Statistical Inference for Stochastic Processes (2023) 26:361–411
https://doi.org/10.1007/s11203-023-09286-y

Threshold estimation for jump-diffusions under small noise asymptotics

Mitsuki Kobayashi[1] · **Yasutaka Shimizu**[1]

Received: 31 August 2022 / Accepted: 19 January 2023 / Published online: 20 February 2023

Abstract
We consider parameter estimation of stochastic differential equations driven by a Wiener process and a compound Poisson process as small noises. The goal is to give a threshold-type quasi-likelihood estimator and show its consistency and asymptotic normality under new asymptotics. One of the novelties of the paper is that we give a new localization argument, which enables us to avoid truncation in the contrast function that has been used in earlier works and to deal with a wider class of jumps in threshold estimation than ever before.

Keywords Small noise asymptotics · Asymptotic distribution · Discrete observations

Mathematics Subject Classification Primary 62M20 · Secondary: 62F12 · 60J74

1 Introduction

This paper is concerned with the following stochastic differential equation (SDE):

$$\mathrm{d}X_t^\varepsilon = a(X_t^\varepsilon, \mu_0)\mathrm{d}t + \varepsilon b(X_t^\varepsilon, \sigma_0)\mathrm{d}W_t + \varepsilon c(X_{t-}^\varepsilon, \alpha_0)\mathrm{d}Z_t^{\lambda_\varepsilon}, \quad X_0^\varepsilon = x_0 \in \mathbb{R}, \tag{1.1}$$

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, $Z^{\lambda_\varepsilon} = (Z_t^{\lambda_\varepsilon})_{t \geq 0}$ is a compound Poisson process given by

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

✉ Mitsuki Kobayashi
mitsuki@fuji.waseda.jp

Yasutaka Shimizu
shimizu@waseda.jp

[1] Department of Pure and Applied Mathematics, Waseda University, 3-4-1 Ohkubo, Shinjuku-ku, Tokyo 169-8555, Japan

where $N^{\lambda_\varepsilon} = (N_t^{\lambda_\varepsilon})_{t\geq 0}$ 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 (2009)]. $W = (W_t)_{t\geq 0}$ is a Wiener process. Here, we denote the filtered probability space by $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P)$. Suppose that we have discrete data $X^\varepsilon_{t_0}, \ldots, X^\varepsilon_{t_n}$ from (1.1) 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.

In the ergodic case, threshold estimation for SDEs with Lévy noise is proposed in Shimizu and Yoshida (2006), and has been considered so far by various researchers [see, *e.g.*, Amorino and Gloter 2019; Gloter et al. 2018; Ogihara and Yoshida 2011; Shimizu 2017, and other references are given in Amorino and Gloter (2021)]. 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 in this paper, we 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., Gloter and Sørensen 2009; Kobayashi and Shimizu 2022; Long et al. 2013; Sørensen and Uchida 2003, and references are given in Long et al. (2017)]. 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 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 this paper is to give a proof by using localization argument [as in, e.g., Remark 1 in Sørensen and Uchida (2003)] in the entire context, though the argument is usually omitted, or instead, Propostion 1 in Gloter and Sørensen (2009) is just referred. As to the proof, we prepare the localization assumptions for jump size densities, i.e., Assumptions 2.9 to 2.12, together with usual localization assumptions for coefficient functions in (1.1), *i.e.*, Assumptions 2.5 and 2.6. Owing to prepare Assumptions 2.9 to 2.12, this paper has more examples of jump size densities than the papers (Ogihara and Yoshida 2011; Shimizu and Yoshida 2006) (see Sect. 5 in this paper, and see, e.g., Ogihara and Yoshida 2011, Example). On the other hand, Assumptions 2.9 to 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 of our paper.

A further attempt of this paper is to simplify the contrast functions used in earlier works (Ogihara and Yoshida 2011; Shimizu and Yoshida 2006) by removing φ_n defined in Ogihara and Yoshida (2011) and Shimizu and Yoshida (2006) 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 Ogihara and Yoshida (2011) and Shimizu and Yoshida (2006). Note that the class of jump size densities in Shimizu (2006) is

also wide (Shimizu 2006 does not assume the twice differentiability of jump size densities, while conversely this paper 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 Shimizu (2006)], but (Shimizu 2006) is concerned with moment estimators in the ergodic case.

In order to see the behavior of our estimator in numerical experiments, we give Table 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 2.7 is one of the crucial points for estimation with jumps, but it is not within the scope of this paper (see, *e.g.*, Shimizu 2008, 2010 for the readers who are interested in the techniques of the way to choose such threshold, and then Lemma 4.8 may also help you estimate the intensity λ_ε). Instead of this discussion, we give another experiment as in Table 2 to see what will occur by using different thresholds.

In Sect. 2, we set up some assumptions and notations. In Sect. 3, we state our main results, *i.e.*, the consistency and the asymptotic normality of our estimator. In Sect. 4, we give a proof of our main results. In Sect. 5, we give some examples of the jump size density for compound Poisson processes in our model. In Sect. 6, we give two numerical experiments to see the finite sample performance of our estimator. In "Appendix A", we state and prove some slightly different versions of well-known results.

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

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

Then, we prepare the following assumptions:

Assumption 2.1 $a(\cdot, \mu_0)$, $b(\cdot, \sigma_0)$ and $c(\cdot, \alpha_0)$ are Lipschitz continuous on $\mathbb{R}$.

Assumption 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 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 2.4 The family $\{f_\alpha\}_{\alpha \in \bar{\Theta}_3}$ satisfies either of the following condtions:

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

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

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

Assumption 2.8 $v_{n1}, \ldots, v_{nn}$ 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

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

for some constants v_1 and v_2.

Assumption 2.9 There exists $\delta > 0$ such that for $(x, y, \alpha) \in I^\delta_{x_0} \times \mathbb{R} \times \Theta$ with $\psi \neq 0$, ψ is differentiable with respect to α_i $(i = 1, \ldots, d_3)$. 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

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

Assumption 2.10 Relative to the choice (i) or (ii) in Assumption 2.4, we assume either of the following conditions (i) or (ii), respectively:

(i) Under Assumption 2.4 (i), there exist constants $C > 0$, $q \geq 1$ and $\delta > 0$ such that

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

(ii) Under Assumption 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

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

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

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

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

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

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

Assumption 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 $\left\{\frac{\partial^2 a}{\partial\theta_i\partial\theta_j}(\cdot, \mu)\right\}_{\mu\in\Theta_1}$, $\left\{\frac{\partial^2 b}{\partial\theta_i\partial\theta_j}(\cdot, \sigma)\right\}_{\sigma\in\Theta_2}$, $\left\{\frac{\partial^2 c}{\partial\theta_i\partial\theta_j}(\cdot, \alpha)\right\}_{\alpha\in\Theta_3}$ $(i, j = 1, \ldots, d)$ 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 $(i = 1, \ldots, d_3)$. For $\alpha \in \Theta$, $i = 1, \ldots, d_3$

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

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

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

Relative to the choice (i) or (ii) in Assumption 2.4, we assume either of the following conditions (i) or (ii), respectively:

(i) Under Assumption 2.4 (i), there exist constants $C > 0$, $q \geq 1$ and $\delta > 0$ such that

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

(ii) Under Assumption 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

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

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

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

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

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

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

Remark 2.1 Instead of Assumptions 2.5 and 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 2003). However, the 'classical' localization argument mentioned in Sørensen and Uchida (2003) is hard to apply for our purpose. Thus, we employ our milder assumptions and show how it works well.

Remark 2.2 Under Assumption 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 $x \in I_{x_0}^{\delta}$.

Remark 2.3 Assumption 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 2.2, 2.9 and 2.10, and is needed for obtaining the convergence (4.16) of the matrix containing the second derivatives of the contrast function.

Furthermore, we introduce the following notations:

Notation 2.3 *Denote*

$$\Delta X_t^{\varepsilon} := X_t^{\varepsilon} - X_{t-}^{\varepsilon} \quad \textit{for } t > 0,$$

where $\varepsilon > 0$.

Notation 2.4 *Denote*

$$\Delta_k^n X^{\varepsilon} := X_{t_k}^{\varepsilon} - X_{t_{k-1}}^{\varepsilon}, \quad \Delta_k^n N^{\lambda_\varepsilon} := N_{t_k}^{\lambda_\varepsilon} - N_{t_{k-1}}^{\lambda_\varepsilon} \quad \textit{for } k = 1, \ldots, n,$$

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

Notation 2.5 *Define random times*

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

Notation 2.6 *Define events* $J_{k,i}^{n,\varepsilon}$ $(k = 1, \ldots, n,\ i = 0, 1, 2)$ *by*

$$J_{k,0}^{n,\varepsilon} := \left\{\Delta_k^n N^{\lambda_\varepsilon} = 0\right\}, \quad J_{k,1}^{n,\varepsilon} := \left\{\Delta_k^n N^{\lambda_\varepsilon} = 1\right\}, \quad J_{k,2}^{n,\varepsilon} := \left\{\Delta_k^n N^{\lambda_\varepsilon} \geq 2\right\}$$

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

Notation 2.7 *Under Assumption 2.8, set events* $C_k^{n,\varepsilon,\rho}$ *and* $D_k^{n,\varepsilon,\rho}$ $(k = 1, \ldots, n)$ *by*

$$C_k^{n,\varepsilon,\rho} := \begin{cases} \left\{ \left| \Delta_k^n X^\varepsilon \right| \le \frac{v_{nk}}{n^\rho} \right\} & \text{under Assumption 2.4 (i),} \\ \left\{ \Delta_k^n X^\varepsilon \le \frac{v_{nk}}{n^\rho} \right\} & \text{under Assumption 2.4 (ii),} \end{cases}$$

$$D_k^{n,\varepsilon,\rho} := \begin{cases} \left\{ \left| \Delta_k^n X^\varepsilon \right| > \frac{v_{nk}}{n^\rho} \right\} & \text{under Assumption 2.4 (i),} \\ \left\{ \Delta_k^n X^\varepsilon > \frac{v_{nk}}{n^\rho} \right\} & \text{under Assumption 2.4 (ii),} \end{cases}$$

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

$$C_{k,i}^{n,\varepsilon,\rho} := C_k^{n,\varepsilon,\rho} \cap J_{k,i}^{n,\varepsilon}, \quad D_{k,i}^{n,\varepsilon,\rho} := D_k^{n,\varepsilon,\rho} \cap J_{k,i}^{n,\varepsilon} \quad \text{for } k = 1, \ldots, n,\ i = 0, 1, 2,$$

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

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

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

Remark 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 2.5 In this paper, 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 Main results

We define the following contrast function $\Psi_{n,\varepsilon}(\theta)$ after the quasi-log likelihood proposed in Shimizu (2017):

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

where for $\rho \in (0, 1/2)$, $\Psi^{(1)}_{n,\varepsilon}(\mu, \sigma)$ and $\Psi^{(2)}_{n,\varepsilon}(\alpha)$ are given by using Notations 2.4 and 2.7 as the following:

$$\Psi^{(1)}_{n,\varepsilon}(\mu,\sigma) := -\frac{1}{n}\sum_{k=1}^{n}\left\{\frac{\left|\Delta_k^n X^\varepsilon - \frac{1}{n}a(X^\varepsilon_{t_{k-1}},\mu)\right|^2}{2\frac{1}{n}\left|\varepsilon b(X^\varepsilon_{t_{k-1}},\sigma)\right|^2} + \frac{1}{2}\log|b(X^\varepsilon_{t_{k-1}},\sigma)|^2\right\} 1_{C_k^{n,\varepsilon,\rho}},$$

$$\Psi^{(2)}_{n,\varepsilon}(\alpha) := \frac{1}{\lambda_\varepsilon}\sum_{k=1}^{n}\psi\left(X^\varepsilon_{t_{k-1}}, \frac{\Delta_k^n X^\varepsilon}{\varepsilon}, \alpha\right) 1_{D_k^{n,\varepsilon,\rho}} \tag{3.1}$$

with

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

Then, the quasi-maximum likelihood estimator is given by

$$\hat{\theta}_{n,\varepsilon} := \underset{\theta\in\Theta}{\operatorname{argmax}}\, \Psi_{n,\varepsilon}(\theta). \tag{3.2}$$

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 4.8:

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

Theorem 3.1 *Under Assumptions 2.1 to 2.10, take ρ as either of the following:*

(i) Under Assumption 2.4 (i), take $\rho \in (0, 1/2)$.
(ii) Under Assumption 2.4 (ii), take $\rho \in (0, \min\{1/2, 1/4q\})$, where q is the constant in Assumption 2.10 Assumption (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 2.6 and 2.8, respectively.

Theorem 3.2 *Under Assumptions 2.1 to 2.12, take ρ as either of the following:*

(i) Under Assumption 2.4 (i), take $\rho \in (0, 1/2)$.
(ii) Under Assumption 2.4 (ii), take $\rho \in (0, \min\{1/2, 1/4q\})$, where q is the constant in Assumptions 2.10 (ii.b) and 2.12 (ii.b).

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

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

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

$$\begin{aligned} I_1^{ij} &:= \int_0^1 \frac{\frac{\partial a}{\partial \mu_i}\frac{\partial a}{\partial \mu_j}(x_t, \mu_0)}{|b(x_t, \mu_0)|^2}\mathrm{d}t && (i, j = 1, \ldots, d_1), \\ I_2^{ij} &:= 2\int_0^1 \frac{\frac{\partial b}{\partial \sigma_i}\frac{\partial b}{\partial \sigma_j}(x_t, \sigma_0)}{|b(x_t, \sigma_0)|^2}\mathrm{d}t && (i, j = 1, \ldots, d_2), \\ I_3^{ij} &:= \int_0^1 \int \frac{\partial \psi}{\partial \alpha_i}\frac{\partial \psi}{\partial \alpha_j}\big(x_t, c(x_t, \alpha_0)z, \alpha_0\big) f_{\alpha_0}(z)\mathrm{d}z\mathrm{d}t && (i, j = 1, \ldots, d_4). \end{aligned} \tag{3.3}$$

Remark 3.1 If $\{f_\alpha\}_{\alpha\in\Theta_3}$ is given by the class of the densities of normal distributions as in Example 5.1, then the range of ρ in Theorems 3.1 and 3.2 is same as in Shimizu and Yoshida (2006) and Ogihara and Yoshida (2011). However, if $\{f_\alpha\}_{\alpha\in\Theta_3}$ is given by the class of the densities of gamma distributions as in Example 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 (2011), where b is the constant defined in the equation (1) in Ogihara and Yoshida (2011).

4 Proofs

4.1 Inequalities

Lemma 4.1 *Under Assumptions 2.1 and 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* $p \ge 2$,

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

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* $k = 1, \ldots, n$ *that*

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

$$E\left[\sup_{t\in[0,1]} \left|X_t^\varepsilon - x_0\right|^p \,\middle|\, \mathcal{F}_{t_0}\right] \le C\left\{1 + \varepsilon^p\lambda_\varepsilon^p\right\}\left(1 + |x_0|^p\right),$$

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

Proof For any $p \geq 2$, we have

$$\begin{aligned}
&\left(E\left[\sup_{u\in[s,t]}\left|X_u^\varepsilon - X_s^\varepsilon\right|^p \,\middle|\, \mathcal{F}_s\right]\right)^{1/p} \\
&\leq \left(E\left[\left|\int_s^t \left|a(X_u^\varepsilon,\mu_0) - a(X_s^\varepsilon,\mu_0)\right| \mathrm{d}u\right|^p \,\middle|\, \mathcal{F}_s\right]\right)^{1/p} \\
&\quad + \varepsilon\left(E\left[\sup_{u\in[s,t]}\left|\int_s^u \left\{b(X_v^\varepsilon,\sigma_0) - b(X_s^\varepsilon,\sigma_0)\right\} \mathrm{d}W_v\right|^p \,\middle|\, \mathcal{F}_s\right]\right)^{1/p} \\
&\quad + \varepsilon\left(E\left[\sup_{u\in[s,t]}\left|\int_s^u \left\{c(X_v^\varepsilon,\alpha_0) - c(X_s^\varepsilon,\alpha_0)\right\} \mathrm{d}Z_v^{\lambda_\varepsilon}\right|^p \,\middle|\, \mathcal{F}_s\right]\right)^{1/p} \\
&\quad + (t-s)\left|a(X_s^\varepsilon,\mu_0)\right| + C\varepsilon\sqrt{t-s}\left|b(X_s^\varepsilon,\sigma_0)\right| \\
&\quad + \varepsilon\left|c(X_s^\varepsilon,\alpha_0)\right|\left(E\left[\sup_{u\in[s,t]}\left|\int_s^u \mathrm{d}Z_v^{\lambda_\varepsilon}\right|^p\right]\right)^{1/p},
\end{aligned} \tag{4.1}$$

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

$$\begin{aligned}
E\left[\left(\int_s^t \left|a(X_u^\varepsilon,\mu_0) - a(X_s^\varepsilon,\mu_0)\right| \mathrm{d}u\right)^p \,\middle|\, \mathcal{F}_s\right] &\leq CE\left[\left(\int_s^t |X_u^\varepsilon - X_s^\varepsilon| \mathrm{d}u\right)^p \,\middle|\, \mathcal{F}_s\right] \\
&\leq C(t-s)^{p-1}\int_s^t E\left[\left|X_u^\varepsilon - X_s^\varepsilon\right|^p \,\middle|\, \mathcal{F}_s\right] \mathrm{d}u,
\end{aligned} \tag{4.2}$$

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 (2009)) that

$$\begin{aligned}
&E\left[\sup_{u\in[s,t]}\left|\int_s^u \left\{b(X_v^\varepsilon,\sigma_0) - b(X_s^\varepsilon,\sigma_0)\right\} \mathrm{d}W_v\right|^p \,\middle|\, \mathcal{F}_s\right] \\
&\leq CE\left[\left|\int_s^t \left|X_u^\varepsilon - X_s^\varepsilon\right|^2 \mathrm{d}u\right|^{p/2} \,\middle|\, \mathcal{F}_s\right] \\
&\leq C(t-s)^{p/2-1}\int_s^t E\left[\left|X_u^\varepsilon - X_s^\varepsilon\right|^p \,\middle|\, \mathcal{F}_s\right] \mathrm{d}u,
\end{aligned} \tag{4.3}$$

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 (2009) that

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

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

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

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

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

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

$$\begin{aligned}
&E\left[\sup_{u\in[s,t]}\left|\int_s^u\left(c(X_v^\varepsilon,\alpha_0) - c(X_s^\varepsilon,\alpha_0)\right)\mathrm{d}Z_v^{\lambda_\varepsilon}\right|^p\,\middle|\,\mathcal{F}_s\right]\\
&\leq C\left(\lambda_\varepsilon^{p/2}(t-s)^{p/2-1} + \lambda_\varepsilon + \lambda_\varepsilon^p(t-s)^{p-1}\right)\int_s^t E\left[\left|X_u^\varepsilon - X_s^\varepsilon\right|^p\,\middle|\,\mathcal{F}_s\right]\mathrm{d}u, \quad (4.4)
\end{aligned}$$

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

$$E\left[\sup_{u\in[s,t]}\left|\int_s^u \mathrm{d}Z_v^{\lambda_\varepsilon}\right|^p\right] \leq C\left(\lambda_\varepsilon^{p/2}(t-s)^{p/2} + \lambda_\varepsilon(t-s) + \lambda_\varepsilon^p(t-s)^p\right), \quad (4.5)$$

where C depends only on p and f_{α_0}. From (4.1), (4.2), (4.3), (4.4) and (4.5),

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

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

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

This implies the conclusion. □

Lemma 4.2 *Under Assumptions 2.1 and 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* $p \geq 2$

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

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

Proof Same as the proof of Lemma 4.1, for any $p \geq 2$, we obtain

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

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

Lemma 4.3 *Under Assumptions 2.1 and 2.3, for* $p \geq 1$

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

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

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

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 4.2. □

Lemma 4.4 *Under Assumptions 2.1 and 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,*

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

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 4.3 and A.2. □

Lemma 4.5 *Under Assumptions 2.1 and 2.3 with Notation 2.5, suppose that* $0 < \varepsilon \le 1$, $\lambda_\varepsilon \ge 1$ *and* $\varepsilon\lambda_\varepsilon \le 1$. *Then, for any* $p \in [1, \infty)$,

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

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

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

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

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

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

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

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

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

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

where C depnds only on p, a and b. From Lemma 4.1, we have

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

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 4.6 *Under Assumptions 2.1 and 2.3, suppose that* $0 < \varepsilon \le 1$, $\lambda_\varepsilon \ge 1$, $\varepsilon\lambda_\varepsilon \le 1$. *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 $p \in (2, \infty)$,

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

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

Proof By using Lemmas 4.4 and 4.5, we have

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

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$ and $\varepsilon\lambda_\varepsilon \to 0$. It follows from Lemma A.3 that

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

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

4.2 Limit theorems

We make a version of Lemma 2.2 in Shimizu (2017) in the sequel.

Lemma 4.7 *Under Assumptions 2.1, 2.3, 2.6 and 2.8 with Notations 2.3 to 2.5 and 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 2.4, because the same argument still works under the case (ii) in Assumption 2.4-. Same as in the proof of Lemma 2.2 in Shimizu and Yoshida (2006), 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

$$
\begin{aligned}
&P\left[\left|X^\varepsilon_{t_k}-X^\varepsilon_{\tau_k}+\Delta X^\varepsilon_{\tau_k}+X^\varepsilon_{\tau_k-}-X^\varepsilon_{t_{k-1}}\right|\le\frac{v_{nk}}{n^\rho}\,\middle|\,\mathcal{F}_{t_{k-1}},\ \Delta^n_k N^{\lambda_\varepsilon}=1\right]\\
&\quad\le P\left[\left|X^\varepsilon_{t_k}-X^\varepsilon_{\tau_k}\right|+\sup_{t\in[t_{k-1},\tau_k)}\left|X^\varepsilon_t-X^\varepsilon_{t_{k-1}}\right|>\frac{v_{nk}}{n^\rho}\,\middle|\,\mathcal{F}_{t_{k-1}},\ \Delta^n_k N^{\lambda_\varepsilon}=1\right]\\
&\qquad+P\left[\left|\Delta Z^{\lambda_\varepsilon}_{\tau_k}\right|\le\frac{4v_{nk}}{c_1n^\rho}\text{ or }\sup_{t\in[t_k,t_{k-1}]}\left|c(X^\varepsilon_t,\alpha_0)-c(x_t,\alpha_0)\right|>\frac{c_1}{2}\,\middle|\,\mathcal{F}_{t_{k-1}},\ \Delta^n_k N^{\lambda_\varepsilon}=1\right],\\
&P\left[\left|X^\varepsilon_{t_k}-X^\varepsilon_{\tau_k}+\Delta X^\varepsilon_{\tau_k}+X^\varepsilon_{\tau_k-}-X^\varepsilon_{t_{k-1}}\right|>\frac{v_{nk}}{n^\rho}\,\middle|\,\mathcal{F}_{t_{k-1}},\ \Delta^n_k N^{\lambda_\varepsilon}=1\right]\\
&\quad\le P\left[\left|X^\varepsilon_{t_k}-X^\varepsilon_{\tau_k}\right|+\sup_{t\in[t_{k-1},\tau_k)}\left|X^\varepsilon_t-X^\varepsilon_{t_{k-1}}\right|>\frac{v_{nk}}{2n^\rho}\,\middle|\,\mathcal{F}_{t_{k-1}},\ \Delta^n_k N^{\lambda_\varepsilon}=1\right]\\
&\qquad+P\left[\left|\Delta Z^{\lambda_\varepsilon}_{\tau_k}\right|>\frac{v_{nk}}{4c_2n^\rho}\text{ or }\sup_{t\in[t_k,t_{k-1}]}\left|c(X^\varepsilon_t,\alpha_0)-c(x_t,\alpha_0)\right|>c_2\,\middle|\,\mathcal{F}_{t_{k-1}},\ \Delta^n_k N^{\lambda_\varepsilon}=1\right]
\end{aligned}
$$

and Lemmas 4.2, 4.5 and A.2 that

$$
\begin{aligned}
P\left[C^{n,\varepsilon,\rho}_{k,1}\,\middle|\,\mathcal{F}_{t_{k-1}}\right]&\le\frac{\lambda_\varepsilon}{n}e^{-\lambda_\varepsilon/n}\left\{C\left(\frac{1}{n^{p(1-\rho)}}+\frac{\varepsilon^p}{n^{p(1/2-\rho)}}\right)\left(1+|X^\varepsilon_{t_{k-1}}|^p\right)\right.\\
&\quad\left.+\int_{|z|\le4v_{nk}/c_1n^\rho}f_{\alpha_0}(z)\mathrm{d}z+C\frac{\varepsilon^p\lambda_\varepsilon}{n}\right\},\\
P\left[D^{n,\varepsilon,\rho}_{k,1}\,\middle|\,\mathcal{F}_{t_{k-1}}\right]&\le\frac{\lambda_\varepsilon}{n}e^{-\lambda_\varepsilon/n}\left\{C\left(\frac{1}{n^{p(1-\rho)}}+\frac{\varepsilon^p}{n^{p(1/2-\rho)}}\right)\left(1+|X^\varepsilon_{t_{k-1}}|^p\right)\right.\\
&\quad\left.+\int_{|z|>v_{nk}/4c_2n^\rho}f_{\alpha_0}(z)\mathrm{d}z+C\frac{\varepsilon^p\lambda_\varepsilon}{n}\right\},
\end{aligned}
$$

where C depends only on p,a,b,c,f_{α_0} and v_1. The other inequalities follow from Lemma 4.5. □

In the proof of Proposition 3.3 (ii) in Shimizu (2017), 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 4.8 *Under Assumptions 2.1, 2.3, 2.4, 2.6 and 2.8, for* $\rho\in(0,1/2)$

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

as $n\to\infty$, $\varepsilon\to0$, $\lambda_\varepsilon\to\infty$, $\lambda_\varepsilon/n\to0$ *and* $\varepsilon\lambda_\varepsilon\to0$. *More precisely, for* $\rho\in(0,1/2)$ *and* $p\in[2/(1-2\rho),\infty)$

$$
\begin{aligned}
&\frac{1}{\lambda_\varepsilon}\sum_{k=1}^n 1_{D^{n,\varepsilon,\rho}_{k,0}}=O_p\left(\frac{1}{\lambda_\varepsilon n^{p(1-\rho)-1}}+\frac{\varepsilon^p}{\lambda_\varepsilon n^{p(1/2-\rho)-1}}\right),\\
&\frac{1}{\lambda_\varepsilon}\sum_{k=1}^n 1_{D^{n,\varepsilon,\rho}_{k,1}}=1+O_p\left(\frac{\lambda_\varepsilon}{n}+\frac{1}{n^{p(1-\rho)}}+\frac{\varepsilon^p}{n^{p(1/2-\rho)}}+\int_{|z|\le4v_{nk}/c_1n^\rho}f_{\alpha_0}(z)\mathrm{d}z\right),\\
&\frac{1}{\lambda_\varepsilon}\sum_{k=1}^n 1_{D^{n,\varepsilon,\rho}_{k,2}}=O_p\left(\frac{\lambda_\varepsilon}{n}\right)
\end{aligned}
$$

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

Proof Since

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

it follows from Lemmas 4.4 and 4.7 that for $p \ge 2$ and $\rho \in (0, 1/2)$

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

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

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

Hence, the conclusion follows from Lemma A.3. □

Lemma 4.9 *Under Assumptions 2.1, 2.3, 2.4, 2.6 and 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* $p \in [2, \infty)$

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

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 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 4.4 and 4.7 that for any $p \in [0, \infty)$

$$\begin{aligned}
\sum_{k=1}^{n} E\left[\left|\frac{1}{n}1_{C_{k,1}^{n,\varepsilon,\rho}}\right| \Big| \mathcal{F}_{t_{k-1}}\right] &= \frac{1}{n}\sum_{k=1}^{n} P\left[C_{k,1}^{n,\varepsilon,\rho} \Big| \mathcal{F}_{t_{k-1}}\right] \\
&\leq C\frac{\lambda_\varepsilon}{n}\left(\frac{1}{n^{p(1-\rho)}} + \frac{\varepsilon^p}{n^{p(1/2-\rho)}} + \frac{\varepsilon^p\lambda_\varepsilon}{n}\right) \\
&\quad \times \frac{1}{n}\sum_{k=1}^{n}\left(1 + |X_{t_{k-1}}^{\varepsilon}|^p\right) + \frac{\lambda_\varepsilon}{n}\int_{|z|\leq 4v_{nk}/c_1 n^\rho} f_{\alpha_0}(z)\mathrm{d}z, \\
\sum_{k=1}^{n} E\left[\left|\frac{1}{n}1_{C_{k,2}^{n,\varepsilon,\rho}}\right| \Big| \mathcal{F}_{t_{k-1}}\right] &= \frac{1}{n}\sum_{k=1}^{n} P\left[C_{k,2}^{n,\varepsilon,\rho} \Big| \mathcal{F}_{t_{k-1}}\right] \leq \frac{\lambda_\varepsilon^2}{n^2}.
\end{aligned}$$

The conclusion follows from Lemma A.3. □

Remark 4.1 From this lemma, under Assumptions 2.1, 2.3, 2.4, 2.6 and 2.8, for $\rho \in (0, 1/2)$ and for any random variables $\xi_{k,\theta}^{n,\varepsilon}$ $(k = 1, \ldots, n,\ n \in \mathbb{N},\ \varepsilon > 0,\ \theta \in \bar{\Theta})$, when

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

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 $\eta > 0$

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

Similarly, from Lemma 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 4.10 *Under Assumptions 2.1, 2.3, 2.4, 2.6 and 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 $\xi_{k,\theta}^{n,\varepsilon}$ $(k = 1, \ldots, n,\ n \in \mathbb{N},\ \varepsilon > 0,\ \theta \in \bar{\Theta})$ be random variables. If

$$\lambda_\varepsilon \to \infty, \quad \varepsilon\lambda_\varepsilon \to 0, \quad \frac{\lambda_\varepsilon^2}{n} \to 0, \quad \lambda_\varepsilon \int_{|z|\leq 4v_2/c_1 n^\rho} f_{\alpha_0}(z)\mathrm{d}z \to 0 \tag{4.6}$$

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 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 4.1 the conclusion. □

Remark 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 (4.6) 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 paper.

Lemma 4.11 *Under Assumptions 2.1, 2.3, 2.4, 2.6 and 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* $p \in [2, \infty)$

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

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 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 $\eta > 0$

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

It follows from Lemmas 4.3, 4.4 and 4.9 that

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

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

Lemma 4.12 *Under Assumptions 2.1, 2.3, 2.4, 2.6 and 2.8, let* $\rho \in (0, 1/2)$. *We assume either of the following conditions (i) or (ii):*

(i) Under Assumption 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 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

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

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

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

(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 $p \geq 2/(1 - 2q\rho)$,

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

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 4.3 Assumption 2.4 is used only for defining $D_k^{n,\varepsilon,\rho}$ in Lemmas 4.7, 4.8, 4.10 and 4.11, while it is essentially used in Lemma 4.12.

Remark 4.4 The assumptions (i.c) and (ii.e) in Lemma 4.12 are ensured if

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

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 (2003)).

Proof of Lemma 4.12 Let $\delta > 0$ be a sufficiently small number satisfying the conditions of the statement and

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

where c_1 and c_2 are the constants from Assumption 2.6. In this proof, we may simply write the maps

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

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

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

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 4.10 that for any non-random $r'_{n,\varepsilon} > 0$ ($n \in \mathbb{N}, \varepsilon > 0$),

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

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

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

$$\frac{1}{\lambda_\varepsilon}\sum_{k=1}^n \left\{ g_k\left(\frac{\Delta^n_k X^\varepsilon}{\varepsilon}, \theta\right) - g_k\left(\frac{\Delta X^\varepsilon_{\tau_k}}{\varepsilon}, \theta\right)\right\} 1_{\tilde{D}^{n,\varepsilon,\rho}_{k,1}} = O_p(r_{n,\varepsilon}), \tag{4.7}$$

$$\frac{1}{\lambda_\varepsilon}\sum_{k=1}^n \left\{ g_k\left(\frac{\Delta X^\varepsilon_{\tau_k}}{\varepsilon}, \theta\right) - g_k\left(c(X^\varepsilon_{t_{k-1}}, \alpha_0)V_{N^{\lambda_\varepsilon}_{\tau_k}}, \theta\right)\right\} 1_{\tilde{D}^{n,\varepsilon,\rho}_{k,1}} = O_p\left(\frac{1}{\varepsilon n^{1-1/p}} + \frac{1}{n^{1/2-1/p}}\right) \tag{4.8}$$

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

Put

$$Y^\varepsilon_k := \frac{X^\varepsilon_{t_k} - X^\varepsilon_{\eta_k}}{\varepsilon} + \frac{X^\varepsilon_{\tau_k -} - X^\varepsilon_{t_{k-1}}}{\varepsilon} \left(= \frac{\Delta^n_k X^\varepsilon}{\varepsilon} - \frac{\Delta X^\varepsilon_{\tau_k}}{\varepsilon} \quad \text{on } D^{n,\varepsilon,\rho}_{k,1} \right).$$

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

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

Here, we remark that $\Delta_k^n X^\varepsilon$ and $\Delta X^\varepsilon_{\tau_k}$ are almost surely positive on $\tilde{D}^{n,\varepsilon,\rho}_{k,1}$ under Assumption 2.4 (ii). To see (4.7), it is sufficient to show that

$$\sup_{\theta\in\Theta}\left|\frac{1}{\lambda_\varepsilon}\sum_{k=1}^n\int_0^1\frac{\partial g_k}{\partial y}\left(\frac{\Delta X^\varepsilon_{\tau_k}}{\varepsilon}+\zeta Y_k^\varepsilon,\theta\right)Y_k^\varepsilon \mathrm{d}\zeta\, 1_{\tilde{D}^{n,\varepsilon,\rho}_{k,1}\cap\{|Y_k^\varepsilon|\le 1\}}\right| = O_p(r_{n,\varepsilon}) \tag{4.9}$$

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

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

and from Lemma 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 2.4. Since for $\zeta\in[0,1]$ we have

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

we obtain from the assumption (ii.b) that

$$\begin{aligned}
&\int_0^1\left|\frac{\partial g_k}{\partial y}\left(\frac{\Delta X^\varepsilon_{\tau_k}}{\varepsilon}+\zeta Y_k^\varepsilon,\theta\right)\right|\mathrm{d}\zeta\, 1_{\tilde{D}^{n,\varepsilon,\rho}_{k,1}\cap\{|Y_k^\varepsilon|\le 1\}}\\
&\le\left\{\left|\frac{\partial g_k}{\partial y}\left(\frac{c_1}{2}V_{N^{\lambda_\varepsilon}_{\tau_k}},\theta\right)\right|+\left|\frac{\partial g_k}{\partial y}\left(\frac{v_1}{n^\rho},\theta\right)\right|+\left|\frac{\partial g_k}{\partial y}\left(2c_2V_{N^{\lambda_\varepsilon}_{\tau_k}}+1,\theta\right)\right|+L\right\}1_{\tilde{D}^{n,\varepsilon,\rho}_{k,1}\cap\{|Y_k^\varepsilon|\le 1\}}\\
&\le\left\{\left|\frac{\partial g_k}{\partial y}\left(\frac{c_1}{2}V_{N^{\lambda_\varepsilon}_{\tau_k}},\theta\right)\right|+\left|\frac{\partial g_k}{\partial y}\left(\frac{v_1}{n^\rho},\theta\right)\right|+\left|\frac{\partial g_k}{\partial y}\left(2c_2V_{N^{\lambda_\varepsilon}_{\tau_k}}+1,\theta\right)\right|+L\right\}1_{J^{n,\varepsilon}_{k,1}}
\end{aligned}$$

Since

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

it follows from Lemma A.3 (ii), Lemmas 4.4 and 4.6 and the assumption (ii.d) that

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

as $n\to\infty$, $\varepsilon\to 0$, $\lambda_\varepsilon\to\infty$ and $\varepsilon\lambda_\varepsilon\to 0$, uniformly in $\theta\in\Theta$, where p is given in the assumption (ii.e). Similarly, it follows from Lemma A.3 (ii), Lemmas 4.4 and 4.6 and the

assumption (ii.d) that

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

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 A.3 (ii), Lemma 4.6 and the assumption (ii.c) that

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

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 (4.9).

Under the case (i) in Assumption 2.4, as in the same argument above, we have

$$\begin{aligned}&\frac{1}{\lambda_\varepsilon}\sum_{k=1}^{n}\sup_{\theta\in\bar\Theta}\left|\int_0^1\frac{\partial g_k}{\partial y}\left(\frac{\Delta X_{\tau_k}^\varepsilon}{\varepsilon}+\zeta Y_k^\varepsilon,\theta\right)Y_k^\varepsilon\mathrm{d}\zeta\right|1_{\tilde D_{k,1}^{n,\varepsilon,\rho}\cap\{|Y_k^\varepsilon|\le 1\}}\\&\quad\le\frac{C}{\lambda_\varepsilon}\sum_{k=1}^{n}\left(2+\left|2c_2V_{N_{\tau_k}^{\lambda_\varepsilon}}\right|^p\right)1_{J_{k,1}^{n,\varepsilon}}\sup_{k=1,\dots,n}\left|Y_k^\varepsilon\right|=O_p\left(\frac{1}{\varepsilon n^{1-1/p}}+\frac{1}{n^{1/2-1/p}}\right)\end{aligned}$$

as $n\to\infty$, $\varepsilon\to 0$, $\lambda_\varepsilon\to\infty$ and $\varepsilon\lambda_\varepsilon\to 0$. Thus, we obtain (4.9).

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

$$(1-\zeta)\frac{\Delta X_{\tau_k}^\varepsilon}{\varepsilon}+\zeta c(X_{t_{k-1}}^\varepsilon,\alpha_0)V_{N_{\tau_k}^{\lambda_\varepsilon}}\ge\frac{c_1}{2}V_{N_{\tau_k}^{\lambda_\varepsilon}}\quad\text{on }\tilde D_{k,1}^{n,\varepsilon,\rho},$$

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

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

so that, (4.8) holds. □

Lemma 4.13 *Let $\rho\in(0,1/2)$. Under Assumptions 2.1, 2.3, 2.4, 2.6 and 2.8, suppose that for $\theta\in\Theta$*

$$x\mapsto\int g(x,c(x,\alpha_0)z,\theta)f_{\alpha_0}(z)\mathrm{d}z,\quad x\mapsto\int|g(x,c(x,\alpha_0)z,\theta)|^2f_{\alpha_0}(z)\mathrm{d}z\tag{4.10}$$

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

$$\begin{aligned}\int\Bigg\{&\sup_{(x,\theta)\in I_{x_0}^\delta\times\Theta}|g(x,c(x,\alpha_0)z,\theta)|\\&+\sum_{j=1}^{d}\sup_{(x,\theta)\in I_{x_0}^\delta\times\Theta}\left|\frac{\partial g}{\partial\theta_j}(x,c(x,\alpha_0)z,\theta)\right|\Bigg\}f_{\alpha_0}(z)\mathrm{d}z<\infty.\end{aligned}\tag{4.11}$$

Then,

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

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 4.4 and the assumption (4.10) that for each $\theta \in \Theta$

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

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

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

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$ and $\varepsilon\lambda_\varepsilon \to 0$. Thus, Lemma 9 in Genon-Catalot and Jacod (1993) shows us that for each $\theta \in \Theta$

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

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

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

Then, by the same argument in the proof of Lemma 4.10, it follows from Lemma 4.3 that

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

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 $\theta \in \Theta$

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

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

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

and we shall use Theorem 5.1 in Billingsley (1999) with the state space $C(\Theta)$, same as in the proofs of Propositions 3.3 and 3.6 in Shimizu and Yoshida (2006) [1]. From the assumption

[1] We cannot use Theorem 20 in Ibragimov and Has'minskii (1981), Appendix I (or Lemma 3.1 in Yoshida (1990)), as in the proof of Lemma 2 in Sørensen and Uchida (2003). In fact, we fail to say that $\{\chi^{n,\varepsilon}\}$ satisfies (1) and (2) in Lemma 3.1 in Yoshida (1990).

(4.11), we obtain

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

and

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

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

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

where C is defined as (4.11) and for $q\geq 1$

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

From Rellich-Kondrachov's theorem (see, *e.g.*, Theorem 9.16 in Brezis (2011)), it follows that the balls B_M, $M>0$ are compact in $C(\Theta)$, and so from Theorem 5.1 in Billingsley (1999) that $\{\chi^{n,\varepsilon}\}$ 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 4.14 *Under Assumptions 2.1, 2.3, 2.4, 2.6 and 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 $I^\delta_{x_0}$ for some small $\delta>0$. Then,*

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

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

Proof At first, we can easily check that

$$\frac{1}{\varepsilon}\sum_{k=1}^{n} g(X^{\varepsilon}_{t_{k-1}},\theta)\left\{\int_{t_{k-1}}^{t_k} a(X^{\varepsilon}_t,\mu_0)\mathrm{d}t - \frac{1}{n}a(X^{\varepsilon}_{t_{k-1}},\mu_0)\right\}1_{C_k^{n,\varepsilon,\rho}} \xrightarrow{p} 0 \tag{4.12}$$

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 4.3, A.2 and A.3 with the equicontinuity of g on I_{x_0} and the following estimate:

$$\begin{aligned}
&\frac{1}{\varepsilon}\sum_{k=1}^{n} E\left[\sup_{\theta\in\Theta}\left|g(X^{\varepsilon}_{t_{k-1}},\theta)\left\{\int_{t_{k-1}}^{t_k} a(X^{\varepsilon}_t,\mu_0)\mathrm{d}t - \frac{1}{n}a(X^{\varepsilon}_{t_{k-1}},\mu_0)\right\}1_{C_k^{n,\varepsilon,\rho}}\right| \,\middle|\, \mathcal{F}_{t_{k-1}}\right] \\
&\quad\le C\left(\frac{1}{n}\sum_{k=1}^{n} E\left[\sup_{\theta\in\Theta}\left|g(X^{\varepsilon}_{t_{k-1}},\theta)\sup_{t\in[t_{k-1},t_k]}\frac{|X^{\varepsilon}_t - X^{\varepsilon}_{t_{k-1}}|}{\varepsilon}\right|^2 \,\middle|\, \mathcal{F}_{t_{k-1}}\right]\right)^{1/2} \\
&\qquad\qquad (\because \text{Schwartz's inequality and 2.1}) \\
&\quad= O_p\left(\frac{1}{\varepsilon n}+\frac{1}{\sqrt{n}}+\frac{\lambda_\varepsilon}{n}\right) \qquad (\because \text{Lemma 4.1 and 4.4})
\end{aligned}$$

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

At second, we show that

$$\begin{aligned}
&\sum_{k=1}^{n} g(X^{\varepsilon}_{t_{k-1}},\theta)\int_{t_{k-1}}^{t_k} b(X^{\varepsilon}_t,\sigma_0)\mathrm{d}W_t 1_{C_k^{n,\varepsilon,\rho}} - \int_0^1 g(X^{\varepsilon}_t,\theta)b(X^{\varepsilon}_t,\sigma_0)\mathrm{d}W_t \\
&\quad= \sum_{k=1}^{n}\int_{t_{k-1}}^{t_k}\left\{g(X^{\varepsilon}_{t_{k-1}},\theta) - g(X^{\varepsilon}_t,\theta)\right\}b(X^{\varepsilon}_t,\sigma_0)\mathrm{d}W_t 1_{C_k^{n,\varepsilon,\rho}} \\
&\qquad - \sum_{k=1}^{n}\int_{t_{k-1}}^{t_k} g(X^{\varepsilon}_t,\theta)b(X^{\varepsilon}_t,\sigma_0)\mathrm{d}W_t 1_{D_k^{n,\varepsilon,\rho}} \xrightarrow{p} 0
\end{aligned} \tag{4.13}$$

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

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

it holds from Morrey's inequality (see, *e.g.*, Theorem 5 in Evans (2010), Section 5.6) that for $q\in(d,\infty)$

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

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

$$\begin{aligned}
&\sum_{k=1}^{n} E\left[\left\|\int_{t_{k-1}}^{t_k}\left\{g(X^{\varepsilon}_{t_{k-1}},\theta)-g(X^{\varepsilon}_t,\theta)\right\}b(X^{\varepsilon}_t,\sigma_0)\mathrm{d}W_t 1_{\tilde{C}^{n,\varepsilon,\rho}_k}\right\|_{L^q(\Theta)}\,\middle|\,\mathcal{F}_{t_{k-1}}\right] \\
&\quad\le C_2\sum_{k=1}^{n}\left(\int_{\Theta}E\left[\int_{t_{k-1}}^{t_k}\left|\left\{g(X^{\varepsilon}_{t_{k-1}},\theta)-g(X^{\varepsilon}_t,\theta)\right\}1_{\tilde{C}^{n,\varepsilon,\rho}_k}b(X^{\varepsilon}_t,\sigma_0)\right|^2\mathrm{d}t\,\middle|\,\mathcal{F}_{t_{k-1}}\right]^{q/2}\mathrm{d}\theta\right)^{1/q} \\
&\qquad\qquad(\because \text{Hölder's and Burkholder's inequalities}) \\
&\quad\le\frac{C_3}{\sqrt{n}}\sum_{k=1}^{n}\left(E\left[\sup_{t\in[t_{k-1},t_k]}|X^{\varepsilon}_{t_{k-1}}-X^{\varepsilon}_t|^2\right.\right. \\
&\qquad\left.\left.\times\left(1+\sup_{t\in[t_{k-1},t_k]}|X^{\varepsilon}_{t_{k-1}}-X^{\varepsilon}_t|^2+|X^{\varepsilon}_{t_{k-1}}|^2\right)\,\middle|\,\mathcal{F}_{t_{k-1}}\right]^{q/2}\right)^{1/q} \\
&\qquad\qquad(\because \text{Hölder's inequality and the equi-Lipschitz continuity of } g \text{ and } b \text{ on } I^{\delta}_{x_0}) \\
&\quad= O_p\left(\frac{1}{\sqrt{n}}+\varepsilon+\varepsilon\sqrt{\lambda_\varepsilon}\right) \qquad (\because \text{Lemmas 4.1and 4.4})
\end{aligned}$$

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 Bhagavatula (1999), it follows that

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

as $n\to\infty$, $\varepsilon\to 0$, $\lambda_\varepsilon\to\infty$ and $\varepsilon\lambda_\varepsilon\to 0$. Thus, it follows from Lemma A.3 that

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

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 4.3 we obtain the convergence of the first term in the left-hand side of (4.13). To obtain (4.13), we remain to prove

$$\sum_{k=1}^{n}\int_{t_{k-1}}^{t_k} g(X^{\varepsilon}_t,\theta)b(X^{\varepsilon}_t,\sigma_0)\mathrm{d}W_t 1_{D^{n,\varepsilon,\rho}_k}\xrightarrow{p}0 \tag{4.14}$$

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 $\tilde{D}^{n,\varepsilon,\rho}_{k,1}:=D^{n,\varepsilon,\rho}_{k,1}\cap\{X^{\varepsilon}_t\in I^{\delta}_{x_0}\text{ for all } t\in[0,1]\}$. We begin with showing that for any $p\in(2,\infty)$ and $q'\in(1,d/(d-1))$

$$\sum_{k=1}^{n}\int_{t_{k-1}}^{t_k} g(X^{\varepsilon}_t,\theta)b(X^{\varepsilon}_t,\sigma_0)\mathrm{d}W_t 1_{\tilde{D}^{n,\varepsilon,\rho}_{k,1}}=O_p\left(\frac{1}{\sqrt{n}}\left(\frac{\varepsilon^p\lambda_\varepsilon^2}{n}+\lambda_\varepsilon\right)^{1/2+1/q'}\right) \tag{4.15}$$

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 (2010), Section 5.6) that for $q \in (d, \infty)$

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

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 (2009)), the equicontinuity of g and Assumption 2.1 that

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

where C_2 depends only on q. By using Lemmas 4.4 and 4.7, for any $p > 2$ we obtain

$$\begin{aligned}
&\sum_{k=1}^n E\left[\left\|\int_{t_{k-1}}^{t_k} g(X_t^\varepsilon,\theta)b(X_t^\varepsilon,\sigma_0)\mathrm{d}W_t\right\|_{L^q(\Theta)} 1_{\tilde{D}_{k,1}^{n,\varepsilon,\rho}} \,\middle|\, \mathcal{F}_{t_{k-1}}\right] \\
&\quad = O_p\left(\sqrt{n}\left\{\frac{\lambda_\varepsilon}{n}\left(\frac{1}{n^{p(1-\rho)}} + \frac{\varepsilon^p}{n^{p(1/2-\rho)}} + \frac{\varepsilon^p\lambda_\varepsilon}{n}\right) + \frac{\lambda_\varepsilon}{n}\right\}^{1/2+1/q'}\right)
\end{aligned}$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$ and $\varepsilon\lambda_\varepsilon \to 0$. Similarly, by using Theorem B.4 in Bhagavatula (1999), we obtain for $j = 1, \ldots, d$

$$\begin{aligned}
&\sum_{k=1}^n E\left[\left\|\int_{t_{k-1}}^{t_k} \frac{\partial g}{\partial\theta_j}\left(X_t^\varepsilon,\theta\right) b(X_t^\varepsilon,\sigma_0)\mathrm{d}W_t\right\|_{L^q(\Theta)} 1_{\tilde{D}_{k,1}^{n,\varepsilon,\rho}} \,\middle|\, \mathcal{F}_{t_{k-1}}\right] \\
&\quad = O_p\left(\sqrt{n}\left\{\frac{\lambda_\varepsilon}{n}\left(\frac{1}{n^{p(1-\rho)}} + \frac{\varepsilon^p}{n^{p(1/2-\rho)}} + \frac{\varepsilon^p\lambda_\varepsilon}{n}\right) + \frac{\lambda_\varepsilon}{n}\right\}^{1/2+1/q'}\right)
\end{aligned}$$

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$ and $\varepsilon\lambda_\varepsilon \to 0$. Since we can take $q' < 2$ small enough, we obtain (4.15) from Remark A.3. Hence, (4.14) holds from (4.15) and Lemma 4.10.

At last, it is an immediate consequence from Lemma 4.9 that

$$\sum_{k=1}^{n} g(X^{\varepsilon}_{t_{k-1}}, \theta) \int_{t_{k-1}}^{t_k} c(X^{\varepsilon}_{t-}, \alpha_0) \mathrm{d}Z^{\lambda_\varepsilon}_t 1_{C^{n,\varepsilon,\rho}_k} \xrightarrow{p} 0$$

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 4.15 *Under Assumptions 2.1 2.3, 2.4, 2.6 and 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 $I^{\delta}_{x_0}$ for some small $\delta > 0$. Then,*

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

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$, $\varepsilon\lambda_\varepsilon \to 0$, $\lambda_\varepsilon^2/n \to 0$ 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 4.9, it is sufficient to show that

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

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

$$\begin{aligned}
&\left| \Delta^n_k X^{\varepsilon} - \frac{1}{n} a(X^{\varepsilon}_{t_{k-1}}, \mu_0) \right|^2 1_{J^{n,\varepsilon}_{k,0}} \\
&= \left\{ \left| \int_{t_{k-1}}^{t_k} \left\{ a(X^{\varepsilon}_t, \mu_0) - a(X^{\varepsilon}_{t_{k-1}}, \mu_0) \right\} \mathrm{d}t \right|^2 + \left| \int_{t_{k-1}}^{t_k} b(X^{\varepsilon}_t, \sigma_0) \mathrm{d}W_t \right|^2 \right. \\
&\quad \left. + 2 \int_{t_{k-1}}^{t_k} \left\{ a(X^{\varepsilon}_t, \mu_0) - a(X^{\varepsilon}_{t_{k-1}}, \mu_0) \right\} \mathrm{d}t \int_{t_{k-1}}^{t_k} b(X^{\varepsilon}_t, \sigma_0) \mathrm{d}W_t \right\} 1_{J^{n,\varepsilon}_{k,0}}.
\end{aligned}$$

Similarly to the proof of (4.12), it follows that

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

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

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

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 2.1, Hölder's inequality and Burkholder's inequality, we obtain

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

where C depends only on a, b. By applying Lemmas 4.3 to 4.5 and A.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 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 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 (1993), 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 $I^\delta_{x_0}$ for some $\delta>0$, the uniform tightness is shown by the same argument in the proof of Lemma 4.13. At second, we shall see

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

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 A.3 and the following estimate:

$$
\begin{aligned}
&\sum_{k=1}^{n} E\left[\sup_{\theta\in\Theta}\left|g(X^\varepsilon_{t_{k-1}},\theta)\left\{\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_{J^{n,\varepsilon}_{k,0}}\right|\,\middle|\,\mathcal{F}_{t_{k-1}}\right] \\
&\quad\le \sum_{k=1}^{n}\sup_{\theta\in\Theta}|g(X^\varepsilon_{t_{k-1}},\theta)|\left(E\left[\left|\int_{t_{k-1}}^{t_k}\{b(X^\varepsilon_t,\sigma_0)+b(X^\varepsilon_{t_{k-1}},\sigma_0)\}\mathrm{d}W_t\right|^2 1_{J^{n,\varepsilon}_{k,0}}\,\middle|\,\mathcal{F}_{t_{k-1}}\right]\right)^{1/2} \\
&\qquad\times\left(E\left[\left|\int_{t_{k-1}}^{t_k}\{b(X^\varepsilon_t,\sigma_0)-b(X^\varepsilon_{t_{k-1}},\sigma_0)\}\mathrm{d}W_t\right|^2 1_{J^{n,\varepsilon}_{k,0}}\,\middle|\,\mathcal{F}_{t_{k-1}}\right]\right)^{1/2} \quad (\because \text{Hölder's inequality}) \\
&\quad\le \sum_{k=1}^{n}\sup_{\theta\in\Theta}|g(X^\varepsilon_{t_{k-1}},\theta)|\left(E\left[\int_{t_{k-1}}^{t_k}|b(X^\varepsilon_t,\sigma_0)+b(X^\varepsilon_{t_{k-1}},\sigma_0)|^2 1_{J^{n,\varepsilon}_{k,0}}\mathrm{d}t\,\middle|\,\mathcal{F}_{t_{k-1}}\right]\right)^{1/2} \\
&\qquad\times\left(E\left[\int_{t_{k-1}}^{t_k}|b(X^\varepsilon_t,\sigma_0)-b(X^\varepsilon_{t_{k-1}},\sigma_0)|^2 1_{J^{n,\varepsilon}_{k,0}}\mathrm{d}t\,\middle|\,\mathcal{F}_{t_{k-1}}\right]\right)^{1/2} \quad (\because \text{Burkholder's inequality}) \\
&\quad\le C\frac{1}{n}\sum_{k=1}^{n}\sup_{\theta\in\Theta}|g(X^\varepsilon_{t_{k-1}},\theta)|\left(E\left[\sup_{t\in[t_{k-1},\tau_k]}(1+|X^\varepsilon_t-X^\varepsilon_{t_{k-1}}|^2+|X^\varepsilon_{t_{k-1}}|^2)\,\middle|\,\mathcal{F}_{t_{k-1}}\right]\right)^{1/2} \\
&\qquad\times\left(E\left[\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} \quad (\because b \text{ is Lipschitz}) \\
&\quad= O_p\left(\frac{1}{n}+\frac{\varepsilon}{\sqrt{n}}\right) \quad (\because \text{Lemma 4.4 and 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 4.1, 4.8 and 4.9 and the linearity of b that

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

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

4.3 Proof of main results

4.3.1 Proof of Theorem 3.1

Proof of Theorem 3.1 It follows from Lemmas 4.11 and 4.14 that

$$
\begin{aligned}
\Phi^{(1)}_{n,\varepsilon}(\mu,\sigma) &:= n\varepsilon^2\left(\Psi^{(1)}_{n,\varepsilon}(\mu,\sigma)-\Psi^{(1)}_{n,\varepsilon}(\mu_0,\sigma)\right) \\
&= \sum_{k=1}^{n}\frac{\left(\Delta^n_k X^\varepsilon - a(X^\varepsilon_{t_{k-1}},\mu_0)/n\right)\left(a(X^\varepsilon_{t_{k-1}},\mu)-a(X^\varepsilon_{t_{k-1}},\mu_0)\right)}{\left|b(X^\varepsilon_{t_{k-1}},\sigma)\right|^2}1_{C^{n,\varepsilon,\rho}_k}
\end{aligned}
$$

$$
\begin{aligned}
&- \frac{1}{2n} \sum_{k=1}^{n} \frac{\left| a(X^{\varepsilon}_{t_{k-1}}, \mu) - a(X^{\varepsilon}_{t_{k-1}}, \mu_0) \right|^2}{\left| b(X^{\varepsilon}_{t_{k-1}}, \sigma) \right|^2} 1_{C_k^{n,\varepsilon,\rho}} \\
&\xrightarrow{p} -\frac{1}{2} \int_0^1 \frac{|a(x_t, \mu) - a(x_t, \mu_0)|^2}{|b(x_t, \sigma)|^2} \mathrm{d}t
\end{aligned}
$$

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 4.11, 4.14 and 4.15 that

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

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 4.12 and 4.13 that

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

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 (2003)), the consistency of $\hat{\theta}_{n,\varepsilon}$ holds under Assumption 2.7. □

4.3.2 Proof of Theorem 3.2

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

$$\sqrt{\varepsilon^{-2}}\frac{\partial \Phi^{(1)}_{n,\varepsilon}}{\partial \mu_i}(\mu,\sigma)\Big|_{\theta=\theta_0} = -\frac{1}{\varepsilon}\sum_{k=1}^{n}\frac{\left\{\Delta^n_k X^\varepsilon - \frac{1}{n}a(X^\varepsilon_{t_{k-1}},\mu)\right\}\frac{\partial a}{\partial \mu_i}\left(X^\varepsilon_{t_{k-1}},\mu\right)}{\left|b(X^\varepsilon_{t_{k-1}},\sigma)\right|^2}1_{C^{n,\varepsilon,\rho}_k}\Big|_{\theta=\theta_0}$$
$$=: \sum_{k=1}^{n}\xi^i_{1,k}\left(\xrightarrow{p}\int_0^1\frac{\frac{\partial a}{\partial \mu_i}(x_t,\mu_0)}{b(x_t,\sigma_0)}\mathrm{d}W_t\right), \quad (\because \text{Lemma 4.14})$$
$$\sqrt{n}\frac{\partial \Psi^{(1)}_{n,\varepsilon}}{\partial \sigma_i}(\mu,\sigma)\Big|_{\theta=\theta_0} = -\frac{1}{\sqrt{n}}\sum_{k=1}^{n}\left\{-\frac{\left|\Delta^n_k X^\varepsilon - \frac{1}{n}a(X^\varepsilon_{t_{k-1}},\mu)\right|^2}{\frac{1}{n}\left|\varepsilon b(X^\varepsilon_{t_{k-1}},\sigma)\right|^2}+1\right\}$$
$$\times\frac{\frac{\partial b}{\partial \sigma_i}\left(X^\varepsilon_{t_{k-1}},\sigma\right)}{b(X^\varepsilon_{t_{k-1}},\sigma)}1_{C^{n,\varepsilon,\rho}_k}\Big|_{\theta=\theta_0}$$
$$=: \sum_{k=1}^{n}\xi^i_{2,k},$$
$$\sqrt{\lambda_\varepsilon}\frac{\partial \Psi^{(2)}_{n,\varepsilon}}{\partial \alpha_i}(\alpha)\Big|_{\alpha=\alpha_0} = \frac{1}{\sqrt{\lambda_\varepsilon}}\sum_{k=1}^{n}\frac{\partial \psi}{\partial \alpha_i}\left(X^\varepsilon_{t_{k-1}},\frac{\Delta^n_k X^\varepsilon}{\varepsilon},\alpha_0\right)1_{D^{n,\varepsilon,\rho}_k} =: \sum_{k=1}^{n}\xi^i_{3,k},$$

and

$$\sum_{k=1}^{n}\tilde{\xi}^i_{1,k} := \sum_{k=1}^{n}\frac{\frac{\partial a}{\partial \mu_i}\left(X^\varepsilon_{t_{k-1}},\mu_0\right)}{b(X^\varepsilon_{t_{k-1}},\sigma_0)}\int_{t_{k-1}}^{t_k}\mathrm{d}W_t 1_{C^{n,\varepsilon,\rho}_{k,0}},$$
$$\sum_{k=1}^{n}\tilde{\xi}^i_{2,k} := -\sqrt{n}\sum_{k=1}^{n}\left\{-\left|\int_{t_{k-1}}^{t_k}\mathrm{d}W_t\right|^2+\frac{1}{n}\right\}\frac{\frac{\partial b}{\partial \sigma_i}\left(X^\varepsilon_{t_{k-1}},\sigma_0\right)}{b(X^\varepsilon_{t_{k-1}},\sigma_0)}1_{C^{n,\varepsilon,\rho}_{k,0}},$$
$$\sum_{k=1}^{n}\tilde{\xi}^i_{3,k} := \sum_{k=1}^{n}\frac{1}{\sqrt{\lambda_\varepsilon}}\frac{\partial \psi}{\partial \alpha_i}\left(X^\varepsilon_{t_{k-1}},c(X^\varepsilon_{t_{k-1}},\alpha_0)V_{N^{\lambda_\varepsilon}_{\tau_k}},\alpha_0\right)1_{J^{n,\varepsilon}_{k,1}}.$$

Lemma 4.16 *Under Assumptions 2.1 to 2.6, 2.8 and 2.10, the following convergences are holds.*

For $\ell = 1, 2$

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

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$, $\varepsilon n \to \infty$, $\varepsilon\lambda_\varepsilon \to 0$, $\lambda^2_\varepsilon/n \to 0$ *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 2.4 (i), take $\rho \in (0, 1/2)$.

(ii) Under Assumption 2.4 (ii), take $\rho \in (0, \min\{1/2, 1/4q\})$, *where* q *is the constant given in Assumption 2.10 Assumption (ii.b).*

Then,

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

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 4.9 and A.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,

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

As in the same argument in Lemma 4.14, it holds from Assumptions 2.1 to 2.3 and Lemmas 4.4 and 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 2.1, Burkholder's inequality, Lemmas 4.4 and 4.5 that

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

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

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

and by the same argument as in the proof of Lemma 4.15, we obtain

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

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 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 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 4.10 4.12 and A.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 $\lambda_\varepsilon\int_{|z|\le 4v_2/c_1 n^\rho} f_{\alpha_0}(z)\mathrm{d}z\to 0$. □

Lemma 4.17 *Under Assumptions 2.1 to 2.3, 2.5, 2.6, 2.8 and 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 4.4 and 4.7 that for any $p \geq 1$

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

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

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

it follows from Lemmas 4.4 and 4.7 that for any $p \geq 1$

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

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

For $\ell = 3$, we may assume $\sup_t |X_t^{\varepsilon} - x_t| < \delta$ for some enough small $\delta > 0$. From Assumption 2.9, we obtain

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

The last equality holds from the fact that

$$\alpha \mapsto \int \psi\left(X_{t_{k-1}}^{\varepsilon}, c(X_{t_{k-1}}^{\varepsilon}, \alpha_0)z, \alpha\right) f_{\alpha_0}(z)\mathrm{d}z$$

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

Lemma 4.18 *Under Assumptions 2.1 to 2.6, 2.8, 2.9 and 2.11,*

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

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

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$ and $\varepsilon\lambda_\varepsilon \to 0$, where $I_1, \ldots, I_3$ are the matrices defined as (3.3).

Proof For $\ell = 1, i, j \in \{1, \ldots, d_1\}$, 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 4.4 and 4.7 for any $p > 1$ we have

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

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

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

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

For $\ell = 2, i, j \in \{1, \ldots, d_2\}$, 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 4.17, it follows from Lemmas 4.4 and 4.7 that for any $p > 1$

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

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$ and $\varepsilon\lambda_\varepsilon \to 0$, we obtain from Lemma 4.4 that

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

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 4.4 and Assumption 2.11 that

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

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 4.4 and 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 4.19 *Under Assumptions 2.1 to 2.3, 2.5, 2.6, 2.8 and 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 4.17. □

Lemma 4.20 *Under Assumptions 2.1 to 2.3, 2.5, 2.6, 2.8 and 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 $i \in \{1, \ldots, d_1\}$, and put $g(x) = |\frac{\partial a}{\partial \mu_i}(x, \mu_0)/b(x, \sigma_0)|^4$. Then, it holds from Lemma 4.4 that

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

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

For $\ell = 2$, let $i \in \{1, \ldots, d_2\}$, and put $g(x) = \left|\frac{1}{b}\frac{\partial b}{\partial \sigma_i}(x, \sigma_0)\right|^4$. Then, it follows from Lemma 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 $\ell = 3$, $i \in \{1, \ldots, d_3\}$, put $g(x, y) = \left|\frac{\partial \psi}{\partial \alpha_i}(x, y, \alpha_0)\right|^4$. Then, similarly to the proof of Lemma 4.18, it follows from Lemma 4.4 and Assumption 2.11 that

$$\sum_{k=1}^{n} E\left[\left|\tilde{\xi}_{3,k}^{i}\right|^4 \,\middle|\, \mathcal{F}_{t_{k-1}}\right] = \frac{1}{\lambda_\varepsilon^2} \sum_{k=1}^{n} E\left[g\left(X_{t_{k-1}}^{\varepsilon}, c(X_{t_{k-1}}^{\varepsilon}, \alpha_0) V_{N_{t_k}^{\lambda_\varepsilon}}\right) 1_{J_{k,1}^{n,\varepsilon}} \,\middle|\, \mathcal{F}_{t_{k-1}}\right] \xrightarrow{p} 0$$

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

Proof of Theorem 3.2 From Theorem A.3 in Shimizu (2007) and Lemmas 4.16 to 4.20,

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

as $n \to \infty$, $\varepsilon \to 0$, $\lambda_\varepsilon \to \infty$, $\varepsilon\lambda_\varepsilon \to 0$, $\lambda_\varepsilon^2/n \to 0$ and $\lambda_\varepsilon \int_{|z| \le 4v_2/c_1 n^\rho} f_{\alpha_0}(z)\mathrm{d}z \to 0$ with $\lim(\varepsilon^2 n)^{-1} < \infty$. Also, it follows from Lemmas 4.11 to 4.15 under Assumption 2.12 that

$$C_{\varepsilon,n}(\theta) := \begin{pmatrix} \varepsilon^2 n \left(\frac{\partial^2 \Psi_{n,\varepsilon}}{\partial \mu_i \partial \mu_j}(\theta)\right)_{i,j} & \varepsilon^2 n \left(\frac{\partial^2 \Psi_{n,\varepsilon}}{\partial \mu_i \partial \sigma_j}(\theta)\right)_{i,j} & 0 \\ \left(\frac{\partial^2 \Psi_{n,\varepsilon}}{\partial \sigma_i \partial \mu_j}(\theta)\right)_{i,j} & \left(\frac{\partial^2 \Psi_{n,\varepsilon}}{\partial \sigma_i \partial \sigma_j}(\theta)\right)_{i,j} & 0 \\ 0 & 0 & \left(\frac{\partial^2 \Psi_{n,\varepsilon}}{\partial \alpha_i \partial \alpha_j}(\theta)\right)_{i,j} \end{pmatrix} \xrightarrow{p} -I_{\theta_0} \quad (4.16)$$

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

$$\begin{aligned}
\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}} \\
&\quad - \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\} \\
&\quad \times \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}} \\
&\quad - \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}} \\
&\quad - \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},
\end{aligned}$$

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 (2003). □

5 Examples

This section is devoted to give some examples of densities which satisfy Assumptions 2.9 to 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 $y \in D_+ (= \mathbb{R} \text{ or } \mathbb{R}_+)$ if and only if $y/c(x, \alpha) \in D_+$ for $(x, \alpha) \in I_{x_0}^\delta \times \Theta_3$ owing to Assumption 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 2.4 (i).

Example 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,

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

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

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

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

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

where C is a constant not depending on (x, y, α). Thus, Assumptions 2.9 to 2.12 are satisfied.

Next, we show examples such that the class of jump size densities satisfies Assumption 2.4 (ii).

Example 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

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

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

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

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

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

where C is a constant not depending on (x, y, α). Thus, Assumptions 2.9 to 2.12 are satisfied, and ρ in Theorems 3.1 and 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 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

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

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

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

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

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

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 2.9 to 2.12 are satisfied, and ρ in Theorems 3.1 and 3.2 can be taken as $\rho \in (0, 1/8)$.

Example 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

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

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

$$\begin{aligned}
f_\alpha'(z) &= \left(\frac{\alpha_2 - 1}{z} - \alpha_2 \left(\frac{z}{\alpha_1} \right)^{\alpha_2 - 1} \right) f_\alpha(z) \quad (z \neq 0), \\
\frac{\partial f_\alpha}{\partial \alpha_1}(z) &= -\frac{\alpha_2}{\alpha_1} \left\{ 1 + \left(\frac{z}{\alpha_1} \right)^{\alpha_2} \right\} f_\alpha(z), \quad \frac{\partial f_\alpha}{\partial \alpha_2}(z) = \left\{ \frac{1}{\alpha_2} + \log \frac{z}{\alpha_1} - \left(\frac{z}{\alpha_1} \right)^{\alpha_2} \log \frac{z}{\alpha_1} \right\} f_\alpha(z)
\end{aligned}$$

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

$$\frac{\partial \psi}{\partial y}(x, y, \alpha) = \frac{(\alpha_2 - 1)}{y} - \frac{\alpha_2}{c(x, \alpha)} \left(\frac{y}{\alpha_1 c(x, \alpha)} \right)^{\alpha_2 - 1}$$

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

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

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

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 2.9 to 2.12 are satisfied, and ρ in Theorems 3.1 and 3.2 can be taken as $\rho \in (0, 1/4)$.

Example 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

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

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

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

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 2.9 to 2.12 are satisfied, and ρ in Theorems 3.1 and 3.2 can be taken as $\rho \in (0, 1/4)$.

Remark 5.1 As in the assumptions of Theorems 3.1 and 3.2, the range of ρ depends on q in Assumption 2.10 (ii.b) and Assumption 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 5.3, $q = 1$ in Examples 5.2 and 5.4, and any $q \in [0, 1)$ in Example 5.5.

6 Numerical experiments

In this section, we show some numerical results of our estimator for the Ornstein-Uhlenbeck processes given by

$$\mathrm{d}X_t^{\varepsilon} = -\mu_0 X_t^{\varepsilon}\mathrm{d}t + \varepsilon\sqrt{\sigma_0}\mathrm{d}W_t + \varepsilon\mathrm{d}Z_t^{\lambda_\varepsilon}, \quad X_0^{\varepsilon} = x_0 \in \mathbb{R}, \tag{6.1}$$

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 5.3.

Table 1 Sample means (with standard deviations in parentheses) of $\hat{\theta}_{n,\varepsilon}$'s, based on 1000 sample paths from the OU process (6.1) with inverse Gaussian f_α as in Example 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)	

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 1, and $N_D = 50, 100, 150$ in Table 2). Then we replace $1_{C_k^{n,\varepsilon,\rho}}$ and $1_{D_k^{n,\varepsilon,\rho}}$ in (3.1) 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 (3.2) from a sample path of (6.1) 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 1 and 2.

Table 2 Sample means (with standard deviations in parentheses) of $\hat{\theta}_{n,\varepsilon}$'s, based on 1000 sample paths from the OU process (6.1) with inverse Gaussian f_α as in Example 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)

The numbers written in bold are computed by using the same threshold as in Table 1, i.e., $N_D = 100$

Remark 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 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 3.1 and 3.2), and in this case, from Lemma 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 $\hat{D}_k^{\lambda_\varepsilon}$.

In Table 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 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.

Acknowledgements This work is partially supported by JSPS KAKENHI Grant-in-Aid for Scientific Research (C) #21K03358 and JST CREST #PMJCR14D7, Japan. Also, the authors would like to thank the anonymous referees, an Associate Editor and the Editor for their constructive comments that improved the quality of this paper.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

A Appendix

In this section, we state and prove some slightly different versions of well-known results. More precisely, we prepare Lemma A.2 as localization of the continuous mapping theorem. Lemma A.3 is a slightly different version of Lemma 9 in Genon-Catalot and Jacod (1993).

Lemma A.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.,*

$$T_{g_\theta}(\tilde{y}_\cdot) := g_\theta(\tilde{y}_\cdot) \quad \text{for } \tilde{y}_\cdot \in L^\infty([0,1];\mathcal{X}).$$

Suppose that $y_\cdot$ 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 $\mathrm{Image}(y_\cdot) := \{y_t \mid t \in [0,1]\}$. *Then, there is a neighborhood $\mathcal{N}_{y_\cdot}$ of $y_\cdot$ in $L^\infty([0,1];\mathcal{X})$ such that $\left\{T_{g_\theta}\right\}_{\theta\in\Theta}$ is a family of operators from $\mathcal{N}_{y_\cdot}$ to $L^\infty([0,1])$, and is equicontinuous at $y_\cdot$.*

Proof Fix an arbirary $\eta > 0$. For each $x \in \mathrm{Image}(y_\cdot)$, there exists $\delta_x > 0$ such that if $\left\|x - x'\right\|_{\mathcal{X}} < \delta_x$, $x, x' \in \mathcal{X}$, then

$$\sup_{\theta\in\Theta} \left|g_\theta(x) - g_\theta(x')\right| < \frac{\eta}{2}.$$

Since $\mathrm{Image}(y_\cdot)$ is compact in $\mathcal{X}$, there are finite points $x_1, \ldots, x_k \in \mathrm{Image}(y_\cdot)$ such that

$$\mathrm{Image}(y_\cdot) \subset \bigcup_{i=1}^{k} B(x_i, \delta_{x_i}/2),$$

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

$$\sup_{\theta\in\Theta} |g_\theta(\tilde{y}_t) - g_\theta(y_t)| \le \sup_{\theta\in\Theta} \left|g_\theta(\tilde{y}_t) - g_\theta(x_{i_t})\right| + \sup_{\theta\in\Theta} \left|g_\theta(x_{i_t}) - g_\theta(y_t)\right| < \eta,$$

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 A.2 *Under the same assumptions as in Lemma A.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 $(Y^\iota_\cdot)_{\iota\in I}$ is a net*

of $\mathcal{X}$-valued bounded random processes on $[0,1]$ with a directed set I. If the net $(Y^\iota_\cdot)_{\iota\in I}$ converges in probability to $y_\cdot$ in $L^\infty([0,1;\mathcal{X}])$, i.e.,

$$\left\|Y^\iota_\cdot - y_\cdot\right\|_{L^\infty([0,1];\mathcal{X})} \xrightarrow{p} 0,$$

then

$$\sup_{\theta\in\Theta}\left\|g(Y^\iota_\cdot,\theta) - g(y_\cdot,\theta)\right\|_{L^\infty([0,1])} \xrightarrow{p} 0.$$

Proof Take an arbitrary $\eta > 0$. It follows from Lemma A.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,

$$P\left(\sup_{\theta\in\Theta}\left\|g(Y^\iota_\cdot,\theta) - g(y_\cdot,\theta)\right\|_{L^\infty([0,1])} > \eta\right) \le P\left(\left\|Y^\iota_\cdot - y_\cdot\right\|_{L^\infty([0,1];\mathcal{X})} > \delta\right).$$

This implies the conclusion. □

Remark A.1 By the proof of Lemma A.2, it also follows that for any $C_1 > 0$,

$$P\left(\sup_{\theta\in\Theta}\left\|g(Y^\iota_\cdot,\theta) - g(y_\cdot,\theta)\right\|_{L^\infty([0,1])} > C_2\right) \le P\left(\left\|Y^\iota_\cdot - y_\cdot\right\|_{L^\infty([0,1];\mathcal{X})} > C_1\right),$$

where C_2 depends only on C_1, g and $\mathrm{Image}(y_\cdot)$.

Lemma A.3 *Suppose that $(\mathcal{X}, \|\cdot\|)$ is a Banach space, $\{(n,\varepsilon)\}$ is a directed set and $\left\{\mathcal{G}_i^{n,\varepsilon}\right\}_i$ 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 $\eta > 0$

$$\lim_{n,\varepsilon} P\left(\sum_{i=1}^n E\left[\|\chi_i^{n,\varepsilon}\| \,\middle|\, \mathcal{G}_{i-1}^{n,\varepsilon}\right] > \eta\right) = 0,$$

then for any $\eta > 0$

$$\lim_{n,\varepsilon} P\left(\left\|\sum_{i=1}^n \chi_i^{n,\varepsilon}\right\| > \eta\right) = 0.$$

(ii) If

$$\lim_{M\to\infty}\sup_{n,\varepsilon} P\left(\sum_{i=1}^n E\left[\|\chi_i^{n,\varepsilon}\| \,\middle|\, \mathcal{G}_{i-1}^{n,\varepsilon}\right] > M\right) = 0,$$

then

$$\lim_{M\to\infty}\sup_{n,\varepsilon} P\left(\left\|\sum_{i=1}^n \chi_i^{n,\varepsilon}\right\| > M\right) = 0.$$

Proof Since for any $\eta, \eta' > 0$

$$
\begin{aligned}
&P\left(\left\|\sum_{i=1}^{n}\chi_i^{n,\varepsilon}\right\| > \eta,\ \sum_{i=1}^{n}E\left[\|\chi_i^{n,\varepsilon}\|\,\middle|\,\mathcal{G}_{i-1}^{n,\varepsilon}\right] \le \eta'\right)\\
&\le \frac{1}{\eta}E\left[\sum_{i=1}^{n}\|\chi_i^{n,\varepsilon}\|\,1_{\left\{\sum_{i=1}^{n}E\left[\|\chi_i^{n,\varepsilon}\|\,\middle|\,\mathcal{G}_{i-1}^{n,\varepsilon}\right]\le\eta'\right\}}\right]\\
&\le \frac{1}{\eta}E\left[\left(E\left[\|\chi_n^{n,\varepsilon}\|\,\middle|\,\mathcal{G}_{n-1}^{n,\varepsilon}\right]+\sum_{i=1}^{n-1}\|\chi_i^{n,\varepsilon}\|\right)1_{\left\{\sum_{i=1}^{n}E\left[\|\chi_i^{n,\varepsilon}\|\,\middle|\,\mathcal{G}_{i-1}^{n,\varepsilon}\right]\le\eta'\right\}}\right]\\
&\le \frac{1}{\eta}E\left[\eta'+\sum_{i=1}^{n-1}\left(\|\chi_i^{n,\varepsilon}\|-E\left[\|\chi_i^{n,\varepsilon}\|\,\middle|\,\mathcal{G}_{i-1}^{n,\varepsilon}\right]\right)1_{\left\{E\left[\|\chi_i^{n,\varepsilon}\|\,\middle|\,\mathcal{G}_{i-1}^{n,\varepsilon}\right]\le\eta'\right\}}\right] < \frac{\eta'}{\eta},
\end{aligned}
$$

we obtain

$$
P\left(\left\|\sum_{i=1}^{n}\chi_i^{n,\varepsilon}\right\| > \eta\right) \le \frac{\eta'}{\eta} + P\left(\sum_{i=1}^{n}E\left[\|\chi_i^{n,\varepsilon}\|\,\middle|\,\mathcal{G}_{i-1}^{n,\varepsilon}\right] > \eta'\right).
$$

Thus, the assertions (i) and (ii) follows. □

Remark A.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 (1993). 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 Jacod and Shiryaev (2003)) when $\mathcal{X}$ is a Banach space.

Remark A.3 We have an immediate consequence from this lemma that

$$
\begin{aligned}
\sum_{i=1}^{n}E\left[\|\chi_i^{n,\varepsilon}\|\,\middle|\,\mathcal{G}_{i-1}^{n,\varepsilon}\right] = o_p(r_{n,\varepsilon}) \quad&\Longrightarrow\quad \sum_{i=1}^{n}\chi_i^{n,\varepsilon} = o_p(r_{n,\varepsilon}),\\
\sum_{i=1}^{n}E\left[\|\chi_i^{n,\varepsilon}\|\,\middle|\,\mathcal{G}_{i-1}^{n,\varepsilon}\right] = O_p(r_{n,\varepsilon}) \quad&\Longrightarrow\quad \sum_{i=1}^{n}\chi_i^{n,\varepsilon} = O_p(r_{n,\varepsilon}),
\end{aligned}
$$

where $r_{n,\varepsilon} \in \mathbb{R}$.

References

Amorino C, Gloter A (2019) Contrast function estimation for the drift parameter of ergodic jump diffusion process. Scand J Stat 47(2):279–346

Amorino C, Gloter A (2021) Joint estimation for volatility and drift parameters of ergodic jump diffusion processes via contrast function. (English summary). Stat Infer Stoch Process 24(1):61–148

Applebaum D (2009) Lévy processes and stochastic calculus, 2nd edn. Cambridge studies in advanced mathematics. Cambridge University Press, Cambridge

Billingsley P (1999) Convergence of probability measures, 2nd edn. Wiley series in probability and statistics: probability and statistics. Wiley, New York

Brezis H (2011) Functional analysis. Sobolev spaces and partial differential equations. Universitext, Springer, New York

Evans LC (2010) Partial differential equations, vol 19, 2nd edn. Graduate studies in mathematics. American Mathematical Society, Providence, RI

Genon-Catalot V, Jacod J (1993) On the estimation of the diffusion coefficient for multi-dimensional diffusion processes. Ann Inst Henri Poincaré Probab Stat 29(1):119–151

Gloter A, Loukianova D, Mai H (2018) Jump filtering and efficient drift estimation for Lévy-driven SDEs. the Ann Stat 46(4):1445–1480
Gloter A, Sørensen M (2009) Estimation for stochastic differential equations with a small diffusion coefficient. Stoch Process Appl 119(3):679–699
Ibragimov IA, Has'minskii RZ (1981) Statistical estimation. asymptotic theory. translated by Samuel Kotz, Applications of mathematics, vol 16. Springer, New York
Jean J, Shiryaev Albert N (2003) Limit theorems for stochastic processes, vol 288, 2nd edn. Springer, Berlin
Kobayashi Mitsuki, Shimizu Yasutaka (2022) Least-squares estimators based on the Adams method for stochastic differential equations with small Lévy noise. Jpn J Stat Data Sci
Long H, Shimizu Y, Sun W (2013) Least squares estimators for discretely observed stochastic processes driven by small Lévy noises. J Multivariate Anal 116:422–439
Long H, Ma C, Shimizu Y (2017) Least squares estimators for stochastic differential equations driven by small lévy noises. Stoch Process Appl 127(5):1475–1495
Ogihara T, Yoshida N (2011) Quasi-likelihood analysis for the stochastic differential equation with jumps. Stat Infer Stoch Process 14(3):189–229
Prakasa Rao BLS (1999) Semimartingales and their statistical inference, monographs on statistics and applied probability, vol 83. Chapman & Hall/CRC, Boca Raton
Shimizu Y (2006) m-estimation for discretely observed ergodic diffusion processes with infinitely many jumps. Stat Infer Stoch Process 9(2):179–225
Shimizu Y (2007) Asymptotic inference for stochastic differential equations with jumps from discrete observations and some practical approaches, Ph.D. thesis, University of Tokyo
Shimizu Y (2008) A practical inference for discretely observed jump-diffusions from finite samples. J Jpn Stat Soc 38(3):391–413
Shimizu Y (2010) Threshold selection in jump-discriminant filter for discretely observed jump processes. Stat Methods Appl 19(3):355–378
Shimizu Y (2017) Threshold estimation for stochastic processes with small noise. Scand J Stat 44(4):951–988
Shimizu Y, Yoshida N (2006) Estimation of parameters for diffusion processes with jumps from discrete observations. Stat Infer Stoch Process 9(3):227–277
Sørensen M, Uchida M (2003) Small-diffusion asymptotics for discretely sampled stochastic differential equations. Bernoulli 9(6):1051–1069
Yoshida N (1990) Asymptotic behavior of m-estimator and related random field for diffusion process, pp 221–251