in the assumptions of Theorems 3.3.1 and 3.3.2, the range of* ρ *depends on* q *in Assumption 3.2.10 (ii.b) and Assumption 3.2.12 (ii.b). So, the differences of the ranges of* ρ *in the examples above are caused by the differences of* q: $q = 2$ *in Example 3.5.3,* $q = 1$ *in Examples 3.5.2 and 3.5.4, and any* $q \in [0, 1)$ *in Example 3.5.5.*

3.6 Numerical experiments

In this section, we show some numerical results of our estimator for the Ornstein-Uhlenbeck processes given by

$$\mathrm{d}X_t^\varepsilon = -\mu_0 X_t^\varepsilon \,\mathrm{d}t + \varepsilon\sqrt{\sigma_0}\,\mathrm{d}W_t + \varepsilon\,\mathrm{d}Z_t^{\lambda_\varepsilon}, \quad X_0^\varepsilon = x_0 \in \mathbb{R}, \tag{3.21}$$

where $Z_t^{\lambda_\varepsilon}$ is a compound Poisson process with the Lévy density f_{α_0} and with the intensity λ_ε. In particular, we fix $x_0 = 0.8$ and $\lambda_\varepsilon = 100$, and we employ the inverse Gaussian densities f_α's as in Example 3.5.3.

To avoid the discussion about how we find some 'appropriate' v_{nk} and ρ, we suppose that the intensity $\lambda_\varepsilon = 100$ is known, and we set

$$\hat{C}_k^{N_D} := \left\{\Delta_k^n X^\varepsilon \text{ is not contained in the } \lceil N_D \rceil \text{ largest positive numbers of } \left\{\Delta_j^n X^\varepsilon\right\}_{j=1,\dots,n}\right\},$$
$$\hat{D}_k^{N_D} := \left\{\Delta_k^n X^\varepsilon \text{ is one of the } \lceil N_D \rceil \text{ largest positive values of } \left\{\Delta_j^n X^\varepsilon\right\}_{j=1,\dots,n}\right\}$$

where $N_D > 0$ and $\lceil\cdot\rceil$ is the ceil function (we take $N_D = \lambda_\varepsilon$ in Table 3.6.1, and $N_D = 50, 100, 150$ in Table 3.6.2). Then we replace $1_{C_k^{n,\varepsilon,\rho}}$ and $1_{D_k^{n,\varepsilon,\rho}}$ in (3.2) with

$$1_{\hat{C}_k^{N_D}} \text{ and } 1_{\hat{D}_k^{N_D}}, \text{ respectively,}$$

and we calculate our estimator $\hat{\theta}_{n,\varepsilon} = (\hat{\mu}_{n,\varepsilon}, \hat{\sigma}_{n,\varepsilon}, \hat{\alpha}_{n,\varepsilon,1}, \hat{\alpha}_{n,\varepsilon,2})$ as in (3.3) from a sample path of (3.21) under the true parameter $(\mu_0, \sigma_0, \alpha_{01}, \alpha_{02})$. We iterate this calculation 1000 times with $n = 200, 500, 1500, 5000$ and $\varepsilon = 1, 0.1, 0.01$. and we summarize the averages and the standard deviations of $\hat{\theta}_{n,\varepsilon}$'s in Tables 3.6.1 and 3.6.2.

Remark 3.6.1. *Note that $\hat{D}_k^{N_D}$ (and $\hat{C}_k^{N_D}$) are defined by using the whole data $\{X_{t_j}^\varepsilon\}_{j=1,\dots,n}$, which conflicts Assumption 3.2.8, however, for simplicity of our numerical experiment we replace $D_k^{n,\varepsilon,\rho}$ with $\hat{D}_k^{\lambda_\varepsilon}$ above. We give an intuitive explanation of the reason why we use this setting as follows: The continuous increments go to zero and the jumps are remained as $n \to \infty$ with ε fixed (recall that in our asymptotics n increases much faster than $1/\varepsilon$ and λ_ε as in Theorems 3.3.1 and 3.3.2), and in this case, from Lemma 3.4.8, $\{\Delta_k^n X^\varepsilon \mid \Delta_k^n X^\varepsilon > v_{nk}/n^\rho\}$ with 'appropriate' v_{nk} and ρ would be the λ_ε largest numbers of $\{X_{t_j}^\varepsilon\}_j$ in probability. Hence, we replace $D_k^{n,\varepsilon,\rho}$ with $\hat{D}_k^{\lambda_\varepsilon}$. Also, it follows from Lemma 3.4.8 that*[2]

[2]Proof of $\frac{1}{\lambda_\varepsilon}\sum_{k=1}^n 1_{\hat{D}_k^{\lambda_\varepsilon}} \xrightarrow{p} 1$: Take an arbitrary $\eta \in (0,1)$. If $\sum_{k=1}^n 1_{\hat{D}_k^{\lambda_\varepsilon}} < (1-\eta)\lambda_\varepsilon$, then the $\lceil\lambda_\varepsilon\rceil$-th largest number of $\{\Delta_j^n X^\varepsilon\}_{j=1,\dots,n}$ is negative, and so $\sum_{k=1}^n 1_{\hat{D}_k^{\lambda_\varepsilon}} = \sum_{k=1}^n 1_{\{\Delta_k^n X^\varepsilon > 0\}} \geq \sum_{k=1}^n 1_{D_k^{n,\varepsilon,\rho}}$. Thus, it follows from Lemma 3.4.8 that

$$P\left(\frac{1}{\lambda_\varepsilon}\sum_{k=1}^n 1_{\hat{D}_k^{\lambda_\varepsilon}} - 1 < -\eta\right) \leq P\left(\frac{1}{\lambda_\varepsilon}\sum_{k=1}^n 1_{D_k^{n,\varepsilon,\rho}} - 1 < -\eta\right) \to 0$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$, $\lambda_\varepsilon/n \to 0$ and $\varepsilon\lambda_\varepsilon \to 0$. Since $\sum_{k=1}^n 1_{\hat{D}_k^{N_D}} \leq \lceil N_D \rceil$ for any $N_D > 0$,

$$P\left(\frac{1}{\lambda_\varepsilon}\sum_{k=1}^n 1_{\hat{D}_k^{\lambda_\varepsilon}} - 1 > \eta\right) \leq P\left(\frac{\lceil\lambda_\varepsilon\rceil}{\lambda_\varepsilon} > 1 + \eta\right) \to 0 \quad \text{as } \lambda_\varepsilon \to \infty.$$

Remark 3.6.2. *For any $N_D > 0$, we have $\sum_{k=1}^{n} 1_{\hat{D}_k^{N_D}} \leq \lceil N_D \rceil$, and the equality holds when the $\lceil N_D \rceil$-th largest value of $\{\Delta_j^n X^\varepsilon\}_{j=1,\dots,n}$ is positive.*

In Table 3.6.1, the averages of $(\mu, \sigma, \alpha_1, \alpha_2)$ becomes close to the true paramter as n grows and ε goes to zero. However, the standard deviation of α_2 for each fixed ε increases as n grows. The reason why it happens is expected as follows: If n is not enough large with fixed ε, then the continuous increments in $\Delta_k^n X^\varepsilon$ is too larger than the jumps. In this case, some of $\Delta_k^n X^\varepsilon$'s including positive jumps may be negative, and furthermore even positive $\Delta_k^n X^\varepsilon$'s may be closer to zero than the jumps included in them. This implies that $\Delta_k^n X^\varepsilon$ with small jumps are ignored and the remained $\Delta_k^n X^\varepsilon$ regared as jumps are underestimated, and therefore, the mean and standard deviations of α_2 are near zero when n is few with fixed ε.

In Table 3.6.2, we consider the following two cases: One is $C_k^{n,\varepsilon,\rho}$ is too loose, *i.e.*, the case $N_D = 50$, and the other is $C_k^{n,\varepsilon,\rho}$ is too tight, *i.e.*, the case $N_D = 150$. In the former case, some small jumps are not removed for the estimation of (μ, σ) and are in short supply for the estimation of α. Thus, it is natural that $\sigma, \alpha_1, \alpha_2$ take bigger values than true values. In the latter case, some Brownian increments are mistakenly regarded as jumps, and so α_1 is closer to zero than the true value.

3.7 Appendix

In this section, we state and prove some slightly different versions of well-known results. More precisely, we prepare Lemma 3.7.2 as localization of the continuous mapping theorem. Lemma 3.7.3 is a slightly different version of Lemma 9 in Genon-Catalot and Jacod [11].

Lemma 3.7.1. *Let $\mathcal{X}$ be a Banach space, and let $\{g_\theta\}_{\theta\in\Theta}$ be a family of functions from $\mathcal{X}$ to $\mathbb{R}$, and let T_{g_θ} be the composition operator on $L^\infty([0,1];\mathcal{X})$ generated by g_θ, i.e.,*

$$T_{g_\theta}(\tilde{y}.) := g_\theta(\tilde{y}.) \quad \text{for } \tilde{y}. \in L^\infty([0,1];\mathcal{X}).$$

Suppose that $y.$ is a version of a function of $C([0,1];\mathcal{X})$ in $L^\infty([0,1];\mathcal{X})$, and that $\{g_\theta\}_{\theta\in\Theta}$ is equicontinuous at every points in Image$(y.) := \{y_t \,|\, t \in [0,1]\}$. *Then, there is a neighborhood $\mathscr{N}_{y.}$ of $y.$ in $L^\infty([0,1];\mathcal{X})$ such that $\{T_{g_\theta}\}_{\theta\in\Theta}$ is a family of operators from $\mathscr{N}_{y.}$ to $L^\infty([0,1])$, and is equicontinuous at $y.$.*

Proof. Fix an arbirary $\eta > 0$. For each $x \in \text{Image}(y.)$, there exists $\delta_x > 0$ such that if $\|x - x'\|_{\mathcal{X}} < \delta_x$, $x, x' \in \mathcal{X}$, then

$$\sup_{\theta\in\Theta} |g_\theta(x) - g_\theta(x')| < \frac{\eta}{2}.$$

Since Image$(y.)$ is compact in $\mathcal{X}$, there are finite points $x_1, \dots, x_k \in \text{Image}(y.)$ such that

$$\text{Image}(y.) \subset \bigcup_{i=1}^{k} B(x_i, \delta_{x_i}/2),$$

Table 3.6.1: Sample means (with standard deviations in parentheses) of $\hat{\theta}_{n,\varepsilon}$'s, based on 1000 sample paths from the OU process (3.21) with inverse Gaussian f_α as in Example 3.5.3 with $(\mu_0, \sigma_0, \alpha_{01}, \alpha_{02}) = (1.0, 2.0, 1.2, 0.5)$ and with $N_D = \lambda_\varepsilon (= 100)$.

		$n = 200$	$n = 500$	$n = 1500$	$n = 5000$	true
μ	$\varepsilon = 1.00$	0.989141	1.016899	1.007082	1.002970	1.0
		(0.068480)	(0.061120)	(0.050175)	(0.048399)	
	$\varepsilon = 0.10$	0.978554	1.024855	1.010180	1.001047	
		(0.060345)	(0.055804)	(0.045878)	(0.043720)	
	$\varepsilon = 0.01$	0.912885	1.005487	1.010121	1.000906	
		(0.030783)	(0.026575)	(0.026601)	(0.023244)	
σ	$\varepsilon = 1.00$	1.920753	1.886210	1.968844	2.002727	2.0
		(0.165435)	(0.086929)	(0.053557)	(0.039613)	
	$\varepsilon = 0.10$	1.942879	1.874213	1.969947	2.002262	
		(0.163257)	(0.085449)	(0.050984)	(0.035897)	
	$\varepsilon = 0.01$	2.379172	1.932349	1.961293	2.000971	
		(0.179838)	(0.074998)	(0.053425)	(0.035739)	
α_1	$\varepsilon = 1.00$	1.326379	1.160758	1.192697	1.178391	1.2
		(0.288811)	(0.200419)	(0.222514)	(0.211085)	
	$\varepsilon = 0.10$	1.381731	1.129643	1.173607	1.188770	
		(0.308788)	(0.205739)	(0.204944)	(0.212477)	
	$\varepsilon = 0.01$	1.731430	1.231199	1.153611	1.182000	
		(0.371265)	(0.259936)	(0.204695)	(0.210043)	
α_2	$\varepsilon = 1.00$	0.099654	0.358790	0.500877	0.533962	0.5
		(0.060910)	(0.095974)	(0.155927)	(0.215146)	
	$\varepsilon = 0.10$	0.109767	0.322201	0.483286	0.527680	
		(0.067662)	(0.096333)	(0.144849)	(0.202786)	
	$\varepsilon = 0.01$	0.266864	0.077035	0.363671	0.490374	
		(0.177911)	(0.055310)	(0.140516)	(0.208392)	

Table 3.6.2: Sample means (with standard deviations in parentheses) of $\hat{\theta}_{n,\varepsilon}$'s, based on 1000 sample paths from the OU process (3.21) with inverse Gaussian f_α as in Example 3.5.3 with $(\mu_0, \sigma_0, \alpha_{01}, \alpha_{02}) = (1.0, 2.0, 1.2, 0.5)$ and with $(n, \varepsilon, \lambda_\varepsilon) = (5000, 0.01, 100)$.

	μ	σ	α_1	α_2
true	1.0	2.0	1.2	0.5
$N_D = 50$	0.851616	3.028911	2.103949	3.025209
	(0.045741)	(0.424354)	(0.392786)	(0.971591)
$N_D = 100$	**1.000666**	**2.001121**	**1.183114**	**0.484350**
	(0.024807)	(0.037437)	(0.211747)	(0.205930)
$N_D = 150$	1.040850	1.933562	0.807600	0.144996
	(0.024657)	(0.022237)	(0.143919)	(0.022638)

where $B(x_i, \delta_{x_i}/2)$ is the ball in $\mathcal{X}$ centered at x_i with radius $\delta_{x_i}/2$. If $\|\tilde{y}_\cdot - y_\cdot\|_{L^\infty([0,1];\mathcal{X})} < \delta$ with $\delta := \min\{\delta_{x_1}/2, \ldots, \delta_{x_k}/2\}$, then for a.e. $t \in [0,1]$ there is $i_t \in \{1, \ldots, k\}$ such that $y_t, \tilde{y}_t \in B(x_{i_t}, \delta_{x_{i_t}})$. Thus, we obtain

$$\sup_{\theta\in\Theta} |g_\theta(\tilde{y}_t) - g_\theta(y_t)| \le \sup_{\theta\in\Theta} |g_\theta(\tilde{y}_t) - g_\theta(x_{i_t})| + \sup_{\theta\in\Theta} |g_\theta(x_{i_t}) - g_\theta(y_t)| < \eta,$$

that is,

$$\sup_{\theta\in\Theta} \|g_\theta(\tilde{y}_\cdot) - g_\theta(y_\cdot)\|_{L^\infty([0,1])} < \eta.$$

This implies the conclusion. □

We prepare the following lemma as localization of the continuous mapping theorem.

Lemma 3.7.2. *Under the same assumptions as in Lemma 3.7.1, suppose that* $\{g(\cdot, \theta)\}_{\theta\in\Theta}$ *is equicontinuous at every points in* $\mathrm{Image}(y_\cdot) := \{y_t \mid t \in [0,1]\}$*, and that* $(Y_\cdot^\iota)_{\iota\in I}$ *is a net of* $\mathcal{X}$*-valued bounded random processes on* $[0,1]$ *with a directed set* I*. If the net* $(Y_\cdot^\iota)_{\iota\in I}$ *converges in probability to* $y_\cdot$ *in* $L^\infty([0,1;\mathcal{X}])$*, i.e.,*

$$\|Y_\cdot^\iota - y_\cdot\|_{L^\infty([0,1];\mathcal{X})} \xrightarrow{p} 0,$$

then

$$\sup_{\theta\in\Theta} \|g(Y_\cdot^\iota, \theta) - g(y_\cdot, \theta)\|_{L^\infty([0,1])} \xrightarrow{p} 0.$$

Proof. Take an arbitrary $\eta > 0$. It follows from Lemma 3.7.1 that there exists a sufficiently small $\delta > 0$ such that if $\|\tilde{y}_\cdot - y_\cdot\|_{L^\infty([0,1];\mathcal{X})} < \delta$, then $\{g(\tilde{y}_\cdot, \theta)\}_{\theta\in\Theta} \subset L^\infty([0,1])$ and

$$\sup_{\theta\in\Theta} \|g(\tilde{y}_\cdot, \theta) - g(y_\cdot, \theta)\|_{L^\infty([0,1])} < \eta,$$

and therefore,

$$P\left(\sup_{\theta\in\Theta} \|g(Y_\cdot^\iota, \theta) - g(y_\cdot, \theta)\|_{L^\infty([0,1])} > \eta\right) \le P\left(\|Y_\cdot^\iota - y_\cdot\|_{L^\infty([0,1];\mathcal{X})} > \delta\right).$$

This implies the conclusion. □

Remark 3.7.1. *By the proof of Lemma 3.7.2, it also follows that for any* $C_1 > 0$,

$$P\left(\sup_{\theta\in\Theta}\|g(Y_\cdot^\iota,\theta)-g(y_\cdot,\theta)\|_{L^\infty([0,1])} > C_2\right) \leq P\left(\|Y_\cdot^\iota - y_\cdot\|_{L^\infty([0,1];\mathcal{X})} > C_1\right),$$

where C_2 *depends only on* C_1, g *and* $\mathrm{Image}(y_\cdot)$.

Lemma 3.7.3. *Suppose that* $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ *is a Banach space,* $\{(n,\varepsilon)\}$ *is a directed set and* $\{\mathcal{G}_i^{n,\varepsilon}\}_i$ *is a filtration for each* n, ε. *Let* $\chi_i^{n,\varepsilon}$, U *be* $\mathcal{X}$*-valued* $\mathcal{G}_i^{n,\varepsilon}$*-measurable random variables.*

(i) If for any $\eta > 0$

$$\lim_{n,\varepsilon} P\left(\sum_{i=1}^n E\left[\|\chi_i^{n,\varepsilon}\|_{\mathcal{X}} \,\middle|\, \mathcal{G}_{i-1}^{n,\varepsilon}\right] > \eta\right) = 0,$$

then for any $\eta > 0$

$$\lim_{n,\varepsilon} P\left(\left\|\sum_{i=1}^n \chi_i^{n,\varepsilon}\right\|_{\mathcal{X}} > \eta\right) = 0.$$

(ii) If

$$\lim_{M\to\infty}\sup_{n,\varepsilon} P\left(\sum_{i=1}^n E\left[\|\chi_i^{n,\varepsilon}\|_{\mathcal{X}} \,\middle|\, \mathcal{G}_{i-1}^{n,\varepsilon}\right] > M\right) = 0,$$

then

$$\lim_{M\to\infty}\sup_{n,\varepsilon} P\left(\left\|\sum_{i=1}^n \chi_i^{n,\varepsilon}\right\|_{\mathcal{X}} > M\right) = 0.$$

Proof. Since for any $\eta, \eta' > 0$

$$\begin{aligned}
P\left(\left\|\sum_{i=1}^n \chi_i^{n,\varepsilon}\right\|_{\mathcal{X}} > \eta,\ \sum_{i=1}^n E\left[\|\|\chi_i^{n,\varepsilon}\|_{\mathcal{X}} \,\middle|\, \mathcal{G}_{i-1}^{n,\varepsilon}\right] \leq \eta'\right) &\leq \frac{1}{\eta} E\left[\sum_{i=1}^n \|\chi_i^{n,\varepsilon}\|_{\mathcal{X}}\, 1_{\{\sum_{i=1}^n E[\|\chi_i^{n,\varepsilon}\|_{\mathcal{X}} \,|\, \mathcal{G}_{i-1}^{n,\varepsilon}] \leq \eta'\}}\right] \\
&\leq \frac{1}{\eta} E\left[\left(E\left[\|\chi_n^{n,\varepsilon}\|_{\mathcal{X}} \,\middle|\, \mathcal{G}_{n-1}^{n,\varepsilon}\right] + \sum_{i=1}^{n-1}\|\chi_i^{n,\varepsilon}\|_{\mathcal{X}}\right) 1_{\{\sum_{i=1}^n E[\|\chi_i^{n,\varepsilon}\|_{\mathcal{X}} \,|\, \mathcal{G}_{i-1}^{n,\varepsilon}] \leq \eta'\}}\right] \\
&\leq \frac{1}{\eta} E\left[\eta' + \sum_{i=1}^{n-1}\left(\|\chi_i^{n,\varepsilon}\|_{\mathcal{X}} - E\left[\|\chi_i^{n,\varepsilon}\|_{\mathcal{X}} \,\middle|\, \mathcal{G}_{i-1}^{n,\varepsilon}\right]\right) 1_{\{E[\|\chi_i^{n,\varepsilon}\|_{\mathcal{X}} \,|\, \mathcal{G}_{i-1}^{n,\varepsilon}] \leq \eta'\}}\right] < \frac{\eta'}{\eta},
\end{aligned}$$

we obtain

$$P\left(\left\|\sum_{i=1}^n \chi_i^{n,\varepsilon}\right\|_{\mathcal{X}} > \eta\right) \leq \frac{\eta'}{\eta} + P\left(\sum_{i=1}^n E\left[\|\chi_i^{n,\varepsilon}\|_{\mathcal{X}} \,\middle|\, \mathcal{G}_{i-1}^{n,\varepsilon}\right] > \eta'\right).$$

Thus, the assertions (i) and (ii) follows. □

Remark 3.7.2. *When* $\mathcal{X} = \mathbb{R}$, *this lemma can be shown by the same argument in the proof of Lemma 9 in Genon-Catalot and Jacod [11]. However, the argument does not work in general, since we may not have Lenglart's inequality (*e.g.*, Lemma 3.30 in Jacod and Shiryaev [18]) when* $\mathcal{X}$ *is a Banach space.*

Remark 3.7.3. *We have an immediate consequence from this lemma that*

$$\sum_{i=1}^{n} E\left[\|\chi_i^{n,\varepsilon}\| \,\middle|\, \mathcal{G}_{i-1}^{n,\varepsilon}\right] = o_p(r_{n,\varepsilon}) \quad \Longrightarrow \quad \sum_{i=1}^{n} \chi_i^{n,\varepsilon} = o_p(r_{n,\varepsilon}),$$
$$\sum_{i=1}^{n} E\left[\|\chi_i^{n,\varepsilon}\| \,\middle|\, \mathcal{G}_{i-1}^{n,\varepsilon}\right] = O_p(r_{n,\varepsilon}) \quad \Longrightarrow \quad \sum_{i=1}^{n} \chi_i^{n,\varepsilon} = O_p(r_{n,\varepsilon}),$$

where $r_{n,\varepsilon} \in \mathbb{R}$.

Bibliography

[1] C. Amorino and A. Gloter, *Contrast function estimation for the drift parameter of ergodic jump diffusion process*, Scand. J. Statist. **47** (2019), no. 2, 279–346.

[2] C. Amorino and A. Gloter, *Joint estimation for volatility and drift parameters of ergodic jump diffusion processes via contrast function. (English summary)*, Stat. Inference Stoch. Process **24** (2021), no. 1, 61-148.

[3] D. Applebaum, *Lévy processes and stochastic calculus*, second ed., Cambridge Studies in Advanced Mathematics, vol. 116, Cambridge University Press, Cambridge, 2009.

[4] P. Billingsley, *Convergence of probability measures*, second ed., Wiley Series in Probability and Statistics: Probability and Statistics, A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York, 1999.

[5] H. Brezis, *Functional analysis, Sobolev spaces and partial differential equations*, Universitext, Springer, New York, 2011.

[6] , D. N. Burghes and M. S. Borrie, *Modelling with differential equations*, Ellis Horwood Ltd., Chichester, New York, 1981.

[7] J. C. Butcher, *Numerical methods for ordinary differential equations*, third edition ed., John Wiley & Sons, Ltd., Chichester, 2016.

[8] P. J. Davis, *Interpolation and approximation. republication, with minor corrections, of the 1963 original, with a new preface and bibliography*, Dover Publications, Inc., New York, 1975.

[9] L. C. Evans, *Partial differential equations*, second ed., Graduate Studies in Mathematics, vol. 19, American Mathematical Society, Providence, RI, 2010.

[10] V. Genon-Catalot, *Maximum contrast estimation for diffusion processes from discrete observations*, Statistics **21** (1990), no. 1, 99–116.

[11] V. Genon-Catalot and J. Jacod, *On the estimation of the diffusion coefficient for multidimensional diffusion processes*, Annales de l'Institut Henri Poincaré Probabilités et Statistiques **29** (1993), no. 1, 119-151.

[12] A. Gloter, D. Loukianova, and H. Mai, *Jump filtering and efficient drift estimation for Lévy-driven SDEs*, Ann. Statist. **46** (2018), no. 4, 1445-1480.

[13] A. Gloter and M. Sørensen, *Estimation for stochastic differential equations with a small diffusion coefficient*, Stochastic Process. Appl. **119** (2009), no. 3, 679-699.

[14] E. Hairer, S. P. Nørsett, and G. Wanner, *Solving ordinary differential equations. i. nonstiff problems*, second ed., Springer Series in Computational Mathematics, vol. 8, Springer-Verlag, Berlin, 1993.

[15] E. Hairer and G. Wanner, *Solving ordinary differential equations. ii. stiff and differential-algebraic problems*, second revised ed., Springer Series in Computational Mathematics, vol. 14, Springer-Verlag, Berlin, 2010.

[16] A. Iserles, *A first course in the numerical analysis of differential equations*, 2 ed., Cambridge Texts in Applied Mathematics, Cambridge University Press, 2008.

[17] I. A. Ibragimov and R. Z. Has'minskii, *Statistical estimation. asymptotic theory. translated by Samuel Kotz*, Applications of Mathematics, vol. 16, Springer-Verlag, New York-Berlin, 1981.

[18] J. Jacod and A. N. Shiryaev, *Limit theorems for stochastic processes*, second ed., vol. 288, Springer-Verlag, Berlin, 2003.

[19] M. Kobayashi and Y. Yamada, *Global existence of weak solutions to forest kinematic model with nonlinear degenerate diffusion*. Adv. Math. Sci. Appl. **29** (2020), no. 1, 187-209.

[20] M. Kobayashi and Y. Shimizu, *Least-squares estimators based on the Adams method for stochastic differential equations with small Lévy noise*, Japanese Journal of Statistics and Data Science **5** (2022), 217-240.

[21] C. F. Laredo, *A sufficient condition for asymptotic sufficiency of incomplete observations of a diffusion process*, Ann. Statist. **18** (1990), no. 3, 1158–1171.

[22] H. Long, Y. Shimizu, and W. Sun, *Least squares estimators for discretely observed stochastic processes driven by small Lévy noises*, J. Multivariate Anal. **116** (2013), 422-439.

[23] H. Long, C. Ma, and Y. Shimizu, *Least squares estimators for stochastic differential equations driven by small lévy noises*, Stochastic Process. Appl. **127** (2017), no. 5, 1475–1495.

[24] X. Mao, *Stochastic differential equations and applications*, second ed., Horwood Publishing Limited, Chichester, 2008.

[25] T. Ogihara and N. Yoshida, *Quasi-likelihood analysis for the stochastic differential equation with jumps*, Stat. Inference Stoch. Process **14** (2011), no. 3, 189-229.

[26] B. L. S. Prakasa Rao, *Semimartingales and their statistical inference*, Monographs on Statistics and Applied Probability, vol. 83, Chapman & Hall/CRC, Boca Raton, FL, 1999.

[27] Y. Shimizu, *m-estimation for discretely observed ergodic diffusion processes with infinitely many jumps*, Stat. Inference Stoch. Process **9** (2006), no. 2, 179–225.

[28] Y. Shimizu, *Asymptotic inference for stochastic differential equations with jumps from discrete observations and some practical approaches*, Ph.D. thesis, University of Tokyo, 2007.

[29] Y. Shimizu, *A practical inference for discretely observed jump-diffusions from finite samples*, Journal of the Japan Statistical Society **38** (2008), no. 3, 391-413.

[30] Y. Shimizu, *Threshold selection in jump-discriminant filter for discretely observed jump processes*, Statistical Methods & Applications **19** (2010), no. 3, 355-378.

[31] Y. Shimizu, *Threshold estimation for stochastic processes with small noise*, Scand. J. Stat. **44** (2017), no. 4, 951–988.

[32] Y. Shimizu and N. Yoshida, *Estimation of parameters for diffusion processes with jumps from discrete observations*, Stat. Inference Stoch. Process **9** (2006), no. 3, 227-277.

[33] M. Sørensen and M. Uchida, *Small-diffusion asymptotics for discretely sampled stochastic differential equations*, Bernoulli **9** (2003), no. 6, 1051–1069.

[34] M. Uchida, *Estimation for discretely observed small diffusions based on approximate martingale estimating functions*, Scand. J. Statist. **31** (2004), no. 4, 553–566.

[35] A. W. van der Vaart, *Asymptotic statistics*, Cambridge Series in Statistical and Probabilistic Mathematics, vol. 3, Cambridge University Press, Cambridge, 1998.

[36] N. Yoshida, *Asymptotic behavior of m-estimator and related random field for diffusion process*, Ann. Inst. Statist. Math. **42** (1990), no. 2, 221-251.

List of research achievements for application of Doctor of Science, Waseda University

Full Name : 小林　光木　　　　seal or signature

Date Submitted(yyyy/mm/dd): 2023/06/21

種類別 (By Type)	題名、 発表・発行掲載誌名、 発表・発行年月、 連名者（申請者含む） (theme, journal name, date & year of publication, name of authors inc. yourself)
Academic paper	○Kobayashi, M. and Shimizu, Y.; Threshold estimation for jump-diffusions under small noise asymptotics; Statistical Inference for Stochastic Processes; 26; 361-411; published 20 February 2023.
Academic paper	○Kobayashi, M. and Shimizu, Y.; Least-squares estimators based on the Adams method for stochastic differential equations with small Levy noise; Japanese Journal of Statistics and Data Science; 5; 217-240; published 13 May 2022.
Academic paper	Kobayashi, M. and Yamada, Y.; Global existence of weak solutions to forest kinematic model with nonlinear degenerate diffusion; Advances in Mathematical Sciences and Applications; 29, (1); 187–209; November 2020.
Lecture	Kobayashi, M.; Threshold estimation for jump-diffusions under small noise asymptotics; MATRIX program “Mathematics of Risk – 2022”; 2 November 2022
Lecture	Kobayashi, M.; Threshold estimation for jump-diffusions under small noise asymptotics; 2022年度統計関連学会連合大会; 5 Septeember 2022.
Lecture	Kobayashi, M.; Least squares estimators based on the Adams method for discretely sampled SDEs with small Lévy noise; 数学･数理科学専攻若手研究者のための異分野･異業種研究交流会2021; 13 November 2021.
Lecture	Kobayashi, M.; Least squares estimators based on the Adams method for discretely sampled SDEs with small Lévy noise; 確率過程の統計推測の最近の展開2021; 11 March 2021.
Lecture	Kobayashi, M.; アダムス法を用いた 常微分方程式の統計的推測; 2019年度統計関連学会連合大会; 11 September 2019.
Lecture	Kobayashi. M.; 非線形退化型拡散を伴う 森林モデルにおける一様有界性; 第12 回非線形変微分方程式と変分問題; 10 February 2018.