Statistical Inference for Stochastic Processes (2006) 9:227–277
DOI 10.1007/s11203-005-8114-x

Estimation of Parameters for Diffusion Processes with Jumps from Discrete Observations

YASUTAKA SHIMIZU[1,*] and NAKAHIRO YOSHIDA[2]
[1]*Division of Mathematical Science, Graduate School of Engineering Science, Osaka University, Toyonaka, Osaka 560-8531 Japan*
[2]*Graduate School of Mathematical Sciences, The University of Tokyo, 3-8-1 Komaba, Meguro-ku, Tokyo 153-8914, Japan*

(Received 9 February 2003; Final version 20 December 2004; Accepted 20 April 2005)

Abstract. In this paper, we consider a multidimensional diffusion process with jumps whose jump term is driven by a compound Poisson process. Let $a(x,\theta)$ be a drift coefficient, $b(x,\sigma)$ be a diffusion coefficient respectively, and the jump term is driven by a Poisson random measure p. We assume that its intensity measure q^θ has a finite total mass. The aim of this paper is estimating the parameter $\alpha=(\theta,\sigma)$ from some discrete data. We can observe $n+1$ data at $t_i^n=ih_n$, $0\leqslant i\leqslant n$. We suppose $h_n\to 0$, $nh_n\to\infty$, $nh_n^2\to 0$.

Key words: diffusion process with jumps, parametric inference, discrete observation, contrast function, asymptotic normality, asymptotic efficiency.

1. Introduction

Let $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\geqslant 0},P)$ be a filtered probability space. We consider a d-dimensional solution process $X.$ to the following stochastic differential equation on this space:

$$\begin{cases} \mathrm{d}X_t=a(X_{t-},\theta)\,\mathrm{d}t+b(X_{t-},\sigma)\,\mathrm{d}W_t+\displaystyle\int_E c(X_{t-},z,\theta)\,(p-q^\theta)(\mathrm{d}t,dz), \\ X_0=x_0, \end{cases} \tag{1.1}$$

where $E=\mathbf{R}^d\setminus\{0\}$, $\theta\in\Theta\subset\mathbf{R}^{m_1}$, $\sigma\in\Pi\subset\mathbf{R}^{m_2}$ are parameters, and $\alpha=(\theta,\sigma)$ belongs to a parameter space $\Xi=\Theta\times\Pi$ which is a compact convex subset of $\mathbf{R}^{m_1}\times\mathbf{R}^{m_2}$. $(W_t)_{t\geqslant 0}$ is an r-dimensional standard $\mathcal{F}_t$-Brownian motion $p(\mathrm{d}t,\mathrm{d}z)$ is a homogeneous Poisson random measure on $\mathbf{R}_+\times\mathbf{R}^d$, and $q^\theta(\mathrm{d}t,\mathrm{d}z)$ is its intensity measure, that is, $E[p(\mathrm{d}t,\mathrm{d}z)]=q^\theta(\mathrm{d}t,\mathrm{d}z)$. We set $q^\theta(\mathrm{d}t,\mathrm{d}z)=f_\theta(z)\,\mathrm{d}z\,\mathrm{d}t$ and $f_\theta(z)=\lambda(\theta)F_\theta(z)$, where $\lambda(\theta)$ is a positive function of θ and F_θ is a probability density. The coefficients a and c are known $\mathbf{R}^d$-valued Borel functions defined on $\mathbf{R}^d\times\Theta$, $\mathbf{R}^d\times E\times\Theta$ respectively, and b is a known $\mathbf{R}^d\otimes\mathbf{R}^r$-valued Borel function defined on $\mathbf{R}^d\times\Pi$.

*Author for correspondence: e-mail: yasutaka@sigmath.es.osaka-u.ac.jp

One of the most simple but important examples of one-dimensional case is written as follows.

$$\mathrm{d}X_t = \tilde{a}(X_{t-},\mu)\,\mathrm{d}t + \tilde{b}(X_{t-},\sigma)\,\mathrm{d}W_t + \tilde{c}(X_{t-},\vartheta)\,\mathrm{d}Z_t^{\vartheta}, \tag{1.2}$$

where $Z_\cdot^{\vartheta}$ is a Lévy process with parameter ϑ. This belongs to the class of (1.1) with $c(x,z,\theta)=\tilde{c}(x,\vartheta)z$, $a(x,\theta)=\tilde{a}(x,\mu)+\int_E c(x,\vartheta)zf_\vartheta(z)\,\mathrm{d}z$ and $\theta=(\mu,\vartheta)$.

In this paper, we studied estimation of the parameter $\alpha=(\theta,\sigma)$ from discrete observations. For this aim, we observe $n+1$ data $(X_{t_i^n})_{0\leqslant i\leqslant n}$, $t_i^n = ih_n$, and show the consistency and asymptotic normality of an estimator under the assumption that $h_n\to 0$, $nh_n\to\infty$, $nh_n^2\to 0$.

So far, estimation of discretely observed diffusion processes has been studied very well by several authors. Although there are some observation schemes, our scheme is so-called rapidly increasing experimental design: $nh_n^2\to 0$. The earlier work on this scheme is seen in Prakasa-Rao [23,24]. He studied a least squares approach. Florense-Zmirou [8] considered estimation of a one-dimensional diffusion process with a constant diffusion coefficient under the less restrictive condition $nh_n^3\to 0$ with convergence rate $\sqrt{nh_n}$ for a diffusion estimation. Yoshida [29] studied the case where drift-diffusion parameter estimation cannot be split, and showed the joint convergence of an adaptive estimator with $\sqrt{n}$ convergence rate for diffusion parameter. After that, Kessler [18,19] improved it to more general case with the design $nh_n^p\to 0$ for all $p\geqslant 2$. For other schemes, for instance $nh_n=$ constant or $h_n=$ constant, see Genon-Catalot and Jacod [10], Dacunha-Castelle and Florens-Zmirou [7], Bibby and Sørensen [4], and so on. On the other hand, estimation of a diffusion processes with jumps is well-known in case the whole path is observed, see, for example, Sørensen [27, 28]. However it is obviously impossible to observe such a path over any given time period in actual situation, therefore inference based on discrete data is of major importance in dealing with practical problems.

Recently, a jump-diffusion model is also used in various fields. It is particularly important in applications to modeling security prices or option prices in financial market, see Aase and Guttorp [1], Mulinacci [21], Scott [25], Gukhal [11], and so on. There are other applications to soil moisture model (Mtundu and Koch [22]), to hydorology (Bodo and Thomoson [6]), and to population model (Hanson and Tuckwell [14], Guttrop and Kulperger [12]), and so on.

In this paper, we construct a single contrast function to estimate the parameter (θ,σ) jointly. The exact likelihood function would be used if we would know the form of transition probabilities. However it is generally impossible to write it down explicitly, therefore we have to approximate it by a suitable function. In Section 2.2, we present an estimating function having two parts: the first part is the log likelihood of a local Gaussian process, and the second one is modeled after the likelihood of Poisson random measures. This estimating function divides increments according to their magnitudes and assigns them those parts. Evaluating the probability of misclassifications, we prove the asymptotic normality of our estimator. Our estimator is of the maximum likelihood type. In order to control this efficient likelihood function, we put regularity conditions such as A?? below. However, if we are satisfied with another less efficient estimator, then we can relax those conditions. For example, it is possible to treat a process with infinitely many jumps on compacts. We discuss such generalizations elsewhere.

We assume the ergodicity of the solution process in this paper. There is a work, e.g., by Kwon and Lee [20] on the conditions of the ergodicity of diffusion processes with jumps.

The plan of this paper is the following. In the next Section 2, we prepare the notations and some assumptions used in this paper. After that, we present a contrast function and our main result. In Section 3, we prepare several propositions which are needed to prove the main theorem. All proofs are presented in Section 4.

2. Discussion and Conclusions

2.1. NOTATIONS AND ASSUMPTIONS

We introduce some notations.

1. We set $\alpha=(\theta,\sigma)$ and $\alpha_0,\theta_0,\sigma_0$ denote the respective true values. For $\lambda(\theta)$ we set $\lambda_0=\lambda(\theta_0)=\int f_{\theta_0}(z)\,\mathrm{d}z$.
2. For $\kappa=(\kappa_1,\ldots,\kappa_d)$, $\partial_{\kappa_j}:=\dfrac{\partial}{\partial\kappa_j}$, $\partial^2_{\kappa_j}:=\dfrac{\partial^2}{\partial\kappa_j^2}$, $\partial^2_{\kappa_i\kappa_j}:=\dfrac{\partial^2}{\partial\kappa_i\partial\kappa_j}$, $\partial_\kappa:=(\partial_{\kappa_1},\ldots,\partial_{\kappa_d})^*$, and $\partial^2_\kappa f(\kappa)=\left(\partial^2_{\kappa_i\kappa_j}f(\kappa)\right)_{1\leqslant i,j\leqslant d}$, where $*$ stands for the transpose.
3. For a function g defined on $\mathbf{R}^d\times\Xi$ we denotes by $g_{i-1}(\alpha)$ the value $g(X_{t^n_{i-1}},\alpha)$. If g is a tensor, then we express its components with upper index, for example, if g is a matrix, then its (k,l)-component is $g^{(k,l)}$.
4. $\mathcal{F}^n_{i-1}:=\mathcal{F}_{t^n_{i-1}}$, $\quad\Delta X^n_i:=X_{t^n_i}-X_{t^n_{i-1}}$, $\quad\Delta X_t:=X_t-X_{t-}$, $\quad\bar{X}_{i,n}(\theta)=\Delta X^n_i-h_n\bar{a}_{i-1}(\theta)$.
5. $\beta(x,\sigma):=b(x,\sigma)b^*(x,\sigma)$.
6. For a tensor A, $|A|^2$ is the sum of squares of the components of A.
7. Let u_n be a real valued sequence. $R:\Xi\times\mathbf{R}\times\mathbf{R}^d\to\mathbf{R}$ denotes a function for which there exists a constant C such that

$$R(\alpha,u_n,x)\leqslant u_nC(1+|x|)^C\quad\text{for all }\alpha\in\Xi,\ x\in\mathbf{R}^d,\ n\in\mathbf{N}.$$

We set $\tilde{R}(\alpha,u_n,x)=1-R(\alpha,u_n,x)$.

8. Let b_n be a real valued sequence such that $b_n \to \infty$, $nh_n^2 b_n \to 0$, $\dfrac{b_n}{nh_n} \to 0$. For example $h_n = n^{-2/3}$, $b_n = n^{1/4}$, etc. Moreover, we set $\varepsilon_n = b_n^{-1/10}$.
9. In the following discussion, if we write X, then it means the solution to (1.1) with $\alpha = \alpha_0$.
10. We often use the notation C (resp. C_k) as a general positive constant (resp. depending on the index k), therefore we sometimes use the same character for different constants from line to line without specially mentioning.

We make the following assumptions.

A.1. For some constant L and function $\zeta(z)$ of polynomial growth in z,

$$|a(x,\theta_0) - a(y,\theta_0)| + |b(x,\sigma_0) - b(y,\sigma_0)| \leqslant L|x-y|,$$

$$|c(x,z,\theta_0) - c(y,z,\theta_0)| \leqslant \zeta(z)|x-y|, \quad |c(x,z,\theta_0)| \leqslant \zeta(z)(1+|x|).$$

A.2. The diffusion process with jumps X is ergodic and stationary for $\alpha = \alpha_0$: if $\pi(dx)$ is its invariant probability measure, the initial distribution is π and

$$\frac{1}{T}\int_0^T f(X_t)\,\mathrm{d}t \xrightarrow{P} \int f(x)\,\mathrm{d}\pi$$

as $T \to \infty$ for any π-integrable function f.

A.3. For every $p \geqslant 1$,

$$\sup_t E[|X_t|^p] < \infty.$$

A.4. For fixed θ and σ, the derivatives $\partial_x^l a(x,\theta)$ and $\partial_x^l b(x,\sigma)$ $(l = 1,2)$ exist on $\mathbf{R}^d$ and they are continuous in x. Moreover, for fixed x, the derivatives $\partial_\theta^l a(x,\theta)$ and $\partial_\sigma^l b(x,\sigma)$ $(l = 1,2)$ exist on Θ and Π respectively, and a, b, and their all derivatives are of polynomial growth in x uniformly in α:

$$|\partial_x^l a(x,\theta)|,\ |\partial_x^l b(x,\sigma)|,\ |\partial_\theta^l a(x,\theta)|,\ |\partial_\sigma^l b(x,\sigma)| \leqslant C(1+|x|)^C.$$

for $l = 0, 1, 2$.

A.5. $\inf_x |c(x,z,\theta_0)| \geqslant c_0|z|$ for some $c_0 > 0$ near the origin

A.6. There exist constants r, $K > 0$ and $\gamma > 3$ such that $f_{\theta_0}(z)\mathbf{1}_{\{|z| \leqslant r\}} \leqslant K|z|^\gamma$, and that

$$\sup_\theta \int |z|^p f_\theta(z)\,\mathrm{d}z < \infty \quad \text{for all } p \geqslant 1.$$

A.7. For each θ and x, the mapping $z \mapsto y = c(x, z, \theta)$ has an inverse $z = c^{-1}(x, y, \theta)$ which is differentiable with respect to y, and we set

$$\Psi_\theta(y, x) := f_\theta\left(c^{-1}(x, y, \theta)\right) J(x, y, \theta),$$

where $J(x, y, \theta)$ is the absolute value of the Jacobian of $c^{-1}(x, y, \theta)$.

A.8. β is a positive definite and $\inf_{x,\sigma} \det \beta(x, \sigma) > 0$.

A.9. The function $\Psi_\theta(y, x)$ is differentiable with respect to x and y, and three times continuously differentiable with respect to θ, Moreover we assume that

$$|\partial_\theta^k \Psi_\theta(y, x)| \leqslant L_1(y)(1 + |x|)^C \quad (k = 0, 1, 2, 3), \tag{2.1}$$

$$|\partial_x \partial_\theta^l \Psi_\theta(y, x)| \leqslant L_2(y) \quad (l = 0, 1, 2), \tag{2.2}$$

where L_1 and L_2 are bounded and dy-integrable functions. Furthermore,

$$|\partial_y \partial_\theta^l \Psi_\theta(y, x)| \leqslant C(1 + |y|)^C (1 + |x|)^C \quad (l = 0, 1, 2), \tag{2.3}$$

$$\int \sup_\theta \left|\partial_\theta^k \log \Psi_\theta(y, x) \cdot \Psi_{\theta_0}(y, x)\right| \mathrm{d}y \leqslant C(1 + |x|)^C \quad (k = 0, 1, 2, 3). \tag{2.4}$$

A.10. $\sigma = \sigma_0$ if and only if $\det \beta(x, \sigma) = \det \beta(x, \sigma_0)$ for almost all x. Moreover, $\theta = \theta_0$ if and only if $a(x, \theta) = a(x, \theta_0)$ and $\Psi_\theta(y, x) = \Psi_{\theta_0}(y, x)$ for almost all x and y.

A.11. There exists a sequence of real valued functions $\{\varphi_n(x, y)\}_{n \in \mathbf{N}}$ possessing following properties: $0 \leqslant \varphi_n \leqslant 1$ and $\varphi_n \to 1$ $\mathrm{d}y \times \mathrm{d}\pi$-$a.s.$ as $n \to \infty$. There exists some $M > 0$ such that

$$\varphi_n(x, y) = \begin{cases} 1 \ (x, y) \in \left\{\inf_{\theta \in \Theta} \Psi_\theta(y, x) > M\right\} \\ 0 \ (x, y) \in \left\{\inf_{\theta \in \Theta} \Psi_\theta(y, x) \leqslant \varepsilon_n\right\} \end{cases}.$$

Moreover,

$$\sup_{x,y} |\partial_x \varphi_n| + \sup_{x,y} |\partial_y \varphi_n| = \mathrm{O}(\varepsilon_n^{-1}), \tag{2.5}$$

and

$$\partial_x \varphi_n = \partial_y \varphi_n = 0 \quad \text{on} \quad \left\{\inf_{\theta \in \Theta} \Psi_\theta(y, x) \leqslant \varepsilon_n\right\}. \tag{2.6}$$

Let us give some examples of $\Psi_\theta(y,x)$ and check conditions in A.9.

EXAMPLE 2.1. We consider the following one-dimensional stochastic differential equation.

$$\mathrm{d}X_t = a(X_{t-})\,\mathrm{d}t + b(X_{t-})\,\mathrm{d}W_t + c(X_{t-},\theta)\,\mathrm{d}Z_t^\theta, \tag{2.7}$$

where $Z_\cdot^\theta$ is a Lévy process with Lévy density $f_\theta(z)$. We assume, for some $M,\ K>0$,

$$c(x,\theta) = \frac{1}{c}, \quad f_\theta(z) = Mz^\alpha(K-z)^\beta \mathbf{1}_{\{0\leqslant z\leqslant K\}},$$

and $\theta = (c,\alpha,\beta)$, where $\alpha,\ \beta > 3$.

$$\Psi_\theta(y,x) = Mc(cy)^\alpha(K-cy)^\beta,$$
$$\log\Psi_\theta(y,x) = \log M + (\alpha+1) \quad \log c + \alpha\log y + \beta\log(K-cy)$$

Noticing the parameter space Θ is compact, it is easy to verify (2.1) – (2.3). Moreover all $\int \sup_\theta \left|\partial_\theta^k \log\Psi_\theta(y,x)\cdot\Psi_{\theta_0}(y,x)\right|\,\mathrm{d}y$ become finite, so (2.4) is satisfied.

EXAMPLE 2.2. For SDE (2.7), we suppose that $\mathrm{supp}\,(f_\theta)\subset\mathbf{R}_+$ and

$$c(x,\theta) = \frac{1}{c}, \quad f_\theta(z)\mathbf{1}_{\{0<z\leqslant M\}} = \mathrm{e}^{-\frac{\gamma}{z}}\mathbf{1}_{\{0<z\leqslant M\}}.$$

and $\theta = (c,\gamma)$, where $c>0,\ \gamma>0$ and $M>0$. It is easy to check A.9 in the neighborhood of the origin since

$$\Psi_\theta(y,x) = e^{-\frac{\gamma}{cy}}, \quad \log\Psi_\theta(y,x) = -\frac{\gamma}{cy}.$$

on the set $\{0<z\leqslant M\}$.

EXAMPLE 2.3. For SDE (2.7), we assume

$$c(x,\theta) = \frac{1}{c}, \quad f_\theta(z) = \frac{\alpha^\beta}{\Gamma(\beta)} z^{\beta-1}\mathrm{e}^{-\alpha z},$$

and $\theta = (c,\alpha,\beta)$, where $\alpha>0,\ \beta>4$. Then

$$\Psi_\theta(y,x) = \frac{\alpha^\beta}{\Gamma(\beta)}(cy)^{\beta-1}\mathrm{e}^{-\alpha cy},$$
$$\log\Psi_\theta(y,x) = \beta\log\alpha - \log\Gamma(\beta) + (\beta-1)\log cy - \alpha cy.$$

Then

$$|\partial_\theta^k \Psi_\theta(y,x)|, \quad |\partial_x \partial_\theta^l \Psi_\theta(y,x)|, \quad |\partial_y \partial_\theta^l \Psi_\theta(y,x)|$$

are all dominated by $C\,(1+|\log y|)\,y^{\beta'}e^{-\alpha' y}$ for some $C>0$, $\beta'>0$, $\alpha'>0$, so Assumptions (2.1)–(2.3) are satisfied. Moreover we obtain

$$\sup_\theta \left|\partial_\theta^k \log \Psi_\theta(y,x)\right| \leqslant C\,(1+|y|+|\log y|)\,.$$

This implies (2.4).

Under Assumptions A.1 and A.6, the stochastic differential equation (1.1) can be rewritten as follows:

$$dX_t = \bar{a}(X_{t-},\theta)\,\mathrm{d}t + b(X_{t-},\sigma)\,\mathrm{d}W_t + \int_E c(X_{t-},z,\theta)\,p(\mathrm{d}t,\mathrm{d}z), \tag{2.8}$$

where $\bar{a}(x,\theta)=a(x,\theta)-\int_E c(x,z,\theta)\,q^\theta(\mathrm{d}t,\mathrm{d}z)$. This expression implies that X follows diffusion process $\mathrm{d}X_t=\bar{a}(X_{t-},\theta)\,\mathrm{d}t+b(X_{t-},\sigma)\,\mathrm{d}W_t$, in the interval in which no jump occurred. We start with the stochastic differential equation (2.8) to construct the contrast function.

2.2. CONTRAST FUNCTION AND MAIN RESULT

Now we present a contrast function for estimating parameters. In Section 2.3, we show how to obtain it.

DEFINITION 2.1. For $\frac{2}{\gamma+1}\leqslant\rho<\frac{1}{2}$, we define the contrast function $l_n(\alpha)$ as follows

$$l_n(\alpha)=\bar{l}_n(\theta,\sigma)+\tilde{l}_n(\theta),$$

where

$$\bar{l}_n(\alpha) = -\frac{1}{2h_n}\sum_{i=1}^n (\bar{X}_{i,n})^*(\theta)\beta_{i-1}^{-1}(\sigma)\bar{X}_{i,n}(\theta)\mathbf{1}_{\{|\Delta X_i^n|\leqslant h_n^\rho\}} - \\ -\sum_{i=1}^n \frac{1}{2}\log\det\beta_{i-1}(\sigma)\mathbf{1}_{\{|\Delta X_i^n|\leqslant h_n^\rho\}},$$

$$\tilde{l}_n(\theta) = \sum_{i=1}^n \left\{\log\Phi_n(\theta, X_{t_{i-1}^n}, \Delta X_i^n)\right\}\varphi_n(X_{t_{i-1}^n}, \Delta X_i^n)\mathbf{1}_{\{|\Delta X_i^n|>h_n^\rho\}} - \\ -h_n\sum_{i=1}^n \int \Phi_n(\theta, X_{t_{i-1}^n}, y)\,\mathrm{d}y,$$

$$\bar{X}_{i,n}(\theta)=X_{t_i^n}-X_{t_{i-1}^n}-h_n\bar{a}_{i-1}(\theta), \quad \Phi_n(\theta,x,y)=\Psi_\theta(y,x)\varphi_n(x,y).$$

Intuitively speaking, this contrast function is very natural since $\bar{l}_n(\alpha)$ corresponds to the contrast for an usual diffusion process, and $\tilde{l}_n(\theta)$ does to the discretization of the likelihood function of an compound Poisson process with Lévy density f_θ.

We used a truncation function φ_n to ensure the P-integrability of the logarithm term and its derivatives with respect to parameters. If we knew $\sup_{i,n} E\left[\partial_y \partial_\theta^l \log \Psi_\theta(\Delta X_i^n, X_{t_{i-1}^n})\right] < \infty$ $(l=0,1,2)$ by some reasons, φ_n would not be needed, however φ_n is needed generally.

Our main theorem is the following. The proof will be presented in Section 4.10

THEOREM 2.1. *Under Assumptions A.1–A.11 and* $nh_n \to \infty$ *as* $n \to \infty$ *the estimator* $\hat{\alpha}_n$ *which satisfies*

$$l_n(\hat{\alpha}_n) = \sup_{\alpha \in \Xi} l_n(\alpha)$$

is consistent:

$$\hat{\alpha}_n \xrightarrow{P} \alpha_0 \quad (n \to \infty).$$

If, in addition, $nh_n^2 \to 0$ *and the true value* α_0 *is in the interior of* Ξ *then*

$$\left(\sqrt{nh_n}(\hat{\theta}_n - \theta_0), \sqrt{n}(\hat{\sigma}_n - \sigma_0)\right) \xrightarrow{d} \mathcal{N}(0, K^{-1}),$$

where

$$K := \begin{pmatrix} K_1 & \mathbf{0} \\ \mathbf{0} & K_2 \end{pmatrix},$$

$$K_1^{(p,q)} = \int (\partial_{\theta_p} \bar{a})^* \beta^{-1} (\partial_{\theta_{p'}} \bar{a})(x, \alpha_0)\, \mathrm{d}\pi + \iint \frac{\partial_{\theta_p} \Psi_{\theta_0} \partial_{\theta_{p'}} \Psi_{\theta_0}}{\Psi_{\theta_0}}(y, x)\, \mathrm{d}y\, \mathrm{d}\pi,$$

$$K_2^{(p,q)} = \frac{1}{2} \int \operatorname{tr}\left[(\partial_{\sigma_q} \beta) \beta^{-1} (\partial_{\sigma_{q'}} \beta) \beta^{-1}\right](x, \sigma_0)\, \mathrm{d}\pi.$$

Remark 2.1. This result can be extended to pure jump type processes with $b(x,\sigma) \equiv 0$, that is, X is a solution process to the following stochastic differential equation:

$$\mathrm{d}X_t = a(X_{t-}, \theta)\, \mathrm{d}t + \int_E c(X_{t-}, z, \theta)\, (p - q^\theta)(\mathrm{d}t, \mathrm{d}z).$$

The contrast function of jump part is similar to the non-degenerate case since we estimate jump parameters from only the number of jumps and their amplitudes. For diffusion part, however, we cannot make use of $\bar{l}_n(\alpha)$ any more because we cannot approximate the path of X by the local

Gaussian approximation in the no jump intervals. We can overcome this difficulty by estimating drift parameters as least square estimators, that is,

$$\bar{l}_n(\theta) = -\frac{1}{2h_n}\sum_{i=1}^{n}(\bar{X}_{i,n})^*(\theta)\bar{X}_{i,n}(\theta)\mathbf{1}_{\{|\Delta X_i^n|\leqslant h_n^\rho\}}.$$

In this case, $\hat{\theta}_n$ has also consistency and asymptotic normality with asymptotic variance K, $K^{(p,q)} = \iint \frac{\partial_{\theta_p}\Psi_{\theta_0}\partial_{\theta_{p'}}\Psi_{\theta_0}}{\Psi_{\theta_0}}(y,x)\,\mathrm{d}y\,\mathrm{d}\pi$. The proof is the same as for non-degenerate case.

Remark 2.2. The asymptotic efficiency for $\hat{\theta}_n$ is obtained since K_1 is the asymptotic variance of the estimator for the continuously observed ergodic diffusion processes with jumps.

Sørensen [28] discusses the inference of diffusion processes with jumps from continuous observations under the setting which includes the non-ergodic case. Particularly, when you compute the asymptotic variance in the ergodic case, it could be more clear to refer Section 3 in Barndorff-Nielsen and Sørensen [3] which is a review of the general likelihood theory.

2.3. CONSTRUCTION OF CONTRAST FUNCTION

As we described, the data we can get are discrete ones, hence we have to decide whether jumps occur or not in an interval from only the increment $|\Delta X_i^n|$, although that is a stochastic decision which may sometimes include some misjudgements. This criterion should be chosen depending on n, and increase the accuracy of judgements as n tends to infinity. The way we will take is that, for $\rho \in [0,1)$, if the increment exceeds h_n^ρ in an interval, we regard it as an interval in which a jump has occurred and if not, as an interval in which no jump. This is because the increment of a diffusion without jumps exceeds h_n^ρ with small probability, and the increment of a diffusion with a single jump also exceeds h_n^ρ with a large probability. Although they are intuitive argument, these are justified by Lemma 2.1 described below.

The value ρ has to be chosen carefully. For instance if ρ is too large, and therefore h_n^ρ is too small, the probability of getting the increment h_n^ρ by continuous diffusion cannot be ignored, on the other hand, if ρ is too small, and therefore h_n^ρ is too large, we cannot ignore the probability of getting an increment less than h_n^ρ when a jump occurs in an interval. Later we will choose ρ as $2/(\gamma+1)\leqslant\rho<1/2$.

LEMMA 2.1. *Define random times* τ_i^n *and* η_i^n *as*

$$\tau_i^n := \inf\left\{t\in[t_{i-1}^n, t_i^n);\, |\Delta X_t|>0\right\},$$

$$\eta_i^n := \sup\left\{t \in [t_{i-1}^n, t_i^n);\ |\Delta X_t| > 0\right\}.$$

If the infimum or supremum on the right hand side does not exist, we define the random times to equal t_i^n. *Assume Conditions A.1, A.3 and A.6. Then, for* $\rho \in [0, 1/2)$ *and any* $p \geqslant 1$,

$$P\left\{\sup_{t\in[t_{i-1}^n,\tau_i^n)} |X_t - X_{t_{i-1}^n}| > h_n^\rho \,\middle|\, \mathcal{F}_{i-1}^n\right\} = R(\alpha, h_n^p, X_{t_{i-1}^n}), \tag{2.9}$$

$$P\left\{\sup_{t\in[\eta_i^n,t_i^n)} |X_{t_i^n} - X_t| > h_n^\rho \,\middle|\, \mathcal{F}_{i-1}^n\right\} = R(\alpha, h_n^p, X_{t_{i-1}^n}). \tag{2.10}$$

Each function R *does not depend on* i.

By these facts, the probability whether the value of $|\Delta X_i^n|$ exceeds h_n^ρ or not is evaluated in the next lemma. In the following discussion, let J_i^n be the number of jumps in the interval $[t_{i-1}^n, t_i^n)$ and we set

$$\left\{|\Delta X_i^n| \leqslant h_n^\rho\right\} = \bigcup_{j=0}^{2} C_{i,j}^n, \quad \left\{|\Delta X_i^n| > h_n^\rho\right\} = \bigcup_{j=0}^{2} D_{i,j}^n,$$

where

$$C_{i,0}^n = \left\{J_i^n = 0,\ |\Delta X_i^n| \leqslant h_n^\rho\right\}, \quad C_{i,1}^n = \left\{J_i^n = 1,\ |\Delta X_i^n| \leqslant h_n^\rho\right\},$$
$$C_{i,2}^n = \left\{J_i^n \geqslant 2,\ |\Delta X_i^n| \leqslant h_n^\rho\right\},$$

$$D_{i,0}^n = \left\{J_i^n = 0,\ |\Delta X_i^n| > h_n^\rho\right\}, \quad D_{i,1}^n = \left\{J_i^n = 1,\ |\Delta X_i^n| > h_n^\rho\right\},$$
$$D_{i,2}^n = \left\{J_i^n \geqslant 2,\ |\Delta X_i^n| > h_n^\rho\right\}.$$

LEMMA 2.2. *Assume Conditions A.1, A.3, A.5 and A.6. Let* $\dfrac{2}{\gamma+1} \leqslant \rho < \dfrac{1}{2}$, *where* γ *is the constant in condition A.6. For any* $p \geqslant 1$, *as* $n \to \infty$

$$P\{C_{i,0}^n|\mathcal{F}_{i-1}^n\} = \mathrm{e}^{-\lambda_0 h_n}\tilde{R}(\alpha, h_n^p, X_{t_{i-1}^n}) \qquad P\{D_{i,0}^n|\mathcal{F}_{i-1}^n\} = \mathrm{e}^{-\lambda_0 h_n}R(\alpha, h_n^p, X_{t_{i-1}^n})$$
$$P\{C_{i,1}^n|\mathcal{F}_{i-1}^n\} = R(\alpha, h_n^3, X_{t_{i-1}^n}) \qquad P\{D_{i,1}^n|\mathcal{F}_{i-1}^n\} = \lambda_0 h_n e^{-\lambda_0 h_n}\tilde{R}(\alpha, h_n^2, X_{t_{i-1}^n})$$
$$P\{C_{i,2}^n|\mathcal{F}_{i-1}^n\} \leqslant \lambda_0^2 h_n^2 \qquad P\{D_{i,2}^n|\mathcal{F}_{i-1}^n\} \leqslant \lambda_0^2 h_n^2$$

This lemma implies that we can judge the interval $[t_{i-1}^n, t_i^n)$ has no jump if $|\Delta X_i^n| \leqslant h_n^\rho$ and the interval has a single jump if $|\Delta X_i^n| > h_n^\rho$, and that we can ignore the events which include more than two jumps in the interval.

In our model, we may consider that the random measure p is generated by d-dimensional compound Poisson process thanks to the assumption A.6. Let $(Z_t)_{t\geqslant 0}$ be a compound Poisson process which is independent of W and has the form $Z_t = \sum_{i=1}^{N_t} \epsilon_i$ where $(N_t)_{t\geqslant 0}$ is a Poisson process with intensity $\lambda(\theta)(\epsilon_i)_{i\in\mathbf{N}}$ is a sequence of d-dimensional random vectors which are independent of each other and identically distributed with density $F_\theta(x)$. N and $(\epsilon_i)_i$ are also independent of each other. In this setting, p can be expressed as follows:

$$p^\theta(\mathrm{d}t, \mathrm{d}z) = \sum_{s\geqslant 0} \mathbf{1}_{\{\Delta Z_s \neq 0\}} \mathbf{1}_{(s,\Delta Z_s)}(\mathrm{d}t, \mathrm{d}z).$$

Hence if Z has a jump of size z at time t, then X will have a jump of size $c(X_{t-}, z, \theta)$ at the same time.

Now let us discuss the approximation of the transition probability. First, we consider the transition probability from $X_{t_{i-1}^n}$ to $X_{t_i^n}$ in the case of single jump in the interval $[t_{i-1}^n, t_i^n)$. We set $\tau_i^n := \inf\{t;\, |\Delta X_t| > 0,\ t_{i-1}^n \leqslant t < t_i^n\}$. Since no jump occurs in $[t_{i-1}^n, \tau_i^n-]$, we approximate the transition by the one of

$$X_{\tau_i^n-} = X_{t_{i-1}^n} + \bar{a}_{i-1}(\tau_i^n - t_{i-1}^n) + b_{i-1}N,$$

where $N \sim \mathcal{N}_d(0, (\tau_i^n - t_{i-1}^n)I)$. The above $X_\cdot$ is not the same as the solution process to (1.1), but the above $X_{t_{i-1}^n}$ has the same value as the one of (1.1). Next, since we suppose that no jump occurs after a jump at τ_i^n, we can take the same approximation as above in $[\tau_i^n, t_i^n)$, that is,

$$X_{t_i^n} = X_{\tau_i^n-} + \bar{a}(X_{\tau_i^n})(t_i^n - \tau_i^n) + b(X_{\tau_i^n})N' + c(X_{\tau_i^n-}, \Delta Z_{\tau_i^n}),$$

where $N' \sim \mathcal{N}_d(0, (t_i^n - \tau_i^n)I)$. Let $\phi(x; A, B)$ be a Gaussian density with mean vector A and variance matrix B. Since the distribution of the jump time τ_i^n conditional on $J_i^n = 1$ becomes a uniform distribution on $[t_{i-1}^n, t_i^n)$,

$$\begin{aligned}
&P\{X_{t_i^n} \in A,\, J_i^n = 1 | \mathcal{F}_{i-1}^n\}/P\{J_i^n = 1\} \\
&\quad = \int_A \int_z \int_{x'} \int_{t_{i-1}^n}^{t_i^n} \frac{1}{h_n} \phi\left(x'; X_{t_{i-1}^n} + \bar{a}_{i-1}(s - t_{i-1}^n), \beta_{i-1}(s - t_{i-1}^n)\right) \times \\
&\qquad \times \phi\left(x; x' + c(x', z) + \bar{a}(x' + c(x', z))(t_i^n - s)\right., \\
&\qquad \left.\beta\ (x' + c(x', z))(t_i^n - s)\right) F_{\theta_0}(z)\,\mathrm{d}s\,\mathrm{d}x'\,\mathrm{d}z\,\mathrm{d}x.
\end{aligned}$$

We denote by $p_{i,n}^d(x)$ the above probability density function. Secondly, using the local Gauss approximation for $J_i^n = 0$

$$P\{X_{t_i^n} \in A,\, J_i^n = 0 | \mathcal{F}_{i-1}^n\}/P\{J_i^n = 0\} = \int_A \phi\left(x; X_{t_i^n} + \bar{a}_{i-1}h_n, \beta h_n\right) \mathrm{d}x.$$

We denote by $p_{i,n}^c(x)$ this probability density function. Finally,

$$P\{X_{t_i^n} \in A,\ \ J_i^n \geqslant 2 | \mathcal{F}_{i-1}^n\} = \mathrm{O}_p(h_n^2).$$

Since

$$
\begin{aligned}
P\{X_{t_i^n}\in A|\mathcal{F}_{i-1}^n\} &= P\{X_{t_i^n}\in A,\ \ J_i^n=0|\mathcal{F}_{i-1}^n\}+P\{X_{t_i^n}\in A,\ \ J_i^n=1|\mathcal{F}_{i-1}^n\}+\\
&\quad + P\{X_{t_i^n}\in A,\ \ J_i^n\geqslant 2|\mathcal{F}_{i-1}^n\}\\
&= \sum_{j=0}^{2}\left[P\{X_{t_i^n}\in A,\, C_{i,j}^n|\mathcal{F}_{i-1}^n\}+P\{X_{t_i^n}\in A,\, D_{i,j}^n|\mathcal{F}_{i-1}^n\}\right],
\end{aligned}
$$

and by Lemma 2.2,

$$
\begin{aligned}
&P\{X_{t_i^n}\in A,\, C_{i,0}^n|\mathcal{F}_{i-1}^n\}=\mathrm{e}^{-\lambda_0 h_n}\int_A \mathbf{1}_{\{|x-X_{t_{i-1}^n}|\leqslant h_n^\rho\}}p_{i,n}^c(x)\,\mathrm{d}x,\\
&P\{X_{t_i^n}\in A,\, C_{i,1}^n|\mathcal{F}_{i-1}^n\}\leqslant P\{C_{i,1}^n|\mathcal{F}_{i-1}^n\}=\mathrm{O}_p(h_n^3),\\
&P\{X_{t_i^n}\in A,\, C_{i,2}^n|\mathcal{F}_{i-1}^n\}=\mathrm{O}_p(h_n^2),\\
&P\{X_{t_i^n}\in A,\, D_{i,0}^n|\mathcal{F}_{i-1}^n\}\leqslant P\{D_{i,0}^n|\mathcal{F}_{i-1}^n\}=\mathrm{O}_p(h_n^3),\\
&P\{X_{t_i^n}\in A,\, D_{i,1}^n|\mathcal{F}_{i-1}^n\}=\lambda_0 h_n\mathrm{e}^{-\lambda_0 h_n}\int_A \mathbf{1}_{\{|x-X_{t_{i-1}^n}|>h_n^\rho\}}p_{i,n}^d(x)\,\mathrm{d}x,\\
&P\{X_{t_i^n}\in A,\, D_{i,2}^n|\mathcal{F}_{i-1}^n\}=\mathrm{O}_p(h_n^2),
\end{aligned}
$$

therefore we can approximate the transition density $p_{i,n}(x)$ by

$$
\begin{aligned}
\log p_{i,n}(x)\approx &\mathbf{1}_{\{|x-X_{t_{i-1}^n}|\leqslant h_n^\rho\}}\log\left(p_{i,n}^c(x)\mathrm{e}^{-\lambda_0 h_n}\right)+\\
&\mathbf{1}_{\{|x-X_{t_{i-1}^n}|>h_n^\rho\}}\log\left(p_{i,n}^d(x)\lambda_0 h_n\mathrm{e}^{-\lambda_0 h_n}\right).
\end{aligned}
$$

By the way, in the expression

$$
\begin{aligned}
p_{i,n}^d(x)=&\int_z\int_{x'}\int_{t_{i-1}^n}^{t_i^n}\frac{1}{h_n}\phi\left(x';\, X_{t_{i-1}^n}+\bar{a}_{i-1}(s-t_{i-1}^n),\, \beta_{i-1}(s-t_{i-1}^n)\right)\times\\
&\times\phi\left(x;\, x'+c(x',z)+\bar{a}(x'+c(x',z))(t_i^n-s)\ ,\right.\\
&\left.\beta\ (x'+c(x',z))(t_i^n-s)\right)F_{\theta_0}(z)\,\mathrm{d}s\,\mathrm{d}x'\,\mathrm{d}z,
\end{aligned}
$$

we can approximate ϕ to δ-function if h_n decreases rapidly, and then

$$
\begin{aligned}
p_{i,n}^d(x)\approx&\int_z\int_{x'}\delta_{X_{t_{i-1}^n}}(x')\delta_{x-x'}(c(x',z))F_{\theta_0}(z)\,\mathrm{d}x'\,\mathrm{d}z\\
=&\int_y\int_{x'}\delta_{X_{t_{i-1}^n}}(x')\delta_{x-x'}(y)F_{\theta_0}\left(c^{-1}(x',y)\right)J(x',y,\theta_0)\,\mathrm{d}x'\,\mathrm{d}y\\
=&F_{\theta_0}\left(c^{-1}(X_{t_{i-1}^n},x-X_{t_{i-1}^n})\right)J(X_{t_{i-1}^n},x-X_{t_{i-1}^n},\theta_0)\\
=&\lambda_0^{-1}\Psi_{\theta_0}(x-X_{t_{i-1}^n},X_{t_{i-1}^n}).
\end{aligned}
$$

Moreover, since $\lambda(\theta)=\iint \Psi_\theta(y,x)\,\mathrm{d}y\,\mathrm{d}\pi$, $\lambda(\theta)$ can be approximated by the data as

$$\frac{1}{n}\sum_{i=1}^{n}\int \Psi_\theta(y, X_{t_{i-1}^n})\,\mathrm{d}y,$$

thanks to the ergodicity of X. This consideration leads the contrast function in Definition 2.1.

3. Moment Inequalities and Some Convergence Theorems

PROPOSITION 3.1. *Assume Conditions A.1, A.3 and A.6. For* $2\leqslant k$, $k\in\mathbf{N}$, $t_{i-1}^n\leqslant t\leqslant t_i^n$

$$E\left[|X_t - X_{t_{i-1}^n}|^k \middle| \mathcal{F}_{i-1}^n\right]\leqslant C_k|t-t_{i-1}^n|(1+|X_{t_{i-1}^n}|)^k. \tag{3.1}$$

If g is a function defined on $\mathbf{R}^d\times\Xi$ and is of polynomial growth in x uniformly in α for each component, then

$$E\left[|g(X_t,\alpha)||\mathcal{F}_{i-1}^n\right]\leqslant C(1+|X_{t_{i-1}^n}|)^C. \tag{3.2}$$

PROPOSITION 3.2. *Assume Conditions A.1, A.3 and A.4 – A.6. Let* $\bar{X}_{i,n}=X_{t_i^n}-X_{t_{i-1}^n}-h_n\bar{a}_{i-1}(\theta_0)$. *For all* $k_j=1,2,\ldots,d$ $(j=1,2,3,4)$,

$$E\left[\bar{X}_{i,n}^{(k_1)}\mathbf{1}_{C_{i,0}^n}|\mathcal{F}_{i-1}^n\right]=R(\alpha,h_n^2,X_{t_{i-1}^n}), \tag{3.3}$$

$$E\left[\bar{X}_{i,n}^{(k_1)}\bar{X}_{i,n}^{(k_2)}\mathbf{1}_{C_{i,0}^n}|\mathcal{F}_{i-1}^n\right]=h_n\mathrm{e}^{-\lambda_0 h_n}\beta_{i-1}^{(k_1,k_2)}(\sigma_0)+R(\alpha,h_n^2,X_{t_{i-1}^n}), \tag{3.4}$$

$$E\left[\bar{X}_{i,n}^{(k_1)}\bar{X}_{i,n}^{(k_2)}\bar{X}_{i,n}^{(k_3)}\mathbf{1}_{C_{i,0}^n}|\mathcal{F}_{i-1}^n\right]=R(\alpha,h_n^2,X_{t_{i-1}^n}), \tag{3.5}$$

$$\begin{aligned}&E\left[\bar{X}_{i,n}^{(k_1)}\bar{X}_{i,n}^{(k_2)}\bar{X}_{i,n}^{(k_3)}\bar{X}_{i,n}^{(k_4)}\mathbf{1}_{C_{i,0}^n}|\mathcal{F}_{i-1}^n\right]\\ &\quad=h_n^2\mathrm{e}^{-\lambda_0 h_n}\left(\beta_{i-1}^{(k_1,k_2)}\beta_{i-1}^{(k_3,k_4)}+\beta_{i-1}^{(k_1,k_3)}\beta_{i-1}^{(k_2,k_4)}+\beta_{i-1}^{(k_1,k_4)}\beta_{i-1}^{(k_2,k_3)}\right)+\\ &\qquad+R(\alpha,h_n^3,X_{t_{i-1}^n}).\end{aligned} \tag{3.6}$$

Remark 3.1. If we take the same argument as in the proof in Section 4.5, then we obtain the following moment inequality: Let $k\geqslant 1$, $t_{i-1}^n\leqslant t\leqslant t_i^n$. For any $p\geqslant 1$,

$$\begin{aligned}E\left[|X_t-X_{t_{i-1}^n}|^k\mathbf{1}_{C_{i,0}^n}\middle|\mathcal{F}_{i-1}^n\right]&\leqslant C_k|t-t_{i-1}^n|^{k/2}\mathrm{e}^{-\lambda_0 h_n}(1+|X_{t_{i-1}^n}|)^k+\\ &\quad+R(\alpha,h_n^p,X_{t_{i-1}^n}).\end{aligned} \tag{3.7}$$

PROPOSITION 3.3. *Assume Conditions A.1–A.3, A.5–A.7, $h_n \to 0$ and $nh_n \to \infty$. Suppose that $g^{(n)}$ is a function: $\mathbf{R}^d \times \Xi \to \mathbf{R}$ which satisfies the following conditions that*

$$|g^{(n)}(x,\alpha)|^4 \leqslant L(x,\alpha), \quad \left|\partial_x g^{(n)}(x,\alpha)\right| \leqslant O(\sqrt{b_n})(1+|x|)^C,$$
$$\left|\partial_\alpha g^{(n)}(x,\alpha)\right| \leqslant C(1+|x|)^C,$$

where L is a π-integrable function for all α, and that there exist a function g for each α such that

$$g^{(n)}(x,\alpha) \longrightarrow g(x,\alpha) \quad \pi\text{-a.s.} \quad (n\to\infty).$$

Then g is a π-integrable function and the following (i), (ii) *and* (iii) *hold:*

$$\text{(i)}\ \sup_{\alpha\in\Xi}\left|\frac{1}{n}\sum_{i=1}^n g_{i-1}^{(n)}(\alpha) - \int g(x,\alpha)\,\pi(dx)\right| \xrightarrow{P} 0 \quad (n\to\infty),$$
$$\text{(ii)}\ \sup_{\alpha\in\Xi}\left|\frac{1}{n}\sum_{i=1}^n g_{i-1}^{(n)}(\alpha)\mathbf{1}_{\{|\Delta X_i^n|\leqslant h_n^\rho\}} - \int g(x,\alpha)\,\pi(dx)\right| \xrightarrow{P} 0 \quad (n\to\infty),$$
$$\text{(iii)}\ \sup_{\alpha\in\Xi}\left|\frac{1}{nh_n}\sum_{i=1}^n g_{i-1}^{(n)}(\alpha)\mathbf{1}_{\{|\Delta X_i^n|> h_n^\rho\}} - \lambda_0\int g(x,\alpha)\,\pi(dx)\right| \xrightarrow{P} 0 \quad (n\to\infty).$$

Remark 3.2. Below, we sometimes use (i) as

$$g^{(n)}(x,\alpha) = \int \partial_\theta^k \Phi_n(\theta,x,y)\,\mathrm{d}y \quad (k=0,1,2).$$

We are able to check the above conditions for these $g^{(n)}$ by A.9 and A.11. Actually, for these $g^{(n)}$,

$$|g^{(n)}(x,\alpha)| \leqslant \left(\int L_1(y)\,\mathrm{d}y\right)(1+|x|)^C.$$

from condition (2.1), and it is obvious that $\left|\partial_\alpha g^{(n)}(x,\alpha)\right| \leqslant C(1+|x|)^C$ similarly. Moreover,

$$\begin{aligned}\left|\partial_x g^{(n)}(x,\alpha)\right| &\leqslant \int |\partial_x\partial_\theta^k\Psi_\theta(y,x)\cdot\varphi_n|\,\mathrm{d}y + \int |\partial_\theta^k\Psi_\theta(y,x)||\partial_x\varphi_n|\,\mathrm{d}y\\ &\leqslant \mathrm{O}(\varepsilon_n^{-1})(1+|x|)^C\\ &\leqslant \mathrm{O}(\sqrt{b_n})(1+|x|)^C,\end{aligned}$$

by (2.1), (2.2) and (2.5).

PROPOSITION 3.4. *Assume Conditions A.1–A.7, $h_n \to 0$ and $nh_n \to \infty$. Suppose that a function $g:\mathbf{R}^d \times \Xi \to \mathbf{R}$ and its derivatives $\partial_\alpha g$ and $\partial_x g$ are of polynomial growth uniformly in α. Then, for $k,l = 1,2,\ldots,d$,*

$$\sup_{\alpha\in\Xi}\left|\frac{1}{nh_n}\sum_{i=1}^{n} g_{i-1}(\alpha)\bar{X}_{i,n}^{(k)}\bar{X}_{i,n}^{(l)}\mathbf{1}_{\{|\Delta X_i^n|\leqslant h_n^\rho\}}- -\int g(x,\alpha)\beta^{(k,l)}(x,\sigma_0)\,\pi(dx)\right| \xrightarrow{P} 0 \quad (n\to\infty).$$

PROPOSITION 3.5. *Under the same assumptions as in Proposition 3.4, for* $k=1,2,\ldots,d$,

$$\sup_{\alpha\in\Xi}\left|\frac{1}{nh_n}\sum_{i=1}^{n} g_{i-1}(\alpha)\bar{X}_{i,n}^{(k)}\mathbf{1}_{\{|\Delta X_i^n|\leqslant h_n^\rho\}}\right| \xrightarrow{P} 0 \quad (n\to\infty).$$

PROPOSITION 3.6. *Assume Conditions A.1 – A.3, A.5 – A.7,* $h_n\to 0$ *and* $nh_n\to\infty$. *Suppose* $g_n(\alpha,y,x):\Xi\times\mathbf{R}^d\times\mathbf{R}^d\to\mathbf{R}$ *satisfies that, for* $y_1,y_2\in\mathbf{R}^d$ *and* $\eta\in[0,1]$,

$$|\partial_\alpha g_n(\alpha,y_1,x)-\partial_\alpha g_n(\alpha,y_2,x)|\leqslant \tilde{g}_n(\alpha,\eta y_1+(1-\eta)y_2,x)|y_1-y_2|, \tag{3.8}$$

where $\tilde{g}_n(\alpha,y,x)\leqslant \mathrm{O}(\sqrt{b_n})(1+|y|)^C(1+|x|)^C$, *and*

$$\left|\partial_\alpha^m g_n(\alpha,y,x)\right|\leqslant \mathrm{O}(\sqrt{b_n})(1+|y|)^C(1+|x|)^C \quad (m=0,1). \tag{3.9}$$

Assume that an integral $G_n(\alpha,x)=\int g_n(\alpha,y,x)\Psi_{\theta_0}(y,x)\,\mathrm{d}y$, *exists for all* x *and* α, *and that*

$$|G_n(\alpha,x)|^4\leqslant L(x,\alpha) \tag{3.10}$$

for a π*-integrable function* $L(x,\alpha)$,

$$|\partial_x G_n(\alpha,x)|\leqslant \mathrm{O}(\sqrt{b_n})(1+|x|)^C. \tag{3.11}$$

Moreover there exists a function g *such that*

$$G_n(\alpha,x)\longrightarrow \int g(\alpha,y,x)\Psi_{\theta_0}(y,x)\,\mathrm{d}y \quad \pi\text{-}a.s. \tag{3.12}$$

and the last integral is a π*-integrable function for all* α. *Furthermore, assume*

$$\int \sup_\alpha |\partial_\alpha g_n(\alpha,y,x)|\Psi_{\theta_0}(y,x)\,\mathrm{d}y\leqslant C(1+|x|)^C, \tag{3.13}$$

for all x. *Then*

$$\sup_{\alpha\in\Xi}\left|\frac{1}{nh_n}\sum_{i=1}^{n} g_n(\alpha,\Delta X_i^n,X_{t_{i-1}^n})\mathbf{1}_{\{|\Delta X_i^n|>h_n^\rho\}}- -\iint g(\alpha,y,x)\Psi_{\theta_0}(y,x)\,\mathrm{d}y\,\mathrm{d}\pi(x)\right| \xrightarrow{P} 0 \quad (n\to 0).$$

Remark 3.3. Condition (3.8) is satisfied if $\partial_\alpha g_n(\alpha, y, x)$ is differentiable with respect to y and

$$\left|\partial_y \partial_\alpha g_n(\alpha, y, x)\right| \leqslant \mathrm{O}(\sqrt{b_n})(1+|y|)^C(1+|x|)^C. \tag{3.14}$$

Remark 3.4. Thanks to Conditions (3.8)–(3.13), we are able to apply Proposition 3.3 (i) to $G_n(\alpha, x)$.

Remark 3.5. Below, we use this proposition as

$$g_n(\alpha, y, x) = \partial_\theta^k \log \Phi_n(\theta, x, y) \cdot \varphi_n(x, y) \quad (k=0,1,2).$$

We are able to check the above conditions for this g_n by A.9. Indeed,

$$\begin{aligned}\left|\partial_y g_n(\alpha, y, x)\right| &\leqslant |\partial_y \partial_\theta^k \log \Phi_n||\varphi_n| + |\partial_y \varphi_n||\partial_\theta^k \log \Phi_n| \\ &\leqslant \mathrm{O}(\varepsilon_n^{-2})(1+|y|)^C(1+|x|)^C + \mathrm{O}(\varepsilon_n^{-3})(1+|y|)^C(1+|x|)^C \\ &\leqslant \mathrm{O}(\sqrt{b_n})(1+|y|)^C(1+|x|)^C.\end{aligned}$$

from (2.1), (2.3), (2.5) and (2.6). Similarly,

$$\begin{aligned}\left|\partial_y \partial_\alpha g_n(\alpha, y, x)\right| &= \left|\partial_y \left[\partial_\theta^{k+1} \log \Psi_\theta(y, x) \cdot \varphi_n(x, y)\right]\right| \\ &\leqslant \mathrm{O}(\varepsilon_n^{-4})(1+|y|)^C(1+|x|)^C \\ &\leqslant \mathrm{O}(\sqrt{b_n})(1+|y|)^C(1+|x|)^C.\end{aligned}$$

Therefore (3.14) is satisfied, and this implies (3.8).

On Inequality (3.9),

$$\begin{aligned}\left|\partial_\alpha^m g_n(\alpha, y, x)\right| &= \left|\partial_\theta^{k+m} \log \Psi_\theta(y, x) \cdot \varphi_n(x, y)\right| \\ &\leqslant \mathrm{O}(\varepsilon_n^{-2-m})(1+|y|)^C(1+|x|)^C \\ &\leqslant \mathrm{O}(\sqrt{b_n})(1+|y|)^C(1+|x|)^C.\end{aligned}$$

On Inequality (3.10),

$$|g_n(\alpha, y, x)\Psi_{\theta_0}(y, x)| \leqslant \{|\partial_\theta^k \log \Psi_\theta(y, x)| + 1\}\Psi_{\theta_0}(y, x), \tag{3.15}$$

therefore (2.1) and (2.4) yield (3.10).

Next,

$$\begin{aligned}|\partial_x G_n(\alpha, x)| &\leqslant \int |\partial_x(\partial_\theta^k \log \Phi_n \cdot \varphi_n)\Psi_{\theta_0}|\,\mathrm{d}y + \int |\partial_\theta^k \log \Phi_n \cdot \varphi_n||\partial_x \Psi_{\theta_0}|\,\mathrm{d}y \\ &\leqslant \int |\partial_x \varphi_n \partial_\theta^k \log \Phi_n||\Psi_{\theta_0}|\,\mathrm{d}y + \int |\varphi_n \partial_x \partial_\theta^k \log \Phi_n||\Psi_{\theta_0}|\,\mathrm{d}y + \\ &\quad + \int |\partial_\theta^k \log \Phi_n \cdot \varphi_n||\partial_x \Psi_{\theta_0}|\,\mathrm{d}y.\end{aligned}$$

Inequalities (2.1), (2.2), (2.5) and (2.6) imply that

$$|\partial_x \varphi_n \partial_\theta^k \log \Phi_n| \leqslant \mathrm{O}(\varepsilon_n^{-3}) L(y)(1+|x|)^C,$$

$$|\varphi_n \partial_x \partial_\theta^k \log \Phi_n| \leqslant \mathrm{O}(\varepsilon_n^{-3}) L(y)(1+|x|)^C$$

and

$$|\partial_\theta^k \log \Phi_n \cdot \varphi_n| \leqslant \mathrm{O}(\varepsilon_n^{-2}) L(y)(1+|x|)^C,$$

where $L(y)$ is a bounded $\mathrm{d}y$-integrable function. Then

$$\begin{aligned}|\partial_x G_n(\alpha, x)| \leqslant \int \mathrm{O}(\varepsilon_n^{-3}) L(y)(1+|x|)^C \Psi_{\theta_0}\, \mathrm{d}y + \\ + \int \mathrm{O}(\varepsilon_n^{-2}) L_2(y) L(y)(1+|x|)^C\, \mathrm{d}y \\ \leqslant \mathrm{O}(\sqrt{b_n})(1+|x|)^C.\end{aligned}$$

Therefore (3.11) is satisfied. The condition (3.12) is obtained from Lebesgue's theorem thanks to (3.15).

Inequality (3.13) is obtained from (2.4) directly with $g(\alpha, x, y) = \partial_\theta^k \log \Phi_\theta(y, x)$.

4. Proofs

In this section, we sometimes omit the true values of the parameters without specially mentioning. For example, we simply write $f(X_{t_{i-1}^n})$ for $f(X_{t_{i-1}^n}, \alpha_0)$ or $q(\mathrm{d}s, \mathrm{d}z)$ for $q^{\theta_0}(\mathrm{d}s, \mathrm{d}z)$ and so on.

4.1. PROOF OF LEMMA 2.1

First we show (2.9) On the interval $[t_{i-1}^n, \tau_i^n)$ X follows a stochastic differential equation

$$dX_t = \bar{a}(X_t)\, \mathrm{d}t + b(X_t)\, \mathrm{d}W_t,$$

hence for $t \in [t_{i-1}^n, \tau_i^n)$

$$\begin{aligned}|X_t - X_{t_{i-1}^n}| &= \left|(t - t_{i-1}^n)\bar{a}(X_{t_{i-1}^n}) + \int_{t_{i-1}^n}^t (\bar{a}(X_s) - \bar{a}(X_{t_{i-1}^n}))\, \mathrm{d}s + \int_{t_{i-1}^n}^t b(X_s)\, \mathrm{d}W_s\right| \\ &\leqslant h_n|\bar{a}(X_{t_{i-1}^n})| + \sup_{u \in [t_{i-1}^n, t_i^n)} \left|\int_{t_{i-1}^n}^u b(X_s)\, \mathrm{d}W_s\right| + L\int_{t_{i-1}^n}^t |X_s - X_{t_{i-1}^n}|\, \mathrm{d}s.\end{aligned}$$

Applying Gronwall inequality,

$$|X_t - X_{t_{i-1}^n}| \leqslant h_n|\bar{a}(X_{t_{i-1}^n})| + \sup_{u\in[t_{i-1}^n, t_i^n)} \left| \int_{t_{i-1}^n}^u b(X_s)\,\mathrm{d}W_s \right| + \\ + Le^{Lh_n}h_n \left(h_n|\bar{a}(X_{t_{i-1}^n})| + \sup_{u\in[t_{i-1}^n, t_i^n)} \left| \int_{t_{i-1}^n}^u b(X_s)\,\mathrm{d}W_s \right| \right).$$

If n is sufficiently large

$$|X_t - X_{t_{i-1}^n}| \leqslant C \left(h_n|\bar{a}(X_{t_{i-1}^n})| + \sup_{u\in[t_{i-1}^n, t_i^n)} \left| \int_{t_{i-1}^n}^u b(X_s)\,\mathrm{d}W_s \right| \right), \tag{4.1}$$

therefore Markov's inequality and Burkholder-Davis-Gundy's inequality yield

$$\begin{aligned} &P\left\{ \sup_{t\in[t_{i-1}^n, \tau_i^n)} |X_t - X_{t_{i-1}^n}| > h_n^\rho \middle| \mathcal{F}_{i-1}^n \right\} \\ &\leqslant P\left\{ Kh_n|\bar{a}(X_{t_{i-1}^n})| > \frac{h_n^\rho}{2} \middle| \mathcal{F}_{i-1}^n \right\} + P\left\{ K \sup_{u\in[t_{i-1}^n, t_i^n)} \left| \int_{t_{i-1}^n}^u b(X_s)\,\mathrm{d}W_s \right| > \frac{h_n^\rho}{2} \middle| \mathcal{F}_{i-1}^n \right\} \\ &\leqslant C_p \left\{ h_n^{p(1-\rho)} E\left[|\bar{a}(X_{t_{i-1}^n})|^p \middle| \mathcal{F}_{i-1}^n \right] + h_n^{-2p\rho} E\left[\sup_{u\in[t_{i-1}^n, t_i^n)} \left| \int_{t_{i-1}^n}^u b(X_s)\,\mathrm{d}W_s \right|^{2p} \middle| \mathcal{F}_{i-1}^n \right] \right\} \\ &\leqslant R(\alpha, h_n^{p(1-\rho)}, X_{t_{i-1}^n}) + C_p h_n^{-2p\rho} E\left[\left| \int_{t_{i-1}^n}^{t_i^n} b^2(X_s)\,\mathrm{d}s \right|^p \middle| \mathcal{F}_{i-1}^n \right] \\ &= R(\alpha, h_n^{p(1-2\rho)}, X_{t_{i-1}^n}). \end{aligned}$$

We used Proposition 3.1 (3.2) in the last equality. We should notice that one can take p arbitrary larger here.

The almost same argument holds for (2.10) Actually for $t \in [\eta_i^n, t_i^n)$

$$|X_{t_i^n} - X_t| = \left| (t_i^n - t)\bar{a}(X_t) + \int_t^{t_i^n} (\bar{a}(X_s) - \bar{a}(X_t))\,\mathrm{d}s + \int_t^{t_i^n} b(X_s)\,\mathrm{d}W_s \right|.$$

By the same argument as (2.9)

$$|X_{t_i^n} - X_t| \leqslant C \left(h_n|\bar{a}(X_{t_i^n})| + \sup_{u\in[t_{i-1}^n, t_i^n)} \left| \int_{t_{i-1}^n}^u b(X_s)\,\mathrm{d}W_s \right| \right), \tag{4.2}$$

and then,

$$\begin{aligned} &P\left\{ \sup_{t\in[\eta_i^n, t_i^n)} |X_{t_i^n} - X_t| > h_n^\rho \middle| \mathcal{F}_{i-1}^n \right\} \\ &\qquad \leqslant C_p \left\{ h_n^{p(1-\rho)} E\left[|\bar{a}(X_{t_i^n})|^p \middle| \mathcal{F}_{i-1}^n \right] + h_n^{-2p\rho} E\left[\left| \int_{t_{i-1}^n}^{t_i^n} b^2(X_s)\,\mathrm{d}s \right|^p \middle| \mathcal{F}_{i-1}^n \right] \right\}. \end{aligned}$$

Proposition 3.1 (3.2) completes the proof. □

4.2. PROOF OF LEMMA 2.2

It is obvious that $P\{C^n_{i,2}|\mathcal{F}^n_{i-1}\} \leqslant h_n^2\lambda_0^2$, and so is $P\{D^n_{i,2}|\mathcal{F}^n_{i-1}\}$. On $C^n_{i,1}$,

$$
\begin{aligned}
P\{C^n_{i,1}|\mathcal{F}^n_{i-1}\} \leqslant \Bigg[P\Bigg\{ &\left|(X_{t_i^n} - X_{\tau_i^n}) + (X_{\tau_i^n-} - X_{t_{i-1}^n}) + \Delta X_{\tau_i^n}\right| \\
&\leqslant h_n^\rho, \left|\Delta Z_{\tau_i^n}\right| > \frac{2h_n^\rho}{c_0} \Big| \mathcal{F}^n_{i-1}, \;\; J_i^n = 1 \Bigg\} + \\
&+ P\left\{ \left|\Delta Z_{\tau_i^n}\right| \leqslant \frac{2h_n^\rho}{c_0} \Big| \mathcal{F}^n_{i-1}, \;\; J_i^n = 1 \right\} \Bigg] P\{J_i^n = 1\},
\end{aligned}
$$

where $\Delta Z_{\tau_i^n}$ has density F_{θ_0}. If $\left|(X_{t_i^n} - X_{\tau_i^n}) + (X_{\tau_i^n-} - X_{t_{i-1}^n}) + \Delta X_{\tau_i^n}\right| \leqslant h_n^\rho$ then

$$
|X_{t_i^n} - X_{\tau_i^n}| + |X_{\tau_i^n-} - X_{t_{i-1}^n}| \geqslant c_0 \left|\Delta Z_{\tau_i^n}\right| - h_n^\rho,
$$

hence by applying Lemma 2.1,

$$
\begin{aligned}
P\{C^n_{i,1}|\mathcal{F}^n_{i-1}\} &\leqslant \lambda_0 h_n \mathrm{e}^{-\lambda_0 h_n} P\left\{ \sup_{t\in[t_{i-1}^n,\tau_i^n)} |X_t - X_{t_{i-1}^n}| + \sup_{t\in[\tau_i^n,t_i^n]} |X_{t_i^n} - X_t| > h_n^\rho \Big| \mathcal{F}^n_{i-1} \right\} + \\
&\quad + \lambda_0 h_n \mathrm{e}^{-\lambda_0 h_n} \int_{-2h_n^\rho/c_0}^{2h_n^\rho/c_0} M|z|^\gamma \,\mathrm{d}z \\
&= R(\alpha, h_n^p, X_{t_{i-1}^n}) + C h_n^{\rho(\gamma+1)+1} \\
&= R(\alpha, h_n^3, X_{t_{i-1}^n}).
\end{aligned}
$$

On $C^n_{i,0}$, again by applying Lemma 2.1

$$
\begin{aligned}
P\{C^n_{i,0}|\mathcal{F}^n_{i-1}\} &= P\{J_i^n = 0|\mathcal{F}^n_{i-1}\} - P\left\{|\Delta X_i^n| > h_n^\rho, J_i^n = 0|\mathcal{F}^n_{i-1}\right\} \\
&= \mathrm{e}^{-\lambda_0 h_n} - P\left\{|X_{\tau_i^n} - X_{t_{i-1}^n}| > h_n^\rho, \tau_i^n = t_i^n|\mathcal{F}^n_{i-1}\right\} \\
&= \mathrm{e}^{-\lambda_0 h_n} \tilde{R}(\alpha, h_n^p, X_{t_{i-1}^n}).
\end{aligned}
$$

Finally,

$$
\begin{aligned}
P\{D^n_{i,0}|\mathcal{F}^n_{i-1}\} &= P\{J_i^n = 0|\mathcal{F}^n_{i-1}\} - P\left\{C^n_{i,0}|\mathcal{F}^n_{i-1}\right\} \\
&= \mathrm{e}^{-\lambda_0 h_n} R(\alpha, h_n^p, X_{t_{i-1}^n}),
\end{aligned}
$$

and

$$
\begin{aligned}
P\{D^n_{i,1}|\mathcal{F}^n_{i-1}\} &= P\{J_i^n = 1|\mathcal{F}^n_{i-1}\} - P\left\{|\Delta X_i^n| \leqslant h_n^\rho, J_i^n = 1|\mathcal{F}^n_{i-1}\right\} \\
&= \lambda_0 h_n \;\mathrm{e}^{-\lambda_0 h_n} - R(\alpha, h_n^3, X_{t_{i-1}^n}) \\
&= \lambda_0 h_n \;\mathrm{e}^{-\lambda_0 h_n} \tilde{R}(\alpha, h_n^2, X_{t_{i-1}^n}).
\end{aligned}
$$

□

4.3. PROOF OF PROPOSITION 3.1

First, we consider the case $p=2^q$, $q\in\mathbf{N}$ From the expression of the stochastic differential equation, we get

$$|X_t - X_{t_{i-1}^n}|^p \leqslant 3^{p-1}\left\{ \left|\int_{t_{i-1}n}^t a(X_s)\,\mathrm{d}s\right|^p + \left|\int_{t_{i-1}^n}^t b(X_s)\,\mathrm{d}W_s\right|^p + \right.$$
$$\left. + \left|\int_{t_{i-1}^n}^t \int c(X_{s-},z)\,(p-q)(\mathrm{d}s,\mathrm{d}z)\right|^p \right\}. \tag{4.3}$$

We denote by H_t the first and second term of (4.3), and we also denote

$$M_t = \int_{t_{i-1}^n}^t \int c(X_{s-},z)\,(p-q)(\mathrm{d}s,\mathrm{d}z),$$
$$N_t = \int_{t_{i-1}^n}^t \int c^2(X_{s-},z)\,(p-q)(\mathrm{d}s,\mathrm{d}z).$$

By using the linear growthness of a,b and Burkholder–Davis–Gundy inequality, we easily obtain

$$E[H_t|\mathcal{F}_{i-1}^n] \leqslant C_p|t-t_{i-1}^n|^{p/2}(1+|X_{t_{i-1}^n}|)^p + C_p\int_{t_{i-1}^n}^t E\left[|X_s - X_{t_{i-1}^n}|^p\middle|\mathcal{F}_{i-1}^n\right]\mathrm{d}s. \tag{4.4}$$

Applying the Lemma 4.1 below and the linear growth of $c(x,z)$, we obtain

$$E\left[|M_t|^p|\mathcal{F}_{i-1}^n\right] \leqslant C_p|t-t_{i-1}^n|(1+|X_{t_{i-1}^n}|)^p + C_p\int_{t_{i-1}^n}^t E\left[|X_s - X_{t_{i-1}^n}|^p\middle|\mathcal{F}_{i-1}^n\right]\mathrm{d}s. \tag{4.5}$$

Inequalities (4.4) and (4.5) and the Gronwall inequality yield for all $q\in\mathbf{N}$

$$E\left[|X_t - X_{t_{i-1}^n}|^p\middle|\mathcal{F}_{i-1}^n\right] \leqslant C_p|t-t_{i-1}^n|(1+|X_{t_{i-1}^n}|)^p.$$

For arbitrary $k\geqslant 2$, if we write $k=\sum_{q=0}^l \delta_q 2^q$ by binary expansion, where $\delta_q=0$ or 1 then

$$E\left[|X_t - X_{t_{i-1}^n}|^k\middle|\mathcal{F}_{i-1}^n\right] = E\left[\prod_{q=0}^l |X_t - X_{t_{i-1}^n}|^{\delta_q 2^q}\middle|\mathcal{F}_{i-1}^n\right].$$

Therefore we obtain Inequality (3.1) by using Cauchy–Schwarz inequality repeatedly.

On (3.2), we can write

$$E\left[|g(X_t,\alpha)||\mathcal{F}^n_{i-1}\right]\leqslant CE\left[(1+|X_t|)^C|\mathcal{F}^n_{i-1}\right]$$
$$\leqslant C\left(1+|X_{t^n_{i-1}}|^C+E\left[|X_t-X_{t^n_{i-1}}|^C|\mathcal{F}^n_{i-1}\right]\right),$$

and (3.1) ends the proof □

4.4. LEMMA 4.1 AND ITS PROOF

LEMMA 4.1. *Assume Conditions A.1, A.3 and A.6. For* $p=2^q$, $q\in\mathbf{N}$, $t^n_{i-1}\leqslant t\leqslant t^n_i$

$$E\left[|M_t|^p|\mathcal{F}^n_{i-1}\right]\leqslant C_pE\left[\iint_{t^n_{i-1}}^t c^p(X_s,z)\,q(\mathrm{d}s,\mathrm{d}z)\Big|\mathcal{F}^n_{i-1}\right].$$

Proof. This proof follows from Bichteler and Jacod [5]. By using Doob's inequality

$$E\left[\sup_{t^n_{i-1}\leqslant s\leqslant t}|M_s|^2|\mathcal{F}^n_{i-1}\right]\leqslant 4E\left[M_t^2|\mathcal{F}^n_{i-1}\right]\leqslant 4E\left[\langle M,M\rangle_t|\mathcal{F}^n_{i-1}\right]$$
$$\leqslant 4E\left[\iint_{t^n_{i-1}}^t c^2(X_s,z)\,q(\mathrm{d}s,\mathrm{d}z)\Big|\mathcal{F}^n_{i-1}\right],$$

hence Lemma 4.1 holds for $q=1$. Next we suppose that Lemma holds for $p=2^q$. We notice that $[M,M]_t=N_t+\langle M,M\rangle_t$ for all $p\in\mathbf{N}$ then

$$[M,M]_t^p\leqslant 2^{p-1}(N_t^p+\langle M,M\rangle_t^p).$$

Using the Burkholder–Davis–Gundy inequality and the above inequality

$$E\left[|M_t|^{2p}|\mathcal{F}^n_{i-1}\right]$$
$$\leqslant C_pE\left[|N_t|^p+\langle M,M\rangle_t^p|\mathcal{F}^n_{i-1}\right]$$
$$\leqslant C_pE\left[\iint_{t^n_{i-1}}^t c^{2p}(X_s,z)\,q(\mathrm{d}s,\mathrm{d}z)+\left(\iint_{t^n_{i-1}}^t c^2(X_s,z)\,q(\mathrm{d}s,\mathrm{d}z)\right)^p\Big|\mathcal{F}^n_{i-1}\right].$$

Applying Jensen's inequality to the last second term, we have

$$E\left[|M_t|^{2p}|\mathcal{F}^n_{i-1}\right]\leqslant C_pE\left[\iint_{t^n_{i-1}}^t c^{2p}(X_s,z)\,q(\mathrm{d}s,\mathrm{d}z)\Big|\mathcal{F}^n_{i-1}\right].$$

Therefore Lemma holds for $p=2^{q+1}$.

4.5. PROOF OF PROPOSITION 3.2

We prove (3.6); the others are done similarly

Let Y be a solution to the following stochastic differential equation

$$dY_t = \bar{a}(Y_t)\,\mathrm{d}t + b(Y_t)\,\mathrm{d}W_t$$

independent of J_i^n. Simple calculation deduce the multidimensional case of Lemma 7 in Kessler [19], that is,

$$\begin{aligned} &E\left[\bar{Y}_{i,n}^{(k_1)}\bar{Y}_{i,n}^{(k_2)}\bar{Y}_{i,n}^{(k_3)}\bar{Y}_{i,n}^{(k_4)}|\mathcal{F}_{i-1}^n\right] \\ &\quad = h_n^2\left(\beta_{i-1}^{(k_1,k_2)}\beta_{i-1}^{(k_3,k_4)} + \beta_{i-1}^{(k_1,k_3)}\beta_{i-1}^{(k_2,k_4)} + \beta_{i-1}^{(k_1,k_4)}\beta_{i-1}^{(k_2,k_3)}\right) + R(\alpha, h_n^3, X_{t_{i-1}^n}). \end{aligned} \tag{4.6}$$

Since J_i^n is independent of $\mathcal{F}_{i-1}^n$, we have

$$\begin{aligned} &E\left[\bar{X}_{i,n}^{(k_1)}\bar{X}_{i,n}^{(k_2)}\bar{X}_{i,n}^{(k_3)}\bar{X}_{i,n}^{(k_4)}\mathbf{1}_{\{J_i^n=0\}}|\mathcal{F}_{i-1}^n\right] \\ &\quad = E\left[\bar{Y}_{i,n}^{(k_1)}\bar{Y}_{i,n}^{(k_2)}\bar{Y}_{i,n}^{(k_3)}\bar{Y}_{i,n}^{(k_4)}\mathbf{1}_{\{J_i^n=0\}}|\mathcal{F}_{i-1}^n\right] \\ &\quad = E\left[\bar{Y}_{i,n}^{(k_1)}\bar{Y}_{i,n}^{(k_2)}\bar{Y}_{i,n}^{(k_3)}\bar{Y}_{i,n}^{(k_4)}|\mathcal{F}_{i-1}^n\right] P_0\left\{J_i^n = 0\right\} \\ &\quad = h_n^2 e^{-\lambda_0 h_n}\left(\beta_{i-1}^{(k_1,k_2)}\beta_{i-1}^{(k_3,k_4)} + \beta_{i-1}^{(k_1,k_3)}\beta_{i-1}^{(k_2,k_4)} + \beta_{i-1}^{(k_1,k_4)}\beta_{i-1}^{(k_2,k_3)}\right) + \\ &\qquad + R(\alpha, h_n^3, X_{t_{i-1}^n}). \end{aligned}$$

On the other hand,

$$\begin{aligned} &E\left[\bar{X}_{i,n}^{(k_1)}\bar{X}_{i,n}^{(k_2)}\bar{X}_{i,n}^{(k_3)}\bar{X}_{i,n}^{(k_4)}\mathbf{1}_{\{J_i^n=0\}}|\mathcal{F}_{i-1}^n\right] \\ &\quad = E\left[\bar{X}_{i,n}^{(k_1)}\bar{X}_{i,n}^{(k_2)}\bar{X}_{i,n}^{(k_3)}\bar{X}_{i,n}^{(k_4)}\mathbf{1}_{C_{i,0}^n}|\mathcal{F}_{i-1}^n\right] + E\left[\bar{X}_{i,n}^{(k_1)}\bar{X}_{i,n}^{(k_2)}\bar{X}_{i,n}^{(k_3)}\bar{X}_{i,n}^{(k_4)}\mathbf{1}_{D_{i,0}^n}|\mathcal{F}_{i-1}^n\right]. \end{aligned}$$

According to (3.1) and Lemma 2.2

$$\begin{aligned} &\left|E\left[\bar{X}_{i,n}^{(k_1)}\bar{X}_{i,n}^{(k_2)}\bar{X}_{i,n}^{(k_3)}\bar{X}_{i,n}^{(k_4)}\mathbf{1}_{D_{i,0}^n}|\mathcal{F}_{i-1}^n\right]\right| \\ &\quad \leqslant \sqrt{E\left[(\bar{X}_{i,n}^{(k_1)}\bar{X}_{i,n}^{(k_2)}\bar{X}_{i,n}^{(k_3)}\bar{X}_{i,n}^{(k_4)})^2|\mathcal{F}_{i-1}^n\right] P\{D_{i,0}^n|\mathcal{F}_{i-1}^n\}} \\ &\quad \leqslant \sqrt{E\left[C\left(|\Delta X_i^n|^8 + |\bar{a}_{i-1}h_n|^8\right)|\mathcal{F}_{i-1}^n\right] P\{D_{i,0}^n|\mathcal{F}_{i-1}^n\}} \\ &\quad \leqslant R(\alpha, h_n^p, X_{t_{i-1}^n}). \end{aligned}$$

This completes the proof. □

4.6. PROOF OF PROPOSITION 3.3

The π-integrability of $g(x)$ is led from the uniform integrability of $g^{(n)}(x,\alpha)$.

Let us prove that each convergence holds for fixed α. We start with the proof of (i)

$$
\begin{aligned}
&P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} g_{i-1}^{(n)}(\alpha)-\int g(x,\alpha)\,\pi(\mathrm{d}x)\right|>\epsilon\right\}\\
&\quad\leqslant P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} g_{i-1}^{(n)}(\alpha)-\frac{1}{nh_n}\int_0^{nh_n} g^{(n)}(X_s,\alpha)\,\mathrm{d}s\right|>\frac{\epsilon}{3}\right\}+\\
&\qquad+P\left\{\left|\frac{1}{nh_n}\int_0^{nh_n} g^{(n)}(X_s,\alpha)\,\mathrm{d}s-\frac{1}{nh_n}\int_0^{nh_n} g(X_s,\alpha)\,\mathrm{d}s\right|>\frac{\epsilon}{3}\right\}+\\
&\qquad+P\left\{\left|\frac{1}{nh_n}\int_0^{nh_n} g(X_s,\alpha)\,\mathrm{d}s-\int g(x,\alpha)\pi(\mathrm{d}x)\right|>\frac{\epsilon}{3}\right\}.
\end{aligned}
$$

This third term on the right-hand side converges to zero by the assumption of ergodicity. Let us call the first and second terms P_n^1 and P_n^2, respectively, then

$$
\begin{aligned}
P_n^1&\leqslant\frac{3}{\epsilon}E\left[\left|\frac{1}{n}\sum_{i=1}^{n} g_{i-1}^{(n)}(\alpha)-\frac{1}{nh_n}\int_0^{nh_n} g^{(n)}(X_s,\alpha)\mathrm{d}s\right|\right]\\
&\leqslant\frac{3}{\epsilon}E\left[\frac{1}{nh_n}\sum_{i=1}^{n}\int_{t_{i-1}^n}^{t_i^n}|g^{(n)}(X_s,\alpha)-g_{i-1}^{(n)}(\alpha)|\,\mathrm{d}s\right].
\end{aligned}
$$

Applying Taylor's formula and Schwarz' inequality,

$$
\begin{aligned}
P_n^1\leqslant&\frac{3}{nh_n\epsilon}\sum_{i=1}^{n}\int_{t_{i-1}^n}^{t_i^n}\left(E\left[|X_s-X_{t_{i-1}^n}|^2\right]\right)^{1/2}\times\\
&\times\left(E\left[\left(\int_0^1\partial_x g^{(n)}(X_{t_{i-1}^n}+u(X_s-X_{t_{i-1}^n}),\alpha)\,\mathrm{d}u\right)^2\right]\right)^{1/2}\mathrm{d}s.
\end{aligned}
$$

The inequalities (3.1) and (3.2) of Proposition 3.1 yield

$$
\begin{aligned}
P_n^1\leqslant&\frac{3}{nh_n\epsilon}\sum_{i=1}^{n}\int_{t_{i-1}^n}^{t_i^n}\left(E\left[C|s-t_{i-1}^n|(1+|X_{t_{i-1}^n}|)^C\right]\right)^{1/2}\times\\
&\times\left(E\left[O(b_n)(1+|X_{t_{i-1}^n}|^C)\right]\right)^{1/2}\mathrm{d}s\\
\leqslant&\frac{\mathrm{O}(\sqrt{b_n})}{nh_n\epsilon}\sum_{i=1}^{n}\left(\int_{t_{i-1}^n}^{t_i^n}|s-t_{i-1}^n|^{1/2}\,\mathrm{d}s\right)\\
\leqslant&\,\mathrm{O}(\sqrt{h_nb_n}).
\end{aligned}
$$

Moreover

$$P_n^2 \leqslant \frac{3}{\epsilon n h_n} \int_0^{nh_n} E|g^{(n)}(X_t, \alpha) - g(X_t, \alpha)|\, \mathrm{d}t$$
$$= \frac{3}{\epsilon} \int |g^{(n)}(x, \alpha) - g(x, \alpha)|\, \mathrm{d}\pi.$$

This converges to zero by Lebesgue's convergence theorem.

Next we show (iii) What we should show are the following (a) and (b) thanks to Lemma 9 in Genon-Catalot and Jacod [10]:

$$\text{(a)} \ \sum_{i=1}^n E\left[\frac{1}{nh_n} g_{i-1}^{(n)}(\alpha) \mathbf{1}_{\{|\Delta X_i^n| > h_n^\rho\}} \middle| \mathcal{F}_{i-1}^n\right] \xrightarrow{P} \lambda_0 \int g(x, \alpha)\, \mathrm{d}\pi(x),$$
$$\text{(b)} \ \sum_{i=1}^n E\left[\frac{1}{n^2 h_n^2} \left(g_{i-1}^{(n)}(\alpha)\right)^2 \mathbf{1}_{\{|\Delta X_i^n| > h_n^\rho\}} \middle| \mathcal{F}_{i-1}^n\right] \xrightarrow{P} 0.$$

(a) By the same argument as for (i), it is sufficient to show that

$$I_n := P\left\{\left|\frac{1}{nh_n} \sum_{i=1}^n g_{i-1}^{(n)}(\alpha) P\left\{|\Delta X_i^n| > h_n^\rho \middle| \mathcal{F}_{i-1}^n\right\} - \right.\right.$$
$$\left.\left. - \lambda_0 \frac{1}{nh_n} \int_0^{nh_n} g^{(n)}(X_s, \alpha)\, \mathrm{d}s\right| > \epsilon\right\}$$
$$\longrightarrow 0.$$

This is easily seen as follows:

$$I_n \leqslant \frac{1}{\epsilon} E\left[\frac{1}{nh_n} \sum_{i=1}^n \int_{t_{i-1}^n}^{t_i^n} \left| g_{i-1}^{(n)}(\alpha) \frac{1}{h_n} P\left\{|\Delta X_i^n| > h_n^\rho \middle| \mathcal{F}_{i-1}^n\right\} - \lambda_0 g^{(n)}(X_s, \alpha)\right| \mathrm{d}s\right]$$
$$\leqslant \frac{1}{nh_n \epsilon} \sum_{i=1}^n \int_{t_{i-1}^n}^{t_i^n} \left[E\left| g_{i-1}^{(n)}(\alpha) \left(\frac{1}{h_n} P\left\{|\Delta X_i^n| > h_n^\rho \middle| \mathcal{F}_{i-1}^n\right\} - \lambda_0\right)\right| + \right.$$
$$\left. + \lambda_0 E|g_{i-1}^{(n)}(\alpha) - g^{(n)}(X_s, \alpha)|\right] \mathrm{d}s$$
$$\leqslant \frac{1}{nh_n \epsilon} \sum_{i=1}^n \int_{t_{i-1}^n}^{t_i^n} \left[\sqrt{E|g_{i-1}^{(n)}(\alpha)|^2} \sqrt{E\left|\frac{1}{h_n} P\left\{|\Delta X_i^n| > h_n^\rho \middle| \mathcal{F}_{i-1}^n\right\} - \lambda_0\right|^2} + \right.$$
$$\left. + \lambda_0 E|g_{i-1}^{(n)}(\alpha) - g^{(n)}(X_s, \alpha)|\right] \mathrm{d}s = \mathrm{O}(\sqrt{h_n b_n}).$$

Here we applied Lemma 2.2 to the term $P\left\{|\Delta X_i^n| > h_n^\rho \middle| \mathcal{F}_{i-1}^n\right\}$ and the same argument as the proof of (i) to the term $E|g_{i-1}^{(n)}(\alpha) - g^{(n)}(X_s, \alpha)|$.

(b)

$$
\begin{aligned}
&P\left\{\left|\frac{1}{n^2h_n^2}\sum_{i=1}^n\left(g_{i-1}^{(n)}(\alpha)\right)^2P\left\{|\Delta X_i^n|>h_n^\rho\Big|\mathcal{F}_{i-1}^n\right\}\right|>\epsilon\right\}\\
&\leqslant\frac{1}{n^2h_n^2\epsilon}\sum_{i=1}^n E\left|\left(g_{i-1}^{(n)}(\alpha)\right)^2P\left\{|\Delta X_i^n|>h_n^\rho\Big|\mathcal{F}_{i-1}^n\right\}\right|\\
&\leqslant\frac{1}{n^2h_n\epsilon}\sum_{i=1}^n\sqrt{E\left|g_{i-1}^{(n)}(\alpha)\right|^4E\left|\frac{1}{h_n}P\left\{|\Delta X_i^n|>h_n^\rho\Big|\mathcal{F}_{i-1}^n\right\}\right|^2}\\
&=\mathrm{O}\left(\frac{1}{nh_n}\right).
\end{aligned}
$$

We can easily deduce (ii) for each fixed α from (i) and (iii) since

$$
\frac{1}{n}\sum_{i=1}^n g_{i-1}^{(n)}(\alpha)\mathbf{1}_{\{|\Delta X_i^n|\leqslant h_n^\rho\}}=\frac{1}{n}\sum_{i=1}^n g_{i-1}^{(n)}(\alpha)-h_n\left(\frac{1}{nh_n}\sum_{i=1}^n g_{i-1}^{(n)}(\alpha)\mathbf{1}_{\{|\Delta X_i^n|>h_n^\rho\}}\right).
$$

Finally we have to show the uniformity of the convergence in α. We only show (i); the uniformity in (ii) can be proved similarly and that in (iii) is shown by the same argument as the proof of more general Proposition 3.6, so we omit the proof here. Let $s_n(\alpha)=\frac{1}{n}\sum_{i=1}^n g_{i-1}^{(n)}(\alpha)$, and we regard this as a random element taking values in $(C(\Xi),\|\cdot\|_\infty)$. It suffices to see the tightness of this sequence. The tightness is implied by $\sup_n E[\sup_\alpha|\partial_\alpha s_n(\alpha)|]<\infty$ (one can easily check a tightness criterion in C space, for example, appeared in Billingsley [2], p. 58), but this is clear if we use the stationarity and the condition A.2.1. □

4.7. PROOF OF PROPOSITION 3.4

We set

$$
\zeta_i^n(\alpha):=\frac{1}{nh_n}g_{i-1}(\alpha)\bar{X}_{i,n}^{(k)}\bar{X}_{i,n}^{(l)}\mathbf{1}_{\{|\Delta X_i^n|\leqslant h_n^\rho\}}.
$$

We show

$$
\begin{aligned}
A_n&:=\sum_{i=1}^n E\left[\zeta_i^n(\alpha)|\mathcal{F}_{i-1}^n\right]\xrightarrow{P}\int g(x,\alpha)\beta^{(k,l)}(x,\sigma_0)\,\mathrm{d}\pi(x),\\
B_n&:=\sum_{i=1}^n E\left[(\zeta_i^n(\alpha))^2|\mathcal{F}_{i-1}^n\right]\xrightarrow{P}0.
\end{aligned}
$$

Using (3.4), we have

$$A_n = \frac{1}{nh_n}\sum_{i=1}^{n} g_{i-1}(\alpha)\sum_{j=0}^{2} E\left[\bar{X}_{i,n}^{(k)}\bar{X}_{i,n}^{(l)}\mathbf{1}_{C_{i,j}^n}\middle|\mathcal{F}_{i-1}^n\right]$$
$$= \frac{1}{nh_n}\sum_{i=1}^{n} g_{i-1}(\alpha)\left\{h_n e^{-\lambda_0 h_n}\beta_{i-1}^{(k,l)}(\sigma_0) + R(\alpha, h_n^2, X_{t_{i-1}^n}) + \right.$$
$$\left. + \sum_{j=1}^{2} E\left[\bar{X}_{i,n}^{(k)}\bar{X}_{i,n}^{(l)}\mathbf{1}_{C_{i,j}^n}|\mathcal{F}_{i-1}^n\right]\right\}.$$

Here, for sufficiently large n, we have

$$\left|\bar{X}_{i,n}^{(k)}\bar{X}_{i,n}^{(l)}\right|\mathbf{1}_{\{|\Delta X_i^n|\leqslant h_n^\rho\}} \leqslant 2\left\{\left(h_n^\rho\right)^2 + |\bar{a}_{i-1}h_n|^2\right\} = R(\alpha, h_n^{2\rho}, X_{t_{i-1}^n}).$$

Hence

$$\left|\sum_{j=1}^{2}\frac{1}{nh_n}\sum_{i=1}^{n} g_{i-1}(\alpha) E\left[\bar{X}_{i,n}^{(k)}\bar{X}_{i,n}^{(l)}\mathbf{1}_{C_{i,j}^n}|\mathcal{F}_{i-1}^n\right]\right|$$
$$\leqslant \sum_{j=1}^{2}\frac{1}{nh_n}\sum_{i=1}^{n} R(\alpha, h_n^{2\rho}, X_{t_{i-1}^n}) P\{C_{i,j}^n|\mathcal{F}_{i-1}^n\}$$
$$= \mathrm{O}_p(h_n^{1+2\rho}).$$

Therefore, by Proposition 3.3(i),

$$A_n = \frac{1}{nh_n}\sum_{i=1}^{n} g_{i-1}(\alpha) h_n \mathrm{e}^{-\lambda_0 h_n}\beta_{i-1}^{(k,l)}(\sigma_0) +$$
$$+ \mathrm{O}_p(h_n) \xrightarrow{P} \int g(x,\alpha)\beta^{(k,l)}(x,\sigma_0)\,\mathrm{d}\pi(x).$$

The convergence of B_n can be proved similarly as for A_n

The proof of the uniformity of convergence is the same as for Proposition 3.3 (i), that is, we set

$$s_n(\alpha) = \frac{1}{nh_n}\sum_{i=1}^{n} g_{i-1}(\alpha)\bar{X}_{i,n}^{(k)}\bar{X}_{i,n}^{(l)}\mathbf{1}_{\{|\Delta X_i^n|\leqslant h_n^\rho\}},$$

and using (3.6)

$$\sup_n E\left[\sup_\alpha |\partial_\alpha s_n(\alpha)|\right] \leqslant \frac{1}{nh_n}\sum_{i=1}^n E\left[C(1+|X_{t_{i-1}^n}|)^C \left|\bar{X}_{i,n}^{(k)}\bar{X}_{i,n}^{(l)}\right| \mathbf{1}_{\{|\Delta X_i^n| \leqslant h_n^\rho\}}\right]$$
$$\leqslant \frac{C}{nh_n}\sum_{i=1}^n \sqrt{\sum_{j=0}^2 E\left[\left|\bar{X}_{i,n}^{(k)}\bar{X}_{i,n}^{(l)}\right|^2 \mathbf{1}_{C_{i,j}^n}\right]}$$
$$\leqslant \frac{C}{nh_n}\sum_{i=1}^n E[R(\alpha, h_n, X_{t_{i-1}^n})] < \infty.$$

This completes the proof. □

4.8. PROOF OF PROPOSITION 3.5

We prove the convergence for each α in a similar way as for Proposition 3.4. Setting

$$\zeta_i^n(\alpha) = \frac{1}{nh_n} g_{i-1}(\alpha)\bar{X}_{i,n}^{(k)}\mathbf{1}_{\{|\Delta X_i^n| \leqslant h_n^\rho\}},$$

we show

$$A_n := \sum_{i=1}^n E\left[\zeta_i^n(\alpha)|\mathcal{F}_{i-1}^n\right] \xrightarrow{P} 0, \quad B_n := \sum_{i=1}^n E\left[(\zeta_i^n(\alpha))^2|\mathcal{F}_{i-1}^n\right] \xrightarrow{P} 0.$$

Using (3.3), we have

$$A_n = \frac{1}{nh_n}\sum_{i=1}^n g_{i-1}(\alpha)\sum_{j=0}^2 E\left[\bar{X}_{i,n}^{(k)}\mathbf{1}_{C_{i,j}^n}\middle|\mathcal{F}_{i-1}^n\right]$$
$$= \frac{1}{nh_n}\sum_{i=1}^n g_{i-1}(\alpha)\left\{R(\alpha, h_n^2, X_{t_{i-1}^n}) + \sum_{j=1}^2 E\left[\bar{X}_{i,n}^{(k)}\mathbf{1}_{C_{i,j}^n}|\mathcal{F}_{i-1}^n\right]\right\}.$$

Here, for sufficiently large n, we have

$$\left|\bar{X}_{i,n}^{(k)}\right| \mathbf{1}_{\{|\Delta X_i^n| \leqslant h_n^\rho\}} \leqslant h_n^\rho + |\bar{a}_{i-1}h_n| = R(\alpha, h_n^\rho, X_{t_{i-1}^n}).$$

Hence

$$\left|\sum_{j=1}^2 \frac{1}{nh_n}\sum_{i=1}^n g_{i-1}(\alpha)E\left[\bar{X}_{i,n}^{(k)}\mathbf{1}_{C_{i,j}^n}|\mathcal{F}_{i-1}^n\right]\right|$$
$$\leqslant \sum_{j=1}^2 \frac{1}{nh_n}\sum_{i=1}^n R(\alpha, h_n^\rho, X_{t_{i-1}^n})P\{C_{i,j}^n|\mathcal{F}_{i-1}^n\}$$
$$= \mathrm{O}_p(h_n^{1+\rho}).$$

Therefore $A_n = \mathrm{O}_p(h_n)$. The convergence of B_n can be proved similarly as for A_n

Next, we show the tightness of

$$\begin{aligned} s_n(\alpha) &:= \sum_{i=1}^n \zeta_i^n(\alpha) \\ &= \sum_{i=1}^n \bar{\zeta}_i^n(\alpha) + \frac{1}{nh_n}\sum_{i=1}^n g_{i-1}(\alpha)\bar{X}_{i,n}^{(k)}\mathbf{1}_{\{|\Delta X_i^n|>h_n^\rho\}}. \end{aligned}$$

where $\bar{\zeta}_i^n(\alpha) = \frac{1}{nh_n} g_{i-1}(\alpha)\bar{X}_{i,n}^{(k)}$. The tightness of the second term in the last right-hand side is shown by the same argument as the proof of Proposition 3.6 below. Therefore we only show the tightness of the first term $\sum_{i=1}^n \bar{\zeta}_i^n(\alpha)$.

According to Theorem 20 in Appendix 1 from Ibragimov and Has'minskii [15], we should verify the following criterion: for any $N \in \mathbf{N}$ and some positive constant H independent of n,

$$E\left[\left(\sum_{i=1}^n \bar{\zeta}_i^n(\alpha)\right)^{2N}\right] \leqslant H, \tag{4.7}$$

$$E\left[\left(\sum_{i=1}^n \bar{\zeta}_i^n(\alpha_1) - \sum_{i=1}^n \bar{\zeta}_i^n(\alpha_2)\right)^{2N}\right] \leqslant H|\alpha_1-\alpha_2|^{2N}. \tag{4.8}$$

Let $G(s,\alpha) = \sum_{i=1}^n g_{i-1}(\alpha)\mathbf{1}_{[t_{i-1}^n, t_i^n)}(s)$, we have

$$\begin{aligned} &\sum_{i=1}^n \bar{\zeta}_i^n(\alpha) \\ &= \frac{1}{nh_n}\Bigg\{\int_0^{nh_n} G(s,\alpha)\bar{a}^{(k)}(X_s)\,\mathrm{d}s + \sum_{j=1}^r \int_0^{nh_n} G(s,\alpha)b^{(k,j)}(X_s)\,\mathrm{d}W_s^{(j)} + \\ &\quad + \int_0^{nh_n}\int G(s,\alpha)c^{(k)}(X_{s-},z)\,(p-q)(\mathrm{d}s,\mathrm{d}z) - \sum_{i=1}^n g_{i-1}(\alpha)\bar{a}_{i-1}^{(k)}h_n\Bigg\}. \end{aligned}$$

Therefore, the left-hand side of (4.7) is evaluated as follows:

$$\begin{aligned} &E\left[\left(\sum_{i=1}^n \bar{\zeta}_i^n(\alpha)\right)^{2N}\right] \\ &\quad \leqslant C_N\left\{E\left[\left(\frac{1}{nh_n}\int_0^{nh_n} G(s,\alpha)\bar{a}^{(k)}(X_s)\,\mathrm{d}s\right)^{2N}\right] + \right. \end{aligned}$$

$$+\sum_{j=1}^{r} E\left[\left(\frac{1}{nh_n}\int_0^{nh_n} G(s,\alpha)b^{(k,j)}(X_s)\,\mathrm{d}W_s^{(j)}\right)^{2N}\right]+$$
$$+E\left[\left(\frac{1}{nh_n}\int_0^{nh_n}\int G(s,\alpha)c^{(k)}(X_{s-},z)\,(p-q)(\mathrm{d}s,\mathrm{d}z)\right)^{2N}\right]+$$
$$+\,E\left[\left(\frac{1}{n}\sum_{i=1}^{n} g_{i-1}(\alpha)\bar{a}_{i-1}^{(k)}\right)^{2N}\right]\Bigg\}.$$

Applying Jensen's and Burkholder–Davis–Gundy's inequality,

$$E\left[\left(\sum_{i=1}^{n}\bar{\zeta}_i^n(\alpha)\right)^{2N}\right]$$
$$\leqslant C_N\Bigg\{\frac{1}{nh_n}\int_0^{nh_n} E\left[G^{2N}(s,\alpha)(\bar{a}^{(k)}(X_s))^{2N}\right]\mathrm{d}s+$$
$$+\frac{1}{(nh_n)^{N+1}}\sum_{j=1}^{r}\int_0^{nh_n} E\left[G^{2N}(s,\alpha)(b^{(k,j)}(X_s))^{2N}\right]\mathrm{d}s+$$
$$+\frac{1}{(nh_n)^{N+1}}\int_0^{nh_n}\int E\left[G^{2N}(s,\alpha)(c^{(k)}(X_{s-},z))^{2N}\right]q(\mathrm{d}s,\mathrm{d}z)+$$
$$+E\left[\left(\frac{1}{n}\sum_{i=1}^{n} g_{i-1}(\alpha)\bar{a}_{i-1}^{(k)}\right)^{2N}\right]\Bigg\}.$$

One can see that these all are bounded because of Assumption A.2.1 On (4.8), by using Jensen's inequality and Burkholder–Davis–Gundy's inequality,

$$E\left[\left(\frac{1}{|\alpha_1-\alpha_2|}\sum_{i=1}^{n}\left\{\bar{\zeta}_i^n(\alpha_1)-\bar{\zeta}_i^n(\alpha_2)\right\}\right)^{2N}\right]$$
$$\leqslant C_N\Bigg\{\frac{1}{nh_n}\int_0^{nh_n} E\left[\left\{\frac{G(s,\alpha_1)-G(s,\alpha_2)}{|\alpha_1-\alpha_2|}\right\}^{2N}(\bar{a}^{(k)}(X_s))^{2N}\right]\mathrm{d}s+$$
$$+\frac{1}{(nh_n)^{N+1}}\sum_{j=1}^{r}\int_0^{nh_n} E\left[\left\{\frac{G(s,\alpha_1)-G(s,\alpha_2)}{|\alpha_1-\alpha_2|}\right\}^{2N}(b^{(k,j)}(X_s))^{2N}\right]\mathrm{d}s+$$
$$+\frac{1}{(nh_n)^{N+1}}\int_0^{nh_n}\int E\left[\left\{\frac{G(s,\alpha_1)-G(s,\alpha_2)}{|\alpha_1-\alpha_2|}\right\}^{2N}(c^{(k)}(X_{s-},z))^{2N}\right]q(\mathrm{d}s,\mathrm{d}z)+$$
$$+\frac{1}{n}\sum_{i=1}^{n}E\left[\left\{\frac{g_{i-1}(\alpha_1)-g_{i-1}(\alpha_2)}{|\alpha_1-\alpha_2|}\right\}^{2N}\{\bar{a}_{i-1}^{(k)}\}^{2N}\right]\Bigg\}. \tag{4.9}$$

Since $\partial_\alpha g$ is of polynomial growth uniformly in α, we have

$$\frac{|G(s,\alpha_1)-G(s,\alpha_2)|}{|\alpha_1-\alpha_2|} \leqslant \sum_{i=1}^{n} \sup_\alpha |\partial_\alpha g_{i-1}(\alpha)| \mathbf{1}_{[t_{i-1}^n, t_i^n)}(s) \leqslant \sum_{i=1}^{n} C(1+|X_{t_{i-1}^n}|)^C \mathbf{1}_{[t_{i-1}^n, t_i^n)}(s),$$

and

$$\frac{|g_{i-1}(\alpha_1)-g_{i-1}(\alpha_2)|}{|\alpha_1-\alpha_2|} \leqslant \sup_\alpha |\partial_\alpha g_{i-1}(\alpha)| \leqslant C(1+|X_{t_{i-1}^n}|)^C.$$

Hence we see that (4.9) is bounded. □

4.9. PROOF OF PROPOSITION 3.6

First, we show the convergence for each α.

Applying Hölder's inequality, for $p>1, \delta \in (0,1/3]$ which satisfies $\frac{1}{p}+\frac{1}{1+\delta}=1$ and $\epsilon>0$,

$$\begin{aligned}
&\sum_{j=0,2} P\left\{\left|\frac{1}{nh_n}\sum_{i=1}^{n} g_n(\alpha, \Delta X_i^n, X_{t_{i-1}^n})\mathbf{1}_{D_{i,j}^n}\right| > \epsilon\right\} \\
&\quad \leqslant \sum_{j=0,2} \frac{1}{\epsilon n h_n} \sum_{i=1}^{n} E\left[\left|g_n(\alpha, \Delta X_i^n, X_{t_{i-1}^n})\mathbf{1}_{D_{i,j}^n}\right|\right] \\
&\quad \leqslant \sum_{j=0,2} \frac{1}{\epsilon n h_n} \sum_{i=1}^{n} \left(E\left|g_n(\alpha, \Delta X_i^n, X_{t_{i-1}^n})\right|^p\right)^{1/p} \left(P\{D_{i,j}^n\}\right)^{1/1+\delta} \\
&\quad = \mathrm{O}\left(h_n^{\frac{1-\delta}{1+\delta}}\sqrt{b_n}\right) \\
&\quad = \mathrm{O}\left(\sqrt{h_n b_n}\cdot h_n^{\frac{1-3\delta}{2+2\delta}}\right) = \mathrm{o}(1).
\end{aligned}$$

From this, we have the following inequality:

$$\begin{aligned}
&P\left\{\left|\frac{1}{nh_n}\sum_{i=1}^{n} g_n(\alpha, \Delta X_i^n, X_{t_{i-1}^n})\mathbf{1}_{\{|\Delta X_i^n|>h_n^\rho\}} - \iint g_n(\alpha, y, x)\Psi_{\theta_0}(y,x)\,\mathrm{d}y\,\mathrm{d}\pi\right| > 3\epsilon\right\} \\
&\quad = P\left\{\left|\frac{1}{nh_n}\sum_{i=1}^{n}\sum_{j=0}^{2} g_n(\alpha, \Delta X_i^n, X_{t_{i-1}^n})\mathbf{1}_{D_{i,j}^n} - \iint g_n(\alpha, y, x)\Psi_{\theta_0}(y,x)\,\mathrm{d}y\,\mathrm{d}\pi\right| > 3\epsilon\right\} \\
&\quad \leqslant P\left\{\left|\frac{1}{nh_n}\sum_{i=1}^{n} g_n(\alpha, \Delta X_i^n, X_{t_{i-1}^n})\mathbf{1}_{D_{i,1}^n} - \iint g_n(\alpha, y, x)\Psi_{\theta_0}(y,x)\,\mathrm{d}y\,\mathrm{d}\pi\right| > \epsilon\right\} +
\end{aligned}$$

$$+\sum_{j=0,2} P\left\{\left|\frac{1}{nh_n}\sum_{i=1}^n g_n(\alpha, \Delta X_i^n, X_{t_{i-1}^n})\mathbf{1}_{D_{i,j}^n}\right| > \epsilon\right\}$$

$$\leqslant \sum_{k=1}^5 I_k + \mathrm{o}(1),$$

where

$$I_1 = P\left\{\left|\frac{1}{nh_n}\sum_{i=1}^n g_n(\alpha, \Delta X_i^n, X_{t_{i-1}^n})\mathbf{1}_{D_{i,1}^n} - \frac{1}{nh_n}\sum_{i=1}^n g_n(\alpha, \Delta X_{\tau_i^n}, X_{t_{i-1}^n})\mathbf{1}_{D_{i,1}^n}\right| > \frac{\epsilon}{5}\right\},$$

$$I_2 = P\left\{\left|\frac{1}{nh_n}\sum_{i=1}^n g_n(\alpha, \Delta X_{\tau_i^n}, X_{t_{i-1}^n})\mathbf{1}_{D_{i,1}^n} - \frac{1}{nh_n}\sum_{i=1}^n g_n(\alpha, \Delta X_{\tau_i^n}, X_{t_{i-1}^n})\mathbf{1}_{\{J_i^n=1\}}\right| > \frac{\epsilon}{5}\right\},$$

$$I_3 = P\left\{\left|\frac{1}{nh_n}\sum_{i=1}^n g_n(\alpha, \Delta X_{\tau_i^n}, X_{t_{i-1}^n})\mathbf{1}_{\{J_i^n=1\}} - \right.\right.$$
$$\left.\left. - \frac{1}{nh_n}\sum_{i=1}^n \int_{t_{i-1}^n}^{t_i^n}\int g_n(\alpha, c_{i-1}(z,\theta_0), X_{t_{i-1}^n})\, p(\mathrm{d}s, \mathrm{d}z)\right| > \frac{\epsilon}{5}\right\},$$

$$I_4 = P\left\{\left|\frac{1}{nh_n}\sum_{i=1}^n \int_{t_{i-1}^n}^{t_i^n}\int g_n(\alpha, c_{i-1}(z,\theta_0), X_{t_{i-1}^n})\, p(\mathrm{d}s, \mathrm{d}z) - \right.\right.$$
$$\left.\left. - \frac{1}{nh_n}\sum_{i=1}^n \int_{t_{i-1}^n}^{t_i^n}\int g_n(\alpha, c_{i-1}(z,\theta_0), X_{t_{i-1}^n})\, q^{\theta_0}(\mathrm{d}s, \mathrm{d}z)\right| > \frac{\epsilon}{5}\right\},$$

$$I_5 = P\left\{\left|\frac{1}{nh_n}\sum_{i=1}^n \int_{t_{i-1}^n}^{t_i^n}\int g_n(\alpha, c_{i-1}(z,\theta_0), X_{t_{i-1}^n})\, q^{\theta_0}(\mathrm{d}s, \mathrm{d}z) - \right.\right.$$
$$\left.\left. - \iint g(\alpha, y, x)\Psi_{\theta_0}(y, x)\,\mathrm{d}y\,\mathrm{d}\pi\right| > \frac{\epsilon}{5}\right\}.$$

Let us evaluate these terms. Applying Hölder's inequality, for $p > 1$ and $q > 1$ with $\frac{1}{p} + \frac{1}{q} = 1$, we see

$$I_1 \leqslant \frac{5}{\epsilon nh_n}\sum_{i=1}^n E\left[|g_n(\alpha, \Delta X_i^n, X_{t_{i-1}^n}) - g_n(\alpha, \Delta X_{\tau_i^n}, X_{t_{i-1}^n})|\mathbf{1}_{D_{i,1}^n}\right]$$
$$\leqslant \frac{5}{\epsilon nh_n}\sum_{i=1}^n E\left[\left|\tilde{g}_n(\alpha, \xi_i^n, X_{t_{i-1}^n})\right|\left(|X_{t_i^n} - X_{\tau_i^n}| + |X_{\tau_i^n -} - X_{t_{i-1}^n}|\right)\mathbf{1}_{\{J_i^n=1\}}\right]$$
$$\leqslant \frac{5}{\epsilon nh_n}\sum_{i=1}^n E\left[\left(E\left[\left|\tilde{g}_n(\alpha, \xi_i^n, X_{t_{i-1}^n})\right|^2 \middle| J_i^n = 1\right]\right)^{1/2} \times\right.$$
$$\left.\times\left(E\left[\left(|X_{t_i^n} - X_{\tau_i^n}| + |X_{\tau_i^n -} - X_{t_{i-1}^n}|\right)^2 \middle| J_i^n = 1\right]\right)^{1/2}\mathbf{1}_{\{J_i^n=1\}}\right],$$

where $\xi_i^n = \eta\Delta X_i^n + (1-\eta)\Delta X_{\tau_i^n}$ for some $[0,1]$-valued random variable η. Here, we notice that X follows the following stochastic differential equation on the set $\{J_i^n = 1\}$;

$$\tilde{X}_t - \tilde{X}_{t_{i-1}^n} = H_t + \int_{t_{i-1}^n}^t \bar{a}(\tilde{X}_s)\,\mathrm{d}s + \int_{t_{i-1}^n}^t b(\tilde{X}_s)\,\mathrm{d}W_s,$$

where $\tilde{X}_{t_{i-1}^n} = X_{t_{i-1}^n}$, $H_t = c(X_{u-}, z)\mathbf{1}_{[u,t_i^n]}(t)$, u is a $[t_{i-1}^n, t_i^n]$-valued uniform random variable which is independent of $(W_t)_{t\geqslant 0}$ and J_i^n, and z is a random variable with density F_{θ_0} which is independent of $(W_t)_{t\geqslant 0}$. Therefore, for example,

$$E\left[\left|X_{\tau_i^n-} - X_{t_{i-1}^n}\right|^2 \middle| J_i^n = 1\right] = E\left[\left|X_{\tau_i^n-} - X_{t_{i-1}^n}\right|^2 \mathbf{1}_{\{J_i^n=1\}}\right] \Big/ P\{J_i^n = 1\} \cdot \mathbf{1}_{\{J_i^n=1\}}$$
$$= E\left[\left|\tilde{X}_{u-} - \tilde{X}_{t_{i-1}^n}\right|^2\right] \mathbf{1}_{\{J_i^n=1\}}.$$

Applying the Burkholder–Davis–Gundy's inequality to (4.1)

$$E\left[\sup_{t\in[t_{i-1}^n,\tau_i^n-]} |\tilde{X}_t - \tilde{X}_{t_{i-1}^n}|^2\right] \leqslant C\left\{h_n^2 + E\left[\int_{t_{i-1}^n}^{t_i^n} b^2(\tilde{X}_s)\,\mathrm{d}s\right]\right\} = \mathrm{O}(h_n).$$

Hence $E\left[\left|X_{\tau_i^n-} - X_{t_{i-1}^n}\right|^2 \middle| J_i^n = 1\right] = \mathrm{O}(h_n)\mathbf{1}_{\{J_i^n=1\}}$. Similarly, by (4.2),

$$E\left[\sup_{t\in[\tau_i^n, t_i^n]} |\tilde{X}_{t_i^n} - \tilde{X}_t|^2\right] = \mathrm{O}(h_n),$$

so $E\left[\left|X_{t_i^n} - X_{\tau_i^n}\right|^2 \middle| J_i^n = 1\right] = \mathrm{O}(h_n)\mathbf{1}_{\{J_i^n=1\}}$. Hence $I_1 = \mathrm{O}\left(\sqrt{h_n b_n}\right)$

$$I_2 = P\left\{\left|\frac{1}{nh_n}\sum_{i=1}^n g_n(\alpha, \Delta X_{\tau_i^n}, X_{t_{i-1}^n})\mathbf{1}_{C_{i,1}^n}\right| > \frac{\epsilon}{5}\right\}$$
$$\leqslant \frac{5}{\epsilon n h_n}\sum_{i=1}^n E\left[\left|g_n(\alpha, \Delta X_{\tau_i^n}, X_{t_{i-1}^n})\mathbf{1}_{C_{i,1}^n}\right|\right]$$
$$\leqslant \frac{C}{nh_n}\sum_{i=1}^n \mathrm{O}(\sqrt{b_n})\sqrt{P\{C_{i,1}^n\}} = \mathrm{O}\left(\sqrt{h_n b_n}\right).$$

$$I_3 \leqslant P\left\{\left|\frac{1}{nh_n}\sum_{i=1}^n g_n(\alpha, \Delta X_{\tau_i^n}, X_{t_{i-1}^n})\mathbf{1}_{\{J_i^n=1\}} - \right.\right.$$
$$\left.\left. - \frac{1}{nh_n}\sum_{i=1}^n g_n(\alpha, c_{i-1}(\Delta Z_{\tau_i^n}, \theta_0), X_{t_{i-1}^n})\mathbf{1}_{\{J_i^n=1\}}\right| > \frac{\epsilon}{10}\right\} +$$
$$+ P\left\{\left|\frac{1}{nh_n}\sum_{i=1}^n g_n(\alpha, c_{i-1}(\Delta Z_{\tau_i^n}, \theta_0), X_{t_{i-1}^n})\mathbf{1}_{\{J_i^n=1\}} - \right.\right.$$

$$-\frac{1}{nh_n}\sum_{i=1}^{n}\int_{t_{i-1}^n}^{t_i^n}\int g_n(\alpha, c_{i-1}(z,\theta_0), X_{t_{i-1}^n})\, p(\mathrm{d}s,\mathrm{d}z)\bigg| > \frac{\epsilon}{10}\bigg\}.$$

The first term of the right-hand side becomes $\mathrm{O}(\sqrt{h_n b_n})$ by the same argument as I_1. Denote the second term by I_3' then

$$\begin{aligned}
I_3' &\leqslant \frac{10}{\epsilon n h_n}\sum_{i=1}^{n} E\Bigg[\bigg|g_n(\alpha, c_{i-1}(\Delta Z_{\tau_i^n},\theta_0), X_{t_{i-1}^n})\mathbf{1}_{\{J_i^n=1\}} - \\
&\quad - \int_{t_{i-1}^n}^{t_i^n}\int g_n(\alpha, c_{i-1}(z,\theta_0), X_{t_{i-1}^n})\, p(\mathrm{d}s,\mathrm{d}z)\bigg|\Bigg] \\
&= \frac{10}{nh_n}\sum_{i=1}^{n} E\Bigg[\bigg|\int_{t_{i-1}^n}^{t_i^n}\int \mathbf{1}_{\{J_i^n\geqslant 2\}} g_n(\alpha, c_{i-1}(z,\theta_0), X_{t_{i-1}^n})\, p(\mathrm{d}s,\mathrm{d}z)\bigg|\Bigg] \\
&\leqslant \frac{10}{nh_n}\sum_{i=1}^{n}\sqrt{P\{J_i^n\geqslant 2\} E\left(\int_{t_{i-1}^n}^{t_i^n}\int g_n(\alpha, c_{i-1}(z,\theta_0), X_{t_{i-1}^n})\, p(\mathrm{d}s,\mathrm{d}z)\right)^2} \\
&\leqslant \frac{C}{n}\sum_{i=1}^{n}\sqrt{E\left[\int_{t_{i-1}^n}^{t_i^n}\int g_n^2(\alpha, c_{i-1}(z,\theta_0), X_{t_{i-1}^n})\, q(\mathrm{d}s,\mathrm{d}z)\right]} = \mathrm{O}(\sqrt{h_n b_n}).
\end{aligned}$$

Hence $I_3 = \mathrm{O}(\sqrt{h_n b_n})$. Furthermore

$$\begin{aligned}
I_4 &\leqslant \frac{25}{\epsilon^2} E\left[\left(\frac{1}{nh_n}\sum_{i=1}^{n}\int_{t_{i-1}^n}^{t_i^n}\int g_n(\alpha, c_{i-1}(z,\theta_0), X_{t_{i-1}^n})\,(p-q)(\mathrm{d}s,\mathrm{d}z)\right)^2\right] \\
&\leqslant \frac{25}{\epsilon^2 n^2 h_n^2}\sum_{i=1}^{n} E\left[\int_{t_{i-1}^n}^{t_i^n}\int g_n^2(\alpha, c_{i-1}(z,\theta_0), X_{t_{i-1}^n}) q(\mathrm{d}s,\mathrm{d}z)\right] + \\
&\quad + \frac{50}{\epsilon^2 n^2 h_n^2}\sum_{i<j} E\bigg[\int_{t_{i-1}^n}^{t_i^n}\int g_n(\alpha, c_{i-1}(z,\theta_0), X_{t_{i-1}^n})\,(p-q)(\mathrm{d}s,\mathrm{d}z)\times \\
&\quad \times E\left[\int_{t_{j-1}^n}^{t_j^n}\int g_n(\alpha, c_{j-1}(z,\theta_0), X_{t_{j-1}^n})\,(p-q)(\mathrm{d}s,\mathrm{d}z)\bigg|\mathcal{F}_{j-1}^n\right]\bigg] \\
&= \mathrm{O}\left(\frac{b_n}{nh_n}\right).
\end{aligned}$$

On I_5, it is obvious that it converges to zero because of Proposition 3.3 (i) (see Remark 3.4).

Let us show the uniformity of convergence. Set

$$s_n(\alpha) = \frac{1}{nh_n}\sum_{i=1}^{n} g_n(\alpha, \Delta X_i^n, X_{t_{i-1}^n})\mathbf{1}_{\{|\Delta X_i^n| > h_n^\rho\}}.$$

We prove the tightness of $\{s_n(\alpha)\}$. The expectation

$$E\left[\sup_\alpha |\partial_\alpha s_n(\alpha)|\right] \leqslant \frac{1}{nh_n}\sum_{i=1}^n \sum_{j=0}^2 E\left[\sup_\alpha |\partial_\alpha g_n(\alpha, \Delta X_i^n, X_{t_{i-1}^n})| \mathbf{1}_{D_{i,j}^n}\right]$$
$$\leqslant \frac{1}{nh_n}\sum_{i=1}^n E\left[\sup_\alpha |\partial_\alpha g_n(\alpha, \Delta X_i^n, X_{t_{i-1}^n})| \mathbf{1}_{D_{i,1}^n}\right] + \mathrm{o}\left(\sqrt{h_n b_n}\right)$$

by condition (3.9) and Hölder 's inequality, and we can show that

$$\frac{1}{nh_n}\sum_{i=1}^n E\left[\sup_\alpha |\partial_\alpha g_n(\alpha, \Delta X_i^n, X_{t_{i-1}^n})| \mathbf{1}_{D_{i,1}^n}\right]$$
$$= \iint \sup_\alpha |\partial_\alpha g_n(\alpha, y, x)| \Psi_{\theta_0}(y, x)\, \mathrm{d}y\, \mathrm{d}\pi + \mathrm{O}\left(\sqrt{h_n b_n}\right).$$

Indeed, by the same argument as above,

$$\Bigg| E\Bigg[\frac{1}{nh_n}\sum_{i=1}^n \sup_\alpha |\partial_\alpha g_n(\alpha, \Delta X_i^n, X_{t_{i-1}^n})| \mathbf{1}_{D_{i,1}^n} -$$
$$- \iint \sup_\alpha |\partial_\alpha g_n(\alpha, y, x)| \Psi_{\theta_0}(y, x)\, \mathrm{d}y\, \mathrm{d}\pi \Bigg]\Bigg| \leqslant \sum_{k=1}^5 H_k,$$

where

$$H_1 = \Bigg| E\Bigg[\frac{1}{nh_n}\sum_{i=1}^n \sup_\alpha |\partial_\alpha g_n(\alpha, \Delta X_i^n, X_{t_{i-1}^n})| \mathbf{1}_{D_{i,1}^n} -$$
$$- \frac{1}{nh_n}\sum_{i=1}^n \sup_\alpha |\partial_\alpha g_n(\alpha, \Delta X_{\tau_i^n}, X_{t_{i-1}^n})| \mathbf{1}_{D_{i,1}^n} \Bigg]\Bigg|,$$

$$H_2 = \Bigg| E\Bigg[\frac{1}{nh_n}\sum_{i=1}^n \sup_\alpha |\partial_\alpha g_n(\alpha, \Delta X_{\tau_i^n}, X_{t_{i-1}^n})| \mathbf{1}_{D_{i,1}^n} -$$
$$- \frac{1}{nh_n}\sum_{i=1}^n \sup_\alpha |\partial_\alpha g_n(\alpha, \Delta X_{\tau_i^n}, X_{t_{i-1}^n})| \mathbf{1}_{\{J_i^n=1\}} \Bigg]\Bigg|,$$

$$H_3 = \Bigg| E\Bigg[\frac{1}{nh_n}\sum_{i=1}^n \sup_\alpha |\partial_\alpha g_n(\alpha, \Delta X_{\tau_i^n}, X_{t_{i-1}^n})| \mathbf{1}_{\{J_i^n=1\}} -$$
$$- \frac{1}{nh_n}\sum_{i=1}^n \int_{t_{i-1}^n}^{t_i^n}\int \sup_\alpha |\partial_\alpha g_n(\alpha, c_{i-1}(z, \theta_0), X_{t_{i-1}^n})|\, p(\mathrm{d}s, \mathrm{d}z) \Bigg]\Bigg|,$$

$$H_4 = \Bigg| E\Bigg[\frac{1}{nh_n}\sum_{i=1}^n \int_{t_{i-1}^n}^{t_i^n}\int \sup_\alpha |\partial_\alpha g_n(\alpha, c_{i-1}(z, \theta_0), X_{t_{i-1}^n})|\, p(\mathrm{d}s, \mathrm{d}z) -$$

$$-\frac{1}{nh_n}\sum_{i=1}^{n}\int_{t_{i-1}^n}^{t_i^n}\int \sup_{\alpha}|\partial_\alpha g_n(\alpha, c_{i-1}(z,\theta_0), X_{t_{i-1}^n})|\, q^{\theta_0}(\mathrm{d}s,\mathrm{d}z)\Bigg]\Bigg|,$$

$$H_5=\Bigg|E\Bigg[\frac{1}{nh_n}\sum_{i=1}^{n}\int_{t_{i-1}^n}^{t_i^n}\int \sup_{\alpha}|\partial_\alpha g_n(\alpha, c_{i-1}(z,\theta_0), X_{t_{i-1}^n})|\, q^{\theta_0}(\mathrm{d}s,\mathrm{d}z)-$$
$$-\iint \sup_{\alpha}|\partial_\alpha g(\alpha, y, x)|\Psi_{\theta_0}(y,x)\,\mathrm{d}y\,\mathrm{d}\pi\Bigg]\Bigg|.$$

We obtain that $H_1=\mathrm{O}\left(\sqrt{h_n b_n}\right)$ by the same argument as for I_1 since

$$\left|\sup_{\alpha}|\partial_\alpha g_n(\alpha,y_1,x)|-\sup_{\alpha}|\partial_\alpha g_n(\alpha,y_2,x)|\right|\leqslant \sup_{\alpha}|\partial_\alpha g_n(\alpha,y_1,x)-\partial_\alpha g_n(\alpha,y_2,x)|$$
$$\leqslant \tilde{g}_n(\alpha,\eta y_1+(1-\eta)y_2,x)|y_1-y_2|.$$

Similarly, we can obtain that $H_2+H_3=\mathrm{O}\left(\sqrt{h_n b_n}\right)$. Moreover it is also easy to see that $H_4=H_5=0$.

Hence, $\sup_n E\left[\sup_\alpha|\partial_\alpha s_n(\alpha)|\right]<\infty$. This completes the proof. □

4.10. PROOF OF THEOREM 2.1

First, we prove the consistency.

Applying Propositions 3.3(i), (ii), 3.4 and 3.6 we can easily obtain that

$$\frac{1}{n}\bar{l}_n(\alpha)\xrightarrow{P}U_1(\sigma,\sigma_0):=-\frac{1}{2}\int\left\{\mathrm{tr}\left(\beta^{-1}(x,\sigma)\beta(x,\sigma_0)\right)+\log\det\beta(x,\sigma)\right\}\mathrm{d}\pi,\quad(4.10)$$
$$\frac{1}{nh_n}\tilde{l}_n(\theta)\xrightarrow{P}U_2(\theta,\theta_0):=\iint\left\{(\log\Psi_\theta(y,x))\,\Psi_{\theta_0}(y,x)-\Psi_\theta(y,x)\right\}\mathrm{d}y\,\mathrm{d}\pi\quad(4.11)$$

uniformly in σ and θ. See Remarks 3.2 and 3.5 on the conditions for the convergence (4.9)

In order to prove the consistency of $\hat{\alpha}_n$, we may assume that the convergences of (4.9) and (4.9) take place almost surely and uniformly in the parameters, and prove that it implies $\hat{\alpha}_n\to\alpha_0$ almost surely since the convergence in probability implies that, for any subsequence, the existence of a subsequence converging almost surely.

For fixed $\omega\in\Omega$, thanks to the compactness of Ξ, there exists a subsequence n_k such that $\hat{\alpha}_{n_k}\to\alpha_\infty=(\theta_\infty,\sigma_\infty)$. Since the mapping $\sigma\to U_1(\sigma,\sigma_0)$ is continuous,

$$\frac{1}{n_k}l_{n_k}(\hat{\alpha}_{n_k})\longrightarrow U_1(\sigma_\infty,\sigma_0),$$

and, by the definition of $\hat{\sigma}_n$, we have $U_1(\sigma_\infty,\sigma_0)\geqslant U_1(\sigma_0,\sigma_0)$. On the other hand, notice the following inequality:

$$\log\frac{\det\beta(x,\sigma_0)}{\det\beta(x,\sigma_\infty)}\leqslant \mathrm{tr}\left[\beta^{-1}(x,\sigma_\infty)\beta(x,\sigma_0)\right]-d,$$

then we have $U_1(\sigma_\infty, \sigma_0) \leqslant U_1(\sigma_0, \sigma_0)$. Hence $U_1(\sigma_\infty, \sigma_0) = U_1(\sigma_0, \sigma_0)$, and Assumption A.10 leads that $\sigma_\infty = \sigma_0$. This implies that any convergent subsequence of $\hat{\sigma}_n$ tends to σ_0. This means the consistency of $\hat{\sigma}_n$.

Next, let us show the consistency of $\hat{\theta}_n$. Since the mapping $\theta \to U_2(\theta, \theta_0)$ is also continuous,

$$\frac{1}{n_k h_{n_k}} \tilde{l}_{n_k}(\hat{\theta}_{n_k}) \longrightarrow U_2(\theta_\infty, \theta_0)$$

for fixed $\omega \in \Omega$. Here we prepare a lemma.

LEMMA 4.2. *Assume Conditions A.1–A.11, $h_n \to 0$ and $nh_n \to \infty$. Then*

$$\frac{1}{nh_n}\bar{l}_n(\theta, \sigma) - \frac{1}{nh_n}\bar{l}_n(\theta_0, \sigma) \xrightarrow{P} -\frac{1}{2}\int (\bar{a}(x, \theta_0) - \bar{a}(x, \theta))^* \beta^{-1}(x, \sigma)(\bar{a}(x, \theta_0) - \bar{a}(x, \theta))\, \mathrm{d}\pi \tag{4.12}$$

uniformly in α.

Thanks to this lemma and the continuity of the limit function (4.12),

$$\frac{1}{nh_n} l_n(\hat{\theta}_{n_k}, \hat{\sigma}_n) - \frac{1}{nh_n} l_n(\theta_0, \hat{\sigma}_n) \xrightarrow{P} -\frac{1}{2}\int (\bar{a}(x, \theta_0) - \bar{a}(x, \theta_\infty))^* \beta^{-1}(x, \sigma_0)(\bar{a}(x, \theta_0) - \bar{a}(x, \theta_\infty))\, \mathrm{d}\pi - -\{U_2(\theta_0, \theta_0) - U_2(\theta_\infty, \theta_0)\}.$$

The above limit is positive because of the definition of $\hat{\theta}_n$. Hence θ_∞ satisfies $\Psi_{\theta_\infty}(x, z) = \Psi_{\theta_0}(x, z)$ and $\bar{a}(x, \theta_0) = \bar{a}(x, \theta_\infty)$ since β^{-1} is a positive definite and $U_2(\theta, \theta_0)$ will be maximized if and only if $\Psi_\theta(x, z) = \Psi_{\theta_0}(x, z)$. Thus the assumption A.10 implies $\theta_\infty = \theta_0$ This ends the proof of consistency.

Second, we proceed with the proof of the asymptotic normality of $\hat{\alpha}_n$ First, let us compute the first and second derivatives of the contrast function. For $p, p' = 1, 2, \ldots, m_1$ and $q, q' = 1, 2, \ldots, m_2$,

$$\partial_{\theta_p} l_n(\alpha) = \sum_{i=1}^n \left\{\delta^p_{i,1}(\alpha) + \delta^p_{i,2}(\alpha)\right\},$$

$$\delta^p_{i,1}(\alpha) = \sum_{k,l=1}^d \partial_{\theta_p} \bar{a}^{(k)}_{i-1}(\theta)(\beta^{-1}_{i-1})^{(k,l)}(\sigma)\bar{X}^{(l)}_{i,n}(\theta)\mathbf{1}_{\{|\Delta X^n_i| \leqslant h^\rho_n\}},$$

$$\delta^p_{i,2}(\alpha) = \partial_{\theta_p}\left\{\log \Phi_n(\theta, X_{t^n_{i-1}}, \Delta X^n_i)\right\} \varphi_n(X_{t^n_{i-1}}, \Delta X^n_i)\mathbf{1}_{\{|\Delta X^n_i| > h^\rho_n\}} - -h_n \int \partial_{\theta_p}\Phi_n(\theta, X_{t^n_{i-1}}, y)\, \mathrm{d}y,$$

$$\partial_{\sigma_q} l_n(\alpha) = \sum_{i=1}^{n} \zeta_i^q(\alpha),$$

$$\zeta_i^q(\alpha) = -\frac{1}{2}\left\{ \sum_{k,l=1}^{d} \frac{\partial_{\sigma_q}(\beta_{i-1}^{-1})^{(k,l)}(\sigma)}{h_n} \bar{X}_{i,n}^{(k)}(\theta)\bar{X}_{i,n}^{(l)}(\theta) + \right.$$

$$\left. + \frac{\partial_{\sigma_q} \det \beta_{i-1}(\sigma)}{\det \beta_{i-1}(\sigma)} \right\} \mathbf{1}_{\{|\Delta X_i^n| \leqslant h_n^\rho\}},$$

$$\partial^2_{\theta_p\theta_{p'}} l_n(\alpha) = \sum_{i=1}^{n} \sum_{k,l=1}^{d} \left\{ \partial^2_{\theta_p\theta_{p'}} \bar{a}_{i-1}^{(k)}(\theta)(\beta_{i-1}^{-1})^{(k,l)}(\sigma)\bar{X}_{i,n}^{(l)}(\theta) - \right.$$

$$\left. -(\partial_{\theta_p} \bar{a}_{i-1}^{(k)}(\theta))(\partial_{\theta_{p'}} \bar{a}_{i-1}^{(l)}(\theta))(\beta_{i-1}^{-1})^{(k,l)}(\sigma) h_n \right\} \mathbf{1}_{\{|\Delta X_i^n| \leqslant h_n^\rho\}} +$$

$$+ \sum_{i=1}^{n} \left\{ \partial^2_{\theta_p\theta_{p'}} \left\{ \log \Phi_n(\theta, X_{t_{i-1}^n}, \Delta X_i^n) \right\} \varphi_n(X_{t_{i-1}^n}, \Delta X_i^n) \mathbf{1}_{\{|\Delta X_i^n| > h_n^\rho\}} - \right.$$

$$\left. - h_n \int \partial^2_{\theta_p\theta_{p'}} \Phi_n(\theta, X_{t_{i-1}^n}, y)\, \mathrm{d}y \right\},$$

$$\partial^2_{\sigma_q\sigma_{q'}} l_n(\alpha) = -\sum_{i=1}^{n} \left\{ \sum_{k,l=1}^{d} \frac{\partial^2_{\sigma_q\sigma_{q'}} (\beta_{i-1}^{-1})^{(k,l)}(\sigma)}{2h_n} \bar{X}_{i,n}^{(k)}(\theta)\bar{X}_{i,n}^{(l)}(\theta) \right.$$

$$\left. + \frac{1}{2} \partial^2_{\sigma_q\sigma_{q'}} \log\det \beta_{i-1}(\sigma) \right\} \mathbf{1}_{\{|\Delta X_i^n| \leqslant h_n^\rho\}},$$

$$\partial^2_{\theta_p\sigma_q} l_n(\alpha) = \partial^2_{\sigma_q\theta_p} l_n(\alpha) = -\sum_{i=1}^{n} \sum_{k,l=1}^{d} \partial_{\theta_p} \bar{a}_{i-1}^{(k)}(\theta) \partial_{\sigma_q} (\beta_{i-1}^{-1})^{(k,l)} \times$$

$$\times (\sigma) \bar{X}_{i,n}^{(l)}(\theta) \mathbf{1}_{\{|\Delta X_i^n| \leqslant h_n^\rho\}}.$$

We define the following notations.

$$M := \begin{pmatrix} \frac{1}{\sqrt{nh_n}} I_{m_1} & \mathbf{0} \\ \mathbf{0} & \frac{1}{\sqrt{n}} I_{m_2} \end{pmatrix},$$

where I_n is an n-dimensional identity matrix. Let

$$S_n := \begin{pmatrix} \sqrt{nh_n}(\hat{\theta}_n - \theta_0) \\ \sqrt{n}(\hat{\sigma}_n - \sigma_0) \end{pmatrix}, \quad L_n(\alpha) := \begin{pmatrix} -\frac{1}{\sqrt{nh_n}} \partial_\theta l_n(\alpha) \\ -\frac{1}{\sqrt{n}} \partial_\sigma l_n(\alpha) \end{pmatrix},$$

and

$$C_n(\alpha) := \begin{pmatrix} \frac{1}{nh_n}\partial_\theta^2 l_n(\alpha) & \frac{1}{n\sqrt{h_n}}\partial_{\theta\sigma}^2 l_n(\alpha) \\ \frac{1}{n\sqrt{h_n}}\partial_{\sigma\theta}^2 l_n(\alpha) & \frac{1}{n}\partial_\sigma^2 l_n(\alpha) \end{pmatrix}.$$

Then

$$M\partial_\alpha^2 l_n = \begin{pmatrix} \frac{1}{\sqrt{nh_n}}\partial_\theta^2 l_n(\alpha) & \frac{1}{\sqrt{nh_n}}\partial_{\theta\sigma}^2 l_n(\alpha) \\ \frac{1}{\sqrt{n}}\partial_{\sigma\theta}^2 l_n(\alpha) & \frac{1}{\sqrt{n}}\partial_\sigma^2 l_n(\alpha) \end{pmatrix} = C_n(\alpha)M^{-1}.$$

Now, by Taylor's formula,

$$\int_0^1 \partial_\alpha^2 l_n(\alpha_0 + u(\hat{\alpha}_n - \alpha_0))\,\mathrm{d}u \begin{pmatrix} \hat{\theta}_n - \theta_0 \\ \hat{\sigma}_n - \sigma_0 \end{pmatrix} = -\partial_\alpha l_n(\alpha_0) \tag{4.13}$$

since $\partial l_n(\hat{\alpha}_n) = 0$. Then, multiplying both sides by M from the left, we have

$$\int_0^1 C_n(\alpha_0 + u(\hat{\alpha}_n - \alpha_0))\,\mathrm{d}u\, S_n = L_n(\alpha_0). \tag{4.14}$$

Thus the asymptotic normality of S_n is proved by Lemmas 4.3 and 4.4 below.

LEMMA 4.3. *Assume A.1–A.11, $h_n \to 0$ and $nh_n \to \infty$. Then the following statements hold.*

(i) $C_n(\alpha_0) \xrightarrow{P} B$, *where $B = -K$ and K is given in Theorem 2.1.*

(ii) *For any positive sequence ϵ_n tending to zero,*

$$\sup_{|\alpha| \leqslant \epsilon_n} |C_n(\alpha + \alpha_0) - C_n(\alpha_0)| \xrightarrow{P} 0 \quad (n \to \infty).$$

LEMMA 4.4. *Assume A.1–A.11, $h_n \to 0$ and $nh_n \to \infty$, in addition, assume $nh_n^2 \to 0$. Then*

$$L_n(\alpha_0) \xrightarrow{d} L \sim \mathcal{N}(0, K).$$

Actually, by (4.14),

$$\left(\int_0^1 \{C_n(\alpha_0 + u(\hat{\alpha}_n - \alpha_0)) - C_n(\alpha_0)\}\,\mathrm{d}u + C_n(\alpha_0)\right) S_n = L_n(\alpha_0).$$

We find that the matrix

$$\int_0^1 \{C_n(\alpha_0+u(\hat{\alpha}_n-\alpha_0))-C_n(\alpha_0)\}\,\mathrm{d}u+C_n(\alpha_0) \tag{4.15}$$

converges in probability to the nonsingular matrix B. Hence, taking the limit on both sides after multiplying by the inverse of (4.15), we have

$$S_n \xrightarrow{d} B^{-1}L \sim \mathcal{N}(0, K^{-1})$$

by the continuous mapping theorem. This is the end of the proof. □

4.11. PROOF OF LEMMA 4.2

By simple computation,

$$\begin{aligned}
&\frac{1}{nh_n}\bar{l}_n(\theta,\sigma)-\frac{1}{nh_n}\bar{l}_n(\theta_0,\sigma)\\
&=-\frac{1}{2n}\sum_{i=1}^{n}(\bar{a}_{i-1}(\theta_0)-\bar{a}_{i-1}(\theta))^*\beta_{i-1}^{-1}(\sigma)(\bar{a}_{i-1}(\theta_0)-\bar{a}_{i-1}(\theta))\mathbf{1}_{\{|\Delta X_i^n|\leqslant h_n^\rho\}}-\\
&\quad-\frac{1}{nh_n}\sum_{i=1}^{n}(\bar{a}_{i-1}(\theta_0)-\bar{a}_{i-1}(\theta))^*\beta_{i-1}^{-1}(\sigma)\left(\Delta X_i^n-\bar{a}_{i-1}(\theta_0)h_n\right)\mathbf{1}_{\{|\Delta X_i^n|\leqslant h_n^\rho\}}.
\end{aligned}$$

Propositions 3.3(ii) and 3.5 end the proof. □

4.12. PROOF OF LEMMA 4.3

(i) We show the convergence of $n^{-1}\partial_\sigma^2 l_n(\alpha_0)$ only. The others are easily shown by using Propositions 3.3(i), (ii), 3.4, 3.5 and 3.6.
Applying Propositions 3.3 (ii) and 3.4, we have

$$\frac{1}{n}\partial^2_{\sigma_q\sigma_{q'}}l_n(\alpha_0)\xrightarrow{P}-\frac{1}{2}\int \mathrm{tr}\left[\partial^2_{\sigma_q\sigma_{q'}}\beta^{-1}(x)\beta(x)\right]\mathrm{d}\pi-\frac{1}{2}\int \partial_{\sigma_q}\partial_{\sigma_{q'}}\log\det\beta(x)\,\mathrm{d}\pi.$$

Noticing that $\partial_{\sigma_q}\log\det\beta(x,\sigma)=-\mathrm{tr}\left[\partial_{\sigma_q}\beta^{-1}(x)\beta(x)\right]$, so also

$$\partial^2_{\sigma_q\sigma_{q'}}\log\det\beta(x)=-\mathrm{tr}\left[\partial^2_{\sigma_q\sigma_{q'}}\beta^{-1}(x)\beta(x)\right]-\mathrm{tr}\left[\partial_{\sigma_q}\beta^{-1}(x)\partial_{\sigma_{q'}}\beta(x)\right],$$

we can obtain that

$$\begin{aligned}
\frac{1}{n}\partial^2_{\sigma_q\sigma_{q'}}l_n(\alpha_0)&\xrightarrow{P}\frac{1}{2}\int \mathrm{tr}\left[\partial_{\sigma_q}\beta^{-1}(x)\partial_{\sigma_{q'}}\beta(x)\right]\mathrm{d}\pi\\
&=-\frac{1}{2}\int \mathrm{tr}\left[(\partial_{\sigma_q}\beta)\beta^{-1}(\partial_{\sigma_{q'}}\beta)\beta^{-1}\right](x)\,\mathrm{d}\pi.
\end{aligned}$$

□

(ii) Let $B(\alpha)$ be the uniform limit of $C_n(\alpha)$, that is,

$$\sup_{\alpha\in H}|C_n(\alpha)-B(\alpha)|\xrightarrow{P}0,$$

and $B(\alpha)$ is easily specified. Then, noticing $B(\alpha_0)=B$, we have

$$\begin{aligned}&\sup_{|\alpha|\leqslant\epsilon_n}|C_n(\alpha+\alpha_0)-C_n(\alpha_0)|\\&\quad\leqslant 2\sup_{|\alpha|\leqslant\epsilon_n}|C_n(\alpha+\alpha_0)-B(\alpha+\alpha_0)|+\sup_{|\alpha|\leqslant\epsilon_n}|B(\alpha+\alpha_0)-B|.\end{aligned}$$

The first term on the right-hand side converges to zero in probability by the uniformity of convergence, The second term also converges to zero in probability by the continuity of $B(\alpha)$. □

4.13. PROOF OF LEMMA 4.4

We apply the central limit theorem for general L^2-discrete processes (see Shiryayev [26], Chapter VII, Section 8, or Hall and Heyde [13], Chapter 3). It suffices to show the following: For $p,p'=1,\dots,m_1,\ q,q'=1,\dots,m_2$ and some $\nu_1,\nu_2>0$,

$$\sum_{i=1}^n\left|E\left[\frac{1}{\sqrt{nh_n}}\delta_{i,v}^p(\alpha_0)|\mathcal{F}_{i-1}^n\right]\right|\xrightarrow{P}0\quad(v=1,2),\tag{4.16}$$

$$\sum_{i=1}^n\left|E\left[\frac{1}{\sqrt{n}}\zeta_i(\alpha_0)^q|\mathcal{F}_{i-1}^n\right]\right|\xrightarrow{P}0,\tag{4.17}$$

$$\sum_{i=1}^n E\left[\frac{1}{nh_n}\delta_{i,1}^p(\alpha_0)\delta_{i,1}^{p'}(\alpha_0)|\mathcal{F}_{i-1}^n\right]\xrightarrow{P}\int(\partial_{\theta_p}\bar{a})^*\beta^{-1}(\partial_{\theta_{p'}}\bar{a}(x,\alpha_0))\,\mathrm{d}\pi,\tag{4.18}$$

$$\sum_{i=1}^n E\left[\frac{1}{nh_n}\delta_{i,2}^p(\alpha_0)\delta_{i,2}^{p'}(\alpha_0)|\mathcal{F}_{i-1}^n\right]\xrightarrow{P}\iint\frac{\partial_{\theta_p}\Psi_{\theta_0}\partial_{\theta_{p'}}\Psi_{\theta_0}}{\Psi_{\theta_0}}(y,x)\,\mathrm{d}y\,\mathrm{d}\pi,\tag{4.19}$$

$$\sum_{i=1}^n E\left[\frac{1}{nh_n}\delta_{i,1}^p(\alpha_0)\delta_{i,2}^{p'}(\alpha_0)|\mathcal{F}_{i-1}^n\right]\xrightarrow{P}0,\tag{4.20}$$

$$\sum_{i=1}^n E\left[\frac{1}{n}\zeta_i^q(\alpha_0)\zeta_i^{q'}(\alpha_0)|\mathcal{F}_{i-1}^n\right]\xrightarrow{P}\frac{1}{2}\int\mathrm{tr}[(\partial_{\sigma_q}\beta)\beta^{-1}(\partial_{\sigma_{q'}}\beta)\beta^{-1}](x,\sigma_0)\,\mathrm{d}\pi,\tag{4.21}$$

$$\sum_{i=1}^n E\left[\frac{1}{n\sqrt{h_n}}\delta_{i,v}^p(\alpha_0)\zeta_i^q(\alpha_0)|\mathcal{F}_{i-1}^n\right]\xrightarrow{P}0\quad(v=1,2),\tag{4.22}$$

$$\sum_{i=1}^{n} E\left[\left|\frac{1}{\sqrt{nh_n}}\delta_{i,v}^{p}(\alpha_0)\right|^{2+\nu_1}|\mathcal{F}_{i-1}^{n}\right]\xrightarrow{P}0 \quad (v=1,2), \tag{4.23}$$

$$\sum_{i=1}^{n} E\left[\left|\frac{1}{\sqrt{n}}\zeta_{i}^{q}(\alpha_0)\right|^{2+\nu_2}|\mathcal{F}_{i-1}^{n}\right]\xrightarrow{P}0. \tag{4.24}$$

Remark 4.1. We use the central limit theorem for triangular arrays to show this lemma, so we have to check the Lindeberg condition for $L_n(\alpha_0)=\sum_{i=1}^{n} X_{ni}$, that is,

$$\sum_{i=1}^{n} E\left[|X_{ni}|^2\mathbf{1}_{\{|X_{ni}|>\epsilon\}}|\mathcal{F}_{i-1}^{n}\right]\xrightarrow{P}0$$

for any $\epsilon>0$. If X_{ni} has an expression $X_{ni}=Y_{ni}+Z_{ni}$ then

$$\begin{aligned}&\sum_{i=1}^{n} E\left[|X_{ni}|^2\mathbf{1}_{\{|X_{ni}|>\epsilon\}}|\mathcal{F}_{i-1}F^{n}\right]\\&\quad\leqslant 4\sum_{i=1}^{n} E\left[|Y_{ni}|^2\mathbf{1}_{\{|Y_{ni}|>\epsilon/2\}}|\mathcal{F}_{i-1}^{n}\right]+4\sum_{i=1}^{n} E\left[|Z_{ni}|^2\mathbf{1}_{\{|Z_{ni}|>\epsilon/2\}}|\mathcal{F}_{i-1}^{n}\right].\end{aligned}$$

Hence, to check the above Lindeberg condition, it suffices to check the following Lyapnov conditions:

$$\sum_{i=1}^{n} E\left[|Y_{ni}|^{2+\nu_1}|\mathcal{F}_{i-1}^{n}\right],\quad \sum_{i=1}^{n} E\left[|Z_{ni}|^{2+\nu_2}|\mathcal{F}_{i-1}^{n}\right]\xrightarrow{P}0$$

for some $\nu_1,\nu_2>0$. Here, it is not necessary that ν_1 and ν_2 are the same, so the above ν_i's of (4.23) and (4.24) can be taken differently.

Proof of (4.16).

For $v=1$,

$$\begin{aligned}&\sum_{i=1}^{n}\left|E\left[\frac{1}{\sqrt{nh_n}}\delta_{i,1}^{p}(\alpha_0)|\mathcal{F}_{i-1}^{n}\right]\right|\\&\quad=\frac{1}{\sqrt{nh_n}}\sum_{i=1}^{n}\left|\sum_{k,l=1}^{d}\partial_{\theta_p}\bar{a}_{i-1}^{(k)}(\theta_0)(\beta_{i-1}^{-1})^{(k,l)}(\sigma_0)\sum_{j=0}^{2}E\left[\bar{X}_{i,n}^{(l)}\mathbf{1}_{C_{i,j}^{n}}|\mathcal{F}_{i-1}^{n}\right]\right|.\end{aligned}$$

Since $|\bar{X}_{i,n}^{(l)}|\mathbf{1}_{\{|\Delta X_i^n|\leqslant h_n^{\rho}\}}=R(\alpha,h_n^{\rho},X_{t_{i-1}^n})$,

$$\frac{1}{\sqrt{nh_n}}\sum_{i=1}^{n}\left|\sum_{k,l=1}^{d}\partial_{\theta_p}\bar{a}_{i-1}^{(k)}(\theta_0)(\beta_{i-1}^{-1})^{(k,l)}(\sigma_0)\sum_{j=1}^{2}E\left[\bar{X}_{i,n}^{(l)}\mathbf{1}_{C_{i,j}^{n}}|\mathcal{F}_{i-1}^{n}\right]\right|$$

$$\leqslant \frac{1}{\sqrt{nh_n}} \sum_{i=1}^{n} R(\alpha, h_n^{\rho}, X_{t_{i-1}^n}) \sum_{j=1}^{2} P\{C_{i,j}^n | \mathcal{F}_{i-1}^n\}$$
$$\leqslant \frac{1}{n} \sum_{i=1}^{n} R\left(\alpha, \sqrt{nh_n^{3+2\rho}}, X_{t_{i-1}^n}\right).$$

Applying (3.3) to the term for $j=0$, we have

$$\sum_{i=1}^{n} \left| E\left[\frac{1}{\sqrt{nh_n}} \delta_{i,1}^{p}(\alpha_0) | \mathcal{F}_{i-1}^n \right] \right|$$
$$= \frac{1}{\sqrt{nh_n}} \sum_{i=1}^{n} \left| \sum_{k,l=1}^{d} \partial_{\theta_p} \bar{a}_{i-1}^{(k)}(\theta_0) (\beta_{i-1}^{-1})^{(k,l)}(\sigma_0) \right| R(\alpha, h_n^2, X_{t_{i-1}^n}) + \mathrm{o}_p(\sqrt{nh_n^3})$$
$$= \frac{1}{n} \sum_{i=1}^{n} R(\alpha, \sqrt{nh_n^3}, X_{t_{i-1}^n}) + o_p(\sqrt{nh_n^3}) \xrightarrow{P} 0.$$

For $v=2$,

$$\sum_{i=1}^{n} \left| E\left[\frac{1}{\sqrt{nh_n}} \delta_{i,2}^{p}(\alpha_0) | \mathcal{F}_{i-1}^n \right] \right|$$
$$= \frac{1}{\sqrt{nh_n}} \sum_{i=1}^{n} \left| E\Big[\sum_{j=0}^{2} \partial_{\theta_p} \left\{ \log \Phi_n(\theta_0, X_{t_{i-1}^n}, \Delta X_i^n) \right\} \varphi_n(X_{t_{i-1}^n}, \Delta X_i^n) \mathbf{1}_{D_{i,j}^n} - \right.$$
$$\left. -h_n \int \partial_{\theta_p} \Phi_n(\theta_0, X_{t_{i-1}^n}, y)\, \mathrm{d}y \Big| \mathcal{F}_{i-1}^n \Big] \right|$$
$$= \frac{1}{\sqrt{nh_n}} \sum_{i=1}^{n} \left| E\Big[\partial_{\theta_p} \left\{ \log \Phi_n(\theta_0, X_{t_{i-1}^n}, \Delta X_i^n) \right\} \varphi_n(X_{t_{i-1}^n}, \Delta X_i^n) \mathbf{1}_{D_{i,1}^n} - \right.$$
$$\left. -h_n \int \partial_{\theta_p} \Phi_n(\theta_0, X_{t_{i-1}^n}, y)\, \mathrm{d}y \Big| \mathcal{F}_{i-1}^n \Big] \right| + o_p\left(\sqrt{nh_n^2 b_n} \right)$$
$$\leqslant \frac{1}{\sqrt{nh_n}} \sum_{i=1}^{n} \sum_{k=1}^{5} I_k^i + o_p\left(\sqrt{nh_n^2 b_n} \right),$$

where

$$I_1^i = \left| E\Big[\partial_{\theta_p} \left\{ \log \Phi_n(\theta_0, X_{t_{i-1}^n}, \Delta X_i^n) \right\} \varphi_n(X_{t_{i-1}^n}, \Delta X_i^n) \mathbf{1}_{D_{i,1}^n} - \right.$$
$$\left. -\partial_{\theta_p} \left\{ \log \Phi_n(\theta_0, X_{t_{i-1}^n}, \Delta X_{\tau_i^n}) \right\} \varphi_n(X_{t_{i-1}^n}, \Delta X_{\tau_i^n}) \mathbf{1}_{D_{i,1}^n} \Big| \mathcal{F}_{i-1}^n \Big] \right|,$$

$$I_2^i = \left| E\Big[\partial_{\theta_p} \left\{ \log \Phi_n(\theta_0, X_{t_{i-1}^n}, \Delta X_{\tau_i^n}) \right\} \varphi_n(X_{t_{i-1}^n}, \Delta X_{\tau_i^n}) \mathbf{1}_{D_{i,1}^n} - \right.$$

$$-\partial_{\theta_p}\left\{\log\Phi_n(\theta_0, X_{t_{i-1}^n}, \Delta X_{\tau_i^n})\right\}\varphi_n(X_{t_{i-1}^n}, \Delta X_{\tau_i^n})\mathbf{1}_{\{J_i^n=1\}}\Big|\mathcal{F}_{i-1}^n\Big]\Big|,$$

$$I_3^i = \Big|E\Big[\partial_{\theta_p}\left\{\log\Phi_n(\theta_0, X_{t_{i-1}^n}, \Delta X_{\tau_i^n})\right\}\varphi_n(X_{t_{i-1}^n}, \Delta X_{\tau_i^n})\mathbf{1}_{\{J_i^n=1\}} -$$
$$-\int_{t_{i-1}^n}^{t_i^n}\int \partial_{\theta_p}\left\{\log\Phi_n(\theta_0, X_{t_{i-1}^n}, c_{i-1}(z,\theta_0))\right\}\times$$
$$\times\varphi_n(X_{t_{i-1}^n}, c_{i-1}(z,\theta_0))\, p(\mathrm{d}s,\mathrm{d}z)\Big|\mathcal{F}_{i-1}^n\Big]\Big|,$$

$$I_4^i = \Big|E\Big[\int_{t_{i-1}^n}^{t_i^n}\int \partial_{\theta_p}\left\{\log\Phi_n(\theta_0, X_{t_{i-1}^n}, c_{i-1}(z,\theta_0))\right\}\varphi_n(X_{t_{i-1}^n}, c_{i-1}(z,\theta_0)) -$$
$$-p(\mathrm{d}s,\mathrm{d}z) - \int_{t_{i-1}^n}^{t_i^n}\int \partial_{\theta_p}\left\{\log\Phi_n(\theta_0, X_{t_{i-1}^n}, c_{i-1}(z,\theta_0))\right\}\times$$
$$\times\varphi_n(X_{t_{i-1}^n}, c_{i-1}(z,\theta_0))q^{\theta_0}(\mathrm{d}s,\mathrm{d}z)\Big|\mathcal{F}_{i-1}^n\Big]\Big|,$$

$$I_5^i = \Big|E\Big[\int_{t_{i-1}^n}^{t_i^n}\int \partial_{\theta_p}\left\{\log\Phi_n(\theta_0, X_{t_{i-1}^n}, c_{i-1}(z,\theta_0))\right\}\varphi_n(X_{t_{i-1}^n}, c_{i-1}(z,\theta_0))\times$$
$$\times q^{\theta_0}(\mathrm{d}s,\mathrm{d}z) - h_n\int \partial_{\theta_p}\Phi_n(\theta_0, X_{t_{i-1}^n}, y)\,\mathrm{d}y\Big|\mathcal{F}_{i-1}^n\Big]\Big|.$$

Since, by Remark 3.5,

$$|\partial_y\partial_{\theta_p}\log\Phi_n(\theta_0,x,y)\varphi_n(x,y)| \leqslant \mathrm{O}(\sqrt{b_n})(1+|y|)^C(1+|x|)^C,$$

if we take the same argument for I_1^i–I_3^i as for H_1–H_3 in the proof of Proposition 3.6 it is easy to obtain that

$$I_1^i = R\left(\alpha, \sqrt{h_n^3 b_n}, X_{t_{i-1}^n}\right), \quad I_2^i = R(\alpha, \sqrt{h_n^3 b_n}, X_{t_{i-1}^n}),$$
$$I_3^i = R\left(\alpha, \sqrt{h_n^3 b_n}, X_{t_{i-1}^n}\right).$$

It is also easy to see that $I_4^i, I_5^i = 0$. Thus

$$\sum_{i=1}^n\left|E\left[\frac{1}{\sqrt{nh_n}}\delta_{i,2}^p(\alpha_0)|\mathcal{F}_{i-1}^n\right]\right| = \mathrm{O}_p\left(\sqrt{nh_n^2 b_n}\right).$$

Proof of (4.17).

Using Proposition 3.2 (3.4), we have

$$\sum_{i=1}^n\left|E\left[\frac{1}{\sqrt{n}}\zeta_i^q(\alpha_0)|\mathcal{F}_{i-1}^n\right]\right|$$

$$
\begin{aligned}
&= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left| \sum_{j=0}^{2} E\left[\sum_{k,l=1}^{d} \frac{\partial_{\sigma_q} (\beta_{i-1}^{-1})^{(k,l)}(\sigma_0)}{2h_n} \bar{X}_{i,n}^{(k)} \bar{X}_{i,n}^{(l)} \mathbf{1}_{C_{i,j}^n} \middle| \mathcal{F}_{i-1}^n \right] + \right. \\
&\quad \left. + \frac{1}{2} \frac{\partial_{\sigma_q} \det \beta_{i-1}(\sigma_0)}{\det \beta_{i-1}(\sigma_0)} P\left\{ |\Delta X_i^n| \leqslant h_n^\rho | \mathcal{F}_{i-1}^n \right\} \right| \\
&= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left| \frac{1}{2} \mathrm{tr}\left[\partial_{\sigma_q} \beta_{i-1}^{-1}(\sigma_0) \beta_{i-1}(\sigma_0) \right] e^{-\lambda_0 h_n} + R(\alpha, h_n, X_{t_{i-1}^n}) + \right. \\
&\quad \left. + \frac{1}{2} \frac{\partial_{\sigma_q} \det \beta_{i-1}(\sigma_0)}{\det \beta_{i-1}(\sigma_0)} \left(e^{-\lambda_0 h_n} + R(\alpha, h_n^3, X_{t_{i-1}^n}) + \lambda_0^2 h_n^2 \right) \right| \\
&= \mathrm{O}_p\left(\sqrt{n h_n^2} \right).
\end{aligned}
$$

We used the relation $\dfrac{\partial_{\sigma_q} \det \beta_{i-1}(\sigma)}{\det \beta_{i-1}(\sigma)} = -\mathrm{tr}\left[\partial_{\sigma_q} \beta_{i-1}^{-1}(\sigma) \beta_{i-1}(\sigma) \right]$. □

Proof of (4.18).

Noticing that $|\bar{X}_{i,n}^{(l)}| \mathbf{1}_{\{|\Delta X_i^n| \leqslant h_n^\rho\}} = R(\alpha, h_n^\rho, X_{t_{i-1}^n})$, we see from Propositions 3.2 (3.4) and Proposition 3.3(i) that

$$
\begin{aligned}
&\sum_{i=1}^{n} E\left[\frac{1}{n h_n} \delta_{i,1}^{p}(\alpha_0) \delta_{i,1}^{p'}(\alpha_0) | \mathcal{F}_{i-1}^n \right] \\
&\quad = \frac{1}{n h_n} \sum_{i=1}^{n} E\Bigg[\left(\sum_{k,l=1}^{d} \partial_{\theta_p} \bar{a}_{i-1}^{(k)}(\theta_0) (\beta_{i-1}^{-1})^{(k,l)}(\sigma_0) \bar{X}_{i,n}^{(l)} \right) \times \\
&\quad\quad \times \left(\sum_{k',l'=1}^{d} \partial_{\theta_{p'}} \bar{a}_{i-1}^{(k')}(\theta_0) (\beta_{i-1}^{-1})^{(k',l')}(\sigma_0) \bar{X}_{i,n}^{(l')} \right) \mathbf{1}_{\{|\Delta X_i^n| \leqslant h_n^\rho\}} \Bigg| \mathcal{F}_{i-1}^n \Bigg] \\
&\quad \xrightarrow{P} \sum_{k,l,k',l'=1}^{d} \int \partial_{\theta_p} \bar{a}^{(k)} \partial_{\theta_{p'}} \bar{a}^{(k')} (\beta^{-1})^{(k,l)} (\beta^{-1})^{(k',l')} (\beta)^{(l,l')} (x, \alpha_0) \, \mathrm{d}\pi \\
&\quad = \int \left(\partial_{\theta_p} \bar{a} \right)^* \left(\beta^{-1} \right) \left(\partial_{\theta_{p'}} \bar{a} \right) (x, \alpha) \, \mathrm{d}\pi.
\end{aligned}
$$

□

Proof of (4.19).

$$
\begin{aligned}
&\sum_{i=1}^{n} E\left[\frac{1}{n h_n} \delta_{i,2}^{p}(\alpha_0) \delta_{i,2}^{p'}(\alpha_0) | \mathcal{F}_{i-1}^n \right] \\
&\quad = \frac{1}{n h_n} \sum_{i=1}^{n} E\Bigg[\Bigg\{ \partial_{\theta_p} \left\{ \log \Phi_n(\theta_0, X_{t_{i-1}^n}, \Delta X_i^n) \right\} \varphi_n(X_{t_{i-1}^n}, \Delta X_i^n) \mathbf{1}_{\{|\Delta X_i^n| > h_n^\rho\}} -
\end{aligned}
$$

$$
\begin{aligned}
&- h_n \int \partial_{\theta_p} \Phi_n(\theta_0, X_{t_{i-1}^n}, y)\,\mathrm{d}y \Big\} \\
&\times \Big\{ \partial_{\theta_{p'}} \left\{ \log \Phi_n(\theta_0, X_{t_{i-1}^n}, \Delta X_i^n) \right\} \varphi_n(X_{t_{i-1}^n}, \Delta X_i^n) \mathbf{1}_{\{|\Delta X_i^n| > h_n^\rho\}} - \\
&- h_n \int \partial_{\theta_{p'}} \Phi_n(\theta_0, X_{t_{i-1}^n}, y)\,\mathrm{d}y \Big\} \Big| \mathcal{F}_{i-1}^n \Big] \\
= {} & \frac{1}{nh_n} \sum_{i=1}^n E\Big[\partial_{\theta_p} \left\{ \log \Phi_n(\theta_0, X_{t_{i-1}^n}, \Delta X_i^n) \right\} \varphi_n(X_{t_{i-1}^n}, \Delta X_i^n) \times \\
&\times \partial_{\theta_{p'}} \left\{ \log \Phi_n(\theta_0, X_{t_{i-1}^n}, \Delta X_i^n) \right\} \varphi_n(X_{t_{i-1}^n}, \Delta X_i^n) \mathbf{1}_{\{|\Delta X_i^n| > h_n^\rho\}} \Big| \mathcal{F}_{i-1}^n \Big] + \\
&+ \mathrm{O}_p(\sqrt{h_n b_n}) \\
\xrightarrow{P} {} & \iint \frac{\partial_{\theta_p} \Psi_{\theta_0} \partial_{\theta_{p'}} \Psi_{\theta_0}}{\Psi_{\theta_0}}(y, x)\,\mathrm{d}y\,\mathrm{d}\pi.
\end{aligned}
$$

The last convergence is proved by the same argument as the proof of (4.16) with $v = 2$ since

$$
\begin{aligned}
&\left| \partial_y^m \left[\partial_{\theta_p} \{ \log \Phi_n(\theta, x, y) \} \partial_{\theta_{p'}} \{ \log \Phi_n(\theta, x, y) \} \varphi_n^2(x, y) \right] \right| \\
&\quad \leqslant O(\sqrt{b_n})(1 + |x|)^C (1 + |y|)^C \quad (m = 0, 1).
\end{aligned}
$$

Proof of (4.20).

Noticing that $|\bar{X}_{i,n}^{(l)}| \mathbf{1}_{\{|\Delta X_i^n| \leqslant h_n^\rho\}} = R(\alpha, h_n^\rho, X_{t_{i-1}^n})$, we see from Proposition 3.2 (3.3) that

$$
\begin{aligned}
&\sum_{i=1}^n E\left[\frac{1}{nh_n} \delta_{i,1}^p(\alpha_0) \delta_{i,2}^{p'}(\alpha_0) | \mathcal{F}_{i-1}^n \right] \\
&\quad - \frac{1}{n} \sum_{i=1}^n E\left[\sum_{k,l=1}^d \left\{ \int \partial_{\theta_p} \Phi_n(X_{t_{i-1}^n}, y)\,\mathrm{d}y \right\} \times \right. \\
&\quad \left. \times \partial_{\theta_p} \bar{a}_{i-1}^{(k)} (\beta_{i-1}^{-1})^{(k,l)} \bar{X}_{i,n}^{(l)} \mathbf{1}_{\{|\Delta X_i^n| \leqslant h_n^\rho\}} \right| \mathcal{F}_{i-1}^n \Big] \\
&= -\frac{1}{n} \sum_{i=1}^n E\left[\sum_{k,l=1}^d \left\{ \int \partial_{\theta_p} \Phi_n(X_{t_{i-1}^n}, y)\,\mathrm{d}y \right\} \times \right. \\
&\quad \left. \times \partial_{\theta_p} \bar{a}_{i-1}^{(k)} (\beta_{i-1}^{-1})^{(k,l)} \bar{X}_{i,n}^{(l)} \mathbf{1}_{C_{i,0}^n} \right| \mathcal{F}_{i-1}^n \Big] + \mathrm{o}_p(h_n^2) \\
&= O(h_n^2).
\end{aligned}
$$

Proof of (4.21).

We used (3.4) and (3.6), and the relation

$$\frac{\partial_{\sigma_q} \det \beta_{i-1}(\sigma_0)}{\det \beta_{i-1}(\sigma_0)} = -\mathrm{tr}\left[\partial_{\sigma_q}\beta_{i-1}^{-1}(\sigma_0)\beta_{i-1}(\sigma_0)\right] = \mathrm{tr}\left[\partial_{\sigma_q}\beta_{i-1}(\sigma_0)\beta_{i-1}^{-1}(\sigma_0)\right]$$

to obtain

$$\begin{aligned}
&\sum_{i=1}^{n} E\left[\frac{1}{n}\zeta_i^q(\alpha_0)\zeta_i^{q'}(\alpha_0)|\mathcal{F}_{i-1}^n\right] \\
&= \frac{1}{4n}\sum_{i=1}^{n} E\Bigg[\left\{\sum_{k,l=1}^{d}\frac{\partial_{\sigma_q}(\beta_{i-1}^{-1})^{(k,l)}(\sigma_0)}{h_n}\bar{X}_{i,n}^{(k)}\bar{X}_{i,n}^{(l)} + \frac{\partial_{\sigma_q}\det\beta_{i-1}(\sigma_0)}{\det\beta_{i-1}(\sigma_0)}\right\}\times \\
&\quad \times\left\{\sum_{k',l'=1}^{d}\frac{\partial_{\sigma_{q'}}(\beta_{i-1}^{-1})^{(k',l')}(\sigma_0)}{h_n}\bar{X}_{i,n}^{(k')}\bar{X}_{i,n}^{(l')} + \frac{\partial_{\sigma_{q'}}\det\beta_{i-1}(\sigma_0)}{\det\beta_{i-1}(\sigma_0)}\right\} \\
&\quad \mathbf{1}_{\{|\Delta X_i^n|\leqslant h_n^\rho\}}\Bigg|\mathcal{F}_{i-1}^n\Bigg] \\
&= \frac{1}{4n}\sum_{i=1}^{n}\Bigg[\sum_{k,l,k',l'=1}^{d} \mathrm{e}^{-\lambda_0 h_n}(\partial_{\sigma_q}\beta_{i-1}^{-1})^{(k,l)}(\partial_{\sigma_{q'}}\beta_{i-1}^{-1})^{(k',l')} \\
&\quad \left(\beta_{i-1}^{(k,l)}\beta_{i-1}^{(k',l')} + \beta_{i-1}^{(k,k')}\beta_{i-1}^{(l,l')} + \beta_{i-1}^{(k,l')}\beta_{i-1}^{(k',l)}\right) + \\
&\quad + \sum_{k',l'=1}^{d}\mathrm{e}^{-\lambda_0 h_n}\frac{\partial_{\sigma_q}\det\beta_{i-1}(\sigma_0)}{\det\beta_{i-1}(\sigma_0)}\partial_{\sigma_{q'}}(\beta_{i-1}^{-1})^{(k',l')}\beta_{i-1}^{(k',l')} + \\
&\quad + \sum_{k,l=1}^{d}\mathrm{e}^{-\lambda_0 h_n}\frac{\partial_{\sigma_{q'}}\det\beta_{i-1}(\sigma_0)}{\det\beta_{i-1}(\sigma_0)}\partial_{\sigma_q}(\beta_{i-1}^{-1})^{(k,l)}\beta_{i-1}^{(k,l)} + \\
&\quad + \mathrm{e}^{-\lambda_0 h_n}\frac{\partial_{\sigma_q}\det\beta_{i-1}(\sigma_0)}{\det\beta_{i-1}(\sigma_0)}\frac{\partial_{\sigma_{q'}}\det\beta_{i-1}(\sigma_0)}{\det\beta_{i-1}(\sigma_0)}\Bigg] + \mathrm{O}_p(h_n^{4\rho}) \\
&\xrightarrow{P} \frac{1}{2}\int \mathrm{tr}[(\partial_{\sigma_q}\beta)\beta^{-1}(\partial_{\sigma_{q'}}\beta)\beta^{-1}]\,\mathrm{d}\pi
\end{aligned}$$

by Proposition 3.3(i).

Proof of (4.22).

For $v=1$,

$$\sum_{i=1}^{n} E\left[\frac{1}{n\sqrt{h_n}}\delta_{i,1}^p(\alpha_0)\zeta_i^q(\alpha_0)|\mathcal{F}_{i-1}^n\right]$$

$$= -\frac{1}{n\sqrt{h_n}}\sum_{i=1}^{n}\sum_{k,l,k'l'=1}^{d}\frac{1}{2h_n}\partial_{\theta_p}\bar{a}_{i-1}^{(k)}(\theta_0)(\beta_{i-1}^{-1})^{(k,l)}(\sigma_0)\partial_{\sigma_q}(\beta_{i-1}^{-1})^{(k',l')}(\sigma_0)\times$$
$$\times E\left[\bar{X}_{i,n}^{(l)}\bar{X}_{i,n}^{(k')}\bar{X}_{i,n}^{(l')}\mathbf{1}_{\{|\Delta X_i^n|\leqslant h_n^\rho\}}\Big|\mathcal{F}_{i-1}^n\right]-$$
$$-\frac{1}{n\sqrt{h_n}}\sum_{i=1}^{n}\frac{1}{2}\sum_{k,l=1}^{d}\partial_{\theta_p}\bar{a}_{i-1}^{(k)}(\theta_0)(\beta_{i-1}^{-1})^{(k,l)}(\sigma_0)\frac{\partial_{\sigma_q}\det\beta_{i-1}(\sigma_0)}{\det\beta_{i-1}(\sigma_0)}\times$$
$$\times E\left[\bar{X}_{i,n}^{(l)}\mathbf{1}_{\{|\Delta X_i^n|\leqslant h_n^\rho\}}\Big|\mathcal{F}_{i-1}^n\right].$$

Noticing that $|\bar{X}_{i,n}^{(l)}|\mathbf{1}_{\{|\Delta X_i^n|\leqslant h_n^\rho\}} = R(\alpha, h_n^\rho, X_{t_{i-1}^n})$, we see from Proposition 3.2 (3.3) and (3.5) that

$$\sum_{i=1}^{n}E\left[\frac{1}{n\sqrt{h_n}}\delta_{i,1}^{p}(\alpha_0)\zeta_i^q(\alpha_0)|\mathcal{F}_{i-1}^n\right] = \frac{1}{n\sqrt{h_n}}\sum_{i=1}^{n}R(\alpha, h_n, X_{t_{i-1}^n}) + \mathrm{O}_p(\sqrt{h_n})$$
$$\xrightarrow{P} 0.$$

For $v=2$, by using Proposition 3.2 (3.4),

$$\sum_{i=1}^{n}E\left[\frac{1}{n\sqrt{h_n}}\delta_{i,2}^{p}(\alpha_0)\zeta_i^q(\alpha_0)|\mathcal{F}_{i-1}^n\right]$$
$$= -\frac{1}{n\sqrt{h_n}}\sum_{i=1}^{n}E\left[\left\{\sum_{k,l=1}^{d}\frac{\partial_{\sigma_q}(\beta_{i-1}^{-1})^{(k,l)}(\sigma_0)}{2h_n}\bar{X}_{i,n}^{(k)}\bar{X}_{i,n}^{(l)}+\frac{1}{2}\frac{\partial_{\sigma_q}\det\beta_{i-1}(\sigma_0)}{\det\beta_{i-1}(\sigma_0)}\right\}\times\right.$$
$$\left.\times\left(h_n\int\partial_{\theta_p}\Phi_n(\theta_0, X_{t_{i-1}^n}, y)\,\mathrm{d}y\right)\mathbf{1}_{\{|\Delta X_i^n|\leqslant h_n^\rho\}}\Big|\mathcal{F}_{i-1}^n\right]$$
$$= \mathrm{o}_p\left(\sqrt{h_n}\right).$$

Proof of (4.23).

For $v=1$,

$$\sum_{i=1}^{n}E\left[\left|\frac{1}{\sqrt{nh_n}}\delta_{i,1}^{p}(\alpha_0)\right|^{2+\nu}|\mathcal{F}_{i-1}^n\right]$$
$$\leqslant \frac{C_\nu}{n^{1+\nu/2}h_n^{1+\nu/2}}\sum_{i=1}^{n}\sum_{k,l=1}^{d}\sum_{j=0}^{2}\left|\partial_{\theta_p}\bar{a}_{i-1}^{(k)}(\theta_0)(\beta_{i-1}^{-1})^{(k,l)}(\sigma_0)\right|^{2+\nu}\times$$
$$\times E\left[\left|\bar{X}_{i,n}^{(l)}\right|^{2+\nu}\mathbf{1}_{C_{i,j}^n}|\mathcal{F}_{i-1}^n\right].$$

Noticing that $E\left[\left|\bar{X}_{i,n}^{(l)}\right|^{2+\nu}\mathbf{1}_{C_{i,0}^n}|\mathcal{F}_{i-1}^n\right]=R(\alpha, h_n^{1+\nu/2}, X_{t_{i-1}^n})$ from (3.7), we have

$$E\left[\left|\frac{1}{\sqrt{nh_n}}\delta_{i,1}^p(\alpha_0)\right|^{2+\nu}|\mathcal{F}_{i-1}^n\right]=\mathrm{O}_p\left(\frac{1}{n^{\nu/2}}\right)+\mathrm{o}_p\left(\frac{h_n}{n^{\nu/2}h_n^{\nu/2}}\right).$$

For $v=2$,

$$\begin{aligned}
&\sum_{i=1}^n E\left[\left|\frac{1}{\sqrt{nh_n}}\delta_{i,2}^p(\alpha_0)\right|^{2+\nu}|\mathcal{F}_{i-1}^n\right]\\
&\quad\leqslant \frac{C_\nu}{n^{1+\nu/2}h_n^{1+\nu/2}}\sum_{i=1}^n E\left[\left|\partial_{\theta_p}\left\{\log\Phi_n(\theta_0, X_{t_{i-1}^n}, \Delta X_i^n)\right\}\varphi_n(X_{t_{i-1}^n}, \Delta X_i^n)\right|^{2+\nu}\times\right.\\
&\quad\quad\left.\times\mathbf{1}_{\{|\Delta X_i^n|>h_n^\rho\}}+h_n^{2+\nu}\left|\int\partial_{\theta_p}\Phi_n(\theta_0, X_{t_{i-1}^n}, y)\,\mathrm{d}y\right|^{2+\nu}\middle|\mathcal{F}_{i-1}^n\right]\\
&\quad\leqslant \frac{C_\nu}{n^{1+\nu/2}h_n^{1+\nu/2}}\\
&\quad\quad\times\sum_{i=1}^n E\left[\left|\partial_{\theta_p}\left\{\log\Phi_n(\theta_0, X_{t_{i-1}^n}, \Delta X_i^n)\right\}\varphi_n(X_{t_{i-1}^n}, \Delta X_i^n)\right|^{2+\nu}\times\right.\\
&\quad\quad\left.\times\left(\mathbf{1}_{D_{i,1}^n}+\mathbf{1}_{D_{i,2}^n}\right)\middle|\mathcal{F}_{i-1}^n\right]+\mathrm{O}_p\left(\frac{h_n^{1+\nu/2}}{n^{\nu/2}}\right).
\end{aligned}$$

Here, it follows from Assumption A.9 (2.1) that

$$\begin{aligned}
\left|\partial_{\theta_p}\left\{\log\Phi_n(\theta_0, X_{t_{i-1}^n}, \Delta X_i^n)\right\}\varphi_n(X_{t_{i-1}^n}, \Delta X_i^n)\right|&\leqslant\varepsilon_n^{-1}L_1(\Delta X_i^n)(1+|X_{t_{i-1}^n}|)^C\\
&=R(\alpha, \varepsilon_n^{-1}, X_{t_{i-1}^n})
\end{aligned}$$

since L_1 is a bounded function given in (2.1). Then we have

$$\begin{aligned}
\sum_{i=1}^n E\left[\left|\frac{1}{\sqrt{nh_n}}\delta_{i,2}^p(\alpha_0)\right|^{2+\nu}|\mathcal{F}_{i-1}^n\right]&\leqslant\frac{\varepsilon_n^{-(2+\nu)}}{n^{1+\nu/2}h_n^{1+\nu/2}}\sum_{i=1}^n R(\alpha, h_n, X_{t_{i-1}^n})+\\
&\quad+\mathrm{O}_p\left(\frac{h_n^{1+\nu/2}}{n^{\nu/2}}\right)\\
&=\mathrm{O}_p\left(\left(\frac{b_n}{nh_n}\right)^{\nu/2}b_n^{(1-2\nu)/5}\right)+\mathrm{O}_p\left(\frac{h_n^{1+\nu/2}}{n^{\nu/2}}\right).
\end{aligned}$$

The last term converges to zero if $\nu\geqslant 1/2$. □

Proof of (4.24).

$$\sum_{i=1}^{n} E\left[\left|\frac{1}{\sqrt{n}}\zeta_i^q(\alpha_0)\right|^{2+\nu}|\mathcal{F}_{i-1}^n\right]$$
$$\leqslant \frac{C_\nu}{n^{1+\nu/2}h_n^{2+\nu}}\sum_{i=1}^{n}\sum_{k,l=1}^{d} E\left[\left|\partial_{\sigma_q}(\beta_{i-1}^{-1})^{(k,l)}(\sigma_0)\right|^{2+\nu}\times\right.$$
$$\left.\times\left|\bar{X}_{i,n}^{(k)}\bar{X}_{i,n}^{(l)}\right|^{2+\nu}\mathbf{1}_{\{|\Delta X_i^n|\leqslant h_n^\rho\}}|\mathcal{F}_{i-1}^n\right]$$
$$+\frac{C_\nu}{n^{1+\nu/2}}\sum_{i=1}^{n}\left|\frac{\partial_q \det\beta_{i-1}(\sigma_0)}{\det\beta_{i-1}(\sigma_0)}\right|^{2+\nu}P\{|\Delta X_i^n|\leqslant h_n^\rho|\mathcal{F}_{i-1}^n\}$$
$$\leqslant \frac{C}{n^{1+\nu/2}h_n^{2+\nu}}\sum_{i=1}^{n}\sum_{k,l=1}^{d}\sum_{j=0}^{2}\left|\partial_{\sigma_q}(\beta_{i-1}^{-1})^{(k,l)}(\sigma_0)\right|^{2+\nu}\times$$
$$\times E\left[\left|\bar{X}_{i,n}^{(k)}\bar{X}_{i,n}^{(l)}\right|^{2+\nu}\mathbf{1}_{C_{i,j}^n}|\mathcal{F}_{i-1}^n\right]+\mathrm{O}_p\left(\frac{1}{n^{\nu/2}}\right).$$

We notice that $E\left[\left|\bar{X}_{i,n}^{(k)}\bar{X}_{i,n}^{(l)}\right|^{2+\nu}\mathbf{1}_{C_{i,0}^n}|\mathcal{F}_{i-1}^n\right]=R(\alpha, h_n^{2+\nu}, X_{t_{i-1}^n})$ from (3.7) and that

$$E\left[\left|\bar{X}_{i,n}^{(k)}\bar{X}_{i,n}^{(l)}\right|^{2+\nu}\mathbf{1}_{C_{i,j}^n}|\mathcal{F}_{i-1}^n\right]\leqslant R(\alpha, h_n^{2\rho(2+\nu)}, X_{t_{i-1}^n})P\{C_{i,j}^n|\mathcal{F}_{i-1}^n\}$$
$$=R(\alpha, h_n^{2\rho(2+\nu)+2}, X_{t_{i-1}^n})$$

for $j=1,2$. Then we have

$$\sum_{i=1}^{n} E\left[\left|\frac{1}{\sqrt{n}}\zeta_i^q(\alpha_0)\right|^{2+\nu}|\mathcal{F}_{i-1}^n\right]=\frac{1}{n^{\nu/2}h_n^{2+\nu}}\mathrm{O}_p\left(h_n^{2\rho(2+\nu)+2}\right)+\mathrm{O}_p\left(\frac{1}{n^{\nu/2}}\right)$$
$$=\mathrm{O}_p\left(\frac{h_n^\mu}{n^{\nu/2}h_n^{\nu/2}}\right)+\mathrm{O}_p\left(\frac{1}{n^{\nu/2}}\right),$$

where $\mu=2\rho(2+\nu)-\nu/2=2(2+\nu)\left(\rho-\frac{\nu}{4(\nu+2)}\right)$. If we take $\nu>0$ sufficiently small, then $\mu>0$ since $2/(\gamma+1)\leqslant\rho<1/2$. This completes the proof. □

References

1. Aase, K. K. and Guttorp, P.: Estimation in models for security prices, *Bernoulli* **4** (1987), 211–224.
2. Billingsley, P.: *Convergence of Probability Measures*, Willey, New York, 1968

3. Barndorff-Nielsen, O. E. and Sørensen, M.: A review of some aspects of asymptotic likelihood theory for stochastic processes, *Int. Stoch. Review* **62** (1994), 133–165.
4. Bibby, B. and Sørensen, M.: Martingale estimation functions for discretely observed diffusion processes, *Bernoulli* **1** (1995), 17–39.
5. Bichteler, K. and Jacod, J.: Calcul de Malliavin pour les diffusions avec sauts: existence d'une densite dans le cas unidimensionnel, *Lecture Notes in Math*, Springer-Verlag, Berlin, 1986, pp. 132–157.
6. Bodo, B. A. and Thomoson, M. E.: A review on stochastic differential equations for applications in hydorology, *Stochastic Hydrol. Hydraul.* **1** (1987), 81–100.
7. Dacunha-Castelle, D. and Florens-Zmirou, D.: Estimation of the coefficient of a diffusion from discrete observations, *Stochastics* **19** (1986) , 263–284.
8. Florens-Zmirou, D.: Approximate discrete schemes for statistics of diffusion processes, *Statistics* **20** (1989), 547–557.
9. Genon-Catalot, V.: Maximum contrast estimation for diffusion processes from discrete observations, *Statistics* **21** (1990), 99–116.
10. Genon-Catalot, V. and Jacod, J.: On the estimation of the diffusion coefficient for multidimensional diffusion process, *Ann. Inst. Henri Poincaré Prob. Statist.* **29** (1993), 119–151.
11. Gukhal, C. R.: Analytical valuation of American options on jump-diffusion processes, *Math. Finance.* **1** (2001), 220–242.
12. Guttrop, P. and Kulperger, R.: Statistical inference for some Volterra population processes in a random environment, *Can. J. Statist.* **12** (1984), 289–302.
13. Hall, P. and Heyde, C.: *Martingale Limit Theory and its Applications*, Academic Press, New York, (1980)
14. Hanson, F. B. and Tuckwell, H. C.: Logistic growth with random density independent disasters, *Theor. Pop. Biol.* **19** (1981), 1–18.
15. Ibragimov, I. A. and Has'minskii, R. Z.: *Statistical Estimation*, Springer-Verlag, Berlin, (1981).
16. Jacod, J. and Shiryayev, A. N.: *Limit Theorems for Stochastic Processes*, Springer-Verlag, Berlin, (1987).
17. Kasonga, R. A.: The consistency of a non-linear least squares estimator from diffusion processes, *Stoch. Proc. Appl.* **24** (1988), 211–229.
18. Kessler, M.: Estimation des paramìtres d'une diffusion par des contrastes corrigés, *C. R. Acad. Sci. Paris Ser. I* **320** (1995), 359–362.
19. Kessler, M.: Estimation of diffusion processes from discrete observations, *Scand. J. Statist.* **24** (1997), 211-229.
20. Kwon, Y. and Lee, C.: Strong Feller property and Ireducibility of diffusions with jumps, *Stoch. Stoch. Rep.* **67** (1999), 147–157.
21. Mulinacci, S.: An approximation of American option prices in a jump-diffusion model, *Stoch. Proc. Appl.* **62** (1996), 1–17.
22. Mtundu, N. D. and Koch, R. W.: A stochastic differential equation approach to soil moisture, *Stochastic Hydrol. Hydraul.* **1** (1987), 101–116.
23. Prakasa-Rao, B. L. S.: Asymptotic theory for non-linear least square estimator for diffusion processes, *Math. Operationsforsch. Statist. Ser. Stat.* **14** (1983), 195–209.
24. Prakasa-Rao, B. L. S.: Statistical inference from sampled data for stochastic processes, *Contemp. Math.* **80** (1988), 249–284.
25. Scott, L. O.: Pricing stock options in a jump-diffusion model with stochastic volatility and interest rates: Applications of Fourier inversion methods, *Math. Finance.* **4** (1997), 413–426.
26. Shiryayev, A. N.: *Probability*. Springer-Verlag, New York-Berlin, (1984).

27. Sørensen, M.: A note on the existence of a consistent maximum likelihood estimator for diffusions with jumps, *Markov Process and Control Theory*, Akademie-Verlag, Berlin, pp. 229–234, 1989.
28. Sørensen, M.: Likelihood methods for diffusions with jumps, *Statistical inference in stochastic processes*, Marcel Dekker, New York, pp. 67–105, 1991.
29. Yoshida, N.: Estimation for diffusion processes from discrete observations, *J. Multivar. Anal.* **41** (1992), 220–242.