Small diffusion asymptotics for discretely sampled stochastic differential equations

[1]Michael Sørensen and [2]Masayuki Uchida

[1]Department of Statistics and Operations Research, University of Copenhagen
Universitetsparken 5, DK-2100 København Ø, Denmark
[2]Faculty of Mathematics, Kyushu University
Ropponmatsu, Fukuoka 810-8560, Japan

Abstract. The minimum contrast estimation of drift and diffusion coefficient parameters for a multi-dimensional diffusion process with a small dispersion parameter ε based on a Gaussian approximation to the transition density is presented in the situation where the sample path is observed at equidistant time points k/n, $k = 0, 1, \cdots, n$. We study asymptotic results for the minimum contrast estimator as ε goes to 0 and n goes to ∞ simultaneously.

AMS Mathematics Subject Classifications: 62F12, 62M05.

Key words and phrases: Diffusion process with small noise, discrete time observation, minimum contrast estimation, parametric inference.

1 Introduction

In this article we consider a family of d-dimensional diffusion processes defined by the following stochastic differential equations

$$\begin{aligned} dX_t &= b(X_t, \alpha)dt + \varepsilon\sigma(X_t, \beta)dw_t,\ t \in [0,1],\ \varepsilon \in (0,1], \\ X_0 &= x_0, \end{aligned} \tag{1}$$

where $(\alpha, \beta) \in \bar{\Theta}_\alpha \times \bar{\Theta}_\beta$ with Θ_α and Θ_β being open bounded convex subsets of $\mathbf{R}^p$ and $\mathbf{R}^q$, respectively. Further, x_0 and ε are known constants, b is an $\mathbf{R}^d$-valued function defined on $\mathbf{R}^d \times \bar{\Theta}_\alpha$, σ is an $\mathbf{R}^d \otimes \mathbf{R}^r$-valued function defined on $\mathbf{R}^d \times \bar{\Theta}_\beta$, and w is an r-dimensional standard Wiener process. We assume that the drift b and the diffusion coefficient σ are known apart from the parameters α and β. The type of data considered in this paper is discrete observations of X at n regularly spaced time points $t_k = k/n$ on the fixed interval $[0,1]$, that is, $(X_{t_k})_{0 \le k \le n}$. We are interested in estimating α and β based on these observations. The type of asymptotics considered is when ε goes to 0 and n goes to ∞ simultaneously.

In case the whole path $X = \{X_t\ ;\ t \in [0,1]\}$ is observed, parametric inference for diffusion type processes with small noises is well developed. The first order asymptotic statistical theory has been studied mainly by Kutoyants [21, 22]. As for higher order asymptotics, Yoshida [33] showed the validity of asymptotic expansions for statistical estimators by means of Malliavin calculus with truncation; see also Yoshida [36, 37, 38], Dermoune and Kutoyants [5], Sakamoto and Yoshida [24], and Uchida and Yoshida [29]. In recent years, also the more realistic case of parametric estimation for discretely observed diffusion processes has been studied by many researchers, see Dacunha-Castelle and Florens-Zmirou [4], Florens-Zmirou [6], Yoshida [35], Genon-Catalot and Jacod [8], Bibby and Sørensen [2, 3], Hansen and Scheinkman [12], Kessler [16, 17], Sørensen [26], Kessler and Sørensen [18], Jacobsen [15] and H. Sørensen [25].

Although small diffusion asymptotics have many applications, (for applications in mathematical finance, see Yoshida [34], Kim and Kunitomo [19], Takahashi [28], Kunitomo and Takahashi [20], Uchida and Yoshida [30]), there is very little work about small noise asymptotics for estimation for diffusion processes from discrete time observations. Genon-Catalot [7] and Laredo [23] studied the efficient estimation of drift parameters of a diffusion process with small noise from discrete observations under the assumptions that diffusion coefficients are known and the asymptotics is when $\varepsilon \to 0$ and $n \to \infty$. Sørensen [27] presented martingale estimation function for discretely observed diffusion processes with small noise, and he showed consistency and asymptotic normality of the estimators of drift and diffusion coefficient parameters when $\varepsilon \to 0$ and n is fixed. Taking account of the above three papers, our goal is to obtain a consistent, asymptotically normal and asymptotically efficient estimator of (α, β) in our setting.

This article is organized as follows. In section 2, we introduce a contrast function based on a Gaussian approximation to the transition density and state several preliminary lemmas. Section 3 presents our main result about the consistency, asymptotic normality and asymptotic efficiency of the minimum contrast estimator obtained from the contrast function constructed in Section 2. Section 4 is devoted to proving the results stated in the previous sections.

2 The contrast function and preliminary lemmas

We first describe the notation and assumptions used in this article.

Suppose that the parameter and the parameter space can be decomposed as follows: $\theta = (\alpha, \beta)$ and $\Theta = \Theta_\alpha \times \Theta_\beta$. Let α_0, β_0 and θ_0 denote the true values of α, β and θ, respectively. Let X_t^0 be the solution of the ordinary differential equation corresponding to $\varepsilon = 0$, i.e. $dX_t^0 = b(X_t^0, \alpha_0)dt$, $X_0^0 = x_0$. For a matrix A, $|A|^2 = \mathrm{tr}(AA^*)$, where "*" indicates the transpose. We denote by $\bar{C}_\uparrow^\infty(\mathbf{R}^d \times \Theta; \mathbf{R}^m)$ the space of all functions f satisfying the following two conditions: (i) $f(x, \theta)$ is an $\mathbf{R}^m$-valued function on $\mathbf{R}^d \times \Theta$ that is smooth in (x, θ), (ii) for $|\mathbf{n}| \geq 0, |\nu| \geq 0$ there exists $C > 0$ such that $\sup_{\theta \in \Theta} |\delta^\nu \partial^\mathbf{n} f| \leq C(1 + |x|)^C$ for all x. Here $\mathbf{n} = (n_1, \cdots, n_d)$ and $\nu = (\nu_1, \cdots, \nu_l)$ are multi-indices, $l = \dim(\Theta)$, $|\mathbf{n}| = n_1 + \cdots + n_d$, $|\nu| = \nu_1 + \cdots + \nu_l$, $\partial^\mathbf{n} = \partial_1^{n_1} \cdots \partial_d^{n_d}$, $\partial_i = \partial/\partial x^i$, $i = 1, \cdots, d$, $\delta^\nu = \delta_1^{\nu_1} \cdots \delta_l^{\nu_l}$, $\delta_j = \partial/\partial \theta^j$, $j = 1, \cdots, l$. Note that ν and δ depend on Θ. For example, $\nu = (\nu_1, \cdots, \nu_p)$, $\delta_j = \partial/\partial \alpha^j$ for $\bar{\Theta}_\alpha$.

In this article, we make the following assumptions.

[A1] Equation (1) has a non-exploding strong solution on $[0, 1]$.

[A2] For all $m > 0$, $\sup_t E[|X_t|^m] < \infty$.

[A3] $b(x, \alpha) \in \bar{C}_\uparrow^\infty(\mathbf{R}^d \times \bar{\Theta}_\alpha; \mathbf{R}^d)$, $\sigma(x, \beta) \in \bar{C}_\uparrow^\infty(\mathbf{R}^d \times \bar{\Theta}_\beta; \mathbf{R}^d \otimes \mathbf{R}^r)$.

[A4] $\inf_{x,\beta} \det[\sigma\sigma^*](x, \beta) > 0$, $[\sigma\sigma^*]^{-1}(x, \beta) \in \bar{C}_\uparrow^\infty(\mathbf{R}^d \times \bar{\Theta}_\beta; \mathbf{R}^d \otimes \mathbf{R}^d)$.

[A5] $\alpha \neq \alpha_0 \Rightarrow b(X_t^0, \alpha) \neq b(X_t^0, \alpha_0)$, $\beta \neq \beta_0 \Rightarrow \sigma\sigma^*(X_t^0, \beta) \neq \sigma\sigma^*(X_t^0, \beta_0)$.

Remark 1 *For [A1], there are several well-known types of sufficient conditions for the existence and uniqueness of a solution of the equation (1). For more details, see Ikeda and Watanabe [14] Chapter IV.*

Moreover, we consider the following assumptions for ε and n.

[B1] $\lim_{\varepsilon \to 0, n \to \infty} (\varepsilon n)^{-1} = 0$.

[B2] $\lim_{\varepsilon \to 0, n \to \infty} (\varepsilon \sqrt{n})^{-1} < \infty$.

Let P_θ be the law of the solution of (1), and L_θ the infinitesimal generator of the diffusion (1):

$$L_\theta f(x) = \sum_{i=1}^d b^i(x, \alpha) \partial_i f(x) + \frac{1}{2} \varepsilon^2 \sum_{i,j=1}^d [\sigma\sigma^*]^{i,j}(x, \beta) \partial_i \partial_j f(x).$$

In order to construct the contrast function, it is natural to consider a Gaussian approximation to the transition density in the same way as in Kessler [16]. Using Lemma 1 in Florens-Zmirou [6], we obtain the following contrast function.

$$U_{\varepsilon,n}(\theta) = \sum_{k=1}^n \{\log \det \Xi_{k-1}(\beta) + \varepsilon^{-2} n P_k^*(\alpha) \Xi_{k-1}(\beta)^{-1} P_k(\alpha)\},$$

where $P_k(\alpha) = X_{t_k} - X_{t_{k-1}} - \dfrac{1}{n} b(X_{t_{k-1}}, \alpha)$, $\Xi_k(\beta) = [\sigma\sigma^*](X_{t_k}, \beta)$. Let R denote a function $(0,1] \times \mathbf{R}^d \to \mathbf{R}$ for which there exists a constant C such that $R(a,x) \leq a(1+|x|)^C$ for all a, x. We define $\mathcal{G}_k^n = \sigma(w_s; s \leq t_k)$, $B_k^i(\alpha_0, \alpha) = b^i(X_{t_k}, \alpha_0) - b^i(X_{t_k}, \alpha)$, and $B(x, \alpha_0, \alpha) = b(x, \alpha_0) - b(x, \alpha)$. Moreover, in order to formulate the preliminary lemmas given later, we prepare several functions and notation as follows. For Lemma 4, we define

$$
\begin{aligned}
U_1(\alpha, \alpha_0, \beta) &= \int_0^1 B^*(X_s^0, \alpha_0, \alpha)[\sigma\sigma^*]^{-1}(X_s^0, \beta) B(X_s^0, \alpha_0, \alpha) ds, \\
U_2(\alpha, \beta, \beta_0) &= \int_0^1 \log\det[\sigma\sigma^*](X_s^0, \beta) ds \\
&\quad + \int_0^1 \mathrm{tr}\left[[\sigma\sigma^*](X_s^0, \beta_0)[\sigma\sigma^*]^{-1}(X_s^0, \beta)\right] ds \\
&\quad + M^2 \int_0^1 B^*(X_s^0, \alpha_0, \alpha)[\sigma\sigma^*]^{-1}(X_s^0, \beta) B(X_s^0, \alpha_0, \alpha) ds,
\end{aligned}
$$

where $M = \lim_{\varepsilon \to 0, n \to \infty} (\varepsilon\sqrt{n})^{-1}$. Note that U_2 is only well-defined under assumption [B2]. For Lemma 5, let

$$
C_{\varepsilon,n}(\theta_0) = \begin{pmatrix} \varepsilon^2 \left(\frac{\partial^2}{\partial\alpha_i \partial\alpha_j} U_{\varepsilon,n}(\theta_0) \right)_{1 \leq i,j \leq p} & \varepsilon \frac{1}{\sqrt{n}} \left(\frac{\partial^2}{\partial\alpha_i \partial\beta_j} U_{\varepsilon,n}(\theta_0) \right)_{1 \leq i \leq p, 1 \leq j \leq q} \\ \varepsilon \frac{1}{\sqrt{n}} \left(\frac{\partial^2}{\partial\beta_i \partial\alpha_j} U_{\varepsilon,n}(\theta_0) \right)_{1 \leq i \leq q, 1 \leq j \leq p} & \frac{1}{n} \left(\frac{\partial^2}{\partial\beta_i \partial\beta_j} U_{\varepsilon,n}(\theta_0) \right)_{1 \leq i,j \leq q} \end{pmatrix},
$$

$$
I(\theta_0) = \begin{pmatrix} \left(I_b^{i,j}(\theta_0) \right)_{1 \leq i,j \leq p} & 0 \\ 0 & \left(I_\sigma^{i,j}(\theta_0) \right)_{1 \leq i,j \leq q} \end{pmatrix},
$$

where

$$
\begin{aligned}
I_b^{i,j}(\theta_0) &= \int_0^1 \left(\frac{\partial}{\partial\alpha_i} b(X_s^0, \alpha_0) \right)^* [\sigma\sigma^*]^{-1}(X_s^0, \beta_0) \left(\frac{\partial}{\partial\alpha_j} b(X_s^0, \alpha_0) \right) ds, \\
I_\sigma^{i,j}(\theta_0) &= \frac{1}{2} \int_0^1 \mathrm{tr}\left[\left(\frac{\partial}{\partial\beta_i} [\sigma\sigma^*] \right) [\sigma\sigma^*]^{-1} \left(\frac{\partial}{\partial\beta_j} [\sigma\sigma^*] \right) [\sigma\sigma^*]^{-1}(X_s^0, \beta_0) \right] ds.
\end{aligned}
$$

For Lemma 6, define

$$
\Lambda_{\varepsilon,n} = \begin{pmatrix} -\varepsilon \left(\frac{\partial}{\partial\alpha_i} U_{\varepsilon,n}(\theta_0) \right)_{1 \leq i \leq p} \\ -\frac{1}{\sqrt{n}} \left(\frac{\partial}{\partial\beta_j} U_{\varepsilon,n}(\theta_0) \right)_{1 \leq j \leq q} \end{pmatrix}.
$$

Lemma 1 *Suppose that [A1]–[A3] hold true. Then*
(i)

$$
E_{\theta_0}[P_k^i(\alpha) | \mathcal{G}_{k-1}^n] = \frac{1}{n} B_{k-1}^i(\alpha_0, \alpha) + R\left(\frac{1}{n^2}, X_{t_{k-1}} \right).
$$

(ii)

$$
\begin{aligned}
E_{\theta_0}[P_k^{i_1}(\alpha) P_k^{i_2}(\alpha) | \mathcal{G}_{k-1}^n] &= \frac{\varepsilon^2}{n} \Xi_{k-1}^{i_1 i_2}(\beta_0) + \frac{1}{n^2} B_{k-1}^{i_1} B_{k-1}^{i_2}(\alpha_0, \alpha) \\
&\quad + R\left(\frac{\varepsilon^2}{n^2}, X_{t_{k-1}} \right) + R\left(\frac{1}{n^3}, X_{t_{k-1}} \right).
\end{aligned}
$$

(iii)

$$
\begin{aligned}
E_{\theta_0}[P_k^{i_1}(\alpha)P_k^{i_2}(\alpha)P_k^{i_3}(\alpha)|\mathcal{G}_{k-1}^n] \;\; = \;\; & \frac{\varepsilon^2}{n^2}\{\Xi_{k-1}^{i_1i_2}(\beta_0)B_{k-1}^{i_3}(\alpha_0,\alpha)+\Xi_{k-1}^{i_1i_3}(\beta_0)B_{k-1}^{i_2}(\alpha_0,\alpha) \\
& +\Xi_{k-1}^{i_2i_3}(\beta_0)B_{k-1}^{i_1}(\alpha_0,\alpha)\} \\
& +\frac{1}{n^3}B_{k-1}^{i_1}B_{k-1}^{i_2}B_{k-1}^{i_3}(\alpha_0,\alpha) \\
& +R\left(\frac{\varepsilon^4}{n^2},X_{t_{k-1}}\right)+R\left(\frac{\varepsilon^2}{n^3},X_{t_{k-1}}\right)+R\left(\frac{1}{n^4},X_{t_{k-1}}\right).
\end{aligned}
$$

(iv)

$$
\begin{aligned}
E_{\theta_0}\left[\prod_{j=1}^4 P_k^{i_j}(\alpha)|\mathcal{G}_{k-1}^n\right] \;\; = \;\; & \frac{\varepsilon^4}{n^2}\left\{\Xi_{k-1}^{i_1i_2}\Xi_{k-1}^{i_3i_4}(\beta_0)+\Xi_{k-1}^{i_1i_3}\Xi_{k-1}^{i_2i_4}(\beta_0)+\Xi_{k-1}^{i_1i_4}\Xi_{k-1}^{i_2i_3}(\beta_0)\right\} \\
& +\frac{\varepsilon^2}{n^3}\left\{\Xi_{k-1}^{i_1i_2}(\beta_0)B_{k-1}^{i_3}B_{k-1}^{i_4}(\alpha_0,\alpha)+\Xi_{k-1}^{i_1i_3}(\beta_0)B_{k-1}^{i_2}B_{k-1}^{i_4}(\alpha_0,\alpha)\right. \\
& +\Xi_{k-1}^{i_1i_4}(\beta_0)B_{k-1}^{i_2}B_{k-1}^{i_3}(\alpha_0,\alpha)+\Xi_{k-1}^{i_2i_3}(\beta_0)B_{k-1}^{i_1}B_{k-1}^{i_4}(\alpha_0,\alpha) \\
& \left.+\Xi_{k-1}^{i_2i_4}(\beta_0)B_{k-1}^{i_1}B_{k-1}^{i_3}(\alpha_0,\alpha)+\Xi_{k-1}^{i_3i_4}(\beta_0)B_{k-1}^{i_1}B_{k-1}^{i_2}(\alpha_0,\alpha)\right\} \\
& +\frac{1}{n^4}B_{k-1}^{i_1}B_{k-1}^{i_2}B_{k-1}^{i_3}B_{k-1}^{i_4}(\alpha_0,\alpha) \\
& +R\left(\frac{\varepsilon^4}{n^3},X_{t_{k-1}}\right)+R\left(\frac{\varepsilon^2}{n^4},X_{t_{k-1}}\right)+R\left(\frac{1}{n^5},X_{t_{k-1}}\right).
\end{aligned}
$$

Lemma 2 *Let $f\in\bar{C}_{\uparrow}^{\infty}(\mathbf{R}^d\times\bar{\Theta};\mathbf{R})$. Assume [A1]–[A3]. Then, under P_{θ_0},*
(i)

$$\frac{1}{n}\sum_{k=1}^n f(X_{t_{k-1}},\theta)\to\int_0^1 f(X_s^0,\theta)ds$$

as $\varepsilon\to 0$ and $n\to\infty$, uniformly in $\theta\in\bar{\Theta}$, and
(ii)

$$\sum_{k=1}^n f(X_{t_{k-1}},\theta)P_k^i(\alpha_0)\to 0$$

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

Lemma 3 *Let $f\in\bar{C}_{\uparrow}^{\infty}(\mathbf{R}^d\times\bar{\Theta};\mathbf{R})$. Assume [A1]–[A3] and [B1]. Then, under P_{θ_0},*
(i)

$$\varepsilon^{-2}\sum_{k=1}^n f(X_{t_{k-1}},\theta)P_k^iP_k^j(\alpha_0) \;\;\to\;\; \int_0^1 f(X_s^0,\theta)[\sigma\sigma^*]^{ij}(X_s^0,\beta_0)ds$$

as $\varepsilon\to 0$ and $n\to\infty$, uniformly in $\theta\in\bar{\Theta}$. Moreover, if [B2] holds true, then, under P_{θ_0},
(ii)

$$
\begin{aligned}
\varepsilon^{-2}\sum_{k=1}^n f(X_{t_{k-1}},\theta)P_k^iP_k^j(\alpha) \;\;\to\;\; & \int_0^1 f(X_s^0,\theta)[\sigma\sigma^*]^{ij}(X_s^0,\beta_0)ds \\
& +M^2\int_0^1 f(X_s^0,\theta)B^iB^j(X_s^0,\alpha_0,\alpha)ds
\end{aligned}
$$

as $\varepsilon \to 0$ and $n \to \infty$, uniformly in $\theta \in \bar{\Theta}$, where $M = \lim_{\varepsilon\to 0, n\to\infty}(\varepsilon\sqrt{n})^{-1}$.

Lemma 4 *Assume [A1]–[A4]. Then, under P_{θ_0}, as $\varepsilon \to 0$ and $n \to \infty$,*
(i)

$$\sup_{\theta\in\bar{\Theta}} \left|\varepsilon^2\{U_{\varepsilon,n}(\alpha,\beta) - U_{\varepsilon,n}(\alpha_0,\beta)\} - U_1(\alpha,\alpha_0,\beta)\right| \to 0.$$

(ii) Moreover, suppose that [B2] holds true. Then, under P_{θ_0}, as $\varepsilon \to 0$ and $n \to \infty$,

$$\sup_{\theta\in\bar{\Theta}} \left|\frac{1}{n}U_{\varepsilon,n}(\alpha,\beta) - U_2(\alpha,\beta,\beta_0)\right| \to 0.$$

Lemma 5 *Assume [A1]–[A4] and [B2]. Then, under P_{θ_0}, as $\varepsilon \to 0$ and $n \to \infty$,*
(i)

$$C_{\varepsilon,n}(\theta_0) \to 2I(\theta_0),$$

(ii)

$$\sup_{|\theta|\le\eta_{\varepsilon,n}} |C_{\varepsilon,n}(\theta_0+\theta) - C_{\varepsilon,n}(\theta_0)| \to 0,$$

where $\eta_{\varepsilon,n} \to 0$.

Lemma 6 *Assume [A1]–[A4] and [B2]. Then*

$$\Lambda_{\varepsilon,n} \to N\left(0, 4I(\theta_0)\right)$$

in distribution, under P_{θ_0}, as $\varepsilon \to 0$ and $n \to \infty$.

3 The main result

Let $\hat{\theta}_{\varepsilon,n} = (\hat{\alpha}_{\varepsilon,n}, \hat{\beta}_{\varepsilon,n})$ be a minimum contrast estimator defined by

$$U_{\varepsilon,n}(\hat{\theta}_{\varepsilon,n}) = \inf_{\theta\in\bar{\Theta}} U_{\varepsilon,n}(\theta). \tag{2}$$

Our main theorem is as follows.

Theorem 1 *Assume [A1]–[A5] and [B2]. Then,*

$$\hat{\theta}_{\varepsilon,n} \to \theta_0$$

in P_{θ_0}-probability as $\varepsilon \to 0$ and $n \to \infty$. Moreover, if $\theta_0 \in \Theta$ and $I(\theta_0)$ is positive definite,

$$\begin{pmatrix} \varepsilon^{-1}(\hat{\alpha}_{\varepsilon,n} - \alpha_0) \\ \sqrt{n}(\hat{\beta}_{\varepsilon,n} - \beta_0) \end{pmatrix} \to N\left(0, I(\theta_0)^{-1}\right)$$

in distribution, under P_{θ_0}, as $\varepsilon \to 0$ and $n \to \infty$.

Remark 2 *(i) Let $P_n^{\alpha,\beta}$ be the restriction of $P_{\alpha,\beta}$ to $\mathcal{F}_n = \sigma(X_{t_k} : 0 \le k \le n)$. In the same way as in Gobet [9, 10], under regularity conditions, we can obtain the Local Asymptotic Normality for the likelihoods as follows: For every $u \in \mathbf{R}^p$ and $v \in \mathbf{R}^q$, under P_{θ_0},*

$$\log\left(\frac{dP_n^{\alpha_0+\varepsilon u,\beta_0+\frac{v}{\sqrt{n}}}}{dP_n^{\alpha_0,\beta_0}}\right)((X_{t_k})_{0\le k\le n}) \to \begin{pmatrix} u \\ v \end{pmatrix}^* \mathcal{N} - \frac{1}{2}\begin{pmatrix} u \\ v \end{pmatrix}^* I(\theta_0)\begin{pmatrix} u \\ v \end{pmatrix}$$

in distribution as $\varepsilon \to 0$ and $n \to \infty$, where $\mathcal{N}$ is a centered Gaussian variable with covariance matrix $I(\theta_0)$. For details, see Uchida [31]. If $I(\theta_0)$ is non-singular, it follows from minimax theorems that $I(\theta_0)^{-1}$ gives the lower bound for the asymptotic variance of regular estimators. This together with Theorem 1 shows that the estimator given by (2) is asymptotically efficient.

(ii) It is worth mentioning that the estimators of the drift and diffusion coefficient parameters in Theorem 1 are asymptotically independent.

(iii) Note also that when $(\varepsilon\sqrt{n})^{-1} \to 0$ the rate of convergence is different for drift and diffusion coefficient parameters. The estimator of the diffusion coefficient parameter converges more quickly than the estimator of the drift parameter because it utilizes information about the diffusion coefficient in the fine-structure of the continuous sample path.

When $\sigma(x,\beta) = \sigma(x)$, Theorem 1 holds under slightly weaken conditions. Let $C_\uparrow^\infty(\mathbf{R}^d; \mathbf{R}^d \otimes \mathbf{R}^r)$ be the set of all functions f of class $C^\infty(\mathbf{R}^d; \mathbf{R}^d \otimes \mathbf{R}^r)$ such that f and all of its derivatives have polynomial growth. Instead of assumptions [A3]–[A5], we suppose the following assumptions.

[A3'] $b(x,\alpha) \in \bar{C}_\uparrow^\infty(\mathbf{R}^d \times \bar{\Theta}_\alpha; \mathbf{R}^d)$, $\sigma(x) \in C_\uparrow^\infty(\mathbf{R}^d; \mathbf{R}^d \otimes \mathbf{R}^r)$.

[A4'] $\inf_x \det[\sigma\sigma^*](x) > 0$, $[\sigma\sigma^*]^{-1}(x) \in C_\uparrow^\infty(\mathbf{R}^d; \mathbf{R}^d \otimes \mathbf{R}^d)$.

[A5'] $\alpha \ne \alpha_0 \Rightarrow b(X_t^0, \alpha) \ne b(X_t^0, \alpha_0)$.

Set $\tilde{I}_b(\alpha_0) = \left(\tilde{I}_b^{i,j}(\alpha_0)\right)_{1\le i,j\le p}$ and $\tilde{I}_b^{i,j}(\alpha_0) = \int_0^1 \left(\frac{\partial}{\partial\alpha_i} b(X_s^0,\alpha_0)\right)^* [\sigma\sigma^*]^{-1}(X_s^0) \left(\frac{\partial}{\partial\alpha_j} b(X_s^0,\alpha_0)\right) ds$. We consider the following contrast function:

$$\tilde{U}_{\varepsilon,n}(\alpha) = \varepsilon^{-2} n \sum_{k=1}^n P_k^*(\alpha)[\sigma\sigma^*]^{-1}(X_{t_{k-1}})P_k(\alpha),$$

and let $\hat{\alpha}_{\varepsilon,n}$ be a minimum contrast estimator defined by

$$\tilde{U}_{\varepsilon,n}(\hat{\alpha}_{\varepsilon,n}) = \inf_{\alpha\in\bar{\Theta}_\alpha} \tilde{U}_{\varepsilon,n}(\alpha).$$

Corollary 1 *Suppose $\sigma(x,\beta) = \sigma(x)$, and assume [A1], [A2], [A3']–[A5'] and [B1]. Then,*

$$\hat{\alpha}_{\varepsilon,n} \to \alpha_0$$

in P_{α_0}-probability as $\varepsilon \to 0$ and $n \to \infty$. Moreover, if $\alpha_0 \in \Theta_\alpha$ and $\tilde{I}_b(\alpha_0)$ is positive definite,

$$\varepsilon^{-1}(\hat{\alpha}_{\varepsilon,n} - \alpha_0) \to N\left(0, \tilde{I}_b(\alpha_0)^{-1}\right)$$

in distribution, under P_{θ_0}, as $\varepsilon \to 0$ and $n \to \infty$.

4 Proofs

Proof of Lemma 1. In the same way as Lemma 7 in Kessler [16], we prove Lemma 1. Let $\phi_j(x,y) = \prod_{l=1}^{j}(y^{i_l} - x^{i_l})$.
(i)

$$\begin{aligned}
E_{\theta_0}[\phi_1(X_{t_{k-1}}, X_{t_k})|\mathcal{G}_{k-1}^n] &= \phi_1(X_{t_{k-1}}, X_{t_{k-1}}) + \frac{1}{n}L_{\theta_0}\phi_1(X_{t_{k-1}}, X_{t_{k-1}}) \\
&\quad + \int_0^{\frac{1}{n}}\int_0^{u_1} E[L_{\theta_0}^2\phi_1(X_{t_{k-1}}, X_{t_{k-1}+u_2})|\mathcal{G}_{k-1}^n]du_2du_1 \\
&= \frac{1}{n}b^{i_1}(X_{t_{k-1}}, \alpha_0) + R\left(\frac{1}{n^2}, X_{t_{k-1}}\right).
\end{aligned}$$

Thus, one has

$$\begin{aligned}
E_{\theta_0}[P_k^{i_1}(\alpha)|\mathcal{G}_{k-1}^n] &= E_{\theta_0}[\phi_1(X_{t_{k-1}}, X_{t_k})|\mathcal{G}_{k-1}^n] - \frac{1}{n}b^{i_1}(X_{t_{k-1}}, \alpha) \\
&= \frac{1}{n}B_{k-1}^{i_1}(\alpha_0, \alpha) + R\left(\frac{1}{n^2}, X_{t_{k-1}}\right).
\end{aligned}$$

∎

(ii)

$$\begin{aligned}
E_{\theta_0}[\phi_2(X_{t_{k-1}}, X_{t_k})|\mathcal{G}_{k-1}^n] &= \phi_2(X_{t_{k-1}}, X_{t_{k-1}}) + \frac{1}{n}L_{\theta_0}\phi_2(X_{t_{k-1}}, X_{t_{k-1}}) \\
&\quad + \frac{1}{2n^2}L_{\theta_0}^2\phi_2(X_{t_{k-1}}, X_{t_{k-1}}) \\
&\quad + \int_0^{\frac{1}{n}}\int_0^{u_1}\int_0^{u_2} E_{\theta_0}[L_{\theta_0}^3\phi_2(X_{t_{k-1}}, X_{t_{k-1}+u_3})|\mathcal{G}_{k-1}^n]du_3du_2du_1 \\
&= \frac{\varepsilon^2}{n}\Xi_{k-1}^{i_1i_2}(\beta_0) + \frac{1}{n^2}b^{i_1}b^{i_2}(X_{t_{k-1}}, \alpha_0) \\
&\quad + R\left(\frac{\varepsilon^2}{n^2}, X_{t_{k-1}}\right) + R\left(\frac{1}{n^3}, X_{t_{k-1}}\right).
\end{aligned}$$

Thus, if $\phi^i(x,y) = (y^i - x^i)$, one has

$$\begin{aligned}
E_{\theta_0}[P_k^{i_1}P_k^{i_2}(\alpha)|\mathcal{G}_{k-1}^n] &= E_{\theta_0}[\phi_2(X_{t_{k-1}}, X_{t_k})|\mathcal{G}_{k-1}^n] \\
&\quad - E_{\theta_0}[\phi^{i_1}(X_{t_{k-1}}, X_{t_k})|\mathcal{G}_{k-1}^n]\frac{1}{n}b^{i_2}(X_{t_{k-1}}, \alpha) \\
&\quad - E_{\theta_0}[\phi^{i_2}(X_{t_{k-1}}, X_{t_k})|\mathcal{G}_{k-1}^n]\frac{1}{n}b^{i_1}(X_{t_{k-1}}, \alpha) \\
&\quad + \frac{1}{n^2}b^{i_1}b^{i_2}(X_{t_{k-1}}, \alpha) \\
&= \frac{\varepsilon^2}{n}\Xi_{k-1}^{i_1i_2}(\beta_0) + \frac{1}{n^2}b^{i_1}b^{i_2}(X_{t_{k-1}}, \alpha_0) + R\left(\frac{\varepsilon^2}{n^2}, X_{t_{k-1}}\right) + R\left(\frac{1}{n^3}, X_{t_{k-1}}\right) \\
&\quad - \left\{\frac{1}{n}b^{i_1}(X_{t_{k-1}}, \alpha_0) + R\left(\frac{1}{n^2}, X_{t_{k-1}}\right)\right\}\frac{1}{n}b^{i_2}(X_{t_{k-1}}, \alpha)
\end{aligned}$$

$$
\begin{aligned}
& -\left\{\frac{1}{n}b^{i_2}(X_{t_{k-1}},\alpha_0)+R\left(\frac{1}{n^2},X_{t_{k-1}}\right)\right\}\frac{1}{n}b^{i_1}(X_{t_{k-1}},\alpha) \\
& +\frac{1}{n^2}b^{i_1}b^{i_2}(X_{t_{k-1}},\alpha) \\
= \; & \frac{\varepsilon^2}{n}\Xi^{i_1i_2}_{k-1}(\beta_0)+\frac{1}{n^2}B^{i_1}_{k-1}B^{i_2}_{k-1}(\alpha_0,\alpha) \\
& +R\left(\frac{\varepsilon^2}{n^2},X_{t_{k-1}}\right)+R\left(\frac{1}{n^3},X_{t_{k-1}}\right).
\end{aligned}
$$

∎

(iii)

$$
\begin{aligned}
E_{\theta_0}[\phi_3(X_{t_{k-1}},X_{t_k})|\mathcal{G}^n_{k-1}] &= \sum_{l=0}^{3}\frac{1}{l!n^l}L^l_{\theta_0}\phi_3(X_{t_{k-1}},X_{t_{k-1}}) \\
&\quad +\int_0^{\frac{1}{n}}\int_0^{u_1}\int_0^{u_2}\int_0^{u_3}E_{\theta_0}[L^4_{\theta_0}\phi_3(X_{t_{k-1}},X_{t_{k-1}+u_4})|\mathcal{G}^n_{k-1}]du_4du_3du_2du_1 \\
&= \frac{\varepsilon^2}{n^2}\Big\{\Xi^{i_1i_2}_{k-1}(\beta_0)b^{i_3}(X_{t_{k-1}},\alpha_0)+\Xi^{i_1i_3}_{k-1}(\beta_0)b^{i_2}(X_{t_{k-1}},\alpha_0) \\
&\quad +\Xi^{i_2i_3}_{k-1}(\beta_0)b^{i_1}(X_{t_{k-1}},\alpha_0)\Big\} \\
&\quad +\frac{1}{n^3}b^{i_1}b^{i_2}b^{i_3}(X_{t_{k-1}},\alpha_0) \\
&\quad +R\left(\frac{\varepsilon^4}{n^2},X_{t_{k-1}}\right)+R\left(\frac{\varepsilon^2}{n^3},X_{t_{k-1}}\right)+R\left(\frac{1}{n^4},X_{t_{k-1}}\right).
\end{aligned}
$$

If $\phi^{ij}(x,y)=(y^i-x^i)(y^j-x^j)$, one has

$$
\begin{aligned}
P^{i_1}_kP^{i_2}_kP^{i_3}_k(\alpha) &= \phi_3(X_{t_{k-1}},X_{t_k}) \\
&\quad -\phi^{i_1i_3}(X_{t_{k-1}},X_{t_k})\frac{1}{n}b^{i_2}(X_{t_{k-1}},\alpha) \\
&\quad -\phi^{i_2i_3}(X_{t_{k-1}},X_{t_k})\frac{1}{n}b^{i_1}(X_{t_{k-1}},\alpha) \\
&\quad +\phi^{i_3}(X_{t_{k-1}},X_{t_k})\frac{1}{n^2}b^{i_1}b^{i_2}(X_{t_{k-1}},\alpha) \\
&\quad -P^{i_1}_kP^{i_2}_k(\alpha)\frac{1}{n}b^{i_3}(X_{t_{k-1}},\alpha).
\end{aligned}
$$

Thus,

$$
\begin{aligned}
E_{\theta_0}[P^{i_1}_kP^{i_2}_kP^{i_3}_k(\alpha)|\mathcal{G}^n_{k-1}] &= E_{\theta_0}[\phi_3(X_{t_{k-1}},X_{t_k})|\mathcal{G}^n_{k-1}] \\
&\quad -\frac{\varepsilon^2}{n^2}\Xi^{i_1i_3}_{k-1}(\beta_0)b^{i_2}(X_{t_{k-1}},\alpha)-\frac{1}{n^3}b^{i_1}b^{i_3}(X_{t_{k-1}},\alpha)b^{i_2}(X_{t_{k-1}},\alpha_0) \\
&\quad -\frac{\varepsilon^2}{n^2}\Xi^{i_2i_3}_{k-1}(\beta_0)b^{i_1}(X_{t_{k-1}},\alpha)-\frac{1}{n^3}b^{i_2}b^{i_3}(X_{t_{k-1}},\alpha)b^{i_1}(X_{t_{k-1}},\alpha_0) \\
&\quad +\frac{1}{n^3}b^{i_1}b^{i_2}(X_{t_{k-1}},\alpha)b^{i_3}(X_{t_{k-1}},\alpha_0)-\frac{\varepsilon^2}{n^2}\Xi^{i_1i_2}_{k-1}(\beta_0)b^{i_3}(X_{t_{k-1}},\alpha) \\
&\quad -\frac{1}{n^3}B^{i_1}_{k-1}B^{i_2}_{k-1}(\alpha,\alpha_0)b^{i_3}(X_{t_{k-1}},\alpha)
\end{aligned}
$$

$$
\begin{aligned}
&\quad +R\left(\frac{\varepsilon^2}{n^3}, X_{t_{k-1}}\right)+R\left(\frac{1}{n^4}, X_{t_{k-1}}\right) \\
= &\quad \frac{\varepsilon^2}{n^2}\Big\{\Xi_{k-1}^{i_1 i_2}(\beta_0)B_{k-1}^{i_3}(\alpha_0,\alpha)+\Xi_{k-1}^{i_1 i_3}(\beta_0)B_{k-1}^{i_2}(\alpha_0,\alpha) \\
&\quad +\Xi_{k-1}^{i_2 i_3}(\beta_0)B_{k-1}^{i_1}(\alpha_0,\alpha)\Big\} \\
&\quad +\frac{1}{n^3}B_{k-1}^{i_1}B_{k-1}^{i_2}B_{k-1}^{i_3}(\alpha_0,\alpha) \\
&\quad +R\left(\frac{\varepsilon^4}{n^2}, X_{t_{k-1}}\right)+R\left(\frac{\varepsilon^2}{n^3}, X_{t_{k-1}}\right)+R\left(\frac{1}{n^4}, X_{t_{k-1}}\right).
\end{aligned}
$$

∎

(iv)

$$
\begin{aligned}
E_{\theta_0}[\phi_4(X_{t_{k-1}}, X_{t_k})|\mathcal{G}_{k-1}^n] &= \frac{\varepsilon^4}{n^2}\Big\{\Xi_{k-1}^{i_1 i_2}\Xi_{k-1}^{i_3 i_4}(\beta_0)+\Xi_{k-1}^{i_1 i_3}\Xi_{k-1}^{i_2 i_4}(\beta_0)+\Xi_{k-1}^{i_1 i_4}\Xi_{k-1}^{i_2 i_3}(\beta_0)\Big\} \\
&\quad +\frac{\varepsilon^2}{n^3}\Big\{\Xi_{k-1}^{i_1 i_2}(\beta_0)b^{i_3}b^{i_4}(X_{t_{k-1}},\alpha_0) \\
&\quad +\Xi_{k-1}^{i_1 i_3}(\beta_0)b^{i_2}b^{i_4}(X_{t_{k-1}},\alpha_0) \\
&\quad +\Xi_{k-1}^{i_1 i_4}(\beta_0)b^{i_2}b^{i_3}(X_{t_{k-1}},\alpha_0) \\
&\quad +\Xi_{k-1}^{i_2 i_3}(\beta_0)b^{i_1}b^{i_4}(X_{t_{k-1}},\alpha_0) \\
&\quad +\Xi_{k-1}^{i_2 i_4}(\beta_0)b^{i_1}b^{i_3}(X_{t_{k-1}},\alpha_0) \\
&\quad +\Xi_{k-1}^{i_3 i_4}(\beta_0)b^{i_1}b^{i_2}(X_{t_{k-1}},\alpha_0)\Big\} \\
&\quad +\frac{1}{n^4}b^{i_1}b^{i_2}b^{i_3}b^{i_4}(X_{t_{k-1}},\alpha_0) \\
&\quad +R\left(\frac{\varepsilon^4}{n^3}, X_{t_{k-1}}\right)+R\left(\frac{\varepsilon^2}{n^4}, X_{t_{k-1}}\right)+R\left(\frac{1}{n^5}, X_{t_{k-1}}\right).
\end{aligned}
$$

If $\phi^{i_1 i_2 i_3}(x,y)=(y^{i_1}-x^{i_1})(y^{i_2}-x^{i_2})(y^{i_3}-x^{i_3})$, one has

$$
\begin{aligned}
E_{\theta_0}[P_k^{i_1}P_k^{i_2}P_k^{i_3}P_k^{i_4}(\alpha)|\mathcal{G}_{k-1}^n] &= E_{\theta_0}[\phi_4(X_{t_{k-1}}, X_{t_k})|\mathcal{G}_{k-1}^n] \\
&\quad -E_{\theta_0}[\phi^{i_1 i_2 i_4}(X_{t_{k-1}}, X_{t_k})|\mathcal{G}_{k-1}^n]\frac{1}{n}b^{i_3}(X_{t_{k-1}},\alpha) \\
&\quad -E_{\theta_0}[\phi^{i_1 i_3 i_4}(X_{t_{k-1}}, X_{t_k})|\mathcal{G}_{k-1}^n]\frac{1}{n}b^{i_2}(X_{t_{k-1}},\alpha) \\
&\quad -E_{\theta_0}[\phi^{i_2 i_3 i_4}(X_{t_{k-1}}, X_{t_k})|\mathcal{G}_{k-1}^n]\frac{1}{n}b^{i_1}(X_{t_{k-1}},\alpha) \\
&\quad +E_{\theta_0}[\phi^{i_1 i_4}(X_{t_{k-1}}, X_{t_k})|\mathcal{G}_{k-1}^n]\frac{1}{n^2}b^{i_2}b^{i_3}(X_{t_{k-1}},\alpha) \\
&\quad +E_{\theta_0}[\phi^{i_2 i_4}(X_{t_{k-1}}, X_{t_k})|\mathcal{G}_{k-1}^n]\frac{1}{n^2}b^{i_1}b^{i_3}(X_{t_{k-1}},\alpha) \\
&\quad +E_{\theta_0}[\phi^{i_3 i_4}(X_{t_{k-1}}, X_{t_k})|\mathcal{G}_{k-1}^n]\frac{1}{n^2}b^{i_1}b^{i_2}(X_{t_{k-1}},\alpha) \\
&\quad -E_{\theta_0}[\phi^{i_4}(X_{t_{k-1}}, X_{t_k})|\mathcal{G}_{k-1}^n]\frac{1}{n^3}b^{i_1}b^{i_2}b^{i_3}(X_{t_{k-1}},\alpha)
\end{aligned}
$$

$$
\begin{aligned}
& -E_{\theta_0}[P_k^{i_1} P_k^{i_2} P_k^{i_3}(\alpha)|\mathcal{G}_{k-1}^n]\frac{1}{n}b^{i_4}(X_{t_{k-1}},\alpha) \\
= \quad & \frac{\varepsilon^4}{n^2}\left\{\Xi_{k-1}^{i_1 i_2}\Xi_{k-1}^{i_3 i_4}(\beta_0) + \Xi_{k-1}^{i_1 i_3}\Xi_{k-1}^{i_2 i_4}(\beta_0) + \Xi_{k-1}^{i_1 i_4}\Xi_{k-1}^{i_2 i_3}(\beta_0)\right\} \\
& +\frac{\varepsilon^2}{n^3}\left\{\Xi_{k-1}^{i_1 i_2}(\beta_0)B_{k-1}^{i_3}(\alpha_0,\alpha)b^{i_4}(X_{t_{k-1}},\alpha_0)\right. \\
& -\Xi_{k-1}^{i_1 i_2}(\beta_0)B_{k-1}^{i_3}(\alpha_0,\alpha)b^{i_4}(X_{t_{k-1}},\alpha) \\
& +\Xi_{k-1}^{i_1 i_3}(\beta_0)B_{k-1}^{i_2}(\alpha_0,\alpha)b^{i_4}(X_{t_{k-1}},\alpha_0) \\
& -\Xi_{k-1}^{i_1 i_3}(\beta_0)B_{k-1}^{i_2}(\alpha_0,\alpha)b^{i_4}(X_{t_{k-1}},\alpha) \\
& +\Xi_{k-1}^{i_1 i_4}(\beta_0)B_{k-1}^{i_3}(\alpha_0,\alpha)b^{i_2}(X_{t_{k-1}},\alpha_0) \\
& -\Xi_{k-1}^{i_1 i_4}(\beta_0)b^{i_2}(X_{t_{k-1}},\alpha)b^{i_3}(X_{t_{k-1}},\alpha_0) \\
& +\Xi_{k-1}^{i_1 i_4}(\beta_0)b^{i_2}(X_{t_{k-1}},\alpha)b^{i_3}(X_{t_{k-1}},\alpha) \\
& +\Xi_{k-1}^{i_2 i_3}(\beta_0)B_{k-1}^{i_1}(\alpha_0,\alpha)b^{i_4}(X_{t_{k-1}},\alpha_0) \\
& -\Xi_{k-1}^{i_2 i_3}(\beta_0)B_{k-1}^{i_1}(\alpha_0,\alpha)b^{i_4}(X_{t_{k-1}},\alpha) \\
& +\Xi_{k-1}^{i_2 i_4}(\beta_0)B_{k-1}^{i_3}(\alpha_0,\alpha)b^{i_1}(X_{t_{k-1}},\alpha_0) \\
& -\Xi_{k-1}^{i_2 i_4}(\beta_0)b^{i_1}(X_{t_{k-1}},\alpha)b^{i_3}(X_{t_{k-1}},\alpha_0) \\
& +\Xi_{k-1}^{i_2 i_4}(\beta_0)b^{i_1}(X_{t_{k-1}},\alpha)b^{i_3}(X_{t_{k-1}},\alpha) \\
& +\Xi_{k-1}^{i_3 i_4}(\beta_0)B_{k-1}^{i_2}(\alpha_0,\alpha)b^{i_1}(X_{t_{k-1}},\alpha_0) \\
& -\Xi_{k-1}^{i_3 i_4}(\beta_0)b^{i_1}(X_{t_{k-1}},\alpha)b^{i_2}(X_{t_{k-1}},\alpha_0) \\
& \left.+\Xi_{k-1}^{i_3 i_4}(\beta_0)b^{i_1}(X_{t_{k-1}},\alpha)b^{i_2}(X_{t_{k-1}},\alpha)\right\} \\
& +\frac{1}{n^4}\left\{b^{i_1}b^{i_2}b^{i_3}b^{i_4}(X_{t_{k-1}},\alpha_0)\right. \\
& -b^{i_1}b^{i_2}b^{i_4}(X_{t_{k-1}},\alpha_0)b^{i_3}(X_{t_{k-1}},\alpha) \\
& -b^{i_1}b^{i_3}b^{i_4}(X_{t_{k-1}},\alpha_0)b^{i_2}(X_{t_{k-1}},\alpha) \\
& -b^{i_2}b^{i_3}b^{i_4}(X_{t_{k-1}},\alpha_0)b^{i_1}(X_{t_{k-1}},\alpha) \\
& +b^{i_1}b^{i_4}(X_{t_{k-1}},\alpha_0)b^{i_2}b^{i_3}(X_{t_{k-1}},\alpha) \\
& +b^{i_2}b^{i_4}(X_{t_{k-1}},\alpha_0)b^{i_1}b^{i_3}(X_{t_{k-1}},\alpha) \\
& +b^{i_3}b^{i_4}(X_{t_{k-1}},\alpha_0)b^{i_1}b^{i_2}(X_{t_{k-1}},\alpha) \\
& -b^{i_4}(X_{t_{k-1}},\alpha_0)b^{i_1}b^{i_2}b^{i_3}(X_{t_{k-1}},\alpha) \\
& \left.-B_{k-1}^{i_1}B_{k-1}^{i_2}B_{k-1}^{i_3}(\alpha_0,\alpha)b^{i_4}(X_{t_{k-1}},\alpha)\right\} \\
& +R\left(\frac{\varepsilon^4}{n^3},X_{t_{k-1}}\right)+R\left(\frac{\varepsilon^2}{n^4},X_{t_{k-1}}\right)+R\left(\frac{1}{n^5},X_{t_{k-1}}\right).
\end{aligned}
$$

∎

Proof of Lemma 2. (i) In view of Theorem B in Genon-Catalot [7] (cf. Theorem 1.3 in Azencott [1]), $\sup_{t\le 1}|f(X_t,\theta)-f(X_t^0,\theta)| = o_p(1)$ for all θ. Thus, under P_{θ_0}, as $\varepsilon\to 0$ and $n\to\infty$, one has

$$
\frac{1}{n}\sum_{k=1}^n f(X_{t_{k-1}},\theta) \to \int_0^1 f(X_s^0,\theta)ds.
$$

Moreover, it follows from the assumption on f and [A2] that

$$\sup_{\varepsilon,n} E_{\theta_0}\left[\sup_{\theta}\left|\frac{\partial}{\partial\theta}\left(\frac{1}{n}\sum_{k=1}^{n} f(X_{t_{k-1}},\theta)\right)\right|\right] < \infty.$$

Therefore, the family of distributions of $\displaystyle\frac{1}{n}\sum_{k=1}^{n} f(X_{t_{k-1}},\cdot)$ on the Banach space $C(\bar{\Theta})$ with sup-norm is tight. ∎

(ii) Let $\xi_k^i(\theta) = f(X_{t_{k-1}},\theta)P_k^i(\alpha_0)$. From Lemmas 1 and 2–(i), under P_{θ_0}, as $\varepsilon \to 0$ and $n \to \infty$, one has

$$\sum_{k=1}^{n} E[\xi_k^i(\theta)|\mathcal{G}_{k-1}^n] = \sum_{k=1}^{n} R\left(\frac{1}{n^2}, X_{t_{k-1}}\right) \to 0,$$

$$\sum_{k=1}^{n} E[(\xi_k^i(\theta))^2|\mathcal{G}_{k-1}^n] = \sum_{k=1}^{n}\left\{R\left(\frac{\varepsilon^2}{n}, X_{t_{k-1}}\right) + R\left(\frac{1}{n^3}, X_{t_{k-1}}\right)\right\} \to 0.$$

It follows from Lemma 9 in Genon-Catalot and Jacod [8] that $\sum_{k=1}^{n}\xi_k^i(\theta) \to 0$ in P_{θ_0}-probability as $\varepsilon \to 0$ and $n \to \infty$. From now on, let C be a generic positive constant independent of ε, n and other variables in some cases (see Yoshida [35] or Kessler [16]). Moreover we may write C_m if it depends on an integer m. In order to prove the tightness of $\sum_{k=1}^{n}\xi_k^i(\cdot)$, it is enough to show the following inequalities (cf. Theorem 20 in Appendix I of Ibragimov and Has'minskii [13] or Lemma 3.1 of Yoshida [32]):

$$E_{\theta_0}\left[\left(\sum_{k=1}^{n}\xi_k^i(\theta)\right)^{2l}\right] \le C, \tag{3}$$

$$E_{\theta_0}\left[\left(\sum_{k=1}^{n}\xi_k^i(\theta_2) - \sum_{k=1}^{n}\xi_k^i(\theta_1)\right)^{2l}\right] \le C|\theta_2-\theta_1|^{2l}, \tag{4}$$

for $\theta,\theta_1,\theta_2 \in \bar{\Theta}$, where $l > (p+q)/2$. We define $A_{k,1}^i(\theta)$, $A_{k,2}^i(\theta)$ and $A_{k,3}^i(\theta)$ by

$$\begin{aligned}
\xi_k^i(\theta) &= f(X_{t_{k-1}},\theta)\int_{t_{k-1}}^{t_k} b^i(X_s,\alpha_0)ds \\
&\quad +\varepsilon f(X_{t_{k-1}},\theta)\int_{t_{k-1}}^{t_k}\sum_{j=1}^{r}\sigma^{ij}(X_s,\beta_0)dw_s^j \\
&\quad -\frac{1}{n}f(X_{t_{k-1}},\theta)b^i(X_{t_{k-1}},\alpha_0) \\
&=: A_{k,1}^i(\theta) + A_{k,2}^i(\theta) - A_{k,3}^i(\theta).
\end{aligned}$$

$$\begin{aligned}
E_{\theta_0}\left[\left|\sum_{k=1}^{n}A_{k,1}^i(\theta)\right|^{2l}\right] &\le n^{2l-1}\sum_{k=1}^{n}E_{\theta_0}\left[\left(\int_{t_{k-1}}^{t_k}|f(X_{t_{k-1}},\theta)b^i(X_s,\alpha_0)|ds\right)^{2l}\right] \\
&\le n^{2l-1}\sum_{k=1}^{n}E_{\theta_0}\left[\left\{\left(\int_{t_{k-1}}^{t_k}|f(X_{t_{k-1}},\theta)b^i(X_s,\alpha_0)|^{2l}ds\right)^{\frac{1}{2l}}\left(\int_{t_{k-1}}^{t_k}ds\right)^{1-\frac{1}{2l}}\right\}^{2l}\right]
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{k=1}^{n} E_{\theta_0}\left[\int_{t_{k-1}}^{t_k} |f(X_{t_{k-1}},\theta)b^i(X_s,\alpha_0)|^{2l} ds\right] \\
&= \sum_{k=1}^{n} \int_{t_{k-1}}^{t_k} E_{\theta_0}\left[|f(X_{t_{k-1}},\theta)|^{2l} E_{\theta_0}[|b^i(X_s,\alpha_0)|^{2l}|\mathcal{G}_{k-1}^n]\right] ds \\
&\le n\cdot\frac{1}{n}\cdot C,
\end{aligned}
$$

where the last estimate is based on Lemma 6 of Kessler [16].

$$
\begin{aligned}
E_{\theta_0}\left[\left|\sum_{k=1}^{n} A_{k,2}^i(\theta)\right|^{2l}\right] &\le C_{2l}\varepsilon^{2l} E_{\theta_0}\left[\left(\sum_{k=1}^{n}\int_{t_{k-1}}^{t_k} f(X_{t_{k-1}},\theta)^2[\sigma\sigma^*]^{ii}(X_s,\beta_0)ds\right)^{l}\right] \\
&\le C_{2l}\varepsilon^{2l} n^{l-1}\sum_{k=1}^{n} E_{\theta_0}\left[\left(\int_{t_{k-1}}^{t_k} f(X_{t_{k-1}},\theta)^2[\sigma\sigma^*]^{ii}(X_s,\beta_0)ds\right)^{l}\right] \\
&\le C_{2l}\varepsilon^{2l} n^{l-1}\left(\frac{1}{n}\right)^{l-1} \\
&\quad \times\sum_{k=1}^{n}\int_{t_{k-1}}^{t_k} E_{\theta_0}[f(X_{t_{k-1}},\theta)^{2l}E_{\theta_0}[[\sigma\sigma^*]^{ii}(X_s,\beta_0))^{l}|\mathcal{G}_{k-1}^n]ds \\
&\le C_{2l}\varepsilon^{2l}C,
\end{aligned}
$$

where the first estimate is based on the Burkholder-Davis-Gundy inequality.

$$
\begin{aligned}
E_{\theta_0}\left[\left|\sum_{k=1}^{n} A_{k,3}^i(\theta)\right|^{2l}\right] &\le \frac{1}{n}\sum_{k=1}^{n} E_{\theta_0}[|f(X_{t_{k-1}},\theta)b^i(X_{t_{k-1}},\alpha_0)|^{2l}] \\
&\le C.
\end{aligned}
$$

Therefore, we deduce the inequality (3). For the inequality (4), we first obtain

$$
\begin{aligned}
& E_{\theta_0}\left[\left(\sum_{k=1}^{n}\{A_{k,1}^i(\theta_2)-A_{k,1}^i(\theta_1)\}\right)^{2l}\right] \\
\le\ & n^{2l-1}\sum_{k=1}^{n} E_{\theta_0}\left[\left(\int_{t_{k-1}}^{t_k}\{f(X_{t_{k-1}},\theta_2)-f(X_{t_{k-1}},\theta_1)\}b^i(X_s,\alpha_0)ds\right)^{2l}\right] \\
\le\ & n^{2l-1}\sum_{k=1}^{n} E_{\theta_0}\left[\left(\int_{t_{k-1}}^{t_k}\left\{\sum_{j=1}^{p+q}\int_0^1\frac{\partial}{\partial\theta_j}f(X_{t_{k-1}},\theta_1+u(\theta_2-\theta_1))du(\theta_2-\theta_1)^j\right\}b^i(X_s,\alpha_0)ds\right)^{2l}\right] \\
\le\ & n^{2l-1}\left(\frac{1}{n}\right)^{2l-1}\sum_{k=1}^{n} E_{\theta_0}\left[\int_{t_{k-1}}^{t_k}\left\{\sum_{j=1}^{p+q}\int_0^1\frac{\partial}{\partial\theta_j}f(X_{t_{k-1}},\theta_1+u(\theta_2-\theta_1))du(\theta_2-\theta_1)^j\right\}^{2l}\right. \\
& \left.\times(b^i(X_s,\alpha_0))^{2l}ds\right] \\
\le\ & \sum_{k=1}^{n}\int_{t_{k-1}}^{t_k} E_{\theta_0}\left[\left(\sum_{j=1}^{p+q}\int_0^1\frac{\partial}{\partial\theta_j}f(X_{t_{k-1}},\theta_1+u(\theta_2-\theta_1))du(\theta_2-\theta_1)^j\right)^{2l}\right.
\end{aligned}
$$

$$
\begin{aligned}
&\quad \times E_{\theta_0}[(b^i(X_s,\alpha_0))^{2l}|\mathcal{G}_{k-1}^n]\Big]\,ds \\
&\leq n\frac{1}{n}C\left(\sum_{j=1}^{p+q}(\theta_2-\theta_1)^j\right)^{2l} \\
&\leq C\left((p+q)\sum_{j=1}^{p+q}[(\theta_2-\theta_1)^j]^2\right)^l \\
&\leq C|\theta_2-\theta_1|^{2l}.
\end{aligned}
$$

Next one has

$$
\begin{aligned}
& E_{\theta_0}\left[\left(\sum_{k=1}^n\{A_{k,2}^i(\theta_2)-A_{k,2}^i(\theta_1)\}\right)^{2l}\right] \\
\leq\ & E_{\theta_0}\left[\left(\sum_{k=1}^n\int_{t_{k-1}}^{t_k}\{f(X_{t_{k-1}},\theta_2)-f(X_{t_{k-1}},\theta_1)\}^2[\sigma\sigma^*]^{ii}(X_s,\beta_0)ds\right)^l\right] \\
\leq\ & n^{l-1}\sum_{k=1}^n E_{\theta_0}\left[\left(\int_{t_{k-1}}^{t_k}\{f(X_{t_{k-1}},\theta_2)-f(X_{t_{k-1}},\theta_1)\}^2[\sigma\sigma^*]^{ii}(X_s,\beta_0)ds\right)^l\right] \\
\leq\ & n^{l-1}\left(\frac{1}{n}\right)^{l-1}\sum_{k=1}^n E_{\theta_0}\left[\int_{t_{k-1}}^{t_k}\{f(X_{t_{k-1}},\theta_2)-f(X_{t_{k-1}},\theta_1)\}^{2l}([\sigma\sigma^*]^{ii}(X_s,\beta_0))^l ds\right] \\
\leq\ & \sum_{k=1}^n\int_{t_{k-1}}^{t_k}E_{\theta_0}\Big[\{f(X_{t_{k-1}},\theta_2)-f(X_{t_{k-1}},\theta_1)\}^{2l}E_{\theta_0}[([\sigma\sigma^*]^{ii}(X_s,\beta_0))^l|\mathcal{G}_{k-1}^n]\Big]\,ds \\
\leq\ & \sum_{k=1}^n\int_{t_{k-1}}^{t_k}E_{\theta_0}\left[\left(\sum_{j=1}^{p+q}\int_0^1\frac{\partial}{\partial\theta_j}f(X_{t_{k-1}},\theta_1+u(\theta_2-\theta_1))du(\theta_2-\theta_1)^j\right)^{2l}\right. \\
& \quad \times E_{\theta_0}[([\sigma\sigma^*]^{ii}(X_s,\beta_0))^l|\mathcal{G}_{k-1}^n]\Big]\,ds \\
\leq\ & n\frac{1}{n}C\left(\sum_{j=1}^{p+q}(\theta_2-\theta_1)^j\right)^{2l}.
\end{aligned}
$$

Finally, it follows that

$$
\begin{aligned}
& E_{\theta_0}\left[\left(\sum_{k=1}^n\{A_{k,3}^i(\theta_2)-A_{k,3}^i(\theta_1)\}\right)^{2l}\right] \\
\leq\ & \frac{1}{n}\sum_{k=1}^n E_{\theta_0}[(\{f(X_{t_{k-1}},\theta_2)-f(X_{t_{k-1}},\theta_1)\}b^i(X_{t_{k-1}},\alpha_0))^{2l}] \\
\leq\ & \frac{1}{n}\sum_{k=1}^n E_{\theta_0}\left[\left(\sum_{j=1}^{p+q}\int_0^1\frac{\partial}{\partial\theta_j}f(X_{t_{k-1}},\theta_1+u(\theta_2-\theta_1))du(\theta_2-\theta_1)^j b^i(X_{t_{k-1}},\alpha_0)\right)^{2l}\right] \\
\leq\ & n\frac{1}{n}C\left(\sum_{j=1}^{p+q}(\theta_2-\theta_1)^j\right)^{2l},
\end{aligned}
$$

which completes the proof. ∎

Proof of Lemma 3. (i) From Lemma 1, one has

$$\begin{aligned}
\sum_{k=1}^{n} E_{\theta_0}[\varepsilon^{-2} f(X_{t_{k-1}},\theta) P_k^i P_k^j(\alpha_0)|\mathcal{G}_{k-1}^n] &= \frac{1}{n}\sum_{k=1}^{n} f(X_{t_{k-1}},\theta)[\sigma\sigma^*]^{ij}(X_{t_{k-1}},\beta_0) \\
&\quad + \sum_{k=1}^{n}\left\{R\left(\frac{1}{n^2},X_{t_{k-1}}\right) + R\left(\frac{\varepsilon^{-2}}{n^3},X_{t_{k-1}}\right)\right\},
\end{aligned}$$

$$\begin{aligned}
\sum_{k=1}^{n} E_{\theta_0}[\varepsilon^{-4} f(X_{t_{k-1}},\theta)^2 (P_k^i P_k^j(\alpha_0))^2|\mathcal{G}_{k-1}^n] &= \sum_{k=1}^{n}\left\{R\left(\frac{1}{n^2},X_{t_{k-1}}\right) + R\left(\frac{\varepsilon^{-2}}{n^4},X_{t_{k-1}}\right)\right. \\
&\quad \left. + R\left(\frac{\varepsilon^{-4}}{n^5},X_{t_{k-1}}\right)\right\}.
\end{aligned}$$

From Lemma 2–(i) and [B1], under P_{θ_0}, as $\varepsilon \to 0$ and $n \to \infty$, one has

$$\begin{aligned}
\sum_{k=1}^{n} E_{\theta_0}[\varepsilon^{-2} f(X_{t_{k-1}},\theta) P_k^i P_k^j(\alpha_0)|\mathcal{G}_{k-1}^n] &\to \int_0^1 f(X_s^0,\theta)[\sigma\sigma^*]^{ij}(X_s^0,\beta_0)ds, \\
\sum_{k=1}^{n} E_{\theta_0}[\varepsilon^{-4} f(X_{t_{k-1}},\theta)^2 (P_k^i P_k^j(\alpha_0))^2|\mathcal{G}_{k-1}^n] &\to 0.
\end{aligned}$$

Thus, it follows from Lemma 9 in Genon-Catalot and Jacod [8] that under P_{θ_0}, as $\varepsilon \to 0$ and $n \to \infty$, one has

$$\varepsilon^{-2}\sum_{k=1}^{n} f(X_{t_{k-1}},\theta) P_k^i P_k^j(\alpha_0) \to \int_0^1 f(X_s^0,\theta)[\sigma\sigma^*]^{ij}(X_s^0,\beta_0)ds.$$

For tightness of the family of distributions of $\varepsilon^{-2}\sum_{k=1}^{n} f(X_{t_{k-1}},\cdot)P_k^i P_k^j(\alpha_0)$, we use that

$$\begin{aligned}
& \sup_{\varepsilon,n} E_{\theta_0}\left[\sup_{\theta}\left|\varepsilon^{-2}\sum_{k=1}^{n}\frac{\partial}{\partial\theta} f(X_{t_{k-1}},\theta) P_k^i P_k^j(\alpha_0)\right|\right] \\
\leq\ & \sup_{\varepsilon,n} E_{\theta_0}\left[\frac{\varepsilon^{-2}}{2}\sup_{\theta}\sum_{k=1}^{n}\left|\frac{\partial}{\partial\theta} f(X_{t_{k-1}},\theta)\right| E_{\theta_0}\left[(P_k^i(\alpha_0))^2 + (P_k^j(\alpha_0))^2\middle|\mathcal{G}_{k-1}^n\right]\right] \\
\leq\ & \frac{1}{2}\sup_{\varepsilon,n} E_{\theta_0}\left[\sum_{k=1}^{n}\sup_{\theta}\left|\frac{\partial}{\partial\theta} f(X_{t_{k-1}},\theta)\right|\left\{\frac{1}{n}\left([\sigma\sigma^*]^{ii}(X_{t_{k-1}},\beta_0) + [\sigma\sigma^*]^{ij}(X_{t_{k-1}},\beta_0)\right)\right.\right. \\
& \left.\left. + R\left(\frac{1}{n^2},X_{t_{k-1}}\right) + R\left(\frac{\varepsilon^{-2}}{n^3},X_{t_{k-1}}\right)\right\}\right] \\
\leq\ & C\sup_{\varepsilon,n}\frac{1}{(\varepsilon n)^2} \\
<\ & \infty,
\end{aligned}$$

where the last estimate is based on $\lim_{\varepsilon\to 0, n\to\infty}(\varepsilon n)^{-1} = 0$. This completes the proof. ∎

(ii) Noting that

$$\begin{aligned}
P_k^i P_k^j(\alpha) &= P_k^i P_k^j(\alpha_0) + \frac{1}{n}P_k^i(\alpha_0)B_{k-1}^j(\alpha_0,\alpha) + \frac{1}{n}P_k^j(\alpha_0)B_{k-1}^i(\alpha_0,\alpha) \\
&\quad + \frac{1}{n^2}B_{k-1}^i B_{k-1}^j(\alpha_0,\alpha),
\end{aligned}$$

it follows from Lemmas 2 and 3-(i) and [B2] that under P_{θ_0}, as $\varepsilon \to 0$ and $n \to \infty$,

$$
\begin{aligned}
& \varepsilon^{-2} \sum_{k=1}^{n} f(X_{t_{k-1}}, \theta) P_k^i P_k^j(\alpha) \\
= \; & \varepsilon^{-2} \sum_{k=1}^{n} f(X_{t_{k-1}}, \theta) P_k^i P_k^j(\alpha_0) + \frac{1}{n^2}\varepsilon^{-2} \sum_{k=1}^{n} f(X_{t_{k-1}}, \theta) B_{k-1}^i B_{k-1}^j(\alpha_0, \alpha) \\
& + \frac{1}{n}\varepsilon^{-2} \sum_{k=1}^{n} f(X_{t_{k-1}}, \theta) \left\{ P_k^i(\alpha_0) B_{k-1}^j(\alpha_0, \alpha) + P_k^j(\alpha_0) B_{k-1}^i(\alpha_0, \alpha) \right\} \\
\to \; & \int_0^1 f(X_s^0, \theta)[\sigma\sigma^*]^{ij}(X_s^0, \beta_0) ds + M^2 \int_0^1 f(X_s^0, \theta) B^i B^j (X_s^0, \alpha_0, \alpha) ds
\end{aligned}
$$

uniformly in $\theta \in \bar{\Theta}$, where $M = \lim_{\varepsilon \to 0, n \to \infty} (\varepsilon\sqrt{n})^{-1}$. This completes the proof. ∎

Proof of Lemma 4. (i) A simple computation yields

$$
\begin{aligned}
\varepsilon^2\{U_{\varepsilon,n}(\alpha, \beta) - U_{\varepsilon,n}(\alpha_0, \beta)\} &= n \sum_{k=1}^{n} (P_k(\alpha) - P_k(\alpha_0))^* \Xi_{k-1}^{-1}(\beta)(P_k(\alpha) + P_k(\alpha_0)) \\
&= \sum_{k=1}^{n} (b(X_{t_{k-1}}, \alpha) - b(X_{t_{k-1}}, \alpha_0))^* \Xi_{k-1}^{-1}(\beta) \\
&\quad \times \left(2\left\{ X_{t_k} - X_{t_{k-1}} - \frac{1}{n} b(X_{t_{k-1}}, \alpha_0) \right\} \right. \\
&\quad \left. + \frac{1}{n}\{b(X_{t_{k-1}}, \alpha_0) - b(X_{t_{k-1}}, \alpha)\} \right).
\end{aligned}
$$

From Lemma 2, under P_{θ_0}, as $\varepsilon \to 0$ and $n \to \infty$, one has

$$
\varepsilon^2\{U_{\varepsilon,n}(\alpha, \beta) - U_{\varepsilon,n}(\alpha_0, \beta)\} \to U_1(\alpha, \alpha_0, \beta)
$$

uniformly in $\theta \in \bar{\Theta}$. This completes the proof. ∎

(ii)It follows from Lemmas 2–(i) and 3-(ii) that under P_{θ_0}, as $\varepsilon \to 0$ and $n \to \infty$,

$$
\frac{1}{n} U_{\varepsilon,n}(\alpha, \beta) \to U_2(\alpha, \beta, \beta_0)
$$

uniformly in $\theta \in \bar{\Theta}$. This completes the proof. ∎

Proof of Lemma 5. We first consider the uniform convergence of $C_{\varepsilon,n}(\theta)$. An easy computation implies

$$
\begin{aligned}
\varepsilon^2 \frac{\partial^2}{\partial\alpha_i \partial\alpha_j} U_{\varepsilon,n}(\theta) &= -2 \sum_{k=1}^{n} \left(\frac{\partial^2}{\partial\alpha_i \partial\alpha_j} b(X_{t_{k-1}}, \alpha) \right)^* \Xi_{k-1}^{-1}(\beta) \left(P_k(\alpha_0) + \frac{1}{n} B(X_{t_{k-1}}, \alpha_0, \alpha) \right) \\
&\quad + 2\frac{1}{n} \sum_{k=1}^{n} \left(\frac{\partial}{\partial\alpha_i} b(X_{t_{k-1}}, \alpha) \right)^* \Xi_{k-1}^{-1}(\beta) \left(\frac{\partial}{\partial\alpha_j} b(X_{t_{k-1}}, \alpha) \right), \\
\varepsilon \frac{1}{\sqrt{n}} \frac{\partial^2}{\partial\alpha_i \partial\beta_j} U_{\varepsilon,n}(\theta) &= \varepsilon^{-1} \frac{1}{\sqrt{n}} (-2) \sum_{k=1}^{n} \left(\frac{\partial}{\partial\alpha_i} b(X_{t_{k-1}}, \alpha) \right)^* \frac{\partial}{\partial\beta_j} (\Xi_{k-1}^{-1}(\beta))
\end{aligned}
$$

$$
\begin{aligned}
&\qquad \times \left(P_k(\alpha_0) + \frac{1}{n} B(X_{t_{k-1}}, \alpha_0, \alpha)\right), \\
\frac{1}{n}\frac{\partial^2}{\partial\beta_i\partial\beta_j} U_{\varepsilon,n}(\theta) &= \frac{1}{n}\sum_{k=1}^n \left(\frac{\partial^2}{\partial\beta_i\partial\beta_j}\log\det\Xi_{k-1}(\beta)\right) \\
&\quad + \varepsilon^{-2}\sum_{k=1}^n P_k(\alpha)^* \left(\frac{\partial^2}{\partial\beta_i\partial\beta_j}\Xi_{k-1}^{-1}(\beta)\right) P_k(\alpha).
\end{aligned}
$$

From Lemma 2, under P_{θ_0}, as $\varepsilon \to 0$ and $n \to \infty$, one has

$$
\begin{aligned}
\varepsilon^2 \frac{\partial^2}{\partial\alpha_i\partial\alpha_j} U_{\varepsilon,n}(\theta) &\to -2\int_0^1 \left(\frac{\partial^2}{\partial\alpha_i\partial\alpha_j} b(X_s^0, \alpha)\right)^* [\sigma\sigma^*]^{-1}(X_s^0, \beta) B(X_s^0, \alpha_0, \alpha) ds \qquad (5) \\
&\quad + 2\int_0^1 \left(\frac{\partial}{\partial\alpha_i} b(X_s^0, \alpha)\right)^* [\sigma\sigma^*]^{-1}(X_s^0, \beta) \left(\frac{\partial}{\partial\alpha_j} b(X_s^0, \alpha)\right) ds, \\
\varepsilon \frac{1}{\sqrt{n}}\frac{\partial^2}{\partial\alpha_i\partial\beta_j} U_{\varepsilon,n}(\theta) &\to -2M\int_0^1 \left(\frac{\partial}{\partial\alpha_i} b(X_s^0, \alpha)\right)^* \left(\frac{\partial}{\partial\beta_j}[\sigma\sigma^*]^{-1}(X_s^0, \beta)\right) B(X_s^0, \alpha_0, \alpha) ds, \qquad (6) \\
\frac{1}{n}\frac{\partial^2}{\partial\beta_i\partial\beta_j} U_{\varepsilon,n}(\theta) &\to \int_0^1 \frac{\partial^2}{\partial\beta_i\partial\beta_j}\log\det[\sigma\sigma^*](X_s^0, \beta) ds \qquad (7) \\
&\quad + \int_0^1 \mathrm{tr}\left[[\sigma\sigma^*](X_s^0, \beta_0)\left(\frac{\partial^2}{\partial\beta_i\partial\beta_j}[\sigma\sigma^*]^{-1}(X_s^0, \beta)\right)\right] ds \\
&\quad + M^2\int_0^1 \mathrm{tr}\left[BB^*(X_s^0, \alpha_0, \alpha)\left(\frac{\partial^2}{\partial\beta_i\partial\beta_j}[\sigma\sigma^*]^{-1}(X_s^0, \beta)\right)\right] ds
\end{aligned}
$$

uniformly in $\theta \in \bar{\Theta}$.

Now (i) follows from (5), (6) and (7). Next, by the assumptions [A3] and [A4], the limits of (5), (6) and (7) are continuous with respect to θ, which completes the proof of (ii). ∎

Proof of Lemma 6. We set

$$
\begin{aligned}
-\varepsilon\frac{\partial}{\partial\alpha_i} U_{\varepsilon,n}(\theta_0) &= 2\varepsilon^{-1}\sum_{k=1}^n \left(\frac{\partial}{\partial\alpha_i} b(X_{t_{k-1}}, \alpha_0)\right)^* \Xi_{k-1}^{-1}(\beta_0) P_k(\alpha_0) \\
&= \sum_{k=1}^n 2\varepsilon^{-1}\sum_{l_1=1}^d \left[\left(\frac{\partial}{\partial\alpha_i} b(X_{t_{k-1}}, \alpha_0)\right)^* \Xi_{k-1}^{-1}(\beta_0)\right]^{l_1} P_k^{l_1}(\alpha_0) \\
&=: \sum_{k=1}^n \xi_k^i(\theta_0),
\end{aligned}
$$

$$
\begin{aligned}
-\frac{1}{\sqrt{n}}\frac{\partial}{\partial\beta_j} U_{\varepsilon,n}(\theta_0) &= -\frac{1}{\sqrt{n}}\sum_{k=1}^n \frac{\partial}{\partial\beta_j}\log\det\Xi_{k-1}(\beta_0) \\
&\quad -\varepsilon^{-2}\sqrt{n}\sum_{k=1}^n P_k(\alpha_0)^*\frac{\partial}{\partial\beta_j}\Xi_{k-1}^{-1}(\beta_0) P_k(\alpha_0) \\
&= -\sum_{k=1}^n \frac{1}{\sqrt{n}}\frac{\partial}{\partial\beta_j}\log\det\Xi_{k-1}(\beta_0)
\end{aligned}
$$

$$
\begin{aligned}
&- \sum_{k=1}^{n} \varepsilon^{-2} \sqrt{n} \sum_{l_1,l_2=1}^{d} \left(\frac{\partial}{\partial \beta_j} \Xi_{k-1}^{-1}(\beta_0) \right)^{l_1 l_2} P_k^{l_1} P_k^{l_2}(\alpha_0) \\
=: \ & \sum_{k=1}^{n} \eta_k^j(\theta_0).
\end{aligned}
$$

In view of Theorems 3.2 and 3.4 of Hall and Heyde [11], it is sufficient to show that under P_{θ_0}, as $\varepsilon \to 0$ and $n \to \infty$, one has

$$\sum_{k=1}^{n} E_{\theta_0}[\xi_k^i(\theta_0)|\mathcal{G}_{k-1}^n] \to 0 \tag{8}$$

$$\sum_{k=1}^{n} E_{\theta_0}[\eta_k^j(\theta_0)|\mathcal{G}_{k-1}^n] \to 0 \tag{9}$$

$$\sum_{k=1}^{n} E_{\theta_0}[\xi_k^{i_1}\xi_k^{i_2}(\theta_0)|\mathcal{G}_{k-1}^n] \to 4 I_b^{i_1 i_2}(\theta_0) \tag{10}$$

$$\sum_{k=1}^{n} E_{\theta_0}[\eta_k^{j_1}\eta_k^{j_2}(\theta_0)|\mathcal{G}_{k-1}^n] \to 4 I_\sigma^{j_1 j_2}(\theta_0) \tag{11}$$

$$\sum_{k=1}^{n} E_{\theta_0}[\xi_k^{i}\eta_k^{j}(\theta_0)|\mathcal{G}_{k-1}^n] \to 0 \tag{12}$$

$$\sum_{k=1}^{n} E_{\theta_0}[(\xi_k^{i}(\theta_0))^4|\mathcal{G}_{k-1}^n] \to 0 \tag{13}$$

$$\sum_{k=1}^{n} E_{\theta_0}[(\eta_k^{j}(\theta_0))^4|\mathcal{G}_{k-1}^n] \to 0. \tag{14}$$

Proof of (8). It follows from Lemma 1–(i) that we obtain

$$\sum_{k=1}^{n} E_{\theta_0}[\xi_k^i(\theta_0)|\mathcal{G}_{k-1}^n] = \sum_{k=1}^{n} R\left(\frac{\varepsilon^{-1}}{n^2}, X_{t_{k-1}} \right) \to 0$$

in P_{θ_0}-probability as $\varepsilon \to 0$ and $n \to \infty$. ∎

Proof of (9). By Lemma 1–(ii), one has

$$
\begin{aligned}
\sum_{k=1}^{n} E_{\theta_0}[\eta_k^j(\theta_0)|\mathcal{G}_{k-1}^n] \ &= \ \sum_{k=1}^{n} \left\{ -\frac{1}{\sqrt{n}} \frac{\partial}{\partial \beta_j} \log\det \Xi_{k-1}(\beta_0) - \frac{1}{\sqrt{n}} \mathrm{tr} \left[\Xi_{k-1}(\beta_0) \left(\frac{\partial}{\partial \beta_j} \Xi_{k-1}^{-1}(\beta_0) \right) \right] \right. \\
&\qquad \left. + R\left(\frac{1}{n\sqrt{n}}, X_{t_{k-1}} \right) + R\left(\frac{\varepsilon^{-2}}{n^2\sqrt{n}}, X_{t_{k-1}} \right) \right\} \\
&= \ \sum_{k=1}^{n} \left\{ R\left(\frac{1}{n\sqrt{n}}, X_{t_{k-1}} \right) + R\left(\frac{\varepsilon^{-2}}{n^2\sqrt{n}}, X_{t_{k-1}} \right) \right\} \\
&\to \ 0
\end{aligned}
$$

in P_{θ_0}-probability as $\varepsilon \to 0$ and $n \to \infty$. ∎

Proof of (10). From Lemma 1–(ii), we have

$$\sum_{k=1}^{n} E_{\theta_0}[\xi_k^{i_1}\xi_k^{i_2}(\theta_0)|\mathcal{G}_{k-1}^n]$$

$$
\begin{aligned}
&= 4\varepsilon^{-2}\sum_{k=1}^{n}\sum_{l_1,l_2=1}^{d}\left[\left(\frac{\partial}{\partial\alpha_{i_1}}b(X_{t_{k-1}},\alpha_0)\right)^{*}\Xi_{k-1}^{-1}(\beta_0)\right]^{l_1}\left[\left(\frac{\partial}{\partial\alpha_{i_2}}b(X_{t_{k-1}},\alpha_0))\right)^{*}\Xi_{k-1}^{-1}(\beta_0)\right]^{l_2} \\
&\quad\times E_{\theta_0}[P_k^{l_1}P_k^{l_2}(\alpha_0)|\mathcal{G}_{k-1}^n] \\
&= 4\frac{1}{n}\sum_{k=1}^{n}\frac{\partial}{\partial\alpha_{i_1}}b(X_{t_{k-1}},\alpha_0)^{*}\Xi_{k-1}^{-1}(\beta_0)\frac{\partial}{\partial\alpha_{i_2}}b(X_{t_{k-1}},\alpha_0) \\
&\quad+\sum_{k=1}^{n}\left\{R\left(\frac{1}{n^2},X_{t_{k-1}}\right)+R\left(\frac{\varepsilon^{-2}}{n^3},X_{t_{k-1}}\right)\right\} \\
&\to 4I_b^{i_1i_2}(\theta_0)
\end{aligned}
$$

in P_{θ_0}-probability as $\varepsilon\to 0$ and $n\to\infty$. ∎

Proof of (11). Using Lemma 1–(ii) and (iv), one has

$$
\begin{aligned}
&\sum_{k=1}^{n}E_{\theta_0}[\eta_k^{j_1}\eta_k^{j_2}(\theta_0)|\mathcal{G}_{k-1}^n] \\
&= \sum_{k=1}^{n}\left\{\frac{1}{n}\left(\frac{\partial}{\partial\beta_{j_1}}\log\det\Xi_{k-1}(\beta_0)\right)\left(\frac{\partial}{\partial\beta_{j_2}}\log\det\Xi_{k-1}(\beta_0)\right)\right. \\
&\quad+\varepsilon^{-2}\frac{\partial}{\partial\beta_{j_1}}\log\det\Xi_{k-1}(\beta_0)\sum_{l_3,l_4=1}^{d}\left(\frac{\partial}{\partial\beta_{j_2}}\Xi_{k-1}^{-1}(\beta_0)\right)^{l_3l_4}E_{\theta_0}[P_k^{l_3}P_k^{l_4}(\alpha_0)|\mathcal{G}_{k-1}^n] \\
&\quad+\varepsilon^{-2}\frac{\partial}{\partial\beta_{j_2}}\log\det\Xi_{k-1}(\beta_0)\sum_{l_1,l_2=1}^{d}\left(\frac{\partial}{\partial\beta_{j_1}}\Xi_{k-1}^{-1}(\beta_0)\right)^{l_1l_2}E_{\theta_0}[P_k^{l_1}P_k^{l_2}(\alpha_0)|\mathcal{G}_{k-1}^n] \\
&\quad\left.+\varepsilon^{-4}n\sum_{l_1,l_2=1}^{d}\sum_{l_3,l_4=1}^{d}\left(\frac{\partial}{\partial\beta_{j_1}}\Xi_{k-1}^{-1}(\beta_0)\right)^{l_1l_2}\left(\frac{\partial}{\partial\beta_{j_2}}\Xi_{k-1}^{-1}(\beta_0)\right)^{l_3l_4}E_{\theta_0}[P_k^{l_1}P_k^{l_2}P_k^{l_3}P_k^{l_4}(\alpha_0)|\mathcal{G}_{k-1}^n]\right\} \\
&= \frac{1}{n}\sum_{k=1}^{n}\left\{\mathrm{tr}\left[\frac{\partial}{\partial\beta_{j_1}}\Xi_{k-1}(\beta_0)\Xi_{k-1}^{-1}(\beta_0)\right]\mathrm{tr}\left[\frac{\partial}{\partial\beta_{j_2}}\Xi_{k-1}(\beta_0)\Xi_{k-1}^{-1}(\beta_0)\right]\right. \\
&\quad-\mathrm{tr}\left[\frac{\partial}{\partial\beta_{j_1}}\Xi_{k-1}(\beta_0)\Xi_{k-1}^{-1}(\beta_0)\right]\mathrm{tr}\left[\frac{\partial}{\partial\beta_{j_2}}\Xi_{k-1}(\beta_0)\Xi_{k-1}^{-1}(\beta_0)\right] \\
&\quad-\mathrm{tr}\left[\frac{\partial}{\partial\beta_{j_2}}\Xi_{k-1}(\beta_0)\Xi_{k-1}^{-1}(\beta_0)\right]\mathrm{tr}\left[\frac{\partial}{\partial\beta_{j_1}}\Xi_{k-1}(\beta_0)\Xi_{k-1}^{-1}(\beta_0)\right] \\
&\quad+\mathrm{tr}\left[\left(\frac{\partial}{\partial\beta_{j_1}}\Xi_{k-1}(\beta_0)\right)\Xi_{k-1}^{-1}(\beta_0)\right]\mathrm{tr}\left[\left(\frac{\partial}{\partial\beta_{j_2}}\Xi_{k-1}(\beta_0)\right)\Xi_{k-1}^{-1}(\beta_0)\right] \\
&\quad\left.+2\mathrm{tr}\left[\left(\frac{\partial}{\partial\beta_{j_1}}\Xi_{k-1}(\beta_0)\right)\Xi_{k-1}^{-1}(\beta_0)\left(\frac{\partial}{\partial\beta_{j_2}}\Xi_{k-1}(\beta_0)\right)\Xi_{k-1}^{-1}(\beta_0)\right]\right\} \\
&\quad+\sum_{k=1}^{n}\left\{R\left(\frac{1}{n^2},X_{t_{k-1}}\right)+R\left(\frac{\varepsilon^{-2}}{n^3},X_{t_{k-1}}\right)+R\left(\frac{\varepsilon^{-4}}{n^5},X_{t_{k-1}}\right)\right\} \\
&\to 4I_\sigma^{j_1j_2}(\theta_0)
\end{aligned}
$$

in P_{θ_0}-probability as $\varepsilon\to 0$ and $n\to\infty$. ∎

Proof of (12). By Lemma 1–(i) and (iii), we obtain

$$
\sum_{k=1}^{n}E_{\theta_0}[\xi_k^{i}\eta_k^{j}(\theta_0)|\mathcal{G}_{k-1}^n]
$$

$$
\begin{aligned}
&= -2\frac{1}{\varepsilon\sqrt{n}}\sum_{k=1}^{n}\sum_{l_1=1}^{d}\left[\left(\frac{\partial}{\partial\alpha_i}b(X_{t_{k-1}},\alpha_0)\right)^*\Xi_{k-1}^{-1}(\beta_0)\right]^{l_1}\frac{\partial}{\partial\beta_j}\log\det\Xi_{k-1}(\beta_0)E_{\theta_0}[P_k^{l_1}(\alpha_0)|\mathcal{G}_{k-1}^n] \\
&\quad -2\varepsilon^{-3}\sqrt{n}\sum_{k=1}^{n}\sum_{l_1,l_2,l_3=1}^{d}\left[\left(\frac{\partial}{\partial\alpha_i}b(X_{t_{k-1}},\alpha_0)\right)^*\Xi_{k-1}^{-1}(\beta_0)\right]^{l_1}\left(\frac{\partial}{\partial\beta_j}\Xi_{k-1}^{-1}(\beta_0)\right)^{l_2l_3} \\
&\quad \times E_{\theta_0}[P_k^{l_1}P_k^{l_2}P_k^{l_3}(\alpha_0)|\mathcal{G}_{k-1}^n] \\
&= \sum_{k=1}^{n}R\left(\frac{1}{\varepsilon n^2\sqrt{n}},X_{t_{k-1}}\right) \\
&\quad +\sum_{k=1}^{n}\left\{R\left(\frac{\varepsilon}{n\sqrt{n}},X_{t_{k-1}}\right)+R\left(\frac{\varepsilon^{-1}}{n^2\sqrt{n}},X_{t_{k-1}}\right)+R\left(\frac{\varepsilon^{-3}}{n^3\sqrt{n}},X_{t_{k-1}}\right)\right\} \\
&\to 0
\end{aligned}
$$

in P_{θ_0}-probability as $\varepsilon\to 0$ and $n\to\infty$. ∎

Proof of (13). Using Lemma 1–(iv), one has

$$
\begin{aligned}
&\sum_{k=1}^{n}E_{\theta_0}[(\xi_k^i(\theta_0))^4|\mathcal{G}_{k-1}^n] \\
&= 16\varepsilon^{-4}\sum_{k=1}^{n}\sum_{l_1,l_2,l_3,l_4=1}^{d}\left[\left(\frac{\partial}{\partial\alpha_i}b(X_{t_{k-1}},\alpha_0)\right)^*\Xi_{k-1}^{-1}(\beta_0)\right]^{l_1}\left[\left(\frac{\partial}{\partial\alpha_i}b(X_{t_{k-1}},\alpha_0)\right)^*\Xi_{k-1}^{-1}(\beta_0)\right]^{l_2} \\
&\quad \times\left[\left(\frac{\partial}{\partial\alpha_i}b(X_{t_{k-1}},\alpha_0)\right)^*\Xi_{k-1}^{-1}(\beta_0)\right]^{l_3}\left[\left(\frac{\partial}{\partial\alpha_i}b(X_{t_{k-1}},\alpha_0)\right)^*\Xi_{k-1}^{-1}(\beta_0)\right]^{l_4} \\
&\quad \times E_{\theta_0}[P_k^{l_1}P_k^{l_2}P_k^{l_3}P_k^{l_4}(\alpha_0)|\mathcal{G}_{k-1}^n] \\
&= \sum_{k=1}^{n}\left\{R\left(\frac{1}{n^2},X_{t_{k-1}}\right)+R\left(\frac{\varepsilon^{-2}}{n^4},X_{t_{k-1}}\right)+R\left(\frac{\varepsilon^{-4}}{n^5},X_{t_{k-1}}\right)\right\} \\
&\to 0
\end{aligned}
$$

in P_{θ_0}-probability as $\varepsilon\to 0$ and $n\to\infty$. ∎

Proof of (14). We first obtain several estimates as follows:

$$
\begin{aligned}
(\eta_k^j(\theta_0))^4 &\le 2^3\left[\frac{1}{n^2}\left(\frac{\partial}{\partial\beta_j}\log\det\Xi_{k-1}(\beta_0)\right)^4\right. \\
&\quad \left.+\varepsilon^{-8}n^2(2d)^3\sum_{l_1l_2=1}^{d}\left[\left(\frac{\partial}{\partial\beta_j}\Xi_{k-1}^{-1}(\beta_0)\right)^{l_1l_2}\right]^4(P_k^{l_1}P_k^{l_2}(\alpha_0))^4\right], \\
E_{\theta_0}[(P_k^{l_1}P_k^{l_2})^4(\alpha_0)|\mathcal{G}_{k-1}^n] &\le 3^3\left\{E_{\theta_0}[(\phi^{l_1}\phi^{l_2})^4(X_{t_{k-1}},X_{t_k})|\mathcal{G}_{k-1}^n]\right. \\
&\quad +\frac{1}{n^4}(b^{l_1}(X_{t_{k-1}},\alpha_0))^4E_{\theta_0}[(\phi^{l_2})^4(X_{t_{k-1}},X_{t_k})|\mathcal{G}_{k-1}^n] \\
&\quad \left.+\frac{1}{n^4}(b^{l_2}(X_{t_{k-1}},\alpha_0))^4E_{\theta_0}[(P_k^{l_1})^4(\alpha_0)|\mathcal{G}_{k-1}^n]\right\}.
\end{aligned}
$$

In the same way as the proof of Lemma 1, we have

$$
E_{\theta_0}[\phi_8(X_{t_{k-1}},X_{t_k})|\mathcal{G}_{k-1}^n] = R\left(\frac{\varepsilon^8}{n^4},X_{t_{k-1}}\right)+R\left(\frac{\varepsilon^6}{n^5},X_{t_{k-1}}\right)+R\left(\frac{\varepsilon^4}{n^6},X_{t_{k-1}}\right) \quad (15)
$$

$$+R\left(\frac{\varepsilon^2}{n^7}, X_{t_{k-1}}\right) + R\left(\frac{1}{n^8}, X_{t_{k-1}}\right).$$

It then follows form (15) and Lemma 1–(iv) that

$$\begin{aligned} E_{\theta_0}[(P_k^{l_1}P_k^{l_2})^4(\alpha_0)|\mathcal{G}_{k-1}^n] &= R\left(\frac{\varepsilon^8}{n^4}, X_{t_{k-1}}\right) + R\left(\frac{\varepsilon^6}{n^5}, X_{t_{k-1}}\right) + R\left(\frac{\varepsilon^4}{n^6}, X_{t_{k-1}}\right) \\ &\quad + R\left(\frac{\varepsilon^2}{n^7}, X_{t_{k-1}}\right) + R\left(\frac{1}{n^8}, X_{t_{k-1}}\right). \end{aligned}$$

Thus, one has

$$\begin{aligned} \sum_{k=1}^n E_{\theta_0}[(\eta_k^j(\theta_0))^4|\mathcal{G}_{k-1}^n] &\leq \sum_{k=1}^n \left\{ R\left(\frac{1}{n^2}, X_{t_{k-1}}\right) + R\left(\frac{\varepsilon^{-2}}{n^3}, X_{t_{k-1}}\right) + R\left(\frac{\varepsilon^{-4}}{n^4}, X_{t_{k-1}}\right) \right. \\ &\quad \left. + R\left(\frac{\varepsilon^{-6}}{n^5}, X_{t_{k-1}}\right) + R\left(\frac{\varepsilon^{-8}}{n^6}, X_{t_{k-1}}\right) \right\} \\ &\to 0 \end{aligned}$$

in P_{θ_0}-probability as $\varepsilon \to 0$ and $n \to \infty$. This completes the proof. ∎

Proof of Theorem 1. First of all, in order to prove consistency of $\hat{\theta}_{\varepsilon,n}$, we note that in view of the compactness of $\bar{\Theta}$, there exists a subsequence (ε_k, n_k) such that $\hat{\theta}_{\varepsilon_k,n_k}$ goes to a limit $\theta_\infty = (\alpha_\infty, \beta_\infty) \in \bar{\Theta}$.

Next, it follows from Lemma 4-(i) and the continuity of $U_1(\alpha, \alpha_0, \beta)$ with respect to (α, β) that under P_{θ_0}, as $\varepsilon_k \to 0$ and $n_k \to \infty$, one has

$$\varepsilon_k^2 U_{\varepsilon_k,n_k}(\hat{\alpha}_{\varepsilon_k,n_k}, \hat{\beta}_{\varepsilon_k,n_k}) - \varepsilon_k^2 U_{\varepsilon_k,n_k}(\alpha_0, \hat{\beta}_{\varepsilon_k,n_k}) \to U_1(\alpha_\infty, \alpha_0, \beta_\infty). \tag{16}$$

From the definition of the minimum contrast estimator $\hat{\theta}_{\varepsilon,n}$ and $\alpha_0 \in \bar{\Theta}_\alpha$,

$$\varepsilon_k^2 U_{\varepsilon_k,n_k}(\hat{\alpha}_{\varepsilon_k,n_k}, \hat{\beta}_{\varepsilon_k,n_k}) - \varepsilon_k^2 U_{\varepsilon_k,n_k}(\alpha_0, \hat{\beta}_{\varepsilon_k,n_k}) \leq 0. \tag{17}$$

By (16), (17), [A4] and [A5], we obtain $\alpha_\infty = \alpha_0$. We now have deduced that any convergent subsequence of $\hat{\alpha}_{\varepsilon,n}$ goes to α_0. Thus, we complete the proof of the consistency of $\hat{\alpha}_{\varepsilon,n}$.

For the consistency of $\hat{\beta}_{\varepsilon,n}$, using Lemma 4-(ii) and the continuity of $U_2(\alpha, \beta, \beta_0)$ with respect to (α, β), under P_{θ_0}, one has

$$\frac{1}{n_k} U_{\varepsilon_k,n_k}(\hat{\alpha}_{\varepsilon_k,n_k}, \hat{\beta}_{\varepsilon_k,n_k}) \to U_2(\alpha_0, \beta_\infty, \beta_0) \tag{18}$$

as $\varepsilon_k \to 0$ and $n_k \to \infty$, where we note that $\alpha_{\varepsilon_k,n_k}$ tends to α_0 by the previous proof. Moreover, in view of the definition of $\hat{\theta}_{\varepsilon,n}$ and $\beta_0 \in \bar{\Theta}_\beta$,

$$\frac{1}{n_k} U_{\varepsilon_k,n_k}(\hat{\alpha}_{\varepsilon_k,n_k}, \hat{\beta}_{\varepsilon_k,n_k}) \leq \frac{1}{n_k} U_{\varepsilon_k,n_k}(\hat{\alpha}_{\varepsilon_k,n_k}, \beta_0). \tag{19}$$

It follows from (18) and (19) that $U_2(\alpha_0,\beta_\infty,\beta_0)\le U_2(\alpha_0,\beta_0,\beta_0)$. Moreover, by a version of Lemma 17 in Genon-Catalot and Jacod [8], one has

$$\log\det[\sigma\sigma^*](X_t^0,\beta_0)+d\le\log\det[\sigma\sigma^*](X_t^0,\beta_\infty)+\mathrm{tr}\left[[\sigma\sigma^*](X_t^0,\beta_0)[\sigma\sigma^*]^{-1}(X_t^0,\beta_\infty)\right]$$

with equality iff $[\sigma\sigma^*](X_t^0,\beta_\infty)=[\sigma\sigma^*](X_t^0,\beta_0)$. Hence $U_2(\alpha_0,\beta_\infty,\beta_0)\ge U_2(\alpha_0,\beta_0,\beta_0)$. Thus [A5] together with the above inequalities for U_2 implies $\beta_\infty=\beta_0$. Therefore, by noting that any convergent subsequence of $\hat\beta_{\varepsilon,n}$ tends to β_0, we finish proving the consistency of $\hat\beta_{\varepsilon,n}$. ∎

Finally we prove the asymptotic normality of $\hat\theta_{\varepsilon,n}$. Let $B(\theta_0;\rho)=\{\theta:|\theta-\theta_0|\le\rho\}$. It follows from $\theta_0\in\Theta$ that for sufficiently small $\rho>0$, $B(\theta_0;\rho)\subset\Theta$. By Taylor's formula, one has, if $\hat\theta_{\varepsilon,n}\in B(\theta_0;\rho)$,

$$D_{\varepsilon,n}S_{\varepsilon,n}=\Lambda_{\varepsilon,n},$$

where $D_{\varepsilon,n}=\int_0^1 C_{\varepsilon,n}\left(\theta_0+u(\hat\theta_{\varepsilon,n}-\theta_0)\right)du$ and $S_{\varepsilon,n}=(\varepsilon^{-1}(\hat\alpha_{\varepsilon,n}-\alpha_0)^*,\sqrt{n}(\hat\beta_{\varepsilon,n}-\beta_0)^*)^*$. It follows from the consistency of $\hat\theta_{\varepsilon,n}$ that for sufficiently small $\rho>0$,

$$P_{\theta_0}[\hat\theta_{\varepsilon,n}\in\Theta]\ge P_{\theta_0}[|\hat\theta_{\varepsilon,n}-\theta_0|<\rho]\to 1$$

as $\varepsilon\to0$ and $n\to\infty$. Moreover, by the consistency of $\hat\theta_{\varepsilon,n}$, there exists a sequence $\{B(\theta_0;\eta_{\varepsilon,n})\}$ such that $\eta_{\varepsilon,n}\to0$ and $P_{\theta_0}[\hat\theta_{\varepsilon,n}\in B(\theta_0;\eta_{\varepsilon,n})]\to1$ as $\varepsilon\to0$ and $n\to\infty$. Since we obtain

$$P_{\theta_0}[\hat\theta_{\varepsilon,n}\in\Theta^c\cup B(\theta_0;\eta_{\varepsilon,n})^c]\le P_{\theta_0}[\hat\theta_{\varepsilon,n}\in\Theta^c]+P_{\theta_0}[\hat\theta_{\varepsilon,n}\in B(\theta_0;\eta_{\varepsilon,n})^c]\to0$$

as $\varepsilon\to0$ and $n\to\infty$, one has $1_{\{\hat\theta_{\varepsilon,n}\in\Theta^c\cup B(\theta_0;\eta_{\varepsilon,n})^c\}}\to0$ in P_{θ_0}-probability as $\varepsilon\to0$ and $n\to\infty$. By Lemma 5-(ii), letting $R_n=D_{\varepsilon,n}-C_{\varepsilon,n}(\theta_0)$,

$$|R_{\varepsilon,n}|\cdot 1_{\{\hat\theta_{\varepsilon,n}\in\Theta\cap B(\theta_0;\eta_{\varepsilon,n})\}}\le\sup_{\theta\in B(\theta_0;\eta_{\varepsilon,n})}|C_{\varepsilon,n}(\theta)-C_{\varepsilon,n}(\theta_0)|\to0$$

in P_{θ_0}-probability as $\varepsilon\to0$ and $n\to\infty$. Thus, under P_{θ_0}, as $\varepsilon\to0$ and $n\to\infty$, one has $R_{\varepsilon,n}\to0$. Using Lemma 5-(i), $D_{\varepsilon,n}\to2I(\theta_0)$ in P_{θ_0}-probability as $\varepsilon\to0$ and $n\to\infty$.

Let $\Gamma(\theta)$ be the limit of $C_{\varepsilon,n}(\theta)$ as $\varepsilon\to0$ and $n\to\infty$. For details of $\Gamma(\theta)$, see (5), (6) and (7) in proof of Lemma 5. Note that $\Gamma(\theta)$ is continuous with respect to θ. Since $I(\theta_0)$ is positive definite, there exists a positive constant C such that $\inf_{|x|=1}|I(\theta_0)x|>2C$. For such $C>0$, there exist $N_1(C)>0$ and $N_2(C)>0$ such that for any $\varepsilon<N_1(C)$ $(\varepsilon>0)$ and $n>N_2(C)$, and for any $\delta\in[0,1]$, $B(\theta_0;\eta_{\varepsilon,n})\subset\Theta$ and $|\Gamma(\theta_0+\delta\eta_{\varepsilon,n})-\Gamma(\theta_0)|<C/2$, where $\eta_{\varepsilon,n}\to0$ as $\varepsilon\to0$ and $n\to\infty$. For such $C>0$, let $\mathcal{C}_{\varepsilon,n}$ be the set defined by

$$\mathcal{C}_{\varepsilon,n}=\left\{\sup_{\theta\in\bar\Theta}|C_{\varepsilon,n}(\theta)-\Gamma(\theta)|<\frac{C}{2},\ \hat\theta_{\varepsilon,n}\in B(\theta_0;\eta_{\varepsilon,n})\right\}.$$

For any $\varepsilon<N_1(C)$ $(\varepsilon>0)$ and $n>N_2(C)$, and for any $|\delta|<1$, one has, on $\mathcal{C}_{\varepsilon,n}$,

$$\begin{aligned}\sup_{|x|=1}|(-D_n+\Gamma(\theta_0))x| &\le \sup_{|x|=1}|(-D_n+\Gamma(\theta_0+\delta\eta_{\varepsilon,n}))x|+\sup_{|x|=1}|(\Gamma(\theta_0)-\Gamma(\theta_0+\delta\eta_{\varepsilon,n}))x|\\ &\le \sup_{|\theta-\theta_0|<\eta_{\varepsilon,n}}|C_{\varepsilon,n}(\theta)-\Gamma(\theta)|+\frac{C}{2}\\ &< C.\end{aligned}$$

Hence, for any $\varepsilon < N_1(C)$ $(\varepsilon > 0)$ and $n > N_2(C)$, we obtain, on $\mathcal{C}_{\varepsilon,n}$,

$$\begin{aligned}\inf_{|x|=1} |D_{\varepsilon,n}x| &\geq \inf_{|x|=1} |\Gamma(\theta_0)x| - \sup_{|x|=1} |(-D_{\varepsilon,n} + \Gamma(\theta_0))x| \\ &> C.\end{aligned}$$

Let $\mathcal{D}_{\varepsilon,n} = \{D_{\varepsilon,n} \text{ is invertible}\}$. It then follows that for any $\varepsilon < N_1(C)$ $(\varepsilon > 0)$ and $n > N_2(C)$, $P_{\theta_0}[\mathcal{D}_{\varepsilon,n}] \geq P_{\theta_0}[\mathcal{C}_{\varepsilon,n}]$. Since it follows from (5), (6) and (7) that under [B2], $P_{\theta_0}[\mathcal{C}_{\varepsilon,n}] \to 1$ as $\varepsilon \to 0$ and $n \to \infty$, one has $P_{\theta_0}[\mathcal{D}_{\varepsilon,n}] \to 1$ as $\varepsilon \to 0$ and $n \to \infty$.

Let $\mathcal{E}_{\varepsilon,n} = \{\hat{\theta}_{\varepsilon,n} \in \Theta\} \cap \mathcal{D}_{\varepsilon,n}$, and $E_{\varepsilon,n} = D_{\varepsilon,n}$ on $\mathcal{E}_{\varepsilon,n}$ and $E_{\varepsilon,n} = J_{p+q}$ on $\mathcal{E}^c_{\varepsilon,n}$, where J_{p+q} is the $(p+q)\times(p+q)$ identity matrix. Note that $P_{\theta_0}[\mathcal{E}_{\varepsilon,n}] \to 1$ as $\varepsilon \to 0$ and $n \to \infty$. Since $|E_{\varepsilon,n} - 2I(\theta_0)|1_{\mathcal{E}_{\varepsilon,n}} \leq |D_{\varepsilon,n} - 2I(\theta_0)|$ and $1_{\mathcal{E}_{\varepsilon,n}} \to 1$ in P_{θ_0}-probability as $\varepsilon \to 0$ and $n \to \infty$, one has $E_{\varepsilon,n} \to 2I(\theta_0)$ in P_{θ_0}-probability as $\varepsilon \to 0$ and $n \to \infty$. Noting that $S_{\varepsilon,n}1_{\mathcal{E}_{\varepsilon,n}} = E^{-1}_{\varepsilon,n}D_{\varepsilon,n}S_{\varepsilon,n}1_{\mathcal{E}_{\varepsilon,n}} = E^{-1}_{\varepsilon,n}\Lambda_{\varepsilon,n}1_{\mathcal{E}_{\varepsilon,n}}$ and by Lemma 6, $S_{\varepsilon,n}1_{\mathcal{E}_n} \to N(0, I(\theta_0)^{-1})$ in distribution as $\varepsilon \to 0$ and $n \to \infty$. And, again using the fact that under P_{θ_0}, as $\varepsilon \to 0$ and $n \to \infty$, $1_{\mathcal{E}_{\varepsilon,n}} \to 1$, we complete the proof. ∎

Proof of Corollary 1. When $\sigma(x,\beta) = \sigma(x)$, it is easy to show that Lemmas 4–(i), 5, 6 hold under the assumptions [A1],[A2],[A3'],[A4'] and [B1]. In the same way as the proof of Theorem 1, we deduce the result. ∎

Acknowledgements

The research was carried out during a yearlong visit by Masayuki Uchida in 2001-2002 to University of Copenhagen in the program for Overseas Researchers of the Japan Ministry of Education, Culture, Sports, Science and Technology. Masayuki Uchida is grateful to Department of Statistics and Operations Research, University of Copenhagen for kind hospitality. The research of Michael Sørensen was supported by MaPhySto-Centre for Mathematical Physics and Stochastics, funded by a grant from The Danish National Research Foundation, and by the European Commission through the Human Potential Programme.

References

[1] Azencott, R.: Formule de Taylor stochastique et développement asymptotique d'intégrales de Feynmann. *Séminaire de Probabilités XVI; Supplément: Géométrie Différentielle Stochastique. Lecture Notes In Math.* **921**, 237-285 Springer Verlag, Berlin (1982)

[2] Bibby, B. M., Sørensen, M.: Martingale estimating functions for discretely observed diffusion processes. Bernoulli **1**, 17-39 (1995)

[3] Bibby, B. M., Sørensen, M.: Simplified estimating functions for diffusion models with a high-dimensional parameter. Scand. J. Statist. **28**, 99-112 (2001)

[4] Dacunha-Castelle, D., Florens-Zmirou, D.: Estimation of the coefficient of a diffusion from discrete observations. Stochastics **19**, 263-284 (1986)

[5] Dermoune, A., Kutoyants, Yu. A.: Expansion of distribution function of maximum likelihood estimate for misspecified diffusion type observations. Stochastics and Stochastic Reports **52**, 121-145 (1995)

[6] Florens-Zmirou, D.: Approximate discrete time schemes for statistics of diffusion processes. Statistics **20**, 547-557 (1989)

[7] Genon-Catalot, V.: Maximum contrast estimation for diffusion processes from discrete observations. Statistics **21**, 99-116 (1990)

[8] Genon-Catalot, V., Jacod, J.: On the estimation of the diffusion coefficient for multidimensional diffusion processes. Ann. Inst. Henri Poincaré Probab. Statist. **29**, 119-151 (1993)

[9] Gobet, E.: Local asymptotic mixed normality property for elliptic diffusion: a Malliavin calculus approach. Bernoulli **7**, 899-912 (2001)

[10] Gobet, E.: LAN property for ergodic diffusions with discrete observations. To appear in Ann. Inst. Henri Poincaré Probab. Statist. (2002)

[11] Hall, P., Heyde, C.: Martingale limit theory and its applications. Academic Press, New York (1980)

[12] Hansen, L. P., Scheinkman, J. A.: Back to the future: generating moment implications for continuous-time Markov processes. Econometrica **63**, 767-804 (1995)

[13] Ibragimov, I. A., Has'minskii, R. Z.: Statistical estimation. Springer Verlag, New York (1981)

[14] Ikeda, N., Watanabe, S.: Stochastic differential equations and diffusion processes, second edition. North-Holland/Kodansha, Tokyo (1989)

[15] Jacobsen, M.: Discretely observed diffusions; classes of estimating functions and small Δ-optimality. Scand. J. Statist. **28**, 123-150 (2001)

[16] Kessler, M.: Estimation of an ergodic diffusion from discrete observations. Scand. J. Statist. **24**, 211-229 (1997)

[17] Kessler, M.: Simple and explicit estimating functions for a discretely observed diffusion process. Scand. J. Statist. **27**, 65-82 (2000)

[18] Kessler, M., Sørensen, M.: Estimating equations based on eigenfunctions for a discretely observed diffusion process. Bernoulli **5**, 299-314 (1999)

[19] Kim, Y. J., Kunitomo, N.: Pricing options under stochastic interest rates. Asia Pacific Financial Markets **6**, 49-70 (1999)

[20] Kunitomo, N., Takahashi, A.: The asymptotic expansion approach to the valuation of interest rate contingent claims. Mathematical Finance **11**, 117-151 (2001)

[21] Kutoyants, Yu. A.: Parameter estimation for stochastic processes. Prakasa Rao, B.L.S. (ed.) Heldermann, Berlin (1984)

[22] Kutoyants, Yu. A.: Identification of dynamical systems with small noise. Kluwer, Dordrecht (1994)

[23] Laredo, C. F.: A sufficient condition for asymptotic sufficiency of incomplete observations of a diffusion process. Ann. Statist. **18**, 1158-1171 (1990)

[24] Sakamoto, Y., Yoshida, N.: Expansion of perturbed random variables based on generalized Wiener functionals. J. Multivariate Anal. **59**, 34-59 (1996)

[25] Sørensen, H.: Discretely observed diffusions: approximation of the continuous-time score function. Scand. J. Statist. **28**, 113-121 (2001)

[26] Sørensen, M.: Estimating functions for discretely observed diffusions: a review. In Basawa, I. V., Godambe, V. P. and Taylor, R. L., editors, Selected Proceedings of the Symposium on Estimating Functions, Hayward: Institute of Mathematical Statistics. IMS Lecture Notes-Monograph Series, **32**, 305-325 (1997)

[27] Sørensen, M.: Small dispersion asymptotics for diffusion martingale estimating functions. Preprint No. 2000-2, Department of Statistics and Operations Research, University of Copenhagen (2000)

[28] Takahashi, A.: An asymptotic expansion approach to pricing contingent claims. Asia Pacific Financial Markets **6**, 115-151 (1999)

[29] Uchida, M., Yoshida, N.: Information criteria for small diffusions via the theory of Malliavin-Watanabe. submitted (1999)

[30] Uchida, M., Yoshida, N.: Asymptotic expansion for small diffusions applied to option pricing. submitted (2001)

[31] Uchida, M.: Local asymptotic normality for discretely observed diffusion processes with small noise. In preparation (2002)

[32] Yoshida, N.: Asymptotic behavior of M-estimator and related random field for diffusion process. Ann. Inst. Statist. Math. **42**, 221-251 (1990)

[33] Yoshida, N.: Asymptotic expansion of maximum likelihood estimators for small diffusions via the theory of Malliavin-Watanabe. Probab. Theory Relat. Fields **92**, 275-311 (1992)

[34] Yoshida, N.: Asymptotic expansion for statistics related to small diffusions J. Japan Statist. Soc. **22**, 139-159 (1992)

[35] Yoshida, N.: Estimation for diffusion processes from discrete observation. J. Multivariate Anal. **41**, 220-242 (1992)

[36] Yoshida, N.: Asymptotic expansion of Bayes estimators for small diffusions. Probab. Theory Relat. Fields **95**, 429-450 (1993)

[37] Yoshida, N.: Asymptotic expansions for perturbed systems on Wiener space: maximum likelihood estimators. J. Multivariate Anal. **57**, 1-36 (1996)

[38] Yoshida, N.: Conditional asymptotics. submitted (2001)